Soybean production is one of the main economic forces driving the expansion of the agricultural frontier in the Brazilian Amazon. To assess the potential for expansion we estimate a model of soybean yield that integrates the major climatic, edaphic, and economic determinants in the Amazon Basin. Yield is modeled as a function of yield as simulated by a crop physiology model that captures the effects of climate and physical attributes on the development of soybean plant; fertilizer applications; and economic/spatial parameters such as credit, transports costs and latitude. Current values of these determinants indicate that roughly 20% of Amazon Region or ∼1,000,000 km² (excluding protected areas) can generate yields greater than 2000 kg/ha. Soybean production may be possible over a wider area of Amazon, but realizing this potential requires improvements in economic determinants such as the transportation infrastructure.

© 2007 Elsevier B.V. All rights reserved.

Keywords: Soybean, Yield model, Amazon, Phenology, Agriculture

1. Introduction

Mad cow disease, Chinese economic growth, and high prices pushed global soybean production from 117 million metric tons in 1992 to 190 million metric tons in 2004 (USDA-FAS, 2004). Much of this increase occurs in tropical regions, including the Amazon Basin. In Brazil, soybean plantings increased from 115,847 km² in 1990 to 215,972 km² in 2004 and annual production increased from 20 to 50 million tons. Because of these increases, Brazil is the world’s second largest producer of soybeans, supplying 27% of the world total, second only to the United States, which supplies about 35% (USDA-FAS, 2004; IBGE, 2005).

Of the Brazilian total, about 33% (16.3 million tons) is harvested in the Legal Amazon¹ (CONAB, 2003; IBGE, 2005), principally in Mato Grosso State, which is covered largely by savannas (Cerrado) and to a less extent by transition forests. In

¹ The Brazilian Amazon region, also known as the North Region, is constituted by tropical rainforest and includes seven states: Amazonas, Acre, Amapá, Pará, Rondônia, Roraima, and Tocantins. Legal Amazon is a larger area including the states of Mato Grosso and Maranhão (the part west of meridian 44°W), which was defined for regional planning purposes. It includes a complex mosaic of tropical rainforest, savannas, inundated lowlands, and steppes (Andersen et al., 2002). In this study we use the terms Amazon region and Legal Amazon alternately but the data set is based on Legal Amazon.
the Cerrado region (Center West Brazil), expanding production is powered by the development of cultivars suited for hot-humid conditions and the application of fertilizer and lime-stone, which raised yields to about 3000 kg/ha (EMBRAPA-COFECARJ, 2002). The Brazilian Government’s last three pluriannual plans (“Brasil em Ação”, “Avança Brasil”, and Brasil um País de Todos”) have improved transport networks and associated infrastructure (e.g. ports, waterways, and hydroelectric power plants), which have increased the economic viability of soybean production in the Amazon Region.

Reducing transport costs encourages private investments in the agricultural frontier (Pfaff, 1996; Nepstad et al., 2000, 2001; Carvalho et al., 2001, 2002). In Brazil, benefits of increased soybean production include increased export earning, increased local employment, and increased productivity (GDP per capita) and welfare. Social costs include the replacement of local communities by capitalized land buyers and an increase in income concentration (Fearnside, 1997, 2001). Environmental costs include the loss of forest area and biodiversity and increased emissions of greenhouse gases.

The importance of these costs and benefits begs the question, how much more land will be converted to soybean agriculture? The answer depends in part on rents for soybean production, which are determined by the price of soybeans, yield, and production costs. Here, we develop an interdisciplinary model that simulates the climatic, edaphic, and economic determinants of soybean yields. When complete, this model can assess spatial variations in the economic viability of soybean production in the Brazilian Amazon and the degree to which expanded plantings can be influenced by the government.

The model for soybean yields in the Brazilian Amazon is described in five sections. Section 2 describes previous efforts to simulate yields, with a special emphasis on interdisciplinary models. The Section 3 describes the data and techniques that are used to estimate our interdisciplinary model. The Section 4 describes the results. The final section evaluates the degree to which our results are consistent with empirical observations.

2. Background

Soybean yields are determined by the climatic, edaphic, and economic environment. Economic conditions in the Amazon Basin may be more important than their role in developed economies because poorly developed transportation networks affect the price of purchased inputs such as fertilizer and the net price of soybeans to farmers. Unfortunately, economic and physical determinants are addressed separately by existing methodologies. Crop simulation models forecast yield by simulating crop physiology, which is driven largely by the climatic and edaphic environment. But these models find it difficult to represent farmer decisions regarding the use of purchased inputs. Regression yield models sometimes include a mix of economic and physical determinants, but the effects of climatic and edaphic determinants rarely are represented in a way that is consistent with crop physiology. Because our method seeks to combine crop simulation and regression yield models, the following paragraphs summarize the strengths and weaknesses of each.

2.1. Soybean crop simulation models

Several crop models simulate physical, chemical, and biological processes in the soybean plant [Glycine max (L.) Merr.] as function of climate, soil, and crop management (Sinclair and Seligman, 2000). The current generation of crop models include: the Soybean Growth Model — SOYBEAN (Sinclair, 1986); the Soybean Simulation Model GLYCIM (Acoc and Trent, 1991); the Soybean Crop Simulator — SOYCRUS (Penning de Vries et al., 1992); and CROPGRO–SOYBEAN (Boote et al., 1998).

SOYBEAN (Sinclair, 1986) uses daily weather data (solar radiation, minimum and maximum temperature, and precipitation) to simulate leaf area index, biomass accumulation, seed growth, and seed weight. The SOYBEAN model is used to study the effect of water and nitrogen limitations on soybean grain production (Muchow and Sinclair, 1986; Sinclair, 1986), nitrogen accumulation and use by soybean plants (Sinclair et al., 2003), and the effects of weather on soybean yield (Sinclair et al., 1992).

GLYCIM simulates yield based on organ-level processes such as photosynthesis, transpiration, carbon partitioning, and organ growth and development. Exogenous variables include data on climate, soil, and management practices, such as planting date, row spacing, and post-planting decisions such as irrigation scheduling, harvest timing, and yield prediction (Acoc and Trent, 1991). The GLYCIM model is used to estimate yields at the county and state levels (Haskett et al., 1999). Soybean growers that use GLYCIM report increases in yield and irrigation use efficiency (USDA-ARS, 2005).

SOYCROS also simulates the growth and production of soybean varieties based on crop physiology, agrometeorology,
soil physics, physiological water stress, and the functioning of roots (Penning de Vries et al., 1992). CROPGRO–SOYBEAN is a deterministic and mechanistic model that is designed to predict yield and the exchange of carbon, water, and nitrogen. To do so, CROPGRO–SOYBEAN simulates primary plant processes (phenological development, photosynthesis, respiration, plant water uptake, biomass growth and partitioning) as a function of environmental variables such as daily temperature, photoperiod, and soil water availability (Boote et al., 1998). CROPGRO–SOYBEAN is used to predict regional yield and production (Jagtap and Jones, 2002) and to study alternative management regimes (Egli and Bruening, 1992), environmental conditions (Curry et al., 1995; Sau et al., 1999; Carbone et al., 2003; Mall et al., 2004), and genetic yield potential (Boote and Tollenaar, 1994). The model also is used to analyze spatial variations in yield (Allen et al., 1996; Paz et al., 1998) and to develop cultivar coefficients (Mavromatis et al., 2001).

Crop simulation models are used to understand how crop systems operate at the daily/hourly temporal scale and at the plot/plant spatial scale. Their main strength is the ability to simulate the interface between the plants and their physical environment and relate these conditions to yield. Despite these strengths, they are not widely used by farmers due largely to differences between public and private sector perceptions and difficulties associated with software designed for use by academic researchers (Welch et al., 2002).

These difficulties are exacerbated by the on-going development of new cultivars, which can be simulated only after long and costly efforts to quantify the cultivar's coefficients. Constraints imposed by time, financial resources, and continuous “improvements” in cultivars prevent scientists from deriving cultivar coefficients from experimental growth analyses before cultivars are marketed (Mavromatis et al., 2001). Another more important restriction concerns the limited number of management practices that the models can represent and the lack of a connection to the macroeconomic environment in which farmers operate (Kaufmann and Snell, 1997).

2.2. Multiple regression yield models

Multiple regression models usually estimate soybean yield as a function of climatic and/or socio-economic variables. Yield response models often are estimated from field data and can be grouped into four categories: (1) those that focus in the effect of environmental conditions, (2) those that emphasize physiological determinants, (3) those that use economic production functions, and (4) those that integrate physical and economic determinants.

One set of regression yield models relate soybean yields to environmental variables such as agrometeorological variables and soil parameters (Ravelo and Decker, 1981; Garcia-Paredes et al., 2000). For example, maximum air temperature, potential evapotranspiration, and soil moisture account for more variation in soybean yields ($R^2 = 0.75$) than other combinations of variables tested (Ravelo and Decker, 1981). For simplicity, these models often specify climatic conditions using monthly values even though the physiological development of crop plants does not follow the human calendar. For example, regression models indicate corn yield is sensitive to temperature and rainfall in July, but only a portion of July corresponds to the phenological stage of seed filling, which is the stage that has an important effect on yield (Kaufmann and Snell, 1997). Moreover, these regression models usually omit variables related to management decisions that influence the yield.

A second set of regression models forecast yield based on physiological characteristics of the soybean plant that are obtained from small plots. At this scale, yield is closely correlated with the length of the seed-filling period, total dry matter, and plant height (Board et al., 1996; Board, 2002). These variables can be obtained easily from small plots; therefore these models are ideal for choosing among cultivars or genotypes. However, inferences from these regression models are restricted spatially (Board, 2002). Furthermore, these regression models exclude the effect of climatic conditions and management decisions on soybean yield.

A third category of regression models use production functions to estimate yield in response to externally applied inputs. These models emphasize the relationship between factor inputs (e.g., capital, labor), purchased materials (e.g., fertilizers, pesticides), and yield. The simplest models specify yield as a linear function of purchased inputs or as a ratio of inputs prices (Houck and Gallagher, 1976). More complex models, like the translog production function, estimate yield as a function of capital, labor, energy, fertilizers, materials, and land (Cooke and Sundquist, 1989). This set of regression yield models is able to simulate an array of management strategies but fail to simulate how climatic conditions may affect yield because climate variables often are omitted from these models. Ignoring the effect of climate implies that the statistical results suffer from omitted variable bias because climate has an important effect on yield (Boyer, 1982).

A fourth category includes multiple regression models that integrate physical, socio-economic, and environment determinants of yield. Hansen (1991) and Kaufmann and Snell (1997) estimate multiple regression models for corn yield. The former uses data for climate, management practices, and land characteristic to estimate corn yield and obtains consistent and statistically significant results. The latter estimates yield by integrating climate data, which are specified for periods that correspond to phenological stages of development of corn, market conditions, technical factors, scale of production, and the policy environment. These models are able to simulate the effect of climate on yield in a way that is consistent with crop physiology while including the effect of the social environment. Chang (2002) investigates the agricultural impact of climate change in Taiwan by integrating climate, technology, management, and land for 60 crops.

3. Simulating soybean yield: integrating physical and economic determinants

We estimate an interdisciplinary model for soybean yield from census tract data for the Brazilian Amazon that integrates the
ability of physiological models to simulate climatic and edaphic determinants of soybean yield with the ability of regression models to simulate economic and spatial determinants. Skipping ahead, the model is given by Eq. (1):

\[
\text{Yield}_i = \beta_0 + \beta_1 \text{MYield}_i + \beta_2 \text{TCost}_i + \beta_3 \text{Credit}_i + \beta_4 \ln(\text{Fertil}_i) + \beta_5 \text{Lat}_i + \beta_6 \text{Long}_i + u_i
\]

(1)

in which, Yield is the average soybean yield (kilogram per hectare) in census tract i for 1995–1996 growing season, MYield is the average soybean yield (kilogram per hectare) simulated by the crop growth model SOYBEAN for census tract i, TCost is the least-accumulative-cost distance (dollars per ton) to ship soybeans from the midpoint of the census tract i to the nearest export port; Credit are total loans granted by grain trading companies and national banks to soybean farmers divided by the area planted in soybeans ($/ha), Fertil is the value of fertilizers including soil correctives such as limestone used to grow soybeans ($/ha), Lat is the latitude at the midpoint of the census tract i, Long is the longitude at the midpoint of the census tract i, and u is the regression error. The motivation and estimation technique for Eq. (1) is described below.

3.1. Motivation

The effect of the climatic and edaphic environment on yield is represented in Eq. (1) using the yield that is forecast by the SOYBEAN model (MYield). For this application, we simulate the SOYBEAN model with daily data for solar radiation, precipitation, minimum temperature and maximum temperature for the 1995–1996 growing season. These data are obtained from the NCEP–NCAR reanalysis project, which uses climate models to interpolate spatially and temporally sparse ground-based measures (NASA/NCEP/NCAR, 2004). Edaphic conditions are represented by rooting depth, which is the estimated depth to which root growth is unrestricted by physical or chemical impediments and is based on discrete classes (FAO, 1990). These data are obtained from a soil map at SOTELAC (A Soils and Terrain Digital Database — SOTER) for Latin and Central America and the Caribbean whose original scale is 1:5,000,000 at 0.5° resolution for four depth categories (very shallow <30 cm; shallow 30–50 cm; moderately deep 50–150 cm; very deep >150 cm) (ISRIC, 1998). These categories are converted to values (e.g. 15 cm, 40 cm, 100 cm, and 150 cm) for use by the SOYBEAN model.

The relationships among climatic and edaphic variables and yield probably are highly non-linear and vary over the phenological development of the plant therefore, using the SOYBEAN model to calculate their effect probably is more effective than specifying climatic and edaphic variables in Eq. (1). For example, the total water requirement of soybean plants is 450–800 mm/cycle, but requirements vary during development. During the R1 (beginning bloom) and R4 (pod filling) reproductive phases, soybean plants need 7–8 mm/day of water and any deficit during these phases affect physiology in a way that reduces yield (EMBRAPA-SOJA, 2002).

Transportation costs play a fundamental role in soybean production because they account for about 30% of soybean production costs. As such, transportation costs affect yields via the marginal production of purchased inputs (Kaufmann and Snell, 1997). High transportation costs reduce the price for soybean that is received by farmers, which reduces the optimal level of applied inputs, such as fertilizers and soil treatments. This effect is reinforced by high prices for purchased inputs in areas with high transportation costs. Nor can these effects be offset easily by substitution among factors of production (Parré and Ferreira Filho, 1998). Together, these effects imply that yield should be negatively related to transportation costs.

The ability of farmers to apply purchased inputs and the economic viability of their use is influenced by the availability of credit, which is issued by transnational grain companies and national banks (Fearnside, 2001; Alencar et al., 2004). At planting, soybean growers often make forward sales to trading companies, such as Cargill, Bunge, ADM and AMaggi Group, in return for seed, fertilizer and crop protection products. These future contracts, which are known locally as green soybean contracts ("contratos de soja verde"), finance about 60–65% of soybean area that is cultivated in the state of Mato Grosso, which is the largest producer in the Amazon (Gasques, 2003; Cadier, 2004). Increasing credits increase the quality and quantity of purchased inputs and promote investments in modern farm machinery. Consistent with these effects, credits are expected have a positive effect on soybean production.

The low fertility of the Brazilian Amazon soils makes impossible to produce soybeans at commercial level without fertilizers and soil treatments. In the Cerrado, limestone is crucial because the soil’s high acidic limits nitrogen fixation, which is a critical determinant of soybean yield. From a physical perspective, the effect of fertilizers and other soil treatments on yield is straightforward — increasing fertilizer applications increase yield. But fertilizer use and yield are jointly determined. Farmers’ decisions about the rate of fertilizer applications depend on its marginal effect on rent, which is determined by the marginal effect on yield, the price of soybeans, and the price of fertilizers. Based on this economic calculus, farmers apply fertilizer to areas only where the value of the increased yield is greater than the cost of the additional fertilizer. This simultaneity is addressed using instrumental variables (Section 3.2).

Crop yield is also affected by photoperiod or day-length. Originally, Brazilian soybeans were cultivated between 20°S and 30°S because US cultivars adapted well to the local climate and soil (EMBRAPA-SOJA, 2002). Recent expansion into low-latitudes (< 25°) including areas near to the Equator is possible due to cultivars that include long-juveniles genes, which delay flowering and maturity. Without long-juvenile genes, soybean...
Fig. 1 – Amazon soybean-producing census tracts from Brazilian Agricultural Census 1995–1996.
plants grown at low-latitudes flower too soon, which generates short plants that are difficult to harvest mechanically and have low yields (Hartwig and Kühl, 1979; Sinclair et al., 2005). Despite gains, soybeans are short-day plants that are less productive in low-latitudes therefore we include latitude to capture the positive effect of photoperiod on soybean yield. Finally, the model specifies longitude to represent spatial variations in an east–west direction or the effect of spatially biased estimates for variables that are included in the SOYBEAN model.

3.2. Estimation technique

The two-way relationship between soybean yield and fertilizer use is exogenous to observed yield. Consistent with this definition, Eq. (2) specifies instrumental variables that represent edaphic conditions, which influence the farmer’s willingness to apply fertilizer:

\[
\ln(\text{Fertil}) = \beta_0 + \beta_1 \text{Rdepth}_i + \beta_2 \text{pH}_i + u_i
\]

in which Fertil is as defined above, Rdepth is the soil rooting depth as described above, pH is the soil’s pH, and \(u\) is the regression error. Data on soil pH are available at 0.5° resolution (ISRIC, 1998). Values range from 3.9 (strongly acidic) to 8.6 (strongly basic). Data on rooting depth are reclassified as 1=effective rooting depth greater than 50 cm and 0=no-effective rooting depth less than 50 cm. These categories are based on empirical studies, which indicate that rooting depth increases with the farmer's willingness to apply fertilizer. The instrumental variable seeks to isolate the movements in fertilizer use that are exogenous to observed yield. Consistent with this definition, Eq. (2) specifies instrumental variables that represent edaphic conditions, which influence the farmer’s willingness to apply fertilizer:

\[
\ln(\text{Fertil}) = \beta_0 + \beta_1 \text{Rdepth}_i + \beta_2 \text{pH}_i + u_i
\]

4. Results and discussion

4.1. Instrumental regression for fertilizer requirements

Farmers in forty-one of the eighty-eight census tracts apply fertilizers and limestone to soybeans, which implies that the natural log of fertilizer use is undefined for forty-seven census tracts. As a result, the instrumental variable equations (Eq. (2)) and the interdisciplinary yield model (Eq. (1)) are estimated from forty-one observations. We recognize that this is a small sample. Unfortunately, there is no empirical evidence that a linear specification can be used to represent the effect of fertilizer use on soybean yields.

The coefficient associated with Rdepth in Eq. (2) is positive (Table 1), which indicates that farmers apply more fertilizer to deep soils. Ceteris paribus deep soils support higher yields than their shallower counterparts; therefore, at the margin an increase in fertilizer use probably generates a larger increase in soybean yield.

Table 1 – Regression results for ln(FERTILIZERS)

<table>
<thead>
<tr>
<th>Independent variables</th>
<th>Coefficient</th>
<th>t-statistic</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>-2.332</td>
<td>-1.210</td>
<td>0.233</td>
</tr>
<tr>
<td>RDEPTH</td>
<td>2.561</td>
<td>2.733</td>
<td>0.009</td>
</tr>
<tr>
<td>pH</td>
<td>0.550</td>
<td>1.693</td>
<td>0.098</td>
</tr>
<tr>
<td>R²=0.21</td>
<td>Regression F(2,38)=5.12[0.01]</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

8 The same argument of simultaneity could be made about the relationship between yields and credits. One alternative could be using as instrument the bank agencies in the municipalities but in the case of soybean the main source of funding comes from private trading companies such Cargill, Grupo AMaggi, ADM, and Bunge, which finance roughly 65% of Mato Grosso’s soybean production. Unfortunately, we do not have these data at the census tract level and could not develop a satisfactory set of instruments.
in yield on deeper soils. The larger marginal product would increase applications on deep soils.

The coefficient associated with pH is positive (Table 1), which indicates that farmers apply more fertilizer to basic soils. A positive relationship with pH can be explained by the higher cation exchange capacity of basic soils. Everything else being equal, the high cation exchange capacity of basic soils allows them to store fertilizer more effectively than acidic soil, and this increases the marginal product of fertilizer applications. Notice that this explanation runs counter to a simple physical explanation in which lower cation exchange capacity of acidic soils lowers their nutrient content. Following this logic, acidic soils “need” more fertilizer to achieve the same yields as a less acidic soil. In other words, these regression results indicate that decisions about fertilizer applications are based on economic criteria, not physical criteria.

Rooting depth and pH account for about 21% of the variation in fertilizer use among census tracts. Consistent with this relatively low R-squared, the F statistic for Eq. (2) is 5.1. Values below 10 generally indicate that the instrumental variables are weak. Under these conditions, the usual t-tests on the second stage regression coefficients cannot be evaluated accurately against a standard normal distribution (Staiger and Stock, 1997). We do not believe that weak instruments cloud the interpretation of estimation results for Eq. (1). As described below, the t-statistic that tests the null hypothesis that fertilizer use has no effect on soybean yields is much larger than the critical value at the five percent threshold.

4.2. Estimate of soybean yield model

The vector of climatic, edaphic, economic, and spatial variables accounts for about 48% of the variation in yield among the forty-one census tracts (Table 2). The coefficients all have signs that are consistent with theory. The positive sign associated with MYield indicates that climatic and edaphic conditions, which increase a plants’ physiological ability to produce seeds, increase yield. The value of $\beta_1$ is 0.073, which indicates that at the margin, changes in climatic and edaphic conditions that increase potential yield by 1 kg per hectare increase observed yield by about 73 g. One explanation for this relatively low crop yield potential is a lack of hybrids that can tolerate high levels of aluminum and iron that are present in local soils and relatively high levels of disease (Leibold et al., 2001). Another explanation focuses on an assumption that is implicit in the SOYBEAN model — optimal management. If farmers do not apply the required inputs, these inputs limit yield in a “Liebigian” sense, and regional variations in non-limiting factors, such as climatic and edaphic conditions, will have a relatively small effect on yield.

The sign on $\beta_2$ in Eq. (1) indicates that transportation costs (TCost) have a negative effect on yield. The poor road network in the Amazon Basin increases the costs of shipping soybean, which reduces the rent that farmers receive, which in turn reduces the economic viability of applying inputs. High transportation costs also increase the price of these agricultural inputs at the place of application, which also discourages their use. As described below, improving the transportation network can raise yields significantly.

Fertilizer use (InFertil) has a positive effect on yield. Applying fertilizers and other soil treatments increase the number of seeds by unit of surface. The application of fertilizer and/or lime is one of the main strategies used by soybean farmers to raise yield because most of soils in the Amazon region are acid and highly deficient in P, N, Ca, K, and Mg (Nicholaides et al., 1983).

Latitude (Lat) also shows a positive effect on yield. Consistent with the physiology of soybeans, the long days that characterize the early portion of the growing season in low-latitudes are not optimal for soybean production. Over the last decade, savannas (Cerrado region) became the main agricultural frontier within Brazil due to the development of soybean grain varieties that are adapted to low-latitudes and hot-humid tropical conditions.

Finally, the positive effect of longitude (Long) on soybean yield indicates that ceteris paribus, soybean yields increase from east to west. We cannot identify a climatic or physiological basis for this effect. Rather, we speculate that this effect represents an omitted variable that varies systematically in an east–west direction or the effect(s) of spatially biased estimates for variables that are included in the SOYBEAN model or Eq. (2).

The relative importance of climatic, edaphic, and economic determinants is quantified by generating four “predicted” values for soybean yield in which observed values are used for all but one of the independent variables. The change in the residual sum of squares is used to represent the fraction of total variation that is “explained” by the variable held at its sample mean. Results indicate that the climatic and edaphic determinants, which are embodied in MYield, account for about 16% of the spatial variation in soybean yield in the 1995–1996 growing season. Fertilizer use, transportation costs, and credit account for another 5, 16 and 10 percentage points respectively.

### 4.3. Forecast of soybean potential area in Amazon

The explanatory power of climatic and edaphic determinants implies that the interdisciplinary model can be used to identify areas where high soybean yields may drive future rates of deforestation. To calculate potential yields, we simulate Eq. (1) with data for the entire Amazon basin that includes average MYield values estimated from average climate data 1950–2001 (NASA/NCEP/NCAR, 2004), transport cost values for 1995–1996

<table>
<thead>
<tr>
<th>Independent variables</th>
<th>Coefficient</th>
<th>t-statistic</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>CONSTANT</td>
<td>-4113.26</td>
<td>-3.27</td>
<td>0.00</td>
</tr>
<tr>
<td>FERTIL</td>
<td>214.96</td>
<td>2.51</td>
<td>0.01</td>
</tr>
<tr>
<td>MYIELD</td>
<td>0.07</td>
<td>2.50</td>
<td>0.01</td>
</tr>
<tr>
<td>TCOST</td>
<td>-5.08</td>
<td>-3.25</td>
<td>0.00</td>
</tr>
<tr>
<td>CREDIT</td>
<td>1.57</td>
<td>3.05</td>
<td>0.00</td>
</tr>
<tr>
<td>LONG</td>
<td>99.72</td>
<td>3.05</td>
<td>0.00</td>
</tr>
<tr>
<td>LAT</td>
<td>47.89</td>
<td>1.68</td>
<td>0.09</td>
</tr>
</tbody>
</table>

$R^2 = 0.48$

FERTIL = predicted values of fertilizers, MYIELD = predicted values of yield (kg/ha), TCOST = transport cost values ($/ton), CREDIT = total credit for soybean farmers ($ per ha of soybean), LONG = longitude, and LAT = latitude.
Fig. 2 – Amazon soybean potential as predicted by yield model (Eq. (1)). Yellow areas represent those places where yields could be greater than 2000 kg/ha. Red polygons represent municipalities currently raising soybean crops. Indigenous lands and protected areas are represented by dot-green polygons (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.).
Fig. 3 – Soybean profitability estimated from the yield model. Yellow areas represent those places where economic returns could be greater than $200/ha (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.).
At a simple level, the Amazon Basin includes the States of Roraima, Acre, Pará, Maranhão, and Tocantins. The Amazon Basin currently is covered by primary forests and the remaining thirty percent (430,000 km²) are open areas. Roughly seventy percent of these open areas are located in the State of Mato Grosso. This concentration confirms the climatic, edaphic and economic suitability of Mato Grosso to raise soybeans. As of 2004, 28% of Mato Grosso’s total area is deforested (256,000 km²), of which soybean production occupies 53,000 km² and cattle ranching predominates in the remaining area (IBGE, 2005; INPE, 2005). This State accounted for 47% of the 26,130 km² of Amazon deforestation that occurred in 2003–2004 (INPE, 2005), most of which was caused by the expansion of these two economic activities. The boom in soybean cultivation is mainly an indirect cause of deforestation in Mato Grosso. Pasturelands often are converted to soybean production due to lower production cost. The displaced cattle are moved to new pasture, which is created by converting forested areas (Alencar et al., 2004). Improvements of roads network and the construction of grain storage facilities in northern Mato Grosso have encouraged direct conversion of forest to soybean crops (Morton et al., 2005).

The model also indicates that an area equal to 360,000 km² with the potential to generate more than 2000 kg/ha of soybean is located on indigenous lands and protected areas of Amazon. Most of these areas currently are “off limits” to soybean production, but infrastructure improvements could make these areas vulnerable. This vulnerability implies that the government may have to enforce regulations and step up monitoring in order to restrain soybean expansion. Excluding these indigenous land and forest reserves, the soybean potential area drops from 1.4 to one million square kilometers or 20% of Amazon Basin. Mato Grosso State holds 44% of this potential area, Amazonas 21%, and Rondônia 13%. The remaining 22% of this soybean potential area is shared among the States of Roraima, Acre, Pará, Maranhão, and Tocantins.

To narrow this area further, we use soybean yield predictions to estimate a surface of soybean profitability, that is, the rent or economic returns obtained by planting soybean crops. At a simple level, rent is the difference between the revenues generated by soybean activity (yield * price) minus total costs.12

Fig. 3 shows that approximately one million square kilometers of Amazon Basin (excluding protected areas) have high rent potential (greater than $200/ha). Although the area with potential for high rents is equivalent to the area of high productivity (2000 kg/ha), its spatial distribution changes drastically. Of the area with a potential for high rents, 60% is located in open lands and 40% in forest lands. Most (85%) of areas with highest economic rents are located in Mato Grosso, Pará, Maranhão, and Tocantins States. These areas coincide with the location of roads, which reduce cost of shipping soybeans.

5. Conclusions

Our results indicate that soybean yields in the Amazon are determined by climatic, edaphic, and economic conditions. Quantifying these effects allows us to assess the effect of soybean production on the current pattern of deforestation and assess the degree to which government investments in transportation infrastructure and agricultural credits can affect future rates of deforestation.

We emphasize that Fig. 2 shows yield under current economic conditions. If the Brazilian government continues to make significant investments in transportation infrastructure, the area where soybean production exceed $2000 kg/ha could expand significantly, and this could accelerate deforestation further. This potential is consistent with previous studies, which indicate that transportation infrastructure is the single most robust predictor of frontier expansion and accompanying deforestation in tropical forest regions (Kaimowitz and Angelsen, 1998). More than two-thirds of Amazon deforestation takes place within 50 km of major paved roads, where agriculture, cattle ranching, and logging activities are economically feasible (Nepstad et al., 2001; Alves, 2002). Indeed, several studies forecast deforestation as a simple function of distance from roads (Laurance et al., 2001).

The potential for enhanced yields and deforestation will be highlighted in future research that quantifies changes associated with the full implementation of the government’s pluriannual plans, which lay out the strategies for economic development in the Amazon (Nepstad et al., 2000, 2001; Carvalho et al., 2001, 2002). Plans call for the construction of roads, ports, and other energy infrastructure that will lower transportation costs and thereby raise soybean rents. Increases in yield and reductions in transportation costs would increase rents, which is one of the main determinants of land-use change.

Identifying the effect of transportation costs and other economic determinants of soybean yields and ultimately, soybean rents, is critical to determining the degree to which decision makers can control deforestation rates. By evaluating specific components of Brazil’s pluriannual plans, policy makers may be able to compare the benefits of increased soybean production against environmental impacts. By shifting such construction efforts, controlling government credits, or taxing soybeans, a “market-based” approach to land-use change ultimately may be more effective and more efficient than a “command and control” approach that seeks to preserve the forest by mandating protected areas and relying on local governments for enforcement.
Acknowledgements

We gratefully acknowledge the financial support for this work that was provided by Gordon and Betty Moore Foundation, National Science Foundation, LBA-NASA Ecology, and US Agency for International Development. We also thank T. R. Sinclair for his valuable comments and for providing us the code of his soybean crop simulation model, Georgia Carvalho for her suggestions on the final version of this paper, and Eustáquio Reis for providing us 1995–1996 Agricultural data at the census tract level.

REFERENCES


