APPLICATION

Stochastic model for setpoint of a rolling mill: an application in the soybean oil production

Magna Paulina de Souza Ferreira[1] · Márcio da Silva Arantes[2] · Jesimar da Silva Arantes[3] · Renan Bonnard[2] · Claudio Fabiano Motta Toledo[1]

Abstract

The present paper proposes a stochastic model for soybean oil extraction, introducing an automated and reliable method for flake thickness control. In this production process, to have a robust solution, we must consider the uncertainty arising, e.g., from inaccurate sensor readings and unforeseen changes in production. The main objective is to maximize the oil extraction by keeping the thicknesses of the soybean flakes within an operating range. Data collection performed in the industrial plant made it possible to analyze the input data with their respective outputs. A methodology based on multiple linear regression was used to describe the relationship between the variables. We propose a mathematical stochastic model to obtain pressure setpoints that produce flakes as ideal as possible for oil extraction. The results reported time and quality improvements within the production process.

Keywords Mathematical modeling · Two-stage stochastic program · Soybean oil extraction · Industry 4.0

1 Introduction

Currently, the growing complexity and the creation of new requirements for efficiency in the industrial environment have changed the way industries operate. In an attempt to adapt to new and improved ways of operating, technological resources have become a strong ally of industrial sectors and give birth to so-called fourth industrial revolution, or industry 4.0. In this context, the fusion of the physical and digital worlds impacts on relevant aspects for the industry such as productivity, cost reduction and process control. The main point is how the information can be used by plant facilities to lead a profound transformation in the whole production chain [1].

Like other activities in the industrial sector, the production of soy is among the economic activities with expressive growth in the last decades. The increase in soybean production can be attributed to several factors with highlights for the oleaginous used for food. The production of soybean oil grows at a significant rate of 3.82% per year. China, the USA, Brazil and Argentina are the countries that produce the most soy oil, being together responsible for 79.01% of the total production [2].

In this paper, we will be dealing with a specific part of the production process of soybean oil within a large oil extraction industry. The main objective is to maximize the oil extraction by keeping the thicknesses of the soybean flakes within an ideal operating range. A rolling mill machine controls the thickness through hydraulic pressures, which have to adjust to guarantee the flakes with the expected thickness. These pressures are adjusted manually by plant operators based on measurements taken using micrometers. There are inaccuracies in the micrometers and manual adjustments;

✉ Magna Paulina de Souza Ferreira
   magna.maximiza@gmail.com

   Márcio da Silva Arantes
   marcio.arantes@sc.senai.br

   Jesimar da Silva Arantes
   jesimar.arantes@ufla.br

   Renan Bonnard
   renan.bonnard@sc.senai.br

   Claudio Fabiano Motta Toledo
   claudio@icmc.usp.br

[1] USP, Institute of Sciences Mathematical and Computer, 400 Trabalhador São Carlense Ave, São Carlos, SP, Brazil

[2] SENAI, Institute of Embedded Systems, 574 Av. Luiz Boiteux Piazza, Florianópolis, SC, Brazil

[3] UFLA, University of Lavras, 855 Av. Antônio Carlos Pinheiro de Alcântara, São Sebastião do Paraíso, MG, Brazil

hence, the samples measured become not reliable compromising the overall adjustments in the rolling mill. However, the primary source of uncertainty in the soybean flaking process is moisture, directly affecting its internal resistance during flaking. These aspects require constant pressure changes in the machine to maintain the thicknesses in the optimal ranges. Considering all these aspects, applying deterministic models to solve the problem does not guarantee a solution that meets the industry's needs.

Stochastic models are widely used in the literature for various industrial problems [3–7]. Thus, the main contribution of this study is a stochastic model for soybean oil extraction, from which we propose an automated and more reliable method for flake thickness control, employing regression analysis and stochastic programming. To the best of our knowledge, there are no literature works dealing with uncertainty in the soybean oil extraction process [8–11]. We report an actual case study with the full development of the decision-making method that automatically controls the thickness of the flake. Several experiments evaluate the method results when compared against the industry's current manual approach.

The paper is organized as follows: Section 2 describes some related works to the research under study. Section 3 contains the complete problem description. Section 4 details the method proposed, while Sect. 5 reports the computational results. Our conclusions follow in Sect. 6.

2 Related work

The present work will propose a method to improve the efficiency of the rolling mill, which is the machine employed in the mechanical processing of flaking the crashed beans. The rolling mill modifies the cross section of crashed beans through the passage between two cylinders, where the distance between the two cylinders is smaller than the initial thickness of the crashed beans.

This technique is widely used in industries, e.g., the authors in [9] conducted experiments to observe the effects of operating conditions on the performance of rolling mills in a steel plant. The objective was to find configurations of the rolling mills that shape the ingots to the desired dimensions. The authors in [12–15] address also related techniques to improve the performance of rolling mills in metallurgical industries. It is proposed a study by [16] on the rolling of plastic sheet in which plastic sheets need to be rolled with different thicknesses. All the mentioned works show how to obtain adjustments to reach some specific thickness through mathematical modeling. However, the models for metallurgic or plastic industries cannot be straight applied to the soybean oil extraction scenario. We must deal with specific constraints not handled by these previous works.

In the soybean context, the authors in [8] propose a study to establish the differences between high pressure and extrusion when determining how high-pressure processing modifies the structures of soy cells. The effects of moisture on soybean oil quality were reported in [10], in which experiments were conducted at a factory in order to establish the optimal moisture content to maximize oil recovery at extraction. These studies give us useful insights to defining our mathematical approach, since aspects such the moisture of soybean will impact over the process of oil extraction.

After the soybeans be crashed, the flake thicknesses produced are another relevant aspect to be considered in the oil extraction. The industries need to optimize the flake thicknesses, and the authors in [17] report the impacts of the physical properties on the thickness of the flakes. The results showed that the increase in the flow rate thrown the machine leads to increase in the thickness of the flakes. They report that the change of pressure parameters in rolling mill only has limitations to adjust the flakes in the ideal operating ranges. The authors conclude that modifications of the metallurgical properties of the rolls and to deal with viscosity properties of the soybeans are possible means for optimizing flake thickness.

A dynamic model of the rolling mill, based on flatness and integrated thickness control, is reported in [18]. The rolling process is simulated, and the dynamic characteristics of thickness and flatness are obtained. The rolling mill rolls wear out over time, and the thicknesses of the materials will not remain in the ideal operating standard. In this sense, to obtain high-quality products, the authors in [11] propose precise methods to predict the deformation of the roll. The aim is to achieve fit parameters such as the bending force of the roll and the displacement of the roll.

There are works focusing in new trends or more easy and precise technology options for the industry as reported in [19]. In this context, mathematical modeling and its optimization in industrial problems have become increasingly effective and able to set parameters of machines. The models usually minimize costs while keeping quality constraints related to the final products [20–22] . Considering these aspects, it should be taken into account that the production process suffers variations from several sources of uncertainties in real environments. The uncertainties are caused by machine malfunction, mechanical wear on the machines, sensor error, sensor inaccuracy, lack of data, among others.

The treatment of uncertainty in models has become increasingly relevant. The authors in [3, 7] approach stochastic modeling techniques to treat the production process in industries considering real contexts. The objective of both works is to represent the reality of the plant considering possible unforeseen events that may actually happen in the real world. The results showed that the approaches were able to obtain solutions compatible with the reality

Fig. 1 Soybean oil extraction process

Raw Soybeans   Cracked Soybeans   Soybean Flakes   Soybean Oil
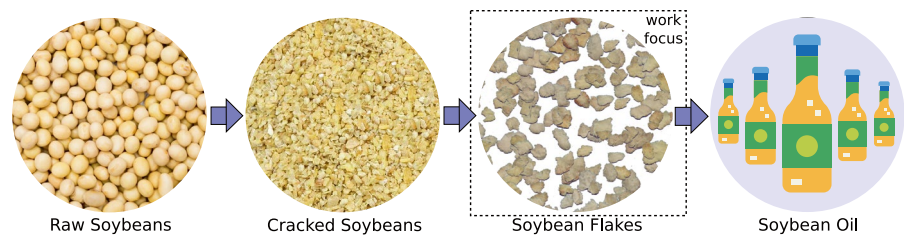
of the production process under study. The works in [4–6] address uncertainties in the production process with stochastic models outperforming deterministic ones. The optimal value of the optimization model was properly approximated when using an enough number of scenarios in the stochastic approach. However, to the best of our knowledge, it was not found in the literature studies proposing treatments of uncertainties within the production process of soybean oil extraction.

3 Problem description

Figure 1 illustrates the transformation flow of the soybean. Initially, we have raw soybeans that go through the creaking mill equipment with two pairs of rollers, reducing the size of the material and facilitating to remove the hull in next steps. In the rolling mill equipment, the creaked beans become small flakes that are easier for the solvent to break oleaginous cells. Finally, the oil is extracted and bottled. The present work focuses on the soybean flaking process, where a better thicknesses must be obtained to maximize oil extraction.

The soybean moisture impacts the thickness of flakes produced by the rolling mill. Thin flakes are obtained from wet beans, while thicker flakes come from dry ones. Thin flakes break easily becoming difficult the circulation of solvent during the extraction, and it can clog the filters causing stoppages in the production process. The thick flakes hinder the solvent's action, decreasing the extraction efficiency. The

soybean temperature can also influence the thickness of the flakes, but the temperature is controlled in the process studied, where the temperature remains as constant as possible. Thus, we did not find a relevant correlations between the temperature and the other variables considered on this study.

Figure 2a illustrates a typical rolling mill equipment used in the soybean flaking process.

The rolling mill has an entrance that controls the flow of the crashed soybeans that will go through the flaking process. If the flow is reduced, the rolling mill will produce thinner flakes; on the other hand, higher flow leads to thicker flakes. There are two smooth cylinders with opposite rotations as shown in Fig. 2b. The distance between these rolls is relevant to produce flakes with adequate thickness. The distance (d) between the rolls results from the hydraulic pressure force and the resistance of the material being rolled. The roll adjustment is made by hydraulic pressure where if the flake is thick, the pressure is increased. In this case, the rollers come closer to generate thinner flakes. Otherwise, the pressure is decreased for thicker flakes be produced.

The moisture control of the soybeans is essential because soybeans with moisture below 9.5% lead to beans with a reduced size that impairs oil extraction. On the other hand, moisture above 11% leads to a more difficult breakage since many grains are crushed, making it difficult to separate the husk. The ideal standard would be a soybean between 9.5 and 11% moisture, both for breaking and flaking the soybean. Moisture is measured using a near-infrared (NIR) sensor, subject to uncertainties arising from measurement failures. These issues prevent effective decision-making in

Fig. 2 Equipment and information about the rolling mill rolls [23]
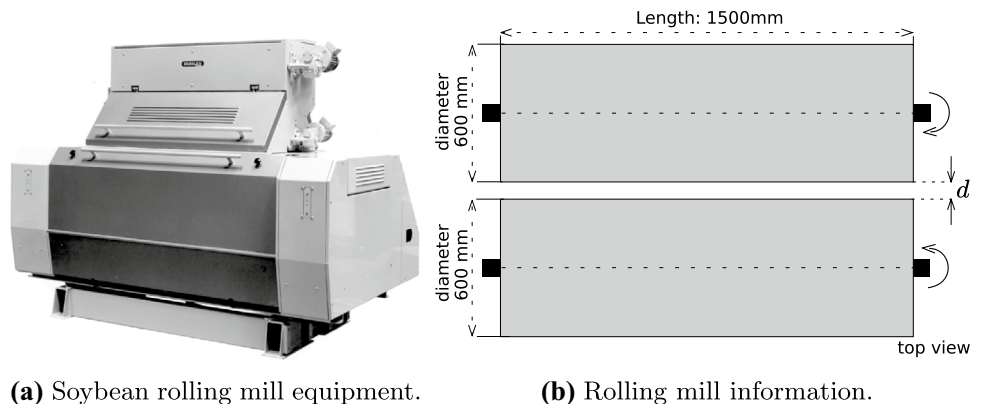
(a) Soybean rolling mill equipment.   **(b)** Rolling mill information.

practice. Details regarding the stochasticity of the problem are reported in Sect. 4.2.

Two pneumatic actuators execute adjustments based on a setpoint that guarantees continuous pressure through closed-loop control. The proposal aims to find the optimal setpoints for the actuators that guarantee suitable flake thickness. In this case, there is no closed-loop control for the sheet thicknesses without sensors capable of obtaining these measurements. In these industries, manual pressure adjustments are standard. Therefore, a structural adaptation in the equipment was necessary for pressure adjustments in an automated way. A rectifier operation of the rolls is performed monthly since there is a wear of the rolls over time.

The authors in [24, 25] suggest as ideal to work with thickness of flakes from 0.2 to 0.4mm, and pressure between 50 and 70 bar in the rolling mill. The authors in [26] state that the ideal thickness of flakes is between 0.25 and 0.35mm. Figure 3 represents the advantages and disadvantages of adopting certain thicknesses ranges of flakes, based on some data collected in the industry. There are two functions where the first describes the impact of the thickness increase on the amount of oil that is lost during the extraction process. The second function describes the problems of filter clogging in relation to the thickness of the flakes. The aim is to combine these two factors in order to obtain efficiency in oil extraction. Therefore, based on this analysis, the ideal thickness of flakes adopted in this work ranges from 0.35mm to 0.40mm.

To sum up, we have a rolling mill machine that receives crashed soybeans, which the level of moisture can vary during the processing to create flakes, and the flow of soybeans as well as the pressure over the cylinders impacts over the thickness of the flakes. The production of flakes is critical with thickness between 0.35mm and 0.40mm, since it will become more efficient the oil extraction. The problem is to reach such efficiency in terms of thickness by executing an automatic adjust of the pressure. The problem itself is physically a complex interaction between the pressure forces on the rolling mill and the internal resistance of the soybean.

However, the ranges practiced in the industry for an adequate flake thickness make the problem restricted, allowing multiple linear approximations to model the problem with minor errors. The proposed method to solve such problem is described next.

4 Methods

The method follows two main steps:

1. Apply multiple linear regression analysis to establish relationships between problem variables.
2. State a stochastic mathematical model to deal with the uncertainty related to the problem variables and data.

The multiple linear regression analysis is defined in the learning phase of the method, where some functions of interest are stated based on data provided by the industry. These functions are handled by the mathematical model to make decisions. Figure 4 illustrates the proposed method when working online. The sensing on the rolling mill will provide the input data that feed the decision stochastic model at each instant t. Next, the stochastic model is solved, and the pressure setpoint will be sent as output to the rolling mill at instant t + 1. At this point, the industry demands that a decision takes less than one minute to be made.

4.1 Multiple linear regression analysis

The main decision variable in the problem is the pressure to be applied by the rolling mill equipment. Hence, data were collected in the industry by changing pressure in the rolling mill and evaluating the impact over the parameter: thickness, flow rate and moisture. This analysis determines the flow rate by opening a valve that controls the production flow following a scale from 4 (minimum flow of raw material in processing) to 7 (maximum flow).

Fig. 3 Influence of flake thickness on the extraction process (approximate behavior) (For confidentiality reasons the original values were omitted, so the graph is merely illustrative showing only the approximate behavior)
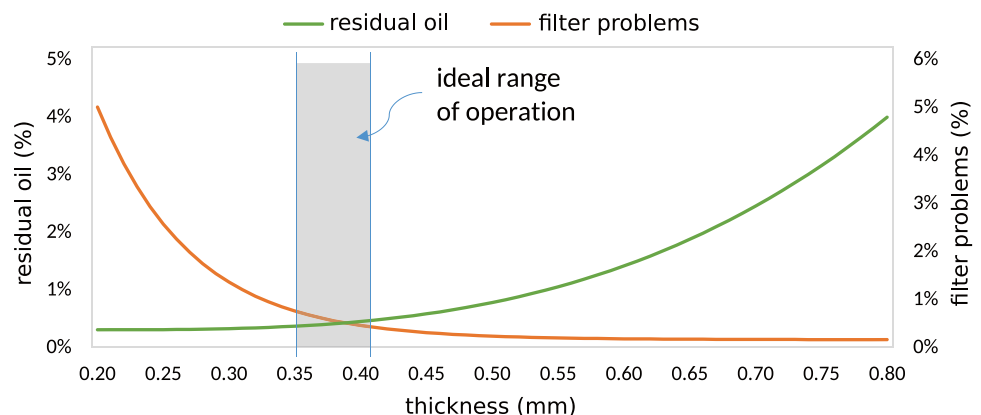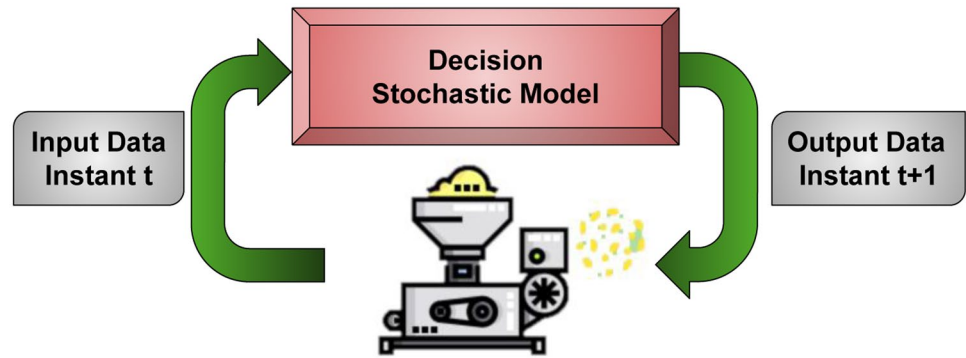
Fig. 4 Method working online

First, a set of 29 experiments was performed in the rolling mill whose results are in Fig. 5.

Next, we decided to apply multiple regression analysis to find electrical current and thickness values based on the pressure and some problem parameters. Multiple regression analysis aims to establish relationships between several variables within a process and their respective results [27]. Thus, the analysis states the relation between dependent variables and independent ones.

Equation (1) represents a multiple linear regression model with n independent variables [28].

$$Y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_{3i} + \cdots + \beta_n x_{ni} + \varepsilon_i \tag{1}$$

- Y_i: response variable (dependent variable);
- x_{ji}: independent variables, where $j = \{0, 1, \cdots n\}$;
- β_0: independent term of regression;
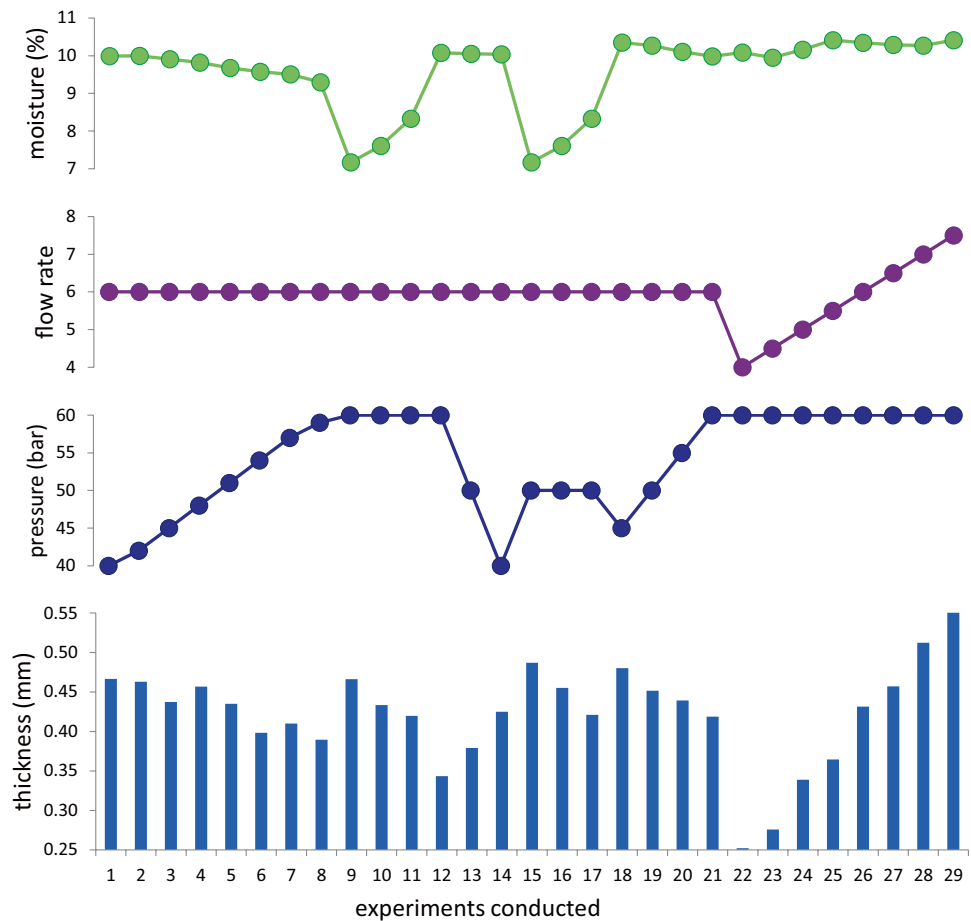- β_j: regression coefficients, in which $j = \{1, \cdots n\}$;

Fig. 5 Experiments performed and flake thickness resulted

Table 1 Estimated parameters through multiple linear regression

	Coefficients	Estimation
Current	β_0	24.6447
	β_1	1.1170
	β_2	−1.4199
	β_3	7.2613

Table 2 Estimated parameters through multiple linear regression

	Coefficients	Estimation
Thickness	β_0	K_t
	β_1	−0.0037
	β_2	−0.0350
	β_3	0.1188

- ε_i: experimental error.

Through the physical characteristics of the rolling mill, the pressure applied to the rollers, the material flow rate and the moisture influence the motor current. In turn, the flake thickness is affected by the motor current that highlights the work done by the rollers, the flow rate and the moisture content. Ideally, the plant works with moistures between 9.5% and 11.5%; high moisture reduces the material's resistance, which leads us to generate thinner flakes, while low moisture produces thicker flakes. The practiced intervals for the flow are between the minimum and maximum values for the flow rate of raw material in processing. Setting the flow rate is a tactical decision to reach a productivity target. For example, increasing the flow rate implies thicker flakes because the work done by the rollers is distributed to a more extensive set of materials.

Thus, the application of multiple regression analysis, over the data collected, allows us to state a relation between the electrical current in the rolling mill engine and pressure (decision variable), and moisture and flow rate (both problem parameters). Equation (2) was estimated to describe the variation of the current based on those inputs. Table 1 shows the estimation of the coefficients obtained by the regression method.

Thus, we have an estimated function for current based on pressure, moisture and flow rate given by:

$$C(P, M, F) = 1.1170 \times P - 1.4199 \times M + 7.2613 \times F + 24.6447 \tag{2}$$

- C: motor electric current measured in amperes (A);
- P: pressure applied to the equipment measured in bar;
- M: soybean moisture measured in percentage (%);
- F: flow rate of processed material measured in cm/s.

Figure 6 compares the real electric current data from industry and the data obtained through the estimated function. The estimated function has an expected behavior with an average error of 2.33%.

Similarly, it was defined a function to estimate the average thickness of flakes having as input: current, moisture and flow rate. Table 2 shows the coefficients obtained through multiple linear regression.

The function is given by Eq. (3):

$$T(C, M, F) = -0.0037 \times C - 0.0350 \times M + 0.1188 \times F + K_t \tag{3}$$

- T: thickness of the flake measured in millimeters (mm);
- K_t: calibration constant for roller wear.

The constant K_t must be calibrated, taking into account the roll's degradation over time once its performance directly influences the outputs. The authors in [11] proposed calculations to predict such wear.

Figure 7 compares the real data and the data obtained through the estimated function (3). The estimated function also has the expected behavior, with an average error of 3.95 %.

Fig. 6 Relation between real and estimated electric current

Thus, the data analysis revealed as factors that contribute most to the final consumed current: pressure (48%), flow rate (43%) and moisture (9%). The factors that most contribute in the thickness are: flow (58%), current (25%) and moisture (17%). Based on this estimation, it is possible to build a stochastic model to make decision on the problem.

4.2 Stochastic model for optimum rolling mill control

We propose a mathematical decision model (decision) to optimize pressure adjustments, where the main constraint is to keep the thickness of the flake within the ideal operating range. To meet these requirements, we first started with the prediction model (prediction) described in the previous section, which estimates the flakes' thickness based on the input parameters. Figure 8 summarizes the relation between these two approaches.

The application of deterministic mathematical models for problems in real production planning is widely used in the literature [29–32]. These models assume exact knowledge of the system and perfect measurements. However, there are many sources of uncertainty in the real production scenario, and that is the case for the problem at hand here. Therefore, the proposed study innovates by applying stochastic models to soybean flaking to maximize oil extraction.

The uncertainty is mainly related to inaccurate sensing in the flaking process. Therefore, we assume the moisture obtained through a near-infrared (NIR) sensor is an uncertain parameter, and, through the data analysis, the moisture follows an approximately normal probability distribution as shown in Fig. 9a. Figure 9b shows the distribution of moisture measurement errors through the NIR sensor.

Thus, the variation of the moisture (M) can be described as a normal distribution with mean $\bar{M}$ and standard deviation ω, following Equation (4).
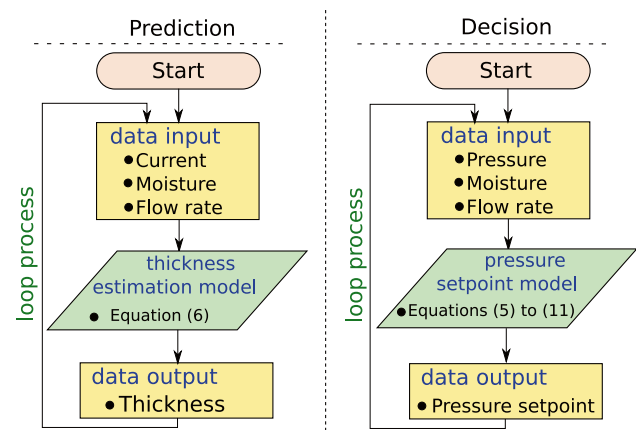

Fig. 8 Flowchart of prediction and decision methods

$$M \sim \mathcal{N}(\bar{M}, \omega) \tag{4}$$

Applying the two-stage-stochastic technique [33], the first decision to be made is about the pressure used on the machine. This decision is made without complete information which characterizes the first stage, and the other decisions are made based on the previous one.

The second stage is divided into a set of discrete scenarios $s = 1 \ldots S$ with a probability of $p_s = \frac{1}{S}$ to happen. Each s scenario represents a possible configuration in terms of moisture (M_s), and we can make optimal subsequent corrective decisions, in terms of flake thickness (T_s) and electric current (C_s), for each one. Each decision variable in the first stage will be considered as fixed for the second stage. This type of stochastic treatment puts all the random information and the corrective action decisions C_s and T_s in the second stage. The scenarios for moisture will be constructed through Eq. (4), where ω is the average variation of the moisture in a given period.

Fig. 7 Relationship between the measured thickness and obtained by the regression
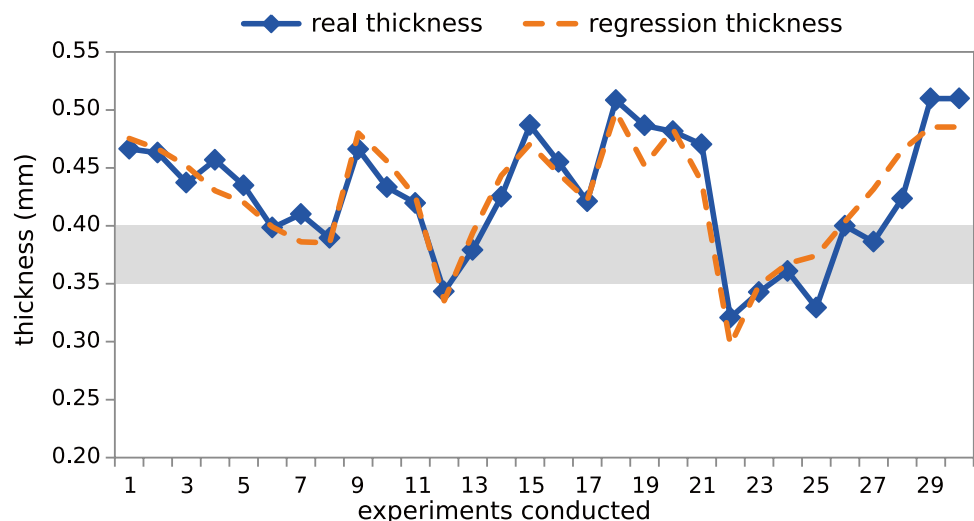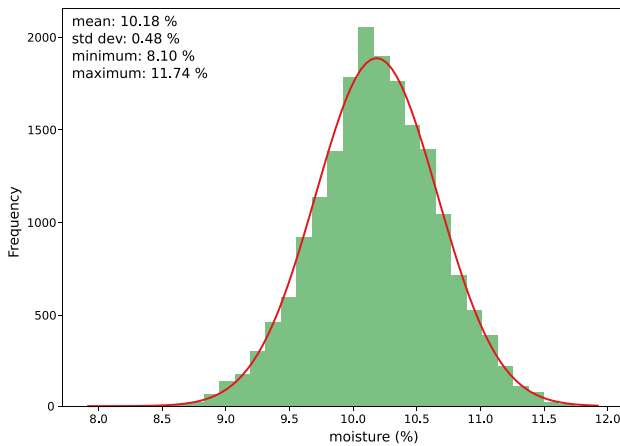
(a) Histogram of moisture distribution.
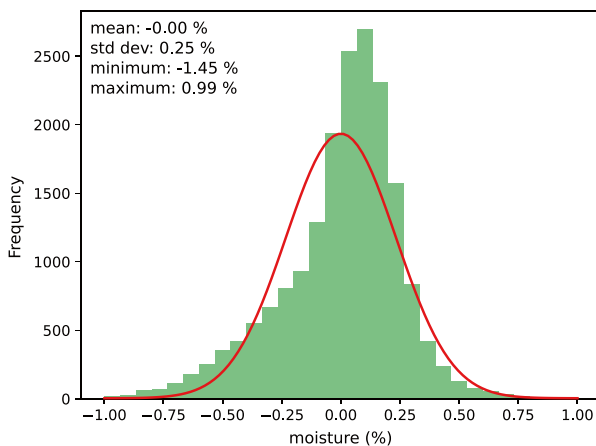
(b) Histogram of error distribution.

Fig. 9 Frequency distribution histograms

Decisions on the first stage inviable at the second stage are not valid. The optimal solution to the problem is to evaluate the best decision on the first stage that is feasible in both stages, minimizing the thickness error in the second stage. Thus, there is a subproblem to be solved in the second stage for each s scenario, where M_s is one of the deterministic parameters and the pressure P is a fixed decision.

We have the following modeling for the problem: Eq. (5) minimizes the module of the difference between variable T_s and the ideal soybean pattern in the value of 0.375, weighted by the occurrence of each scenario. The value 0.375 is the central value in the operating range; therefore, the objective function minimizes non-standard thicknesses in all possible scenarios.

$$min \quad \sum_s p_s \times \|E_s - 0.375\| \tag{5}$$

Equation (6) obtains the estimation of the flake for a scenarios s.

$$T_s(C_s, M_s, F) = -0.0037 \times C_s - 0.0350 \times M_s + 0.1188 \times F + K_t + W_s \tag{6}$$

W_s represents possible modeling errors from this linear equation approximation that estimates the flake thickness. The variation W_s can be described as a normal distribution with zero mean and standard deviation σ according to Eq. (7).

$$W_s \sim \mathcal{N}(0, \sigma) \tag{7}$$

Equation (8) gives us the ideal pressure to generate a electric current that produces a desirable flake thickness, based on the input parameters for each generated scenario.

$$C_s(P, M_s, F) = 1.1170 \times P - 1.4199 \times M_s + 7.2613 \times F + 24.644 + Z_s \tag{8}$$

Following the previous treatment, the modeling errors Z_s can be described by a normal distribution with zero mean and standard deviation ψ as shown in Eq. (9).

$$Z_s \sim \mathcal{N}(0, \psi) \tag{9}$$

Equation (10) gives the boundaries for pressure allowed when operating the rolling mill.

$$40 \leq P \leq 60 \tag{10}$$

Equation (11) determines the domain of the variables (T_s and C_s).

$$T_s, C_s \in \mathbb{R} \tag{11}$$

5 Results

This section will present results about the validation of the prediction step and the performance of the mathematical model. In all experiments, the mathematical stochastic model generates 100 scenarios and the average value of the pressure, solving each scenario, is set in the rolling mill. The related linear programming model for each scenario is coded and solved using the optimization solver CPLEX. The computational tests run in a machine under 64 bits Windows 10 operating system, Intel Core i5 with 2.70 GHz and 8 GB of RAM. Each scenario takes a few milliseconds to be solved, and the overall evaluation of the scenarios takes much less than 1 minute, which is the time limit demanded by the industry for an automatic decision.

5.1 Validation of the flake prediction function

Function (3) estimates the thickness of flakes based on the multiple regression analysis done over a set of 29 real data

collected from the factory. The next step was to validate it over 17 new real data collected during the normal operation of the rolling mill. Figure 10 represents the results obtained from the validation set.

In this case, we have an average error that reaches approximately 3%. Thus, the function was able to predict flakes thicknesses based on engine current conditions, soybean moisture, and the amount of raw material being processed within an acceptable error margin for the industry.

Table 3 represents the minimum (min), average (avg) and maximum (max) thicknesses obtained by both methods. The average estimated using the prediction model has a mean percentage error of 2.87%. This analysis highlights the reliability of the prediction method by obtaining measurements for the flakes very close to the real ones.

5.2 Model validation over the first data set

The first evaluation of the mathematical model is carried on the 29 set of data initially provided by the industry. The pressure control obtained through the proposed model is illustrated in Fig. 11.

In the set of experiments from 0 to 10, the pressure returned by the model is maintained at the maximum values, and the thickness is kept in the ideal standard value until the moisture parameter becomes to decrease in the data set. The soybean's resistance increases as the moisture decreases, becoming necessary to apply more tension to deform the crashed beans in the ideal flake thickness. However, the pressure is already at the maximum which reveals problems during the previous drying process of the beans, or imprecise data provided by the NIR sensor. Since there is no other way to control the thickness with pressure, we have some flakes out of the standard in the experiments 8 and 9.

The increase in the moisture for experiments 11, 12 and 13 leads to decrease in the tension to deform the crashed beans. We had another decrease in moisture for experiments

Table 3 Min, max and avg prediction function

Group	Methods	Real thickness	Estimated thickness	Error(%)
min		0.3184	0.3114	0
avg		0.3513	0.3484	2.8780
max		0.3434	0.3939	10.8696

14, and the model set the pressure again to its maximum value. In the experiments from 22 to 30, the moisture has slight changes, but the flow has a drop. In this case, there are a few crashed beans going through the rolling mill, so the model takes this information from the flow rate parameter and decreased pressure. As soon as the flow begins to increase, the pressure also increases. At the end, the flakes did not remain in the desirable range only because the pressure is already at the maximum value. Thus, it was not possible to handle better the increase in the flow rate given the pressure boundary in the rolling mill. Both a high level of flow rate and moisture contribute to thickness out of the desired value in the last experiments.

5.3 Model validation over a new data set

The next test takes 17 new real data collected in the industry under a regular operating condition. In this case, it was possible to compare the thickness of flakes generated through our decision system and the thickness obtained by manual operation. According to Fig. 12, the manual adjustments were not able to keep the flakes within a range for all samples. The settings returned by the proposed model are able to keep the flake thickness under control. The moisture and flow rate parameters were ideal to obtain a high-quality soybean lamination. These results validate the assertiveness of the model. It is worth mentioning that, even in the optimal conditions provided by the soybean input parameters, the
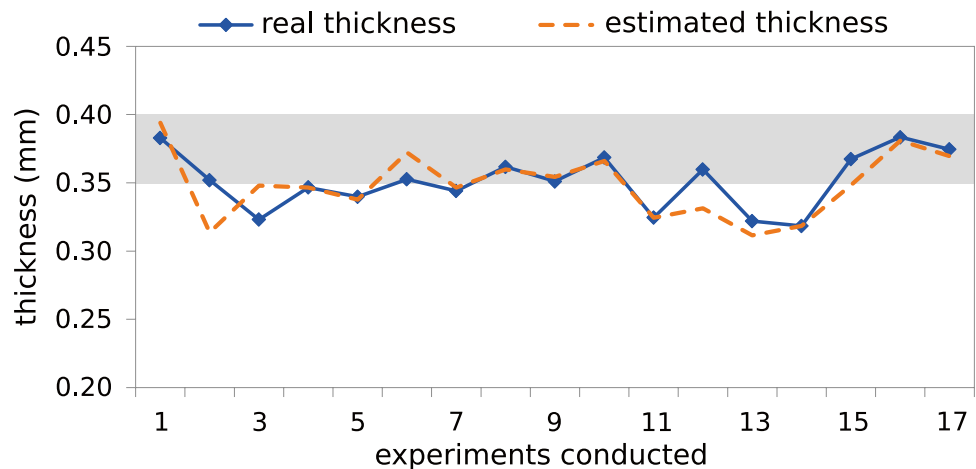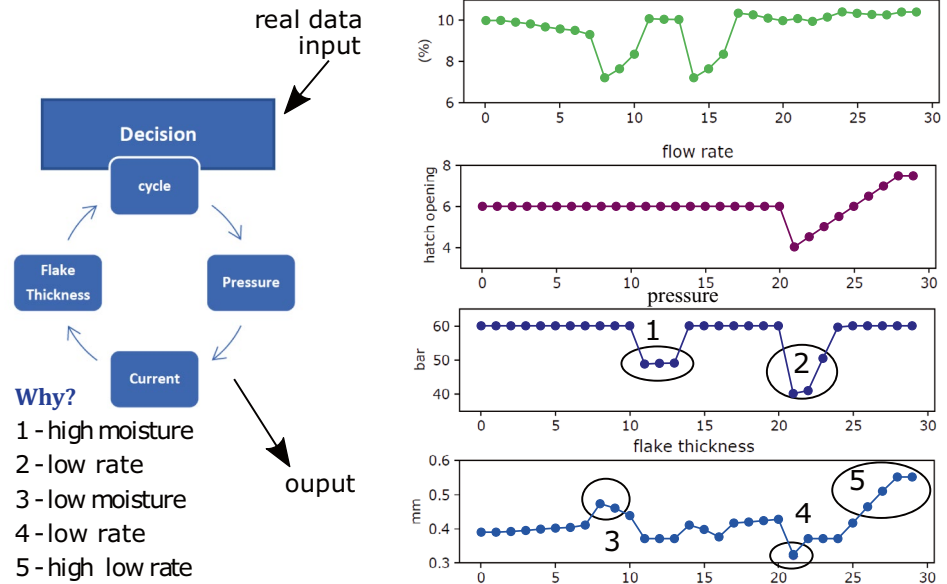
Fig. 10 Validation of flake prediction function

Fig. 11 Relationship between pressure setpoint and input parameters

manual operation could not keep the flakes within the optimal range of operation.

Figure 13 shows the pressure setpoint given by the mathematical model and the pressure setpoint applied manually by the operators. The operators' pressures remain at the upper limit most of the time, which generates flakes in undesirable values since the change in moisture directly influences the thicknesses obtained. On the other hand, the mathematical model makes pressure decisions different from manual operation, based on changes in input parameters, to obtain better results as proposed.

From these results, we can conclude that the solutions obtained through the proposed method are superior to the manual methods used today in the industry. At the moment, the lack of real data prevents us from evaluating more situations for this problem. There is no sensing able to measure flake thickness automatically in the production

process, which becomes hard to build large data sets. However, the idea is to integrate soon the proposed method in an automatic production process where such measurements will feed our method in real time.

5.4 Operation analysis

In this last analysis, a test was performed considering the 24 hours of a specific operation day with fully data collected. We evaluate the operators' decisions made about three times a day, and our method decisions made every one minute of operation.

Figure 14 shows the input data for each time instant analyzed. The flow rate is six during the first 11 hours of operation, and it is decreased to five to satisfy other industrial needs during the whole soybean process on that operation day.

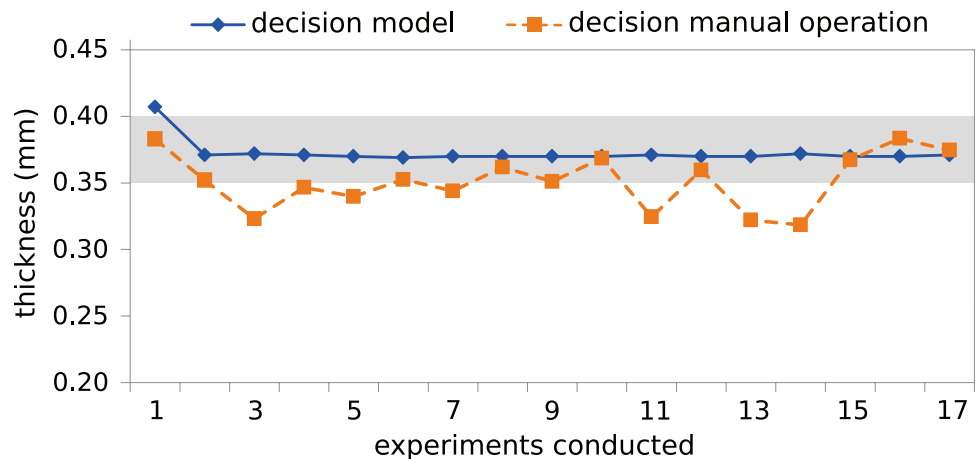**Fig. 12** Thicknesses obtained through manual operation and decision model

Fig. 13 Optimum pressure obtained by model and manual operation pressure
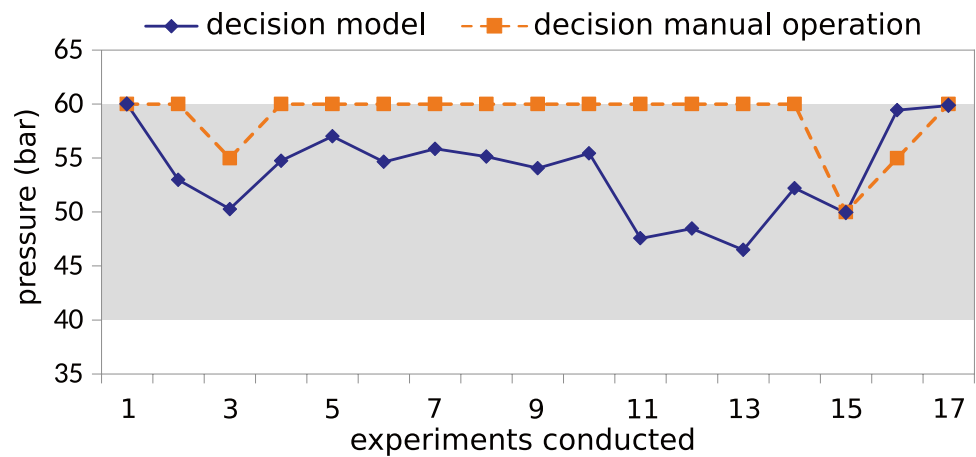

Figure 15 illustrates the pressure decisions made by operators and the pressure decisions made through the our method. In the manual operation, only two pressure decisions were made, while our method suggested several variations. Based on the pressures suggested by the manual setting and our method, the flake thickness is estimated for that period as shown in Fig. 3.

In the first 11 hours of operation, the decisions of both methods are the same. The decision to keep the pressure at 60 is justified by the high flow rate. The moisture changes within the ideal range (Fig. 14), but the flakes still tend to exceed the upper limit (Fig. 16). Between 10:50 and 15:00 hours, the moisture decreases (Fig. 14) and both methods suggest to decrease pressure to keep the ideal flake thickness. However, the manual operation decided to set pressure as 55 bar, leading the flakes to decrease their thickness out of the ideal range of operation in Fig. 16. On the other hand, our method decided for a lower pressure close to 40 bar, keeping the flakes controlled and avoiding thin flakes.

As the manual operation takes only two pressure decisions, because it needs human intervention, the flake thicknesses are not well controllable. They stay out of the ideal

operating standard for approximately 12 hours. Our method would keep the flake controlled and stable during all the operating process. Thus, an advantage of our method is the continuous adjustment, once such settings will change pressure as many times as necessary. It becomes possible to avoid undesirable flakes in longer time intervals.

As shown in Fig. 3, the goal is to keep flakes in the ideal range of operation, avoiding waste and filters problems. Figure 17 shows the estimated percentage of oil lost in the extraction process for both modes. These estimations are based on the knowledge from the decision-makers in the industry, after evaluating the performance of both methods.

As we can see in Fig. 17, our method obtains loss of oil of 1.05% on average during the extraction process, while the manual mode returns loss of 1.15% on average. The deviation obtained through both modes is 0.11%. To sum up, in the operation analysis, considering 24 hours of production, we used the decision model to obtain pressure setpoints that keep the blades optimized. This analysis compares the oil loss of both methods, getting an average deviation of 0.11%. Extrapolating this result, considering, for example, one year of production, we can see that it would be possible to obtain savings of approximately US$100,000.
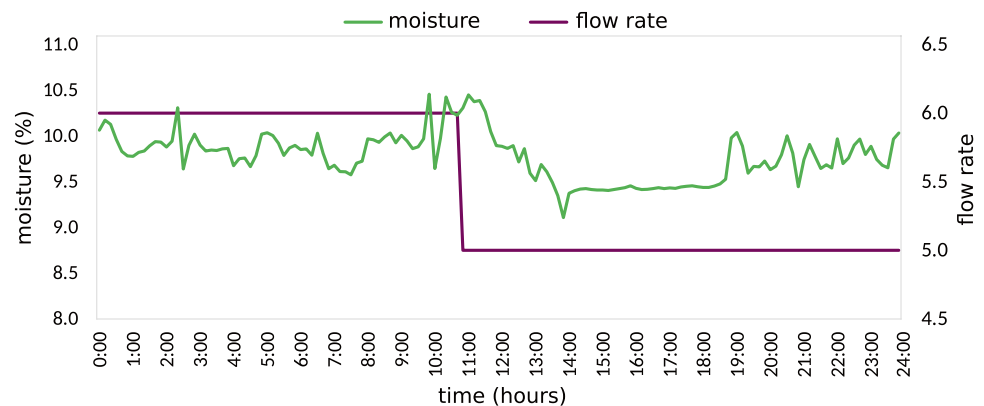
Fig. 14 Input parameters

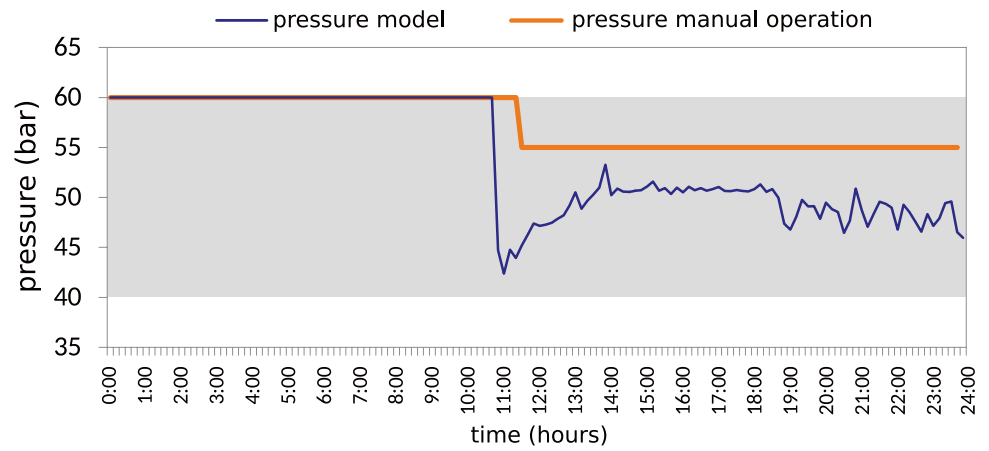**Fig. 15** Optimum pressure obtained by model and manual operation pressure

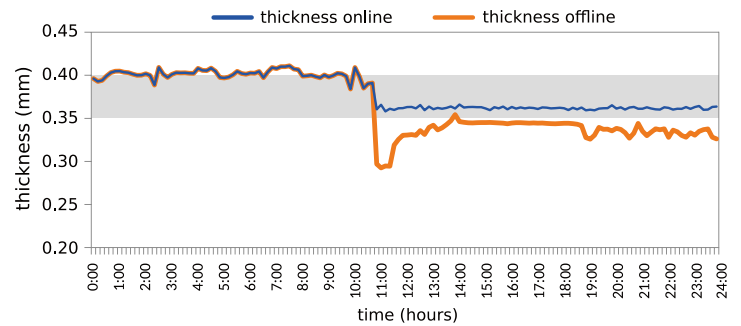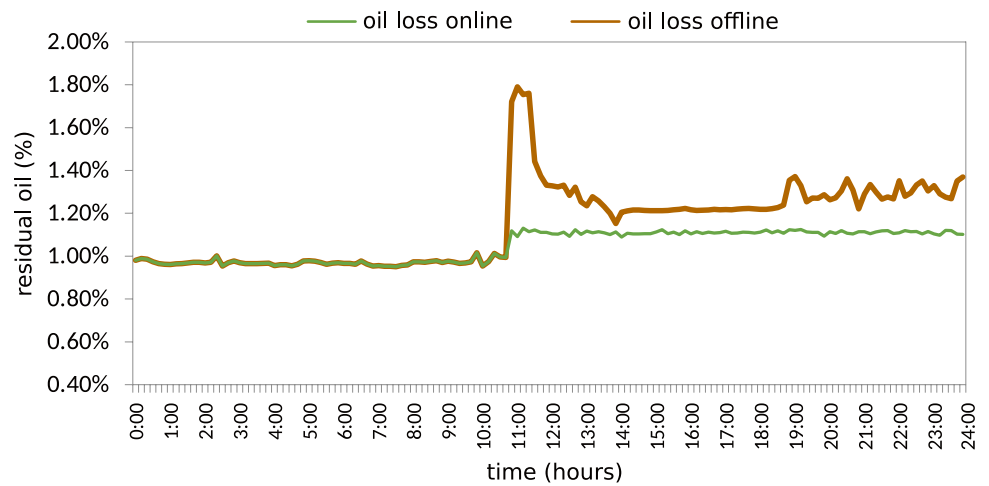**Fig. 16** Thicknesses obtained through both modes

Fig. 17 Oil loss obtained through both methods

6 Conclusions

This paper introduced a stochastic model developed specifically to attend a real industry scenario, where the extraction of oil must be maximized based on the thicknesses of the soybean flakes. Thus, the proposed method adjusted the pressure parameter in the rolling mill, employing regression analysis and stochastic programming to support decision-making. The multiple linear regression analysis and the stochastic model are developed to deal with the uncertainties from sensor measurements. The experiments showed that the method outperformed the manual operational currently used in the industry. The decision made, using the stochastic model, keeps the flakes within the required range of operation. The method improves the whole process in the rolling mill in terms of time and quality. It takes much less than one minute to make a decision, becoming possible to update in short time periods the rolling mill pressure, based on the quality of flakes provided in the input data. The results also reveal some possible problems related to the production chain. For instance, the method suggested a maximum pressure in some of the experiments, but it was not enough to keep the flakes within the standard thickness. This behavior can be caused by a previous process, such as the quality control step for beans provided by suppliers or the drying step that keeps the right level of moisture during the storing. If there is a problem with those previous steps, the rolling mill will have problems to do its job properly. The mathematical modeling developed is being used by the industry for a decision-making process involving the complete soybean oil extraction process. It is still a challenge to get the constant degradation of the roll, a factor that influences the thickness of the flakes. As future works, studies are being developed to include the degradation of the roll, as well as other possible parameters, which can lead to a more precise description by the stochastic model, and better decisions about pressure in the rolling mill.

Declarations

Ethics approval Not applicable.

Consent to participate All the authors involved in the manuscript contributed in some way to its completion and application in industry.

Consent to publication All the authors involved in this manuscript give full consent for publication of this submitted article.

Competing interests The authors declare no competing interests.

References

1. Hofmann E, Rüsch M (2017) Industry 4.0 and the current status as well as future prospects on logistics. Comput Ind 89:23–34
2. Hirakuri MH, Lazzarotto JJ (2014) O agronegócio da soja nos contextos mundial e brasileiro. Embrapa Soja-Documentos (INFOTECA-E)
3. Alem D, Morabito R (2015) Modelos de programação estocãistica no planejamento da produção de empresas moveleiras. Production 25(3):657–677
4. Alem D, Curcio E, Amorim P, Almada-Lobo B (2018) A computational study of the general lot-sizing and scheduling model under demand uncertainty via robust and stochastic approaches. Comput Oper Res 90:125–141
5. Hu Z, Hu G (2016) A two-stage stochastic programming model for lot-sizing and scheduling under uncertainty. Int J Prod Econ 180:198–207
6. Li Y, Hu G (2017) Shop floor lot-sizing and scheduling with a two-stage stochastic programming model considering uncertain demand and workforce efficiency. Comput Ind Eng 111:263–271
7. Tang L, Che P, Liu J (2012) A stochastic production planning problem with nonlinear cost. Comput Oper Res 39(9):1977–1987
8. Jung S, Mahfuz A, Maurer D (2009) Structure, protein interactions and in vitro protease accessibility of extruded and pressurized full-fat soybean flakes. J Am Oil Chem Soc 86(5):475–483
9. Koons G, Perlic B (1977) A study of rolling-mill productivity utilizing a statistically designed simulation experiment. Technical Report. Institute of Electrical and Electronics Engineers (IEEE)
10. Seth S, Agrawal Y, Ghosh P, Jayas D (2010) Effect of moisture content on the quality of soybean oil and meal extracted by isopropyl alcohol and hexane. Food Bioprocess Technol 3(1):121
11. Xie L, Ar He, Liu C (2018) A rapid calculation method for predicting roll deformation of six-high rolling mill. J Iron Steel Res Int 25(9):901–909
12. Alvarado L, Torres-Treviño LM, González F, Nieves L (2014) A mathematical model of a cold rolling mill by symbolic regression alpha-beta. In: Proceedings of the Companion Publication of the 2014 Annual Conference on Genetic and Evolutionary Computation, pp 1347–1352
13. Calvo-Rolle JL, Casteleiro-Roca JL, Quintián H, del Carmen Meizoso-Lopez M (2013) A hybrid intelligent system for pid controller using in a steel rolling process. Expert Syst App 40(13):5188–5196
14. Gasyarov VR, Radionov AA, Loginov BM, Voronin SS, Khramshin VR (2017) Improvement of work roll bending control system installed at plate mill stand. In: Proceedings of the 9th International Conference on Computer and Automation Engineering, pp 269–273
15. Pérez D, Díaz I, Cuadrado AA, Rendueles JL, García D (2018) Interactive data visualization of chatter conditions in a cold rolling mill. Comput Ind 103:86–96
16. Wang H, Wang C, Wang J (2018) Design of roller and roll gap adjustment mechanism for small mill. In: Proceedings of the International Symposium on Big Data and Artificial Intelligence, pp 124–128
17. Singh P, Maier D, Okos M, Cattanach E, Trumble K (1999) Effects of physical properties and operating parameters on soybean flaking. J Am Oil Chem Soc 76(8):981–987

18. Sun JL, Peng Y, Liu HM (2014) Dynamic characteristics of cold rolling mill and strip based on flatness and thickness control in rolling process. J Cent South Univ 21(2):567–576

19. Buetas E, Abad I, Cerrada JA, Cerrada C (2020) A propagation breakdown management model for the industrial internet of things. Comput Ind 123:103305

20. Copil K, Wörbelauer M, Meyr H, Tempelmeier H (2017) Simultaneous lotsizing and scheduling problems: a classification and review of models. OR Spectrum 39(1):1–64

21. Toledo CF, Kimms A, França PM, Morabito R (2015) The synchronized and integrated two-level lot sizing and scheduling problem: Evaluating the generalized mathematical model. Math Prob Eng 2015

22. Tomotani JV, de Mesquita MA (2018) Lot sizing and scheduling: a survey of practices in brazilian companies. Prod Plan Control 29(3):236–246. https://doi.org/10.1080/09537287.2017.1409370

23. Buhler AG (2006) Laminador de copos. Technical Report, Suíça. Buhler Group

24. Mandarino JMG, Roessing AC (2001) Tecnologia para produção do óleo de soja: Descrição das etapas, equipamentos, produtos e subprodutos. Technical Report. Embrapa Soja-Documentos (INFOTECA-E)

25. Masieiro LC (2012) Preparo de grãos de seja para extração. Technical Report. Engenharia Industrial

26. Pereira MA, Castro VA (2015) Cadeia produtiva do farelo de soja: Um enfoque na produção nacional. Technical Report. Faculdade de Engenharia de Produção, Universidade do Rio Verde (UNIRV)

27. Sanchez-Marquez R, Guillem JMA, Vicens-Salort E, Vivas JJ (2020) Diagnosis of quality management systems using data analytics-a case study in the manufacturing sector. Comput Ind 115:103183

28. Silva JPBCD (2017) Modelos de Regressão Linear e Logística utilizando o software R. PhD Thesis

29. Fachini RF, Esposto KF, Camargo VCB (2017) Glass container production planning with warm-ups and furnace extraction variation losses. Int J Adv Manuf Technol 90(1–4):527–543

30. de Souza Amorim FM, Arantes MDS, Toledo CFM, Frisch PE, Almada-Lobo B (2018) Hybrid genetic algorithms applied to the glass container industry problem. In: 2018 IEEE Congress on Evolutionary Computation (CEC), pp 1–8. https://doi.org/10.1109/CEC.2018.8477762

31. Toledo CFM, da Silva Arantes M, Hossomi MYB, França PM, Akartunalı K (2015) A relax-and-fix with fix-and-optimize heuristic applied to multi-level lot-sizing problems. J Heuristics 21(5):687–717

32. Toledo CFM, Arantes M, Hossomi MYB, Almada-Lobo B (2016) Mathematical programming-based approaches for multifacility glass container production planning. Comput Oper Res 74(2016):92–107

33. Birge JR, Louveaux F (2011) Introduction to stochastic programming. Springer Science & Business Media, U.S.A