Estimating non-additive within-season temperature effects on maize yields using Bayesian approaches

Jisang Yu* and Gyuhyeong Goh+

*Kansas State University, Dept. of Agricultural Economics
+Kansas State University, Dept. of Statistics

SCC-76 2019 Annual Meeting

Background

- Many studies documented negative impacts of extreme heat on crop yields: e.g. Schlenker and Roberts (2009)
- Only several studies focus on within-season temperature variability: e.g. Ortiz-Bobea et al. (2018); Butler and Huybers (2015); Tack et al. (2015)
- We attempt to contribute to this literature by proposing estimation approaches that are more suitable in the context of high-dimensional data.

Key Research Questions

- Are detrimental heat impacts additive within growing season?
- If not, what are the implications on the warming effects?
- What are the implications on the cost of crop insurance?
  e.g. Tack et al. (2018); Perry et al. (2017)

Model Specifications

Table: Four alternative specifications (Dep. Var = ln $yield_{it}$)

Models	Explanatory Variables
M1	Growing season avg. Growing Degree Days (GDD) and Heating Degree Days (HDD), quadratic growing season avg. precipitation, quadratic state-specific time trends, and county fixed effects
M2	Growing season avg. GDD and HDD, quadratic growing season avg. precipitation, quadratic state-specific time trends, and county fixed effects + Interaction terms of GDD and HDD with quadratic precipitation variables
M3	Monthly GDDs and HDDs, quadratic monthly precipitation, quadratic state-specific time trends, and county fixed effects
M4	Monthly GDDs and HDDs, quadratic monthly precipitation, quadratic state-specific time trends, and county fixed effects + Interaction terms of GDD and HDD with quadratic precipitation variables (64 weather-related variables)

High-dimensional Data

Penalized Regressions versus Bayesian Approaches

- The OLS estimates often lead to poor estimation and prediction accuracy with a large number of explanatory variables (Tibshirani, 1996). M4 has 64 weather-related variables.
- One of the alternatives is to use penalized least squares (PLS): the determination of tuning parameters, which control the degree of the sparsity, is a big challenge.
- In a Bayesian framework, the tuning parameter selection problem can be resolved by integrating out the tuning parameter through Markov Chain Monte Carlo method (Narisetty and He, 2014).

Bayesian Variable Selection and Bayesian Modeling Average

Let γ be one of the candidate models. Bayesian variable selection (BVS) can be done by finding the highest posterior probability of γ:

$$p(\gamma|\text{data}) = \frac{p(\gamma)\int\int f(y|\beta_\gamma,\sigma^2)p(\beta_\gamma,\sigma^2)d\beta_\gamma d\sigma^2}{\sum_\gamma p(\gamma)\int\int f(y|\beta_\gamma,\sigma^2)p(\beta_\gamma,\sigma^2)d\beta_\gamma d\sigma^2}.$$

To address the uncertainty associated with the estimated model $\hat{\gamma}$, Bayesian model averaging (BMA) uses

$$p(\beta,\sigma^2|\text{data}) = \sum_\gamma p(\beta,\sigma^2|\text{data},\gamma)p(\gamma|\text{data}).$$

Dealing with computational issues: We use MC^3.

Data

- We use the USDA NASS corn yield data from corn belt counties in Iowa, Illinois, and Indiana for the period of 1989 - 2014.
- We use the weather data from the Parameter-elevation Regressions on Independent Slopes Model (PRISM).
- Similar to Schlenker and Roberts (2009), using the minimum and maximum temperatures from the PRISM data, we approximate a distribution of temperatures for each day based on a sinusoidal curve of Snyder (1985).
- And then, we calculate the growing degree days (GDDs) and the heating degree days (HDDs) for each month.

Recall that the candidate models are...

Table: Four alternative specifications (Dep. Var = ln $yield_{it}$)

Models	Explanatory Variables
M1	Growing season avg. Growing Degree Days (GDD) and Heating Degree Days (HDD), quadratic growing season avg. precipitation, quadratic state-specific time trends, and county fixed effects
M2	Growing season avg. GDD and HDD, quadratic growing season avg. precipitation, quadratic state-specific time trends, and county fixed effects + Interaction terms of GDD and HDD with quadratic precipitation variables
M3	Monthly GDDs and HDDs, quadratic monthly precipitation, quadratic state-specific time trends, and county fixed effects
M4	Monthly GDDs and HDDs, quadratic monthly precipitation, quadratic state-specific time trends, and county fixed effects + Interaction terms of GDD and HDD with quadratic precipitation variables (64 weather-related variables)

Out-of-sample Prediction Performances

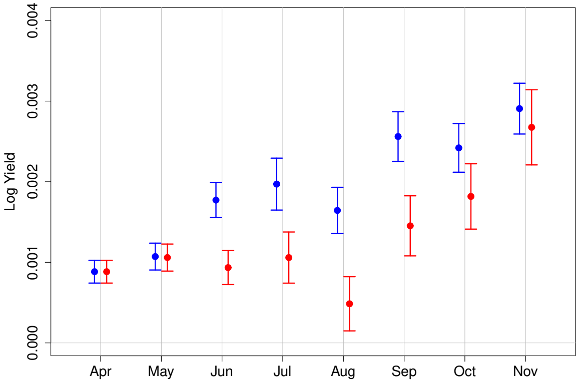
Table: Out-of-sample prediction performances

Specifications	RMSE	MAPE	PCC
M1 - OLS	0.2481	0.0365	0.7480
	(34.14)	(32.41)	(61.80)
M2 - OLS	0.2816	0.0426	0.7486
	(25.24)	(21.11)	(61.93)
M3 - OLS	0.2136	0.0323	0.7850
	(43.30)	(40.19)	(69.80)
M4 - OLS	0.1976	0.0286	0.7946
	(47.54)	(47.04)	(71.88)
M4 - BVS	0.1911	0.0283	0.8100
	(49.27)	(47.59)	(75.21)
M4 - BMA	0.1905	0.0282	0.8111
	(49.43)	(47.78)	(75.45)

Note: Changes compared to the model without weather variables (RMSE=0.3767, MAPE=0.0540, PCC=0.4623) are reported in parenthesis (% reductions for RMSE and MAPE, % increases for PCC)

Conditional Marginal Effects

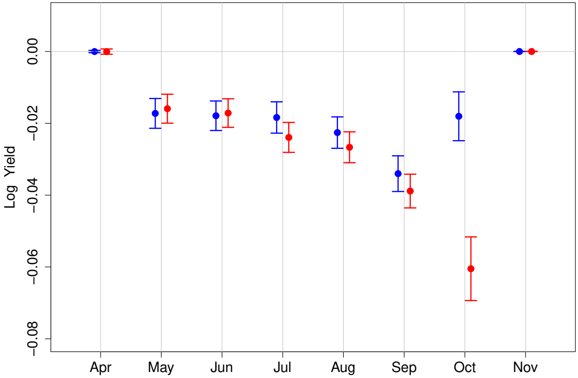
M4 - BMA, Growing Degree Days



Note: Blue represents the marginal effects at 75% percentile of precipitation and Red represents the marginal effects at 25% percentile of precipitation.

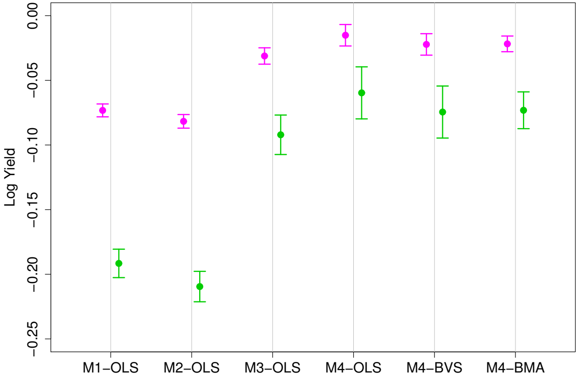
Conditional Marginal Effects

M4 - BMA, Degree Days above 29°C



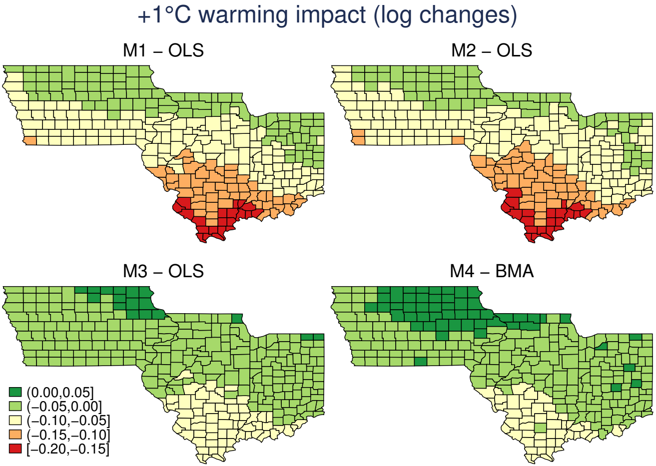
Note: Blue represents the marginal effects at 75% percentile of precipitation and Red represents the marginal effects at 25% percentile of precipitation.

Uniform Warming Effects



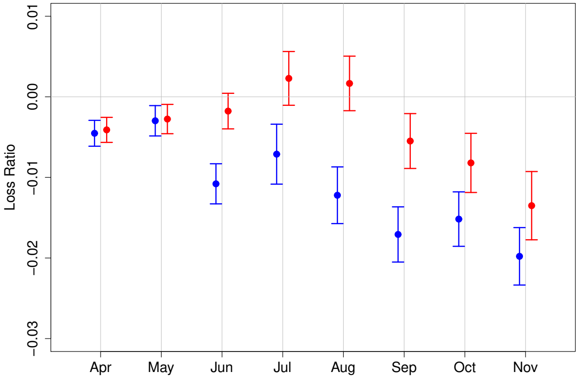
Note: Pink represents + 1°C and Green represents + 2°C.

Geographic Heterogeneity



Application to Crop Insurance Loss Ratios

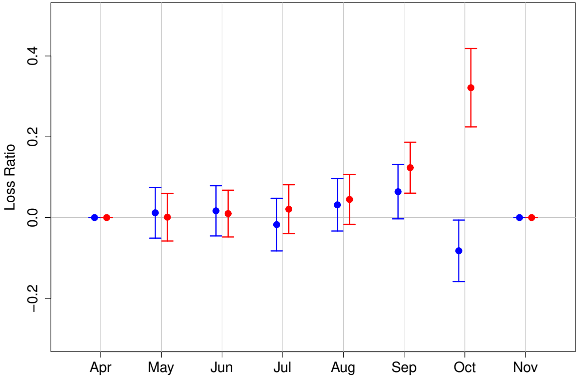
M4 - BMA, Growing Degree Days



Note: Blue represents the marginal effects at 75% percentile of precipitation and Red represents the marginal effects at 25% percentile of precipitation.

Application to Crop Insurance Loss Ratios

M4 - BMA, Degree Days above 29°C



Note: Blue represents the marginal effects at 75% percentile of precipitation and Red represents the marginal effects at 25% percentile of precipitation.

Remaining Questions and Next Steps

- What is the appropriate/efficient level of aggregation time window?
- What are the implications on climate change adaptation?
- Spatial correlations
- Warming impacts on the cost of crop insurance programs (Analyze using Loss-cost ratios)

References I

Butler, E. E. and P. Huybers (2015). Variations in the sensitivity of us maize yield to extreme temperatures by region and growth phase. *Environmental Research Letters 10*(3), 034009.

Narisetty, N. N. and X. He (2014, 04). Bayesian variable selection with shrinking and diffusing priors. *The Annals of Statistics 42*(2), 789–817.

Ortiz-Bobea, A., H. Wang, C. Carrillo, and T. Ault (2018). Intra-seasonal soil moisture and temperature effects indicate severe vulnerability of U.S. crop yields under climate change. *Working Paper*.

Perry, E., J. Yu, and J. B. Tack (2017). Estimating temperature effects on the cost of the federal crop insurance program.

References II

Schlenker, W. and M. J. Roberts (2009). Nonlinear temperature effects indicate severe damages to us crop yields under climate change. *Proceedings of the National Academy of Sciences 106*(37), 15594–15598.

Snyder, R. L. (1985). Hand calculating degree days. *Agricultural and forest meteorology 35*(1-4), 353–358.

Tack, J., A. Barkley, and L. L. Nalley (2015). Effect of warming temperatures on us wheat yields. *Proceedings of the National Academy of Sciences*, 201415181.

Tack, J., K. Coble, and B. Barnett (2018). Warming temperatures will likely induce higher premium rates and government outlays for the us crop insurance program. *Agricultural Economics 49*(5), 635–647.

References III

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society. Series B (Methodological)*, 267–288.