Communications in Statistics - Simulation and Computation

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/lssp20

Hierarchical Bayes estimation of small area means under a spatial nonstationary Fay–Herriot model

Priyanka Anjoy & Hukum Chandra

To cite this article: Priyanka Anjoy & Hukum Chandra (2021): Hierarchical Bayes estimation of small area means under a spatial nonstationary Fay–Herriot model, Communications in Statistics - Simulation and Computation, DOI: 10.1080/03610918.2021.1926501

To link to this article: https://doi.org/10.1080/03610918.2021.1926501

Published online: 24 May 2021.

COMMUNICATIONS IN STATISTICS - SIMULATION AND COMPUTATION®
https://doi.org/10.1080/03610918.2021.1926501

Taylor & Francis
Taylor & Francis Group

Check for updates

Hierarchical Bayes estimation of small area means under a spatial nonstationary Fay–Herriot model

Priyanka Anjoy and Hukum Chandra

Department of Agricultural Statistics, ICAR-Indian Agricultural Statistics Research Institute, New Delhi, India

ABSTRACT

The Fay–Herriot (FH) model is widely used in small area estimation (SAE) for aggregated level data, but in several applications presence of spatial effect between contiguous or neighboring region cannot be denied which is not handled by this model. Conditional Autoregressive and Simultaneous Autoregressive specifications do incorporate spatial associationship while taking into account the spatial correlation effects among areas. However, none of these approaches implement the idea of spatially varying covariates through spatially dependent fixed effect parameters. Such approach in statistics is known as spatial nonstationarity. This article introduces spatial nonstationary version of FH model considering hierarchical Bayesian paradigm and then deliberates estimation of small area means. The proposed SAE approach is evaluated through extensive simulation studies. The empirical results from simulation studies demonstrate the superiority of proposed spatial nonstationary SAE method over the non-spatial and stationary alternatives. The method is also applied to estimate paddy (green) crop yield at district level in the state of Uttar Pradesh in India using survey data from the improvement of crop statistics scheme and linked with Census data. A spatial map presents a quick view to the regional variations or disparity in district level yield estimates and are certainly helpful to the decision makers for identifying the regions and areas requiring more attention for designing targeted interventions and policy development.

ARTICLE HISTORY
Received 30 July 2020
Accepted 1 May 2021

KEYWORDS
Hierarchical Bayes; Small area estimation; Spatial nonstationarity

MATHEMATICAL SUBJECT CLASSIFICATION
62H30

1. Introduction

Sample surveys are designed to produce reliable and concrete statistical inference about the target population. Direct survey estimates based on area specific sample data are known to represent large target regions or aggregate of small areas (such as national, state, province etc.). But small area inference based on these direct estimates typically fails due to nonavailability of sufficient area-specific sample sizes. By small area we mean subpopulations or domains which were unplanned during designing of large-scale sample surveys. Contextually, model-based approaches have continually gained attention to provide acceptable estimates for such small area or small domain, popularly known as small area estimation (SAE) approach. In the context of the 2030 agenda of sustainable development goals (SDGs), there is continual emphasis on decentralized level statistics for micro level planning, policy formulation and targeted upliftment. To reconcile the need for reliable and representative disaggregate level official statistics, SAE is very relevant and need of the day. Two type of small area models are basically practiced in various real life

CONTACT Priyanka Anjoy anjoypriyanka90@gmail.com Department of Agricultural Statistics, ICAR-Indian Agricultural Statistics Research Institute, Library Avenue, New Delhi, India.

applications of SAE approach. Unit level models are implemented wherever we have unit-specific variable information; area level models are utilized with aggregated variable information on target and auxiliary variates. Small area inferences drawn from these models are essentially based on the idea of borrowing information/strength from related areas and sources (such as census or administrative record) to improve effective sample sizes of particular small domain (Rao and Molina 2015). Hence, small area model-based approach finally results in precise and reliable estimates, that is, smaller percentage coefficient of variation (CV) compared to those direct survey estimates (You and Zhou 2011). Since a decade economic planning is becoming more decentralized, therefore importance of micro-level statistics at lower level of administration cannot be undermined. Micro-level statistics is also essential to target social and spatial heterogeneity in the programmes and strategies aimed at alleviating the inter-personal and inter-regional inequalities.

This article focuses on area (or aggregate) level small area models to improve the direct survey estimates. The pioneering work of Fay and Herriot (1979) has yielded Fay–Herriot (FH) model which is implemented at a great scale to draw needful area level small area inference. The mixed modeling framework of area level FH model allows us to incorporate fixed effect as well as area random effects. A good number of covariates in the fixed effect part certainly influences the parameter estimation, but random effect component captures the unexplained heterogeneity between areas beyond that is revealed by auxiliary information (Rao 2003). However, a restrictive assumption on area random effects is that random errors are independent, identical and normally distributed. Such restrictive assumption is necessary for mean squared errors (MSE) estimation working under a frequentist perspective. But difficult to justify the validity of such postulation in various real life situations particularly variables involving spatial association among geographical units or areas (You and Zhou 2011; Chandra, Salvati, and Chambers 2017). In agricultural, environmental or health estimation problems application of spatial models are therefore quite reasonable because of the presence of spatial correlation among areas. Area level version of Conditional Autoregressive (CAR) and Simultaneous Autoregressive (SAR) are popular and widely implemented to provide domain-specific reliable estimates in case of spatial dependency (Pratesi and Salvati 2008; Chandra 2013). However, it's worth noting that, one common consideration in the discussed area level FH, SAR, CAR or other spatial models are that simple 'global model' is advocated to explain any kind of relationship that exists between the given set of variables. Such approach is basically spatial stationarity, where fixed effect parameters of the model do not vary spatially. Whether, in some study cases we cannot restrict to a single global model and nature of the model must vary across spaces to reflect the structure within the data (Brunsdon, Fotheringham, and Charlton 2010). This is the case of spatial nonstationarity. This approach is quite analogous to geographically weighted regression (GWR) in a multiple regression model which allows different relation to exist between study and auxiliary variates to exist at different points in space. The attempt in this article is to conceptualize the GWR version of area level of small area model to yield spatial nonstationary FH model. A hierarchical Bayes (HB) paradigm is proposed to obtain small area or domain level estimates through this model. Developed approach is motivated by a study aimed to obtain district level estimates of paddy (green) yield in the state of Uttar Pradesh in India using survey data from the Improvement of crop statistics (ICS) scheme and linked with Indian Population Census. Earlier, Chandra, Salvati, and Chambers (2015) has proposed nonstationary empirical best linear unbiased predictor (NSEBLUP) to obtain precise area level small area estimates in presence of spatial nonstationarity. In contrast, this article discusses a HB framework to attain spatially smoothen Bayes estimates at subpopulation level. Bayesian approach is somewhat more flexible than frequentist framework yielding quick and easier MSE computation which is posterior variance; additionally posterior mean or point estimate known to include more reasonable credible interval region.

Rest of the article is organized as follows. Next section provides description of paddy (green) crop yield data collected under the ICS scheme as motivational example. Section 3 delineates methodological discussion and development. Simulation studies are furnished in Sec. 4 followed by an application to obtain district level estimates of paddy (green) yield in Uttar Pradesh using proposed SAE method. The article concludes with relevant concluding remarks.

2. Descriptions of data

In India, most of the large scale surveys are planned at higher aggregation level and provide valid direct estimates for state and nation, whereas any planning at smaller administrative units like districts, municipalities, gram panchayats require survey designing at this stage which are both costly and time consuming. Therefore, SAE method can be a crucial and acceptable alternative to provide reliable statistics at disaggregate or micro level (e.g. districts, municipalities, gram panchayats etc.) from the existing surveys. Agriculture is one of the key drivers of Indian economy; this sector is such a crucial that prosperity of agrarian community is essential for even Govt./ institutional stability. Accurate estimation of yield and productivity of different crops hold utmost importance therefore to formulate policy actions undertaken by the government departments in order to monitor the progress of agriculture sector and deliver insurance support. Crop-cutting experiments (CCEs) conducted under the scheme of general crop estimation surveys (GCES) accurately estimate crop yield during cultivation cycle. The data gathered from CCE are useful to the multiple stakeholders in the agricultural value chain, especially to the Govt. and financial institutions to extend insurance and loan coverage to the farmers in case of poor harvest or failure. But due to huge spread and volume of field level and compilation work under GCES, quality of such data is objectionable. Therefore, a scheme entitled ICS has been introduced by Government of India to carry out quality check and supervision of around 30,000 CCEs every year. But this comes with the compromise of reduced sample sizes under ICS whereas of better quality. As a consequence, direct survey estimates of yield (based on ICS data) produced at disaggregate level like districts are not acceptable due to high degree of sampling variability (i.e. CV). The endeavor of SAE methodology is a practical and proficient alternative in this context to provide district level estimate of crop yield with reasonable precision.

An inadequate sample size under ICS has been one of the significant hindrances to provide reasonable estimates of crop yield at district level. For this study, in the state of Uttar Pradesh ICS data of paddy yield collected during the year 2009–2010 is available for 58 districts only and there is no sample data for the remaining 12 districts. Study variable is yield rate for paddy (green) crop recorded as gram per 43.12 m^2 based on equilateral triangle CCE plot of side 10 m each. District specific sample sizes for the 58 sampled districts ranges from 4 to 28 with median of sample sizes 10. Figure 1 is portrayed for visual scrutiny of district specific sample size distribution. With the few district specific sample sizes traditional survey estimation approach leads to imprecise estimates, further there is no design based solution to obtain estimates for 12 out-of-sample districts. This motivates to carry forward SAE approach instead of pertaining to traditional design based option. The production of reliable small area estimates is based on the availability of accurate auxiliary information. The auxiliary variables for this study at small area (i.e. district) level comes from Indian Population Census 2011. In the original data file, there are more than 121 available covariates. Initial scrutiny of these variables identified a group of potential auxiliary variables to be used for the study. This has been done based on measuring correlation between direct survey estimates and pool of available variables from census database. Finally, step-wise regression method was used to select auxiliary variables for SAE which significantly explained the model. See for example, Chandra (2013). The final selected auxiliary variables for the small area model were average household size (AHS) and female population of marginal household (FPMH); checks on spatial nonstationarity on these two variables were also done. Refer Table 1 for descriptive measure of the auxiliary variable values and sample sizes over

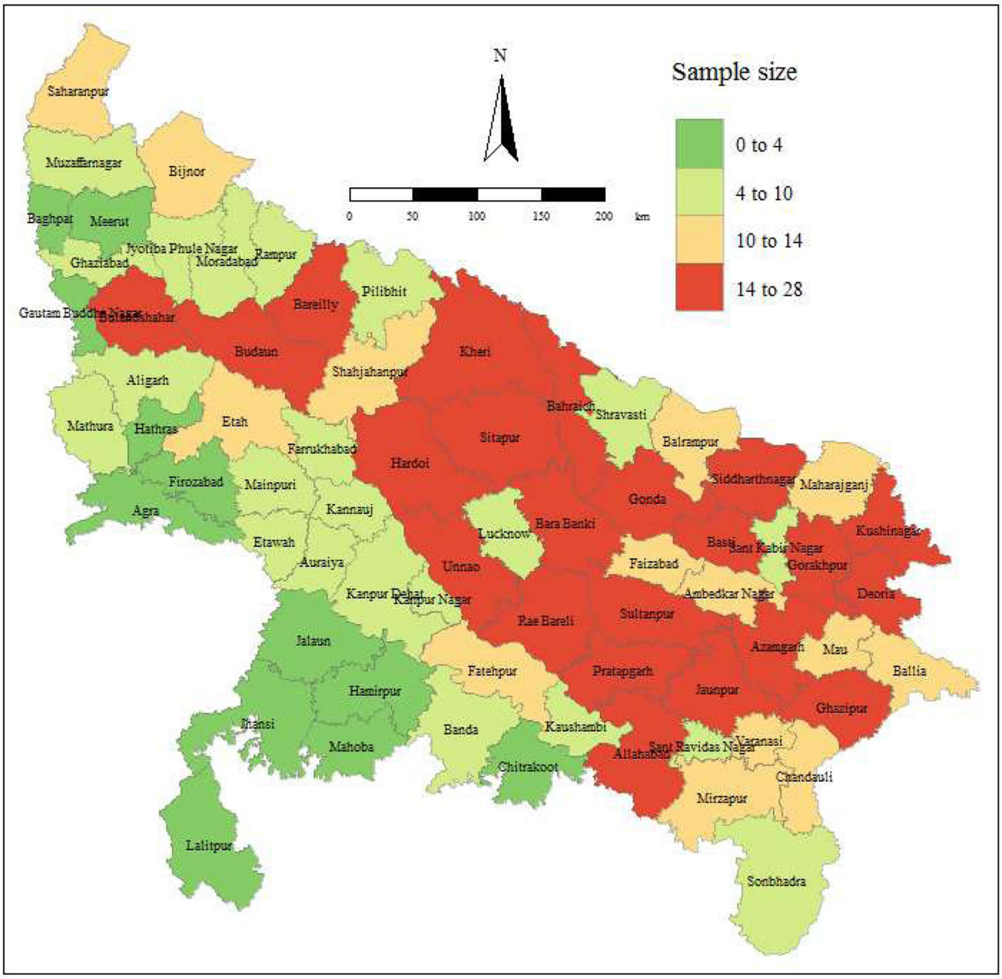

Figure 1. Map showing distribution of district specific sample sizes.

Table 1. Descriptive statistics for the auxiliary variables and district specific sample sizes.

Descriptive statistics	AHS	FPMH	Sample size
Minimum	5.66	1425.0	0
Q1	6.11	3791.8	4
Mean	6.46	8636.3	9
Median	6.49	7230.0	10
Q3	6.66	10422.8	14
Maximum	8.36	39002.0	28

all the districts or small areas. Note that for SAE of 12 out-of-sampled districts same two covariates were used, since the underlying model for sample areas also holds for out-of-sample districts.

3. Methodology setup

In small area applications, area (or aggregate) level models are widely used when unit-level data are unavailable, or, as is often the case, where auxiliary variables are only available in aggregate form. The area level models also offer flexibility in combining different sources of information with different error structures. After the pioneering work of Fay and Herriot (1979) on area level small area model (popularly referred as Fay Herriot (FH) model), till date the volume of small area literatures and methodological inventions has taken a gigantic form. Basis structure of the FH model includes a sampling model for the direct survey estimates and a linking model to incorporate auxiliary information as well as area specific random effect which probably explains unstructured variations among areas not countered by fixed effect part (auxiliary variables). But, in FH model an implicit independence assumption is also imposed on the random effect component which implies different small areas are simply uncorrelated. However, in agricultural, environmental surveys spatial dependence between neighboring areas cannot be denied. Thus, to incorporate the neighboring effect, it is reasonable to construct spatial model to capture the spatial association between areas. In this context Pratesi and Salvati (2008); Chandra (2013) has extended the FH model to incorporate spatially correlated random effects using CAR and SAR specifications. These models define the dependence between areas by using certain contiguity matrix, which can be obtained by using coordinates of the centroid of each small area, its geometric properties (extension, perimeter, etc.) and the neighborhood structure (Baldermann, Salvati, and Schmid 2018). Additionally, You and Zhou (2011); Anjoy and Chandra (2019) have pertained to the same concept of using spatial model by SAR specification under Bayesian framework. Chandra, Salvati, and Chambers (2015) and Chandra, Salvati, and Chambers (2017) have investigated the spatial association between neighboring areas via spatial nonstationary process under frequentist framework. In contrast, this article introduces spatial nonstationary version of FH model (NSFH) under a hierarchical Bayesian (HB) framework to estimate small area means. The key feature of spatial nonstationary process is that here spatial effect is added via spatially varying covariates, that is, regression parameters vary spatially. Again, one of the strategic advantages of using Bayes framework is that here estimations are described by assuming particular probability distributions, which render the opportunities to analyze the uncertainties involved in the decision process. Bayesian approach of SAE leads to more reasonable interval estimates (Anjoy, Chandra, and Basak 2019). What follows, we first delineate the FH model followed by the spatial version of FH (SFH) model of Anjoy and Chandra (2019) and then proposed NSFH model in hierarchical Bayesian framework.

3.1. Hierarchical Bayes Fay–Herriot (HBFH) method of SAE

Let D be the number of small areas (or simply areas) in the population. We use a subscript i to index the quantities belonging to area i. Let y_i denotes the direct survey estimate of population parameter (e.g. the population mean, total or some derived function of mean or total) θ_i of a variable of interest y for area i. Let $\boldsymbol{x}_i = (1, x_{i1}, \dots, x_{ip-1})'$ be the p-vector of auxiliary variables for area i, often obtained from administrative and census records, related to the population parameter θ_i. The simple area specific two stage model suggested by Fay and Herriot (1979) is

$$y_i = \theta_i + e_i \text{ and } \theta_i = \mathbf{x}'_i\boldsymbol{\beta} + v_i. \tag{1}$$

The first part (also referred as sampling model) of model (1) accounts for the sampling variability of the direct survey estimates y_i of population parameter θ_i and the second part (i.e. linking model) links the population parameter θ_i to a vector of known auxiliary variables $\mathbf{x}_\mathrm{i}$. Combining the two components of model (1), the FH model can be expressed as a random effect model of form

$$y_i = \mathbf{x}'_i\boldsymbol{\beta} + v_i + e_i, \quad i = 1, \dots, D \tag{2}$$

where $\boldsymbol{\beta} = (\beta_0, \dots, \beta_{\mathrm{p}-1})'$ is the p-vector of unknown of regression coefficients and v_i being the area specific random effect which is independent and identically (i.i.d) distributed with $E(v_i)= 0$ and $\mathrm{var}(v_i) = \sigma_v^2$. Here e_i is independent sampling error associated with direct survey estimator

y_i. It is assumed that $E(e_i|\theta_i) = 0$ and $var(e_i|\theta_i) = \sigma_{ei}^2$. The two random errors are independent of each other within and across areas. Usually the sampling variances $\sigma_{ei}^2 (i = 1, ..., D)$ are assumed to be known and these are obtained from survey data considering the underlying survey design. However, various Bayesian SAE literatures also reports the cases where sampling variances are assumed to be unknown and derived out following χ^2 distribution or through using design effect (You and Zhou 2011; Liu, Lahiri, and Kalton 2014). Aggregating D area level model (2) leads to population level FH model of form

$$\mathbf{y} = \boldsymbol{\theta} + \mathbf{e} = \mathbf{X}\boldsymbol{\beta} + \mathbf{v} + \mathbf{e}, \tag{3}$$

where $\mathbf{y} = (y_1, ..., y_D)'$ is the $D \times 1$ vector of direct survey estimates, $\boldsymbol{\theta} = (\theta_1, ..., \theta_D)'$ is the $D \times 1$ vector of population parameters, $\mathbf{X} = (\mathbf{x'}_1, ..., \mathbf{x'}_D)'$ is the $D \times p$ matrix of auxiliary variables whose i-th row is given by $\mathbf{x'}_i$, $\mathbf{v} = (v_1, ..., v_D)'$ is the D-vector of random area effects with $\mathbf{v} \sim N(\mathbf{0}, \sigma_v^2 \mathbf{I}_D)$ and $\mathbf{e} = (e_1, ..., e_D)'$ is the D-vector of sampling errors with $\mathbf{e} \sim N(\mathbf{0}, \boldsymbol{\Sigma}_e)$, where $\boldsymbol{\Sigma}_e = \text{diag}\{\sigma_{ei}^2; 1 \le i \le D\}$ is the known matrix of design variances. Further, it is assumed that the vector of area effects $\mathbf{v}$ is distributed independently of the sampling errors $\mathbf{e}$, so that the covariance matrix of the vector $\mathbf{y}$ is $\text{Var}(\mathbf{y}) = \mathbf{V} = \sigma_v^2 \mathbf{I}_D + \boldsymbol{\Sigma}_e$, where $\mathbf{I}_D$ is the identity matrix of order D. The parameters σ_v^2 and $\boldsymbol{\Sigma}_e$ are often referred to as the variance components of model (3). With this, we attempt to draw small area inference for population parameter vector $\boldsymbol{\theta}$ (equivalently each $\theta_i, i = 1, ..., D$) through HB approach by implementing Gibbs sampling method. The HB version of FH (HBFH) model can be expressed as

$$\begin{array}{l} \text{Sampling model} : \mathbf{y}|\boldsymbol{\theta} \sim \text{N}(\boldsymbol{\theta}, \boldsymbol{\Sigma}_e) \quad \text{and} \\ \text{Linking model} : \boldsymbol{\theta}|\boldsymbol{\beta}, \sigma_v^2 \sim \text{N}(\mathbf{X}\boldsymbol{\beta}, \sigma_v^2 \mathbf{I}_D). \end{array} \tag{4}$$

Following standard literature prior choice for $\boldsymbol{\beta}$ is usually taken to be $N(0, \sigma_0^2)$ and for σ_v^2 *Inverse Gamma*(a_0, b_0) where σ_0^2 is set to be very large (say, 10^6) and very small values for a_0 and b_0 (usually $a_0 = b_0 \to 0$) to reflect lack of prior knowledge about variance parameters (Rao 2003; You and Zhou 2011; Liu, Lahiri, and Kalton 2014; Anjoy, Chandra, and Basak 2019). Hereafter, this method of SAE is referred as HBFH.

3.2. Hierarchical Bayes spatial Fay Herriot (HBSFH) method of SAE

The FH or HBFH model implicitly assumes that direct survey estimates from different small areas are uncorrelated. However, in practice the boundaries that define a small area are typically arbitrary, and there appears to be no good reason why neighboring areas should not be correlated. It is therefore often reasonable to assume that the effects of neighboring small areas, defined via a contiguity criterion, are correlated (Pratesi and Salvati 2008). In small area modeling incorporating the information of spatial dependence between neighboring areas often improves the model accuracy. Therefore, to incorporate spatial information linking model with spatial dependence in error structure, so called SAR error process is often used. Let, define the random area effect $\mathbf{u}$ satisfy

$$\mathbf{u} = \rho \mathbf{W}\mathbf{u} + \mathbf{v}, \tag{5}$$

where ρ is the spatial autoregressive coefficient measuring the strength of spatial relationship and $\mathbf{W}$ is the proximity or contiguity matrix defining how random effects from neighboring areas are related. Contiguity matrix $\mathbf{W}$ provides a simplest way to define spatial interaction between adjoining small areas. Different choices of $\mathbf{W}$ matrix haven been in practice in the literature (Chandra 2013). In this article, we consider the contiguity matrix with element $w_{jk} (j, k = 1, ..., D)$ taking the value 1 if area j shares an edge with area k and 0 otherwise. In particular a row-standardized form of contiguity matrix is used. We can also rewrite, $\mathbf{u} = (\mathbf{I} - \rho\mathbf{W})^{-1}\mathbf{v}$ with $\mathbf{v} \sim N(\mathbf{0}, \sigma_v^2 \mathbf{I}_D)$ so $E(\mathbf{u}) = 0$ and $Var(\mathbf{u}) = \sigma_v^2 [(\mathbf{I}_D - \rho\mathbf{W})(\mathbf{I}_D - \rho\mathbf{W}')]^{-1}$. Following Anjoy and Chandra (2019), the spatial dependent HBFH (HBSFH) model is given by

$$\begin{aligned} &\text{Sampling model}: \mathbf{y}|\boldsymbol{\theta} \sim \mathrm{N}(\boldsymbol{\theta}, \boldsymbol{\Sigma}_e) \quad \text{and} \\ &\text{Linking model}: \boldsymbol{\theta}|\boldsymbol{\beta}, \rho, \sigma_v^2 \sim \mathrm{N}\Big(\mathbf{X}\boldsymbol{\beta}, \sigma_v^2\big[(\mathbf{I}_D - \rho\mathbf{W})(\mathbf{I}_D - \rho\mathbf{W}')\big]^{-1}\Big). \end{aligned} \tag{6}$$

In HBSFH model (6), prior choice for β is $N(0,\ 10^{-6})$; prior for hyperparameter σ_v^2 taken as *Inverse Gamma*(a_0, b_0) and prior for spatial autoregressive coefficient ρ is *Uniform* (-1,1).

3.3. Hierarchical Bayes nonstationarity Fay Herriot (HBNSFH) method of SAE

The FH model (1) postulates that fixed-effect parameter or regression coefficient vector $\boldsymbol{\beta}$ does not vary spatially, that is, $\boldsymbol{\beta}$ is spatially invariant, this is the case of spatial stationarity. The HBSFH model (6) allows for spatial correlation in the area effects but it also assumes the same invariant form of $\boldsymbol{\beta}$ (Anjoy and Chandra 2019). There may be data situation where model parameter varies spatially which referred as spatial nonstationarity (Opsomer et al. 2008; Baldermann, Salvati, and Schmid 2018). Regression coefficients in the small area model therefore may be expressed as explicit functions of the spatial locations of the sample observations instead of defining one single global model with fixed parameter. Brunsdon, Fotheringham, and Charlton (2010) was pioneering in forwarding the concept for handing such situation of spatial nonstationarity in regression model, which is through GWR model. In area level model, Chandra, Salvati, and Chambers (2015, 2017) has contributed NSEBLUP and nonstationary generalized linear mixed model (NSGLMM). This article adds another step to deal with spatial nonstationarity in SAE field through area level HBNSFH model. Analytic MSE expression of NSEBLUP model is quite complex and based on very some approximation (Chandra, Salvati, and Chambers 2015). In contrast, the strategic advantage in considering HB approach is that, here estimations are described by taking particular probability distributions which render the opportunities to analyze the uncertainties involved in the decision process. In the HB method, together with prior distribution of the parameters, prior of the hyper-parameters (model parameters) are also specified then inferences are made from the posterior distributions. A parameter is estimated by posterior mean and posterior variance is taken as the measure of the error or uncertainty of the estimates. The HB approach can effectively deal with complex small area models using Monte Carlo Markov Chain (MCMC), which overcomes the computational difficulties of high-dimensional integrations of posterior densities (You and Rao 2002).

We now define a spatial nonstationary extension of FH model. Let l_i denote the spatial location of area i which corresponds to the coordinates (longitude and latitude) of an arbitrarily defined spatial location in the area. Typically, this will be its centroid. Let $L(l_i, l_j)$ be an appropriate measure of the distance between the spatial locations of areas i and j, and define the spatial contiguity of these two locations to be $\omega_{ij} = (1 + L(l_i, l_j))^{-1}$. Let $\mathbf{W} = (\omega_{ij})$ denote the positive definite $D \times D$ matrix of spatial contiguities defined by the l_i. This spatial contiguity matrix is assumed to be known. Following Chandra, Salvati, and Chambers (2015), a spatial nonstationary version of FH (NSFH) model for area i is given by

$$y_i = \mathbf{x}'_i\boldsymbol{\beta}(l_i) + v_i + e_i = \mathbf{x}'_i\boldsymbol{\beta} + \mathbf{x}'_i\boldsymbol{\gamma}(l_i) + v_i + e_i, \tag{7}$$

where $\boldsymbol{\beta}(l_i) = \boldsymbol{\beta} + \boldsymbol{\gamma}(l_i)$, v_i is the area-specific random effect, assumed to follow a normal distribution with zero mean and variance σ_v^2, that is, $v_i \sim N(0, \sigma_v^2)$ and e_i is independent sampling error associated with y_i, assuming that $e_i \sim N(0, \sigma_{ei}^2)$. Again, independence of these two error terms e_i and v_i are also assumed. Here $\boldsymbol{\gamma}(l_i) = (\gamma_k(l_i); k = 1, \ldots, p)$ is a spatially correlated vector-valued random process of dimension p with $E(\boldsymbol{\gamma}(l_i)) = \mathbf{0}_{p\times 1}$ and $\mathrm{cov}(\gamma_k(l_i), \gamma_m(l_j)) = a_{km}(1 + L(l_i, l_j))^{-1}; k, m = 1, \ldots, p$, where $\mathbf{a} = (a_k)$ is a p-vector of unknown positive constants that satisfies the conditions for the $pD \times pD$ matrix $\boldsymbol{\Sigma}_\gamma = \mathbf{W} \otimes (\mathbf{aa}')$ to be a covariance matrix, where $\otimes$ denotes Kronecker product. Let $\mathbf{l} = (l_1, \ldots, l_D)'$ be the D-vector of spatial locations, that is, the set of locations for the D areas, $\mathbf{Z} = \{diag(\mathbf{x}_1), \ldots, diag(\mathbf{x}_D)\}'$ be the $D \times pD$ matrix of known auxiliary data, and $\mathbf{\Gamma} = (\boldsymbol{\gamma}'(l_1), \ldots, \boldsymbol{\gamma}'(l_D))'$ be a $pD \times 1$ vector of spatial normal random effects that capture the spatial nonstationarity in the data. We assume that $\mathbf{\Gamma}$ has a zero mean vector and a covariance matrix $\mathbf{\Sigma}_\gamma$. That is, $\mathrm{E}(\mathbf{\Gamma}|\mathbf{Z}, \mathbf{l}) = \mathbf{0}_{pD\times 1}$ and $\mathrm{Var}(\Gamma|\mathbf{Z}, \mathbf{l}) = \mathbf{\Sigma}_\gamma$. Recollecting different terms, we can express the population level version of NSFH (7) as

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{\Gamma} + \mathbf{v} + \mathbf{e}, \tag{8}$$

with

$$\mathrm{E}(\mathbf{y}|\mathbf{Z}, \mathbf{l}) = \mathbf{X}\boldsymbol{\beta}, \quad Var(\mathbf{y}|\mathbf{Z}, \mathbf{l}) = \mathbf{V} = \mathbf{Z}\mathbf{\Sigma}_\gamma\mathbf{Z}' + \sigma_v^2\mathbf{I}_D + \mathbf{\Sigma}_e \quad \text{and} \quad \mathrm{Cov}(Y_i, \mathbf{y}|\mathbf{Z}, \mathbf{l}) = \mathbf{Z}_i\mathbf{\Sigma}_\gamma\mathbf{Z}' + \sigma_v^2\boldsymbol{\delta}_i,$$

where $\mathbf{Z}_i$ is the i^{th} row of $\mathbf{Z}$, $\boldsymbol{\delta}_i$ denotes the i^{th} row of $\mathbf{I}_D$ and $\mathbf{\Sigma}_e = diag\{\sigma_{ei}^2; i = 1, \ldots, D\}$. In practice, the variance component parameters σ_v^2 and $\mathbf{a}$ are unknown and have to be estimated from the data. Following Chandra, Salvati, and Chambers (2015), in this article we restrict to the simple specification $\mathbf{a} = \sqrt{\eta}\mathbf{1}_p$ so that $\mathrm{cov}(\gamma_k(d_i), \gamma_l(d_j)) = \eta(1 + L(d_i, d_j))^{-1}$, where $\eta \geq 0$ and $\mathbf{1}_p$ denotes the unit vector of order p. In this case, we assume that the distance metric used to define $L(d_i, d_j)$ is such that the matrix $\mathbf{\Sigma}_\gamma = \eta\mathbf{W} \otimes (\mathbf{1}_p\mathbf{1}'_p)$ is positive semidefinite, with the parameter η then reflecting the "intensity" of spatial clustering in the data, so $\eta = 0$ corresponds to the situation where the model is spatially homogeneous. The HB version of NSFH model (8) is expressed as

$$\begin{array}{l}\text{Sampling Model}: \mathbf{y}|\boldsymbol{\theta} \sim \mathrm{N}(\boldsymbol{\theta}, \mathbf{\Sigma}_e) \quad \text{and} \\ \text{Linking model}: \boldsymbol{\theta}|\boldsymbol{\beta}, \eta, \sigma_v^2 \sim \mathrm{N}(\mathbf{X}\boldsymbol{\beta}, \mathbf{Z}\mathbf{\Sigma}_\gamma\mathbf{Z}' + \sigma_v^2\mathbf{I}_D).\end{array} \tag{9}$$

The prior choice for hyper-parameter $\boldsymbol{\beta}$ is usually taken to be $N(0, \sigma_0^2)$ and for variance parameter η and σ_v^2 *Inverse Gamma*(a_0, b_0) where σ_0^2 is set to be very large (say, 10^6) and very small for a_0 and b_0 (usually $a_0 = b_0 \to 0$) to reflect lack of prior information. Gibbs sampling method is implemented to estimate posterior mean$E(\theta_i|\mathbf{y})$ and posterior variance$var(\theta_i|\mathbf{y})$. Now onwards, we refer this method of SAE as HBNSFH. The required full conditional distributions for the Gibbs sampler under HBNSFH model (9) are given as,

$$\begin{aligned}
&\boldsymbol{\theta}|\boldsymbol{\beta}, \eta, \sigma_v^2, \mathbf{y} \sim \mathrm{MVN}\left[\mathbf{X}\boldsymbol{\beta} + (\mathbf{Z}\mathbf{\Sigma}_\gamma\mathbf{Z}' + \sigma_v^2\mathbf{I}_D)\mathbf{V}^{-1}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}), (\mathbf{Z}\mathbf{\Sigma}_\gamma\mathbf{Z}' + \sigma_v^2\mathbf{I}_D)\mathbf{V}^{-1}\mathbf{\Sigma}_e\right], \\
&\boldsymbol{\beta}|\boldsymbol{\theta}, \eta, \sigma_v^2 \sim \mathrm{MVN}\left[(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}(\mathbf{X}'\mathbf{V}^{-1}\boldsymbol{\theta}), (\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}(\sigma_v^2\mathbf{I}_D + \mathbf{Z}\mathbf{\Sigma}_\gamma\mathbf{Z}')\right], \\
&\sigma_v^2|\boldsymbol{\beta}, \eta, \boldsymbol{\theta} \sim \mathrm{IG}\left[a_1 + \frac{D}{2}, b_1 + \frac{(\boldsymbol{\theta} - \mathbf{X}\boldsymbol{\beta} - \mathbf{Z}\mathbf{\Gamma})'(\boldsymbol{\theta} - \mathbf{X}\boldsymbol{\beta} - \mathbf{Z}\mathbf{\Gamma})}{2}\right], \text{and} \\
&\eta|\boldsymbol{\beta}, \sigma_v^2, \boldsymbol{\theta} \sim \mathrm{IG}\left[a_0 + \frac{D}{2}, b_0 + \frac{(\boldsymbol{\theta} - \mathbf{X}\boldsymbol{\beta} - \mathbf{v})'(\boldsymbol{\theta} - \mathbf{X}\boldsymbol{\beta} - \mathbf{v})}{2}\right].
\end{aligned}$$

4. Empirical evaluation

In this section we illustrate model based simulation studies to compare the performance of small area estimates produced by HBFH, HBSFH, and HBNSFH model. The scenarios in the model based simulations are settings under spatial stationarity and nonstationarity with different prior choices. Data were generated using both stationary and nonstationary methods for $D = 49$ and 100 areas, respectively. Further, in nonstationary method, two approaches of data generation have been considered. The simulation study involves sensitivity analysis for distribution of variance parameter σ_v^2 with respect to shape and scale parameters of *Inverse Gamma* (*IG*) distribution. Accordingly, we have taken three different prior form, for example, *IG* (0.001, 0.001), *IG* (0.01, 0.01), *IG* (0.1, 0.1). In all cases prior for $\boldsymbol{\beta}$ has taken to be $N(0, 10^6)$ and prior choice for η in HBNSFH is same as σ_v^2. Hereafter, words "model" and "method" will be used interchangeably in the text.

4.1. Stationary data generation process

In stationary data generation process (SDGP) regression coefficients are spatially invariant, hence the aim is to explore how HBNSFH method performs as compared when the data follow usual HBFH model. Here the data has been generated using area level model:

$$y_i = 10 + 2x_i + v_i + e_i, \quad i = 1, \ldots, D,$$

where $x_i \sim Uniform[0,1]$; $v_i \sim N(0, \sigma_v^2 = 1)$ and independent sampling errors e_i generated from $N(0,\sigma_{ei}^2)$ with σ_{ei}^2 taking values 7,6,5,4,3, respectively, for equal number of areas (Datta, Rao, and Smith 2005).

For $D = 49$ areas $\{\sigma_{ei}^2\}_{i=1}^{10} = 7$; $\{\sigma_{ei}^2\}_{i=11}^{20} = 6$; $\{\sigma_{ei}^2\}_{i=21}^{30} = 5$; $\{\sigma_{ei}^2\}_{i=31}^{40} = 4$; $\{\sigma_{ei}^2\}_{i=41}^{49} = 3$.

For $D = 100$ areas $\{\sigma_{ei}^2\}_{i=1}^{20} = 7$; $\{\sigma_{ei}^2\}_{i=21}^{40} = 6$; $\{\sigma_{ei}^2\}_{i=41}^{60} = 5$; $\{\sigma_{ei}^2\}_{i=61}^{80} = 4$; $\{\sigma_{ei}^2\}_{i=81}^{100} = 3$.

4.2. Nonstationary data generation process

In nonstationary data generation process (NSDGP), regression parameters vary spatially, that is, spatially variant. Here two methods of DGP denoted, respectively, as NSDGP1 and NSDGP2 are illustrated. In NSDGP1 data is generated via GWR model adding an area specific random effect. The underpinning model for NSDGP1 is,

$$y_i = \beta_{0i} + \beta_{1i}x_i + v_i + e_i, i = 1, \ldots, D,$$

with

$\beta_{0i} = 10 + (2 \times \text{longitude}_i) + (0.5 \times \text{latitude}_i)$ and $\beta_{1i} = 4 \times \cos\left\{\text{sqrt}\left\{(1.2\pi \times \text{longitude}_i)^2 + (1.2\pi \times \text{latitude}_i)^2\right\}\right\}$. The distribution of auxiliary variable x_i, random effect v_i and sampling error e_i are same as defined in SDGP. To define longitude_i and latitude_i, it is assumed that observations has been drawn from a two-dimensional grid consist of a $\left(\sqrt{D}x\sqrt{D}\right)$ points uniformly spaced between -1 and 1 with a distance of $2/\left(\sqrt{D} - 1\right)$ between any two neighboring points along the vertical and horizontal axes. When $D = 49$, the lattice points where the observations are taken are$(\text{latitude}_i, \text{longitude}_i) = (k_1, k_2)$ where $\{k_1, k_2 = -1, -0.66, -0.33, 0, 0.33, 0.66, 1\}$; for $D = 100$, the set (k_1, k_2) is $\{k_1, k_2 = -1, -0.77, -0.55, -0.33, -0.11, 0.11, 0.33, 0.55, 0.77, 1\}$. The D points or spatial locations are therefore arranged in such a way that k_1 varies from -1 to 1 for each given k_2, which also then varies from -1 to 1.

For NSDGP2 data is generated via the following model,

$$y_i = 10 + 2x_i + \sqrt{\eta}\left(\gamma_0(l_i) + \gamma_1(l_i)x_i\right) + v_i + e_i, \quad i = 1, \ldots, D.$$

The values of η has been used as 2, 4, 6 in this study. The vector $(\gamma_0(l_i), \gamma_1(l_i))'$ has been defined as a random draw from $N(0, \mathbf{W} \otimes \mathbf{I}_2)$ with $\mathbf{W}$ being the distance matrix between lattice points or generated spatial locations(l_i, l_j). The lattice points for $D = 49$ and 100 areas are same as defined in NSDGP-1 with $l_i = (\text{latitude}_i, \text{longitude}_i)$, $i = 1, \ldots, D$. All other aspects of data generation with respect to distribution of x_i, v_i and e_i remains the same.

We generated $K = 500$ independent data sets in each specified scenario illustrated above under different prior set up, different number of areas and different DGP, then estimated small area population means using HBFH, HBSFH, and HBNSFH methods. We then compare the empirical performance and relative efficiency of proposed HBNSFH approach as compared to other nonspatial as well as spatial stationary alternatives. Under SDGP and NSDGP how the performance varies for HBFH, HBSFH, and HBNSFH estimates are noticed, along with performance of the small area estimators under each model are compared with respect to different prior cases. R and JAGS software has been used for implementation of the discussed models. To implement the Gibbs sampler, three independent chains are used each of length 10000. The

Table 2. Mean values for RB% and RRMSE% over $D = 49$ and 100 areas under different scenarios of data generation process and priors for HBFH, HBSFH, and HBNSFH methods of SAE.

		IG (0.001, 0.001)		*IG* (0.01, 0.01)		*IG*(0.1, 0.1)	
Priors	Criterion	RB%	RRMSE%	RB%	RRMSE%	RB%	RRMSE%
$D = 49$							
SDGP	HBFH	−0.142	9.285	−0.133	9.238	−0.138	9.168
	HBSFH	−0.185	9.344	−0.143	9.243	−0.125	9.596
	HBNSFH	−0.122	9.319	−0.125	9.273	−0.119	9.294
NSDGP1	HBFH	2.968	17.891	2.935	17.875	2.940	17.836
	HBSFH	1.811	16.247	1.808	16.220	1.780	16.153
	HBNSFH	1.593	15.684	1.605	15.684	1.597	15.653
NSDGP2	HBFH	0.389	11.010	0.364	10.846	0.367	10.858
(η=2)	HBSFH	0.366	11.001	0.290	10.839	0.312	10.842
	HBNSFH	0.360	10.991	0.268	10.831	0.273	10.823
NSDGP2	HBFH	0.734	12.083	0.735	12.057	0.734	12.080
(η=4)	HBSFH	0.729	12.005	0.674	11.994	0.662	11.992
	HBNSFH	0.604	11.978	0.579	11.936	0.592	11.938
NSDGP2	HBFH	0.805	12.652	0.787	12.571	0.795	12.599
(η=6)	HBSFH	0.805	12.604	0.773	12.491	0.784	12.495
	HBNSFH	0.804	12.582	0.765	12.453	0.779	12.457
$D = 100$							
SDGP	HBFH	−0.015	8.865	−0.010	8.813	0.006	8.759
	HBSFH	0.013	8.892	0.009	8.852	−0.016	8.785
	HBNSFH	0.001	8.906	0.005	8.876	0.004	8.886
NSDGP1	HBFH	2.259	16.161	2.255	16.162	2.267	16.159
	HBSFH	1.070	14.560	1.072	14.561	1.079	14.667
	HBNSFH	0.781	13.342	0.789	13.345	0.810	13.368
NSDGP2	HBFH	0.793	10.159	0.779	10.037	0.780	10.091
(η=2)	HBSFH	0.736	10.151	0.745	10.028	0.720	10.077
	HBNSFH	0.613	10.137	0.501	10.019	0.517	10.060
NSDGP2	HBFH	1.180	11.046	1.185	11.011	1.193	11.053
(η=4)	HBSFH	1.117	11.018	1.104	11.009	1.092	11.023
	HBNSFH	0.842	10.968	0.799	10.903	0.818	10.949
NSDGP2	HBFH	1.987	12.780	1.978	12.762	1.989	12.771
(η=6)	HBSFH	1.536	12.507	1.532	12.494	1.524	12.479
	HBNSFH	1.321	12.445	1.247	12.353	1.234	12.337

first 5000 iterations are deleted as "burn-in" periods. Based on $K = 500$ samples, the performance indicators calculated for comparison of models for each area i are:

- $\mathrm{RB}_i = \left(K^{-1}\sum_{k=1}^{K}\theta_i^{(k)}\right)^{-1}\left\{K^{-1}\sum_{k=1}^{K}(\hat{\theta}_i^{(k)} - \theta_i^{(k)})\right\} \times 100$ is the Relative Bias Percentage (RB%) for i^{th} domain, where $\hat{\theta}_i^{(k)}$ is the estimate of true population mean $\theta_i^{(k)}$ for i^{th} for small area at k^{th} simulation.
- $\mathrm{RRMSE}_i = \left(K^{-1}\sum_{k=1}^{K}\theta_i^{(k)}\right)^{-1}\left\{\sqrt{K^{-1}\sum_{k=1}^{K}(\hat{\theta}_i^{(k)} - \theta_i^{(k)})^2}\right\} \times 100$ is the Relative Root Mean Squared Error Percentage (RRMSE%) for i^{th} for small area.
- $\mathrm{TRMSE}_i = \sqrt{K^{-1}\sum_{k=1}^{K}(\hat{\theta}_i^{(k)} - \theta_i^{(k)})^2}$ is True or Simulation Root MSE (TRMSE) for i^{th} area.
- $\mathrm{ERMSE}_i = \sqrt{K^{-1}\sum_{k=1}^{K} mse_i^{(k)}}$ is the Estimated RMSE (ERMSE), where $mse_i^{(k)}$ is the posterior variance based on particular HB model pertinent to k^{th} simulation.
- $\mathrm{CR}_i = K^{-1}\sum_{k=1}^{K} I(LB(\hat{\theta}_i^{(k)}) \le \theta_i^{(k)} \le UB(\hat{\theta}_i^{(k)})) \times 100$ is the Coverage Rate (CR%) for i^{th} small area, where $LB(\hat{\theta}_i^{(k)})$ and $UB(\hat{\theta}_i^{(k)})$ are, respectively, Lower Bound (LB) and Upper Bound (UB) of the estimated population mean $\hat{\theta}_i^{(k)}$. Here $I(.)$ denotes an indicator function which takes values 1 if true parameter value $\theta_i^{(k)}$ is within the computed interval, otherwise

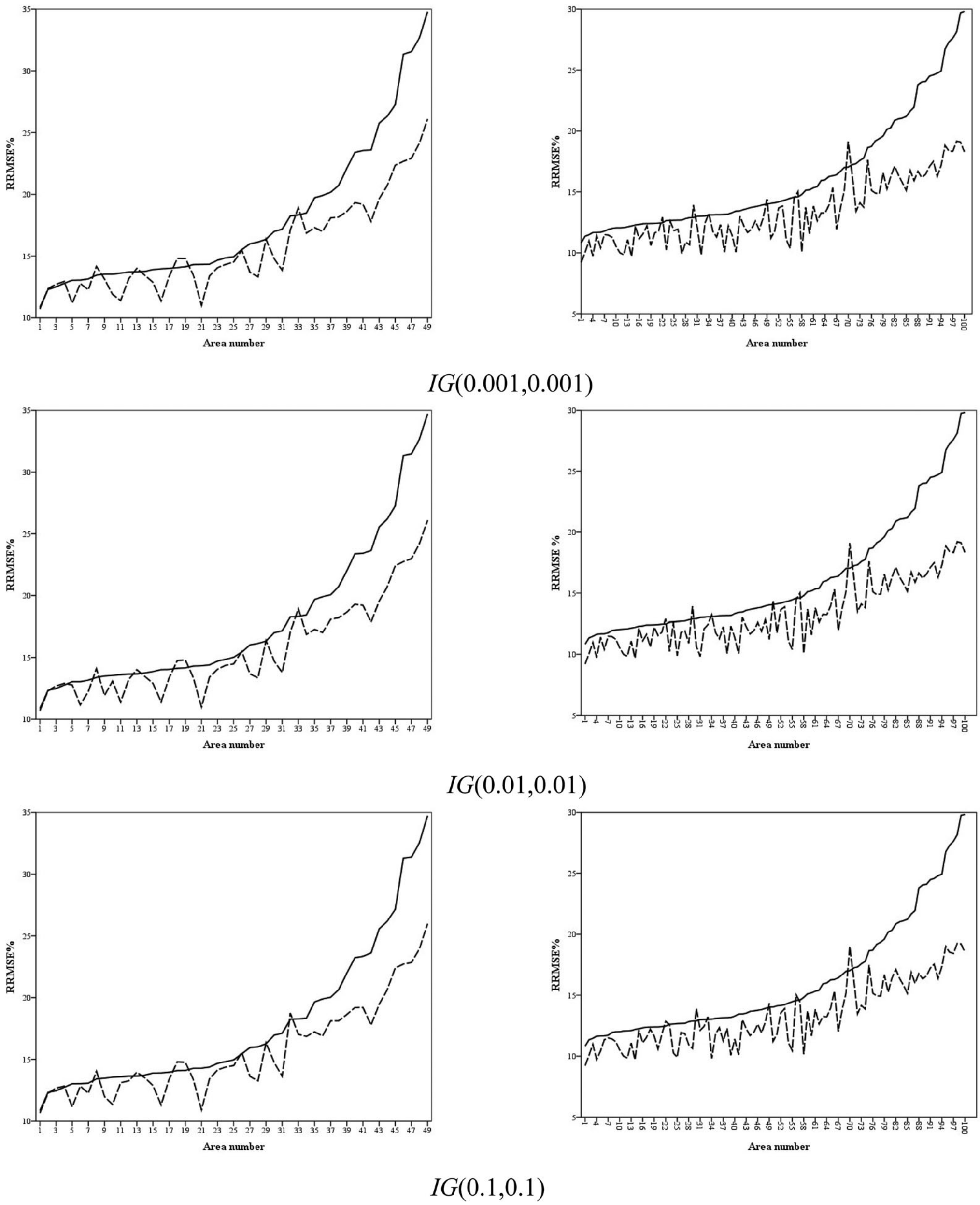

Figure 2. Plot of RRMSE% values over $D = 49$ (Right) and $D = 100$ (Left) small areas for NSDGP1 under different priors for HBFH (solid line) and HBNSFH (dash line) methods of SAE.

it takes value 0. This CR% particularly will demonstrate the credible interval property of HB models.

- $\text{ARB}(v)_i = TRMSE_i^{-1}|(TRMSE_i - ERMSE_i)| \times 100$ is the Absolute Relative Bias Percentage (ARB_v%) for variance or MSE terms.

A better model should show smaller values of all the above statistics expect CR%. Higher the CR% better is the model.

Table 3. Mean values of ARB$_v$%, CR%, TRMSE, ERMSE over $D = 49$ and 100 areas for NSDGP1 and NSDGP2 (η=6) under different priors for variance estimation of HBFH, HBSFH, and HBNSFH methods of SAE.

	$D = 49$				$D = 100$			
Areas Measure	ARB$_v$%	CR%	TRMSE	ERMSE	ARB$_v$%	CR%	TRMSE	ERMSE
IG(0.001, 0.001) Prior								
NSDGP1								
HBFH	13.39	91	1.569	1.481	13.18	94	1.410	1.409
HBSFH	9.18	92	1.437	1.452	8.28	94	1.283	1.281
HBNSFH	8.75	94	1.389	1.395	6.95	95	1.176	1.198
NSDGP2(η=6)								
HBFH	15.69	91	1.463	1.420	15.79	93	1.399	1.371
HBSFH	15.59	91	1.460	1.415	14.15	93	1.372	1.324
HBNSFH	15.55	92	1.456	1.409	14.11	93	1.366	1.317
IG(0.01, 0.01) Prior								
NSDGP1								
HBFH	13.38	91	1.568	1.483	13.17	94	1.411	1.410
HBSFH	9.14	92	1.435	1.479	8.23	94	1.283	1.279
HBNSFH	8.98	94	1.389	1.407	7.36	95	1.176	1.207
NSDGP2(η=6)								
HBFH	15.65	93	1.454	1.432	15.80	93	1.397	1.374
HBSFH	15.57	93	1.448	1.428	14.20	93	1.371	1.325
HBNSFH	15.54	94	1.442	1.424	14.01	94	1.357	1.336
IG (0.1, 0.1) Prior								
NSDGP1								
HBFH	13.39	92	1.565	1.483	13.24	94	1.410	1.378
HBSFH	9.78	93	1.429	1.486	8.60	94	1.347	1.281
HBNSFH	9.47	95	1.387	1.427	8.49	95	1.208	1.228
NSDGP2(η=6)								
HBFH	15.84	92	1.457	1.424	15.82	93	1.398	1.370
HBSFH	15.80	94	1.450	1.434	14.24	93	1.369	1.327
HBNSFH	15.77	93	1.442	1.425	14.10	93	1.355	1.329

4.3. Simulation results and discussion

This section presents the results of model based simulation study with respect to different DGP and different prior situations as described above. Results have been produced for $D = 49$ and 100 small areas, respectively. Table 2 reports the mean values for RB% as well as RRMSE% over the areas in different scenarios of DGP and prior cases. Table 2 shows that HBFH demonstrates relatively lower RRMSE% than HBSFH and HBNSFH method in case of SDGP. When the underlying data is stationary, it is expected that spatial stationary HBFH would perform better. This follows for all the cases of priors and $D = 49$ and 100 areas, respectively. Similarly, when the underlying data is nonstationary as in case of NSDGP1 and NSDGP2 as one would expect HBSFH and HBNSFH should perform better than HBFH, as both the models utilize spatial information. The result follows the same in terms of both RB% and RRMSE%. Additionally, as the number of areas increases ($D = 49$–100), impact of nonstationarity in the data becomes stronger. Therefore, gain in RRMSE% of HBNSFH model over HBFH improves. In particular, gain in RRMSE% is significantly higher for NSDGP1 than NSDGP2. Further, the HBNSFH consistently performs better over the HBSFH. Figure 2 portrays the plot of RRMSE% values for $D = 49$ and 100 small areas over all the priors for NSDGP1. Considering NSDGP1, for $D = 49$ areas the percentage gain in mean RRMSE% of HBNSFH model over HBFH model is 14.07, 13.97, and 13.94 for *IG*(0.001, 0.001), *IG*(0.01, 0.01), *IG*(0.1, 0.1) priors, respectively. Again, for $D = 100$ small areas the percentage gain in mean RRMSE% of HBNSFH over HBFH is 21.13, 21.11, and 20.88 for *IG*(0.001, 0.001), *IG*(0.01, 0.01), *IG*(0.1, 0.1) priors, respectively. Percentage improvement in mean RB% of HBNSFH over HBFH also considerably increases by increasing number of areas for both NSDGP1 and NSDGP2, also as we move from η =2 to higher value for NSDGP2. In these DGPs,

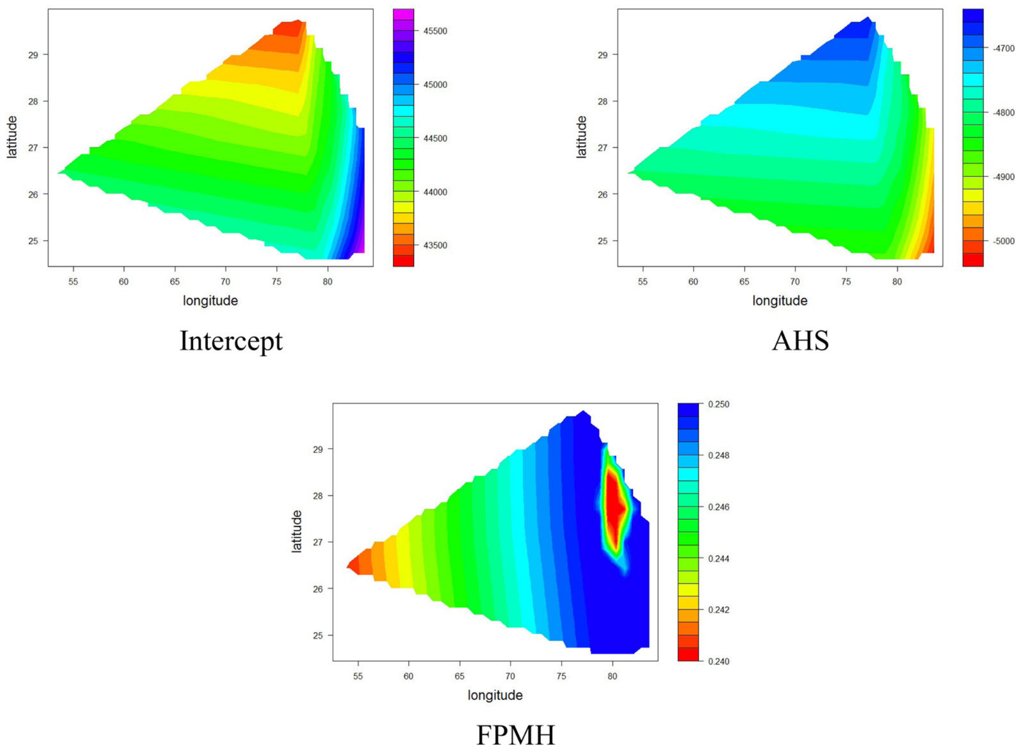

Figure 3. Contour maps showing the spatial variation in the district specific regression coefficients generated through GWR model fitting to the ICS data.

Table 4. Summary of %CV generated by the direct and different SAE methods for 58 sample districts.

Values	Direct	HBFH	HBSFH	HBNSFH
Minimum	3.01	3.00	3.00	2.99
Q1	10.04	9.57	9.74	9.45
Mean	15.14	13.02	12.71	12.30
Median	13.42	12.46	12.37	11.81
Q3	19.46	16.48	15.90	15.45
Maximum	49.15	29.14	26.24	22.78

the performance of HBSFH is in between HBFH and HBNSFH, it is definitely better than HBFH in terms of RB% and RRMSE% for all prior cases but performs poorly than HBNSFH. Table 2 and Figure 2 ensure the fact that HBNSFH is essentially better over HBFH for spatially nonstationary data. Further, it can be observed from Table 2 that the mean RRMSE% is not affected much by the use of different form of vague priors for variance parameter σ_v^2. Simulation results under different DGP are not influenced by the form of vague priors taken for the models.

Table 3 represents the mean values of ARB_v%, CR%, TRMSE, ERMSE over $D = 49$ and 100 areas for NSDGP1 and NSDGP2 (η=6) under different priors. The NSDGP1 shows considerably lower mean values of ARB_v% for HBNSFH as compared to HBSFH and HBFH in all prior situations. This indicates the smaller bias in estimating posterior variance for HBNSFH when comparing the values of TRMSE and ERMSE. The mean values of TRMSE and ERMSE are also reported in the Table 3, but ARB_v% shows a clear picture of better performing model. Further, gain in mean ARB_v% values for HBNSFH over HBFH improves by increasing the number of small areas from 49 to 100. Under NSDGP2 (η=6) for $D = 100$ areas, the improvement with respect to

Table 5. District wise estimates of paddy (green) crop yield (gram per 43.12 m^2) along with 95% credible interval and %CV for direct and HBNSFH methods of SAE.

		Direct				HBNSFH			
Districts	Sample Size	Estimates	Lower	Upper	%CV	Estimates	Lower	Upper	%CV
Saharanpur	10	19575	14574	24576	13.04	17852	13865	22005	11.58
Muzaffarnagar	6	23483	14035	32932	20.53	19050	13233	25208	15.90
Bijnor	12	19442	16669	22214	7.28	19089	16581	21741	6.91
Moradabad	8	17700	11916	23484	16.67	17944	13305	22438	13.24
Rampur	8	17250	16234	18266	3.01	17220	16195	18216	2.99
Jyotiba Phule Nagar	4	10850	7940	13760	13.68	11635	9022	14421	11.95
Ghaziabad	4	16800	6581	27019	31.03	14664	9059	20319	19.51
Bulandshahar	14	17418	13443	21393	11.64	17146	13349	20964	11.18
Aligarh	8	12419	7605	17232	19.77	12881	8675	16997	16.61
Mathura	4	10483	4880	16085	27.27	12069	7682	16516	18.92
Etah	10	12125	9813	14437	9.73	12344	10115	14471	9.15
Mainpuri	8	14019	7814	20224	22.58	14039	9558	18548	16.66
Budaun	14	12721	8968	16475	15.05	13060	9764	16379	12.71
Bareilly	14	13511	10021	17000	13.18	13825	10651	16888	11.82
Pilibhit	8	14938	9098	20777	19.94	15312	10930	19956	15.16
Shahjahanpur	12	18863	16560	21165	6.23	18431	16280	20597	6.12
Kheri	16	14975	11638	18312	11.37	15211	12079	18198	10.26
Sitapur	20	15986	11880	20093	13.11	15851	12304	19338	11.31
Hardoi	18	19286	16494	22078	7.39	18926	16317	21692	7.29
Unnao	14	12843	9841	15844	11.92	13440	10724	16200	10.38
Lucknow	8	17331	10170	24492	21.08	15466	10486	20457	16.54
Rae Bareli	18	19506	16053	22958	9.03	18186	15006	21431	8.88
Farrukhabad	5	8880	5582	12178	18.95	10193	7046	13352	15.75
Kannauj	4	34050	30416	37684	5.45	33250	29516	36809	5.61
Etawah	4	15463	13925	17000	5.07	15400	13928	16867	4.91
Auraiya	6	23717	19085	28348	9.96	21508	17641	25529	9.28
Kanpur Dehat	8	21200	16705	25695	10.82	19082	15271	22959	10.31
Kanpur Nagar	8	15375	10172	20578	17.27	15514	11377	19662	13.53
Banda	4	8888	326	17449	49.15	12321	6827	17838	22.78
Fatehpur	10	14612	8853	20371	20.11	14281	9853	18562	15.57
Pratapgarh	14	16304	11665	20942	14.52	15959	12086	19815	12.52
Kaushambi	8	15450	7295	23605	26.93	15095	9878	20509	17.80
Allahabad	20	19465	14994	23936	11.72	19909	15597	24180	10.98
Barabanki	14	18668	14600	22736	11.12	17528	14038	21155	10.16
Faizabad	12	16379	11802	20957	14.26	15755	11979	19647	12.44
Ambedkar Nagar	12	17692	14417	20966	9.44	17361	14317	20422	8.72
Sultanpur	18	16609	13493	19725	9.57	16604	13662	19551	8.98
Bahraich	14	14714	13593	15835	3.89	14658	13574	15753	3.81
Shrawasti	4	15075	9490	20660	18.9	14238	9943	18530	15.46
Balrampur	10	11975	8541	15409	14.63	12489	9526	15575	12.33
Gonda	16	16981	14828	19134	6.47	16675	14605	18751	6.37
Siddharthnagar	14	12829	9422	16235	13.55	13189	10213	16186	11.70
Basti	14	14268	9736	18800	16.21	14458	10657	18145	13.33
Sant Kabir Nagar	8	13319	11660	14978	6.35	13373	11825	14988	6.07
Mahrajganj	10	21690	16526	26854	12.15	18386	14258	22719	11.80
Gorakhpur	18	12164	9129	15199	12.73	12704	9913	15477	11.26
Kushinagar	14	19343	13702	24984	14.88	17076	12634	21683	13.63
Deoria	18	8364	5482	11246	17.58	9226	6495	11936	14.78
Azamgarh	28	11957	9961	13953	8.52	11875	9923	13730	8.09
Mau	10	9820	6039	13601	19.64	10230	6697	13690	17.30
Ballia	12	7029	4167	9892	20.78	8318	5653	11017	16.29
Jaunpur	20	16990	13571	20409	10.27	16267	13034	19448	9.97
Ghazipur	18	10858	8029	13687	13.29	11456	8784	14191	11.74
Chandauli	10	12000	7382	16618	19.63	12638	8859	16512	15.27
Varanasi	10	17665	12341	22989	15.38	16358	11421	21340	15.44
Sant Ravidas Nagar	6	6693	1943	11443	36.21	9856	5517	14178	22.53
Mirzapur	10	15625	12039	19211	11.71	15467	12123	18724	10.79
Sonbhadra	6	15283	7347	23220	26.49	12833	7627	18325	21.18
Meerut*	0					13570	9906	17234	13.78
Baghpat*	0					13962	9018	18905	18.06

(*continued*)

Table 5. Continued.

		Direct				HBNSFH			
Districts	Sample Size	Estimates	Lower	Upper	%CV	Estimates	Lower	Upper	%CV
Gautam Buddha Nagar*	0					11420	7626	15214	16.95
Hathras*	0					14229	10620	17839	12.94
Agra*	0					15150	10808	19492	14.62
Firozabad*	0					12748	9069	16426	14.72
Jalaun*	0					13610	10193	17027	12.81
Jhansi*	0					12712	8554	16871	16.69
Lalitpur*	0					11280	8662	13898	11.84
Hamirpur*	0					11257	8823	13691	11.03
Mahoba*	0					12434	10183	14685	9.24
Chitrakoot*	0					13067	10955	15178	8.24

*Out-of-sample districts.

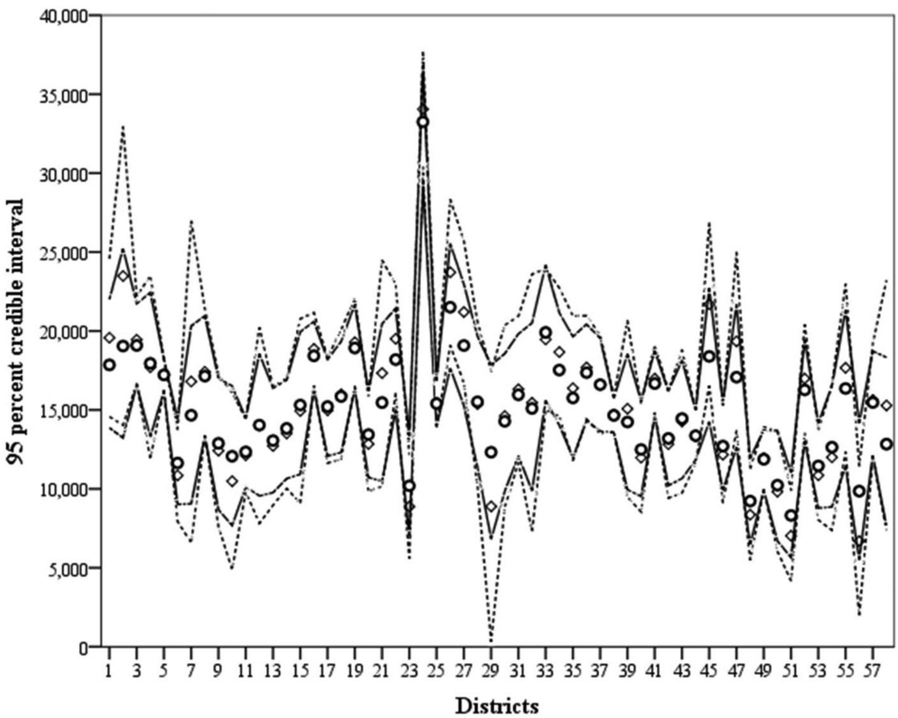

Figure 4. District-wise 95% credible interval (lower and upper) plot of paddy yield for the direct estimates (dash line, ◇) and the HBNSFH estimates (solid line, ○).

percentage gain in ARB$_v$% of HBNSFH over HBFH is 11.90, 12.77, and 12.19 for $IG(0.001, 0.001)$, $IG(0.01, 0.01)$, $IG(0.1, 0.1)$ priors, respectively. Under NSDGP1, such improvement in mean ARB$_v$% of HBNSFH over HBFH is even more. Table 3 also shows our investigation on coverage properties of both the models. The noncoverage rate is marginally higher for HBFH as compared to the other. Again, as number of areas increases all the models show the better coverage percentage.

5. Empirical results

This section presents the implementation of FH, SFH and NSFH approach in producing HB small area estimates of paddy yield for different districts of the state Uttar Pradesh in India.

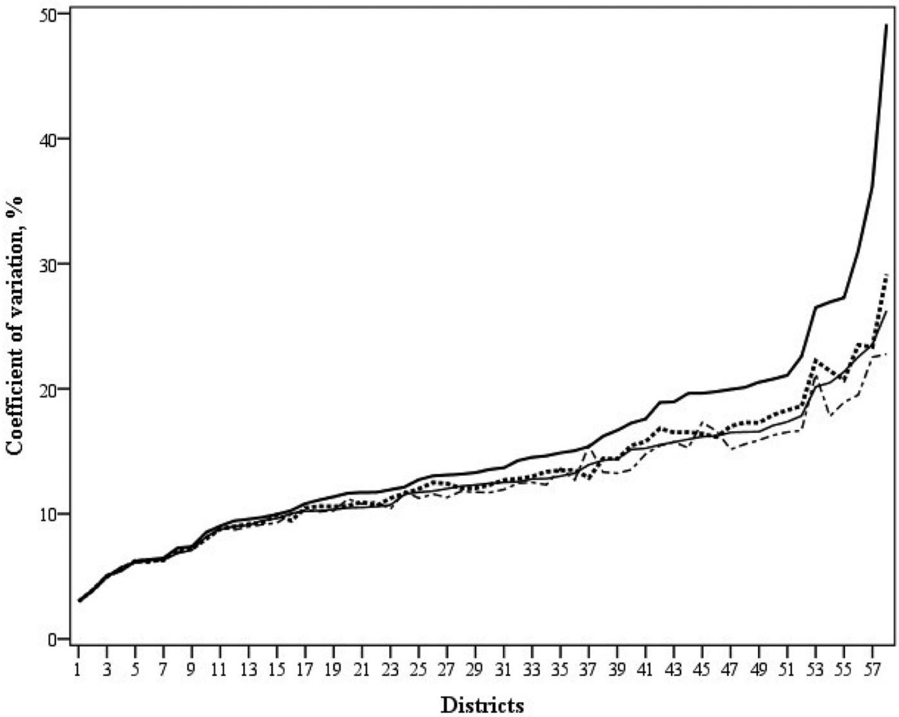

Figure 5. District-wise %CV for direct (solid thick line), HBFH (solid dash line), HBSFH (thin line) and HBNSFH (dash line) methods of SAE.

Traditional direct survey estimates for paddy yield are also computed to carry out comparison of small area model based method *vs.* direct estimation approach. Although, such assessment for model based *vs.* design based method is quite dwelling in survey literature; we just try to portray, what estimates are available in public domain for disaggregated level paddy yield in Uttar Pradesh and what we have generated. Percentage coefficient of variation (% CV) is the criteria which have been used to indicate the better performing model with stable estimates. However, before we present detailed empirical results, it is necessary to explore whether the described ICS data set exhibit spatial nonstationarity or not. For this purpose, district specific regression coefficients are computed by fitting GWR model. In the fitted model we have two covariates, AHS and FPMH; therefore we have three regression coefficients (i.e. intercepts and two slope parameters with respect to AHS and FPMH). Figure 3 shows surface plot of estimated regression coefficients for ICS data from a GWR fit (Fotheringham, Brunsdon, and Charlton 2002) to direct estimates over different sample (58 sample and 12 out-of-sample) districts. This contour map confirms the evidence of spatial nonstationarity in the ICS data; hence we may expect a better performance of small area estimates with the newly developed method of SAE, that is, HBNSFH method.

Table 4 shows the descriptive statistics of %CV for direct estimates as well as small area model based estimates for sample districts generated by HBFH, HBSFH and HBNSFH methods of SAE. Estimates with smaller %CV are more reliable than others. Comparing all the HB models, it is to be noted that the precision level of HBNSFH is better than the other model based alternative. In direct estimation approach %CV is ranging from 3.01 to 49.15, whereas, in HBNSFH the range of %CV is 2.99–22.78. This result reveals that the application of HBNSFH method for the data exhibiting spatial nonstationarity will lead to significant gains in efficiency of small area estimates over direct method and other model based alternative method. Again it is noteworthy that, for

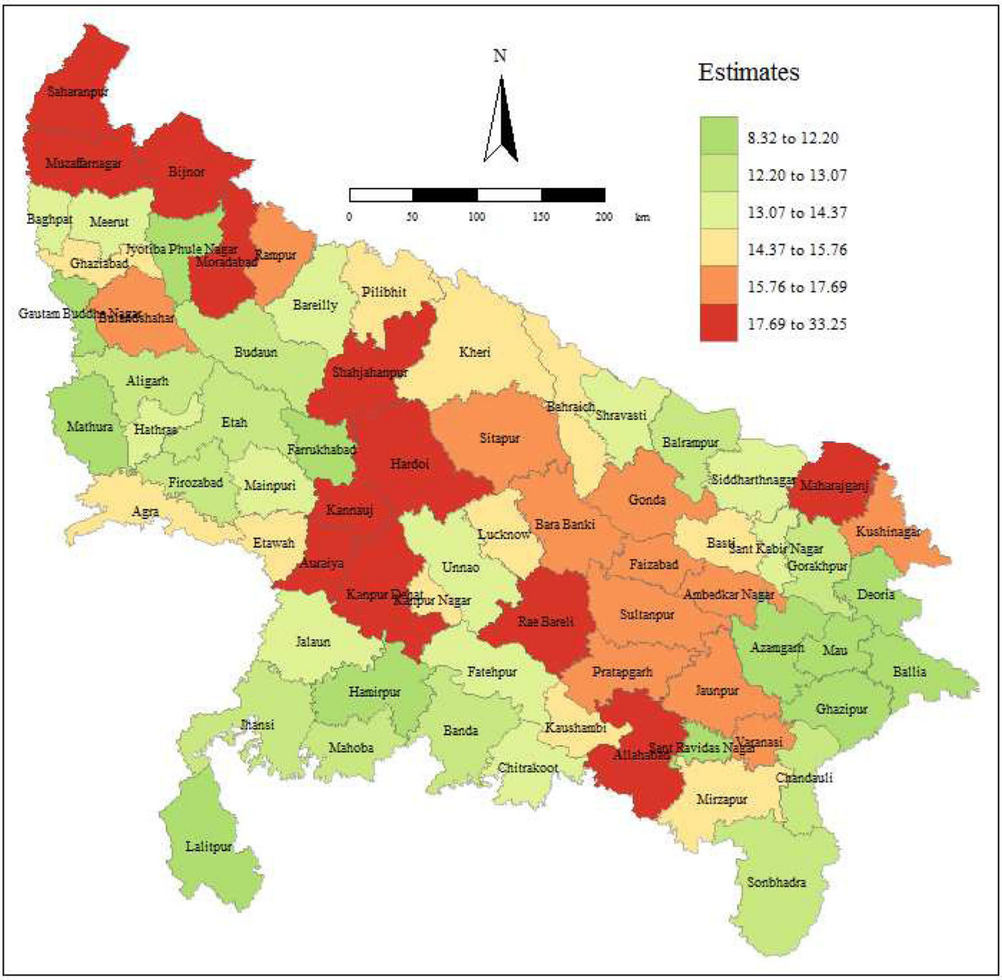

Figure 6. Spatial map showing distribution of paddy yield (in kg. per 43.12 m^2) across districts of Uttar Pradesh generated by the HBNSFH method of SAE.

no sample districts direct estimates cannot be produced. Whereas, SAE approach still can produce estimates for such districts with %CV in a reasonable limit.

Table 5 presents district wise estimates of paddy yield (gram per 43.12 m^2) along with 95% credible interval (CI) and %CV for direct and HBNSFH estimation approach. Figure 4 portrays the comparative illustration of 95% CIs of the model based HBNSFH and the direct estimates. In general, 95% CIs for the direct estimates are wider than the 95% CIs for the HBNSFH estimates. Further, 95% CIs for the HBNSFH estimates are more precise and contain both direct and model based estimates of the yield. Figure 5 is a visual picture of the district wise %CV, respectively, implementing direct, HBFH, HBSFH and HBNSFH methods. The HB models have also been compared through Bayesian model evaluation or comparison criteria DIC (Deviance information criterion). Smaller value of DIC is generally expected, which is indicative of better fit. The DIC value of HBFH, HBSFH and HBNSFH was found, respectively, as 1104, 1100, and 1097. This Bayesian model comparison result also confirms our estimation result, as HBNSFH turns out to be relatively better model. Figure 6 presents the spatial map showing the distribution of paddy yield (in kg. per 43.12 m^2) across districts of Uttar Pradesh generated by the HBNSFH method of SAE. Figure 7 is the spatial map of district wise %CV generated by the HBNSFH. Spatial map

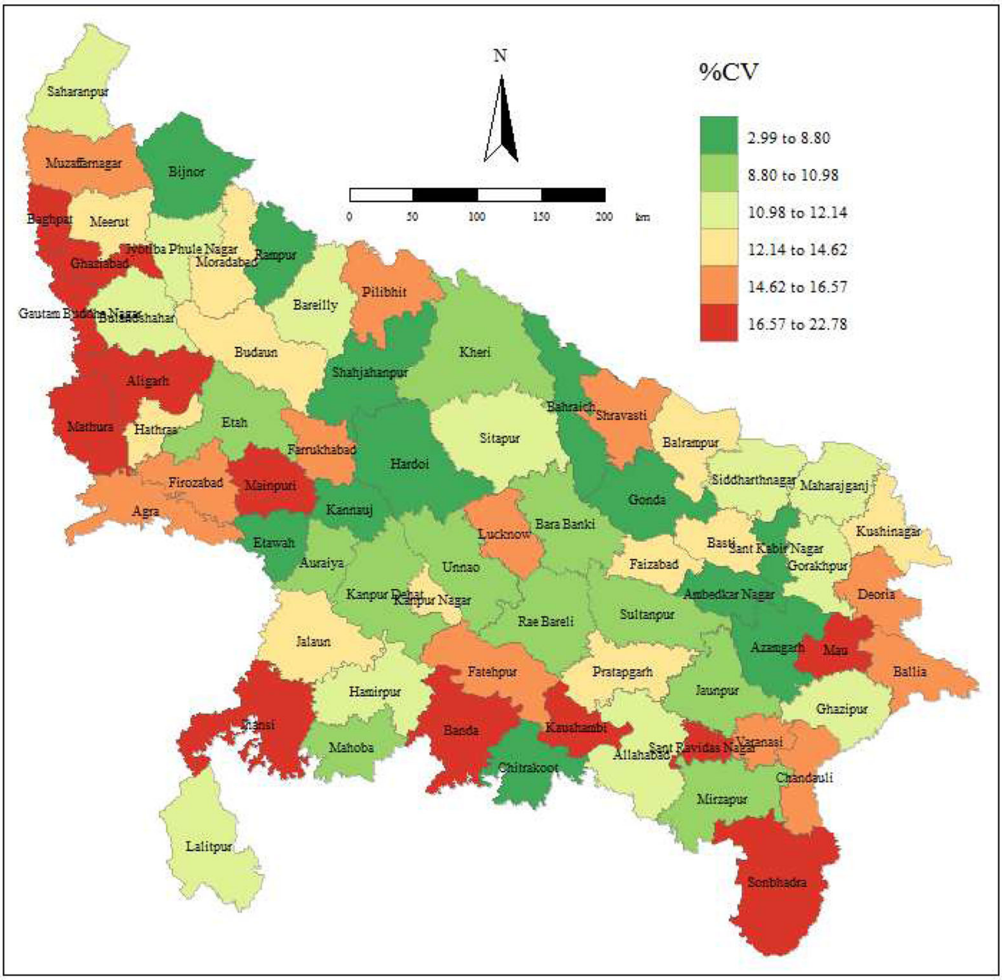

Figure 7. Spatial map of district wise %CV generated by the HBNSFH method of SAE.

produced from the model based HBNSFH estimates of paddy yield presents a quick view to the regional variations or disparity in district level yield estimates. Such spatial maps are certainly useful to the policy makers to frame targeted plans eying to the upliftment of deprived regions of the population. As a profound application, the suitability of this study can be found in insurance schemes like Pradhan Mantri Fasal Bima Yojana (PMFBY) in India to deliver insurance and input support to the needy farmers.

5. Concluding remarks

The potentiality of SAE methodologies to generate reliable small domain inference is now quite established fact from varied theoretical researches, what needed is its real life implementation and applications. To strengthen the micro level planning, disaggregate level estimates are often required and small area models serve this purpose both adequately and efficiently. The current study encompasses the development of spatial nonstationary version of HBFH (HBNSFH) SAE method and the performance of such method has been found to be promising both in simulated data and application. Implementation of HBFH model to the data exhibiting spatial nonstationarity may not provide the proficient estimates. Hence in presence of spatial nonstationarity, the application of HBNSFH model and associated method should be encouraged regarding estimation problem of population mean or total. SAE method is officially used in many countries to produce several official estimates and even more spread of such approach is need of the day with emerging necessities for micro level data. However, to fully trap the potentiality of this approach it is prerequisite to check the basic diagnostics of survey data and related auxiliary variables. Application of spatial nonstationary or other spatial models are good enough to yield promising estimates but application should be need based too.

Acknowledgments

The authors would like to acknowledge the valuable comments and suggestions of the Editor and an anonymous referee. These led to a considerable improvement in the article. This article is a tribute to **Dr. Hukum Chandra** who was a passionate researcher in the field of Survey statistics and "Small Area Estimation". He has left for heavenly abode before the final acceptance of the article.

References

Anjoy, P., H. Chandra, and P. Basak. 2019. Estimation of disaggregate-level poverty incidence in Odisha under area-level Hierarchical Bayes small area model. *Social Indicators Research* 144 (1):251–73. doi:10.1007/s11205-018-2050-9.

Anjoy, P., and H. Chandra. 2019. Hierarchical Bayes aggregated level spatial model for crop yield estimation. *Journal of the Indian Society of Agricultural Statistics* 73:143–52.

Baldermann, C., N. Salvati, and T. Schmid. 2018. Robust small area estimation under spatial non-stationarity. *International Statistical Review* 86 (1):136–59. doi:10.1111/insr.12245.

Brunsdon, C., A. S. Fotheringham, and M. E. Charlton. 2010. Geographically weighted regression: A method for exploring spatial nonstationarity. *Geographical Analysis* 28 (4):281–98. doi:10.1111/j.1538-4632.1996.tb00936.x.

Chandra, H. 2013. Exploring spatial dependence in area-level random effect model for disaggregate-level crop yield estimation. *Journal of Applied Statistics* 40 (4):823–42. doi:10.1080/02664763.2012.756858.

Chandra, H., N. Salvati, and R. Chambers. 2015. A spatially nonstationary Fay-Herriot model for small area estimation. *Journal of Survey Statistics and Methodology* 3 (2):109–35. doi:10.1093/jssam/smu026.

Chandra, H., N. Salvati, and R. Chambers. 2017. Small area prediction of counts under a non-stationary spatial model. *Spatial Statistics* 20:30–56. doi:10.1016/j.spasta.2017.01.004.

Datta, G. S., J. N. K. Rao, and D. D. Smith. 2005. On measuring the variability of small area estimators under a basic area-level model. *Biometrika* 92 (1):183–96. doi:10.1093/biomet/92.1.183.

Fay, R. E., and R. A. Herriot. 1979. Estimates of income for small places: An application of James-Stein procedures to census data. *Journal of the American Statistical Association* 74 (366a):269–77. doi:10.1080/01621459.1979.10482505.

Fotheringham, A. S., C. Brunsdon, and M. E. Charlton. 2002. *Geographically weighted regression*. New York: John Wiley & Sons.

Liu, B., P. Lahiri, and G. Kalton. 2014. Hierarchical Bayes modeling of survey-weighted small area proportions. *Survey Methodology* 40:1–13.

Opsomer, J. D., G. Claeskens, M. G. Ranalli, G. Kauermann, and F.J. Breidt. 2008. Nonparametric small area estimation using penalized spline regression. *Journal of the Royal Statistical Society B* 70 (1):265–86. doi:10.1111/j.1467-9868.2007.00635.x.

Pratesi, M., and N. Salvati. 2008. Small area estimation: The EBLUP estimator based on spatially correlated random area effects. *Statistical Methods and Application* 17:114–31.

Rao, J. N. K. 2003. *Small area estimation*. New York: John Wiley & Sons.

Rao, J. N. K., and I. Molina. 2015. *Small area estimation*: 2nd ed. New York: John Wiley & Sons.

You, Y., and J. N. K. Rao. 2002. Small area estimation using unmatched sampling and linking models. *The Canadian Journal of Statistics* 20:3–15.

You, Y., and M. Q. Zhou. 2011. Hierarchical Bayes small area estimation under a spatial modelwith application to health survey data. *Survey Methodology* 37:25–37.