Contents lists available at ScienceDirect

Heliyon

Research article

Unraveling birth weight determinants: Integrating machine learning, spatial analysis, and district-level mapping

Rubaiya, Mohaimen Mansur *, Md. Muhitul Alam, Md. Israt Rayhan

Institute of Statistical Research and Training, University of Dhaka, Bangladesh

A R T I C L E   I N F O

A B S T R A C T

Despite a decrease in the prevalence of low birth weight (LBW) over time, its ongoing significance as a public health concern in Bangladesh remains evident. Low birth weight is believed to be a contributing factor to infant mortality, prolonged health complications, and vulnerability to non-communicable diseases. This study utilizes nationally representative data from the Multiple Indicator Cluster Surveys (MICS) conducted in 2012-2013 and 2019 to explore factors associated with birth weight. Modeling birth weight data considers interactions among factors, clustering in data, and spatial correlation. District-level maps are generated to identify high-risk areas for LBW. The average birth weight has shown a modest increase, rising from 2.93 kg in 2012-2013 to 2.96 kg in 2019. The study employs a regression tree, a popular machine learning algorithm, to discern essential interactions among potential determinants of birth weight. Findings from various models, including fixed effect, mixed effect, and spatial dependence models, highlight the significance of factors such as maternal age, household head's education, antenatal care, and few data-driven interactions influencing birth weight. District-specific maps reveal lower average birth weights in the southwestern region and selected northern districts, persisting across the two survey periods. Accounting for hierarchical structure and spatial autocorrelation improves model performance, particularly when fitting the most recent round of survey data. The study aims to inform policy formulation and targeted interventions at the district level by utilizing a machine learning technique and regression models to identify vulnerable groups of children requiring heightened attention.

1. Introduction

Low birth weight (LBW), defined by the World Health Organization (WHO) as a birth weight less than 2500 grams, poses a considerable global public health challenge [1]. In the year 2020, 19.8 million newborns, accounting for approximately 14.7 percent of all babies born worldwide, experienced LBW [2]. This metric serves as a critical indicator of both the health condition of newborns and the complexities faced by mothers during pregnancy [3,4]. LBW is identified as a significant risk factor for mortality, particularly neonatal mortality, and contributes to higher morbidity and malnourishment in children. Infant mortality rates are six times higher among babies with LBW compared to those with normal birth weight [5,6]. Furthermore, LBW exposes newborns to an elevated risk of mortality within the initial 28 days of life and increases the likelihood of long-term complications, including challenges in physical development, cerebral palsy, blindness, deafness, developmental delay, and a heightened susceptibility to noncommunicable diseases

---

such as diabetes and cardiovascular conditions [7]. Moreover, LBW increases the risk of various developmental issues, including respiratory distress, brain haemorrhage, heart problems, and intestinal abnormalities that can significantly affect the infant's quality of life [7–10]. Children with low birth weight are also found to exhibit a higher prevalence of mental health issues in the later stages of their lives [11]. In response to the persistent global challenge of LBW, the World Health Assembly (WHA) approved a Comprehensive Implementation Plan on Maternal, Infant, and Young Child Nutrition in 2012. This initiative established six global nutrition targets, including a goal to reduce the count of LBW live births by 30% from 2012 to 2025 [12]. While there has been progress in South Asia, where the occurrence of LBW decreased from 32.3% to 26.4% between 2000 and 2015, this rate of reduction falls short of the objectives set by the WHA [13]. Addressing the complexities of LBW remains an ongoing priority to ensure healthier outcomes for both mothers and newborns worldwide.

Extensive research has been conducted globally to explore the factors associated with LBW. Numerous studies have highlighted the variability of LBW, manifesting differences not only within individual countries but also across nations, a pattern often interconnected with the developmental status of the respective regions [14–18]. Utilizing machine learning techniques, it has been unveiled that crucial factors correlated with LBW encompass gestational age, count of abortions, and maternal age during delivery [19].The consideration of interactions among factors emerges as a crucial aspect in studying the variability in LBW. For example, rather than separately assessing birth weight for mothers with lower educational attainment and those in economically disadvantaged groups, a more advantageous approach may involve investigating whether birth weights significantly differ among children born to mothers with lower educational levels and belonging to low-wealth households. This exploration can be best conducted by considering the interactions between maternal education and wealth. Neglecting interactions and focusing solely on main effects can lead to overly generalized conclusions regarding the impact of individual risk factors on LBW [20].

The high prevalence of low birth weight is a long-standing public health problem in Bangladesh, with almost 15% of the newly born at risk of being underweight [21]. While there has been a decline in the LBW rate in Bangladesh from 2011 to 2017, it remains high compared to many other developing nations [22]. Several studies have examined socio-economic and demographic factors associated with LBW in recent years. Factors such as maternal socio-economic status, frequency of antenatal care utilization, delivery location, and the quality of delivery assistance emerge as significant covariates associated with LBW. The advancement of economic conditions, intertwined with the elevation of women's educational attainment, remains a central focus in addressing the prevalent incidence of LBW [23–25]. Certain studies have additionally demonstrated that a correlation exists between the maternal history of encountering intimate partner violence (IPV) throughout their lifetime and the chance of delivering infants with LBW [26]. Within the specific context of Bangladesh, spatial interdependence is observable in the prevalence of LBW, with selected districts in the northern and southern regions facing an elevated risk [27].

Although numerous studies have highlighted the intra-country variability in LBW prevalence, many investigations conducted in Bangladesh did not focus in identifying the regions with heightened risk [14,15]. Moreover, despite several studies illustrating the significance of considering interactions, most research endeavors within Bangladesh have overlooked this aspect [20]. Additionally, most studies have not implemented spatial adjustments [28]. Among the few that did, the analysis was performed at the district level, potentially leading to information loss compared to an analysis conducted at the household level [27]. Furthermore, we posit that a more nuanced approach lies in studying birth weight rather than dichotomizing it into LBW. This allows for a comprehensive examination of factors contributing to reduced birth weights, even on a finer scale.

The study aims to leverage data from nationally representative surveys to reveal complex interactions among crucial socio-economic determinants, enhancing the understanding of birth weight. By integrating identified interactions and accounting for clustering and spatial correlation, the study seeks to generate model-based estimates for birth weight. Additionally, the research aims to provide district-level estimates of average birth weights to reveal districts at higher risk of LBW.

2. Methods

2.1. Source of data and sample size

The study data are derived from the Bangladesh Multiple Indicator Cluster Survey (MICS) conducted during two periods: MICS 2012-2013, held from December 2012 to April 2013, and MICS 2019, carried out between 19 January and 1 June 2019. These surveys were administered by the Bangladesh Bureau of Statistics (BBS) under the Ministry of Planning [29,30]. Employing a two-stage stratified cluster sampling method, administrative districts served as strata during the sample selection process. In the MICS 2012-2013, a total of 55,200 sample households from 2760 sample clusters in 64 districts were incorporated. Our study focused on a subset of 2319 mothers who provided birth weight information out of the 7866 women (aged 15-49 years) with recent live births. Similarly, the MICS 2019 survey included 64,400 sample households from 2586 sample clusters in 64 districts. Among the 9210 women (aged 15-49 years) with recent live births, our analysis concentrated on the sub-sample of 4652 mothers providing birth weight information. Fig. 1 shows how the final sample was obtained.

2.2. Outcome variable

The outcome variable of the study is birth weight which is a continuous variable measured in Kg.
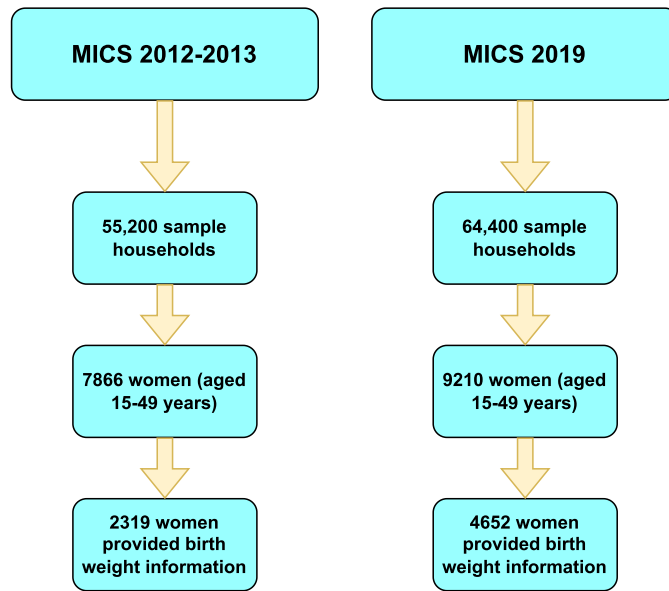
Fig. 1. Final sample size for MICS 2012-2013 and MICS 2019 in this study.

2.3. Independent variables

Following an extensive literature review, specific explanatory variables have been chosen as potential factors associated with birth weight. These variables are: mother's age in years ("≤20", "21-30", "30+"), mother's education ("secondary complete or higher", "other"), household heads education ("secondary complete or higher", "other"), received anti-natal care ("yes", "no"), number of anti-natal care visits ("1", "2", "3+"), place of delivery ("home", "other"), assistant during delivery ("doctor/nurse/midwife/auxiliary midwife", "other"), delivery by caesarean method ("yes", "no"), parity ("1 (Primiparous)", "1+"), wealth index ("low", "middle", "high"), area of residence ("urban", "rural"), division ("Barisal", "Chittagong", "Dhaka", "Khulna", "Rajshahi", "Rangpur", "Sylhet")

2.4. Statistical analysis

The analysis begins by implementing a regression tree on MICS data for identifying potential interactions among birth weight predictors. Renowned in the machine learning literature, the regression tree is a non-parametric method known for capturing interactions among variables in a data-driven manner [31]. The tree mechanism entails recursively partitioning the predictor space into several small, non-overlapping regions based on simple rules. The mean outcome (e.g., birth weight) within a region serves as the predicted value for a new observation in that specific region. Recursive binary splitting is utilized to construct the regression tree, with decision rules for splitting, the sequence of crucial predictors, and their interactions succinctly summarized in a visually intuitive manner.

A regression tree presents various advantages over parametric regression models, characterized by minimal distributional assumptions and ease of interpretation. Its most notable strength lies in its inherent capacity to capture complex interactions, including non-linear ones, among variables in high-dimensional settings. Parametric regression models prove less adept at capturing and analyzing interactions, particularly when the crucial ones are unknown in advance. The potential number of interaction terms tends to escalate with the inclusion of variables in a classical regression model, leading to the curse of dimensionality [32]. This, in turn, results in over-parameterization of models and subsequent challenges in estimation.

It is crucial to note that continual splitting in a regression tree can lead to overfitting data, resulting in poor outcome predictions. To mitigate such overfitting issues and ensure statistically valid estimations, we establish a stopping rule for tree-growing by setting the minimum sample size in each node to be 50.

In the subsequent phase of the analysis, four distinct linear regression models are employed to identify significant covariates associated with birth weight. The important rules (interactions) extracted from exploratory regression trees are turned into binary variables (1, if defined by the rule and 0, otherwise) and included in the set of explanatory variables for each model. Given that MICS data exhibit a built-in multilevel structure with individuals nested under geographical divisions (clusters), it becomes essential to address potential cluster-induced correlations when modeling children's low birth weight. Additionally, accounting for potential spatial correlations that reflect similarities in estimates among neighboring districts is of equal significance. The specifications of the regression models vary depending on the levels of such complexities they account for.

The first model is called a fixed effect (FE) model, which allows no clustering or spatial effect and considers only fixed covariate effects. Let y_{ij} is the birth weight of an individual j in district $i : i = 1, 2, \ldots, 64$. This model links the mean response and p potential predictors $(x_{ij1}, x_{ij2}, \ldots, x_{ijp})'$, including interactions, using an identity-like function.

FE model: $E(y_{ij}) = \beta_0 + \sum_{k=1}^{p} \beta_k x_{ijk}$,

where $\{\beta_k\}_{i=1}^{p}$ are the regression coefficients for observed predictors and interaction variables obtained from panel regression tree.

Some unobserved predictors were not considered in the FE model. Some of these predictors are deep-rooted within the districts and induce correlations among individuals from the same district having some common characteristics. The mixed effect (ME) model includes an identically independent distributed (IID) spatially unstructured random effect (u_i) for within-district correlation to account for the effect of some unobserved predictors. With unstructured random effect u_i at the district level, the FE model can be extended to the ME model with an identity link function:

ME model: $E(y_{ij}) = \beta_0 + \sum_{k=1}^{p} \beta_k x_{ijk} + u_i$,

where u_i's are independent, identically and normally distributed with mean 0 and constant variance σ_u^2.

Again, some unobserved predictors vary spatially across the districts and prompt spatial auto-correlation due to geographical proximity. If two districts are neighbors, then observed predictors and responses may have similar values in those districts. The Besag model includes a spatially structured random component (s_i) to account for spatial variations with the fixed predictors [33].

Besag model: $E(y_{ij}) = \beta_0 + \sum_{k=1}^{p} \beta_k x_{ijk} + s_i$,

where the probabilistic behavior of $s = \{s_i\}_{i=1}^{64}$ are defined by specifying the distribution of s_i conditional on the set $s_{-i} = \{s_{i'}\}_{i' \neq i}$. A widely used scheme of representation for such conditional distribution under irregularly shaped areas is the intrinsic conditional autoregressive (ICAR) model, where the conditional distribution of s_i given s_{-i} is given by

$s_i|s_{-i} \sim N(\bar{s}_i, \frac{\sigma_s^2}{w_i^+})$,

where $\bar{s}_i = \sum_j \frac{w_{ij}s_j}{w_i^+}$, $w_i^+ = \sum_{j \neq i} w_{ij}$ and $\{w_{ij}\}_{i,j=1}^{64}$ is a 64×64 matrix which represents the spatial proximity structure such that $w_{ij} = 1$ suggesting the i-th and j-th districts are the neighbors if they share a common boundary, and zero otherwise. Setting $w_{ii} = 0$ ensures that a location will not be used to predict itself. σ_s^2 is the conditional variance that determines the amount of spatial variation [34–36].

The Besag model assumes only a spatially structured component, and the existence of unstructured random effect within area i is modeled as spatial correlation, which leads to a biased estimate of the model parameters [37]. Besag-York-Mollie (BYM) had addressed this issue [34]. The BYM model allows convolution of both an unstructured (IID component) and a structured spatial component as a random effect, which is given by

BYM model: $E(y_{ij}) = \beta_0 + \sum_{k=1}^{p} \beta_k x_{ijk} + s_i + u_i$.

We utilized the Integrated Nested Laplace Approximation (INLA) technique using the R package R-INLA. This approach, which offers a computationally efficient and rapid Bayesian approximation, was employed to estimate parameters within the Besag and BYM models [38,39]. For the fixed effect parameters β's, it is established in the literature that the β's are well estimated, and the associated bias is less sensitive to the prior choice. Therefore, for a complete specification of the Bayesian framework, an independent zero-mean normal prior with large variance was considered for each β_k. However, special care is required for the variance parameters of the random effects. The $gamma(\epsilon, \epsilon)$ priors are the popular choices in some software (WinBUGS, OpenBUGS, JAGS, R-INLA). For example, INLA specifies $gamma(1, 0.001)$ and Win-BUGS specifies $gamma(0.01, 0.01)$ prior as default prior for the precision [40].

To identify the most suitable model for the data, a comparison of all these models will be conducted using the mean squared error (MSE), deviance information criterion (DIC), widely applicable Bayesian information criterion (WAIC), conditional predictive ordinate (CPO) and k-fold cross-validation. The optimal fitted model is determined by selecting the one with the lowest DIC, WAIC, MSE, $-\sum \log(CPO_i)$ and cross-validation error [41–43].

2.5. Ethical approval

This study made use of data from the publicly accessible Multiple Indicator Cluster Survey (MICS) conducted at two distinct periods. The dataset can be accessed via the following link: https://mics.unicef.org/. As such, this study did not necessitate supplementary ethical approval, given that its analysis relied on freely available secondary data.

Table 1
Summary of response.

Variable	MICS 2012 - 2013					MICS 2019				
	1st Q.	SD	Mean	Median	3rd Q.	1st Q.	SD	Mean	Median	3rd Q.
Birth Weight	2.50	0.76	2.93	3.00	3.50	2.50	0.63	2.96	3.00	3.50

3. Results

3.1. Socio-economic and demographic characteristics

The summary of the response variable, birth weight (in Kg), is presented in Table 1. While the mean birth weight rose from 2.93 Kg in 2012-2013 to 2.96 Kg in 2019, the first, second, and third quartiles remained constant.

Table 2 shows the distribution of respondents by the explanatory variables and the mean birth weight under different categories of the variables. In both rounds of the survey, most of the mothers belonged to the age bracket of 21-30 years and had relatively higher mean birth weights of babies. Merely 28% of the mothers attained secondary or higher level education in 2012-2013, a proportion that was much higher (82%) in 2019. Across both datasets, mothers with higher educational backgrounds exhibited noticeablyhigher average birth weights. A similar pattern in the distribution of respondents and birthweight can also be observed in the case of the education of household heads.

In the sample data, 87% and 95% of mothers received antenatal care in 2012-2013 and 2019, respectively. In both time periods, this group exhibited an average birth weight of nearly 3 kg, which marginally surpasses that of women who did not receive antenatal care. Moreover, most of the respondents sought ANC more than twice and demonstrated higher average birth weight compared to their counterparts who had only one ANC visit. The proportion of mothers who delivered outside their homes was substantially higher (by nearly 22%) in 2019 than in 2012-2013. During this period, there was a notable 20% increase in the proportion of mothers who opted for caesarean delivery. Notably, across both periods, these women exhibited a higher average birth weight in comparison to those who did not undergo caesarean delivery.

For mothers belonging to the low-wealth class, the mean birth weight stood at 2.81 kg and 2.95 kg in 2012-2013 and 2019, respectively, indicating values below the national average. In the year 2012-2013, the district Rangpur exhibited the lowest average birth weight at approximately 2.77 kg. However, in 2019, this unfortunate distinction shifted to the region of Sylhet, which has an average birth weight of nearly 2.83 kg.

3.2. Regression tree

Fig. 2 shows the results from the regression tree for MICS 2012-2013 data. The regression tree attempts to discriminate groups of mothers whose offspring exhibit the lowest birth weight from those with relatively higher birth weight through the partitioning of the socio-demographic predictors. Notably, the occurrence of delivery via cesarean section constituted the first and foremost split at the root and, thereby, emerged as the principal predictor influencing low birth weight. Children born through cesarean delivery were allocated to the left branch of this primary classification. On this particular branch of the tree, parity appeared to be the second most influential predictor. Children of primiparous mothers weighed, on average, 3.13 kg, a considerably higher birth weight compared to the nationwide average of 2.93 kg. Offsprings who were not delivered through cesarean section were grouped under the right branch at the root of the tree. Within this branch, wealth was identified as the second most prominent determinant. Children originating from households with low and middle income exhibited an average birth weight of approximately 2.75 kg, a figure lower than the nation's average. Moreover, among the higher wealth strata, a further division becomes apparent with respect to parity. Specifically, children born to primiparous mothers from households with higher income levels had an average birth weight of roughly 2.78 kg, representing a marginal increment in comparison to the previous branch. We identified five interactions among socio-demographic characteristics that were derived from the tree structure. These interactions may demonstrate potentially significant associations with birth weight. The interactions are represented as rules below:

Rule 1: Delivered by cesarean {yes} and parity {1} (average birth weight 3.13 kg, relevant node 3)
Rule 2: Delivered by cesarean {yes} and parity {1+} (average birth weight 2.95 kg, relevant node 4)
Rule 3: Delivered by cesarean {no} and wealth index {low, middle} (average birth weight 2.75 kg, relevant node 6)
Rule 4: Delivered by cesarean {no}, wealth index {high} and parity {1} (average birth weight 2.78 kg, relevant node 8)
Rule 5: Delivered by cesarean {no}, wealth index {high} and parity {1+} (average birth weight 3.04 kg, relevant node 9)

Fig. 3 depicts the outcomes obtained from the regression tree analysis conducted on the MICS 2019 dataset. Similar to the findings in the MICS 2012-2013 study, the incidence of a cesarean section emerged as the predominant predictor exerting influence on birth weight, thus representing the initial split. Offsprings delivered through the cesarean method were assigned to the left branch of the classification. Within this specific branch of the tree, parity emerged as the second most influential predictor, constituting the next split. A third split based on the education of the household head occurred among primiparous mothers (parity 1). Thus, mothers who underwent cesarean deliveries were primiparous and belonged to households led by heads with below secondary level education, exhibiting an average birth weight of 2.97 kg, a value closely approximating the national average. In the context of higher parity mothers and maternal residence in urban areas, an additional split was observed based on maternal education. Mothers who experienced cesarean deliveries, had higher than unit parity, resided in urban areas, and acquired primary or lower level education

Table 2
Frequency distribution of explanatory variables with percentages and mean of birth weights.

	MICS 2012-2013		MICS 2019	
Variables	Frequency (%)	Mean Birth Weight	Frequency (%)	Mean Birth Weight
Mothers Age				
< = 20	569 (22)	2.88	1028 (22)	2.93
21-30	1604 (63)	2.95	2767 (59)	2.98
30+	383 (15)	2.89	887 (19)	2.95
Mothers Education				
Other	1838 (72)	2.89	844 (18)	2.91
Secondary or Higher	719 (28)	3.01	3838 (82)	2.98
Head Education				
Other	1875 (73)	2.92	2249 (48)	2.93
Secondary Complete or Higher	679 (27)	2.93	2432 (52)	3.00
ANC				
Yes	2225 (87)	2.95	4433 (95)	2.97
No	332 (13)	2.78	249 (5.3)	2.86
ANC Times				
1	239 (9.3)	2.85	339 (7.2)	2.94
2	378 (15)	2.94	755 (16)	2.94
3+	1939 (76)	2.93	3588 (77)	2.97
Delivery Place				
Home	752 (29)	2.78	305 (6.5)	2.99
Other	1804 (71)	2.99	4376 (93)	2.96
Delivery Assist				
Doctor/Nurse/Midwife/ Auxiliary Midwife	1950 (76)	2.97	4361 (93)	2.97
Other	607 (24)	2.78	321 (6.9)	2.94
Caesarean				
Yes	1184 (46)	3.05	3086 (66)	2.99
No	1372 (54)	2.82	1596 (34)	2.91
Parity				
1 (Primiparous)	1169 (46)	2.95	1992 (43)	2.98
1+	1387 (54)	2.91	1690 (57)	2.95
Wealth Index				
Low	723 (28)	2.81	1169 (25)	2.95
Middle	395 (15)	2.88	894 (19)	2.96
High	1438 (56)	3.00	2619 (56)	2.99
Area				
Urban	786 (31)	2.97	1307 (28)	2.93
Rural	1771 (69)	2.91	3374 (72)	2.98
Division				
Barisal	69 (2.7)	2.89	180 (3.8)	3.00
Chittagong	354 (14)	2.90	892 (19)	2.90
Dhaka	940 (37)	2.94	1347 (29)	2.92
Khulna	328 (13)	3.00	644 (14)	3.05
Rajshahi	276 (11)	3.09	571 (12)	3.05
Rangpur	502 (20)	2.77	544 (12)	3.03
Sylhet	86 (3.4)	2.97	256 (5.5)	2.83
Mymenshing			248 (5.3)	2.97

gave birth to infants with an average weight of 2.66 kg, a weight significantly lower than the national average of 2.96 kg. While maintaining all prior splits unchanged, a modification limited to the maternal education level being secondary or higher resulted in an average birth weight of 2.92 kg for children. Six interactions from the regression tree are listed below:

Rule 1: Delivered by cesarean {yes}, parity {1} and household heads education {other} (average birth weight 2.98 kg, relevant node 4)

Rule 2: Delivered by cesarean {yes}, parity {1} and household heads education {secondary or higher} (average birth weight 3.01 kg, relevant node 5)

Rule 3: Delivered by cesarean {yes}, parity {1+}, area of residence {urban} and mothers education {other} (average birth weight 2.66 kg, relevant node 8)

Rule 4: Delivered by cesarean {yes}, parity {1+}, area of residence {urban} and mothers education {secondary or higher} (average birth weight 2.92 kg, relevant node 9)

Rule 5: Delivered by cesarean {yes}, parity {1+}, area of residence {rural} (average birth weight 2.99 kg, relevant node 10)

Rule 6: Delivered by cesarean {no} (average birth weight 2.91 kg, relevant node 11)
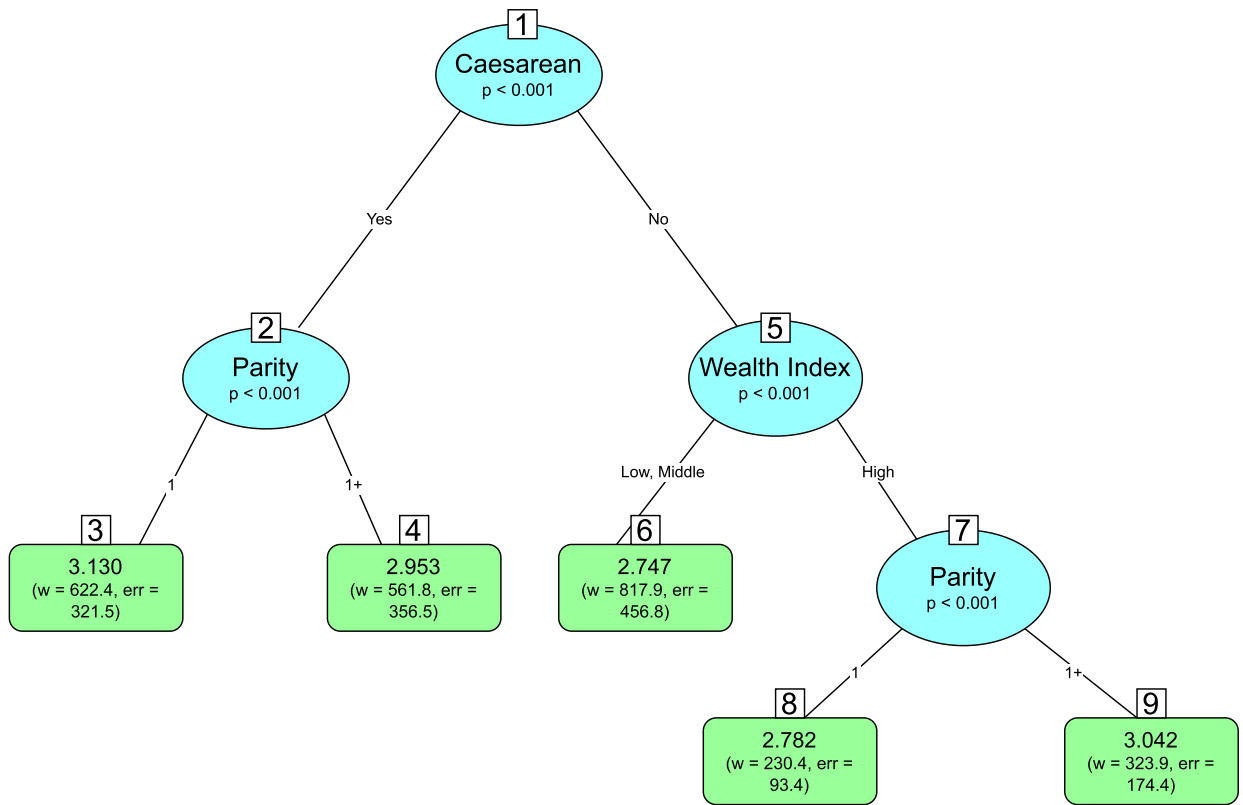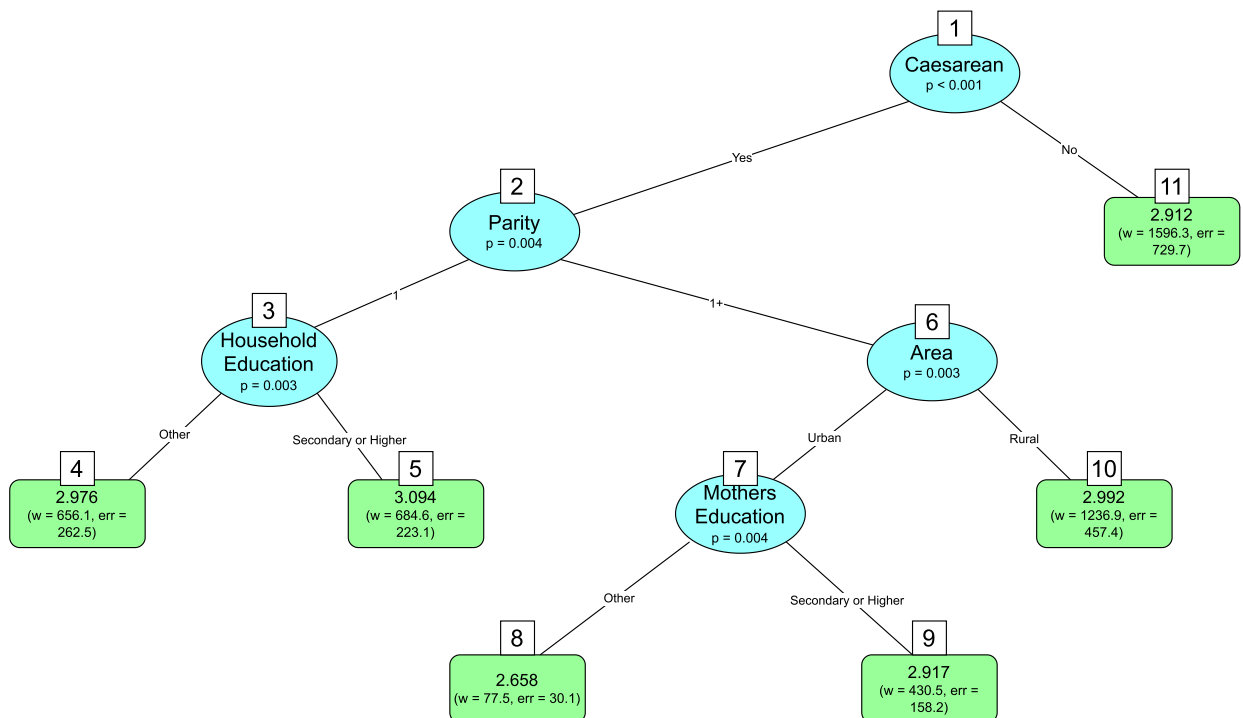
Fig. 2. Regression tree for MICS 2012-2013.

Fig. 3. Regression tree for MICS 2019.

Table 3
Summary of regression models for MICS 2012-2013.

Variables	FE Model Without Interaction $\hat{\beta}$ (Conf Int)	FE Model With Interaction $\hat{\beta}$ (Conf Int)	ME Model $\hat{\beta}$ (Conf Int)	BESAG Model $\hat{\beta}$ (Cred Int)	BYM Model $\hat{\beta}$ (Cred Int)
Mothers Age					
< = 20					
21 - 30	0.06 (-0.03, 0.14)	0.05 (-0.03, 0.14)	0.05 (-0.03, 0.14)	0.05 (-0.03, 0.13)	0.05 (-0.03, 0.13)
31+	0.02 (-0.10, 0.14)	0.02 (-0.10, 0.14)	0.02 (-0.1, 0.14)	0.03 (-0.09, 0.14)	0.03 (-0.09, 0.14)
Mothers Education					
Others					
Secondary or Higher	0.07 (-0.02, 0.15)	0.07 (-0.02, 0.15)	0.07 (-0.02, 0.15)	0.07 (-0.01, 0.15)	0.07 (-0.009, 0.15)
Head Education					
Others					
Secondary Complete or Higher	**-0.13 (-0.22, -0.05)**	**-0.12 (-0.21, -0.04)**	**-0.11 (-0.02, -0.03)**	**-0.12 (-0.21, -0.04)**	**-0.12 (-0.2, -0.04)**
ANC					
Yes					
No	-0.06 (-0.17, 0.04)	-0.07 (-0.17, 0.03)	-0.06 (-0.16, 0.05)	-0.05 (-0.16 , 0.04)	-0.06 (-0.16, 0.04)
ANC Times					
1					
2	0.08 (-0.05, 0.20)	0.09 (-0.04, 0.22)	0.08 (-0.05, 0.2)	0.09 (-0.04, 0.21)	0.08 (-0.04, 0.2)
3+	0.07 (-0.04, 0.18)	0.09 (-0.02, 0.20)	0.08 (-0.03, 0.19)	0.08 (-0.02, 0.19)	0.09 (-0.02, 0.19)
Delivery Place					
Home					
Others	0.05 (-0.07, 0.17)	0.03 (-0.09, 0.15)	0.03 (-0.09, 0.15)	0.03 (-0.08, 0.14)	0.03 (-0.08, 0.14)
Delivery Assist					
Doctor/Nurse/Midwife/ Auxiliary Midwife					
Others	-0.02 (-0.13, 0.10)	-0.03 (-0.14, 0.09)	-0.01 (-0.13, 0.1)	-0.02 (-0.13, 0.09)	-0.02 (-0.13, 0.09)
Caesarean					
Yes					
No	**-0.16 (-0.24, -0.08)**	**0.11 (0.00, 0.23)**	**0.12 (0.00, 0.23)**	**0.11 (0.001, 0.22)**	**0.11 (0.003, 0.22)**
Parity					
1 (Primiparous)					
1+	-0.01 (-0.09, 0.06)	0.05 (-0.07, 0.17)	0.05 (-0.07, 0.17)	0.05 (-0.07, 0.16)	0.05 (-0.07, 0.16)
Wealth Index					
Low					
Middle	0.04 (-0.05, 0.14)	0.03 (-0.07, 0.13)	0.03 (-0.07, 0.13)	**0.03 (0.06, 0.12)**	0.03 (-0.06, 0.12)
High	**0.11 (0.02, 0.19)**	-0.01 (-0.13, 0.11)	0.01 (-0.11, 0.13)	-0.001 (-0.12, 0.12)	0.00 (-0.12, 0.11)
Area					
Urban					
Rural	0.02 (-0.05, 0.09)	0.01 (-0.06, 0.09)	-0.01 (-0.09, 0.07)	0.001 (-0.07, 0.07)	-0.003 (-0.07, 0.07)
Rule - 1		**0.23 (0.08, 0.37)**	**0.23 (0.09, 0.37)**	**0.23 (0.09, 0.36)**	**0.23 (0.09, 0.36)**
Rule - 3		**-0.28 (-0.43, -0.12)**	**-0.25 (-0.4, -0.09)**	**-0.27 (-0.42, -0.11)**	**-0.26 (-0.41, -0.11)**
Rule - 4		**-0.22 (-0.39, -0.05)**	**-0.21 (-0.38, -0.04)**	**-0.22 (-0.38, -0.05)**	**-0.22 (-0.38, -0.05)**
Intercept	**2.8 (2.6, 3.0)**	**2.8 (2.5, 3.0)**	**2.8 (2.6, 3)**	**2.78 (2.56, 3.01)**	**2.79 (2.56, 3.01)**
Max Log Likelihood	-2856.75	-2836.96	-2871.12	-3017.53	-2953.06

Note: Estimates in bold fonts are statistically significant at 5% level of significance.

3.3. Results from different regression models

We have converted the rules from the regression tree to dummy variables. The dummy variable takes a value of one for individuals belonging to the associated rule and zero otherwise. Some rules have been ignored in the model to avoid multicollinearity. Table 3 shows the results of five different regression models for the MICS 2012-2013 dataset. The FE model without interactions is considered as a benchmark for comparison. The rest of the models include rules (interactions) identified from the tree. The results of the models are similar. Education of household head and whether a newly born baby was delivered through C-section were significantly associated with children's birth weight at 5% level of significance across all the competing models. Babies born to mothers who resided in households where the household head possessed at least a secondary education were, on average, 0.11 to 0.13 kg lighter than those babies born to mothers who came from households led by more educated household heads. While making this interpretation, all the other variables are held fixed. Interestingly, the benchmark fixed effects model without interactions indicates a significant reduction in birth weight for children born without the cesarean method ($\hat{\beta}$ = -0.16, 95% CI = [-0.24, -0.08]), the remaining four models demonstrate lower birth weight for children born through the cesarean procedure. Also, the fixed effects model without interactions and the Besag model both reveal a significant association between higher wealth status and higher levels of birth weight.

Notably, the interaction terms obtained from the regression tree as rules (namely, rule-1, rule-3 and rule-4) are all significantly and consistently associated with birth weight at 5% significance level across all the four models with interactions. Mothers who

Table 4
Summary of regression models for MICS 2019.

Variables	FE Model Without Interaction $\hat{\beta}$ (Conf Int)	FE Model With Interaction $\hat{\beta}$ (Conf Int)	ME Model $\hat{\beta}$ (Conf Int)	BESAG Model $\hat{\beta}$ (Cred Int)	BYM Model $\hat{\beta}$ (Cred Int)
Mothers Age					
< = 20					
21 - 30	**0.09 (0.03, 0.14)**	**0.08 (0.02, 0.13)**	**0.08 (0.03, 0.13)**	**0.08 (0.03, 0.13)**	**0.08 (0.03, 0.13)**
31+	**0.08 (0.01, 0.15)**	**0.07 (0.00, 0.14)**	**0.08 (0.01, 0.15)**	0.07 (0.00, 0.14)	**0.07 (0.01, 0.14)**
Mothers Education					
Others					
Secondary or Higher	0.03 (-0.02, 0.08)	0.01 (-0.04, 0.06)	0.02 (-0.04, 0.07)	0.001 (-0.05, 0.05)	0.00 (-0.05, 0.05)
Head Education					
Others					
Secondary Complete or Higher	**0.07 (0.03, 0.11)**	**0.05 (0.01, 0.10)**	0.03 (-0.02, 0.07)	-0.04 (-0.003, 0.09)	0.04 (0.00, 0.09)
ANC					
Yes					
No	**-0.10 (-0.19, -0.02)**	**-0.11 (-0.19, -0.03)**	**-0.10 (-0.18, -0.02)**	**-0.11 (-0.02, -0.03)**	**-0.11 (-0.20, -0.03)**
ANC Times					
1					
2	0.00 (-0.08, 0.08)	0.00 (-0.08, 0.08)	-0.01 (-0.09, 0.07)	-0.01 (-0.09, 0.07)	-0.01 (-0.09, 0.07)
3+	0.04 (-0.03, 0.11)	0.04 (-0.04, 0.11)	0.03 (-0.04, 0.10)	0.03 (-0.04, 0.11)	0.03 (-0.04, 0.11)
Delivery Place					
Home					
Others	**-0.15 (-0.26, -0.03)**	**-0.15 (-0.26, -0.03)**	-0.11 (-0.22, 0.00)	**-0.12 (-0.24, -0.01)**	**-0.12 (-0.24, -0.01)**
Delivery Assist					
Doctor/Nurse/Midwife/ Auxiliary Midwife					
Others	-0.06 (-0.17, 0.05)	-0.06 (-0.18, 0.05)	-0.05 (-0.16, 0.05)	-0.06 (-0.17, 0.05)	-0.06 (-0.17, 0.05)
Caesarean					
Yes					
No	**-0.08 (-0.12 -0.04)**	0.00 (-0.09, 0.09)	0.02 (-0.08, 0.11)	0.02 (-0.07, 0.11)	0.02 (-0.07, 0.11)
Parity					
1 (Primiparous)					
1+	**-0.07 (-0.12, -0.03)**	0.02 (-0.05, 0.09)	0.00 (-0.07, 0.07)	0.01 (-0.05, 0.08)	0.01 (-0.05, 0.08)
Wealth Index					
Low					
Middle	0.02 (-0.03, 0.08)	0.03 (-0.03, 0.08)	0.03 (-0.03, 0.08)	0.04 (-0.01, 0.10)	0.04 (-0.01, 0.10)
High	-0.02 (-0.07, 0.03)	-0.02 (-0.07, 0.03)	0.03 (-0.02, 0.08)	0.03 (-0.02, 0.08)	0.03 (-0.03, 0.08)
Area					
Urban					
Rural	**0.07 (0.02, 0.11)**	0.03 (-0.04, 0.10)	0.01 (-0.06, 0.09)	0.02 (-0.05, 0.09)	**0.14 (0.02, 0.26)**
Rule - 1		**0.14 (0.01, 0.26)**	**0.11 (0.01, 0.24)**	**0.14 (0.02, 0.26)**	**0.14 (0.02, 0.23)**
Rule - 2		**0.20 (0.08, 0.32)**	**0.18 (0.06, 0.31)**	**0.19 (0.07, 0.32)**	**0.19 (0.07, 0.32)**
Rule - 3		**-0.23 (-0.39, -0.07)**	**-0.18 (-0.35, -0.01)**	**-0.24 (-0.40, -0.08)**	**-0.24 (-0.40, -0.08)**
Rule - 4		-0.02 (-0.12, 0.08)	-0.05 (-0.16, 0.05)	-0.03 (-0.13, 0.07)	**-0.03 (-0.13, -0.07)**
Rule - 5		0.06 (-0.04, 0.16)	0.07 (-0.03, 0.17)	0.07 (-0.03, 0.17)	0.07 (-0.03, 0.17)
Intercept	**2.8 (2.6, 3.0)**	**2.8 (2.5, 3.0)**	**2.9 (2.7, 3.1)**	**2.89 (2.71, 3.06)**	**2.89 (2.72, 3.07)**
Max Log Likelihood	-4770.93	-4757.73	-4442.94	-4636.71	-4572.33

Note: Estimates in bold fonts are statistically significant at 5% level of significance.

undergo caesarean deliveries and are primiparous (rule 1) are observed to have children with a birth weight that is, on average, 0.23 kg higher ($\hat{\beta}$ = 0.23, CI = [0.08, 0.37]) in comparison to those in different circumstances, while maintaining all other variables constant. Mothers who did not undergo caesarean deliveries and are affiliated with the low or middle class (rule 3) exhibit a birth weight for their children that is, on average, 0.25 kg to 0.28 kg lower compared to those in differing situations. The combination of no C-section and low parity but higher wealth (rule 4) also resulted in significantly lower average birth weight.

The outcomes of the five regression models applied to the more recent MICS 2019 dataset are presented in Table 4. In contrast to the 2012-13 data, a different set of variables, mother's age, uptake of ANC and place of delivery appear to be the most consistent predictors of birth weight across models. Mother's older than 20 years of age have higher average birth weight for their newborns in comparison to younger mothers. No uptake of ANC is associated with around 0.10 kg reduction in birth weight, and delivery in places other than home has been linked to 0.12 kg to 0.15 kg lower birth weight, on average. There are certain differences in the models' results. Non-cesarean births and births to non-primiparous mothers were significantly associated with lower birth weight ($\hat{\beta}$ = -0.08, CI = [-0.12, -0.04] and $\hat{\beta}$ = -0.07, CI = [-0.12, -0.03, respectively])only in the benchmark FE model without interaction, but insignificant in other models. Higher education of the household head shows significantly higher birth weight only in FE models, with and without interaction. Interestingly, residence in rural areas relates to elevated birth weights in comparison to residence in urban areas only in the simple benchmark FE model and the most complex BYM model.
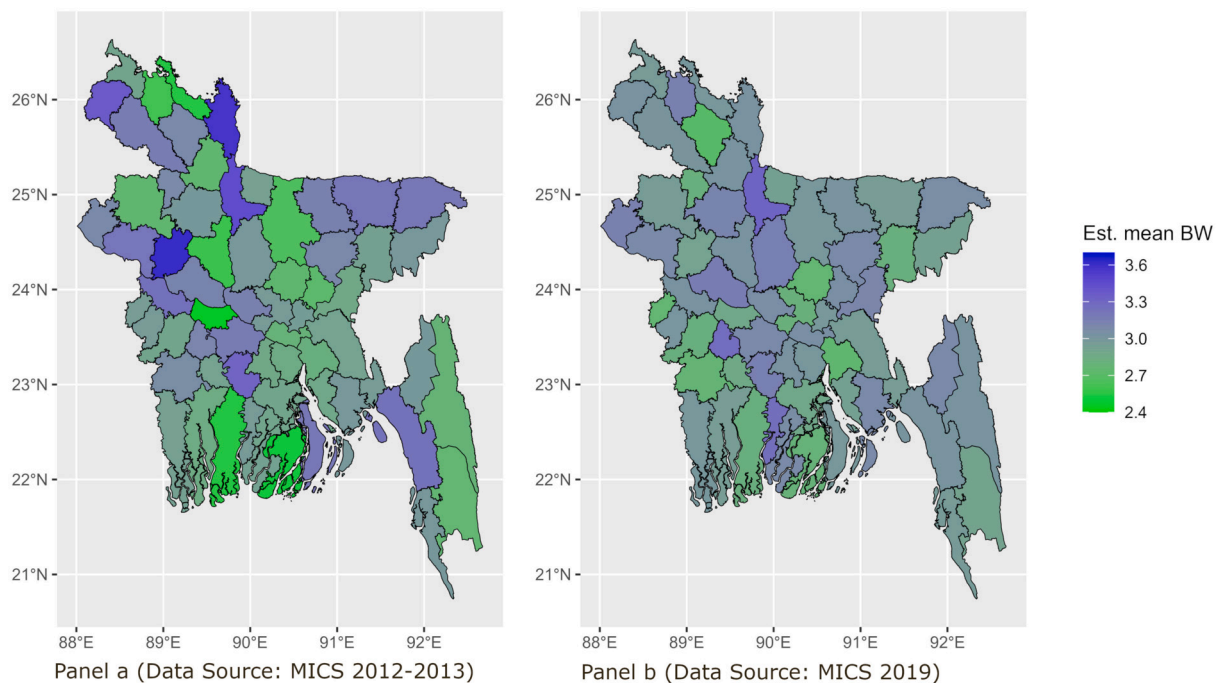
Fig. 4. District-wise observed average birth weight. *Panel a* refers to district-level mapping of average birth weights computed from observed MICS 2012-2013 data and *Panel b* refers to district-level mapping of average birth weights computed from observed MICS 2019 data.

Like in the case of the MICS 2012-2013 dataset, several tree-based rules (rules 1, 2 and 3) capturing important interactions among socio-demographic characteristics of babies appear to be important predictors of their birth weight. For example, rule 2 combining birth by C-section, primiparous mothers and above secondary level education for household heads is associated with about 0.20 higher average birth weight across all models. On the contrary, babies born via C-section, born to mothers with higher parity and born to households residing in urban areas (rule 3) are expected to weigh 0.18 kg to 0.24 kg lower at birth than babies characterized differently.

In addition to four specifications of the regression models we also applied the Leorex model [44] which defines the spatial correlation in a different structure than Besag and BYM models. The results of the Leorex model are not quantitatively or qualitatively very different from that of those four competing models and are presented in Table 6 in the Appendix.

3.4. Geographical mapping of birth weight

Fig. 4 illustrates district-wise observed average birth weights. The visualization indicates an overall increase in average birth weight from 2012-2013 (panel a) to 2019 (panel b). However, the districts with the lowest average birth weights have remained relatively consistent. When comparing the years 2012-2013 with 2019, improvement in birth weight has been noted in central and northwestern districts. However, it is noteworthy that districts in the southwest region continue to exhibit lower birth weights, suggesting a persistent area of concern.

In contexts with limited district-level sample sizes, model-based estimates offer a valuable means to comprehend average birth weights per district in contrast to empirical observations. Fig. 5 depicts predicted average birth weights within districts, employing the ME model. While the general trend persists, variations in intensity suggest shifts in the values of the covariates over time in Bangladesh. In comparison to the observed birth weights, the predicted birth weights derived from the ME model indicate a higher level of uniformity across the various districts. Districts in the central, southern-western and some northern regions continue to exhibit elevated susceptibility to low birth weight occurrences.

The maps generated using the predicted values derived from the Besag model, as depicted in Fig. 6, exhibit outcomes analogous to those of the ME model. A significant rise in birth weight has been noted in the districts of the southern-east region over the observed periods.

Concurrently, as portrayed in Fig. 7, outcomes indicate that districts with a heightened risk of low birth weight remain consistent with prior observations. Nonetheless, the forecasts for the southwestern districts suggest a reduced risk of low birth weight in comparison to the preceding maps.
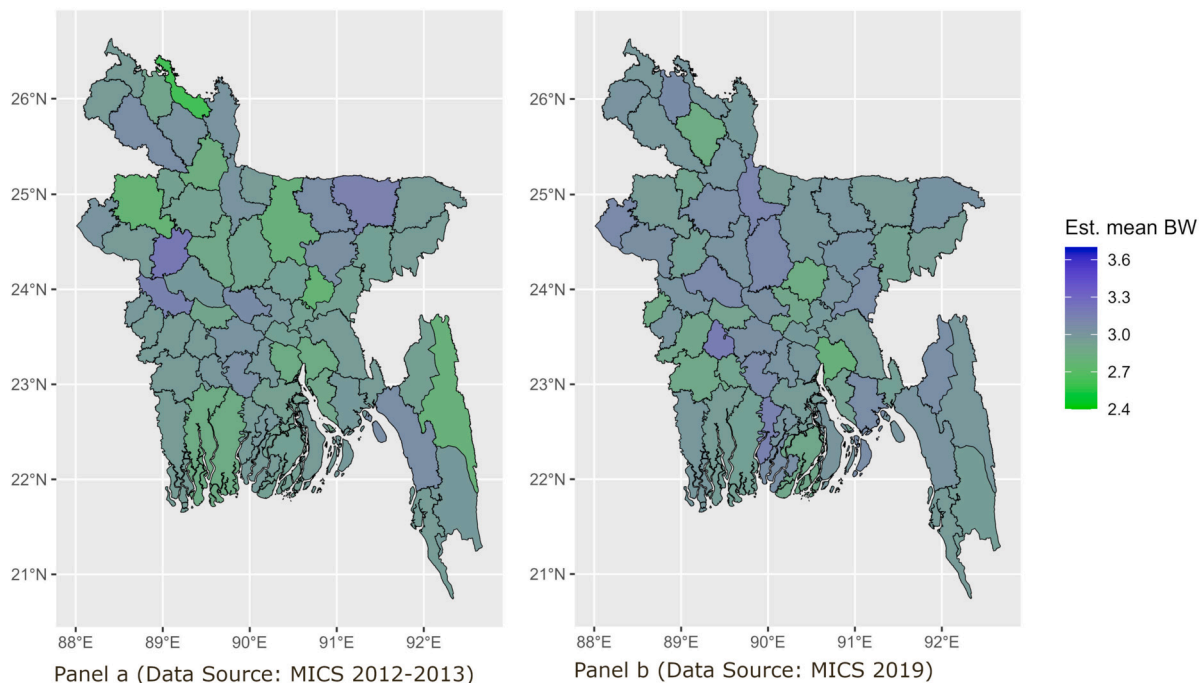
Fig. 5. District-wise predicted average birth weight using ME model. *Panel a* refers to district-level mapping of average birth weights predicted from applying ME model on MICS 2012-2013 data and *Panel b* refers to district-level mapping of average birth weights predicted from applying ME model on MICS 2019 data.
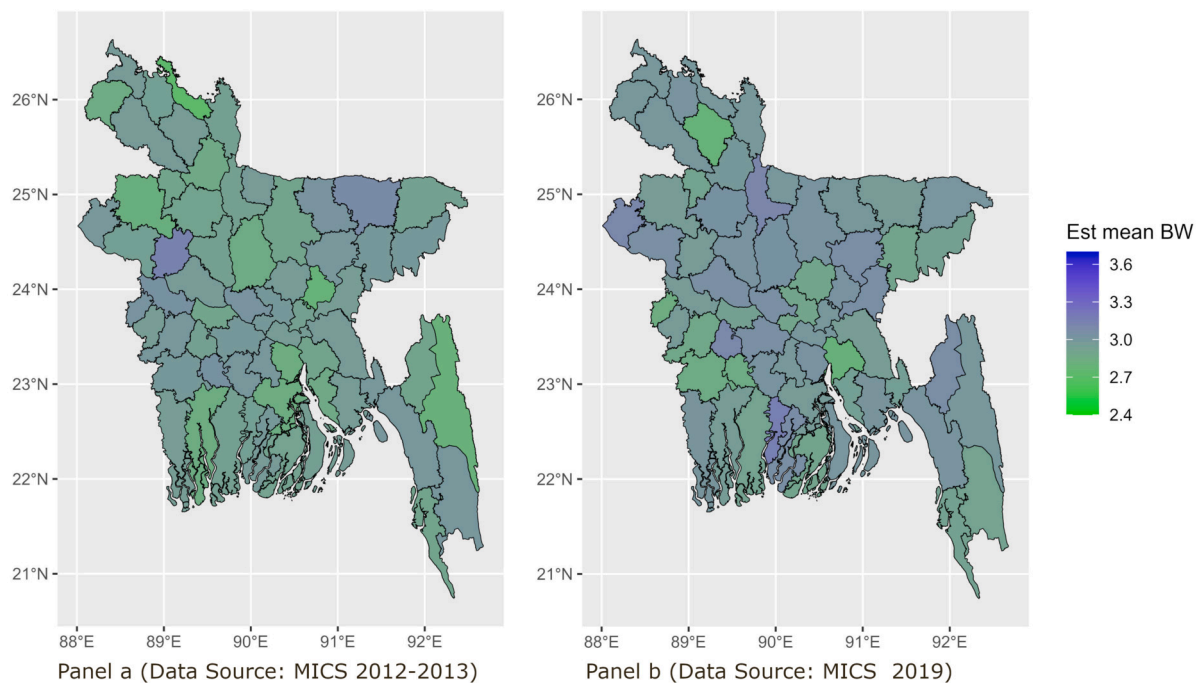

Fig. 6. District-wise predicted average birth weight using Besag model. *Panel a* refers to district-level mapping of average birth weights predicted from applying Besag model on MICS 2012-2013 data and *Panel b* refers to district-level mapping of average birth weights predicted from applying Besag model on MICS 2019 data.

3.5. Comparison of predictive performance of the models

In analyzing the MICS 2012-2013 dataset, it becomes evident from Table 5 that in the context of in-sample fit of data, the ME model emerges as the most optimal performer when considering MSE and DIC as metrics of performance. Interestingly, the Besag and BYM models, which incorporate spatial correlation, exhibit comparable performance to the ME model, surpassing the performance of
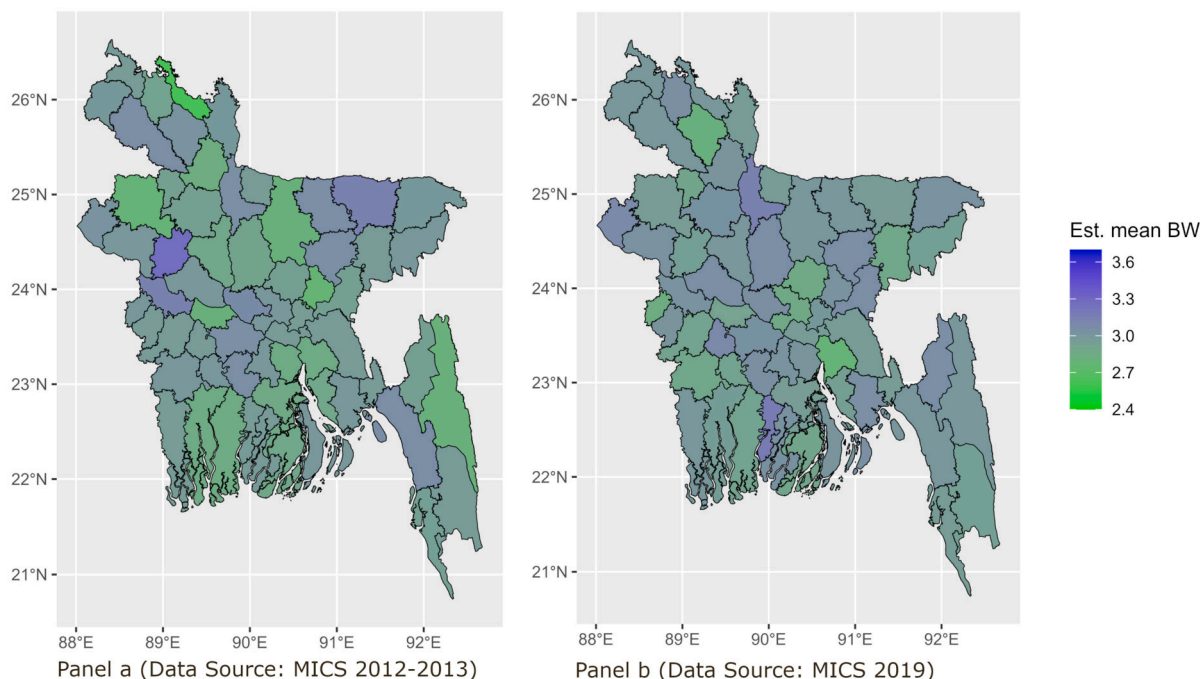
Fig. 7. District-wise predicted average birth weight using BYM model. *Panel a* refers to district-level mapping of average birth weights predicted from applying BYM model on MICS 2012-2013 data and *Panel b* refers to district-level mapping of average birth weights predicted from applying BYM model on MICS 2019 data.

Table 5
Predictive performance of the models.

Measure of Performance	MICS 2012-2013				MICS 2019			
	FE model	ME model	Besag model	BYM model	FE model	ME model	Besag model	BYM model
In sample estimate								
MSE	0.585	0.570	0.568	0.569	0.400	0.389	0.381	0.380
DIC	7060.82	5585.38	5724.94	5723.24	11360.95	9366.91	8905.47	8902.70
WAIC			5764.16	5758.73			8924.88	8922.17
Out sample estimates								
10 fold CV	0.576	0.564	0.585	0.581	0.394	0.386	0.391	0.390
$-\sum \log(\text{CPO}_i)$			2586.95	2592.39			4335.39	4334.80

the FE model (with interaction) by a big margin. Upon employing WAIC for the purpose of comparing the Besag and BYM models, it becomes evident that the Besag model demonstrates a slight improvement in performance. When considering predictive performance out-of-sample, an evaluation of the 10-fold cross-validation error reveals a consistent outcome: the ME model stands out as the superior choice. The aggregated value of CPO indicates a recurring pattern: the Besag model consistently outperforms the BYM model.

Within the MICS 2019 dataset, the findings exhibit intriguing disparities. Specifically, when considering in-sample performance, the models incorporating spatial adjustments, namely the Besag model and BYM model, demonstrate superior performance compared to the ME and FE models (with interaction). Interestingly, the BYM model emerges as the stronger performer between these two spatially adjusted models, suggesting a more pronounced spatial and cluster dependence evident within this dataset. The out-of-sample performance of the ME, Besag, and BYM models demonstrates a relatively consistent similarity, with all three models surpassing the performance of the FE model.

4. Discussion

This study seeks to identify factors associated with birth weight by integrating data-driven interactions and acknowledging plausible spatial and cluster effects. The statistical models used to predict birth weight are presented in the order of increasing complexities captured from the data. The MICS data constitute a hierarchical structure where individual respondents (mothers) are nested within districts. The naive FE model looks at the potential fixed effects of the determinants of birth weight but does not consider possible district-level cluster effects inherent in the data. Thus, it ignores similarities in the birth weight of babies among women in the same geographical districts that may arise due to unobserved common characteristics (e.g., culture, ritual, general health-seeking behavior, etc.). The omission of such random cluster effects, if they exist, is known to introduce problems in statistical

inference. The ME model accounts for these unobserved district-specific random effects but fails to incorporate potential spatial correlation that suggests that women from neighboring districts may exhibit further similarities (e.g., broader regional poverty or provision of healthcare, etc.). Two variants of a spatial model, namely Besag and BYM, are then introduced to adjust for such spatial correlations in a specific structured manner. The former, however, ignores the unstructured random clustering effects within districts, but the latter considers both cluster effects and spatial correlation simultaneously. While the spatial models are sophisticated in terms of the higher complexities they integrate, their estimation poses a considerable challenge. The relatively simpler FE and ME models are estimated by the well-established maximum likelihood method, but more complex approximations to the Bayesian inference method are adopted for the estimation of spatial models.

The study reveals that the mean birth weight stands at 2.93 kg in 2012-2013 and rises to 2.96 kg in 2019. Among its determinants, household heads' education, whether birth occurred via caesarean delivery or not, and household wealth emerged as significant factors of birth weight in 2012-2013. In particular, babies delivered through C-section had significantly lower birth weight than those born via alternative methods. These results are consistent with the findings from a paper that utilized the same dataset [23]. In 2019, predictive dominance shifted towards a different set of determinants of birth weight, namely mothers' age, ANC uptake, and place of delivery. Utilization of ANC and delivery at home were found to be associated with higher birth weights, which is in line with findings from existing literature [25,45,46]. A few models suggest that in MICS 2012-2013 birth weight is observed to be significantly higher among mothers in higher wealth class as observed in other previous research [47,21] but in 2019 wealth class is observed to be an insignificant factor. Model comparison metrics show that in 2012-13, there are clear gains in the goodness of fit of sample data and in predictive accuracy when considering random cluster effects but marginal losses from adding further complexity like spatial correlation. In 2019, however, incorporating both cluster effects and spatial correlation resulted in improved performance both in-sample and out-of-sample.

A unique contribution of this study is the application of a regression tree, a machine learning technique, for detecting complex but important interactions among determinants of birthweight. These interactions set out as simple rules based on socio-demographic characteristics help to identify groups of mothers who are the most vulnerable to giving birth to low-weight babies. The regression tree analysis carried out in 2012-2013 unveils a notably low average birth weight (approximately 2.75 kg) in mothers who did not undergo caesarean deliveries and belonged to households in low or middle-wealth categories. In contrast, mothers who experienced caesarean sections and are primiparous exhibit an average birth weight of 3.1 kg, surpassing the national average of 2.93 kg. In the year 2019, the mean birth weight stands at only 2.66 kg for women who underwent caesarean deliveries, possess a parity of more than 1, reside in urban areas, and who completed less than secondary level education.

Importantly, the tree-based interactions are consistently significant across all model specifications, highlighting the importance of looking at the interplay among different sociodemographic factors as determinants of low birth weight rather than solely looking at factors in isolation. The combination of non-cesarean delivery and low wealth constituted the group at the highest risk of low birth weight in 2012-13. In contrast, interactions of C-section delivery, urban residence, higher parity and low education defined the most vulnerable group of women in the year 2019.

Mapping the average birth weight on a district level uncovers geographical disparities [21]. Notably, the districts previously identified as having a high risk of LBW in 2012-2013 continue to exhibit the same vulnerability in 2019. Model-based estimations of birth weights identify specific districts situated in the southern-western and northern regions as being at elevated risk of LBW. These results are in line with the findings of several previous studies [27,13].

This study has a number of limitations, which mostly originate from the data collection and sampling procedures. First, around 70% and 50% of the babies born to mothers of reproductive age during the last 2 years were not weighed at birth in the 2012-13 and 2019 surveys, respectively [29,30]. These large proportions of missing values may indicate that the average birth weight derived from the severely curtailed samples may not be unbiased estimates of population averages. Moreover, since the absence of weighing at birth was found to occur largely in certain socio-demographic groups of women (e.g., elderly, poor, lowly educated, etc.), any association between a variable defining such a sub-group and LBW may be unreliable due to possible selection bias. Secondly, around 90% of the birth weight data are obtained from mothers' recall as opposed to from health card records. This memory-based self-reporting is a possible source of additional error. Finally, any relationship established between socio-demographic factors and birth weight in this study are correlations rather than causal and interpreted as such.

5. Conclusion

Despite a decline in the prevalence of low birth weight over time, its significance as a public health concern in Bangladesh remains pronounced. The findings of this study unveil a persistent trend wherein districts with low average birth weights in 2012-2013 continue to exhibit relatively lower values even after six years. This observation suggests that concerted efforts to enhance birth weights in these districts might have been insufficiently implemented. This study holds the potential to inform policy formulation and provide targeted interventions at the district level to enhance birth weights. The outcomes derived from applying machine learning techniques can aid in pinpointing the groups that require heightened attention. The interactions of socio-demographic characteristics in determining birth weight call for considering multiple factors simultaneously when designing related health policies. Enhancing female education and campaigning against early childbearing might contribute to an increase in birth weights on a national scale. The fact that receiving ANC emerged as an important predictor of birth weight in the recent survey highlights the importance of encouraging appropriate utilization of healthcare facilities. At the end, it is worthy to summarize that the methodology showcased in this study for investigating birth weight can be applied to identify risk factors contributing to other public health problems such as infant mortality, child immunization, early marriage and childbearing, etc. This approach holds potential not only for Bangladesh

but also for other nations, particularly those seeking precise estimates at the local level to inform internal policy formulation and foster harmonious local development.

Ethical statement

The authors did not collect any data but rather used secondary sources. UNICEF provided financial support for MICS 2019 and MICS 2012-2013. An oral consent was taken from each of the respondents before the interview. All the respondents are unidentified.

CRediT authorship contribution statement

Rubaiya: Writing – original draft, Formal analysis, Conceptualization. **Mohaimen Mansur:** Writing – review & editing, Supervision, Methodology, Conceptualization. **Md. Muhitul Alam:** Writing – review & editing, Writing – original draft, Visualization, Formal analysis. **Md. Israt Rayhan:** Writing – review & editing, Supervision.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

This paper uses nationally representative data from the Multiple Indicator Cluster Survey (MICS) 2019 and MICS 2012-2013. The data are available upon registration and request from the MICS UNICEF website at https://mics.unicef.org/surveys.

Appendix A

In evaluating spatial dependence in data, a practical model is presented in the work by Leroux, Lei, and Breslow [44]. Within this model, the precision matrix's structure is characterized by a convex combination of an identity matrix I (representing an independent and identically distributed spatial effect) and the precision matrix of an intrinsic conditional autoregressive (CAR) model Q (indicative of a spatial pattern). The precision matrix has the following form,

$$(1 - \lambda)I + \lambda Q,$$

where the variable λ can assume any value within the range of 0 to 1, signifying the degree of spatial structure present in the data [40].

The results from the Leroux model are shown below:

Table 6
Results from Leroux model.

Variables	MICS 2012-2013 $\hat{\beta}$ (Cred Int)	MICS 2019 $\hat{\beta}$ (Cred Int)
Mothers Age		
< = 20		
21 - 30	0.04 (-0.04, 0.13)	**0.08 (0.03, 0.13)**
31+	0.02 (-0.09, 0.14)	**0.07 (0.00, 0.14)**
Mothers Education		
Others		
Secondary or Higher	0.07 (-0.01, 0.15)	0.00 (-0.05, 0.05)
Head Education		
Others		
Secondary Complete or Higher	**-0.11 (-0.20, -0.03)**	**0.04 (0.00, 0.09)**
ANC		
Yes		
No	-0.04 (-0.15, 0.06)	**-0.11 (-0.20, -0.03)**
ANC Times		
1		
2	0.07 (-0.05, 0.19)	-0.01 (-0.09, 0.07)
3+	0.07 (-0.03, 0.18)	0.03 (-0.04, 0.10)
Delivery Place		
Home		
Other	**0.04 (-0.24, -0.01)**	**-0.13 (-0.24, -0.01)**
Delivery Assist		

Table 6 (*continued*)

Variables	MICS 2012-2013 $\hat{\beta}$ (Cred Int)	MICS 2019 $\hat{\beta}$ (Cred Int)
Doctor/Nurse/Midwife/ Auxiliary Midwife		
Others	0.00 (-0.17, 0.05)	-0.06 (-0.17, 0.05)
Caesarean		
Yes		
No	0.10 (-0.11, 0.11)	0.02 (-0.07, 0.11)
Parity		
1 (Primiparous)		
1+	0.05 (-0.07, 0.17)	0.01 (-0.05, 0.08)
Wealth Index		
Low		
Middle	0.04 (-0.06, 0.14)	0.04 (-0.01, 0.09)
High	0.04 (-0.08, 0.16)	0.03 (-0.02, 0.08)
Area		
Urban		
Rural	-0.03 (-0.10, 0.05)	0.02 (-0.05, 0.08)
Rule - 1	**0.22 (0.08, 0.36)**	**0.14 (0.02, 0.25)**
Rule - 2		**0.19 (0.07, 0.31)**
Rule - 3	**-0.22 (-0.37, -0.06)**	**-0.24 (-0.40, -0.08)**
Rule - 4	**-0.20 (-0.36, -0.03)**	-0.03 (-0.13, 0.07)
Rule - 5		0.07 (-0.03, 0.16)
(Intercept)	-0.54 (-1569830.04, 1569828.44)	**2.89 (2.71, 3.07)**
Predictive Performance		
MSE	0.54	0.38
DIC	5741.10	8903.48
WAIC	5780.32	8923.06
10 fold CV	0.59	0.39
$-\sum \log(CPO_i)$	2885.27	4457.44

Note: Estimates in bold fonts are statistically significant at 5% level of significance.

References

[1] World Health Organization, Global nutrition targets 2025: low birth weight policy brief, https://www.who.int/publications/i/item/WHO-NMH-NHD-14.5, 2014, World Health Organization, Geneva.

[2] UNICEF, Low birthweight, https://data.unicef.org/topic/nutrition/low-birthweight/, 2023.

[3] M.-R. Järvelin, P. Elliott, I. Kleinschmidt, M. Martuzzi, C. Grundy, A.-L. Hartikainen, P. Rantakallio, Ecological and individual predictors of birthweight in a northern Finland birth cohort 1986, Paediatr. Perinat. Epidemiol. 11 (3) (1997) 298–312.

[4] A.B.d. Moraes, R.R. Zanini, J. Riboldi, E.R.J. Giugliani, Risk factors for low birth weight in Rio Grande do Sul State, Brazil: classical and multilevel analysis, Cad. Saude Publica 28 (2012) 2293–2305.

[5] M.S. Kramer, Determinants of low birth weight: methodological assessment and meta-analysis, Bull. World Health Organ. 65 (5) (1987) 663.

[6] M.T.A. Sampurna, K.D. Handayani, M.T. Utomo, D. Angelika, R. Etika, A. Harianto, M.P. Mapindra, M.P. Mahindra, F. Efendi, R.K. Kaban, et al., Determinants of neonatal deaths in Indonesia: a national survey data analysis of 10,838 newborns, Heliyon 9 (1) (2023).

[7] J. Tian, W. Tu, S. Tedders, D. Chen, A spatial–temporal analysis of low birth weight prevalence in Georgia, USA, GeoJournal 78 (2013) 885–895.

[8] J.D. Boardman, D.A. Powers, Y.C. Padilla, R.A. Hummer, Low birth weight, social factors, and developmental outcomes among children in the United States, Demography 39 (2002) 353–368.

[9] H.E. Nasreen, Z.N. Kabir, Y. Forsell, M. Edhborg, Low birth weight in offspring of women with depressive and anxiety symptoms during pregnancy: results from a population based study in Bangladesh, BMC Public Health 10 (2010) 1–8.

[10] S. Dinesen, G. Greisen, Quality of life in young adults with very low birth weight, Arch. Dis. Child., Fetal Neonatal Ed. 85 (3) (2001) F165–F169.

[11] L.K. Lund, T. Vik, S. Lydersen, G.C. Løhaugen, J. Skranes, A.-M. Brubakk, M.S. Indredavik, Mental health, quality of life and social relations in young adults born with low birth weight, Health Qual. Life Outcomes 10 (1) (2012) 1–10.

[12] W.H. Organization, et al., Comprehensive implementation plan on maternal, infant and young child nutrition, Tech. Rep., World Health Organization, 2014.

[13] H. Blencowe, J. Krasevec, M. de Onis, R.E. Black, X. An, G.A. Stevens, S. Guo, A.M. Faal, L.C. Mullany, C.e.a. Robert, National, regional, and worldwide estimates of low birthweight in 2015, with trends from 2000: a systematic analysis, Lancet Glob. Health 7 (7) (2019) e849–e860.

[14] L.A. Thompson, D.C. Goodman, C.-H. Chang, T.A. Stukel, Regional variation in rates of low birth weight, Pediatrics 116 (5) (2005) 1114–1121.

[15] V. Fuster, P. Zuluaga, S.E. Colantonio, J. Román-Busto, Regional differences in low birth weight in Spain: biological, demographic and socioeconomic variables, J. Biosoc. Sci. 47 (1) (2015) 90–104.

[16] I. Marete, O. Ekhaguere, C.M. Bann, S.L. Bucher, P. Nyongesa, A.B. Patel, P.L. Hibberd, S. Saleem, R.L. Goldenberg, S.S. Goudar, et al., Regional trends in birth weight in low-and middle-income countries 2013–2018, Reprod. Health 17 (3) (2020) 1–8.

[17] M.J. Rana, R. Kim, S. Ko, L.K. Dwivedi, K. James, R. Sarwal, S. Subramanian, Small area variations in low birth weight and small size of births in India, Matern. Child Nutr. 18 (3) (2022) e13369.

[18] L.S. Álvarez-Casta no, B. Caicedo-Velásquez, C. Castaño-Díez, M. Marí-Dell'Olmo, M. Gotsens, Full-term low birth weight and its relationship with the socioeconomic conditions of municipalities in Antioquia: spatio-temporal analysis, Biomédica 38 (3) (2018) 345–354.

[19] M. Arayeshgari, S. Najafi-Ghobadi, H. Tarhsaz, S. Parami, L. Tapak, Machine learning-based classifiers for the prediction of low birth weight, Healthc. Inform. Res. 29 (1) (2023) 54–63.

[20] M.H. Abel, Low birth weight and interactions between traditional risk factors, J. Genet. Psychol. 158 (4) (1997) 443–456.

[21] M.S. Ahmed, Mapping the prevalence and socioeconomic predictors of low birth weight among Bangladeshi newborns: evidence from the 2019 multiple indicator cluster survey, Int. Health 14 (5) (2022) 485–491.

[22] WHO, Sixty-fifth World Health Assembly, wha65/2012/rec/1, 2012, pp. 1–3.
[23] J.R. Khan, M.M. Islam, N. Awan, O. Muurlink, Analysis of low birth weight and its co-variants in Bangladesh based on a sub-sample from nationally representative survey, BMC Pediatr. 18 (1) (2018) 1–9.
[24] S.A. Islam Pollob, M.M. Abedin, M.T. Islam, M.M. Islam, M. Maniruzzaman, Predicting risks of low birth weight in Bangladesh with machine learning, PLoS ONE 17 (5) (2022) e0267190.
[25] S. Khatun, M. Rahman, Socio-economic determinants of low birth weight in Bangladesh: a multivariate approach, Bangladesh Med. Res. Counc. Bull. 34 (3) (2008) 81–86.
[26] J. Ferdos, M.M. Rahman, Maternal experience of intimate partner violence and low birth weight of children: a hospital-based study in Bangladesh, PLoS ONE 12 (10) (2017) e0187138.
[27] M.S. Alam, S.S. Hossain, F.F. Sheela, Spatial smoothing of low birth weight rate in Bangladesh using Bayesian hierarchical model, J. Appl. Stat. 46 (10) (2019) 1870–1885.
[28] G. Monawar Hosain, N. Chatterjee, A. Begum, S.C. Saha, Factors associated with low birthweight in rural Bangladesh, J. Trop. Pediatr. 52 (2) (2005) 87–91.
[29] Bangladesh Bureau of Statistics (BBS), United Nations Children's Fund (UNICEF), Bangladesh multiple indicator cluster survey 2019, https://mics.unicef.org/surveys, 2020.
[30] Bangladesh Bureau of Statistics (BBS), United Nations Children's Fund (UNICEF), Bangladesh Multiple Indicator Cluster Survey 2012–2013. Progotir Pathey. Final Report, BBS and UNICEF Bangladesh, Dhaka, Bangladesh, 2014.
[31] L. Breiman, Classification and Regression Trees, Routledge, 2017.
[32] D.W. Andrews, Y.-J. Whang, Additive interactive regression models: circumvention of the curse of dimensionality, Econom. Theory 6 (4) (1990) 466–479.
[33] J. Besag, Spatial interaction and the statistical analysis of lattice systems, J. R. Stat. Soc., Ser. B, Methodol. 36 (2) (1974) 192–225.
[34] J. Besag, J. York, A. Mollié, Bayesian image restoration, with two applications in spatial statistics, Ann. Inst. Stat. Math. 43 (1991) 1–20.
[35] S. Banerjee, B.P. Carlin, A.E. Gelfand, Hierarchical Modeling and Analysis for Spatial Data, Chapman and Hall/CRC, 2003.
[36] N. Cressie, C.K. Wikle, Statistics for Spatio-Temporal Data, John Wiley & Sons, 2015.
[37] N. Breslow, B. Leroux, R. Platt, Approximate hierarchical modelling of discrete data in epidemiology, Stat. Methods Med. Res. 7 (1) (1998) 49–62.
[38] H. Rue, S. Martino, N. Chopin, Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations, J. R. Stat. Soc., Ser. B, Stat. Methodol. 71 (2) (2009) 319–392.
[39] F. Lindgren, H. Rue, Bayesian spatial modelling with R-INLA, J. Stat. Softw. 63 (19) (2015).
[40] R. Bivand, V. Gómez-Rubio, H. Rue, Spatial data analysis with R-INLA with some extensions, J. Stat. Softw. 63 (20) (2015) 1–31.
[41] V. Gómez-Rubio, Bayesian Inference with INLA, CRC Press, 2020.
[42] H. Sutanto, H. Pramoedyo, W. Wardhani, S. Astutik, The selection of Bayesian polynomial regression with INLA by using DIC, WAIC and CPO, J. Phys. Conf. Ser. 1747 (2021) 012029, IOP Publishing.
[43] A. Vehtari, A. Gelman, J. Gabry, Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC, Stat. Comput. 27 (2017) 1413–1432.
[44] B.G. Leroux, X. Lei, N. Breslow, Estimation of disease rates in small areas: a new mixed model for spatial dependence, in: Statistical Models in Epidemiology, the Environment, and Clinical Trials, Springer, 2000, pp. 179–191.
[45] M.N.A. Siddiqi, A. Muyeed, M.N. Haque, M.A. Goni, S.C. Shadhana, Low birth weight of newborns and its association with demographic and socio-economic determinants: findings from multiple indicator cluster survey (MICS) Bangladesh 2019, Int. J. Health Stud. 7 (1) (2021) 37–42.
[46] M.J. Alam, M.M. Islam, M. Maniruzzaman, N.F. Ahmed, M. Tawabunnahar, M.J. Rahman, D.C. Roy, J. Mydam, Socioeconomic inequality in the prevalence of low birth weight and its associated determinants in Bangladesh, PLoS ONE 17 (10) (2022) e0276718.
[47] N.J. Sathi, B. Ahammed, K. Alam, R. Hashmi, K.Y. Lee, S.A. Keramat, Socioeconomic inequalities in low birth weight in South Asia: a comparative analysis using Demographic and Health Surveys, SSM-Popul. Health 20 (2022) 101248.