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, KS

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 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 North Central US, we assigned the observed planting date of the survey data into optimal or non-optimal planting. Using the backdoor criterion, we identified the minimal sufficient adjustment set of covariates accounting for confounding. Propensity score (PS) matching was conducted to estimate the average marginal effect of adhering to early planting recommendations on yield increase. The PS estimated using a probit regression of the treatments on the covariates yielded an adequate balance. A linear regression model with yield as outcome, and treatment and covariates as predictors including the full matching weights, was used for estimation. Results revealed a significant yield increase of 270 kg/ha (SE = 45.52, p < 0.001) in 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 aims to develop causal models interpreting how planting dates may influence soybean disease incidence and severity, as well as management requirements.