Colombia Purdue Partnership

CPP_LP19-2

Orinoquía Agricultural Linear Programming Model – Documentation

Paul V. Preckel
Professor, Department of Agricultural Economics
Purdue University

Carlos Fontanilla
PhD Student, Department of Agricultural Economics
Purdue University

J. Lowenberg-DeBoer
Professor Emeritus of Agricultural Economics, Purdue University
Elizabeth Creak Chair in Agri-Tech Economics, Harper Adams University

John H. Sanders
Professor, Department of Agricultural Economics
Purdue University

Email Contact: orinoquia@purdue.edu

February 5, 2019

This report documents the linear programming model developed for the Orinoquía Initiative. The Orinoquía is a diverse region with varying natural resources (e.g., land types, weather, access to irrigation water, etc.) and farms of different scales and resource endowments (e.g., land, family labor, access to hired labor, etc.). The model was designed in a general manner that allows the same model structure for different regions and farm types within the Orinoquía.

I. Introduction

This report is structured as follows. A first section describes provides an introduction. The second section describes the conceptual framework for the model organization. The third section describes the model variables, and the fourth section describes the structure of the model equations. A fifth section describes the organization of files, and how to run the model using the GAMSIDE software. Finally, a sixth section describes how model results can be used to determine farm-level recommendations. An appendix describes the layout of the spreadsheet file that is used to store the data for the model.

II. Conceptual Framework and Data

The model, Orinoquía LP, is a one-year steady-state model of farm-level agricultural production possibilities. Orinoquía LP is an optimization model that has the goal of determining the portfolio of production enterprises that maximizes the revenues less variable costs on the farm, subject to limitations on the resources available on the farm. In order to track the use of resources throughout the year, the year is divided into monthly *Periods*. By steady-state, we mean that the model determines a set of enterprises that use resources in a manner that is consistent with repeating the plan year after year. While this is straightforward for a continuous annual crop that is grown year after year, for rotations. By way of illustration, we take a two-year rotation of corn and soybeans as an example.

Because this is a one-year model, area in each crop in the rotation (e.g. corn and soybeans) must be planted in our steady-state year. Rotations typically need to be treated differently from simply planting fractions of the land area in each of the continuous crops (e.g. half a hectare of continuous corn and half a hectare of continuous soybeans per hectare of rotation corn and beans) because the agronomics of the crops grown in rotation are different from the continuous case. The idea is that a fraction of each acre that is planted in the rotation is in each of the crops in the rotation, but the inputs and outputs for the rotation crop are adjusted to reflect the agronomics of the rotation.

This rotation concept also extends to tree crops. Thus cacao, with a ten-year productive life, is viewed as a ten-year rotation. For each hectare planted to cacao, one tenth of a hectare is treated as first year cacao, one tenth of a hectare is treated as second year cacao, and so on with, at the end, one tenth of a hectare treated as tenth year cacao. Thus, the uses of inputs and the production of outputs are averages over the ten years in each period. This means that the per hectare use of labor in January is equal to the average of per hectare labor use in January across the ten years. Similarly, per hectare output of cacao in October is equal to the average October output per hectare across the ten years.

This steady-state perspective implies that the model results should be viewed as a goal – a long-run ideal portfolio for organization of the farm enterprises. This portfolio should be viewed as a target. The model does not give results that specify how a farm can get from their current situation to the ideal portfolio. Ideas for how to use the model results in conjunction with additional analysis to determine the transition path towards the target are addressed in the sixth section of this document.

The productive activities in this model are referred to as *Enterprises*. Examples of enterprises would be: continuous corn, a two-year rotation of corn and soybeans, a ten-year rotation of cacao, beef fattening, dairy with processing of a fraction of the milk to yogurt. Because some of the enterprises produce multiple outputs, we introduce the concept of *Commodities* into the model. Commodities are the outputs of enterprises. Their principle use is sales to generate revenue on the farm, but they may also be used as inputs to other enterprises.

Farm-level *Resources* include land, labor, and monetary capital. Uses of these resources are constrained by their availability on a period-by-period basis. Many of the aspects of region within the Orinoquía and farm size are defined by the farm's endowments of these resources. *Family Labor* is defined as a number of workers (adult male equivalent). Similarly, if the farm has existing commitments to *Permanent Hired Labor*, that is also expressed in a number of workers. The difference between these is that in the accounting of revenues net of variable costs in the model objective, the salaries for the permanent hired labor are treated as costs, while there is no deduction from revenues for family labor. In addition to permanent labor, *Temporary Labor* can also be hired. The distinction between these is that, while hiring permanent laborer generates salary costs and provides labor resources in every period, temporary labor can be hired in just the periods when it is needed. To reflect the fact that labor markets may have limited workers available, maximums may be specified for the number of permanent or temporary workers that can be hired. The limit for permanent workers is expressed in workers (adult male equivalent) per year, while the limit for temporary workers is expressed in man-days (adult male equivalent) per month.

An important aspect of labor availability is that labor is only effective when field conditions are such that work can be done. This gives rise to the concept of *Good Field Days* the number of days per period that historical records indicate should be available for fieldwork. For example, if a crop must be planted in a given period with a requirement of 1 man-day per hectare, and if only one family worker and no hired workers are available, then if the number of good field days in the period is 8, then at most 8 hectares of that crop can be planted.

Land is another important endowment of the farm. The model is designed to treat land as heterogeneous. In our initial model datasets (documented in companion reports), just two land types are considered: land equipped for irrigation, and non-irrigated land. The same crop grown on different types of land are treated as different enterprises. For example, oil palm grown on irrigated land is a separate enterprise from oil palm grown on non-irrigated land. These crops have different input requirements and produce different amounts of output.

The final resource that is explicitly tracked in the model is monetary capital or *Cash*. The farm is endowed with an initial amount of cash at the beginning of the one-year time horizon (January 1 in the initial versions of the model). Each period, monetary capital is increased by commodity sales, decreased by cash expenses for enterprises and hiring expenses. Any accumulation of revenues in excess of expenses is passed forward to the next period as retained earnings. To facilitate cash flow, borrowing is permitted. Borrowing is treated as a period-wise operating loan, for which the principle plus interest is repaid the next period. To reflect possible credit constraints, there is also a limit on the maximum amount that can be borrowed in each period.

The data that describe the enterprises reflect their use of resources (and potentially commodities) and their production of commodities. These are specified in tables on worksheets in a data spreadsheet. For example, each enterprise on each land type has a level of *Labor Use* each period. Similarly, each enterprise on each land type has a level of *Cash Use* each period. Because of the climate in the Orinoquía, it may be possible to grow more than one crop on the same piece of land in the same year in sequence. For example, corn or rice may be grown in the early rainy seasons (roughly March to July) followed by soybeans in the latter part of the rainy season (roughly August to November). For this reason, each enterprise on each land type has a level of *Land Use* in each period. So, if a crop only occupies the land from March through July, the land would be available August through February to grow another crop.

Another table of data makes the link between commodity outputs and enterprises. This table specifies for each enterprise on each land type the production of each commodity in each period. Note that many enterprise/land type combinations produce commodities in multiple periods. In addition, some enterprise/land type combinations produce multiple commodities. For example, the pineapple enterprise (on either irrigated or non-irrigated land) produces high quality pineapples and low quality pineapples. These must be treated as different commodities because the quality is an important determinant of pineapple price.

A parallel table links commodity inputs and enterprises. This table specifies for each enterprise on each land type the input of each commodity in each period. This could be important if it was desired to incorporate on-farm milk processing into yogurt or cheese. While the raw milk would be a commodity produced by a dairy enterprise, the raw milk would be an input to the yogurt-making enterprise.

The last piece of data relevant to commodities are their selling prices. Commodities are treated as uniform. That is, the price of the commodity is not dependent upon the enterprise that produced it. The yield of the commodity will be affected by the choice of enterprise and land type, but the selling price is the same no matter how or on what land it is produced. Thus, there is one selling price for each commodity, which is independent of time. (Commodity purchasing and storage are not permitted in the current version of the model.)

III. Model Variables

Model variables, or activities, are included for the enterprise production, commodity selling, retaining cash from period to period, borrowing from period to period, permanent hiring for the entire year, temporary hiring by period, and for end of year net return. Enterprise production is handled through the variables called *produce*, and is indexed by land type (l) and enterprise (e). In our datasets, the inputs and outputs for these activities are expressed per hectare of land. Sales are handled through variables called *sell*, which are indexed by commodity (c) and period (t).

The unit of the *sell* variables depend on the commodity (e.g. in our datasets this is kilograms for citrus, ton for cacay, carga for coffee, and cajita for guayaba). These units are documented in the spreadsheet in the table on the worksheet that indicates the outputs (commodities) produced by the different enterprises on alternative land types by period. Prices of commodities must be in consistent units so that the product of the commodity units (e.g. cargas) times the price units (e.g. thousand pesos per carga) equals thousands of pesos.

There is a single variable for hiring of permanent labor hiring (*phire*). This variable is expressed in man-years. The level of this variable in the solution is the *total* number of permanent workers hired including both existing permanent labor and additional hires. The number of existing permanent workers is the lower bound on this variable. The maximum number of additional permanent workers plus the number of existing permanent workers is the upper bound on this variable.

There is a variable for hiring temporary labor (*thire*) for each model period, indicated by the period index (*t*). The units of these variables are man-days in the specified period. Lower bounds on these variables are zero, and the user can specify the maximum amount of temporary labor hiring in any period through the spreadsheet data.

Because the various enterprises will have inflows and outflows of cash, potentially every period, there is a variable (*save*) indexed by time period (*t*) that keeps track of the retained earnings (cumulative revenues less current expenses plus cumulative borrowings less loan repayments) retained from period *t* to the next period. These variables are measured in thousands of pesos. Borrowing is handled via a variable (*borrow*) indexed by time period (*t*) the represents the amount of cash borrowed in period *t* and repaid with interest in the next period.

The final variable is the net return or revenues less all cash costs (*netrev*) at the end of the year. This is equal to retained earnings at the end of the planning horizon minus repayment of final period loans plus interest. The units of this variable are thousands of pesos.

For any of the model variables, if the level of the variable equals its lower bound, the marginal (shadow price) on that variable is the amount that net return would decrease per unit of increase in that variable. If the upper bound is active, the marginal (shadow price) on that variable is the amount that net return would increase per unit of increase in the upper bound.

IV. Model Equations

Model equations ensure that the uses of resources and commodities do not exceed their sources, and define the objective for the problem. The first resource is land. The uses and sources of land are governed by the *land* equation, which is indexed by land type (*l*) and time period (*t*). The sources of land are fixed endowments by land type. There are no provisions in the current model to all for renting land, and there is no facility for converting land from one type to another (e.g. making investments to equip a non-irrigated parcel for irrigation). The uses of land are for the different enterprises. That is, we add across enterprises the level of the produce activity times the land use parameter (*lu*) for that land type (*l*), enterprise (*e*), and period (*t*). Because a short-season annual crop may not use the land for all periods within the year, this accounting for the land uses allows for multiple crops to be grown in sequence on the same plot of land, providing that their growing periods do not overlap.

The second resource is labor. Sources and uses of labor are governed by the *labor* equation, which is indexed by time period. The sources of family labor (*flab*), permanent hired labor (*phire*), and temporary hired labor (*thire*). Family and permanent hired labor are both measured in workers, and so for each worker there is a man-day of labor available for each good field day (*gfd*) for the period (*t*). Temporary hired labor is expressed in man-days of labor, and so the sum of man-days from family and permanent hired labor plus temporary hired labor gives us the total man-days available for allocation. As noted above, the level of the variable *phire* in the solution is the *total* number of permanent workers hired including both existing permanent labor and additional hires. The number of existing permanent workers is the lower bound on this variable. The maximum number of additional permanent workers plus the number of existing permanent workers is the upper bound on this variable. The uses of labor are for the different enterprises. That is, we add across land types (*l*) and enterprises (*e*) the hectares of the enterprise times the labor use (*wu*) for that land type, enterprise and time period (*t*).

The third resource is commodities. Sources and uses of commodities are governed by the *comuse* equation, which is indexed by commodity (*c*) and time period (*t*). Unlike land and labor, there is no possibility of a fixed endowment of commodities. The source of commodities is production. Total production of a commodity (*c*) is obtained by summing across land types (*l*) and enterprises (*e*) the output (*entcom*) of commodity (*c*) in period (*t*) from enterprise (e) on land type (*l*). There are two potential uses of commodities – sales or use as inputs to enterprises. There is a sales activity (*sell*) for each commodity and time period to reflect the first use. For the use as an intermediate input, we sum over land type (*l*) and enterprise (*e*) the intermediated input use (*fu*) of commodity (*c*) in period (*t*) for land type (*l*) for enterprise (*e*) per hectare of production times the hectares of production of enterprise (*e*) on land type (*l*).

The fourth and final resource is cash, for which uses and sources are governed by the *cash* equation. There are several sources of cash. The first is initial cash available at the beginning of the planning horizon, or first period of the model. This source appears only as a source in the first model period. The second source of cash in period (*t*) is from commodity sales. These are obtained by summing across commodities the selling price (*sprc*) for the commodity (*c*) times the quantity sold (*sell*) of the commodity (*c*) sold in the period (*t*). As noted earlier, the units of commodities may vary, but the units of the prices vary in a consistent way so that the product of commodity price times quantity equals thousands of pesos. The third source of cash is retained net returns (*save*) from the previous period ($t – 1$). The fourth and final source of cash is loans (*borrow*) originated in the current period (*t*), provided that it is not the final period in the model. There are also several uses of cash. The first use of cash is for variable costs, excluding labor, (*cu*) per hectare enterprise (*e*) on land type (*l*) in period (*t*) times the hectares of enterprise (*e*) on land type (*l*). The second use is equal to permanent hired laborers (*phire*) times the annual salary of the laborer (*pwlab*) measured in thousands of pesos per year divided by the number of model periods. This distributes the expenses for hired labor equally across the time periods. The third use is equal to the man-days of temporary hired labor (*thire*) times the daily wage for temporary labor (*twlab*) measured in thousands of pesos per day. The fourth use is for repayment of the loan (*borrow*) from the previous period ($t - 1$) times one plus the rate of interest (*intrst*) reflecting repayment of the principle plus interest. The fifth and final use is for replacing the initial cash (*initcash*) that was available at the beginning of the year. This term appears only in the final period of the model and ensures that the production plan will be sustainable from year to year – i.e. that the plan does not draw down the initial cash from year to year.

The final equation defines the objective variable for the linear programming problem. This variable is the net return (*netret*) left at the end of the year and equals the retained returns in the final model period (*fnt*).

In addition to these general equations, a number of bounds on individual variables are included. These include lower bounds (*lobd*) and upper bounds (*upbd*) on the enterprises by land type (*l*) and enterprise (*e*). In addition, there are upper bounds (*thlab*) on temporary hired labor. As mentioned above, the lower bound on permanent hired labor (*phire*) is equal to the existing permanent hired labor (*plab*), and the upper bound on permanent hired labor is the sum of existing permanent hired labor and the maximum allowable additional permanent hired labor (*plab* + *phlab*). Finally, there is an upper bound on borrowing in each period, which is in thousands of pesos.

V. Managing Data and Running the Model with GAMSIDE

The organization of the Orinoquía Agricultural LP Model software divides the inputs into two parts. The first of these is the GAMS source program. The source program should not be modified except for the third line of the program as will be explained below. The second part of the inputs is in the form of a spreadsheet that contains a series of worksheets that contain the data for the linear program.

There are four main types of GAMS files, and each has a different file extension. These are: (1) GAMS project files (extension .gpr) which contain information related to your current session; (2) GAMS program files (extension .gms) which contain GAMS source programs; (3) GAMS listing files (extension .lst) which contains the standard output of a GAMS program; and (4) GAMS log files (extension .log) which contain summary information from GAMS and the solver. The Windows folder in which the GAMS project file is located is especially important. This is because this is where GAMS will look for the data spreadsheet for the Orinoquía Agricultural LP Model. GAMS will also place the listing files, log files, and a number of temporary files that are not important to the user during a model run.

It is important to develop a file naming convention that will facilitate keeping track of the work. One issue is that whenever a user requests GAMS to run a model, it is saved with the current file name. That means if the user starts with a working program, modifies it and reruns it, it may be difficult to reproduce the original working program. Thus, when starting a new experiment, it is best to (a) open the GAMS source file, (b) immediately save it with a different name, (c) begin the process of modification and debugging. To make the files manageable, it is worthwhile to develop a file naming convention. The naming convention that we have settled on is a short name for the region and a farm size letter (S,M for small/medium and L for large) followed by the date in the form yyyymmdd, for example AltillanuraM20171210.gms would be the GMS source file name for the Altillanura region, medium farm size created on December 10, 2017. We name the data spreadsheet similarly, with the spreadsheet that goes with the indicated source program for the Altillanura and indicated date named AltillanuraM20171210.xlsx.

As mentioned, the GAMS source file should not be modified with one exception – the third line of the program. The purpose of this line is to ensure that the GAMS program uses the correct data spreadsheet. For the example above where the GAMS source file is named AltillanuraM20171210.gms and the data spreadsheet is named AltillanuraM20171210.xlsx, this line should read:

```
$call copy AltillanuraM20171210.xlsx Orinoquia_Tables.xlsx
```

Any time the relevant spreadsheet name changes, this line must also be changed so that the correct data is used. Again, this is the only line in the GAMS source program that should be changed.

VI. Considerations for Using Model Results

Care is needed for interpreting and using the results of the Orinoquía Agricultural Linear Programming Model. The first consideration is that this model is designed to simulate the steady state situation. This is important because the feasibility of the use of farm resources may depend on a constant portfolio of enterprises being pursued each year.

So, conceptually the model has been designed to simulate the situation at a point in time far enough in the future that all rotations are fully operational. This means that a hectare of a 50-50 rotation of corn and soybeans would plant half a hectare in corn and half a hectare in soybean each year. This is of special concern for the tree crops that may have very long cycles. Recall that a ten-year tree crop is modeled as a ten-year rotation, and it would take ten years to put such a rotation in place (with one tenth of the area planted in the first year, one tenth in the second year, and so on).

In this sense, the Orinoquía Agricultural Linear Programming Model produces are portfolio of enterprises that “fit together” in terms of resource use that can be viewed as a goal or target. Additional analysis is needed to determine how to make the transition from the current production system on the farm to this long-run target. There are several considerations in developing a transition plan to achieve the target. One is that there may be one-time start-up costs that are not included in the steady-state rotation costs that should be taken into account before investing, such as building roads for internal transportation in the farm, irrigation pump stations, or other type of investments that can last for several cycles of the agricultural alternative. For example in the development of an irrigated crop the installment of the irrigation pump station will be a one-time expense that should not be included in the variable costs in the steady-state model. On the other hand, maintenance of the pump station is an on-going periodic expense and should be included in the variable costs in the model. This means that when considering whether to invest in an enterprise, these one-time start-up costs must be taken into account and compared with the long-run discounted cash flow coming from the enterprise to ensure the start-up costs are not so high that the long-run profitability of the investment is compromised. In addition, the cash flow requirements of these start-up costs needs to be taken into account. If the target portfolio of enterprises includes a crop that has substantial start-up costs, it may be necessary to phase this crop in over time so that other enterprises on the farm can generate the cash flow needed to cover the start-up costs.

Another consideration is that the Orinoquía Agricultural Linear Programming Model is a deterministic model. This means that yields and prices in the model are treated as known with certainty and not subject to variation due to weather, pest or disease pressure, or uncertain future market conditions. Ideally, yields and prices should represent long-term averages. However, when the yield data are obtained by producer interviews, it is difficult to get this average perspective. Agricultural producers are inherently optimists, since otherwise it would be unlikely they would be in farming. Even if average yields are entered in the model taking into account all manner of yield risk, the results may still suggest a plan that cannot be sustained if a year where a substantial adverse event occurs.

Price data is likewise problematic. While there may be historical price series for crop products, the future prices may be quite uncertain. This is especially true for crops that are new to the region. The typical situation is that producers recognize that a new crop is highly profitable at current and recent past prices, making this crop attractive for investment. However, all producers see these same crop prices, and if a large number of producers dedicate substantial land to the crop, the product prices may collapse, potentially even leading to losses for what seemed to be a highly profitable venture.

These aspects of riskiness of agricultural production cannot be directly addressed in the Orinoquía Agricultural Linear Programming Model, and thus caution should be used in interpreting the results. In particular, if initial model runs indicate the desirability of large investments in a crop that is particularly new or that bears a substantial risk of large losses, it may be prudent to place limits on the area allocated to that crop so as to limit the downside of returns.

Notwithstanding these needs for caution in interpreting and using the model results, history has shown that this type of model can be useful for identifying portfolios of enterprises that can increase farm profitability and sustainability. The resulting plans can find ways to make complementary use of resources, leading the farm closer to full utilization of resources at the most profitable uses. In addition, because the approach to data development focuses on individual enterprises that are sustainable, the resulting plans should also be sustainable.

Appendix – Data Spreadsheet Organization

All data for the Orinoquía Agricultural Linear Programming Model are contained in a single spreadsheet. To illustrate how this data is organized in the spreadsheet and explain the meaning of the various numbers, we will use a relatively small dataset, which is an early version of a model of a large farm in the Altillanura. This dataset was created October 25, 2017 and has since been revised. Using our naming convention the name of the GAMS source program is AltillanuraL20171025.gms (for Altillanura region, large farm, created in 2017 on October 25). The parallel name for the data spreadsheet is AltillanuraL20171025.xlsx. Instructions for editing the spreadsheets are found in "Parameter Editing for the Orinoquía Agricultural Linear Programming Model" (Fontanilla, 2019).