DOI: https://doi.org/10.1016/j.ejor.2022.12.029

European Journal of Operational Research 309 (2023) 330–344

Contents lists available at ScienceDirect

European Journal of Operational Research

journal homepage: www.elsevier.com/locate/ejor

Innovative Applications of O.R.

Multi-objective optimization for integrated sugarcane cultivation and harvesting planning

Angelo Aliano Filho[a], Washington A. Oliveira[b], Teresa Melo[c,*]

[a] Universidade Tecnológica Federal do Paraná, Departamento de Matemática, Apucarana, PR, 868000-20, Brazil
[b] Universidade Estadual de Campinas, Faculdade de Ciências Aplicadas, Limeira, 13483-531, Brazil
[c] Business School, Saarland University of Applied Sciences, D 66123, Saarbrücken, Germany

ARTICLE INFO

Article history:
Received 29 April 2022
Accepted 19 December 2022
Available online 23 December 2022

Keywords:
OR in agriculture
Integrated cultivation and harvesting planning
Multi-objective optimization
Integer non-linear programming
Sugar-energy industry

ABSTRACT

Sugarcane and its by-products make a relevant contribution to the world economy. In particular, the sugar-energy industry is affected by the timing of sugarcane cultivation and harvesting from which sucrose and bio-energy are produced. We address this issue by proposing a mixed-integer non-linear programming model to schedule planting and harvesting operations for different varieties of sugarcane. The decisions to be made include the choice of sugarcane varieties to be grown on a given set of plots, the periods for their cultivation, the subsequent harvesting periods, and the type of harvesting equipment. These decisions are subject to various constraints related to matching cultivation periods with harvesting periods according to the maturity cycles of the selected sugarcane varieties, the availability of harvesting machinery, the demand for sucrose and fiber, and further technical requirements. The tactical cultivation and harvesting plans to be determined account for three conflicting objectives, namely maximization of the total sucrose and fiber production, minimization of the total time devoted to harvesting, and minimization of the total cost of transporting the harvesting equipment. We develop a tailored exact method based on the augmented Chebyshev scalarization technique extended with a mechanism for identifying an initial feasible integer solution that greatly helps reduce the computational effort for obtaining Pareto-optimal solutions. Our computational study with instances that reflect the current cultivation and harvesting practices in Brazil demonstrate the effectiveness of the proposed methodology. In addition, a comparative analysis reveals the trade-offs achieved by alternative planting and harvesting schedules, thereby facilitating the decision-making process.

1. Introduction

World sugarcane production is led by Brazil, which produced 654.8 million tonnes in the 2020/2021 crop season (CONAB, 2021), accounting for approximately 40 percent of the world total. Raw sugar and ethyl alcohol (ethanol) are the main products obtained from sugarcane in specialized processing units. Brazil is also the world's largest producer of sugar and the second largest producer of ethanol (after the USA), with a share of 21 percent and 27 percent, respectively (Fava Neves & Kalaki, 2020). Unsurprisingly, the sugarcane industry is a significant economic driver for Brazil, representing about 2 percent of the country's gross domestic product (UNICA, 2017). In addition, a number of by-products obtained from sugarcane (e.g., bagasse, straw, molasses, vinasse, filter cake, and yeasts) also contribute to an economically and environmentally significant utilization of these residues (Santos, Eichler, Machado, De Mattia, & De Sousa, 2020).

Driven by recent technological advances, the sugarcane industry is in the midst of a transformation process that makes it possible to manufacture a number of value-added products from sugarcane and its residues. These include a wide selection of bio-fuels, bio-materials, bio-chemicals, and bio-energy that are obtained through bio-refining (Katakojwala, Naresh Kumar, Chakraborty, & Venkata Mohan, 2019; Santos et al., 2020). The potential of this emerging technology has been recognized for its contribution to the sustainable production of new products. For the Brazilian industry, this trend has proven to be economically relevant and has led to the development of new varieties of sugarcane over the years, with improved sucrose yield or increased biomass productivity (Cursi et al., 2022). In this paper, we address this trend by proposing a multi-objective mathematical model that supports planning the cultivation and harvesting of different sugarcane varieties to meet a growing demand for sugar and fiber, with the latter intended for energy production. Specifically, key decisions include determining

* Corresponding author.
E-mail address: teresa.melo@htwsaar.de (T. Melo).

https://doi.org/10.1016/j.ejor.2022.12.029
0377-2217/

the time periods and varieties of sugarcane that should be planted on a given set of plots, the subsequent time periods for harvesting the sugarcane, and the deployment of harvesting equipment. These decisions are affected by several factors, such as the expected demand for sucrose and fiber, the different crop cycles that depend on the selected sugarcane varieties and the period when they reach their peak maturity, the availability of different types of harvesting machinery, and additional technical limitations. The aim is to identify a schedule that maximizes the total amount of sucrose and fiber produced, minimizes the total time devoted to harvesting, and minimizes the total cost of transporting the harvesting equipment. As will be shown, these objectives are conflicting, which means that any sugarcane cultivation and harvesting schedule will have to consider trade-offs to achieve them.

The sugarcane production chain has attracted the attention of scholars, researchers, and practitioners from a wide range of disciplines. The recent literature review by Teixeira, Rangel, Florentino, & Araújo (2021) provides evidence of the interest and relevance of this subject within the Operations Research community. Various mathematical models and solution methods have been developed for planning problems arising at different stages of the production chain, including sugarcane cultivation, harvesting, transportation, and processing. As noted by Teixeira et al. (2021), harvest scheduling problems are the most studied, since these are at the core of the production chain as for any other agricultural product (Ahumada & Villalobos, 2009). Such problems involve determining a plan for carrying out harvesting operations, taking into account the availability of machinery and labor as well as a number of technical requirements regarding the execution of these operations. The latter depend on the practices in force in the country under consideration. For example, at the tactical planning level, the mathematical models proposed by Jarumaneeroj, Dusadeerungsikul, Chotivanich, & Akkermang (2021) and Thuankaewsing, Khamjan, Piewthongngam, & Pathumnakul (2015) address the particular conditions under which harvesting operations take place in Thailand, while Kong, Kuriyan, Shah, & Guo (2019) and Stray, van Vuuren, & Bezuidenhout (2012) focus on identifying optimal harvesting plans at the operational level in South Africa.

Harvesting planning for the Brazilian sugarcane industry has also received extensive attention (Teixeira et al., 2021). Next, we review some relevant references related to this industry. At the operational planning level, Lamsal, Jones, & Thomas (2017) develop a model for scheduling harvesting operations over the course of one day. The resulting schedule is then used to determine the fleet size for transporting sugarcane to a processing facility. Sugarcane harvesting and transportation are also considered by Aliano Filho, Melo, & Pato (2021) by means of a comprehensive bi-objective mathematical model involving the deployment and scheduling of different types of equipment for deliveries to multiple facilities over a multi-period planning horizon. The model developed by Junqueira & Morabito (2019) simultaneously addresses tactical and operational harvesting decisions, while Jena & Poggi (2013) adopt a hierarchical approach that determines a weekly harvesting schedule in the first phase and details daily operations in the second phase. The goal programming model proposed by Florentino et al. (2018) identifies harvesting plans that reflect the trade-off between cutting the sugarcane at the time closest to peak maturity and the effort put into harvesting so as to reduce machinery maintenance and repair costs. Recently, Morales-Chávez, Sarache, Costa, & Soto (2020) addressed the harvest scheduling problem from the economic, environmental, and social perspectives, and applied their model to a real case in Peru.

The planning of harvesting operations is greatly affected by the choice of sugarcane varieties and the timing of their cultivation. However, the integration of planting and harvesting decisions into a single model has not received much attention in the literature as reported by Teixeira et al. (2021). Our work contributes to filling this gap by proposing a novel multi-objective mixed-integer nonlinear programming (MINLP) model that captures the current practices in Brazil for planting different varieties of sugarcane to produce sucrose and fiber, and their intertwining with the harvesting operations. Selecting the right variety of sugarcane to plant on each plot and choosing the appropriate period for its cultivation are decisions that depend on agricultural, environmental, and economic factors as well as the production capacity of the processing facilities (Colin, 2009; Florentino & Pato, 2014). Florentino et al. (2020) address this issue in their integrated model by considering these decisions along with determining a harvest schedule so as to maximize the total production of sugarcane. Our work is more comprehensive, since we also decide on the assignment of the harvesting machines to the different plots and schedule their operating times. Moreover, we consider multiple objectives instead of a single objective. Carvajal, Sarache, & Costa (2019) present a two-stage stochastic model that integrates decisions on sowing and growing sugarcane, and scheduling harvesting operations over a 20-year planning horizon. Uncertainties in crop yield, sowing and harvesting equipment capacity, and demand for sugarcane are captured by a finite set of scenarios. The model is solved by the sample average approximation method. Unlike our model, a single variety of sugarcane is considered as well as a single objective, namely the maximization of the net present value of the total expected profit. Our work shares some features with the recent mixed-integer linear programming (MILP) model proposed by Poltroniere, Aliano Filho, Balbo, & Florentino (2021), as we also consider different types of sugarcane varieties for sucrose and fiber production like these authors. However, in contrast to our work, Poltroniere et al. (2021) do not take into account the importance of selecting distinct sugarcane varieties in adjacent plots to avoid crop damage caused by pests and diseases. Moreover, we also include in our model decisions about the choice of harvesting equipment to be deployed. While Poltroniere et al. (2021) aim at determining a planting and harvesting schedule that maximizes the total amount of sucrose and fiber produced, we consider this goal along with two other objectives concerning the minimization of the harvesting effort over the planning horizon and the minimization of the total cost of moving the harvesting machinery. These features result in a more comprehensive model that is therefore better suited to support farm managers, but at the same time is computationally more challenging to solve.

The main contributions of the present study are summarized as follows: (1) We develop a multi-objective MINLP formulation for integrated planning of cultivation and harvesting decisions over a multi-period planning horizon. Compared to the literature, our model better depicts current practices in the Brazilian sugarcane production chain at the tactical level. (2) We use linearization techniques to obtain a computationally tractable formulation for real-life applications. (3) We develop a tailored exact method to accelerate the generation of Pareto-optimal solutions. As will be discussed in Section 3, no exact method to solve this problem has been presented so far. (4) We assess the validity of the new model and the proposed solution technique by reporting and analyzing the results of a computational study using semi-randomly instances of realistic size. To facilitate the decision-making process, we also provide managerial insights that reveal the trade-offs that are achieved by alternative schedules for planting and harvesting different varieties of sugarcane.

The remainder of this paper is organized as follows. In Section 2, we formally describe our problem and present a multi-objective MINLP formulation. Linearization techniques are also introduced that yield an equivalent model. In Section 3, the solution methodology is described. Computational results are reported

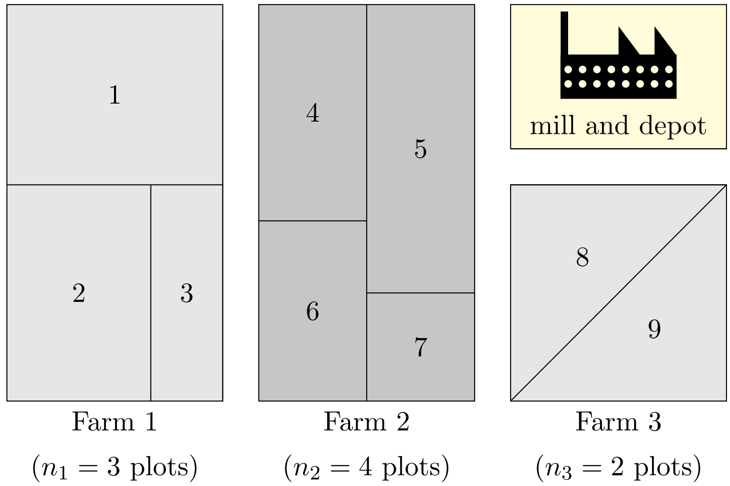

Fig. 1. Example showing 9 plots across 3 farms, the mill, and the depot for the harvesters. Using the notation introduced in Section 2.1 and assuming that different varieties of sucrose-cane are planted on farms 1 and 2, while energy-cane is planted on farm 3, we have $F_S = \{1, 2\}$, $F_E = \{3\}$, $k_S = 7$, $k_E = 2$, $J_S = \{1, 2, 3, 4, 5, 6, 7\}$, $J_E = \{8, 9\}$, $\widetilde{J}_1 = \{1, 2, 3\}$, $\widetilde{J}_2 = \{4, 5, 6, 7\}$, and $\widetilde{J}_3 = \{8, 9\}$.

in Section 4. Finally, Section 5 presents a summary of our findings and gives directions for future research.

2. Problem description and mathematical formulation

Our problem concerns the integrated cultivation and harvest planning of several farms over a time horizon divided into a finite number of periods, usually of equal length. Typically, a period represents one month. Each farm is divided into multiple growing areas, called *plots* (see illustrative example in Fig. 1), each of which must be harvested in a single period. Two types of sugarcane are considered, namely sucrose-cane, which is intended for the production of sugar and ethanol, and energy-cane, which provides fiber (dry mass) used in thermal energy generation. Each farm is dedicated to producing an individual product, either sucrose-cane or energy-cane.

A relevant but non-trivial decision for the farm managers is to determine the time period and sugarcane variety to be planted as well as the harvest period on each plot. The selection of sugarcane varieties depends on many factors, including resistance to pests and diseases, adaptability to climate and soil, productivity as well as planting and harvesting periods. Using the same sugarcane variety over a large area may damage the crop if, e.g., a new disease breaks out. The more distinct varieties grown on the same farm, the less likely they are to be damaged by adverse conditions. This strategy contributes to crop diversification, an agricultural approach that has proven beneficial as it increases crop resilience to diseases, insect pests, and weeds, in addition to improving soil quality and enhancing productivity (Shah et al., 2021). For example, dos Santos, Michelon, Arenales, & Santos (2011) and dos Santos, Munari, Costa, & Santos (2015) explicitly include constraints related to this issue in their optimization models for growing horticultural produce. In our case, to achieve high yields of sucrose-cane and energy-cane it is necessary to carefully select the varieties to be grown on each farm, to determine the plot and the ideal time period for planting each variety, and to identify the best period for harvesting each individual plot.

There are multiple varieties of sucrose-cane that reach their peak maturity at different periods. Ideally, the sugarcane should be harvested in the period in which it attains peak sucrose content. We consider various varieties of sucrose-cane, some with a 12-period cycle and others with an 18-period cycle. Harvesting before or after the peak maturity period negatively affects the yield of sucrose-cane. According to Ramos, Isler, Florentino, Jones, & Nervis (2016), the productivity s_{id} of the ith sucrose-cane variety that is harvested with a deviation d from its peak maturity period is determined by

$$s_{id} = \left(-0.0243\, d^2 + 1\right) s_i^0. \tag{1}$$

Observe that the highest production level s_i^0 is reached for $d = 0$.

Regarding energy-cane, all varieties considered reach maturity at 12 periods. Matsuoka & Rubio (2019) model the productivity m_{id} of the ith variety of energy-cane with harvesting deviation d as a linear growth function as follows:

$$m_{id} = (0.0041\, d + 1)\, m_i^0, \tag{2}$$

with m_i^0 denoting the fiber production level at the end of the 12-period cycle. In this case, delaying the harvest results in a higher yield level. We also note that the deviation parameter d in (1) and (2) can be positive (postponing the harvest d periods), zero (harvesting at maturity), or negative (bringing the harvest forward d periods). Figure 6 in the supplementary material associated with this article displays the deviations d in a Pareto-optimal solution to an illustrative problem.

After harvesting, the different sugarcane varieties are processed in a mill. The sucrose produced is a commodity with a much higher market value than the fiber extracted from the energy-cane. Hence, the sucrose can be stored for later use or sale, while the fiber is burned to generate energy for the mill, and is therefore not stored.

Fluctuations in sucrose and fiber demands at the mill as well as different planting times do not always make it possible to harvest the plots during the ideal maturation period for each variety. These issues reflect negatively on the quality of the harvested cane and cause variations in sucrose and fiber yields. In addition, the maturation of sugarcane in each plot does not occur in the same period. Thus, determining the periods for planting and harvesting that maximize total sucrose and fiber production levels are decisions with a relevant economic impact. These decisions also interact with the choice of harvesting equipment to be deployed. Not only is it necessary to decide on the number and type of harvesting machines to assign to each plot, but also the time required to operate each machine, as the latter is affected by the available labor force. The machinery is stored in a depot (see Fig. 1), which location is adjacent to the mill or at a convenient site in the milling area. Harvesting machines are transported by special vehicles to the farms and return at the end of the harvest to the depot for cleaning, inspection, and maintenance, before being assigned to the next farms. Reducing the number of harvesting fronts, that is, concentrating the harvesting effort on a minimum number of farms in each time period, is a way to make better use of the available equipment and labor resources as well as to complete the harvest sooner. However, these decisions also depend on the sugarcane varieties chosen and the time at which they were planted on the farms.

The mathematical formulation to be presented in the next section models the aforementioned decisions under a given set of constraints. These rule the scheduling of planting and harvesting operations on each plot over the planning horizon, the matching of cultivation periods with the different cycles for harvesting the selected sugarcane varieties, the use of harvesting equipment and labor resources according to their availability, the satisfaction of the demand for sucrose and fiber at the mill, and the compliance with other technical requirements prevailing in the Brazilian sugarcane industry. The aim is to determine an integrated schedule for cultivation and harvesting operations that achieves a trade-off between reaching high production levels of sucrose and fiber, completing the harvest as quickly as possible, and moving the machinery between the depot and the farms at reduced cost. This last issue has practical relevance because transport and maintenance of harvesters is very expensive, and a lower rate of machine utilization results in lower expenses. Three objective functions are defined to model these goals, and their conflicting nature will be discussed in the next section.

Finally, we point out that the scope of our model lies at the tactical planning level and comprises a number of aspects that are usually overlooked in the planning of farm and mill operations. In particular, the relationship between crop cultivation and crop harvesting is often disregarded, despite its practical and economic relevance (Teixeira et al., 2021).

2.1. Notation

In this section, we introduce the notation that will be used hereafter.

Sets, indices, and other parameters:

F_S	set of farms dedicated to cultivation of sucrose-cane (index f);
F_E	set of farms dedicated to cultivation of energy-cane (index f);
F	set of all farms, $F = F_S \cup F_E$;
n_f	total number of plots on farm f ($f \in F$);
k_S	total number of plots dedicated to sucrose-cane cultivation, $k_S = \sum_{f\in F_S} n_f$;
k_E	total number of plots dedicated to energy-cane cultivation, $k_E = \sum_{f\in F_E} n_f$;
J_S	set of all plots dedicated to sucrose-cane cultivation, $J_S = \{1,\ldots,k_S\}$;
J_E	set of all plots dedicated to energy-cane cultivation, $J_E = \{k_S+1,\ldots,k_S+k_E\}$;
J	set of all plots, $J = J_S \cup J_E$;
$\widetilde{J}_f$	set of plots (index j) on farm f ($f \in F$); the plots are numbered consecutively, starting with the plots of farms where sucrose-cane can be cultivated and ending with the plots of farms where energy-cane can be grown (see example in Figure 1); $\widetilde{J}_1 = \{1,\ldots,n_1\}$ represents the set of plots on farm 1, $\widetilde{J}_2 = \{n_1+1,\ldots,n_1+n_2\}$ is the set of plots on farm 2, and so forth; for farm f, we have $\widetilde{J}_f = \left\{1+\sum_{\ell=1}^{f-1} n_\ell,\ldots,\sum_{\ell=1}^{f} n_\ell\right\}$;
$N(j)$	neighborhood of plot j, i.e. set of plots (index j') adjacent to j ($j \in \widetilde{J}_f$, $f \in F$); plot j is adjacent to itself;
V_S^1	set of sucrose-cane varieties with a 12-period cycle (index i);
V_S^2	set of sucrose-cane varieties with an 18-period cycle (index i);
V_S	set of all sucrose-cane varieties, $V_S = V_S^1 \cup V_S^2$;
V_E	set of energy-cane varieties with a 12-period cycle (index i);
V	set of all sugarcane varieties, $V = V_S \cup V_E$;
D_S	set of (negative and positive) time deviations from the peak maturity period for sucrose-cane (index d);
D_E	set of (negative and positive) time deviations from the peak maturity period for energy-cane (index d);

(*continued on next page*)

D	set of all deviations from the peak maturity periods, $D = D_S \cup D_E$;
$\underline{d}$	smallest deviation from the peak maturity period for sugarcane varieties with a 12-period cycle (i.e., $i \in V_S^1 \cup V_E$), $\underline{d} = \min\{d : d \in D\}$;
$\bar{d}$	largest deviation from the ideal harvesting period for sugarcane varieties with an 18-period cycle (i.e., $i \in V_S^2$), $\bar{d} = \max\{d : d \in D\}$;
T_P	set of time periods for sugarcane planting (index t);
$\underline{t}$	first time period allowed for sugarcane to be harvested;
$\bar{t}$	last time period allowed for sugarcane to be harvested;
T_H	set of time periods for sugarcane harvesting (index t), $T_H = \{\underline{t}+\underline{d},\ldots,\bar{t}+\bar{d}\}$, $\lvert T_P\rvert < \underline{t} < \bar{t}$;
H	set of different types of harvesting machines (index h).

Parameters concerning resources and machine transportation costs:

a_j	area of plot j (in ha), $j \in J$;
p_{ijd}^S	amount of sucrose (in tonnes) produced from the ith variety of sucrose-cane harvested on plot j with deviation d from the peak maturity period, with $p_{ijd}^S = a_j\, s_{id}$ and s_{id} given by (1) ($i \in V_S$, $j \in J_S$, $d \in D_S$);
p_{ijd}^E	amount of fiber (in tonnes) produced from the ith variety of energy-cane harvested on plot j with deviation d from the peak maturity period, with $p_{ijd}^E = a_j\, m_{id}$ and m_{id} given by (2) ($i \in V_E$, $j \in J_E$, $d \in D_E$);
d_t^S	demand for sucrose in time period t at the mill (in tonnes), $t \in T_H$;
d_t^E	demand for fiber in time period t at the mill (in tonnes), $t \in T_H$;
c	cost of transporting a harvesting machine between the depot and a farm (in R$ per km);
d_f^0	travel distance between the depot and farm f (in km), $f \in F$;
r_f	total cost of moving a harvesting machine between the depot and farm f (in R$), $r_f = c\, d_f^0$, $f \in F$;
$capM_t$	milling capacity (in tonnes) available at time period t ($t \in T_H$);
$\bar{z}_{ht}$	total number of harvesters of type h available at time period t ($h \in H$, $t \in T_H$);
$capH_h$	harvesting capacity (in tonnes/hour) of a machine of type h ($h \in H$);
$\overline{w}$	total number of working hours of a harvesting machine available in a given time period;
η_i	proportion of the ith sugarcane variety that is allowed to be planted ($0 < \eta_i < 1$, $i \in V$);
θ_S	factor for converting sucrose production into sucrose-cane to be processed at the mill;
θ_E	factor for converting fiber production into energy-cane to be processed at the mill.

2.2. Multi-objective mixed-integer non-linear formulation

We present in this section a formulation of the problem that includes non-linearities in one of the objective functions and in some constraints. These allow us to model certain technical requirements in place in the Brazilian sugarcane industry in a clear way. Later, in Section 2.3, a transformation into an equivalent MILP formulation will be described.

The proposed multi-objective MINLP formulation uses binary decision variables for the assignment of different sugarcane varieties to plots over the planting time horizon and the determination of the periods for their harvest. In addition, continuous variables are defined to represent the inventory level of sucrose-cane at the mill and to measure the time devoted to harvesting the plots. Finally, integer variables are associated with the number and type of harvesting machines deployed.

Decision variables:

$$x_{ijt} = \begin{cases} 1, & \text{if sugarcane variety } i \text{ is planted on plot } j \text{ in time period } t \\ 0, & \text{otherwise} \end{cases}$$

$(i \in V, j \in J, t \in T_P)$;

$$y_{ijtd} = \begin{cases} 1, & \text{if sugarcane variety } i \text{ planted on plot } j \text{ is harvested in time period } t \text{ with deviation } d \text{ from its peak maturity period} \\ 0, & \text{otherwise} \end{cases}$$

$(i \in V, j \in J, t \in T_H, d \in D)$;

$$u_{ft} = \begin{cases} 1, & \text{if any plot on farm } f \text{ is harvested in time period } t \\ 0, & \text{otherwise} \end{cases}$$

$(f \in F, t \in T_H)$;

e_t : inventory level of sucrose-cane in the mill at the end of time period t $(t \in T_H)$;

w_{jt} : number of working hours spent on harvesting plot j in time period t $(j \in J, t \in T_H)$;

z_{hjt} : number of harvesting machines of type h working on plot j in time period t $(h \in H, j \in J, t \in T_H)$.

The multi-objective MINLP formulation is as follows:

$$\text{maximize } v_1 = \sum_{i \in V_S} \sum_{j \in J_S} \sum_{t \in T_H} \sum_{d \in D_S} p^S_{ijd}\, y_{ijtd} + \sum_{i \in V_E} \sum_{j \in J_E} \sum_{t \in T_H} \sum_{d \in D_E} p^E_{ijd}\, y_{ijtd} \quad (3)$$

$$\text{minimize } v_2 = \sum_{f \in F} \sum_{t \in T_H} u_{ft} \quad (4)$$

$$\text{minimize } v_3 = \sum_{j \in J} \sum_{h \in H} r_f \left(z_{hj(\underline{t}+\underline{d})} + z_{hj(\bar{t}+\bar{d})} \right) + \sum_{f \in F} \sum_{t \in T_H \setminus \{\bar{t}+\bar{d}\}} r_f \left| \sum_{j \in \tilde{J}_f} \sum_{h \in H} \left(z_{hj(t+1)} - z_{hjt} \right) \right| \quad (5)$$

subject to

$$\sum_{i \in V} \sum_{t \in T_P} x_{ijt} = 1 \qquad j \in J \quad (6)$$

$$\sum_{j \in J_S} \sum_{t \in T_P} x_{ijt} \le \eta_i k_S \qquad i \in V_S \quad (7)$$

$$\sum_{j \in J_E} \sum_{t \in T_P} x_{ijt} \le \eta_i k_E \qquad i \in V_E \quad (8)$$

$$\sum_{j' \in N(j)} \sum_{t \in T_P} x_{ij't} \le 1 \qquad i \in V,\ j \in J \quad (9)$$

$$\sum_{i \in V} \sum_{t \in T_H} \sum_{d \in D} y_{ijtd} = 1 \qquad j \in J \quad (10)$$

$$\sum_{i \in V} \sum_{t \in T_P} i\, x_{ijt} = \sum_{i \in V} \sum_{t' \in T_H} \sum_{d \in D} i\, y_{ijt'd} \qquad j \in J \quad (11)$$

$$\sum_{i \in V_S^1} \sum_{t \in T_P} (t+12)\, x_{ijt} = \sum_{i \in V_S^1} \sum_{t' \in T_H} \sum_{d \in D_S} (t'-d)\, y_{ijt'd} \qquad j \in J_S \quad (12)$$

$$\sum_{i \in V_S^2} \sum_{t \in T_P} (t+18)\, x_{ijt} = \sum_{i \in V_S^2} \sum_{t' \in T_H} \sum_{d \in D_S} (t'-d)\, y_{ijt'd} \qquad j \in J_S \quad (13)$$

$$\sum_{i \in V_E} \sum_{t \in T_P} (t+12)\, x_{ijt} = \sum_{i \in V_E} \sum_{t' \in T_H} \sum_{d \in D_E} (t'-d)\, y_{ijt'd} \qquad j \in J_E \quad (14)$$

$$\sum_{i \in V_S} \sum_{j \in J_S} \sum_{d \in D_S} p^S_{ijd}\, y_{ijtd} + e_{t-1} = d^S_t + e_t \qquad t \in T_H \quad (15)$$

$$\sum_{i \in V_E} \sum_{j \in J_E} \sum_{d \in D_E} p^E_{ijd}\, y_{ijtd} \ge d^E_t \qquad t \in T_H \quad (16)$$

$$\theta_S \sum_{i \in V_S} \sum_{j \in J_S} \sum_{d \in D_S} p^S_{ijd}\, y_{ijtd} + \theta_E \sum_{i \in V_E} \sum_{j \in J_E} \sum_{d \in D_E} p^E_{ijd}\, y_{ijtd} \le capM_t \qquad t \in T_H \quad (17)$$

$$\sum_{i \in V} \sum_{j \in \tilde{J}_f} \sum_{d \in D} y_{ijtd} \le n_f\, u_{ft} \qquad f \in F,\ t \in T_H \quad (18)$$

$$u_{ft} \le \sum_{i \in V} \sum_{j \in \tilde{J}_f} \sum_{d \in D} y_{ijtd} \qquad f \in F,\ t \in T_H \quad (19)$$

$$1 + \sum_{t \in T_H \setminus \{\bar{t}+\bar{d}\}} u_{ft}\, u_{f(t+1)} = \sum_{t \in T_H} u_{ft} \qquad f \in F \quad (20)$$

$$\sum_{j \in J} z_{hjt} \le \bar{z}_{ht} \qquad h \in H,\ t \in T_H \quad (21)$$

$$\sum_{h \in H} z_{hjt} \le \sum_{h \in H} \bar{z}_{ht} \sum_{i \in V} \sum_{d \in D} y_{ijtd} \qquad j \in J,\ t \in T_H \quad (22)$$

$$w_{jt} \le \overline{w} \sum_{i \in V} \sum_{d \in D} y_{ijtd} \qquad j \in J,\ t \in T_H \quad (23)$$

$$\sum_{h \in H} \sum_{t \in T_H} capH_h\, w_{jt}\, z_{hjt} = \sum_{i \in V_S} \sum_{t \in T_H} \sum_{d \in D_S} p^S_{ijd}\, y_{ijtd} \qquad j \in J_S \quad (24)$$

$$\sum_{h \in H} \sum_{t \in T_H} capH_h\, w_{jt}\, z_{hjt} = \sum_{i \in V_E} \sum_{t \in T_H} \sum_{d \in D_E} p^E_{ijd}\, y_{ijtd} \qquad j \in J_E \quad (25)$$

$$x_{ijt} \in \{0,1\} \qquad i \in V,\ j \in J,\ t \in T_P \quad (26)$$

$$y_{ijtd} \in \{0,1\} \qquad i \in V,\ j \in J,\ t \in T_H,\ d \in D \quad (27)$$

$$u_{ft} \in \{0,1\} \qquad f \in F,\ t \in T_H \quad (28)$$

$$e_t \ge 0 \qquad t \in T_H \quad (29)$$

$$w_{jt} \ge 0 \qquad j \in J,\ t \in T_H \quad (30)$$

$$z_{hjt} \ge 0 \text{ and integer} \qquad h \in H,\ j \in J,\ t \in T_H. \quad (31)$$

The first objective function (3) maximizes the total amount of sucrose and fiber produced. The objective function (4) minimizes the total number of harvesting fronts over the planning horizon. The purpose is to avoid unnecessary machinery movements during harvesting, forcing the harvest on each farm to last a minimum number of periods. The non-linear objective function (5) minimizes the total cost of transporting the harvesting machines between the depot and the farms. The first component of (5) gives the total cost associated with the first and last harvesting periods, while the second component measures the total cost associated with the intermediate periods. Machinery movements from one plot to another plot on the same farm are not accounted for as they incur a negligible cost. For illustration purposes, suppose there is a single farm with three plots, where the first (last) harvesting period can occur in period $\underline{t} = 13$ ($\bar{t} = 17$). Moreover, let us assume that a single machine is deployed, which remains on each plot for exactly one period, starting on plot 1 in period 14 and ending on plot 3 in period 16 (i.e., $z_{1,1,14} = z_{1,2,15} = z_{1,3,16} = 1$). According to (5), the total number of movements of this machine from periods 13 through 17 is given by $|1|+|-1+1|+|-1+1|+|-1|=2$. This means that two movements are accounted for, namely the transport of the harvester from the depot to plot 1 in period 13, and its return to the depot at the end of period 16, after it has completed harvesting on plot 3.

We note that the three objectives are conflicting. The smaller the deviation of the harvest period on each plot from the ideal period, the higher the level of sucrose and fiber production (cf. (3)). Reducing the number of harvesting fronts according to objective (4) implies harvesting the farms in less time, resulting in an increase in the total number of harvesting machines to be used and, consequently, in higher transportation costs. Clearly, this strategy conflicts with objective (5). In turn, less machine movement means harvesting many of the plots outside their ideal periods, thus reducing production yields and as a result worsening objective (3).

Constraints (6) ensure that each plot is planted only once and with a single variety of sugarcane during the planting time horizon. Constraints (7) enforce the number of plots dedicated to the cultivation of a specific sucrose-cane variety does not exceed the pre-specified threshold. Similar conditions are imposed by constraints (8) for the various varieties of energy-cane. Constraints (9) make sure that different sugarcane varieties are grown on adjacent plots of a farm. In other words, it is not allowed to sow the same variety of sugarcane on neighboring plots of a farm, as this is intended to mitigate the risk of the crop being damaged by a pest affecting a specific variety. To illustrate this requirement, let us consider plot 4 belonging to farm 2 in the example shown in Fig. 1. Accordingly, $N(4)=\{4,5,6\}$ is the set of neighbors of plot 4. Over the planting horizon T_P, variety $i=1$ cannot be grown on all the plots of $N(4)$, a condition that is imposed by the inequality $\sum_{t\in T_P}(x_{14t}+x_{15t}+x_{16t})\leq 1$. Constraints (10) state that each plot is harvested only once over the time horizon. Constraints (11) guarantee that the sugarcane variety planted on a plot is also the variety that is later harvested. Constraints (12)–(14) relate the cultivation period of a specific sugarcane variety to its harvest period, taking into account the deviation from the peak maturity period. For example, if variety i of energy-cane is planted in period $t=1$ on a given plot j (i.e., $x_{ij1}=1$), then it achieves peak maturity twelve periods later, i.e. in period 13. As a result, the left-hand side of (14) is $(1+12)\,x_{ij1}=13$. Furthermore, suppose the harvest on plot j can be postponed up to two periods and brought forward at most one period, i.e. $D=\{-1,0,1,2\}$. Accordingly, plot j can be harvested in period $t'=12$ $(d=-1)$, in period $t'=13$ $(d=0)$, in period $t'=14$ $(d=1)$, or in period $t'=15$ $(d=2)$. These four options are stated on the right-hand side of constraints (14) as follows: $\sum_{t'\in T_H}\sum_{d=-1}^{2}\left(t'-d\right)y_{ijt'd}=(12-(-1))\,y_{ij12(-1)}+(13-0)\,y_{ij13(0)}+(14-1)\,y_{ij14(1)}+(15-2)\,y_{ij15(2)}$. Due to constraints (10), only one of these four variables y can be equal to 1. Constraints (15) are inventory balance and demand satisfaction conditions for sucrose at the mill. Demand for fiber is satisfied according to constraints (16). Constraints (17) impose a capacity limit on the total amount of sugarcane processed by the mill at each period. Inequalities (18) and (19) link the variables y_{ijtd} and u_{ft}. Clearly, if no harvest takes place on farm f at period t $(u_{ft}=0)$ then no plot will be harvested $(y_{ijtd}=0)$, cf. (18). Moreover, harvesting will take place on at least one plot of farm f when it is decided to harvest on this farm at a given period (cf. (19)). For practical reasons, the non-linear constraints (20) require that harvesting be carried out in consecutive periods on each farm. In other words, once harvesting is started on a farm it cannot be discontinued until all the plots have been harvested. For example, suppose that harvesting on farm $f=1$ requires three periods, which translates into the right-hand side of (20) being equal to 3. If these three periods were non-consecutive, e.g. $u_{11}=u_{12}=u_{14}=1$ and $u_{1t}=0$ for all $t\notin\{1,2,4\}$, then constraint (20) would be violated as its left-hand side would be equal to 2 $(1+u_{11}\cdot u_{12}+u_{12}\cdot u_{13}+u_{13}\cdot u_{14}+u_{14}\cdot u_{15}+0=1+1\cdot 1+1\cdot 0+0\cdot 1+1\cdot 0+0=2)$. The only way to satisfy (20) is to have, for instance, $u_{11}=u_{12}=u_{13}=1$ and $u_{1t}=0$ for $t\geq 4$. In other words, farm 1 has its plots harvested in periods 1, 2, and 3.

At each time period, the total number of deployed machines of a given type is limited by inequalities (21). According to constraints (22) and (23), if at a given period a plot is not harvested $(y_{ijtd}=0)$ then no equipment can be assigned to it $(w_{jt}=z_{hjt}=0)$. The non-linear constraints (24) and (25) impose the amount of sugarcane grown on each plot be fully harvested. In addition, the bilinear term $w_{jt}\,z_{htj}$ on the left-hand side of (24) and (25) allows the harvesting effort to be modeled in terms of the number of hours worked and the number of machines used on each plot and in each period. For example, having $w_{jt}=2$ and $z_{htj}=4$ for some plot j, period t, and machine type h results in a total of 8 hours of work. If $w_{jt}=4$ and $z_{htj}=2$, also 8 hours of labor are spent, but at a lower cost because fewer machines need to be moved (recall objective function (5)). Finally, constraints (26)–(31) state the integer and non-negativity conditions for the different sets of variables.

Our formulation differs from the model recently developed by Poltroniere et al. (2021) in that it (i) considers multiple objectives; (ii) takes into account neighborhood conditions for planting; and (iii) schedules harvesting operations, including working hours of machines. Clearly, these features make our model more difficult to solve. In particular, the various non-linear components in the model make it challenging to develop an efficient and effective methodology to obtain Pareto-optimal solutions. For this reason, we develop in the next section an equivalent formulation obtained through linearization techniques. This approach has proven useful in solving other problems related to the tactical planning of operations in the sugarcane industry, e.g., Aliano Filho et al. (2021).

2.3. A linear and equivalent mathematical formulation

Recall that the objective function (5) and constraints (20), (24) and (25) have non-linear terms. First, we transform the absolute value function in (5) into a linear function by introducing the non-negative continuous variables ρ^+_{ft} and ρ^-_{ft} as follows:

$$\rho^+_{ft}-\rho^-_{ft}=\sum_{h\in H}\sum_{j\in \bar{J}_f} z_{hj(t+1)}-z_{hjt}\qquad f\in F,\ t\in T_H\setminus\{\bar{t}+\bar{d}\}. \quad (32)$$

As a result, the linear objective function (41) is obtained.

Next, we replace each product of binary variables in (20) by the new binary variable γ_{ft}, and add the following linear constraints:

$$\gamma_{ft}\leq u_{ft}\qquad f\in F,\ t\in T_H\setminus\{\bar{t}+\bar{d}\} \quad (33)$$

$$\gamma_{ft}\leq u_{f(t+1)}\qquad f\in F,\ t\in T_H\setminus\{\bar{t}+\bar{d}\} \quad (34)$$

$$u_{ft}+u_{f(t+1)}\leq 1+\gamma_{ft}\qquad f\in F,\ t\in T_H\setminus\{\bar{t}+\bar{d}\}. \quad (35)$$

These inequalities impose $\gamma_{ft}=0$ when $u_{ft}=0$ or $u_{f(t+1)}=0$, otherwise $\gamma_{ft}=1$. Hence, constraints (20) are replaced by (42).

The left-hand sides of constraints (24) and (25) involve bilinear terms of non-negative continuous and integer variables. We rewrite these terms using a new set of continuous variables δ_{hjt} $(h\in H,\ j\in J,\ t\in T_H)$. Furthermore, we model the binary expansion of the integer variables z_{hjt} by defining new binary variables α_{hjtr} as follows:

$$z_{hjt}=\sum_{r\in R_{ht}}2^{r-1}\,\alpha_{hjtr}\qquad h\in H,\ j\in J,\ t\in T_H, \quad (36)$$

with $R_{ht} = \{1, \ldots, \lfloor \log_2(\bar{z}_{ht}) \rfloor + 1\}$. Therefore, each term $w_{jt}\, z_{hjt}$ in (24)–(25) can be rewritten as

$$\delta_{hjt} = w_{jt} \sum_{r \in R_{ht}} 2^{r-1} \alpha_{hjtr} = \sum_{r \in R_{ht}} 2^{r-1} w_{jt}\, \alpha_{hjtr} \qquad h \in H,\ j \in J,\ t \in T_H.$$

We now introduce a new set of non-negative continuous variables $\beta_{hjtr} = w_{jt}\, \alpha_{hjtr}$ ($h \in H$, $j \in J$, $t \in T_H$, $r \in R_{ht}$), and state the following relationships:

$$\delta_{hjt} = \sum_{r \in R_{ht}} 2^{r-1} \beta_{hjtr} \qquad h \in H,\ j \in J,\ t \in T_H. \tag{37}$$

Since the continuous variables w_{jt} are bounded by $\overline{w}$, the linearization of the term $w_{jt}\, \alpha_{hjtr}$ in (24)–(25) is obtained by introducing the following three sets of constraints:

$$\beta_{hjtr} \le \overline{w}\, \alpha_{hjtr} \qquad h \in H,\ j \in J,\ t \in T_H,\ r \in R_{ht} \tag{38}$$

$$\beta_{hjtr} \le w_{jt} \qquad h \in H,\ j \in J,\ t \in T_H,\ r \in R_{ht} \tag{39}$$

$$\beta_{hjtr} \ge w_{jt} - \overline{w}(1 - \alpha_{hjtr}) \qquad h \in H,\ j \in J,\ t \in T_H,\ r \in R_{ht}. \tag{40}$$

The technique used to develop constraints (36)–(40) along with the values that the variables δ_{hjt}, α_{hjtr}, and β_{hjtr} can take, provides an exact linearization of the original bilinear terms as proven by Gupte, Ahmed, Cheon, & Dey (2013).

The reformulation of the MINLP model defined by (3)–(31) into an equivalent MILP model is given as follows:

$$\text{maximize } \nu_1 = \sum_{i \in V_S} \sum_{j \in J_S} \sum_{t \in T_H} \sum_{d \in D_S} p^S_{ijd}\, y_{ijtd} + \sum_{i \in V_E} \sum_{j \in J_E} \sum_{t \in T_H} \sum_{d \in D_E} p^E_{ijd}\, y_{ijtd} \tag{3}$$

$$\text{minimize } \nu_2 = \sum_{f \in F} \sum_{t \in T_H} u_{ft} \tag{4}$$

$$\text{minimize } \nu_3 = \sum_{f \in F} \sum_{j \in \tilde{J}_f} \sum_{h \in H} r_f \left(z_{hj(\underline{t}+\underline{d})} + z_{hj(\bar{t}+\bar{d})} \right) + \sum_{f \in F} \sum_{t \in T_H \setminus \{\bar{t}+\bar{d}\}} r_f \left(\rho^+_{ft} + \rho^-_{ft} \right) \tag{41}$$

subject to
(6)-(19), (21)-(23), (26)-(40)

$$1 + \sum_{h \in T_H \setminus \{\bar{t}+\bar{d}\}} \gamma_{ft} = \sum_{t \in T_H} u_{ft} \qquad f \in F \tag{42}$$

$$\sum_{h \in H} \sum_{t \in T_H} capH_h\, \delta_{hjt} = \sum_{i \in V_S} \sum_{t \in T_H} \sum_{d \in D_S} p^S_{ijd}\, y_{ijtd} \qquad j \in J_S \tag{43}$$

$$\sum_{h \in H} \sum_{t \in T_H} capH_h\, \delta_{hjt} = \sum_{i \in V_E} \sum_{t \in T_H} \sum_{d \in D_E} p^E_{ijd}\, y_{ijtd} \qquad j \in J_E \tag{44}$$

$$\rho^+_{ft},\ \rho^-_{ft} \ge 0 \qquad f \in F,\ t \in T_H \setminus \{\bar{t}+\bar{d}\} \tag{45}$$

$$\gamma_{ft} \in \{0, 1\} \qquad f \in F,\ t \in T_H \setminus \{\bar{t}+\bar{d}\} \tag{46}$$

$$\delta_{hjt} \ge 0 \qquad h \in H,\ j \in J,\ t \in T_H \tag{47}$$

$$\alpha_{hjtr} \in \{0, 1\} \qquad h \in H,\ j \in J,\ t \in T_H,\ r \in R_{ht} \tag{48}$$

$$\beta_{hjtr} \ge 0 \qquad h \in H,\ j \in J,\ t \in T_H,\ r \in R_{ht} \tag{49}$$

Naturally, the linearization of the original formulation results in a larger model, especially with respect to the total number of constraints, which grow significantly. This aspect will be illustrated in Section 4 (cf. Table 2) for the instances considered in our computational study.

3. Solution methodology

Several solution approaches have been proposed for optimization problems that include sugarcane crop and harvest planning decisions. Florentino et al. (2020) develop a genetic algorithm (GA) and a hybrid metaheuristic, combining variable neighborhood search and GA, while Poltroniere et al. (2021) design a heuristic procedure that uses relax-and-fix and fix-and-optimize strategies. Driven by the tactical nature of our problem, we chose to develop an exact method rather than a heuristic algorithm. In practice, since planning is carried out well in advance (usually many weeks before the planting season), it is not paramount to obtain Pareto-optimal solutions in a short computing time. This fact allows us to invest in a method for which computational performance does not have to be stringent. Moreover, the availability of optimal (or near-optimal) solutions is extremely important for being able to evaluate the quality of the feasible solutions returned by a (meta-) heuristic that may eventually be developed.

The tailored exact method that we present in this section is based on the augmented Chebyshev scalarization technique. Several reasons motivated our choice. On the one hand, our approach has theoretical advantages over the classical ε-constraint method (Ehrgott, 2005) and the Progressive Bounded Constraint procedure (Aliano Filho et al., 2021; Gonçalves et al., 2019). In addition, it requires a single optimization step to obtain an efficient solution (other than a lexicographic solution), which is also advantageous. Furthermore, preliminary tests with the ε-constraint method revealed that it is very challenging to obtain feasible integer solutions within reasonable computing time. By contrast, the method that we propose uses an initial feasible integer solution that greatly helps reduce the computational effort required to determine the optimal solutions to the various subproblems throughout the optimization process. Moreover, our solution approach is also easy to implement. Finally, a further advantage is that this is a general procedure that can be used to solve linear programming problems with three objectives, either with integer variables or not.

The method that we describe in this section consists of three phases. Phase 1 is designed to obtain a feasible solution to the problem defined by the objective functions (3), (4), and (41), subject to constraints (6)–(19), (21)–(23), (26)–(40), and (42)-(49). In Phase 2, lexicographic ordering is applied, and the ideal and anti-ideal (nadir) vectors are determined. Finally, in the third phase, a linear version of the min-max augmented Chebyshev scalarization method (Ehrgott, 2005; Miettinen, 1999) is employed to identify further Pareto-optimal solutions to our problem. In particular, this phase enables non-supported efficient solutions to be found. The initial feasible solution (Phase 1) is used in the subsequent phases in an attempt to reduce the computational effort. Our scheme requires solving multiple single objective subproblems to optimality. As will be shown in Section 4, this is achieved by using general-purpose optimization software. Figure 2 illustrates the three phases of the developed method, along with their interconnection.

Let ν_k denote the kth objective function ($k \in \{1, 2, 3\}$) and let $\mathbf{s}$ be a feasible solution to the (linearized) problem, i.e., $\mathbf{s} = (\mathbf{x}, \mathbf{y}, \mathbf{u}, \mathbf{e}, \mathbf{w}, \mathbf{z}, \boldsymbol{\rho}^+, \boldsymbol{\rho}^-, \boldsymbol{\gamma}, \boldsymbol{\delta}, \boldsymbol{\alpha}, \boldsymbol{\beta})$. The feasible space $\mathcal{S}$ is defined by the constraints given above. For convenience, we set $\nu_1 = -\nu_1$

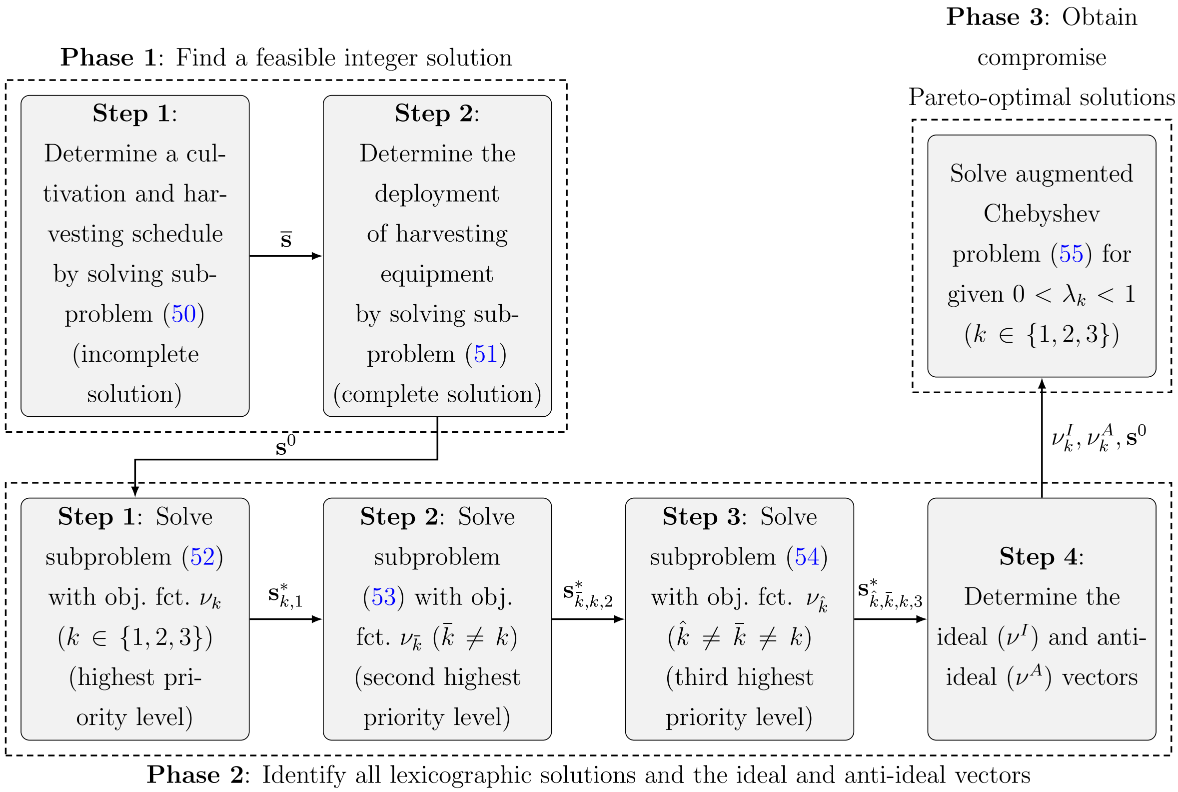

Fig. 2. Flowchart of the proposed methodology. $\bar{\mathbf{s}}$, $\mathbf{s}^0$, $\mathbf{s}^*_{k,1}$, $\mathbf{s}^*_{\bar{k},k,2}$, and $\mathbf{s}^*_{\hat{k},\bar{k},k,3}$ indicate different solutions that are obtained in the course of the algorithm.

so that maximizing the objective (3) is equivalent to minimizing ν_1. Moreover, $\nu_2 = \nu_2$ (cf. (4)) and $\nu_3 = \nu_3$, with ν_3 given by the linear objective function (41).

Phase 1: *Determining a feasible integer solution*

An initial feasible solution is identified by decomposing the problem into two subproblems and ignoring the three objective functions. Solving the first subproblem (Step 1) results in a schedule for planting and harvesting sugarcane on farms that satisfies the capacity and demand constraints as well as other technical requirements. The deployment of the harvesting equipment, represented by variables z_{hjt} and w_{jt}, is deferred to the second subproblem (Step 2).

Step 1 *Cultivation and harvesting schedule*

$$\begin{aligned} &\text{minimize } \nu_4 = q \\ &\text{subject to } (6)-(19), (26)-(29), (33)-(35), (42), (46) \\ &\qquad q = 1 \end{aligned} \tag{50}$$

Let us denote the optimal solution to problem (50) by $\bar{\mathbf{s}} = (\bar{\mathbf{x}}, \bar{\mathbf{y}}, \bar{\mathbf{u}}, \bar{\mathbf{e}}, \bar{\boldsymbol{\gamma}})$. Using this incomplete, yet feasible, solution $\bar{\mathbf{s}}$ to the original problem, we now select the harvesting equipment to be deployed and determine the associated working time by solving the following reduced problem.

Step 2 *Deployment of harvesting machines*

$$\begin{aligned} &\text{minimize } \nu_4 = q \\ &\text{subject to } (21)-(23), (36)-(40), (43)-(44), (47)-(49) \\ &\qquad q = 1 \end{aligned} \tag{51}$$

Since the three objective functions are not explicitly considered, Steps 1 and 2 return a feasible solution $\mathbf{s}_0$ to the original problem with relatively low computational effort.

Phase 2 is designed to identify the lexicographic solutions. This involves solving 15 single objective subproblems by ranking the three objective functions according to their importance. This task is performed in three steps. Furthermore, in the last step, the ideal vector and the anti-ideal vector are obtained.

Phase 2: *Determining lexicographic solutions*

Lexicographic ordering starts by minimizing each objective individually over the feasible space.

Step 1 *Highest priority level*

$$\nu^*_{k,1} = \min\{\nu_k(\mathbf{s}) : \mathbf{s} \in \mathcal{S}\}, \qquad k \in \{1,2,3\}. \tag{52}$$

Each subproblem is solved from the initial feasible solution, $\mathbf{s}_0$, which results in a significant decrease in computational effort. For each k, let $\mathbf{s}^*_{k,1}$ be the optimal solution to subproblem (52) and let $\nu^*_{k,1}$ be its optimal value ($k \in \{1,2,3\}$). In addition, the objective functions $\nu_\ell(\mathbf{s}^*_{k,1})$ are also evaluated ($\ell \in \{1,2,3\}$, $\ell \neq k$) and their values retained as they will be required in Step 4.

Step 2 *Second highest priority level*

In this step, the second most important objective is optimized by adding a new constraint that guarantees that the first objective function preserves its optimal value. Therefore, six subproblems are solved.

$$\nu^*_{\bar{k},k,2} = \min\left\{\nu_{\bar{k}}(\mathbf{s}) : \nu_k(\mathbf{s}) \le \nu^*_{k,1},\ \mathbf{s} \in \mathcal{S}\right\}, \quad \bar{k}, k \in \{1,2,3\},\ \bar{k} \neq k. \tag{53}$$

For each $\bar{k}$, the optimal solution $\mathbf{s}^*_{\bar{k},k,2}$ to subproblem (53) is determined starting from the optimal solution $\mathbf{s}^*_{k,1}$ to the subproblem it originated from at the previous level, i.e., in which the objective function ν_k was given the highest preference ($k \neq \bar{k}$). This strategy helps expedite the process of solving each subproblem. Moreover, for each $\mathbf{s}^*_{\bar{k},k,2}$, the associated value of the objective function ν_ℓ is also determined for $\ell \in \{1,2,3\}$, $\ell \neq \bar{k}$, and $\ell \neq k$.

Step 3 *Third highest priority level*

Finally, the least important objective is optimized in the feasible space, with the latter extended with two constraints that ensure the optimal values of the two most important objectives are preserved. In total, six single objective subproblems are solved.

$$\nu^*_{\hat{k},\bar{k},k,3} = \min \left\{ \nu_{\hat{k}}(\mathbf{s}) : \nu_k(\mathbf{s}) \le \nu^*_{k,1},\ \nu_{\bar{k}}(\mathbf{s}) \le \nu^*_{\bar{k},k,2},\ \mathbf{s} \in \mathcal{S} \right\}, \quad \hat{k}, \bar{k}, k \in \{1,2,3\},\ \hat{k} \neq \bar{k} \neq k. \tag{54}$$

Again, we use the optimal solutions $\mathbf{s}^*_{\bar{k},k,2}$ obtained in Step 2 to solve subproblems (54) to optimality as a way to reduce the computational effort. Let $\mathbf{s}^*_{\hat{k},\bar{k},k,3}$ be the optimal solutions to subproblems (54) and $\nu^*_{\hat{k},\bar{k},k,3}$ their optimal values. Note that $\mathbf{s}^*_{\hat{k},\bar{k},k,3}$ are the lexicographic solutions to the original problem ($\hat{k}, \bar{k}, k \in \{1,2,3\}$, $\hat{k} \neq \bar{k} \neq k$). Even though they reflect 'extreme' cases due to the predominance of one objective over the other two, their images in the objective space provide critical reference points for the decision-maker (i.e., lower bounds on the different criteria).

Step 4 *Determining the ideal and anti-ideal vectors*

Let $\nu^I = \left(\nu^I_1, \nu^I_2, \nu^I_3\right)^T$ be the ideal vector. Its components are available from Step 1, i.e., $\nu^I = \left(\nu^*_{1,1}, \nu^*_{2,1}, \nu^*_{3,1}\right)^T$. In addition, let $\nu^A = \left(\nu^A_1, \nu^A_2, \nu^A_3\right)^T$ be the anti-ideal vector. The components of ν^A are given by the worst objective values that were obtained in the course of determining the set of lexicographic solutions. In this way, upper bounds on all three objectives are available, which are also relevant for the decision-maker.

In the last phase of the method, additional Pareto-optimal solutions are identified with a linear version of the augmented Chebyshev scalarization method. Our approach enables the generation of compromise solutions that represent non-supported non-dominated solutions.

Phase 3: *Determining compromise Pareto-optimal solutions*

The two vectors ν^I and ν^A obtained in Phase 2 (Step 4) are required to define the constraints of subproblem (55). To reduce the computational effort, the latter is solved from the initial solution $\mathbf{s}_0$ obtained in Phase 1.

$$\begin{aligned} \text{minimize} \quad & \nu_5(\mathbf{s}) = \mu + \varepsilon(\nu_1(\mathbf{s}) + \nu_2(\mathbf{s}) + \nu_3(\mathbf{s})) \\ \text{subject to} \quad & \lambda_k \frac{\nu_k(\mathbf{s}) - \nu^I_k}{\nu^A_k - \nu^I_k} \le \mu, \qquad k \in \{1,2,3\}, \\ & \mu \ge 0,\ \mathbf{s} \in \mathcal{S}, \end{aligned} \tag{55}$$

where μ is an auxiliary variable, the parameter $\lambda_k > 0$ is pre-specified for $k \in \{1,2,3\}$ such that $\lambda_1 + \lambda_2 + \lambda_3 = 1$, and ε is a suitable small positive constant that prevents weakly efficient solutions from being generated. Observe that subproblem (55) is a linear version of the non-linear min-max Chebyshev distance model, i.e.,

$$\min \left\{ \max_{1 \le k \le 3} \left\{ \lambda_k \frac{\nu_k(\mathbf{s}) - \nu^I_k}{\nu^A_k - \nu^I_k} \right\} : \mathbf{s} \in \mathcal{S} \right\}.$$

We emphasize that by varying the values of parameters λ_k, $k \in \{1,2,3\}$, distinct non-dominated solutions are obtained to the original problem that achieve different trade-offs among the three objectives.

4. Computational study

In this section, we present and discuss the results of an extensive computational study. In Section 4.1, we describe how the data were generated. Numerical results are reported in Section 4.2 and relevant insights into the characteristics of the Pareto-optimal solutions identified are discussed in Section 4.3.

4.1. Test instances

The basis of our study consists of four semi-randomly test instances of realistic size that represent the current cultivation and harvesting practices in the center-south region of Brazil, which is the heart of the Brazilian sugarcane industry. The instances differ in the number of farms, plots, and cultivation areas as shown in Table 1. The smallest instance has 30 plots across 5 farms, while the largest instance has 240 plots across 37 farms. A simple procedure is used to define the neighborhoods of the plots, namely for each plot $j \in J$ we set $N(j) = \{j, p_1, p_2, p_3\}$, with $p_i = j + i$ if $j + i \le |J|$ and $p_i = j - i$ otherwise, for $i \in \{1,2,3\}$.

As in the study by Poltroniere et al. (2021), we consider 25 different sugarcane varieties, of which 20 are for sucrose-cane and the remaining five for energy-cane. Their productivity levels (s_{id}, m_{id}) are specified according to Poltroniere et al. (2021). Each sugarcane variety can be planted on 30 percent of the plots ($\eta_i = 0.3$, $i \in V$). For example, for the largest instance, this means that $0.3\,k_S = 48$ plots and $0.3\,k_E = 24$ plots are grown with a specific variety of sucrose-cane and energy-cane, respectively. The area of each plot is randomly selected from a normal distribution $\mathcal{N}(\mu, \sigma)$ with mean $\mu = 25$ ha and standard deviation $\sigma = 15$ ha, which are realistic parameters. Taking into account the productivity of each individual sugarcane variety, the areas of the plots (a_j), and the number of plots dedicated to sucrose-cane (k_S) and energy-cane (k_E) in each instance, we have estimated the mean demand and the standard deviation per product and period. These estimates are used to randomly draw the demand for each type of sugarcane per period from a normal distribution, see d^S_t and d^E_t in Table 1. The last column of the table also gives the mill capacity ($capM_t$) in each instance, which is 5 percent greater than the total amount of sugarcane that the plots could produce if they were all harvested at the peak maturity period.

The sucrose content of traditional sucrose-cane varieties ranges from 10 percent to 15 percent (Florentino et al., 2020; Matsuoka, Bressiani, Maccheroni, & Fouto, 2012). Recent studies have indicated that the more fibrous energy-cane varieties exceed 35 percent of dry mass (fiber and Brix, the latter being a measure of sugar content) (Matsuoka, dos Santos, & Tomazela, 2017; Santchurn, Ramdoyal, Badaloo, & Labuschagne, 2012). Based on these values, we have assumed an average of 12.5 percent sucrose for sucrose-cane varieties and 37.5 percent fiber for energy-cane varieties. As a result, we have set $\theta_S = \frac{100}{12.5} = 8$ and $\theta_E = \frac{100}{37.5} = 2.67$ as conversion factors in constraints (17).

The overall planning horizon spans 23 periods, with each period representing one month. Sucrose-cane varieties with an 18-month cycle can be planted in the first four months (i.e., from January to April), and therefore $T_P = \{1,2,3,4\}$ for every $i \in V^2_S$. Thus, depending on the month of cultivation, the highest sucrose content is reached in month 19, 20, 21, or 22. Sucrose-cane varieties with a 12-month cycle can be planted in September and October, i.e., $T_P = \{9,10\}$ for every $i \in V^1_S$. In this case, maturity is attained either in month 21 or month 22. The planting season for energy-cane varieties is $T_P = \{1,2,3,4,9,10\}$ ($i \in V_E$). For any given sugarcane variety, harvesting can occur up to three months before or after the peak maturity month, i.e., $\underline{d} = -3$, $\bar{d} = 3$, and $D_S = D_E = \{-3, -2, \ldots, 2, 3\}$. This implies that for 18-month varieties, the harvest time horizon starts in month 16 (April) and ends

Table 1
Parameters related to farms, plots, demands for sucrose-cane and energy-cane, and mill capacity.

$\|F\|$	$\|F_S\|$	$\|F_E\|$	$\|J\|$	k_S	k_E	η_i	a_j (ha)	d_t^S (ton.)	d_t^E (ton.)	$capM_t$ (ton.)
5	3	2	30	20	10	0.3	$\mathcal{N}(25, 15)$	$\mathcal{N}(9000, 100)$	$\mathcal{N}(7000, 100)$	$\mathcal{N}(100000, 100)$
10	6	4	60	40	20			$\mathcal{N}(18000, 100)$	$\mathcal{N}(14000, 100)$	$\mathcal{N}(200000, 100)$
22	14	8	120	80	40			$\mathcal{N}(40000, 100)$	$\mathcal{N}(30000, 100)$	$\mathcal{N}(400000, 100)$
37	24	13	240	160	80			$\mathcal{N}(80000, 100)$	$\mathcal{N}(60000, 100)$	$\mathcal{N}(800000, 100)$

in month 23 (November). Note that due to the total length of the planning horizon, the maximum deviation from peak maturity is reduced to $\bar{d} = 2$ and $\bar{d} = 1$ when sugarcane is cultivated in March ($t = 3$) and April ($t = 4$), respectively. For varieties with a 12-month cycle, the largest deviation is $\bar{d} = 2$ ($\bar{d} = 1$) for cane planted in September (October) of the preceding year (i.e., $t = 9$ or $t = 10$). In short, $T_H = \{16, \ldots, 23\}$.

Furthermore, two types of harvesters ($|H| = 2$) that can cut up to $capH_1 = 20$ and $capH_2 = 25$ tonnes of sugarcane per hour, respectively, are considered. Each harvesting machine can be operated up to 16 hours per day, corresponding to two 8-hour work shifts, on 25 days per month, leading to a monthly availability of $\bar{w} = 400$ hours. In the small instances with 5 and 10 farms, the total number of harvesters of type h available in month t is given by $\bar{z}_{ht} = \lceil \mathcal{N}(15, 1) \rceil$, while for the larger instances we set $\bar{z}_{ht} = \lceil \mathcal{N}(25, 1) \rceil$ ($h \in H$, $t \in T_H$). The total cost of moving a harvester between the depot and farm f is $r_f = c\, d_f^0$, with unit cost $c = 10$ R\$ per km and travel distance d_f^0 randomly generated from a normal distribution with mean $\mu = 50$ km and standard deviation $\sigma = 20$ km.

4.2. Numerical results

The multi-objective MILP formulation developed in Section 2.3 was implemented with the JuMP modeling language (version v0.20, Dunning, Huchette, & Lubin, 2017) and embedded in the Julia programming language (version 1.5.0, Bezanson, Edelman, Karpinski, & Shah, 2017). All experiments were performed on a laptop computer with a 2.5 GHz Intel Core i7-2450M processor, 8 GB RAM, and running a 64-bit operating system. Subproblems (50)–(55) were solved with IBM ILOG CPLEX 12.8. A limit of 3,600 seconds of CPU time and an optimality gap of 0.01% were set for each solver run.

The size of the Pareto-optimal set returned by the method described in Section 3 greatly depends on the number of different combinations chosen for parameters λ_k in Phase 3. In our computational study, we have opted to focus our analysis on a subset of non-dominated solutions that are not only good representatives of the conflicting nature of the objectives (3), (4), and (41), but also provide relevant information about the trade-offs achieved by each alternative. Through this subset, a decision-maker will become aware of the practical implications of different cultivation and harvesting schedules. To this end, the results that will be presented in this section and the next section refer to a subset with four Pareto-optimal solutions, three of which are lexicographic solutions and the fourth is the outcome of solving the linearized Chebyshev min-max subproblem (55) taking $\varepsilon = 10^{-6}$ and $\lambda_k = 1/3$ for $k \in \{1, 2, 3\}$. This choice represents a balanced compromise between the three objectives and is supported by the studies of Florentino et al. (2018) and Aliano Filho, Florentino, Pato, Poltroniere, & Costa (2022) for other multi-objective optimization problems in the Brazilian sugarcane production chain. In practice, other values for λ_k could be fixed following a discussion with the decision-maker. As will be shown, our choice provides an acceptable trade-off between computational effort and number of efficient solutions generated by the proposed method while avoiding an overly extensive comparative analysis. Among the lexicographic solutions returned by Phase 2, we will analyze the solutions for $[k, \bar{k}, \hat{k}] = [1, 2, 3]$ (Pareto-optimal solution 1), $[k, \bar{k}, \hat{k}] = [2, 1, 3]$ (Pareto-optimal solution 2), and $[k, \bar{k}, \hat{k}] = [3, 1, 2]$ (Pareto-optimal solution 3). We note that the characteristics of the remaining lexicographic solutions do not significantly differ from the ones we have selected. For example, the efficient solution associated with $[k, \bar{k}, \hat{k}] = [1, 3, 2]$ shares many of the features of solution 1. Similar observations hold for $[k, \bar{k}, \hat{k}] = [2, 3, 1]$ compared to solution 2 and for $[k, \bar{k}, \hat{k}] = [3, 2, 1]$ compared to solution 3. Since 12 subproblems need to be solved to obtain the set of selected efficient solutions for an individual instance, in total 48 subproblems are solved.

For each number of farms and plots considered in a test instance, Table 2 presents the total number of variables and constraints in the non-linear formulation (3)–(31) (columns 3–4) and the equivalent linear formulation (3)–(4), (6)–(19), (21)–(23), (26)–(49) (columns 5–6). Furthermore, the average optimality gap (column 7) and the average CPU time (column 8) returned by CPLEX are also reported for each solution. For simplicity, we will continue to use the term Pareto optimality even when the optimality of a particular solution is not guaranteed because the pre-specified time limit was reached.

Unsurprisingly, the linearization techniques greatly affect the size of the equivalent formulation. In particular, the total number of constraints increases, on average, by a factor of 13.4, while 15 percent more variables are included in the linear model. Despite the large size of the linear formulation, our method identifies high-quality solutions, as demonstrated by the optimality gaps of less than one percent in nine of the 16 solutions. The remaining seven solutions have larger, but still satisfactory, optimality gaps taking into account the short time limit. Compared to the other lexicographic solutions, this feature suggests that it is computationally more expensive to assign the highest preference to minimizing the total number of harvesting fronts (4), followed by minimizing the total transportation cost of the harvesters (41), and finally maximizing the total amount of sucrose and fiber produced (3).

Table 3 gives the objective values of the different efficient solutions (columns 3–5). To facilitate the comparative analysis, we also present for each solution and objective function, the percent deviation of its objective value from the best value among the four solutions available (see columns 6–8). The smallest deviation (0%) is highlighted in bold type. The conflicting nature of the three objectives is clearly reflected in the results obtained, particularly in solutions 1–3. Most notably, to achieve a higher level of productivity (solutions 1) it is necessary to harvest the sugarcane at or very close to its peak maturity, especially for the sucrose-cane varieties that are the majority of the varieties considered. This involves concentrating the harvesting effort on certain months. As a result, more equipment is needed in some months on different farms, thus significantly increasing the number of fronts (ν_2) and machine transportation costs (ν_3). As expected, placing preference on minimizing the number of farms that are harvested simultaneously (solutions 2) requires the deployment of a larger number of harvesting machines, thereby negatively affecting their transportation costs (ν_3). At the same time, it is not possible to harvest all the plots at their ideal time, which reduces total productivity, especially for sucrose-cane varieties. Accordingly, the worst perfor-

Table 2
Size of test instances and CPLEX performance.

$\|F\|/\|J\|$	Sol.	MINLP formulation		MILP formulation		CPLEX performance	
		# var.	# const.	# var.	# const.	Gap (%)	CPU time (sec.)
5/30	1	38,268	1,540	43,653	18,485	< 0.01	1,830
	2					1.25	3,600
	3					< 0.01	1,445
	4					< 0.01	1,224
10/60	1	76,528	3,015	87,298	36,905	0.00	2,341
	2					2.23	3,600
	3					< 0.01	2,876
	4					0.14	3,600
22/120	1	153,064	5,983	178,486	87,261	< 0.01	3,376
	2					4.21	3,600
	3					1.65	3,600
	4					0.86	3,600
37/240	1	306,064	11,838	356,761	174,191	< 0.01	3,552
	2					5.25	3,600
	3					2.33	3,600
	4					3.35	3,600

Table 3
Performance of the solutions identified.

$\|F\|/\|J\|$	Sol.	Objective values			Dev. to best obj. value (%)		
		Production ν_1 (ton.)	# harv. fronts ν_2	Transp. cost ν_3 (R$)	ν_1	ν_2	ν_3
5/30	1	160,589.33	23	16,500.00	**0.0**	43.8	223.5
	2	156,486.27	16	16,300.00	-2.6	**0.0**	219.6
	3	160,395.19	27	5,100.00	-0.1	68.8	**0.0**
	4	158,849.76	21	9,700.00	-1.1	31.3	90.2
10/60	1	317,829.63	45	21,400.00	**0.0**	125.0	100.0
	2	295,728.16	20	36,500.00	-7.0	**0.0**	241.1
	3	307,080.79	53	10,700.00	-3.4	165.0	**0.0**
	4	304,502.30	40	25,700.00	-4.2	100.0	140.22
22/120	1	633,893.06	82	94,600.00	**0.0**	173.3	192.9
	2	595,503.08	30	132,600,00	-6.1	**0.0**	310.5
	3	591,552.99	94	32,300.00	-6.7	213.3	**0.0**
	4	601,496.50	38	107,300.00	-5.1	26.7	232.2
37/240	1	1,065,431.72	130	217,980.00	**0.0**	60.5	53.8
	2	996,694.33	81	287,440.00	-6.5	**0.0**	102.8
	3	1,002,142.47	139	141,720.00	-5.9	71.6	**0.0**
	4	1,026,490.76	99	224,660.00	-3.7	22.2	58.5

mance with respect to objectives ν_1 and ν_3 is achieved with solutions 2 for almost all test instances. In contrast to this alternative, minimizing the total transportation cost (solutions 3) clearly explains the greatest deterioration of objective ν_2 through scheduling the largest number of harvesting fronts. Moreover, as the size of the instances increases, the level of sucrose and fiber productivity drops markedly. As for solutions 4, the compromise between the three objectives is noticeable in Table 3. The number of harvesting fronts always takes second place in the ranking of the identified solutions, while productivity and total transportation cost appear either in the second or third position. In conclusion, the information provided by Table 3 is very useful from a practical viewpoint, as it enables the decision-maker to understand the trade-offs between the three criteria, and thus supports his/her choice of the schedule that could be implemented. A more detailed analysis of the different characteristics of each solution is presented in the supplementary material associated with this article (see Figures 5 and 6) for the instance with 5 farms and 30 plots. Accordingly, the harvesting schedules are specified for the individual sugarcane types grown, and compared using several key performance measures. In addition, a solution created manually using simple rules commonly found in practice is also shown in Table 7 in the supplementary material for this instance. This solution is evaluated against the available Pareto-optimal solutions, and its inferior quality is discussed. Finally, a brief analysis is also carried out on the total number of different sugarcane varieties planted on a farm in all instances generated, see Table 8 in the supplementary material.

4.3. Managerial insights

Further insights into the features of the Pareto-optimal solutions identified for each instance are provided in this section. The performance measures to be analyzed include: (i) the choice of equipment and its use with respect to time worked; (ii) harvest efficiency; (iii) machine transportation spending; and (iv) harvesting deviations from the peak maturity month. At the end of the section, we give a summary of the main findings of our analysis, thus providing a decision-maker with a broad view of the main trade-offs in the alternative planting and harvesting schedules.

Table 4 reports the average number of harvesters of each type allocated to a plot (columns 3 and 4) as well as the total number of machines deployed per plot. Columns 6–8 give the average number of hours worked on each plot, farm, and per harvester, respectively. The last column presents the average number of months devoted to harvesting on a farm. For each instance, the best values among the four Pareto-optimal solutions are highlighted in bold type. In

Table 4
Number of harvesters and hours worked.

$\|F\|/\|J\|$	Sol.	# harvesters per plot (avg.)			Hours worked (avg.)			# of months to
		$h = 1$	$h = 2$	Total	Per plot	Per farm	Per machine	finish the harvest on a farm (avg.)
5/30	1	1.91	2.10	4.01	169.56	1,400.64	42.29	4.60
	2	1.92	2.22	4.15	**138.39**	**1,386.48**	**33.38**	**3.20**
	3	1.08	1.06	2.14	228.68	1,420.67	106.92	5.40
	4	1.07	1.05	**2.12**	208.29	1,392.41	98.18	4.20
10/60	1	1.27	1.53	2.80	196.79	1,422.14	70.38	4.50
	2	1.07	1.21	2.28	201.56	**1,309.85**	88.51	**2.00**
	3	1.05	1.08	**2.12**	213.80	1,338.41	100.64	5.30
	4	1.44	1.51	2.95	**173.64**	1,353.88	**58.77**	4.00
22/120	1	2.00	1.75	3.75	**170.37**	1,243.82	**45.43**	3.73
	2	2.05	1.45	3.49	183.31	1,179.35	52.48	**1.36**
	3	1.00	1.00	**2.00**	195.58	**1,081.73**	97.79	4.27
	4	1.71	1.46	3.17	180.83	1,190.65	57.04	1.73
37/240	1	1.43	1.47	2.89	183.80	1,445.15	63.52	3.51
	2	1.43	1.47	2.90	**166.26**	1,349.94	**57.37**	**2.19**
	3	1.34	1.19	**2.53**	180.96	**1,312.49**	71.62	3.76
	4	1.29	1.39	2.68	179.09	1,372.10	66.88	2.68

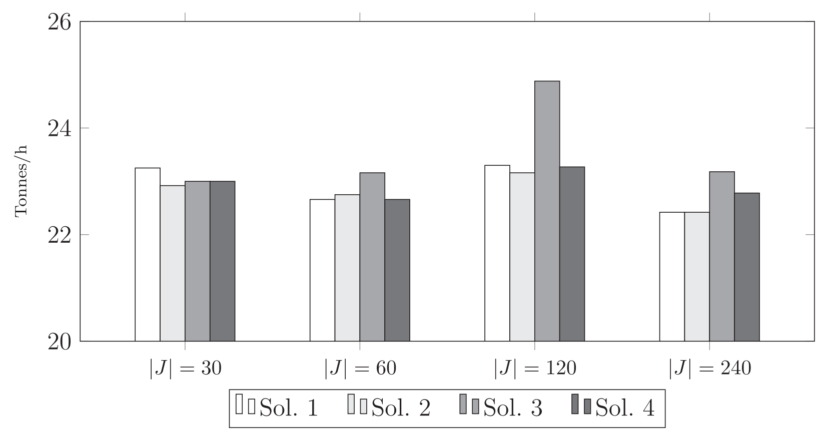

Fig. 3. Average amount of sugarcane harvested per hour.

almost all solutions, there is a preference for operating more harvesting machines of type 2 than of type 1, as the former have larger capacity, which allows for increased production in a shorter period of time. As expected, on average, fewer harvesters are used per plot in those solutions that give more relevance to minimizing the total cost of transporting the machines between the depot and the farms (solutions 3). However, each machine is operated longer and thus subject to more wear. In turn, the maintenance effort and even the risk of machine failure increase. These aspects are comparatively less marked in the other solutions. When minimizing the number of harvesting fronts is the most important criterion (solutions 2), harvesting is performed faster on each farm, as indicated in columns 7 and 9. This feature is especially striking in the instance with 22 farms and 120 plots. We note that the small fluctuations in the average number of hours worked on each farm (column 7) are due to differences in the production levels from one solution to another and the types of machines used.

Harvesting efficiency is related to the quantity of sugarcane that is cut per hour. Figure 3 illustrates this performance measure for the different Pareto-optimal solutions. It can be seen that the harvesting schedules generated by solutions 3 offer greater efficiency, which can be attributed in part to a more intensive use of the machines with higher harvesting capacity over the planning horizon.

Figure 4 displays the relative cost of machine transports per 100 tonnes of sucrose and fiber produced. The costs rise as the number of fronts, and therefore also the number of farms, where harvesting is carried out in the same month, increase because more harvesting machines need to be deployed. Hence, confirming our previous analysis, solutions 2 are the most expensive, whereas solutions 3 have the lowest cost. Compromise solutions 4 are likely to have a slightly higher cost than solutions 1, since the former use a larger number of harvesters on average (recall Table 4).

The month in which a cane variety is harvested affects not only the level of sucrose and fiber yields, but also the deployment of machinery. Recall from (1) that anticipating or delaying the harvest of any sucrose-cane variety is detrimental. On the contrary, delaying the harvest of energy-cane varieties (in our case, up to 2 months so that the harvest takes place within the planning horizon) results in higher fiber production, see (2). Table 5 addresses this issue by reporting in columns 3 and 5 the percent of plots that are harvested at the ideal month. Columns 4 and 6 give the mean of all absolute deviations. For the larger instances, solutions 1 that favor maximizing sucrose and fiber yields provide the best results for both types of sugarcane. Surprisingly, this effect is not observed in the smaller instances. Regardless of the Pareto-optimal solution, the proportion of plots with energy-cane harvested at its peak maturity is always higher than the proportion of sucrose-cane plots. However, more variability occurs in the harvest months for the energy-cane varieties, thus explaining why the averages are higher for this type of sugarcane.

Based on the previous analysis and to facilitate the decision-making process, we summarize in Table 6 the main advantages

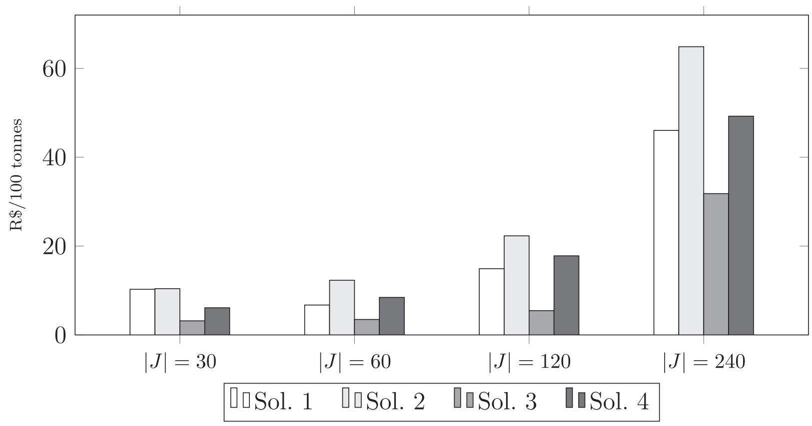

Fig. 4. Transportation cost per 100 tonnes of sugarcane harvested.

Table 5
Deviations from the ideal month for harvesting sugarcane.

$\|F\|/\|J\|/k_S/k_E$	Sol.	Sucrose-cane varieties		Energy-cane varieties	
		% of deviations equal to 0	Absolute avg. dev. (months)	% of deviations equal to 3	Absolute avg. dev. (months)
5/30/20/10	1	25.0	1.4	30.0	1.6
	2	25.0	1.5	**50.0**	1.8
	3	**35.0**	**1.3**	40.0	**2.1**
	4	25.0	1.5	30.0	1.8
10/60/40/20	1	17.5	**1.4**	25.0	1.4
	2	**25.0**	1.6	30.0	1.8
	3	7.5	1.7	**40.0**	**1.8**
	4	15.0	1.6	25.0	1.4
22/120/80/40	1	**28.8**	**1.3**	**50.0**	**2.0**
	2	16.3	1.7	32.5	1.9
	3	22.5	1.7	22.5	1.8
	4	15.0	1.6	32.5	2.0
37/240/160/80	1	**16.9**	**1.8**	**23.8**	**1.8**
	2	10.0	2.0	12.5	1.6
	3	10.6	2.0	15.0	1.7
	4	11.3	2.0	13.8	1.6

Table 6
Summary of advantages and disadvantages of lexicographic solutions.

Feature	Solution 1		Solution 2		Solution 3	
	Adv.	Disadv.	Adv.	Disadv.	Adv.	Disadv.
Total production (sucrose and fiber) (ν_1)	✓			✓		
Deviations regarding maturation	✓					✓
Number of varieties planted		✓	✓			
Number of harvesting fronts (ν_2)			✓			✓
Effort of harvesters			✓			✓
Harvest duration on a plot			✓			✓
Harvest duration on a farm			✓			✓
Harvest efficiency				✓	✓	
Machine transportation costs (ν_3)				✓	✓	
Number of harvesters				✓	✓	

(columns 2, 4, and 6) and drawbacks (columns 3, 5 and 7) of the three lexicographic solutions for a number of relevant features (column 1). Although these solutions reflect ‘extreme’ situations, they frame the ranges of the efficient alternative schedules, and are therefore useful for gaining insight into the complex trade-offs that occur. Our evaluation disregards Pareto-solution 4 because its characteristics are not pronounced due to the balance between the three objectives achieved by the values assigned to λ_k in (55). The features listed in the table are organized into three categories, each related to an individual objective, to provide a better understanding of the influence of different preference information. For example, a decision-maker favoring a short harvest becomes aware that this perspective has a lower productivity level, incurs higher machine transportation costs, requires greater equipment maintenance effort, and has decreased harvesting efficiency.

5. Conclusions

We studied a problem in the Brazilian sugarcane production chain that integrates decisions on sugarcane planting, sugarcane harvesting, and the deployment of different types of harvesting machines over a multi-period planning horizon. The aim is to maximize the level of sucrose and fiber production, to complete harvesting on all farms as early as possible, and to minimize the expenditure on transporting harvesting equipment. Due to the conflicting nature of these goals, we proposed a multi-objective mixed-integer program that includes non-linear constraints and non-linear terms in one of the objective functions. Using linearization techniques we transformed the original formulation into a computationally tractable and equivalent mixed-integer linear program. By integrating key decisions into a single model that are typically made separately, our approach improves coordination between cultivation and harvesting operations, and consequently increases overall performance at the tactical planning level.

To solve the problem at hand, we developed an exact method based on augmented Chebyshev scalarization. Finding a feasible solution early in the procedure has proven to be very useful in accelerating the identification of Pareto-optimal solutions. Besides being a general method that can be applied to any problem with three objectives, it also has the advantage of being able to identify efficient solutions in the non-convex region of the Pareto front.

To validate the mathematical formulation and methodology developed, we conducted a computational study on a set of semi-random instances that reflect current sugarcane cultivation and harvesting practices in Brazil. Through our comparative analysis of a representative subset of Pareto-optimal solutions, we have examined in depth the trade-offs between the three competing objectives, which is a crucial step to support the decision-maker and thus bring transparency to the decision process. Our study revealed that, on average, the productivity level of sucrose and fiber varies 6 percent among the solutions analyzed, there is a 53.2 percent variation regarding the number of harvesting fronts, and a 66.5 percent variation for the total transportation cost of the harvesters. These figures give evidence of the relevance of using a multi-objective approach to this problem. Maximizing sucrose and fiber production yields leads to an increase in sucrose stocks, a decrease in the number of different sugarcane varieties grown (which makes the risk of crop damage higher), greater harvest effort, and therefore also higher machine transportation costs. In a scenario where the number of harvesting fronts is minimal, there is a reduction in machine effort and harvesting time. However, these positive features are offset by a reduction in total production, an increased total transportation cost, and a slight loss in harvesting efficiency. The advantages of giving more emphasis to low machine transportation costs include deploying fewer harvesters, building less sucrose stock, and improving harvest efficiency. But at the same time more deviations between the harvesting periods and maximum sugarcane maturity occur (being more noticeable for the sucrose-cane varieties), the total harvesting time and the number of harvesting fronts increase, and each machine is operated for a longer time. As a result, harvesting equipment is subject to more wear, and therefore more prone to breakdowns and greater maintenance effort. Furthermore, our method can also return solutions that provide some balance between the three objectives, thus quantifying the gains and losses in each as a result of the parameters chosen in the scalarization function. These valuable insights support value-added decision-making and set the framework for detailing the planting and harvesting schedules at the operational level.

The problem we have studied is challenging because it involves a very large number of variables and constraints, especially the linear formulation. Therefore, a future line of research could be to design a tailored heuristic method in order to find approximate Pareto-optimal solutions for very large-sized instances within a reasonable time limit. For example, the heuristic algorithm developed by Poltroniere et al. (2021) for a particular case of our problem, involving a single farm, a single objective function (maximization of total production), and fewer technical requirements, could be an interesting starting point to be extended to our more complex multi-objective problem. Another future research venue could be the development of a stochastic model to explicitly capture the uncertainty associated with some parameters, for example, the weather conditions during the cane growing periods and the demand for sucrose and fiber at the mill.

Acknowledgments

The first author is grateful to the Federal Technological University of Paraná (UTFPR) for the computational and financial support provided through project no. 2019–010001582. The work of the second author was supported by the Brazilian National Council for Scientific and Technological Development (CNPq) under grant no. 309925/2021-5. This support is gratefully acknowledged. We also thank three anonymous reviewers for their valuable comments.

Supplementary material

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.ejor.2022.12.029.

References

Ahumada, O., & Villalobos, J. (2009). Application of planning models in the agri-food supply chain: A review. *European Journal of Operational Research, 196*, 1–20.
Aliano Filho, A., Florentino, H., Pato, M., Poltroniere, S., & Costa, J. (2022). Exact and heuristic methods to solve a bi-objective problem of sustainable cultivation. *Annals of Operations Research, 314*, 347–376.
Aliano Filho, A., Melo, T., & Pato, M. (2021). A bi-objective mathematical model for integrated planning of sugarcane harvesting and transport operations. *Computers & Operations Research, 134*, 105419.
Bezanson, J., Edelman, A., Karpinski, S., & Shah, V. (2017). Julia: A fresh approach to numerical computing. *SIAM Review, 59*, 65–98.
Carvajal, J., Sarache, W., & Costa, Y. (2019). Addressing a robust decision in the sugarcane supply chain: Introduction of a new agricultural investment project in Colombia. *Computers and Electronics in Agriculture, 157*, 77–89.
Colin, E. (2009). Mathematical programming accelerates implementation of agro-industrial sugarcane complex. *European Journal of Operational Research, 199*, 232–235.
CONAB (2021). Companhia Nacional de Abastecimento - Acompanhamento da safra brasileira - cana-de-açucar, v. 7-SAFRA 2020/21 - n. 4 - Quarto levantamento. (In Portuguese.) Available online: https://www.conab.gov.br/info-agro/safras/cana/boletim-da-safra-de-cana-de-acucar (last accessed on 18 April 2022).
Cursi, D., Hoffmann, H., Barbosa, G., Bressiani, J., Gazaffi, R., Chapola, R., Fernandes Junior, A., Balsalobre, T., Diniz, C., Santos, J., & Carneiro, M. (2022). History and current status of sugarcane breeding, germplasm development and molecular genetics in Brazil. *Sugar Tech, 24*, 112–133.
dos Santos, L., Michelon, P., Arenales, M., & Santos, R. (2011). Crop rotation scheduling with adjacency constraints. *Annals of Operations Research, 190*, 165–180.
dos Santos, L., Munari, P., Costa, A., & Santos, R. (2015). A branch-price-and-cut method for the vegetable crop rotation scheduling problem with minimal plot sizes. *European Journal of Operational Research, 245*, 581–590.
Dunning, I., Huchette, J., & Lubin, M. (2017). JuMP: A modeling language for mathematical optimization. *SIAM Review, 59*, 295–320.
Ehrgott, M. (2005). *Multicriteria optimization* (2nd ed.). Berlin, Heidelberg: Springer.
Fava Neves, M., & Kalaki, R. (2020). *Bionergy from sugarcane* (1st ed.). Guariba, SP: Socicana. Available online: https://unica.com.br/wp-content/uploads/2021/03/Bioenergy-from-Sugar-Cane-by-Fava-Neves-Kalaki-2021.pdf (last accessed on 18 April 2022).
Florentino, H., Cantane, D., Jones, D., Ouelhadj, D., Khosravi, B., & Irawan, C. (2020). An optimization model for combined selecting, planting and harvesting sugarcane varieties. *Annals of Operations Research, 314*, 451–469.
Florentino, H., Irawan, C., Aliano Filho, A., Jones, D., Cantane, D., & Nervis, J. (2018). A multiple objective methodology for sugarcane harvest management with varying maturation periods. *Annals of Operations Research, 267*, 153–177.
Florentino, H., & Pato, M. (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*, 842–854.
Gonçalves, E., Balbo, A., Nunes da Silva, D., Nepomuceno, L., Baptista, E., & Soler, E. (2019). Deterministic approach for solving multi-objective non-smooth environmental and economic dispatch problem. *Electrical Power and Energy Systems, 104*, 880–897.
Gupte, A., Ahmed, S., Cheon, M., & Dey, S. (2013). Solving mixed integer bilinear problems using MILP formulations. *SIAM Journal on Optimization, 23*, 721–744.
Jarumaneeroj, P., Dusadeerungsikul, P., Chotivanich, T., & Akkermang, R. (2021). A multi-objective modeling approach to harvesting resource scheduling: Decision support for a more sustainable Thai sugar industry. *Computers & Industrial Engineering, 162*, 107694.
Jena, S., & Poggi, M. (2013). Harvest planning in the Brazilian sugar cane industry via mixed integer programming. *European Journal of Operational Research, 230*, 374–384.

Junqueira, R., & Morabito, R. (2019). Modeling and solving a sugarcane harvest front scheduling problem. *International Journal of Production Economics, 213*, 150–160.

Katakojwala, R., Naresh Kumar, A., Chakraborty, D., & Venkata Mohan, S. (2019). Valorization of sugarcane waste: Prospects of a biorefinery. In M. Narasimha, V. Prasad, P. de Campos Fava, M. Vithanage, & S. Venkata Mohan (Eds.), *Industrial and municipal sludge - Emerging concerns and scope for resource recovery* (pp. 47–60). Butterworth-Heinemann. Chapter 3

Kong, Q., Kuriyan, K., Shah, N., & Guo, M. (2019). Development of a responsive optimisation framework for decision-making in precision agriculture. *Computers & Chemical Engineering, 131*, 106585.

Lamsal, K., Jones, P., & Thomas, B. (2017). Sugarcane harvest logistics in Brazil. *Transportation Science, 51*, 771–789.

Matsuoka, S., Bressiani, J., Maccheroni, W., & Fouto, I. (2012). Bioenergia da cana. *Cana-de-açúcar: Bioenergia, açúcar e álcool, 1*, 548–576.

Matsuoka, S., dos Santos, E., & Tomazela, A. (2017). Free fiber level drives resilience and hybrid vigor in energy cane. In M. Khan, & I. Khan (Eds.), *Sugarcane biofuels - Status, potential, and prospects of the sweet crop to fuel the world* (2, pp. 39–66). Springer Cham.

Matsuoka, S., & Rubio, L. (2019). Energy cane: A sound alternative of a bioenergy crop for tropics and subtropics. In M. Khan, & I. Khan (Eds.), *Sugarcane biofuels - Status, potential, and prospects of the sweet crop to fuel the world* (pp. 39–66). Springer Cham.

Miettinen, K. (1999). *Nonlinear multiobjective optimization*. Boston: Kluwer Academic Publishers.

Morales-Chávez, M., Sarache, W., Costa, Y., & Soto, J. (2020). Multiobjective stochastic scheduling of upstream operations in a sustainable sugarcane supply chain. *Journal of Cleaner Production, 276*, 123305.

Poltroniere, S., Aliano Filho, A., Balbo, A., & Florentino, H. (2021). Integrated planning for planting and harvesting sugarcane and energy-cane for the production of sucrose and energy. *Computers and Electronics in Agriculture, 134*, 105956.

Ramos, R., Isler, P., Florentino, H., Jones, D., & Nervis, J. (2016). An optimization model for the combined planning and harvesting of sugarcane with maturity considerations. *African Journal of Agricultural Research, 11*, 3950–3958.

Santchurn, D., Ramdoyal, K., Badaloo, M., & Labuschagne, M. (2012). From sugar industry to cane industry: Investigations on multivariate data analysis techniques in the identification of different high biomass sugarcane varieties. *Euphytica, 135*, 543–558.

Santos, F., Eichler, P., Machado, G., De Mattia, J., & De Sousa, G. (2020). By-products of the sugarcane industry. In F. Santos, S. Rabelo, M. De Matos, & P. Eichler (Eds.), *Sugarcane biorefinery, technology and perspectives* (pp. 21–48). Elsevier. Chapter 2

Shah, K., Modi, B., Pandey, H., Subedi, A., Aryal, G., Pandey, M., & Shrestha, J. (2021). Diversified crop rotation: An approach for sustainable agriculture production. *Advances in Agriculture, 8924087*.

Stray, B., van Vuuren, J., & Bezuidenhout, C. (2012). An optimisation-based seasonal sugarcane harvest scheduling decision support system for commercial growers in South Africa. *Computers and Electronics in Agriculture, 83*, 21–31.

Teixeira, E., Rangel, S., Florentino, H., & Araújo, S. (2021). A review of mathematical optimization models applied to the sugarcane supply chain. *International Transactions in Operational Research*. https://doi.org/10.1111/itor.13056

Thuankaewsing, S., Khamjan, S., Piewthongngam, K., & Pathumnakul, S. (2015). Harvest scheduling algorithm to equalize supplier benefits: A case study from the Thai sugar cane industry. *Computers and Electronics in Agriculture, 110*, 42–55.

UNICA (2017). União da Indústria de Cana–de–Açúcar (Brazilian Sugarcane Industry Association). Available online: https://unica.com.br/ (last accessed on 18 April 2022).