The Not So Big Squeeze: Modeling Future Land Use Patterns of

Miscanthus for Bioenergy Using Fine Scale Data

Lunyu Xie, Sarah Lewis, Maximilian Auffhammer, Deepak Jaiswal, Peter Berck[1]

Date of this Draft: August 10, 2013

Abstract

We predict land coverage in the Midwest for a new bioenergy crop, *Miscanthus x Giganteus (miscanthus)*, using fine scale data and an econometric model. To explain farmers' current crop choices we use a local, limited dependent variable regression based on physical criteria. To this model we add miscanthus as a new crop, based on its place dependent *BioCro* model-predicted yield. We then predict where miscanthus will be grown given that it will be grown on five percent of the landscape. We find that 80 percent of the land that miscanthus will occupy is not currently occupied by food crops.

JEL Codes: Q15, Q42

Keywords: Biofuels, Land Use, Food Crops

The Not So Big Squeeze: Modeling Future Land Use Patterns of

Miscanthus for Bioenergy Using Fine Scale Data

1 Introduction

Concerns about energy security and climate change have led to a variety of federal and state level regulatory interventions designed to reduce the domestic consumption of liquid petroleum products. The two main regulatory strategies to achieve this goal have been to significantly strengthen corporate average fuel economy standards (CAFÉ) and ramping up the production of second generation biofuels via the renewable fuels standard (RFSII). This production standard envisions 39.7 billion liters of cellulosic ethanol produced in the US in 2020.[2]

If that standard were met solely with *Miscanthus x Giganteus* (hereafter miscanthus), it would take approximately 8 million hectares of crop land or 12.7 million ha of CRP land and 5 million ha of crop land (Scown, et al., 2012). The potential use of large quantities of land for biofuels raises questions about the effects of this policy on food prices and the effect on biodiversity. The magnitude of the impact on food prices is likely proportional to the amount crop land used for miscanthus instead of food crops. The main goal of this paper is to provide evidence, based on past farmer crop choices and projected miscanthus yields, on what types of lands the market will allocate to miscanthus.

Our paper fits into a crop choice literature discussed in the next section, which has largely used computable general equilibrium models. We use a novel econometric approach to model crop choice. We use an estimated proportion model using relevant weather, soil, and other data. The model is used to explain historical crop coverage in parts of six Midwestern states. We then add the bioenergy crop,

---

[2] Energy Independence and Security Act of 2007 sets the second Renewable Fuel Standard, known as RFSII. The standard includes years to 2020. However the standard has proved infeasible in the short term, see American Petroleum Institute v. EPA. The US Court of Appeals for the District of Columbia Circuit. January, 2013.

miscanthus, by simulating its yield with BioCro[3] and making its proportion on the landscape function of its yield in mega grams per hectare (Mg/ha). This method of examining the allocation of land to another crops builds on a long tradition of estimating crop coverage (e.g. Nerlove 1956) and on McFadden's (1974) method of adding new alternatives to a discrete choice model. This model predicts where farmers will put bioenergy crops on the landscape based upon previous choice data rather than on engineering-economic estimates of costs and profits.

Specifically, we impose a required average level of miscanthus coverage, five percent of the landscape, and use our model to find the percent of land allocated to major crops, pasture and wild land, and miscanthus before and after the biofuel mandate. To preview the findings, we project that 80 percent of the land that miscanthus will occupy is not currently occupied by food crops.

The remainder of the paper is organized as follows. Section 2 reviews related literatures. Section 3 summarizes the data on land use, soil conditions, weather, and miscanthus yield for the states along the Mississippi-Missouri river corridor. Section 4 describes estimation issues and establishes the econometric system. Section 5 presents the estimation results. Section 6 simulates the spatial distribution of miscanthus. Section 7 concludes.

2 Literature Review

In the economic literature, there are three main types of ways to estimating what type of land will be allocated to growing advanced bioenergy crops under RFSII:

- Computable General Equilibrium Models

- Modeled Profit based

- Past choice based models

---

[3] BioCro is a generic vegetation model based on the previously published WIMOVAC (Humphries & Long, 1995) as adapted for miscathus in Miguez et al. (2009, 2012).

All three share the same underlying logic that land is allocated at least partially based on profits and all three model miscanthus (or other bioenergy crops) based on limited current information.

The computable general equilibrium (CGE) approach considers the whole economy and makes all outputs and prices variable. Taheripour et al. (2011) implemented a worldwide CGE based on the GTAP model and included biofuel and biofuel processing. The model treated land at a very aggregate level, allocating land to different uses based upon the profitability of the use and a set of elasticities of substitution among the land use classes. The advantage of this approach is that all problems are solved at once: how much new land to allocate to food and bioenergy crops and how much prices of food increase.

Leaving the very aggregate level of CGE's, the land allocation process is made to depend more on local conditions and expected profits. The underlying disaggregated approach is a model of miscanthus (or other bioenergy crops) yield and profitability. Khanna et al. (2008) examined the likely placement of miscanthus in Illinois. They used a crop-productivity model, MISCANMOD, to predict yields of miscanthus and estimated the costs of harvesting and transportation. For other crops they used actual yield and University of Illinois costs of production data. Using historical price information, the yield information, and the cost information, the authors worked out an opportunity cost of land for other crops. They then parametrically varied the price of miscanthus and found the locations where miscanthus would produce more profit than the alternative crop. The resultant map (their figure 4) showed that miscanthus breaks even first in southern Illinois, which is not prime corn growing ground. In a later paper (Khanna, et al., 2011) they used this supply technology in a multi-market agricultural model to see where in the United States bioenergy crops would be grown. While the purview here is only the United States, the multi-market model, like the CGE models, apportions land and solves for prices. The approach in these papers starts with a detailed place specific model of biofuel supply, which is an improvement over the broad brush elasticity approach of the large CGE models.

Scown et al. (2012) examined where miscanthus should be grown, considering both economic profits and environmental factors. They are most different from the previous papers in their emphasis on

energy as a by-product and their broader explicit objectives-environment as well as economics. Their model included the optimal location of bio-refineries and the value of power offsets based on non-baseload greenhouse gas emissions. The power offsets arise because bio-refineries can burn lignin and sell the power on the grid. The model was calibrated to meet the RFSII 2020 goal of 39.7 billion liters of fuel. They had six scenarios depending on what constraints are imposed. Without constraint they estimated 7.9 million ha of active cropland are required to make the required biofuel. Other scenarios showed that, avoiding drought prone counties and insuring a large power offset, the fuel can be made with 5 million ha of active cropland and 12.7 million ha of CRP land. Figure 1 in their paper showed the results for their six scenarios, which are limited to the east and Midwest. The scenarios are strikingly different, showing how the power sales from the lignin can powerfully influence location.

As discussed in the introduction, we use a novel econometric approach to model crop choice based on historical data. We use an estimated proportion model using relevant weather, soil, and other data. The model is used to explain historical crop coverage in parts of six Midwestern states. We then add the bioenergy crop, miscanthus, by simulating its yield with BioCro (a successor to MISCANMOD) and making its proportion on the landscape function of its yield.

3   Data

We match the currently most fine scale geospatially explicit data on land cover, soil characteristics, weather, and miscanthus yield on a 4km by 4km grid to create the primary data set. The states included in the analysis are those along the Mississippi-Missouri river corridor for which there are at least 10 years of land cover data: Iowa, Illinois, Mississippi, and part of Wisconsin, Missouri, and Arkansas. There is currently insufficient land cover data to extend our analysis to other states. Summary statistics are provided in Table 1. Each variable in the table is described in detail below.

Land Use

5

Land cover data is derived from the Cropland Data Layer (CDL) available annually from 2000 to 2010 (USDA NASS) for the six states. The CDL is generated based on Resourcesat-1 AWiFS, Landsat 5 TM, and Landsat 7 ETM+ satellites and has a ground resolution of 56 or 30 meters, depending on the year and sensors used (Mueller & Seffrin, 2006).

We divide land cover into major crops, other crops, non-crop and wild, urban, and water bodies. The major crops include corn and soybean for Iowa, Wisconsin, and Illinois; and corn, soybean, rice, and cotton for Missouri, Arkansas, and Mississippi. The category of non-crop and wild land includes pasture, forest, improved pasture, etc. We define agricultural land as the land you could potentially grow something on, including major crops, other crops, and non-crop and wild land. Because urban and water bodies are very difficult to convert into crop land, we do not include them in the category of agricultural land nor discuss them in this paper. Therefore, we define the share of major crops as the area of major crops divided by area of agricultural land.

Figure 1 shows the shares of corn, soybean, rice, and cotton in our study area. Corn grows mainly in the colder north, while soy crops are more widely distributed. Rice and cotton concentrate along the river in Missouri and Arkansas. For corn, the average percent coverage from 2002 to 2010 is 34.3% in the north (the three northern states: Wisconsin, Iowa, and Illinois), while it is only 2.9% in the south (the three southern states: Missouri, Arkansas, and Mississippi). For soy, the coverage is 26.4% and 14.1% in the north and the south, respectively. There is little cotton and rice in the north, while in the south, cotton takes 4.5% of the agricultural land and rice takes 5%.

Soil Characteristics

For soil data we focus on two types of variables, both derived from the USDA's U.S. General Soil Map (STATSGO2). First, the underlying soil data include percent clay, sand, and silt, water holding capacity, pH value, electrical conductivity, slope, frost-free days, depth to water table, and depth to restrictive layer. Soil variable averages are spatially weighted from irregular polygons for each grid cell.

Second, we use a classification system generated by the USDA – Land Capability Class (LCC). A LCC value of one defines the best soil with the fewest limitations for production, and progressively lower LCC classifications signify more limitations on the land for agricultural production. The LCC integer scores decline incrementally to eight, where soil conditions are such that agricultural planting is nearly impossible. The use of LCC codes add explanatory power to the raw soil characteristics because these codes were assigned with knowledge of past yields that depend on characteristics not present in our data set. The distribution of LCC levels is shown in Figure 2. Together with Figure 1, we see that prime agricultural soils are absent in southern Iowa and so largely is the corn-soy complex. Similarly, more optimal soils hug the river in Missouri and Arkansas, and so do rice and cotton.

Weather Variables

For weather data we use PRISM data processed by Schlenker and Roberts (2009) to a 4km by 4km spatial resolution, with a daily level of temporal resolution. The dataset includes both temperature (daily min and max) and precipitation. Figure 3 shows the observed weather condition in the planting season (from April to June)[4] from 2002 to 2010 and the growing seasons (from April to November) from 2002 to 2009. The observed temperatures are warmer in the south and the precipitation levels are appreciably larger. Average temperature in the growing season ranges from 12° to 25°c from the top of Iowa to the bottom of Mississippi, a distance of 1600 km. Total rainfall in a growing season is also variable across this landscape with a high of 130 cm and a low of 30 cm, highest in the southeast and lowest in the northwest.

Because this study has so many cross-sectional data cells, we are able to use a great deal of detail from the weather data. Two time periods of weather data are used for each crop year. (1) The planting season data, which farmers know before they actually plant. A cold wet spring, for instance, would delay planting and make a shorter season crop more desirable than a longer season crop. Compared to corn, soy

---

[4] Planting season and growing season vary across crops. In the six states along the Mississippi-Missouri river corridor, the planting season is from April to May for corn, rice, and cotton, and from May to June for soybean. The harvest season is October for rice and corn, and November for cotton and soybean. Growing season is defined as the period between planting season and harvest season.

is more tolerant of being planted late and more dependent on daylight hours, so it can make up time easily. When the planting season is late, farmers are more inclined to plant soy. (2) Past weather is used as a proxy for expected weather. We do not find much gain from including past weather beyond one season, though, in terms of predicting current weather, quite a few lags of past weather are statistically significant. For parsimony, we limit the lags of past weather to one.

Degree days are calculated from daily highs and lows using a fitted sine curve to approximate the amount of hours the temperature is at or above a given threshold (Baskerville & Emin, 1969). As in Schlenker and Roberts (2009), we bin the weather data into degree days at a given temperature and above. We draw on their work and other literature to reduce the number of bins to just those at critical thresholds. However, we expand the number of classifications of temperature to account for the month in which it occurs. We expect, for instance, that hot temperatures are not as harmful in autumn as they are in the middle of the growing season.

Miscanthus Yield Model

The miscanthus yield predictions that we use in this paper are produced using the R package BioCro, a semi-mechanistic dynamic crop growth and production model. The model and its calibration are described in Miguez et al. (2009; 2012). The model is based on solar radiation intercepted by the plants leaf area, with consideration of moisture, soil, and other physical variables. The model predictions are matched to the 4km square grid used for the rest of the data. The simulated yield is the rainfed end-of-growth-season harvestable biomass.[5]. For each 4km square, soil variables include field capacity, wilting pint, available depth, and initial soil moisture; weather variables include precipitation, solar radiation, wind speed, temperature, and relative humidity, extracted from corresponding North American Regional Reanalysis (NARR) grid.

---

[5] This is significantly less than maximum biomass and accounts for losses while waiting for harvest.

Figure 4 shows the predictions of miscanthus yield along the corridor. Compared with Figure 1 and Figure 2, we see that most of the places with high miscanthus yield are where the soil is good and corn share is high. Southern Mississippi is the obvious exception, with high miscanthus yield and low corn shares.

As with all experimental crop data there are caveats. Experimental fields are more carefully managed and controlled than farmer's fields, but these conditions are not necessarily profit maximizing. Experimental fields do not show the full extent of potential insect and disease problems. There is also a potential upside for full scale producers in that farmers eventually will find unforeseen ways of increasing yields and profits.

4 Empirical Model

First, we estimate the land coverage shares by a limited dependent variable local regression generalization of Nerlove's (1956) classic estimating equations. Nerlove's estimating equations are of the form that coverage is a function of lagged coverage, crop price, input prices and other variables. In its original form, the estimating equations do not force land use to add up to the land that there is. Some authors (Lichtenberg, 1989; Wu & Segerson, 1995) used discrete choice models for crop share decisions. In this way, the adding up constraint is satisfied automatically. Berry's logit (1994) is an appealing discrete choice model for shares that are not zero or one, because it is linear in the parameters and errors. However, our crop coverage data has many data points with zero coverage. To simultaneously deal with the adding up constraint and a great number of zero shares, we use a limited dependent variable regression with a transformation function as described below.

Within each of the 4km squares, n, we observe the fraction of land in year t that was allocated to crop (or other use) i: S_{int}. There are M crops. If we imagine that each hectare of our squares has a crop choice, then on that hectare the crop with the highest revenue will be chosen. As a result, the fraction of the crop chosen in a square will be a proportion type model.

(1)
$$S_{int} = \phi\big((\boldsymbol{\beta_1}'\boldsymbol{X_{1nt}} + d_{1nt}), \dots, (\boldsymbol{\beta_M}'\boldsymbol{X_{Mnt}} + d_{Mnt})\big)$$

where $\boldsymbol{X_{int}}$ is a vector of determinate factors of revenue from planting crop i on plot n at year t, $\boldsymbol{\beta_i}$ is a vector of coefficients and d_{int} is an error term. $\phi()$ is a suitable transformation with its domain on the unit interval. When all of the shares are strictly within the unit interval, using logit as the transformation and rearranging terms gives a linear estimation equation (Berry, 1994): $log(S_{int}) - log(S_{0nt}) = \boldsymbol{\beta_i}'\boldsymbol{X_{int}} + d_{int}$. To deal with the fact that many plots do not have a certain crop (i.e., many S_{int} are zeros), we use a ratio transformation and we get

(2)
$$\frac{S_{int}}{S_{0nt}} = \boldsymbol{\beta_i}'\boldsymbol{X_{int}} + d_{int}$$

In order to predict shares as a function of the independent variables, we sum the share ratio over S_i (recall that the shares sum to one) and solve for S_{0nt}

(3)
$$S_{0nt} = \frac{1}{1 + \sum_{j=1}^{M}(\boldsymbol{\beta_j}'\boldsymbol{X_{jnt}} + d_{jnt})}$$

Substituting (3) into (2), we get

(4)
$$S_{int} = \frac{\boldsymbol{\beta_i}'\boldsymbol{X_{int}} + d_{int}}{1 + \sum_{j=1}^{M}(\boldsymbol{\beta_j}'\boldsymbol{X_{jnt}} + d_{jnt})}$$

Second, we simulate crop shares assuming a biofuel mandate is adopted, requiring 5%, for example, of the landscape planted with miscanthus (or other bioenergy crops). Suppose the yearly returns to crop i is R_i and the return to miscanthus is kR_b, where k is the shadow value of the constraint that 5% of land must grow miscanthus. In this setup k has to be high enough so that for exactly 5% of the land $, kR_b > \max(R_i, i = 1, \dots M)$. Assuming crop coverage is linear in its yearly return, we take the incentive to grow miscanthus as a constant times yield (kY_{bnt}). When the mandate is adopted, the share of a traditional crop becomes,

(5)
$$S_{int} = \frac{\boldsymbol{\beta_i}'\boldsymbol{X_{int}} + d_{int}}{1 + kY_{bnt} + \sum_{j=1}^{M}(\boldsymbol{\beta_j}'\boldsymbol{X_{jnt}} + d_{jnt})}$$

and the share of miscanthus is

(6) $$S_{bnt} = \frac{kY_{bnt}}{1 + kY_{bnt} + \sum_{j=1}^{M} (\boldsymbol{\beta_j'X_{jnt}} + d_{jnt})}$$

The estimation strategy is that first we estimate equation (2) by Tobit, accounting for the zero shares. Then we simulate d_{jnt} ($j = 1, \dots, M$) by taking from a left truncated normal distribution with mean 0, standard deviation σ_{jnt} and truncation at $-\boldsymbol{\beta_i'X_{jnt}}$. Finally, we calculate S_{int} and S_{bnt} for each draw and take the averages.

Because the scale of this study encompasses more than a thousand kilometers, there are conditions that are unaccounted for in our variables that change across the landscape. This spatial correlation can induce heteroskedasticity, which would make straightforward tobit estimation inconsistent. We know of two feasible estimation strategies. One strategy is to estimate a linear probability model with a Spatial Error Model (SEM) correction for the errors. In the linear probability model, OLS would be consistent and the SEM would serve to produce the correct standard errors and a more efficient estimate of the coefficients. The limitation is that the prediction is not guaranteed to be between 0 and 1. The other solution is to estimate a local regression framework: the data used for the estimation for a given place is only a part of the data that his near that place. The spatial correlation is addressed because the coefficients, including the constant, are free to vary across the landscape. We define neighbors of county i to be counties whose centroids are within 70 km distance of the centroid of county i. 70 km is chosen based on Moran's I tests. The tests show that the spatial correlation in error decrease exponentially and beyond 70 km it is lower than 10^{-3}. Within 70 km, a county has 8 neighbors on average and each county has about 100 4km squares. Therefore, each regression has about 900 different squares for 9 years.

Next, we consider what explanatory variables should be included. The Nerlovian adaptive price expectations model (Nerlove, 1956) assumed that farmers have rational price expectations based on their information set, and described it in three equations. Braulke (1982) derived a reduced form from the three equations by removing the unobserved variables. Choi and Helmberger (1993) combined this reduced form and farmer's demand functions, and based on their work, Huang and Khanna (2010) described the

crop share as a function of the lagged share, climate variables, economic variables, risk variables, population density, and time trend. Hausman (2012) included most of these explanatory variables, and also futures prices, substitute crop share and crop yield. To follow the literature,[6] we include lagged crop share, lagged substitute crop share, weather in the current planting season and the last growing season, and soil conditions as explanatory variables. We include the interaction term of heat and moisture to account for the possibility that dry warming is much more harmful than warming with moisture (Lobell, et al., 2011) . Given that the interest in this study is in the introduction of a new crop, rather than the more common price elasticity, we are able to use fixed effects to account for many variables that are common to the observations across space. (1) In many countries (e.g., the United States and European Union), the incentive to grow crops in addition to the price is government payments. As these programs change year to year and have different marginal effects for different farmers, it is not possible to have a fully satisfactory treatment of the payments variable. We use year fixed effects to account for both prices and government programs. The year fixed effects also would account for differences in input prices. (2) Many authors (Just, 1974; Chavas & Holt, 1990; Lin & Dismukes, 2007) argue that the risk of growing a crop, perhaps the variance or lower semi-variance, is an important determinant of crop choices. So long as the risk of growing a crop is taken as constant, which is a good approximation in a short time series, crop fixed effects account for this factor. Therefore, we include year fixed effects to account for both output and input prices and government programs. This leads to the following specification:

(7) $\frac{S_{int}}{S_{0nt}} = \alpha_i + \beta_i S_{int-1} + \gamma_i' SS_{int-1} + \varphi_i' Soil_n + \theta_{1i}' GDD_{nt-1} + \theta_{2i}' PDD_{nt} + \theta_{3i}' GP_{nt-1} +$

$\theta_{4i}' PP_{nt} + \theta_{5i}' PreDD_{nt} + \theta_{6i}' PreDD_{nt-1} + \mu_t + \varepsilon_{int}$

where S_{int} is the share of crop i planted at square n in year t. SS_{int-1} is a vector of substitute crop shares planted in year $t - 1$. $Soil_n$ is a vector of soil conditions, including all the soil characteristics described in the data section. GDD_{nt-1} is a vector of degree days by month in the last growing season

---

[6] For reviews of share response literature, see Askari and Cummings (1977) and Nerlove and Bessler (2001) .

(April through November in year $t-1$). $\boldsymbol{PDD}_{nt}$ is a vector of degree days by month in the current planting season (April through June in year t). The critical temperatures in a planting season include $10\,^{\circ}$c and $15\,^{\circ}$c. $10\,^{\circ}$c is the base temperature limit of rice, corn, and soybean development, while $15\,^{\circ}$c is the base temperature limit of cotton development. The critical temperatures in a growing season include $10\,^{\circ}$c, $15\,^{\circ}$c, $20\,^{\circ}$c, $25\,^{\circ}$c, $29\,^{\circ}$c, and $32\,^{\circ}$c. Temperatures higher than $29\,^{\circ}$c are harmful to corn, $30\,^{\circ}$c to soybean, and $32\,^{\circ}$c to rice and cotton (Schlenker & Roberts, 2009). $\boldsymbol{GP}_{nt-1}$ is a vector of precipitation by month in the last growing season. $\boldsymbol{PP}_{nt}$ is a vector of precipitation by month in the current planting season. $\boldsymbol{PreDD}$ are vectors of interactions of degree days above $30\,^{\circ}$c and precipitation levels in the same month. All months in the current planting season and the last growing season are included.

Finally, given that miscanthus is a perennial, we allow five years for miscanthus to grow and predict the shares of miscanthus and other land use five years after the introduction of miscanthus. By assuming miscanthus is introduced in 2010, we simulate

(8)
$$S_{bn2015} = \frac{kY_{bn2010}}{1+kY_{bn2010}+\sum_{j=1}^{M}(\boldsymbol{\beta_j'X_{jn2015}}+d_{jn2015})}$$

and

(9)
$$S_{in2015} = \frac{\boldsymbol{\beta_i'X_{in2015}}+d_{in2015}}{1+kY_{bn2010}+\sum_{j=1}^{M}(\boldsymbol{\beta_j'X_{jn2015}}+d_{jn2015})}$$

We do not observe temperature and precipitation in future, so we assume that they are the same as 2010. The independent variable which is evolving is the lagged crop shares. That is, based on observed $\boldsymbol{X_{in2010}}$, we predict $\frac{S_{in2010}}{S_{0n2010}}$ and simulate S_{bn2010} and S_{in2010}. Then we substitute S_{in2010} into equation (5) and predict $\frac{S_{in2011}}{S_{0n2011}}$, assuming weather conditions in 2011 are the same as 2010. Next, we repeat the steps above until we gain S_{bn2015} and S_{in2015}.

5 Estimation Results

We run separate regressions for each crop and each county. In sum, we have 1022 sets of estimates (368 counties; 2 main crops for the northern states and 4 main crops for the southern states). We test the significance of soil, precipitation, and degree days. The F-test results are shown in Table 2. Soil, precipitation and temperature are significant at the 1% significance level in most of the regressions for corn, soy, and cotton, while they are significant in half of the regressions for rice. Rice only covers about 4% of the land in the southern states, while the land for other use covers about 80% of the land. It is not surprising that the coefficients for rice are not statistically significant, given that the dependent variable is the ratio of rice share and the share of other land use. In a linear probability model, using just rice share, all coefficient groups are significant, so the lack of significance is likely because of the inability to predict the "other" category. Cotton covers a small portion of land as well, however cotton responds more strongly to weather than rice. Therefore, the coefficients are significant in the regressions for cotton, while they are not in the regressions for rice.

6 Adding a New Bioenergy Crop with a Mandate

Based on equation (6), we see that bioenergy crop coverage is proportionately smaller where the other crops have a high index value and proportionately larger where the bioenergy crop has a high yield. From equation (5), we also see that the new crop takes proportionately from all other uses in each 4km square .

The constant k solves the problem of making the average fraction of the area (average taken over plots n) equal the desired fraction of the landscape in the bioenergy crop. We first find that a k of 0.0105 resulted in 5% of the landscape devoted to miscanthus. Being a bioenergy crop, the movement of the biomass is a large factor in costs, so we run the experiment again, this time requiring that for any land to be allocated to miscanthus at least 10% of the land in a 4km square must be predicted for miscanthus. We find that when $k = 0.018$, 5% of the total agricultural land will be converted to miscanthus, with any 4km square that is converted containing at least 10% miscanthus. This serves to concentrate miscanthus on the landscape.

Figure 5 shows the spatial distributions of miscanthus and the changes in crop shares. It indicates that miscanthus will take land mainly from minor crops, pasture, or forest. The main crops are not affected significantly. If miscanthus is introduced in 2010, the predicted corn share in 2015 is 28.82%, while it is 29.22% if miscanthus is not introduced. That is, among the 5% of the landscape devoted to miscanthus, only 0.4% comes from corn. Similarly, 0.34% comes from soy (soy shares decreases from 29.26% to 28.92%), 0.004% from rice (rice shares decreases from 4.1% to 4.096%), and 0.04% from cotton (cotton shares decreases from 3.99% to 3.95%). To contrast, 4.17% of the landscape comes from minor crops, pasture, or forest (other land use decreases from 33.43% to 29.26%).

The predicted spatial distributions of miscanthus coverage and yield are shown in Figure 6. Miscanthus is most prevalent in Eastern and Southern Mississippi, where none of the major crops are prevalent. It is also grown in the southern part of Iowa, where the soils have a LCC of 2 or lower. Miscanthus is planted at the expense of corn and soy in Southern Wisconsin, however, even on the better soils.

7   Conclusion

Using fine scale data on weather and soil, this paper shows the likely geographical distribution of miscanthus, taking into account the competition with existing crops. The predictions of miscanthus land use are based on yield from the *BioCro* model adopted to miscanthus and the assumption that the desirability of growing miscanthus would be proportional to yield. For existing crops we use actual land coverage, which are farmer's decisions, to find predicted crop coverage desirability as a function of weather and soil. We then combine these two sources of information on likely land coverage into one model and use it to simulate the crop coverage when miscanthus is constrained to take up 5% of the relevant landscape.

The model presented here is a new hybridization of deterministic and statistical modeling for the introduction of a new crop. The advantage over a purely process model is that it is "normed" to hundreds of thousands of actual farm land cover decisions, rather than being calibrated to a few plots. The

advantage over the usual ways of introducing a "new alternative" (see McFadden BART study (1977) for a famous example) is that a process model is great improvement over assuming that a calibrating constant (alternative specific constant in McFadden) is quite like the constant in an existing alternative.
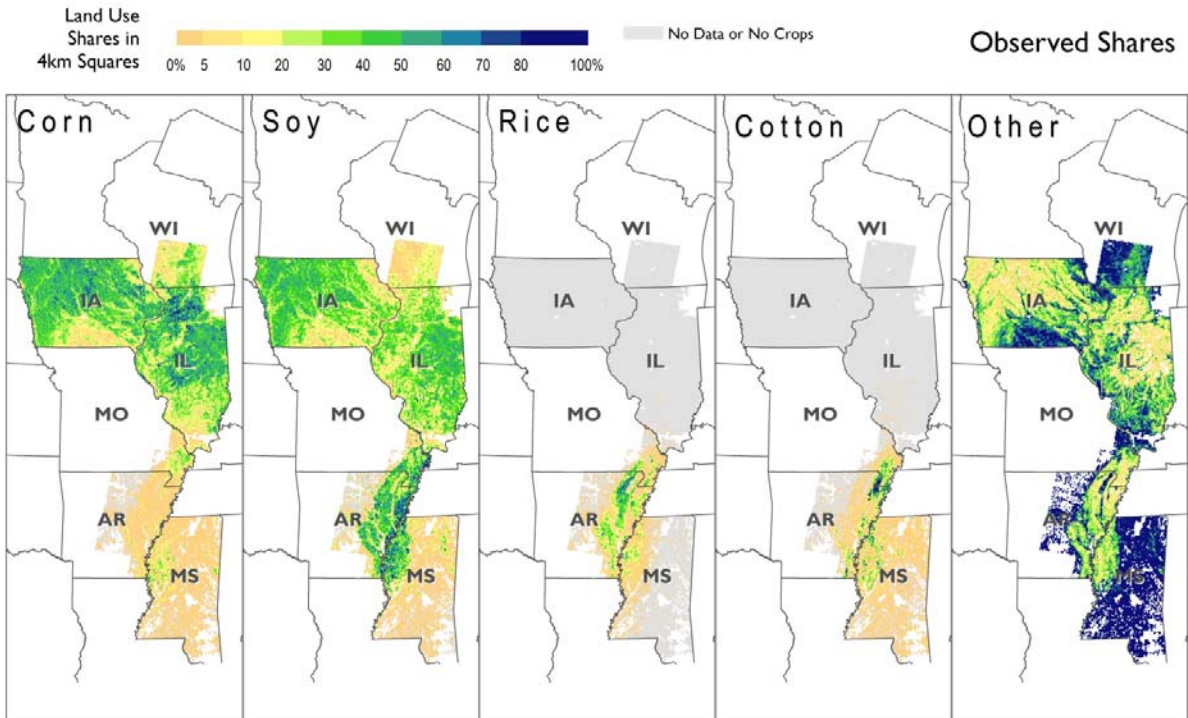
Growing a bioenergy crop can compete directly for land with crops grown for food, as is the case when corn is used for ethanol. Empirical estimates of how greatly this raises food prices are given in Hausman, Auffhammer and Berck (2012). One possible superiority of miscanthus to corn based ethanol is that farmers may choose to grow it on land not currently used for food crops. Indeed Figure 6 shows that this is largely the case, even for miscanthus grown within the corn states of Illinois and Iowa. For Iowa and Arkansas particularly, variation in soil quality is the defining characteristic for where miscanthus is grown: it is grown on the lower quality soils. The use of the lower Mississippi region for miscanthus production is likely a function of temperature and moisture, which are not suitable for the main crops in this study. While we find that miscanthus potentially has a low requirement for land currently used for corn, soy, cotton, and rice, we also find that it has a high requirement for land that is currently in the other category: pasture, forest, and grassland.

.

Reference

Askari, H. & Cummings, J. T., 1977. Estimating Agricultural Supply Response with the Nerlove Model: a Survey. *International Economic Review,* 18(2).

Baskerville, G. & Emin, P., 1969. Rapid Estimation of Heat Accumulation from Maximum and Minimum Temperatures. *Ecology,* pp. 514-517.

Berry, S. T., 1994. Estimating Discrete-choice Models of Product Differentiation. *The RAND Journal of Economics,* pp. 242-262.

Braulke, M., 1982. A Note on the Nerlove Model of Agricultural Supply Response. *International Economic Review,* 23(1), pp. 241-244.

Chavas, J.-P. & Holt, M. T., 1990. Acreage Decisions Under Risk: the Case of Corn and Soybeans. *American Journal of Agricultural Economics,* 72(3), pp. 529-538.

Choi, J.-s. & Helmberger, P. G., 1993. How Sensitive are Crop Yields to Price Changes and Farm Programs?. *Journal of Agricultural and Applied Economics,* Volume 25, pp. 237-244.

Hausman, C., 2012. Biofuels and Land Use Change: Sugarcane and Soybean Acreage Response in Brazil. *Environmental and Resource Economics,* 51(2), pp. 163-187.

Hausman, C., Auffhammer, M. & Berck, P., 2012. Farm Acreage Shocks and Food Prices: An SVAR Approach to Understanding the Impacts of Biofuels. *Environmental and Resource Economics,* 53(1), pp. 117--136.

Huang, H. & Khanna, M., 2010. *An Econometric Analysis of US Crop Yield and Cropland Acreage: Implications for the Impact of Climate Change.* Denver, Colorado, s.n., pp. 25-27.

Just, R. E., 1974. An Investigation of the Importance of Risk in Farmers' Decisions. *American Journal of Agricultural Economics,* 56(1), pp. 14-25.

Khanna, M., Chen, X., Huang, H. & Onal, H., 2011. Supply of Cellulosic Biofuel Feedstocks and Regional Production Pattern. *American Journal of Agricultural Economics,* 93(2), pp. 473--480.

Khanna, M., Dhungana, B. & Clifton-Brown, J., 2008. Costs of Producing Miscanthus and Switchgrass for Bioenergy in Illinois. *Biomass and Bioenergy,* 32(6), pp. 482--493.

Lichtenberg, E., 1989. Land Quality, Irrigation Development, and Cropping Patterns in the Northern High Plains. *American Journal of Agricultural Economics,* 71(1), pp. 187-194.

Lin, W. & Dismukes, R., 2007. Supply Response Under Risk: Implications for Counter-cyclical Payments' Production Impact. *Applied Economic Perspectives and Policy,* 29(1), pp. 64-86.

Lobell, D. B., Banziger, M., Magorokosho, C. & Vivek, B., 2011. Nonlinear Heat Effects on African Maize as Evidenced by Historical Yield Trials. *Nature Climate Change,* 1(1), pp. 42-45.
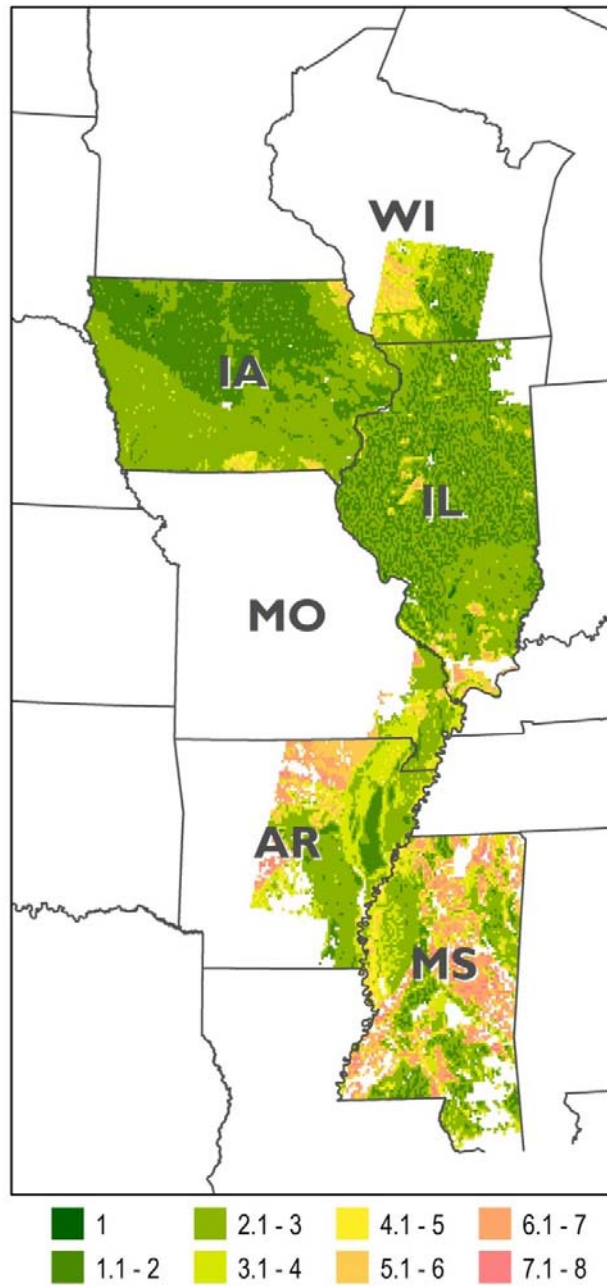
McFadden, D., 1974. The Measurement of Urban Travel Demand. *Journal of Public Economic,* 3(4), p. 303–328.

McFadden, D. et al., 1977. *Demand Model Estimation and Validation.* s.l.:Institute of Transportation Studies.

Miguez, F. E., Maughan, M., Bollero, G. A. & Long, S. P., 2012. Modeling Spatial and Dynamic Variation in Growth, Yield, and Yield Stability of the Bioenergy Crops Miscanthus x giganteus and Panicum Virgatum across the Conterminous United States. *GCb Bioenergy,* pp. 509-520.

Miguez, F. E. et al., 2009. A Semimechanistic Model Predicting the Growth and Production of the Bioenergy Crop Miscanthus x giganteus: Description, Parameterization and Validation. *GCB Bioenergy,* 1(4), pp. 282-296.

Mueller, R. & Seffrin, R., 2006. New Methods and Satellites: A Program Update on the NASS Cropland Data Layer Acreage Program. *Remote Sensing Support to Crop Yield Forecast and Area Estimates, ISPRS Archives,* 36(8), p. W48.

Nerlove, M., 1956. Estimates of the Elasticities of Supply of Selected Agricultural Commodities. *Journal of Farm Economics,* 38(2), pp. 496-509.

Nerlove, M. & Bessler, D. A., 2001. Expectations, Information and Dynamics. *Handbook of Agricultural Economics,* Volume 1, pp. 155-206.

Schlenker, W. & Roberts, M. J., 2009. Nonlinear Temperature Effects Indicate Severe Damages to US Crop Yields under Climate Change. *Proceedings of the National Academy of Sciences,* 106(37), pp. 15594-15598.

Schlenker, W. & Roberts, M. J., 2009. Nonlinear Temperature Effects Indicate Severe Damages to US Crop Yields under Climate Change. *Proceedings of the National Academy of Sciences,* 106(37), pp. 15594-15598.

Schlenker, W. & Roberts, M. J., 2009. Nonlinear Temperature Effects Indicate Severe Damages to US Crop Yields under Climate Change. *Proceedings of the National Academy of Sciences,* 106(37), pp. 15594-15598.

Scown, C. D. et al., 2012. Corrigendum: Lifecycle Greenhouse Gas Implications of US National Scenarios for Cellulosic Ethanol Production. *Environmental Research Letters,* 7(1), p. 9502.

Taheripour, F., Tyner, W. E. & Wang, M. Q., 2011. Global Land Use Changes due to the US Cellulosic Biofuel Program Simulated with the GTAP Model. *Argonne National Laboratory, http://greet. es. anl. gov/files/luc_ethanol.*

Wu, J. & Segerson, K., 1995. The Impact of Policies and Land Characteristics on Potential Groundwater Pollution in Wisconsin. *American Journal of Agricultural Economics,* 77(4), pp. 1033-1047.

Figures 1: Observed Crop Coverage along the Mississippi-Missouri River System

Notes: Graphs display observed coverage shares for corn, soy, rice, cotton, and other land use, in the six states along the Mississipppi-Missouri river corridor. They are average shares over 2002-2010.

Figure 2: Distribution of Land Capabilty Classification (LCC) Levels

Notes: Land Capability Class (LCC) 1 is the best soil, which has the fewest limitations. Progressively lower classifications signify more limitations on the land for agricultural production.
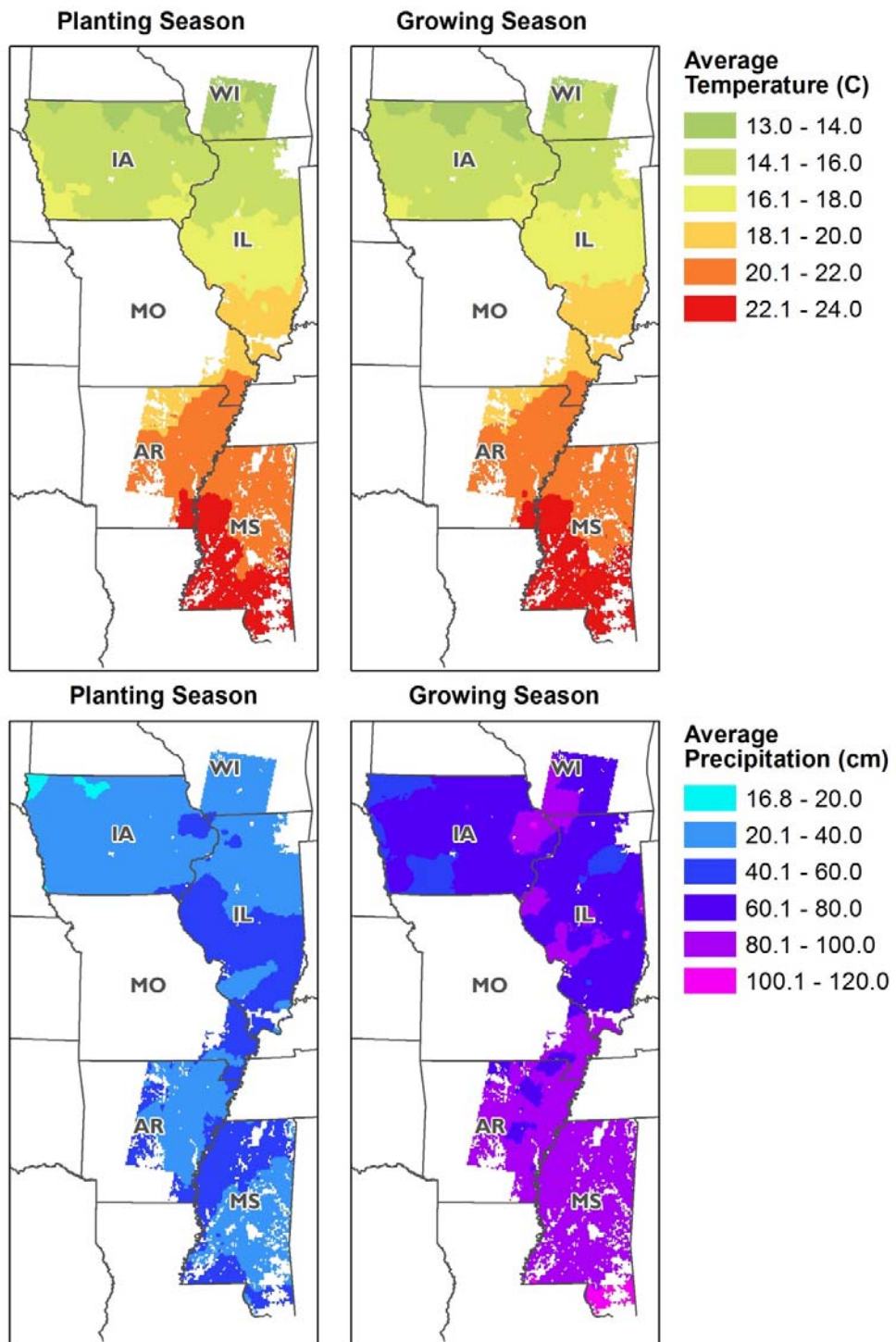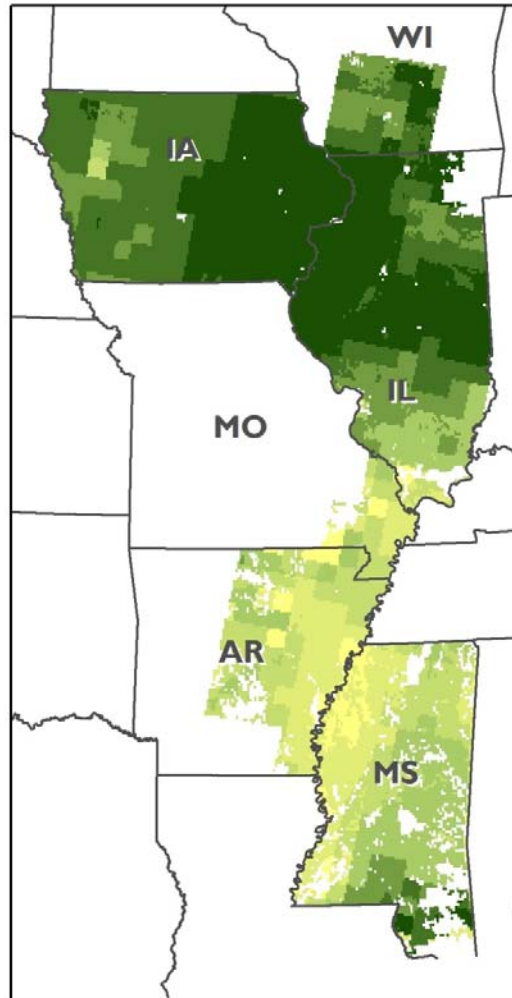
Figure 3: Weather Conditions

Figure 4: Miscanthus Yield

Miscanthus Yield 2010 (Mg/Ha)

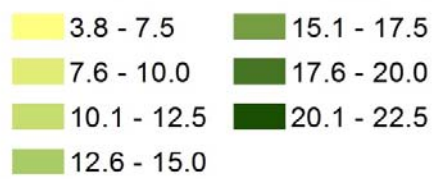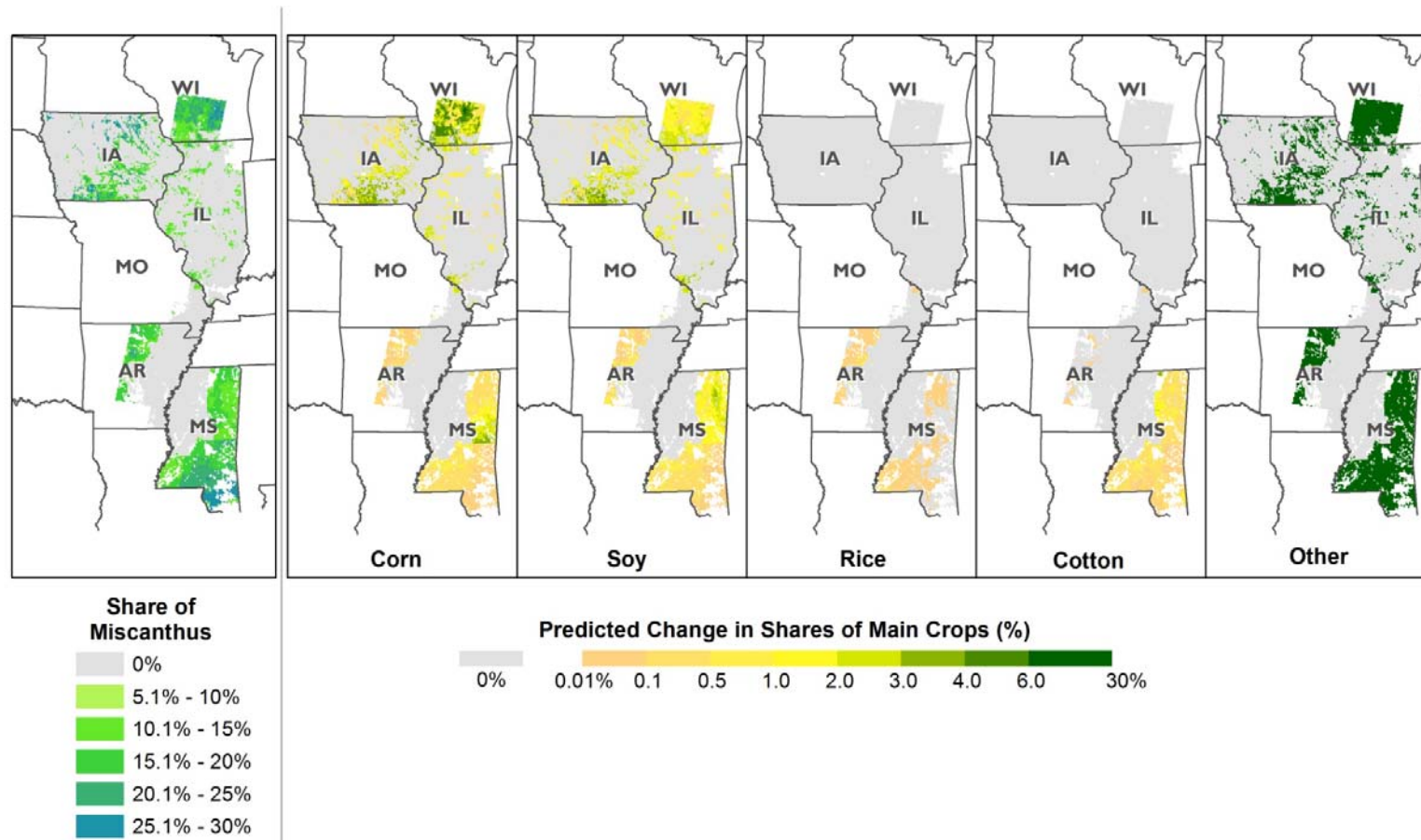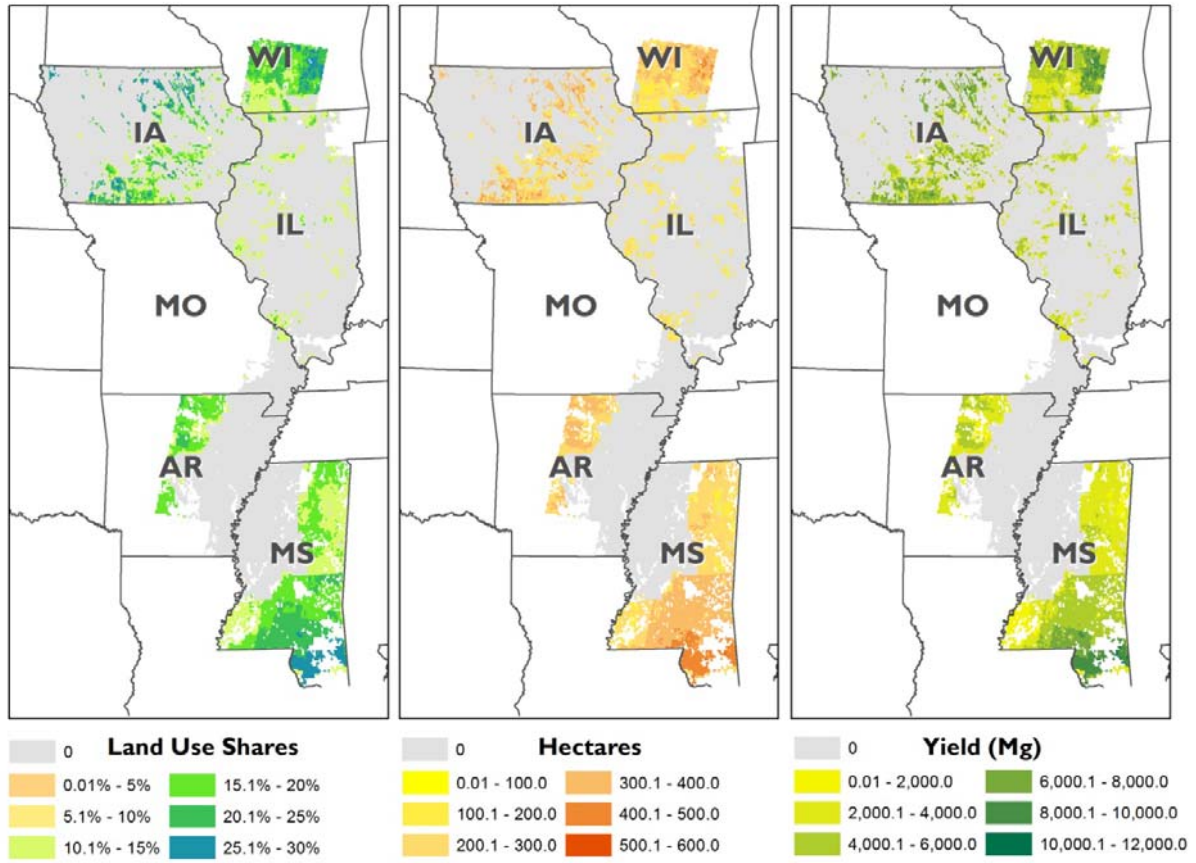|	
3.8 - 7.5	15.1 - 17.5
7.6 - 10.0	17.6 - 20.0
10.1 - 12.5	20.1 - 22.5
12.6 - 15.0	

Figure 5: Predicted Changes in Shares of Miscanthus and Main Crops

Notes: a 20% change reported here means corn (for example) share decreases from a to a-0.2.

Figure 6: Predicted Miscanthus Coverage and Yield

Notes: a 20% reported here means miscanthus share increases from 0 to 20% of the agricultural land.

Table 1. Summary Statistics

Variable	Mean	St.dev.	Min	Max	N. of Obs.
Land Use (Unit: ha; Period: 2002-2009)					
Corn	348	318	0	1480	266409
Soy	335	266	0	1640	266409
Rice	28	122	0	1690	266409
Cotton	25	103	0	1370	266409
Agricultural Land	1520	213	0	1840	266409
Soil Condition					
Percent clay (%)	27.61	8.16	0	62.90	29601
Percent sand (%)	23.41	13.46	0	95.30	29601
Percent silt (%)	48.32	12.23	0	73.60	29601
Water holding capacity	0.17	0.03	0	0.33	29601
pH	6.15	0.70	0	7.70	29601
Slope	4.73	5.55	0	48.00	29601
Electrical conductivity	0.01	0.10	0	5.10	29601
Frost free days	176.48	48.93	0	302.70	29601
Depth to water table	74.05	36.84	0	201.00	29601
Depth to restrictive layer	180.36	44.45	3.30	201.00	29601
Percent of Land in Class 1	0.01	0.06	0	1	29601
Percent of Land in Class 2	0.56	0.43	0	1	29601
Percent of Land in Class 3	0.25	0.36	0	1	29601
Percent of Land in Class 4	0.04	0.16	0	1	29601
Percent of Land in Class 5	0.03	0.16	0	1	29601
Percent of Land in Class 6	0.07	0.22	0	1	29601
Percent of Land in Class 7	0.02	0.12	0	1	29601
Percent of Land in Class 8	0.00	0.01	0	0.77	29601
Weather Variables					
Planting Season (April through June from 2002 to 2010)					
Temperature (Daily Average, Celsius)	17.82	2.92	11.98	24.87	266409
Precipitation (Total, CM)	33.93	8.77	5.64	74.76	266409
Growing Season (April through November from 2002 to 2009)					
Temperature (Daily Average, Celsius)	17.66	2.85	12.26	24.16	236808
Precipitation (Total , CM)	78.43	14.48	32.54	127.29	236808
Miscanthus Yield					
Harvested Biomass (Unit: Mg/ha; in 2010)	16.36	4.69	3.76	22.38	29567

Table 2: F-tests for Soil, Precipitation and Temperature

Regressions with 1% Significance Level		Corn	Soy	Rice	Cotton
	Soil	93%	89%	54%	82%
	Precipitation	76%	72%	45%	72%
	Temperature	93%	90%	66%	92%
Regressions with 5% Significance Level		Corn	Soy	Rice	Cotton
	Soil	97%	92%	57%	85%
	Precipitation	82%	81%	49%	80%
	Temperature	95%	92%	67%	93%
Regressions with 10% Significance Level		Corn	Soy	Rice	Cotton
	Soil	97%	93%	60%	89%
	Precipitation	85%	84%	54%	83%
	Temperature	95%	93%	67%	93%
Number of Regressions in Total		368	368	143	143