Presentation type: Poster

**Evaluating planting date recommendation effects on soybean yield using causal inference**

Santosh Sanjel1, Denis A. Shah2, Shawn P. Conley3, Spyridon Mourtzinis3 and Paul Esker1

1Department of Plant Pathology and Environmental Microbiology, The Pennsylvania State University, State College, Pennsylvania, USA

2Department of Plant Pathology, Kansas State University, Manhattan, Kansas, USA

3Department of Agronomy, University of Wisconsin–Madison, Madison, Wisconsin, USA

Email: [sks7388@psu.edu](mailto:sks7388@psu.edu)

Selecting the optimal planting date is a crucial decision for soybean farmers aiming to maximize yield. Planting date recommendations are usually made through several years of field research data and modeling. This study introduces an observational causal inference framework to assess the impact of adhering to early planting recommendation in the North Central United States (NCUS) on soybean yield gains. A grower-derived field survey collected from approximately 5,000 farms in the NCUS over three seasons (2014, 2015 and 2016) was used as the input dataset. A causal graph of the soybean farming system was developed based on agronomic and crop production practices in current use in the NCUS. Based on early planting recommendations made by a group of soybean researchers from 11 different universities in the NCUS, we assigned the observed planting date in the data into optimal or non-optimal planting (a binary categorization, taken as the planting date ‘treatment’). Using the backdoor criterion, we identified the minimal sufficient adjustment set of covariates accounting for confounding. Propensity score (PS) full matching was done using a probit regression of the treatment on the covariates; checks indicated that adequate balance was achieved. A weighted (by the full matching weights) linear regression model with yield as outcome, and the binary planting date treatment and covariates as predictors was fitted. Performing g-computation and using cluster-robust variance, we estimated the average treatment effect (ATE) of 270 kg/ha and 95% CI of 185 to 363, for fields adhering to recommended early planting dates. Overall, this study presents a comprehensive observational causal inference framework that is beyond the scope of predictive accuracy with potential applications in decision support systems across various fields. Future research aim is to develop causal models interpreting how planting dates may influence soybean disease incidence and severity, as well as management requirements.