# Sustainable Vegetable Crop Supply Problem

Lana Mara R. dos Santos<sup>1</sup>, Alysson M. Costa<sup>2</sup>, Marcos N. Arenales<sup>2</sup>, Ricardo Henrique S. Santos<sup>3</sup>

<sup>1</sup>Departamento de Matemática, Universidade Federal de Viçosa, Brazil

<sup>2</sup>Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, Brazil

<sup>3</sup>Departamento de Fitotecnia, Universidade Federal de Viçosa, Brazil

### Abstract

We consider an agricultural production problem, in which one must supply a given demand of crops while respecting ecologically-based production constraints. The problem is twofold: in order to supply the demand, one must determine the division of the available heterogeneous arable areas in plots and, for each plot, obtain an appropriate crop rotation schedule. Rotation plans must respect ecologically-based constraints such as the interdiction of certain crop successions, and the regular insertion of fallows and green manures. We propose a linear formulation for this problem, in which each variable is associated with a crop rotation schedule. The model contains an exponential number of variables and is, therefore, solved by means of a column-generation approach. We also discuss some extensions to the model, in order to incorporate additional characteristics found in field conditions. A set of computational tests using instances based on real-world data confirms the efficacy of the proposed methodology.

**Keywords**: linear programming, crop demand, crop rotation, column generation.

### 1. Introduction

The conventional agricultural production and distribution model is based mainly on large monocultures, and on an intensive use of capital, pesticides, synthetic fertilizers and other non-renewable and polluting oil-based resources. Although this model allows for a reduction in food production costs, there are important and usually non-computed side-effects (Altieri, 1995; Gliessman, 2000). Examples of these are the environmental costs of water contamination by pesticides and other polluting inputs or the social costs associated with the exclusion of small farmers due to the increasing requirements of capital. These extra costs which until very recently did not influence farmers or societal choices concerning production methods (Tillman et al., 2002), threaten the sustainability of these food production models and highlight the need for new paradigms from agriculture.

In view of the above, sustainable agricultural production and distribution models have recently gained attention. Indeed, an increasing number of consumers and governments are becoming more conscious about these topics and there is now an extra monetary value associated with products originating from more ecologically-based and social-friendly agricultural systems, such as organic, biodynamical, equitable, fair-trade systems, etc. (Makatouni, 2000; Seyfang, 2006; Lyon, 2006). This premium value has propelled the

1

development of these alternative models, giving rise to several new planning problems in which one must consider both technical and ecological production aspects, as well as the access of small farmers to the market. On the production side, for example, one is now concerned with the diversity of the production fields, the preservation of the productive resources, the effective use of these resources in a local or regional level and the adoption of cultural and biological methods in order to control the population of weeds, herbivores and pathogens. On the distribution side, one must ensure that small producers can cost-effectively offer their productions to consumers and that the generated benefits are distributed fairly among the supply-chain participants.

In this article, we consider a situation commonly faced by small family farmers of vegetable crops in Brazil. These producers own small cropping areas yielding small and discontinuous productions. This discontinuity is due mainly to climatic conditions which restrict the growth of some crops in certain areas and periods. If on one hand, the small production volumes do not need investments in cleaning, packing and distribution structures, on the other hand, the discontinuous aspect of production makes the establishment of stable links difficult between vegetable farmers and markets since the latter requires a continuous supply of products.

In face of these problems, some small vegetable farmers unite in collective movements such as cooperatives or associations. These organizations promote the gathering of many small productions in packing houses for standardization and distribution, according to long term contracts with consumers and supermarkets. In order to achieve this, these associations must organize the size of areas to be grown with each crop in order to meet demands, and decide where and when this production will occur, due to the fact that each small producer's land might be located in different climatic regions and, therefore, be subjected to different production constraints. In the case of ecologically-based agriculture, additional technical-ecological constraints must be considered. Among these constraints, one can cite (a) the undesirability of growing two crops from the same botanic family in sequence on the same piece of land in order to reduce the propagation of pests, (b) the periodical growth of the so-called green manure, usually some legume species, which help protect the soil and increase its fertility, and (c) the use of fallows, when the natural vegetation is left to grow in order to help reestablish the structure and fauna of the soil and reduce pest damage (Altieri, 1995; Gliessman, 2000). Constraints (a)–(c) suggest the problem of determining in which order crops should be grown on a piece of land. A crop rotation plan is a solution for this problem and will be studied in detail in Section 2.

In view of the above description, we define the *Sustainable Vegetable Crop Supply Problem* (SVCSP), which can be described as the problem of determining the best division of various heterogeneous pieces of land in order to meet a known demand and optimize a given metric, such as maximize production volume or revenues. Moreover, the crop rotation plan on each piece of land must respect the constraints listed in (a)–(c). Note that the complexity of this problem is increased by the fact that it concerns vegetable crop growers. These growers usually cultivate a large number of crop species in side-by-side plots. These vegetable species belonging to different botanic families, present very different production times and for the most part, have planting seasons determined by climate conditions.

The problem of deciding the optimal distribution of arable land is certainly not new to operation researchers. Indeed, as early as 1939, Kantorovich had already mentioned this problem as one that could be dealt with by mathematical tools (Kantorovich, 1960 – translated from the Russian original, dated 1939). The author proposes a mathematical model that selects the optimal partition of land areas in order to maximize the production while the proportion of the different crop production respects a given plan. Hildreth and Reiter (1951), in one of the first conferences on applications of operations research, state that crop rotation can be beneficial since it can increase crop yields (due to the fact that distinguished rotations produce different effects on the soil). Moreover, the authors mention that dividing the available land among different rotations might reduce the need for resources (such as water, labor, etc) since different crops can have different needs throughout the year, and provide some sort of security against failure of a given crop in one year. The authors use a list of pre-determined rotations, which simplify the problem.

El-Nazer and McCarl (1986) eliminate the need for pre-determined rotations with the assumption that the yield of a crop only depends upon the crops grown in the same land in the previous years. Haneveld and Stegeman (2005) use the idea of crop succession requirements, which are expressed in terms of infeasible sequence of crops. The proposed model allows the use of a solution method based on a max-flow algorithm. Detlefsen and Jensen (2007) take into account the fact that crop rotation influences the needs of nitrogen and the yield of the fields. Their model considers that the amount of land to be planted with each crop is given, for each year, which enables the formulation of the crop rotation problem as a transportation problem. In this and in the previous mentioned work, the crop production times are assumed to be one year, the available cropping area is homogeneous and the authors explicitly forbid infeasible crop sequences. Clarke (1989) considers crops with different production times and planting seasons of the year. However, there are no constraints in the sequence of crops.

Other work in the literature has presented decision-aid support systems that evaluate the effect of diverse rotations (Dogliotti et al., 2003; Jones et al., 2003; Stöckle et al., 2003, Bachinger and Zander, 2006). In this work, the computational implementations usually serve as aid tools for agricultural specialists, which can provide a given crop rotation and obtain the effects on yields and soil conditions from the tool.

The short review presented above exemplifies the richness of the research in this area. Indeed, the same basic situation might originate various distinguished problems, depending on the particular objectives or constraints that are considered. In the practical problem that motivated this study, one can highlight the need of supplying a known demand from production coming from heterogeneous pieces of land, while respecting some ecologically-based agriculture constraints. These main particularities can be summarized in the following five characteristics:

- 1. A known demand for each crop (associated with the existing contracts with consumers, for example) should be met.
- 2. Each crop has an earliest and a latest planting time that must be respected.
- 3. Crops have different production times, in other words, the time delay from planting to harvesting varies.

- 4. Ecologically-based production constraints are modeled with the inclusion, in each rotation, of fallow and special crops for green manuring. Moreover, to reduce the occurrence of pests, crops belonging to the same botanic family can not be grown in sequence.
- 5. Available pieces of land can be heterogeneous due to climactic characteristics. For example, in a specific location, a subset of available crops can be grown and these crops have certain yields. This set of crops and their respective yields can be different in another location with different climatic and or soil characteristics.

In Santos et al. (2008), similar production constraints are considered (characteristics 2-4), but there are no demands associated with the crops, and both climate and yields for each available piece of land are the same.

The remainder of this paper is organized as follows. In the next section, we propose a linear mixed-integer formulation for the SVCSP. First, we detail the concept of crop rotation and propose linear constraints that model the set of feasible rotations which respect ecologically-based production conditions (characteristic 4 listed above) while considering that crops can be grown in specific periods of the year and that they have different growing times (characteristics 2 and 3). This crop rotation schedule model is used within a mathematical formulation for the complete sustainable vegetable crop demand supply problem. Extensions and additional comments concerning the model are presented in Section 3. The proposed formulation has an exponential number of variables and is, therefore, solved with a column-generation approach, which is presented in Section 4, as well as heuristic approaches to cope with the extensions discussed in the previous section. Section 5 presents computational results of the proposed methodology when applied to instances based on real data obtained from an ecologically-based vegetable grower established in the countryside in Brazil. Section 6 ends this paper with some conclusions.

# 2 Mathematical formulation

In this section we propose a mathematical formulation for the SVCSP. The proposed model has an exponential number of variables, each one associated with a feasible crop rotation plan. We first define linear constraints modeling the feasible crop rotations in the following subsection. Then, in subsection 2.2, the complete model is presented and is illustrated with a numerical example.

#### 2.1 Sustainable crop rotation

In an agricultural production system, a crop rotation schedule can be defined as the sequence of crops that should be planted, one after the other, in a given area. Many of the articles presented in Section 1 deal with grain crops. In this case, a crop rotation schedule is simply a list of crops in a specific sequence. For example, the rotation "corn-potatoes-potatoes" indicate that in year 1, corn should be planted. Then, in years 2 and 3, the land should be occupied with potatoes, returning to corn in year 4.

In the case of the SVCSP, we first define the time interval in which the sequence of crops will be repeated, T, which is divided into in M equal time periods. A crop rotation schedule (or simply crop schedule) is a planting calendar for each of the crops cultivated in the rotation.

Ecologically-based production systems impose additional constraints on the crop rotations. In this study, among many possible sustainability practices, we concentrate on those mentioned in Section 1 which are presented in detail below:

- (i) Botanic family: crops belonging to the same botanic family can not be planted in sequence;
- (ii) Green manure: a number of crops for green manuring must be planted in each crop rotation. Moreover, green manures must be adequately spaced in time.
- (iii) Fallow: in each crop rotation, a time of fallow must be respected. The required fallows might be subjected to specific durations and must be adequately spaced in time.

As an example, let us consider a crop rotation schedule of one year, divided into twelve periods of one month (M = 12). There are three crops, X, Y and Z with production times of 5, 4 and 2, respectively. Crops X and Y belong to the same botanic family while Z is a crop for green manuring. Crop X can be planted between January (month 1) and July (month 7), while crops Y and Z can be planted all year round. Finally, let us assume that one green manure and one month of fallow are mandatory per schedule. Figure 1 presents a feasible crop rotation schedule for this example:

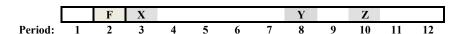


Figure 1. A feasible one year crop rotation schedule

In the crop rotation schedule presented in Figure 1, crop X is planted in month 3 and occupies the land until period 7 (since its production time is 5 months). In period 8, crop Z is planted, staying in the land until period 9. In period 10, crop Y is planted and its final harvesting occurs in period 1 (of the next production year). In period 2, there is one month fallow (F). Note that each crop respects its planting times and crops X and Y (which belong to the same botanic family) are never planted in sequence. Moreover, as desired, in each crop rotation a green manure crop (Z) is cultivated and there is a period associated with fallow (F).

In the remainder of this section, we propose a linear model to describe crop rotations that respect the characteristics listed above. This is done with the help of the following decision variables:

$$x_{ij} = 1$$
, if crop i is planted in period j, 0 otherwise,  $i = 1..n, j \in I_i$ 

where  $I_i$  is the set of periods in which the planting of crop i is allowed. This and other parameters are detailed below:

*M* number of periods in each crop rotation (in a predefined unit of time);

C set of crops that can be selected for planting, not including those associated with green manuring;

G set of crops that can be selected for green manuring;

*N* cardinality of  $C \cup G$ ;

n = N+1 hypothetical crop associated with fallow;

*NF* number of botanic families;

F(p) set of crops of botanic family p, p = 1..NF;

 $t_i$  production time of crop i, including the time estimated for preparing the soil and harvesting;

 $I_i$  set of periods in which crop i can be planted. For the fallow, we adopt  $I_n = \{1, ..., M\}$ 

The constraints that describe the set of feasible crop rotations can then be written as

$$\sum_{i=1}^{n} \sum_{r=0}^{t_i-1} x_{i,j-r} \le 1, \ j = 1..M$$
 (1)

$$\sum_{i \in F(p)} \sum_{r=0}^{l_i} x_{i,j-r} \le 1, \ p = 1..NF, j = 1..M$$
 (2)

$$\sum_{i \in G} \sum_{j \in I_i} x_{ij} = 1 \tag{3}$$

$$\sum_{i=1}^{M} x_{nj} = 1 \tag{4}$$

$$x_{ij} \in \{0,1\}, i = 1..n, j \in I_i$$
 (5)

where, in equations (1)–(2), j - r is replaced by j - r + M, whenever it is less or equal to zero.

Constraints (1) ensure that at most a crop is being cultivated at each period while constraints (2) forbid crops belonging to the same botanic family of being cultivated one immediately after the other. Constraints (3) and (4) force the planting of single green manure and a time of fallow at each crop rotation.

The model above allows for solutions in which a given period has no associated crop. This is equivalent to allowing more times of fallow in each schedule, which poses no problem.

In formulation (1)-(5), the number of mandatory green manures and fallows has been set at one for each rotation period. Depending on the crop rotation length, these numbers might need to be increased. In this case, constraints (3) and (4) can then be written as:

$$\sum_{i \in G} \sum_{j \in I_i} x_{ij} = ngm \tag{7}$$

$$\sum_{i=1}^{M} x_{nj} = ntf \tag{8}$$

where ngm is the number of crops in G that must be planted per crop rotation, and ntf is the number of fallows per crop rotation schedule. In this case, one might need to ensure a minimum time delay between the planting of two green manures or between two fallows. This can be done with the following additional constraints:

$$\sum_{i=G} \sum_{r=0}^{\lg m} x_{i,j-r} \le 1, \ j = 1..M$$
 (9)

$$\sum_{i \in G} \sum_{r=0}^{tgm} x_{i,j-r} \le 1, \ j = 1..M$$

$$\sum_{r=0}^{tf} x_{i,j-r} \le 1, \ j = 1..M$$
(9)

where tgm and tf are the minimum time delays for green manures and fallows.

The proposed formulation can be easily adapted to the case where one is not concerned with the planning of a crop rotation but with the definition of a crop sequence schedule. Formulation (1)–(5) does not include demand constraints. In the next section, a model including them is developed. This new formulation contains an exponential number of variables and is, therefore, solved within a column generation approach.

### 2.2 Crop demand supply problem

A crop rotation schedule is basically a planting calendar indicating when each crop should be planted. Since the planting of a crop in a given period implies one or more subsequent harvesting periods, with expected productivities, one might easily convert the crop rotation schedule (or planting calendar) into a crop production schedule, which is a harvesting calendar indicating the amount (in mass, volume, units or other appropriate measure) of each crop that is being produced in each period.

Consider again the example presented in the previous section and illustrated in Figure 1. Assume that crop X has its first partial harvesting three months after planting and other partial harvesting occurs monthly, until the final harvesting, five months after the time of planting. The productions in months 3, 4 and 5 after planting are 1, 2 and 1 kg per square meter of planted area. Crop Y has no partial harvestings and its production is estimated at 3 kg per square meter of planted area with the harvesting occurring in the last period of its production time, i.e., four months after planting. With the use of this data, and assuming the planting calendar presented in Figure 1 has been used in an area of 2m<sup>2</sup>, we can express the harvesting calendar (or production Schedule) in Figure 2. The productions are presented in italics and are associated to the last planted crop.

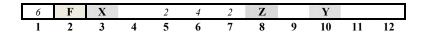


Figure 2. representation of a feasible one year crop production schedule.

As before, crop X is planted in month 3, crop Z in month 8, and crop Y in month 10. The new information in the figure concerns the yields of each crop in each period. Indeed, we can now observe that in months 5, 6 and 7, there is a production of 2, 4 and 2 kg of crop A, respectively. Analogously, in month 1 there is a production of 6 kg of crop Y. The production of crop Z is not indicated since it is being used just for green manuring and has no associated demand.

In order to model the SVCSP, one must match, at each period, the production schedule with the demands to be supplied of each crop. Let C be a set of crops with a positive demand and L be the number of different areas. Due to possible heterogeneities in climate and soil conditions, the set of available crops and the associated productivity might differ in each cropping area. Let:

 $C_k$ , be the set of crops that can be planted in cropping area k;

 $S_k$ , the set of feasible production schedules associated with cropping area k;

 $\overline{I}_i$ , the set of possible harvesting periods of crop i.

Now, let  $a_{ijk}^s$  be the amount of crop  $i \in C_k$  produced in period  $j \in \overline{I_i}$  per square meter of land in area k to which is assigned a crop rotation schedule. Defining variables:

 $\lambda_{ks}$  size of the plot associated with crop schedule s in area k.

We can write the SVCSP as:

$$z_{CSP} = \max \sum_{k=1}^{L} \sum_{s \in S_k} c_{ks} \lambda_{ks}$$
 (11)

$$s.t. \qquad \sum_{k=1}^{L} \sum_{s \in S_k} a_{ijk}^s \lambda_{ks} \ge d_{ij} \quad i \in C_k , k = 1..L, j \in \overline{I}_i$$
 (12)

$$\sum_{s \in S_k} \lambda_{ks} \le A_k , \ k = 1..L \tag{13}$$

$$\lambda_{ks} \ge 0, \ s \in S_k, \ k = 1..L \tag{14}$$

where  $d_{ij}$  is the demand of crop i in period j, given in appropriate units (kg, pack, etc.).

Parameter  $c_{ks}$  is the return obtained when production schedule  $s \in S_k$  is used in one square meter of cropping area k. The goal of model (11)–(14) is to maximize this return which can be related, for example, to the total production volumes or revenues. Constraints (12) ensure that the demands for all crops, at all periods are satisfied while constraints (13) ensure that only the available land at each area is used.

As mentioned earlier, a production schedule is completely defined once a crop rotation schedule has been chosen. Indeed, consider:

 $o_i$  number of periods between planting and harvesting of crop i.

 $p_{ikr}$   $r^{th}$  partial harvesting of crop i in cropping area k,  $r = 1...t_i - o_i - 1$ .

Given the crop rotation schedule  $\mathbf{x}_{k}^{s} = (x_{ijk}^{s})$ , we have:

$$a_{i(j+r+o_i)k}^s = p_{ikr} x_{ijk}^s, i \in C_k, j+o_i+r \in \overline{I}_i, r=1...t_i-o_i-1, s \in S_k, k=1..L.$$

For the case where one wishes to maximize the cropping area production, return values  $c_{ks}$  obtained for a production schedule  $\mathbf{a}_k^s = \mathbf{a}(x_{ijk}^s)$  can be written as:

$$c_{ks} = \sum_{i \in C_k} \sum_{j \in I_i} \sum_{r=1}^{t_i - o_i - 1} p_{ikr} \hat{x}_{ijk}^s$$

Therefore, in terms of a crop rotation schedule  $\mathbf{x}_{k}^{s} = (x_{ijk}^{s})$ , problem (11)–(14) becomes:

$$(CSP-p) \qquad z_{CSP-p} = \sum_{k=1}^{L} \sum_{s \in S_k} \sum_{i \in C_k} \sum_{j \in I_i} \sum_{r=1}^{t_i - o_i - 1} p_{ikr} \hat{x}_{ijk}^s \lambda_{ks}$$

$$s.t \qquad \sum_{k=1}^{L} \sum_{s \in S_k} p_{ir} \hat{x}_{i(j - o_i - r)k}^s \lambda_{ks} \ge d_{ij}, i \in C_k, j = 1..M, r = 1...t_i - o_i - 1$$

$$\sum_{s=1}^{S_k} \lambda_{ks} \le A_k, k = 1..L$$

$$\lambda_{ks} \ge 0, s \in S_k, k = 1..L$$

where,  $j - o_i - r$  is replaced by  $j - o_i - r + M$ , whenever it is less or equal to zero.

**Example**: Consider 3 cropping areas with sizes  $A_1 = 6$ ,  $A_2 = 4$  and  $A_3 = 5$  m<sup>2</sup>. There are 8 crops, which are listed in Table 2. For each crop i, Table 2 presents the botanic family to which it belongs  $(F_i)$ , the set of possible planting dates  $(I_i)$ , the production times  $(t_i)$ , the number of periods before partial harvestings are allowed  $(o_i)$  and the r expected partial harvestings values,  $p_{ir}$ . The table also presents the demands for each crop 1–6, at each period. Crops 5 and 6 are associated with green manuring and have no associated demand.

Due to climatic and soil conditions, crop 1 cannot be planted in area 1. Analogously, crops 2, 3 and 5 are not compatible with area 2 while area 3 can not grow crops 2 and 4. Moreover, assume that the productivity in areas 1, 2 and 3 are 110%, 100%, and 80% of the values presented in Table 2, respectively.

Table 2. problem data: crop data and demands.

			Cro	p data				demand per period											
	$F_{i}$	$I_i$	$t_i$	$o_i$	$p_{i1}$	$p_{i2}$	$p_{i3}$	1	2	3	4	5	6	7	8	9	10	11	12
1	1	1-7; 1-12	4	2	1	1				1	1								
2	1	1-9	3	2	1						1								
3	2	3-11	3	1	1	1		2			1	1			1	1			
4	2	2-11	4	1	1	2	1	4				8	6	8			2	8	10
5	3	7-12; 1-5	3	1	1	1				1	1								
6	3	3-10	3	1	2	1					2	2	2	1	1	2	2	2	2
7	4	2-12	3	2	1														
8	4	8-12; 1-6	2	1	1														

Forcing one green manure and a fallow per schedule and maximizing production, the obtained optimal solution is presented in Figure 5.

A woo 2	3	$(1.25 \text{ m}^2)$	8		3	1	1	6	2	1	9	6	2	1
Area 3 (5 m <sup>2</sup> )	2	$(1.25 \text{ m}^2)$	9	5	1	1	8		3	1	1	6	2	1
(3 III )	1	$(2.5m^2)$	2	9	6	4	2	8		6	4	2	3	2
Area 2	2	(3m <sup>2</sup> )	3	9		4	3	6	3	8		4	3	6
(4 m <sup>2</sup> )	1	(1m <sup>2</sup> )	1		1	1	9	8		4	1	2	1	
Area 1	2	(5m <sup>2</sup> )	5.5	9		4	5.5	10.1	5.5	8		4	5.5	10.1
(6 m <sup>2</sup> )	1	(1m²)		2		1.1	6	2.2	1.1	9	6	1.1	1.1	8
	•		1	2	3	4	5	6	7	Q	Q	10	11	12

Figure 5. Crop production schedules for areas  $A_1$ ,  $A_2$  and  $A_3$ .

The sizes of cropping areas 1 and 2 have been divided between two production schedules while the area of area 3 has been divided into three production schedules. Each production schedule is presented in Figure 5 following the notation used in Figure 3. The values of variables in the optimal solution are:  $\lambda_{21} = 1$ ,  $\lambda_{22} = 3$ ,  $\lambda_{31} = 2.5$ ,  $\lambda_{32} = \lambda_{33} = 1.25$ . In each production schedule there is a fallow, which is represented as crop 9. In the production schedules of area 2 and in production schedule 2 of area 1, an additional fallow has been used. Finally, one can observe that green manure is always present in each rotation.

Clearly, model (11)–(14) contains a very large number of variables due to the fact that the number of possible crop schedules, at each area, grows exponentially with the considered number of crops and periods. For this reason, we propose a column-generation approach to solve this problem. The method is described in the Section 4.

#### 3 Model discussion and extensions

In the previous section, a linear program has been presented for the SVCDP. In the following Next, we further analyze the proposed model. We first note, in subsection 3.1, that (11)–(14) can be simplified in the case of homogeneous areas. Then, in subsection 3.2, we discuss additional characteristics of the SVCDP found in practice and how they could be integrated in the model. Finally, subsection 3.3 discusses a practical question concerning the number of production schedules generated.

### 3.1 Homogeneous areas

If there is no distinction among plots, i.e., if any crop can be planted in any plot yielding the same productivity, model (11)–(14) can be simplified and the plots can be treated as a single continuous arable area. Consider  $A = \bigcup A_k$ . Since a given crop production schedule now yields the same production schedule, regardless of the actual plot on which it is used, we can ignore index k in (11)–(14), obtaining:

(CSP-h) 
$$z_{CSP-h} = \max \sum_{s \in S} c_s \lambda_s$$
 (15)

s.t. 
$$\sum_{k=1}^{L} \sum_{S \in S} a_{ij}^{S} \lambda_{S} \ge d_{ij}, i \in C, j \in \overline{I}_{i}$$
 (16)

$$\sum_{s=1}^{S} \lambda_s \le A,\tag{17}$$

$$\lambda_{s} \ge 0, \ s \in S, k = 1..L \tag{18}$$

The new variables  $\lambda_s$  indicate the amount of the available area in which crop schedule s is used. S is the set of all possible schedules. Let  $\hat{\lambda}_1, \dots, \hat{\lambda}_K$  be an optimal solution to (15)–(18), the actual land division on each plot can be obtained with the help of the following linear system:

$$\sum_{k=1}^{L} \lambda_{ks} = \hat{\lambda}_{s} , s = 1..K$$

$$\sum \lambda_{ks} \le A_k$$
,  $k = 1..L$ 

$$\lambda_{ks} \ge 0$$

## 3.2 Non-supplied demand and upper bounds on demands

In practice, it might be interesting to allow solutions which do not supply all demand but, instead, make a better use of the productive land. This situation can be incorporated into model (11)–(14) with the inclusion of artificial variables in constraints (12). To limit the amount non-supplied demand, these variables can be bounded or, instead, penalty terms can be associated with them in the objective functions. Considering this latter case, model (11)–(14) can be rewritten as:

(CSP-r) 
$$z_{CSP-r} = \max \sum_{k=1}^{L} \sum_{S=1}^{S_k} c_{kS} \lambda_{kS} - \sum_{i \in C} \sum_{j \in \bar{I}_i} r_{ij} \theta_{ij}$$
 (19)

$$S.t. \qquad \sum_{k=1}^{L} \sum_{S \in S_{k}} \mathbf{a}_{k}^{S} \lambda_{kS} \ge \mathbf{d} - \mathbf{\theta}$$
 (20)

$$\sum_{S \in S_k} \lambda_{kS} \le A_k , \ k = 1..L \tag{21}$$

$$\mathbf{\theta} = (\theta_{ij}) \ge 0 \; , \; \lambda_{ks} \ge 0 \; , \; s \in S_k \; , \; k = 1..L,$$
 (22)

where  $\theta_{ij}$ , for  $i \in C$ ,  $j \in \overline{I_i}$  is the amount of non-supplied demand of crop i in period j, and  $r_{ij}$  is the associated penalty.

Besides allowing part of the demand to be non-supplied, it might be convenient to establish an upper bound on the production of the crops. The motivation for this would be to limit the production to levels that can be absorbed by the farmers' clients. This can easily be done with the introduction of upper bounds on constraints (12). In Section 5, the effect of allowing non-supplied demands and upper bounds on production is studied.

#### 3.3 Considerations on the number of plots

Model (11)–(14) imposes no constraints on the number of production schedules than can be used in a given area, allowing solutions with a very large number of plots. In order to reduce the occurrence of these solutions, which might be impractical in real-life situations, a cost  $z_k$  can be associated with each variable  $\lambda_{ks} > 0$ . The original objective function is then replaced by:

$$\max \sum_{k=1}^{L} \sum_{s \in S_k} c_{ks} \lambda_{ks} + z_{ks} \delta_{ks}$$
 (23)

with the following additional constraints:

$$\lambda_{ks} \le A_k \delta_{ks}, \ \delta_{ks} = \{0,1\}, s \in S_k, k = 1..L$$
 (24)

The penalty terms  $z_k$  are selected according to the desired trade-off between the number of plots in the solution and the obtained return. Formulation (23), subject to (12)–(13) and (24) is no longer a linear programming model but a more complex mixed-integer problem. In Section 4, we discuss heuristic strategies for its resolution.

#### 4 Solution approaches

Model (11)–(14) can be seen as a master program in the column generation framework, since it is a linear optimization problem with an exponential number of columns, which can be obtained by solving subproblems (Lübbecke and Desrosiers, 2005). The column-generation approach uses a *restricted* version of this master program, in which only a small subset of variables (or columns) is considered. New columns are iteratively added to the restricted master problem, while the procedure improves the current solution. When no improving columns can be found, the current restricted master solution is also the solution of the master problem.

In our case, since each column in the master problem is a production schedule, one must find a production schedule that maximizes the corresponding reduced cost. Associating dual variables  $\pi_{ij}$  with

constraints (12) and dual variables  $\alpha = [\alpha_1, \dots, \alpha_L]$  with constraints (13), one can write the reduced cost of a column for cropping area k as:

$$\hat{c}_k = z_R(k) - \alpha_k$$

where  $z_R(k)$  is given by solving the subproblem (or pricing problem),

$$z_R(k) = \max \{c_{ks} - \pi a_k^s, s \in S_k\}$$

Since we only want to consider production schedules that respect ecologically-based production conditions listed in (i)–(iii), we can write the pricing problem as the problem of finding a crop rotation schedule, respecting (1)–(5).

A pseudo-algorithm for the used column-generation approach is presented below:

- 1. Choose a set of initial columns to the restricted master problem.
- 2. Solve the restricted master problem and obtain  $(\pi, \alpha)$ , the dual variables associated with constraints (12) and (13), respectively.
- 3. For each k = 1.. L, compute the reduced costs  $\hat{c}_k = z_R(k) \alpha_k$ , by solving the pricing subproblem

$$z_p(k) = \max \{c(\mathbf{x}) - \pi a(\mathbf{x})\},$$

and obtain the associated columns  $\mathbf{x}_k^s$ ,  $\mathbf{a}_k^s = \mathbf{a}(\mathbf{x}_k^s)$  and  $c_{ks} = c(\mathbf{x}_k^s)$ .

4. If, for k = 1.. L,  $\hat{c}_k = 0$ , then **stop**. The current solution is optimal. Otherwise, insert column  $\mathbf{a}_k^s$  in the restricted master problem and go to step 2.

At the end of the column-generation algorithm, we have a set of feasible production schedules  $(\mathbf{a}_k^s)$  associated with given plot sizes  $(\hat{\lambda}_{ks})$ . There might be a large number of  $\hat{\lambda}_{ks} > 0$ , a significant number of which having tiny values. Next, we propose some strategies to deal with this problem.

Given ( $\mathbf{a}_k^s$ ,  $\hat{\lambda}_{ks}$ ), a solution for (11)–(14) containing a large number of values  $\hat{\lambda}_{ks} > 0$ , a simple way to reduce the number plots is to set each  $\hat{\lambda}_{ks} < \alpha$  at zero, where  $\alpha$  is a minimum desired plot size. We name this simple strategy  $\alpha$ –plot elimination. Another heuristic, named plot-minimization, is to select, among the columns (production schedules) present in the optimal restricted master problem, the minimum number which is enough to supply the demand. This can be done with the aid of the following mixed-integer problem:

(CSP-min) 
$$z_{CSP-\min} = \min \sum_{k=1}^{L} \sum_{s=1}^{\hat{S}_k} \delta_{ks}$$
s.a. 
$$\sum_{k=1}^{L} \sum_{s \in \hat{S}_k} \mathbf{a}_k^s \lambda_{ks} \ge \mathbf{d}$$

$$\begin{split} &\sum_{s \in \hat{S}_k} \lambda_{ks} \leq A_k \;,\; k=1..L \\ &\lambda_{ks} \leq A_k \delta_{ks} \;,\; k=1..L \\ &\delta_{ks} = \{0,1\} \;,\; \lambda_{ks} \geq 0 \;,\; s \in \hat{S}_k \;, k=1..L \end{split}$$

Both  $\alpha$ -plot elimination and plot-minimization procedures lead to an increase in unused area and non-supplied demand. General residual strategies for obtaining a good use of the unused area (and possibly reducing the non-supplied demand) can be developed. For example, let  $\hat{\lambda}_{ks}$ ,  $s=1...S_k$ , k=1...L, be an optimal solution of the restricted master problem and define the residual and relative residual of crop i at period j as:

$$r_{ij} = d_{ij} - \sum_{k=1}^{L} \sum_{s \in \hat{S}_k} a_{ijk}^s \hat{\lambda}_{ks}$$
$$rr_{ij} = \frac{r_{ij}}{d_{ii}}$$

The resulting unused area due can be used to increase some plot sizes in order to reduce the non-supplied demands. Priority can be given to crops with higher (relative) residual demands. Another option is to solve a modified version of model (11)–(14) in which the demand values are changed for their residual counterparts,  $r_{ij}$ . This procedure can be used recursively, until there is no area left to be assigned or another appropriate stopping criterion is reached.

Naturally, other procedures for dealing with the problem mentioned in subsection 3.3 exist. Among them, one can include alternative heuristics or even the exact resolution of the proposed mixed-integer formulation with the aid of a branch-and-price algorithm, obtained by adapting the proposed column-generation within a branch-and-bound technique.

In Section 5, we present results for both  $\alpha$ -plot elimination and plot-minimization heuristics. The results are presented in terms of a worst-case analysis, where residual heuristics are not used.

# 5. Computational experiments

We ran several computational experiments in order to verify the performance of the column-generation approach proposed in Section 3, as well as the efficacy of the  $\alpha$ -plot elimination and plot-minimization heuristics presented in the same section. All methods were implemented in C++ and linear and mixed-integer models were solved with the aid of CPLEX 11.0 (ILOG 2007). The tests were run on an Intel Core Duo machine, with 2.33GHz clock and 4Gb of RAM.

A planning horizon of 2 years, divided into weeks, was used. A single green manure and one month's fallow were adopted at each rotation schedule. Four groups of instances were tested, with a varying number of crops (N = 12, 16, 20 and 24). Crop data, listed in Tables A.1 and A.2 in the appendix), as well as

The combination of the different parameters on the number of crops (N = 12, 16, 20 and 24), number of cropping areas (L = 1, 3 and 5) and upper bounds on productions ( $d\infty$ , d200) gives rise to 24 different classes, each one containing 50 instances. In all cases, we have included artificial variables to allow for non-supplied demands (see subsection 3.2), assuming equal penalties for the non-supplied demands of the different crops. The objective function used was to maximize crop production.

Tables 3, 4 and 5 present the results for the column-generation algorithm. For each class of instances, the tables list the execution times (*time*), the average number of plots in the optimal solution (# *plots*), the average percentage of non-supplied demand (*nsdem*), and the percentage of the total area that is used (*oarea*). In these tests, the computational times were limited to one hour.

Table 3: Results for instances with L = 1.

		<i>L</i> =	= 1		
crop	demand	time	# plots	nsdem	oarea
12	d200	250.20	224.30	0.73	72.85
1,2	$d\infty$	757.25	119.80	0.00	100.00
16	d200	910.35	395.40	0.47	86.88
10	$d\infty$	2331.47	237.00	0.00	100.00
20	d200	3283.76	494.70	0.50	100.00
20	$d\infty$	3075.58	426.30	0.00	100.00
24	d200	3230.28	549.40	0.60	100.00
24	$d\infty$	3601.36	528.22	0.00	100.00

Table 4: Results for instances with L = 3.

		<i>L</i> =	= 3		
crop	demand	time	# plots	nsdem	oarea
12	d200	257.37	266.10	0.51	84.53
12	$d\infty$	646.62	116.89	0.00	100.00
16	d200	811.76	335.40	0.63	91.82
10	$d\infty$	1283.94	244.50	0.00	100.00
20	d200	2630.80	497.90	0.49	100.00
20	$d\infty$	3430.26	457.56	0.00	100.00
24	d200	3230.28	549.40	0.60	100.00
24	$d\infty$	3596.96	594.50	0.00	100.00

Table 5: Results for instances with L = 5.

		<i>L</i> =	= 5		
crop	demand	time	# plots	nsdem	oarea
12	d200	129.83	249.60	0.68	64.81
12	$d\infty$	708.97	135.60	0.00	100.00
16	d200	766.70	394.50	0.48	94.30
10	$d\infty$	2722.03	241.50	0.00	100.00
20	d200	2062.26	502.40	0.67	100.00
20	$d\infty$	2947.01	424.80	0.00	100.00
24	d200	3168.13	546.67	0.60	100.00
44	$d\infty$	3612.39	635.89	0.00	100.00

As shown in the tables, the algorithm was able to obtain solutions which respected the demand constraints (in up to 99% in all cases) in the allowed time limit. The computational times increase with the number of crops but show no clear dependency on the number of cropping areas. In general, classes with an upper bound on the production are solved faster. In cases labeled  $d\infty$ , the constraints on available arable land are active (the area is completely used) while when the production is limited to 200%, there is an amount of unused area for cases with fewer crops.

The number of cropping areas has no clear effect on the number of plots in the optimal solution, but these are strongly dependent on the number of crops. Indeed, the larger the number of crops, the larger the number of different crop rotation schedules used, which can reach, on average, 600 schedules for the larger instances (N=24). In order to cope with this problem, the procedures presented in Section 4 were used.

Applying the  $\alpha$ -plot elimination with  $\alpha = 1 \text{m}^2$ , yields the results in Table 6, where for each class the resulting number of plots (#plots) are presented, the area associated with discarded plots (discarded area), in percentage terms of the total area, and the percentage of non-supplied demand due to these plots (lost demand).

Table 6: Results for the  $\alpha$ -plot elimination with  $\alpha = 1 \text{m}^2$ 

			<i>L</i> = 1			L=3			L = 5	
crop	dmax	# plots	discarded area	lost demand	# plots	discarded area	lost demand	# plots	discarded area	lost demand
12	d200	201.40	0.11	0.86	233.10	0.15	1.18	224.80	0.12	1.54
12	$d\infty$	114.70	0.03	0.15	110.00	0.03	0.50	128.30	0.04	0.38
16	d200	336.60	0.26	0.68	303.90	0.14	1.19	351.10	0.19	0.62
10	$d\infty$	221.50	0.07	0.42	232.00	0.06	0.42	230.10	0.06	0.38
20	d200	432.00	0.30	1.30	442.30	0.25	0.62	443.50	0.27	1.06
20	$d\infty$	378.80	0.23	1.22	420.33	0.18	0.93	390.60	0.16	0.70
24	d200	478.20	0.33	1.32	478.20	0.33	1.32	487.89	0.28	1.40
24	$d\infty$	469.44	0.26	0.93	529.80	0.31	1.49	561.67	0.35	1.68

As can be seen in Table 6, the sole application of the simple  $\alpha$ –plot elimination strategy was able to reduce the number of plots by 10%. However, since the areas associated with the discarded plots are small, the net effect on the demand supply is not very significant (around 1%, in average).

Table 7 shows the results for the *plot–minimization* heuristic. The Table presents the heuristic solving times (limited to 0.5h), the resulting number of plots (# *plots*), and the final occupied area (*oarea*).

Table 7: Results for the *plot-minimization* 

			<i>L</i> = 1			L=3			L=5	
crop	dmax	time	# plots	oarea	time	# plots	oarea	time	# plots	oarea
12	d200	259.99	73.70	38.03	591.09	86.70	46.97	333.66	75.40	35.65
12	$d\infty$	409.33	33.80	83.06	1026.59	36.22	81.89	1442.74	37.00	69.48
16	d200	463.94	137.00	53.56	844.12	120.70	52.07	764.38	138.50	58.76
10	$d\infty$	1340.40	55.30	80.84	1800.03	61.40	88.63	1800.06	69.90	95.23
20	d200	1551.62	181.20	73.48	1323.63	192.40	82.38	1638.11	190.90	79.36
20	$d\infty$	1800.09	169.60	100.00	1800.10	195.33	95.34	1800.22	189.50	90.63
24	d200	1540.84	247.80	89.49	1540.84	247.80	89.49	1800.18	301.78	97.50
24	$d\infty$	1800.41	278.00	100.00	1800.56	272.90	100.00	1800.46	302.89	91.11

The reduction in the number of plots now reaches 60% on average, in comparison to the initial solution. Since we force the demand supply in this heuristic, it might occur that there are still plots associated with tiny areas. The  $\alpha$ –plot elimination could, therefore, be applied to the obtained solution.

#### 6. Conclusions

We have considered the Sustainable Vegetable Crop Demand Supply Problem (SVCSP), in which one must supply known crop demands while respecting ecologically-based production restrictions. The problem has been modeled as a linear program with an exponential number of variables, each one representing the area associated with a given feasible rotation schedule. A linear pricing problem to obtain these feasible schedules was proposed, which allowed the solving of the SVCSP by means of a column-generation approach. Post-optimization methodologies were discussed in order to improve the applicability of the obtained solutions. Computational tests on instances adapted from real-data of up to 24 crops indicated the efficacy of the proposed methodologies.

### Acknowledgements

This work was partially supported by Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), and Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP).

# References

- [1] Altieri, M.A., 1995. Agroecology: the science of sustainable agriculture. Boulder: Westview press.
- [2] Bachinger J., Zander, P., 2006. ROTOR, a tool for generating and evaluating crop rotations for organic farming systems. European Journal Agronomy, 26, 130–143.
- [3] Castellazzi, M.S., Wood, G.A., Burgess, P.J., Morris, J., Conrad, K.F., Perry, J.N., 2008. A systematic representation of crop rotations. Agricultural Systems, 97, 26–33.

- [4] Clarke, H.R., 1989. Combinatorial aspects of cropping pattern selection in agriculture. European Journal of Operational Research, 40, 70-77.
- [5] Detlefsen, N., Jensen, A.L., 2007. Modelling optimal crop sequences using network flows. Agricultural Systems 94, 566–572.
- [6] Dogliotti, S., Rossing, W.A.H., Van Ittersum, M.K., 2003. ROTAT, a tool for systematically generating crop rotations. European Journal of Agronomy, 19, 239–250.
- [7] Dogliotti, S., Rossing, W.A.H., Van Ittersum, M.K., 2004. Systematic design and evaluation of crop rotations enhancing soil conservation, soil fertility and farm income: a case study for vegetable farms in South Uruguay. Agricultural Systems, 80, 277–302.
- [8] El-Nazer, T., McCarl. B.A., 1986. The choice of crop rotation: A modeling approach and case study. American Journal of Agricultural Economics, 68, 1, 127–136.
- [9] Gliessman, S.R., 2000. Agroecology: ecological processes in sustainable agriculture. Chelsea: Ann Arbor.
- [10] Haneveld, W. K., Stegeman, A.W., 2005. Crop succession requirements in agricultural production planning. European Journal of Operations Research, 166, 406–429.
- [11] Hildreth, C.G., Reiter, S., 1951. On the choice of a crop rotation plan. In T.C. Koopmans ed., Proceedings of the Conference on Linear Programming held in Chicago in 1949, 177-188.
- [12] ILOG CPLEX 10.1 Reference Manual. 10.1. Copyright by ILOG, France, 2006.
- [13] Jones, J., Hoogenboom, G., Porter, C., Boote, K., Batchelor, W., Hunt, L., Wilkens, P., Singh, U., Gijsman, A., Ritchie, J., 2003. The DSSAT cropping system model. European Journal of Agronomy, 18(3), 235-265.
- [14] Kantorovich, L., 1960. Mathematical methods of organizing and planning production (translated from a report in Russian, dated 1939). Management Science, 6, 366–422.
- [15] Lübbecke, M.E., Desrosiers, J., 2005. Selected topics in column generation. Operations Research, 53(6), 1007-1023.
- [16] Lyon, S., 2006. Evaluating fair trade consumption: politics, defetishization and producer participation. International Journal of Consumers Studies, 30, 452-464.
- [17] Makatouni, A., 2002. What motivates consumers to buy organic food in the UK? Results from a qualitative study. British Food Journal, 104, 345-352.
- [18] Santos, L.M.R., Michelon, P., Arenales, M.N., Santos, R.H.S., 2008. Crop rotation scheduling with adjacency constraints. Annals of Operations Research, doi: 10.1007/s10479-008-0478-z.
- [19] Seyfang, G., 2006. Ecological citizenship and sustainable consumption: Examining local organic food networks. Journal of Rural Studies, 22, 383–395.
- [20] Stöckle, C.O.; Donatelli, M., Nelson, R., 2003. CropSyst, a cropping systems simulation model. European Journal of Operations Research, 18, 289–307.
- [21] Tillman D., Cassman, K. G., Matson, P. A., Naylor, R., Polask, S., 2002. Agricultural sustainability and intensive production practices. Nature, 418, 671–677.

## **APPENDIX**

For the computational experiments, we selected 24 crops which are cultivated at an organic farming unit in Barbacena, Brazil. Crops 21-24 are used for green manuring. The discretization time is one week. Data listed in Tables A.1 and A.2 are approximations of the real values. The production time for each crop includes the estimated planting and harvesting periods.

Table A.1. Data for the 24 crops: name, botanic family, appropriate planting periods and production time.

	CROP	D	OTANIC FAMILY	PLANTING	G PERIOD	PRODUCTION
	CKOF	DO	JIANIC FAMILI	beginning	end	TIME
1	Crisp head lettuce	1	Compositae	all year	round	7
2	Loose leaf lettuce	1	Compositae	all year	7	
3	Butter head lettuce	1	Compositae	all year	round	7
4	Watercress	1	Compositae	all year	round	7
5	Chinese cabbage	2	Brassicaceae	February	September	32
6	Broccoli	2	Brassicaceae	February	October	20
7	Cauliflower	2	Brassicaceae	March	October	18
8	Beet	3	Chenopodiaceae	February	September	11
9	Spinach	3	Chenopodiaceae	February	September	20
10	Zucchini	4	Cucurbitaceae	October	February	14
11	Pumpkin	4	Cucurbitaceae	November	January	19
12	Cucumber	4	Cucurbitaceae	September	March	13
13	Garlic	5	Liliaceae	March	April	24
14	Onion	5	Liliaceae	March	July	24
15	Leek	5	Liliaceae	April	April	12
16	Okra	6	Malvaceae	November	January	27
17	Tomato	7	Solanaceae	all year	round	24
18	Carrot	8	Umbelliferae	all year	round	16
19	Parsley	8	Umbelliferae	October	February	21
20	Bean	9	Leguminosae	October	February	12
21	Black velvet bean	Black velvet bean 9 Leguminos		October	January	16
22	Jack bean	9 Leguminosae		October February		12
23	Lupine	9 Leguminosae		March July		18
24	24 Hairy vetch		Leguminosae	March	July	20

Table A.2. Data for the 20 crops with associated demand: name, measuring unit, number of periods between planting and first partial harvesting ( $o_i$ ) and partial expected harvestings per square meter of planted area.

	CD OD	measuring										HAI	RVEST	ING									
	CROP	unit	$o_i$	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	Crisp head lettuce	unities	5	9	3																		
2	Loose leaf lettuce	unities	5	9	3																		
3	Butter head lettuce	unities	5	9	3																		
4	Watercress	unities	5	9	3																		
5	Chinese cabbage	pack	12	1	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	1	1
6	Broccoli	pack	10	1	1	2	2	3	3	3	2	2	1										
7	Cauliflower	unities	16	1	3																		
8	Beet	kg	8	1	2	1																	
9	Spinach	pack	5	3	4	4	4	4	4	4	4	4	4	4	4	3	3	2					
10	Zucchini	kg	9	0.2	0.3	0.3	0.3	0.2															
11	Pumpkin	kg	16	0.5	0.5	0.5																	
12	Cucumber	kg	8	1	2	2	1	1															
13	Garlic	kg	23	0.3																			
14	Onion	kg	23	3																			
15	Leek	kg	8	3	3	3	3																
16	Okra	kg	14	0.2	0.3	0.3	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.2	0.2	0.2							
17	Tomato	kg	15	0.8	0.8	1	1	1.2	1.2	1	0.8	0.2											
18	Carrot	kg	13	1.5	2	1.5																	
19	Parsley	pack	8	14	16	16	16	16	16	16	16	16	16	16	14	10							
20	Bean	kg	11	0.3																			