

# An optimization model for combined selecting, planting and harvesting sugarcane varieties

Helenice de O. Florentino · Dylan F. Jones · Chandra Ade Irawan · Djamila Ouelhadj · Banafesh Khosravi · Daniela R. Cantane

Received: date / Accepted: date

**Abstract** The problem of selecting sugarcane varieties has been widely discussed due to its computational complexity and its great impact for the sugar and ethanol industry. This paper proposes a new integrated mathematical programming model to deal with the selection of sugarcane varieties to be planted and the determination of the optimal period for planting and harvesting in order to increase production in the sugarcane industry. The proposed model optimizes the production of sugarcane and improves the quality of biomass whilst satisfying the main constraints imposed by sugarcane companies. The problem is modelled as an integer linear program (ILP) and solved using an exact method to generate optimal solutions for small and medium problems. For large problems, metaheuristic approaches based on Genetic Algorithm (GA) and Variable Neighbourhood Search (VNS) are proposed. According to the results, the proposed methodology provides sugarcane company managers with decision support in selecting the most suitable varieties and in determining the best period to plant and harvest their sugarcane.

**Keywords** Agriculture · heuristics · integer programming · sugarcane production

## 1 Introduction

The quantity and quality of the sugarcane crop can be improved by optimizing the agricultural planning process. This optimal planning can also increase profits whilst decreasing costs, as well as facilitating the management of the mills.

---

Helenice de O. Florentino and Daniela R. Cantane  
Departamento de Bioestatística IB UNESP  
Botucatu, SP Brasil  
E-mail: helenice.silva@unesp.br

Dylan F. Jones and Djamila Ouelhadj and Banafesh Khosravi  
School of Mathematics and Physics, Centre for Operational Research and Logistics, University of Portsmouth,  
Lion Gate Building, Lion Terrace, Portsmouth, PO1 3HF, UK

Chandra Ade Irawan  
Nottingham University Business School China, University of Nottingham Ningbo China,  
199 Taikang East Road, Ningbo 315100, China

However, obtaining these plans is not an easy task, the mill managers hence need mathematical optimization techniques to provide decision support in the planning process. In this context, many researchers have presented different mathematical and computational tools aiming to optimize sugarcane plantation and harvest planning processes, as well as to choose the varieties to be planted (Higgins et al, 2004; Higgins and Laredo, 2006; Scarpari and Beauclair, 2010).

Among the important factors that influence the quality and quantity of sugarcane produced, the right choice of the sugarcane variety to be planted for each plot and the correct period of the sugarcane to be planted and harvested can be considered as the most important factors. The selection of the sugarcane variety to be planted depends on a number of aspects including the resistance to pests and diseases, the adaptability to climate and soil, the productivity and the period of planting and harvesting. Imprecise decision on the sugarcane variety selection may reduce the total production and the sustainable sugarcane cultivation. Moreover, the use of the same sugar variety on the same place over time can reduce the chance for the crop to grow and damage the crop if a new disease for a sugarcane variety grows in particular areas.

Piewthongngam et al (2009) propose an optimization model for planning and cultivating sugarcane where the model selects the best period and the best variety to be planted in order to avoid oversupply during the peak of the harvest. It can be noted that their model optimizes the overall sugar production. Jena and Poggi (2013) present a case study of a sugarcane harvest planning system in Brazil. They discuss the benefits of the proposed plan, and the results obtained. Moreover, computational difficulties and the need for further research are reported. Higgins et al (1998) propose an integer linear programming model to maximize the sugar yield and the net revenue relating to the harvest period and the age of the crop. The model aims to improve the vehicles schedule and reduces queues and downtime of vehicles at the mill. Metaheuristics techniques to solve the problem are also proposed.

Extensive research has been conducted to address the choice of sugarcane variety and the plan of the sugarcane crop in order to minimize the cost and residue of harvesting while maximizing profit and sugar production. The selection of sugarcane varieties problem (SSVP) is very important due to its implications for the economic and environmental planning of sugar and ethanol companies (Colin, 2009; Florentino and Pato, 2014; Florentino et al, 2011, 2008; Leboreiro and Hilaly, 2011; Piewthongngam et al, 2009; Sartori et al, 2001; Scarpari and Beauclair, 2010). According to Florentino and Pato (2014), choosing the right varieties of sugarcane is not straightforward, as it depends on fundamental agronomic, industrial, environmental and economic factors. The authors proposed a bi-objective binary linear model for the SSVP in order to minimize collection and transport costs and maximize the energy balance obtained from residues of the sugarcane harvest. In their paper, computational results are also presented and discussed. The application to real problems requires the decision-maker to choose one solution to be implemented from amongst the multiple Pareto optimal set of solutions generated. An interesting work related to agriculture optimization is presented by Shukla and Jharkharia (2014) where an optimization model is designed in order to maximize the overall profit by reducing the waste of agri-fresh produce via syn-

chronization of the demand with supply through an harvest schedule obtained by a heuristic method.

Ramos et al (2016) study a problem similar to the one dealt with in this work and propose a genetic algorithm to solve the optimization problem. However, the methodology proposed by these authors is able to solve small problems only (20 sugarcane varieties and 21 plots) as the proposed model is binary and nonlinear. Moreover, the proposed genetic algorithm requires a long computational time and a large memory space to solve the problem. In a big country such as Brazil, the planning of sugarcane planting and harvesting is not an easy task due to the presence of large mills. It is not unusual that large sugar-energy companies have to manage from 1000 to 3000 plots (areas) for each season, which makes the planning process very difficult. Therefore, in this study a linear model that can solve any size of instances (small, medium and realistic large problems) is proposed. The methodology presented here has a great advantage compared to the approach proposed by Ramos et al (2016). In summary, the main difference between our proposed methodology and the approach presented in Ramos et al (2016) is that in this work we propose a linear model that can be solved by the exact method for the medium scale problem. Moreover, we introduce a different chromosome structure for GA, as well as a hybrid heuristic incorporating GA and VNS to solve large scale problems.

Evolutionary-approach-based meta-heuristics have been widely used for solving agrarian problems, and Genetic Algorithm (GA) has been extensively implemented for addressing these problems (Calija et al, 2001; Paiva, 2009; Bolboaca et al, 2010; Florentino et al, 2017). Florentino et al (2017) propose an extended goal programming model to determine an optimal planning for the sugarcane harvesting only. A genetic algorithm is developed to obtain the approximated solution for the realistically large harvest problem within an appropriate computational time. The success of this work encourages us to apply GA in this work.

Most of the authors cited above give examples pertaining solely to simple cases and moderate sizes of the SSVP. New techniques are required to solve larger and more complex instances that arise in practice. Moreover, an important factor, the quality of sugarcane harvested, is not considered. The quality of sugarcane can be achieved if the harvest is conducted in the correct maturity period of the sugarcane. It is very important to plan the period of planting and harvesting the sugarcane by considering its maturity peak. The main contributions of this study are as follows:

- i. the development of an integer linear model for the SSVP integrated with planting and strategic harvest planning, in order to optimize the production of sugarcane whilst satisfying the main constraints imposed by the mill.
- ii. The construction of metaheuristic approaches for solving larger size models in order to obtain good solutions in a reasonable computational time.

The model introduces new qualitative characteristics in the production and planning of the cane, which makes the proposed model harder to solve as compared to the simple, pure SSVP models available in the literature. The proposed methodology can be adapted to different situations by updating the parameters of the model.

The remainder of the paper is organized as follows: First, a brief description of planting and harvesting sugarcane is given in Section 2. The formulation of the mathematical model and solutions methods based on Genetic Algorithm (GA) and a hybrid metaheuristic integrating a GA with Variable Neighborhood Search (VNS) are described in Section 3. Thereafter, Section 4 presents the computational experiments and the discussion of the results. The last section provides a summary of findings and some avenues for future research.

## 2 Planting and harvesting sugarcane

The sugarcane planting system is generally divided into three types, namely year-sugarcane, year-and-half-sugarcane and winter-sugarcane. The planting is undertaken at the beginning of the rainy season and the plant has its development paralyzed in the cooler months. After about 12 months the sugarcane has reached the required maturity and it can hence be harvested. So in this planting system, the sugarcane is harvested every 12 months. This planting method has low productivity in the first year of planting which then improves in the subsequent years. In the year-and-half-sugarcane planting system, the planting is conducted in the months that have the ideal conditions of temperature and humidity for the plant development. The first harvesting is performed approximately 18 months after the planting. This planting system is the most favorable system for mill managers due to high productivity in the first harvest/cut. After the first cut, the sugarcane is then harvested every 12 months (Colin, 2009). The planting of the winter-sugarcane is usually undertaken in the dry season. Therefore, this planting system can be carried out if irrigation is available. In this system, sugarcane can be harvested every 12 months. The sugarcane productivity is high in the first year due to the facility provided to control the water in the soil.

If the cultivation of sugarcane is well conducted, the reforming or replanting the sugarcane will be conducted 5 or 6 years after the initial planting. High productivity and good quality can be achieved in this period if the correct management and technical recommendations are followed by the mills. Engelke (2002) describes the measurements to evaluate the quality of the sugarcane, such as *pol* (% of sucrose in sugarcane), percentage of brix (total soluble solids), percentage of fiber in cane, commercial cane sugar (CCS) and purity. Among these measurements, the *pol* (%cane) can be considered as the most important parameter. The *pol* percentage in cane is directly related to the point of sugarcane maturation. At the point of maximum maturation (see  $t^*$  in Fig. 1), the content of sucrose is at its highest level; however, if the sugarcane is harvested before or after this phase, a significant decrease in the sugarcane and *pol* production can occur, as illustrated by Figure 1.

Other factors that need to be considered during the planting and harvesting are the diversification of varieties to be planted and the technical capacities of the mill (Higgins et al, 1998). Note that a plot can be planted with a single variety of sugarcane only. To summarize, the period of planting and harvesting sugarcane depends on the variety of sugarcane used whereas the quality and quantity of the sugarcane to be harvested rely on the development of the sugarcane and the period

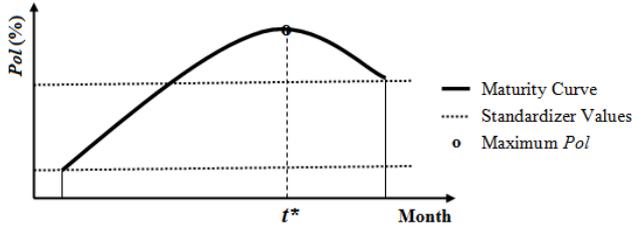


Fig. 1 Maturity curve of the sugarcane (Ramos et al, 2016).

when it is planted, as well as the monthly demand and the capacities of the mill. These facts culminate to result in a challenging decision problem to be solved.

### 3 The model and solution method

In this section, the mathematical model for selecting the sugarcane variety together with the optimal period to plant and harvest is first presented. This is followed by the description of the solution methods proposed to solve the model.

#### 3.1 Mathematical programming model

The selection of sugarcane varieties problem (SSVP) with plant and harvest planning presented in this section consists of the choice of variety  $i$ , among the  $n$  varieties adaptable to climate and local soil,  $i = 1, \dots, n$ , to be planted in each plot  $j$ ,  $j = 1, \dots, k$ . The periods (month) to plant ( $t_0$ ) and to harvest ( $t_c$ ) the sugarcane for each year ( $c = 1, \dots, C$ ) are also determined for each plot. The model aims to maximize the total sugarcane production during the  $C$  years whilst satisfying the constraints and demands imposed by the mill.

The main constraints considered in the model are as follows:

- Only one variety can be planted in a plot to make disease and pest propagation harder to spread;
- The harvesting should be performed as close as possible to the point of maximum sugarcane maturation in each year;
- The *pol* should be maintained above a recommended level;
- The crushing capacity should not be violated;
- The harvesting should be carried out only once in each plot per year;
- A single variety cannot be planted in more than F% of the total area.

Let  $n$  and  $k$  as defined above and  $C$  the number of the years of the horizon for the planning of the sugarcane crop, which is divided into  $T$  months. Let  $M$  denote a set of suitable months for planting, and  $(m + t^*)$  represent the ideal month for harvesting,  $m \in M$ . If it is the first harvest of the year-and-half-sugarcane then  $t^* = 18$ , or  $t^* = 12$  otherwise. It should be noted that  $(m + t^*) \in Q$ , where  $Q$  is a set of periods recommended by the sugarcane companies to harvest the

sugarcane.  $Dev$  denotes the predefined set of the deviations that will be allowed, i.e.,  $Dev = \{d_l, l = 1, \dots, a\}$  where  $d_l$  can be a zero, positive or negative deviation, depending if the sugarcane is harvested in the ideal month ( $m + t^*$ ), after the ideal month or before the ideal month, respectively.

Based on real reduction of sugarcane productivity, when it is harvested outside the ideal period (Ramos et al, 2016), a correction function  $f(d)$  is proposed to reduce the productivity of sugarcane variety  $i$  in the  $c^{th}$  cut,  $Prod_{ic}$ , when it is harvested with  $d$  (zero, positive or negative) months deviation:

$$f(d) = \begin{cases} -0.0243 \cdot d^2 + 1 & \text{if } d_1 \leq d \leq d_a, \\ 0 & \text{otherwise.} \end{cases} \quad (1)$$

The productivity of sugarcane variety  $i$  in the  $c^{th}$  cut and harvested with deviation  $d$  from the ideal period can be calculated using (1), as follows

$$P_{ic}(d) = f(d) \cdot Prod_{ic} \quad (2)$$

The correction function  $f(d)$  penalizes the sugarcane productivity when the harvesting is not conducted at the maximum maturation point. Let  $x_{ijmd}$  be a binary decision variable,

$$x_{ijmd} = \begin{cases} 1 & \text{if sugarcane variety } i \text{ is planted in plot } j \text{ in month } m \\ & \text{and is harvested in month } (m + t_{ij}^* + d) \in Q_i, \\ 0 & \text{otherwise.} \end{cases}$$

where  $i = 1, \dots, n; j = 1, \dots, k; m \in M; d \in Dev$  and  $Q_i$  is the set of periods recommended for harvesting variety  $i$ .

In this study,  $t_{ij}^*$  is set to 18 if sugarcane variety  $i$  planted in plot  $j$  is a year-and-half-sugarcane and it is the first harvest, or  $t_{ij}^* = 12$  otherwise. In addition, the variable  $t_{0j}$  is the month for planting the sugarcane in plot  $j$  and  $t_{cj}$  is the month for the  $c^{th}$  harvesting of the sugarcane in plot  $j$ . Let  $t_{1ijh}$  be a binary decision variable,

$$t_{1ijh} = \begin{cases} 1 & \text{if in the first year of the planning, the variety } i \text{ planted in} \\ & \text{plot } j \text{ is harvest in month } h, \\ 0 & \text{otherwise.} \end{cases}$$

where  $i = 1, \dots, n; j = 1, \dots, k; h \in H$ ,  $H$  is the set of months in which there is demand for sugarcane.

The proposed optimization model for planting and harvesting planning is expressed as follows.

$$\text{Maximise } \sum_{i=1}^n \sum_{j=1}^k \sum_{m \in M} \sum_{d \in Dev} \left( P_{i1}(d) + \sum_{c=2}^C R_{ic} \cdot P_{ic}(d) \right) L_j \cdot x_{ijmd} \quad (3)$$

Subject to

$$\sum_{i=1}^n \sum_{m \in M} \sum_{d \in Dev} x_{ijmd} = 1, \quad j = 1, \dots, k \quad (4)$$

$$t_{0j} = \sum_{i=1}^n \sum_{m \in M} \sum_{d \in Dev} m \cdot x_{ijmd}, \quad j = 1, \dots, k \quad (5)$$

$$t_{1ij} = t_{0j} + \sum_{m \in M} \sum_{d \in Dev} t_{ij}^* \cdot x_{ijmd} + \sum_{m \in M} \sum_{d \in Dev} d \cdot x_{ijmd}, \quad i = 1, \dots, n; j = 1, \dots, k \quad (6)$$

$$t_{1ij} = t_{0j} + \sum_{h \in H} h \cdot t_{1ijh}, \quad i = 1, \dots, n; j = 1, \dots, k \quad (7)$$

$$\sum_{h \in H} t_{1ijh} = 1, \quad i = 1, \dots, n; j = 1, \dots, k \quad (8)$$

$$t_{1ij} = \sum_{m \in M} \sum_{d \in Dev} \sum_{\bar{t}_i \in Q_i} \bar{t}_i \cdot x_{ijmd}, \quad i = 1, \dots, n; j = 1, \dots, k \quad (9)$$

$$\sum_{i=1}^n \sum_{j=1}^k (A_i - D) \cdot t_{1ijh} \geq 0, \quad \forall h \in H \quad (10)$$

$$Ml \leq \sum_{i=1}^n \sum_{j=1}^k Prod_i \cdot L_j \cdot t_{1ijh} \leq Mu, \quad \forall h \in H \quad (11)$$

$$\sum_{j=1}^k \sum_{m \in M} \sum_{d \in Dev} L_j \cdot x_{ijmd} \leq F \cdot \sum_{j=1}^k L_j, \quad i = 1, \dots, n \quad (12)$$

$$t_{cj} = t_{(c-1)j} + t_j^*, \quad c = 1, \dots, C \quad (13)$$

$$t_{1ij} \in Q_i, \quad i = 1, \dots, n; j = 1, \dots, k \quad (14)$$

$$x_{ijmd} = \{0, 1\}, \quad i = 1, \dots, n; j = 1, \dots, k; \forall m \in M, d \in Dev \quad (15)$$

$$t_{1ijh} = \{0, 1\}, \quad i = 1, \dots, n; j = 1, \dots, k; \forall h \in H \quad (16)$$

where

$A_i$  is the *pol* productivity of sugarcane variety  $i$ ;

$c$  is the index associated with the years in the planning horizon;

$C$  is the number of years in the planning horizon;

$d$  is the index associated with the deviations;

$D$  is the value of the *pol* recommended by the mill;

$Dev$  is the set of permissible deviations;

$F$  is the maximum percentage of the total area to be planted with a single variety;

$h$  is the index associated with the months in which demand must be satisfied;

$H$  is the set of months in which the demand must be satisfied;

$i$  is the index associated with the sugarcane varieties;

$j$  is the index associated with the plots;

$k$  is the number of plots;

$L_j$  is the area of the plot  $j$  (ha);

$m$  is the index associated with the months when planting is possible;

$M$  is the set of months when planting is possible;

$Ml$  and  $Mu$  are the monthly lower and upper bounds for sugarcane crushing capacity of the mill (tonnes);  
 $n$  is the number of sugarcane varieties adaptable to the local climate and soil;  
 $Pro_i$  is the average of productivity of sugarcane variety  $i$  in  $C$  years ( $Pro_i = \sum_{c=1}^C Prod_{ic}/C$ );  
 $Q_i$  is the set of months recommended for harvesting the sugarcane variety  $i$  and  $\bar{t}_i \in Q_i$ ;  
 $Q_i = \{\bar{t}_1^i, \dots, \bar{t}_{|Q_i|}^i\}$ ,  $i = 1, \dots, n$ ,  $|Q_i|$  is the cardinality of  $Q_i$ ;  
 $R_{ic}$  is the average percentage change in sugarcane productivity in the  $c^{th}$  cut of the sugarcane variety  $i$ ;  
 $t_{0j}$  is the month for planting the sugarcane in plot  $j$ ;  
 $\bar{t}_1 = \bar{t}_1^i, \dots, \bar{t}_{|Q_i|}^i$ ;  
 $t_{cj}$  is the month of the  $c^{th}$  harvest in plot  $j$ ;

The objective is to select the sugarcane varieties to be planted in  $k$  plots of the mill and to plan the time for planting and harvesting of the sugarcane during the  $C$  year planning period, in order to maximize the total sugarcane production, as described in Equation (3) of the proposed model. Constraints (4) impose that each plot must be planted with one variety only. Equations (5) define the month for planting in each plot. Constraints (6) define the month for the first harvest in each plot. Equations (7) and (8) guarantee that the harvesting is performed in months in which there is sufficient demand. Constraints (9) impose that the harvest of sugarcane is conducted in the period requested by the company. Constraints (10) guarantee that the demand of  $pol$  is met. Constraints (11) impose the crushing capacity constraints of the mill for grinding the sugarcane. Constraints (12) impose the maximal number of areas that can be planted by a sugarcane variety whereas Constraints (13) define the month for the  $c^{th}$  harvest for each plot ( $c > 1$ ). Constraints (14) ensure that the harvest of sugarcane is performed in the period recommended by company and, finally, Constraints (15) and (16) define the binary variables.

To solve the resulting linear, binary programming model, an exact method is used to determine the optimal solutions in a reasonable computational time for small and medium problems (up to 500 plots). However, large problems are hard to solve using the exact method, therefore metaheuristic optimization algorithms based on GA and VNS are proposed to address the problem, which is discussed in the following subsection.

### 3.2 The proposed solution methods

This subsection describes the solution methods that we propose to solve the integrated SSVP problem. The first method is designed based on Genetic Algorithm (GA) whereas the second one is hybridization of GA and VNS.

#### 3.2.1 The proposed Genetic Algorithm (GA)

GA is a search heuristic that generates solutions to NP-hard problems using techniques inspired by natural evolution, such as selection, crossover and mutation

(called genetic operators) (Holland, 1992). In GA, a set of the solutions, called individuals or chromosomes, are usually randomly generated forming a population. A new generation of candidate solutions is obtained by applying the genetic operators on the population. This process is repeated maintaining good genetic characteristics and promoting the evolution of the individuals, and then optimizing the problem. In each generation, the fitness of every individual in the population is evaluated; the fitness is usually the value of the objective function in the optimization problem. Commonly, the algorithm terminates with pre-defined stopping condition, such as a maximum number of generations.

In the GA proposed in the paper, the chromosome has 3 rows and  $k$  columns, as shown in Figure 2. The columns represent the plots whereas the first row denotes the chosen variety to be planted in the corresponding plot, row 2 represents the month for planting, and row 3 the deviation in months of the first harvest from the ideal period for each plot. The main steps of the proposed GA are presented in Algorithm 1.

	$J=1$	$J=2$	...	$J=k$
Variety to be planted $i \in \{1,2, \dots, n\}$	3	1	...	10
Planting month $m \in \mathcal{M}=\{1,2,3,9,10\}$	3	10	...	2
Deviation $d_j \in Dev=\{-2,-1,0,1,2\}$	0	-2	...	1

**Fig. 2** Example for chromosome structure.

*Algorithm 1: The main procedure of the proposed GA*

*Step 1:* Create the initial population.

A random/constructive population of  $N$  chromosomes, as shown in Fig. 2, is generated as follows.

Repeat  $N$  times the following steps:

- Define  $c = 1$  (first year for harvesting),  $H = \{h_1, h_2, \dots, h_R\}$  ( $R$  months with demand),  $J = \{1, 2, \dots, k\}$  ( $k$  plots),  $Dev = \{-5, -4, -3, -2, -1, 0, 1, 2, 3, 4, 5\}$  (deviation),  $M = \{m_1, m_2, \dots, m_r\}$  ( $r$  months for planting) and for each  $h \in H$  do  $Q_{hr} = \{i_1, i_2, \dots, i_l\}$  (set of the sugarcane variety  $i$  that can be harvested in month  $hr$ ),  $i = 1, \dots, n$ .
- Randomly select a month  $t_{1j} = h_{rj} \in H$  for harvesting the sugarcane in each plot  $j \in J$ . This can be achieved by dividing randomly the set  $J$  into  $R$  groups with similar numbers of elements, i.e., the cardinality of  $R$  sets differ in at most one unit. For example, for 9 plots,  $k = 9$ , and 4 months for harvesting in a month in which the demand must be satisfied,  $H = \{\text{January, February, March, October}\}$ ,  $R = 4$ . With a random selection, there are 4 sets constructed:  $P_{Jan} = \{3, 9\}$ ,  $P_{Feb} = \{5, 7\}$ ,  $P_{Mar} = \{1, 4, 8\}$  and  $P_{Oct} = \{2, 6\}$ , implying in January, set  $P_{Jan}$ , harvesting is planned in plots 3 and 9, and so on.
- Create the individual as illustrated in Fig. 2. Randomly select variety  $i$  to be planted in each plot  $j$ ,  $i \in Q_{h_{rj}}$ , and determine the month for planting this sugarcane variety as  $t_{0j} = h_{rj} - t^* + d$ . Where  $d = 0$  if

$(h_{rj} - t^*) \in M$ , otherwise  $d \in (Dev - \{0\})$  is randomly selected such that  $(h_{rj} - t^* + d) \in M$ .

*Step 2:* Evaluate each individual of the initial population.

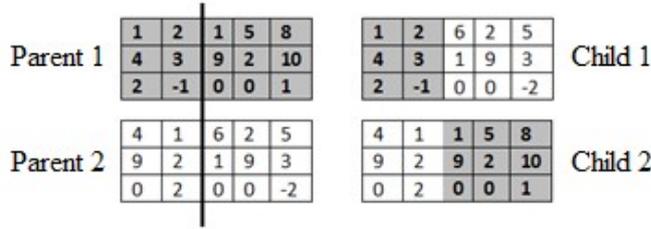
The fitness  $F_{ind}$  of each chromosome or individual in the population is calculated as the value of the objective function (3) for current solution ( $ind$ ).

The best individual ( $Best_{ind}$ ) also needs to be determined and stored.

*Step 3:* Apply the genetic operators.

Repeat the following steps  $G$  times:

- (a) *Select pairs of parent chromosomes from the current population.* The parents are selected among existing solutions in the population using the Roulette Wheel Selection method giving preference to the individual with better fitness, (Holland, 1992).
- (b) *Crossover the pairs at randomly selected points.* A pair of chromosomes are chosen (parent 1 and parent 2) with uniform probability to form two offspring, as well as the number of points at which a crossover takes place. The crossover is performed by cutting the columns of the matrices representing the two parents and the separation of the genes to produce two new individuals (child 1 and child 2), as shown in Figure 3.



**Fig. 3** Crossover operator.

- (c) *Mutate the offspring at each locus with probability  $pm$ .* The mutation is randomly applied with probability  $pm$  for each individual, in order to modify the information contained in some genes of the chromosome. If the number drawn is less than  $pm$  the change will occur, and the element and new value to be changed are randomly chosen.
- (d) *Evaluate each new individual.* The fitness  $F_{ind}$  of each new chromosome created with genetic operators is calculated as the value of the objective function (3) for this solution ( $ind$ ). Update the best individual ( $Best_{ind}$ ).
- (e) *Create the new population.* The best  $N$  individuals evaluated among the current population and new individuals created with genetic operators make up the new population.

*Step 4:* Take  $Best_{ind}$  as the best solution of the problem description.

### 3.2.2 Hybrid Metaheuristic Optimization (HMO)

This section proposes a hybrid metaheuristic optimization algorithm (HMO) to solve the integrate SSVP problem. This hybrid approach is designed by integration

of Genetic Algorithm (GA) and Variable Neighbourhood Search (VNS). VNS was introduced by Brimberg and Mladenovic (1996) for solving continuous location-allocation problems. It was formally formulated by Hansen and Mladenovic (1997) when solving the  $p$ -median problem.

A VNS is an algorithm in which a set of neighborhood structures  $N_k$  ( $k = 1, \dots, k_{max}$ ) is defined. The first (construction) step creates an initial solution. The next (improvement) step applies perturbations, local search and movements within the defined neighborhood to find a better solution based on the established evaluation, which may be the objective function value. Here, at each iteration of the algorithm a solution  $S$  is perturbed in the current neighborhood  $N_k$ , i.e., a feasible solution  $S'$  is generated randomly (perturbation). A local search procedure is applied in neighborhood  $N_k(S')$ , improving the solution  $S'$ . The current solution  $S$  is replaced by the new locally optimal solution  $S'$  if and only if  $S'$  is better than  $S$ . Then, the same search procedure is restarted from solution  $S'$  in the first neighborhood  $N_1$ . Otherwise, if  $S$  is better than  $S'$ , the algorithm moves to the next neighborhood, and attempts to improve the current solution ( $S$ ). The neighborhood structure of the proposed VNS Algorithm is presented as follows.

*Neighborhood structure considered:*

The structure of the solution method in this algorithm is relatively similar to the chromosome considered in the GA illustrated in Figure 2. The solution is constructed as a  $3 \times k$  matrix where the columns are associated with the  $k$  plots and the rows are associated respectively with varieties ( $i$ ), planting months ( $m$ ) and deviations ( $d$ ).

Let  $S$  be a solution described as follows:

$$S = \begin{bmatrix} i_1 & i_2 & \dots & i_k \\ m_1 & m_2 & \dots & m_k \\ d_1 & d_2 & \dots & d_k \end{bmatrix} = \begin{bmatrix} i_j \\ m_j \\ d_j \end{bmatrix},$$

where  $j = \{1, 2, \dots, k\}$ ,  $i_j \in \{1, 2, \dots, n\}$ ,  $m_j \in M$  e  $d_j \in Dev$ .

A neighborhood  $N_1(S)$  is defined by fixing the values of  $i_j$ ,  $m_j$  and  $d_j$  in every column  $j$  of  $S$  such that  $|d_j| \neq 5$  and by varying randomly these values when  $|d_j| = 5$ . This can be illustrated as follows:

$N_1(S) = \{S' = [i'_j, m'_j, d'_j]^t$  such that  $i'_j = i_j, m'_j = m_j$  and  $d'_j = d_j$  for all  $j$  such that  $|d_j| \neq 5$  and  $i'_j \neq i_j$  for all  $j$  such that  $|d_j| = 5\}$ ;

In the same way, we can define:

$N_2(S) = \{S' = [i'_j, m'_j, d'_j]^t$  such that  $i'_j = i_j, m'_j = m_j$  and  $d'_j = d_j$  for all  $j$  such that  $|d_j| \neq 4$  and  $i'_j \neq i_j$  for all  $j$  such that  $|d_j| = 4\}$ ;

$N_3(S) = \{S' = [i'_j, m'_j, d'_j]^t$  such that  $i'_j = i_j, m'_j = m_j$  and  $d'_j = d_j$  for all  $j$  such that  $|d_j| \neq 3$  and  $i'_j \neq i_j$  for all  $j$  such that  $|d_j| = 3\}$ .

*Perturbation:*

To construct a matrix  $S' \in N_1(S)$  set elements  $S'_{aj} = S_{aj}$  if  $|d_j| \neq 5$  and  $a = 1, 2, 3$ . If  $|d_y| = 5$  in some column  $j = y$ , select randomly a variety  $i'_y$  to be planted in plot  $y$ ,  $i'_y \in Q_{h_{ry}}$  and  $i'_y \neq i_y$ , and determine the month for planting this sugarcane variety as  $m'_y = t_{0y} = h_{ry} - t^* + d'_y$ , where  $d'_y = 0$  if  $(h_{ry} - t^*) \in M$ , otherwise  $d'_y \in \{Dev - \{0\}\}$  is randomly selected such that  $(h_{ry} - t^* + d'_y) \in M$ . Here,  $S' = [i'_j, m'_j, d'_j]^t$ . In the same way  $S' \in N_2(S)$  and  $S' \in N_3(S)$  can be created.

The main steps of the HMO are presented in Algorithm 2 which consists of two steps. The first step aims to generate a relatively good initial solution where the proposed GA presented in Algorithm 1 is implemented. The initial solution is then fed to Step 2 which is the proposed VNS algorithm.

*Algorithm 2: The main procedure of the proposed HMO*

*Step 1:* Construct the initial solution  $S_0$ .

Solve the integrated SSVP problem using *Algorithm 1* and let  $Best_{ind}$  be the obtained solution.

Set  $S_0 \leftarrow Best_{ind}$ .

Avaliate  $S_0$ :  $F_{S_0} \leftarrow F_{Best_{ind}}$

Set  $Best_{.s} \leftarrow S_0$  and  $F_{Best_{.s}} \leftarrow F_{S_0}$ .

*Step 2:* VNS Algorithm (perturbation, local search and movement)

Set  $k_{max} = 3$  and the neighborhood  $N_k(Best_{.s}), k = 1, 2, 3$  is defined, do the following steps:

(a)  $k \leftarrow 1$ ;

(b) While  $k \leq k_{max}$

i. Find  $S'$  (the best neighbour of  $Best_{.s}, S' \in N_k(Best_{.s})$ ) using the following steps:

Generate randomly  $S' \in N_K(Best_{.s}), l \leftarrow 1$ ;

While  $l \leq l_{max}$  ( $l_{max}$  is the maximum number of iterations in local search)

A. Perturb  $S'$  (generate randomly  $S'' \in N_k(S'), (S'' \neq S')$ );

B. If  $(F_{S''} > F_{S'})$  then  $S' \leftarrow S''$ ;

C.  $l \leftarrow l + 1$ ;

End-While.

ii. If  $(F_{S'} > F_{Best_{.s}})$  then  $Best_{.s} \leftarrow S'$  and  $k \leftarrow 1$ ;

Else  $k \leftarrow k + 1$ ; (movement)

End-If.

End-While.

(c) Return the best solution ( $Best_{.s}$ ).

## 4 Computational results

In this section, the results of the proposed model and solution methods for the SSVP problem integrated with planting and harvesting planning are presented. An exact method and two metaheuristics, GA and HMO algorithm, were applied to solve the model (3)-(16). The computational experiments were carried out on eight instances with 16, 50, 180, 300, 500, 660, 850 and 1000 plots. The first instance is a real case data from a small farm in the region of Botucatu, São Paulo State,

Brazil whereas the second one is taken from Ramos et al (2016). From the third instance onwards are newly constructed instances which are randomly generated to represent a medium and large farms in the State of So Paulo, Brazil. The small and medium problems are solved by the exact method using IBM ILOG CPLEX version 12.7. The proposed GA and HMO algorithms were implemented with MATLAB 7.14.0.739 (R2012a) software. In the experiments, a computer with an Intel Core i7 - 8.0GB RAM was used.

Twenty sugarcane varieties adaptable to soil of the region of Botucatu city, São Paulo state, Brazil are used. Table 1 presents the average data of sugarcane and  $pol$  productivity for each sugarcane variety in five harvests and the harvest period recommended by the research institutions that develop the varieties.

**Table 1** Sugarcane varieties adaptable to local climate and soil of the region of Botucatu city São Paulo state Brazil with the average data of the sugarcane and the  $pol$  productivity for five harvests, and the period recommended for harvesting these varieties.

$i$	Variety	Sugarcane Productivity (ton.ha <sup>-1</sup> )					$pol(\%)$ Productivity ( $A_i$ )	Period for Harvesting* ( $Q_i$ )
		$c = 1$	$c = 2$	$c = 3$	$c = 4$	$c = 5$		
1	CTC 2	129.10	110.46	96.66	87.71	82.06	13.70	$6 \leq t_1 \leq 10$
2	CTC 4	130.90	112.00	98.01	88.93	83.20	12.96	$4 \leq t_2 \leq 6$
3	CTC 6	136.00	116.37	101.83	92.40	86.45	14.34	$8 \leq t_3 \leq 12$
4	CTC 9	100.00	85.56	74.88	67.94	63.56	15.17	$4 \leq t_4 \leq 7$
5	CTC 15	132.80	113.63	99.43	90.22	84.41	13.88	$7 \leq t_5 \leq 12$
6	CTC 17	112.30	96.09	84.09	76.30	71.38	14.34	$5 \leq t_6 \leq 8$
7	CTC 20	165.00	141.18	123.54	112.10	104.88	12.93	$4 \leq t_7 \leq 5$
8	RB855156	117.80	100.79	88.20	80.03	74.88	13.88	$5 \leq t_8 \leq 6$
9	RB855453	121.93	104.33	91.30	82.84	77.50	13.28	$7 \leq t_9 \leq 12$
10	RB85755	137.90	117.99	103.25	93.69	87.65	14.19	$8 \leq t_{10} \leq 11$
11	RB86755	148.20	126.80	110.97	100.69	94.20	10.93	$5 \leq t_{11} \leq 8$
12	RB925211	89.29	76.40	66.86	60.66	56.76	14.05	$9 \leq t_{12} \leq 11$
13	RB928062	113.00	96.69	84.61	76.77	71.83	15.08	$5 \leq t_{13} \leq 8$
14	RB966928	123.10	105.33	92.17	83.63	78.25	12.75	$5 \leq t_{14} \leq 9$
15	RB922579	142.40	121.84	106.62	96.75	90.51	15.03	$6 \leq t_{15} \leq 8$
16	RB855453	133.35	114.10	99.85	90.60	84.76	13.31	$6 \leq t_{16} \leq 9$
17	SP80-1842	112.80	96.52	84.46	76.64	71.70	14.27	$6 \leq t_{17} \leq 10$
18	SP80-3280	121.70	104.13	91.12	82.68	77.36	14.17	$7 \leq t_{18} \leq 12$
19	SP81-3250	140.60	120.30	105.28	95.52	89.37	14.38	$7 \leq t_{19} \leq 12$
20	SP83-2847	126.70	108.41	94.87	86.08	80.53	12.64	$7 \leq t_{20} \leq 9$

\*The numbers 1 to 12 are related to months from January to December.

Source: Ramos et al (2016).

Table 2 presents the parameters used by the proposed GA to solve the model (3)-(16) where  $P_c$  and  $P_m$  are the percentage of the individual selected for mating pool and the mutation rate respectively. The values of parameters  $P_c$  and  $P_m$  are set based on the preliminary study which are also in line with the recommendations provided by Goldberg (1989).

In the southeastern region of Brazil where the case study arises, sugarcane is generally planted in three periods:

- from January through April and harvested 18 months later ( $t^* = 18$ ),  $M_1 = \{1, 2, 3, 4\}$ ;

**Table 2** Parameters used for the implementation of the heuristics for all instances.

$P_c$	$P_m$	$n$	$D$	$F$	$l_{max}$	$Dev$	$M$	$H$
0.8	0.05	20	13%	25%	20	$\{-5, -4, -3, -2, -1, 0, 1, 2, 3, 4, 5\}$	$\{1, 2, \dots, 10\}$	$\{4, \dots, 12\}$

- from April through August when the winter sugarcane can be planted, but the irrigation is needed and it can be harvested in 12 months, ( $t^* = 12$ ),  $M_2 = \{4, 5, 6, 7, 8\}$ ;
- in months September and October and harvested 12 months later ( $t^* = 12$ );  $M_3 = \{9, 10\}$ , here  $M = M_1 \cup M_2 \cup M_3$ .

The periods of the 12 or 18 months are used for the first harvest as it is the best period for maturity of sugarcane. For practical harvesting reasons, sugarcane industries aim to allow a deviation up to two months at this time,  $Dev = \{-2, -1, 0, 1, 2\}$ , in this period there is a relatively small loss of sucrose. However, in reality it is not always possible to keep the deviation within this range for all plots. Therefore, the deviations is set to  $\{-5, -4, -3, -2, -1, 0, 1, 2, 3, 4, 5\}$  to maintain the feasibility of the model. The model is designed to seek a solution with a deviation as close as possible to zero at each harvest. In this region of Brazil, the harvest of sugarcane is conducted from April to December as in these months rainfall level is low,  $H = \{4, 5, \dots, 12\}$ . Additionally, one variety of sugarcane can typically be planted with maximum range of 15% to 30% of the total area where 25% is used in this study as it is commonly used by the mills. Finally, the minimum value of the *pol* recommended by the mill is set to 13%. In order to analyze the performance of the methodology, winter sugarcane is not considered in the computational experiments. However, if the optimal solution indicates that harvesting with absolute deviation is greater than 2, winter sugarcane can be planted in this plot in order to harvest this sugarcane with zero deviation.

Table 3 shows the 8 instances that are used to evaluate the proposed method. In the table,  $N$  represents the number of individuals in population whereas  $G$  is the maximum number of generations.  $Ml$  and  $Mu$  indicate the lower and upper bounds for the sugarcane crushing capacity of the mill ( $\text{ton.month}^{-1}$ ) respectively. The parameters  $N$  and  $G$  in the GA were chosen after several tests by observing the influence of these parameters on the computational time required to generate the solution. The number of the 0-1 variables and number of constraints are presented for each instance.

Tables 4, 5 and 6 present the results obtained by the exact method, the proposed GA and HMO respectively. The tables provide total sugarcane production in 5 years, sum of the absolute value of deviation from the ideal month that the sugarcane needs to be harvested in the first year, maximum absolute value of deviation, mean of the absolute value of deviation, and CPU time.

For the exact method, the parameters setting is set based on the default setting provided by CPLEX. In this setting, CPLEX will terminate once the %Gap (the guaranteed maximal percentage from optimality) reached 0.01%. We have attempted to set the %Gap to 0%, however a long computational time is required

**Table 3** Characteristics of the instances.

Instance	Number of the Plots $k$	Area (ha) $\sum_{j=1}^k L_j$	$N$	$G$	$Ml$	$Mu$	Number of 0-1 Variables	Number of Constraints
I-1	16	346.23	200	200	2576	7141	38080	1354
I-2	50	1092.35	200	200	8128	22530	119000	4142
I-3	180	4298.92	200	300	31987	88665	428400	14802
I-4	300	7060.60	300	300	52537	145620	714000	24642
I-5	500	11785.41	400	400	87693	243074	1190000	41042
I-6	660	15885.98	500	500	118204	327648	1570800	54162
I-7	850	20669.30	500	500	153796	426304	2023000	69742
I-8	1000	24249.67	500	500	180440	500150	2380000	82042

to generate the optimal solutions. Moreover, CPLEX terminated due to memory issue when solving some instances. Table 4 shows the solutions for the proposed model on instances with 16, 50, 180, 300, 500, 660 and 850 plots. For the large problem (1000 plots), CPLEX was not able to obtain the solution due to being out of memory. However, the small and medium instances can be solved in an acceptable computational time.

**Table 4** The exact method results for model (3)-(16) using CPLEX.

Instance	Sugarcane Production (ton)	Sum of $ Dev $	Maximum $ Dev $	Average $ Dev $	CPU Time (s)
I-1	186479.46	17	3	1.06	101
I-2	609406.56	46	3	0.92	164
I-3	2406894.46	140	3	0.78	229
I-4	3955196.65	235	3	0.78	235
I-5	6602512.09	403	3	0.81	551
I-6	8900882.99	556	3	0.80	666
I-7	11581417,87	666	3	0.78	1835
I-8	NA	NA	NA	NA	NA

NA: CPLEX cant solve the problem due to being out of memory.

The proposed GA was executed 10 times when solving the problem for each instance. Based on the results shown in Table 5, it is found that all solutions generated by GA are feasible and can be considered as good solutions. Moreover, the computational time needed to solve the problem is also acceptable.

Table 6 shows the summary results when the problems are solved by the proposed HMO. According to the table, HMO produced the average deviation of 1.32 which is better than the one obtained by GA (1.4). This indicates that the proposed harvest planning period generated by HMO is closer to the maximum maturity period than the one obtained by GA.

**Table 5** Results of application of the genetic algorithm to solve the model (3)-(16).

Instance	Sugarcane	Sum of	Maximum	Average	CPU
	Production (ton)	$ Dev $	$ Dev $	$ Dev $	Time (s)
I-1	186216.7	25	4	1.56	16.79
I-2	599613.05	64	5	1.28	140.18
I-3	2342471.32	246	5	1.37	104.30
I-4	3875376.03	413	5	1.38	204.00
I-5	6437400.24	709	5	1.42	918.83
I-6	8561042.29	915	5	1.38	1067.81
I-7	11131631.7	1174	5	1.38	3085.38
I-8	12707899.25	1444	5	1.44	3155.77

**Table 6** Results of application of the hybrid algorithm to solve the model (3)-(16).

Instance	Sugarcane	Sum of	Maximum	Average	CPU
	Production (ton)	$ Dev $	$ Dev $	$ Dev $	Time (s)
I-1	186405.61	23	4	1.44	74.11
I-2	609223.99	60	5	1.20	386.84
I-3	2391563.14	220	5	1.22	503.47
I-4	3890764.57	397	5	1.32	835.33
I-5	6457583.83	658	5	1.32	1390.98
I-6	8687566.89	901	5	1.36	2607.42
I-7	11261343.8	1167	5	1.37	5099.75
I-8	13141310.2	1363	5	1.36	5135.72

The computational time required by the methods to solve the problem is dependent to the difficulty in meeting the demand and the problem size. The hybrid method needs a longer computational time than the GA as int the HMO, the preliminary GA is incorporated.

Table 7 compares the results from different methods, namely the exact, GA, and HMO methods. The table presents the value of %Gap which represents the relative deviation between the objective function value ( $Z$ ) obtained by the method (the exact, GA, and HMO methods) and the lower bound obtained by the exact method. The exact method yields very small optimality gap values, showing that the problem was solved with good performance. As for heuristics, based on the %GAP, we can conclude that the HMO algorithm performs better than the GA as the HMO produces near optimal solutions.

Tables 4, 5, 6 and 7 show that the methodology proposed to determine the optimized planning for the sugarcane crop is able to solve problems with different number of plots, even from 500 to 1000 plots. This means that the model is able to help the decision makers in small, medium, and large sugarcane companies. It can be noted that the proposed model can be solved by exact methods for a large number of plots. However, when the number of plots reached 1000 plots, the exact method was not able to solve the problem. Therefore, the proposed heuristics are excellent alternatives to solve such problems, since they generate solutions not far

from the optimal ones. We can conclude that the exact method is the best solution method for solving the small and medium problems whereas the HMO is the most suitable for the large problem.

**Table 7** Performance of the CPLEX, GA and HMO.

Instance	Sugarcane Production (ton)			GAP (%)		
	CPLEX	GA	HMO	CPLEX	GA	HMO
I-1	186479.46	186216.7	186405.61	0.010	0.14	0.04
I-2	609406.56	599613.05	609223.99	0.010	1.61	0.03
I-3	2406894.46	2342471.32	2391563.14	0.010	2.68	0.64
I-4	3955196.65	3875376.03	3890764.57	0.004	2.02	1.63
I-5	6602512.09	6437400.24	6457583.83	0.009	2.50	2.20
I-6	8900882.99	8561042.29	8687566.89	0.009	3.82	2.40
I-7	11581417.87	11131631.7	11261343.8	0.002	3.88	2.76
I-8	NA	12707899.25	13141310.2	-	-	-

NA: CPLEX cant solve the problem due to being out of memory.

Instance I-1 with 16 plots is presented in more detail in order to study the performance of the proposed methodology to determine the optimal planning of the sugarcane crop in new or replanting areas. The data of these 16 plots are shown in Table 8. The optimal planning for sugarcane crop of the instance I-1 is presented in Table 9. Table 9 shows the optimal planning, obtained by the exact method, for the instance I-1 over a 6 year planning horizon, where the variety of sugarcane to be planted in each plot and the month of the planting period in year 1 are given in the third and fourth columns. The month for harvesting in years 2 to 6 is also presented in the fifth column. It can be noted that plots 1 and 4-16 are harvested in the period of industrial utilization (PIU). The harvest period of plots 2 and 3 is close to PIU (deviation -3), but a winter sugarcane can be used with irrigation in these plots. This indicates that planting in April for the first year and harvesting in April for the following years will yield zero deviation.

**Table 8** The area of each plot in Instance I-1 ( $k = 16$ ).

Plot $j$	1	2	3	4	5	6	7	8
Area (ha) ( $L_j$ )	18.49	14.52	28.18	14.22	15.74	16.61	38.41	15.08
Plot $j$	9	10	11	12	13	14	15	16
Area (ha) ( $L_j$ )	12.01	54.95	28.66	13.78	10.43	16.15	18.79	30.21

The methodology was applied and compared with a real case using data of a farm in the region of the Botucatu, São Paulo state, Brazil. This farm plants seven kinds of the sugarcane varieties in ten plots, as shown in Table 10 where the data of sugarcane variety, the area where each variety was planted and the sugarcane production for 4 years are given.

**Table 9** Optimal planning of sugarcane crop for the instance I-1.

Plot	Variety to be plated	Sugarcane planning			Deviation
		Year 1 Planting	Year 2 Cut 1	Years 3-6 Cuts 2-5	
1	7. CTC 20	March	September	September	0
2	15. RB922579	January	April	April	-3
3	2. CTC 4	January	April	April	-3
4	15. RB922579	January	May	May	-2
5	10. RB857515	January	May	May	-2
6	19. SP81-3250	January	May	May	-2
7	19. SP81-3250	April	November	November	1
8	15. RB922579	February	August	August	0
9	8. RB855156	February	August	August	0
10	1. CTC 2	January	July	July	0
11	15. RB922579	January	July	July	0
12	15. RB922579	January	June	June	-1
13	4. CTC 9	February	August	August	0
14	7. CTC 20	October	November	November	1
15	11. RB867515	October	October	October	0
16	19. SP81-3250	October	December	December	2

**Table 10** The farm data in region of the Botucatu city, São Paulo state, Brazil.

Plot	Variety	Area of the Plots	Sugarcane Production (ton)			
			$c = 1$	$c = 2$	$c = 3$	$c = 4$
1	SP813250	53.18	7290.45	5762.58	4869.69	4488.39
2	SP813250	49.64	6805.15	5378.99	4545.53	4189.62
3	SP813250	41.06	5628.92	4449.26	3759.86	3465.46
4	SP803280	50.36	5260.61	4597.87	4915.14	4196.50
5	PO 8862	17.76	2333.99	2085.92	1548.38	1583.11
6	PO 8862	31.65	4528.78	3232.15	2675.07	2533.82
7	SP89-1115	25.55	3655.94	2609.21	2159.50	2045.47
8	RB 855453	26.36	3326.92	2908.91	2529.36	2350.70
9	RB 835486	1.69	160.08	160.61	183.53	207.12
10	RB 855536	28.32	4119.57	2617.25	2356.70	2367.78
Total		325.57	43110.39	33802.76	29542.77	27427.96
Total sugarcane production (ton)			133,883.89			

The proposed methodology (GA) has been applied on the data given in Tables 1 and 10 where the experiment results are presented in Table 11. The optimal solution presented in Table 11 shows that the proposed methodology is able to determine the best plan for the farm; by allocating varieties to plots and providing the period when planting and harvesting should be conducted in order to get an average annual percentage growth of sugarcane production of 9.5% and a percentage growth of 11.4% in four years, when compared to the practice used by the mill, as shown in Table 10.

**Table 11** Results of the application of the proposed methodology using the data presented in Tables 1 and 10.

Plot	Variety	Area of the plots	Sugarcane Production (ton)			
			$c = 1$	$c = 2$	$c = 3$	$c = 4$
1	CTC6	53.18	7232.48	6188.56	5415.32	4913.83
2	CTC20	49.64	8190.6	7008.17	6132.53	5564.64
3	RB857515	41.06	5662.17	4844.67	4239.44	3846.91
4	RB857515	50.36	6944.64	5941.98	5199.67	4718.23
5	RB922579	17.76	2529.02	2163.88	1893.57	1718.28
6	RB922579	31.65	4506.96	3856.24	3374.52	3062.14
7	SP81-3250	25.55	3592.33	3073.66	2689.90	2440.54
8	SP81-3250	26.36	3706.22	3171.11	2775.18	2517.91
9	SP89-1115	1.69	241.82	172.58	142.84	135.30
10	SP89-1115	28.32	4052.30	2892.09	2393.62	2267.23
Total		325.57	46658.55	39312.94	34256.60	31185.00
Total sugarcane production (ton)			151,095.61			

## 5 Conclusions

In this paper a novel integer linear model together with exact and metaheuristic solution methods are proposed for the selection of sugarcane varieties problem (SSVP) integrated with the planning of planting and harvesting in new or re-planting areas. This model can assist decision makers to select the best sugarcane variety to be planted and to determine the best periods for planting and harvesting in order to maximize the production of sugarcane.

The proposed model has been solved using an exact method for small and medium instances. However, it cannot produce an optimal solution for large instances due to memory issue. Therefore metaheuristic approaches based on hybrid optimization algorithm by integrating a GA with VNS are proposed to generate good quality solutions in a relatively small average computational time compared to the exact method and the gap between heuristic and exact values is small. Given the strategic nature of harvest planning, the recommendation is therefore to use the exact solution method where possible (i.e. small and medium scale farms) and the heuristic based method otherwise (i.e. large farms). Another advantage of the proposed methodology is that the planning period of the harvest can be carried out in the correct period, i.e., close to the point of the maximum sugarcane maturity. In general, the exact method is the best solution method for solving the small and medium problems whereas the HMO is the most suitable for the large one. The real case study indicates that the methodology could improve the sugarcane production by more than 11%. This methodology can be applied for sugarcane planting and harvesting planning for any sugarcane company in the world. Therefore, the proposed model and its solution method can be an appropriate decision support tool for the mill manager.

**Acknowledgements** To Brazilian foundations: Conselho Nacional de Desenvolvimento Científico e Tecnológico - CNPq (Grant Number 302454/2016-0), Pró-Reitoria de Pesquisa da

UNESP - PROPE, Fundação para o Desenvolvimento da UNESP - FUNDUNESP and Fundação de Amparo à Pesquisa do Estado de São Paulo - FAPESP (Grant Numbers 2014/01604-0 and 2014/04353-8) for their financial support.

## References

- Bolboaca S, Jantschi L, Balan V, Diudea M, Sestras R (2010) State of art in genetic algorithms for agricultural systems. *Notulae Botanicae Horti Agrobotanici Cluj-Napoca* 38(3):51–63
- Brimberg J, Mladenovic N (1996) A variable neighbourhood algorithm for solving the continuous location-allocation problem. *Studies of Locational Analysis* 10:1 – 12
- Calija A Vand Higgins, Jackson L PAand Bielig, Coomans D (2001) An operations research approach to the problem of the sugarcane selection. *Annals of Operations Research* 108:123–142
- Colin EC (2009) Mathematical programming accelerates implementation of agro-industrial sugarcane complex. *European Journal of Operational Research* 199(1):232 – 235
- Engelke J (2002) Sugarcane: Measuring commercial quality. Tech. rep., FARM-NOTE Department of Agriculture, Kununurra, Australia, n. 23
- Florentino HdO, Irawan C, Aliano AF, Jones DF, Cantane DR, Nervis JJ (2017) A multiple objective methodology for sugarcane harvest management with varying maturation periods. *Annals of Operations Research* DOI 10.1007/s10479-017-2568-2
- Florentino HO, Pato MV (2014) A bi-objective genetic approach for the selection of sugarcane varieties to comply with environmental and economic requirements. *Journal of the Operational Research Society* 65(6):842–854
- Florentino HO, Moreno EV, Sartori MMAP (2008) Multiobjective optimization of economic balances of sugarcane harvest biomass. *Scientia Agricola* 65:561 – 564
- Florentino HO, Lima ADD, Carvalho LRD, Balbo AR, Homem TPD (2011) Multiobjective 01 integer programming for the use of sugarcane residual biomass in energy cogeneration. *International Transactions in Operational Research* 18(5):605–615
- Goldberg DE (1989) *Genetic Algorithms in Search, Optimization and Machine Learning*. Addison-Wesley, Reading, Massachusetts, DOI 10.1023/A:1022602019183
- Hansen P, Mladenovic N (1997) Variable neighbourhood search for the p-median. *Location Science* 5:207 – 225
- Higgins A, Laredo LA (2006) Improving harvesting and transport planning within a sugar value chain. *Journal of the Operational Research Society* 57:367 – 376
- Higgins A, Muchow R, Rudd A, Ford A (1998) Optimising harvest date in sugar production: A case study for the mossman mill region in australia: I. development of operations research model and solution. *Field Crops Research* 57(2):153 – 162
- Higgins A, Antony G, Sandell G, Davies I, Prestwidge D, Andrew B (2004) A framework for integrating a complex harvesting and transport system for sugar production. *Agricultural Systems* 82(2):99 – 115

- Holland JH (1992) *Adaptation in natural and artificial systems*. MIT Press, MA, USA, DOI 10.1137/1018105
- Jena SD, Poggi M (2013) Harvest planning in the brazilian sugar cane industry via mixed integer programming. *European Journal of Operational Research* 230(2):374 – 384
- Leboreiro J, Hilaly AK (2011) Biomass transportation model and optimum plant size for the production of ethanol. *Bioresource Technology* 102(3):2712 – 2723
- Paiva R RPOand Morabito (2009) An optimization model for the aggregate production planning of a brazilian sugar and ethanol milling company. *Annals of Operations Research* 169:117
- Piewthongngam K, Pathumnakul S, Setthanan K (2009) Application of crop growth simulation and mathematical modeling to supply chain management in the thai sugar industry. *Agricultural Systems* 102(1):58 – 66
- Ramos RP, Isler PR, Florentino HO, Jones D, Nervis JJ (2016) An optimization model for the combined planning and harvesting of sugarcane with maturity considerations. *African Journal of Agricultural Research* 11:3950–3958
- Sartori MMP, de Oliveira Florentino H, Basta C, Leo AL (2001) Determination of the optimal quantity of crop residues for energy in sugarcane crop management using linear programming in variety selection and planting strategy. *Energy* 26(11):1031 – 1040
- Scarpari M, Beauclair E (2010) Optimized agricultural planning of sugarcane using linear programming. *Investigacion operacional* 31:126–132
- Shukla M, Jharkharia S (2014) Harvest scheduling to reduce waste in agri-fresh produce supply chains: an artificial immune system-based solution approach. *International Journal of Planning and Scheduling* 2(1):14–39