

# SPATIO-TEMPORAL MODELING OF AGRICULTURAL YIELD DATA WITH AN APPLICATION TO PRICING CROP INSURANCE CONTRACTS

VITOR A. OZAKI, SUJIT K. GHOSH, BARRY K. GOODWIN, AND RICARDO SHIROTA

This article presents a statistical model of agricultural yield data based on a set of hierarchical Bayesian models that allows joint modeling of temporal and spatial autocorrelation. This method captures a comprehensive range of the various uncertainties involved in predicting crop insurance premium rates as opposed to the more traditional ad hoc, two-stage methods that are typically based on independent estimation and prediction. A panel data set of county-average yield data was analyzed for 290 counties in the State of Paraná (Brazil) for the period of 1990 through 2002. Posterior predictive criteria are used to evaluate different model specifications. This article provides substantial improvements in the statistical and actuarial methods often applied to the calculation of insurance premium rates. These improvements are especially relevant to situations where data are limited.

*Key words:* crop insurance, hierarchical Bayesian models, spatio-temporal models.

Historically, crop insurance in Brazil has been offered by the government at both the federal and state levels. In spite of the government's efforts, the experience with crop insurance in Brazil has generally not been satisfactory. The absence of a suitable actuarial method to price crop insurance contracts is one of the main reasons for the poor performance and ultimate failure of this agricultural risk management program. High premium rates inhibited demand for the insurance by producers and, at the same time, selected only those growers with a higher probability of receiving the indemnity. This is the classic problem of adverse selection, which has characterized historical efforts at developing crop insurance in Brazil.

In recent years, efforts have been made to improve the performance of the programs and to make crop insurance more popular among producers. In December 2003, premium subsidies were introduced by the federal government and the program was focused on providing coverage to those engaged in activities considered to be risk reducing or technology enhancing. State governments have also undertaken actions to stimulate produc-

ers' demand for crop insurance. These actions have included a number of additional subsidy programs and expansions to cover a wide variety of crops. Covered crops now include cotton, peanuts, irrigated rice, cassava, soybeans, sorghum, wheat, pineapples, plums, kaki, guava, passion fruit, peaches, and cabbage and premium subsidies exceed 50% in some cases.

This article concentrates on statistical and actuarial methods with the objective of pricing an alternative crop insurance contract based on county yields. This type of insurance is widely available in the United States (the group risk plan or GRP) as well as in India, Sweden, and Canada (Miranda, Skees, and Hazell 1999). It is also offered in Brazil in the state of Rio Grande do Sul. It is important to point out that the methods proposed in this article can also be applied to pricing others forms of insurance contracts, such as those based on individual yields, as long as there are sufficient data to do the analysis.

## Historical Background

Agricultural insurance was introduced in the form of hail coverage in Brazil in 1938. The early performance of the program was poor, with loss ratios (i.e., the ratio of indemnities paid out to premiums collected) above 3.8 being observed in the early years. In January 1954, the federal government of Brazil

---

Vitor A. Ozaki and Ricardo Shirota are professors, respectively, in the Department of Exact Sciences and Economics, Business, and Sociology, University of São Paulo; Sujit K. Ghosh and Barry K. Goodwin are full and William Neal Reynolds Distinguished Professor, respectively, in the Department of Statistics and Agricultural and Resource Economics, North Carolina State University.

Research support from the North Carolina Agricultural Research Service and CAPES (Brazil) are gratefully acknowledged.

established the Agrarian Insurance Stability Fund to guarantee insurance market stability, allow the gradual adjustment of premium rates, cover catastrophic risks, and to provide a number of other initiatives to improve crop insurance. However, the program was not successful due to extreme centralization of the program's administration and ignorance of the specific features and peculiarities of local environments. In December 1973, the Farming Activity Guarantee Program, called PRO-AGRO, was created to protect the financial system in case of large-scale defaults on loan obligations by producers. From its beginnings through 1993, the program accumulated large deficits (\$1.6 billion) and suffered from a number of operational problems, including fraud and abuse. New legislation in 1996 stipulated specific seeding rates and introduced differentiated insurance premium rates for farmers that adhered to recommended practices. Several private insurance companies currently offer crop insurance in Brazil.<sup>1</sup> Although the amount of business is still small, some pilot projects have been implemented, including area-wide plans based on county yields in the State of Rio Grande do Sul. Other types of crop insurance including individual farm coverage and cooperative loss-sharing arrangements can also be found in the country.

In the analysis that follows, a number of alternative Bayesian hierarchical models are considered for Brazilian corn yield data with the objective of modeling the stochastic-generating process of yield data and, in particular, properly recognizing the temporal, spatial, and spatio-temporal dynamics underlying crop yields. To select among a large number of potential candidate models, a minimum mean square prediction error criterion is used.

### Statistical Modeling Framework

The management of agricultural risks presents a set of important economic and statistical problems. Myers (1988) demonstrated that crop insurance programs, taken in conjunction with futures markets, may provide an important contingency market framework for the management of risk and the enhancement of

economic welfare.<sup>2</sup> A fundamental parameter of any insurance contract is the premium rate. An actuarially fair premium rate is a rate that is set such that premiums collected are equal to expected indemnities. An inaccurate premium rate results in distortions to the insurance pool and thus may result in program losses as agents adversely select against the insurance provider. In particular, low-risk agents may be overcharged and high-risk agents may be undercharged. This will distort participation in favor of the higher risks, and thus premiums will not be sufficient to cover indemnity payments. This condition of adverse selection has been well documented for a number of crop insurance plans. The eventual failure of an insurance program as a result of such selection is often called the "death spiral of adverse selection." Optimally, an insurance provider would prefer to calculate individual premium rates for each farmer on the basis of that farmer's risks and expected yields. However, individual data are rare at best, and thus crop insurance plans are often based upon more aggregate data—such as data at the county level. Such index-based crop insurance plans were developed to overcome the problem of short or nonexistent individual crop yield series.

Another important aspect of insurance contract design pertains to the actuarial procedures used in the calculation of insurance premium rates. In particular, the derivation of such rates generally requires a statistical analysis of crop yields. A wide variety of statistical methods are often adopted in the estimation of crop insurance rates and a number of issues relating to the modeling of crop yields are pertinent to these methods. For example, one often must address issues related to the fact that yields tend to have substantial trends and tend to be significantly correlated across space due to the systemic nature of weather. One subtlety often overlooked in crop insurance yield models pertains to the fact that a degree of uncertainty also applies to the parameters of any model used to describe the uncertainty of yields. For example, it is common to detrend yields using standard regression models and then use the detrended yields to measure yield uncertainty. However, a certain degree of uncertainty is also inherent in the models used to detrend yields. This

<sup>1</sup> In contrast to the U.S. crop insurance program, the Brazilian government neither fixes premium rates or other parameters of the contract nor subsidizes the administration or operational costs of the insurers.

<sup>2</sup> Innes and Rausser (1989) discuss the welfare impacts of various agricultural policies under conditions of uncertainty and demonstrate how these effects may differ from what "first-best" implications may suggest.

is analogous to the common problems associated with treating generated or predicted data as though it were directly observable without error. In this analysis, we adopt a Bayesian inferential framework that accounts for all such sources of uncertainty while estimating the appropriate premium rate.

Over many years, the statistical issues underlying agricultural yields have been a controversial point in the crop insurance literature. Several statistical approaches have been considered, including parametric yield models, semiparametric methods (Ker and Coble 2003), nonparametric models (Goodwin and Ker 1998; Turvey and Zhao 1999), and empirical Bayes nonparametric approaches (Ker and Goodwin 2000).

Within the parametric modeling approach, some researchers have concluded that crop yields tend to follow a normal distribution (Just and Weninger 1999). However, a large number of other researchers including Day (1965), Taylor (1990), Ramírez (1997) and Ramírez, Misra, and Field (2003) have found evidence against normality. Other suggestions included the use of a beta distribution (Nelson and Preckel 1989), inverse hyperbolic sine transformations (Moss and Shonkwiler 1993), and gamma distributions (Gallagher 1987). Sherrick et al. (2004) used several parametric distributions including the normal, lognormal, beta, weibull, and logistic distributions to model individual yield data. Of course, the characteristics of crop yields may be idiosyncratic and may vary by location, crop, and production practice. Thus, it is unlikely that any single parametric approach will be universally supported across different applications. A subtle point regarding conditioning variables is also relevant to crop insurance applications. One only wishes to condition on those variables that are deterministic or that can be observed before insurance contracts are offered. For example, weather clearly influences crop yields but is generally unpredictable, and thus crop yields are not conditioned on past weather observations. Insurance contracts typically assume that “best production practices” are followed (i.e., that no moral hazard exists), and thus optimal levels of input usage are assumed and yields are not typically conditioned on inputs.<sup>3</sup>

<sup>3</sup> An important component of the loss-adjustment process in any insurance program relates to the verification that the loss is due to an insurable event and not due to negligence or deliberate actions of the insured agent.

As we have pointed out, a related problem pertains to the limited number of yield observations typically available for empirical models. This is true even when aggregated data are considered. This limitation typically precludes the use of individual farm-level data for the purposes of modeling yields and rating insurance contracts. The choice of a statistical model that adequately reflects the conditional density of yields is an important consideration in the actuarial calculation of an accurate premium rate. In doing this, one must try to recover the probability-generating process of the yield data. Agricultural yields follow a spatio-temporal process, in the sense that, if we take the average in a region conditional on the underlying temporal process, one can recover the conditional yield density generated by the information known at moment  $t$  (Ker and Goodwin 2000).

In most empirical work, the only information known at time  $t$  is the time index and previously realized yields. Thus, in these analyses, the conditional density is based only on the temporal-generating process of the data. Our work addresses this temporal aspect of the data-generating process, but we also give attention to the spatial dimension. In particular, we explicitly recognize the fact that the events that underlie yield realizations (e.g., weather, disease, and pest damages) tend to affect large areas at any single time. Thus, adjacent regions may experience substantial spatial correlations of yields over time. Thus, our models combine the two aspects of space and time in order to construct a spatio-temporal model of crop yields.

In this article, we simultaneously model the time trend and temporal and spatial autocorrelation and obtain premium rate estimates directly (within the model) in contrast to two-stage methods. A typical two-stage method will first detrend the time series and then treat the detrended yield data (often referred to as “normalized yields”) as “observed” data to estimate the premium rate.<sup>4</sup> Thus, this method fails to adequately capture the uncertainty of the premium rate estimate. Our approach makes the premium rate calculation less ad hoc, in the sense that rates are derived directly from a predictive distribution obtained by a simulation-based method known as the Markov Chain Monte Carlo algorithm (MCMC). Moreover, when calculating the rate we are able to

<sup>4</sup> For further details see Goodwin and Ker (1998) and Goodwin and Mahul (2004).

capture its model-based uncertainty through a standard error that is estimated within the model.

The fact that our data set is not large in the time dimension creates additional difficulties regarding the forecast or prediction of crop yields in future years.<sup>5</sup> In the construction of crop insurance contracts, it is typically the case that the terms and parameters of the contract must be available one to two years prior to the insurance cycle. This reflects the fact that crop yield data may take some time to be adequately measured following the harvest.<sup>6</sup> Further, an insurance provider will not offer coverage after the insurance buyers have information about their yields. For example, contracts must typically be signed before planting season. If not, farmers may have an information advantage over insurers, who had to specify contract parameters at a much earlier date. In addition, administrative issues relating to the operation of any program require substantial lead time in providing the parameters of the contract offering. In our case, the last observation recorded was for the year 2002. We will assume that there is a two-year lag between the receipt of historical yield data and the deadline required for filing new contract terms.<sup>7</sup> In this context, we must attempt to choose the best possible statistical model to predict yields for the following two years. In light of this objective, we model the structure of the yield mean and assume that the precision of our models is conditionally constant throughout the analysis. Gelfand et al. (1998) point out that modeling the mean component rather than the precision in forecasting models results in more effective results.

Before continuing the analysis we describe the basic notation. Let  $y_{it}$  be the agricultural yield in county  $i$  in year  $t$ , where  $i = 1, 2, \dots, S$  and  $t = 1, 2, \dots, N$ . Conditionally on a stochastic mean, we assume that the observed data follow a normal distribution, such that,  $y_{it} \sim N(\mu_{it}, \sigma^2)$ . The objective is to model the stochastic mean component, so that  $\mu_{it}$  reflects the temporal effects, spatial variation, and the spatio-temporal relationships relevant to agri-

cultural yields. In some applications, statistical models may be comprised of a large number of parameters. This is especially true in analyses of data that have been pooled over time and over cross-sections. In such cases, a natural way of modeling the parameters is through hierarchical models. Under such an approach, the dependence structure between the parameters can be represented by the joint probability distribution. Consequently, we can define a prior distribution for these parameters, assuming that they can be considered as a sample from a common population distribution. In our case a version of the model can be represented by the following hierarchical structure:

$$\begin{aligned} y_{it} &\sim N(\mu_{it}, \sigma^2) \\ \mu_{it} &= \psi_t + \phi_i \\ (1) \quad \psi_t &= \sum_{l=0}^p \beta_l t^l + u_t \\ u_t &= \rho u_{t-1} + v_t \\ \phi &\sim N_S(0, \sigma_\phi^2 W) \end{aligned}$$

where  $\phi = (\phi_1, \dots, \phi_S)^T$  denotes the vector of spatial random effects that is assigned a conditionally autoregressive prior (CAR), which is a multivariate normal distribution with mean vector 0 and variance matrix  $\sigma_\phi^2 W$ , in which  $W$  is the weighting matrix, and the parameter vector  $\beta = (\beta_0, \dots, \beta_p)^T$  consists of the regression coefficients for the deterministic temporal trend. The  $u_t$ 's represent the autoregressive errors that are assumed to follow a mean zero first-order autoregressive AR(1) process for simplicity. More details of the model building and prior specifications are provided in a later section of empirical analysis.

Modeling the structure underlying the mean yield realization by adopting hierarchical models is intuitive and facilitates the visualization of each component in the analysis instead of modeling such structure directly through the  $y_{it}$ .<sup>8</sup> A limitation of representing correlation structures through the use of hierarchical models is that all of the pairwise correlations must be positive. When modeling through

<sup>5</sup> In this article, the terms "forecast" and "prediction" and "density" and "distribution" are used interchangeably.

<sup>6</sup> For example, in the case of the U.S. Group Risk Program (GRP), aggregate yield data arrive with a two-year lag. This lag reflects the substantial amount of time required to derive accurate aggregate yield measures from farm-level surveys.

<sup>7</sup> Again, as we have noted, such a two-year lag exists in the U.S. Group Risk Program, which uses data from the National Agricultural Statistics Service (NASS).

<sup>8</sup> For this alternative version, Anselin (1988) shows several spatial and spatial-temporal models, such as SUR (seemingly unrelated regression), where the beta coefficients are allowed to vary in one of the two dimensions and the error term is correlated in the other dimension. In those models the dependence structure is modeled through the error term  $\varepsilon_{it}$ , where  $y_{it} = x_{it}\beta_{it} + \varepsilon_{it}$ .

this framework one must be careful to recognize that improper priors may yield improper posterior distributions.<sup>9</sup> It is possible to obtain proper posterior distributions even when improper priors are used (e.g., see Berger and Strawderman 1996; Sun, Tsutakawa, and Speckman 1999; Berger, Oliveira, and Sanso 2001; Hodges, Carlin, and Fan 2003). However one limitation on the use of improper priors is that Bayes factors cannot be used for model selection since it may depend on an undefined multiplicative factor (Kass and Raftery 1995).

In the case of improper priors, it is possible that the joint posterior distribution is improper, although the final results based on numerical output seem reasonable and the analyst may not realize the problem. In such a case, the analyst will make inferences about a nonexistent posterior distribution. In a practical sense, as shown in Gelfand and Smith (1990), this problem can be prevented by considering proper prior distributions that assure that the Gibbs sampling estimation process will be well behaved, where ignorance can be represented as values for the precision parameter close to zero.<sup>10</sup>

Initially, extending the work of Ker and Goodwin (2000), we modeled  $\mu_{it}$  as coming from two subpopulations or groups, a catastrophic group and a noncatastrophic group. A catastrophic event can be defined by an adverse climatic event that occurs in a determined period of time (such as drought, hail, etc.). Consequently, if such an adverse event occurs, the agricultural yield will be drawn from the catastrophic group. Alternatively, yields are considered to be drawn from the noncatastrophic group when normal weather events are realized. One can think of yield realizations as being drawn from a finite mixture of two distributions.

Alternatively, we assume that  $\psi_t = \sum_{l=0}^p \beta_l t^l + u_t$ . For this type of deterministic trend model, we center the variable  $t$  in order to improve the speed of convergence of our Markov Chain Monte Carlo (MCMC) algorithm. Thus, we have  $t^* = [t - (N + 1) \times 0.5]$ . We consider  $p = 1, 2$  in the model estimation and use the normal prior distributions for the intercept and trend parameters.

<sup>9</sup> In this context, Hobert and Casella (1996) estimated the parameters of a hierarchical linear mixed model using the Gibbs sampler and warned about using a noninformative prior distribution that can lead to an improper posterior distribution.

<sup>10</sup> However, even in this case Gelman et al. (2003) raises some computational and numerical issues.

As an initial data exploration technique, we use empirical plots to evaluate the type of trend that might be present in the data. This evaluation indicated that a quadratic trend was sufficient to capture deterministic trend effects in the yield data. Beyond the deterministic trend models, one can also analyze complementarily, stochastic trend models and the interactions between stochastic and deterministic models. The stochastic trend component follows a first-order autoregressive model AR(1).<sup>11</sup> We also adopt two assumptions regarding the exact specification of the model. First, the correlation parameter  $\rho$  in the stochastic trend models can be allowed to vary according to the region. Second, an exchangeable truncated normal prior can be assigned to the parameter  $\rho$  when it is allowed to vary by region and normal and inverse gamma hyper-distributions for the mean and variance parameters, respectively. The general temporal model is:

$$(2) \quad \psi_{it} = \rho_i y_{i,t-1} + \sum_{l=0}^p \beta_{li} t^{*l}.$$

The spatial correlation is modeled in equation (2) considering conditional autoregressive prior distributions (CAR) for the parameters of the deterministic trend, where  $\beta_{li} = \xi_i + c_{li}$ , such that  $\xi$  is assigned a CAR prior distribution and  $c_{li} \sim N(\mu_c, \sigma_c^2)$  (Besag 1974; Clayton and Kaldor 1987; Cressie and Chan 1989; Besag, York, and Mollié 1991; Bernardinelli, Clayton, and Montomoli 1995a). We also allow the spatial effects to be nested within the temporal process resulting in spatio-temporal models, such that the parameters of the deterministic trend ( $\beta$ 's) are modeled using the CAR prior. Intuitively, one can think of the trend parameters as being correlated across space, given time (Bernardinelli et al. 1995b; Waller et al. 1997; Dreassi 2003).

Several models emerge as potential candidates for our particular problem. A basic question is thus how to select the best model, taking into account one of the objectives of this

<sup>11</sup> In light of the small sample size, a more sophisticated temporal model was not possible. For example, Ker and Goodwin (2000, p. 465) proposed an IMA(1,1) process, represented by  $y_t = y_{t-1} + \theta_0 + \theta_1 e_{t-1} + e_t$ . The number of observations used in their article was small as well, though larger than in our case. Modeling an IMA(1,1) process with few observations can be troublesome with regard to the stability and convergence of the parameters. In this manner, because one can express an MA(1) process as an AR( $\infty$ ) process, they modeled the temporal process as an AR(4), such that  $y_t = y_{t-1} + \beta_0 + \beta_1(y_{t-1} - y_{t-2}) + \beta_2(y_{t-2} - y_{t-3}) + \beta_3(y_{t-3} - y_{t-4}) + \beta_4(y_{t-4} - y_{t-5}) + e_t$ .



work—prediction of agricultural yields. Traditional model selection criteria, such as the Bayes factor, are not applicable in cases like ours where noninformative or CAR prior distributions are used. Carlin and Louis (2000, p. 220), have shown that the use of improper priors results in improper conditional predictive distributions, limiting the use of Bayes factor as a model selection criterion in these cases. There are several alternative modifications of the Bayes factors that can be used in the presence of improper priors (e.g., see Section 5.3 of Kass and Raftery 1995). However, such alternatives are difficult to implement in our type of application, and we thus will consider such alternative methods as a part of our future work. Criteria based on cross-validation are also difficult to implement when more sophisticated models are considered due to the inclusion of heterogeneity and clustering variables defined only by the prior (Waller et al. 1997).

In this article, we select our model specification by adopting a criteria based on predictive densities. As Laud and Ibrahim (1995) pointed out, these criteria are easy to interpret since they are not based on asymptotic analysis and they allow for the incorporation of prior distributions. Working in the predictive space, the penalty appears without the necessity of asymptotic definitions. Intuitively, one can think that good models must result in predictions close to what is observed in identical experiments.

In this context, Gelfand and Ghosh (1998) formalized a predictive criterion using a general form of loss function. The objective is to minimize the posterior predictive loss. The posterior predictive distribution is given by:

$$(3) \quad f(y_{new} | y_{obs}) \\ = \int f(y_{new} | M) p(M | y_{obs}) dM$$

where  $M$  represents the set of all parameters in a given model and  $y_{new}$  is the replicate of the vector of observed data  $y_{obs}$ .

The penalty is considered in the criteria regardless of the model dimension. In this work, a slightly different version of the model selection criterion will be utilized. Instead of using the quadratic predicted error, the mean squared predictive error will be considered relative to the number of regions used in the analysis. Note that the inclusion of a common denominator in all models does not affect the criterion.

## Data Description

The agricultural yield data used in this study were provided by the Statistical and Geography Brazilian Institute and correspond to the period of 1990 through 2002 for corn in the state of the Paraná, located in the southern region of Brazil.<sup>12</sup> The state of Paraná is the largest producer of corn in the country, with a total amount produced in 2002 equal to 9,797,816 tons, which is slightly more than 27% of all Brazilian production. Corn yields in Paraná are generally the fourth highest in Brazil (3,987 kilograms per hectare–kg/ha in 2002).

The state is made up of 399 counties. Annual yield observations for all thirteen years are only available in 290 counties. Consequently, we carry out the analysis with only those counties with the largest number of observations. The five largest counties in terms of average yields are Castro (6,142 kg/ha), Ponta Grossa (5,629 kg/ha), Marilândia do Sul (5,488 kg/ha), Tibagi (5,346 kg/ha), and Catanduvas (4,923 kg/ha).

## Empirical Application

We begin our analysis by choosing the model that minimizes the posterior predictive loss. We analyze several different models (25 in all). From among these alternatives, the model that minimizes the predictive error criteria was the one including the deterministic and stochastic trend. Basically, this model can be expressed according to the following hierarchical structure:

$$(4) \quad y_{it} \sim N(\mu_{it}, \tau) \\ \mu_{it} = \rho_i y_{it-1} + \beta_{1i} + \beta_{2i}^C t^*.$$

In this framework, the prior distributions for the parameters are as follows:

$$\begin{aligned} \rho_i &\sim N(\mu_\rho, \tau_\rho) \\ \beta_{1i} &\sim N(\mu_{\beta_1}, \tau_{\beta_1}) \\ (5) \quad \beta_{2i}^C &= \xi_i + c \\ \xi_i &\sim N(\bar{\xi}_i, \sigma_\xi^2/n_i) \\ c &\sim N(\mu_c, \tau_c). \end{aligned}$$

<sup>12</sup> The data may be freely downloaded at [www.ibge.gov.br](http://www.ibge.gov.br).

At this point, a further comment will be useful considering the prior distribution for  $\beta_{2i}^C$ . We assume that the spatial structure  $\xi_i$  conditional on  $\xi_j$  ( $j \neq i$ ), is proportional to

(6)

$$\xi_i \mid \xi_j \sim \exp \left\{ -n_i / 2\sigma_\xi^2 \left( \xi_i - \sum_{j \neq i} \omega_{ij} \xi_j \right)^2 \right\}$$

where  $n_i \geq 0$  is a “sample size” associated with region  $i$  and  $\omega_{ij} \geq 0$  is the weight reflecting the influence of  $\xi_j$  on the conditional mean of  $\xi_i$ . We let  $\omega_{ij} = 1$  if  $j$  is a neighbor of  $i$  and 0 otherwise and set  $n_i$  equal to the number of neighbors of  $i$ . Thus, the conditional distribution  $\xi_i \mid \xi_j$  simplifies to  $\xi_i \sim N(\bar{\xi}_i, \sigma_\xi^2/n_i)$ , where  $\bar{\xi}_i$  is the average of the  $\xi_j$ ’s, where  $j$  indexes the neighboring sites of  $i$ . The variance parameter  $\sigma_\xi^2$  receives an inverse gamma prior distribution. Finally, the precision parameter  $\tau$  receives an inverse gamma distribution, such that  $\tau \sim IG(\nu, \kappa)$ .

The hyper-prior distributions are:

(7)

$$\begin{aligned} \mu_\rho &\sim N(\mu_0, \tau_0) \\ \tau_\rho &\sim IG(\nu_0, \kappa_0) \end{aligned}$$

where:

$$\begin{aligned} \mu_{\beta_1} &= 0, \quad \tau_{\beta_1} = 10^{-6}, \quad \mu_c = 0, \\ \tau_c &= 10^{-3}, \quad \nu = 10^{-1}, \quad \kappa = 10^{-3}, \\ \mu_0 &= 0, \quad \tau_0 = 10^{-1}, \\ \nu_0 &= 10^{-1} \kappa_0 = 10^{-1}. \end{aligned}$$

We ran three chains to check the mixing of the Markov sequence and also check for all the parameters the graphical diagnostics of convergence. Results showed that all parameters achieved good convergence and mixing. One of the main advantages of the Bayesian analysis is that one can incorporate uncertainty when estimating the parameters. Taking this fact into account, table 1 shows the estimated value of the parameters. For these counties, the average standard deviations for  $\beta_1$ ,  $\beta_2$ , and  $\rho$  are 582, 3.9, and 0.11, respectively. We show only descriptive statistics of the 290 counties in order to save space. Thus, the maximum predicted values of  $\beta_1$ ,  $\beta_2$ , and  $\rho$  are respectively, 2,410, 46.85, and 0.83. The minimum values are 550, 46.73, and 0.30 and the average, 1,174,

**Table 1. Predicted Parameter Values for Selected Counties**

County	Parameter	Posterior Mean
Castro	$\beta_1$	1,366
	$\beta_2$	46.83
	$\rho$	0.8073
Catanduvas	$\beta_1$	1,545
	$\beta_2$	46.78
	$\rho$	0.7147
Marilândia do Sul	$\beta_1$	1,446
	$\beta_2$	46.78
	$\rho$	0.7703
Ponta Grossa	$\beta_1$	1,511
	$\beta_2$	46.82
	$\rho$	0.7553
Rolândia	$\beta_1$	2,109
	$\beta_2$	46.79
	$\rho$	0.5579
Tibagi	$\beta_1$	1,380
	$\beta_2$	46.82
	$\rho$	0.7751

46.79, and 0.61. The average standard deviation is 430, 3.95, and 0.13.

Because the series are relatively short, we do not correct for conditional heteroskedasticity. Instead we assume that the series are conditionally homoskedastic. If the series were longer, a procedure that could be used to verify heteroskedasticity would allow the parameter to vary in time and space and subsequently monitor the parameter to verify the variation in the precision and correct it when necessary. In table 2 we show the predicted values of yields for Castro, Ponta Grossa, Marilândia do Sul, Tibagi, Catanduvas, and Rolândia counties. The variance of the predicted value tends to increase as the time lag increases.

*Rating the Crop Insurance Contract*

Pricing an insurance contract accurately is essential for the viability and existence of an agricultural insurance market. Premium rates that are too high result in an insurance pool made up of only high-risk individuals. Similarly, rates that are universally too low will result in insurance losses since premiums are not adequate to cover indemnity outlays. The selection problem that is brought about by inaccurate rates is known as adverse selection. In the literature of insurance economics, this is often also referred to as the hidden information problem

**Table 2. Predicted Yield Values (kg/ha), of Selected Counties in 2003 and 2004**

County	Year	Predicted Yield
Catanduvas	2003	5,968
	2004	5,833
Castro	2003	8,301
	2004	8,455
Marilândia do Sul	2003	7,499
	2004	7,624
Ponta Grossa	2003	6,553
	2004	6,638
Rolândia	2003	7,336
	2004	7,461
Tibagi	2003	7,730
	2004	7,779

since agents tend to know more about their risks than does the insurance provider.

The insurance premium rate (PR) represents expected payouts as a proportion (or percentage) of total liability. In the simple case where a proportion  $\lambda$  ( $0 \leq \lambda \leq 1$ ) of the expected crop yield  $y^e$  is used to form the basis of insurance, the premium rate is given by:

$$(8) \quad PR = \frac{F(\lambda y^e) E_Y[\lambda y^e - y \mid y < \lambda y^e]}{\lambda y^e} \\ = \frac{\int_0^{\lambda y^e} P^G(\lambda y^e - y) f(y) dy}{P^G \lambda y^e}$$

where  $E$  is the expectation operator,  $P^G$  is the price at which losses will be paid,  $f$  is the probability density for yields, and  $F$  is the cumulative distribution function of yields. Note that the premium rate is completely transparent to the price at which yields are valued since the price term would appear in both the numerator and denominator of the premium rate expression. Rates can be derived directly from our Bayesian hierarchical model. A slightly different derivation of the premium rate is convenient for our purposes.

If we reparameterize  $y$ , such that,  $y^* = y/\lambda y^e$ , then (8) becomes:

$$(9) \quad PR = P(y^* < 1) E_{y^*}[1 - y^* \mid y^* < 1]$$

Note that the support of the random variable  $Y$  remains the same in this transformation. If we consider  $w = 1 - y^*$ , then (9) can be rewritten such that:

$$(10) \quad PR = P(w > 0)[1 - E_w(1 - w \mid w > 0)] \\ = P(w > 0) E_w[w \mid w > 0].$$

After some simplification, the premium rate equation (10) reduces to:

$$(11) \quad PR = \int_0^1 w f(w) dw.$$

We can similarly write (11) as  $PR = E[wI(0 < w < 1)]$ . Because of the change of variable, the support is also changed such that  $w$  now lies between 0 and 1. In our model, we can easily computationally implement (10) using the predicted yields. This expression represents the mean of  $w$ , or more specifically, the “posterior mean” of  $w$ , which is the PR calculated for each county and for each level of coverage. Moreover, the Bayesian approach allows one to derive standard error estimates of the premium rates which, in our context, are the Monte Carlo (MC) standard errors of the mean (Spiegelhalter et al. 2003). In table 3, we show some rates and their MC errors. Note that Antonina County was included to illustrate the variability in rates. The standard deviation and consequently the MC error are much higher in this county compared to the others. In this case, the uncertainty about premium rates is much higher.

A natural advantage of having a viable measure of the uncertainty associated with an individual premium rate estimate can be found in the common insurance practice known as “loading.” Loading refers to markups or additive factors that are often applied to premium rates to account for uncertainty.<sup>13</sup> These adjustments are typically ad hoc and are based upon the actuary’s confidence in the estimate. The standard errors of the premium rate estimates provide a natural metric to guide such loading practices. In particular, higher load adjustments can be applied to those rates that reflect a greater degree of uncertainty. The provision of such standard errors is one important innovation offered by our research. The standard errors account for all of the uncertainty that are associated with the model, including the estimation of yield trend effects and spatial correlation factors.

<sup>13</sup> Loading is also commonly used to build reserves, to cover administrative and operating costs, and to ensure a positive profit for the insurer.



**Table 3. Premium Rates for Selected Counties by Level of Coverage and MC Error Estimates**

County	Level of Coverage (%)	Premium Rate (%)	MC Error of Premium Rate
Antonina	70	7.295	0.001542
	75	9.478	0.001706
	80	11.810	0.001863
	85	14.230	0.002002
	90	16.690	0.002120
Castro	70	0.014	0.00002721
	75	0.084	0.00009512
	80	0.318	0.00026260
	85	0.897	0.00051740
	90	2.041	0.00083410
Catanduvas	70	0.017	0.00003252
	75	0.096	0.00009258
	80	0.342	0.00020670
	85	0.905	0.00037390
	90	1.920	0.00057560
Marilândia do Sul	70	0.013	0.00002727
	75	0.076	0.00007440
	80	0.314	0.00017840
	85	0.899	0.00037790
	90	2.007	0.00068260
Ponta Grossa	70	0.006	0.00001766
	75	0.039	0.00005644
	80	0.178	0.00014510
	85	0.553	0.00030040
	90	1.326	0.00053250
Rolândia	70	0.001	0.00000634
	75	0.013	0.00002677
	80	0.063	0.00006643
	85	0.220	0.00013670
	90	0.593	0.00024130
Tibagi	70	0.016	0.00002892
	75	0.096	0.00010000
	80	0.356	0.00024060
	85	0.980	0.00045110
	90	2.120	0.00070750

**Conclusions**

We have discussed a statistical and actuarial method of pricing a crop insurance contract that is based upon hierarchical Bayesian models. Our models of the probability-generating process of yield data consider temporal and spatial effects as well as the interaction between these two effects, resulting in spatio-temporal models. The contracts are based upon a regional crop yield index. Such crop insurance plans have been adopted in many areas, including in the United States. Area-wide plans of this sort are now being implemented as an alternative risk management tool in the

South of Brazil. We point out that this methodology can also be applied to contracts based on individual yields as long as there are enough data to conduct the statistical analysis. Conventional methods of pricing this type of individual contract using aggregate yield data such as county averages are not recommended because they do not reflect accurately the risk structure of an individual producer, thus increasing the problem of adverse selection. The use of these new risk management tools, together with the approval of new legislation in December 2003, provides support for the continued development of a crop insurance market in Brazil. Similarly, these developments

improve incentives for the entry of new private insurance companies in this market. Finally, the new legislation includes incentives for agricultural producers to buy crop insurance contracts in the form of premium subsidies.

The methodology developed in this article was used to forecast corn yields for selected counties in the State of Paraná using data covering 1990 through 2002. Using the posterior predictive criteria of Gelfand and Ghosh (1998), we chose from among several models appropriate for this forecasting and insurance pricing problem. The optimal model was used in the calculation of premium rates for insurance coverage based on regional yield indexes. Our analysis considered not only the temporal aspect of yield movements but also the spatial correlation that exists between counties. The resulting spatio-temporal model is thus more flexible compared to other potential specifications that have been considered in the literature. In light of the rather small sample of data available, we demonstrate the sensitivity of premium rates to the yield observed in 2002. In particular, higher rates were found in the regions where yields were lower in this year. We discuss the potential application of our methods to the general problem of pricing insurance contracts for individual coverage. We note that, to the extent that sufficient data are available, these methods may be applicable to the problem of pricing crop insurance contracts with individual coverage. Future research will evaluate methods of pricing insurance contracts for individual yields using the methods developed in this analysis.

[Received February 2006;  
accepted January 2008.]

## References

- Anselin, L. 1988. *Spatial Econometrics: Methods and Models*. Boston: Kluwer Academic Publishers.
- Berger, J.O., and W. Strawderman. 1996. "Choice of Hierarchical Priors: Admissibility in Estimation of Normal Means." *Annals of Statistics* 24:931–51.
- Berger, J.O., V. Oliveira, and B. Sanso. 2001. "Objective Bayesian Analysis of Spatially Correlated Data." *Journal of the American Statistical Association* 96:247–67.
- Bernardinelli, L., D. Clayton, and C. Montomoli. 1995a. "Bayesian Estimates of Disease Maps: How Important Are Priors?" *Statistics in Medicine* 14:2411–31.
- Bernardinelli, L., D. Clayton, C. Pascutto, C. Montomoli, M. Ghisland, and M. Songini. 1995b. "Bayesian Analysis of Space-time Variation in Disease Risk." *Statistics in Medicine* 14:2433–43.
- Besag, J. 1974. "Spatial Interaction and the Statistical Analysis of Lattice Systems." *Journal of the Royal Statistical Society: Series B* 36:192–236.
- Besag, J., J. York, and A. Mollié. 1991. "Bayesian Image Restoration, with Applications in Spatial Statistics." *Annals of the Institute of Statistical Mathematics* 43:1–59.
- Carlin, B.P., and T.A. Louis. 2000. *Bayes and Empirical Bayes Methods for Data Analysis*, 2nd ed. Boca Raton, FL: Chapman & Hall/CRC.
- Clayton, D.G., and J. Kaldor. 1987. "Empirical Bayes Estimates of Age-Standardized Relative Risks for Use in Disease Mapping." *Biometrics* 43:671–81.
- Cressie, N.A.C., and N.H. Chan. 1989. "Spatial Modeling of Regional Variables." *Journal of the American Statistical Association* 84:393–401.
- Day, R.H. 1965. "Probability Distributions of Field Crop Yields." *Journal of Farm Economics* 47:713–41.
- Dreassi, E. 2003. "Space-Time Analysis of the Relationship Between Material Deprivation and Mortality for Lung Cancer." *Environmetrics* 14:511–21.
- Gallagher, P. 1987. "U.S. Soybean Yields: Estimation and Forecasting with Nonsymmetric Disturbances." *American Journal of Agricultural Economics* 69:796–803.
- Gelfand, A.E., and S.K. Ghosh. 1998. "Model Choice: A Minimum Posterior Predictive Loss Approach." *Biometrika* 85:1–11.
- Gelfand, A.E., and A.F.M. Smith. 1990. "Sampling-Based Approaches to Calculating Marginal Densities." *Journal of the American Statistical Association* 85:398–409.
- Gelfand, A.E., S.K. Ghosh, J.R. Knight, and C.F. Sirmans. 1998. "Spatio-temporal Modeling of Residential Sales Data." *Journal of Business & Economic Statistics* 16:312–21.
- Gelman, A., J.B. Carlin, H.S. Stern, and D.B. Rubin. 2003. *Bayesian Data Analysis*. Boca Raton, FL: Chapman & Hall/CRC.
- Goodwin, B.K., and A.P. Ker. 1998. "Nonparametric Estimation of Crop Yield Distributions: Implications for Rating Group-Risk Crop Insurance Contracts." *American Journal of Agricultural Economics* 80:139–53.
- Goodwin, B.K., and O. Mahul. 2004. "Risk Modeling Concepts Relating to the Design and Rating of Agricultural Insurance Contracts." Policy Research Working Papers (IBRD), no. 3392, Washington DC: World Bank (IRDB).

- Hobert, J.P., and G. Casella. 1996. "The Effect of Improper Priors on Gibbs Sampling in Hierarchical Linear Mixed Models." *Journal of the American Statistical Association* 91:1461–73.
- Hodges, J.S., B.P. Carlin, and Q. Fan. 2003. "On the Precision of the Conditionally Autoregressive Prior in Spatial Models" *Biometrics* 59:317–22.
- Innes, R.D., and G.C. Rausser. 1989. "Incomplete Markets and Government Agricultural Policy." *American Journal of Agricultural Economics* 74:915–31.
- Just, R.E., and Q. Weninger. 1999. "Are Crop Yields Normally Distributed?" *American Journal of Agricultural Economics* 81:287–304.
- Kass, R.E., and A.E. Raftery. 1995. "Bayes Factors." *Journal of the American Statistical Association* 90:773–95.
- Ker, A.P., and K. Coble. 2003. "Modeling Conditional Yield Densities." *American Journal of Agricultural Economics* 85:291–304.
- Ker, A.P., and B.K. Goodwin. 2000. "Nonparametric Estimation of Crop Insurance Rates Revisited." *American Journal of Agricultural Economics* 83:463–78.
- Laud, P.W., and J.G. Ibrahim. 1995. "Predictive Model Selection." *Journal of the Royal Statistical Society: Series B* 57:247–62.
- Miranda, M., J. Skees, and P. Hazell. 1999. "Innovations in Agricultural and Natural Disaster Insurance for Developing Countries." Working paper, Dep. of Agr., Envir. and Development. Econ., The Ohio State University.
- Moss, C.B., and J.S. Shonkwiler. 1993. "Estimating Yield Distributions with a Stochastic Trend and Nonnormal Errors." *American Journal of Agricultural Economics* 75:1056–62.
- Myers, R.J. 1988. "The Value of Contingency Markets in Agriculture." *American Journal of Agricultural Economics* 70:255–67.
- Nelson, C.H., and P.V. Preckel. 1989. "The Conditional Beta Distribution as a Stochastic Production Function." *American Journal of Agricultural Economics* 71:370–8.
- Ramírez, O.A. 1997. "Estimation and Use of a Multivariate Parametric Model for Simulating Heteroskedastic, Correlated, Nonnormal Random Variables: The Case of Corn Belt Corn, Soybean and Wheat Yields." *American Journal of Agricultural Economics* 79:191–205.
- Ramírez, O.A. S. Misra, and J. Field. 2003. "Crop-Yield Distributions Revisited." *American Journal of Agricultural Economics* 85:108–20.
- Sherrick, B.J., F.C. Zanini, G.D. Schnitkey, and S.H. Irwin. 2004. "Crop Insurance Valuation Under Alternative Yield Distributions." *American Journal of Agricultural Economics* 86:406–19.
- Spiegelhalter, D., A. Thomas, N. Best, and D. Lunn. 2003. *Winbugs User Manual*. Medical Research Council Biostatistics Unit: Cambridge, U.K.
- Sun, D., R.K. Tsutakawa, and P.L. Speckman. 1999. "Posterior Distribution of Hierarchical Models Using CAR(1) Distributions." *Biometrika* 86:341–50.
- Taylor, C.R. 1990. "Two Practical Procedures for Estimating Multivariate Nonnormal Probability Density Functions." *American Journal of Agricultural Economics* 72:210–7.
- Turvey, C., and C. Zhao. 1999. "Parametric and Nonparametric Crop Yield Distributions and their Effects on All-risk Crop Insurance Premiums." Working paper, Dept. of Agricultural Economics and Business, University of Guelph, Ontario.
- Waller, L.A., B.P. Carlin, H. Xia, and A.E. Gelfand. 1997. "Hierarchical Spatio-Temporal Mapping of Disease Rates." *Journal of the American Statistical Association* 92:607–17.