Supplementary information

Mapping child growth failure across low- and middle-income countries

authors and unedited

In the format provided by the **Local Burden of Disease Child Growth Failure Collaborators***

Supplementary Information

Table of Contents

Supplementary Figures

Supplementary Tables

0.0.1 Supplementary Discussion

The two main immediate causes of child growth failure (CGF) are the insufficient quantity and nutritional quality of food, and infections^{1,2}. Stunting is a chronic condition marked by inadequate linear growth, often indicative of long-term inadequate nutrition and recurrent infections^{3,4}. Wasting encompasses moderate and severe acute malnutrition, indicative of short-term weight loss often due to insufficient food intake or incident infectious disease such as diarrhoea^{3,4}. Underweight is a composite of both stunting and wasting, and thus has fewer direct implications for public health decision making^{3,4}.

Stunting is recognised as a reliable proxy of physical well-being in children and can provide insight into health inequalities faced by a population^{4,5}. Cross-cutting community-level solutions are needed to prevent or ameliorate stunting and its multigenerational economic and productivity outcomes on mothers and their children⁶⁻⁸. Peru's results-based budgeting strategy (El Presupuesto por Resultados (PpR)), which includes community-level vaccination campaigns, infant and child growth monitoring, and promotion of improved hygiene and feeding practices, has been praised as one of the key drivers in successfully halving stunting levels in less than a decade⁹. In order to achieve reduction of child stunting prevalence to 25% nationally by 2022, India has expanded its National Nutrition Mission programme (POSHAN Abhiyaan) to an additional 315 districts, focusing on areas with particularly high stunting and poor socioeconomic conditions¹⁰. India's national nutrition policy¹¹ and nutrition projects in India's Bellary district in Karnataka¹²; the Tamil Nadu and Dular scheme in Bihar¹²; and child malnutrition management in the drought-affected Rajasthan district¹³ have shown varied success in improving child growth. Beyond investments supporting proximal solutions for nutritional deficiencies and infectious diseases in pregnant women and children, such as nutrition supplement programmes, breastfeeding and complimentary feeding support, vaccination, and sanitation campaigns^{14,7}, it is critical to simultaneously invest in breaking the cyclical patterns of poverty that risk affecting all future generations. CGF is implicated as both a cause and consequence of entrenched cycles of poverty, as the wider socioeconomic, environmental, and political contexts of nations and communities are underlying factors that perpetuate intergenerational cycles of poor health and converge with poverty to increase risks of stunting^{6,15}. Local estimates aid in identifying communities that are most likely experiencing extreme multidimensional disadvantage.

Wasting is an acute condition associated with food shortages and famine that accompany drought and conflict^{5,16}. In spite of years of humanitarian interventions in low- and middle-income countries (LMICs), wasting has persisted along the African Sahel and in South Asia, with some areas exceeding the critical, emergency-level threshold (≥15%)⁵; these levels have spanned more than a decade in areas of Niger, Chad, South Sudan, Pakistan, and India. Arid and semi-arid areas of Somalia, northeastern Kenya, and Ethiopia's Afar and Somali regions continue to experience endemic wasting as increasingly erratic climatic conditions 17 , competition for resources, and political or civil instability constantly threaten livelihoods and food security^{18,19}. Political unrest, armed conflict, and violence disrupt food production and distribution, contributing to increased risks of wasting^{19–22}; areas in northern Nigeria, DRC, southern Pakistan, and Pakistan-border areas in Afghanistan experienced conflict and high wasting (≥10%)²³ throughout the 2000–2017 period. Particular first administrative-level units have consistently experienced climatic events or conflict, as well as persistently high wasting, including Zinder (Niger), Kanem (Chad), Unity (South Sudan), Maharashtra and Telangana (India), where wasting reached critical levels every year from 2000 to 2017. The persistent nature of wasting in these countries and areas calls for policy and programming prioritization. Organisations have assisted in reducing wasting prevalence through community-level awareness-raising, active screening, and therapeutic feeding programmes, such as the French Red Cross in villages throughout Niger's Niamey, Agadez, and Zinder provinces²⁴.

These maps highlight countries and specific locations that have maintained elevated levels of CGF over time, enabling policy makers, public health practitioners, and donors with tools to support efficient directing of investments and development assistance. These estimates could also aid in developing strategically-placed improved nutrition surveillance systems across countries and regions to inform the most appropriate responses for optimum impact and return on investment²⁵. The effort needed to achieve WHO GNTs will depend on current rates of progress, country-specific population growth rates, nutrition policies and programmes, and the amount of resources allocated toward implementation. We believe the findings of this paper are important to further the field of demographic research and to guide global, national, and administrative-level decision making.

We intend to regularly update and eventually expand these analyses to include high-income country estimates via our user-friendly online visualisation tools. In future analyses, we plan to determine how to incorporate sub-model uncertainty in our results, stratify our CGF estimates by sex and age, assess the double burden of child undernutrition and overweight, analyse important maternal indicators that impact child nutritional status outcomes (such as anemia), and continue to monitor progress toward 2025 WHO GNTs. In doing so, we aim to provide the necessary platform to examine past, current, and future trajectories of malnutrition comprehensively to support evaluation of programme and policy success and the precise targeting of resources to highly-affected populations for the greatest impact.

1.0 GATHER Compliance

Supplementary Table 1: Guidelines for Accurate and Transparent Health Estimates Reporting (GATHER) checklist.

2.0 Data

2.1 CGF Indicator Definitions, Socio-demographic Index (SDI) Classification

We modelled the prevalence of child growth failure (CGF) indicators (stunting, wasting, and underweight; defined in Supplementary Table 2) in 105 low- and middle-income countries (LMICs) from 2000 to 2017. These countries were determined by their Socio-demographic Index (SDI), a summary measure of development which combines education, fertility, and poverty, which was developed and computed by the Global Burden of Disease (GBD) study²⁶. Countries were assigned stages based on their SDI quintile, as well as their geographic continuity. The 105 LMICs (Stage 1 and Stage 2 countries) we modelled in this study are described below, along with their SDI and SDI quintile (e.g, low, low-middle, middle) (Supplementary Table 3). Stage 1 countries were included in our previous study estimating fine geospatial CGF prevalence throughout the continent of Africa²⁷, and Stage 2 countries encompass LMICs in other continents. French Guiana and Western Sahara were not modelled by GBD and therefore do not have a calculated SDI, but were included in this study for geographic continuity. China, Iran, Libya, and Malaysia were included despite high-middle SDI status to create better geographic continuity. Albania and Moldova were excluded despite their Middle SDI status due to geographic discontinuity with other included countries and lack of available survey data. We did not estimate for the island nations of American Samoa, Federated States of Micronesia, Fiji, Kiribati, Marshall Islands, North Korea, Samoa, Solomon Islands, or Tonga, where no available survey data could be sourced.

Supplementary Table 2: Definitions of CGF indicators.

Supplementary Table 3: Socio-demographic Index (SDI) of countries included in model.

2.2 Data Sources

The data sources used to model CGF indicators are described below. Information on survey locations, years, source, and number of individuals, polygons, and/or geo-positioned clusters can be found in Supplementary Table 4–5. Excluded datasets and reasons for their exclusion from our analysis are detailed in Supplementary Table 6.

Supplementary Table 4: Household surveys used in mapping.

*Number identification (NID) can be used to locate a particular data source in the Global Health Data Exchange (GHDx) at [http://ghdx.healthdata.org/.](http://ghdx.healthdata.org/) *Indicates a survey from a country we previously modelled that has been added since the first publication.*

†Data source is not publicly available due to restrictions by the data provider and was used under license for the current study.

Supplementary Table 5: Survey Reports added to model.

*Number identification (NID) can be used to locate a particular data source in the Global Health Data Exchange (GHDx) at [http://ghdx.healthdata.org/.](http://ghdx.healthdata.org/) *Indicates a survey from a country we previously modelled that has been added since the first publication. †Data source is not publicly available due to restrictions by the data provider and was used under license for the current study.*

Supplementary Table 6: Data excluded from model.

2.3 Data Process

Supplementary Figure 1: Flowchart for data processes.

The data processing pipeline began with raw survey microdata and ended with the input data for the model. We extracted and standardised names and measurement units of relevant CGF data and matched the corresponding survey clusters with the finest geographies possible. Observations (representing children) were dropped due to insufficient or implausible data for age, height (≤0 cm or ≥180 cm), or weight (≤0 kg or ≥45 kg). We calculated z-scores using the height, weight, and age data, and implausible z-scores^{3,29,30} (according to WHO reference population²⁸) were dropped; 3.30% of children were dropped due to implausible HAZ scores (<-6 or >6), 2.37% due to implausible WHZ scores (<-5 or >5), and 1.09% due to implausible WAZ scores (<-6 or >5). Children that met the definitions of stunted, wasted, or underweight were identified, and data collapsed by survey, year, and geography. Children that could not be matched to a geography were also dropped. Survey reports that were manually extracted at an aggregated level, most often at the first administrative (Admin 1) or second administrative (Admin 2) level, were appended to the collapsed data, and all data attributed to a polygon were resampled to points. After examining diagnostic plots, a small number of surveys were dropped for exhibiting implausible trends. The final cleaned and vetted data were used as input data for the modelling.

2.4 Data Availability by Region

Supplementary Figures 2–16 show the data availability for stunting, wasting, and underweight indicators in the regions we modelled. We incorporated data from a number of survey series, which are represented in the figures. These included: the Demographic and Health Survey by Macro International (Macro DHS), the Multiple Indicator Cluster Survey by the United Nations International Children's Emergency Fund (UNICEF MICS), the Pan Arab Programme on Family Health survey by League of Arab States (PAPFAM), the Core Welfare Indicators Questionnaire by World Bank (CWQI), the Living Standards Measurement Study by World Bank (LSMS) and the Integrated Surveys on Agriculture (LSMS ISA), the Priority Survey series by World Bank (Priority Survey), the Reproductive Health Survey by Center for Disease Control (CDC RHS), the Demographic and Health Survey for Pacific countries funded by the Asian Development Bank (ADB DHS), the Family Life Survey by RAND Corporation (RAND FLS), the Global Enteric Multicenter Study by the Center for Vaccine Development (CVD GEMS), and a number of other country-specific surveys that were not clearly associated with an international survey series (Other). We greatly reduced our risk of incorporating duplicative data in our model by using almost exclusively raw survey microdata. We regularly monitored larger survey series such as Macro DHS and UNICEF MICS for new data, and worked with GBD collaborators from LMICs to identify and obtain other surveys with useful data. While it is possible that there was some overlap in patient populations for surveys that were conducted in the same countries and similar time periods, it is not possible to identify if and how frequently this occurred.

The database for stunting consists of 142,468 clusters and 18,010 polygons with a sample size totaling over 3.9 million children in LMICs. The database for wasting consists of 142,017 clusters and 17,997 polygons with a sample size totaling over 3.9 million children in LMICs. The database for underweight consists of 142,528 clusters and 18,192 polygons with a sample size of over 4.0 million children in LMICs.

Supplementary Figure 2: Stunting data availability by type and country, 2000–2017 in Africa.

Supplementary Figure 3: Stunting data availability by type and country, 2000–2017 in Central America and the Caribbean and South America.

Supplementary Figure 4: Stunting data availability by type and country, 2000–2017 in East and Southeast Asia.

Supplementary Figure 5: Stunting data availability by type and country, 2000–2017 in South Asia*.*

Stunting: Middle East and Central Asia

Supplementary Figure 6: Stunting data availability by type and country, 2000–2017 in Middle East and Central Asia*.*

Supplementary Figure 7: Wasting data availability by type and country, 2000–2017 in Africa.

Supplementary Figure 8: Wasting data availability by type and country, 2000–2017 in Central America and the Caribbean and South America.

Supplementary Figure 9: Wasting data availability by type and country, 2000–2017 in East and Southeast Asia.

Supplementary Figure 10: Wasting data availability by type and country, 2000–2017 in South Asia.

Wasting: Middle East and Central Asia

Supplementary Figure 11: Wasting data availability by type and country, 2000–2017 in Middle East and Central Asia.

Supplementary Figure 12: Underweight data availability by type and country, 2000–2017 in Africa.

Supplementary Figure 13: Underweight data availability by type and country, 2000–2017 in Central America and the Caribbean and South America.

Supplementary Figure 14: Underweight data availability by type and country, 2000–2017 in East and Southeast Asia.

Supplementary Figure 15: Underweight data availability by type and country, 2000–2017 in South Asia.

Underweight: Middle East and Central Asia

Supplementary Figure 16: Underweight data availability by type and country, 2000–2017 in Middle East and Central Asia.

2.5 Covariates

A variety of environmental and socioeconomic variables were used to predict CGF outcomes. Where available, the finest spatiotemporal resolution of gridded datasets were used. In addition to the covariates detailed below, some country-level variables were included: lag distributed income per capita, and the proportion of the population with access to adequate sanitation, were included in models for stunting, wasting, and underweight.

Supplementary Table 7: Covariates used in mapping.

Supplementary Figure 17: Covariates.

Twelve covariate raster layers of possible environmental and socioeconomic correlates of CGF in LMICs were used as inputs for the stacking modelling process. Time-varying covariates are presented for the year 2017. For the year of production of non-time-varying covariates and additional details, please refer to the individual covariate citation in Supplementary Table 7.

3.0 Supplementary Methods

3.1 Seasonality Adjustment

Weight-for-height z-scores (WHZ) were used to calculate an individual child wasting status. As a data preprocessing step, we performed a seasonality adjustment on individual-level child weights in order to account for differences in observed child weight that may have been due to food scarcity around the month in which the survey was conducted. To adjust weight measurements, we fit a model for each region (Extended Data Fig. 10) with a 12-month seasonal spline, a country-level fixed effect, and a smooth spline over the duration of our data collection using the *mgcv* package in R and the following formula:

 W HZ ~ s_{cc} (month) + $s_{tp}(t)$ + as. factor(country).

 WHZ is a child's weight-for-height z-score, month is the integer-valued month of the year (1, ..., 12), t is the time of the interview in integer months since the earliest observation of any child in the dataset, and country is a factor variable representing the country where the observation was recorded. We modelled the periodic component on months using 12 cyclic cubic regression splines basis functions (cc) and we accounted for a smooth longer time temporal trend using four thin-plate splines (tp) . The country effects and the long-term temporal spline were included only to help avoid confounding during fitting of the seasonal spline fit and neither country effects nor the long-term trend were used in the seasonal adjustment. We then adjusted all observations to account for the difference in the seasonal period between the month of the interview and an average day of the year as determined by which days align with the mean of the periodic spline.

Supplementary Figure 18: Periodic seasonality adjustment.

The fitted seasonal periodic spline for wasting for Central sub-Saharan Africa region with the marked mean of the periodic function and an example of the seasonality adjustment that would be applied to weight-for-height z-scores collected in this region in the month of July.

3.2 Geostatistical Model

3.2.1 Model geographies

A total of three sub-models were run for each CGF indicator based on continuous geographic regions within LMICs chosen to align with the 14 regions as shown the Extended Data Figure 10. These regions were determined based on both proximity and epidemiological similarity. All data within the spatial region, and within a one-degree buffer from the boundaries of each region, were included in each regional model to minimize edge effects.

3.2.2 Ensemble covariate modelling

An ensemble covariate modelling method was implemented in order to select covariates and capture possible non-linear effects and complex interactions between them³¹. For each region, three sub-models were fit to our dataset, using all of our covariate data as explanatory predictors: generalised additive models (GAM), boosted regression trees (BRT), and lasso regression. Country-level fixed effects were also included in the BRT model as dummy-coded covariates. Sample weights were used in sub-models, where applicable, such that cluster locations with latitude and longitude had a sample weight of 1, while cluster locations where the latitude and longitude were generated by the polygon resampling process had a weight based on the K-means clustering process.

Each sub-model's predictive performance was tested using five-fold cross-validation to avoid overfitting. We removed 20% of the data to create five out-of-sample predications and complied these into a single comprehensive set of predictions. Additionally, the same sub-models were also run using 100% of the data, and a full set of in-sample predictions were created. The five sets of out-of-sample sub-model predictions were fed into the full geostatistical model as the explanatory covariates when performing the model fit. The in-sample predictions from the sub-models were used as the covariates when generating predictions using the fitted full geostatistical model. A recent study has shown that this ensemble approach can improve predictive validity by up to 25% over an individual model³¹.

Predictions from each sub-model are generated based on patterns and relationships between the raw covariates and prevalence survey data, while predictions from the full geostatistical model are generated based on patterns and relationships between the predictions from the ensemble of submodels and prevalence survey data. To discover the relationships between the sub-model prediction layers (used as covariates in the full geostatistical model) and the prevalence data, the only values of the covariates (sub-model prediction layers) "seen" by the model are the values underlying the locations of surveys. As such, it is possible that estimates will be generated in areas where the values of the covariates exceed the minimum and maximum values observed by the model. In these areas, the estimates are generated by extrapolating from the patterns observed within the range of covariates underlying the survey data. More information on the list of the covariates and plots of all covariates, can be found in Supplementary Table 7 and Supplementary Fig. 17.

The primary goal of using the stacking procedure in our analyses was to maximise the predictive power of the raster covariates by capturing the non-linear effects and complex interactions between covariates to optimise the model performance. Bhatt *et al*. (2017)³¹ contend that the primary purpose of the sub-model predictions is to improve the mean function of the Gaussian process. While we have determined a way to include the uncertainty from two of our sub-models (GAM and Lasso regression), we have not determined a way to include uncertainty from the BRT sub-model into our final estimates. Whereas GAM and Lasso regression seek to fit a single model that best describes the relationship between response variable and some set of predictors, BRT method fits a large number of relatively simple models whose predictions are then combined to give robust estimates of the response. While this feature of BRT model makes it a powerful tool for analysing complex data, quantifying the relative uncertainty contributed by each simple model as well as uncertainty from the complex interactions of the predictor variables is challenging $32,33$.

Supplementary Figure 19: Ensemble predicted rasters.

Predicted 2017 rasters, for use as covariates in the INLA (integrated nested Laplace approximation) modelling, shown for the Eastern sub-Saharan Africa region. The *gam* plot shows the predictions from a generalised additive model fit, the *gbm* plot shows the predictions from a boosted regression tree fit, the *lasso* plot shows the predictions from a lasso penalised regression model fit.
3.2.3 Model description

Binomial count data are modelled within a Bayesian hierarchical modelling framework using a logit link function and a spatially and temporally explicit hierarchical generalised linear regression model to fit prevalence of each of CGF indicators in 14 modelling regions³⁴: Andean South America, Central America and the Caribbean, Central sub-Saharan Africa (SSA), East Asia, Eastern SSA, Middle East, North Africa, Oceania, Southeast Asia, South Asia, Southern SSA, Central Asia, Tropical South America, and Western-SSA (as seen in Extended Data Fig 10). For each region, we explicitly write the hierarchy that defines our Bayesian method:

For each binomial CGF indicator, we modelled mean children with stunting, wasting, or who were underweight in each survey cluster, d . Survey clusters are precisely located by their GPS coordinates and year of observation, which we map to a spatial raster location, i , at time, t . We observed the number of children reported to be stunted, wasted, or underweight, respectively, as binomial count data, C_d , among an observed sample size, N_d . As we may have observed several data clusters within a given location, i, at time, t, we refer to the probability of stunting, wasting, or underweight, p , within a given cluster, d , by its indexed location, i, and time, t, as $p_{i(d),t(d)}$.

$$
C_{d}|p_{i(d),t(d)}, N_{d} \sim \text{Binomial}(p_{i(d),t(d)}, N_{d}) \space \forall \text{ observed clusters } d
$$
\n
$$
\text{logit}(p_{i,t}) = \beta_{0} + \mathbf{X}_{i,t}\boldsymbol{\beta} + Z_{i,t} + \epsilon_{\text{ctr}(i)} + \epsilon_{i,t} + Z_{i,t} \space \forall i \in \text{spatial domain } \forall t \in \text{time domain}
$$
\n
$$
\sum_{h=1}^{3} \beta_{h} = 1
$$
\n
$$
\epsilon_{\text{ctr}} \sim \text{iid Normal}(0, \gamma^{2})
$$
\n
$$
\epsilon_{i,t} \sim \text{iid Normal}(0, \sigma^{2})
$$
\n
$$
\mathbf{Z} \sim \text{GP}(0, \Sigma^{\text{space}} \otimes \Sigma^{\text{time}})
$$
\n
$$
\Sigma^{\text{space}} = \frac{\omega^{2}}{\Gamma(\nu)2^{\nu-1}} \times (\kappa D)^{\nu} \times \text{K}_{\nu}(\kappa D)
$$

For indices d , i, and t , *(index) is the value of * at that index. The probabilities, $p_{i,t}$ represent both the annual prevalence at the space-time location and the probability that an individual child was afflicted with the risk factor given that they lived at that particular location. The annual prevalence, $p_{i,t}$, of each indicator was modelled as a linear combination of the three sub-models (generalised additive model (GAM), boosted regression trees (BRT), and lasso regression), rasterised covariate values $\mathbf{X}_{i,t}$, a correlated spatiotemporal error term, $Z_{i,t}$, country random effects, $\epsilon_{ctr(i)}$ with one unstructured country random effect fit for each country in the modelling region and all ϵ_{ctr} sharing a common variance parameter, γ^2 , and an independent nugget effect, $\epsilon_{i,t}$ with variance parameter, σ^2 . Coefficients in β_h in the three sub-models $h=1,2,3$ represent their respective predictive weighting in the mean logit link, while the joint error term, $Z_{i,t}$, accounts for residual spatiotemporal autocorrelation between individual data points that remains after accounting for the predictive effect of the sub-model covariates, the country-level random effect, $\epsilon_{\text{ctr}(i)}$, and the nugget independent error term, $\epsilon_{i,t}$. The residuals, $Z_{i,t}$, are modelled as a three-dimensional Gaussian process (GP) in space-time centered at

 $\Sigma_{j,k}^{\text{time}} = \rho^{\vert k-j \vert}$

zero and with a covariance matrix constructed from a Kronecker product of spatial and temporal covariance kernels. The spatial covariance, Σ^{space}, is modelled using an isotropic and stationary Matérn function³⁵, and temporal covariance, Σ^{time} fun, as an annual autoregressive (AR1) function over the 18 years represented in the model. In the stationary Matérn function, Γ is the Gamma function, K_v is the modified Bessel function of order $v > 0$, $\kappa > 0$ is a scaling parameter, D denotes the Euclidean distance, and ω^2 is the marginal variance. The scaling parameter, κ , is defined to be $\kappa=\sqrt{8v}/\delta$ where δ is a range parameter (which is about the distance where the covariance function approaches 0.1) and v is a scaling constant, which is set to 2 rather than fit from the data $36,37$. This is parameter is difficult to reliably fit, as documented by many other analyses^{36,38,39} that set this to 2. The number of rows and the number of columns of the spatial Matérn covariance matrix are both equal to the number of spatial mesh points for a given modelling region. In the AR1 function, ρ is the autocorrelation function (ACF), and k and j are points in the time series where $|k - j|$ defines the lag. The number of rows and the number of columns of the AR1 covariance matrix are both equal to the number of temporal mesh points (18). The number of rows and the number of columns of the space-time covariance matrix, $\Sigma^{\text{space}} \otimes$ Σ^{time} , for a given modelling region are both equal to: (the number of spatial mesh points times the number of temporal mesh points).

This approach leveraged the data's residual correlation structure to more accurately predict prevalence estimates for locations with no data, while also propagating the dependence in the data through to uncertainty estimates⁴⁰. The posterior distributions were fit using computationally efficient and accurate approximations in R-INLA^{41,42} (integrated nested Laplace approximation) with the stochastic partial differential equations (SPDE)³⁷ approximation to the Gaussian process residuals using R project v.3.5.1. The SPDE approach using INLA has been demonstrated elsewhere, including the estimation of health indicators, particulate air matter, and population age structure^{27,43–46}. Uncertainty intervals (UIs) were generated from 1,000 draws (i.e., statistically plausible candidate maps)⁴⁷ created from the posteriorestimated distributions of modelled parameters.

3.2.4 Priors

The following priors were used for all three of our CGF models:

$$
\beta_0 \sim N(\mu = 0, \sigma^2 = 3^2),
$$

\n
$$
\beta \sim \text{iid } N\left(\mu = \frac{1}{\#\text{ ensemble models}}, \sigma^2 = 3^2\right),
$$

\n
$$
\log\left(\frac{1+\rho}{1-\rho}\right) \sim N(\mu = 0, \sigma^2 = 1/0.15),
$$

\n
$$
\log\left(\frac{1}{\sigma_{nug}^2}\right) \sim \log{gamma(\alpha = 1, \gamma = 2)}.
$$

\n
$$
\theta_1 = \log(\tau) \sim N(\mu_{\theta_1}, \sigma_{\theta_1}^2)
$$

\n
$$
\theta_2 = \log(\kappa) \sim N(\mu_2, \sigma_{\theta_2}^2).
$$

Given that our covariates used in INLA (i.e. the predicted outputs from the ensemble models) should be on the same scale as our predictive target, we believe that the intercept in our model should be close to zero and that the regression coefficients should sum to 1. As such, we have chosen the prior for our intercept to be $N(0, \sigma^2 = 3^2)$, and the prior for the fixed effect coefficients to be $N(\frac{1}{\# \text{ensemble models}}, \sigma^2 = 3^2)$. The prior on the temporal correlation parameter, ρ , is chosen to be mean zero, showing no prior preference for either positive or negative auto-correlation structure, and with a distribution that is wide enough such that within three standard deviations of the mean, the prior includes values of ρ ranging from -0.95 to 0.95. The priors on the random effect variances were chosen to be relatively loose given that we believe our fixed effects covariates should be well-correlated with our outcome of interest, which might suggest relatively small random effects values. At the same time, we wanted to avoid using a prior that was so diffuse as to actually put high prior weight on large random effect variances. For stability, we used the uncorrelated multivariate normal priors that INLA automatically determines (based on the finite elements mesh) for the log-transformed spatial hyperparameters, κ and τ . In our parameterisation, we represent α and γ in the *loggamma* distribution as scale and shape, respectively.

3.2.5 Mesh construction

We constructed the finite elements mesh for the stochastic partial differential equation approximation to the Gaussian process regression using a simplified polygon boundary (in which coastlines and complex boundaries were smoothed) for each of the regions within our model. We set the inner mesh triangle maximum edge length (the mesh size for areas over land) to be 0.75 degrees, and the buffer maximum edge length (the mesh size for areas over the ocean) to be 5.0 degrees. An example finite elements mesh constructed for Eastern sub-Saharan mesh can be found in [Supplementary Fig.](#page-76-0) 20*.*

3.2.6 Model fitting and estimate generation

Models were fit in INLA with methods consistent with those used in geospatial modelling of CGF, under-5 mortality, and educational attainment in Africa, published previously^{27,43,48}.

Where possible, the point data (GPS-positioned data) were used in the analyses. In instances where this was not possible, the data were matched to the smallest possible areal unit. The areal data were then resampled to generate pseudo-point data based on the underlying population distribution within the polygon. The methods for the resampling are consistent with those previously used in geospatial modelling of under-5 mortality⁴³. Resampling K-means weights were used within the INLA fit by multiplying the corresponding log-likelihood evaluation for the specific observation by the observation's K-means weight. These weights were used to ensure that we did not artificially inflate the amount of information in the dataset by effectively using them to inflate the dispersion in the log-likelihood for resampled-polygon points. While the model this induces is not necessarily generative, it does yield a well-defined target distribution. This is analogous to how weighting is often done in generalised additive models⁴⁹. Data points that could be geo-referenced to latitude-longitude locations were assigned a weight of 1, ensuring that when the log-likelihood contribution from an observation was evaluated it contributed only to the log-likelihood at that observation's space-time location. For cluster locations generated based on the polygon resampling process, the log-likelihood of those points contributed proportionate to the K-means weights, effectively diffusing the evaluation of the observation across the polygon.

As part of the ensemble modelling process, prediction surfaces from the out-of-sample ensemble submodels were used as covariates in the spatiotemporal model. Estimates of the fixed effects' beta coefficients were derived from the contribution of each of the sub-models to INLA's predicted prevalence estimates, in conjunction with parameter estimates of the contribution of location and time. To create final estimates, the in-sample prediction surfaces of prevalence from the sub-models (serving as covariates) were used as covariates in conjunction with the fitted random effects from INLA to predict and calculate estimates of prevalence for each grid cell in each year.

Our implementation of INLA using the R-INLA software relies on a Gaussian approximation of the full conditional distribution of latent variables, and uses the empirical Bayes approximation for the hyperparemeters⁴¹. We have tried the full hyperparameter grid integration and central composite design (CCD) integration in various settings and have found our models to be nearly indistinguishable. Due to its computing resource efficiency, we used the empirical Bayes procedure. In a very similar setting with malaria household survey data, other authors (including the senior author here) compared the INLA results directly with those from Hamiltonian Markov Chain Monte Carlo and found nearly identical results between the two fits 50 .

All estimates were generated by taking 1,000 draws from the posterior distribution, which yielded 1,000 candidate maps used to summarise the grid cell- and aggregated-level statistics. For estimates at the grid cell level, these draws were used directly to generate estimates and uncertainty. Aggregated estimates, in which grid cell-level estimates were summarised to administrative boundaries, were generated by creating population-weighted averages for each administrative boundary, for each draw. 95% uncertainty intervals around the mean of our estimates were generated.

Finite elements mesh over Eastern Sub-Saharan Africa

Supplementary Figure 20: Finite elements mesh.

The finite elements mesh used to fit the space-time correlated error for the Eastern sub-Saharan Africa (ESSA) region overlaid on the countries in ESSA. Both the fine-scale mesh over land in the modelling region and the coarser buffer region mesh are shown. The simplified region polygon used to determine the boundary for the modelling region is shown in blue.

3.3 Model Results

Supplementary Table 9: Stunting fitted parameters.

Lower, median, and upper quantiles (percentiles 0.025, 0.50, 0.975) are displayed for the main parameters from the stunting models by region. The fixed effects covariates corresponding to the predicted ensemble rasters are shown in the first five columns, while fitted values for the spatiotemporal field hyperparameters and the precisions (inverse variance) for our random effects are shown in the last five columns.

Supplementary Table 10: Wasting fitted parameters.

Lower, median, and upper quantiles (percentiles 0.025, 0.50, 0.975) are displayed for the main parameters from the wasting models by region. The fixed effects covariates corresponding to the predicted ensemble rasters are shown in the first five columns, while fitted values for the spatiotemporal field hyperparameters and the precisions (inverse variance) for our random effects are shown in the last five columns.

Supplementary Table 11: Underweight fitted parameters.

Lower, median, and upper quantiles (percentiles 0.025, 0.50, 0.975) are displayed for the main parameters from the underweight models by region. The fixed effects covariates corresponding to the predicted ensemble rasters are shown in the first five columns, while fitted values for the spatiotemporal field hyperparameters and the precisions (inverse variance) for our random effects are shown in the last five columns.

Supplementary Figure 21: Stunting posterior means and upper and lower 95% uncertainty intervals for 2017.

Supplementary Figure 22: Wasting posterior means and upper and lower 95% uncertainty intervals for 2017.

Supplementary Figure 23: Underweight posterior means and upper and lower 95% uncertainty intervals for 2017.

4.0 Supplementary Results

4.1 Additional information for figure descriptions

For Figure 1d (overlapping population-weighted quartiles for child stunting and relative 95% uncertainty for 2017):

Quartile cut-offs were 16.3% (25th percentile), 27.3% (50th percentile), 39.9% (75th percentile) for the stunting prevalence axis, and 0.758 (25th) percentile), 1.094 (50th percentile), and 1.352 (75th percentile) for the relative uncertainty axis (calculated as the absolute range of the uncertainty intervals divided by the estimate).

For Figure 2d (overlapping population-weighted quartiles for child wasting and relative 95% uncertainty for 2017):

Quartile cut-offs were 3.5% (25th percentile), 6.8% (50th percentile), 12.6% (75th percentile) for the wasting prevalence axis, and 1.178 (25th percentile), 1.439 (50th percentile), and 1.758 (75th percentile) for the relative uncertainty axis (calculated as the absolute range of the uncertainty intervals divided by the estimate).

For Extended Data Figure 5d (overlapping population-weighted quartiles for child underweight and relative 95% uncertainty for 2017): Quartile cut-offs were 5.5% (25th percentile), 10.8% (50th percentile), 26.3% (75th percentile) for the underweight prevalence axis, and 0.858 (25th percentile), 1.122 (50th percentile), and 1.703 (75th percentile) for the relative uncertainty axis (calculated as the absolute range of the uncertainty intervals divided by the estimate).

For Figures 1g, 2g, and Extended Data Figure 5g (2000-2017 annualised decreases in CGF indicators relative to rates needed during 2017–2025 to meet WHO GNT):

100% indicates the annualised decrease from 2000 to 2017 is equivalent to the pace of progress required during 2017–2025 to meet WHO GNT by 2025 (40% reduction in stunting or underweight; wasting prevalence less than 5%) relative to 2010. Blue and green grid cells exceeded this pace; yellow grid cells proceeded at a slower rate than required; orange grid cells were non-decreasing; and purple grid cells were estimated to have met the target by 2017 ('Met GNT').

For Figures 1h, 2h, and Extended Data Figure 5h:

Grid-cell-level (5 × 5-km resolution) predicted prevalence in 2025 is based on annualised decrease achieved from 2000 to 2017 and projected from 2017.

For Figures 1, 2, and Extended Data Figures 5, 7, and 8:

Maps reflect administrative boundaries, land cover, lakes, and population; grey-coloured grid cells had fewer than ten people per 1 × 1-km grid cell and were classified as "barren or sparsely vegetated"^{51–57}, or were not included in these analyses. Interactive visualisation tools are available at [https://vizhub.healthdata.org/lbd/cgf.](https://vizhub.healthdata.org/lbd/cgf)

4.2 Countries estimated to meet WHO GNT in 2017 and 2025 at various spatial levels

Supplementary Table 12: Countries predicted to have met WHO GNTs by 2017, with >95% probability.

"Yes" indicates that the country was estimated to have met WHO GNT for either stunting, wasting, or underweight in 2017 at the national level (Admin 0), in all first administrative-level units (Admin 1), and/or in all second administrative-level units (Admin 2), with >95% probability.

Supplementary Table 13: Countries predicted to meet WHO GNTs by 2025, with >95% probability.

"Yes" indicates that the country is predicted to meet WHO GNT for either stunting, wasting, or underweight in 2025 at the national level (Admin 0), in all first administrative-level units (Admin 1), and/or in all second administrative-level units (Admin 2), with >95% probability.

5.0 Model Validation

5.1 In-sample metrics

In order to assess the in-sample performance of our models and compare to national-level estimates produced by GBD, we generated a suite of diagnostic plots for each CGF indicator estimates in each of the regions and countries modelled.

To explore residual error over space and time, absolute error (data minus predicted posterior mean estimates at the corresponding grid cells) were produced.

5.2 Metrics of predictive validity

In order to assess the predictive validity of our estimates, we validated our models using spatiallystratified five-fold out-of-sample cross-validation⁵⁸. To construct each spatial fold, we used a modified bi-tree algorithm to spatially aggregate data points. This algorithm recursively partitions twodimensional space, alternating between horizontal and vertical splits on the weighted data sample size medians, until the data contained within each spatial partition are of a similar sample size. The depth of recursive partitioning is constrained by the target sample size within a partition and the minimum number of clusters or pseudo-clusters allowed within each spatial partition (in this case, a minimum sample size of 500 was used). These spatial partitions are then allocated to one of five folds for crossvalidation. For validation, each geostatistical model was run five times, each time holding out data from one of the folds, generating a set of out-of-sample predictions for the held-out data. For each indicator, a full suite of out-of-sample predictions over the entire dataset was generated by combining the out-ofsample predictions from the five cross-validation runs.

Using these out-of-sample predictions, we then calculated mean error (ME, or bias), root-meansquared-error (RMSE, which summarises total variance), coefficient of variation (CoV, defined to be the standard deviation divided by the mean and multiplied by 100, which is a measure of relative variability), and 95% coverage of our predictive intervals (the proportion of observed out-of-sample data that fall within our predicted 95% credible intervals) aggregated up to different administrative levels (levels 0, 1, and 2) as defined by Database of Global Administrative Areas (GADM)⁵⁴. Administrative level 0 (Admin 0) borders correspond to national boundaries, first administrative level (Admin 1) borders generally correspond to regions, provinces, or state-level boundaries within a country, and second administrative level (Admin 2) borders correspond to the next finer unit-level, often districts, within regions. These metrics are summarised in Supplementary Tables 14–22 for each CGF indicator and are calculated across all regions. Included in the sample tables for comparison are the same metrics calculated on in-sample predictions.

5.2.1 Stunting validation metrics

The out-of-sample (OOS) column indicates whether the metric was calculated using in-sample or out-ofsample predictions. Mean error, root-mean-squared-error (RMSE), Correlation (Corr), and coefficient of variation (Cov) are in proportion.

Supplementary Table 14: Predictive metrics for stunting aggregated to the national level (Admin 0).

Supplementary Table 15: Predictive metrics for stunting aggregated to the first administrative level (Admin 1).

Supplementary Table 16: Predictive metrics for stunting aggregated to the second administrative level (Admin 2).

Supplementary Figure 24: Stunting national-level (Admin 0) aggregation in-sample.

Comparison of in-sample stunting predictions aggregated to Admin 0 with 95% uncertainty intervals plotted against Admin 0 aggregated data observations.

Supplementary Figure 25: Stunting national-level (Admin 0) aggregation out-of-sample.

Comparison of out-of-sample stunting predictions aggregated to Admin 0 with 95% uncertainty intervals plotted against Admin 0 aggregated data observations.

Supplementary Figure 26: Stunting first-administrative-level (Admin 1) aggregation in-sample.

Comparison of in-sample stunting predictions aggregated to Admin 1 with 95% uncertainty intervals plotted against Admin 1 aggregated data observations.

Supplementary Figure 27: Stunting first-administrative-level (Admin 1) aggregation out-ofsample.

Comparison of out-of-sample stunting predictions aggregated to Admin 1 with 95% uncertainty intervals plotted against Admin 1 aggregated data observations.

Supplementary Figure 28: Stunting second-administrative-level (Admin 2) aggregation insample.

Comparison of in-sample stunting predictions aggregated to Admin 2 with 95% uncertainty intervals plotted against Admin 2 aggregated data observations.

Supplementary Figure 29: Stunting second-administrative-level (Admin 2) aggregation out-ofsample.

Comparison of out-of-sample stunting predictions aggregated to Admin 2 with 95% uncertainty intervals plotted against Admin 2 aggregated data observations.

5.2.2 Wasting validation metrics

The out-of-sample (OOS) column indicates whether the metric was calculated using in-sample or out-ofsample predictions. Mean error, root-mean-squared-error (RMSE), Correlation (Corr), and coefficient of variation (Cov) are in proportion.

Supplementary Table 17: Predictive metrics for wasting aggregated to the national level (Admin 0).

Supplementary Table 18: Predictive metrics for wasting aggregated to the first administrative level (Admin 1).

Supplementary Table 19: Predictive metrics for wasting aggregated to the second administrative level (Admin 2).

Supplementary Figure 30: Wasting national-level (Admin 0) aggregation in-sample.

Comparison of in-sample wasting predictions aggregated to Admin 0 with 95% uncertainty intervals plotted against Admin 0 aggregated data observations.

Supplementary Figure 31: Wasting national-level (Admin 0) aggregation out-of-sample.

Comparison of out-of-sample wasting predictions aggregated to Admin 0 with 95% uncertainty intervals plotted against Admin 0 aggregated data observations.

Supplementary Figure 32: Wasting first-administrative-level (Admin 1) aggregation in-sample.

Comparison of in-sample wasting predictions aggregated to Admin 1 with 95% uncertainty intervals plotted against Admin 1 aggregated data observations.

Supplementary Figure 33: Wasting first-administrative-level (Admin 1) aggregation out-ofsample.

Comparison of out-of-sample wasting predictions aggregated to Admin 1 with 95% uncertainty intervals plotted against Admin 1 aggregated data observations.

Supplementary Figure 34: Wasting second-administrative-level (Admin 2) aggregation insample.

Comparison of in-sample wasting predictions aggregated to Admin 2 with 95% uncertainty intervals plotted against Admin 2 aggregated data observations.

Supplementary Figure 35: Wasting second-administrative-level (Admin 2) aggregation out-ofsample.

Comparison of out-of-sample wasting predictions aggregated to Admin 2 with 95% uncertainty intervals plotted against Admin 2 aggregated data observations.

5.2.3 Underweight validation metrics

The out-of-sample (OOS) column indicates whether the metric was calculated using in-sample or out-ofsample predictions. Mean error, root-mean-squared-error (RMSE), Correlation (Corr), and coefficient of variation (Cov) are in proportion.

Supplementary Table 21: Predictive metrics for underweight aggregated to the first administrative level (Admin 1).

Supplementary Table 22: Predictive metrics for underweight aggregated to the second administrative level (Admin 2).

Supplementary Figure 36: Underweight national-level (Admin 0) aggregation in-sample.

Comparison of in-sample underweight predictions aggregated to Admin 0 with 95% uncertainty intervals plotted against Admin 0 aggregated data observations.

Supplementary Figure 37: Underweight national-level (Admin 0) aggregation out-of-sample.

Comparison of out-of-sample underweight predictions aggregated to Admin 0 with 95% uncertainty intervals plotted against Admin 0 aggregated data observations.

Supplementary Figure 38: Underweight first-administrative-level (Admin 1) aggregation insample.

Comparison of in-sample underweight predictions aggregated to Admin 1 with 95% uncertainty intervals plotted against Admin 1 aggregated data observations.

Supplementary Figure 39: Underweight first-administrative-level (Admin 1) aggregation outof-sample.

Comparison of out-of-sample underweight predictions aggregated to Admin 1 with 95% uncertainty intervals plotted against Admin 1 aggregated data observations.

Supplementary Figure 40: Underweight second-administrative-level (Admin 2) aggregation insample.

Comparison of in-sample underweight predictions aggregated to Admin 2 with 95% uncertainty intervals plotted against Admin 2 aggregated data observations.

Supplementary Figure 41: Underweight second-administrative-level (Admin 2) aggregation out-of-sample.

Comparison of out-of-sample underweight predictions aggregated to Admin 2 with 95% uncertainty intervals plotted against Admin 2 aggregated data observations.

6.0 Post-estimation Calibration to National and Subnational Estimates

In order to leverage national-level data included in GBD 2017²⁶, but outside the scope of our current geospatial modelling framework, and to ensure perfect calibration between these estimates and GBD 2017 estimates, we performed a post hoc calibration to each of our 1,000 candidate maps. For each posterior draw, we calculated population-weighted grid cell aggregations at the level of GBD estimates (at national or subnational level) and compared these estimates in each year to the analogous and available GBD 2017 estimates from 2000 to 2017. We defined the calibration factor to be the ratio between the GBD 2017 estimates and our current estimates and linearly interpolated calibration factors in each country between the available years. Finally, we multiplied each of our grid cells in a countryyear by its associated calibration factor. This ensures perfect calibration between our geospatial estimates and GBD 2017 estimates, while preserving our estimated within-country geospatial and temporal variation.

7.0 Supplementary References

- 1. Caulfield, L. E., Richard, S. A., Rivera, J. A., Musgrove, P. & Black, R. E. Stunting, Wasting, and Micronutrient Deficiency Disorders. in *Disease Control Priorities in Developing Countries* (eds. Jamison, D. T. et al.) (World Bank, 2006).
- 2. Black, R. E. Would control of childhood infectious diseases reduce malnutrition? *Acta Paediatr.* **80**, 133–140 (1991).
- 3. Waterlow, J. C. *et al.* The presentation and use of height and weight data for comparing the nutritional status of groups of children under the age of 10 years. 489–498 (1977).
- 4. WHO. *Physical status: the use and interpretation of anthropometry: report of a WHO Expert Committee*.

http://apps.who.int/iris/bitstream/handle/10665/37003/WHO_TRS_854.pdf;jsessionid=2FE8F4F177 D025F6433656F4F7577F3F?sequence=1 (1995).

- 5. WHO. *Nutrition Landscape Information System (NLIS) Country Profile Indicators: Interpretation Guide*. https://www.who.int/nutrition/nlis/en/ (2010).
- 6. Stewart, C. P., Iannotti, L., Dewey, K. G., Michaelsen, K. F. & Onyango, A. W. Contextualising complementary feeding in a broader framework for stunting prevention. *Matern. Child. Nutr.* **Suppl 2**, 27–45 (2013).
- 7. Bhutta, Z. A. *et al.* What works? Interventions for maternal and child undernutrition and survival. *Lancet* **371**, 417–440 (2008).
- 8. *2017 Global Nutrition Report: Nourishing the SDGs*. http://globalnutritionreport.org/wpcontent/uploads/2017/11/Report_2017-2.pdf (2017).
- 9. Marini, A., Rokx, C. & Gallagher, P. *Standing Tall: Peru's Success in Overcoming its Stunting Crisis*. http://documents.worldbank.org/curated/en/815411500045862444/pdf/FINAL-Peru-Nutrition-Book-in-English-with-Cover-October-12.pdf (2017).
- 10. National Institution for Transforming India (NITI Aayog); Government of India. Nourishing India: National Nutrition Strategy. http://niti.gov.in/writereaddata/files/document_publication/Nutrition_Strategy_Booklet.pdf (2017).
- 11. Government of India, Department of Women and Child Development, Ministry of Human Resource Development. National Nutrition Policy. http://wcd.nic.in/sites/default/files/nnp_0.pdf (1993).
- 12. Sahu, S. K. *et al.* Malnutrition among under-five children in India and strategies for control. *J. Nat. Sci. Biol. Med.* **6**, 18–23 (2015).
- 13. Kumar, S. & Bhawani, L. Managing child malnutrition in a drought affected district of Rajasthan--a case study. *Indian J. Public Health* **49**, 198–206 (2005).
- 14. Black, R. E. & Heidkamp, R. Causes of Stunting and Preventive Dietary Interventions in Pregnancy and Early Childhood. *Nestle Nutr. Inst. Workshop Ser.* **89**, 105–113 (2018).
- 15. Black, R. E. *et al.* Maternal and child undernutrition: global and regional exposures and health consequences. *Lancet* **371**, 243–260 (2008).
- 16. Neufeld, L. M. & Osemdar, S. J. M. World Nutrition Situation: Global, Regional and Country Trends in Underweight and Stunting as Indicators of Nutrition and Health of Populations. in *International Nutrition: Achieving Millennium Goals and Beyond* vol. 78 11–19 (2013).
- 17. Chotard, S., Mason, J. B., Oliphant, N. P., Mebrahtu, S. & Hailey, P. Fluctuations in wasting in vulnerable child populations in the Greater Horn of Africa. *Food Nutr. Bull.* **31**, S219–S233 (2010).
- 18. FAO, IFAD, UNICEF, WFP & WHO. *The State of Food Security and Nutrition in the World 2017. Building Resilience for Peace and Food Security*. http://www.fao.org/3/a-I7695e.pdf (2017).
- 19. Kinyoki, D. K. *et al.* Conflict in Somalia: impact on child undernutrition. *BMJ Glob. Health* **2**, e000262 (2017).
- 20. Dunn, G. The impact of the Boko Haram insurgency in Northeast Nigeria on childhood wasting: a double-difference study. *Confl. Health* **12**, (2018).
- 21. Akseer, N. *et al.* Geospatial inequalities and determinants of nutritional status among women and children in Afghanistan: an observational study. *Lancet Glob. Health* **6**, PE447-E459 (2018).
- 22. ACLED. The Armed Conflict Location and Event Data Project (ACLED). https://www.acleddata.com/dashboard/ (2017).
- 23. de Onis, M. *et al.* Prevalence thresholds for wasting, overweight and stunting in children under 5 years. *Public Health Nutr.* 1–5 (2018).
- 24. IFRC. Tackling malnutrition in Niger. https://www.ifrc.org/PageFiles/113731/Tackling%20malnutrition%20in%20Niger-EN.pdf.
- 25. *2018 Global Nutrition Report: Shining a light to spur action on nutrition*. https://globalnutritionreport.org/reports/global-nutrition-report-2018/ (2018).
- 26. Dicker, D. *et al.* Global, regional, and national age-sex-specific mortality and life expectancy, 1950– 2017: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet* **392**, 1684–1735 (2018).
- 27. Osgood-Zimmerman, A. *et al.* Mapping child growth failure in Africa between 2000 and 2015. *Nature* **555**, 41–47 (2018).
- 28. WHO Multicentre Growth Reference Study Group. WHO Child Growth Standards based on length/height, weight and age. *Acta Paediatr.* **95**, 76–85 (2006).
- 29. WHO & UNICEF. *WHO Child Growth Standards and the Identification of Severe Acute Malnutrition in Infants and Children: A Joint Statement*. http://www.who.int/nutrition/publications/severemalnutrition/9789241598163/en/ (2009).
- 30. Wang, Y. & Chen, H.-J. Use of percentiles and z-scores in anthropometry. in *Handbook of Anthropometry* (ed. Preedy, V. R.) 29–48 (Springer New York, 2012).
- 31. Bhatt, S. *et al.* Improved prediction accuracy for disease risk mapping using Gaussian process stacked generalization. *J. R. Soc. Interface* **14**, 20170520 (2017).
- 32. Elith, J., Leathwick, J. R. & Hastie, T. A working guide to boosted regression trees. *J. Anim. Ecol.* **77**, 802–813 (2008).
- 33. Leathwick, J. R., Elith, J., Francis, M. P., Hastie, T. & Taylor, P. Variation in demersal fish species richness in the oceans surrounding New Zealand: an analysis using boosted regression trees. *Mar. Ecol. Prog. Ser.* **321**, 267–281 (2006).
- 34. Murray, C. J. L. *et al.* GBD 2010: design, definitions, and metrics. *Lancet* **380**, 2063–2066 (2012).
- 35. Stein, M. L. *Interpolation of Spatial Data*. (Springer New York, 1999).
- 36. Lindgren, F. & Rue, H. Bayesian Spatial Modelling with R-INLA. *J. Stat. Softw.* **63**, (2015).
- 37. Lindgren, F., Rue, H. & Lindström, J. An explicit link between Gaussian fields and Gaussian Markov random fields: the stochastic partial differential equation approach. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **73**, 423–498 (2011).
- 38. Rozanov, Y. A. *Markov Random Fields*. (Springer-Verlag, 1982).
- 39. Whittle, P. On Stationary Processes in the Plane. *Biometrika* **41**, 434–449 (1954).
- 40. Hotez, P. J. & Ribeiro, P. J. *Model-based Geostatistics*. (Springer New York, 2007). doi:10.1007/978- 0-387-48536-2.
- 41. Rue, H., Martino, S. & Chopin, N. Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **71**, 319–392 (2009).
- 42. Martins, T. G., Simpson, D., Lindgren, F. & Rue, H. Bayesian computing with INLA: new features. *Comput Stat Data Anal* **67**, 68–83 (2013).
- 43. Golding, N. *et al.* Mapping under-5 and neonatal mortality in Africa, 2000‒15: a baseline analysis for the Sustainable Development Goals. *Lancet* 390, 2171-2182 (2017).
- 44. Cameletti, M., Lindgren, F., Simpson, D. & Rue, H. Spatio-temporal modeling of particulate matter concentration through the SPDE approach. *AStA Adv. Stat. Anal.* **97**, 109–131 (2013).
- 45. Alegana, V. A. *et al.* Fine resolution mapping of population age-structures for health and development applications. *J. R. Soceity Interface* **12**, 20150073 (2015).
- 46. Kinyoki, D. K. *et al.* Assessing comorbidity and correlates of wasting and stunting among children in Somalia using cross-sectional household surveys: 2007 to 2010. *BMJ Open* **6**, e009854 (2016).
- 47. Patil, A. P., Gething, P. W., Piel, F. B. & Hay, S. I. Bayesian geostatistics in health cartography: the perspective of malaria. *Trends Parasitol.* **27**, 246–253 (2011).
- 48. Graetz, N. *et al.* Mapping local variation in educational attainment across Africa. *Nature* **555**, 48 (2018).
- 49. Wood, S. N. Package 'mgcv'. http://cran.r-project.org/web/packages/mgcv/mgcv.pdf (2019).
- 50. Bhatt, S. *et al.* The effect of malaria control on Plasmodium falciparum in Africa between 2000 and 2015. *Nature* **526**, 207–211 (2015).
- 51. WorldPop. WorldPop dataset. http://www.worldpop.org.uk/data/get_data/ (2017).
- 52. Tatem, A. J. WorldPop, open data for spatial demography. *Sci. Data* **4**, 170004 (2017).
- 53. GeoNetwork. The Global Administrative Unit Layers (GAUL). http://www.fao.org/geonetwork/srv/en/main.home (2015).
- 54. Global Administrative Areas (GADM). GADM database of global administrative areas. http://www.gadm.org (2018).
- 55. Land Processes Distributed Active Archive Center (LPDAAC). Combined MODIS 5.1 dataset. *MCD12Q1 | LP DAAC  :: NASA Land Data Products and Services*.
- 56. Lehner, B. & Döll, P. Development and validation of a global database of lakes, reservoirs and wetlands. *J. Hydrol.* **296**, 1–22 (2004).
- 57. World Wildlife Fund & Center for Environmental Systems Research, University of Kassel, Germany. Global Lakes and Wetlands Database, Level 3. https://www.worldwildlife.org/pages/global-lakesand-wetlands-database (2004).
- 58. Roberts, D. R. *et al.* Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure. *Ecography* **40**, 913–929 (2017).

8.0 Author Contributions

Managing the estimation or publications process

Damaris K Kinyoki, Lauren E Schaeffer, Laurie B Marczak, Megan F Schipp, Nicholas J Kassebaum, and Simon I Hay.

Writing the first draft of the manuscript

Damaris K Kinyoki, Lauren E Schaeffer, Laurie B Marczak, Megan F Schipp, and Simon I Hay.

Primary responsibility for this manuscript focused on: applying analytical methods to produce estimates

Damaris K Kinyoki, Aaron E Osgood-Zimmerman, Michael Collison, Nathaniel J Henry, and Natalia V Bhattacharjee.

Primary responsibility for this manuscript focused on: seeking, cataloguing, extracting, or cleaning data; designing or coding figures and tables

Damaris K Kinyoki, Brandon V Pickering, Alice Lazzar-Atwood, and Lucas Earl.

Managing the overall research enterprise

Laurie B Marczak, Lalit Dandona, Tsegaye Gebrehiwot, Aubrey J Levine, George Mensah, Ali Mokdad, Mohsen Naghavi, Benn Sartorius, Nicholas J Kassebaum, and Simon I Hay.

Providing data or critical feedback on data sources

Damaris K Kinyoki, Fares Alahdab, Mehran Alijanzadeh, Khalid Altirkawi, Catalina Liliana Andrei, Jalal Arabloo, Olatunde Aremu, Bahram Armoon, Marcel Ausloos, Marco Avila, Ashish Awasthi, Beatriz Paulina Ayala Quintanilla, Samad Azari, Neeraj Bedi, Bayu Begashaw Bekele, Boris Bikbov, Somayeh Bohlouli, Gabrielle Britton, Roy Burstein, Carlos Castañeda-Orjuela, Lucia Cuevas-Nasu, Lalit Dandona, Rakhi Dandona, Aso Darwesh, Rajat Das Gupta, Jan-Walter De Neve, Kebede Deribe, Manisha Dubey, Lucas Earl, Andem Effiong, Maysaa El Sayed Zaki, Anwar Faraj, Mohammad Fareed, Andre Faro, Irina Filip, Nataliya Foigt, Takeshi Fukumoto, Tsegaye Gebrehiwot, Alireza Ghajar, Sameer Gopalani, Ayman Grada, Arvin Haj-Mirzaian, Arya Haj-Mirzaian, Mario Herrero, Claudiu Herteliu, Long Hoang Nguyen, Michael Hole, Mehdi Hosseinzadeh, Guoqing Hu, Sheikh Mohammed Shariful Islam, Mihajlo Jakovljevic, Jost B Jonas, Jacek Jozwiak, Tanuj Kanchan, Manoochehr Karami, Amir Kasaeian, Yousef Khader, Ejaz Khan, Yun Jin Kim, Kewal Krishan, Barthelemy Kuate Defo, G Anil Kumar, Manasi Kumar, Sheetal D Lad, Huong Lan Nguyen, Paul Lee, Chi Linh Hoang, Reza Malekzadeh, Abdullah Mamun, Francisco Rogerlândio Martins-Melo, Anthony Masaka, Benjamin Massenburg, Benjamin Mayala, Fabiola Mejía-Rodriguez, Mulugeta Melku, Walter Mendoza, Tomasz Miazgowski, Erkin Mirrakhimov, Babak Moazen, Shafiu Mohammed, Ali Mokdad, Maziar Moradi-Lakeh, Seyyed Meysam Mousavi, Ulrich Mueller, Mehdi Naderi, Mohsen Naghavi, Ionut Negoi, Malihe Nourollahpour Shiadeh, Felix Ogbo, Andrew T Olagunju, Bolajoko Olusanya, Jacob Olusanya, Eduardo Ortiz-Panozo, Mahesh P A, Adrian Pana, Anamika Pandey, Sanghamitra Pati, Ellen G. Piwoz, Akram Pourshams, Hedley Quintana, Amir Radfar, Alireza Rafiei,

Fatemeh Rajati, David Rawaf, Salman Rawaf, Rahul Rawat, Carlos Rios-González, Leonardo Roever, Ali Rostami, Siamak Sabour, Nasir Salam, Abdallah M Samy, Benn Sartorius, Brijesh Sathian, David C Schwebel, Sadaf Sepanlou, Masood Ali Shaikh, Teresa Shamah-Levy, Diego Augusto Santos Silva, Dhirendra Narain Sinha, Agus Sudaryanto, Mohamad-Hani Temsah, Abdullah Terkawi, Zemenu Tessema, Andrew Thorne-Lyman, Marcos Roberto Tovani-Palone, Bach Tran, Irfan Ullah, Olalekan Uthman, Masoud Vaezghasemi, Pascual Valdez, Vasiliy Vlassov, Yasir Waheed, Gelin Xu, Tomohide Yamada, Naohiro Yonemoto, Mustafa Younis, Chuanhua Yu, Yunquan Zhang, Nicholas J Kassebaum, and Simon I Hay.

Development of methods or computational machinery

Damaris K Kinyoki, Michael Collison, Suleman Atique, Samad Azari, Natalia Bhattacharjee, Roy Burstein, Aso Darwesh, Aniruddha Deshpande, Lucas Earl, Anwar Faraj, Mehdi Hosseinzadeh, Ali Mokdad, Jonathan Mosser, Carlos Rios-González, Abdallah M Samy, Zemenu Tessema, John VanderHeide, Nicholas J Kassebaum, and Simon I Hay.

Providing critical feedback on methods or results

Damaris K Kinyoki, Lauren E Schaeffer, Michael Collison, Zegeye Abebe, Victor Adekanmbi, Keivan Ahmadi, Fares Alahdab, Mehran Alijanzadeh, Vahid Alipour, Rajaa Al-Raddadi, Khalid Altirkawi, Catalina Liliana Andrei, Carl Abelardo T Antonio, Jalal Arabloo, Olatunde Aremu, Suleman Atique, Marcel Ausloos, Ashish Awasthi, Beatriz Paulina Ayala Quintanilla, Samad Azari, Alaa Badawi, Kaleab Baye, Bayu Begashaw Bekele, Michelle Bell, Natalia Bhattacharjee, Krittika Bhattacharyya, Boris Bikbov, Somayeh Bohlouli, Gabrielle Britton, Roy Burstein, Zahid Butt, Ester Cerin, Dinh Toi Chu, Michael A Cork, Lalit Dandona, Rakhi Dandona, Farah Daoud, Aso Darwesh, Rajat Das Gupta, Diego De Leo, Jan-Walter De Neve, Kebede Deribe, Beruk Desalegn, Aniruddha Deshpande, Melaku Desta, Daniel Diaz, Manisha Dubey, Laura Dwyer-Lindgren, Andem Effiong, Maysaa El Sayed Zaki, Maha El Tantawi, Anwar Faraj, Mohammad Fareed, Andre Faro, Seyed-Mohammad Fereshtehnejad, Irina Filip, Florian Fischer, Nataliya Foigt, Morenike Folayan, Takeshi Fukumoto, Tsegaye Gebrehiwot, Kebede Embaye Gezae, Alireza Ghajar, Paramjit Gill, Philimon Gona, Sameer Gopalani, Ayman Grada, Yuming Guo, Arvin Haj-Mirzaian, Arya Haj-Mirzaian, Samer Hamidi, Bernardo Hernandez, Claudiu Herteliu, Long Hoang Nguyen, Michael Hole, Naznin Hossain, Mehdi Hosseinzadeh, Guoqing Hu, Sheikh Mohammed Shariful Islam, Mihajlo Jakovljevic, Ravi Prakash Jha, Jost B Jonas, Jacek Jozwiak, Amaha Kahsay, Tanuj Kanchan, Manoochehr Karami, Amir Kasaeian, Yousef Khader, Ejaz Khan, Mona Khater, Yun Jin Kim, Ruth Kimokoti, Sonali Kochhar, Hamidreza Komaki, Ai Koyanagi, Kewal Krishan, Barthelemy Kuate Defo, G Anil Kumar, Manasi Kumar, Sheetal D Lad, Huong Lan Nguyen, Qiangyi Li, Shanshan Li, Chi Linh Hoang, Rakesh Lodha, Marek Majdan, Reza Malekzadeh, Abdullah Mamun, Mohammad Ali Mansournia, Francisco Rogerlândio Martins-Melo, Anthony Masaka, Benjamin Massenburg, Benjamin Mayala, Fabiola Mejía-Rodriguez, Mulugeta Melku, Walter Mendoza, George Mensah, Tomasz Miazgowski, Ted R Miller, GK Mini, Erkin Mirrakhimov, Babak Moazen, Shafiu Mohammed, Farnam Mohebi, Ali Mokdad, Yoshan Moodley, Ghobad Moradi, Maziar Moradi-Lakeh, Paula Moraga, Shane Morrison, Jonathan Mosser, Seyyed Meysam Mousavi, Ulrich Mueller, Ghulam Mustafa, Mehdi Naderi, Mohsen Naghavi, Farid Najafi, Duduzile Ndwandwe, Ionut Negoi, Jing Nie, Chukwudi Nnaji, Malihe Nourollahpour Shiadeh, Peter

Nyasulu, Felix Ogbo, Andrew T Olagunju, Bolajoko Olusanya, Jacob Olusanya, Eduardo Ortiz-Panozo, Stanislav Otstavnov, Mahesh P A, Adrian Pana, Anamika Pandey, Sanghamitra Pati, George Patton, Norberto Perico, David Pigott, Meghdad Pirsaheb, Ellen G. Piwoz, Hedley Quintana, Amir Radfar, Alireza Rafiei, Vafa Rahimi-Movaghar, Rajesh Kumar Rai, David Rawaf, Salman Rawaf, Rahul Rawat, Giuseppe Remuzzi, Andre Renzaho, Carlos Rios-González, Leonardo Roever, Jennifer Ross, Ali Rostami, Siamak Sabour, Yahya Safari, Roya Safari-Faramani, Amirhossein Sahebkar, Payman Salamati, Yahya Salimi, Hamideh Salimzadeh, Abdallah M Samy, Benn Sartorius, Brijesh Sathian, David C Schwebel, Sadaf Sepanlou, Masood Ali Shaikh, Kiomars Sharafi, Rajesh Sharma, Aziz Sheikh, Jasvinder Singh, Dhirendra Narain Sinha, Agus Sudaryanto, Muawiyyah Babale Sufiyan, Rafael Tabarés-Seisdedos, Birkneh Tilahun Tadesse, Mohamad-Hani Temsah, Zemenu Tessema, Andrew Thorne-Lyman, Marcos Roberto Tovani-Palone, Bach Tran, Irfan Ullah, Olalekan Uthman, Masoud Vaezghasemi, Pascual Valdez, Yousef Veisani, Vasiliy Vlassov, Yasir Waheed, Yafeng Wang, Yuan-Pang Wang, Gelin Xu, Tomohide Yamada, Engida Yisma, Naohiro Yonemoto, Mustafa Younis, Chuanhua Yu, Telma Zahirian Moghadam, Sojib Bin Zaman, Mohammad Zamani, Yunquan Zhang, Nicholas J Kassebaum, and Simon I Hay.

Drafting the manuscript or revising is critically for important intellectual content

Damaris K Kinyoki, Lauren E Schaeffer, Victor Adekanmbi, Fares Alahdab, Catalina Liliana Andrei, Carl Abelardo T Antonio, Jalal Arabloo, Olatunde Aremu, Mehran Asadi, Marcel Ausloos, Ashish Awasthi, Quique Bassat, Kaleab Baye, Neeraj Bedi, Michelle Bell, Krittika Bhattacharyya, Zahid Butt, Carlos Castañeda-Orjuela, Ester Cerin, Dinh Toi Chu, Aso Darwesh, Rajat Das Gupta, Jan-Walter De Neve, Kebede Deribe, Beruk Desalegn, Melaku Desta, Daniel Diaz, Manisha Dubey, Laura Dwyer-Lindgren, Lucas Earl, Andem Effiong, Maysaa El Sayed Zaki, Maha El Tantawi, Anwar Faraj, Andre Faro, Seyed-Mohammad Fereshtehnejad, Irina Filip, Florian Fischer, Nataliya Foigt, Morenike Folayan, Takeshi Fukumoto, Tsegaye Gebrehiwot, Philimon Gona, Sameer Gopalani, Ayman Grada, Yuming Guo, Claudiu Herteliu, Long Hoang Nguyen, Michael Hole, Naznin Hossain, Mehdi Hosseinzadeh, Sheikh Mohammed Shariful Islam, Mihajlo Jakovljevic, Ravi Prakash Jha, Jost B Jonas, Jacek Jozwiak, Tanuj Kanchan, Manoochehr Karami, Amir Kasaeian, Yousef Khader, Ejaz Khan, Mona Khater, Yun Jin Kim, Hamidreza Komaki, Ai Koyanagi, Barthelemy Kuate Defo, Manasi Kumar, Huong Lan Nguyen, Paul Lee, Shanshan Li, Chi Linh Hoang, Shai Linn, Abdullah Mamun, Francisco Rogerlândio Martins-Melo, Anthony Masaka, Benjamin Massenburg, Mulugeta Melku, Walter Mendoza, Tomasz Miazgowski, Ted R Miller, Babak Moazen, Shafiu Mohammed, Farnam Mohebi, Ali Mokdad, Ghobad Moradi, Maziar Moradi-Lakeh, Paula Moraga, Shane Morrison, Jonathan Mosser, Mehdi Naderi, Ionut Negoi, Chukwudi Nnaji, Felix Ogbo, Andrew T Olagunju, Bolajoko Olusanya, Jacob Olusanya, Eduardo Ortiz-Panozo, Mahesh P A, Adrian Pana, Sanghamitra Pati, George Patton, David Pigott, Ellen G. Piwoz, Hedley Quintana, Amir Radfar, Alireza Rafiei, Vafa Rahimi-Movaghar, Rahul Rawat, Andre Renzaho, Carlos Rios-González, Leonardo Roever, Jennifer Ross, Siamak Sabour, Mahdi Safdarian, Nasir Salam, Hamideh Salimzadeh, Abdallah M Samy, Megan Schipp, David C Schwebel, Masood Ali Shaikh, Aziz Sheikh, Diego Augusto Santos Silva, Jasvinder Singh, Dhirendra Narain Sinha, Muawiyyah Babale Sufiyan, Mohamad-Hani Temsah, Zemenu Tessema, Marcos Roberto Tovani-Palone, Bach Tran, Khanh Bao Tran, Irfan Ullah, Olalekan Uthman, Masoud Vaezghasemi, Francesco S Violante, Vasiliy Vlassov, Yuan-Pang Wang, Gelin Xu, Mustafa Younis, Sojib Bin Zaman, Nicholas J Kassebaum, and Simon I Hay.