****DRAFT**** March 2011**

Prepared Paper for the 55th Annual AARES Conference

Melbourne VIC February 8th-11th 2011

An Empirical Model of Crop Rotations

Duncan MacEwan and Richard Howitt

University of California, Davis

Abstract

Crop rotation systems have played a key role in agricultural production for thousands of years, dating back to the biennial grain-fallow rotations employed by the Ancient Greeks. Fundamentally, rotations are rooted in intertemporal spillover effects between crops, the economic consequences of which depend on relative input and output prices. We contribute to the literature by developing a dynamic, field-level model of crop rotations using a geo-referenced panel dataset that covers 12 years and over 14,000 individual fields. We identify empirical rotations using a Sequence Analysis procedure from the bio-informatics literature, and calibrate a dynamic field-level profit function that satisfies the underlying Euler dynamic first-order conditions using Generalized Maximum Entropy. The resulting model is based entirely on empirical data, and exhibits a stable rotational cycle which responds to changes in expected prices and costs. We illustrate the mechanics of the model with a four-crop rotation of alfalfa, cotton, grain, and fallow, and simulate field-level changes resulting from changes in relative prices.

Introduction

Rotation systems are an important part of agricultural planting decisions for breaking pest and disease cycles and managing soil fertility. Rotations increase yields and save on fertilizer and pesticide costs. The science of rotations is well established and includes managing soil fertility, reducing weed populations, and managing soil water content. The economics of crop rotations is less formally developed due, in large part, to a lack of detailed data, as well as the inherent difficulty of the rotation problem. Rotations are a result of behavior, space, and time through management decisions, field specific physical capital, and the sequence of crop plantings, respectively. Incorporating these effects simultaneously is difficult and further complicated by farm-wide considerations such as risk and input availability. However these considerations are important for understanding the nature of agricultural production and supply response. Treating production as a static process omits the true underlying dynamics which may lead to erroneous supply elasticity estimates and policy response.

We contribute to the literature by using a unique panel dataset of geo-referenced field data to empirically identify observed rotations and develop a field-level dynamic model of agricultural production. The model is calibrated based on empirical data and satisfies the underlying Euler conditions for profit maximization. We demonstrate the usefulness of the model by simulating a four crop rotation and applying the model under price shocks to the study region, Kern County California.

Relevant Literature

Heady (1948) was the first to consider the crop rotation problem with an application to the hay-grain rotations observed in the Corn Belt. Subsequent work focuses on the dynamics of the rotation problem. A seminal paper on the dynamics of crop rotations is Burt and Allison (1963), who consider a dynamic programming approach to crop planting decisions and explicitly model the dynamics of alternative crop rotations. In general, research on the economics of crop rotations falls into four main areas: (i) linear programming type models of production with rotation constraints (El-Nazer and McCarl 1986), (ii) models that lend themselves to econometric analysis and control for lagged crop choice (Wu et al. 2004, Hennessy 2006), (iii) dynamic analysis where crop rotation is modeled as a control variable consisting of the proportion of total land use (Thomas 2003) (Kennedy 1986) (Orazem and Miranowski 1994), and (iv) multiple-phase optimal control models which model the switch point between two successive crops (Doole 2008, Doole 2009).

Field Level Model Specification

Consider a farmer's sequential planting decision on a single field, where a field is defined as an area to which only one crop is planted at any given time. The farmer is interested in determining an optimal management plan for sequential planting of crops on the field. Rotating crops in sequence has an effect on yields and costs, depending on the crop planted. We assume only a one year memory for the rotation. The farmer has expectations about future prices, which vary over time, and has direct knowledge of the soil type and water quality associated with the individual field. We define the dynamic model as follows.

For simplicity, normalize the size of the field to one acre. Let c_t denote planting of crop c in period t for $t = 1, 2, ..., T$. Denote $p_t(c)$ as the price of a unit of output (yield in tons) of crop c in period t which yields $\bar{y}(c)$ average yield in tons per acre. Assume that average yield, $\bar{y}(c)$, is constant over time. The costs of production are constant over time and denoted $A(c)$. Thus, the average farmer profits generated from the field, in the absence of rotational effects, are:

$$\pi_t = p_t(c)\bar{y}(c) - A(c) \tag{1}$$

We introduce rotational effects, as represented by the effect of crop rotation on the state of the field. The underlying dynamics of the state of the field are a combination of factors including soil quality, water storage capacity, pests, disease, and exogenous weather factors. Let "fertility" of the field be the representative variable which captures the net effect of all these considerations. Furthermore, assume that the underlying fertility of the field at any point in time can be described entirely based on the previous crop history. Let s_t be the state variable that represents the underlying "fertility" of the field, which depends solely on a function of the crop planted in the previous period, $s_t = g(c_{t-1})$

At any point in time yields, and thus profits, are affected by the fertility of the field. Following the suggestions by Hennessey (2006), crop rotations result in yield increases and input savings.

To capture yield effects (output boost) we introduce a function that adjusts average yield by crop (positive or negative) depending on the state of the system (a function of previous crop choice). Let $\gamma(c_t, s_t)$ denote the yield adjustment function for crop c in period t given state s_t. Assume that this functional form is stationary. Input savings is included in a similar fashion, denoted $\psi(c_t, s_t)$.

In addition to rotational effects, yields are affected by soil quality and water quality. Thus, we introduce two coefficients to capture the effect of salt and soil on yields. Let these be denoted by $\beta_1(c)$ and $\beta_2(c)$ representing the marginal effect of salt and soil on crop yields, respectively. We assume that these are stationary and unaffected (directly) by crop rotation.

Farmers form expectations about current and future prices in order to make current production decisions. To model prices that are uncertain at any point in time we allow prices to follow a first-order Markov process, with five states per crop, to represent farmer price expectations. We later estimate the corresponding transition matrix based on a time series of price data. This simplification allows us to model prices as being time independent. Taking this into account and letting the discount factor be denoted as δ, the farmers' optimal rotation problem is written as:

$$\begin{aligned}
&\max_{c_t} \sum_{t=0}^{T} \delta^t \left\{ p_t(c_t)\left(\bar{y}(c_t) - \gamma(c_t, s_t) - \beta_1(c_t) - \beta_2(c_t)\right) - \left(A(c_t) - \psi(c_t, s_t)\right)\right\} \\
&s.t. \qquad\qquad (2)\\
&s_{t+1} = g(c_t) \\
&p_t \ \ follows\, a\ \ first\, order\, Markov\, Process
\end{aligned}$$

Following Bellman (1957) we can express this problem using the Bellman Equation. Let $V_t(s)$ be the value function, which is the maximum attainable sum of current and future profits from cropping activities given that the field is in state *s* at time *t*.

$$\begin{aligned}
&V(s_{t-1}, p_{t-1}(c_{t-1}), \bar{y}(c_{t-1})) = \qquad (3)\\
&\max_{c_t} E_{p_t} \left\{ p_t(c_t)\left(\bar{y}(c_t) - \gamma(c_t, s_t) - \alpha(c_t) - \beta(c_t)\right) - \left(A(c_t) - \psi(c_t, s_t)\right) + \beta V(s_t, p_t(c_t), \bar{y}(c_t))\right\}
\end{aligned}$$

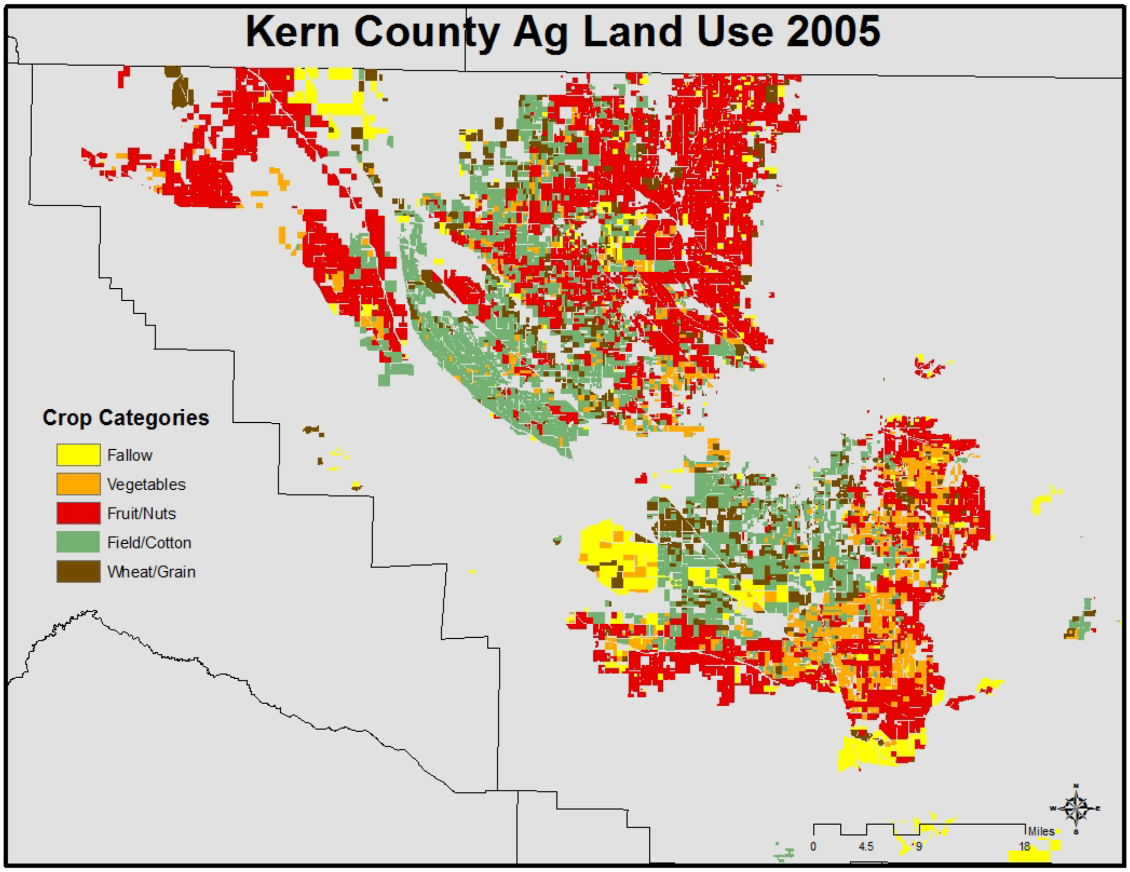
This formulation lends itself to the solution methods of Dynamic Programming. Specifically, we assume an infinite time horizon for the problem and use value function iteration to find a fixed point of the Bellman Equation. From there we determine the optimal policy function, the optimal crop planting decision, at any point in time, given the state of the system.

Data and Study Region

The data requirements for a model of the form specified above are significant. We use a unique panel dataset of geo-referenced land use and field specific conditions. The study region is Kern County, California.

Kern County California is located at the southern end of the San Joaquin Valley. Agriculture in the region is primarily irrigated with water coming from State and Federal surface water projects and groundwater in addition to local surface supplies. Our data includes all agricultural land use in Kern County between 1997 and 2009. On each field and year we observe the crop grown, field size in acres, farm owner, farm manager of the field, soil type, shallow groundwater salinity, actual ET, and dry biomass production. We are able to uniquely identify and track fields across time using a geo-referenced dataset provided by the Kern County Agricultural Commissioners Office. Aggregated proportions for 2005 agricultural land use are shown below in Figure 1 which details the spatial heterogeneity in agricultural production across Kern County. The spatial distribution of production is clearly non-random. The implication is that micro level, field specific heterogeneity affect production which can be modeled within a dynamic field level framework.

Figure 1. 2005 Land Use in Kern County, CA



Identifying Rotations

Rotations are important for farmers to achieve agronomic, financial, or environmental objectives. The various objectives of rotations are unlikely to be met simultaneously, thus rotations observed empirically are the result of several unobservable constraints and management decisions. Agronomic research on Kern County crops[1] offer insights into the recommended rotations, but these fail to incorporate observed farmer behavior. In order to empirically identify rotations, we adopt a methodology initially developed in the field of bio-informatics research called sequence alignment (Needleman and Wunsch 1970). We employ a simple version of a sequence alignment algorithm called Optimal Matching in order to empirically identify crop rotations. Optimal Matching is another method for identifying commonalities across sequences which has several versions of the algorithm and has a range of applications (Abbott and Tsay 2000). We

[1] http://www.ipm.ucdavis.edu

employ a package, SQ-Ados, developed in Stata by Brzinsky-Fay, Kohler, and Luniak (2006). We use a sequence "suppression" option that condenses multiple sequential crops, of the same type, into a single observation and identifies commonalities across these reduced form sequences. For example, AABC is the same as ABBC and ABCC, and so on. We justify this by noting that rotations are a dynamic process, subject to external shocks, and we intend to formally model the underlying process. The goal of the sequence analysis is solely to identify observed base rotations.

The results of the analysis are summarized in Table 1, below. We show the ten most commonly observed sequences. We select alfalfa-cotton-grain-fallow as the rotation that we will reproduce and simulate in what follows in the paper. This rotation is selected because it offers an interesting rotation of crops with different salt tolerance, profitability, and includes fallow. Fallowing a field is a zero profit (absent of rotational effects) event and is an interesting addition to the dynamic model.

Table 1. Identified Rotations

Rotation	Number of Fields	Percent of Top 10	Percent of Total
Alfalfa-Grain-Fallow	379	13.3	6.42
Alfalfa-Cotton-Fallow	375	13.16	6.36
Alfalfa-Cotton-Grain-Fallow	355	12.46	6.02
Vegetable-Grain-Fallow	321	11.27	5.44
Cotton-Grain-Fallow	279	9.79	4.73
Alfalfa-Cotton	271	9.51	4.59
Vegetable-Cotton-Grain-Fallow	271	9.51	4.59
Cotton-Fallow	244	8.56	4.14
Alfalfa-Cotton-Grain	182	6.39	3.08
Vegetable-Grain	172	6.04	2.92

Estimating Model Parameters

We simulate the alfalfa-cotton-grain-fallow rotation and assume a one year crop lag as detailed in Equation (2). Instead of specifying functions for the soil, salt, yield carryover, and cost carryover effects we specify these as individual crop specific parameters, constant across time. Given the above model definition, there are 106 parameters to be estimated. After imposing restrictions on second, third, and fourth year alfalfa, the dynamic economic model reduces to 61 parameters. These parameters are as follows.

$\underset{4\times1}{\beta_1}$ *soil*

$\underset{4\times1}{\beta_2}$ *salinity*

$\underset{7\times7}{\gamma}$ *yield carryover effect*

$\underset{7\times7}{\psi}$ *cost carryover effect*

β_1 and β_2 are 4 by 1 vectors of crop specific soil and salinity yield adjustment parameters, respectively. We anticipate that these are positive, reflecting a negative effect of decreasing soil quality and increasing shallow groundwater salinity. The parameters ψ and γ represent cost and yield carryover effects due to crop rotation, respectively. The *i,j* entry of each matrix represents the yield or cost adjustment from planting crop *i* today given crop *j* was planted in the previous year. We anticipate that these parameters can take any sign, representing both the positive and negative agronomic effects from rotating crops. For example, nitrogen depleting crops planted in succession would have negative yield and cost effects, representing depleted soil quality, increased fertilizer costs, or both.

For the rotation under consideration, alfalfa-cotton-grain-fallow (ACGF), we observe this rotation across 355 fields in the data. The time frame for the empirical data is 2000 – 2009. We derive the necessary dynamic conditions that must hold. We define crop $i \text{ and } j \in \{a1, a2, a3, a4, c, g, f\}$., prices as p_i, and average yields as $\overline{y}_i$. Additionally, there are variation in yields, unobserved in the panel dataset, but observed in County level data across years. We estimate the yield variance, σ_i^2, in the four crops using the County Agricultural Comissioner's database for Kern County between 2000 and 2009. We define the crop specific profits on a single field from growing crop *i* following crop *j* in any year *t* as:

$$\pi_{t,i|j} = p_t(i)\left(\left[\overline{y}(i) + \varepsilon_i\right] - \gamma(i,j) - Soil * \beta_1(i) - Salt * \beta_2(i)\right) - \left(F(i) - \psi(i,j)\right)$$

where $\varepsilon_i \sim N(0, \sigma_i^2)$ and *Salt* and *Soil* measure shallow groundwater salinity and soil quality, as defined previously.

Assuming farmers are profit maximizing agents, behaving according to the model specified above, there are a set of 42 dynamic first-order conditions that must hold in order for ACGF to be observed. For example, if alfalfa is observed on a field in the current year then the field will rotate into cotton in the subsequent period and was fallowed in the previous period (prior to year 1 alfalfa). This implies six conditions that must be satisfied for each year of the alfalfa crop (four years):

$$\pi_{c|a} \geq \pi_{i|a} \text{ for all } i \neq c$$

Similar logic holds for fields that, at a point in time, are observed in cotton, grain, or fallow. Thus there are 42 first-order conditions that must hold in order for the ACGF rotation to be observed on the set of fields. At any point in time, if a given crop is observed we know the crop that must have preceded it as well as the crop that must succeed it. These conditions will not hold exactly in all cases, thus we specify an error term and use this system of equations for the estimation procedure. There are 42 first-order conditions and 61 parameters, thus there are infinitely many solutions to the system of equations so the problem is underdetermined. The problem is ill-posed and one solution is to use Generalized Maximum Entropy (GME) (Jaynes 1963) (Shannon 1948) (Mittelhammer, Judge and Miller 2003). Given that you have incomplete observations about a statistical process, the best way to recover parameters for inference is to impose probabilistic structure on the model in such a way that it is consistent with observed data and imposes as little additional information as possible.

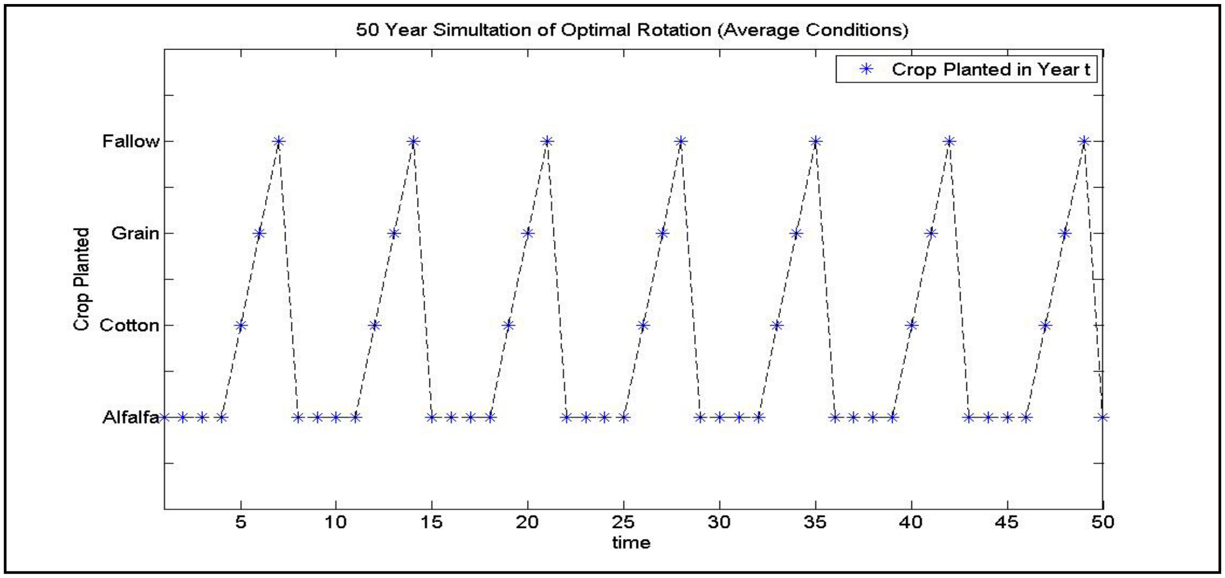
We estimate the GME program in the General Algebraic Modeling Software (GAMS) using the conopt3 non-linear solver. An iteration of the program solves in just over 8 minutes. We bootstrap standard errors for the parameters with 500 boot-strap iterations.

Calibrated Model Application

We substitute the estimated parameters from the GME program back into the dynamic program given by Equation (2), and define the Bellamn equation described by Equation (3). We solve the program using Value Function Iteration which solves for the fixed point of the Bellman equation. The program solves in slightly over 200 iterations.

First we demonstrate the base results of the model in Figure 3, below. We assume a field starting in fallow, over a 50 year time horizon with average salinity and soil conditions. As shown, we can reproduce the base alfalfa-cotton-grain-fallow rotation as an infinite cycle. This shows that, under average price, salinity, and soil conditions the model reproduces the observed rotation. Shocks to prices or other variables will cause the model to switch cycles, possibly temporarily, and rotate different crops.

Figure 3. Average Conditions - Base Results



One interesting application of the model is to impose price shocks and see how different fields respond. To demonstrate the model for these situations we create a grain price shock from $120/ton to $380/ton beginning in year 15 and lasting ten years. We consider a 30 year horizon for this example.

Figure 4 shows the field described above, over average salt and soil, except this field is initially fallowed. With a price spike in year 15 the farmer has already made the decision to plant in alfalfa. Thus the optimal decision is to cycle through four years of alfalfa, into cotton, and then into a grain monoculture for the duration of the price spike. The dynamically optimizing farmer managing this field does not immediately shift into grain, although does after several years

Figure 4. Average Conditions – Grain Price Spike

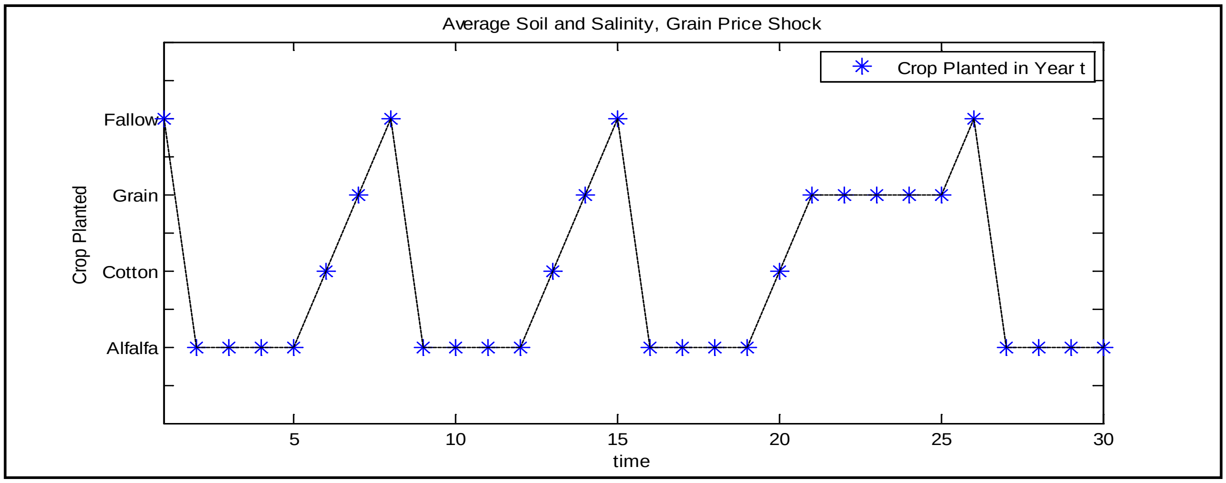


Figure 5 shows a field situated over either high salinity or poor soil quality, the results are the same for both. The dynamically optimizing farmer initially plants the field to a continuous alfalfa-fallow rotation. This is a result of the effect of poor soil or high salinity on crop yields, the cost of keeping the field in the ACGF rotation (or other rotations) is too high given the poor soil and water quality. In year 15, when the price spike is realized, the farmer has just finished a four year alfalfa rotation. Realizing the grain price spike, it is optimal to shift the field into grain monoculture for the duration of the price spike. After the spike, the field is put back into an alfalfa-fallow rotation. Again, the timing of the spike is important. A different outcome would be expected for a field at a different point in the alfalfa rotation, for example year 2.

Figure 5. Poor Soil or High Salinity – Grain Price Spike

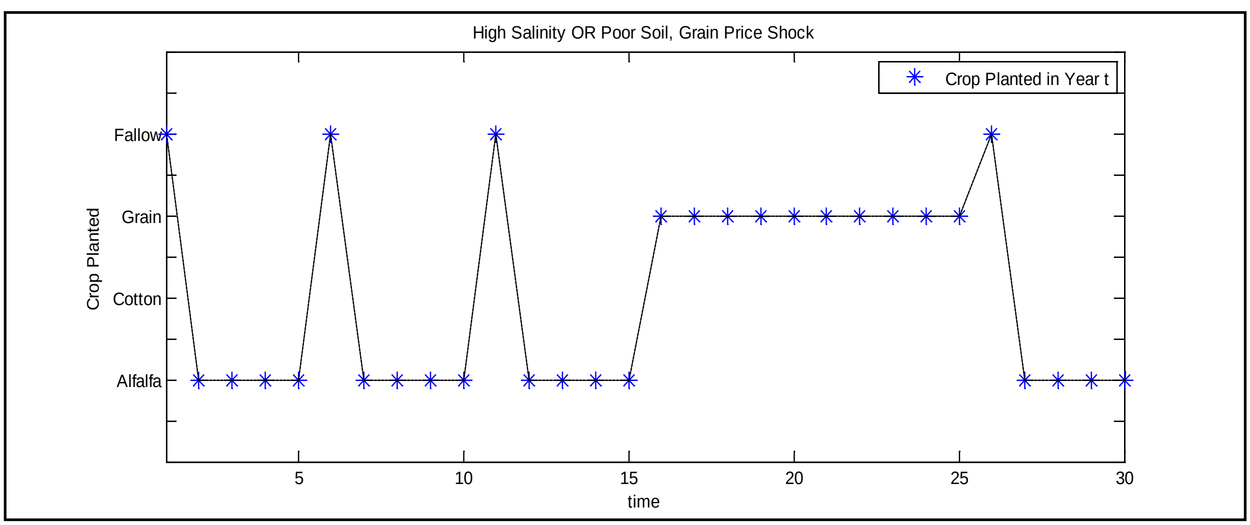
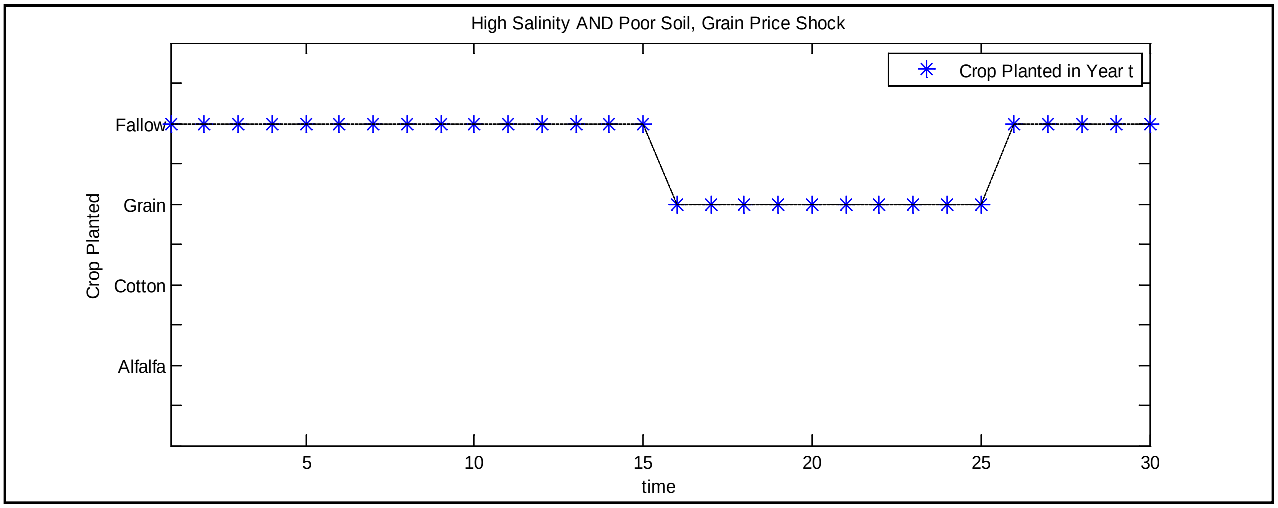


Figure 6 shows a field over very marginal land, with high salinity and poor soil quality. At average prices, it is never optimal to plant this field. This represents marginal agricultural land that may be chronically fallow but has developed irrigation or other infrastructure. With a grain price spike in year 15, grain monoculture becomes profitable for the duration of the price spike, but after the spike, the field is fallowed again.

Figure 6. Poor Soil Quality and High Salinity – Grain Price Spike



Conclusion

In this paper we determined empirically observed rotations, formulated a dynamic model to reproduce these rotations, estimated the parameters for the model based on observed farmer behavior, and simulated the resulting solution. We applied the model, generated a ten year grain price shock example, and evaluated the fit of the model.

This analysis is relevant for supply and policy response modeling. Treating agricultural production as a static process misses the dynamic features of production due to rotations. Additionally, aggregate regional analysis of agricultural production omits the heterogeneity in soil, salinity, and farm management conditions. A natural extension of this paper, which we are currently working on, is to incorporate these results into larger regional models of production. From there we can make comparisons between supply response when rotations are considered and when rotations are omitted. Additionally, we are working on simulating additional policies including salinity effects, drought, and climate change.

References

Abbott, A. & A. Tsay (2000) Sequene Analysis and Optimal Matching Methods in Sociology: Review and Prospect. *Sociological Methods and Research,* 29.

Bellman, R. 1957. *Dynamic Programming*. Princeton, N.J.: Princeton University Press.

Brzinsky-Fay, C., U. Kohler & Luniak (2006) Sequence Analysis with Stata. *The Stata Journal,* 6**,** 435-460.

Burt, O. & J. Allison (1963) Farm Management Decisions with Dynamic Programming. *Journal of Farm Economics,* 45**,** 1963.

Doole, G. J. (2008) Optimal Management of Annual Ryegrass (*Lolium rigidum* Guad.) in Phase Rotations in the Western Australian Wheatbelt. *The Australian Journal of Agricultural and Resource Economics,* 52**,** 339-362.

--- (2009) A Practical Algorithm for Multiple-Phase Control Systems in Agricultural and Natural Resource Economics. *Journal of Agricultural and Resource Economics,* 34**,** 109-129.

El-Nazer, T. & B. McCarl (1986) The Choice of Crop Rotation: A Modeling Approach and Case Study. *American Journal of Agricultural Economics,* 68**,** 127-136.

Heady, E. O. (1948) The Economics of Rotations with Farm and Production Policy Applictions. *Journal of Farm Economics,* 30**,** 645-664.

Hennessy, D. A. (2006) On Monoculture and the Structure of Crop Rotations. *American Journal of Agricultural Economics,* 88**,** 900-914.

Jaynes, E. T. 1963. Information Theory and Statistical Mechanics. In *Statistical Physics,* ed. K. W. Ford. New York: W A Benjamin Inc.

Kennedy, J. (1986) Rules for Optimal Fertilizer Carryover: An Alternative Explanation. *Review of Marketing and Agricultural Economics,* 54**,** 3-10.

Mittelhammer, R. C., G. G. Judge & D. J. Miller. 2003. Ill-Posed Underdetermined Inverse Problems. In *Econometric Foundations*. New York: Cambridge University Press.

Needleman, S. & C. Wunsch (1970) A General Method Applicable to the Search for Similarities in the Amino Acid Sequence of Two Proteins. *Journal of Molecular Biology,* 48**,** 443-453.

Orazem, P. & J. Miranowski (1994) A Dynamic Model of Acreage Allocation with General and Crop Specific Soil Capital. *American Journal of Agricultural Economics,* 76**,** 285-295.

Shannon, C. E. (1948) A Mathematical Theory of Communication. *Bell System Technical Journal,* 27**,** 379-423.

Thomas, A. (2003) A Dynamic Model of On-Farm Integrated Nitrogen Management. *European Review of Agricultural Economics,* 30**,** 439-460.

Wu, J., R. M. Adams, C. L. Kling & K. Tanaka (2004) From Microlevel Decisions to Landscape Changes: An Assessment of Agricultural Conservation Policies. *American Journal of Agricultural Economics,* 86**,** 26-41.