ACCEPTED VERSION

Stochastic season-wide optimal production planning of virgin olive oil

P. Cano Marchal*, D. Martínez Gila, J. Gámez García, J. Gómez Ortega

Robotics, Automation and Computer Vision Group, University of Jaén, Spain

ABSTRACT

The quality and obtained quantity of Virgin Olive Oil is determined by the influence of the process variables during the production process based on the properties of the olives being processed. Since the quality of the olives evolve during the harvesting season, a relevant question is to systematically suggest when to harvest the olives in order to maximize the profit over the whole season. This work proposes a method to determine an optimal production plan for the whole harvesting season explic-itly considering the stochastic nature of the problem, the uncertainty in the problem parameters and the information available at each step, and presents the results obtained in its application to different production scenarios.

1. Introduction

Virgin olive oil (VOO) production is an important, growing and global economic activity. Currently, more than 30 countries produce VOO, and the world average production increased 5.9% in the 2015–2010 period over the 2005–2010 figures. The production increase is expected to continue in the near future, as young olive orchards are still expected to increase their production.

The economic return of the activity depends on the amount of VOO recovered from the olives – usually referred to as *industrial yield* –, the quality of that VOO, and the incurred costs during the extraction process. Unfortunately but unsurprisingly, obtaining high industrial yields, high quality VOO and low costs are conflicting objectives. Thus, the different decisions to be made through the process must take into account these inherent trade-offs [13,7].

The evolution of the properties of the olives in the orchards, the harvesting method, and the values of the process variables during the extraction process influence the final value obtained of each of these objectives [1]. Hence, before processing a batch of olives to produce VOO, a decision must be made about what point in the trade-off surface between yield, quality and cost to aim for. This decision defines a production objective that entails a set of values of the process variables to fulfill this pursued objective.

The influence of the process variables in the quality and quantity of the produced VOO and its on-line measurement have been a topics of extensive research, see, for instance [10,12,19,25,30,22,14,29,24], respectively.

The effort devoted to this research is justified in the great importance of the topic, since it is precisely the interaction of the different process variables that will finally determine the output of the process, i.e. the actual quality and quantity of VOO obtained.

Building upon this research, a relevant question is to systematically suggest what would be the *best* production objective for a given batch of olives, assuming that we intend to maximize the economic profit of the activity. In [8,6,5], this problem is considered in the context of a general system for defining and updating suitable set points for the VOO elaboration process, and some considerations about how to approach it are hinted.

However, assuming that the batch of olives to be processed is already in the factory means that the properties of these olives are already fixed. Given that the properties of the olives evolve through the season, and that they are a key factor in the process [17,20], it is very relevant to consider the question of what values of these olive properties would be optimal to have in order to maximize the profit over the whole season. Explicitly considering the economic aspect of the production is an important issue, as recent research effort devoted to these aspect show, for instance [3,23,31] or [11].

The objective of this work is to further elaborate on the preliminary work included in [9] and propose a method to determine an optimal production plan for the whole harvesting season, i.e., define the amount for each VOO quality that maximize the company profit and when it should be produced, given pertinent restrictions and considering the stochastic nature of the problem. The contributions of this work include a reformulation of the problem so that the influence of the specifics of the extraction process are condensed into a single parameter, the consideration of fixed costs related to the amount of time that the factory is open and the treatment of the problem as a stochastic one, explicitly addressing the uncertainty in some parameters and the information available at each step.

* Corresponding author.
E-mail address: pcano@ujaen.es (P. Cano Marchal).

https://doi.org/10.1016/j.jprocont.2018.08.001

The rest of the paper is organized as follows: Section 2 covers the theoretical part of the work, presenting the proposed objective function, constraints and models, as well as the different approaches for obtaining the value of the parameters of these models. Section 3 shows the results obtained using the proposed method for different scenarios and Section 4 presents the conclusions of the work and the future research lines.

2. Optimization problem definition

The following sections discuss in detail the definition of the optimization problem, the alternatives to obtain the value of the different parameters included in the optimization problems and the convenience and implications of considering the stochastic nature of the problem. A brief description of the olive oil production process can be found in [8] and more detailed ones are available in [2,12].

2.1. Optimization problem definition

The total production of VOO carried out through the harvesting season can be thought of as the collection of VOO batches produced at different production periods, where each batch is characterized by the properties of the corresponding olives (ν), the specific values of all the variables involved in the transformation of olives into oil from the orchard to the factory (ξ), and the particular quality of the produced VOO (ϕ).

In the bulk VOO market, the price of the VOO is primarily determined by its regulatory quality, as defined by the EC 2568/91 – *Extra Virgin*, *Virgin* and *Lampante*, the lowest quality of VOO–, with possibly minor variations of the price according to the specific value of some parameters of special interest, but typically small compared to the price gap between the different categories.

A relevant aspect to consider is that two VOO of the same quality may provide different profit depending on the way they are marketed. A clear example is the difference between bottled and bulk VOO commercialization: typically, VOO producers have a higher profit selling bottled VOO, as the price is higher and compensates the associated higher marketing costs. However, restrictions typically exist on the amount of bottled VOO that a company can sell, while that constraint does not apply to the bulk market.

These considerations suggest the convenience of introducing the concept of *product* for the formalization of the optimization problem. A *product* is characterized by a quality requirement and an associated commercialization method. For instance, *Bulk Extra VOO* and *Bottled Extra VOO* are two different products whose quality requirement is the same, but whose commercialization method – and consequently, their selling price – differ. This way, constraints on the total amount of each type of VOO can be easily introduced.

The production plan, thus, requires prescribing the amount of each product to be produced during the whole season. If we let the index $i = 1, 2, \ldots, I$ enumerate the different production periods – typically, weeks or days –, and $k = 1, 2, \ldots K$ refer to the different products, then the production plan can be formalized as a matrix $P \in \mathbb{R}^{I \times K}$:

$$P = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{i1} & \ldots & a_{1K} \\ a_{21} & a_{22} & \ldots & a_{i2} & \ldots & a_{2K} \\ \vdots & \vdots & \ldots & \vdots & \ldots & \vdots \\ a_{i1} & a_{i2} & \ldots & a_{ik} & \ldots & a_{iK} \\ \vdots & \vdots & \ldots & \vdots & \ldots & \vdots \\ a_{I1} & a_{I2} & \ldots & a_{Ik} & \ldots & a_{IK} \end{bmatrix}. \quad (1)$$

Here, each entry a_{ik} of the matrix represent the amount of olives devoted to the production of product k at period i, each column describes the specific production plan through time for a single product, and each row represent the production of each product for the corresponding time period.

The following Sections detail the different aspects considered in the construction of the optimization problem.

2.1.1. Analysis of a production stage

We begin the analysis of the problem considering a given specific production stage i. In this stage, the properties of the olives ν_i can be considered to be fixed, as they are the ones available during that specific time period.

The profit obtained when processing a specific batch of a certain product and selling the obtained VOO is simply the amount of VOO produced times the price at which it was sold, minus the total production costs. Mathematically, this can be expressed as:

$$J_{ik} = s_{ik}\,\pi_k - a_{ik}\,\kappa_{ik}, \quad (2)$$

where s_{ik} denotes the total produced amount of VOO, π_k stands for the selling price of the product, κ_{ik} represents the total cost of the operation per unit of processed olives, and a_{ik} symbolize the amount of olives that were processed in the batch to obtain the product k.

The amount of obtained oil (s_{ik}) can be expressed as the amount of olives processed times the industrial yield of the extraction process (ρ_{ik}). In turn, π_k depends on the quality of the VOO (ϕ_k), and κ_{ik} depends on the way that the olives were processed from the grove to the end of the process, that is, on the value of the different process variables (ξ_{ik}). Taking into account these considerations, Eq. (2) can be rewritten as:

$$J_{ik} = a_{ik}\,(\rho_{ik}\,\pi_k(\phi_k) - \kappa_{ik}(\xi_{ik})). \quad (3)$$

It is relevant to note that ϕ_k is a property of the product and it is fixed throughout the whole season, so it does not have a i index. For simplicity, the sale price π_k is also considered to be constant through the season and to depend only on the specific product. As discussed above, ρ_{ik} is a function of the olive properties (ν_i) and the process variables (ξ_{ik}).

If we explicitly introduce this dependency in Eq. (3) we have:

$$J_{ik} = a_{ik}\,(\rho_{ik}(\xi_{ik}, \nu_i, \phi_k)\,\pi_k(\phi_k) - \kappa_{ik}(\xi_{ik})). \quad (4)$$

This equation states that the profit for a certain batch of olives is a function of the properties of those olives (ν_i), the way that they are transformed into VOO (ξ_{ik}) and the quality of the product (ϕ_k), that is, $J_{ik} = J(a_{ik}, \nu_i, \xi_{ik}, \phi_k)$.

However, ϕ_k is a fixed value, and the well known trade-off between quality and yield means that fixing a specific value of ϕ_k restricts the freedom of choice of ξ_{ik}, thus influencing the value of ρ_{ik}. Furthermore, ν_i is also fixed, as it is the value at the considered time instant. To make explicit that the only real variable of choice is ξ_{ik}, we may use the notation $J_{ik} = J_{ik}(\xi_{ik}|\nu_i, \phi_k)$. For convenience, we may define:

$$\gamma_{ik}(\xi_{ik}\,|\,\nu_i, \phi_k) = \rho_{ik}(\xi_{ik}\,|\,\nu_i, \phi_k)\,\pi_k(\phi_k) - \kappa_{ik}(\xi_{ik}). \quad (5)$$

This quantity γ_{ik} has the straightforward interpretation of being the profit per unit of processed olives for product k, and allows expressing Eq. (4) more compactly as:

$$J_{ik}(a_{ik}, \xi_{ik}\,|\,\nu_i, \phi_k) = a_{ik}\,\gamma_{ik}(\xi_{ik}\,|\,\nu_i, \phi_k). \quad (6)$$

A straightforward inspection of the above equation shows that the obtained profit for a given batch of a given product can be factored into two terms: the amount of olives processed (a_{ik}) and a factor that depends on the type of product (π_k), the properties of the olives (ν_i) and the way they are processed (ξ_{ik}), but does not explicitly depend on the amount of olives processed. Since, given their definitions, these two terms are independent, it is clear that the optimum of Eq. (6) can be obtained independently optimizing its two terms: a_{ik} and γ_{ik}.

The optimization of the second term is the problem of finding the value of ξ_{ik} that maximizes ρ_{ik} taking into account the selling price and the production costs, subject to the properties of the available olives ν_i and the requirement of attaining a VOO quality of ϕ_k. Mathematically:

$$\underset{\xi_{ik}}{\text{maximize}} \quad \gamma_{ik} = \rho_{ik}(\xi_{ik} \mid \nu_i, \phi_k)\pi_k(\phi_k) - \kappa_{ik}(\xi_{ik}).$$

This problem can be approached using the techniques proposed in [8,6,5] and will be studied in more detail in Section 2.2. However, the key remark is that it can be solved independently of the general production planning problem, and its solution γ^*_{ik} can be considered a fixed parameter in the general problem.

Another important remark is that the requirement of attaining a specific value of ϕ_k may render the optimization problem unfeasible. The associated intuitive idea is that if the properties of the available olives are just not good enough, then it is not possible to obtain the required VOO. If this is the case, then we must assure that no production of product k is assigned to this period. This can be achieved imposing an explicit constraint on the general optimization problem $a_{ik} \le 0$, or just setting $\gamma^*_{ik} = 0$.

Once that γ^*_{ik} is available, the value of a_{ik} must be determined. If we consider that we can only produce product k, then the obvious solution is that a_{ik} must be as large as possible. If, on the other hand, we can indeed produce different products, then we must choose which of those provide the higher return, subject to the total amount of olives that we have for the period:

$$\begin{aligned} &\underset{a_{ik}}{\text{maximize}} && \sum_k a_{ik}\gamma^*_{ik} \\ &\text{subjectto} && \sum_k a_{ik} \le \bar{a}_i. \end{aligned} \tag{7}$$

The solution to this problem is clearly to assign the production to those products whose γ^*_{ik} is greater. However, it is not true in general that the optimal solution to the season-wide production problem requires that there is production in every considered production period, so each production stage cannot be optimized independently of each other. Next Section deals with the formulation of the complete production problem.

2.1.2. Basic general production planning problem

According to the discussion of the previous section, we can suppose that γ^*_{ik} are available for each product and time instant, since their value is independent of the final production plan.

This way, the production problem is an extension of (7) and its objective function can be formalized as:

$$\underset{a_{ik}}{\text{maximize}} \quad J = \sum_i \sum_k a_{ik}\gamma^*_{ik} \tag{8}$$

Several constraints must be included to define a well-posed optimization problem. First, the total number of olives for the season is clearly bounded, so a constraint limiting the total production of the year must be included. As the season advances, olives lose moisture, thus decreasing their weight. This effect must be accounted for introducing a factor λ_i inversely proportional to the moisture content of the olives at each time instant:

$$\sum_i \sum_k \lambda_i a_{ik} \le \bar{a}. \tag{9}$$

Two aspects limit the amount of olives to be processed per time stage: the production capacity of the facilities of the company and the availability of olives. Typically, the production capacity can be considered to be fixed for the whole season, however, the availability of olives does indeed change through the season, as different number of growers are willing to harvest the olives. These considerations can be modeled with the constraint:

$$\sum_k a_{ik} \le \bar{a}_i. \tag{10}$$

Limitations on the availability of olives fit for the production of a precise quality can also be easily included, analogously to the discussion of the previous Section. The bound of these constraints can be precomputed taking into account the amount of olives of the different qualities available at each time instant. These constraints are:

$$a_{ik} \le \bar{a}_{ik}. \tag{11}$$

Finally, the already mentioned possible constraints on the selling capacity of the company of different products implies that the production of the product during the whole season should not exceed the selling capacity. The total bound on sold products is naturally thought of in terms of olive oil, not olives; however, a_{ik} represents olives. In order to not introduce further variables into the optimization problem, a_{ik} can be weighted by ρ_{ik}, thus being transformed into oil. This way, the constraint can be modeled as:

$$\sum_i \rho_{ik}\, a_{ik} \le \bar{a}_k. \tag{12}$$

It is worth noting that ρ_{ik} is available whenever γ^*_{ik} is available, as it is required for its computation. Eqs. (8)–(12) configure a simple linear optimization problem, that constitutes the basic foundation of the problem at hand.

2.1.3. Inclusion of fixed costs

Solutions of the above problem tend to assign the production of *lampante* olive oil to the latest possible time periods. This is provoked by the fact that olives reduce their moisture content along the season, and thus increase their oil content on wet basis, giving higher profit by means of lower production costs, as less kg of olives have to be processed to obtain the same amount of VOO.

However, opening the factory and keeping it open makes the company incur in costs. Personnel and electric power supply must be hired, and part of these costs are fixed, as they are independent of whether there is production during a certain time period or not. There is, thus, a term of pseudo-fixed costs that is proportional to the amount of time that the factory is open, but independent of the total production during these time slots.

In order to model this effect, several variables need to be introduced. First, an integer variable δ_i encodes whether there is production in a certain period. The corresponding constraints associated with these variables, employing the well-known big-M formulation [27], are just:

$$\sum_k a_{ik} \le M\delta_i \tag{13}$$

However, the inclusion of these variables and constraints is not enough to consider the number of time periods elapsed since production is started until it is finished. To accomplish this, two new sets of binary variables (ϵ^A_i and ϵ^B_i) are introduced, along with the constraints:

$$\begin{aligned} &\sum_{k,j\ge i} a_{jk} \le M\epsilon^A_i \\ &\sum_{k,j\le i} a_{jk} \le M\epsilon^B_i. \end{aligned} \tag{14}$$

These constraints force ϵ_i^A to be equal to one whenever there is production in or *after* the specified time instant i, while ϵ_i^B is one whenever there is production during or *before* i.

For any given time instant, there are four possibilities:

- Production has not started yet. In this scenario, ϵ^A equals 1, while ϵ^B equals 0.
- There is production assigned to the period. Here, both ϵ^A and ϵ^B equal 1.
- Production has already started and has not yet finished, but there is no production assigned to the period. In this case, again both ϵ^A and ϵ^B equal 1.
- Production has already finished, and, logically, there is no production assigned to the node. In this scenario, ϵ^A equals 0, with ϵ^B being 1.

So, the absolute value of the difference between ϵ^A and ϵ^B is 1 when the factory either has not started production or has already finished, and 0 when either there is production in the time instant or both before and after the considered period. Thus, the amount of time period considered minus the sum of the absolute value of the difference of ϵ^A and ϵ^B provide the number of time instants that the factory is open.

This discussion can be encoded using two new sets of binary variables (ξ^A and ξ^B) associated to the following constraints:

$$\begin{aligned} \xi_i^A - \xi_i^B &= \epsilon_i^A - \epsilon_i^B, \\ \xi_i^A + \xi_i^B &\leq 1. \end{aligned} \tag{15}$$

This way, the total number of open time periods is given by $t = n - \sum_i(\xi_i^A + \xi_i^B)$, and the cost function can be augmented with a term penalizing this variable:

$$\underset{a_{ik}}{\text{maximize}} \qquad J = \sum_i \sum_k a_{ik}\gamma_{ik}^* - C\,t \tag{16}$$

Here, C represents the fixed cost assigned to having the facility open for one time step, while t represents the total number of time periods when the facility is opened. Unfortunately, the inclusion of these constraints suppose that the problem is no longer a simple linear optimization problem; since integer variables have been introduced the problem is now a Mixed-Integer Linear Problem. However, the size of the typical problem and the fact that there are no real hard requirements on the computation time needed to solve it allow considering these constraints without prohibitively increasing the difficulty of solving the problem.

2.2. Obtaining the parameters of the optimization problem

In the previous Sections we have formulated the optimization problem to be solved to obtain the optimal production plan for the whole season under the assumption that the values of the problem parameters are known. However, we have not yet addressed how to assign appropriate values to the different parameters included in the optimization problem.

We may distinguish two types of parameters: those whose value depends on the characteristics of the available olives (ν_i) and those whose value does not. Parameters that depend on ν_i are regarded to as *production-related* parameters, while those that do not are addressed as *business-related* parameters.

Business-related parameters include company-specific aspects such as the total availability of olives for the season and the production capacity per time stage, and market-related parameters, such as the prices of each product and their associated maximum selling capacity. The company-specific values are fairly straightforward to estimate just analyzing the available equipment of the company and historical data regarding the daily reception of olives. In turn, the selection of optimal values for market-related parameters so that the revenue of the company is maximized constitutes an interesting problem in its own right. However, that problem requires considering the marketing and business strategy of the company and falls out of the scope of this work. We suppose that decisions such as the selling price of the products and their associated marketing costs have already been decided and are available for our problem.

On the other hand, production-related parameters are connected to the specifics of the VOO production process itself, and are affected by decisions about the process variables. According to the problem definition presented in Section 2.1, the parameters that explicitly appear in the problem formulation are γ_{ik}^* and $\bar{a}_{ik}$. However, as discussed in that Section, the values of these parameters are the result of all the technological factors that influence the VOOPP, and the whole process must be analyzed to determine them.

Olive properties evidently play a fundamental role in the optimization problem. Since the characteristics of the olives evolve in time, a model of this evolution is required to provide the value of the variable exclusively as a function of the time period considered. The characteristics of the olives relevant to the problem are the ripeness (R_i), the fat content (F_i^D), the humidity (H_i) and the overall state of the olives, i.e., whether they are damaged due to hail, plagues, etc.

Harvesting methods also influence the characteristics that olives exhibit when they arrive to the factory. Harvesting methods can be classified in two major groups:

- Methods that separate olives coming from the tree from olives already in the ground, and
- Methods that mix olives coming from the tree and the ground.

Olives that have fallen to the ground present poor quality characteristics, due to the chemical reactions that begin to take place [15]. Therefore, methods that mix olives cause a decrease of the potential quality that could be obtained if only olives coming from the tree were to be harvested. The ratio of fallen/tree olives is a parameter of importance, as determines the decrease of quality due to the mixture of qualities. The amount of fallen drupes increases as the harvesting season advances, due to the reduction of the retention force of the olives as they ripen. Meteorological phenomena, such as high intensity wind, may increment the amount of fallen olives in stages where they would normally still be on the tree. However, methods that mix olives tend to offer lower costs, since they require lower manual labor [32].

The conditions of the olives that arrive to the factory, as affected by the aforementioned factors, constitute the starting point for the processes carried out in the factory. The fundamental trade-off between oil recovery yield and oil quality require choosing adequate values for the set-points of the main process variables so that optimal revenue can be achieved for a given batch of olives [7].

The paste preparation is the stage of the VOOPP that covers the operations from the reception of the fruit until the so-called *olive paste* is fed into the decanter to separate the oil from the rest of the paste components, and is the key step in the trade-off between yield and quality. Fig. 1 shows a simplified model of this paste preparation stage of the VOOPP proposed in [6].

The output variables of the model are Kneading State, a variable that is related with the total yield that can be achieved, and Fruity and Defect, two parameters that determine the quality of the obtained VOO. These variables are determined by the properties of the olives, namely, Ripeness, Incoming Fruit State and Incoming Olive Moisture, and the way the process is carried out, as defined by the value of the process variables considered in the model, namely Storage Time in Hopper, Coadjuvant Addition, Kneading Time and

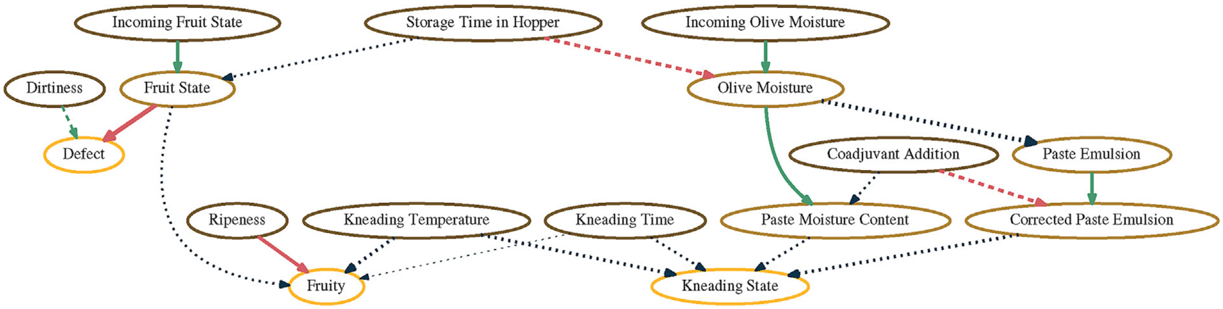

Fig. 1. Simplified model of the paste preparation stage in the virgin olive oil elaboration process.

Kneading Temperature. The variable Dirtiness accounts for whether the factory is in proper condition to not contribute to having a large value of Defect. The rest of the nodes in the model are internal variables that help define the relations between these input and output nodes.

This model was constructed using Fuzzy Cognitive Maps [21,26] and can be used as a constraint for an optimization problem whose objective is to determine the maximum plausible yield, subject to the properties of the batch and obtaining the required quality for product k. Mathematically:

$$\begin{aligned} \underset{\xi}{\text{maximize}} \quad & J = \rho_{ik} \\ \text{subject to} \quad & \mathbf{y} = f(\xi, \nu) \\ & \nu = \nu_i \\ & \phi = \phi_k \\ & \xi_{\min} \leq \xi \leq \xi_{\max} \end{aligned} \tag{17}$$

where $\mathbf{y} = [\rho_{ik}, \phi_k]^T$ and $f(\cdot)$ denotes the fuzzy model that relates the different process variables and olive properties to the outputs.

The model of the process relations introduces several linear, a SOS2 and a nonconvex constraint per node to the problem. The solution of this optimization problem provides ρ_{ik} and ξ_{ik}, which can be used to compute the production costs κ_{ik}, and that, along with π_k, can be used to compute γ_{ik}. Further details on the model, the optimization problem and its solution can be found in [6] and [5].

Finally, the factorized nature of Eq. (6) gives rise to an alternative approach for estimating the production-relation parameters. Optimizing the production of VOO for the season involves two aspects: handling each batch of olives such that the maximum profit for the batch is obtained, and allocating the production in those time periods that provide the highest profits. As discussed before, these two actions are independent of each other; in particular, an optimal allocation of production can be planned even if the production of the batches is not carried out optimally, i.e., the coefficients $\hat{\gamma}_{ik}$ are not the actual optimum values γ^*_{ik}. If this is the case, the solution of the optimization problem having $\hat{\gamma}_{ik}$ as parameters won't provide the truly absolute optimal value for the production, which would be obtained if the actual γ^*_{ik} were used; however, it will provide the optimal solution for the production planning problem *given* that the transformation of olives into VOO is carried out such that the parameters are $\hat{\gamma}_{ik}$. In other words, we can solve the optimal production allocation problem for given current operating practices of a company.

An alternative approach to obtaining the ρ_{ik}, κ_{ik} and $\bar{a}_{ik}$ is to employ historical data obtained in past harvesting seasons. The drawback is that using this data does not guarantee that the coefficients will be optimal, in the sense that it is possible that the operation of the factory were not fully optimized and better yields or VOO quality could have been obtained for the given olive characteristics. However, the advantage of using these data is twofold: on the one hand, the required data is usually quite easily recovered from already available records of the company, thus reducing the amount of work required to set up the optimization problem. On the other hand, solving the problem with historical data means optimizing the season-wide production with the actual current knowledge and technology of the company, which is an interesting problem in its own. Tables 1 and 2 respectively summarize the business-related and production-related parameters considered.

Table 1
Business-related parameters used in the optimization problem. The value of these parameters does not depend on the properties of the olives. Typical values of these parameters can be found in Fig. 2 and Table 3.

Symbol	Parameter
$\bar{a}$	Total availability of olives per season
$\bar{a}_i$	Total production capacity per time stage
$\bar{a}_k$	Total selling capacity of product k
π_k	Selling price of product k
ϕ_k	Required quality for product k

Table 2
Production-related parameters used in the optimization problem. The value of these parameters depends on the properties of the olives.

Symbol	Parameter
$\bar{a}_{ik}$	Bound on product k for stage i due to availability of olives capable of providing quality ϕ_k
γ_{ik}	Profit per kg of processed olives devoted to product k at time step i
ρ_{ik}	Industrial yield for product k at time step i
κ_{ik}	Total production cost for product k at time step i

2.3. Uncertainty on the problem parameters

The planning of the production of VOO for the whole season is subject to several sources of uncertainty that must be taken into account in the definition of the optimization problem. This uncertainty has implications both in the value of the problem parameters and on the structure of the problem itself.

First, the evolution of the properties of the olives clearly constitutes a random variable, whose expected behavior can be inferred from the research literature – see, for instance [17,20] – and historical data in each company, but whose precise realization for each harvesting season is unknown. This uncertainty constitutes the major issue in the optimization problem, since the main idea of the method is to choose when to harvest according to these olive properties.

In order to handle this uncertainty, different olive properties evolution scenarios can be considered, and the solution to each of these can be studied. However, simply considering the different scenarios for the computation of the parameters and solving the problem posed in Section 2.1.2 is not enough, as it is equivalent of considering that the future evolution of the properties is known

Parameter File

```
set I := 1 2 3 4 5 6
7 8 9 10 11 12 13 14;
set K := extra_sup extra
 virgin lampante;

param gamma :=
1 extra_sup 0.938348393544
2 extra_sup 0.950053186212
3 extra_sup 0.980554571724
4 extra_sup 1.02332729528
5 extra_sup 1.042056
1 extra 0.560079558215
2 extra 0.567065892024
3 extra 0.585271499493
4 extra 0.610801599273
5 extra 0.621980352;

param quality_capacity :=
1 extra_sup 1000
2 extra_sup 1000
3 extra_sup 0
4 extra_sup 0
5 extra_sup 0
1 extra 1000
2 extra 1000
3 extra 1000
4 extra 1000
5 extra 1000;
```

```
param sale_capacity:=
extra_sup 0
extra 20000
virgin 20000
lampante 20000
;

param harvest_capacity:=
1 30.0
2 60.0
3 90.0
4 120.0
5 150.0
;

param capacity:=
1000
;

param coef_humidity:=
1 1.0888356008
2 1.0885804146
3 1.0648482088
4 1.0401684932
5 0.9887648554
;
```

Fig. 2. Excerpt from the file that defines the values for the optimization problem parameters for Scenario A-D0.

in the planning instant, which is obviously not true. An alternative is to solve the problem considering the expected value of the uncertain problem parameters, which provides the *expected value solution* of the problem. Finally, to cope with the fact that the value of the properties of the olives for each time instant is discovered as the season advances, a full multi-stage stochastic programming problem, including non-anticipatory constraints can be considered [4]. This way, in each stage the production is planned according to the data that is realistically available: the actual evolution of the properties so far, and the expectation of the future evolution. It is not trivial, in general, to see whether the benefits obtained by considering the full multi-stage problem are worth the additional complexity introduced, compared to the expected value solution. Section 3 presents some results regarding this discussion.

Another important set of parameters that can clearly be regarded as random are VOO prices. Changes in the spread between the different products do indeed alter the profitability of each of the products, and thus, may have an impact in the solution of the problem. Here, it is not straightforward to model when is the price information revealed, and whether there is room for decisions based on this new information. If new information is disclosed as the season evolves, we would be facing a multi-stage optimization problem, similar to the one faced regarding the evolution of the quality properties. If, on the other hand, we consider that the precise value of the prices is known once that the harvesting is completed, then we do not face a multistage stochastic problem, as all production decision are already made when the actual value of the prices are revealed.

3. Results and discussion

This Section presents the solutions obtained using the proposed approach for a set of scenarios. The optimization problems were set up using Pyomo [18] and PySP [33], and involved defining a deterministic base model and several scenario tree models that defined the problem stages and the value of the uncertain parameters in each scenario. The solution to the problems were carried out using PySP and Gurobi [16].

All the scenarios consider a progressive increase in the harvesting capacity ($\bar{a}_i$) during the first few weeks, modeling the gradual incorporation of farmers to the harvesting tasks. Historical evolution data for olive moisture and fat content evolution were provided by an expert laboratory, and the production related parameters were obtained using the models presented in [6]. Fig. 2 includes an excerpt of the parameter file used in the Scenario A-D0 presented in Section 3.3.

3.1. Inclusion of fixed costs and moisture decay

The first analysis is a comparison of the production plans obtained considering a deterministic evolution of the olive properties and prices when fixed costs and moisture evolution are considered in the model. Fig. 3 shows the results obtained. Top-left plot shows the results when neither fixed costs nor moisture evolution is considered. As depicted in the plot, production is assigned based exclusively on the highest values of ρ^*_{ik}, being concentrated at the end of the season, but assigning sporadic production for high

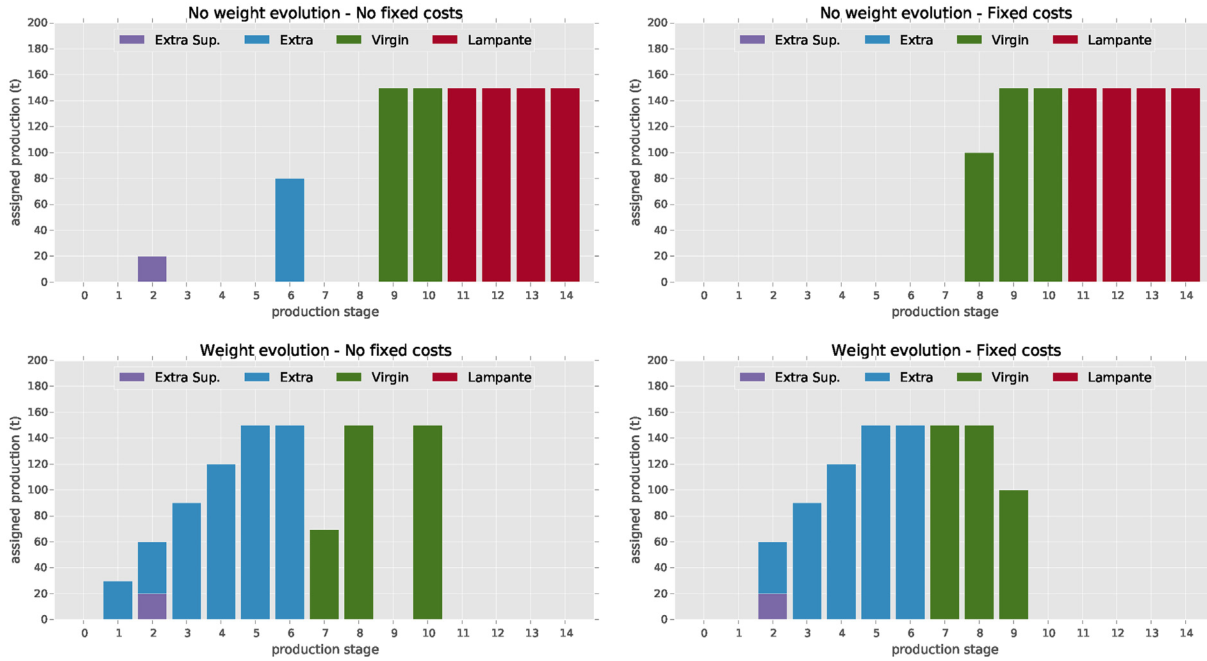

Fig. 3. Production Plan for fixed olive quality properties evolution and prices in four different cases. Plots on the left (right) column don't (do) include fixed cost. Plots on the top (bottom) row don't (do) include moisture evolution consideration. The inclusion of fixed costs congruently minimizes the span of days devoted to production, while moisture evolution provokes production to be transferred from latter season to earlier periods.

Table 3
Prices in €/kg for the different scenarios in 3.2.

Product	Scenario I (2014)	Scenario II (2015)	Scenario III (2016)
Extra Sup.	6	6	6
Extra	2.317	3.496	3.279
Virgin	2.140	3.186	3.133
Lampante	2.036	2.991	2.961

quality VOO. Top-right plot considers the inclusion of fixed costs, and congruently minimizes the span of days devoted to production. With the particular set of prices considered in the scenario (Scenario I in Table 3), production is completely assigned to the end of the season, and no high-quality olive oil is produced. When we take into account the lose in weight of olives in the groves due to their decrease in moisture content (left-bottom plot), we see how production is transferred from the latter season to earlier periods. This reflects the fact that although there is a higher fat content in latter-season olives, there are also less kilograms of olives on a wet basis to be harvested. Finally, the inclusion of fixed costs in the right-bottom plot again reduces the span of production days, compared to the solution that does not include them.

3.2. *Stochastic prices*

This analysis supposes fixed and known evolution of the olive properties, but uncertainty in the prices of the different products considered. The uncertainty is modeled assuming that there is a equal chance of having the 2014, 2015 or 2016 average Spanish bulk market prices, as included in Table 3 [28]. The prices are supposed to be revealed when all production is done, as the average sale price for the season is not known while production is being carried out, as typically producers sell throughout the year and only a very small amount is actually sold while being produced. Since the prices are revealed at the end, there is no decision to be made based on the actual outcome of the prices. Since there is no second-stage decision to be made, the recourse solution of the problem coincides with the expected value solution. If the prices were updated for each time period, then the production could be adjusted based on that information, just like the information about the olive properties is incorporated.

Fig. 4 shows the production plans for each of the production scenarios considering perfect knowledge of the prices, and the expected value solution. As depicted in the Figure, Scenarios I and II, despite noticeable differences in the price levels of the products, provide the same production plan. However, Scenario III, showing similar prices to Scenario II, provides a different solution. This can be explained by the spreads between the prices of the different products. Price spreads are similar in Scenarios I and II, while Scenario III provides narrower spreads, which encourage the production of products with a higher yield. The value of perfect information [4], that is, the difference between the profits obtained when we execute the production plan optimal to the specific realization of the prices and those obtained executing the stochastic solution is zero for Scenarios I and II, as the expected value solution coincides with the optimal production plan, and is 0.15% in Scenario III.

3.3. *Stochastic prices and olive quality evolution*

In this Section a stochastic evolution of the olive properties is considered in addition to the stochastic prices. Six different scenarios are considered: two price levels (corresponding to years 2015 and 2016, assigned equal probability) and three olive property evolution Scenarios: regular evolution as considered in the previous Sections, a case when olives are damaged in time step 1, and another in time step 3, with probabilities given by a parameter p, according to Table 4. These scenarios have been chosen so that we consider a regular season with no abnormal quality evolution, a situation when there is a incident very early in the season and another when a later incident may still have a significant impact in the production.

In this case, the information about the quality of the olives is revealed as the season advances, so decisions based on the actual outcome of the random variables so far can be made. Fig. 5 shows the different production plans provided for this Section. The

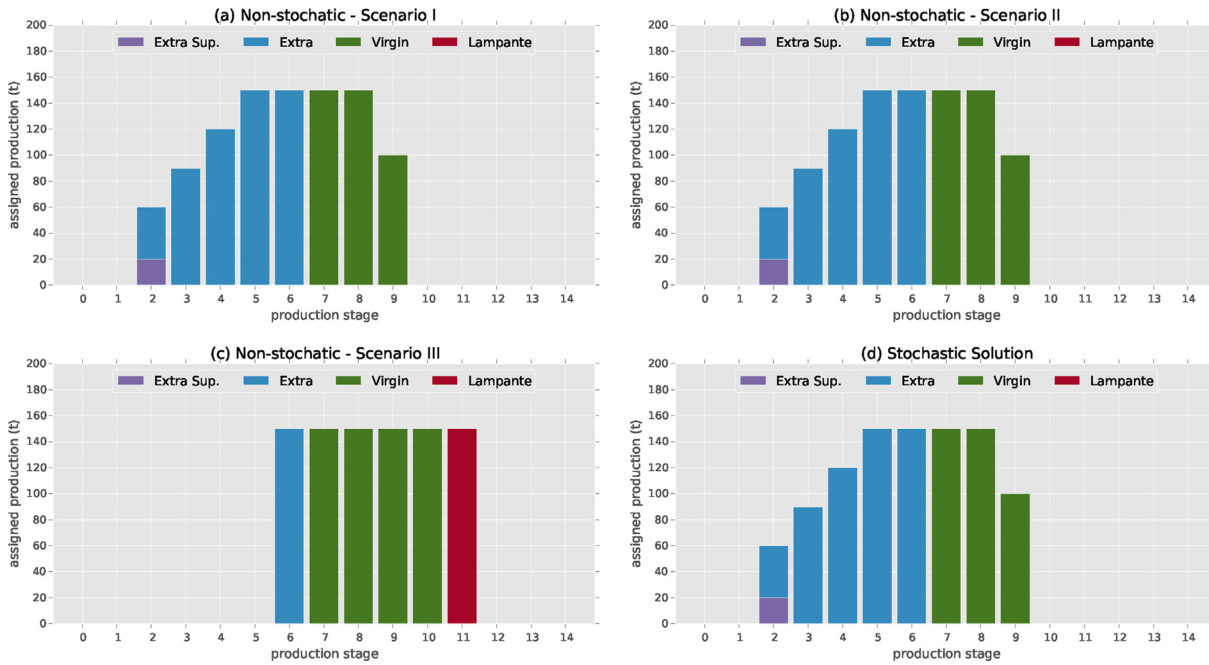

Fig. 4. Production plans for the scenarios considered in Section 3.2. Figures (a), (b) and (c) show the optimal production plan, assuming perfect information, for Scenarios I, II and III, respectively. Figure (d) shows the expected value solution when the stochastic nature of the prices is considered.

Table 4
Scenarios considered in Section 3.3.

	No Damage	Damage In Time Step 1	Damage In Time Step 3
2015 Prices	A-D0	A-D1	A-D3
2016 Prices	B-D0	B-D1	B-D3
Probability	$(1-p)^2$	p	$p(1-p)$

Table 5
Value of Perfect Information (VPI) and Value of the Stochastic Solution (VSS) for the different values of p considered in Section 3.3.

	$p=0.1$	$p=0.3$	$p=0.5$	$p=0.7$
VPI%	0.32	0.27	0.12	0.04
VSS%	0.34	0.32	0.31	0.20

columns correspond to the three different quality evolution scenarios. The first and second rows show the optimal production plans supposing perfect information for price Scenario A and B respectively. As depicted in the figure, the plans coincide for scenarios D1 and D3, differing in case that no damage to the olive occurs. The third row represents the recourse problem solution. Rows 4 to 6 represent the expected value solution to the problem for $p=0.1$, $p=0.3$ and $p=0.7$, respectively.

It is worth noting that the expected value solution does not take into account information being revealed through the season, so it provides the same production plan for the three quality scenarios. As depicted in the plot, these proposed production plans suggest the same time instants for production for $p\geq 0.3$, with the products suggested changing due to the decrease of the average quality evolution. Strictly speaking, the product plans provided by the expected value solution for D1 and D2 are unfeasible, since the actual quality properties found do not allow to produce the specified products.

On the other hand, the solution to the recourse problem does not depend on p. It provides different production plans for the different Scenarios, as it is capable of incorporating the information that is being revealed as the season advances, and these are always feasible. Furthermore, the proposed plans can always be implemented, as the non-anticipatory constraints make sure that the plans are consistent.

Table 5 includes the value of perfect information (VPI) and the value of the stochastic solution (VSS, i.e. the difference between the profits obtained when executing the full stochastic solution and the expected value one) for the different values of p considered. As included in the Table, the VSS is modest, but on the same order of magnitude of the VPI for all values of p. Although modest percentages, a 0.34% increase means around 27,000 € for a factory processing 10 million kg of olives per season. The typical production plan carried out by VOO producers the authors have relation with is actually quite similar to the one proposed in the expected value solution with $p=0.1$, so the VSS also constitutes a good estimation of the potential gains of using stochastic production planning versus the manual planning.

Finally, it is worth noting that the computational cost of solving the optimization problems, carried out in a mid 2012 MacBook Pro with a Pentium i7 processor with 8 Gb of RAM, proved to be fairly modest. As an example of the order of magnitude, the solution time for the stochastic problem with $p=0.3$ was 5.9 s, with the expected value solution requiring less than 2 s.

4. Conclusion

In this paper the problem of obtaining an optimal production plan for the VOO elaboration has been regarded. The different factors that must be considered and the relations between variables have been pointed out, resulting in the formulation of an optimization problem including convenience of the inclusion of fixed costs and the stochastic nature of the problem. In particular, the differences between considering perfect knowledge, the expected value solution and the full stochastical setup have been discussed. Then, the practical consideration of obtaining the parameters for the problem has been addressed from two points of view. Finally, based on data provided by bibliography and experts, the optimization problem proposed has been solved for a variety of different scenarios and the results have been discussed.

The proposed approach serves as a fundamental tool for the construction of a comprehensive decision support system for vir-

Fig. 5. Production plans for the scenarios considered in Section 3.3. Columns correspond to the three different quality evolution scenarios. First and second row show the optimal production plans supposing perfect information for price Scenario A and B, respectively. Third row represents the recourse problem solution. Rows 4 to 6 represent the expected value solution to the problem for $p=0.1$, $p=0.3$ and $p=0.7$, respectively.

gin olive oil production, that, together with the system proposed in [6,5], may help operators in each decision to be made in the production process. Since the decision of when to harvest is a key decision in the process, the proposed approach enriches and completes systems focusing on the process once olives are already at the factory.

Further work will include the consideration of different olive cultivar varieties might be of interest, along with a more sophisticated modeling of the stochastic distribution of prices and demand. Other aspects, such as the stochastic nature of the weather conditions and their restrictions on the harvesting capacity, leading to both changes in the properties of the olives and extending the amount of time required to carry out the production, can also be of interest to be included in the model.

Acknowledgments

This work was partially supported by the projects DPI-2011-27284, PI10-AGR-6616 and DPI-2016-78290-R. P. Cano Marchal has been in receipt of a F.P.U. grant from the Spanish Ministry of Education. The authors are grateful to CM Europa for the historical data of olive properties evolution provided.

References

[1] M.P. Aguilera, G. Beltran, S. Sanchez-Villasclaras, M. Uceda, A. Jimenez, Kneading olive paste from unripe 'Picual' fruits: I. Effect on oil process yield, J. Food Eng. 97 (4) (2010) 533–538.

[2] J. Alba, Elaboración de aceite de oliva virgen, in: D. Barranco, R. Fernández-Escobar, L. Rallo (Eds.), El Cultivo Del Olivo, 2nd ed., Mundi-Prensa, Madrid, 1997.

[3] P. Báez-González, A.J. del Real, M.A. Ridao Carlini, C. Bordons, Day-ahead economic optimization of energy use in an olive mill, Control Eng. Pract. 54 (2016) 91–103.

[4] J.R. Birge, F. Louveaux, Introduction to Stochastic Programming, 2nd ed., Springer, New York, 2011.

[5] P. Cano Marchal, J.G. García, J.G. Ortega, Application of fuzzy cognitive maps and run-to-run control to a decision support system for global set-point determination, IEEE Trans. Syst. Man Cybern.: Syst. PP (99) (2017) 1–12.

[6] P. Cano Marchal, J. GámezGarcía, J. GómezOrtega, Decision support system based on fuzzy cognitive maps and run-to-run control for global set-point determination, 2015 IEEE International Conference on Systems, Man, and Cybernetics (2015) 1745–1751.

[7] P. Cano Marchal, J. GómezOrtega, D. AguileraPuerto, J. GámezGarcía, Current situation and future perspectives on virgin olive oil elaboration process control, Rev. Iberoam. Autom. Inform. Ind. RIAI 8 (3) (2011) 258–269.

[8] P. Cano Marchal, D. Martínez Gila, J. Gámez García, J. Gómez Ortega, Fuzzy decision support system for the determination of the set points of relevant variables in the virgin olive oil elaboration process, 2013 IEEE International Conference on Systems, Man, and Cybernetics (2013) 3489–3494.

[9] P. Cano Marchal, D. Martínez Gila, J. Gámez García, J. Gómez Ortega, Optimal production planning for the virgin olive oil elaboration process, IFAC World Congress, vol. 19 (2014) 8921–8926.

[10] M.L. Clodoveo, Malaxation: influence on virgin olive oil quality. Past, present and future – an overview, Trends Food Sci. Technol. 25 (1) (2012) 13–23.

[11] R. De-Luca, F. Bezzo, Q. Béchet, O. Bernard, Exploiting meteorological forecasts for the optimal operation of algal ponds, J. Process Control 55 (2017) 55–65.

[12] L. Di Giovacchino, N. Costantini, M.L. Ferrante, A. Serraiocco, Influence of malaxation time of olive paste on oil extraction yields and chemical and organoleptic characteristics of virgin olive oil obtained by a centrifugal decanter at water saving, Grasas Aceites 53 (2) (2002) 179–186.

[13] L. Di Giovacchino, S. Sestili, D. Di Vicenzo, Influence of olive processing on virgin olive oil quality, Eur. J. Lipid Sci. Technol. 104 (2002) 587–601.

[14] R. Furferi, M. Carfagni, M. Daou, Artificial neural network software for real-time estimation of olive oil qualitative parameters during continuous extraction, Comput. Electron. Agric. 55 (2) (2007) 115–131.

[15] J.M. García, K. Yousfi, The postharvest of mill olives, Grasas Aceites 57 (1) (2007).

[16] I. Gurobi Optimization, Gurobi Optimizer Reference Manual, 2015.

[17] F. Gutiérrez, B. Jiménez, A. Ruiz, M.A. Albi, Effect of olive ripeness on the oxidative stability of virgin olive oil extracted from the varieties picual and hojiblanca and on the different components involved, J. Agric. Food Chem. 47 (1) (1999) 121–127.

[18] W.E. Hart, C. Laird, J.-P. Watson, D.L. Woodruff, Pyomo – Optimization Modeling in Python Springer Optimization and Its Applications, vol. 67, Springer US, Boston, MA, 2012.

[19] A.M. Inarejos-García, A. Gómez-Rico, M.D. Salvador, G. Fregapane, Influence of malaxation conditions on virgin olive oil yield, overall quality and composition, Eur. Food Res. Technol. 228 (4) (2009) 671–677.

[20] B. Jiménez Herrera, A. Rivas Velasco, A. Sánchez-Ortíz, M.L. Lorenzo Tovar, M. Ubeda Munoz, R.M. Callejon, E. Ortega Bernaldo de Ouiros, Influence of fruit maturation process on the sensory quality of virgin olive oils from picual, hojiblanca and picudo cultivars, Grasas Aceites 63 (4) (2012) 403–410, WOS:000311708500007.

[21] B. Kosko, Fuzzy cognitive maps, Int. J. Man-Mach. Stud. 24 (1) (1986) 65–75.

[22] A. Leone, A. Tamborrino, R. Zagaria, E. Sabella, R. Romaniello, Plant innovation in the olive oil extraction process: a comparison of efficiency and energy consumption between microwave treatment and traditional malaxation of olive pastes, J. Food Eng. 146 (2015) 44–52.

[23] X. Liu, J. Cui, Economic model predictive control of boiler-turbine system, J. Process Control 66 (2018) 59–67.

[24] D. Martínez Gila, P. Cano Marchal, J. Gámez García, J. Gómez Ortega, On-line system based on hyperspectral information to estimate acidity, moisture and peroxides in olive oil samples, Comput. Electron. Agric. 116 (2015) 1–7 http://www.sciencedirect.com/science/article/pii/S0168169915001593.

[25] M. Moya, F. Espínola, D.G. Fernández, A. de Torres, J. Marcos, J. Vilar, J. Josue, T. Sánchez, E. Castro, Industrial trials on coadjuvants for olive oil extraction, J. Food Eng. 97 (1) (2010) 57–63.

[26] E. Papageorgiou, J. Salmeron, A review of fuzzy cognitive maps research during the last decade, IEEE Trans. Fuzzy Syst. 21 (1) (2013) 66–79.

[27] Y. Pochet, L.A. Wolsey, Production Planning by Mixed Integer Programming, in: Springer Series in Operations Research and Financial Engineering, Springer, New York, 2006.

[28] PoolRed, POOLred-Sistema de Información de Precios en Origen del Mercado de Contado del Aceite de Oliva, 2016.

[29] L. Salguero-Chaparro, B. Palagos, F. Peña-Rodríguez, J. Roger, Calibration transfer of intact olive NIR spectra between a pre-dispersive instrument and a portable spectrometer, Comput. Electron. Agric. 96 (2013) 202–208 http://www.sciencedirect.com/science/article/pii/S0168169913001269.

[30] A. Tamborrino, S. Pati, R. Romaniello, M. Quinto, R. Zagaria, A. Leone, Design and implementation of an automatically controlled malaxer pilot plant equipped with an in-line oxygen injection system into the olive paste, J. Food Eng. 141 (2014) 1–12.

[31] M. Vaccari, G.M. Mancuso, J. Riccardi, M. Cantù, G. Pannocchia, A sequential linear programming algorithm for economic optimization of hybrid renewable energy systems, J. Process Control (2017).

[32] J. Vilar Hernandez, M.d.M. Velasco Gamez, R. Puentes Poyatos, Influence of the olive exploitation system on the net profit of the producer. Strategies for extensive olive fields in the context of the absence of subsidies, Grasas Aceites 61 (4) (2010) 430–440, WOS:000282085300013.

[33] J.-P. Watson, D.L. Woodruff, W.E. Hart, PySP: modeling and solving stochastic programs in Python, Math. Progr. Comput. 4 (2) (2012) 109–149.