Dynamic Modeling of Crop Yield Distributions: Implications for Crop Yield Insurance

Wenbin Wu, Ximing Wu, David J. Leatham

Department of Agricultural Economics

Introductory

The federal crop insurance program has been an important part of U.S. agricultural policy to stabilize farmers' income and protect against unpredictable risks for several decades. Among the policies provided by the federal crop insurance program, area-based policies are one of the most popular kinds of crop insurance policies, most of current research is also conducted at area (county) level. In order to design an insurance contract or estimate the premium rate of the area-based yield insurance policy, reliable modeling of crop yield distributions is of fundamental importance. In the literature, a two-stage procedure for crop yield risk modeling is the most widely accepted method. In the first stage regression analysis is applied to detrend the yield series, and in the second stage it is assumed the residuals follow a certain type of distribution, parametrically or nonparametrically. Most of researchers in this field center on the appropriate shape of the yield distribution, trying to capture the potential skewness or kurtosis in the distribution. Even though the suitable assumption of yield distribution is important, there is another important assumption people usually ignore or take for granted. This important assumption is that the yield distribution is static, which means the shape of yield distribution will not change over time (only location of the distribution will change). If this assumption does not hold, then the implementation of two-stage procedure is doubtable. However, this vital assumption has been discredited due to the progress of technology and change of environment we have seen in the past decades. Instead of assuming the yield distribution is static, we allow the yield distribution evolves over time in this paper.

Methodology

In this research we propose a dynamic modeling framework for crop yield distribution based on Gaussian Processes. Gaussian processes are one powerful non-parametric machine learning tool for regression and classification, and it has also been widely used in time series forecasting. Formally, a Gaussian process generates data from some domain such that any finite subset of the range follows a multivariate Gaussian distribution, in particularly arbitrary point will follow a Gaussian distribution. We utilize Gaussian processes as a distribution estimation and prediction tool for crop yield distribution. In this research we propose two estimators based on Gaussian processes. The first one is a Gaussian processes estimator and the second one is a Bayesian weighted Gaussian processes estimator. For the first estimator, we estimate the crop yield series using Gaussian Processes and then treat the estimated distributions at different time points as crop yield distributions at according times. For the second estimator, we utilize the Bayesian model averaging method to improve the estimation accuracy of Gaussian Processes estimators and to allow for more flexible tail behaviors of the estimated distributions.

Data and Results

In this paper we use corn yield observations from Iowa, which is the largest corn producing state in United States. Our data consists of annual county corn yields data of 99 Iowa counties from year 1960 through 2010, obtained from the National Agricultural Statistics Service.

We design several simulations to evaluate the effectiveness of proposed dynamic Gaussian Processes based estimators. To fully investigate the performance of our dynamic estimators, we implement two scenarios of simulations. The first scenario we assume the underlying ``true'' crop yield distribution is static, which will not change through time. The second scenario we assume the underlying ``true'' crop yield distribution is dynamic, which will have different shapes as time evolves. We refer these two scenarios as ``static scenario'' and ``dynamic scenario''. In our simulations, we consider three sample sizes: T=20, 30, 40. We focus on the Mean Squared Error (MSE) of crop insurance premiums for different estimators from their ``true'' values. Simulation results show better performance of our proposed dynamic estimators compared to conventional estimators.

In the last step we implement our proposed method in the real U.S. insurance market, we conduct an out-of-sample crop insurance rating game.

The potential for generating discussion during the Symposium

Crop yield distribution modeling has always been a popular topic in agricultural economics, is important for estimation of crop insurance premiums and farm risk-management decisions. In this paper we propose a novel dynamic framework to estimate crop yield distributions. Simulation and application results show promising properties of our proposed estimators compared to conventional estimators.