GeoHealth

RESEARCH ARTICLE

10.1029/2024GH001027

Key Points:

- Low height-for-age has significant intra-area local variation
- The adopted approach showed superiority in characterizing the spatial effects and nonlinear effects of relevant height-for-age risk factors
- The results uncovered positive, negative, bell-shaped, and U-shaped non-linear associations between height-for-age and its related risk factors

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

A. Mansourian,
ali.mansourian@nateko.lu.se

Citation:

Nduwayezu, G., Kagoyire, C., Zhao, P., Eklund, L., Pilesjo, P., Bizimana, J. P., & Mansourian, A. (2024). Spatial machine learning for exploring the variability in low height-for-age from socioeconomic, agroecological, and climate features in the Northern Province of Rwanda. *GeoHealth*, *8*, e2024GH001027. https://doi.org/10.1029/2024GH001027

Received 5 FEB 2024
Accepted 22 AUG 2024

Author Contributions:

Conceptualization: Gilbert Nduwayezu, Ali Mansourian
Data curation: Gilbert Nduwayezu
Formal analysis: Gilbert Nduwayezu
Funding acquisition: Jean Pierre Bizimana, Ali Mansourian
Investigation: Gilbert Nduwayezu
Methodology: Gilbert Nduwayezu, Clarisse Kagoyire, Pengxiang Zhao, Lina Eklund, Ali Mansourian
Project administration: Ali Mansourian
Resources: Gilbert Nduwayezu

Spatial Machine Learning for Exploring the Variability in Low Height-For-Age From Socioeconomic, Agroecological, and Climate Features in the Northern Province of Rwanda

Gilbert Nduwayezu[1,2], Clarisse Kagoyire[1,3], Pengxiang Zhao[1], Lina Eklund[1], Petter Pilesjo[1], Jean Pierre Bizimana[4], and Ali Mansourian[1,5]

[1]Department of Physical Geography and Ecosystem Science, GIS Centre, Lund University, Lund, Sweden, [2]Department of Civil, Environmental and Geomatics Engineering, University of Rwanda, Kigali, Rwanda, [3]Centre for Geographic Information Systems and Remote Sensing, University of Rwanda, Kigali, Rwanda, [4]Department of Spatial Planning, University of Rwanda, Kigali, Rwanda, [5]Lund University's Profile Area: Nature-based Future Solutions, Lund University, Lund, Sweden

Abstract Childhood stunting is a serious public health concern in Rwanda. Although stunting causes have been documented, we still lack a more in-depth understanding of their local factors at a more detailed geographic level. We cross-sectionally examined 615 height-for-age prevalence observations in the Northern Province of Rwanda, linked with their related covariates, to explore the spatial heterogeneity in the low height-for-age prevalence by fitting linear and non-linear spatial regression models and explainable machine learning. Specifically, complemented with generalized additive models, we fitted the ordinary least squares (OLS), a standard geographically weighted regression (GWR), and multiscale geographically weighted regression (MGWR) models to characterize the imbalanced distribution of stunting risk factors and uncover the nonlinear effect of significant predictors, explaining the height-for-age variations. The results reveal that 27% of the children measured were stunted, and that likelihood was found to be higher in the districts of Musanze, Gakenke, and Gicumbi. The local MGWR model outperformed the ordinary GWR and OLS, with coefficients of determination of 0.89, 0.84, and 0.25, respectively. At specific ranges, the study shows that height-for-age decreases with an increase in the number of days a child was left alone, elevation, and rainfall. In contrast, land surface temperature is positively associated with height-for-age. However, variables like the normalized difference vegetation index, slope, soil fertility, and urbanicity exhibited bell-shaped and U-shaped non-linear associations with the height-for-age prevalence. Identifying areas with the highest rates of stunting will help determine the most effective measures for reducing the burden of undernutrition.

Plain Language Summary Local variations exist between height-for-age prevalence and its related risk factors. Global spatial regression methods, therefore, make it more difficult to locally revisit ongoing strategies and nutrition initiatives, particularly in areas where the burden of stunting was shown to be substantially higher. The main contribution of the present study lies in employing household-level information aggregated at a fine scale to model stunting using a local multiscale geographically weighted regression with generalized additive model (GAM) as interpretable machine learning to bridge traditional global linear models' gaps. Locally geographically weighted regressions assessed the spatial effects, and GAMs characterized the nonlinear effect of relevant height-for-age risk factors to potentially satisfy the needs of all end users. These findings revealed that low height-for-age has significant intra-area local variation and uncovered positive, negative, bell-shaped, and U-shaped non-linear associations between height-for-age and its related risk factors. The generated spatial maps highlight areas with a high prevalence of stunting, which can help the government and donor organizations allocate resources efficiently.

1. Introduction

Stunting, or low height-for-age, is a measure of linear growth retardation and cumulative growth deficits (UNICEF, 2017). A child is considered stunted if their height z-score is less than -2, and severe stunting is defined as a z-score less than -3 standard deviations below the median of the distribution of normal heights for children of the same age and sex (WHO, 2018). Child nutrition deficits have significant consequences, particularly in developing countries (UNICEF et al., 2021). Stunting is a matter of concern due to its wide-ranging impact on productivity and development (UNICEF, 2017). Stunting has multiple causes and is associated with

Supervision: Ali Mansourian
Validation: Gilbert Nduwayezu
Visualization: Gilbert Nduwayezu
Writing – original draft: Gilbert Nduwayezu
Writing – review & editing: Gilbert Nduwayezu, Pengxiang Zhao, Lina Eklund, Petter Pilesjo, Jean Pierre Bizimana, Ali Mansourian

severe short- and long-term health and psychological consequences, including impaired brain and physical development, lower school performance, and reduced future economic prospects (WHO, 2018). Globally, the prevalence of stunting has decreased from 33.1% (203.6 million) to 22% (149.2 million) between 2000 and 2020 (United Nations, 2021). However, in sub-Saharan Africa, 57.3% of children under the age of five are affected by stunting, a statistic that has remained largely unchanged for the past two decades (UNICEF et al., 2021). Stunting is a severe public health problem in Rwanda, where it affects 33% of children under the age of five, accounting for approximately 50% of all under-five mortality (NISR et al., 2021). The World Health Organization (WHO) considers a prevalence of more than 30% to be a very high stunting rate (WHO, 2018). According to the Rwanda Demographic and Health Survey (RDHS) for 2019–2020, one in three children under the age of five in Rwanda is stunted. The nutritional health of children varies significantly by province, with the highest stunting rate in the North (41%) and the lowest in the city of Kigali (21%) (NISR et al., 2021). While the general causes of stunting have been documented, there is still a lack of in-depth understanding of the local effects of its risk factors at a more detailed geographic level (Uwiringiyimana, Veldkamp, & Amer, 2019). To achieve the targets for 2020 and 2050, aimed at ending all forms of malnutrition, it is necessary to find innovative approaches to reduce the high prevalence of stunting (Republic of Rwanda, 2020). Rwanda has implemented significant country-specific nutrition strategies, including the national food and nutrition policy (MINISANTE, 2014), the one cow per poor household program (MINAGRI, 2006), and the childhood-based nutrition program (NECDP, 2017), to eliminate stunting. However, despite these initiatives, there is still a significant disparity between the target and the current prevalence (NISR et al., 2021; UN-HABITAT, 2018). Using a robust analytical framework, it is necessary to identify the local risk factors for stunting and understand the roles these determinants play in the geographical variability of stunting.

In Rwanda, there is limited research exploring the association between child stunting and socioeconomic, agroecological, and climate factors (Habyarimana et al., 2017; Mukabutera, Jamie, et al., 2016; Uwiringiyimana, Veldkamp, & Amer, 2019). Although the causes of stunting are well understood, no studies have been conducted to assess strategies that integrate both nutrition-specific and individual environmental and agroecological determinants of stunting at the local level. Some studies, such as Habyarimana et al. (2017) and Uwiringiyimana, Veldkamp, and Amer (2019) have touched upon spatial aspects by examining stunting predictions at the national level. Other studies have predominantly focused on the socioeconomic factors influencing child stunting, with limited attention given to the impact of agroecological variables (Habyarimana et al., 2017). Certain studies (Mukabutera, Thomson, Hedt-Gauthier et al., 2016; Uwiringiyimana, Ocké, et al., 2019; Weatherspoon et al., 2019) have utilized multivariate logistic regression models that rely on linear and parametric smooth functions. However, most predictors and response variables do not necessarily have linear relationships (Fotheringham et al., 2017). As a result, simple linear or traditional regression models fail to accurately capture these nonlinear relationships due to their susceptibility to local collinearity, which can lead to unreliable results (Li et al., 2020). These linear regression models estimate a global statistic that assumes a stationary and constant relationship over space, resulting in the same estimated parameters for the entire study area (Brunsdon et al., 1996). Unfortunately, they do not consider the spatial heterogeneity in the variables where relationships are measured (Li et al., 2020). Therefore, these models do not account for the spatial dynamics of the explanatory variables used (Fotheringham et al., 2017). In real life, the relationship between risk factors and disease incidence is often nonlinear (Li et al., 2020). The traditional linear model, although simple, often fails in these situations (Hastie et al., 2017). Geographically weighted regression (GWR) was developed to examine the non-stationary relationship between predictor and response variables (Brunsdon et al., 1996; Fotheringham et al., 2017). However, GWR has several drawbacks. First, the uniform bandwidth specified in a standard GWR may not be appropriate in cases where different predictors operate across different spatial scales (Li et al., 2020). Second, this approach assumes that all predictors have a uniform spatial bandwidth (Fotheringham et al., 2017). However, GWR overlooks the likelihood that various ecological predictors affect disease prevalence at various spatial scales (Li et al., 2020). This could potentially introduce bias into the model's results and compromise its performance (Oshan et al., 2019). The multiscale geographically weighted regression (MGWR), a derivative of GWR, addresses this issue by calibrating the regression model using different spatial scales (Li et al., 2020). Additionally, the MGWR model provides additional spatially weighted information about the relationship between covariates and response variables (Fotheringham et al., 2017). Most machine learning models are often considered "black boxes," which makes it challenging to understand the rationale behind the model, particularly in the field of healthcare (Molnar, 2022). To uncover precise relationships in the data, a set of explainable artificial intelligence features derived from these systems helps users understand how the model works and makes predictions for critical decision-making (Hastie et al., 2017). Generalized additive models (GAMs) have proven to be the most commonly used explainable artificial intelligence approach in the statistical community for exploring the nonlinear effects between disease incidence and associated risk factors, owing to their ability to model complex and nonlinear relationships (Hastie & Tibshirani, 1995). Unlike generalized linear models (GLMs), fitting GAMs involves a straightforward and flexible backfitting process that allows for the selection of the best fitting method for each input variable (Hastie et al., 2017).

The association between socioeconomic, agroecological, and climate variables and the low childhood height-for-age is complex and non-linear, involving both associative and causal perspectives (Molnar & Freiesleben, 2024; Rasmussen et al., 2016; Tellings, 2017). As documented by Woodward (2013), exploring associations from observational studies forms the starting point for causal hypothesizing, which is crucial in controlled experiments and longitudinal studies. Socioeconomic household related factors such as parental education (Habyarimana et al., 2017), household wealth (Weatherspoon et al., 2019), water and sanitation (Fink et al., 2011), and child feeding practices (Uwiringiyimana, Ocké, et al., 2019), can directly influence household food security, which can then impact dietary diversity and contribute to the low child height-for-age (Mukabutera, Jamie, et al., 2016). However, the causal relationships between these variables and the childhood height-for-age distribution are not uniformly discernible (Yeboah et al., 2022), as they can be influenced by multifaceted interactions and mediating agroecological and climatic variables (Johnson & Brown, 2014). Hence, socio-economic and demographic factors alone are not sufficient to fully understand the intricate issue of childhood stunting (Uwiringiyimana, Veldkamp, & Amer, 2019). Therefore, to fully understand the potential associations and causal pathways of childhood stunting, an integrated approach complimenting these factors with climatic and agroecological conditions is crucial for a more comprehensive analysis (Tusting et al., 2020; van der Merwe et al., 2022; Vilcins et al., 2018). There is increasing evidence that climate and agroecology are directly and indirectly effecting the health and well-being of children (R. E. Baker & Anttila-Hughes, 2020; Dasgupta & Robinson, 2023). However, such influence is substantial in developing countries where a largely number of population is depending on rain fed agriculture (Hagos et al., 2014). These metrics which have been extensively explored elsewhere (Amondo et al., 2023; Bangelesa et al., 2023; Blom et al., 2022; Brown et al., 2014; Grace et al., 2012; Johnson & Brown, 2014; Niles et al., 2021), play a physical, mechanical, or biological influence on animal and plant sourced foods via crop health (Hummel et al., 2018), disease transmission or metabolically child growth and cognitive development (Kismul et al., 2017). The climate affects the nutritional outcomes through agroecosystems pathways with diverse impacts on crops and diseases (Lobell & Field, 2007), livestock and aquatic food sources (Weatherspoon et al., 2019), impacting the components of food security and diet diversity including the availability and quality of food (Dror & Allen, 2011; Murphy & Allen, 2003). Particularly in Rwanda, the agriculture sector serves as a primary source of income and food for rural households (NISR, 2019). Each of these channels can drastically lowering the childhood nutrition status and health (Niles et al., 2021).

A few recent studies done in Africa have underlined the importance of the local geographic location in which a child resides (De Sherbinin, 2011). Furthermore, the evidence associating climate and child stunting, particularly at different geographic scales, remains scarce (Lopez-Carr et al., 2016). However, due to the scarcity of georeferenced data, remote sensing indicators can serve as proxies when associated with disease data, facilitating the extrapolation of model outputs across extensive geographical regions (Wimberly et al., 2021). To date, no research has empirically examined the relationship between child stunting and climate determinants using spatial machine learning at lower spatial scales. Additionally, no study has investigated the spatial scale in terms of (a) the influential range of stunting variables at different scales, and (b) the quantitative aspect of such influence across space. To address this gap, the present study aims to comprehend the significant spatial variations in infant stunting at a fine scale. The primary contribution of this study is the utilization of individual-level information aggregated at a fine scale to model stunting, employing a local MGWR with GAM as an interpretable machine learning approach to bridge the gaps of traditional linear parametric models. To the best of our knowledge, there is a lack of small-area spatial stunting modeling using more robust models like MGWR in scientific literature. Moreover, local spatial machine learning analyses for stunting with agroecological predictors at the fine-scale level have not been explored in Rwanda. Notably, this study is the first to conduct an in-depth investigation of spatial patterns of stunting using a combination of variables at the lower scale. The findings of this study will help identify areas that require interventions to ensure fair and effective utilization of resources in reducing all forms of undernourishment and achieving sustainable development.

2. Materials and Methods

2.1. Study Setting and Sample Selection

We conducted this study in the Northern Province of Rwanda, which is mainly characterized by mountains and hills with steep slopes as well as fragile and degraded soil types as a result of the dominance of small-scale farming (NISR, 2018). It is composed of five districts, namely Burera, Gakenke, Gicumbi, Musanze, and Rulindo. Most of the settlements in the Northern Province are located on hills and steep slopes (NISR et al., 2021). This consequently triggers a high level of vulnerability to floods and landslides (NISR, 2018). This cross-sectional survey targeted children under 3 years of age whose mothers were 18 years of age or older. We administered the survey questionnaire to the primary caretaker (biological mother) or guardian of the eligible child at the household level. We used the two-stage cluster random sampling strategy to determine the sample population. To calculate the required sample size, according to NISR et al. (2021), we considered the estimated under-5 stunting prevalence in the Northern Province (41%). The mathematical formula used in prevalence studies (Kish, 1965), is as follows:

$$n = \frac{Z^2 \times p(1-p)}{\varepsilon^2} \times \mathrm{DEFF} \tag{1}$$

where n is the sample size, Z is the z-score or the critical value associated with the 95% confidence interval (1.96), p is the estimated proportion of stunting among children in the Northern Province (0.4), ε is the desired level of precision or margin of error (0.05), and DEFF is the design effect of 1.5. Applying the formula, we got an n equal to 553.19 as the estimated sample size. After considering the non-response (NR) of 10% or 0.1, we obtained 615 households as the adjusted sample size. We chose these 615 households randomly from 137 sampling units (villages), which represents 5% of all 2,744 villages in the study area. First, to randomly select these 137 villages, we created a grid map overlaid on the study area map to make sure that all sampled villages were evenly distributed across the entire study area and to get at least one sample from every grid cell. This spatial method was applied to get lower-scale data from household data. Next, we determined the number of individual households selected from each village based on the population density. Villages with low and higher population density had 3–4 and 5–6 households interviewed, respectively. After that, using a list of all candidate households obtained from community health workers, we applied systematic random sampling to obtain the eligible households in each selected village (sampling unit). In case the first and the next selected household were flocked together, a replacement was done for the nearest eligible household. Figure 1 illustrates the distribution of all surveyed households in the study area. We gathered information on households' socio-economic, demographic, and household-related decision-making characteristics, the child's health status, feeding practices, dietary characteristics, maternal health conditions, dietary habits, and violence experience. In accordance with the study's purpose, we collected the location coordinates of surveyed households to be able to spatialize undernourishment and its associated factors. We focused on height-for-age as the main outcome variable and other selected sociodemographic covariates from the data set, as shown in Table 1. The height-for-age z-scores were estimated and categorized with reference to the WHO child growth standards (WHO & UNICEF, 2021) and the Rwanda Demographic and Health Survey (RDHS) (NISR et al., 2021). We used four continuous variables from the surveyed data set: the diet diversity score, the household food insecurity score, the number of times the mother has left the child alone for more than 1 hr in the past week, and the number of days a child was left in the care of another child, that is, someone less than 10 years old, for more than an hour during the last week, to complement a set of agroecological, climate, and geographic risk factors obtained from multiple data sets. We calculated the dietary diversity indicator for children based on maternal and child dietary intake information collected using 24-hr dietary recall information according to the WHO and UNICEF (2021) guidelines. We determined the minimum dietary diversity based on the food consumed from various food groups over the past 24-hr period. We again calculated the household food insecurity score variable on the basis of the answers to nine questions about the household's access to and food consumption during the preceding 7 days.

2.2. Exploratory Data Analysis and Spatial Dependency

Before performing spatial autocorrelation analysis, we first applied the grid-based aggregation approach to aggregate surveyed household point data to obtain fine-scale areal data (D. M. Baker & Valleron, 2014; Hassler et al., 2024; Souris, 2019). Areal health data is often aggregated by converting point data into continuous

10.1029/2024GH001027

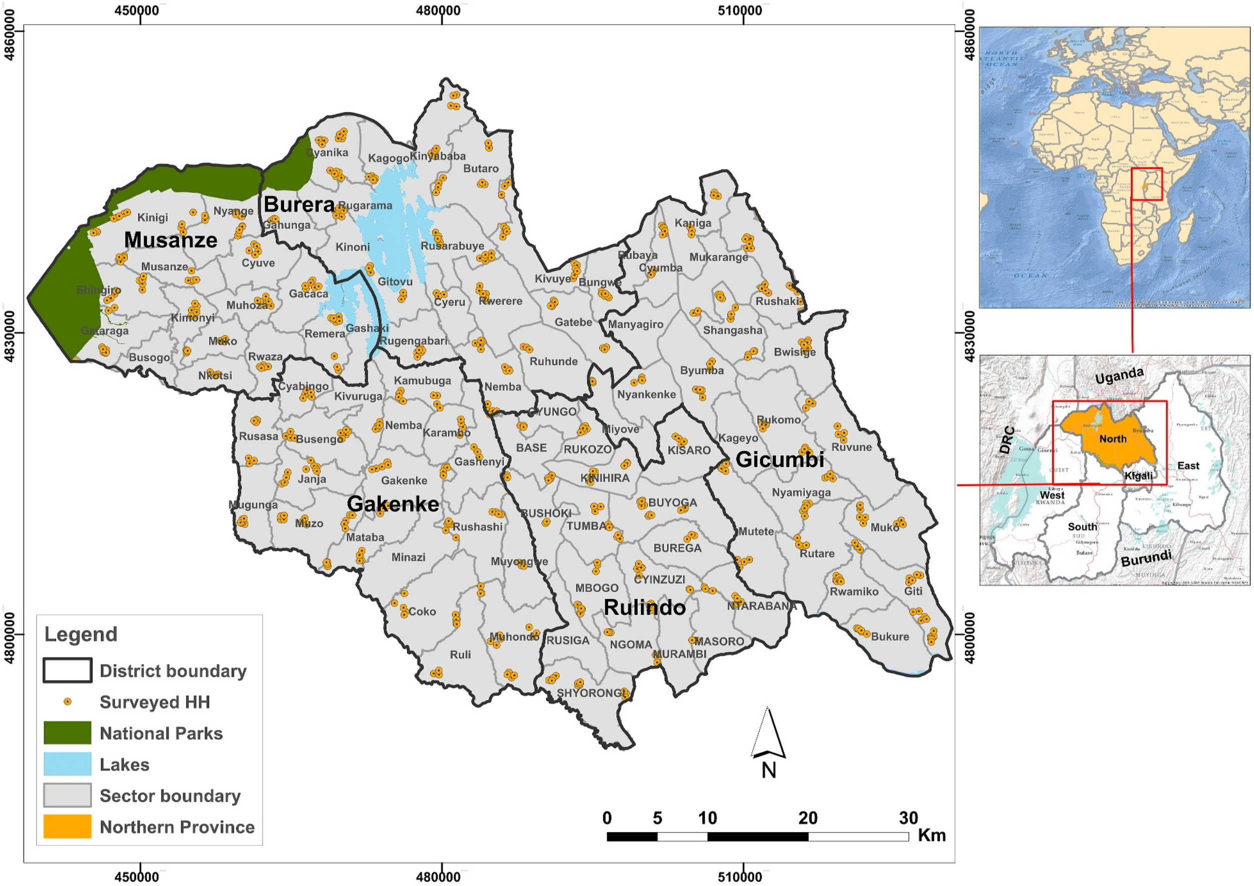

Figure 1. Study area location with the administrative boundary and surveyed household data.

information at geopolitical zones or disease catchments due to various reasons (Nduwayezu et al., 2023). These reasons include protecting patient confidentiality and safeguarding sensitive disease information (Moraga, 2019). To ensure the privacy of the survey respondents and mask private health information, we adopted a grid-based method by aggregating households' information over a geopolitical zone in Rwanda. Then, we computed global and local spatial autocorrelation analyses of height-for-age by estimating the *z*-scores and *p*-values for each areal unit (Anselin, 1995). We used Moran's I to check the global autocorrelation. After that, we computed the Getis-Ord Gi* statistic Anselin (1995) to identify the locations of statistically significant hot spots and cold spots in the data. Finally, to provide an in-depth exploratory data analysis, we computed the density and violin/box plots to characterize the concentration of height-for-age across districts.

2.3. Preparation of Predictors: Feature Screening and Selection

We compiled a comprehensive set of agroecological, climate, and geographic risk factors as input variables, including socioeconomic health covariates selected from the surveyed data. We utilized multiple data sets and joined them together to develop an integrated data set to assess child height-for-age outcomes. We used a modified zonal statistics-based approach developed by Nduwayezu et al. (2023) to record the average value of each variable raster cell in the attribute table of the height-for-age areal polygons. Table 1 and Text S3 in Supporting Information S1 provide grounds for the climatic and agroecological features used in this study to shed light onto its inclusion relevancy in the model, while Figure 2 and Figure S2 in Supporting Information S1 depict the spatial distribution of variables utilized in the modeling process.

Prior to standard GWR and multiscale GWR models, we fitted multilinear regression models to examine the multicollinearity between the predictor variables, which might introduce redundancy into the model. We used the variance inflation factor (VIF) to assess multicollinearity, which is a problem in spatial regression modeling

Table 1
The List of Responses and Covariates Used

Candidate indicator	Description/Hypothesis	Unit	Data source
Socio-economic			
Height-for-age (HAZ)	Height-for-age z-score estimates of children under 3 years of age according to the WHO and UNICEF (2021) standard	Number (−5.94) to (+4.79)	The cross-sectional survey
Diet diversity score	The minimum dietary diversity (MDD) is a scale from 0 to 10 that takes into account the food consumed from various food groups during the period of the previous 24 hr, including cereal grains, white tubers and root foods, dark leafy greens, vitamin A-rich vegetables and tubers, vitamin A-rich fruits, other fruits and vegetables, meat and fish foods, eggs, legumes, nuts, and seeds, as well as milk and milk products (Hussien et al., 2021; UNICEF et al., 2021). The more access to various types of food, the less stunting level attainment	Number 0–10	The cross-sectional survey
Household food insecurity score	This variable was calculated based on the answers to nine questions on the household's food intake and access during the previous 7 days. The higher the score, the more insecure a household is	Number 0–27	The cross-sectional survey
Days the child was left alone	Number of times the mother has left the child alone for more than an hour in the past week. The likelihood of having children with a higher stunting rate increases with the number of days a child is left alone	Number 0–10	The cross-sectional survey
Days the child was left with another child	The number of days a child was left in the care of another child, that is, someone less than 10 years old, for more than an hour during the last week. The assumption is that this caretaker lacks maturity and parenting experience. Thus, the stunting rate increases with the number of days left with another child	Number 0–8	The cross-sectional survey
Households using an improved water source	The estimated percentage of the population living in households whose main source of drinking water is an improved source. Improved sanitation and water access have been associated with a lower risk of stunting. With access to a better water source, the stunting level decreases (Fink et al., 2011; Sahiledengle et al., 2022)	Values 0–1	The DHS model Surface (Burgert-Brucker et al., 2018)
Households using open defecation	The estimated percentage of the population living in households using open defecation. Households using open defecation have a high likelihood of having stunted children. Improved sanitation and water access have been associated with a lower risk of stunting	Values 0–0.3	The DHS model Surface (Burgert-Brucker et al., 2018)
Climate			
NDVI	The values of the vegetation index vary from the least to the most vegetation. One of the crucial factors affecting stunting is the state of the vegetation. This is so that it can supply different ecosystem services and serve as a source of food and income from the forest (Bangelesa et al., 2023; Lopez-Carr et al., 2016; Niles et al., 2021; Vilcins et al., 2018)	Values (−1) to (+1)	https://scihub.copernicus.eu/dhus/#/home
Rainfall	Rainfall is associated with a higher prevalence of vector-borne diseases such as malaria due to the favorable conditions created for disease vectors. However, inadequate rainfall is associated with poorer agricultural yield, which worsens undernourishment. In addition, in arid areas, water scarcity due to a lack of rain may also result in poor sanitation (Fenta et al., 2021; Mukabutera, Thomson, Murray, et al., 2016; Niles et al., 2021)	Millimeters	https://www.worldclim.org/data/worldclim21.html

10.1029/2024GH001027

Table 1
Continued

Candidate indicator	Description/Hypothesis	Unit	Data source
Maximum and minimum air temperature	Temperature variations contribute to the direct health implications of climate change, including heat stress, illnesses, and air quality (Niles et al., 2021; van der Merwe et al., 2022)	Celsius degrees	https://www.worldclim.org/data/worldclim21.html
LST	The LST anomalies may impact childhood stunting through their influence on agricultural productivity, disease dynamics and heat stress (Lobell & Field, 2007; Prillaman, 2022). First, warmer LST stresses crops (Johnson & Brown, 2014), leading to food shortages and nutritional deficits (Hatfield & Prueger, 2015). Moreover, warmer LST can create more favorable conditions for the transmission of infectious diseases (Gage et al., 2008)	Celsius degrees	https://scihub.copernicus.eu/dhus/#/home
Agroecological			
Soil moisture	Soil moisture has an impact on plants both directly and indirectly since it increases evapotranspiration and water stress	Values 0–0.5	Zhang et al. (2023, https://zenodo.org/record/7172664#.ZAtcwh_MKUk)
Soil type	Shallow plant roots and poor plant growth resulting from compacted and high bulk density soils will influence crop yield and lower the amount of vegetative cover available to protect soil from erosion. It also impacts soil aeration, which affects the uptake of water and nutrients (Dong et al., 2021; Weatherspoon et al., 2019)	Values 0–1.51	MINAGRI
Livestock density	Counts of cattle in 2022 for each sector were used to calculate the livestock density per square kilometer. Livestock contributes directly to livelihoods and food security, and it affects diet and health (Niles et al., 2021; Robinson et al., 2014)	Values 0–933	MINAGRI
Geographic			
DEM	Children who live in lowlands have a lower risk of stunting than those who grow up in mountainous regions (Dang et al., 2008; Uwiringiyimana, Veldkamp, & Amer, 2019)	Meters	National Land Authority
Slope	Computed from DEM. Natural disasters such as flooding, and landslides are more likely in areas with steep slopes. The landslide-affected community is located on steep mountainous terrain, restricting accessibility to maternal and other child health care services (Nahalomo et al., 2022)	Degrees	National Land Authority
Population density	Population density estimates are calculated using total population counts for each sector. Urban areas have a lower probability of infectious disease incidence because they are associated with better living conditions, such as easier access to healthcare, compared to rural areas. Higher population density in urban areas facilitates the sharing of health-related information and resources	Values 74–17494	The fifth national population and housing census in Rwanda (NISR, 2022)
Urbanicity	Global human footprint index (from extremely rural to extremely urban). People living in urban areas have more access to infrastructure, including roads and transportation, which can provide market access and potentially improve incomes (Niles et al., 2021)	Values 0–274	https://hub.worldpop.org/geodata/summary?id=17232
Distance to OSM	Distance per grid centroid to the major roads. People close to the main road have high access to health facilities, which in turn facilitates the interaction of the population in need (children and pregnant women) with lifesaving interventions and treatment that prevent mortality (Aoun et al., 2015; Karra et al., 2017)	Values 0–20	https://hub.worldpop.org/geodata/summary?id=17481, which was derived from OpenStreetMap

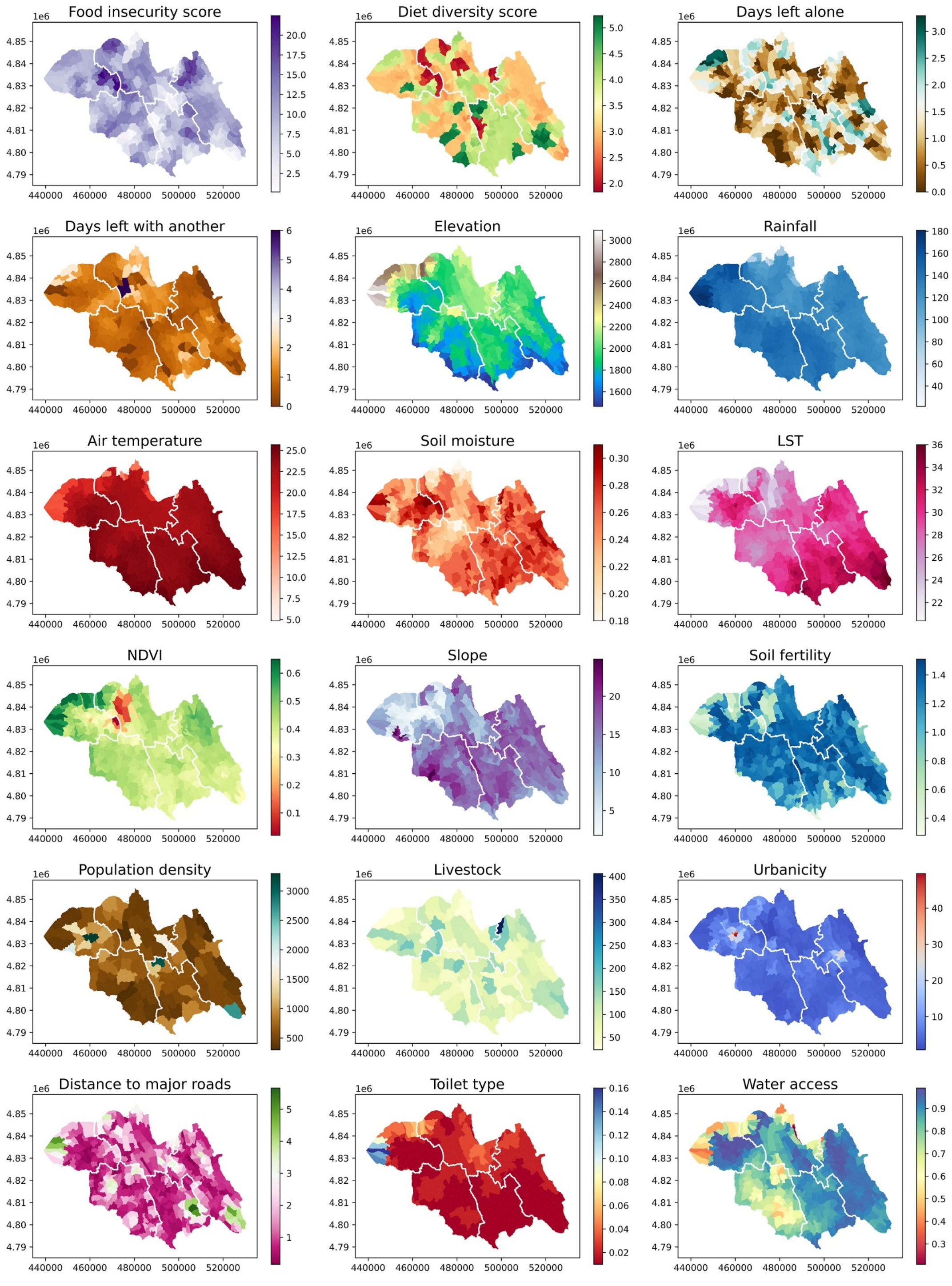

Figure 2. Spatial distribution of variables used.

(Cheng, 2003). VIF is a measure of how much the variance of the estimated regression coefficient increases if the explanatory variables are correlated. From the modeling perspective, any correlated variable can be selected as a predictor without significantly affecting the model's predictive performance. Once one of the correlated variables is used, the importance of the others is reduced. The higher the value of VIF, the greater the degree of collinearity. With a VIF greater than 10, there is strong evidence that collinearity is affecting the regression coefficients. During analysis, variables with a VIF higher than 10 were removed from the model. Later, to ensure that there is no strong correlation between predictors, we computed the Pearson correlation coefficient between the variables retained by the VIF test. The Pearson correlation coefficient describes the linear correlation between two features.

2.4. Multiscale Geographically Weighted Regression Model

Prior to GWR and MGWR, the ordinary least squares (OLS) model was fitted as baseline model for analyzing the data to examine the multicollinearity between the predictor variables and to detect the presence of spatially autocorrelated residuals, implying that local geographically weighted models may be appropriate. The MGWR algorithm calibrates the regression model using different spatial scales as opposed to GWR, which calibrates the regression model at the same spatial scale, and OLS, which estimates the global statistic that assumes a stationary and constant relationship over space. Further, OLS and GWR mask interesting spatial variations in the way that predictors influence the spatial distribution of relevant variables (Fotheringham et al., 2017). However, the MGWR model provides additional spatially weighted information about the relationship between covariates and response variables. The locally created MGWR model only uses a limited number of neighboring data points to train the model. The maximum distance between a data point and its kernel is called the bandwidth, and the area in which the local model operates is the neighborhood (or kernel) (Brunsdon et al., 1996). In essence, either the number of nearest neighbors (adoptive kernel) or a distance threshold value (bandwidth-fixed kernel) is used to build the neighborhood or kernel (Oshan et al., 2019). Its model is expressed mathematically as:

$$y_i = \beta_0(u_i, v_i) + \sum_{j=1}^{m} \beta_{bwj}(u_i, v_i) x_{ij} + \epsilon_i \tag{2}$$

where $\beta_0(u_i, v_i)$ is the intercept; β_{bwj} is the local coefficient at location i with coordinates (u_i, v_i), x_{ij} is the jth feature at location i, m is the number of predictor variables, and ϵ_i is the random error at location i. Variables resulting from OLS with a p-value less than 0.05, indicating variables with a statistically significant relationship with height-for-age were fed into the GWR and MGWR models. First, for approximating the kernel bandwidth for GWR and MGWR models, we employed the adaptive bi-square spatial kernel weighted method (Fotheringham et al., 2017). Next, we chose the default golden bandwidth search approach for computing uniform (GWR) and locally varying (MGWR) bandwidths. Then, we considered Akaike's Information Corrected Criterion (AICc) metric as an optimization criteria for model bandwidth selection (Fotheringham et al., 2017). After that, we computed the Monte Carlo tests to examine the variability of the parameter estimates for the local regression models (Oshan et al., 2019). If the p-value was lower than 0.05, then the variability of the influence of the dependent variable was confirmed, and that variable was identified as a local variable for the given type of experience; if not, then the parameter estimates indicated a relatively stationary pattern across the space, and accordingly, the independent variable could be interpreted as a global variable. Finally, we measured the bandwidth confidence intervals at different levels of probability to ensure reliable spatially varying bandwidths derived from both GWR and MGWR (Fotheringham et al., 2017). Bandwidth can also be seen as an indicator that reflects spatial scale. A higher bandwidth value indicates a broader spatial scale of influence. We used the AICs and the coefficients of determination to assess the goodness of fit of the MGWR model (Oshan et al., 2019). For these parameter estimates, a higher R^2 and a smaller AIC indicated the best fit of the model. In addition, we mapped and compared the spatial distribution of the standardized residuals for all models to better understand how local GWR models addressed spatial heterogeneity (Li et al., 2020). We also estimate spatial autocorrelation through local Moran's I to trace potential clustering in the residuals for OLS, GWR, and MGWR's residuals, with its associated z-score and residual p-values, for assessing the model's performance. A model with random pattern residuals is better. Another metric to evaluate the model's goodness of fit in regression models is the model parameter's estimates. The value of the parameter estimates indicated the intensity of the influence of the predictor on the response variable. We used MGWR2.2 software to calibrate all of the GWR and MGWR models (Oshan et al., 2019). These models were then imported into the Python 3.12 and ArcGIS Pro 3.2 environments for analysis and visualization.

2.5. GAMs

The GAM is a nonparametric interpretable machine learning model developed by Hastie and Tibshirani (1995) as an extension to the conventional GLMs that allow for nonparametric relationships between a response variable and predictors (Hastie et al., 2017). The approach uses regression splines to provide a flexible way of approximating the underlying regression functions with polynomials. Splines are often centered around the mean prediction, so a point on the curve represents the deviation from the mean prediction (Hastie & Tibshirani, 1995). Backfitting is a method used to fit GAMs using smoothing splines. This approach utilizes a cross-validation approach to fine-tune a smoothness parameter that is commonly used to control the flexibility of the curve (Molnar, 2022). GAMs offer a general framework for extending a standard linear model by enabling non-linear functions of each of the variables while maintaining additivity (Hastie & Tibshirani, 1995). First, GAMs allow to automatically model non-linear relationships that normal linear regression neglects by fitting a non-linear f_j to each x_j. The non-linear fits could potentially make more accurate predictions for the response y. Second, due to the additive nature of the model, the effect of each x_j on y could be examined individually while keeping all other variables fixed. Finally, the smoothness of the function f_j for the variable x_j can be summarized via degrees of freedom. The GAM model is expressed mathematically as follows:

$$y_i = \beta_0 + f_1(x_{i1}) + f_2(x_{i2}) + \ldots + f_p(x_{ip}) + \epsilon_i \tag{3}$$

where, β_0 is the intercept; $f_p(x_{ip})$ is a (smooth) nonlinear function, and ϵ_i is the random error at location i. This is known as an additive model because we calculate a separate f_p for each x_p and then add together all of their contributions. Because of the observed non-linear trends in height-for-age variations, the GAM algorithm was implemented. Subsequently, we used a GAM to examine the nonlinear and spatial effects of the potential height-for-age risk factors. We plotted the relationship between height-for-age with its relevant variables obtained from the MGWR model, taking into account their probability values. A detailed outline of the study's methodological framework is depicted in Figure 3.

3. Results

3.1. Exploratory Data Analysis and Spatial Dependency

The results obtained from the exploratory data analysis suggest that there are variations in the height-for-age rate within the study area. Out of the 601 observations on height-for-age among households that were retained, it was found that 27% of the children were stunted. The district with the highest prevalence of stunting was Musanze, whereas Rulindo District had the lowest stunting prevalence rate (refer to Figure 4a). Furthermore, the analysis conducted using the Getis-Ord Gi* Hot Spot method revealed a high clustering of height-for-age rates in the western sectors of Musanze District and the middle sectors of Gakenke and Gicumbi districts. On the other hand, the southern part of Rulindo District and the northern part of Gicumbi District exhibited cold (low) clusters of height-for-age. In Burera Districts, there was mainly no significant relationship observed (refer to Figure 4b).

Figure 5a shows a density plot, while Figure 5b displays violin/box plots, depicting the height-for-age distribution across five districts of the Northern Province. The figures reveal that Musanze District has a high density of individuals with low height-for-age, whereas Rulindo District exhibits a high prevalence of individuals with high height-for-age. The height-for-age pattern in the remaining districts is characterized by heterogeneity.

3.2. Local Geographically Weighted Regression Spatial Effects

The final OLS model was fitted using 17 variables, after removing the minimum and maximum air temperature variables to address the issue of multicollinearity. Figure S1, Tables S1, and S2 in Supporting Information S1 reveal that the VIF values for all selected variables were lower than 10, indicating the absence of severe multicollinearity. However, the model fit was found to be unsatisfactory, as evidenced by an adjusted R^2 of 0.25. The AIC for the linear model was 1065.672. Notably, the OLS residuals exhibited strong spatial clustering, as indicated by the results of Moran's I test: Moran's I = 0.893, z-score = 4.016, and p-value < 0.000 (Table 4). The presence of autocorrelated residuals in the OLS violates the assumption of independence of errors, necessitating caution in the interpretation of the estimated coefficients. Eight variables resulting from OLS regression analysis with a p-value less than 0.05, including the number of days a child was left alone, elevation, rainfall, soil fertility, LST, NDVI, slope, and urbanicity, were included in the GWR and MGWR models. We adopted this approach to ensure that only statistically significant variables are included in the regression models, thus avoiding the dilution of the significance of important variables and maintaining the overall accuracy of the models (Brunsdon et al., 1996). The selected bandwidths for the GWR and MGWR models are presented in Tables 2 and 3, respectively. The GWR model had a universal bandwidth of 59, whereas the bandwidths selected by the MGWR model varied for different variables. Notably, the bandwidths for most variables, except for the number of days a

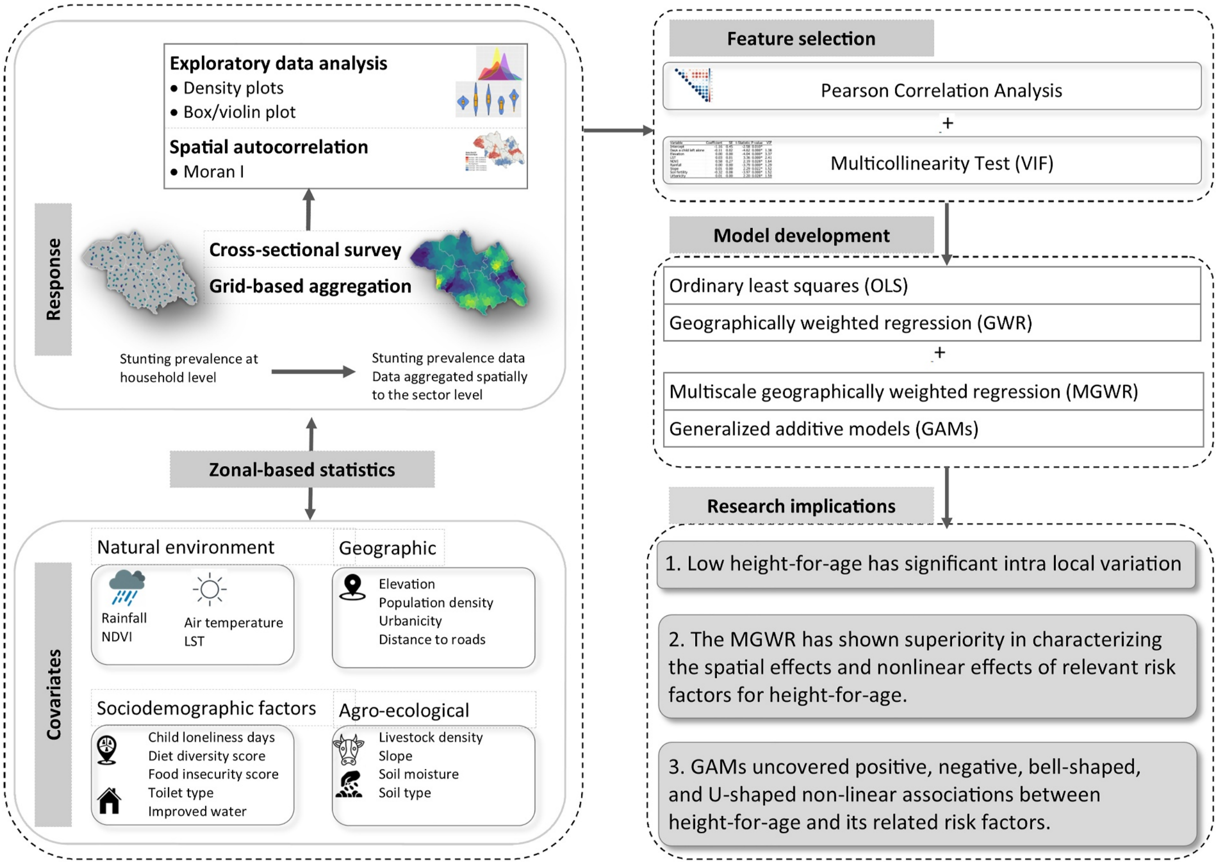

Figure 3. Methodological framework of this study.

child was left alone and slope, were smaller than the GWR bandwidth. This suggests that the influence of these variables on height-for-age is highly localized. On the other hand, the association between the number of days a child was left alone, slope, and height-for-age ratio is less localized. Based on the parameter estimates from the GWR model shown in Table 2, the variables such as the number of days a child was left alone, elevation, rainfall, and soil fertility exhibited a local negative association with height-for-age. Conversely, the findings from this

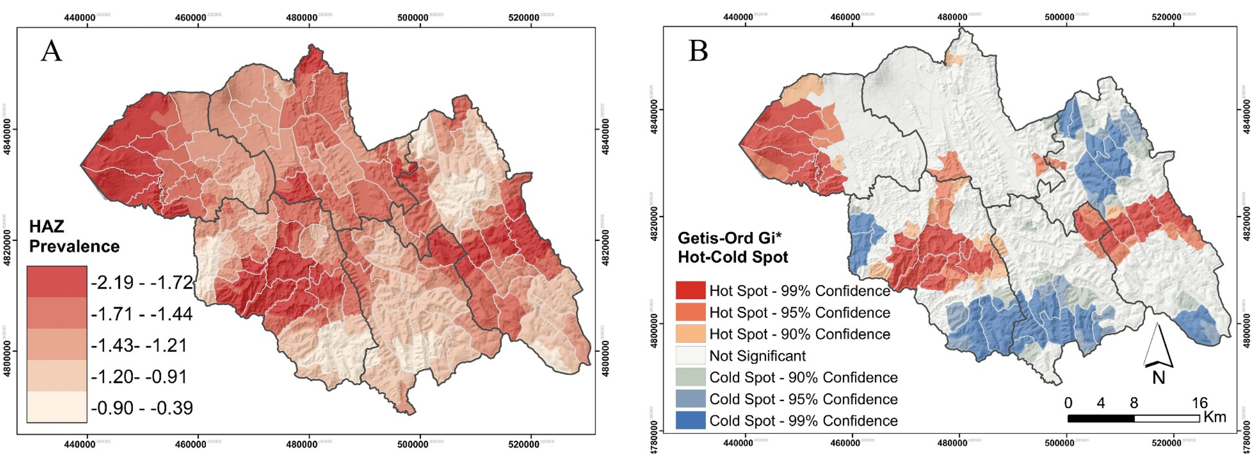

Figure 4. (a) Spatial distribution of height-for-age aggregated from household data at the sector level; (b) Geographical clusters of sectors from Getis-Ord Gi* statistics of height-for-age. The geographical clusters of sectors with significant-high (hot spot-statistically significant positive *z* scores, red color) or low (cold spot-statistically significant negative *z*-scores, dark blue color) values of the Getis-Ord Gi* statistics for the height-for-age.

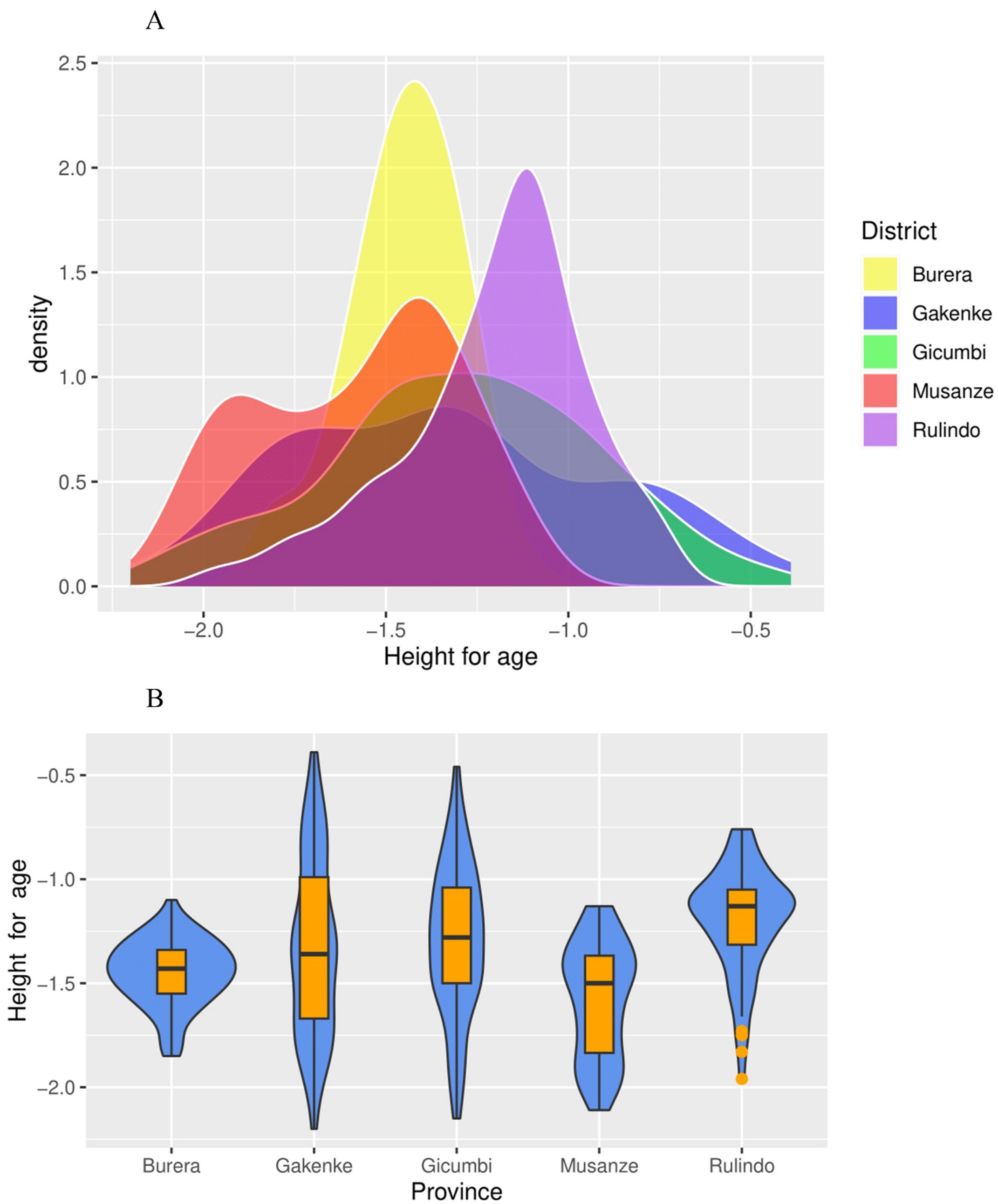

Figure 5. Density plot (a) and violin/box plots (b) of height-for-age prevalence distributed according to districts, geopolitical administrative zones in Rwanda. The horizontal line inside the box displays the cutoff height-for-age mean.

local model revealed a positive local relationship between height-for-age and LST, NDVI, slope, and urbanicity. However, the MGWR model yielded slightly different results. It indicated that the variables including the number of days a child was left alone, elevation, rainfall, slope, soil fertility, and urbanicity had a localized negative association with height-for-age, while NDVI and LST showed a local positive relationship with height-for-age. This difference could be attributed to the complex relationship between height-for-age and its risk factors within specific ranges, as illustrated in Figure 11. Furthermore, a *p*-value less than 0.05 was obtained from Monte Carlo tests conducted on the GWR and MGWR models, thus confirming the significance of the observed spatial variability of coefficients at a 95% confidence level.

Table 2
Summary Statistics for Geographically Weighted Regression Parameter Estimates

Variable	Mean	STD	Min	Median	Max	Bandwidth	Bandwidth confidence interval (95%)	Monte Carlo test for spatial variability
Intercept	0.02	0.666	−2.156	0.043	1.556	59	(59, 61)	0.000
Days left alone	−0.191	0.284	−0.998	−0.133	0.304	59	(59, 61)	0.000
Elevation	−0.363	0.587	−2.236	−0.277	1.462	59	(59, 61)	0.000
Rainfall	−0.082	0.854	−3.281	−0.036	2.17	59	(59, 61)	0.000
LST	0.048	0.605	−2.879	0.057	1.703	59	(59, 61)	0.000
NDVI	0.154	0.493	−0.997	0.14	2.325	59	(59, 61)	0.000
Slope	0.017	0.429	−1.48	0.013	1.13	59	(59, 61)	0.000
Soil fertility	−0.053	0.248	−1.656	−0.05	1.298	59	(59, 61)	0.000
Urbanicity	0.18	0.473	−1.179	0.115	1.673	59	(59, 61)	0.000

In comparison to OLS, both GWR and MGWR have achieved better fits with improved adjusted R^2. The local models explain 84% (GWR) and 89% (MGWR) of the variance in the height-for-age rate. Table 4 and Figure 6 demonstrate that the adjusted R^2 and RSS from GWR and MGWR are very similar, suggesting that the local regression models successfully describe the relationships between height-for-age and its risk factors. In terms of AIC, MGWR is more parsimonious than OLS and GWR. Moreover, MGWR achieves the lowest Residual Sum of Squares (RSS), followed by GWR and OLS. However, GWR exhibits lower performance in explaining height-for-age variability in some sectors (0.29) compared to MGWR (0.53). Therefore, the MGWR model should be used to test local variations in these relationships. Figure 6 displays the predicted height-for-age for both GWR and MGWR. The study findings reveal that the MGWR model-based estimates are more accurate in predicting height-for-age than the direct survey-based estimates shown in Figure 4. Concerning the spatial distribution of residuals in Figures 7 and 8, the OLS model produces a statistically significant ($p < 0.05$) spatially clustered pattern of residuals, while both GWR and MGWR produce a random distribution of residuals ($p > 0.05$). This suggests that GWR and MGWR effectively address the spatial autocorrelation or clustering of residuals in height-for-age. Consequently, GWR and MGWR have successfully accounted for spatial heterogeneity in most locations.

Figures 9 and 10, as well as Tables 2 and 3, display the intercepts of GWR and MGWR coefficient estimates, along with the predominance of coefficient estimates for each covariate. Comparing coefficient surfaces can enhance our understanding of spatial and scale variations. It can be observed that the covariate coefficient estimates for all GWR and MGWR models transition from negative (dark purple) to positive (yellow), suggesting the existence of both local negative and positive associations with height-for-age. This implies that the coefficient estimates for the majority of variables exhibit both positive and negative relationships with height-for-age. In relation to the MGWR model, the coefficient estimates indicate a negative relationship between the number of days a child was left alone and height-for-age in Gicumbi, the southern part of Burera, and the western part of

Table 3
Summary Statistics for Multiscale Geographically Weighted Regression Parameter Estimates

Variable	Mean	STD	Min	Median	Max	Bandwidth	Bandwidth confidence interval (95%)	Monte Carlo test for spatial variability
Intercept	0.231	0.349	−0.545	0.178	0.937	43	(43, 45)	0.000
Days left alone	−0.169	0.196	−0.636	−0.135	0.156	63	(55, 68)	0.030
Elevation	−0.263	0.312	−1.113	−0.172	0.285	43	(43, 47)	0.000
Rainfall	−0.02	0.497	−0.956	−0.076	1.525	45	(44, 50)	0.000
LST	0.145	0.204	−0.495	0.132	0.632	43	(43, 47)	0.000
NDVI	0.103	0.239	−0.505	0.12	0.727	43	(43, 50)	0.000
Slope	−0.059	0.197	−0.787	−0.016	0.235	78	(63, 96)	0.001
Soil fertility	−0.078	0.123	−0.509	−0.087	0.179	43	(43, 50)	0.004
Urbanicity	−0.004	0.271	−0.711	0.004	0.637	45	(44, 50)	0.013

Table 4
Comparison of the Models in Terms of Goodness-Of-Fit and Residual Dependency

Evaluation metric	AIC	Adj. R^2	RSS	Log-likelihood	Moran's I	z score	Residuals p-value	Residuals pattern
OLS	1065.672	0.25	304.478	−523.836	0.893	4.016	0.000	Clustered
GWR	496.552	0.848	44.552	−125.993	−0.125	−0.551	0.581	Random
MGWR	328.631	0.898	30.515	−47.66	0.033	0.162	0.871	Random

Note. RSS: residual sum of squares; AIC: corrected Akaike's information criterion; OLS: ordinary least squares; GWR: geographically weighted regression; MGWR: multiscale geographically weighted regression.

Gakenke. The median coefficient value is −0.169. However, in Musanze District, Rulindo, and the northern part of Burera, there is a positive relationship. Similarly, there is a negative relationship between elevation and height-for-age in all districts, with a median coefficient value of −0.263. In contrast, there is a positive association in the southern part of Gicumbi District, the eastern part of Rulindo, and the eastern part of Gakenke districts. On the other hand, the MGWR coefficient estimate for LST shows a positive relationship with height-for-age, with a median coefficient value of 0.145. This positive relationship is observed in Gicumbi (except its middle regions), Burera, Gakenke, Rulindo, and Musanze. The coefficient estimates for the NDVI variable in the MGWR model are positive, with a median coefficient value of 0.103, indicating a local positive variation with height-for-age in all districts. The findings suggest that rainfall has both negative and positive correlations with height-for-age, with a median coefficient value of −0.02. The spatial variation is significant, with an acceptable bandwidth of 45. The median coefficient of −0.059 demonstrates a clear pattern of primarily negative association between height-for-age and slope. However, there is a portion of positive relationship between height-for-age and slope with high variability in Musanze, Burera, the northern part of Gicumbi, and the southern part of Rulindo, but a negative association in Gakenke. The MGWR results in Table 3 and Figure 10 reveal that soil fertility has a moderate

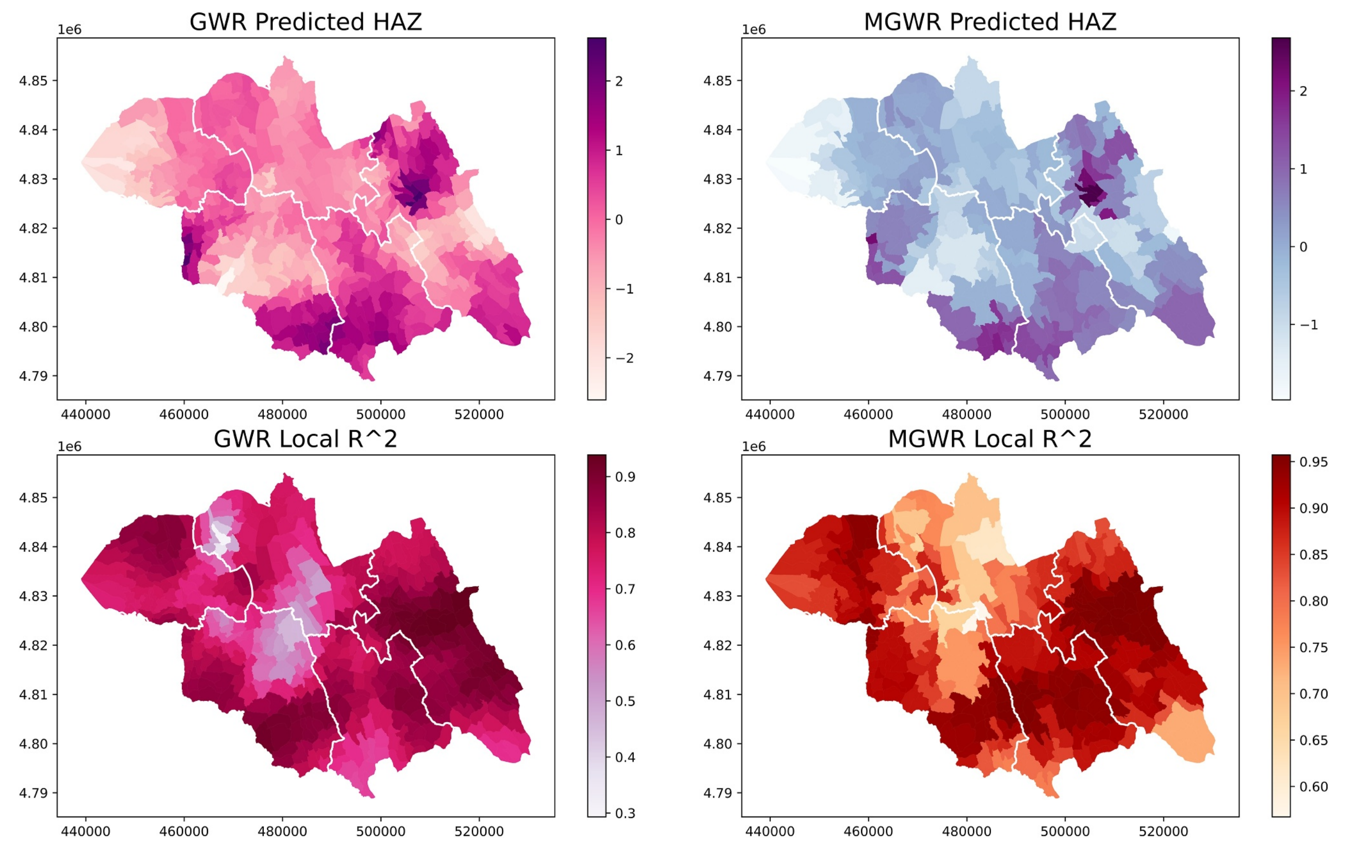

Figure 6. Predicted HAZ along with the distribution of the local coefficient of determination for both geographically weighted regression and multiscale geographically weighted regression models.

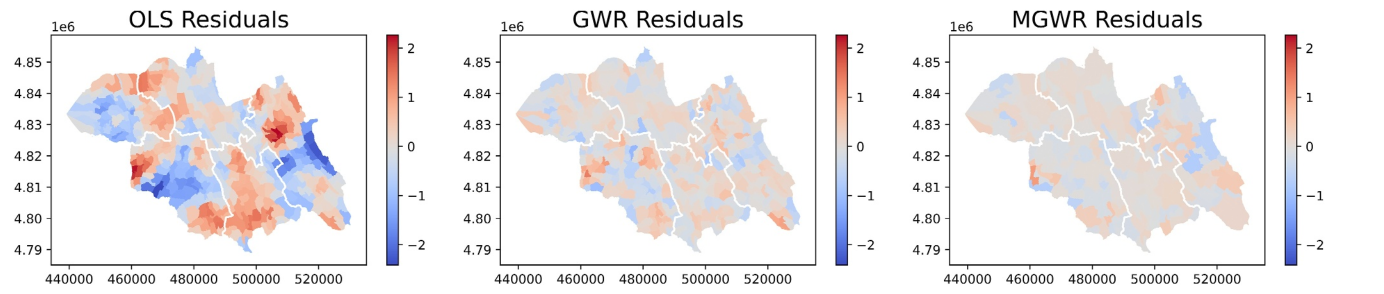

Figure 7. Spatial distribution of the standardized models' residuals; ordinary least squares, geographically weighted regression, and multiscale geographically weighted regression. A smaller standardized error indicates higher model performance.

relationship with height-for-age, with a median coefficient estimate of −0.078. The relationship is localized, with a bandwidth of 43. There is a negative relationship in Gakenke and Gicumbi, while the rest of the study area regions show a positive association. Similarly, urbanicity shows a weak negative correlation with height-for-age, with a negative median coefficient estimate of −0.004. This negative association is significant in Burera, Gakenke, and the southern part of Rulindo and Gicumbi districts. In contrast, the rest of the districts exhibit a positive association with height-for-age.

3.3. GAMs Non-Linear Effects

We conducted an analysis using GAM plots to assess the nonlinear relationship between predictor variables and height-for-age. The findings indicated that all predictors exhibit a curvilinear relationship with height-for-age. While a nonlinear effect between height-for-age and its risk factors exists, it is limited to specific intervals. The results for the non-linear effects of different factors on height-for-age, such as days a child was left alone, elevation, rainfall, LST, NDVI, slope, soil fertility, and urbanicity, are presented in Figure 11. Each figure displays the posterior means and 95% credible intervals. The results depicted in Figure 11a show a downward pattern for the days a child was left alone, suggesting that as the number of days a child was left alone increased, the height-for-age values decreased. This implies a higher likelihood of stunting. In the case of elevation (Figure 11b), the findings show that as elevation increases up to approximately 2,200 m, the likelihood of stunting increases. However, after reaching around 2,500 m, the height-for-age starts to decrease. Figure 11c demonstrates the non-linear effects of rainfall on height-for-age. The results indicate that the likelihood of stunting is highest as rainfall increases up to around 110 mm. Thereafter, the likelihood of stunting decreases up to around 130 mm, before gradually rising again. Height-for-age decreases when rainfall is either below 110 mm or above 130 mm. Regarding LST (Figure 11d), the study findings reveal a positive association between higher LST values and greater height-for-age. This implies that as LST increases, the likelihood of stunting decreases. The U-shaped functions displayed in Figure 11e indicate that as NDVI rises to 0.4, the chances of stunting increase. However, between 0.4 and 0.6, the chances of stunting decrease before rising again. Figure 11f illustrates the non-linear effects of slope on height-for-age, showing a somewhat bell-shaped curve. This suggests that as slope values

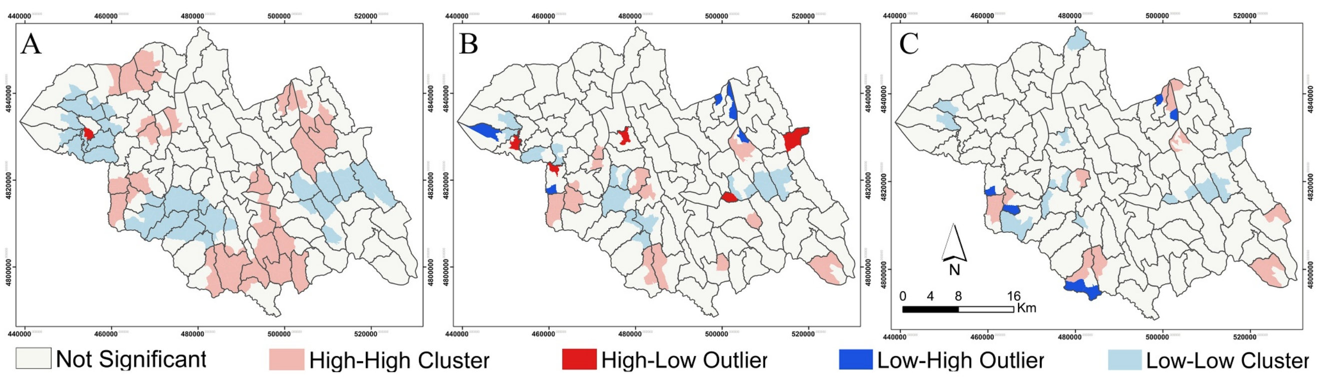

Figure 8. Local Moran's I of residuals; ordinary least squares in panel (a), geographically weighted regression in panel (b) and multiscale geographically weighted regression in panel (c).

10.1029/2024GH001027

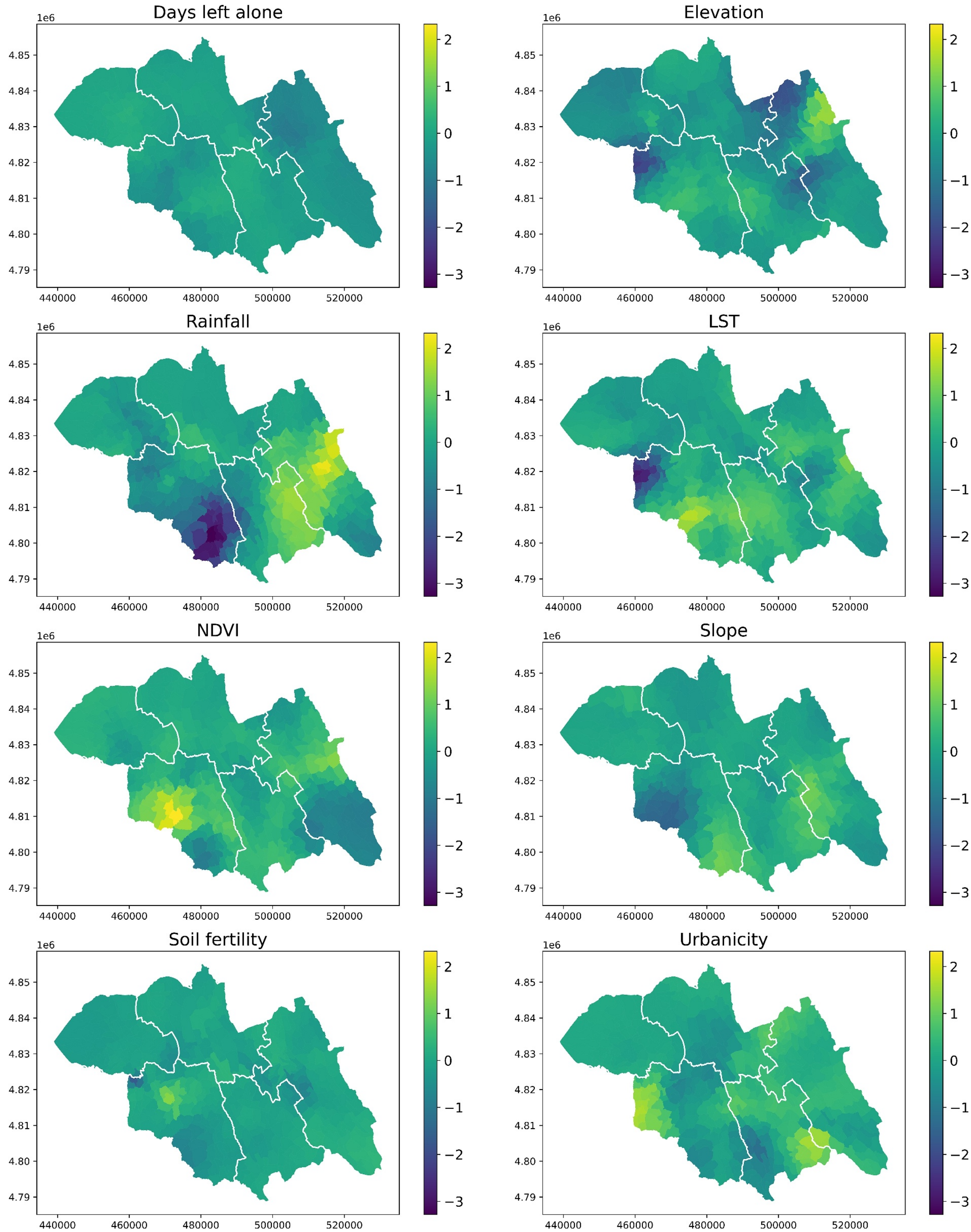

Figure 9. Spatial distribution of coefficients of the geographically weighted regression model. A positive sign denotes that the explanatory variable increases the probability of the outcome, whereas a negative sign indicates that the variable lowers the likelihood of the outcome.

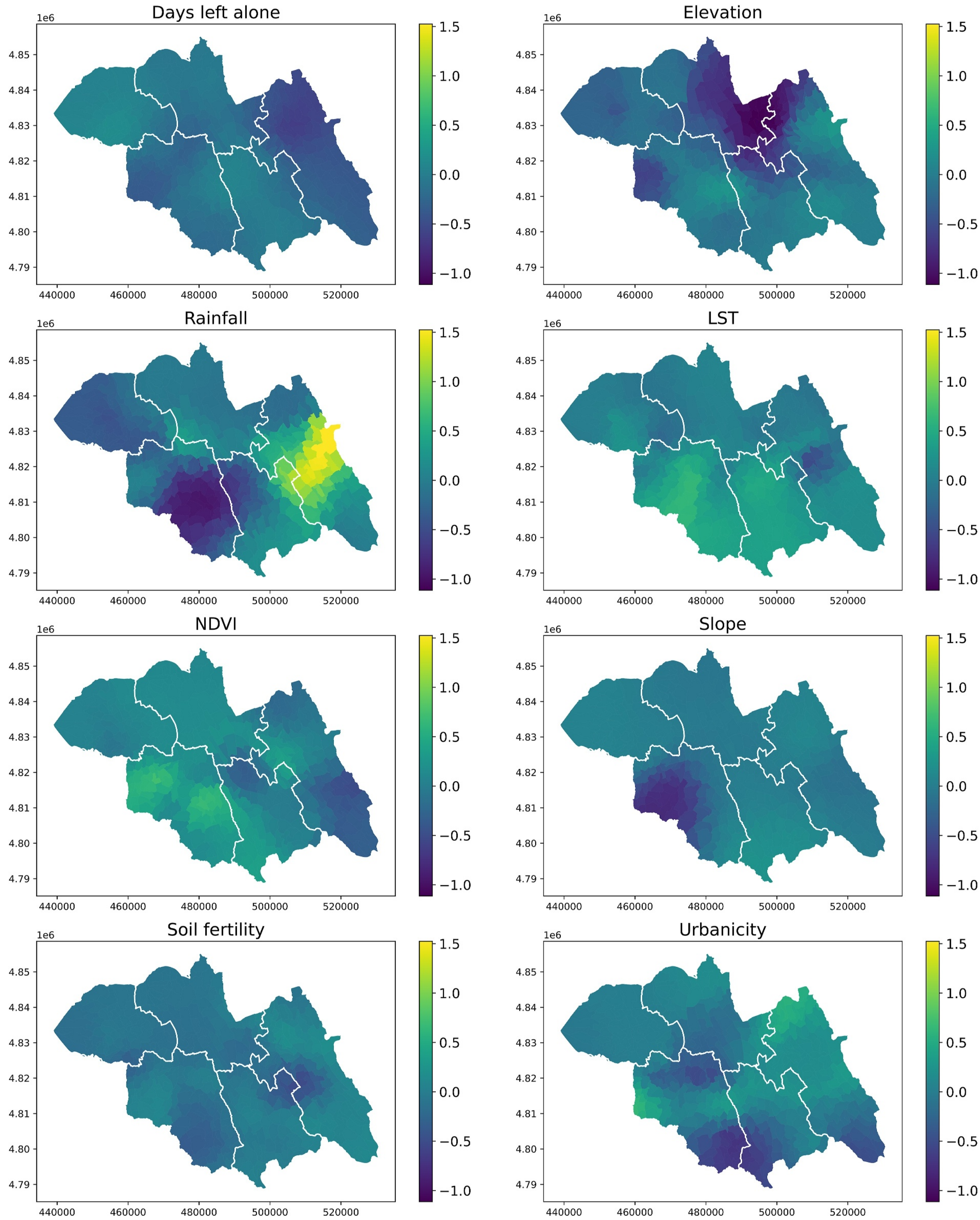

Figure 10. Spatial distribution of coefficients of the multiscale geographically weighted regression model. A positive sign denotes that the explanatory variable increases the probability of the outcome, whereas a negative sign indicates that the variable lowers the likelihood of the outcome.

increase up to 15°, height-for-age also increases. However, there is a downward pattern between 15 and 25°. In a similar fashion, as soil fertility increases up to 1.0, the likelihood of a child being stunted decreases. However, the likelihood starts to increase thereafter, reaching its highest point at around 1.2. The credible interval widens

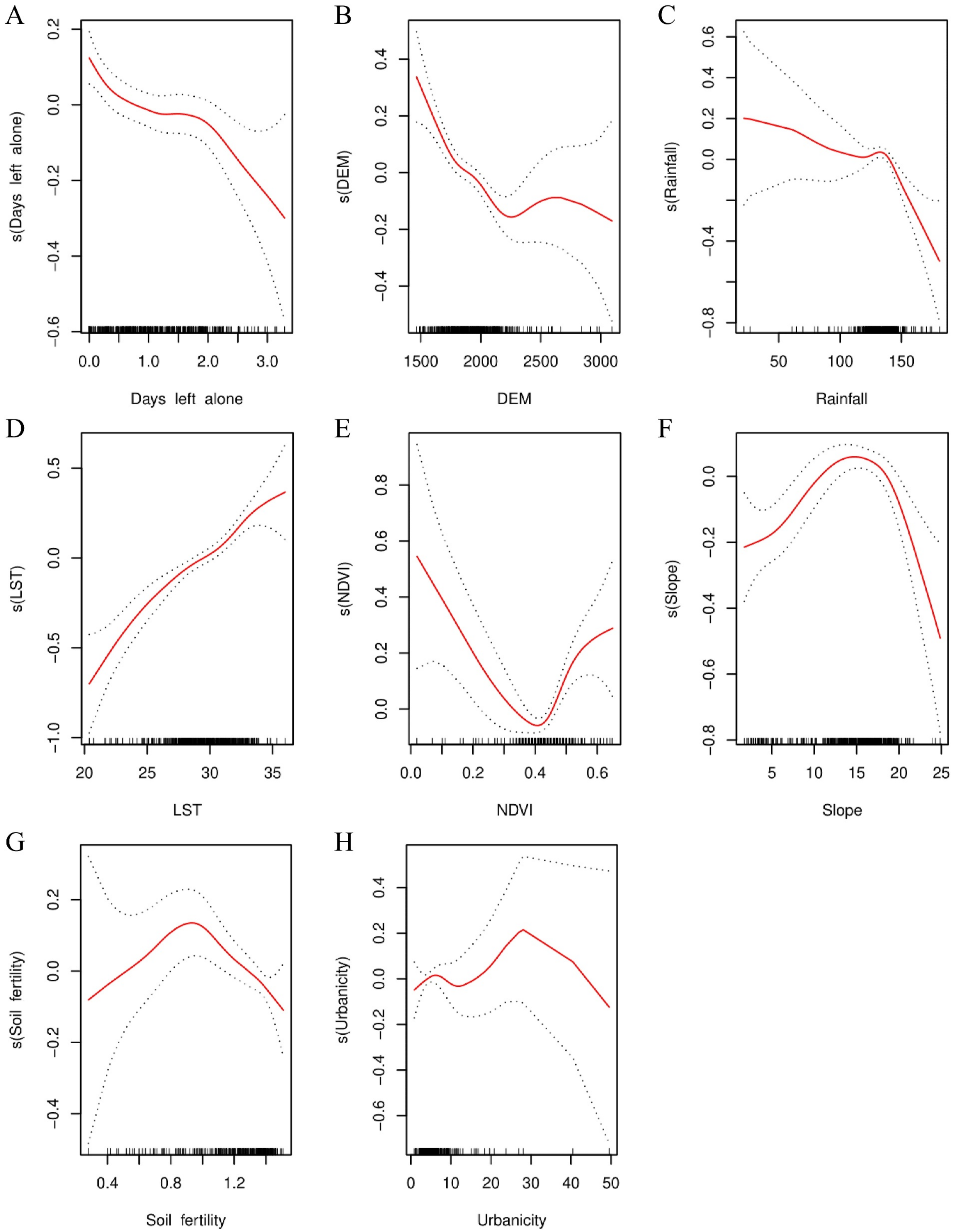

Figure 11. Generalized additive model plots of the nonlinear effects of relevant risk factors on height-for-age. The solid line is the estimated effect, with a 95% confidence limit as a dashed line. A declining curve signifies increasing stunting prevalence, and a rising curve signifies lowering stunting prevalence.

toward the end of the graph (Figure 11g). Lastly, the results for Figure 11h indicate that as average urbanicity values increase, the likelihood of stunting reduces until it reaches around 30. After that point, the likelihood of stunting starts to increase again.

4. Discussion

The aim of the current study was to analyze the spatial variability of height-for-age rates using fine-scale level data from a cros-sectional household survey, in conjunction with agroecological and climate data sets, in the Northern Province of Rwanda. The approach used combined spatial, linear, and non-linear effects to provide realistic estimates of the coefficients of various covariates. Local geographically weighted regressions were employed to assess spatial effects, while GAMs were utilized to characterize the non-linear impact of relevant height-for-age risk factors. The results indicated that the local MGWR model outperformed all other models in terms of prediction accuracy and generated residuals with a significant degree of randomness. The local MGWR model explained 89.8% of the total variance of height-for-age, compared to 84.8% and 25.0% for GWR and OLS, respectively. These findings suggest that, similar to previous spatial epidemiological studies (Chen et al., 2023; Lotfata & Tomal, 2022; Oshan et al., 2020), MGWR has the potential to capture local variability by identifying risk factors that influence the local variation of height-for-age. The analysis of spatial effects revealed significant intra-area variation in low height-for-age within the Northern Province region. The areas experiencing significant variations in stunting among children under the age of three are primarily located in the northern part of the study area, particularly in Musanze, Gakenke, and Gicumbi districts. Our findings align with those of Sekiyama et al. (2020), who suggested that the persistent stunting observed in the Northern Province may be mainly attributed to a reliance on starchy and plant-based protein foods, with limited consumption of nutrient-dense animal-source foods that provide high-quality proteins. Additionally, the increased prevalence of stunting among children under the age of three may also be linked to the higher rate of multiple pregnancies in the Northern Province of Rwanda (NISR, 2022), which could potentially lead to decreased parental caregiving for infants. Furthermore, the findings from the GAMs indicate that an increased number of days when a child is left alone is associated with a higher likelihood of nutritional deficiencies among children under the age of three. Among the climate variables examined, rainfall was found to be associated with a higher risk of stunting, while LST showed a positive association with height-for-age. Conversely, elevation had a detrimental impact on the nutritional status of children under three years of age, as evidenced by its negative association with height-for-age. However, variables such as NDVI, slope, soil fertility, and urbanicity exhibited both negative and positive relationships with height-for-age. These findings are largely consistent with existing literature (Balk et al., 2005; Lieber et al., 2022; Sununtnasuk, 2013; Tusting et al., 2020; Westerterp, 2001).

The variable, the number of days a child was left alone, which was not previously used in any undernourishment research, was found to be significant with low height-for-age. This association may be attributed to variations in traditional living habits, cultural differences, and agricultural practices within the study area. Numerous studies conducted in North America have shown that children who are left home alone often experience feelings of loneliness, worry, fear, and are susceptible to engaging in antisocial behaviors such as truancy, stealing, and drinking (Aizer, 2004; Ruiz-Casares et al., 2012; Ruiz-Casares & Rousseau, 2010). In their study, Doi et al. (2018) revealed that leaving children home alone was linked to higher total difficulty scores, particularly in the areas of conduct problems, hyperactivity, and difficulties in peer relationships. Similarly, Mertens et al. (2015), discovered that middle school students in the United States who were left alone for 3 hr or more exhibited higher levels of depression, behavior problems, low self-esteem, and reduced academic efficacy. Furthermore, Matsuyama et al. (2023) found that leaving children alone at home for less than an hour each week was a risk factor for dental caries among 6-7-year-old children. Additionally, Zhou et al. (2021) found that the absence of parents is associated with lower psychological resilience and related behavioral issues. On the other hand, allowing a mature and well-prepared child to stay home alone can be a positive experience, fostering their confidence and promoting independence and responsibility. Nevertheless, it is important to acknowledge that unsupervised children face real risks (Child Welfare Information Gateway, 2018). Numerous studies have established that from a very early age, infants seek interaction with the adults who care for them. If this interaction is lacking, it can hinder the child's brain development, potentially impacting their future educational experiences (Foley, 2017).

Consistent with findings from previous studies conducted in Uganda and Nepal (Dang et al., 2004), our research also found a higher prevalence of stunting among children aged 1–36 months living in hilly areas within the study location. The correlation between living in highland areas and childhood stunting in Rwanda has been previously documented by Uwiringiyimana, Veldkamp, and Amer (2019) and Weatherspoon et al. (2019). Dang et al. (2008), similarly found a negative relationship between height-for-age and households residing in higher-altitude regions. Uwiringiyimana, Veldkamp, and Amer (2019) further revealed that children living in lowlands are less likely to experience stunting compared to those raised in hilly areas. However, it is possible that this relationship is closely tied to the remoteness of these regions, which often face higher levels of dietary insecurity and limited access to healthcare services. Physiologically, populations in these areas lack the capacity to maintain energy balance at high altitudes due to the disparity between energy intake and expenditure. As a result, living at a high altitude slightly accelerates metabolism, leading to increased calorie expenditure (Westerterp, 2001). Our findings align with the research conducted by Egan (2013), who identified the optimal altitude range for human habitation as between 2,100 and 2,500 m. Additionally, our findings found a positive association between height-for-age and lower slope values (>15), followed by a subsequent negative relationship with a bell-shaped graph. Although there is no direct relationship between slope and stunting, previous studies have linked steep slope terrain with soil erosion, which can lead to reduced agricultural production (Sununtnasuk, 2013), increased risk of flooding and landslides (Siswanto & Sule, 2019), and limited access to essential health infrastructure (Lieber et al., 2022), all of which contribute to lower rates of height-for-age. Research conducted in rural India has provided evidence that children from households affected by flooding are more likely to experience stunting compared to those from non-flooded households (Gaire et al., 2016). Due to the challenges related to limited infrastructure and services, areas situated on steep slopes are characterized by inadequate healthcare facilities, poor sanitation, and restricted access to safe and clean water, all of which contribute to a higher stunting prevalence in comparison to areas on moderate slopes (Kismul et al., 2017).

Furthermore, when examining the NDVI and its impact on height-for-age rates, a negative relationship is observed between lower values of the NDVI (less than 0.4) and subsequent positive relationships. This influence is illustrated in the graph, which displays a distinct U-shaped pattern. Similar findings have been reported in the literature (Bangelesa et al., 2023; Galway et al., 2018; Lopez-Carr et al., 2016; Sununtnasuk, 2013). For example, Lopez-Carr et al. (2016) showed that crop production may decline in areas experiencing a negative change in vegetation index, despite potentially favorable climate conditions, resulting in infant stunting. Limited agricultural productivity is associated with low vegetation index cover, which in turn leads to an increase in childhood stunting (Balk et al., 2005). Particularly in rural areas where forests are in close proximity to most people, forest ecosystems serve as the primary source of local food (Richardson, 2010). This finding aligns with Bangelesa et al. (2023), who found that children living in areas with higher leaf area index (LAI) values have a lower risk of stunting compared to those in areas with lower LAI values. Additionally, Sununtnasuk (2013) revealed that the NDVI significantly decreases the likelihood of stunting in Nepal. Our findings have revealed a positive association between average rainfall and HAZ, but a negative association between HAZ and both low and high rainfall. These findings are consistent with previous studies conducted by Lieber et al. (2022), Mukabutera, Jamie, et al. (2016), Sununtnasuk (2013), and Yeboah et al. (2022). In their study, Boah et al. (2022) found that younger children who are more vulnerable are more likely to be negatively affected by increased rainfall, while older children tend to cope better with such conditions. Lieber et al. (2022) also found a significant association between rainfall as a proxy for climate change and child health and nutrition. Specifically, they found that children, especially those under the age of two in Burkina Faso, are more likely to have stunted growth when exposed to above-average rainfall. Moreover, extreme rainfall and vegetation density in Rwanda have been linked to a higher prevalence of vector-borne diseases like malaria, which contributes to the mortality of under five of the age children due to the favorable conditions created for disease vectors (Nduwayezu et al., 2023). Previous research conducted in Rwanda has also shown that heavy rainfall can contaminate surface water by carrying waste and sediment into drinking water sources (Mukabutera, Thomson, Murray, et al., 2016). The Northern Province, with its steep-sloped landscape, is particularly susceptible to floods and landslides, which result in fatalities during the rainy season (Bizimana & Nduwayezu, 2021). However, it is important to acknowledge the significant role that average rainfall plays in agricultural productivity (Sununtnasuk, 2013).

In line with previous studies, the results of this study also show a positive correlation between LST and height-for-age (Sarker et al., 2012; Serrat, 2014; Tusting et al., 2020). A study conducted on children living in hotter regions of sub-Saharan Africa found that warm temperatures are associated with stunting (Tusting et al., 2020). Similarly, a study carried out in Bangladesh revealed that higher temperatures have a positive influence on rice growth, potentially leading to increased food productivity in that specific area (Sarker et al., 2012). These findings are further supported by various potential biological explanations (Mora et al., 2022; Patz et al., 2005; Serrat, 2014; Tusting et al., 2020). Research has shown that individuals raised in warmer climates have permanently longer limbs and bones compared to their siblings raised in cooler environments (Serrat, 2014). Tusting et al. (2020) also suggest that genetic factors may contribute to the decline in stunting associated with higher temperatures. They further propose the possibility of unknown epigenetic factors interacting with environmental temperature to impact height in African populations. However, this hypothesis still requires confirmation through additional research. In areas prone to drought, where poor agricultural production exacerbates hunger and extreme poverty, the detrimental effects of consistently warm climates cannot be underestimated. For instance, extreme temperatures influence the lifecycle of malaria mosquitoes, increasing the transmission of the disease (Mora et al., 2022). Frequent illness from malaria hampers child well-being and nutritional status, which can consequently impair child growth and development (Patz et al., 2005).

This study also uncovered a positive correlation between height-for-age and soil fertility (<1.0), followed by a negative correlation. These findings align with previous studies (De Sherbinin, 2011; Weatherspoon et al., 2019). Weatherspoon et al. (2019) found significantly lower rates of stunting in households located in areas with more fertile soils. Although this study established a positive relationship between areas with soil fertility less than 1.0 and height-for-age, it is perplexing that soil fertility greater than 1.0 has a detrimental impact on height-for-age. De Sherbinin (2011) found that mountainous regions in Africa possess temperate climates and potentially volcanic soils, resulting in high agricultural productivity. Conversely, other areas exhibit lower productivity, particularly in steep slope areas, leading to persistent soil erodibility (Weatherspoon et al., 2019). These findings may be supported by the steep terrain of the northern region, which experiences fragile and damaged soils due to the prevalence of small-scale farming (NISR, 2018). However, further research is necessary to investigate the mediating effect of this association on childhood stunting.

This study has further revealed a positive association between height-for-age and areas with a moderate urbanicity rate (less than 30), and a negative association onwards. Our findings indicate higher low height-for-age rates in rural areas compared to urban areas, supporting existing literature (Balk et al., 2005; Jones et al., 2016; Van de Poel et al., 2007). Urban areas are associated with better living conditions, such as improved healthcare access, which lowers the likelihood of infectious disease contraction compared to rural areas (Balk et al., 2005). The higher population density in urban areas facilitates the sharing of health-related information and resources (Jones et al., 2016). Conversely, rural areas in most developing countries are characterized by higher poverty rates, poor healthcare systems, and a lack of household hygiene (Van de Poel et al., 2007). The negative relationship between height-for-age and urbanicity rate may be explained by the prevalent informal urban morphology in many developing cities (Nduwayezu et al., 2021). Previous studies conducted in Rwanda have shown a high prevalence of stunting in informal urban settlements (NISR, 2018; NISR et al., 2021). It is understandable that informal settlements in developing cities have limited access to healthcare services, poor sanitation, and inadequate hygiene, resulting in a higher risk of diseases like malaria, pneumonia, and diarrhea (Balk et al., 2005). However, further investigations are necessary to empirically verify these findings.

The strength of this study lies in its state-of-the-art machine learning framework, which simultaneously considers various predictors to facilitate the analysis of the linear, nonlinear, and spatially heterogeneous relationships between height-for-age and its relevant risk factors. This study is the first to examine the geographic variation in height-for-age at a fine-scale level in response to socioeconomic, agroecological, and climate determinants, using locally weighted regression combined with interpretable machine learning. It is essential to note that this study focused on identifying association rather than proving the causality, which is important when interpreting the findings of this cross-sectional study (Cofield et al., 2010). In essence, cross-sectional studies limited in their ability to ascertain causality (Tellings, 2017), are carried out during a short interval at a specific time (Levin, 2006; Mann, 2003), which help in estimating the prevalence of the outcome and identifying its key risk factors and potential areas for intervention and formulating causal hypotheses for further in-depth analysis (Molnar & Freiesleben, 2024). However, longitudinal studies, which track changes over time and identify causal associations (Levin, 2005) could be further designed to investigate the effects of spatial and seasonal risk factors on height-for-age variability, and to compare these associations more closely at the household level (Setia, 2016). This can help to better elucidate the driving factors of the variability in low height-for-age distribution in the Northern province of Rwanda. Furthermore, future studies should consider additional factors such as household violence, maternal health conditions, and socio-cultural practices, which play a significant role in influencing childhood stunting, in addition to the socio-economic, agro-ecological, and climate-related factors considered in the current study.

5. Conclusion

This study aimed to analyze the impact of socioeconomic, agroecological, and climate factors on the spatial variations of height-for-age using locally geographically weighted regressions with the GAMs algorithm as a spatial analysis approach in the Northern Province of Rwanda. Identifying areas with the highest rates of stunting will help determine the most effective measures for reducing the burden of undernutrition. The maps generated for the height-for-age risk variables have identified specific areas with high and low likelihoods of stunting prevalence, which may assist policymakers in understanding the undernourishment status and needs of each area and enable tailored policy measures. Additionally, by comparing these maps to socioeconomic and epidemiological indices, further explanations for the observed spatial patterns could potentially be uncovered. For example, based on the maps, areas with high stunting prevalence may be associated with their respective levels of agricultural production, educational attainment, or wealth levels. These significant findings about undernourishment in the study area would not have been known without refined spatial data on infant nutritional status. To our knowledge, there has been a lack of small-area spatial modeling of height-for-age in Rwanda. Therefore, the results from this study will help revisit ongoing strategies and nutrition initiatives, particularly in areas where the burden of stunting was shown to be substantially higher. A finer-scale assessment helps accurately identify the effects of relevant variables on height-for-age, paving the way for more in-depth analysis and the formulation of effective nutrition initiatives in Rwanda to precisely address childhood undernourishment.

Global Research Collaboration Statement

We thank all the mothers and children who participated in the study for their valuable time and shared life experiences with us. We also extend our gratitude to all community health workers and village leaders who worked with us to identify the eligible households that constituted our village sampling frame. We also thank the experienced interviewers, including local registered nurses, who helped us conduct the survey under the supervision of Rwandan and Swedish supervisors. The research protocol and all study tools were checked and approved for scientific and ethical integrity by the Institutional Review Board (IRB) in the College of Medicine and Health Sciences of the University of Rwanda. We also obtained authorization to undertake our fieldwork from the Ministry of Local Governance (reference No. 0779/07.01) and the National Health Research Committee (no. NHRC/2020/PROT/047).

Conflict of Interest

The authors declare no conflicts of interest relevant to this study.

Data Availability Statement

The covariates used can be accessed freely through the links provided in Table 1. The modeled surface covariates can be obtained from the DHS Program for a credible research project via an application process available at https://spatialdata.dhsprogram.com. The nutritional height-for-age data sets used in the current study are geo-located data that contain some private health information about the survey respondents, such as HIV, etc. However, the data can be requested from the project coordinator at https://www.gu.se/om-universitetet/hitta-person/gunillakrantz2. The used MGWR algorithm in the Python version, along with all the codes, is available for free (Fotheringham et al., 2023).

Acknowledgments
We convey our sincere gratitude to the Swedish International Development Agency (SIDA) under the undernutrition sub-program, with the collaboration of Lund University and the University of Rwanda, for their support. The SIDA-Rwanda Bilateral Programme Contribution No. 11277 provided funding for this study.

References

Aizer, A. (2004). Home alone: Supervision after school and child behavior. *Journal of Public Economics*, *88*(9–10), 1835–1848. https://doi.org/10.1016/S0047-2727(03)00022-7

Amondo, E. I., Nshakira-Rukundo, E., & Mirzabaev, A. (2023). The effect of extreme weather events on child nutrition and health. *Food Security*, *15*(3), 571–596. https://doi.org/10.1007/s12571-023-01354-8

Anselin, L. (1995). Local Indicators of Spatial Association—LISA. *Geographical Analysis*, *27*(2), 93–115. https://doi.org/10.1111/J.1538-4632.1995.TB00338.X

Aoun, N., Matsuda, H., & Sekiyama, M. (2015). Geographical accessibility to healthcare and malnutrition in Rwanda. *Social Science & Medicine*, *130*, 135–145. https://doi.org/10.1016/J.SOCSCIMED.2015.02.004

Baker, D. M., & Valleron, A. J. (2014). An open source software for fast grid-based data-mining in spatial epidemiology (FGBASE). *International Journal of Health Geographics*, *13*(1), 1–8. https://doi.org/10.1186/1476-072x-13-46

Baker, R. E., & Anttila-Hughes, J. (2020). Characterizing the contribution of high temperatures to child undernourishment in Sub-Saharan Africa. *Scientific Reports*, *10*(1), 18796. https://doi.org/10.1038/S41598-020-74942-9

Balk, D., Storeygard, A., Levy, M., Gaskell, J., Sharma, M., & Flor, R. (2005). Child hunger in the developing world: An analysis of environmental and social correlates. *Food Policy*, *30*(5–6), 584–611. https://doi.org/10.1016/J.FOODPOL.2005.10.007

Bangelesa, F., Hatloy, A., Mbunga, B. K., Mutombo, P. B., Matina, M. K., Akilimali, P. Z., et al. (2023). Is stunting in children under five associated with the state of vegetation in the Democratic Republic of the Congo? Secondary analysis of Demographic Health Survey data and the satellite-derived leaf area index. *Heliyon*, *9*(2), e13453. https://doi.org/10.1016/J.HELIYON.2023.E13453

Bizimana, J. P., & Nduwayezu, G. (2021). Spatio-temporal patterns of malaria incidence in Rwanda. *Transactions in GIS*, *25*(2), 751–767. https://doi.org/10.1111/TGIS.12711

Blom, S., Ortiz-Bobea, A., & Hoddinott, J. (2022). Heat exposure and child nutrition: Evidence from West Africa. *Journal of Environmental Economics and Management*, *115*, 102698. https://doi.org/10.1016/J.JEEM.2022.102698

Boah, M., Issah, A. N., Yeboah, D., Kpordoxah, M. R., & Sira, J. (2022). Association between compliance with the new WHO-recommended frequency and timing of antenatal care contacts and receiving quality antenatal care in Cameroon. *SAGE Open*, *12*(3), 21582440221117807. https://doi.org/10.1177/21582440221117807

Brown, M. E., Grace, K., Shively, G., Johnson, K. B., & Carroll, M. (2014). Using satellite remote sensing and household survey data to assess human health and nutrition response to environmental change. *Population and Environment*, *36*(1), 48–72. https://doi.org/10.1007/s11111-013-0201-0

Brunsdon, C., Fotheringham, A. S., & Charlton, M. E. (1996). Geographically weighted regression: A method for exploring spatial nonstationarity. *Geographical Analysis*, *28*(4), 281–298. https://doi.org/10.1111/J.1538-4632.1996.TB00936.X

Burgert-Brucker, C. R., Dontamsetti, T., & Gething, P. W. (2018). The DHS program's modeled surfaces spatial datasets. *Studies in Family Planning*, *49*(1), 87–92. https://doi.org/10.1111/SIFP.12050

Chen, H., Cao, Y., Feng, L., Zhao, Q., & Torres, J. R. V. (2023). Understanding the spatial heterogeneity of COVID-19 vaccination uptake in England. *BMC Public Health*, *23*(1), 895. https://doi.org/10.1186/s12889-023-15801-w

Cheng, J. (2003). *Modeling spatial and temporal of urban growth*. Utrecht University.

Child Welfare Information Gateway. (2018). Leaving your child home alone.

Cofield, S. S., Corona, R. V., & Allison, D. B. (2010). Use of causal language in observational studies of obesity and nutrition. *Obesity Facts*, *3*(6), 353–356. https://doi.org/10.1159/000322940

Dang, S., Yan, H., & Yamamoto, S. (2008). High altitude and early childhood growth retardation: New evidence from Tibet. *European Journal of Clinical Nutrition*, *62*(3), 342–348. https://doi.org/10.1038/SJ.EJCN.1602711

Dang, S., Yan, H., Yamamoto, S., Wang, X., & Zeng, L. (2004). Poor nutritional status of younger Tibetan children living at high altitudes. *European Journal of Clinical Nutrition*, *58*(6), 938–946. https://doi.org/10.1038/SJ.EJCN.1601915

Dasgupta, S., & Robinson, E. J. Z. (2023). Climate, weather and child health in Burkina Faso. *The Australian Journal of Agricultural and Resource Economics*, *67*(4), 576–602. https://doi.org/10.1111/1467-8489.12530

De Sherbinin, A. (2011). The biophysical and geographical correlates of child malnutrition in Africa. *Population, Space and Place*, *17*(1), 27–46. https://doi.org/10.1002/PSP.599

Doi, S., Fujiwara, T., Isumi, A., Ochi, M., & Kato, T. (2018). Relationship between leaving children at home alone and their mental health: Results from the A-CHILD study in Japan. *Frontiers in Psychiatry*, *9*, 192. https://doi.org/10.3389/FPSYT.2018.00192

Dong, Q., Li, N., Zhang, Y., & Sun, Z. (2021). RZWQM simulation of soil moisture and maize yield in newly cultivated land with different soil bulk density. *Communications in Soil Science and Plant Analysis*, *52*(8), 832–842. https://doi.org/10.1080/00103624.2020.1869767

Dror, D. K., & Allen, L. H. (2011). The importance of milk and other animal-source foods for children in low-income countries. *Food and Nutrition Bulletin*, *32*(3), 227–243. https://doi.org/10.1177/156482651103200307

Egan, E. (2013). Notes from higher grounds: An altitude training guide for endurance athletes.

Fenta, H. M., Zewotir, T., & Muluneh, E. K. (2021). Spatial data analysis of malnutrition among children under-five years in Ethiopia. *BMC Medical Research Methodology*, *21*(1), 232. https://doi.org/10.1186/S12874-021-01391-X

Fink, G., Günther, I., & Hill, K. (2011). The effect of water and sanitation on child health: Evidence from the demographic and health surveys 1986–2007. *International Journal of Epidemiology*, *40*(5), 1196–1204. https://doi.org/10.1093/IJE/DYR102

Foley, K. V. (2017). The care effect: Why no child should grow up alone.

Fotheringham, A. S., Oshan, T. M., & Li, Z. (2023). Multiscale geographically weighted regression: Theory and practice. *Multiscale Geographically Weighted Regression: Theory and Practice*, 1–176. https://doi.org/10.1201/9781003435464

Fotheringham, A. S., Yang, W., & Kang, W. (2017). Multiscale Geographically Weighted Regression (MGWR). *Annals of the American Association of Geographers*, *107*(6), 1247–1265. https://doi.org/10.1080/24694452.2017.1352480

Gage, K. L., Burkot, T. R., Eisen, R. J., & Hayes, E. B. (2008). Climate and vectorborne diseases. *American Journal of Preventive Medicine*, *35*(5), 436–450. https://doi.org/10.1016/J.AMEPRE.2008.08.030

Gaire, S., Delbiso, T. D., Pandey, S., & Guha-Sapir, D. (2016). Impact of disasters on child stunting in Nepal. *Risk Management and Healthcare Policy*, *9*, 113. https://doi.org/10.2147/RMHP.S101124

Galway, L. P., Acharya, Y., & Jones, A. D. (2018). Deforestation and child diet diversity: A geospatial analysis of 15 Sub-Saharan African countries. *Health & Place*, *51*, 78–88. https://doi.org/10.1016/J.HEALTHPLACE.2018.03.002

Grace, K., Davenport, F., Funk, C., & Lerner, A. M. (2012). Child malnutrition and climate in Sub-Saharan Africa: An analysis of recent trends in Kenya. *Applied Geography*, *35*(1–2), 405–413. https://doi.org/10.1016/J.APGEOG.2012.06.017

Habyarimana, F., Zewotir, T., Ramroop, S., & Ayele, D. G. (2017). Spatial distribution of determinants of malnutrition of children under five years in Rwanda: Simultaneous measurement of three anthropometric indices. *Kamla Raj Enterprises*, *54*(3), 138–149. https://doi.org/10.1080/09709274.2016.11906996

Hagos, S., Lunde, T., Mariam, D. H., Woldehanna, T., & Lindtjørn, B. (2014). Climate change, crop production and child under nutrition in Ethiopia; A longitudinal panel study. *BMC Public Health*, *14*(1), 884. https://doi.org/10.1186/1471-2458-14-884

Hassler, J., Andersson Granberg, T., Steins, K., & Ceccato, V. (2024). Towards more realistic measures of accessibility to emergency departments in Sweden. *International Journal of Health Geographics*, *23*(1), 1–14. https://doi.org/10.1186/s12942-024-00364-9

Hastie, T., & Tibshirani, R. (1995). Generalized additive models for medical research. *Statistical Methods in Medical Research*, *4*(3), 187–196. https://doi.org/10.1177/096228029500400302

Hastie, T., Tibshirani, R., & Friedman, J. (2017). *The elements of statistical learning: Data mining, inference, and prediction* (2nd ed.). Springer. https://doi.org/10.1007/978-0-387-84858-7

Hatfield, J. L., & Prueger, J. H. (2015). Temperature extremes: Effect on plant growth and development. *Weather and Climate Extremes*, *10*, 4–10. https://doi.org/10.1016/J.WACE.2015.08.001

Hummel, M., Hallahan, B. F., Brychkova, G., Ramirez-Villegas, J., Guwela, V., Chataika, B., et al. (2018). Reduction in nutritional quality and growing area suitability of common bean under climate change induced drought stress in Africa. *Scientific Reports*, *8*(1), 1–11. https://doi.org/10.1038/s41598-018-33952-4

Hussien, F. M., Mebratu, W., Ahmed, A. Y., Mekonnen, T. C., Hassen, A. M., Asfaw, Z. A., et al. (2021). Performance of individual dietary diversity score to identify malnutrition among patients living with HIV in Ethiopia. *Scientific Reports*, *11*(1), 18681. https://doi.org/10.1038/s41598-021-98202-6

Johnson, K., & Brown, M. E. (2014). Environmental risk factors and child nutritional status and survival in a context of climate variability and change. *Applied Geography*, *54*, 209–221. https://doi.org/10.1016/J.APGEOG.2014.08.007

Jones, A. D., Acharya, Y., & Galway, L. P. (2016). Urbanicity gradients are associated with the household- and individual-level double burden of malnutrition in sub-Saharan Africa. *The Journal of Nutrition*, *146*(6), 1257–1267. https://doi.org/10.3945/JN.115.226654

Karra, M., Fink, G., & Canning, D. (2017). Facility distance and child mortality: A multi-country study of health facility access, service utilization, and child health outcomes. *International Journal of Epidemiology*, *46*(3), 817–826. https://doi.org/10.1093/IJE/DYW062

Kish, L. (1965). *Survey sampling*. Wiley. https://doi.org/10.2307/2091346

Kismul, H., Acharya, P., Mapatano, M. A., & Hatløy, A. (2017). Determinants of childhood stunting in the Democratic Republic of Congo: Further analysis of Demographic and Health Survey 2013–14. *BMC Public Health*, *18*(1), 1–14. https://doi.org/10.1186/s12889-017-4621-0

Levin, K. A. (2005). Study design I. *Evidence-Based Dentistry*, *6*(3), 78–79. https://doi.org/10.1038/sj.ebd.6400355

Levin, K. A. (2006). Study design III: Cross-sectional studies. *Evidence-Based Dentistry*, *7*(1), 24–25. https://doi.org/10.1038/sj.ebd.6400375

Li, Z., Fotheringham, A. S., Goodchild, M. F., & Li, W. (2020). Multiscale geographically weighted regression computation, inference, and application.

Lieber, M., Chin-Hong, P., Kelly, K., Dandu, M., & Weiser, S. D. (2022). A systematic review and meta-analysis assessing the impact of droughts, flooding, and climate variability on malnutrition. *Global Public Health*, *17*(1), 68–82. https://doi.org/10.1080/17441692.2020.1860247

Lobell, D. B., & Field, C. B. (2007). Global scale climate–crop yield relationships and the impacts of recent warming. *Environmental Research Letters*, *2*(1), 014002. https://doi.org/10.1088/1748-9326/2/1/014002

Lopez-Carr, D., Mwenda, K. M., Pricope, N. G., Kyriakidis, P. C., Jankowska, M. M., Weeks, J., et al. (2016). Climate-related child undernutrition in the Lake Victoria basin: An integrated spatial analysis of health surveys, NDVI, and precipitation data. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *9*(6), 2830–2835. https://doi.org/10.1109/JSTARS.2016.2569411

Lotfata, A., & Tomal, M. (2022). Exploring housing determinants of obesity prevalence using multiscale geographically weighted regression in Chicago, Illinois. *The Professional Geographer*, *75*(3), 335–344. https://doi.org/10.1080/00330124.2022.2111692

Mann, C. J. (2003). Observational research methods. Research design II: Cohort, cross sectional, and case-control studies. *Emergency Medicine Journal*, *20*(1), 54–60. https://doi.org/10.1136/EMJ.20.1.54

Matsuyama, Y., Isumi, A., Doi, S., & Fujiwara, T. (2023). Being left alone at home and dental caries of children aged 6–7 years. *Journal of Epidemiology*, *33*(6), 303–310. https://doi.org/10.2188/JEA.JE20210321

Mertens, S. B., Flowers, N., & Mulhall, P. F. (2015). Research on middle school renewal: Should middle grades students be left alone after school? *Middle School Journal*, *34*(5), 57–61. https://doi.org/10.1080/00940771.2003.11494517

MINAGRI. (2006). The one cow per poor family programme (Girinka) in Rwanda.

MINISANTE. (2014). Rwanda National Food and Nutrition Policy. Retrieved from https://www.moh.gov.rw/news-detail/national-food-and-nutrition-policy

Molnar, C. (2022). Interpretable machine learning: A guide for making black box models explainable.

Molnar, C., & Freiesleben, T. (2024). Supervised Machine Learning for Science: How to stop worrying and love your black box—Causality. Retrieved from https://ml-science-book.com/

Mora, C., McKenzie, T., Gaw, I. M., Dean, J. M., von Hammerstein, H., Knudson, T. A., et al. (2022). Over half of known human pathogenic diseases can be aggravated by climate change. *Nature Climate Change*, *12*(9), 869–875. https://doi.org/10.1038/s41558-022-01426-1

Moraga, P. (2019). *Geospatial health data, modeling and visualization with R-INLA and Shiny*. Chapman and Hall/CRC. Retrieved from https://www.routledge.com/Geospatial-Health-Data-Modeling-and-Visualization-with-R-INLA-and-Shiny/Moraga/p/book/9780367357955

Mukabutera, A., Jamie, I. F., Laetitia, N., Habimana, M., & Paulin, B. (2016). Associations of rainfall with childhood under-nutrition in Rwanda: An ecological study using the data from Rwanda meteorology agency and the 2010 demographic and health survey. *Asian Journal of Agriculture and Food Sciences*, *4*(01), 35–45.

Mukabutera, A., Thomson, D., Murray, M., Basinga, P., Nyirazinyoye, L., Atwood, S., et al. (2016). Rainfall variation and child health: Effect of rainfall on diarrhea among under 5 children in Rwanda, 2010. *BMC Public Health*, *16*(1), 1–9. https://doi.org/10.1186/s12889-016-3435-9

Mukabutera, A., Thomson, D. R., Hedt-Gauthier, B. L., Basinga, P., Nyirazinyoye, L., & Murray, M. (2016). Risk factors associated with underweight status in children under five: An analysis of the 2010 Rwanda Demographic Health Survey (RDHS). *BMC Nutrition*, *2*(1), 1–12. https://doi.org/10.1186/s40795-016-0078-2

Murphy, S. P., & Allen, L. H. (2003). Nutritional importance of animal source foods. *The Journal of Nutrition*, *133*(11), 3932S–3935S. https://doi.org/10.1093/JN/133.11.3932S

Nahalomo, A., Iversen, P. O., Andreassen, B. A., Kaaya, A. N., Rukooko, A. B., Tushabe, G., et al. (2022). Malnutrition and associated risk factors among children 6–59 months old in the landslide-prone Bududa District, Eastern Uganda: A cohort study. *Current Developments in Nutrition*, *6*(2), nzac005. https://doi.org/10.1093/CDN/NZAC005

Nduwayezu, G., Manirakiza, V., Mugabe, L., & Malonza, J. M. (2021). Urban growth and land use/land cover changes in the post-genocide period, Kigali, Rwanda. *Environment and Urbanization Asia*, *12*(1_suppl), S127–S146. https://doi.org/10.1177/0975425321997971

Nduwayezu, G., Zhao, P., Kagoyire, C., Eklund, L., Bizimana, J. P., Pilesjo, P., & Mansourian, A. (2023). Understanding the spatial non-stationarity in the relationships between malaria incidence and environmental risk factors using Geographically Weighted Random Forest: A case study in Rwanda. *Geospatial Health*, *18*(1), 1184. https://doi.org/10.4081/gh.2023.1184

NECDP. (2017). Rwanda National Early Childhood Development Program Strategic Plan 2018–2024.

Niles, M. T., Emery, B. F., Wiltshire, S., Brown, M. E., Fisher, B., & Ricketts, T. H. (2021). Climate impacts associated with reduced diet diversity in children across nineteen countries. *Environmental Research Letters*, *16*(1), 015010. https://doi.org/10.1088/1748-9326/ABD0AB

NISR. (2018). Rwandan Integrated Household Living Conditions Survey (EICV5) 2007/2018. Main Indicators Report.

NISR. (2019). *GDP National Accounts, 2019*. National Institute of Statistics Rwanda. Retrieved from https://www.statistics.gov.rw/publication/gdp-national-accounts-2019

NISR. (2022). Fifth population and housing census - 2022: Main indicators report. Retrieved from http://www.statistics.gov.rw/datasource/47

NISR, MoH-Rwanda, & ICF. (2021). *Rwanda Demographic and Health Survey 2019–2020*. NISR, MOH, and ICF International.

Oshan, T. M., Li, Z., Kang, W., Wolf, L. J., & Stewart Fotheringham, A. (2019). MGWR: A Python implementation of multiscale geographically weighted regression for investigating process spatial heterogeneity and scale. *ISPRS International Journal of Geo-Information*, *8*(6), 269. https://doi.org/10.3390/IJGI8060269

Oshan, T. M., Smith, J. P., & Fotheringham, A. S. (2020). Targeting the spatial context of obesity determinants via multiscale geographically weighted regression. *International Journal of Health Geographics*, *19*(1), 1–17. https://doi.org/10.1186/s12942-020-00204-6

Patz, J., Campbell-Lendrum, D., Holloway, T., & Foley, J. A. (2005). Impact of regional climate change on human health. *Nature*, *438*(7066), 310–317. https://doi.org/10.1038/nature04188

Prillaman, M. (2022). Climate change is making hundreds of diseases much worse. *Nature*. https://doi.org/10.1038/D41586-022-02167-Z

Rasmussen, S. A., Jamieson, D. J., Honein, M. A., & Petersen, L. R. (2016). Zika virus and birth defects — Reviewing the evidence for causality. *New England Journal of Medicine*, *374*(20), 1981–1987. https://doi.org/10.1056/nejmsr1604338

Republic of Rwanda. (2020). Vision 2050. Retrieved from https://vision2050.minecofin.gov.rw/

Richardson, R. B. (2010). Ecosystem services and food security: Economic perspectives on environmental sustainability. *Sustainability*, *2*(11), 3520–3548. https://doi.org/10.3390/SU2113520

Robinson, T., Wint, G. R. W., Conchedda, G., Van Boeckel, T. P., Ercoli, V., Palamara, E., et al. (2014). Mapping the global distribution of livestock. *PLoS One*, *9*(5), e96084. https://doi.org/10.1371/journal.pone.0096084

Ruiz-Casares, M., & Rousseau, C. (2010). Between freedom and fear: Children's views on home alone. *The British Journal of Social Work*, *40*(8), 2560–2577. https://doi.org/10.1093/BJSW/BCQ067

Ruiz-Casares, M., Rousseau, C., Currie, J. L., & Heymann, J. (2012). "I hold on to my teddy bear really tight": Children's experiences when they are home alone. *American Journal of Orthopsychiatry*, *82*(1), 97–103. https://doi.org/10.1111/J.1939-0025.2011.01141.X

Sahiledengle, B., Petrucka, P., Kumie, A., Mwanri, L., Beressa, G., Atlaw, D., et al. (2022). Association between water, sanitation and hygiene (WASH) and child undernutrition in Ethiopia: A hierarchical approach. *BMC Public Health*, *22*(1), 1943. https://doi.org/10.1186/s12889-022-14309-z

Sarker, M. A. R., Alam, K., & Gow, J. (2012). Exploring the relationship between climate change and rice yield in Bangladesh: An analysis of time series data. *Agricultural Systems*, *112*, 11–16. https://doi.org/10.1016/J.AGSY.2012.06.004

Sekiyama, M., Matsuda, H., Mohan, G., Yanagisawa, A., Sudo, N., Amitani, Y., et al. (2020). Tackling child malnutrition by strengthening the linkage between agricultural production, food security, and nutrition in Rural Rwanda. *Sustainability Challenges in Sub-Saharan Africa II: Insights from Eastern and Southern Africa*, 3–28. https://doi.org/10.1007/978-981-15-5358-5_1

Serrat, M. A. (2014). Environmental temperature impact on bone and cartilage growth. *Comprehensive Physiology*, *4*(2), 621–655. https://doi.org/10.1002/CPHY.C130023

Setia, M. S. (2016). Methodology series module 3: Cross-sectional studies. *Indian Journal of Dermatology*, *61*(3), 261. https://doi.org/10.4103/0019-5154.182410

Siswanto, S. Y., & Sule, M. I. S. (2019). The impact of slope steepness and land use type on soil properties in Cirandu Sub-Sub Catchment, Citarum Watershed. *IOP Conference Series: Earth and Environmental Science*, *393*(1), 012059. https://doi.org/10.1088/1755-1315/393/1/012059

Souris, M. (2019). Epidemiology and geography: Principles, methods and tools of spatial analysis. In *Epidemiology and geography: Principles, methods and tools of spatial analysis*. Wiley. https://doi.org/10.1002/9781119528203

Sununtnasuk, C. (2013). *Child nutrition, agricultural production and satellite remotely sensed images: Exploring linkages in Nepal* (Theses and Dissertations). Purdue University. Retrieved from https://docs.lib.purdue.edu/dissertations/AAI1547981

Tellings, A. (2017). Evidence-based practice in the social sciences? A scale of causality, interventions, and possibilities for scientific proof. *Theory and Psychology*, *27*(5), 581–599. https://doi.org/10.1177/0959354317726876

Tusting, L. S., Bradley, J., Bhatt, S., Gibson, H. S., Weiss, D. J., Shenton, F. C., & Lindsay, S. W. (2020). Environmental temperature and growth faltering in African children: A cross-sectional study. *The Lancet Planetary Health*, *4*(3), e116–e123. https://doi.org/10.1016/S2542-5196(20)30037-1

UN-HABITAT. (2018). Transforming our world: The 2030 agenda for sustainable development, *A New Era in Global Health*. https://doi.org/10.1891/9780826190123.ap02

UNICEF. (2017). *Reducing stunting in children under 5 years of age: A comprehensive evaluation of UNICEF's strategies and programme performance – Global synthesis report*. UNICEF. Retrieved from https://www.unicef.org/evaldatabase/index_102384.html

UNICEF, WHO, & World Bank. (2021). Levels and trends in child malnutrition: UNICEF/WHO/The World Bank Group joint child malnutrition estimates: Key findings of the 2021 edition. Retrieved from https://apps.who.int/iris/handle/10665/341135

United Nations. (2021). 2021 Global Nutrition Report: The state of global nutrition. Retrieved from https://globalnutritionreport.org/reports/2022-global-nutrition-report/

Uwiringiyimana, V., Ocké, M. C., Amer, S., & Veldkamp, A. (2019). Predictors of stunting with particular focus on complementary feeding practices: A cross-sectional study in the northern province of Rwanda. *Nutrition*, *60*, 11–18. https://doi.org/10.1016/j.nut.2018.07.016

Uwiringiyimana, V., Veldkamp, A., & Amer, S. (2019). Stunting spatial pattern in Rwanda: An examination of the demographic, socio-economic and environmental determinants. *Geospatial Health*, *14*(2), 329–339. https://doi.org/10.4081/GH.2019.820

Van de Poel, E., O'Donnell, O., & Van Doorslaer, E. (2007). Are urban children really healthier? Evidence from 47 developing countries. *Social Science & Medicine*, *65*(10), 1986–2003. https://doi.org/10.1016/J.SOCSCIMED.2007.06.032

van der Merwe, E., Clance, M., & Yitbarek, E. (2022). Climate change and child malnutrition: A Nigerian perspective. *Food Policy*, *113*, 102281. https://doi.org/10.1016/J.FOODPOL.2022.102281

Vilcins, D., Sly, P. D., & Jagals, P. (2018). Environmental risk factors associated with child stunting: A systematic review of the literature. *Annals of Global Health*, *84*(4), 551. https://doi.org/10.29024/AOGH.2361

Weatherspoon, D. D., Miller, S., Ngabitsinze, J. C., Weatherspoon, L. J., & Oehmke, J. F. (2019). Stunting, food security, markets and food policy in Rwanda. *BMC Public Health*, *19*(1), 1–13. https://doi.org/10.1186/s12889-019-7208-0

Westerterp, K. R. (2001). Energy and water balance at high altitude. *Physiology*, *16*(3), 134–137. https://doi.org/10.1152/PHYSIOLOGYONLINE.2001.16.3.134

WHO. (2018). Reducing stunting in children: Equity considerations for achieving the Global Nutrition Targets 2025.

WHO, & UNICEF. (2021). Indicators for assessing infant and young child feeding practices: Definitions and measurement methods. Retrieved from https://www.who.int/publications/i/item/9789240018389

Wimberly, M. C., de Beurs, K. M., Loboda, T. V., & Pan, W. K. (2021). Satellite observations and malaria: New opportunities for research and applications. *Trends in Parasitology*, *37*(6), 525–537. https://doi.org/10.1016/j.pt.2021.03.003

Woodward, M. (2013). *Epidemiology: Study design and data analysis* (3rd ed.). CRC Press. https://doi.org/10.1201/B16343/EPIDEMIOLOGY-MARK-WOODWARD

Yeboah, E., Kuunibe, N., Mank, I., Parisi, D., Bonnet, E., Lohmann, J., et al. (2022). Every drop matters: Combining population-based and satellite data to investigate the link between lifetime rainfall exposure and chronic undernutrition in children under five years in rural Burkina Faso. *Environmental Research Letters*, *17*(5), 054027. https://doi.org/10.1088/1748-9326/AC661C

Zhang, Y., Liang, S., Ma, H., He, T., Wang, Q., Li, B., et al. (2023). *Generation of global 1-km daily soil moisture product from 2000 to 2020 using ensemble learning*. Earth System Science Data. Retrieved from https://www.researchgate.net/publication/367116850_Generation_of_global_1-km_daily_soil_moisture_product_from_2000_to_2020_using_ensemble_learning

Zhou, C., Lv, Q., Yang, N., & Wang, F. (2021). Left-behind children, parent-child communication and psychological resilience: A structural equation modeling analysis. *International Journal of Environmental Research and Public Health*, *18*(10), 5123. https://doi.org/10.3390/IJERPH18105123