Supplementary materials for this article are available at https://doi.org/10.1007/s13253-018-0329-6.

Small Area Estimation of Proportions with Constraint for National Resources Inventory Survey

Xin WANG, Emily BERG, Zhengyuan ZHU, Dongchu SUN, and Gabriel DEMUTH

Motivated by the need to produce small area estimates for the National Resources Inventory survey, we develop a spatial hierarchical model based on the generalized Dirichlet distribution to construct small area estimators of compositional proportions in several mutually exclusive and exhaustive landcover categories. At the observation level, the standard design-based estimators of the proportions are assumed to follow the generalized Dirichlet distribution. After proper transformation of the design-based estimators, beta regression is applicable. We consider a logit mixed model for the expectation of the beta distribution, which incorporates covariates through fixed effects and spatial effect through a conditionally autoregressive process. In a design-based evaluation study, the proposed model-based estimators are shown to have smaller root-mean-square error and relative root-mean-square error than design-based estimators and multinomial model-based estimators.

Supplementary materials accompanying this paper appear online.

Key Words: Generalized Dirichlet distribution; Spatial hierarchical model; Sampling variance modeling; Small area estimation; Survey statistics.

1. INTRODUCTION

The National Resources Inventory (NRI) is a longitudinal survey that monitors status and trend in numerous characteristics, primarily related to natural resources and agriculture, on nonfederal US land. It is the largest and one of the longest longitudinal surveys in the USA and provides critical information on soil erosion, land management, and landcover change,

This research has been supported in part by cooperative Agreement 68748214502 between the USDA Natural Resources Conservation Service and Iowa State University. Sun's research is also supported by NSF Grant MDS-1007874 and SES-1260806, and the 111 Project (B14019) in China.

Xin Wang, Department of Statistics, Miami University, Oxford, OH 45056, USA. Emily Berg, Zhengyuan Zhu (✉) and Gabriel Demuth, Department of Statistics, Iowa State University, Ames, USA (E-mail: *zhuz@iastate.edu*). Dongchu Sun, Department of Statistics, University of Missouri-Columbia, Columbia, USA and Department of Statistics, East China Normal University, Shanghai, China.

Journal of Agricultural, Biological, and Environmental Statistics, Volume 23, Number 4, Pages 509–528
https://doi.org/10.1007/s13253-018-0329-6

which is important for the evaluation of climate change and effects of land conservation practices. One of the parameters of interest in the NRI is the proportion of area for a set of mutually exclusive and exhaustive land categories termed broaduses. Examples of broaduses are cultivated cropland, pasture, forest, and developed land. The NRI sample design is a two-phase stratified design, and data collection is largely through interpretation of aerial photography. Section 2 reviews the essential features of the NRI sample design, data collection, and estimation procedures for our application. Nusser and Goebel (1997) and Breidt and Fuller (1999) provide further detail. Traditionally, the NRI publishes estimates of broaduse proportions at state and national levels.

Accurate information on local landcover compositions is essential for developing conservation policies and land management plans. In particular, monitoring cultivated cropland is important for agricultural planning and ensuring a sustainable food supply. Motivated by such demand, Natural Resources Conservation Service (NRCS) asked us to develop county-level estimates of broaduse proportions for NRI. Because of small sample sizes, standard NRI estimators can have relatively large estimated coefficient of variation at the county level. Additional sources of information, particularly auxiliary variables and explicit model assumptions, are needed to improve the precision of the county-level estimators.

We would like the model applied to NRI estimators of county-level broaduse proportions to have several characteristics. Estimators based on the model should respect the parameter space for the proportions and satisfy a sum-to-one constraint. The model should allow incorporation of covariates and spatial dependence structures to provide more information to improve the estimators. As we will demonstrate in Sect. 3, including spatial dependence as well as auxiliary information is important because the auxiliary variables do not fully explain the spatial structure in the data. Additionally, it is desirable to incorporate the estimated variance of the original design-based NRI estimators; since the NRI is a complex survey, variance estimator can reflect the complexity of the design. The NRI county-level estimation application is an example of a more general problem of estimating a vector of proportions that sum to one for each small area.

Fay and Herriot (1979) and Battese et al. (1988) introduce the approach of using linear mixed effects models to obtain more precise small area estimators. Rao and Molina (2015), Jiang and Lahiri (2006b) and Pfeffermann (2013) review extensions to more complex models, including models with correlated random components and nonlinear expectation functions. One approach for binary response variables is to model the small area counts (Rao and Molina 2015). For example, He and Sun (2000) use a hierarchical Bayesian model with spatial correlations that treats the realized counts as binomial random variables to estimate hunting success rates. An alternative method is to model the proportions directly (Datta et al. 1999). Jiang and Lahiri (2006a) use a beta linking model for the expectation of design-based estimators of proportions. Liu et al. (2007) compare several hierarchical Bayesian models for proportions in the context of small area estimation. In particular, models where the design-based estimators are assumed to have beta distributions are compared to models where the design-based estimators are assumed to have normal distributions. The models of Liu et al. (2007) respect the sampling design and include covariates, but they do not incorporate spatial dependence structures. These methods for binary data do not apply readily to a vector of proportions with a sum-to-one constraint.

Analyzing a vector of proportions with a constraint can start from a multinomial or Dirichlet distribution. Agresti and Hitchcock (2005) and Congdon (2005) review Bayesian estimation of multinomial parameters. Molina et al. (2007) and López-Vizcaíno et al. (2013) use multinomial-based mixed models to analyze labor force participation without considering spatial dependence. In López-Vizcaíno et al. (2015), the model was extended to a time-correlated model with area random effects. Jin et al. (2013) use spatial multinomial regression models for the purpose of understanding relationships between land ownership history and forest landscape structure, an objective that is analytical in nature and differs from small area prediction. Berg and Fuller (2014) use the covariance structure of the Dirichlet distribution as a working model for small area prediction of vectors of proportions that satisfy a restriction. Models based on the multinomial or Dirichlet distribution assume negative correlations between different categories, a structure that the real data may not satisfy. The relationships between means and variances of different categories for these two distributions may not hold for the survey estimators. The generalized Dirichlet (GD) distribution is a flexible distribution for vectors of proportions that satisfy a sum-to-one constraint. We can incorporate the sampling variances and preserve the sampling variance structure. Connor and Mosimann (1969) discuss general properties of the GD distribution. Wong (1998) discusses the use of the GD distribution as a prior for the multinomial distribution. One convenient property of the GD distribution is that independent beta distributions with different parameters are obtained after proper transformation. Ferrari and Cribari-Neto (2004) propose beta regression to model rates and proportions. In their model, they reparameterize the beta density so that the parameters of the beta density function are an expectation parameter and a dispersion parameter. Simas et al. (2010) extend beta regression to allow a nonlinear term in the regression and model the dispersion parameter as well. Gamerman and Cepeda-Cuervo (2013) consider spatial effects in both the mean and dispersion models for beta regression models.

To specify a model appropriate for the NRI application, we begin with an assumption that the observed county-level proportions are realizations from the GD distribution. The GD assumption permits a transformation of the county-level proportions to independent beta random variables with distinct mean and dispersion parameters. Spatial information can also be used to improve small area estimation (Militino et al. 2006; Petrucci and Salvati 2006). Spatial hierarchical Bayesian models are specified for the transformed variables here. The expectation of the beta distribution is modeled as a logit-linear mixed model with covariates describing large-scale structure and spatially correlated random effects for counties. The spatial structure is specified through a spatial conditionally autoregressive (CAR) model, as in Banerjee et al. (2014). The complexities of the NRI design and the availability of the auxiliary information make area-level modeling preferable to unit-level models. We discuss how the NRI motivates our model choice in more detail in Sects. 2 and 3.

Modeling the sampling variances of the survey estimators is essential in many small area estimation applications because the survey-based variance estimators often provide approximately unbiased variance estimators, but can have large variances due to small sample sizes. Appropriately specified variance models can retain the information about the sample design and estimation procedures contained in the design-based variance estimator, while reducing the variance of the variance estimator by borrowing information across areas.

Cho et al. (2002) model the sampling variances with log-normal distributions. Maples et al. (2009), Dass et al. (2012), and Maiti et al. (2014) discuss the use of Chi-squared distributions to model sampling variance. Gomez-Rubio et al. (2010) discuss both spatial models and modeling the variance in small areas in a Bayesian setting. Our model for the design-based estimators of the variances, described in more detail in Sect. 3, accomplishes two goals. The first is to treat the survey-based variance estimators as approximately unbiased. The second is to incorporate auxiliary information. Our variance model exploits both the Chi-square distribution and the log-normal distribution to achieve these two ends. Our sampling variance model extends that of Maiti et al. (2014) to incorporate covariates and uses Bayesian instead of frequentist procedures for inference.

A design-based simulation study is conducted in Sect. 4 to compare our proposed model-based estimators to design-based estimators and the estimators based on the multinomial model proposed in López-Vizcaíno et al. (2013). We treat the NRI "foundation sample," the sample obtained in the first phase of the NRI's two-phase design, as the finite population for the design-based simulation study and use a sample design similar to the NRI sample design to select subsamples for the Monte Carlo (MC) study. The MC relative root-mean-square error and mean square error are computed for both design-based estimators and model-based estimators, using the complete foundation sample as the reference for constructing true parameters. The results show that our proposed model-based estimators can reduce relative RMSE and RMSE by 15% or more on average compared with design-based estimators and perform better than the estimators based on the multinomial model.

In Sect. 2, we introduce the National Resources Inventory survey in detail. In Sect. 3, the proposed models are described. Then, we compare design-based and model-based estimators through a design-based Monte Carlo study, in which we treat the foundation data as the target finite population and sample it using a sampling design that reflects the properties of the NRI sampling design in Sect. 4. In Sect. 5, the proposed models are applied to estimate the proportions of area in several broaduses for Iowa counties in 2012. Section 6 summarizes and identifies areas for future work.

2. NATIONAL RESOURCES INVENTORY

The NRI is supported by the US Department of Agriculture Natural Resources Conservation Service and conducted in cooperation with Iowa State University. The NRI sample design has two phases of sample (Nusser and Goebel 1997). The first-phase sample, called the "foundation sample," consists of approximately 300,000 segments (primary sampling units), each of which contains 2 or 3 sampled points (secondary sampling units). From 1982 to 1997, the full NRI foundation sample was observed in 5-year intervals (1982, 1987, 1992, and 1997). In 2000, the NRI transitioned to an annual sample design to facilitate special studies and spread the workload more evenly. Because observing all 300,000 segments in the foundation sample on an annual basis is too expensive, the annual samples are subsamples of original foundation sample. In the annual samples, approximately 40,000 segments, called "core" segments, are observed every year. The rest of the foundation sample is divided into several supplemental panels, each with approximately 30,000 segments that are observed periodically. About 70,000 segments are observed every year since 2000.

Data collection in the NRI is primarily through visual interpretation of aerial photography supplemented by local data collection and integration of administrative records. Some information is collected at the point level, such as the type of crop planted in the field containing a point. Other information is collected at the segment level, such as the urban area in the segment. An estimation procedure creates imputed points to represent segment information and imputes data to create a complete time series for points not observed in a particular year. All the information is transformed to points with associated weights to represent the sample design and adjustments to administrative control totals. In the final data set, each record corresponds to a real or imputed point, each of which contains a complete time series and an associated weight. Weighted sums of characteristics for points are considered approximately unbiased for the corresponding population parameters.

Each year, a point y_j is classified into a set of mutually exclusive and exhaustive landcover categories called broaduses. The standard NRI design-based estimator of the proportion of area in broaduse k ($k = 1, \ldots, 12$) for the state in a particular year is defined as

$$\hat{p}_k = \frac{\sum_{j=1}^{n_{\text{state}}} w_j I[y_j = k]}{\sum_{j=1}^{n_{\text{state}}} w_j}, \tag{1}$$

where w_j is the weight for point j in the state, $I[y_j = k]$ is the indicator that point j is classified in category k, n_{state} is the number of points in the state, and the subscript for year is omitted because we focus on a single time point. Because the NRI uses the area of the state as control, $\sum_{j=1}^{n_{\text{state}}} w_j$ is equal to the area of the state. Similarly, the NRI design-based estimator of the proportion of county i in broaduse k is defined as

$$\hat{p}_{ki} = \frac{\sum_{j=1}^{n_i} w_{ij} I[y_{ij} = k]}{\sum_{j=1}^{n_i} w_{ij}}, \tag{2}$$

where n_i is the number of points in county i, and ij indexes the jth point in the ith county. Because the first-phase strata are contained in counties, it is reasonable to treat the sampling errors for estimators for two different counties as independent. Because of the complexity of the NRI design, jackknife method is used for variance estimation. The jackknife variance estimator for category k and county i is defined as

$$\hat{V}\left(\hat{p}_{ki}\right) = \frac{B-1}{B} \sum_{b=1}^{B} (\hat{p}_{ki}^{(b)} - \hat{p}_{ki}^{\text{est}})^2, \tag{3}$$

where $\hat{p}_{ki}^{(b)}$ is the estimate based on the bth set of replicate weights, $\hat{p}_{ki}^{\text{est}}$ is the mean of the B replicate estimates, and in the NRI, $B = 29$. To define the replicate weights, the NRI sample is sorted geographically and divided into 29 groups. The weight for an element assigned to group b is set to zero in replicate b, and ratio adjustments similar to those used to construct the original weights are repeated to construct the replicate weights.

The design-based county estimates defined in (2) are not published. Since the survey is designed for state-level estimates, the summation of all weights in one county may not match the area of the county. In addition, due to the relatively small sample sizes for counties, particularly in the annual samples, the county-level estimators are often judged unreliable in terms of estimated coefficient of variation (CV). For example, the estimated CV for cultivated cropland at the state level is 0.57% in 2012. However, estimates of CV for counties range from 10 to 30%. For pastureland, the average estimated CV across counties is 40%. Thus, we consider model-based estimators to improve the precision of the county-level estimators.

The cropland data layer (CDL) is a classification of square pixels into several mutually exclusive and exhaustive landcover categories, which classified satellite readings into landcover categories using NASS survey data and administrative data as ground truth. See Han et al. (2012) for further details on the CDL. The CDL has been released annually from 2006 through the present. We decided to obtain auxiliary information from the cropland data layer because it is nationally consistent and timely. NRI and CDL categories are similar enough that we are able to build a map between CDL and NRI categories based on the definitions. And CDL contains categories that are relatively straightforward to map to NRI categories. However, the goals of the NRI and CDL projects are different, and a one-to-one mapping between the NRI and CDL does not exist. After mapping NRI categories to CDL categories, the specific covariate used in the models is the proportion of pixels in a county classified in a particular NRI broaduse. Because the NRI and CDL use different definitions and data collection procedures, the correlations between the covariate and the NRI estimators defined in (2) vary across categories. For example, the linear correlation is 0.9651 for cultivated cropland, but for pastureland, the linear correlation is 0.5029. For pastureland, the spatial dependence has the potential to help improve the model-based estimators, even after considering auxiliary information.

3. SPATIAL HIERARCHICAL MODELS FOR PROPORTIONS

As mentioned in Sect. 2, the NRI county-level estimates have large values of estimated coefficient of variation. Thus in this section, we propose to use the spatial Bayesian hierarchical models for small area estimation of NRI county-level proportions. In Sect. 3.1, we introduce the generalized Dirichlet (GD) distribution and present the transformation to independent beta random variables. In Sect. 3.2, we specify the hierarchical models used for small area estimation, which begin with an assumption that the NRI estimators of proportions are realizations from GD distributions and variance estimators are realizations from Chi-squared distributions. Section 3.3 presents specific details required for the Bayesian inference.

3.1. Generalized Dirichlet Distribution

Let $0 < p_k < 1, k = 1, \ldots, K$ be the proportion of the kth category with $\sum_{k=1}^{K} p_k = 1$. The probability density function of the generalized Dirichlet distribution (Connor and Mosimann 1969) is

$$f(\boldsymbol{p}, |\boldsymbol{\eta}_1, \boldsymbol{\eta}_2) = \left[\prod_{k=1}^{K-1} B(\eta_{1k}, \eta_{2k})\right]^{-1} p_K^{\eta_{2,K-1}-1} p_1^{\eta_{11}-1} \prod_{k=2}^{K-1} \times \left[p_k^{\eta_{1k}-1} \left(\sum_{j=k}^{K} p_j\right)^{\eta_{2,k-1}-(\eta_{1k}+\eta_{2k})}\right], \tag{4}$$

where $\boldsymbol{p} = (p_1, \ldots, p_{K-1}, p_K)^{\mathrm{T}}, \boldsymbol{\eta}_1 = (\eta_{11}, \ldots, \eta_{1,K-1})^{\mathrm{T}}, \boldsymbol{\eta}_2 = (\eta_{21}, \ldots, \eta_{2,K-1})^{\mathrm{T}}$, and $B(\eta_{1i}, \eta_{2i})$ is the beta function. We denote $f(\boldsymbol{p}, |\boldsymbol{\eta}_1, \boldsymbol{\eta}_2)$ as $GD(\boldsymbol{\eta}_1, \boldsymbol{\eta}_2)$. From Connor and Mosimann (1969), the GD distribution has a useful connection to a collection of independent beta random variables. To define this relationship, let

$$z_k = \frac{p_k}{\sum_{j=k}^{K} p_j}, \text{ for } k = 1, \ldots, K-1. \tag{5}$$

It can be shown (Connor and Mosimann 1969) that $\{z_k, k = 1, \ldots, K-1\}$ is a collection of independent random variables with beta distributions, where the parameters governing the distribution of z_k are η_{1k} and η_{2k}. Let $\alpha_k = E(z_k)$, from the transformation in (5) and the independence of z_k's, the expectations of p_k's are given by

$$E\left(p_k\right) = \begin{cases} \alpha_1 & k = 1, \\ \alpha_k \prod_{j=1}^{k-1}\left(1-\alpha_j\right) & k = 2, \ldots, K-1, \\ \prod_{j=1}^{K-1}\left(1-\alpha_j\right) & k = K. \end{cases} \tag{6}$$

The properties of the GD distribution are useful for the NRI application. Compared with the Dirichlet distribution, the GD distribution has more parameters, permitting greater flexibility. For Dirichlet distribution, the variance and the mean have a specific relationship and this relationship can be more flexible and is controlled by an extra parameter for the GD distribution. The functional restrictions of the Dirichlet distribution also lead us to prefer the flexibility that the GD distribution allows.

3.2. MODEL SPECIFICATION

Let $i = 1, \ldots I$ be the index for area and $k = 1, \ldots K$ be the index for category. In the NRI application, the categories correspond to different landcover classes (broaduses), and the small areas are counties. Let $\hat{p}_{ki}$ be the design-based estimator of the proportion of the kth category in county i. Assume that a design-based estimator of the sampling variance, denoted $\hat{V}(\hat{p}_{ki})$, is available. In the NRI, $\hat{V}(\hat{p}_{ki})$ is obtained by jackknife variance.

Assume $(\hat{p}_{1i}, \ldots, \hat{p}_{Ki})^{\mathrm{T}}$ follows GD $\left(\boldsymbol{\eta}_{1i}, \boldsymbol{\eta}_{2i}\right)$, where $\boldsymbol{\eta}_{1i} = (\eta_{1,1i}, \ldots, \eta_{1,(K-1)i})^{\mathrm{T}}$ and $\boldsymbol{\eta}_{2i} = (\eta_{2,11}, \ldots, \eta_{2,(K-1)i})^{\mathrm{T}}$. Based on the properties of the GD distribution, we use the following transformations,

$$z_{ki} = \frac{\hat{p}_{ki}}{\hat{p}_{ki} + \cdots + \hat{p}_{Ki}}, \tag{7}$$

where $z_{ki} \overset{\text{ind}}{\sim} \text{Beta}(\eta_{1,ki}, \eta_{2,ki})$ for $k = 1, \dots, K-1$. Then the problem becomes a beta regression problem. We follow Ferrari and Cribari-Neto (2004) and Simas et al. (2010) and model the expectation parameters $\alpha_{ki} = E(z_{ki}) = \eta_{1,ki}/(\eta_{1,ki} + \eta_{2,ki})$ and dispersion parameters $\phi_{ki} = \eta_{1,ki} + \eta_{2,ki}$.

The model structure we consider for the transformed expectation α_{ki} has following form,

$$\text{logit}(\alpha_{ki}) = \beta_{k0} + x_{ki}\beta_{k1} + U_{ki}. \tag{8}$$

Since $0 < \alpha_{ki} < 1$, we use a logit link function, that is $\text{logit}(x) = \log(x/(1-x))$. x_{ki} is a covariate, which can have more than one dimension in general. β_{k0} and β_{k1} are regression coefficients. We also consider the spatial information U_{ki}, since adjacent counties may have similar characteristics which are not fully explained by the auxiliary information.

In NRI, the covariate is obtained from the CDL, as discussed in Sect. 2. To define the covariate x_{ki}, we begin by defining a set of CDL proportions to have the same categories as the NRI proportions. The same transformation defined in (7) is applied to the CDL proportions to obtain proportions $\hat{p}_{c,ki}$'s for $k = 1, \dots, K-1$. The covariate x_{ki} is obtained by applying a logit transformation to the $\hat{p}_{c,ki}$. The CAR model (Banerjee et al. 2014) is used to model the spatial effect in (8), which is defined by conditional distributions $(U_{ki}|U_{kj}, j \neq i) \sim N(\rho_k \sum_{j\neq i} C_{ij}U_{kj}/C_{i+}, \delta_k/C_{i+})$, where ρ_k is the spatial dependence parameter, δ_k is the variance component of category k, $\boldsymbol{C}$ is the adjacency matrix with diagonal element $C_{ii} = 0$ and ijth off-diagonal element $C_{ij} = I[\text{ counties i and j share a common boundary}]$, and $C_{i+} = \sum_{j\neq i} C_{ij}$. The joint distribution of $\boldsymbol{U}_k = (U_{k1}, \dots, U_{kI})'$ is $N(0, \delta_k(\boldsymbol{D} - \rho_k\boldsymbol{C})^{-1})$, where $\boldsymbol{D} = \text{diag}(C_{1+}, \dots, C_{n+})$. In order to guarantee $\boldsymbol{D} - \rho_k\boldsymbol{C}$ is positive definite, ρ_k should satisfy $\lambda_{\min}^{-1} < \rho_k < \lambda_{\max}^{-1}$, where $\lambda_{\min}$ and $\lambda_{\max}$ are the minimal negative and maximal positive eigenvalues of $\boldsymbol{D}^{-1/2}\boldsymbol{C}\boldsymbol{D}^{-1/2}$, respectively. The ability of this formulation to accommodate different spatial dependence parameters and different regression parameters is particularly important for the NRI application. As explained in Sect. 2, the correlations between NRI proportions and CDL proportions vary by category, as well as the strength of the spatial dependence. To allow this flexibility, we consider the general model in which different categories have different spatial dependence parameters ρ_k.

Next, we will build a model on the dispersion parameters ϕ_{ki}. The most general assumption for dispersion parameter allows a different ϕ_{ki} for each area and category. Under this assumption, one approach of estimating ϕ_{ki} is to estimate ϕ_{ki} with the design-based estimators and variance estimators of the proportions and then to treat the estimated ϕ_{ki} as fixed quantities in the models. Treating a design-based estimate of the variance as the true variance is an approach that has been used in the small area estimation literature (Jiang and Lahiri 2006a). For this approach, according to the definition of beta distribution, we have

$$\hat{\phi}_{ki} = \frac{z_{ki}(1-z_{ki})}{\hat{V}(z_{ki})} - 1, \tag{9}$$

where z_{ki} is obtained from the estimators of proportions as defined in (7) and $\hat{V}(z_{ki})$ can be approximated by Taylor series approximation given in (10),

$$\begin{aligned}\hat{V}(z_{ki}) = &\left(\frac{\sum_{j=k+1}^{K} \hat{p}_{ji}}{\left(\sum_{j=k}^{K} \hat{p}_{ji}\right)^2}\right)^2 \hat{V}(\hat{p}_{ki}) + \sum_{j=k+1}^{K} \left(\frac{\hat{p}_{ki}}{\left(\sum_{j=k}^{K} \hat{p}_{ji}\right)^2}\right)^2 \hat{V}(\hat{p}_{ji}) \\ &- 2\sum_{j=k+1}^{K} \frac{\hat{p}_{ki} \sum_{j=k+1}^{K} \hat{p}_{ji}}{\left(\sum_{j=k}^{K} \hat{p}_{ji}\right)^4} \widehat{\mathrm{Cov}}(\hat{p}_{ki}, \hat{p}_{ji}) \\ &+ \sum_{j=k+1}^{K} \sum_{g \neq j} \frac{\hat{p}_{ki}^2}{\left(\sum_{j=k}^{K} \hat{p}_{ji}\right)^4} \widehat{\mathrm{Cov}}(\hat{p}_{ji}, \hat{p}_{gi}).\end{aligned} \tag{10}$$

The covariance is calculated as $\widehat{\mathrm{Cov}}(\hat{p}_{ki}, \hat{p}_{ji}) = \rho_{kj,i}\sqrt{\hat{V}(\hat{p}_{ki}) \cdot \hat{V}(\hat{p}_{ji})}$, where $\rho_{kj,i}$ is replaced by the sample correlation between $\hat{\boldsymbol{p}}_k$ and $\hat{\boldsymbol{p}}_j$ with $\hat{\boldsymbol{p}}_k = (\hat{p}_{k1}, \ldots, \hat{p}_{kI})^{\mathrm{T}}$. The model based on fixed quantities of ϕ_{ki} in (9) is denoted as "M1." That is, in "M1," the expectation model is $\mathrm{logit}(\alpha_{ki}) = \beta_{k0} + x_{ki}\beta_{k1} + U_{ki}$ and ϕ_{ki}'s are estimated from (9).

The other approach is to model ϕ_{ki} through modeling the sampling variance. As in Maiti et al. (2014), we can build a joint model with both means and variances. The variance model is built using both Chi-squared and log-normal distributions in (11) and (12), allowing us to use the unbiased NRI variance estimators in the model, respect the mean–variance relationship in the beta distribution, and incorporate covariates. For the variance model, assume

$$\hat{V}(z_{ki}) \sim \frac{V(z_{ki})}{q_i}\chi^2(q_i), \tag{11}$$

where $V(z_{ki}) = \alpha_{ki}(1 - \alpha_{ki})/(\phi_{ki} + 1)$. The dispersion parameter ϕ_{ki} is modeled as

$$\log(\phi_{ki}) = \gamma_{k0} + \gamma_{k1}u_{ki} + e_{ki}, \tag{12}$$

where γ_{k0} and γ_{k1} are coefficients to estimate, u_{ki} is a covariate, and $e_{ki} \overset{\text{iid}}{\sim} N(0, \delta_{\phi k})$. Here we use the same covariate in both mean model and variance model. In the NRI, we use the number of the primary sampling units (PSUs) in a county as the value of q_i, where we use the number of PSUs instead of the number of sampled points because we expect that a positive intracluster correlation would cause the variance estimate based on the number of sampled points to be too small. The model combining variance model in (11) is denoted as "M2." In our application, we prefer to use the model "M2," which combines the mean model in (8) and the variance model (11) together. That is, in "M2," the expectation model is $\mathrm{logit}(\alpha_{ki}) = \beta_{k0} + x_{ki}\beta_{k1} + U_{ki}$, and ϕ_{ki}'s are modeled through the variance model $\hat{V}(z_{ki}) \sim \chi^2(q_i)V(z_{ki})/q_i$.

3.3. BAYESIAN ESTIMATION

We specify the following priors: $\pi(\boldsymbol{\beta}) \propto 1$, $\pi(\boldsymbol{\gamma}) \propto 1$, and $\rho_k \sim \mathrm{Uniform}(\lambda_{\min}^{-1}, \lambda_{\max}^{-1})$. As in Gelman (2006) and Polson and Scott (2012), the inverted-beta prior is used as the prior distribution for variance parameters, which is equivalent to the use of half-Cauchy prior for the standard deviation parameter. We use Gibbs sampling to simulate from the posterior distributions. Because the full conditional distribution for $v_{ki} = \text{logit}(\alpha_{ki})$ does not have a closed form, Metropolis–Hastings algorithm is used to sample from the posterior distributions of these parameters using the techniques in Diggle et al. (1998). For ρ_k, since the posterior distribution of ρ_k is proportional to a log-concave function, adaptive rejection sampling (Gilks and Wild 1992) is used here. To diagnose convergence, we use the scale reduction factor (Gelman and Rubin 1992). The burn-in value is 5000 iterations, and the next 5000 iterations are used to approximate the posterior distributions in the simulation study. In real data analysis, the 50,000 iterations after burn-in size 10,000 iterations are used.

4. DESIGN-BASED SIMULATION STUDY

In this section, we conduct a design-based Monte Carlo (MC) study to evaluate the properties of the estimators under conditions reflective of the NRI. The first-phase NRI sample (the foundation sample) serves as the finite population for the simulation study. The parameters of interest are the county-level proportions in the categories cultivated cropland, pastureland, and the remainder (a combined category containing all other 9 categories) for the year 1997, the last year in which the full foundation sample was observed. The full foundation sample is considered as the finite population. Samples are drawn from the population. For each sample, we calculate the design-based estimates and model-based estimates. We compare estimators based on the models proposed in Sect. 3 (“M1” and “M2”) to design-based estimators in (2) and estimators based on the multinomial model of López-Vizcaíno et al. (2013).

The sample design for the simulation study is a stratified single-stage cluster sample with counties as strata. The primary sampling unit is an NRI segment, and all points in a selected segment are included in the sample. Iowa has 99 counties, and the number of segments per county in the finite population for Iowa ranges from 46 to 259. We use pivotal sampling (Deville and Tille 1998) implemented in the R package Tille and Matei (2016) (*sampling*) to select a without-replacement sample of segments with specified selection probabilities. The initial inclusion probability for each segment is proportional to the weight of the segment, which is the summation of all the point weights in the segment. For each county, the sampling fraction is 0.2, which is close to that in the NRI annual sample. The design-based estimators (Horvitz–Thompson estimator) (Särndal et al. 2003) and the corresponding variance estimators (Stehman and Overton 1989) are calculated. The estimation procedure is implemented by the R package Lumley (2011) (*survey*) .

In the foundation sample (the finite population for the simulation), the area of each of the three categories (cultivated cropland, pastureland, and the remainder) that we consider is greater than zero in every county. A design-based estimate for a random sample, however, may equal zero. Because the support of the beta distribution does not contain zero, we use a simple procedure to replace zero estimates with positive values. The weight of each point is rounded to 100 acres, and we know the area of each county. A zero estimate for a category therefore means that estimate of the area of that category in the county, without rounding the weights, would fall between 0 and $1/T_i$ acres, where T_i is the known area of county i in units of 100 acres, which ranges from 2523 to 6331. We replace a zero design-based estimate with the small proportion $0.5/T_i$, which is the midpoint between the lower bound of 0 and the upper bound of $1/T_i$. The percentage of zero estimates is around 4% .

Figures 1 and 2 show the MC root-mean-square error (RMSE) and the MC relative root-mean-square error (RRMSE), respectively, for model-based estimators and the design-based estimator. The relative RMSE is calculated as $\mathrm{RMSE}_{ki}/p_{ki}$, where $\mathrm{RMSE}_{ki} = \sqrt{R^{-1}\sum_{r=1}^{R}(\hat{p}_{ki}^{(r)} - p_{ki})^2}$, $\hat{p}_{ki}^{(r)}$ is the estimator (model-based or design-based) of the proportion in category k and county i in MC sample r, and p_{ki} is the finite population proportion for the simulation. The number of MC samples is $R = 200$. The estimator based on the multinomial model of López-Vizcaíno et al. (2013) is denoted "Mult."

In terms of mean and median RMSE and RRMSE, the proposed model-based estimators perform better than the design-based estimators and the estimators based on the multinomial model. As mentioned above, all the components in multinomial distribution are assumed to be negatively correlated. For the GD distribution, only the first component is negatively correlated with other components. For this finite population, the estimators for cultivated cropland are negatively correlated with other two categories, while the estimators for pasture and the remainder are positively correlated with each other. The ability of the GD distribution to describe this correlation structure may explain why the estimators based on the GD distribution are more efficient than estimators based on the multinomial distribution. Estimators based on M1 have the smallest mean and median RMSE and RRMSE. The estimators for the pasture domain have larger RRMSE than the estimators for the remainder category because the finite population proportions for pasture are typically small.

We also evaluate the posterior variance as an estimator of the MC variance. The MC variance is defined as the variance of R estimates, $V(\hat{p}_{ki}) = \sum_{r=1}^{R}\left(\hat{p}_{ki}^{r} - \hat{p}_{ki}^{\mathrm{avg}}\right)^2/(R-1)$, where $\hat{p}_{ki}^{\mathrm{avg}}$ is the mean of the R estimates. In order to evaluate the bias of the posterior variance as an estimator of the MC variance, we calculate the MC relative bias as $E(\hat{V}(\hat{p}_{ki}))/V(\hat{p}_{ki}) - 1$, where $E(\hat{V}(\hat{p}_{ki}))$ is the MC mean of the posterior variance. For "M1," posterior variances of the GD-based estimators have positive bias for cultivated cropland and the remainder category, but a negative bias for pastureland. The posterior variance based on the "M2" model has a positive bias for the MC variance for all categories. Even though the posterior variance is not expected to be an unbiased estimator of the design-variance of the estimators, the average of posterior variances is smaller than the variance of the design-based estimators, demonstrating that the posterior variance captures the efficiency gain due to the use of the spatial hierarchical Bayesian model.

5. APPLICATION TO 2012 NRI

In this section, we apply models M1 and M2 to 2012 NRI data to obtain estimates of county-level proportions. The parameters of interest in the application are the proportion of area in each of Iowa's 99 counties in the categories of cultivated cropland, pastureland, and the remainder, which is a set containing all other 9 categories, in 2012. All of these three categories have nonzero estimates. As discussed in Sect. 2, the estimated coefficients of variation for design-based NRI estimates at the county level are often large. Model-based estimators are considered here to improve the reliability of the design-based estimators. The

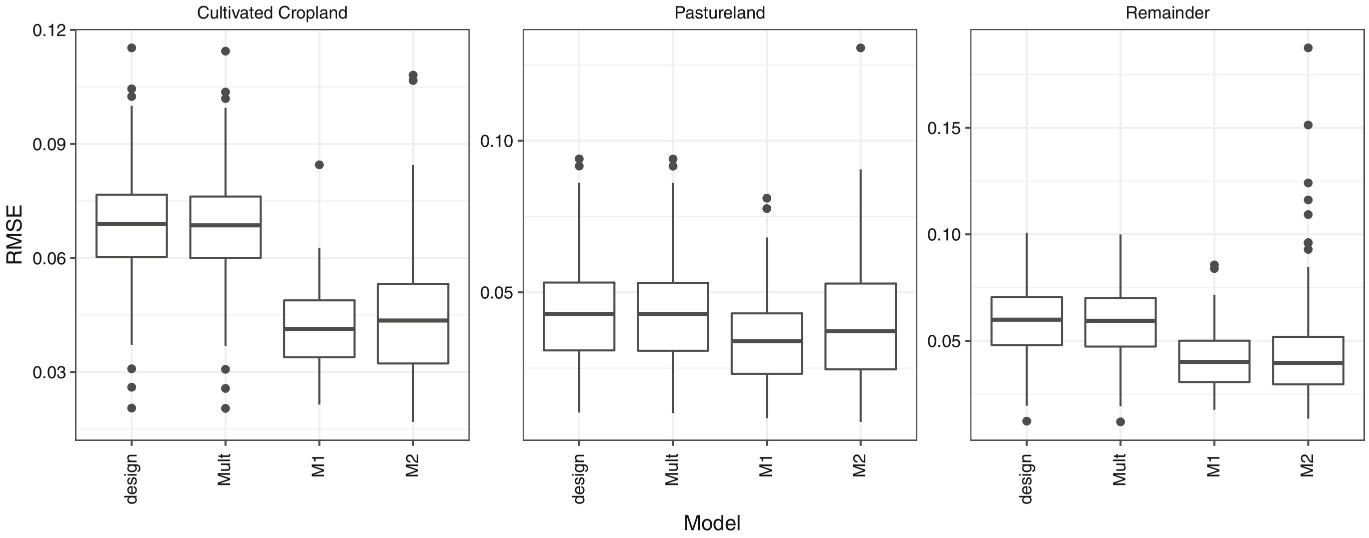

Figure 1. Design RMSE of small area estimators based on different models.

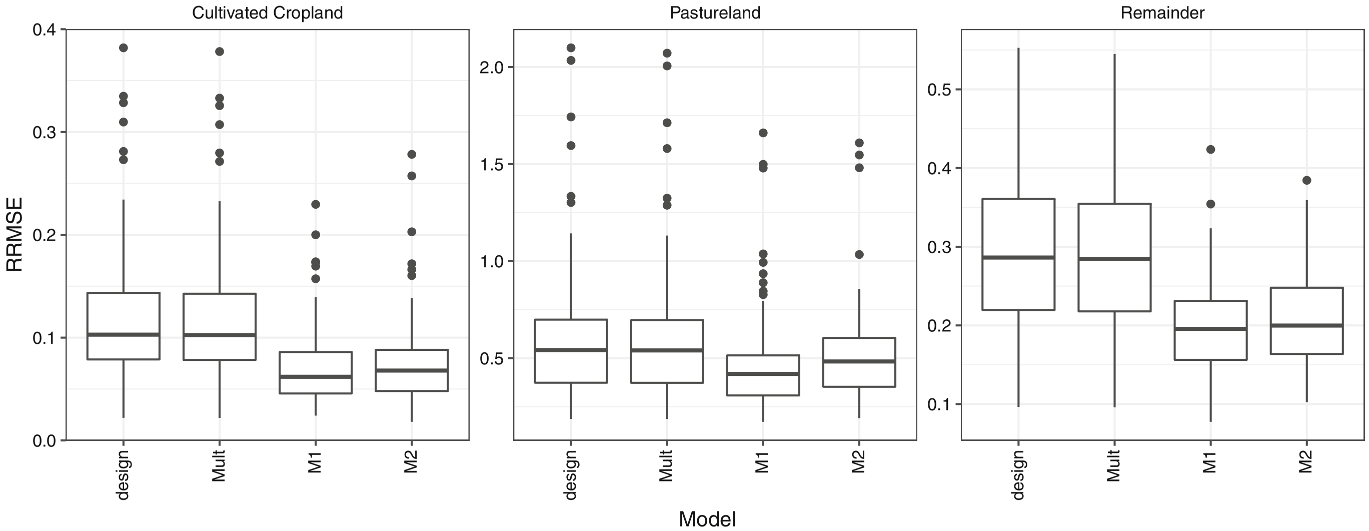

Figure 2. Design relative RMSE of small area estimators based on different models.

Table 1. Model assessment p values and DIC values.

	p value			
	Cultivated cropland	Pastureland	Others	DIC
M1	0.0470	0.8802	0.9501	− 357.8808
M2	0.2204	0.2387	0.9087	− 340.8546

estimated variance of z_{ik} is calculated using jackknife replicate weights prepared for the 2012 NRI.

The model assessment is based on the posterior predictive distribution (Gelman et al. 2014). Because state-level estimates are considered stable in the NRI, we choose state-level estimates as the characteristics of the data to which we compare data generated from the model. The following method is used to assess the model.

1. Calculate NRI design-based proportions $\hat{p}_k$ for $k = 1, \ldots, 3$ at state level.
2. Simulate $z_{ki}^{(m)}$ from the posterior predictive distribution for $i = 1, \ldots, I, m = 1, \ldots, M$ from M1 and M2, where m denotes the Gibbs iteration.
3. Transform $z_{ki}^{(m)}$ back to $\hat{p}_{ki}^{(m)}$.
4. Calculate state-level proportions $\hat{p}_k^{(m)} = \sum_{i=1}^{n} \hat{p}_{ki}^{(m)} \sum_{j=1}^{n_i} w_{ij} / \sum_{i=1}^{n} \sum_{j=1}^{n_i} w_{ij}$.
5. Calculate p values as $\frac{1}{M} \sum_{m=1}^{M} I(\hat{p}_k^{(m)} > \hat{p}_k)$,

where n is the number of counties, n_i is the number of points in county i, and w_{ij} is the weight of point ij. If we have really small p values or really large p values, that means the county-level estimates based on models cannot capture the characteristics, state-level estimates, of the data set. We also use DIC (Spiegelhalter et al. 2002) to compare different models. The model with smaller DIC is preferred.

Table 1 shows the results of the model assessment for considered models. In terms of p values, M2 is better, while M1 is better according to DIC. Based on these results, M2 reproduces the state-level proportions better than M1. The predictive posterior distribution and DIC give different preferred models. In our application, we want the selected model to respect the original data structure and characteristics. Thus, we prefer M2 over M1 for this application, since we consider the state-level estimates as important characteristics.

Table 2 shows posterior means and standard deviations for different parameters. For the spatial effect ρ_k, the 95% credible intervals based on 2.5 and 97.5% quantiles are (− 1.575, 0.967) and (0.799, 0.999) for cultivated cropland and pastureland, respectively. For cultivated cropland ($k = 1$), the spatial effect does not differ significantly from zero. The reason is that the covariate CDL itself has a strong spatial effect (Moran's I p values less than 2^{-16}), and the NRI and CDL cultivated croplands also have a strong correlation (95% credible interval of β_1 is (0.833, 1.116) in Table 2). Thus, the CDL explains the spatial structure in the NRI cultivated cropland estimates. In contrast, for pastureland ($k = 2$), the relationship between the NRI and CDL is not very strong (95% credible interval of β_1 is (− 0.106, 0.687), and the

Table 2. Estimates of parameters.

	$k = 1$ est (sd)	$k = 2$ est (sd)
β_0	0.234 (0.062)	– 1.033 (0.48)
β_1	0.973 (0.072)	0.288 (0.201)
γ_0	2.822 (0.072)	2.963 (0.101)
γ_1	– 1.041 (0.086)	– 0.708 (0.206)
δ	0.071 (0.06)	0.756 (0.238)
ρ	– 0.075 (0.755)	0.952 (0.056)
δ_ϕ	0.271 (0.047)	0.575 (0.096)

spatial effect becomes highly significant, which is used to reduce the uncertainty in county estimates.

Because of the assumption of generalized Dirichlet distribution, the sum-to-one constraint is satisfied automatically. Since NRI is designed for state estimates, we also want that the aggregated county-level model-based estimates are equal to the design-based survey state estimates. Thus, benchmarked estimates are considered, which satisfy both the sum-to-one constraint and the aggregated state-level estimates based on the county-level estimates equal to the NRI state-level estimates. Specifically, the benchmarking constraints are,

$$\sum_{k=1}^{3} \hat{p}_{ki} = 1, \quad \text{for } i = 1, \ldots, n, \tag{13}$$

$$\sum_{i=1}^{n} \hat{p}_{ki} A_i = A_0 \hat{p}_k, \quad \text{for} k = 1, \ldots, 3, \tag{14}$$

where A_i is the known area of county i, and $A_0 = \sum_{i=1}^{n} A_i$, which is the administrative state area. (13) and (14) are for the sum-to-one constraint and state-level estimates constraint, respectively. We use raking method (Kalton 1983) to benchmark the estimates. Figure 3 shows the estimates of different categories.

According to You et al. (2004), the posterior mean square error (PMSE) of the benchmarked estimator can be calculated as,

$$\widehat{\text{PMSE}}\left(\hat{p}_{ki}^{(\text{bench})}\right) = \hat{V}\left(\hat{p}_{ki}^{\text{post}}\right) + \left(\hat{p}_{ki}^{\text{post}} - \hat{p}_{ki}^{(\text{bench})}\right)^2, \tag{15}$$

where $\hat{p}_{ki}^{\text{post}}$ is the model-based estimator, $\hat{V}(\hat{p}_{ki}^{\text{post}})$ is the posterior variance of $\hat{p}_{ki}$, and $\hat{p}_{ki}^{(\text{bench})}$ is the benchmarked estimator. The PMSE includes the corrections due to the benchmarking process. The benchmarked estimates do not differ much from the original model-based estimates. For cultivated cropland and pastureland, the benchmarked estimates are larger than the original model-based estimates. But for the remainder, the benchmarked estimates are smaller.

Figure 4 shows the estimates and 95% confidence intervals of the design-based estimates and 95% posterior intervals of the benchmarked model-based estimates for cultivated cropland. The confidence intervals are defined as $\hat{p}_{ki} \pm 1.96\sqrt{\hat{V}(\hat{p}_{ki})}$, where $\hat{p}_{ki}$ is the NRI

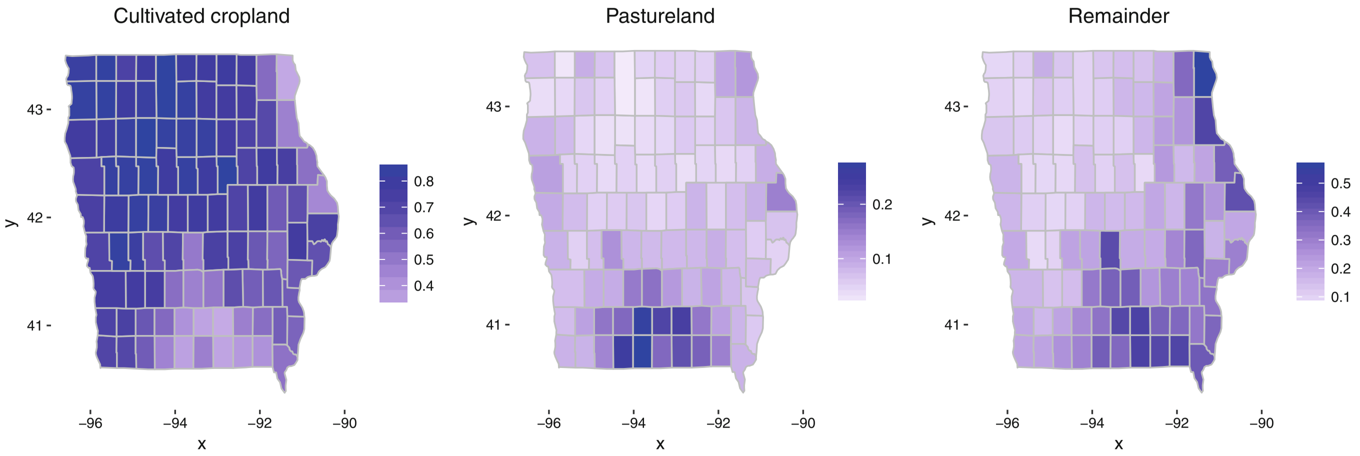

Figure 3. Estimated maps.

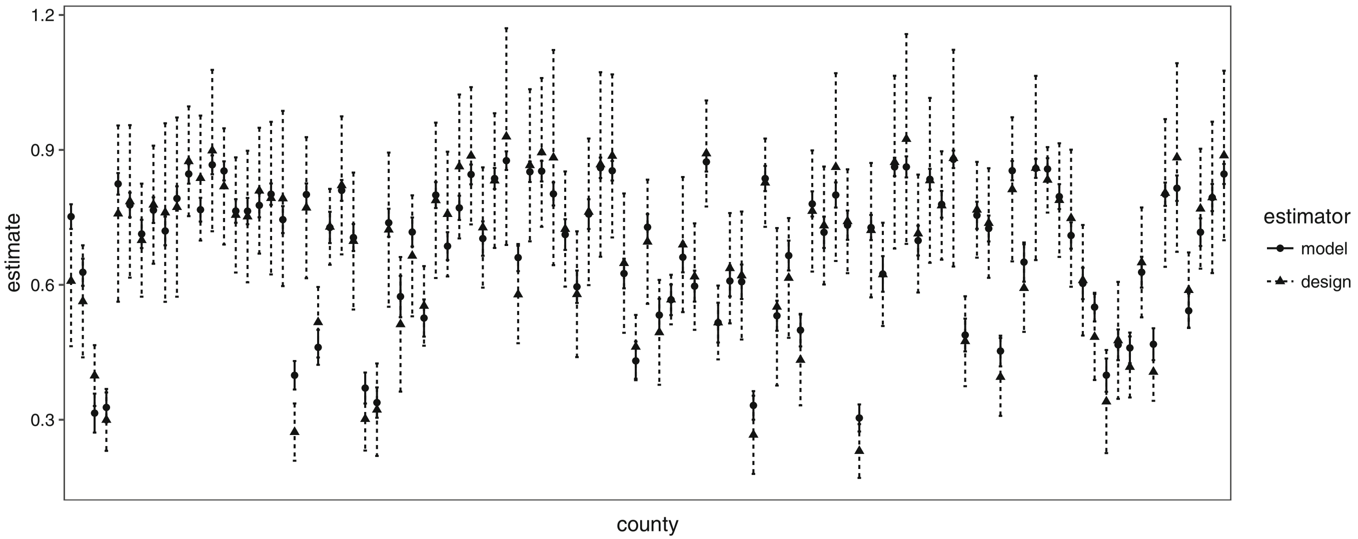

Figure 4. Intervals of cultivated cropland for NRI design-based estimates and benchmarked estimates at county level.

design-based estimate and $\hat{V}(\hat{p}_{ki})$ is the jackknife variance. And the posterior intervals are defined as $\hat{p}_{ki}^{(\text{bench})} \pm 1.96\sqrt{\widehat{\text{PMSE}}(\hat{p}_{ki}^{(\text{bench})})}$. Figure 4 demonstrates the efficiency gain due to the spatial hierarchical Bayesian model. This has important implication for policy because county estimates with better accuracy can provide better guides for land management planning at county level.

6. DISCUSSION

This paper uses the generalized Dirichlet distribution to model design-based estimates of proportions and obtain small area estimators of compositional proportions. Based on the relationship between the GD distribution and the beta distribution, a spatial Bayesian hierarchical model with beta regression is formulated and applied to NRI data. Another innovation is the introduction of a model for the dispersion parameter of the beta distribution that utilizes both Chi-squared and log-normal distributions. In a design-based Monte Carlo study that represents the NRI data, the model-based estimators are superior to design-based estimators and multinomial model estimators in terms of RMSE and relative RMSE. The use of the posterior predictive distribution validates the use of the variance model for the NRI application.

The approach based on the GD distribution has several advantages for the NRI application. The GD distribution allows greater flexibility than both the multinomial distribution and the Dirichlet distribution. The variance model allows us to incorporate auxiliary information in the design-based variance estimators. The model allows different covariates, regression parameters, and spatial effects for different categories.

The study generates several questions for future work. The proposed models assume that all proportions are greater than 0. While this is not an important limitation for this application, in the future we will consider a zero-inflated model that allows zeros for both estimated proportions and true values. An extension to include a temporal component has the potential utility for forecasting and estimation of change.

[Received February 2018. Accepted June 2018. Published Online July 2018.]

REFERENCES

Agresti, A. and Hitchcock, D. B. (2005). Bayesian inference for categorical data analysis. *Statistical Methods and Applications*, 14(3):297–330.

Banerjee, S., Carlin, B P., and Gelfand, A. E. (2014). *Hierarchical modeling and analysis for spatial data*. Crc Press.

Battese, G. E., Harter, R. M., and Fuller, W. A. (1988). An error-components model for prediction of county crop areas using survey and satellite data. *Journal of the American Statistical Association*, 83(401):28–36.

Berg, E. J. and Fuller, W. A. (2014). Small area prediction of proportions with applications to the Canadian Labour Force Survey. *Journal of Survey Statistics and Methodology*, 2(3):227–256.

Breidt, F. J. and Fuller, W. A. (1999). Design of supplemented panel surveys with application to the National Resources Inventory. *Journal of Agricultural, Biological, and Environmental Statistics*, 4(4):391–403.

Cho, M. J., Eltinge, J. L., Gershunskaya, J., and Huff, L. (2002). Evaluation of generalized variance function estimators for the US Current Employment Survey. In *Proceedings of the American Statistical Association, Survey Research Methods Section*, pages 534–539.

Congdon, P. (2005). *Bayesian models for categorical data*.

Connor, R. J. and Mosimann, J. E. (1969). Concepts of independence for proportions with a generalization of the Dirichlet distribution. *Journal of the American Statistical Association*, 64(325):194–206.

Dass, S. C., Maiti, T., Ren, H., and Sinha, S. (2012). Confidence interval estimation of small area parameters shrinking both means and variances. *Survey Methodology*, 38(2):173–187.

Datta, G. S., Lahiri, P., Maiti, T., and Lu, K. L. (1999). Hierarchical Bayes estimation of unemployment rates for the states of the US. *Journal of the American Statistical Association*, 94(448):1074–1082.

Deville, J.-C. and Tille, Y. (1998). Unequal probability sampling without replacement through a splitting method. *Biometrika*, 85(1):89–101.

Diggle, P. J., Tawn, J., and Moyeed, R. (1998). Model-based geostatistics. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 47(3):299–350.

Fay, R. E. and Herriot, R. A. (1979). Estimates of income for small places: an application of James-Stein procedures to census data. *Journal of the American Statistical Association*, 74(366a):269–277.

Ferrari, S. and Cribari-Neto, F. (2004). Beta regression for modelling rates and proportions. *Journal of Applied Statistics*, 31(7):799–815.

Gamerman, D. and Cepeda-Cuervo, E. (2013). Generalized Spatial Dispersion Models. In *Technical report, Universidade Federal do Rio de Janeiro*.

Gelman, A. (2006). Prior distributions for variance parameters in hierarchical models (comment on article by Browne and Draper). *Bayesian analysis*, 1(3):515–534.

Gelman, A., Carlin, J. B., Stern, H. S., and Rubin, D. B. (2014). Bayesian data analysis, volume 2. Taylor & Francis.

Gelman, A. and Rubin, D. B. (1992). Inference from iterative simulation using multiple sequences. *Statistical science*, 7(4):457–472.

Gilks, W. R. and Wild, P. (1992). Adaptive rejection sampling for Gibbs sampling. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 41(2):337–348.

Gomez-Rubio, V., Best, N., Richardson, S., Li, G., and Clarke, P. (2010). Bayesian statistics small area estimation. Technical report, Imperial College London (Unpublished). http://eprints.ncrm.ac.uk/1686/.

Han, W., Yang, Z., Di, L., and Mueller, R. (2012). CropScape: A Web service based application for exploring and disseminating US conterminous geospatial cropland data products for decision support. *Computers and Electronics in Agriculture*, 84:111–123.

He, Z. and Sun, D. (2000). Hierarchical Bayes estimation of hunting success rates with spatial correlations. *Biometrics*, 56(2):360–367.

Jiang, J. and Lahiri, P. (2006a). Estimation of finite population domain means: A model-assisted empirical best prediction approach. *Journal of the American Statistical Association*, 101(473):301–311.

——— (2006b). Mixed model prediction and small area estimation. *Test*, 15(1):1–96.

Jin, C., Zhu, J., Steen-Adams, M. M., Sain, S. R., and Gangnon, R. E. (2013). Spatial multinomial regression models for nominal categorical data: a study of land cover in Northern Wisconsin, USA. *Environmetrics*, 24(2):98–108.

Kalton, G. (1983). Compensating for Missing Survey Data. Technical report.

Militino, A.F., Ugarte, M.D., Goicoa, T., and González-Audícana, M. (2006). Using small area models to estimate the total area occupied by olive trees. *Journal of Agricultural, Biological, and Environmental Statistics*, 11(4):450–461.

Liu, B., Lahiri, P., and Kalton, G. (2007). Hierarchical Bayes modeling of survey-weighted small area proportions. In *Proceedings of the Survey Research Methods Section, American Statistical Association*, pages 3181–3186.

López-Vizcaíno, E., Lombardía, M. J., and Morales, D. (2013). Multinomial-based small area estimation of labour force indicators. *Statistical modelling*, 13(2):153–178.

——— (2015). Small area estimation of labour force indicators under a multinomial model with correlated time and area effects. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 178(3):535–565.

Lumley, T. (2011). *Complex surveys: a guide to analysis using R*. John Wiley & Sons.

Maiti, T., Ren, H., and Sinha, S. (2014). Prediction error of small area predictors shrinking both means and variances. *Scandinavian Journal of Statistics*, 41(3):775–790.

Maples, J., Bell, W., and Huang, E. T. (2009). Small area variance modeling with application to county poverty estimates from the American community survey. In *Proceedings of the Section on Survey Research Methods Section, American Statistical Association*, pages 5056–5067.

Molina, I., Saei, A., and José Lombardía, M. (2007). Small area estimates of labour force participation under a multinomial logit mixed model. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 170(4):975–1000.

Nusser, S. and Goebel, J. (1997). The National Resources Inventory: a long-term multi-resource monitoring programme. *Environmental and Ecological Statistics*, 4(3):181–204.

Petrucci, A. and Salvati, N. (2006). Small area estimation for spatial correlation in watershed erosion assessment. *Journal of agricultural, biological, and environmental statistics*, 11(2):169.

Pfeffermann, D. (2013). New important developments in small area estimation. *Statistical Science*, 28(1):40–68.

Polson, N. G. and Scott, J. G. (2012). On the half-Cauchy prior for a global scale parameter. *Bayesian Analysis*, 7(4):887–902.

Rao, J. N. and Molina, I. (2015). *Small area estimation*. John Wiley & Sons.

Särndal, C.-E., Swensson, B., and Wretman, J. (2003). Model assisted survey sampling. Springer Science & Business Media.

Simas, A. B., Barreto-Souza, W., and Rocha, A. V. (2010). Improved estimators for a general class of beta regression models. *Computational Statistics & Data Analysis*, 54(2):348–366.

Spiegelhalter, D. J., Best, N. G., Carlin, B. P., and Van Der Linde, A. (2002). Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 64(4):583–639.

Stehman, S. and Overton, W. (1989). Pairwise inclusion probability formulas in random-order, variable probability, systematic sampling. *Oregon State University, Technical Report*, 131:28.

Tille, Y. and Matei, A. (2016). *sampling: Survey Sampling. R package version 2.8.* https://cran.r-project.org/package=sampling.

Wong, T.-T. (1998). Generalized Dirichlet distribution in Bayesian analysis. *Applied Mathematics and Computation*, 97(2):165–181.

You, Y., Rao, J., and Dick, P. (2004). Benchmarking hierarchical Bayes small area estimators in the Canadian census undercoverage estimation. *Statistics in Transition*, 6(5):631–640.