Journal of Agricultural and Resource Economics 47(3):634–655 ISSN: 1068-5502 (Print); 2327-8285 (Online)

doi: 10.22004/ag.econ.313319

Estimating Crop Yield Densities for Counties with Missing Data

Eunchun Park, Ardian Harri, and Keith H. Coble

Crop yield densities are often estimated at the county level. However, county-level yield data providers often omit county records due to low participation or other reasons. The data omission can undermine insurance premiums' credibility and thereby lead to restrictions on the provision of area insurance products in specific locations. To address this problem, we propose a novel Bayesian spatial interpolation method to estimate crop yield densities for counties with missing data. Empirical results indicate that our approach is consistently superior to the benchmark approaches. Importantly, our approach offers noticeable estimation accuracy even at a significant level of data omission.

Key words: Bayesian spatial interpolation, crop yield density estimation, data omission, spatial dependence

Introduction

In this study, we investigate how to use spatial dependence of crop yields to estimate crop yield densities for counties with a significant level of missing yield data or no data at all. Although the spatial dependence of crop yields has been broadly discussed in the literature, a methodological approach to estimate yield densities in the presence of missing data has not yet been thoroughly investigated.

Crop yield densities are often estimated at the county-level due to a lack of long series of individual farm-level data and to support area-based commodity programs and crop insurance policy designs. For instance, commodity programs such as the Agricultural Risk Coverage (ARC) and the Supplemental Coverage Option (SCO) programs need county-level crop yield data. Therefore, many studies have used county-level crop yield data (Coble et al., 1996; Annan et al., 2014; Ker, Tolhurst, and Liu, 2016; Zhang, 2017; Park, Brorsen, and Harri, 2019; Ker and Tolhurst, 2019; Ramsey and Goodwin, 2019; Liu and Ker, 2020a,b; Ramsey, 2020).

However, two primary data providers, the National Agricultural Statistics Service (NASS) and the Risk Management Agency (RMA), often omit certain counties due to confidentiality concerns that would only be resolved with greater sampling and respondent burden. In the case of NASS data, over 24% of the corn yield records and 35% of the soybean records are omitted from 1950 to 2017.[1] Based on the 2019 NASS corn production dataset, the production level of the counties with missing

Eunchun Park (park@uark.edu) is an assistant professor in the Department of Agricultural Economics and Agribusiness at the University of Arkansas. Ardian Harri is a professor and Keith H. Coble is the vice president of the Division of Agriculture, Forestry, and Veterinary Medicine at Mississippi State University.
This work was supported by the Office of the Chief Economist of the USDA.

Review coordinated by Anton Bekkerman.

[1] NASS reports complete series of county-level corn yield data from 1955 to 2017 for 57 out of 87 Minnesota counties, 64 out of 92 Indiana counties, and 48 out of Missouri 114 counties. The data omission is more severe in non-Corn Belt states. For instance, only 7 out of 82 Mississippi counties and 2 out of Colorado 64 have complete yield series. In the case of winter wheat, NASS reports complete series of county-level data for 23 out of 114 Missouri counties and 13 out of 56 Montana counties.

records is approximately 19% of the total production in the United States, valued at more than $10 billion. The omission of the most recent data makes the data omission problem more significant. This is primarily due to a declining survey response rate. According to Ridolfo, Boone, and Dickey (2013), the response rates on NASS county agricultural production surveys have declined in the last 2 decades. Response rates were 80%–85% in the early 1990s but have fallen to around 60% in some counties in recent years. This problem is especially challenging for crop insurance premium rating since the use of the most recent data is essential for estimating crop yield density (Liu and Ker, 2020b).

Because of generally high crop insurance program participation, the Agriculture Improvement Act of 2018 (also known as the 2018 Farm Bill) directed the USDA to prioritize using RMA data over NASS data (Li et al., 2020). However, the missing data problem also affects the RMA dataset. For example, around 20% of county-level irrigated corn yield records are omitted, and almost 26% of the nonirrigated corn yields are omitted. The problem is magnified when soybean data are examined. Over 36% of irrigated yields and almost 40% of the nonirrigated soybean yield reports are omitted.

County-level data generally have a short time dimension. With a few exceptions, most of the previous approaches for crop yield density estimations—including the Harri–Coble–Ker–Goodwin (HCKG) approach, on which the current RMA rating is based on (Harri et al., 2011)—estimate densities of individual counties separately. These separately estimated densities are less accurate due to the short history and limited information (Park, Brorsen, and Harri, 2019, 2020). Thus, such inaccuracy is a significant problem of the counties with missing data. There are counties where the crop insurance programs are not offered, specifically for minor crops and marginal producing areas, due to missing data.

Our approach also has broader applications in the context of developing countries' risk management programs, where the missing data problem can be more problematic. Currently, most insurance programs in developing countries are established based on index-based insurance instead of area-based insurance due to the limited data provision. However, index insurance contains naturally inherent design risk (Miranda, 1991; Jensen, Barrett, and Mude, 2016), which is the difference between the index and the underlying covariate losses.[2] Hence, the risk management ability of such programs might be limited.

We use an out-of-sample premium rating (cede/retain) game developed by Ker and McGowan (2000) to examine our approach's performance. The rates derived using our approach are compared with the rates from the HCKG approach (Harri et al., 2011) that the current RMA rating is based on and the Bayesian model averaging (BMA) approach (Ker, Tolhurst, and Liu, 2016). For an empirical application, we employ county-level corn yield data from NASS for the states of Iowa, Maryland, and Colorado and winter wheat yield data for the states of Kansas, Indiana, and Colorado.

Our results show that the missing data problem significantly reduces the premium rate-making credibility of current federal crop insurance programs. We also find that the proposed approach offers a more accurate and less sensitive premium rating than the benchmark approaches in the presence of the data omission. Our results consistently indicate a preference for our approach in all crop/state combinations.

Literature Review

A growing body of literature has investigated ways to obtain density estimation accuracy by incorporating a spatial dependence structure of crop yields into the density estimation process. This is because crop yield densities are likely to be spatially dependent since nearby locations can share similar climate, geological features, and other unknown factors that could affect crop yields (Annan et al., 2014; Du et al., 2015).

[2] Miranda (1991) separates basis risk into two parts: the idiosyncratic (systematic) and design components. The idiosyncratic risk is the difference between an individual's losses and average losses for an area (group), and the design risk is the difference between the index and the underlying covariate losses.

Li and Racine (2003) used a nonparametric kernel approach to smooth parameters for a joint probability density function. Racine and Ker (2006) employed the nonparametric kernel approach for crop insurance rating. Zhang (2017) proposed a density ratio estimator method to estimate crop yield densities when the number of observations is small. The density ratio estimator method regards individual density as a distortion from the baseline density. The technique offers accuracy gains when the historical data dimension is short.

There are also many studies based on Bayesian statistics. Ozaki et al. (2008) suggested a Bayesian method to generate crop yield densities by importing data from the neighboring counties under a normality assumption. However, the method gives equal weights to the data from the neighboring counties and gives zero weights beyond the adjoining counties. Ker, Tolhurst, and Liu (2016) estimated crop yield distribution using the BMA method, which imports information from outside counties' yield history. They found that the BMA approach is strongly preferred to the individual estimation approaches (e.g., the HCKG approach) when the historical dataset is limited. Woodard (2016) employed the BMA approach to obtain a weighted average of trend parameters in crop insurance rating. Liu and Ker (2020a) extended the BMA approach for borrowing information across both time and space.

Bayesian kriging is another type of Bayesian approach. Park, Brorsen, and Harri (2019) used Bayesian kriging to estimate spatially smoothed tail densities of crop yield under the extreme value theory (EVT). The approach assumes a spatial process among the tail density parameters and updates the structure under the Bayesian updating algorithm. They found that their approach generates accuracy gains by taking into account the spatial structure of crop yield data. More recently, Ramsey and Goodwin (2019) used a Bayesian quantile regression approach that borrows information across space and quantile levels under the Bayesian kriging framework.

Despite efforts to explore possible options to get accuracy gains, previous studies have only used balanced datasets. Indeed, counties with missing data have been discarded from the final dataset. Therefore, the superiority of previous studies' approaches might be limited since they measure accuracy gains only in counties with complete data. Unlike previous studies, we include all counties in the original dataset, regardless of data omission level, and recover crop yield densities of counties with missing data.

Bayesian Modeling Framework

We use the Bayesian hierarchical structure to obtain estimates for our approach, referred to as the Park–Harri–Coble (PHC) approach from here. The Bayesian hierarchical structure incorporates three layers: the likelihood layer, the process layer, and the prior layer. The likelihood layer forms a parametric probability density under normality.[3]

Likelihood Layer

We define the likelihood layer that models the crop yield density. Let y_{it} be the crop yield of county i at year t, where $i = 1,\ldots,N$ and $t = 1,\ldots,T$. Since we assume normality, the likelihood layer can be formed as

$$P_1(\boldsymbol{Y} \mid \boldsymbol{\beta},\boldsymbol{\gamma},\boldsymbol{\Theta}) = \prod_{t=1}^{T} \frac{1}{\sqrt{2\pi\,|\Lambda_t|}} \exp \frac{(\boldsymbol{y}_t - \boldsymbol{\beta}\boldsymbol{x}_t)\,,\Lambda_t^{-1}\,(\boldsymbol{y}_t - \boldsymbol{\beta}\boldsymbol{x}_t)}{2}, \tag{1}$$

[3] Theoretically, any parametric distributional assumption, such as Gamma or Beta, can be applied in our PHC approach. However, if there is a significant level of data omission (30% or 60%), the model with nonnormality does not converge well. The main objective of the research is to suggest a practical way to estimate county-level crop yield density when there are considerable levels of data omission. Therefore, we mainly use the model with the normality assumption in the study. However, we report the results from the Beta distribution in the case of full data and a moderate level (10%) of the data omission in Online Supplement A (see www.jareonline.org).

where $\boldsymbol{y}_t$ is a vector of crop yield at year t that spans all counties, $\boldsymbol{y}_t = [y_{1t},\ldots,y_{Nt}]'$, and thus $Y = [\boldsymbol{y}_1,\ldots,\boldsymbol{y}_T]$, $\boldsymbol{x}_t$ is a 2×1 vector of explanatory variables at year t that includes intercept and a linear trend variable, $\boldsymbol{\beta} = [\boldsymbol{\beta}_1,\boldsymbol{\beta}_2]$ is a $N\times 2$ vector of the mean equation coefficients, where $\boldsymbol{\beta}_1 = [\beta_{1i},\ldots,\beta_{1N}]'$ and $\boldsymbol{\beta}_2 = [\beta_{2i},\ldots,\beta_{2N}]'$, Λ_t is a variance matrix at year t, $\Lambda_t = \text{diag}(\boldsymbol{\omega}_t)$, which is structured by a vector of standard deviation equations, $\boldsymbol{\omega}_t = \boldsymbol{\gamma}\boldsymbol{x}_t$, where $\boldsymbol{\gamma} = [\boldsymbol{\gamma}_1,\boldsymbol{\gamma}_2]$ is a $N\times 2$ vector of the standard deviation equation coefficients, $\boldsymbol{\gamma}_1 = [\gamma_{1i},\ldots,\gamma_{1N}]'$, and $\gamma_2 = [\gamma_{2i},\ldots,\gamma_{2N}]'$.

Note that posteriors of $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ are conditionally drawn by hyper parameters priors $\boldsymbol{\Theta} = [\boldsymbol{\delta}_1,\boldsymbol{\delta}_2,\boldsymbol{\vartheta}_1,\boldsymbol{\vartheta}_2,\theta_{\beta 1},\theta_{\beta 2}\theta_{\gamma 1},\theta_{\gamma 2},\rho_{\beta 1},\rho_{\beta 2},\rho_{\gamma 1},\rho_{\gamma 2}]'$ through the process layer. The hyper parameters consist of the deterministic constant and kriging parameters that determine the spatial smoothing structure of the parameters of $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$.

Process Layer

The process layer updates posteriors of $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$. These county-specific coefficients are assumed to follow a multivariate Gaussian spatial process.[4] Then the coefficients are spatially smoothed by kriging parameters of sill (ρ) and range (θ) via a spatial covariance function Σ, which is an $N\times N$ matrix structured by a function of standardized Euclidean distances (D_{ij}) between counties i and j calculated from longitude/latitude coordinates.[5]

Under this setting, we accommodate not only county-specific trends in the mean equation (β_{2i}) but also county-specific heteroskedasticity from the standard deviation trend coefficient γ_{2i}. The county-specific coefficients are obtained by assuming the following multivariate Gaussian spatial process such that

$$\begin{aligned}\boldsymbol{\beta}_k \mid \boldsymbol{\delta}_k,\theta_{\beta k},\rho_{\beta k} &\sim \text{MVGP}\left(\boldsymbol{\delta}_k,\Sigma_{\beta k}\right)\\ \boldsymbol{\gamma}_k \mid \boldsymbol{\vartheta}_k,\theta_{\gamma k},\rho_{\gamma k} &\sim \text{MVGP}\left(\boldsymbol{\vartheta}_k,\Sigma_{\gamma k}\right),\end{aligned} \tag{2}$$

where $k = 1,2$; $\boldsymbol{\beta}_k$ and $\boldsymbol{\gamma}_k$ are vectors of the county-specific intercepts and trend parameters defined in the previous layer; $\boldsymbol{\delta}_k$ and $\boldsymbol{\vartheta}_k$ are vectors of the deterministic part of each coefficients that are uniform across all counties where $\boldsymbol{\delta}_k = [\delta_k,\ldots,\delta_k]'$ and $\boldsymbol{\vartheta}_k = [\vartheta_k,\ldots,\vartheta_k]'$; and $\Sigma_{\beta k}$ and $\Sigma_{\gamma k}$ are corresponding spatial covariance matrices for these coefficients.[6] Intuitively, these spatially smoothed parameters ($\boldsymbol{\beta}_k$ and $\boldsymbol{\gamma}_k$) are modeled as a deterministic constant plus a spatial random effect. The spatial random effect generated from the Gaussian spatial process is included in the mean so that the final posteriors of the parameters are county specific.

---

[4] Under the multivariate Gaussian spatial process, any spatially distributed random variables (i.e., mean and standard deviation coefficient of counties β_i and γ_i) are correlated and multivariate normally distributed with correlation dependent on distance. Specifically, the level of correlation between coefficients for locations i and j is determined by the distance between the two locations. Since these coefficients are spatially correlated, the crop yield densities of counties, generated by a posterior predictive distribution, are spatially correlated as well.

[5] In the study, we use an exponential-type spatial matrix where

$$\text{cov}\left(\beta_i,\beta_j\right) = \Sigma = \rho e^{-D_{ij}/\theta} = \rho\begin{bmatrix} 1 & & e^{-D_{1N}/\theta} \\ \vdots & \ddots & \vdots \\ e^{-D_{N1}/\theta} & & 1 \end{bmatrix}.$$

[6] Since the spatial process is based on a Bayesian framework, posteriors of $\boldsymbol{\beta}$ are multivariate normally distributed in terms of space. Suppose $\boldsymbol{\beta} \sim \text{MVGP}(\boldsymbol{\delta},\Sigma)$ and we draw K number of random samples. Then, for any kth $N\times 1$ sampled vector $\boldsymbol{\beta}_k = [\beta_{1k},\ldots,\beta_{Nk}]'$, the average of the vector elements should be close to δ, where $\frac{1}{N}\sum_{i=1}^{N}\beta_{ik} \approx \delta$.

We can now define the second layer of the hierarchy by multiplying the Gaussian spatial processes, such that

$$
\begin{aligned}
P_2(\boldsymbol{\beta},\boldsymbol{\gamma} \mid \boldsymbol{\Theta}) = & \frac{1}{\sqrt{(2\pi)^N \left|\Sigma_{\beta 1}\right|}} \exp\left[-\frac{1}{2}\left(\boldsymbol{\beta}_1 - \boldsymbol{\delta}_1\right)' \Sigma_{\beta 1}^{-1}\left(\boldsymbol{\beta}_1 - \boldsymbol{\delta}_1\right)\right] \\
& \times \frac{1}{\sqrt{(2\pi)^N \left|\Sigma_{\beta 2}\right|}} \exp\left[-\frac{1}{2}\left(\boldsymbol{\beta}_2 - \boldsymbol{\delta}_2\right)' \Sigma_{\beta 2}^{-1}\left(\boldsymbol{\beta}_2 - \boldsymbol{\delta}_2\right)\right] \\
& \times \frac{1}{\sqrt{(2\pi)^N \left|\Sigma_{\gamma 1}\right|}} \exp\left[-\frac{1}{2}\left(\boldsymbol{\gamma}_1 - \boldsymbol{\vartheta}_1\right)' \Sigma_{\gamma 1}^{-1}\left(\boldsymbol{\gamma}_1 - \boldsymbol{\vartheta}_1\right)\right] \\
& \times \frac{1}{\sqrt{(2\pi)^N \left|\Sigma_{\gamma 2}\right|}} \exp\left[-\frac{1}{2}\left(\boldsymbol{\gamma}_2 - \boldsymbol{\vartheta}_2\right)' \Sigma_{\gamma 2}^{-1}\left(\boldsymbol{\gamma}_2 - \boldsymbol{\vartheta}_2\right)\right].
\end{aligned} \tag{3}
$$

Prior Layer

Prior selection is one of the most critical parts to sample good posteriors, specifically for a model with a complex marginal posterior structure, like the model in this study. For most of the parameters in the model, noninformative priors are used for the posterior sampling, yet some parameters are sampled via informative priors. All vectors of the deterministic mean/variance coefficients $(\boldsymbol{\delta}_1,\boldsymbol{\delta}_2,\boldsymbol{\vartheta}_1,\boldsymbol{\vartheta}_2)$ are sampled using noninformative multivariate Gaussian priors, $MVN(\mathbf{0},10^2\boldsymbol{I})$. Similarly, we impose general noninformative inverse Gamma priors, $IG(0.1,0.1)$, for the sill parameters $((\rho_{\beta 1},\rho_{\beta 2},\rho_{\gamma 1},\rho_{\gamma 2})$.

However, setting up priors for the range parameters $(\theta_{\beta 1},\theta_{\beta 2}\theta_{\gamma_1},\theta_{\gamma_2})$, which determine spatial dependence structure (i.e., the maximum distance of spatial correlation) among mean/variance coefficients, is more difficult than others. Following Bayesian statistic studies (Banerjee, Carlin, and Gelfand, 2004; Cooley, Naveau, and Poncet, 2006; Cooley, Nychka, and Naveau, 2007) and agricultural economics studies (Park, Brorsen, and Harri, 2019), we impose informative priors on the range parameters based on the spatial information of the empirical dataset. First, we normalize the longitude/latitude coordinates and calculate all possible distance D_{ij} between counties i and j. Then we give uniform priors for all range parameters $(\theta_{\beta 1},\theta_{\beta 2}\theta_{\gamma 1},\theta_{\gamma 2})$ such that Uniform Uniform$(0,2 * \max(D_{ij}))$, where $\max(D_{ij})$ is the maximum distance among all D_{ij}. Then the final prior layer can be structured as

$$
P_3(\Theta) = p(\delta_1)\, p(\delta_2)\, p(\vartheta_1) \ldots p\left(\rho_{\gamma 1}\right) p\left(\rho_{\gamma 2}\right). \tag{4}
$$

Therefore, by Bayes's theorem, we now have the joint posterior distribution $P(\boldsymbol{\beta},\boldsymbol{\gamma},\boldsymbol{\Theta} \mid \boldsymbol{Y})$ by multiplying three densities from each layer— $P_1(\boldsymbol{Y} \mid \boldsymbol{\beta},\boldsymbol{\gamma},\boldsymbol{\Theta})$, $P_2(\boldsymbol{\beta},\boldsymbol{\gamma} \mid \boldsymbol{\Theta})$, and $P_3(\boldsymbol{\Theta})$—such that

$$
P(\boldsymbol{\beta},\boldsymbol{\gamma},\boldsymbol{\Theta} \mid \boldsymbol{Y}) \propto P_1(\boldsymbol{Y} \mid \boldsymbol{\beta},\boldsymbol{\gamma},\boldsymbol{\Theta}) P_2(\boldsymbol{\beta},\boldsymbol{\gamma} \mid \boldsymbol{\Theta}) P_3(\boldsymbol{\Theta}). \tag{5}
$$

We sample the final posteriors using the Metropolis–Hastings (MH) steps within a Gibbs sampler algorithm written in R. Online Supplement B (see www.jareonline.org) describes the derivation of the joint marginal posterior and computational details.

Review of Premium Calculation Procedures

This section reviews differences in the premium rating procedures of the HCKG, BMA, and PHC approaches. In the following section, we conduct the out-of-sample premium rating game (Ker and

McGowan, 2000) to provide an out-of-sample prediction comparison of the PHC with the HCKG (Harri et al., 2011) and BMA (Ker, Tolhurst, and Liu, 2016) approaches.

HCKG Approach

The first benchmark approach, suggested by Harri et al. (2011), is the current approach used by US government agencies. Under this approach, a premium rating is conducted using a two-step procedure. That is, a trend is estimated in the first stage (two-knots linear spline function) followed by the second-stage heteroskedasticity adjustment of the residuals from the first-stage estimation. Let $\hat{\varepsilon}_{it}$ and $\hat{y}_{it}$ denote the detrended residuals and fitted values from the first-stage estimation. Then the HCKG approach accounts for the heteroskedasticity via the following auxiliary regression:

$$\ln(\hat{\varepsilon}_{it}^2) = \alpha_0 + \alpha_1 \ln(\hat{y}_{it}) + \epsilon_{it}. \tag{6}$$

The heteroskedasticity-adjusted yields, $\hat{y}_{it}^*$, can be obtained from a one-step-ahead forecast yield, $\hat{y}_{i,T+1}$. and the estimated coefficient $\hat{\alpha}_1$ from equation (6), such that

$$\hat{y}_{it}^* = \hat{y}_{i,T+1} + \hat{\varepsilon}_{it} \left(\frac{\hat{y}_{i,T+1}}{\hat{y}_{it}} \right)^{\frac{\hat{\alpha}_1}{2}}, \tag{7}$$

where $\hat{\varepsilon}_{it} \left(\frac{\hat{y}_{i,T+1}}{\hat{y}_{it}} \right)^{\frac{\hat{\alpha}_1}{2}}$ is the heteroskedasticity-adjusted residuals.

After obtaining the heteroskedasticity-adjustment yield, $\hat{y}_{it}^*$, an empirical premium rate is calculated from the following equation:

$$\text{prem}_{i,T+1}^{RMA} = \frac{1}{T} \sum_{t=1}^{T} \max \left[0, \frac{\lambda \hat{y}_{i,T+1} - \hat{y}_{it}^*}{\hat{y}_{i,T+1}} \right], \tag{8}$$

where λ is the coverage level.

Under the empirical rating procedure, a lack of observations could be problematic because the premiums are calculated from the residuals. Since the approach estimates premiums independently using each county's data, a county with a small number of observations due to missing data issues (thus, a small number of residuals) could not provide accurate premium rates due to limited information in the small sample.[7] Moreover, no premiums would be calculated for counties with yield histories that are less than the minimum sample size (the sample size needed to have nonzero degrees of freedom) required to estimate the first-stage trend model.

BMA Approach

The BMA approach has, in general, been used to account for model uncertainty. Ker, Tolhurst, and Liu (2016) adapted it to get accuracy gains by modeling unknown spatial similarity among a set of county-level densities by importing information from other counties. The density generating scheme of the approach is entirely data-driven. For instance, information from other counties with some explanatory power to fit the target county's data would receive weights to generate the final target county's density. The BMA approach is a two-step estimation procedure. In the first step, county-level densities (ϕ_i) under a normal distribution (can also work for a normal-mixture) are estimated individually from their own historical observations and the Bayesian information criterion (BIC), BIC_i^i, is calculated. In the next step, the approach utilizes the individual parameter estimates

[7] We should note that the final premiums offered by the RMA have some additional adjustments beyond the HCKG approach. These adjustments could partly reduce the effect of the missing data. For instance, RMA uses some restrictions on the trend coefficient estimates at the region level. However, the premium rating is still based on independent estimation by county.

of the target county i from the first step to explain other counties' (county j) yield realizations and calculates the BICs, say BIC_j^i. Then, the weight of county j's density to reproduce county i's yield realizations (w_j^i) is estimated by using the following formula:

$$w_j^i = \frac{\exp\left[-\frac{1}{2}\text{BIC}_j^i\right]}{\sum_{k=1}^{K}\exp\left[-\frac{1}{2}\text{BIC}_k^i\right]}. \tag{9}$$

Therefore, the final county-level density, f_i, is a weighted average of individual density estimates (ϕ_i) and the density estimates from outside of the county (ϕ_j), such that

$$f_i = \sum_{j=1}^{K} w_j^i \phi_j. \tag{10}$$

After estimating the final density, f_i, the premium rate for county i in year $T+1$ for the area-based insurance program with coverage level λ (Ker and Coble, 2003) is obtained from the following equation:

$$\text{prem}_{i,T+1} = \frac{P\left(y_{i,T+1} < \lambda\hat{y}_{i,T+1}\right)\left(\lambda\hat{y}_{i,T+1} - E\left[\left(y_{i,T+1} \mid y_{i,T+1} < \lambda\hat{y}_{i,T+1}\right)\right]\right)}{\hat{y}_{i,T+1}}, \tag{11}$$

where $0 \le \lambda \le 1$ is the coverage level and $\hat{y}_{i,T+1}$ is the predicted yield for county i in year $T+1$. All measures in equation (11), such as predictions and expectations, are obtained from the BMA density function, f_i.

Similar to the current empirical rating scheme used by RMA, the BMA approach can only generate a density for a county with more yield observations than the minimum sample size needed to estimate the first-stage individual density estimation. Therefore, counties that fail to meet this threshold do not have density estimates under the BMA approach.

PHC Approach

Under the PHC approach, the premiums can also be calculated using equation (11). The predicted yield, the expectation, and all probability measures in the equation are calculated from the conditional predictive posterior distribution, $\text{P}(\boldsymbol{y}_{T+1} \mid \boldsymbol{Y})$, such that

$$P\left(\boldsymbol{y}_{T+1} \mid \boldsymbol{Y}\right) = \int_{\boldsymbol{\Theta}}\int_{\boldsymbol{\gamma}}\int_{\boldsymbol{\beta}} P_1\left(\boldsymbol{y}_{T+1} \mid \boldsymbol{\beta},\boldsymbol{\gamma},\boldsymbol{\Theta}\right) P_2(\boldsymbol{\beta},\boldsymbol{\gamma} \mid \boldsymbol{\Theta}) P(\boldsymbol{\beta},\boldsymbol{\gamma},\boldsymbol{\Theta} \mid \boldsymbol{Y}) d\boldsymbol{\beta} d\boldsymbol{\gamma} d\boldsymbol{\Theta}, \tag{12}$$

where $\boldsymbol{Y}$ is a matrix of the whole dataset that spans all counties $i = 1,\ldots,N$ and years $t = 1,\ldots,T$; $\boldsymbol{y}_{T+1}$ is a vector of predicted yield for year $T+1$; $P_1(\boldsymbol{y}_{T+1} \mid \boldsymbol{\beta},\boldsymbol{\gamma},\boldsymbol{\Theta})$ is the likelihood layer in equation (1); $P_2(\boldsymbol{\beta},\boldsymbol{\gamma} \mid \boldsymbol{\Theta})$ is the process layer in equation (3); and $P(\boldsymbol{\beta},\boldsymbol{\gamma},\boldsymbol{\Theta} \mid \boldsymbol{Y})$ is the posteriors in equation (5) estimated from the Markov chain Monte Carlo (MCMC) sampling.

We can interpret equation (11) as the value that divides expected indemnity by the expected yield. Both denominator and numerator values are calculated from the predicted posterior samples. We calculate values that divide the indemnity calculated from each posterior predictive sample by the expected yield and take an average of the values to get the premiums. We have 40,000 posterior samples, and thus we have 40,000 posterior predictive yield samples. Therefore, we have 40,000 calculated premium values, and the final premium is the average of these values.

Empirical Analysis

We use county-level corn and winter wheat yield data obtained from NASS for the empirical analysis. We select corn and winter wheat since they are US major crops. Corn yield has a clear

upward time trend, while winter wheat does not. Thus, our crop choice would help compare and contrast the model's performance for a crop with and without a trend. We do not include soybean since the reporting region and crop yield models for the soybean largely overlap with corn.

The data contain annual yields (bu/acre) for 1955–2017 for Iowa (99 counties) and Maryland (23 counties) and 1963–2017 for Colorado (53 counties) for corn and for 1955–2017 for Kansas (105 counties) and Indiana (92 counties) and for 1963–2017 for Colorado (53 counties) for winter wheat. We discard counties with no yield report at all from the dataset. We choose these states to test the performance under different characteristics (e.g., production level, climate or geographical features, and level of the data omission).

We include Iowa and Kansas because they are the largest US producers of corn and winter wheat, respectively. They also have relatively small levels of data omission due to their large production level. One difference between Iowa and Kansas is their geographical characteristics. Iowa's soil characteristics and climate are very uniform, whereas Kansas does not. Indiana's winter wheat is used since it represents a minor production area but has uniform geography. Also, it would highlight the characteristics of the winter wheat yield model in the Corn Belt region. Maryland is a minor corn production area, but it has a surprisingly low level of data omission. Additionally, Maryland has two different climates, highland west and coastal east. We use both corn and Colorado winter wheat of to see the performance of the PHC with different crops in the same region. Colorado also has significantly heterogeneous geographic characteristics from the eastern plains to the western mountains.

In summary, the areas chosen represent major (Iowa and Kansas) and marginal (Maryland, Indiana, and Colorado) production areas, homogeneous (Iowa and Indiana) and heterogeneous (Maryland, Kansas, and Colorado) geographical features, and small (Iowa, Kansas, and Maryland) and large (Colorado) portions of missing data.[8]

We use three benchmark approaches: HCKG, BMA-Normal, and BMA-Mixture. The BMA-Normal is the BMA approach under the assumption of normality, and the BMA-Mixture is the BMA approach with a bivariate-normal mixture distribution. We randomly drop 10%, 30%, and 60% of each state's yield history to see how the data omission affects each approach's crop density estimation. We then calculate 70% and 90% coverage level loss ratios for each data omission scenario. The following formula calculates the loss ratio:

$$\text{loss ratio}_i = \frac{\sum_{t=1}^{T} \max\left[\lambda \hat{y}_{it} - y_{it}, 0\right]}{\sum_{t=1}^{T} \text{prem}_{it}^{\lambda} \hat{y}_{it}}, \tag{13}$$

where λ is coverage level, $\hat{y}_{it}$ is predicted yield of county i at year t, $4y_{it}$ is the actual yield for county i at year t, and $\text{prem}_{it}^{\lambda}$ is the premium rate of λ coverage level for county i at year t. The loss ratio is interpreted as total indemnity divided by total premium. If the loss ratio is higher (lower) than 1, the premium is underpriced (overpriced). Therefore, the loss ratio should be 1 under actuarially fair premium rates.

We estimate crop yield densities using data up to 1997 and obtain predicted yields and premiums for 1998. We repeat the procedure from 1998 to 2017 and calculate premium gains and losses. Note that although the PHC can estimate densities for counties with no yield data reported, we include only 27 counties in Colorado for the loss ratio calculation since neither the benchmark HCKG nor BMA approach can provide accurate enough empirical premium rates for counties with a significant data omission level.[9]

[8] In the case of corn yield data, Iowa has only 8 missing observations out of 6,237 observations (0.1% omission) and Maryland has 44 missing observations out of 1,449 observations (3% omission). In contrast, Colorado has 1,361 missing observations out of 2,913 observations, which is approximately a 47% omission rate. In the winter wheat case, the number of missing and total observations are, respectively, 153 out of 6,613 for Kansas (2% omission), 605 out of 5,796 for Indiana (10% omission), and 548 out of 2,913 for Colorado (26% omission). We estimate the PHC approach for each state separately since the estimation gets exponentially slower as the number of locations increases.

[9] Counties with fewer than 20 yield observations are discarded from the out-of-sample rating game.

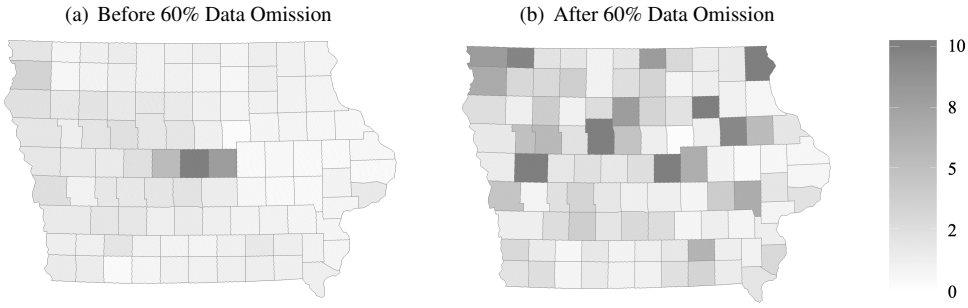

Figure 1. Estimated Loss Ratios from the BMA-Mixture Approach for Iowa Corn Before and After 60% Data Omission

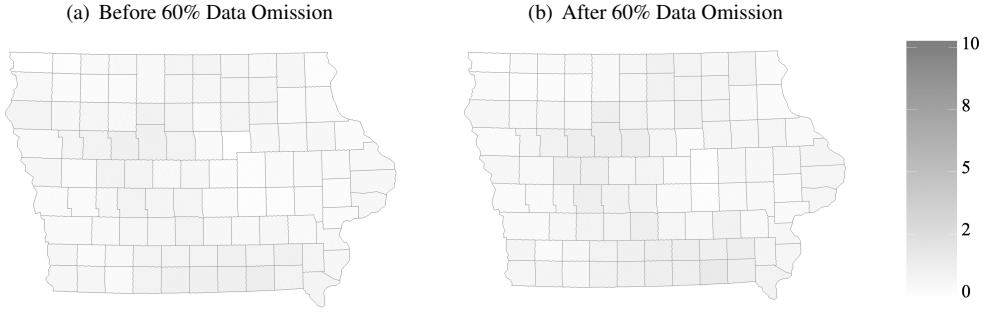

Figure 2. Estimated Loss Ratios from the PHC Approach for Iowa Corn Before and After 60% Data Omission

Tables 1–4 report median, maximum, minimum, and 1st and 3rd quantiles of county-level loss ratios for 20 years for each crop/state/coverage/omission level combination. These reported statistics describe the within-state accuracy and inequality in the loss ratios. The maximum and minimum statistics help to identify extreme cases of regional disparity in loss ratio calculation. In all cases, as indicated by the narrow interquantile intervals, the PHC has fewer within-state disparities than the HCKG and BMA approaches. As the data omission level increases and the coverage level moves deeper in the tail, the benchmark approaches' rating accuracy is dramatically diminished. They also do not offer reasonable premiums in Colorado, with a considerable level of missing data reported for corn and winter wheat. Loss ratio estimation results favor PHC over both HCKG and BMA approaches, and more so when there is a large portion of missing data.

Figures 1 and 2 illustrate the 90% coverage level loss ratios for Iowa corn before (panel A) and after (panel B) the 60% data omission for the BMA-Mixture and the PHC, respectively. Figures 3 and 4 are the loss ratios from the two approaches for Maryland corn. These figures illustrate that the PHC is less sensitive to data omission. As stated in Table 1, in Iowa and Maryland corn, the median loss ratios of the PHC's 90% coverage level increased only from 0.51 to 0.61 and from 0.93 to 1.05, respectively, even after 60% of data were omitted. In the case of the BMA-Mixture approach, however, the median loss ratios increased much more, from 1.10 to 4.29 and from 1.38 to 3.51, respectively, after the 60% of data omission. In detail, the BMA-Mixture's 90% coverage loss ratio for Allegany County in Maryland increased dramatically from 3.09 to 40.19, whereas the PHC's loss ratio changed only from 0.93 to 0.82.

Figure 5 illustrates Colorado corn's loss ratios from the original dataset for the BMA-Mixture and the PHC. In Colorado corn, the state with significant data omission, the median loss ratio of 90%

Figure 3. Estimated 90% Coverage Level Loss Ratios from the BMA-Mixture Approach for Maryland Corn Before and After 60% Data Omission

Figure 4. Estimated 90% Coverage Level Loss Ratios from the PHC Approach for Maryland Corn Before and After 60% Data Omission

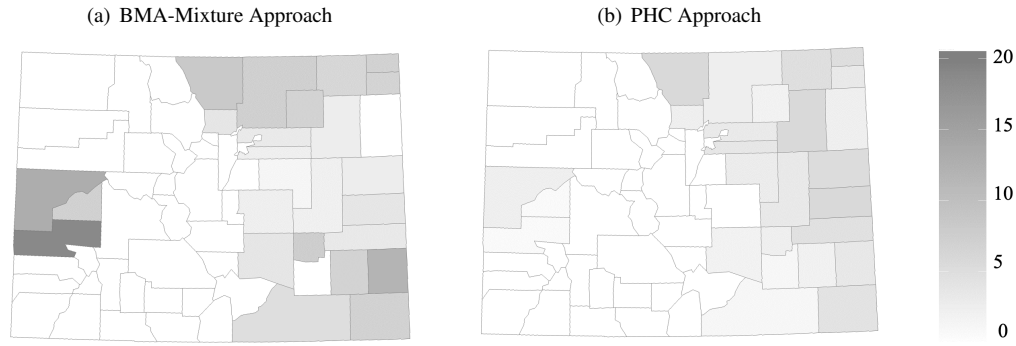

Figure 5. Estimated 90% Coverage Level Loss Ratios from the BMA-Mixture and PHC Approaches for Colorado Corn

coverage level of the PHC is 2.74, whereas that of the BMA-Mixture approach is 6.69. Notably, in some counties like Yuma County, the loss ratio estimation from the BMA-Mixture is 60.90, while the PHC is 2.28. Like the 90% coverage loss ratio for Iowa, the BMA-Normal is less sensitive to the data omission in Colorado. However, the BMA-Normal shows less accuracy in the 70% coverage loss ratio.

The data omission sensitivity of the BMA approach is attributed to its way of importing information. The BMA approach produces the density by importing information from counties

Table 1. 90% Coverage Level Out-of-Sample Loss Ratios of Corn from the HCKG, BMA, and PHC Approaches

	No Omission				10% Omission				30% Omission				60% Omission			
Model	HCKG	BMA-Normal	BMA-Mixture	PHC	HCKG	BMA-Normal	BMA-Mixture	PHC	HCKG	BMA-Normal	BMA-Mixture	PHC	HCKG	BMA-Normal	BMA-Mixture	PHC
Iowa																
Median	1.01	0.81	1.10	0.51	4.04	0.86	1.11	0.49	3.49	0.79	1.29	0.56	4.29	0.98	1.95	0.61
Min.	0.21	0.00	0.24	0.01	0.31	0.00	0.21	0.01	0.00	0.00	0.35	0.02	0.24	0.00	0.13	0.00
1st quantile	0.58	0.42	0.78	0.34	1.62	0.45	0.78	0.32	2.02	0.39	0.83	0.34	1.96	0.35	1.14	0.40
3rd quantile	1.24	1.17	1.66	0.76	10.26	1.18	1.63	0.78	7.53	1.23	2.16	0.78	14.51	1.76	3.28	0.90
Max.	6.58	2.15	11.85	1.27	222.39	2.63	13.71	1.31	245.09	8.39	12.39	1.35	228.73	14.72	54.32	1.54
Maryland																
Median	1.38	1.72	1.38	0.93	1.92	1.78	1.40	0.95	2.01	2.00	1.48	0.98	3.51	2.17	2.02	1.05
Min.	0.57	0.53	0.54	0.33	0.61	0.46	0.52	0.32	0.78	0.36	0.83	0.28	0.83	0.09	0.72	0.31
1st quantile	1.09	1.29	1.13	0.80	1.13	1.28	1.16	0.79	1.97	1.46	1.27	0.70	2.16	1.62	1.48	0.75
3rd quantile	1.86	2.09	1.60	1.59	1.95	2.33	1.70	1.59	5.13	2.51	1.93	1.47	8.49	2.96	2.75	1.73
Max.	4.06	3.75	3.09	2.46	4.00	4.25	3.28	2.46	12.04	6.13	6.18	2.32	20.67	7.86	293.91	2.69
Colorado																
Median	10.74	3.92	6.69	2.74	–	–	–	–	–	–	–	–	–	–	–	–
Min.	1.23	0.00	1.49	0.00	–	–	–	–	–	–	–	–	–	–	–	–
1st quantile	3.66	2.12	1.02	1.94	–	–	–	–	–	–	–	–	–	–	–	–
3rd quantile	4,420.21	5.78	7.50	4.39	–	–	–	–	–	–	–	–	–	–	–	–
Max.	34.13	9.73	60.90	5.92	–	–	–	–	–	–	–	–	–	–	–	–

Notes: Each omission level indicates that the empirical data from 1955 to 2017 are randomly omitted for each respective percentage to test the performance. Insurance buyers collect indemnity when the actual county yield is lower than the county production guarantee (coverage level × projected county yield). HCKG refers to the Harri–Coble–Ker–Goodwin approach (Harri et al., 2011). BMA-Normal refers to the Bayesian model averaging approach under the normality assumption (Ker, Tolhurst, and Liu, 2016). BMA-Mixture refers to the Bayesian model averaging approach under the normal mixture (Ker, Tolhurst, and Liu, 2016). PHC refers to the Park–Harri–Coble approach, the proposed Bayesian kriging approach presented in this paper.

Table 2. 70% Coverage Level Out-of-Sample Loss Ratios of Corn from the HCKG, BMA, and PHC Approaches

	No Omission				10% Omission				30% Omission				60% Omission			
Model	**HCKG**	**BMA-Normal**	**BMA-Mixture**	**PHC**	**HCKG**	**BMA-Normal**	**BMA-Mixture**	**PHC**	**HCKG**	**BMA-Normal**	**BMA-Mixture**	**PHC**	**HCKG**	**BMA-Normal**	**BMA-Mixture**	**PHC**
Iowa																
Median	3.57	44.04	0.00	0.00	0.00	41.40	4.01	0.00	0.00	42.98	6.10	0.00	0.00	32.57	21.37	0.00
Min.	0.42	0.00	0.00	0.00	0.00	0.00	0.35	0.00	0.00	0.00	0.71	0.00	0.00	0.00	0.38	0.00
1st quantile	2.27	16.20	0.00	0.00	0.00	16.67	1.92	0.00	0.00	15.74	2.55	0.00	0.00	11.47	7.34	0.00
3rd quantile	11.77	141.43	0.15	0.37	0.00	157.57	11.56	0.44	0.23	130.73	29.52	0.48	0.74	219.38	77.57	0.84
Max.	869.54	555.17	3.10	2.64	9.71	679.62	576.41	3.06	169.66	8,943.2	1,518.89	8.47	119.18	9,225.31	8,665.86	8.09
Maryland																
Median	1.79	20.13	7.64	0.89	1.91	20.39	8.51	0.74	1.87	25.50	11.89	0.69	1.63	32.58	36.19	0.67
Min.	0.00	8.15	2.49	0.00	0.00	8.16	1.95	0.00	0.00	5.01	2.74	0.00	0.00	1.29	2.35	0.00
1st quantile	0.63	16.98	5.62	0.29	0.93	15.73	6.57	0.26	0.99	17.71	10.06	0.32	0.00	20.84	12.99	0.16
3rd quantile	7.12	23.07	9.64	2.06	3.48	24.93	12.88	2.03	4.77	48.06	24.97	1.75	5.77	55.88	176.29	1.69
Max.	1,889.17	85.86	849.85	3.47	12.84	92.52	3,853.09	4.94	22.55	374.27	4,225.35	3.02	48.95	336.46	9,871.11	3.23
Colorado																
Median	3.77	341.86	543.26	2.03	–	–	–	–	–	–	–	–	–	–	–	–
Min.	0.00	0.00	1.96	0.00	–	–	–	–	–	–	–	–	–	–	–	–
1st quantile	0.47	42.81	28.82	0.00	–	–	–	–	–	–	–	–	–	–	–	–
3rd quantile	6.81	3,300.74	792.63	5.14	–	–	–	–	–	–	–	–	–	–	–	–
Max.	54.62	5,249.91	1,245.99	29.47	–	–	–	–	–	–	–	–	–	–	–	–

Notes: Each omission level indicatespercentage omission refers that the empirical data from 1955 to 2017 are randomly omitted for each respective percentage to test the performance. Insurance buyers collect indemnity when the actual county yield is lower than the county production guarantee (coverage level × projected county yield). HCKG refers to the Harri–Coble–Ker–Goodwin approach (Harri et al., 2011). BMA-Normal refers to the Bayesian model averaging approach under the normality assumption (Ker, Tolhurst, and Liu, 2016). BMA-Mixture refers to the Bayesian model averaging approach under the normal mixture (Ker, Tolhurst, and Liu, 2016). PHC refers to the Park–Harri–Coble approach, the proposed Bayesian kriging approach presented in this paper.

Table 3. 90% Coverage Level Out-of-Sample Loss Ratios of Winter Wheat from the HCKG, BMA, and PHC Approaches

	No Omission				10% Omission				30% Omission				60% Omission			
Model	HCKG	BMA-Normal	BMA-Mixture	PHC	HCKG	BMA-Normal	BMA-Mixture	PHC	HCKG	BMA-Normal	BMA-Mixture	PHC	HCKG	BMA-Normal	BMA-Mixture	PHC
Kansas																
Median	2.76	1.40	1.53	0.68	2.84	1.42	1.56	0.68	3.07	1.42	1.57	0.60	4.83	1.59	1.64	0.56
Min.	0.34	0.54	0.31	0.09	0.38	0.59	0.25	0.11	0.29	0.63	0.29	0.11	0.27	0.49	0.15	0.07
1st quantile	1.21	1.07	0.93	0.37	1.17	1.08	0.90	0.40	1.15	1.09	0.85	0.37	2.05	1.07	0.93	0.35
3rd quantile	4.72	2.07	2.15	1.04	4.50	2.07	2.18	1.03	5.56	2.02	2.17	0.89	8.73	2.25	2.34	0.78
Max.	10.09	3.60	3.64	2.02	10.55	3.81	4.14	1.99	12.68	3.72	4.50	1.77	2,103.67	3.35	4.10	1.65
Indiana																
Median	2.00	1.01	1.32	0.80	2.03	1.11	1.27	0.75	2.59	1.08	1.48	0.72	4.19	1.03	1.91	0.47
Min.	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
1st quantile	1.29	0.39	0.70	0.38	1.22	0.37	0.68	0.35	1.28	0.41	0.79	0.25	1.66	0.39	0.84	0.22
3rd quantile	3.37	1.54	1.93	1.25	3.69	1.57	2.06	1.22	5.20	1.70	2.27	1.02	8.87	1.77	3.12	0.75
Max.	17.47	25.63	9.90	7.86	41.23	20.29	14.39	8.41	416.37	16.65	12.80	7.34	122.65	15.93	16.22	5.67
Colorado																
Median	4.40	1.99	2.54	1.01	–	–	–	–	–	–	–	–	–	–	–	–
Min.	0.72	0.10	0.59	0.00	–	–	–	–	–	–	–	–	–	–	–	–
1st quantile	2.42	1.35	1.62	0.52	–	–	–	–	–	–	–	–	–	–	–	–
3rd quantile	12.12	2.51	4.06	2.20	–	–	–	–	–	–	–	–	–	–	–	–
Max.	34.02	8.10	27.86	5.96	–	–	–	–	–	–	–	–	–	–	–	–

Notes: Each omission level indicates that the empirical data from 1955 to 2017 are randomly omitted for each respective percentage to test the performance. Insurance buyers collect indemnity when the actual county yield is lower than the county production guarantee (coverage level × projected county yield). HCKG refers to the Harri–Coble–Ker–Goodwin approach (Harri et al., 2011). BMA-Normal refers to the Bayesian model averaging approach under the normality assumption (Ker, Tolhurst, and Liu, 2016). BMA-Mixture refers to the Bayesian model averaging approach under the normal mixture (Ker, Tolhurst, and Liu, 2016). PHC refers to the Park–Harri–Coble approach, the proposed Bayesian kriging approach presented in this paper.

Table 4. 70% Coverage level Out-of-Sample Loss Ratios of Winter-Wheat from the HCKG, BMA, and PHC Approaches

	No Omission				10% Omission				30% Omission				60% Omission			
Model	**HCKG**	**BMA-Normal**	**BMA-Mixture**	**PHC**	**HCKG**	**BMA-Normal**	**BMA-Mixture**	**PHC**	**HCKG**	**BMA-Normal**	**BMA-Mixture**	**PHC**	**HCKG**	**BMA-Normal**	**BMA-Mixture**	**PHC**
Kansas																
Median	6.84	9.01	4.94	0.47	6.84	8.67	5.11	0.58	8.65	8.59	5.37	0.43	2.74	9.05	7.80	0.44
Min.	0.00	2.85	0.65	0.00	0.00	2.57	0.46	0.00	0.00	2.94	0.40	0.00	0.00	2.79	0.33	0.00
1st quantile	2.15	7.18	2.69	0.17	1.79	6.70	3.05	0.26	1.46	6.93	3.03	0.20	0.16	6.34	3.46	0.19
3rd quantile	23.66	11.77	8.83	0.95	22.68	11.71	8.60	1.01	23.77	11.51	9.17	0.77	7.12	13.03	12.89	0.80
Max.	776.76	22.18	38.30	2.43	752.43	25.17	74.00	2.75	207.13	22.99	8,338.95	1.90	134.10	60.33	160.81	1.99
Indiana																
Median	0.00	71.96	18.71	0.00	0.00	73.25	22.61	0.00	0.00	69.48	28.89	0.00	0.00	62.69	54.01	0.00
Min.	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
1st quantile	0.00	24.44	6.05	0.00	0.00	21.17	7.18	0.00	0.00	21.06	10.19	0.00	0.00	12.71	9.58	0.00
3rd quantile	0.42	119.99	52.52	0.00	0.273	128.30	74.73	0.00	0.07	187.58	105.38	0.00	0.00	160.29	437.76	0.00
Max.	186.25	2,069.83	1,336.17	16.92	178.94	2,775.07	5,262.83	20.27	16.39	1,286.12	1,209.50	14.66	56.48	3,707.93	3,988.99	8.08
Colorado																
Median	6.31	12.82	23.73	0.84	–	–	–	–	–	–	–	–	–	–	–	–
Min.	0.00	4.67	2.69	0.00	–	–	–	–	–	–	–	–	–	–	–	–
1st quantile	1.11	7.29	9.59	0.10	–	–	–	–	–	–	–	–	–	–	–	–
3rd quantile	12.74	21.82	51.99	3.37	–	–	–	–	–	–	–	–	–	–	–	–
Max.	233.97	1,938,16	4,470.24	100.02	–	–	–	–	–	–	–	–	–	–	–	–

Notes: Each omission level indicates that the empirical data from 1955 to 2017 are randomly omitted for each respective percentage to test the performance. Insurance buyers collect indemnity when the actual county yield is lower than the county production guarantee (coverage level × projected county yield). HCKG refers to the Harri–Coble–Ker–Goodwin approach (Harri et al., 2011). BMA-Normal refers to the Bayesian model averaging approach under the normality assumption (Ker, Tolhurst, and Liu, 2016). BMA-Mixture refers to the Bayesian model averaging approach under the normal mixture (Ker, Tolhurst, and Liu, 2016). PHC refers to the Park–Harri–Coble approach, the proposed Bayesian kriging approach presented in this paper.

with similar historical outcomes. Hence, it provides an accurate density estimation when the dataset contains good references. However, if there are not enough outside references to import information to estimate the target county's density (i.e., counties with similar yield history), the BMA approach would not be accurate.[10] Therefore, it tends to provide relatively good estimates in states with homogeneous yield history due to similar soil and climate conditions such as Iowa corn but not in states with heterogeneous conditions.

On the other hand, the PHC estimates each county's density as a variation from an integrated yield density structure across space. The variation is determined by spatial structural parameters (i.e., kriging parameters) and Euclidean distances among counties. Therefore, the PHC can minimize the estimation distortion from the missing observations as long as the integrated yield density structure across space is well identified.

Premium Rating Game

In this section, we test the performance of the PHC in a more formal manner. To do that, we employ the out-of-sample premium rating game suggested by Ker and McGowan (2000). For the final comparison, we select two types of BMA approach (BMA-Normal and BMA-Mixture) as the primary benchmarks; several studies (Ker, Tolhurst, and Liu, 2016; Liu and Ker, 2020a) have proved that the BMA is statistically superior to the HCKG approach when the time dimension of the data is short. We use the loss ratios in the previous section, from 1998 to 2017, to conduct the premium rating game.

We first use the premium rating games to test across approaches (i.e., BMA vs. PHC). This is referred to as the *between* test. In an alternative test, the comparison is between the case where the complete original dataset is used, versus the case where an omission level is applied to the original data. This is referred to as the *within* test. For Iowa and Maryland corn as well as Kansas and Indiana winter wheat, states with no significant data omissions, we perform both the *between* and the *within* test. For Colorado, a state with a large portion of the data omission, we only perform the *between* test.

To avoid a possible advantage of the private company over the Federal Crop Insurance Corporation (FCIC) that can determine retain/cede decisions after observing premiums from the FCIC, we employ the relative loss ratio index suggested by Ker, Tolhurst, and Liu (2016). We utilize the index to perform both the *between* test(BMA vs. PHC) and the *within* test(original vs. omitted data), such that

$$RL^{between} = \frac{LR_C^P / LR_R^P}{LR_C^B / LR_R^B};$$

$$RL^{within} = \frac{LR_C^O / LR_R^O}{LR_C^M / LR_R^M}; \tag{14}$$

where LR_C^P and LR_R^P are the average loss ratios across ceded and retained policies, respectively, from the PHC; LR_C^B and LR_R^B are the ceded and retained loss ratio, respectively, from the BMA approach; LR_C^O and LR_R^O are the ceded and retained loss ratio, respectively, using the original dataset; and LR_C^M and LR_R^M are the ceded and retained loss ratio, respectively, using the dataset with missing observations.

[10] As noted previously, this article only considers a single state estimation. Therefore, the BMA approach might not have valuable references to recover the target county's density in a single state dataset. If we extend the dataset to a sufficiently larger locational boundary beyond the state level, results for the BMA approach might improve. However, for states like Colorado and Montana, where there is a significant level of missing data in the entire state, it might be hard to get valuable references even when extending the boundary beyond the state level.

Table 5. Out-of-Sample Rating Game Results of the *between* Test

Crop/State	No. of Counties	Data Omission	BMA-Normal vs. PHC 90% Coverage (*p*-Value)	BMA-Normal vs. PHC 70% Coverage (*p*-Value)	BMA-Mixture vs. PHC 90% Coverage (*p*-Value)	BMA-Mixture vs. PHC 70% Coverage (*p*-Value)
Corn						
Iowa	99	No omission	0.132	$2 \times 10^{-5***}$	0.002^{***}	$2 \times 10^{-6***}$
		10%	0.227	$2 \times 10^{-6***}$	0.132	$2 \times 10^{-5***}$
		30%	0.132	$9 \times 10^{-7***}$	0.010^{***}	$9 \times 10^{-7***}$
		60%	$2 \times 10^{-5***}$	$9 \times 10^{-7***}$	$2 \times 10^{-5***}$	$9 \times 10^{-7***}$
Maryland	23	No omission	0.001^{**}	$2 \times 10^{-4***}$	0.021^{**}	$2 \times 10^{-4***}$
		10%	0.021^{**}	$2 \times 10^{-4***}$	0.001^{***}	$2 \times 10^{-5***}$
		30%	0.000^{***}	$2 \times 10^{-5***}$	$2 \times 10^{-4***}$	$2 \times 10^{-5***}$
		60%	0.000^{***}	$9 \times 10^{-7***}$	$2 \times 10^{-4***}$	$9 \times 10^{-7***}$
Colorado	27	No omission	0.084^{*}	$3 \times 10^{-5***}$	0.090^{*}	$9 \times 10^{-4***}$
Wheat						
Kansas	105	No omission	0.001^{***}	$2 \times 10^{-4***}$	0.001^{***}	$2 \times 10^{-5***}$
		10%	$2 \times 10^{-4***}$	$2 \times 10^{-5***}$	$2 \times 10^{-5***}$	$9 \times 10^{-7***}$
		30%	$2 \times 10^{-4***}$	$2 \times 10^{-5***}$	$2 \times 10^{-5***}$	$9 \times 10^{-7***}$
		60%	$2 \times 10^{-5***}$	$9 \times 10^{-7***}$	$2 \times 10^{-5***}$	$9 \times 10^{-7***}$
Indiana	92	No omission	0.058^{*}	$2 \times 10^{-4***}$	0.010^{**}	$2 \times 10^{-5***}$
		10%	0.010^{**}	$2 \times 10^{-5***}$	0.015^{**}	$2 \times 10^{-5***}$
		30%	$2 \times 10^{-4***}$	$2 \times 10^{-5***}$	$2 \times 10^{-5***}$	$9 \times 10^{-7***}$
		60%	$2 \times 10^{-4***}$	$2 \times 10^{-5***}$	$2 \times 10^{-5***}$	$2 \times 10^{-5***}$
Colorado	27	No omission	0.132	$2 \times 10^{-5***}$	0.015^{**}	$2 \times 10^{-4***}$

Notes: A lower p-value indicates that the PHC approach dominates the BMA approach and vice versa. Single, double, and triple asterisks (*, **, ***) indicate the statistical significance of lower private (government) loss ratios at the 10%, 5%, and 1% level, respectively. Each omission level indicates that the empirical data from 1955 to 2017 are randomly omitted for each respective percentage to test the performance. Insurance buyers collect indemnity when the actual county yield is lower than the county production guarantee (coverage level × projected county yield). BMA-Normal refers to the Bayesian model averaging approach under the normality assumption (Ker, Tolhurst, and Liu, 2016). BMA-Mixture refers to the Bayesian model averaging approach under the normal mixture (Ker, Tolhurst, and Liu, 2016). PHC refers to the Park–Harri–Coble approach, the proposed Bayesian kriging approach presented in this paper.

First, under the setting of the relative loss ratio for the *between* test, $RL^{between}$, two players (a private company and the FCIC) use different premium rating approaches; $RL^{between}$ identifies the accuracy of using the PHC relative to the BMA approach. That is, a higher $RL^{between}$ indicates that the PHC approach is more accurate than the BMA approach. In terms of the relative loss ratio for the *within* test (original and omitted data), RL^{within} identifies the accuracy of using the original dataset relative to using the dataset with missing observations. Therefore, an approach with higher RL^{within} indicates higher levels of accuracy loss in the presence of missing data.

We first discuss the results for the *between* test. Under the null hypothesis that both approaches estimate the yield density equally well, yield predictions and premiums calculated from the approaches must be equally accurate. Therefore, the ceded-to-retained loss ratios from the two approaches must be identical (i.e., under the null, the median of the distribution for $RL^{between}$ is 1). We obtain $RL^{between}$ across counties for each year from 1998 to 2017 for each state in the presence of the missing data. Like Ker, Tolhurst, and Liu (2016), we define the random variable RL^{*}, indicating the number of $RL^{between}$ greater than 1 within the 20-year period, where $RL^{*} \sim \text{Binomial}(0.5,20)$.[11]

Table 5 presents the results of the *between* test. The *p*-values are obtained from the binomial distribution of RL^{*}. A lower *p*-value indicates that the PHC dominates the benchmark approaches. The PHC is preferred to both BMA approaches in all combinations of crop, state, coverage level, and omission level and is statistically superior in 67 of the 72 combinations. Notably, the PHC is preferred to both BMA approaches in a significant data omission (60% omission) and deeper coverage levels (i.e., 70% coverage level). For instance, the PHC rejects the null in all the cases of 70% coverage and 60% omission. The BMA-Normal approach works comparatively better than the BMA-Mixture in Iowa corn and Colorado winter wheat cases for the 90% coverage level, but the PHC significantly dominates all other combinations.

As a complementary comparison of accuracy, we examine how missing data affects the accuracy of premium rating via the *within* test. If an approach imported valuable information from other counties and produced sufficiently accurate densities of counties with missing data, there would be no notable changes in loss ratios after omitting data. Like the index for the *between* test, the comparison index RL^{within} is distributed with a median of 1 under the null hypothesis, meaning that the ceded-to-retained loss ratios calculated from the two datasets are identical. We repeat the same process as for the *between* test for both the BMA approaches and the PHC. We then estimate the random variable RL^{*} for 20 years, where $RL^{*} \sim \text{Binomial}(0.5,20)$.

Table 6 reports the *within* test results. Unlike the *between* test, the *within* test results should not be statistically significant if the approaches adequately estimate densities via omitted datasets. A lower *p*-value indicates that results from the original dataset outperform the results from the missing data. As expected, in both approaches, yield predictions and premium calculations using the original dataset provide better estimations (p-value < 0.5) in most cases. Neither BMA approach rejects the null hypothesis in the 90% coverage/10% data omission combination. However, as the data omission level increases, the number of rejections increases. For instance, the two BMA approaches reject the null for 8 out of 16 combinations at the 30% data omission level and reject the null in all 60% data omission cases.

In contrast, the accuracy of the PHC does not significantly decline with data omission. The PHC does not reject the null hypothesis in all combinations of crop, state, and coverage level at the 10% data omission level and rejects the null in 1 out of 12 combinations at the 30% data omission level and 2 out of 12 combinations at the 60% data omission level. The results demonstrate that the PHC is remarkably accurate even with a large portion of data omitted.

We further measure the approximate economic gains (losses) of using the PHC compared to the HCKG and the BMA approaches when there is a data omission issue. The economic gains and losses are based on the standpoint of a policy maker. That is, if a county's loss ratio is higher than 1, we

[11] A year that does not have indemnity payments (i.e., no counties with actual yield outcomes below the trigger level) across all counties is discarded since the *RL* indices in equation (13) cannot be calculated. Therefore, if there are k years of no indemnity payments, then the random variable follows $RL^{*} \sim \text{Binomial}(0.5, 20 - k)$.

Table 6. Out-of-Sample Rating Game Results for the *within* Test (no omission vs. dataset with omission)

State	Model	Data Omission	90% Coverage (p-Value)	70% Coverage (p-Value)
Corn				
Iowa	BMA-Normal	No vs. 10%	0.696	0.059*
		No vs. 30%	0.006***	0.038**
		No vs. 60%	0.004***	0.059*
	BMA-Mixture	No vs. 10%	0.500	0.059*
		No vs. 30%	0.084*	0.084*
		No vs. 60%	2×10^{-5}***	2×10^{-6}***
	PHC	No vs. 10%	0.212	0.313
		No vs. 30%	0.227	0.345
		No vs. 60%	0.313	0.524
Maryland	BMA-Normal	No vs. 10%	0.133	0.040**
		No vs. 30%	2×10^{-5}***	0.010**
		No vs. 60%	2×10^{-5}***	0.010**
	BMA-Mixture	No vs. 10%	0.119	0.407
		No vs. 30%	0.240	0.119
		No vs. 60%	0.004***	0.015**
	PHC	No vs. 10%	0.377	0.345
		No vs. 30%	0.805	0.254
		No vs. 60%	0.377	0.500
Wheat				
Kansas	BMA-Normal	No vs. 10%	0.407	2×10^{-5}***
		No vs. 30%	0.118	0.015**
		No vs. 60%	0.002***	2×10^{-5}***
	BMA-Mixture	No vs. 10%	0.500	0.676
		No vs. 30%	0.032**	0.178
		No vs. 60%	0.021**	0.058*
	PHC	No vs. 10%	0.274	0.773
		No vs. 30%	0.928	0.605
		No vs. 60%	0.015**	0.304
Indiana	BMA-Normal	No vs. 10%	0.227	0.105
		No vs. 30%	0.105	0.227
		No vs. 60%	0.095*	0.002***
	BMA-Mixture	No vs. 10%	0.315	0.166
		No vs. 30%	0.500	0.166
		No vs. 60%	0.001***	0.006***
	PHC	No vs. 10%	0.252	0.820
		No vs. 30%	0.171	0.891
		No vs. 60%	0.748	0.748

Notes: A lower p-value indicates that the PHC approach dominates the BMA approach and vice versa. Single, double, and triple asterisks (*, **, ***) indicate the statistical significance of lower private (government) loss ratios at the 10%, 5%, and 1% level, respectively. Each omission level indicates that the empirical data from 1955 to 2017 are randomly omitted for each respective percentage to test the performance. Insurance buyers collect indemnity when the actual county yield is lower than the county production guarantee (coverage level × projected county yield). BMA-Normal refers to the Bayesian model averaging approach under the normality assumption (Ker, Tolhurst, and Liu, 2016). BMA-Mixture refers to the Bayesian model averaging approach under the normal mixture (Ker, Tolhurst, and Liu, 2016). PHC refers to the Park–Harri–Coble approach, the proposed Bayesian kriging approach presented in this paper.

Table 7. Total Economic Gains (per year) of Using the PHC Approach

Model	Program	No Omission	10% Omission	30% Omission	60% Omission
PHC vs. HCKG	ARP	$19,108,725	$23,197,557	$73,850,230	$88,791,356
	AYP	$664,138	$870,289	$1,213,021	$924,206
	GRP	$3,977,403	$5,211,161	$6,231,629	$5,533,767
	Total	$23,750,266	$29,279,007	$81.294,880	$95,249,329
PHC vs. BMA-Mixture	ARP	$14,858,421	$17,847,984	$21,108,487	$61,234,994
	AYP	$128,410	$118,255	$115,833	$277,964
	GRP	$770,361	$709,775	$694,688	$1,665,448
	Total	$15,757,192	$18,676,014	$21,919,008	$63,178,406

Notes: ARP stands for Area Revenue Protection, AYP stands for Area Yield Protection, and GRP stands for Group Risk Plan. Each omission level indicates that the empirical data from 1955 to 2017 are randomly omitted for each respective percentage to test the performance. HCKG refers to the Harri–Coble–Ker–Goodwin approach (Harri et al., 2011). BMA-Mixture refers to the Bayesian model averaging approach under the normal mixture (Ker, Tolhurst, and Liu, 2016). PHC refers to the Park–Harri–Coble approach, the proposed Bayesian kriging approach presented in this paper.

consider the loss ratio deviation from 1 to be an economic loss of overpaying indemnity to insured in the county. In contrast, a county with a loss ratio smaller than 1 is considered to be an economic loss due to overrated premiums and, thus, a loss in higher subsidy payments. We use the county-level loss ratios calculated from each approach in the previous section and obtain monetary values of the economic significance of the PHC relative to other approaches by using the RMA Summary of Business database from 2005 to 2015.

We use county-level indemnity payments, premiums, subsidies, and insured acreage of the Area Revenue Protection (ARP), the Area Yield Protection (AYP), and the Group Risk Plan (GRP) from the RMA data to calculate the economic gains and losses. Since the dataset merely covers the federal insurance programs for corn, we only measure corn insurance products' economic gains and losses. We also choose to use the 90% coverage level premiums because roughly 95% of the current policies sold are at this coverage level.

In order to calculate the overpaid indemnities and subsidies from the underpriced and overpriced premium calculations, we measure the average indemnity/subsidy overpayments per unit loss ratio difference from one due to the inaccuracy of the premium rating procedure.[12] We then compare the PHC's economic gains/losses with two other benchmark approaches for different data omission levels. Online Supplement C presents descriptive statistics of the RMA data, the estimated overpayment measures, and the estimated per acre economic gains/losses.

Finally, we calculate the total economic gains by using the information of the total insured acreage and the estimated (per acre) overpayment measures. Table 7 shows the total economic gains per year of using the PHC compared to the HCKG and BMA-Mixture approaches. For the 10% data omission scenario, using the PHC results in total economics gains of $29,279,007/year and $18,676,014/year compared to the HCKG and BMA-Mixture approaches respectively. For the 30% data omission scenario, using the PHC results in the total economic gains of $81,294,880/year and $21,919,008/year compared to the HCKG and BMA-Mixture approaches, respectively. For the 60% data omission scenario, total economic gains are are $95,249,329/year and $63,178,406/year compared to the HCKG and BMA-Mixture approaches, respectively.

[12] First, we sort counties with underpriced (loss ratios > 1) and overpriced (loss ratios < 1) premiums by using the actual county-level loss ratios obtained from the RMA Summary of Business database. Suppose an average indemnity payment per acre and an average loss ratio of the counties with underpriced premiums (loss ratios > 1) are $100 and 1.2, respectively. Then the deviation of the loss ratio from 1 is 0.2. In that case, the indemnity overpayment per unit is $500, which can be calculated by dividing $100 by 0.2. Therefore, we consider that one unit of the overestimation of the loss ratio results in a $500 overpayment of the indemnity per acre compared to the case with actuarially fair case (loss ratio =1), on average. Similarly, we can measure a unit subsidy payment loss due to overpriced premiums by using the subsidy payments dataset of counties with overpriced premiums (loss ratios < 1).

However, one must carefully interpret the economic significance in the study since the superiority in the premium rating of the PHC approach is only based on the empirical dataset used here. Therefore, if we extend the scope of the empirical dataset, the value might be different. Also, RMA does not offer area-based insurance products in counties with insufficient yield history. Therefore, our data omission scenario might overestimate the economic gains in some cases. The fact remains that there are significant economic gains in offering accurate premiums via the PHC approach. Thus, we have demonstrated that the PHC would be significantly beneficial in some states/crops, particularly where data provision is a concern.

Limitations and Further Research

This study investigates the problem of missing yield data and proposes a new approach to estimate county-level densities in the presence of a significant level of missing data. We find that missing data can result in a severe problem in the federal crop insurance premium rating. Our results show that the proposed approach provides more accurate estimation and is less sensitive to the data omission problem than the benchmark premium rating approaches in estimating county-level crop yield densities.

The study has some limitations. The primary distributional assumption in the study is based on normality. The Gaussian distribution has some advantages over other types of distribution (e.g., having computational brevity and thus fast convergence) but the distribution cannot accommodate higher moments adjustment, such as skewness and kurtosis of crop yield densities. Still, even with the normality assumption, our approach is computationally intensive. For example, the approach takes approximately 36 hours to obtain 50,000 posteriors for Iowa with 99 counties via *Intel® Xeon® W-2133 Processor (8.25M Cache, 3.60 GHz)*.

In this context, a possible extension of this research is to use a more efficient MCMC algorithm, such as the Hamiltonian Monte Carlo (HMC) algorithm, to reduce the computational burden. Updating missing observations under the current MCMC algorithm gets exponentially slower when the dataset includes many locations. Therefore, we considered a single state estimation in the study, but we may extend the scope to a broader area by resolving this computational limitation. The HMC algorithm allows one to easily apply non-Gaussian distributions to the PHC, especially with a high level of data omissions.

One valuable future research question is investigating how to calculate counties' insurance payouts with no data reported. One possible option is to use alternative data sources, such as the RMA data, when NASS data are not available. This raises another question about how much accuracy will be achieved when integrating two primary data sources in insurance premium rating to reduce missing data problems.

[First submitted March 2021; accepted for publication August 2021.]

References

Annan, F., J. Tack, A. Harri, and K. Coble. "Spatial Pattern of Yield Distributions: Implications for Crop Insurance." *American Journal of Agricultural Economics* 96(2014):253–268. doi: 10.1093/ajae/aat085.

Banerjee, S., B. P. Carlin, and A. E. Gelfand. *Hierarchical Modeling and Analysis for Spatial Data*. Boca Raton, FL: CRC Press, 2004.

Coble, K. H., R. O. Knight, R. D. Pope, and J. R. Williams. "Modeling Farm-Level Crop Insurance Demand with Panel Data." *American Journal of Agricultural Economics* 78(1996):439–447. doi: 10.2307/1243715.

Cooley, D., P. Naveau, and P. Poncet. "Variograms for Spatial Max-Stable Random Fields." In P. Bertail, P. Soulier, and P. Doukhan, eds., *Dependence in Probability and Statistics,* No. 187 in Lecture Notes in Statistics. New York, NY: Springer, 2006, 373–390.

Cooley, D., D. Nychka, and P. Naveau. "Bayesian Spatial Modeling of Extreme Precipitation Return Levels." *Journal of the American Statistical Association* 102(2007):824–840. doi: 10.1198/016214506000000780.

Du, X., C. L. Yu, D. A. Hennessy, and R. Miao. "Geography of Crop Yield Skewness." *Agricultural Economics* 46(2015):463–473. doi: 10.1111/agec.12174.

Harri, A., K. H. Coble, A. P. Ker, and B. J. Goodwin. "Relaxing Heteroscedasticity Assumptions in Area-Yield Crop Insurance Rating." *American Journal of Agricultural Economics* 93(2011): 707–717. doi: 10.1093/ajae/aar009.

Jensen, N. D., C. B. Barrett, and A. G. Mude. "Index Insurance Quality and Basis Risk: Evidence from Northern Kenya." *American Journal of Agricultural Economics* 98(2016):1450–1469. doi: 10.1093/ajae/aaw046.

Ker, A. P., and K. Coble. "Modeling Conditional Yield Densities." *American Journal of Agricultural Economics* 85(2003):291–304. doi: 10.1111/1467-8276.00120.

Ker, A. P., and P. McGowan. "Weather Based Adverse Selection: The Private Insurance Company Perspective." *Journal of Agricultural and Resource Economics* 25(2000):386–410. doi: 10.22004/ag.econ.30907.

Ker, A. P., and T. N. Tolhurst. "On the Treatment of Heteroscedasticity in Crop Yield Data." *American Journal of Agricultural Economics* 101(2019):1247–1261. doi: 10.1093/ajae/aaz004.

Ker, A. P., T. N. Tolhurst, and Y. Liu. "Bayesian Estimation of Possibly Similar Yield Densities: Implications for Rating Crop Insurance Contracts." *American Journal of Agricultural Economics* 98(2016):360–382. doi: 10.1093/ajae/aav065.

Li, Q., and J. Racine. "Nonparametric Estimation of Distributions with Categorical and Continuous Data." *Journal of Multivariate Analysis* 86(2003):266–292. doi: 10.1016/S0047-259X(02)00025-8.

Li, X., Z. Shen, A. Harri, and K. H. Coble. "Comparing Survey-Based and Programme-Based Yield Data: Implications for the U.S. Agricultural Risk Coverage-County Programme." *Geneva Papers on Risk and Insurance* 45(2020):184–202. doi: 10.1057/s41288-019-00148-4.

Liu, Y., and A. P. Ker. "Rating Crop Insurance Contracts with Nonparametric Bayesian Model Averaging." *Journal of Agricultural and Resource Economics* 45(2020a):244–264. doi: 10.22004/ag.econ.302453.

———. "When Less Is More: On the Use of Historical Yield Data with Application to Rating Area Crop Insurance Contracts." *Journal of Agricultural and Applied Economics* 52(2020b):194–203. doi: 10.1017/aae.2019.40.

Miranda, M. J. "Area-Yield Crop Insurance Reconsidered." *American Journal of Agricultural Economics* 73(1991):233–242. doi: 10.2307/1242708.

Ozaki, V. A., S. K. Ghosh, B. K. Goodwin, and R. Shirota. "Spatio-Temporal Modeling of Agricultural Yield Data with an Application to Pricing Crop Insurance Contracts." *American Journal of Agricultural Economics* 90(2008):951–961. doi: 10.1111/j.1467-8276.2008.01153.x.

Park, E., B. W. Brorsen, and A. Harri. "Using Bayesian Kriging for Spatial Smoothing in Crop Insurance Rating." *American Journal of Agricultural Economics* 101(2019):330–351. doi: 10.1093/ajae/aay045.

———. "Spatially Smoothed Crop Yield Density Estimation: Physical Distance versus Climate Similarity." *Journal of Agricultural and Resource Economics* 45(2020):533–548. doi: 10.22004/ag.econ.302461.

Racine, J. S., and A. P. Ker. "Rating Crop Insurance Policies with Efficient Nonparametric Estimators that Admit Mixed Data Types." *Journal of Agricultural and Resource Economics* 31(2006):1–13. doi: 10.22004/ag.econ.10146.

Ramsey, A. F., and B. K. Goodwin. "Value-at-Risk and Models of Dependence in the U.S. Federal Crop Insurance Program." *Journal of Risk and Financial Management* 12(2019):65. doi: 10.3390/jrfm12020065.

Ridolfo, H., J. Boone, and N. Dickey. "Will They Answer the Phone if They Know It's Us? Using Caller ID to Improve Response Rates." RDD Research Report Number 234302, US Department of Agriculture, National Agricultural Statistics Service, 2013.

Woodard, J. D. "Determining Optimal Data Aggregation: An Application of Out-of-Sample Mixture Models." 2016. Paper presented at the annual meeting of the Agricultural and Applied Economics Association, Boston, Massachusetts, July 31–Aug 2.

Zhang, Y. Y. "A Density-Ratio Model of Crop Yield Distributions." *American Journal of Agricultural Economics* 99(2017):1327–1343. doi: 10.1093/ajae/aax021.

Journal of Agricultural and Resource Economics 47(3):S1–S10 ISSN: 1068-5502 (Print); 2327-8285 (Online)

doi: 10.22004/ag.econ.313319

Online Supplement: Estimating Crop Yield Densities for Counties with Missing Data

Eunchun Park, Ardian Harri, and Keith H. Coble

Supplement A: Bayesian Interpolation Model under the Beta Distribution.

The Hierarchical Structure of the Proposed Approach with Beta Distribution

The estimation of the proposed approach is based on the Bayesian hierarchical framework. The Bayesian hierarchical structure incorporates three layers: the likelihood layer, the process layer, and the prior layer. However, a different distributional assumption on the crop yield changes forms of each hierarchy.

Likelihood Layer

The likelihood layer forms the crop yield density. Let y_{it} be the crop yield of county i at year t, where $i = 1, \dots, N$ and $t = 1, \dots, T$. We now assume that the crop yields follow Beta distribution, instead of Gaussian, the likelihood layer is formed as,

$$P_1(\boldsymbol{Y}|\,\boldsymbol{\beta},\boldsymbol{\gamma},\boldsymbol{\Theta}) \;=\; \prod_{t=1}^{T} \frac{y_t^{\boldsymbol{\beta}\boldsymbol{x}_t-1}(\boldsymbol{y}^M-\boldsymbol{y}_t)^{\boldsymbol{\gamma}\boldsymbol{x}_t-1}}{B(\boldsymbol{\beta}\boldsymbol{x}_t,\boldsymbol{\gamma}\boldsymbol{x}_t)(\boldsymbol{y}^M)^{\boldsymbol{\beta}\boldsymbol{x}_t+\boldsymbol{\gamma}\boldsymbol{x}_t-1}} \tag{A1}$$

where $\boldsymbol{y}_t$ is a vector of crop yield at year t for all counties, $\boldsymbol{y}_t = [y_{1t}, \dots, y_{Nt}]'$, and thus $\boldsymbol{Y} = [\boldsymbol{y}_1, \dots, \boldsymbol{y}_T]$, $\boldsymbol{y}^M$ is the vector of yield ceiling, $\boldsymbol{y}^M = [y^M, \dots, y^M]'$, which is 20% greater than the highest historical yield following Norwood, Roberts, and Lusk (2004), $\boldsymbol{x}_t$ is a 2×1 vector of explanatory variables at year t that includes intercept and a linear trend variable, $\boldsymbol{\beta} = [\boldsymbol{\beta}_1, \boldsymbol{\beta}_2]$ is a $N \times 2$ vector of the mean equation coefficients, where $\boldsymbol{\beta}_1 = [\beta_{1i}, \dots, \beta_{1N}]'$ and $\boldsymbol{\beta}_2 = [\beta_{2i}, \dots, \beta_{2N}]'$, $\boldsymbol{\gamma} = [\boldsymbol{\gamma}_1, \boldsymbol{\gamma}_2]$ is a $N \times 2$ vector of the standard deviation equation coefficients, $\boldsymbol{\gamma}_1 = [\gamma_{1i}, \dots, \gamma_{1N}]'$ and $\boldsymbol{\gamma}_2 = [\gamma_{2i}, \dots, \gamma_{2N}]'$, $B(\cdot)$ is the Beta function, and $\boldsymbol{\Theta}$ is a vector of hyper parameters, $\boldsymbol{\Theta} = \left[\boldsymbol{\delta}_1, \boldsymbol{\delta}_2, \boldsymbol{\vartheta}_1, \boldsymbol{\vartheta}_2, \theta_{\beta 1}, \theta_{\beta 2}\theta_{\gamma 1}, \theta_{\gamma 2}, \rho_{\beta 1}, \rho_{\beta 2}, \rho_{\gamma 1}, \rho_{\gamma 2}\right]'$.

Process Layer

The process layer forms the two shape parameters of the Beta distribution defined in the likelihood layer. In the layer, a spatial interpolation for the coefficients in the two shape parameters equations via a Gaussian spatial process.

The county-specific coefficients are obtained by assuming the following multivariate Gaussian spatial process such that,

---

The material contained herein is supplementary to the article named in the title and published in the *Journal of Agricultural and Resource Economics* (*JARE*).

$$\begin{aligned} \boldsymbol{\beta}_k \mid \boldsymbol{\delta}_k, \theta_{\beta k}, \rho_{\beta k} &\sim MVGP\left(\boldsymbol{\delta}_k, \Sigma_{\beta k}\right) \\ \boldsymbol{\gamma}_k \mid \boldsymbol{\vartheta}_k, \theta_{\gamma k}, \rho_{\gamma k} &\sim MVGP\left(\boldsymbol{\vartheta}_k, \Sigma_{\gamma k}\right) \\ &\text{where } k = 1, 2 \end{aligned} \tag{A2}$$

where $\boldsymbol{\beta}_k$ and $\boldsymbol{\gamma}_k$ are vectors of the intercepts and trend parameters defined in the previous layer, $\boldsymbol{\delta}_k$ and $\boldsymbol{\vartheta}_k$ are vectors of the deterministic part of each coefficients that are uniform across all locations where $\boldsymbol{\delta}_k = [\delta_k, \dots, \delta_k]'$ and $\boldsymbol{\vartheta}_k = [\vartheta_k, \dots, \vartheta_k]'$, $\Sigma_{\beta k}$ and $\Sigma_{\gamma k}$ are corresponding spatial covariance matrices for each coefficient, which are $N \times N$ matrices structured by a function of standardized Euclidean distances (D_{ij}) between counties i and j calculated from longitude/latitude coordinates.

We now can define the second layer of the hierarchy such that,

$$\begin{aligned} P_2(\boldsymbol{\beta}, \boldsymbol{\gamma} | \boldsymbol{\Theta}) = & \frac{1}{\sqrt{(2\pi)^N \left|\Sigma_{\beta 1}\right|}} \exp\left[-\frac{1}{2}(\boldsymbol{\beta}_1 - \boldsymbol{\delta}_1)' \Sigma_{\beta 1}^{-1} (\boldsymbol{\beta}_1 - \boldsymbol{\delta}_1)\right] \\ & \times \frac{1}{\sqrt{(2\pi)^N \left|\Sigma_{\beta 2}\right|}} \exp\left[-\frac{1}{2}(\boldsymbol{\beta}_2 - \boldsymbol{\delta}_2)' \Sigma_{\beta 2}^{-1} (\boldsymbol{\beta}_2 - \boldsymbol{\delta}_2)\right] \\ & \times \frac{1}{\sqrt{(2\pi)^N \left|\Sigma_{\gamma 1}\right|}} \exp\left[-\frac{1}{2}(\boldsymbol{\gamma}_1 - \boldsymbol{\vartheta}_1)' \Sigma_{\gamma 1}^{-1} (\boldsymbol{\gamma}_1 - \boldsymbol{\vartheta}_1)\right] \\ & \times \frac{1}{\sqrt{(2\pi)^N \left|\Sigma_{\gamma 2}\right|}} \exp\left[-\frac{1}{2}(\boldsymbol{\gamma}_2 - \boldsymbol{\vartheta}_2)' \Sigma_{\gamma 2}^{-1} (\boldsymbol{\gamma}_2 - \boldsymbol{\vartheta}_2)\right]. \end{aligned} \tag{A3}$$

Prior Layer

Unlike Gaussian distribution, Beta distribution parameters are bounded at zero, $\alpha, \delta > 0$. Also, due to the non-linearity of the Beta function $B(\cdot)$, informative priors are required to achieve convergence. Therefore, we first fit Beta distribution via Maximum Likelihood Estimation (MLE) and use the estimated parameter information to impose informative priors. We use multivariate log-normal priors, Multi-lognormal ($\mathbf{0}$, $\boldsymbol{I}$), for the shape parameters coefficients ($\boldsymbol{\delta}_1, \boldsymbol{\delta}_2, \boldsymbol{\vartheta}_1, \boldsymbol{\vartheta}_2$). Identical to the Gaussian model in the manuscript, we impose general non-informative inverse gamma priors, $IG(0.1, 0.1)$ for the sill parameters ($\rho_{\beta 1}, \rho_{\beta 2}, \rho_{\gamma 1}, \rho_{\gamma 2}$) and informative priors for the range parameters ($\theta_{\beta 1}, \theta_{\beta 2} \theta_{\gamma 1}, \theta_{\gamma 2}$) , $\text{Uniform}\left(0,\ 2 * \max\left(D_{ij}\right)\right)$, where $\max\left(D_{ij}\right)$ is maximum distance among all D_{ij}.

Then the final prior layer can be structured as

$$P_3(\boldsymbol{\Theta}) = p(\boldsymbol{\beta}_1) p(\boldsymbol{\beta}_2)\, p(\boldsymbol{\gamma}_1) \dots p\left(\rho_{\gamma 1}\right) p\left(\rho_{\gamma 2}\right). \tag{A4}$$

Finally, we have the joint posterior distribution $P(\boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{\Theta} \mid \boldsymbol{Y})$ by multiplying three densities from each layer, $P_1(\boldsymbol{Y} | \boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{\Theta})$, $P_2(\boldsymbol{\beta}, \boldsymbol{\gamma} | \boldsymbol{\Theta})$, and $P_3(\boldsymbol{\Theta})$ such that

$$P(\boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{\Theta} \mid \boldsymbol{Y}) \propto P_1(\boldsymbol{Y} | \boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{\Theta}) P_2(\boldsymbol{\beta}, \boldsymbol{\gamma} | \boldsymbol{\Theta}) P_3(\boldsymbol{\Theta}). \tag{A5}$$

Premium Rating Game

We conduct the out-of-sample prediction game for the PHC with the Beta distribution. As stated in Footnote 4 in the manuscript, any type of parametric distribution can be applied in the PHC

approach. However, a substantial level of data omission often results in convergence issues. Therefore, we only report the results from the Beta distribution in the case of full data and a moderate level (10%) of the data omission here.

Table A1 presents the *between* test results. The "*p*-value" indicates the type 1 error obtained from the binomial distribution of RL^* and a lower *p*-value indicates that the PHC with the Beta distribution dominates the BMA approach. Like the Gaussian model in the manuscript, the PHC outperforms both BMA approaches in all crop/state/coverage-level/omission-level combinations. The results show that the PHC tends to be less sensitive to the data missingness and offers more adequate risk measures in deeper tail probability.

Table A2 shows the *within* test results. A lower *p*-value indicates that estimation results from the original dataset outperform the results from the data with missing observations. Therefore, the *within* test results are not statistically significant if the model suitably estimates densities from the omitted dataset. Similar to the comparison conducted in the manuscript for the *within* test, all approaches do not reject the null hypothesis in the case of 10% data omission in 90% coverage level. However, two BMA-based approaches show rejections in the 70% coverage level when 10% data omission happened.

Table A1. Out-of-Sample Rating Game Results of the *Between* Test

				BMA-Normal[1] vs PHC[3]		**BMA-Mixture[2] vs PHC**	
Crop	**State**	**No. of Counties**	**Data omission**	**90% Coverage (*p*-value)**	**70% Coverage (*p*-value)**	**90% Coverage (*p*-value)**	**70% Coverage (*p*-value)**
Corn	Iowa	99	No omission	0.084*	0.000***	0.048**	0.000***
			10%	0.021**	0.000***	0.073*	0.015**
	Maryland	23	No omission	0.084**	0.005***	0.090**	0.015**
			10%	0.000***	0.000***	0.001**	0.000***
	Colorado	27	No omission	0.000***	0.000***	0.001***	0.000***
Wheat	Kansas	105	No omission	0.005***	0.005***	0.015**	0.000***
			10%	0.000***	0.000***	0.001***	0.000***
	Indiana	92	No omission	0.015**	0.000***	0.010**	0.000***
			10%	0.000***	0.025**	0.001***	0.018**
	Colorado	27	No omission	0.000***	0.000***	0.000***	0.000***

Note : A lower *p*-value indicates the proposed approach dominates the BMA approach, vice versa. Statistical significance of lower private (government) loss ratios indicated by ∗–10%, ∗∗–5%, and ∗∗∗–1%.

[1] Bayesian Model Averaging approach under the normality assumption (Ker, Tolhurst, and Liu 2016).

[2] Bayesian Model Averaging approach under the normal mixture (Ker, Tolhurst, and Liu 2016).

[3] Park-Harri-Coble (PHC), the proposed Bayesian Kriging approach in the paper.

Table A2. Out-of-Sample Rating Game Results for the *Within* Test (no omission vs dataset with omission)

Crop	State	Model	Data omission	90% Coverage (*p*-value)	70% Coverage (*p*-value)
Corn	Iowa	BMA-Normal[1]	No vs 10%	0.696	0.059*
		BMA-Mixture[2]	No vs 10%	0.500	0.059*
		PHC[3]	No vs 10%	0.212	0.313
	Maryland	BMA-Normal	No vs 10%	0.133	0.040**
		BMA-Mixture	No vs 10%	0.119	0.407
		PHC	No vs 10%	0.377	0.345
Wheat	Kansas	BMA-Normal	No vs 10%	0.407	0.000***
		BMA-Mixture	No vs 10%	0.500	0.676
		PHC	No vs 10%	0.274	0.773
	Indiana	BMA-Normal	No vs 10%	0.227	0.105
		BMA-Mixture	No vs 10%	0.315	0.166
		PHC	No vs 10%	0.252	0.820

Note : A lower *p*-value indicates the proposed approach dominates the BMA approach, vice versa. Statistical significance of lower private (government) loss ratios indicated by ∗–10%, ∗∗–5%, and ∗∗∗ – 1%.

[1] Bayesian Model Averaging approach under the normality assumption (Ker, Tolhurst, and Liu 2016).

[2] Bayesian Model Averaging approach under the normal mixture (Ker, Tolhurst, and Liu 2016).

[3] Park-Harri-Coble (PHC), the proposed Bayesian Kriging approach in the paper.

Supplement B: Computational Details.

The hierarchy of the model is structured as following three stages,

$$P_1(\boldsymbol{Y}|\,\boldsymbol{\beta},\boldsymbol{\gamma},\boldsymbol{\Theta})$$

$$P_2(\,\boldsymbol{\beta},\boldsymbol{\gamma}\,|\boldsymbol{\Theta})$$

$$P_3(\boldsymbol{\Theta}).$$

By Bayes' theorem, the prior distribution for the likelihood layer, say $P(\boldsymbol{\beta},\boldsymbol{\gamma},\boldsymbol{\Theta}\,)$, can be separated into two components,

$$P(\boldsymbol{\beta},\boldsymbol{\gamma},\boldsymbol{\Theta}\,) = P_2(\,\boldsymbol{\beta},\boldsymbol{\gamma}\,|\boldsymbol{\Theta})P_3(\boldsymbol{\Theta})$$

Under this setting, the joint posterior distribution of the hierarchical model can be estimated as

$$P(\boldsymbol{\beta},\boldsymbol{\gamma},\boldsymbol{\Theta}\mid \boldsymbol{Y}) = \frac{P(\boldsymbol{\beta},\boldsymbol{\gamma},\boldsymbol{\Theta},\boldsymbol{Y})}{P(\boldsymbol{Y})} = \frac{P_1(\boldsymbol{Y}\mid \boldsymbol{\beta},\boldsymbol{\gamma},\boldsymbol{\Theta})P(\boldsymbol{\beta},\boldsymbol{\gamma},\boldsymbol{\Theta})}{\iiint P_1(\boldsymbol{Y}\mid \boldsymbol{\beta},\boldsymbol{\gamma},\boldsymbol{\Theta})P(\boldsymbol{\beta},\boldsymbol{\gamma},\boldsymbol{\Theta})d\boldsymbol{\Theta}d\boldsymbol{\gamma}d\boldsymbol{\beta}}$$

and thus the joint posterior distribution $P(\boldsymbol{\beta},\boldsymbol{\gamma},\boldsymbol{\Theta}\mid \boldsymbol{Y})$ is proportional to the multiplication of the likelihood $P_1(\boldsymbol{Y}\mid \boldsymbol{\beta},\boldsymbol{\gamma},\boldsymbol{\Theta})$ and $P(\boldsymbol{\beta},\boldsymbol{\gamma},\boldsymbol{\Theta})$

$$P(\boldsymbol{\beta},\boldsymbol{\gamma},\boldsymbol{\Theta}\mid \boldsymbol{Y}) \propto P_1(\boldsymbol{Y}\mid \boldsymbol{\beta},\boldsymbol{\gamma},\boldsymbol{\Theta})P(\boldsymbol{\beta},\boldsymbol{\gamma},\boldsymbol{\Theta}).$$

Next, we plug in $P_2(\boldsymbol{\beta},\boldsymbol{\gamma}\mid\boldsymbol{\Theta})P_3(\boldsymbol{\Theta})$ instead of $P(\boldsymbol{\beta},\boldsymbol{\gamma},\boldsymbol{\Theta})$, resulting in the final formula

$P(\boldsymbol{\beta},\boldsymbol{\gamma},\boldsymbol{\Theta}\mid \boldsymbol{Y}) \propto P_1(\boldsymbol{Y}\mid \boldsymbol{\beta},\boldsymbol{\gamma},\boldsymbol{\Theta})P_2(\boldsymbol{\beta},\boldsymbol{\gamma}\mid\boldsymbol{\Theta})P_3(\boldsymbol{\Theta})$.

We sample posteriors using the Metropolis-Hastings (MH) steps within a Gibbs sampler algorithm written in `R`. We use Gibbs sampling to update vectors of all coefficients of mean and standard deviation equations that spans all counties ($\boldsymbol{\beta}_1$, $\boldsymbol{\beta}_2$, $\boldsymbol{\gamma}_1$, and $\boldsymbol{\gamma}_2$), including missing observations y_{it}^m, via a Gaussian candidate density. The sill and range (ρ and θ) parameters, also known as Kriging parameters, are also updated by Gibbs sampling. The mean and variance of the Gaussian candidate are obtained by maximum likelihood estimation under the specification in (2). Missing observations are also updated within the Markov Chain Monte Carlo (MCMC) draws. We use a Gaussian candidate for missing observations $y_{it}^m \sim N(\mathrm{E}[\boldsymbol{y}_t], \mathrm{var}[\boldsymbol{y}_t])$ where y_{it}^m is a missing observation for county i in year t, $\mathrm{E}[\boldsymbol{y}_t]$ and $\mathrm{var}[\boldsymbol{y}_t]$ are the cross-sectional (across counties) mean and variance of yields in year t. The spatial smoothing to get county-specific parameter estimates is conducted when updating the coefficients within the MCMC. For the final estimation, we generate 50,000 MCMC samples to get posterior densities and drop the first 10,000 observations as burn-in.

We follow a general Bayesian spatial interpolation algorithm. The spatial smoothing to get county-specific parameter estimates is conducted when the model updates the coefficients within the MCMC. We first generate a random Gaussian spatial process $\boldsymbol{z}_k = [z_{1k}, \dots, z_{Nk}]'$, $\boldsymbol{z}_k \sim N(0,1)$, where $k = 1, \dots, K$. Note that since the model assumes the Gaussian process, $\lim_{n\to\infty} \frac{\sum_{i=1}^{N} z_{ik}}{n} = 0$ for every kth MCMC draw. The model then conducts a Cholesky decomposition with the spatial correlation matrix, $\Sigma_k = \psi(D_{ij}; \theta_k, \rho_k) = \mathbf{L}_\mathrm{k}\mathbf{L}_\mathrm{k}'$, where $\mathbf{L}_\mathrm{k}$ is a lower triangular matrix from the Kriging parameters θ and ρ. Then we draw $\boldsymbol{\beta}_k = \boldsymbol{\beta}_k + \mathbf{L}_\mathrm{k}\boldsymbol{z}_k$ from the random Gaussian process $\boldsymbol{z}_k$ and then update location specific coefficient parameters $\boldsymbol{\beta}_k$.

The Metropolis-Hastings within Gibbs sampler algorithm is used to draw posterior samples implemented in `R`. Gibbs sampling updates all coefficient parameters and Kriging parameters. For the final estimation, we generate 50,000 MCMC samples to get posterior densities and drop the first 10,000 observations as burn-in. A trace plot is used to monitor a graphical convergence of the MCMC samples. We also check the convergence of the selected counties via Geweke (1992) convergence diagnostic for Markov chains, which is based on a test for equality of the means of the first 20% and last 50% part of a Markov chain. In general, the range parameters $\theta_{\beta 1}$, $\theta_{\beta 2}$, $\theta_{\gamma 1}$ and $\theta_{\gamma 2}$ are the parameters with the convergence issue. Therefore, we also check the convergence of the range parameters via both the traceplot and the Geweke diagnostic test. All range parameters satisfy the convergence criteria. The following figures are the traceplots of the range parameters.

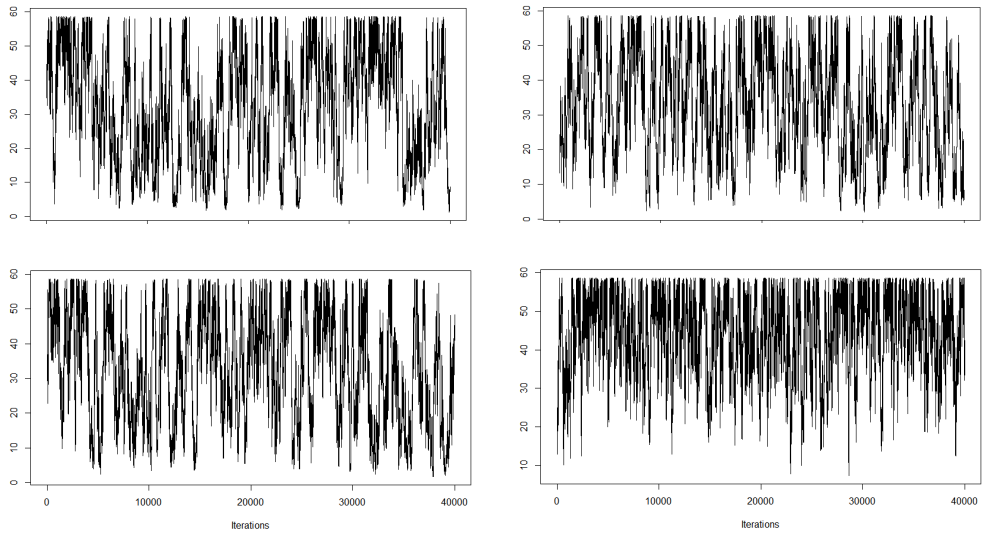

Figure B1. Posterior Range Parameters of Mean (β_1, β_2) and Variance (γ_1, γ_2) Equations in Iowa Corn

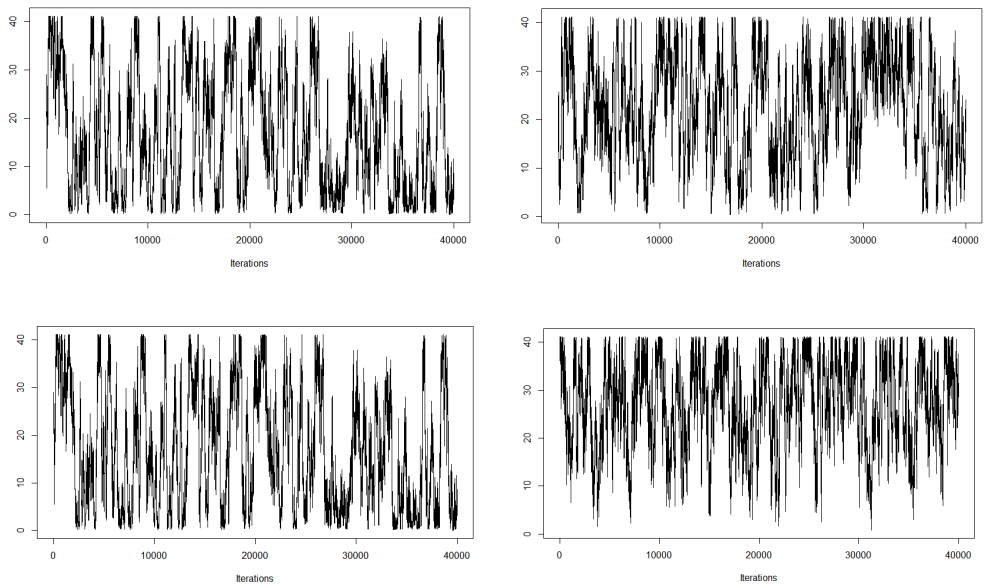

Figure B2. Posterior Range Parameters of Mean (β_1, β_2) and Variance (γ_1, γ_2) Equations in Maryland Corn

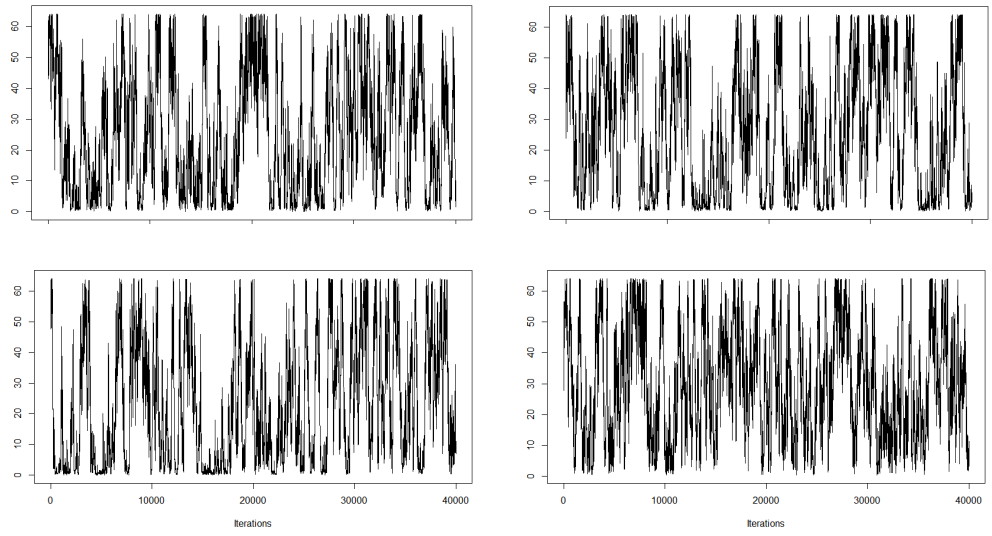

Figure B3. Posterior Range Parameters of Mean (β_1, β_2) and Variance (γ_1, γ_2) Equations in Colorado Corn

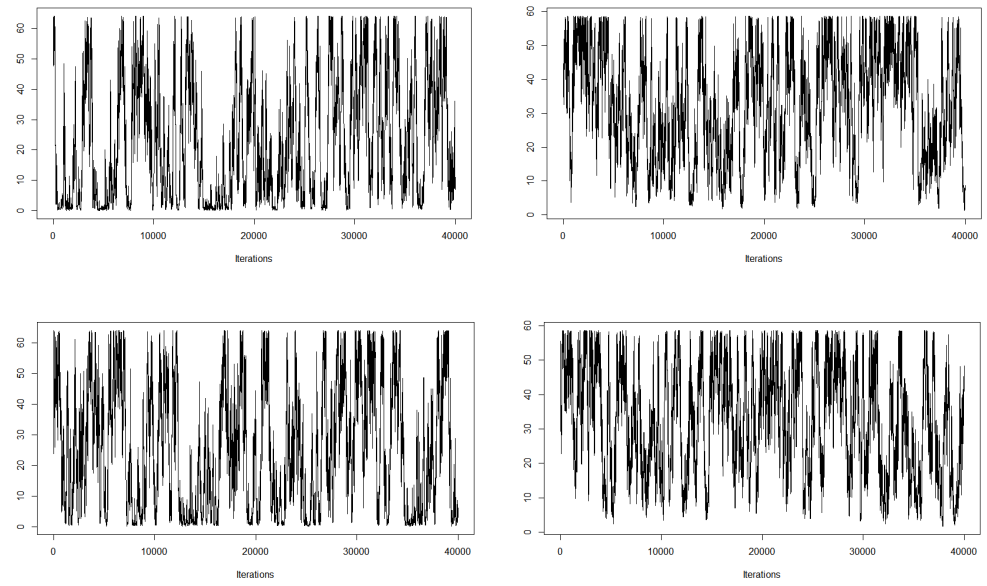

Figure B4. Posterior Range Parameters of Mean (β_1, β_2) and Variance (γ_1, γ_2) Equations in Kansas Wheat

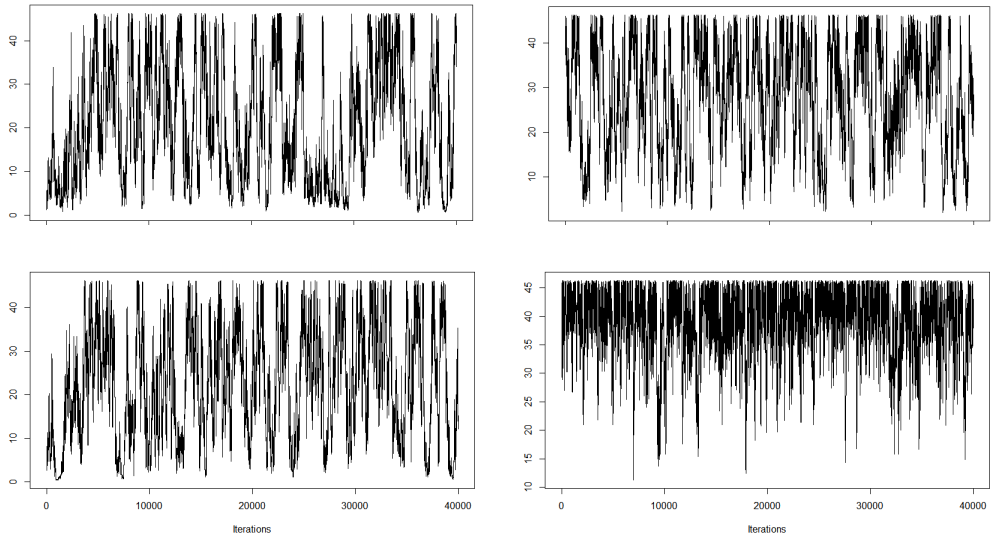

Figure B5. Posterior Range Parameters of Mean (β_1, β_2) and Variance (γ_1, γ_2) Equations in Indiana Wheat

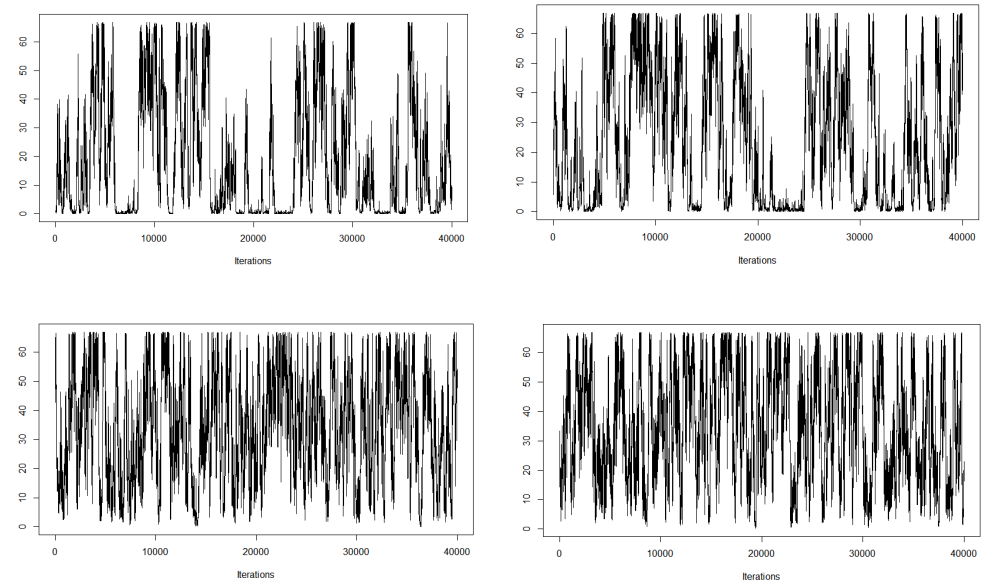

Figure B6. Posterior Range Parameters of Mean (β_1, β_2) and Variance (γ_1, γ_2) Equations in Colorado Wheat

Supplements C: Economic Significance of Adopting the Proposed Approach

Table C1. Descriptive Statistics of the RMA County-level Business Database

Programs (Years)	Variable	Mean ($, acre)	Maximum ($, acre)	Minimum ($, acre)	Total ($, acre)
ARP (2014 - 2015)	Indemnity	75,555	7,530,528	50,608	38,004,188
	Premium	96,562	6,258,555	0	48,571,095
	Subsidy	42,768	2,753,766	0	21,512,308
	Acreage	1,108	50,608	0	557,811
AYP (2014 - 2015)	Indemnity	2,265	212,838	0	1,094,053
	Premium	8,619	438,395	0	4,163,260
	Subsidy	4,631	223,582	0	2,236,803
	Acreage	207	5,259	0	130,574
GRP (2005-2013)	Indemnity	5,885	1,806,200	0	23,742,318
	Premium	11,430	1,041,273	0	46,111,154
	Subsidy	6,289	531,049	0	25,371,615
	Acreage	645	23,821	0	2,603,692

Note: ARP stands for Area Revenue Protection, AYP stands for Area Yield Protection, and GRP stands for Group Risk Plan.
The sample is covering 1,590 counties over 11 years (2005−2015), with a total of 5,020 observations.

Table C2. Estimated Overpayment Measures

Programs	Variable	Value
ARP	Unit indemnity loss per acre	$54.56
	Unit subsidy loss per acre	$43.11
AYP	Unit indemnity loss per acre	$3.72
	Unit subsidy loss per acre	$17.29
GRP	Unit indemnity loss per acre	$2.10
	Unit subsidy loss per acre	$9.84

Note: ARP stands for Area Revenue Protection, AYP stands for Area Yield Protection, and GRP stands for Group Risk Plan.
The unit indemnity/subsidy losses in the table represent that an increment of one loss ratio deviation leads to overpayments of the respective indemnity/subsidy per acre.

Table C3. Estimated Overpayment Savings (per acre) of Using the PHC Approach

Model	Program	Variable	No omission	10% omission	30% omission	60% omission
PHC vs HCKG	ARP	Indemnity saving per acre	$59.58	$75.60	$259.04	$311.46
		Subsidy saving per acre	$8.93	$7.57	$5.75	$6.68
	AYP	Indemnity saving per acre	$9.29	$12.58	$15.45	$13.48
		Subsidy saving per acre	$0.87	$0.74	$0.56	$0.67
	GRP	Indemnity saving per acre	$12.56	$17.00	$19.21	$18.21
		Subsidy saving per acre	$1.19	$1.00	$0.76	$0.92
PHC vs BMA-Mixture	ARP	Indemnity saving per acre	$45.03	$54.78	$69.32	$212.05
		Subsidy saving per acre	$8.62	$9.21	$6.36	$7.05
	AYP	Indemnity saving per acre	$0.75	$0.91	$1.15	$3.52
		Subsidy saving per acre	$0.84	$1.23	$0.62	$0.73
	GRP	Indemnity saving per acre	$1.01	$1.01	$1.56	$4.76
		Subsidy saving per acre	$1.14	$1.22	$0.85	$0.99

References

Geweke, J. 1992. Evaluating the Accuracy of Sampling-Based Approaches to Calculating Posterior Moments. *In Bayesian Statistics 4*. Clarendon Press, Oxford, UK.

Norwood, B., M. C. Roberts, J. L. Lusk (2004), Ranking Crop Yield Models Using Out-of-Sample Likelihood Functions, *American Journal of Agricultural Economics*, 86(4): 1032–43.