Days Suitable for Fieldwork in the US Corn Belt: Climate, Soils and Spatial Heterogeneity

Benjamin M. Gramig

Associate Professor, Purdue University | bgramig@purdue.edu

Seong Do Yun

Postdoctoral Research Associate, Yale University | seongdo.yun@yale.edu

Selected paper prepared for presentation at the 2016 Agricultural & Applied Economics Association Annual Meeting, Boston, Massachusetts, July 31-Auguest 2, 2016.

Days Suitable for Fieldwork in the US Corn Belt: Climate, Soils and Spatial Heterogeneity

Benjamin M. Gramig and Seong Do Yun

Abstract

Days suitable for field work (DSFW) is an important piece of data for production agriculture and agricultural extension focused on practical decision making about investment in farm machinery and cropping systems management. It is, however, noteworthy that there has been limited attention paid to DSFW. To fill this gap, this study tries to answer two research questions: (1) what is the trend in DSFW during the planting and harvest period from 1980-2010? (2) what is the accuracy of a predictive econometric model of DSFW based on agro-environmental data? To tackle the economic dimensions of DSFW, we model DSFW consistent with two major approaches in climate change impacts on agriculture: the Ricardian approach and the panel estimation approach. We first specify the regression model of DSFW in panel model from two conceptual approaches: the response function and the factor demand function of cost minimization. Both approaches provide consistent regression specification of fixed and random effects models. We construct an unbalanced panel of weekly DSFW observations, historic weather data, and soil data in five Corn Belt States for 1980-2010 at the Crop Reporting District (CRD) level to implement out-of-sample and in-sample prediction analysis. The results show that the random effects model is the most suitable model to perform climate change response analysis for our data. This paper contributes to the literature in three ways. First, the analytical derivation of two econometric interpretations of DSFW and link them to econometric model specification strategies are easily extended to other agro-environmental analysis. Second, the estimation results for panel models empirically demonstrate that random effects model could be proper model specification taking into account soil effects. Lastly, we discuss that DSFW could be an important constraints for policy corresponding to climate change and its adaptation.

Key words: Days Suitable for Fieldwork, Soil Trafficability, Heterogeneity, Panel Estimation Approach

JEL codes: Q10, Q15, Q54,

Introduction

Days suitable for fieldwork (DSFW) is defined as the number of days in a week that soil moisture conditions allow farmers to perform work (e.g. tillage, planting, side-dressing fertilizer, harvest) in agricultural fields. The USDA National Agricultural Statistics Service (NASS) collects this data weekly during the annual agricultural growing season from spring through the time when harvest is completed each autumn, with the exact weeks varying across states due to differences in climatology and annual weather experience that may affect planting dates or harvest.[1] The DSFW in a given week are determined by weather, soil characteristics and drainage, and antecedent soil moisture conditions. This is a particularly important piece of information for production agriculture and agricultural extension focused on practical decision-making about farm machinery investment and different facets of cropping systems management. Days suitable are also an important agro-environmental measure because they influence decisions about tillage practices, tile drainage installation and fertilizer application timing that influence water quality, are influenced by extreme rainfall events, and are relevant to adapting farm management to climate change.

Days suitable for fieldwork are a constraint on timely field operations that directly influence farm profits in years when planting is delayed to the point where potential yield is reduced, the cost of grain drying increases because field dry down periods before harvest are shortened, or physiological maturity of a crop is delayed preventing harvest that maximizes crop quality. There has been limited attention paid to DSFW in the peer-reviewed literature with most research focused on a basic understanding of the usefulness of these data for farm management

[1] By the NASS, a suitable day is defined as "one where weather and field conditions allow producers to work in fields a major portion of that day." A number from 0-7 should be entered for the number of days suitable for field work for the past week (Monday through Sunday).

and planning (Griffin, 2009), or the applied use of historical DSFW data to support farm machinery sizing (Rotz, Muhtar, & Black, 1983; Rotz & Harrigan, 2005) and more general farm management optimization (Doster et al., 1983; Dillon, Mjelde, and McCarl, 1989; Etyang et al, 1998) in different production systems and locations. The practical importance of DSFW is evidenced by extension publications (e.g. Parsons and Doster, 1980; Massey, Carpenter, and Gerit, 2007; Edwards, 2015), and regional (Griffin, 2016) and national (USDA-NASS, 2016) online data resources devoted to farm management decision support.

This research compiles a novel data set using historical data on DSFW and weekly weather conditions at the geographic scale of a single crop reporting district (CRD) that takes key soil characteristics into account to estimate a predictive econometric model of this important farm management variable. Most states only report DSFW data at the state level, but, given the large spatial variation in precipitation, this scale of aggregation is of limited usefulness for a farmer in a single location that may not be very correlated with the statewide mean DSFW. The objective is to quantify the relationship between key agro-climatic variables and the suitability of field conditions for farm work. A statistical approach is developed and subjected to predictive performance tests with the aim of understanding the relationship between DSFW and widely available weather variables and soil characteristics. This is distinct from prior research that takes a soil-water balance or process-based modeling approach to simulate soil moisture conditions in an individual location to estimate whether fields are too saturated to perform fieldwork. The advantages of the proposed approach are that it requires much less data than a detailed simulation model and that the necessary inputs to estimate DSFW are easily understood and obtained by researchers and practitioners alike.

The paper proceeds by describing the data and econometric methods, estimating an empirical model, and then performing an out-of-sample prediction performance analysis using the disaggregated data available from five states in the US Corn Belt region. The observed trends in DSFW over the period from 1980-2010 are reported for the states with CRD-level days suitable data. The estimated model parameters are combined with CRD-level variable data in the seven Corn Belt states that only report state level data to generate a historical data set containing fitted values based on the parameter estimates from the states with observed CRD-level days suitable data. An illustration of the use of the data to evaluate the economics and risk of split-Nitrogen fertilizer application follows.

Economic and Econometric Framework

We follow two parallel conceptual frameworks in the empirical climate change economics literature today to develop an econometric model of DSFW. We model DSFW as a response function (Schlenker and Roberts 2009) consistent with the panel estimation approach of Deschênes and Greenstone (2007, hereafter DG).[2] The fixed effects and random effects model specification strategies followed in the later sections are based on both approaches that have been well developed in a panel structure.

The response function approach has been followed in agricultural economics and has the benefit of clear intuition along the lines of a biological production function. The most familiar type of response function measures crop yield response (Boyer et al. 2013; Dixon et al. 1994; Schlenker and Roberts 2009; Tack et al. 2015) and assumes a direct relationship between crop

[2] This study adopts the model specification approach in DG (2007). We begin our discussion in the context of cost minimization rather than profit maximization behavior, and, thus, our approach is based on the duality of the derivations in DG (2007).

growth and agro-environmental factors like temperature, water availability, and soil fertility conditions. From an economic viewpoint, yield response is an output from a (not necessarily explicit) combination of management and inputs together with agro-environmental conditions. For example, Motamed et al. (2016) adopt a corn acreage response function as a measure of ethanol biorefining capacity. To model DSFW as a response function, it is helpful to keep in mind the USDA-NASS definition of a suitable day[3]: "One where weather and field conditions allow producers to work in fields a major portion of day." This definition can be written as a function *y* = *f* (agro-environmental factors), where *y* is a suitable day and the agro-environmental factors are temperature, precipitation, and soil conditions. This conceptual model of DSFW does not model the agricultural output yield itself, but does have explanatory variables in common with yield response.

Following the specification strategy suggested in Schlenker and Roberts (2009), we can specify a DSFW response function as a fixed effects panel model:

(1) $y_{it} = \boldsymbol{X}_{it}\boldsymbol{\beta} + \boldsymbol{Z}_{it}\boldsymbol{\delta} + c_i + \varepsilon_{it},$

where y_{it} is DSFW of i-th region at time t, $\boldsymbol{X}_{it}$ is a vector of agro-environmental factors, $\boldsymbol{Z}_{it}$ is a vector of other factors including time trend and seasonal dummies to capture intra-seasonal variation, c_i is region-specific variation, $\boldsymbol{\beta}$ and $\boldsymbol{\delta}$ are response parameters, and ε_{it} is a random disturbance term with $\varepsilon_{it} \sim N(0, \sigma^2)$. Region-specific fixed effects in equation (1) are assumed to capture soil conditions in Schlenker and Roberts (2009), such that equation (1) does not include any soil variables in their model. Though soil is a fixed factor, soils in different locations are in fact similar and may be correlated. It is important to note that in applying this model to DSFW, the time-step *t* between each observation at location *i* is one week, such that the time trend would

[3] The definition comes from the survey instructions of USDA NASS: https://cpcsweb.nass.usda.gov/html/cpcs.instruction.html [Accessed on Jan. 19. 2016].

capture weekly human-contributed factors that influence day suitable rather than the standard interpretation of a trend as reflecting technological progress that occurs over a longer time scale; the time trend then should be annual to capture any trend over the 30 year timespan of the weekly DSFW data so that the seasonality and any long term trend are taken into account.

The second alternative following DG (2007) would adopt a panel estimation approach and start by addressing issues of possible identification and specification in the fixed effects model of Equation (1) above. The fixed effects panel model of Equation (1) is based upon the assumptions of exogenous agro-environmental factors and independent effects of human-contributed factors on them. It is, however, well known that human contributions, e.g. irrigation and tillage, influence the input-output relationship inherent in agricultural production. This means that, by assuming the independence of human-contributed factors and agro-environmental factors, Equation (1) could be improperly specified. If DSFW is possibly determined by human economic behavior y with cost p, then we can write a farmer's cost minimization problem as:

$$\min_{L,K,y} wL + rK + py$$

$$\text{s.t. } Q = f(L, K, y, A),$$

where L is labor input and K is capital input with region-specific prices w and r. A denotes a vector of exogenous agro-environmental factors. If we know the proper production function, f, then we can find the optimized cost function $C^*(w, r, p, Q, y, A)$ and DSFW become a conditional factor demand. Following Shephard's Lemma, we can derive $y^* = \frac{\partial C^*}{\partial p} = y^*(w, r, p, Q, A)$.

To specify an econometric model of $y^* = y^*(w, r, p, Q, A)$, we note that the factor prices ($w, r,$ and p) and agricultural output (Q) vary across regions and are possibly correlated through factor markets. This leads us to introduce random effects terms in a panel model specification.

As previously mentioned, it is necessary to model soil properties as being fixed because they are exogenously given regional conditions. Taking all of this into account, we specify a random effects (RE) panel model based on the structure of Equation (1) as:

(2) $y_{it} = \boldsymbol{X}_{it}\boldsymbol{\beta} + \boldsymbol{Z}_{it}\boldsymbol{\delta} + s_i + u_i + \varepsilon_{it}$,

where all variables are the same as the FE panel model except for region-specific soil conditions s_i and.the i-th region-specific random heterogeneity component u_i. In the standard RE model, the variance-covariance of $\tau_{it} = u_i + \varepsilon_{it}$ becomes $var(\tau_{it}) = \sigma_{\varepsilon}^2 + \sigma_u^2$, $cov(\tau_{it}, \tau_{is})$ for $t \neq s$, and $cov(\tau_{it}, \tau_{js}) = 0$ for all t and s if $i \neq j$. Because it is rarely possible to control for region-specific agro-environmental factors with fixed effect terms alone, this study argues that the RE specification of Equation (2) is more reasonable to deal with (possibly correlated) unobserved spatial heterogeneity caused by region-specific factors.

The two model specifications in equations (1) and (2) assume a known functional form. However, the true data generating process of DSFW is unknown and there has not been much attention paid to this issue in the literature. From this standpoint, it is noteworthy that these two specifications have so consistently been the econometric base models used to analyze climate change impacts on agriculture and tend to be referred to as the Ricardian (or hedonic) approach following Medelsohn, et al. (1994) and the panel estimation approach (DG, 2007). Therefore, the suggested model is arguably free from the debate about specification issues following the aforementioned approached but not from identification issues (Dell et al. 2014). Since the purpose of this study is to describe the trend of DSFW and predictive performances, we focus on the model fit and prediction capability of the suggested models by implementing in-sample and out-of-sample simulations.

Data

To implement the empirical analysis, we construct an unbalanced panel of weekly DSFW observations and historic weather data in twelve Corn Belt States for 1980-2010 at the Crop Reporting District level. In these data, DSFW is only measured during the crop growing season, i.e. mainly from March to November. We use DSFW observations from the USDA-NASS, weather data developed by the Midwestern Regional Climate Center (MRCC), and the USDA-NRCS gridded Soil Survey Geographic (gSSURGO) database. Table 1 summarizes geographical ranges and availability of data.

-- Table 1 about here --

First of all, we have no geographical mismatch issue to use CRD-level weather and soil variables. Since the provided MRCC data is all matched up with the original DSFW, we had no problem to have CRD-level weather variable over all twelve states. The gSSURGO soil data has no temporal variation and it is originally 10m X 10m resolution rater data. So, only gSSURGO cells over agricultural areas are aggregated up to the CRD-level with area-weighted averages. To check state-level data validity, we calculate area-weighted averages of both weather and soil variables as well.

Unlike weather or soil variables, the original DSFW data has messy structure in geographical level as shown in Table 1. The group G1 is five states having CRD-level DSFW whereas the group G2 is seven states only providing state-level DSFW. If we use area weighted averages over the group G1, we can calculate state-level DSFW. However, the other direction to have CRD-level from the state-level for the group G2 is rarely figured out. Since state-level is generally too large to depict spatial heterogeneity of weather and soil variables, we fit the G2 data into the CRD-level by using the standard econometric methods and demonstrate its validity. To this end, we follow the three steps below.

First, we adjust DSFW data in Indiana for 4/1/1980 – 11/10/1994 as weekly numbers. During that period, Indiana collected DSFW with 10-day intervals. By the formula, (DSFW/10)×7, we construct the expected weekly DSFW. The rest of period in Indiana and other states is weekly measure. Second, we merge weekly historic weather data obtained from the MRCC with the constructed DSFW data. Since the DSFW data has two types of interval schemes, Sunday to Monday and Monday to Sunday, we merge the MRCC data with the exactly matched days-interval for each observations.). Lastly, we combine the area-weighted mean soil variables from the gSSURGO database into the final dataset used for estimation. All weighted area includes 30m resolution agricultural lands only.

Considering the model specifications in Equation (1) and (2), we summarize the descriptive statistics of the data used in estimation in Table 2.

-- Table 2 about here --

As described, 10-days interval DSFW in Indiana is adjusted to the expected weekly DSFW and thus, all DSFW in the group 1 is distributed between zero and seven. Since the geographical levels (G1: CRD vs. G2: state) between two groups are different, we cannot directly compare statistics in Table 2 across groups. It is, however, obvious that there exists variations within and between groups in all variables.

To construct CRD-level DSFW observations across all twelve Corn Belt states, we first estimate two specifications of Equation (1) and Equation (2) with unbalanced panel of the group G1. To select a better fit model between two specifications, we implement 1,000 out-of-sample prediction simulations to evaluate out-of-sample model performance. Comparing the root means squared errors (RMSE) of in-sample and out-of-sample simulation, we choose the random effects model of Equation (2) as the most suitable model to generate CRD-level DSFW values (fitted values, $\widehat{y_{it}}$) using less spatially aggregated agro-climatic variables for the historic period in the seven additional states that only report state-level DSFW.

Results

With the five states DSFW observation data in the group G1, we plot the 95% confidence intervals round the means across CRDs and times as shown in Figure 1.

-- Figure 1 about here --

The left panel of Figure 1 provides the changes of DSFW across CRD in the G1 group states. As expected, the trend of means and confidence intervals fluctuate greatly. This means there exists serious spatial CRD-level heterogeneity in the DSFW. In an empirical analysis, these results supports the fact that the aggregated averages, e.g., area weighted average, up to state-level can lose spatial heterogeneity due to the disappeared spatial variation in DSFW. The right panel of Figure 1 describes the changes of DSFW across time sequences. This graph also shows large time heterogeneity across time. These two, therefore, figures heterogeneity in space and time to be considered in our estimation process.

Table 3 shows the estimation results of Equation (1) and (2). The estimation results of pooled regression are also presented as a reference.

-- Table 3 about here --

From the spatial and temporal heterogeneity described in Figure 1, two-way fixed and random panel estimation is appropriate for the analysis. In the second column of fixed effect model

(Equation (1)), however, the space-specific one-way estimation is implemented to avoid the dummy variable trap. If we consider this is 45 CRDs by 1,173 weeks unbalanced panel model, including time-specific heterogeneity produces bad condition number in the demeaned estimation. In the last column of random effects model of Equation (2), two-way Wansbeek and Kapteyn variance components estimation is applied. Therefore, it contains both space and time heterogeneities. All estimates of three models show statistically significant and expected signs. The last two rows of R-squared and the RMSE show that the random effects specification is the best fit in in-sample prediction performances.

To select a better prediction model, we additionally implement out-of-sample prediction simulations. Since this model includes spatial heterogeneity affected by geographical locations and every year can have different lengths of weeks of growing seasons, we randomly select our samples as yearly base rather than complete random sampling. Among total 31 years, we sample 25 random years as base years of estimation and 6 (19.35%) years for prediction. We replicate this simulation 1,000 times and Figure 2 shows the boxplots of the RMSEs.

-- Figure 2 about here --

From Figure 2, we can conclude that the random effects model specification of Equation (2) has the better fit than the fixed effects model specification of Equation (1). This result is consistent with the in-sample prediction performance as well. Therefore, we select the Equation (2) as our base model to generate CRD-level DSFW in the seven states of G2 that report state-level DSFW only.

Using the estimation results of the random effects model in Table 3, we generate the predicted CRD-level DSFW for the group G2. To test robustness of the results, we calculate the sensitivity of state and CRD-level DSFWs for each G1 and G2 as shown in Figure 3.

-- Figure 3 about here --

For the group G1, the horizontal axis is the observed CRD-level DSFW and the vertical axis is area weighted state-level DSFW. In the group G2, the horizontal axis is the predicted CRD-level DSFW derived from the estimates of random effects model and the vertical axis is the observed state-level DSFW. The linear regression results of these two figures have the similar magnitudes on both intercept and slope. Therefore, we can infer that the predicted CRD-level DSFW in G2 is robust if G2 has the similar agro-environmental factor sensitivity with G1.

To analyze agro-environmental factor sensitivity in CRD-level DSFW, we draw boxplots of DSFW across different variable levels in Figure 4.

-- Figure 4 about here --

It is noteworthy that 25%-75% quantiles in G2 has shorter length as expected in predicted values. Importantly, the trend of the median in both G1 and G2 show reasonable patterns: DSFW increases as maximum or minimum temperature increases while it decreases as precipitation increase. In addition, their patterns in G1 and G2 are synchronized and we can conclude that the DSFW are reasonable prediction supported by the similarity of agro-environmental factor sensitivity. In case of drainage class, it varies very shorter ranges and reflects location-specific variation, we can only suppose G2 has lower drainage producing the lower median DSFW compared to the G1's. In the standpoint of out-bounds values of the predicted DSFW (great than 7 or less than 0 days) in G2, we can see that our predictions produce underestimation always and they are not serious because only a few number of out-bounding values are generated.

Figure 5 presents the averaged RMSE of G1 and G2 across each geographical levels. In the figure, the RMSE greater than 1 is assigned as 1 for a better legibility.

-- Figure 5 about here --

Overall, the RMSE levels are relatively small values between zero and one in both G1 and G2.

Figure 6 shows the CRD-level (averaged) observed predicted DSFW over all twelve Corn Belt states.

-- Figure 6 about here --

The western CRDs have longer DSFW than the eastern CRDs. We note that almost of Indiana observations are the expected weekly DSFW and thus, the presented DSFW is smoothed rather than other states. In addition, we find a statistically significant decrease [increase] in DSFW during the planting period in Illinois, Indiana and Iowa [Kansas], and a significant increase in DSFW during the harvest period in Illinois, Indiana, Iowa and Missouri.[4]

Conclusion

Despite of its importance, relatively less attention is paid to analyze DSFW in agro-climate and economics literature. To fill this gap, this paper derives a panel regression models consistent to the Ricardian approach and the panel estimation approach. Based on the response function framework, we specify a fixed effects model comparable to Schlenker and Roberts (2009). From the cost minimization context, we derive a random effects model that includes a different modeling rule on soil characteristics from the fixed effects model. Since both models are consistent with general agro-climatic models, the suggested model is arguably free from the debate about specification issues. With constructed DSFW dataset, we implement out-of-sample

[4]The same conclusion can be found from a blog post: Gramig, B.M. 2014. Days Suitable for Field Work in the Corn Belt. *AgriClimate Connection, http://sustainablecorn.org/blog/?p=477* [Accessed on 22 May 2016].

and in-sample prediction analysis and conclude that the random effects models have better explanation for our data.

This research contributes to the literature in three ways. First, the analytical derivation of two econometric interpretations of DSFW and link them to econometric model specification strategies are easily extended to other agro-environmental analysis. Second, the estimation results for panel models empirically demonstrate that random effects model could be proper model specification taking into account soil effects. Lastly, we discuss that DSFW could be an important constraints for policy corresponding to climate change and its adaptation.

In the next step of research, we will develop more abundant analysis and policy implications on climate change and its adaptation. This could include change of historical patterns in DSFW and future expectations. For this purpose, adopting various regional climate model (RCM) is under consideration.

Acknowledgements

The authors are grateful for funding support from the United States Department of Agriculture (USDA) National Institute for Food and Agriculture (NIFA) through competitive award no. 2011-68002-30220 to the Useful to Usable (U2U): Transforming Climate Variability and Change Information for Cereal Crop Producers (www.agclimate4u.org) project. Additional funding from the Purdue Research Foundation and the W-3133 Multi-State Hatch project supported this work. Ray Massey and Larry Biehl provided extensive help identifying anomalies and gaps in the DSFW data that contributed to the quality of the analysis and end product. Mike Timmins at the Midwest Regional Climate Center graciously provided historical CRD-level weekly weather data. Philip Owens and Edwin Winzeler provided the soil property data for initial analysis and the methodology followed to extract these data for the entire region. Historical DSFW data were compiled from state NASS offices prior to their consolidation into regional offices—without these offices' archives of hard copies of these data we could not have constructed the electronic data set that is the basis for this research. We thank Jeff Andresen, Chad Hart, Andy Higgins, Carson Reeling, Gary Schnitkey, Douglas Todey, and Ray Massey for their work with state NASS office personnel to obtain these data, in some cases entering the data from 30 years of weekly paper copies of the NASS Crop Progress and Condition report issued in their states.

References

Boyer, C. N., J. A. Larson, R. K. Roberts, A. T. McClure, D. D. Tyler, and V. Zhou. 2013. Stochastic Corn Yield Response Functions to Nitrogen for Corn after Corn, Corn after Cotton, and Corn after Soybeans. *Journal of Agricultural and Applied Economics* 45(4):669-681.

Dell, M., B. F. Jones, and B. A. Olken. 2014. What do We Learn from the Weather? The New Climate-Economy Literature. *Journal of Economic Literature* 52(3): 740-798.

Deschênes, O., and M. Greenstone. 2007. The Economic Impacts of Climate Change: Evidence from Agricultural Output and Random Fluctuations in Weather. *The American Economic Review* 97(1): 354-385.

Dillon, C., J. Mjelde, and B. McCarl. 1989. Biophysical simulation in support of crop production decisions: A case study in the blacklands region of Texas. *Southern Journal of Agricultural Economics 21*(1): 73-86.

Dixon, B. L., S. E. Hollinger, P. Garcia, and V. Tirupattur. 1994. Estimating Corn Yield Response Models to Predict Impacts of Climate Change. *Journal of Agricultural and Resource Economics* 19(1):55-68.

Doster, D. H., D. R. Griffith, J. V. Mannering, and S. D. Parsons. 1983. Economic returns from alternative corn and soybean tillage systems in Indiana. *Journal of Soil and Water Conservation 38*(6): 504-508.

Edwards, W. 2015. Farm Machinery Selection. Ag Decision Maker File A3-28, Iowa State University Extension. http://www.extension.iastate.edu/agdm/crops/html/a3-28.html

Etyang, M. N., P. V. Preckel, J. K. Binkley, and D. H. Doster. 1998. Field time constraints for farm planning models. *Agricultural Systems 58*(1): 25-37. doi:http://dx.doi.org/10.1016/S0308-521X(98)00029-8

Griffin, T. 2009. Acquiring and Applying Days Suitable for Fieldwork for your State. *Journal of the American Society of Farm Managers and Rural Appraisers* 35-42.

Griffin, T. 2016. Kansas Days Suitable for Fieldwork. AgManager, Kansas State University Extension. [Accessed on 22 May 2016] http://www.agmanager.info/farmmgt/machinery/FWD.asp

Massey, R., B. Carpenter, and S. Gerit. 2007. *Fieldwork Days and Machinery Capacity*. Retrieved from http://extension.missouri.edu/p/G363

Mendelsohn, R., W. D. Nordhaus, and D. Shaw. 1994. The Impact of Global Warming on Agriculture: A Ricardian Analysis. *The American Economic Review* 84(4): 753–771.

Motamed, M., L. McPhail, and R. Williams. 2016. Corn Area Response to Local Ethanol Markets in the United States: A Grid Cell Level Analysis. *American Journal of Agricultural Economics* (Advance Access).

Parsons, S. D. and D. H. Doster. 1980. *Days Suitable for Fieldwork in Indiana with Emphasis on Machinery Sizing*. Purdue University Extension, Station Bulletin No. 293. West Lafayette, IN.

Rotz, C. A., and T. M. Harrigan. 2005. Predicting Suitable Days for Field Machinery Operations in a Whole Farm Simulation 21(4): 563-571.

Rotz, C. A., H. A. Muhtar, and J. R. Black. 1983. A multiple crop machinery selection algorithm. *Transactions of the ASAE* 26(6): 1644-1649.

Schlenker, W. and M. J. Roberts. 2009. Nonlinear Temperature Effects Indicate Severe Damage to U.S. Crop Yields under Climate Change. *Proceedings of the National Academy of Sciences* (PNAS) 106(37): 15594-15598.

Tack, J., A. Barkley, and L. L. Nalley. 2015. Estimating Yeild Gaps with Limited Data: An Application to United States Weat. *American Journal of Agricultural Economics* 97(5): 1461-1477.

United States Department of Agriculture, National Agricultural Statistics Service (USDA-NASS). (2016) Days Suitable for Fieldwork and Soil Moisture. [Accessed 22 May 2016] https://www.nass.usda.gov/Charts_and_Maps/Crop_Progress_&_Condition/soilmap.php

Table 1. Data Availability

Group	FIPS	State	DSFW		Weather variables (tmax, tmin, prec)		Soil drainage class	
			CRD	State	CRD	State	CRD	State
G1	17	Illinois	O	◌	O	◌	◌	◌
	18	Indiana	O	◌	O	◌	◌	◌
	19	Iowa	O	◌	O	◌	◌	◌
	20	Kansas	O	◌	O	◌	◌	◌
	29	Missouri	O	◌	O	◌	◌	◌
G2	26	Michigan	X	O	O	◌	◌	◌
	27	Minnesota	X	O	O	◌	◌	◌
	31	Nebraska	X	O	O	◌	◌	◌
	38	North Dakota	X	O	O	◌	◌	◌
	39	Ohio	X	O	O	◌	◌	◌
	46	South Dakota	X	O	O	◌	◌	◌
	55	Wisconsin	X	O	O	◌	◌	◌

G1: States having both crop reporting district(CRD)-level and State-level DSFW data

G2: State having only State-level DSFW data

O: observed / X: unobserved / ◌: area-weighted average

Table 2. Summary Statistics of Dataset for DSFW Analyses

Group 1: 44,888 CRD-level weekly observations of five states				
Variables	Mean	S.D.	Min.	Max.
DSFW (days)	4.6567	1.8467	0.0000	7.0000
Maximum temperature (ºF)	74.5595	13.0504	11.9000	107.6000
Minimum temperature (ºF)	51.5682	12.5845	-1.8000	77.6000
Total precipitation (in)	0.8667	0.8857	0.0000	11.6300
A week lag of maximum temperature (ºF)	74.7176	12.7628	21.8000	107.6000
A week lag of minimum temperature (ºF)	51.6411	12.4402	2.4000	77.6000
A week lag of total precipitation (in)	0.8681	0.8811	0.0000	11.6300
(A week lag of total precipitation) x (minimum temperature)	45.9765	49.4556	-0.0180	785.4500
Yearly trend	16.6672	8.9870	1.0000	31.0000
Spring dummy	0.3410	0.4740	0.0000	1.0000
Fall dummy	0.2729	0.4454	0.0000	1.0000
Winter dummy	0.0122	0.1099	0.0000	1.0000
Drainage class	3.9149	0.5869	2.8235	5.1487
Group 2: 5,086 State-level weekly observations of seven states				
Variables	Mean	S.D.	Min.	Max.
DSFW (days)	4.8717	1.5023	0.0000	7.0000
Maximum temperature (ºF)	70.1027	13.9322	5.0078	99.5789
Minimum temperature (ºF)	46.1448	12.2751	-8.2493	70.3078
Total precipitation (in)	0.6549	0.5594	0.0000	4.6885
A week lag of maximum temperature (ºF)	70.3839	13.4250	11.0534	99.5789
A week lag of minimum temperature (ºF)	46.2854	11.9950	-4.3691	70.3078
A week lag of total precipitation (in)	0.6571	0.5572	0.0000	4.6885
(A week lag of total precipitation) x (minimum temperature)	31.8278	29.3684	-3.5960	269.3912
Yearly trend	19.6966	8.0392	1.0000	31.0000
Spring dummy	0.3266	0.4690	0.0000	1.0000
Fall dummy	0.2774	0.4478	0.0000	1.0000
Winter dummy	0.0012	0.0343	0.0000	1.0000
Drainage class	4.3360	0.5887	3.5159	5.0278

Table 3. Estimation Results

Variables	Pooled (1)		Fixed Eff. (2)		Random Eff. (3)	
Intercept	-1.3838	***	-0.5719	***	-5.5629	**
	(0.1423)		(0.1412)		(2.6180)	
Maximum temperature	0.2592	***	0.1862	***	0.2103	***
	(0.0084)		(0.0081)		(0.0087)	
Maximum temperature squared	-0.0037	***	-0.0021	***	-0.0024	***
	(0.0002)		(0.0002)		(0.0002)	
Minimum temperature	-0.2275	***	-0.2088	***	-0.2337	***
	(0.0086)		(0.0083)		(0.0089)	
Minimum temperature squared	-0.0040	***	-0.0024	***	-0.0025	***
	(0.0002)		(0.0002)		(0.0002)	
(Maximum x Minimum) temperature	0.0076	***	0.0048	***	0.0052	***
	(0.0003)		(0.0003)		(0.0003)	
Total precipitation	-1.2362	***	-1.2241	***	-1.1255	***
	(0.0139)		(0.0132)		(0.0129)	
Total precipitation squared	0.1298	***	0.1305	***	0.1139	***
	(0.0033)		(0.0031)		(0.0030)	
A week lag of maximum temperature	0.0313	***	0.0551	***	0.0550	***
	(0.0014)		(0.0014)		(0.0015)	
A week lag of minimum temperature	-0.0093	***	-0.0283	***	-0.0359	***
	(0.0016)		(0.0015)		(0.0017)	
A week lag of total precipitation	-1.1131	***	-1.1460	***	-1.0135	***
	(0.0291)		(0.0276)		(0.0281)	
A week lag of total precipitation	0.0093	***	0.0105	***	0.0087	***
x minimum temperature	(0.0005)		(0.0005)		(0.0005)	
Yearly trend	0.0174	***	0.0261	***	0.0215	***
	(0.0024)		(0.0023)		(0.0082)	
Yearly trend squared	-0.0001	**	-0.0004	***	-0.0003	
	(0.0001)		(0.0001)		(0.0002)	
Spring	-0.7126	***	-0.6163	***	-0.4537	***
	(0.0175)		(0.0170)		(0.0252)	
Fall	0.4499	***	0.5314	***	0.3748	***
	(0.0199)		(0.0195)		(0.0266)	
Winter	-1.2195	***	-0.8406	***	-0.4251	***
	(0.0534)		(0.0515)		(0.0547)	
Drainage class					3.0652	**
					(1.3190)	
Drainage class squared					-0.4332	***
					(0.1640)	
CRD Fixed Effects	NO		YES		NO	
# of Observations	44,888		44,888		44,888	
R-squared	0.6604		0.6958		0.5663	
Root mean squared error (RMSE)			1.0193		0.9989	

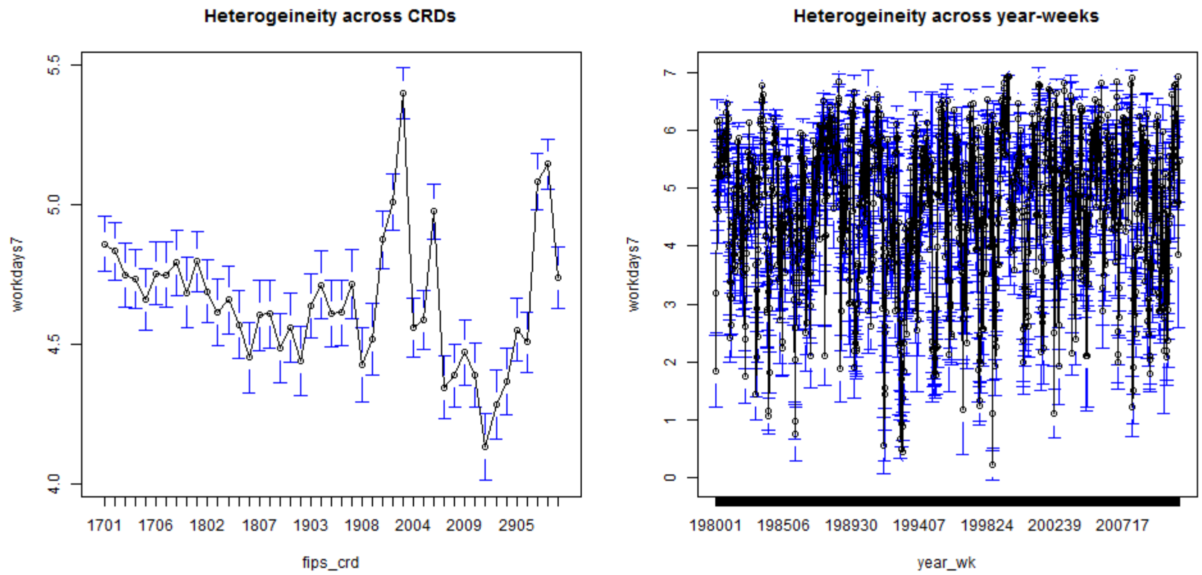


Figure 1. Heterogeneity Diagram

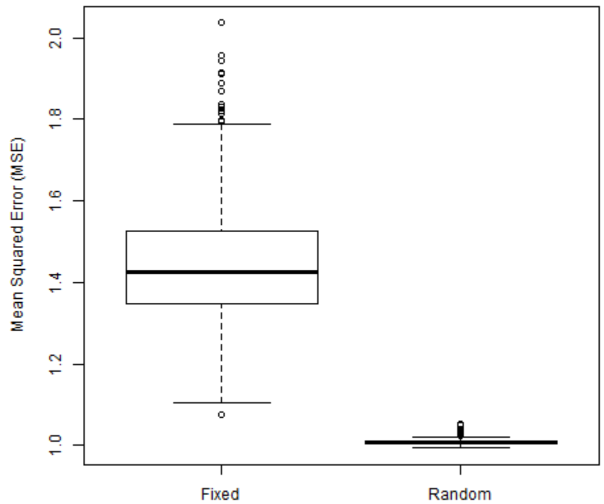
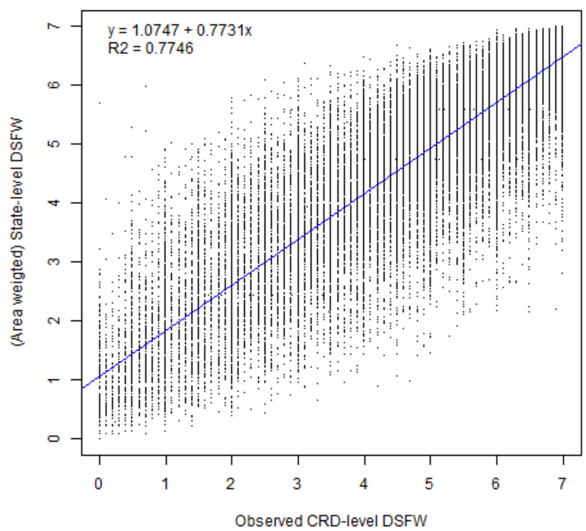
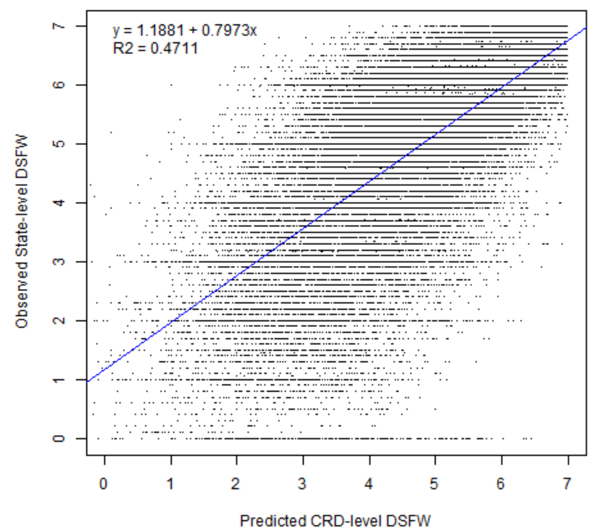


Figure 2. Mean Squared Error (MSE) of 1,000 Out-of-Sample Predictions

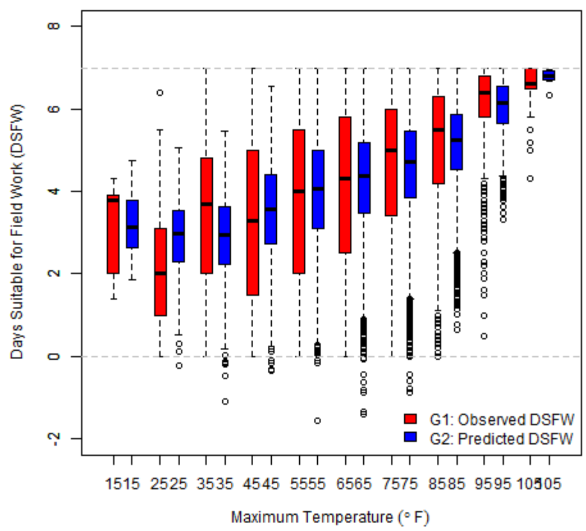


(a) Group1

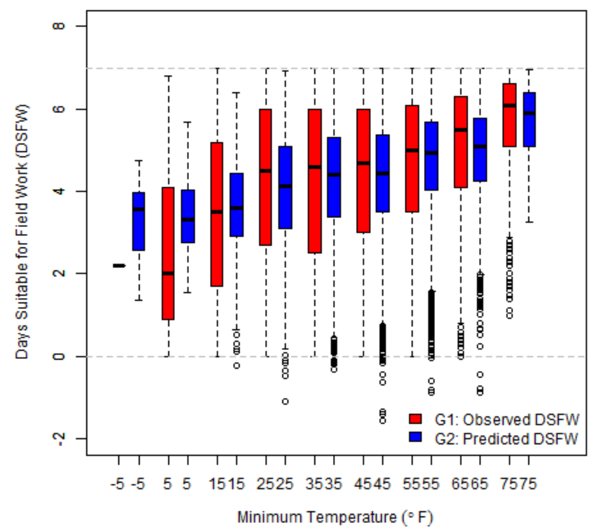


(b) Group 2

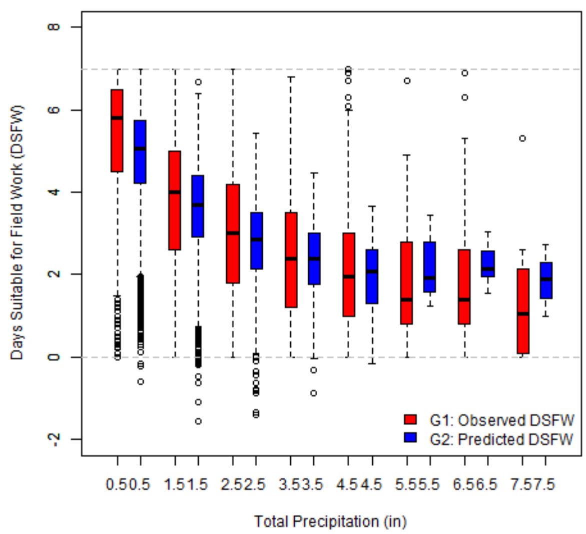
Figure 3: Aggregation Sensitivity Analysis: CRD-level vs. State-level



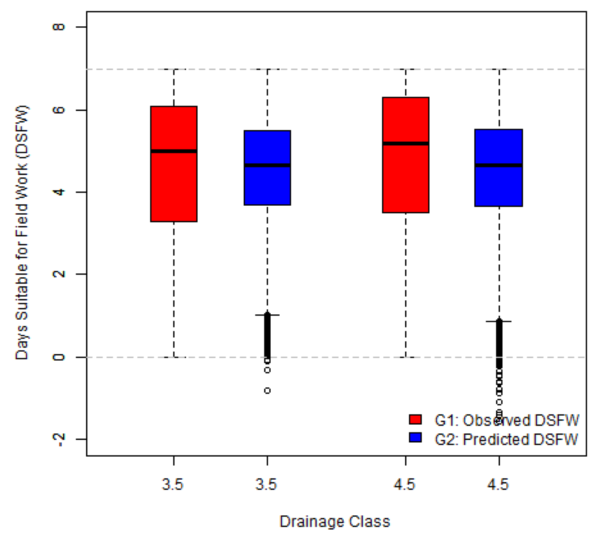
(a) Maximum Temperature



(b) Minimum Temperature

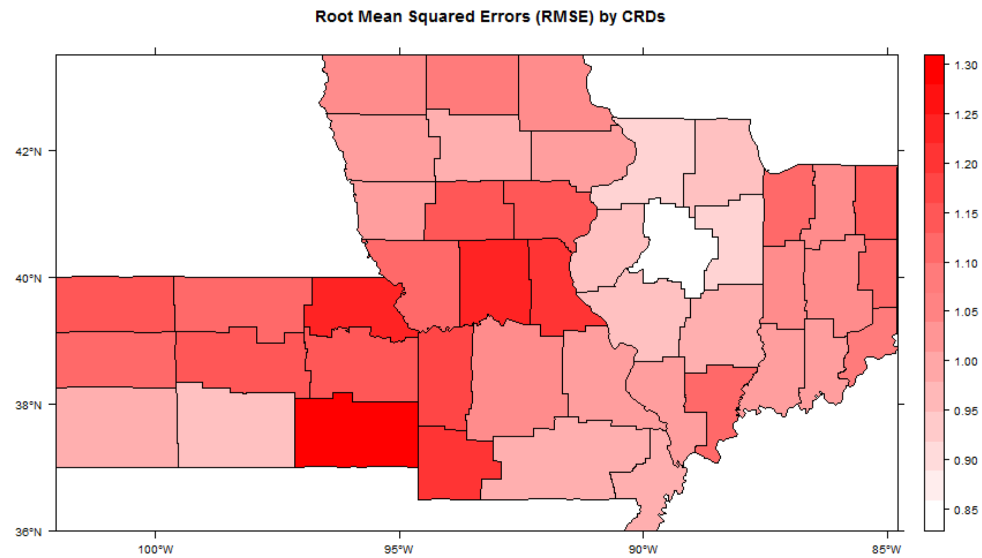


(c) Total Precipitation

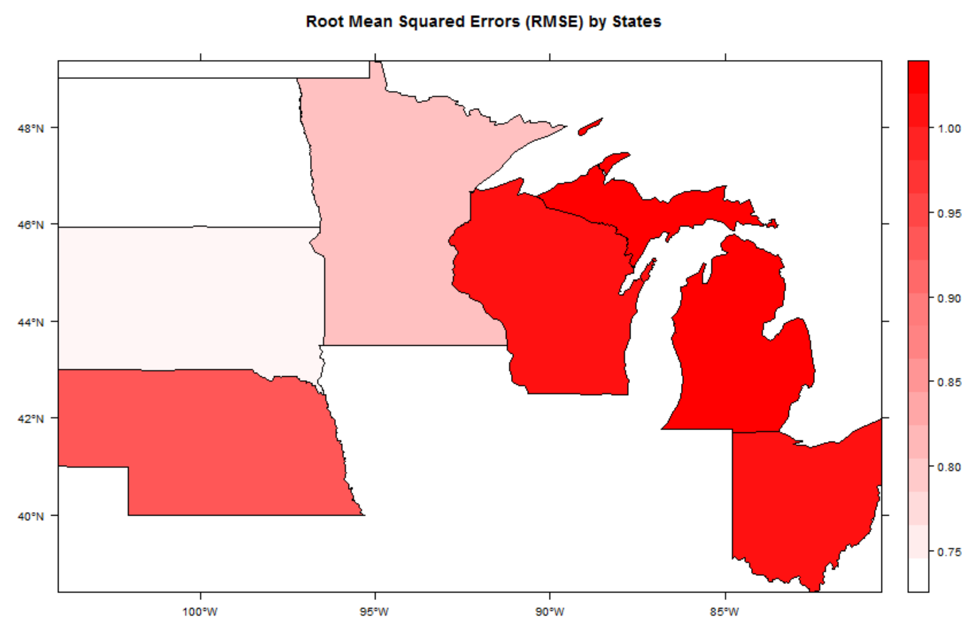


(d) Drainage Class

Figure 4. CRD-level Comparisons between Observed DSFW (Group 1) and Predicted DSFW (Group 2) across Variables



(a) Root Mean Squared Errors (RMSE) of Group 1: CRD-level



(b) Root Mean Squared Errors (RMSE) of Group 2: State-level

Figure 5. Prediction Errors by Groups: Averaged RMSE by Geographical Level

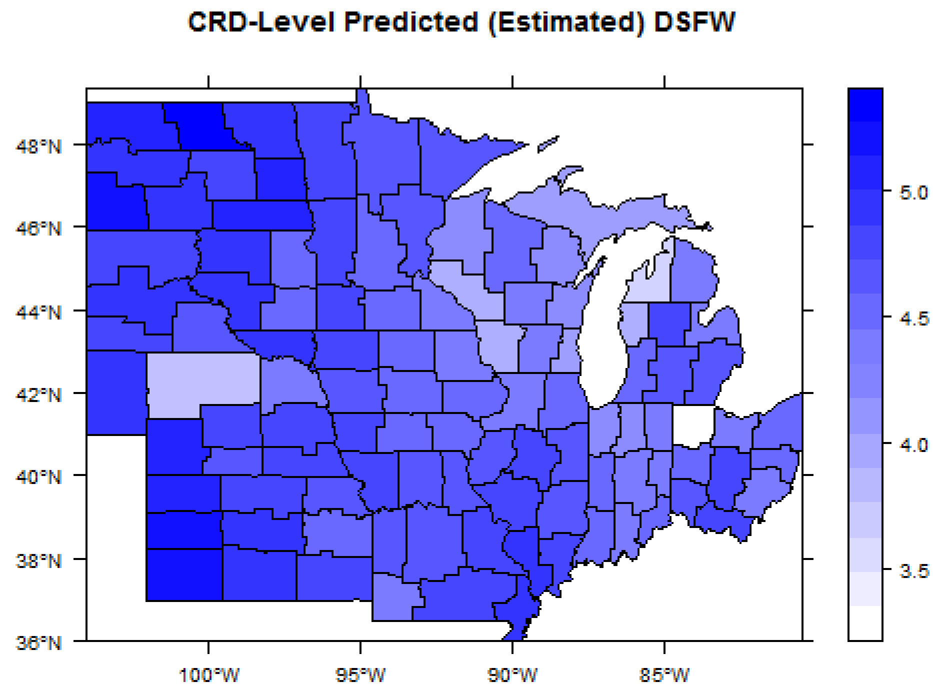


Figure 6. Predicted (Group1) and Estimated (Group2) DSFW: Averaged DSFW by CRD