

Research Article

Small Area Estimation of Undernutrition under Age Five Children Based on Spatial-Temporal Models

Seyifemickael Amare Yilema ^{1,2}, Yegnanew A Shiferaw,³ Temesgen Zewotir,⁴ and Essey Kebede Muluneh⁵

¹Department of Statistics, College of Science, Bahir Dar University, P.O Box 79, Bahir Dar, Ethiopia

²Department of Statistics, College of Natural and Computational Science, Debre Tabor University, P.O Box 272, Debre Tabor, Ethiopia

³Department of Statistics, University of Johannesburg, P.O Box 524, Johannesburg 2006, South Africa

⁴Department of Statistics, School of Mathematics, Statistics and Computer Science, University of KwaZulu-Natal, Durban, South Africa

⁵School of Public Health, Bahir Dar University, P.O Box 79, Bahir Dar, Ethiopia

Correspondence should be addressed to Seyifemickael Amare Yilema; samarey1981@gmail.com

Received 17 January 2022; Revised 16 April 2022; Accepted 18 April 2022; Published 28 April 2022

Academic Editor: Mehrbakhsh Nilashi

Copyright © 2022 Seyifemickael Amare Yilema et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The mean flow of direct survey estimates is mainly concerning the sample adequacy fulfillment unless it has been produced large variance estimates, and therefore, the small area estimations are developed to manage this flaw of the path. Small area estimation improved the direct survey estimates by borrowing strength from the census data and at the same time by using historical data from consecutive surveys. In this paper, we applied the spatiotemporal Fay–Herriot (STFH) model for producing fairly reliable disaggregate-level estimates of undernutrition indicators across all zones. The STFH model is an appropriately fitted model to the undernutrition data since it has the lowest information criteria (IC) value. The spatiotemporal estimates improved both the direct and spatial estimates of undernutrition under the FH model and have brought efficiency gain in the percent coefficient of variation (CV). These results may provide useful information to the government's planners, policymakers, and legislative organs for effective policy formulation and budget allocation in all zones.

1. Introduction

The complete enumeration of surveys in geographically small areas with adequate sample sizes is too expensive and time-consuming; it is utterly unthinkable for developing countries like Ethiopia. It is known that censuses are conducted once in a decade, while surveys are conducted within five years of intervals. A survey is often planned to provide reasonable estimates at large geographical areas like national and regional levels [1]. However, the sample sizes are seldom large enough for small areas to produce direct estimates of adequate precision for the domains characteristics of interest [2]. Large estimation errors are produced in such cases, and the inferences are unreliable and useless for policymakers [1, 2].

The demand for small area statistics at disaggregated levels is increasing across the globe for policy interventions [1]. In addition to survey estimates produced efficiently at large scales (national and regional levels), these surveys also contribute to the country's economic, health, social, and political decisions, and policy implementations. The legislative organs of the government of Ethiopia ratified laws, implemented policies, and made political decisions are based on only the information received from the national and regional levels. The estimated information is not decentralized as the government's structures are decentralized. The accessibility of disaggregate-level statistics for target-oriented effective policy planning and monitoring is essential for Ethiopia's decentralized administrative planning systems.

Different researchers have suggested several model-based techniques of small area estimation to improve the direct survey estimates for domains with small samples sizes [3–5]. The ordinary Fay–Herriot (FH) model was developed and studied by [1, 6–9] to improve the small areas with small sample sizes. The spatial small area estimates are studied by [4, 10–13] with area correlations of the characteristics interest under the FH model. The need for spatial information in neighborhood areas is stated by [14] as “everything is related to everything else, but near things are more related than distant things.” Thus, closer areas tend to have similar socioeconomic characteristics of interest than distant small areas. Such extensions of the spatial FH model are further studied by [5, 10, 15].

Spatiotemporal small area estimation incorporates time-related historical data. It simultaneously includes the spatial correlation among data from the neighboring areas (zones in our cases) with fixed regression parameters across time.

In our study, we considered the zones to be small areas. This study focuses on child undernutrition indicators stunting, wasting, and underweight among children under five years of age. Among undernutrition indicators; stunting (height-for-age), underweight (weight-for-age), and wasting (weight-for-height) have been considered. Children whose height-for-age z-score is below minus two standard deviations (-2 SD) from the reference population’s median are considered stunted. Stunting is also called shortness, which means low height relative to age. Children whose weight for height z-score is below minus two standard deviations (-2 SD) are considered as wasted [16–18]. This study used the z-scores of the standard forms of stunting, wasting, and underweight as continuous variables to utilize the maximum amount of information available in the data set.

Globally, estimated 144 million and 47 million under five-year-old children were stunted and wasting, respectively, according to a research in [17]. The majority of the world’s stunted, underweight, and wasted children under the age of five lived in Asia and Africa [17]. Furthermore, undernutrition is associated with 45 percent of deaths in children under the age of five worldwide [17]. In Ethiopia, 38%, 10%, and 24% of children under the age of five were stunted, wasting, or underweight, respectively [18].

Decentralization is the most important administrative element in Ethiopian healthcare system [19]. Complementary to government institutions, the federal ministry of health decentralized the health service (regions, zones, and woredas). These administrative hierarchies are the most important entities in the country’s healthcare delivery [19, 20]. Between regional and woreda (district) governments, zonal governments act as a link (milestone). The health efforts in the districts are monitored and evaluated by the zonal health department [19]. As a result, estimates of undernutrition indicators at the zonal level are a considerable benefit for legislative bodies, policymakers, and monitors at all levels of government.

The focus of this paper is exploiting the spatial information obtained via the neighborhood area characteristics of interest for improving the direct survey estimates for unplanned domains, which are zones. Besides this, the

spatiotemporal model has been adopted to further improve the direct survey estimates by simultaneously incorporating the four years from 2000 to 2016 Ethiopian demographic and health survey (DHS) to strengthen the direct survey estimates of the last survey data. The Ethiopian DHS has been carried out within five-year intervals for the large geographical areas of regions and national levels. This study generated a spatiotemporal zonal level estimate using the surveys taken in 2000, 2005, 2011, and 2016. In this study, the researcher applied spatial FH and STFH models to obtain reliable and precise estimates of undernutrition (stunting, wasting, and underweight) by linking characteristics of interest from the 2000, 2005, 2011, and 2016 DHS data, which are considered to be temporal data, with the 2007 census data.

The remainder of the paper has been arranged as follows: in Section 2, the study discussed the methods and materials of the study, the spatial FH model with spatial correlation among the small areas (zones in our case), and the STFH model, which incorporates both area effects and time-related random effects. We report the results of the data analysis in Section 3 and discuss them in Section 4. Finally, we present conclusions in Section 5.

1.1. Literature Review. The spatiotemporal model in small area estimations is proposed by [5] and further studied by [21–23]. Using survey data from 2004 to 2008, a spatial-temporal Fay–Herriot model is applied with Spanish data to estimate poverty indicators for Spanish provinces in 2008 [5]. The spatial autocorrelation among the neighboring areas might be exploited to improve the direct survey estimates; however, incorporating time-related historical data further improves the direct survey estimates and spatial small area estimates [5, 21]. The spatiotemporal small area estimation is not studied in the country, yet it has not received attention in the undernutrition literature.

Estimation results for spatial and spatiotemporal small area models were compared in different types of literature [5, 21, 23, 24]. The spatiotemporal small area estimates of income in Poland data were applied by [25]. According to the studies in [25], the spatiotemporal models that used spatial correlation between neighboring areas as well as historical data were compared to EBLUPs based on spatial models derived separately for each year and with EBLUPs [5, 22, 23, 26]. The findings, the Polish data coming from the household budget survey and the administrative data, show that spatiotemporal small area models has been realized a noticeable reduction in estimation errors, especially when strong spatial and time autocorrelations were detected [25].

Spatiotemporal Fay–Herriot models are one of the SAE approaches that have incorporated spatial and time effects and have been utilized in poverty at the district level in west Sulawesi province [22]. The studies in [27] investigated area-level time models for small area estimate of poverty indicators and borrowed strength from time by employing area-level linear time models. The Spanish living conditions survey’s poverty indicators are evaluated using spatiotemporal models [27]. Spatiotemporal Fay–Herriot model

applied with Spanish EU-SILC data is carried out to obtain estimates of poverty indicators for Spanish provinces in 2008, making use of survey data from years 2004–2008 [21]. Spatiotemporal data contain a diverse set of variables, posing distinct problems and possibilities for professionals attempting to maximize its full potential.

2. Methods and Materials

2.1. Data Sources. The data were taken from the nationally representative cross-sectional study design of 2000, 2005, 2011, and 2016 Ethiopian DHS for the characteristics of interest and the 2007 census data for auxiliary variables. The Ethiopian DHSs are designed nationally representative, probabilistic, and household surveys that include a wide range of key demographic and health indicators used to monitor and evaluate population, health, and nutrition programs [18, 28–30]. The 1994 and 2007 population and housing census were considered for the sampling frame classifications, and the collection of the surveys is based on standardized questionnaires that yield different data files [18, 28–30]. Within 83 zones in Ethiopia, this study was conducted on under five-year-old children consisting of 8590 under five-year-old children from 2000, 3874 under five-year-old children from 2005, 9611 under five-year-old children from 2011, and (8505 stunting, 8675 wasting, and 8556 underweight) from 2016 Ethiopian DHS data. The height and weight measurements were collected from children 0–59 months [18, 28–30] in all the selected households for all survey years.

2.2. Study Variables. For this analysis, there are 41 area (zonal)-level proportions of covariates taken from the 2007 population and housing census. Stepwise regression analysis was used for all variables to filter out some of the best explanatory variables. Women aged 15–24, children aged 4–5, parents without disabilities, marital status (separated, widowed, divorced, and others), illiterate mothers, mothers with babies younger than one, and families with only one daughter who has died are selected for stunting under five-year-old children. Females, children under one year old, different marital statuses (separated, widowed, divorced, etc.), parents without disabilities, children ages 2–3, and mothers working for governmental organizations are targeted for wasting under age five children. Families with less than age five children, other marital statuses (separate, widowed, divorced, etc.), married, improved water facilities, and other occupations are selected for underweight children under five years of age.

2.3. Spatial Data. For administration purposes, Ethiopia has been divided into nine regions and two administrative cities, which, in turn, are divided into 83 zones. The global positioning system (GPS) point data were linked to each sampled urban-rural cluster residence to all household attributes. The GPS urban/rural locations have been masked [31] for confidentiality reasons. The GPS latitude/longitude position for DHSs data is randomly displaced to keep the

respondent’s confidentiality. In small area administrative units, the displacement is randomly carried out with 2 kilometers, and 5 kilometers for urban and rural residence clusters, respectively, and also 1% of rural clusters were displaced up to 10 kilometers [31]. The GPS point data as shapefiles are also available and obtained from <https://www.dhsprogram.com>. And also, the shapefiles for Ethiopian administrative boundaries are available on the website <https://africaopendata.org>.

2.4. Spatial and Spatiotemporal Small Area Estimation. The FH model has been extended in various works of literature. The multivariate FH models are investigated by multiple researchers [3, 6, 32–36]. The temporal FH model, which borrows strength from historical data, past time instants, and correlations, is studied by [27, 37] to produce reliable area-level estimates. The STFH model borrowed strength from census data and similar small areas through the time effect in historical data and spatial effects [5, 21, 23]. In the STFH model, the spatial and time-related dependencies have had between domains (zones in this case).

Let θ_{it} be the variable of characteristics of interest for area i and time t , where i, \dots, m and t, \dots, T . If the direct estimator of this quantity is denoted by $\hat{\theta}_{it}^{dir}$, the sampling errors can be expressed as ε_{it} . Similar with the FH model, the extended STFH model has two stages. The first stage can be expressed as follows:

$$\hat{\theta}_{it}^{dir} = \theta_{it} + \varepsilon_{it}, \tag{1}$$

where the sampling error ε_{it} is assumed to be independent and normally distributed with variance σ_{it}^2 known for all i and t .

The small area model, which incorporates the spatio-temporal relationships in the second stage, is given as

$$\theta_{it} = X_{it}^T + \nu_{1i} + \nu_{2i}, \tag{2}$$

where X_{it} is the vector of p auxiliary variables dependent linearly with θ_{it} for area i at time instant t , and β is the regression coefficients of auxiliary variables. Finally, the area effects vector $(\nu_{11}, \dots, \nu_{1m})^T$ follows a first-order simultaneous autoregressive process with variance parameter σ_1^2 , and spatial autocorrelation $|\rho_1| \leq 1$ and row-standardized proximity matrix W . And also the vectors of area-time random effects $(\nu_{21}, \dots, \nu_{2m})^T$ will follow identically and independently distributed for each area i and follow an first order autoregressive (AR(1)) with autocorrelation parameter $|\rho_2| \leq 1$, that is,

$$\nu_{2i} = \rho_2 \nu_{2i,t-1} + \varepsilon_{2it}, \tag{3}$$

where $\varepsilon_{2it} \sim ii \text{ d } N(0, \sigma_2^2)$.

The matrix notations of STFH general linear mixed model can be written as

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{v} + \boldsymbol{\varepsilon}. \tag{4}$$

Using stacking notations for vectors and matrices, the following relationship is considered as

$\mathbf{y} = \text{col}_{1 < i < m}(\text{col}_{1 < t < T}(\hat{\theta}_{it}))$, $\mathbf{X} = \text{col}_{1 < i < m}(\text{col}_{1 < t < T}(X_{it}))$, $\boldsymbol{\varepsilon} = \text{col}_{1 < i < m}(\text{col}_{1 < t < T}(\varepsilon_{it}))$, $\boldsymbol{\nu}_1 = \text{col}_{1 < i < m}(\text{col}_{1 < t < T}(\nu_{1i}))$, and $\boldsymbol{\nu}_2 = \text{col}_{1 < i < m}(\text{col}_{1 < t < T}(\nu_{2i}))$, where $\boldsymbol{\nu} = (\boldsymbol{\nu}_1^T, \boldsymbol{\nu}_2^T)^T$, and \mathbf{Z} is the constant unit matrix. Let the random component parameter $\boldsymbol{\tau} = (\sigma_1^2, \rho_1, \sigma_2^2, \rho_2)^T$ be the vector of unknown parameters involved in the covariance structure of the STFH model. The random sampling errors $\boldsymbol{\varepsilon} \sim N(0, \mathbf{V}_\varepsilon)$, where $\mathbf{0}$ vectors of zeros, and \mathbf{V}_ε is the diagonal matrix $\mathbf{V}_\varepsilon = \text{diag}_{1 < t < T}(\sigma_{it}^2)$. In addition, $\boldsymbol{\nu} \sim N(0, \mathbf{V}_\nu(\boldsymbol{\tau}))$ with covariance matrix given by the block diagonal matrix has the following form $\mathbf{V}_\nu(\boldsymbol{\tau}) = \text{diag}(\sigma_1^2 \Omega_1(\rho_1), \sigma_2^2 \Omega_2(\rho_2))$, the matrices Ω_1 and Ω_2 have the following relationships $\Omega_1 = [(I_m - \rho_1 W)^T (I_m - \rho_1 W)]^{-1}$ and $\Omega_2 = \text{diag}_{1 < i < m} \Omega_{2i}(\rho_2)$, and

$$\Omega_{2i}(\rho_2) = \frac{1}{1 - \rho_2^2} \begin{bmatrix} 1 & \rho_2 & \dots & \rho_2^{T-2} & \rho_2^{T-1} \\ \rho_2 & 1 & \dots & \rho_2^{T-3} & \rho_2^{T-2} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \rho_2^{T-1} & \rho_2^{T-2} & \dots & \rho_2 & 1 \end{bmatrix}. \quad (5)$$

Give all the above expressions, the covariance matrix for the full model (the sampling error plus the random components) can be written as $\text{var}(\mathbf{y}) = \mathbf{ZV}_\nu \mathbf{Z}^T + \mathbf{V}_\varepsilon$.

Note that the STFH model is a more general model and the ordinary spatial FH model can be obtained by ignoring the random time effects. When the spatial autocorrelation ρ_1 and area-time random effects ρ_2 are zero, the STFH model absolutely becomes the complementary FH model [21, 26].

2.4.1. Parameters Estimation of Spatiotemporal Model. Predicting and measuring the variability of the random components is one of the main issues in small area estimations. In this paper, we dealt with the problem of predicting θ_{it} by using empirical best linear unbiased prediction (EBLUP) of $\hat{\theta}_{it}$. By adopting the STFH model analogous to the Prasad and Rao [8,38], the mean square error (MSE) of EBLUP estimator under spatiotemporal FH model is as follows:

$$\text{MSE}(\hat{\theta}_{it}) = g_1(\hat{\boldsymbol{\tau}}) + g_2(\hat{\boldsymbol{\tau}}) + 2g_3(\hat{\boldsymbol{\tau}}), \quad (6)$$

where $\hat{\boldsymbol{\tau}} = (\hat{\sigma}_1^2, \hat{\rho}_1, \hat{\sigma}_2^2, \hat{\rho}_2)^T$ is the vector of estimated random variance components in the STFH model, $g_1(\hat{\boldsymbol{\tau}})$ is due to the estimation of random area effects with the order of $O(1)$ for large m , $g_2(\hat{\boldsymbol{\tau}})$ is due to the estimates of β with order $O(m^{-1})$, and the third term $g_3(\hat{\boldsymbol{\tau}})$ is due to the estimates of variance component.

The STFH model is fitted by restricted maximum likelihood (REML) methods, and also, the parametric bootstrap techniques are used for estimating the spatiotemporal EBLUP and its MSE [5, 21, 26]. The R package sae provides small area estimation methods based on the area-level models extended Fay–Herriot model, which allows for spatiotemporal correlation [39].

3. Result

3.1. Diagnostic Measures. The random component parameter estimator of the STFH model is reported in Table 1. The spatial correlations under the simultaneous autoregressive process are 96.5, 90.5, and 98 for stunting, wasting, and underweight, respectively. In addition, the time-related autoregressive are -63, -68.7, and -73.3 for stunting, wasting, and underweight, respectively. These results show that both the spatial area effect and the time-related random effects are in the FH model, and therefore, STFH model is appropriate model of this analysis. The random component variance of both the spatial and times effects is also presented in Table 1 for all target variables.

The relative best model was identified using the information criterion (IC) in Table 2. The IC (-2 log-likelihood (-2LL), Akaike information criterion (AIC), and Bayesian information criterion (BIC)) of the STFH model are smaller than the spatial Fay–Herriot model (SFH) for all undernutrition indicators. Thus, the STFH model is better than the SFH model.

The p -values for stunting, wasting, and underweight are a lot larger than 0.05 for the Kolmogorov–Smirnov test. Thus, we can conclude that the distribution of the STFH model for the undernutrition indicator does not differ significantly from a normal distribution. As a result, the STFH model meets its normality assumption satisfactorily.

Figure 1 shows the STFH model bias diagnostics measures. The direct survey estimates are plotted on the y -axis, and the spatiotemporal estimates are plotted on the x -axis. The diagnostic measure examines the validities of model-based spatiotemporal small area estimates. The regression of direct survey estimates is analogous to the spatiotemporal small area estimates since the spatiotemporal estimates are adjacent to the actual values. The graph shows that the model-based spatiotemporal small area estimates are not very far off from the fitted values of the regression line. As a result, the spatiotemporal estimates of small areas are not very different from direct survey estimates, indicating that the model-based estimates are valid. Overall, the bias diagnostic measures indicate that the model-based spatiotemporal small area estimates are likely to agree with direct survey estimates for all target variables of undernutrition indicators less than five years.

Figure 2 shows the STFH-based EBLUP estimates and the direct estimates. According to the figure, EBLUPs and direct survey estimates are equivalent, so EBLUPs are stable in all undernutrition indicators (stunting, wasting, and underweight).

The percent coefficient of variation of spatiotemporal small area estimates, spatial small area estimates, and direct estimates are presented in Figure 3. The percent coefficient of variation (CV) of spatiotemporal EBLUP is smaller than the corresponding direct estimates and spatial EBLUP estimates. From the figure, we observed that the CV (%) of direct survey estimates of undernutrition indicators is larger than the corresponding spatial small area estimates. The spatial small area estimates are more precise and reliable than the direct survey estimates, meaning that spatial estimates

TABLE 1: Parameters estimator of STFH model.

Parameters	Stunting	Wasting	Underweight
σ_1^2	0.00019	0.0026	0.0014
σ_2^2	0.0106	0.044	0.08
ρ_1	0.965	0.905	0.98
ρ_2	-0.63	-0.687	-0.733

TABLE 2: Model comparison based on information criteria.

IC	Spatial FH model			STFH model		
	Stunting	Wasting	Underweight	Stunting	Wasting	Underweight
-2LL	384.52	240.72	367.1	171.56	135.7	288.26
AIC	210.27	262.72	389.09	195.57	159.69	312.24
BIC	248.32	304.57	430.95	241.23	205.35	357.90

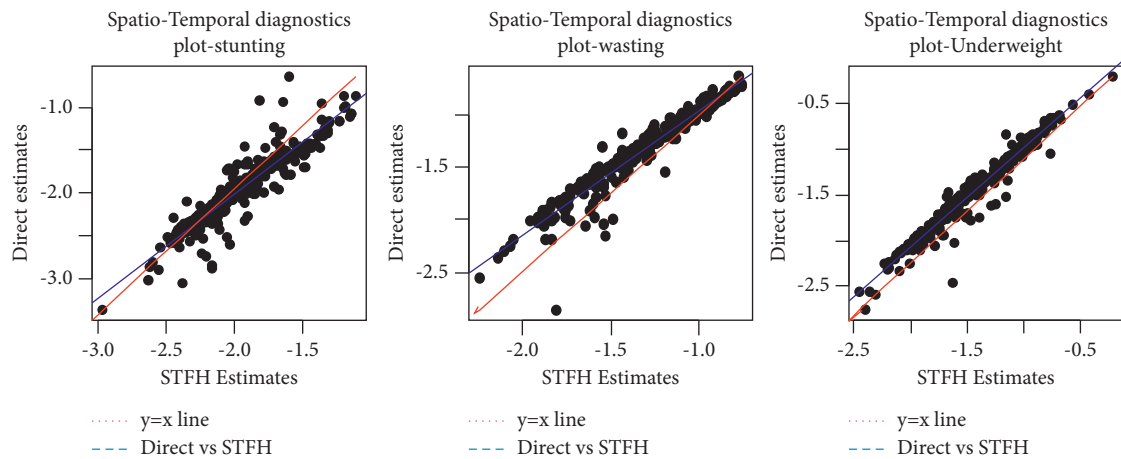


FIGURE 1: Bias diagnostic plot with $y = x$ line (red line) and regression line (blue line) for stunting, wasting, and underweight: Model-based STFH estimates versus direct estimates.

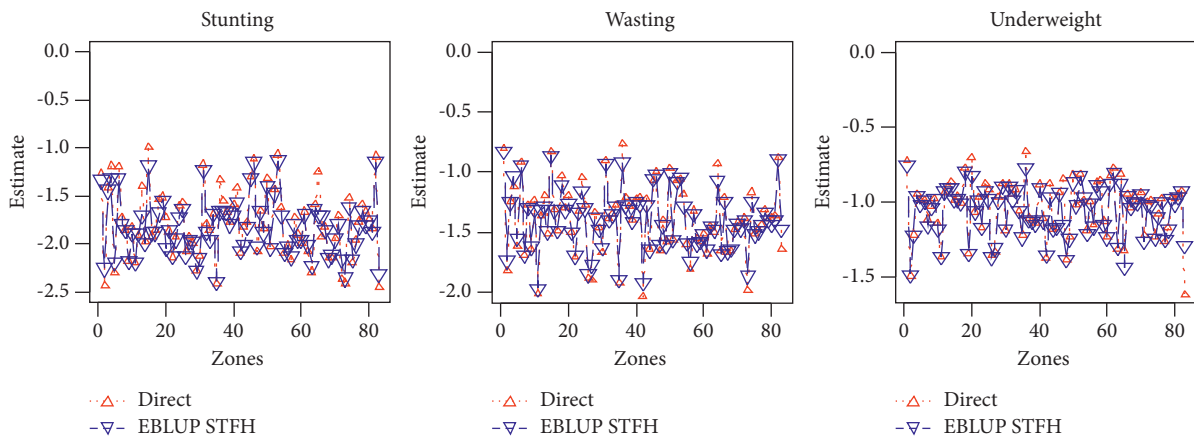


FIGURE 2: EBLUPs based on STFH model and direct estimates of children under five years of age for each zones.

improve the direct survey estimates because of the correlation among the neighboring zones.

Compared with the spatial estimates, the spatiotemporal estimates are deemed more precise and reliable. The

spatiotemporal small area estimates improve the precision of spatial estimates for all undernutrition indicators. In addition to being more precise than direct survey estimates, spatiotemporal small area estimates also improve them. As a

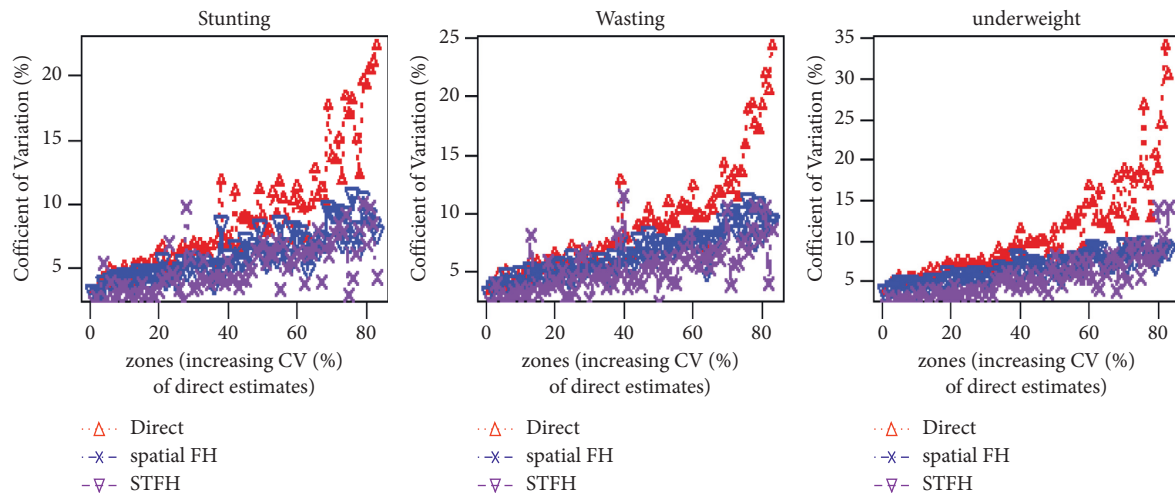


FIGURE 3: Zones percentage of coefficient of variation (CV) of direct, spatial FH, and STFH estimators of undernutrition indicators.

result, the direct survey estimates and spatial small area estimates are improved by spatiotemporal small area estimates. Generally, using temporal data (2000, 2005, 2011, and 2016 Ethiopian DHS) with fixed regression parameters from census data across all surveys improves the direct estimates and spatial small area estimates. Due to the inclusion of time-related autoregressive AR (1) correlations across small areas, STFH models have an advantage over spatial FH models.

Figure 4 show the zonal-wise root MSE of stunting, wasting, and underweight for children under five years of age. Direct survey estimates have the largest root MSE, while spatiotemporal EBLUPs have the smallest root MSE. Because the root MSE means are small, spatiotemporal estimates are the most precise, followed by spatial small area estimates. By contrast, direct survey estimates are the least accurate. Consequently, we can confirm that the spatiotemporal EBLUP is the most reliable and precise estimate of undernutrition indicators in children under age five.

The summary results of spatiotemporal EBLUP efficiency gain in CV over the spatial EBLUP and direct survey estimates are presented in Table 3. This table examined the magnitude to which the spatiotemporal small area estimates of stunting, wasting, and underweight improved in precision than the spatial EBLUP and direct survey estimates. The efficiency gains in CV due to spatiotemporal EBLUP over the direct and spatial EBLUP are improved for all target variables. Compared to direct survey estimates, the spatiotemporal EBLUP methods showed a median value of 53.74% with a maximum value of 90.34% for stunting, a median value of 49.60% with a maximum value of 91.30% for wasting, and a median value of 49.23% with a maximum value of 89.64% for underweight. The minimum, the first quartile, the mean, and the third quartiles of spatiotemporal EBLUP efficiency gain in CV over direct survey estimate for all undernutrition indicators are reported in Table 3. A few zones have a loss in efficiency for all target variables since the minimum values of gain in efficiency are recorded as negative values.

The spatiotemporal small area estimates efficiency gain over the spatial small area estimates are also investigated and reported in Table 3. The results of spatiotemporal EBLUP have a maximum value of 78.41%, 77.76%, and 68.77% efficiency gain for stunting, wasting, and underweight, respectively, over the corresponding spatial EBLUP. These results clearly show that the spatiotemporal small area estimates are more precise, efficient, and reliable than corresponding to the spatial small area for stunting, wasting, and underweight due to the incorporations of temporal effects on the spatially correlated zones. The spatiotemporal small area estimates improved the direct survey estimates and the spatially correlated zonal estimates. Therefore, the spatiotemporal small area estimates are the best reliable, precise, and efficient estimates for all undernutrition indicators.

4. Discussion

This part discussed the spatiotemporal small area estimates of undernutrition indicators: stunting, wasting, and underweight for children under age five. This article provides the zonal estimates of undernutrition indicators: stunting, wasting, and underweight for children under age five in Ethiopia using four consecutive surveys (2000, 2005, 2011, and 2016) and the 2007 population and housing census data. The standardized z-scores of undernutrition indicators, stunting, wasting, and underweight were used to exploit the maximum amount of information. The STFH model was applied to obtain zonal level estimates of undernutrition indicators in Ethiopia. It is a STFH model that accounts for the spatial correlation between neighboring areas and that simultaneously incorporates the time-related (four consecutive surveys from T time instants) to enhance small area estimates at the current time (in this case, the 2016 survey) [5, 21, 23].

The validity, reliability, and precision of model-based spatiotemporal small area estimates of undernutrition indicators were examined using bias diagnostics, Kolmogorov-Smirnov test, CVs, and root MSEs [2, 5, 21–23]. These

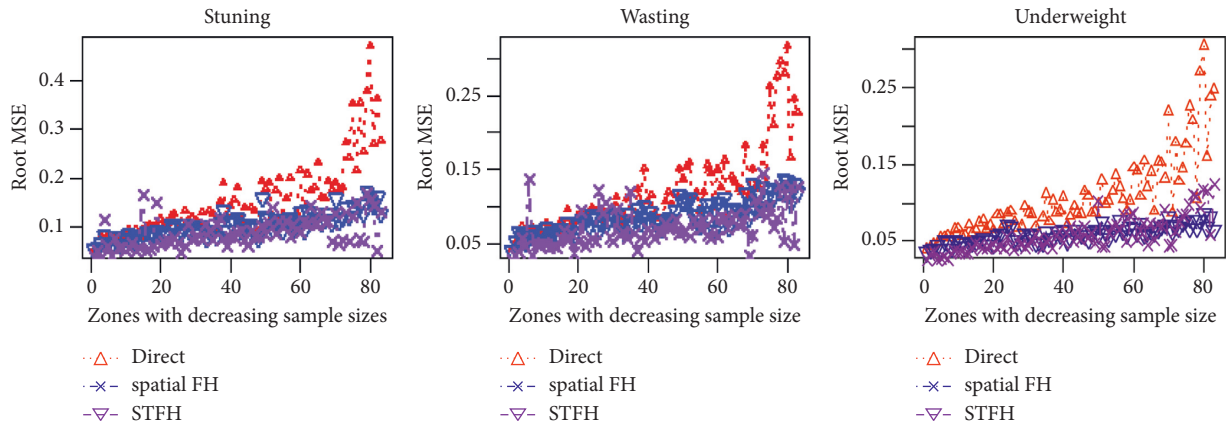


FIGURE 4: Zones (increasing sample size) Root MSE of direct and STFH estimators of stunting, wasting, and underweight.

TABLE 3: Summary statistics of efficiency gain in CV (%) for spatiotemporal EBLUPs over direct survey estimates and spatial EBLUPs.

	Efficiency gains of STF model over direct estimates			Efficiency gains of STFH model over the spatial FH model		
	Stunting	Wasting	Underweight	Stunting	Wasting	Underweight
Min	-132.84	-118.31	-117.60	-139.28	-145.30	-181.122
Q1	34.03	35.42	29.65	11.75	8.01	-10.44
Mean	47.61	43.96	39.39	25.68	20.39	4.79
Median	53.74	49.6	49.23	33.78	28.07	15.66
Q3	71.09	66.64	64.11	47.36	44.73	37.16
Max	90.34	91.30	89.64	78.41	77.76	68.77

measures indicate that the spatiotemporal small area estimates are superior to the direct survey estimates since the STFH model borrows strength from time-related temporal data from the four consecutive surveys (2000, 2005, 2011, and 2016) [5, 21–23]. In addition, the Kolmogorov–Smirnov test and the bias diagnostics of spatiotemporal small area estimates versus direct survey estimates (Figure 1) were used to test the model assumptions. Therefore, the model assumptions have been met satisfactorily.

For comparison, we computed the percentage CVs and root MSEs of direct survey estimates, spatial small area estimates, and spatiotemporal small area estimates of undernutrition indicators. In comparing the percentage CV of direct estimates and model-based spatiotemporal estimates (Figure 3), the spatiotemporal estimates were lower than the direct estimates for all indicators of undernutrition. Besides comparing spatiotemporal small area estimates with direct survey estimates, we also compared the percentage CV of the spatiotemporal estimates with the spatially correlated zonal estimates (Figure 3). Consequently, the spatiotemporal estimates have a lower CV than the spatial small area estimates. Similarly, the comparisons of root MSEs follow a CV-like approach.

As a result of the model assumptions and small area estimate diagnostic, the spatiotemporal small area estimates were more stable and precise than the corresponding direct survey estimates and spatial small area estimates for all undernutrition indicators of children under age five. And in turn, the spatial small area estimates have greater precision

and reliability than the direct survey estimates [4, 12, 40]. Direct survey estimates of undernutrition have improved using a spatial FH model with reasonably large spatial autocorrelations. And also, the spatiotemporal model further enhances the direct survey estimates, taking into account the temporal data considerations.

As a measure of the performance estimators under the most general model, the STFH model, efficiency gains in the CV of the STFH model over the spatial FH model, and direct estimates are computed. The average median improvements in the CV of the STFH model over the direct survey estimates are 53.74, 49.60, and 49.23 for stunting, wasting, and underweight, respectively (Table 3). Similarly, the average median efficiency gains in the CV of the STFH model over the spatial FH model are 33.78, 28.07, and 15.66 for stunting, wasting, and underweight, respectively (Table 3). Thus, it is evident that using temporal data in the STFH model offers advantages over spatial estimates of undernutrition indicators such as stunting, wasting, and underweight [5, 21].

5. Conclusion

This paper applied the STFH methods in four consecutive surveys and Ethiopia’s 2007 population and housing census data to improve the direct survey estimates of undernutrition indicators for children under age five across all zones. The CSA conducted regular surveys on several essential health, demographic, and socioeconomic indicators, but the results were limited to national and regional estimates. By

contrast, the administrative levels below regional levels are not examined due to small sample sizes. As this paper demonstrated, spatiotemporal small area estimations can be used as a cost-effective and efficient method for estimating undernutrition indicators.

The improvement of direct survey estimates and spatial small area estimates of zones are achieved in root MSE and CV for all target variables. It is evident that the use of temporal data in the ST FH model brings efficiency gain in CV over the spatial small area estimates of undernutrition indicators stunting, wasting, and underweight. Therefore, the use of temporal data is adequate. Consecutively, STFH models have an advantage over spatial FH models with temporal consideration due to the inclusion of time-related correlations across the zones. These results may provide useful information to the government's planners, policy-makers, and legislative organs for effective policy formulation and budget allocation in all zones.

Abbreviations

AIC:	Akaike information criteria
BIC:	Bayesian information criteria
CV:	Coefficient of variation
DHS:	Demographic and health survey
EBLUP:	Empirical best linear unbiased predictor
FH:	Fay–Herriot
GPS:	Global positioning system
IC:	Information criteria
LL:	Log likelihood
MSE:	Mean square error
REML:	Restricted maximum likelihood
SD:	Standard deviation
STFH:	Spatiotemporal Fay–Herriot.

Data Availability

The 2000–2016 Ethiopian DHS data were accessed after a request for registration from the DHS program website <https://www.dhsprogram.com>. Similarly, the GPS enumeration area shape files were obtained from <https://www.dhsprogram.com>, and also the shape files for Ethiopian zonal-level administrative boundaries are available on the website <https://africaopendata.org>. The Central Statistical Agency in Ethiopia provides access to 10 percent sample of the 2007 census for research purposes.

Conflicts of Interest

There are no conflicts of interest among the authors.

Authors' Contributions

SAY was involved in this study from data management, data analysis, drafting, and revising the final manuscript. YAS, TZ, and EKM contributed to the conception, design, and interpretation of data, as well as to manuscript reviews and revisions. All authors have read and approved the manuscript.

Acknowledgments

The authors thank Debre Tabor University, Debre Tabor, Ethiopia, and Bahir Dar University, Bahir Dar, Ethiopia.

References

- [1] J. N. K. Rao and I. Molina, *Small Area Estimation*, Wiley, Hoboken, 2nd edition, 2015.
- [2] S. A. Yilema, Y. A. Shiferaw, T. Zewotir, and E. K. Muluneh, "Multivariate small area estimation of undernutrition for children under five using official statistics," *Statistical Journal of the IAOS*, vol. 38, pp. 1–12, 2022.
- [3] R. Benavent and D. Morales, "Multivariate Fay-Herriot models for small area estimation," *Computational Statistics & Data Analysis*, vol. 94, pp. 372–390, 2016.
- [4] H. Chandra, "Exploring spatial dependence in area-level random effect model for disaggregate-level crop yield estimation," *Journal of Applied Statistics*, vol. 40, no. 4, pp. 823–842, 2013.
- [5] B. B. Singh, G. K. Shukla, and D. Kundu, "Spatio-temporal models in small area estimation spatio-temporal models in small area estimation," *Survey Methodology*, vol. 31, pp. 183–195, 2005.
- [6] R. E. Fay and R. A. Herriot, "Estimates of income for small places: an application of james-stein procedures to census data," *Journal of the American Statistical Association*, vol. 74, no. 366a, pp. 269–277, 1979.
- [7] I. Molina, G. S. Datta, and J. N. K. Rao, "Small area estimation under a Fay-Herriot model with preliminary testing for the presence of random area effects," *Survey Methodology*, vol. 41, pp. 1–19, 2015.
- [8] G. S. Datta and P. Lahiri, "A unified measure of uncertainty of estimated best linear unbiased predictor in small area estimation problems," *Statistica Sinica*, vol. 10, pp. 613–627, 2000.
- [9] G. S. Datta, J. N. K. Rao, and D. D. Smith, "On measuring the variability of small area estimators under a basic area level model," *Biometrika*, vol. 92, no. 1, pp. 183–196, 2005.
- [10] A. Petrucci, N. Salvati, A. Petrucci, and N. Salvati, "Small area estimation for spatial correlation in watershed erosion assessment," *Journal of Agricultural, Biological, and Environmental Statistics*, vol. 11, no. 2, pp. 169–182, 2006.
- [11] S. Guha and H. Chandra, "Measuring disaggregate level food insecurity via multivariate small area modelling: evidence from rural districts of Uttar Pradesh, India," *Food Security*, vol. 13, no. 3, pp. 597–615, 2021.
- [12] M. J. Hossain, S. Das, H. Chandra, and M. A. Islam, "Disaggregate level estimates and spatial mapping of food insecurity in Bangladesh by linking survey and census data," *PLoS One*, vol. 15, pp. e0230906–e0230916, 2020.
- [13] A. Risal, "Expenditure per capita model with spatial small area estimation," *Parameters: Journal of Statistics*, vol. 1, no. 2, pp. 38–47, 2021.
- [14] W. R. Tobler, "A computer movie simulating urban growth in the detroit region," *Economic Geography*, vol. 46, pp. 234–240, 1970.
- [15] A. Petrucci, M. Pratesi, and N. Salvati, "Geographic information in small area estimation: small area models and spatially correlated random area effects," *Stat Transit*, vol. 7, pp. 609–623, 2005.
- [16] M. De Onis, *WHO Child Growth Standards*, WHO, Geneva, Switzerland, 2006.

- [17] UNICEF WHO World Bank, "Levels and trends in child malnutrition: key findings of the 2020 edition of the joint child malnutrition estimates," *Geneva WHO*, vol. 24, pp. 1–16, 2020.
- [18] CSA and ICF, *Ethiopia Demographic Health Survey. Addis Ababa, Ethiopia, and Rockville, CSA and ICF, Maryland, USA*, 2016.
- [19] M. Woldie, C. Jirra, and G. Azene, "Presence and use of legislative guidelines for the distribution of decentralized decision making authority in the Jimma zone health system, Southwest Ethiopia," *Ethiopian journal of health sciences*, vol. 21, pp. 29–38, 2011.
- [20] Y. Kitaw, G.-E. Teka, H. Meche, H. Damen, and M. Fentahun, *The Evolution of Public Health in Ethiopia*, Ethiopian Public Health Association, Addis Ababa, Ethiopia, 2nd edition, 2012.
- [21] Y. Marhuenda, I. Molina, and D. Morales, "Small area estimation with spatio-temporal Fay-Herriot models," *Computational Statistics & Data Analysis*, vol. 58, pp. 308–325, 2013.
- [22] A. T. Rumiati and H. Kuswanto, "Spatio-temporal fay-herriot models in small area estimation to obtain factors that affecting poverty in polewali mandar district spatio-temporal fay-herriot models," *IOP Conference Series: Materials Science and Engineering*, vol. 546, no. 5, Article ID 052076, 2019.
- [23] J. Kubacki, "Small area estimation of income using spatio-temporal models," in *Proceedings of the 2nd Congress of Polish Statistics*, Warsaw, Poland, June 2018.
- [24] G. Vicente, T. Goicoa, and M. D. Ugarte, "Bayesian inference in multivariate spatio-temporal areal models using INLA: analysis of gender-based violence in small areas," *Stochastic Environmental Research and Risk Assessment*, vol. 34, no. 10, pp. 1421–1440, 2020.
- [25] A. Jędrzejczak and J. Kubacki, "Estimation of income characteristics for regions in Poland using spatio-temporal small area models," *Stat Transit*, vol. 20, 2019.
- [26] L. N. Pereira, "Small area estimation using a spatio-temporal linear mixed model," *REVSTAT-Statistical Journal*, vol. 10, pp. 285–308, 2012.
- [27] M. D. Esteban, D. Morales, A. Pérez, and L. Santamaría, "Area-level time models for small area estimation of poverty indicators," *Advances in Intelligent Systems and Computing*, vol. 77, 2010.
- [28] CSA, *Ethiopia Demographic and Health Survey. Addis Ababa, Ethiopia and Calverton*, Central Statistical Authority and ORC Macro, Maryland, USA, 2000.
- [29] CSA, *Ethiopia Demographic and Health Survey. Addis Ababa, Ethiopia and Calverton*, Central Statistical Agency and ORC Macro, Maryland, USA, 2005.
- [30] CSA, *Ethiopia Demographic and Health Survey. Addis Ababa, Ethiopia and Calverton*, Central Statistical Agency and ICF International, Maryland, USA, 2011.
- [31] C. R. Burgert, J. Colston, T. Roy, and B. Zachary, "Geographic displacement procedure and georeferenced data release policy for the demographic and health surveys," DHS Spatial Analysis Reports No. 7, ICF International, Maryland, USA, 2013.
- [32] S. Guha and H. Chandra, "Measuring and mapping disaggregate level disparities in food consumption and nutritional status via multivariate small area modelling," *Social Indicators Research*, vol. 154, no. 2, pp. 623–646, 2021.
- [33] T. Ito and T. Kubokawa, "Empirical best linear unbiased predictors in multivariate nested-error regression models," *Communications in Statistics - Theory and Methods*, vol. 50, no. 10, pp. 2224–2249, 2021.
- [34] I. Ngaruye, J. Nzabanita, D. v. Rosen, and M. Singull, "Small area estimation under a multivariate linear model for repeated measures data," *Communications in Statistics - Theory and Methods*, vol. 46, no. 21, pp. 10835–10850, 2017.
- [35] A. T. Porter, C. K. Wikle, and S. H. Holan, "Small area estimation via multivariate fay-herriot models with latent spatial dependence," *Australian & New Zealand Journal of Statistics*, vol. 57, no. 1, pp. 15–29, 2015.
- [36] A. Ubaidillah, K. A. Notodiputro, A. Kurnia, and I. W. Mangku, "Multivariate Fay-Herriot models for small area estimation with application to household consumption per capita expenditure in Indonesia," *Journal of Applied Statistics*, vol. 46, no. 15, pp. 2845–2861, 2019.
- [37] M. D. Esteban, D. Morales, A. Pérez, and L. Santamaría, "Small area estimation of poverty proportions under area-level time models," *Computational Statistics & Data Analysis*, vol. 56, no. 10, pp. 2840–2855, 2012.
- [38] N. G. N. Prasad and J. N. K. Rao, "The estimation of the mean squared error of small-area estimators," *Journal of the American Statistical Association*, vol. 85, no. 409, pp. 163–171, 1990.
- [39] I. Molina and Y. Marhuenda, "Sae: an R package for small area estimation," *The R Journal*, vol. 7, no. 1, p. 81, 2015.
- [40] H. Chandra, N. Salvati, and R. Chambers, "Small area estimation under a spatially non-linear model," *Computational Statistics & Data Analysis*, vol. 126, pp. 19–38, 2018.