Benchmarking a triplet of official estimates

Andreea L. Erciulescu, Nathan B. Cruze & Balgobin Nandram

Environmental and Ecological Statistics

ISSN 1352-8505
Volume 25
Number 4

Environ Ecol Stat (2018) 25:523-547
DOI 10.1007/s10651-018-0416-4

Springer

Environmental and Ecological Statistics (2018) 25:523–547
https://doi.org/10.1007/s10651-018-0416-4

CrossMark

Benchmarking a triplet of official estimates

Andreea L. Erciulescu[1] · Nathan B. Cruze[2] · Balgobin Nandram[3]

Received: 4 April 2018 / Revised: 18 September 2018 / Published online: 21 November 2018

Abstract
The publication of official statistics at different levels of aggregation requires a benchmarking step. Difficulties arise when a benchmarking method needs to be applied to a triplet of related estimates, at multiple stages of aggregation. For ratios of totals, external benchmarking constraints for the triplet (numerator, denominator, ratio) are that the weighted sum of denominator/numerator/ratio estimates equals to a constant. The benchmarking weight, applied to the ratio estimates, is a function of the denominator estimates. For example, the United States Department of Agriculture's National Agricultural Statistics Service's county-level, end-of-season acreage, production and yield estimates need to aggregate to the corresponding agricultural statistics district-level estimates, which also need to aggregate to the corresponding prepublished state-level values. Moreover, the definition of yield, as the ratio of production to harvested acreage, needs to hold at the county level, the agricultural statistics district level and the state level. We discuss different methods of applying benchmarking constraints to a triplet (numerator, denominator, ratio), at multiple stages of aggregation, where the denominator and the ratio are modeled and the numerator is derived. County-level and agricultural statistics district-level, end-of-season acreage, production and yield estimates are constructed and compared using the different methods. Results are illustrated for 2014 corn and soybean in Indiana, Iowa and Illinois.

Keywords Auxiliary information · Corn and soybean estimates · End-of season yield · Hierarchical Bayes · Ratio

Handling Editor: Pierre Dutilleul.

✉ Andreea L. Erciulescu
alerciulescu@gmail.com

[1] National Institute of Statistical Sciences, 1750 K Street, NW, Suite 1100, Washington, DC 20006-2306, USA

[2] USDA National Agricultural Statistics Service, 1400 Independence Avenue, SW, Room 6412 A, Washington, DC 20250-2054, USA

[3] Department of Mathematical Sciences, Worcester Polytechnic Institute and USDA National Agricultural Statistics Service, Stratton Hall, 100 Institute Road, Worcester, MA 01609, USA

1 Introduction

Statistical agencies and policy makers require that the official statistics for nested levels are consistent. For the case where estimates for an area containing small areas, are published before the estimates for the small areas are constructed, it is desirable that the estimates for the small areas aggregate to the prepublished values; the aggregation of total estimates is a simple sum, however, the aggregation of ratio estimates is a weighted sum. In this paper, we consider estimation for a set of related parameters, a *triplet* (numerator, denominator, ratio), at two levels of aggregation. This is a difficult problem because the final triplet estimates need to satisfy identity constraints (ratio of numerator to denominator) as well as benchmarking constraints at multiple levels.

The process of adjusting the estimates to satisfy external upper-level constraints is known as external, top–down, benchmarking. The agreement between lower-level and upper-level estimates is necessary and it provides confidence for policy makers when utilizing official estimates. Also, the benchmarking constraint provides external protection against an imperfect model, when the final estimates are produced using model-based methods; see (Pfeffermann and Barnard 1991).

Estimation and benchmarking methods produce reliable and reproducible official estimates at nested levels. Benchmarking methods were developed for estimates of a single quantity of interest, at one level of aggregation, by a number of authors, see for example (Pfeffermann and Barnard 1991; Wang et al. 2008; Nandram et al. 2014; Rao and Molina 2015). We note that (Ghosh and Steorts 2013) proposed a two-stage benchmarking methodology, with extensions to multiple, *unrelated*, parameters of interest. We benchmark the triplet (numerator, denominator, ratio) using draws from the posterior distribution that are constructed for the denominator total and for the ratio of two totals, as a result of independent model fitting. Estimates are produced using the draws from the posterior distribution of the two quantities, the third (the numerator total) being produced as a result. Our goal is to discuss the application of the benchmarking adjustment to the triplet of estimates and its consequences to maintaining the ratio identity at different aggregation levels.

The motivation for this research is county-level and agricultural statistics district-level yield estimation at United States Department of Agriculture's (USDA's) National Agricultural Statistics Service (NASS), where yield is defined as the ratio of production to harvested acreage. Hence the triplet (numerator, denominator, ratio) of interest is (production, harvested acreage, yield). NASS produces county-level and agricultural statistics district-level estimates (groups of neighboring counties within a state, hereafter denoted by ASDs) of acreage, production and yield that may contribute to the magnitude of payout in some agricultural programs. The variances for the total (acreage and production) survey estimates are calculated using the delete-a-group Jackknife method with 15 replicate groups, and include adjustments for nonsampling errors. The ratio (yield) survey estimator and its corresponding variance estimator are constructed using a second-order Taylor series expansion for the ratio. See (Kott 1989) for more details on the variance estimation for the survey estimates.

Currently, NASS publishes end-of-season county-level estimates using a top–down approach, for a unique year-state-commodity combination. The official statistics are a synthesis of survey and other data for small grains crops and major row crops, including corn and soybeans, achieved in a narrow timeframe, approximately twenty days between the survey final date and the publication date. More details about NASS production timelines can be found in Cruze et al. (2016). As a result, the official lower-lever (county and ASD) statistics aggregate to the official upper-level (ASD and state) statistics. National and state end-of-season crop estimates are produced using NASS's quarterly crops Acreage, Production and Stocks (APS) surveys, and published prior to county-level estimates. Agricultural statistics district-level estimates are then produced such that they aggregate to the state-level values. End-of-season crop estimates at the county and ASD levels are produced using the APS surveys, and their supplement County Agricultural Production Survey (CAPS). The county-level estimates need to aggregate to the prepublished state-level values, while the intermediate agreements with the ASDs need to hold. The survey county-level estimates may not aggregate to the prepublished state-level values (targets). The extent of discrepancy can vary by state and commodity crop; we note that, on average, in three cornbelt states (Illinois, Indiana, and Iowa), survey county totals for harvested area of corn and soybeans aggregate to approximately 92.8% of the state target, and county production totals aggregate to approximately 91.3% of the state target. In this paper, the prepublished state-level values are treated as targets, to which the county-level and the ASD-level estimates will be benchmarked to.

Because the current NASS estimation procedures use expert opinion to combine multiple sources of data including the aforementioned survey estimates, uncertainty measures for the official estimates are not available for publication. Models are being investigated to formalize the aforementioned composition of information, while giving rise to estimates and associated measures of uncertainty, for use as official statistics. A model-based estimation approach is proposed in Erciulescu et al. (2018), using a subarea-level model that incorporates multiple sources of information, to produce county-level and ASD-level acreage estimates and associated measures of uncertainty. The authors investigated different benchmarking methods for acreage totals and (Nandram et al. 2018) proposed a robust Bayesian benchmarking method for model-based estimates. Although the model-based methods in Erciulescu et al. (2018) and Nandram et al. (2018) could also be applied to yield, it is important to evaluate the effects of the different benchmarking adjustments for yield, as the ratio of production and acreage. The difficulty of the problem arises mainly from the need to estimate, preserve agreement between multiple levels of aggregation and preserve agreement between the three quantities of interest simultaneously.

To introduce the problem setting and notation, let $\theta_{ij}^{T1}, \theta_{ij}^{T2}$ and θ_{ij}^{R} denote the true (unknown) parameters of interest for the jth subarea, in the ith area (a group of subareas), corresponding to two totals and their ratio, respectively. In the application study, the subarea represents the county, the area represents the ASD, and the triplet $(T1, T2, R)$ represents the triplet (production, harvested acreage, yield). Let a_{T1}, a_{T2} and $a_R := (a_{T2})^{-1} a_{T1}$ denote the fixed targets for the totals and for the ratio, respectively. Then, the benchmarking constraints to be satisfied are

$$\sum_{j=1}^{n_{c_i}} w_{ij}^{\zeta} \hat{\theta}_{ij}^{\zeta,B} = \hat{\theta}_i^{\zeta,B},$$

$$\sum_{i=1}^{m} w_i^{\zeta} \hat{\theta}_i^{\zeta,B} = a_{\zeta}, \tag{1}$$

where $\hat{\theta}_{ij}^{\zeta,B}$ and $\hat{\theta}_i^{\zeta,B}$ denote the final estimates for the subareas and for the areas, respectively, for $\zeta \in \{T1, T2, R\}$, $j = 1, \ldots, n_{c_i}$ and $i = 1, \ldots, m$.

Without loss of generality, the methodology is illustrated for the subarea-to-target constraints, using subscript j to denote the subarea,

$$\sum_{j=1}^{n_c} w_j^{\zeta} \hat{\theta}_j^{\zeta,B} = a_{\zeta}, \tag{2}$$

where $n_c = \sum_{i=1}^{m} n_{c_i}$ is the total number of subareas. Area-level estimates are constructed using the information for all subareas $j = 1, \ldots, n_{c_i}$ in area i. Two cases are considered. The first case is inspired by Nandram et al. (2014) and Berg et al. (2014), where model-based estimates for a ratio (in their applications, yield or cash rental rate, the ratio of total rent to rented acreage) are benchmarked conditioning on the denominator estimates. For this case, the authors construct estimates for the ratio only, not for the triplet, the benchmarking weights for the ratios are functions of the denominator estimates, treated as fixed and known, and estimates for the numerator are constructed as a result. The numerator estimates are the product of the ratio estimates and the denominator estimates, $\hat{\theta}_j^{T1,B} = \hat{\theta}_j^{R,B} \hat{\theta}_j^{T2,B}$, with estimated variances conditional on the denominator estimates. Benchmarking adjustments applied to the final estimates of one quantity of interest are discussed in Sect. 6.4.6 in Rao and Molina (2015), for example, for a mean, a total or a ratio. Our contributions relative to these approaches are that the benchmarking adjustment is applied at the (posterior distribution) draw-level, instead of being applied to the posterior means, and that we benchmark a triplet of estimates, instead of estimates for one quantity only, for example, the ratio. Alternatively, we introduce a second case, when the benchmarking weights for the ratios are functions of the draws from the posterior distribution of the denominator. For this case, the numerator estimates are constructed using the product of the draws from the posterior distribution for the ratio and for the denominator. The variance estimation for the numerator estimates is improved since conditioning on the denominator estimates is not necessary.

The rest of the paper is organized as follows. In Sect. 2, we introduce the methods of applying the benchmarking constraints to the triplet (numerator, denominator, ratio) estimates and discuss the advantages and disadvantages for each method. Estimates for subarea and area levels are constructed for the methods described in Sect. 2. In Sect. 3, we illustrate our methods to produce end-of-season, model-based, county-level and ASD-level harvested acreage, production and yield estimates, for 2014 corn and soybeans in three selected states. Results for the application study are presented in Sect. 4. A discussion is provided in Sect. 5 and technical details are presented in "Appendix A".

2 Benchmarking a ratio using numerical approximation methods

The three parameters of interest are two totals and their ratio. Therefore, it is sufficient to construct Bayes estimates for two of the three parameters of interest, the third being constructed as a result. Moreover, since the benchmarking weights for the ratio depend on the denominator, Bayes estimates for the denominator need to be constructed first; only numerator and ratio may be derived. In this section, we introduce two estimation methods, by modeling the denominator total $T2$ and the ratio R independently, and deriving the numerator total $T1$ as a result. Estimating two quantities and deriving the third is a method consistent with the data collection, as a survey respondent may provide acreage and production (or yield) on the survey questionnaire. Although the methods apply to any benchmarking adjustments, in this paper, we provide illustrations of a ratio benchmarking [see (Rao and Molina 2015), Sect. 6.4.6, for a traditional ratio benchmarking approach].

Following the notation defined for (2), we assume a general model specification. Let $\tilde{\boldsymbol{\theta}}$ denote the vector of observed data (survey direct estimates) and let the vector of subarea parameters of interest be $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_{n_c})$. Let $\boldsymbol{\lambda}$ be a vector of nuisance parameters and let $\mathbf{X}$ be a vector of covariates. Assume the conditional distribution of $\tilde{\boldsymbol{\theta}}$ is $f(\tilde{\boldsymbol{\theta}}|\boldsymbol{\theta}, \boldsymbol{\lambda}, \mathbf{X})$ and that the prior is $\pi(\boldsymbol{\theta}, \boldsymbol{\lambda})$. Then, the posterior distribution of interest is $p(\boldsymbol{\theta}, \boldsymbol{\lambda}|\tilde{\boldsymbol{\theta}}, \mathbf{X}) \propto f(\tilde{\boldsymbol{\theta}}|\boldsymbol{\theta}, \boldsymbol{\lambda}, \mathbf{X})\pi(\boldsymbol{\theta}, \boldsymbol{\lambda})$, which needs to be proper.

Two independent Bayesian models are fit to the denominator total and the ratio survey estimates, respectively. Results are constructed using draws from the posterior distributions for the two quantities, $T2$ and R. Initial fitting of the models results in draws that do not satisfy (2): let $\theta_{jk}^{T2,noadj}$ and $\theta_{jk}^{R,noadj}$ denote the draws of θ_j^{T2} and θ_j^R, respectively, where k denotes the draw from the posterior distribution and $k = 1, \ldots, K$. For each of the two parameters $\zeta \in \{T2, R\}$, at every subarea $j = 1, \ldots, n_c$, the draws from the posterior distributions are:

$$
\begin{array}{c c}
j\backslash k & \begin{array}{ccccc} 1 & \quad 2 & \quad \ldots & \quad K \end{array} \\
\begin{array}{c} 1 \\ \\ 2 \\ \\ \vdots \\ \\ n_c \end{array} &
\begin{bmatrix}
\theta_{11}^{\zeta,noadj} & \theta_{12}^{\zeta,noadj} & \ldots & \theta_{1K}^{\zeta,noadj} \\
\theta_{21}^{\zeta,noadj} & \theta_{22}^{\zeta,noadj} & \ldots & \theta_{2K}^{\zeta,noadj} \\
\vdots & \vdots & \vdots & \vdots \\
\theta_{n_c1}^{\zeta,noadj} & \theta_{n_c2}^{T1/T2/R,noadj} & \ldots & \theta_{n_cK}^{\zeta,noadj}
\end{bmatrix},
\end{array}
$$

where (j, k) denotes the kth draw for subarea $j, k = 1, \ldots, K$. The jth subarea estimator is $\hat{\theta}_j^{\zeta,noadj} = K^{-1} \sum_{k=1}^{K} \theta_{jk}^{\zeta,noadj}$. Hereafter, it is assumed that the benchmarking weights for the denominator total, w_j^{T2}, are fixed and known (do not depend on k). So $w_{jk}^{T2} = w_j^{T2}$, for all $k = 1, \ldots, K$. The benchmarking constraint $\sum_j w_{jk}^{\zeta}\theta_{jk}^{\zeta,noadj} = a_\zeta$ may not be automatically satisfied, for fixed k, using the model output prior to adjustments.

Note that the pair $(\hat{\theta}_j^{T2,noadj}, \hat{\theta}_j^{R,noadj})$, $j = 1, \ldots, n_c$, is a set of estimators for the denominator and ratio, but the benchmarking constraints in (2) may not hold. Moreover, the benchmarking constraints in (2) may not hold at the (posterior distribution) draw level. We will use different methods to construct the weights w_{jk}^R that ensure the constraints in (2) and the identity between the ratio and the totals, at the (posterior distribution) draw level, after benchmarking adjustment, in Sects. 2.1 and 2.2 In each of the two sections, we will construct the estimator for the numerator, as a result of modeling the denominator and the ratio.

The estimates of the denominator and the ratio, $E(\theta_j^{T2}|data)$ and $E(\theta_j^R|data)$, are constructed using the traditional arithmetic mean of the draws:

$$K^{-1}\sum_{k=1}^{K}\theta_{jk}^{T2,noadj} = \hat{\theta}_j^{T2,noadj},$$
$$K^{-1}\sum_{k=1}^{K}\theta_{jk}^{R,noadj} = \hat{\theta}_j^{R,noadj}, \tag{3}$$

where $\hat{\theta}_j^{T2,noadj}$ and $\hat{\theta}_j^{R,noadj}$ are the Monte Carlo estimators of θ_j^{T2} and θ_j^R, respectively, before the benchmarking adjustment.

The benchmarking constraints, given in (2), are used to adjust the Bayes estimates, computed at the kth draw. For example, for a fixed k and using a *ratio* benchmarking adjustment,

$$\theta_{ik}^{T2,noadj}, \theta_{jk}^{T2} = \theta_{jk}^{T2,noadj}\frac{a_{T2}}{\sum_{j=1}^{n_c} w_{jk}^{T2}\theta_{jk}^{T2,noadj}},$$
$$\theta_{jk}^{R} = \theta_{jk}^{R,noadj}\frac{a_{R}}{\sum_{j=1}^{n_c} w_{jk}^{R}\theta_{jk}^{R,noadj}}. \tag{4}$$

Note that for each draw k, the subarea-level draws aggregate to the target values,

$$\sum_{j=1}^{n_c} w_{jk}^{T2}\theta_{jk}^{T2} = \sum_{j=1}^{n_c} w_{jk}^{T2}\theta_{jk}^{T2,noadj}\frac{a_{T2}}{\sum_{j=1}^{n_c} w_{jk}^{T2}\theta_{jk}^{T2,noadj}} = a_{T2},$$
$$\sum_{j=1}^{n_c} w_{jk}^{R}\theta_{jk}^{R} = \sum_{j=1}^{n_c} w_{jk}^{R}\theta_{jk}^{R,noadj}\frac{a_{R}}{\sum_{j=1}^{n_c} w_{jk}^{R}\theta_{jk}^{R,noadj}} = a_{R}. \tag{5}$$

If θ_{jk}^{T2} and θ_{jk}^R are the adjusted draws, then the posterior subarea means are estimated by

$$\widehat{E}(\theta_j^{T2}|\text{data and constraint})] = K^{-1}\sum_{k=1}^{K}\theta_{jk}^{T2} := \hat{\theta}_j^{T2,B},$$
$$\widehat{E}(\theta_j^{R}|\text{data and constraint}) = K^{-1}\sum_{k=1}^{K}\theta_{jk}^{R} := \hat{\theta}_j^{R,B}, \tag{6}$$

and the posterior subarea variances are estimated by

$$\widehat{Var}(\theta_j^{T2}|\text{data and constraint}) = (K-1)^{-1}\sum_{k=1}^{K}\left(\theta_{jk}^{T2} - K^{-1}\sum_{k=1}^{K}\theta_{jk}^{T2}\right)^2,$$
$$\widehat{Var}(\theta_j^{R}|\text{data and constraint}) = (K-1)^{-1}\sum_{k=1}^{K}\left(\theta_{jk}^{R} - K^{-1}\sum_{k=1}^{K}\theta_{jk}^{R}\right)^2. \tag{7}$$

Using the adjusted subarea-level draws θ_{jk}^{T2} and θ_{jk}^{R} [for example, defined in (4)], for all $j = 1, \ldots, n_{c_i}$ within area i, define the adjusted area-level draws

$$\theta_{ik}^{T2} = \sum_{j\in D_i} w_{jk}^{T2}\theta_{jk}^{T2},$$
$$\theta_{ik}^{R} = \sum_{j\in D_i} w_{jk}^{R}\theta_{jk}^{R}, \tag{8}$$

where D_i is the set of all subareas within area i. Then the posterior area means are estimated by

$$\widehat{E}(\theta_i^{T2}|\text{data and constraint}) = K^{-1}\sum_{k=1}^{K}\theta_{ik}^{T2} := \hat{\theta}_i^{T2,B},$$
$$\widehat{E}(\theta_i^{R}|\text{data and constraint}) = K^{-1}\sum_{k=1}^{K}\theta_{ik}^{R} := \hat{\theta}_i^{R,B}, \tag{9}$$

and the posterior area variances are estimated by

$$\widehat{Var}(\theta_i^{T2}|\text{data and constraint}) = (K-1)^{-1}\sum_{k=1}^{K}\left(\theta_{ik}^{T2} - K^{-1}\sum_{k=1}^{K}\theta_{ik}^{T2}\right)^2,$$
$$\widehat{Var}(\theta_i^{R}|\text{data and constraint}) = (K-1)^{-1}\sum_{k=1}^{K}\left(\theta_{ik}^{R} - K^{-1}\sum_{k=1}^{K}\theta_{ik}^{R}\right)^2. \tag{10}$$

We consider two choices for w_{jk}^{R} in Sects. 2.1 and 2.2 For Sect. 2.1 (method R11), $w_{jk}^{R} \propto \hat{\theta}_j^{T2,B}$, for $k = 1, \ldots, K$, and $\hat{\theta}_j^{T2,B}$ are treated as fixed and known. For Sect. 2.2 (method R12), $w_{jk}^{R} \propto w_{jk}^{T2}\theta_{jk}^{T2}$, for $k = 1, \ldots, K$. Both cases ensure the benchmarking constraints at the (posterior distribution) draw level, but only the first case at the subarea level. Similar choices of weights are investigated for the area-level estimates.

2.1 Fixed benchmarking weights (R11)

For this case, we define the subarea-level adjusted draws for the ratio, θ_{jk}^{R}, as in (4) with $w_{jk}^{R} := w_j^{R} = (a_{T2})^{-1}\hat{\theta}_j^{T2,B}$, for all $k = 1, \ldots, K$. Similarly, the area-level adjusted draws for the ratio, θ_{ik}^{R}, are defined as in (8) with $w_{ik}^{R} := w_{i}^{R} = (a_{T2})^{-1}\hat{\theta}_{i}^{T2,B}$, for all $k = 1, \ldots, K$.

Subarea-level draws for the numerator are constructed as

$$\theta_{jk}^{T1} := \hat{\theta}_{j}^{T2,B}\theta_{jk}^{R}, \tag{11}$$

and the estimator $\hat{\theta}_{j}^{T1,B}$ is constructed as

$$\widehat{E}(\theta_{j}^{T1}|\hat{\theta}_{j}^{T2,B}, \text{data and constraint}) = K^{-1}\sum_{k=1}^{K}\hat{\theta}_{j}^{T2,B}\theta_{jk}^{R} = \hat{\theta}_{j}^{T2,B}\hat{\theta}_{j}^{R,B} := \hat{\theta}_{j}^{T1,B},$$

satisfying automatically the benchmarking constraint (see A1 in "Appendix A"), with

$$\begin{aligned}
&\widehat{Var}(\theta_{j}^{T1}|\hat{\theta}_{j}^{T2,B}, \text{data and constraint}) \\
&= (K-1)^{-1}\sum_{k=1}^{K}\left(\hat{\theta}_{j}^{T2,B}\theta_{jk}^{R} - K^{-1}\sum_{k=1}^{K}\hat{\theta}_{j}^{T2,B}\theta_{jk}^{R}\right)^{2} \\
&= (K-1)^{-1}\left(\hat{\theta}_{j}^{T2,B}\right)^{2}\sum_{k=1}^{K}\left(\theta_{jk}^{R} - K^{-1}\sum_{k=1}^{K}\theta_{jk}^{R}\right)^{2} \\
&= \left(\hat{\theta}_{j}^{T2,B}\right)^{2}\widehat{Var}(\theta_{j}^{R}|\text{data and constraint}).
\end{aligned}$$

Area-level draws for the numerator are constructed as

$$\theta_{ik}^{T1} := \hat{\theta}_{i}^{T2,B}\theta_{ik}^{R}, \tag{12}$$

and the estimator $\hat{\theta}_{i}^{T1,B}$ is constructed as

$$\widehat{E}(\theta_{i}^{T1}|\hat{\theta}_{i}^{T2,B}, \text{data and constraint}) = K^{-1}\sum_{k=1}^{K}\hat{\theta}_{i}^{T2,B}\theta_{ik}^{R} = \hat{\theta}_{i}^{T2,B}\hat{\theta}_{i}^{R,B} := \hat{\theta}_{i}^{T1,B},$$

satisfying automatically the benchmarking constraint (see A1 in "Appendix A"), with

$$\begin{aligned}
&\widehat{Var}(\theta_{i}^{T1}|\hat{\theta}_{i}^{T2,B}, \text{data and constraint}) \\
&= (K-1)^{-1}\sum_{k=1}^{K}\left(\hat{\theta}_{i}^{T2,B}\theta_{ik}^{R} - K^{-1}\sum_{k=1}^{K}\hat{\theta}_{i}^{T2,B}\theta_{ik}^{R}\right)^{2} \\
&= (K-1)^{-1}\left(\hat{\theta}_{i}^{T2,B}\right)^{2}\sum_{k=1}^{K}\left(\theta_{ik}^{R} - K^{-1}\sum_{k=1}^{K}\theta_{ik}^{R}\right)^{2} \\
&= \left(\hat{\theta}_{i}^{T2,B}\right)^{2}\widehat{Var}(\theta_{i}^{R}|\text{data and constraint}).
\end{aligned}$$

Note that, while the benchmarking constraints are satisfied for all three parameters of interest, the posterior summaries for the numerator are conditional on the denominator estimates; see (11).

2.2 Random benchmarking weights (R12)

For this case, subarea-level benchmarking weights w_{jk}^{R} are set equal to the ratio of the weighted denominator total draws to the denominator total target $(a_{T2})^{-1} w_{jk}^{T2} \theta_{jk}^{T2} = (a_{T2})^{-1} w_{j}^{T2} \theta_{jk}^{T2}$. The benchmarking weights w_{jk}^{R} are applied to the draws $\theta_{jk}^{R,noadj}$, leading to the benchmarked draws θ_{jk}^{R} and to the adjusted ratio estimator $\hat{\theta}_{j}^{R,B}$. Draws for the numerator total θ_{jk}^{T1} are constructed as the product of the draws for the denominator total θ_{jk}^{T2} and the draws for the ratio θ_{jk}^{R},

$$\theta_{jk}^{T1} := \theta_{jk}^{T2} \theta_{jk}^{R}, \tag{13}$$

and the estimator $\hat{\theta}_{j}^{T1,B}$ is constructed as

$$\widehat{E}(\theta_{j}^{T1} | \text{data and constraint}) = K^{-1} \sum_{k=1}^{K} \theta_{jk}^{T1} := \hat{\theta}_{j}^{T1,B},$$

satisfying automatically the benchmarking constraint (see A2 in "Appendix A"), with

$$\begin{aligned}\widehat{Var}(\theta_{j}^{T1} | \text{data and constraint}) &= (K-1)^{-1} \sum_{k=1}^{K} \left(\theta_{jk}^{T1} - K^{-1} \sum_{k=1}^{K} \theta_{jk}^{T1} \right)^2 \\ &= (K-1)^{-1} \sum_{k=1}^{K} \left(\theta_{jk}^{T2} \theta_{jk}^{R} - K^{-1} \sum_{k=1}^{K} \theta_{jk}^{T2} \theta_{jk}^{R} \right)^2 .\end{aligned}$$

Under the choice of random weights, the benchmarking constraints are satisfied, at the (posterior distribution) draw level, for the three parameters of interest and the posterior summaries for the numerator are not conditional on the denominator estimates; see (13). However, the benchmarking constraints for the ratio are not (exactly) satisfied at the subarea and area levels, and the equality between the subarea-level ratio estimator and the ratio of the subarea-level estimators for the totals is not (exactly) satisfied; the differences in the subarea-level estimates are illustrated in A2 of "Appendix A".

The area-level estimators for the numerator are derived in a manner similar to the subarea-level estimators. The area-level weights w_{ik}^{R} are set equal to $(a_{T2})^{-1} \theta_{ik}^{T2}$ and applied to the draws $\theta_{ik}^{R,noadj}$, leading to the benchmarked draws θ_{ik}^{R} and to the adjusted ratio estimator $\hat{\theta}_{i}^{R,B}$. Draws θ_{ik}^{T1} are constructed as the product of θ_{ik}^{T2} and θ_{ik}^{R},

$$\theta_{ik}^{T1} := \theta_{ik}^{T2} \theta_{ik}^{R}, \tag{14}$$

and the estimator $\hat{\theta}_i^{T1,B}$ is constructed as

$$\widehat{E}(\theta_i^{T1}|\text{data}) = K^{-1}\sum_{k=1}^{K}\theta_{ik}^{T1} := \hat{\theta}_i^{T1,B},$$

satisfying automatically the benchmarking constraint (see A2 in "Appendix A"), with

$$\begin{aligned}\widehat{Var}(\theta_i^{T1}|\text{data and constraint}) &= (K-1)^{-1}\sum_{k=1}^{K}\left(\theta_{ik}^{T1} - K^{-1}\sum_{k=1}^{K}\theta_{ik}^{T1}\right)^2 \\ &= (K-1)^{-1}\sum_{k=1}^{K}\left(\theta_{ik}^{T2}\theta_{ik}^{R} - K^{-1}\sum_{k=1}^{K}\theta_{ik}^{T2}\theta_{ik}^{R}\right)^2.\end{aligned}$$

3 Application to the end-of-season crop estimates at NASS

Model-based estimates for 2014 end-of-season crop harvested acreage, production and yield are constructed for corn and soybeans in three selected states. Survey direct estimates are available from the NASS's summary of pooled data from the quarterly crops Acreage, Production and Stocks (APS) surveys and its supplement, the County Agricultural Production Survey (CAPS). Official county-level, ASD-level and state-level estimates of acreage, production and yield are obtained from NASS QuickStats at USDA (2018); recall that the official estimates are constructed using multiple data sources, hence they are different from the survey direct estimates. As stated earlier, the goal of this section is to use model-based estimation methods to produce county-level and ASD-level estimates for the triplet (harvested acreage, production, yield), such that the benchmarking constraints between the three levels (county, ASD and state) hold. The state-level values are published prior to substate-level estimation and considered as fixed targets in the benchmarking. Estimation is conducted state by state and commodity by commodity.

3.1 Survey direct estimates

Three states that are major United States producers of corn and soybeans are considered in this application: Illinois, Indiana and Iowa. The number of counties in each state equals 102, 92 and 99, respectively. The number of ASDs is 9 in each of the three states. The survey summary provides the county and ASD survey direct estimates and their corresponding sampling variances, for the three parameters of interest, the ratio definition being satisfied for the three sets of survey estimates; i.e., yield survey estimates equal to the ratio of production survey estimates to harvested acreage survey estimates. However, the benchmarking constraints in (2) do not hold for the survey estimates. The county-level survey estimates will be used as inputs into the models and for model assessement, and the ASD-level estimates will be used for model assessement only.

The county sample size, defined here by the number of records used to construct the county-level survey estimates, differs from state to state, commodity to commodity, and parameter to parameter. The county sample sizes range from 1 to 98, for corn in Illinois, Indiana and Iowa. The production and yield sample sizes are equal and are less than or equal to the acreage sample size. The estimated coefficients of variation (CVs) for the survey estimates increase as the county sample size decreases, and their ranges also differ from state to state, commodity to commodity, and parameter to parameter. For harvested acreage (HV) survey estimates, the CVs range from 8.1 to 92.3%, for production (PD) survey estimates, the CVs range from 8.6 to 100.0% and for yield (YD) survey estimates, the CVs range from 0.0 to 11.7%. While the CVs for yield survey estimates are smaller than 12%, the estimates do not satisfy the benchmarking constraint to the state-level yield. Figure 1 shows the distribution of the CVs of the county-level corn survey estimates, relative to the county-level corn sample sizes. Similar CVs and sample sizes are identified for soybeans.

3.2 Auxiliary sources of information

We select three sources of auxiliary information for covariates: the USDA Farm Service Agency's planted acreage data (FSA.PL), the USDA Natural Resources Conservation Service's National Commodity Crop Productivity Index (NCCPI) and the National Oceanic and Atmospheric Administration's weather data. The county-level covariates, FSA.PL and NCCPI, were selected from a larger pool of variables, using correlation analysis and model comparison criteria. As expected, exploratory data analysis shows that strong linear relationships exist between the survey acreage estimates and the FSA.PL, and that strong linear relationships exist between the survey yield estimates and the NCCPI values. The ASD-level covariate, NOAA March Precipitation (PCPN), was also selected based on correlation analysis and model comparison criteria, from an initial pool of 204 weather variables, including temperature, precipitation and drought indices. It is also known that March precipitation may impact the corn planting practices (see Erciulescu et al. 2018) and March is the first month of planting activity for corn and soybeans, across the United States, see USDA (2010).

3.3 A hierarchical Bayesian model

Following (Erciulescu et al. 2018), the proposed model is a subarea-level model, where the area represents the ASD and the subarea represents the county. Let $i = 1, \ldots, m$ be an index for the m ASDs in the state, $j = 1, \ldots, n_{c_i}$, be an index for the n_{c_i} counties in ASD i, and n_{ij} be the county sample size of the jth county in the ith ASD. The total number of counties in the state is $\sum_{i=1}^{m} n_{c_i} = n_c$ and the state sample size is $\sum_{i=1}^{m} \sum_{j=1}^{n_{c_i}} n_{ij} = n$. The county-level covariate values are x_{ij} and the ASD-level covariate values are z_i.

Let $\tilde{\theta}_{ij}$ be the survey estimate in county i and ASD j, σ_{ij}^2 be the sampling variance in county i and ASD j, and c_{ij} be known constants (to be specified). Illustrated for one state, one commodity and one parameter, let the hierarchical Bayesian subarea-level model be

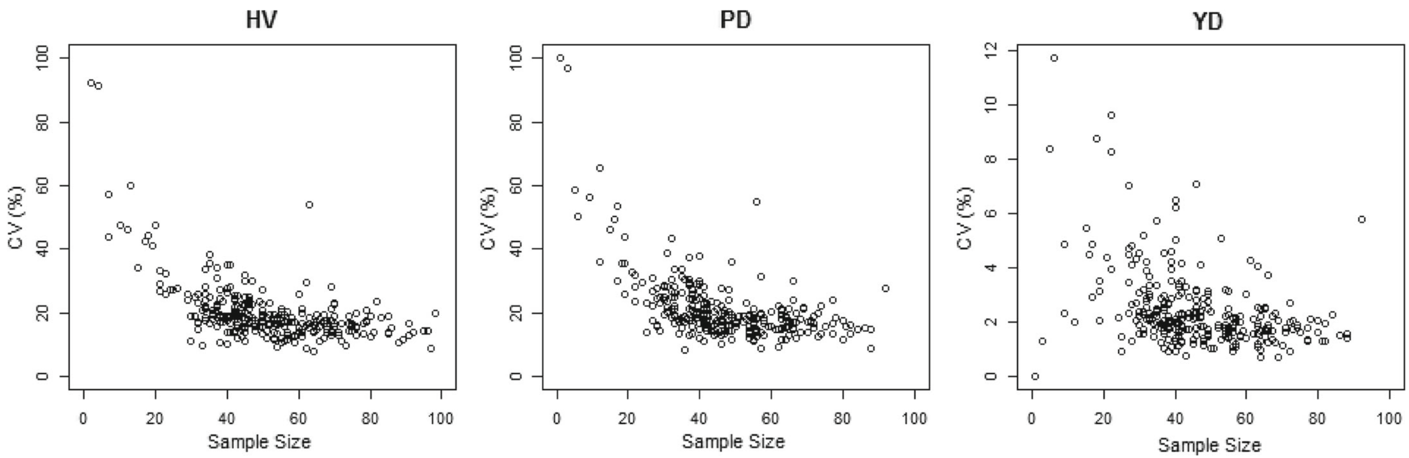

Fig. 1 *County*-level Survey Direct Estimates, CVs versus Sample Sizes 2014 *Corn* in Illinois, Indiana and Iowa. The plots illustrate the relationship between the coefficients of variation (CVs) of the county-level corn survey estimates and the county-level corn sample size, for harvested acreage (HV), production (PD) and yield (YD). The ranges of CV and sample size *differ* from state to state, from parameter to parameter and from commodity to commodity

$$
\begin{aligned}
c_{ij}^{-1}\tilde{\theta}_{ij}\Big| \left(\theta_{ij}, \sigma_{ij}^2\right) &\overset{ind}{\sim} N\left(\theta_{ij}, c_{ij}^{-2}\sigma_{ij}^2\right), \\
\theta_{ij}\Big| \left(\boldsymbol{\beta}, v_i, \sigma_u^2\right) &\overset{ind}{\sim} N\left(\left(1, x_{ij}, z_i\right)\boldsymbol{\beta} + v_i, \sigma_u^2\right), \\
v_i\big|\sigma_v^2 &\overset{iid}{\sim} N\left(0, \sigma_v^2\right).
\end{aligned} \tag{15}
$$

Model (15) borrows information from all the counties in an ASD and from all the ASDs in the state, while combining auxiliary information available at different levels of aggregation. To improve the symmetry of the distribution of the survey estimates, the model in (15) fitted to total survey estimates (acreage) uses $c_{ij} = n_{ij}$. The distribution of the county-level yield estimates is approximately symmetric, so the survey estimates are modeled without scaling. That is, when the model is fit to yield, $c_{ij} = 1$.

Exploratory data analysis suggested strong relationships between the survey direct estimates for totals and their associated sampling errors. To utilize these relationships, the model for total (acreage), has an additional hierarchical level, corresponding to a shrinkage model for sampling variances,

$$
\begin{aligned}
\left(n_{ij} - 1\right)\frac{\tilde{\sigma}_{ij}^2}{\sigma_{ij}^2}\Big|\sigma_{ij}^2 &\overset{ind}{\sim} \chi_{(n_{ij}-1)}^2, \\
log\left(c_{ij}^{-2}\sigma_{ij}^2\right)\Big|\left(\boldsymbol{\alpha}, \sigma^2\right) &\overset{ind}{\sim} N\left(\left(1, log\left(x_{ij}\right)\right)\boldsymbol{\alpha}, \sigma^2\right).
\end{aligned} \tag{16}
$$

A priori independent parameters and noninformative proper priors are considered for $(\boldsymbol{\alpha}', \boldsymbol{\beta}', \sigma_u^2, \sigma_v^2, \sigma^2)$. Specifically, the prior distributions for the model parameters $(\boldsymbol{\beta}', \boldsymbol{\alpha}')$ are normal distributions with mean and variance equal to the mean and 10^3 times the variance of the least squares estimates of $(\boldsymbol{\beta}', \boldsymbol{\alpha}')$, obtained from simple linear fits as described in Erciulescu et al. (2018). The prior distributions for the model variance components σ_u^2, σ_v^2, and σ^2 are $Uniform(0, 10^{13})$, $Uniform(0, 10^{13})$, and Inverse-Gamma$(10^3, 10^3)$, respectively.

3.3.1 Model fit and estimation

For each state and for each commodity, model (15, 16) is fit to the survey direct estimates of harvested acreage per unit and model (15) is fit to the CAPS survey estimates of yield, independently. The harvested acreage total, the production total and the yield ratio correspond to the $T2$, $T1$ totals and to the R ratio in Sect. 2, respectively. As a result of variable selection, each model is fit using one county-level covariate and one ASD-level covariate. The county-level covariate for acreage is the administrative acreage, divided by the county-level sample size. The county-level covariate for yield is the NCCPI. The ASD-level covariate is the March precipitation available from NOAA.

The models are fit using R JAGS and the posterior distributions constructed using MCMC simulation, using 10,000 Monte Carlo samples, 1000 burn-in samples, 3 chains, each thinned every 9 samples. Hence, the number of draws from the posterior distribution used for inference is 3000. Standard methods discussed in Gelman and Rubin (1992) and Geweke (1992) are used to monitor convergence: trace plots, multiple potential scale reduction factors (values less than 1.1) and the Geweke test of stationarity for each chain. Also, to monitor simulation accuracy, the effective sample size is constructed once the simulated chains have mixed.

We construct posterior summaries from the posterior distributions of the county-level and ASD-level estimates, under the different benchmarking scenarios presented in Sect. 2, treating the ij index as the j index and $K = 3000$ as the number of draws.

The ratio benchmarking adjustment is applied to county-level estimates, under the constraint to the fixed, prepublished state-level values. The benchmarking weights for yield are functions of the harvested acreage estimates. Let θ_j^R denote the yield parameter at the county-level, let θ_j^{T2} denote the harvested acreage parameter at the county-level and let θ_j^{T1} denote the production parameter at the county-level. Then the benchmarking weights for the acreage w_{jk}^{T2} are set equal to the county-level sample sizes, for all $k = 1, \ldots, K$, because the models are fit to the scaled totals. Hence, the benchmarking weights for production equal to the benchmarking weights for acreage $w_j^{T1} = w_j^{T2} = w_{jk}^{T2}$, for all $k = 1, \ldots, K$. The model-based point estimates are compared to the corresponding NASS official estimates (denoted by ASB), and the model-based standard errors of the point estimates are compared to the NASS survey errors.

4 Results for the application study

We present results for model-based estimates for the triplet (HV, PD, YD), for corn and soybean. The draws for HV and YD are used to construct posterior summaries for the denominator and the ratio, under the two choices of benchmarking weights for YD, as described in Sect. 2. The relative differences between the model-based estimates (MERB) and the ASB estimates, and the relative differences between the standard errors of the model-based estimates (MERBSE) and of the survey estimates (CAPSSE), are quantified using three metrics that protect the confidentiality of the data.

Results for the 2014 county-level HV estimates in Illinois, Indiana and Iowa, combined, are presented in Table 1, and are consistent with the results in Erciulescu et al. (2018). The median relative differences between the HV model-based estimates, ratio adjusted to satisfy benchmarking constraints, and the ASB estimates, are 3.54% for corn and 4.46% for soybeans, which are approximately equal to 0.43 and 0.54 model standard errors for corn and soybeans, respectively. Moreover the estimated error in the model-based HV estimates is lower than the estimated error in the survey HV estimates; the model-based estimates have variances approximately 50% lower than the variances of the survey estimates for both corn and soybeans. As a result, the estimated CVs for the model-based county-level HV estimates are approximately 40−45% of the corresponding estimated CVs for the county-level survey estimates. This is a result of the model's strength of borrowing information across counties and ASDs, and from auxiliary data.

Results for the 2014 county-level PD and YD estimates in Illinois, Indiana and Iowa are presented in Tables 2 and 3, and Fig. 2. Table 2 contains median absolute relative differences of county-level yield estimates, where the median is applied over all the

Table 1 Median absolute relative differences of *County-Level HV* 2014 *Corn and Soybeans* estimates in Illinois, Indiana and Iowa combined

	$\left\lvert\frac{MERB-ASB}{ASB}\right\rvert$ (%)	$\left\lvert\frac{MERB-ASB}{MERBSE}\right\rvert$	$\left\lvert\frac{MERBSE-CAPSSE}{CAPSSE}\right\rvert$ (%)
Corn	3.54	0.43	51.37
Soybeans	4.46	0.54	49.65

Table 2 Median absolute relative differences of *County-Level YD* 2014 *Corn and Soybeans* estimates in Illinois, Indiana and Iowa

State	Estimator	$\left\lvert\frac{MERB-ASB}{ASB}\right\rvert$ (%)	$\left\lvert\frac{MERB-ASB}{MERBSE}\right\rvert$	$\left\lvert\frac{MERBSE-CAPSSE}{CAPSSE}\right\rvert$ (%)
Corn				
IL	R11	0.50	0.30	5.37
	R12	0.50	0.29	5.32
IN	R11	0.59	0.27	6.08
	R12	0.59	0.27	5.95
IA	R11	0.56	0.34	7.84
	R12	0.56	0.34	7.78
Soybeans				
IL	R11	0.55	0.26	3.69
	R12	0.55	0.26	3.55
IN	R11	0.82	0.44	3.47
	R12	0.82	0.44	3.64
IA	R11	0.97	0.52	6.16
	R12	0.97	0.52	6.20

counties in a particular state. The point estimates are similar when benchmarked using fixed and random weights because the variance of the random weights is negligible (in the range $(10^{-6}, 10^{-7})$), see the results in the first column, in Table 2. Although, under random weights, the equality between the yield estimates and the ratio of production estimates to harvested acreage estimates is not exactly satisfied, the percentage medians of relative differences, relative to the ASB estimates are equal for the two cases (R11 and R12), up to two significant digits. Recall that R11 is the method of Sect. 2.2 and R12 is the method of Sect. 2.2 The results in the first two columns in Table 2 show that most of the yield county-level estimates are within $0-1\%$ from the corresponding ASB estimates, or within $0-0.52$ standard errors from the corresponding ASB estimates; 99.7% of the ASB estimates fall inside the 95% credible intervals of the corresponding model-based estimates. The reduction in the county-level standard errors of the survey estimates is illustrated in the last column in Table 2, the percentage medians of relative differences are in the range $3.47-7.84\%$.

Table 3 contains median absolute relative differences of county-level productions estimates, where the median is applied over all the counties in a particular state. Sim-

Table 3 Median absolute relative differences of *County-Level PD* 2014 *Corn and Soybeans* estimates in Illinois, Indiana and Iowa

State	Estimator	$\left\lvert\frac{MERB-ASB}{ASB}\right\rvert$ (%)	$\left\lvert\frac{MERB-ASB}{MERBSE}\right\rvert$	$\left\lvert\frac{MERBSE-CAPSSE}{CAPSSE}\right\rvert$ (%)
Corn				
IL	R11	2.17	1.14	88.52
	R12	2.17	0.31	58.13
IN	R11	4.49	2.12	88.11
	R12	4.48	0.43	38.83
IA	R11	4.58	2.61	88.39
	R12	4.58	0.65	53.59
Soybeans				
IL	R11	3.12	1.37	86.47
	R12	3.13	0.37	49.63
IN	R11	6.76	3.05	86.84
	R12	6.76	0.61	38.31
IA	R11	5.91	3.21	87.34
	R12	5.91	0.83	54.22

ilarly to the results for yield, the point estimates are similar when benchmarked using fixed and random weights because the variance of the random weights is negligible (in the range $(10^{-6}, 10^{-7})$), see the results in the first column, in Table 3. The results in the first two columns in Table 3 show that most of the production county-level estimates are within $3-7\%$ from the corresponding ASB estimates; 52% of the ASB estimates fall inside the 95% credible intervals of the corresponding model-based estimates. However, in the last column in Table 3 we show that the variance estimates for production (R11) are artificially decreased, as a result of conditioning on the acreage estimates: the percentage medians of relative differences in the standard errors of the estimates are in the range $86.47-88.52\%$ for the (R11) method, and in the range $38.31-58.13\%$ for the (R12) method. Hence, the percentage medians of relative differences in the point estimates, given in the second column in Table 3 are as high as 3.21%, for method (R11).

Results for the 2014 ASD-level corn PD and YD estimates in Illinois, Indiana and Iowa, are presented in Tables 4 and 5, and Fig. 3. The organization of results in Tables 4 and 5, and Fig. 3 is similar to the organization of results in Tables 2 and 3, and Fig. 2, respectively. Table 4 contains median absolute relative differences of ASD-level yield estimates, where the median is applied over all the ASDs in a particular state. Because the ASD-level estimates are aggregates of the county-level estimates, the percentage medians of relative differences, relative to the ASB estimates are equal for the two cases (R11 and R12), up to two significant digits, as was the case for the county-level estimates. The results in the first two columns in Table 4 show that most of the yield ASD-level estimates are within 0.5% from the corresponding ASB estimates, or within 0.5% standard errors from the corresponding ASB estimates, better agreement than it

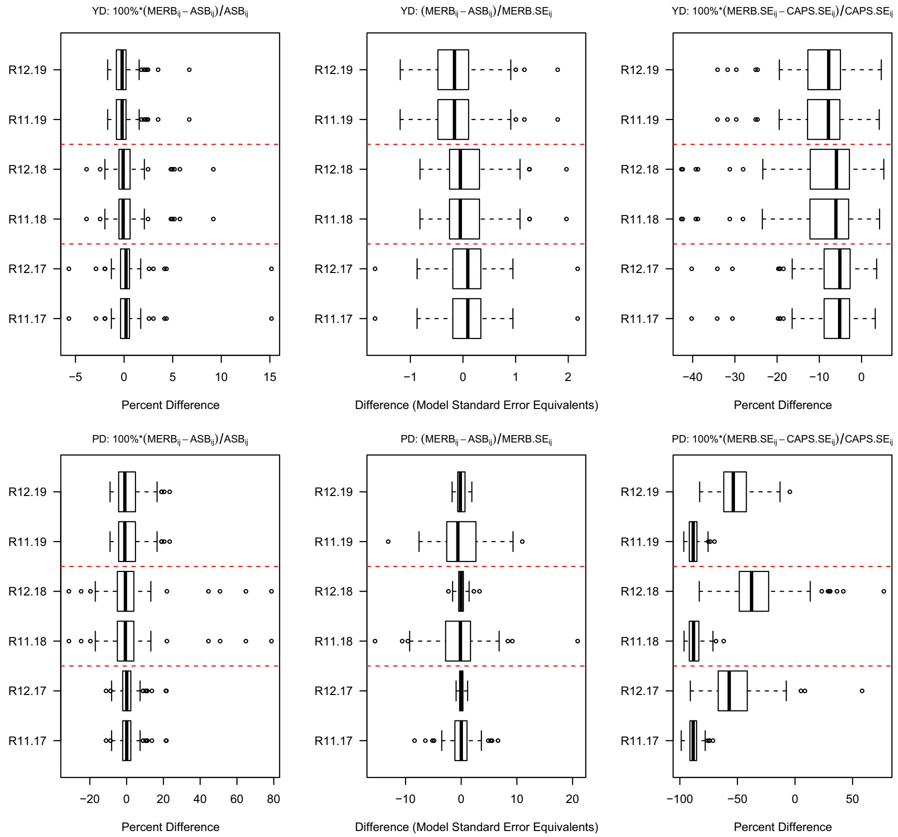

Fig. 2 *County*-level Model-based Estimates versus ASB Estimates 2014 *Corn* in Illinois, Indiana and Iowa. The box plots illustrate the relationship between the county-level survey direct estimates, the model-based estimates (R11, R12) and the ASB estimates, for production and yield. The model-based estimates are ratio benchmarked using the methods described in Sect. 2. The three horizontal sets of box plots correspond to the three states, denoted by their FIPS codes (17 for Illinois, 18 for Indiana and 19 for Iowa). One county with sample size equal to one was removed from the visual representations because its (approximately zero) corresponding survey variance decreased the visual precision. Results are similar for soybeans

was observed at the county-level. The reduction in the ASD-level standard errors of the survey estimates is illustrated in the last column in Table 4, the percentage medians of relative differences are in the range 7.68−28.98%, greater than it was observed at the county-level. It is expected that the absolute relative differences between the model-based ASD-level estimates and the corresponding official estimates be smaller than the absolute relative differences for the county-level because the survey direct estimates are also more reliable at the ASD-level than at the county-level due to larger sample sizes at the ASD-level than at the county-level.

Table 5 contains median absolute relative differences of ASD-level productions estimates, where the median is applied over all the ASDs in a particular state. The results in the first column in Table 5 show that most of the production ASD-level estimates are within 1.37−4.35% from the corresponding ASB estimates. However, in the last

Table 4 Median absolute relative differences of *ASD-Level YD* 2014 *Corn and Soybeans* estimates in Illinois, Indiana and Iowa

State	Estimator	$\left\lvert\frac{MERB-ASB}{ASB}\right\rvert$ (%)	$\left\lvert\frac{MERB-ASB}{MERBSE}\right\rvert$	$\left\lvert\frac{MERBSE-CAPSSE}{CAPSSE}\right\rvert$ (%)
Corn				
IL	R11	0.25	0.31	13.61
	R12	0.25	0.31	12.77
IN	R11	0.27	0.32	23.93
	R12	0.27	0.32	21.31
IA	R11	0.20	0.33	9.08
	R12	0.20	0.32	7.68
Soybeans				
IL	R11	0.31	0.49	13.60
	R12	0.31	0.48	12.76
IN	R11	0.44	0.67	28.98
	R12	0.44	0.64	26.86
IA	R11	0.29	0.44	20.42
	R12	0.29	0.43	19.48

Table 5 Median absolute relative differences of *ASD-Level PD* 2014 *Corn and Soybeans* estimates in Illinois, Indiana and Iowa

State	Estimator	$\left\lvert\frac{MERB-ASB}{ASB}\right\rvert$ (%)	$\left\lvert\frac{MERB-ASB}{MERBSE}\right\rvert$	$\left\lvert\frac{MERBSE-CAPSSE}{CAPSSE}\right\rvert$ (%)
Corn				
IL	R11	1.40	2.48	89.18
	R12	1.40	0.44	49.33
IN	R11	2.66	3.77	88.60
	R12	2.66	0.61	24.20
IA	R11	3.54	6.97	87.60
	R12	3.54	1.04	13.69
Soybeans				
IL	R11	1.37	1.68	86.59
	R12	1.37	0.48	33.86
IN	R11	3.94	4.09	88.21
	R12	3.94	0.64	26.19
IA	R11	4.35	7.95	86.69
	R12	4.35	1.27	13.88

column in Table 5 we show that the variance estimates for production (R11) are artificially decreased, as a result of conditioning on the acreage estimates: the percentage medians of relative differences in the standard errors of the estimates are in the range 86.69−89.18% for the (R11) method, and in the range 13.69−49.33% for the (R12)

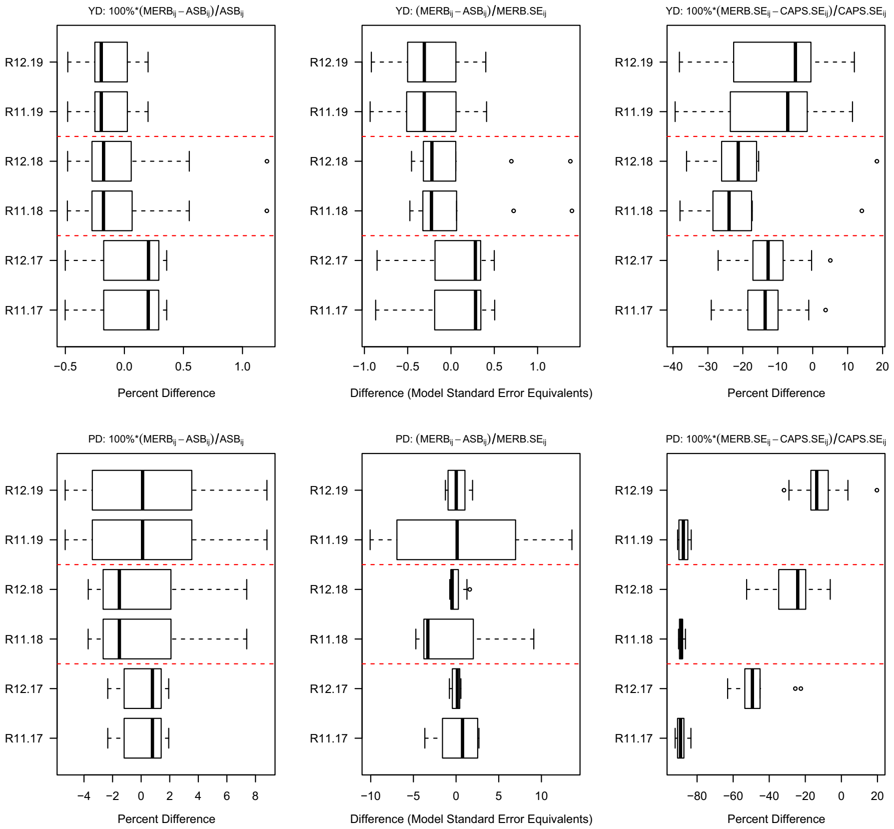

Fig. 3 *ASD*-level Model-based Estimates versus ASB Estimates 2014 *Corn* in Illinois, Indiana and Iowa. The box plots illustrate the relationship between the ASD-level survey direct estimates, the model-based estimates (R11, R12) and the ASB estimates, for production and yield. The model-based estimates are ratio benchmarked using the methods described in Sect. 2. The three horizontal sets of box plots correspond to the three states, denoted by their FIPS codes (17 for Illinois, 18 for Indiana and 19 for Iowa). Results for soybeans are similar

method. Hence, the percentage medians of relative differences in the point estimates, given in the second column in Table 5 are as high as 7.95%, for method (R11).

Median results for the three choices of metrics are presented in Tables 2, 3, 4 and 5. For completeness, we present distributions of the three metrics (as percentages) in Figs. 2 and 3, for corn. There is good agreement between the model-based estimates and the official statistics and the model-based estimates are more precise than the corresponding survey-estimates.

5 Discussion

NASS's interest in producing reliable end-of-season yield estimates and associated standard errors, subject to constraints at multiple aggregation levels, has motivated the research presented in this paper. The problem of estimating and benchmarking three related quantities of interest leads to the investigation of the effect of benchmarking constraints for the (numerator, denominator, ratio) triplet, for example the (production, harvested acreage, yield) triplet. We presented different methods of constructing model-based estimates for two totals and their ratio at lower-level areas that aggregate to fixed values at upper-levels. In this paper, we introduce and compare methods of applying the benchmarking adjustments to model-based estimates at multiple stages. The methodology presented applies to *any* benchmarking method. We chose to illustrate the methods using the ratio benchmarking adjustment, one of the two classic benchmarking adjustments: ratio and difference. The reason we did not pick the difference adjustment is that it may lead to negative estimates for the ratio of positive totals. The applicability of these methods is not restricted to yield estimation, as the methods can be applied to any ratio estimation, for example, wage rates, poverty rates or body mass index.

The (R11) application of benchmarking adjustment to ratio estimates, where the estimates of the denominator are treated as fixed and known, is common practice at NASS for projects where only one quantity is of interest (such as yield forecasting or cash rental rate estimation). We showed that the (R11) estimates are conditional on the denominator estimates, resulting in an underestimated variance for production estimates. As a solution to produce reliable numerator estimates, we introduced method (R12). For the application study, the construction (R12) provides a valuable set of estimates for the triplet, the inconsistencies in the ratio identity may be avoided with a simple rounding. Moreover, the estimated CVs for the county-level yield (R12) estimates are approximately $80-90\%$ of the corresponding estimated CVs for the county-level survey estimates and the estimated CVs for the county-level production (R12) estimates are approximately $40-45\%$ of the corresponding estimated CVs for the county-level survey estimates.

The methods illustrated in this paper are a good fit for a selected state-year-commodity combination. Some minor adjustments are needed, including state-year-commodity-specific variables selection. Specific to the benchmarking methods proposed in this manuscript, additional adjustments would be needed to account for county-district-state-specific structures, since we can only discuss triplet estimation for a county with survey data available on both totals (acreage and production) and their ratio (yield).

We presented harvested acreage, production and yield estimates for 2014 end-of-season corn and soybeans in Illinois, Indiana and Iowa, three states that are major United States producers of corn and soybeans. Posterior summaries are available for the county-level estimates and ASD-level estimates, for the triplet that closely satisfies the benchmarking constraints at both levels. Extensions of the methods, to produce posterior summaries for any intermediate aggregation level between county-level and state-level, such as groups of counties or groups of ASDs, is straightforward. In this paper, we presented posterior means and posterior variances and compared the results to the survey estimates and to the official estimates. We showed consistency between the *modeled* YD estimates (R11, R12) with survey and official estimates. In an operation environment, either (R11) or (R12) may be used, with a caution that (R11) is conditional on the denominator. If the user is concerned with production and yield estimation only, assuming that good harvested acreage values are already available, then the (R11) construction may be used, keeping the harvested acreage values for the denominator as fixed and known.

The emerging remote sensing estimation procedures provide good sources of yield estimates, available for selected commodities and for selected states. See, for example, Johnson (2014). However, the remote sensing estimates are subject to coverage errors due to the environment, they are not benchmarked to the state-level values, and measures of uncertainty, associated with the estimates, are not available. Applying a benchmarking adjustment to such estimates may or may not fall under one of the methods illustrated in this paper, and it is left for future investigation. Our preliminary results show that the remote sensing estimates follow a different pattern than the model-based yield estimates (R11, R12). Nevertheless, the remote sensing yield estimates and other remote sensing data (such as Normalized Difference Vegetation Indexes) may be used as potential covariates in the yield models.

Acknowledgements The findings and conclusions in this preliminary publication have not been formally disseminated by the U.S. Department of Agriculture and should not be construed to represent any agency determination or policy. This research was supported in part by the intramural research program of the U.S. Department of Agriculture, National Agriculture Statistics Service. Dr. Nandram's work was supported by a Grant from the Simons Foundation (#353953, Balgobin Nandram). The authors thank the reviewers for their comments and suggestions. The authors are grateful to Wendy Barboza, Valbona Bejleri, Dave Johnson and Linda Young from the National Agricultural Statistics Service.

Appendix A

A1. Fixed benchmarking weights

For a ratio benchmarking adjustment, the R11 subarea-level numerator estimates satisfy the benchmarking constraint. The proof is as follows

$$\begin{aligned}
\sum_{j=1}^{n_c} \hat{\theta}_j^{T1,B} &= \sum_{j=1}^{n_c} \hat{\theta}_j^{T2,B} \hat{\theta}_j^{R,B}, \\
&= \sum_{j=1}^{n_c} \hat{\theta}_j^{T2,B} K^{-1} \sum_{k=1}^{K} \theta_{jk}^{R}, \text{ by (6)} \\
&= \sum_{j=1}^{n_c} \hat{\theta}_j^{T2,B} K^{-1} \sum_{k=1}^{K} \theta_{jk}^{R,noadj} a_R \left(\sum_{j=1}^{n_c} w_{jk}^{R} \theta_{jk}^{R,noadj} \right)^{-1}, \text{ by (4)} \\
&= a_R K^{-1} \sum_{k=1}^{K} \sum_{j=1}^{n_c} \hat{\theta}_j^{T2,B} \theta_{jk}^{R,noadj} \left(\sum_{j=1}^{n_c} w_{jk}^{R} \theta_{jk}^{R,noadj} \right)^{-1}, \\
&\quad \text{since } a_R \text{ is constant} \\
&= a_R K^{-1} \sum_{k=1}^{K} \sum_{j=1}^{n_c} \hat{\theta}_j^{T2,B} \theta_{jk}^{R,noadj} \left(\sum_{j=1}^{n_c} a_{T2}^{-1} \hat{\theta}_j^{T2,B} \theta_{jk}^{R,noadj} \right)^{-1},
\end{aligned}$$

$$
\begin{aligned}
&\text{by definition of } w_{jk}^R \\
&= a_R a_{T2} K^{-1} \sum_{k=1}^{K} \left(\sum_{j=1}^{n_c} \hat{\theta}_j^{T2,B} \theta_{jk}^{R,noadj} \right) \left(\sum_{j=1}^{n_c} \hat{\theta}_j^{T2,B} \theta_{jk}^{R,noadj} \right)^{-1}, \\
&\text{since } a_{T2} \text{ is constant} \\
&= a_R a_{T2} = a_{T1}.
\end{aligned}
$$

Similarly, the R11 area-level numerator estimates satisfy the benchmarking constraints. The proof follows immediately,

$$
\begin{aligned}
\sum_{i=1}^{m} \hat{\theta}_i^{T1,B} &= \sum_{i=1}^{m} \hat{\theta}_i^{T2,B} \hat{\theta}_i^{R,B}, \\
&= \sum_{i=1}^{m} K^{-1} \sum_{k=1}^{K} \hat{\theta}_i^{T2,B} \theta_{ik}^{R}, \text{ by (6)} \\
&= \sum_{i=1}^{m} K^{-1} \sum_{k=1}^{K} a_{T2} w_i^R \theta_{ik}^{R}, \text{ by definition of } w_{ik}^R \\
&= a_{T2} a_R = a_{T1}.
\end{aligned}
$$

The ratio estimator for the set of all subareas is the weighted sum of the subarea-level ratio estimators,

$$
\begin{aligned}
\sum_{j=1}^{n_c} w_j^R \hat{\theta}_j^{R,B} &= \sum_{j=1}^{n_c} w_j^R K^{-1} \sum_{k=1}^{K} \theta_{jk}^{R}, \text{ by (6)} \\
&= K^{-1} \sum_{k=1}^{K} \sum_{j=1}^{n_c} w_j^R \theta_{jk}^{R}, \text{ exchanging summations} \\
&= K^{-1} \sum_{k=1}^{K} \sum_{j=1}^{n_c} w_j^R \theta_{jk}^{R,noadj} a_R \left(\sum_{j=1}^{n_c} w_j^R \theta_{jk}^{R,noadj} \right)^{-1}, \text{ by (4)} \\
&= K^{-1} \sum_{k=1}^{K} a_R, \text{ since } a_R \text{ is constant} \\
&= a_R.
\end{aligned}
$$

A2. Random benchmarking weights

Note that, from the derivation of production total estimates, the benchmarking weights for the two totals are equal. Hence, $w_j^{T1} = w_j^{T2} = w_{jk}^{T2}$, for all $k = 1, \ldots, K$.

For a ratio benchmarking adjustment, the R12 subarea-level numerator estimates satisfy the benchmarking constraint. The proof is as follows

$$
\begin{aligned}
\sum_{j=1}^{n_c} w_j^{T1}\hat{\theta}_j^{T1,B} &= \sum_{j=1}^{n_c} w_j^{T1} K^{-1} \sum_{k=1}^{K} \theta_{jk}^{T1} \\
&= \sum_{j=1}^{n_c} w_j^{T1} K^{-1} \sum_{k=1}^{K} \theta_{jk}^{T2}\theta_{jk}^{R}, \text{ by (13)} \\
&= \sum_{j=1}^{n_c} w_j^{T1} K^{-1} \sum_{k=1}^{K} \sum_{j=1}^{n_c} \theta_{jk}^{T2}\theta_{jk}^{R,noadj} a_R \left(\sum_{j=1}^{n_c} w_{jk}^{R}\theta_{jk}^{R,noadj}\right)^{-1}, \\
&\quad \text{by (4)} \\
&= a_R K^{-1} \sum_{k=1}^{K} \sum_{j=1}^{n_c} w_j^{T1} a_{T2} w_{jk}^{R} \left(w_{jk}^{T2}\right)^{-1} \theta_{jk}^{R,noadj} \\
&\quad \left(\sum_{j=1}^{n_c} w_{jk}^{R}\theta_{jk}^{R,noadj}\right)^{-1}, \text{ since } a_R \text{ is constant and} \\
&\quad \text{by definition of } w_{jk}^{R} \\
&= a_R a_{T2} K^{-1} \sum_{k=1}^{K} \left(\sum_{j=1}^{n_c} w_j^{T1} \left(w_{jk}^{T2}\right)^{-1} w_{jk}^{R}\theta_{jk}^{R,noadj}\right) \\
&\quad \left(\sum_{j=1}^{n_c} w_{jk}^{R}\theta_{jk}^{R,noadj}\right)^{-1}, \text{ since } a_{T2} \text{ is constant} \\
&= a_R a_{T2} K^{-1} \sum_{k=1}^{K} \left(\sum_{j=1}^{n_c} w_{jk}^{R}\theta_{jk}^{R,noadj}\right) \left(\sum_{j=1}^{n_c} w_{jk}^{R}\theta_{jk}^{R,noadj}\right)^{-1}, \\
&\quad \text{since } w_j^{T1} = w_{jk}^{T2}, \text{ for all } k = 1, \ldots, K \\
&= a_R a_{T2} = a_{T1}.
\end{aligned}
$$

The ratio estimator for the set of all subareas is the weighted sum of the subarea-level ratio estimators,

$$
\begin{aligned}
\sum_{j=1}^{n_c} w_j^{R}\hat{\theta}_j^{R,B} &= \sum_{j=1}^{n_c} w_j^{R} K^{-1} \sum_{k=1}^{K} \theta_{jk}^{R}, \text{ by (8)} \\
&= \sum_{j=1}^{n_c} \left(K^{-1}\sum_{k'} w_{jk'}^{R}\right)\left(K^{-1}\sum_{k=1}^{K} \theta_{jk}^{R}\right),
\end{aligned}
$$

$$\text{since } K^{-1}\sum_{k'} w^R_{jk'} = (a_{T2})^{-1}\,\hat{\theta}^{T2,B}_j = w^R_j$$

$$\neq K^{-2}\sum_{k=1}^{K}\sum_{k=1}^{K}\sum_{j=1}^{n_c} w^R_{jk}\theta^R_{jk}$$

$$= K^{-2}\sum_{k=1}^{K}\sum_{k=1}^{K} a_R = a_R.$$

The subarea-level numerator total estimator is

$$\hat{\theta}^{T1,B}_j = K^{-1}\sum_{k=1}^{K}\theta^{T1}_{jk} = K^{-1}\sum_{k=1}^{K}\theta^{T2}_{jk}\theta^R_{jk}, \text{ by (13)}$$

$$\neq \left(K^{-1}\sum_{k'}^{K}\theta^{T2}_{jk'}\right)\left(K^{-1}\sum_{k=1}^{K}\theta^R_{jk}\right).$$

References

Berg E, Cecere W, Ghosh M (2014) Small area estimation for county-level farmland cash rental rates. J Surv Stat Methodol 2:1–37

Cruze NB, Erciulescu AL, Nandram B, Barboza WJ, Young LJ (2016) Developments in model-based estimation of county-level agricultural estimates. In: ICES V proceedings. American Statistical Association, Alexandria

Erciulescu AL, Cruze NB, Nandram B (2018) Model-based county-level crop estimates incorporating auxiliary sources of information. J R Stat Soc Ser A. https://doi.org/10.1111/rssa.12390

Gelman A, Rubin DB (1992) Inference from iterative simulation using multiple sequences. Stat Sci 7:457–511

Geweke J (1992) Evaluating the accuracy of sampling-based approaches to calculating posterior moments. In: Bernardo JM, Berger JO, Dawid AP, Smith AFM (eds) Bayesian statistics. Clarendon Press, Oxford, p 4

Ghosh M, Steorts RC (2013) Two-stage benchmarking as applied to small area estimation. TEST: Off J Span Soc Stat Oper Res 22(4):668–670

Johnson D (2014) An assessment of pre- and within-season remotely sensed variables for forecasting corn and soybean yields in the United States. Remote Sens Environ 141:116–128

Kott P (1989) Mathematical formulae for the 1989 survey processing system (SPS) summary, p 26. https://www.nass.usda.gov/Education_and_Outreach/Reports,_Presentations_and_Conferences/Survey_Reports/Mathematical%20Formulae%20for%20the%201989%20Survey%20Processing%20System%20(SPS)%20Summary.pdf. Accessed 11 Nov 2018

Nandram B, Berg E, Barboza W (2014) A hierarchical Bayesian model for forecasting state-level corn yield. Environ Ecol Stat 21(3):507–530

Nandram B, Erciulescu AL, Cruze N (2018) Bayesian benchmarking of the Fay-Herriot model using random deletion. Surv Methodol (in press)

Pfeffermann D, Barnard C (1991) Some new estimators for small area means with application to the assessment of farmland values. J Bus Econ Stat 9:31–42

Rao JNK, Molina I (2015) Small area estimation. Wiley, Hoboken

USDA NASS (2010) Field crops. Usual planting and harvesting dates. http://usda.mannlib.cornell.edu/usda/current/planting/planting-10-29-2010.pdf. Accessed 11 Nov 2018

USDA NASS (2018) QuickStats. https://quickstats.nass.usda.gov/. Accessed 11 Nov 2018

Wang J, Fuller W, Qu Y (2008) Small area estimation under a restriction. Surv Methodol 34:29–36

Andreea L. Erciulescu is Research Associate at the National Institute of Statistical Sciences, working at the United States Department of Agriculture's National Agricultural Statistics Service. Her research interests include small area estimation, mixed models, bootstrap, survey statistics, official statistics.

Nathan B. Cruze is a mathematical statistician in the Research and Development Division at the United States Department of Agriculture's National Agricultural Statistics Service.

Balgobin Nandram is Senior Professor of Statistics, Mathematical Sciences, Worcester Polytechnic Institute, MA. His research is on hierarchical Bayesian models and small area estimation focusing on data analysis in numerous applications such as Agriculture, Health Sciences, Social Sciences and Business.