ISBN: 978-628-95207-4-3. ISSN: 2414-6390. Digital Object Identifier: https://dx.doi.org/10.18687/LACCEI2023.1.1.860

Productivity improvement in an agribusiness dedicated to the export of snow peas using Lean Manufacturing and Mathematical Optimization tools

Jonatan Rojas Polo, Mg[1], and Alexia Cáceres Cansaya, Mg[1]
[1]Pontificia Universidad Católica del Perú, Perú, jrojasp@pucp.pe, alexia.caceres@pucp.pe

Abstract*– This research focuses on the situation faced by an agribusiness that processes Snow Peas in the planning of its supply chain in the phases of the crop field, transportation to the production plant, and the processing plant. The planning is carried out in the short term horizons, which deals with the allocation of the operators' tasks in order to fulfill the customers' orders; and the medium term horizon that determines the monthly planning based on the harvest forecast data provided by the crop fields. In previous research, a vehicle routing algorithm was developed to optimize the collection of raw material [1]. Once the raw material arrives at the processing plant, there is a workforce of approximately 200 people in different production areas that have an 8-hour workday, and whose production capacity is about 5.6 tons of raw material entering the plant.

The principal problem detected in the research is the lack of reliability in the forecasting method for the supply of raw material; for example, on several occasions, an entry of 5 tons of material is estimated when at the end of the day only 2 tons enter the plant, which generates a very high variation in the use of the plant's capacity. Also, it was determined that there is a deficient standardization of processes and a low operational capacity in the selection, cutting, and packing areas, which is reflected in a cost of US$1.2 per kilogram processed.

Therefore, this research determined a three-phase methodology to improve the processes using lean manufacturing and mathematical optimization tools. In the first phase, data was collected from the crop fields in order to create a multivariable mathematical-statistical model; the model obtained has an accuracy level of 85%. In the second phase, a value stream mapping was carried out to determine value-added times, non-value-added times, and activities that do not generate value; this was the basis for restructuring process operations and developing a new layout that maximizes the production flow. In the third phase, a mathematical S&OP model was developed to determine the number of operators per week to maximize production capacity. The application of the research resulted in an 80% increase in workers' salaries, a 150% increase in capacity, and a reduction in production cost to US$ 0.38 per kilogram processed.

Keywords-- Snow peas processing, Lean manufacturing and mathematical optimization synergies, Agribusiness process optimization, Multivariate forecasting for harvesting, Optimization in S&OP.

Digital Object Identifier: (only for full papers, inserted by LACCEI).
ISSN, ISBN: (to be inserted by LACCEI).
DO NOT REMOVE

I. INTRODUCTION

Agribusiness in Peru focused on exports is growing without previous precedent during the last decades, because Peru is a megadiverse country with 3 regions: coast, highlands, and jungle. The agriculture and livestock sector represents approximately 6.4% of the GDP [2]. During 2021, the highest level of exports in the last 20 years was reached; according to MIDAGRI, the main crops were: In Tumbes with banana, lemon and cocoa crops; in Piura with grape, banana and mango crops; in Lambayeque with avocado, mango and blueberry; La Libertad with asparagus, avocado and blueberry; Ancash with avocado; Lima with mandarin and avocado; Ica with asparagus and grapes; Moquegua with avocado and grapes; Arequipa with onion and potato; Tacna with olives; Cajamarca with coffee and potato; Huánuco with potato and cocoa; Pasco with coffee, potato and rocoto; Junín with coffee, potato and cocoa; Huancavelica with potato, starchy corn and avocado; Ayacucho with potato, cocoa and quinoa; Apurímac with potato, starchy corn and kiwicha; Cusco with potato, coffee and starchy corn; Puno with potato, fodder oats and quinoa; Amazonas with coffee and rice; San Martin with rice, coffee and cocoa; Loreto with cassava, plantain and rice; Ucayali with oil palm, cocoa and plantain; and Madre de Dios with brachiaria, hard yellow corn and plantain [3, 4, 5]. However, competition in this agro-industrial sector has a greater scope than at the regional level, with competition at the global level. There are major competitors such as China (canned asparagus, artichokes, peppers, paprika peppers), Kenya (Holantao), and Mexico (fresh asparagus), among others. It is a necessary obligation for Peruvian agribusiness to strengthen its image as a supplier of high-quality agricultural products, and to add value through innovative products. In addition, they must offer prices that are competitive with other international markets. Therefore, these companies must possess techniques and knowledge to exceed production standards to reduce operating costs and optimize the use of raw materials. They must also have excellent planting and adequate post-harvest management.

To achieve this, companies must develop medium and long-term adaptability strategies in the face of a very unstable consumer market. For this reason, this research is carried out in an agro-industrial company located south of Lima, which produces mainly two types of products, Snow peas (Holantao) and Sugar daddy, in the presentations required by the client. This company has fields certified under the European

GLOBAL G.A.P. standard. However, profit contribution margins are low, given that they currently have a raw material use yield of 70%. It is worth mentioning that Snow Peas also called holantao or snow pea, with the scientific name Pisum Sativum of the saccharatum variety [6], a legume native to the Mediterranean and East Africa, and belong to the Fabaceae family, of the Fabales order [7]. The fruit is a pod, somewhat compressed and ending in a small curve.

II. STATE OF THE ART

This section details the principal tools that serve as the foundation for this research.

A. Multivariable regression model

Francis Galton was the one who coined the term regression, he argued that, despite the tendency of tall parents to produce tall children and short parents to produce short children, the average height of children of parents of a given height tended to return to the average height of the total population. Concluding that the height of the children of unusually tall or unusually short parents tended to go toward the average height of the population. This law of universal regression was confirmed by Karl Pearson, who collected over a thousand records of family group heights. Pearson discovered that this was a phenomenon whereby tall children and short children "regress" equally to the average height of all others [8, 9, 10]. The modern interpretation of regression states that regression analysis deals with the study of the dependence of a variable (dependent variable) on one or more variables (explanatory variables) with the objective of estimating or predicting a population average value of the former in terms of known values. Although regression analysis is concerned with the dependence of a variable on other variables, it does not necessarily imply causality. That is, regression analysis involves situations that have more than one regressor variable, which is called a multiple regression model. Equation 1 details a multiple regression model:

$$y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_n X_n + \varepsilon \qquad (1)$$

Where y represents the dependent variable, X_i represents the independent variables, and ε is the random error [11].

B. Lean manufacturing

Lean manufacturing is a management methodology that has been applied in many manufacturing plants these decades. It started in the automotive industry [12], then sequential improvement initiatives were carried out to improve changes in manufacturing practices. It is a methodology that seeks to increase the value to the customer. According to Womack & Jones, lean manufacturing is a five-step process: define customer value, define the value stream, make it flow, pull it from the end (customer) and pursue excellence [13, 14, 15].

C. Aggregate planning – S&OP

An aggregate production plan is a tool that integrates the plan. That is all the aggregated units, for normally monthly periods, considering the production variables, trying to fulfill the long-term plan [16, 17].

The objective of aggregate planning also called sales and operations planning (S&OP) is to meet the forecasted demand and minimize planning costs [18]. The aggregate sales and operations plan is useful because it focuses on a general course of action, consistent with the company's strategic goals and objectives. Aggregate planners must determine whether they can meet the budget objectives by scheduling each of the company's products and employees individually. In general, companies conduct aggregation along three dimensions: services or products, workforce, and time [19]. Four basic strategies can be used to determine the number of resources: Chase strategy, Stable workforce—variable work hours, level strategy, and mathematical modeling [20].

III. CASE STUDY

In this section, the research describes and evaluates the current situation, followed by the methodology and implementation of the process improvement proposal.

A. Current situation

The research was carried out in a company that has more than 100 hectares planted with Snow Peas and Sugar Daddy, which are distributed in the town of Humay belonging to Pisco - Ica, as shown in Figure 1.

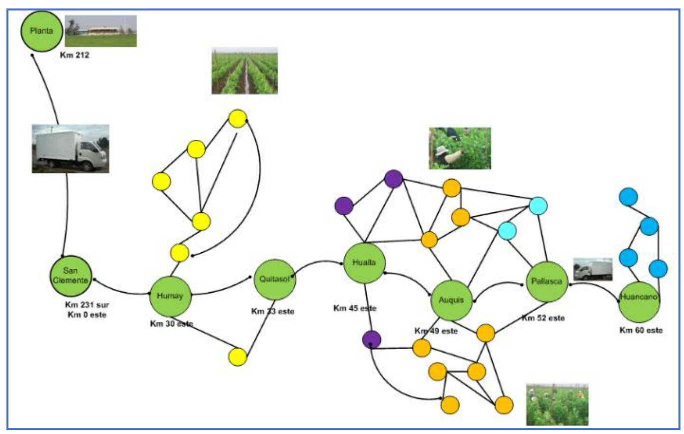


Fig. 1 Diagram of the crop fields.

Snow Peas have a planting/harvest cycle time of approximately 60 to 80 days, depending on the seed variety. Of this planting cycle, only 25 to 45 days are useful for harvesting, so harvesting, storage, and post-harvest handling must be properly planned to synchronize their flow during the value chain and maximize the quality of raw material obtained in the field.

A.1. Raw material procurement forecast

Currently, the field harvest forecast is developed by the engineer in charge of the field, where his main criterion to determine the kilograms to be harvested per type of field, is based on distributing the theoretical yields of the type of crop (according to the company selling the seeds are 8 tons per hectare) among the harvest days. The total daily harvest amount is the sum of all the fields authorized for harvest in the planning period.

This way of forecasting has brought with it many planning drawbacks since not all sowing fields have the same harvest yield in reality. The locations of the planting fields are shown in Table I.

Table I
Notation of each lot - crop field

Location	Lot	Location	Lot
Humay	Nestares 1	Auquis	Huamani 2
Humay	Nestares 2	Auquis	Huamani 3
Humay	Nestares 3	Auquis	Saavedra 2
Humay	Nestares 4	Pallasca	Sebastián 1
Humay	Nestares 5	Pallasca	Sebastián 2
Humay	Nestares 6	Huancano	Huancano 1
Hualla	Pantaleón 1	Huancano	Huancano 2
Hualla	Pantaleón 2	Huancano	Huancano 3
Hualla	Pantaleón 3	Huancano	Huancano 4
Auquis	Valdivieso 1	Humay	
Auquis	Valdivieso 2	Hualla	
Auquis	Huamani 1	Auquis	
Auquis	Blanco 1	Pallasca	
Auquis	Saavedra 1	Huancano	
Auquis	Blanco 2	Quitasol	Yauricasa

Figure 2 shows the daily behavior of kilograms provided by the plant during the harvest stage. It can be seen that during the first 2 weeks, its maximum harvest is 200 kilograms per day, and during weeks 3 and week 4 it increases up to 1200 kilograms per day (peak production week), then from day 29 to day 45 the production for harvest is notoriously reduced.

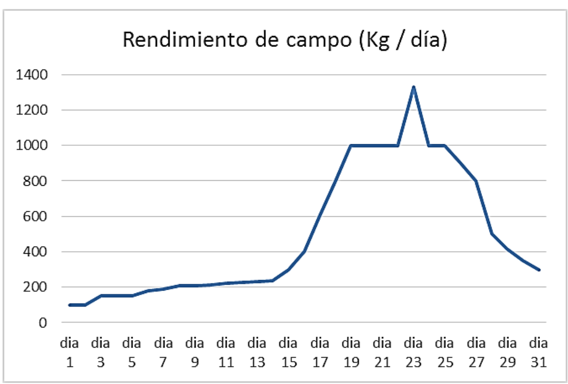


Fig. 2 The normal behavior of a snow peas harvesting lot.

A.2 Production and processes in the operating plant:

The variation in the raw material harvest forecast significantly impacts the capacity of the production plant. In other words, production capacity is often greater than the necessary capacity, given that the availability of raw materials is highly variable. In the analysis, we also observed that there is no correct labor force hiring policy. In other words, the production department is not sure how many people should be working in a given period. We also observed idle time in the workers' activities, one of the reasons being the lack of raw materials, thus incurring higher costs. Table II shows the production capacities of each task.

Table II
Production capacity per area - operator

Capacity		Kg/hour-operator	Hour/day	Kg/day-operator
Selection	Selection	30	8	240
	Selection supplier	600	8	4800
	Selection weight control	600	8	4800
Cutting	Cutting supplier	300	8	2400
	Punnet cutter	5.7	8	45.6
	Weighing of cut raw material	500	8	4000
Packing	Punnet coder	125	8	1000
	Punnet filling	30	8	240
	Weighing of punnet	40	8	320
	Punnet closing	120	8	960
	Bulk bag filling	60	8	480
	coder/weigher	120	8	960
	USA Packaging - Selection/Weighing	80	8	640
Cold storage	Box labeling	768	8	6144
	Stacking and palletizing	288	8	2304
	Storage	600	8	4800
Maintenance	Cleaning	4 personas		
	Maintenance	2 personas		

The production system is composed of five areas: selection, cutting, packing, cold storage and maintenance.

a) Selection

In the selection area we find the person who receives the raw material, the supplier of the selection belt, the selection personnel, and the personnel who collect and weigh the raw material that has been selected. See figure 3

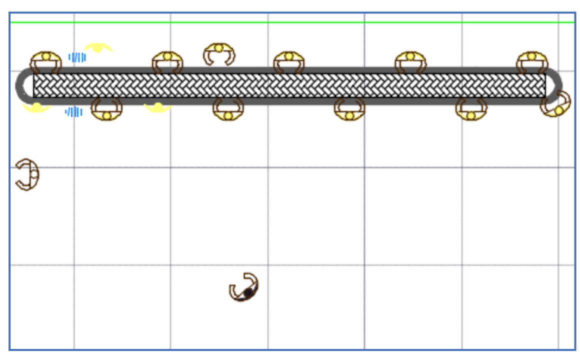


Fig. 3 Selection area.

In this area, first, the raw material collected in the field is weighed, then the sorting supplier places the raw material on the belt and the sorters are in charge of separating the products that present defects and foreign bodies that may be immersed in the raw material. There are 23 selectors, 1 controller, and 1 supplier on the line. The pace of this area is determined by the progress of the selectors.

b) Cutting

In the cutting area we find the person that supplies the selected raw material to each of the cutters. The cutting operation consists of separating the ends of the snow peas or sugar daddy, using small scissors, approximately 10 cm long. The length of the Snow Peas should be in the range of 7.5 to 11 cm. The cutters work on a table. The supplier supplies the selected product to be cut, then the cutters place the cut snow peas in containers of approximately 10 kilograms. There are 117 cutters, 1 controller, and 2 suppliers in the cutting area. The cutting operation is shown in Figure 4.

Fig. 4 Cutting operators.

c) Packaging:

After the Snow Peas have been cut and weighed, approximately 6 containers of the product are placed on a table, where operators place the product in trays called punnets. Two types of trays are used, the B-124 and the B-138, and their use depends on the presentations to be made. The presentations are as follows: Box 12 x 250 gr, Box 8 x 150 gr, Box 24 x 200 gr, Box 28 x 200 gr, and Box 10 pounds. These presentations are for both Snow Peas and sugar daddy. Figure 5 shows two presentations.

Fig. 5 Left: Presentation of 250 grams of Snow Peas. Right: Presentation 150 grams of sugar daddy.

In the packing area, there are 22 punnet fillers, 16 weighers, 5 sealers, and 5 punnet coders or bulk bags. When the punnets are filled, the weighing operation is congested, since the weigher is the one who often has to arrange the presentation of the tray. Figure 5 shows the punnet tray presentation. Finally, the weighted tray is sealed with a porous microfilm. See Figure 6. In this area the packaging for the 10-pound boxes destined for the United States is also developed, the 10-pound format is a simple process to fill since the presentation of the filling does not matter. In the packing area, the weighers are the ones that set the pace of production.

Fig. 6 Film sealing a 150-gram format - Snow Peas.

d) Cold storage:

The cold room is one of the most important spaces in the process, given that the product is kept at a temperature of 3 to 6 °C to prolong its shelf life. The tasks in this area include palletizing the boxes, and coding and strapping the pallets, as shown in Figure 7. The person who determines the pace of production is the palletizer.

Fig. 7 Strapping of a pallet 11x20 boxes - 8x200gr. Snow Peas.

e) Maintenance:

The maintenance area is a transversal area that watches over the daily operation of the electrical flow, machinery, cold air flow, cooling vats, and machinery used for production and shipping. It also ensures the cleanliness of the process plant. A total of 2 metal mechanics technicians and 4 people work in cleaning. The number of personnel in each area is sufficient to process 5.6 tons of raw material per day.

The following is the analysis of the processes using the value stream mapping tool. First, we detail the iterations of the supply chain; the relationship between the customer, our production system, and our supplier. The zigzag lines represent the flow of information, orders, demand forecasting, and product types. This planning occurs in short and medium-term periods, see Figure 8.

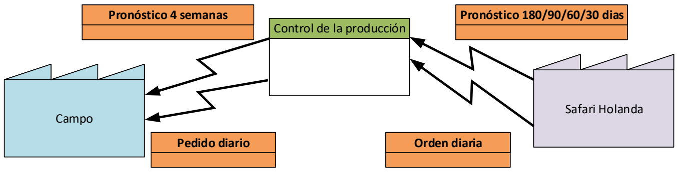


Fig.8 Data interaction in the supply chain.

In the picking area, the cycle time is 17 seconds, the setup time is 9 minutes, the process efficiency is 95%, the available time per working day is 28800 seconds per day, the planning time is one week, and the number of operators is 25. This data is shown in figure 9. The same procedure is performed for the cutting, packing, and cooling (cold storage) operations.

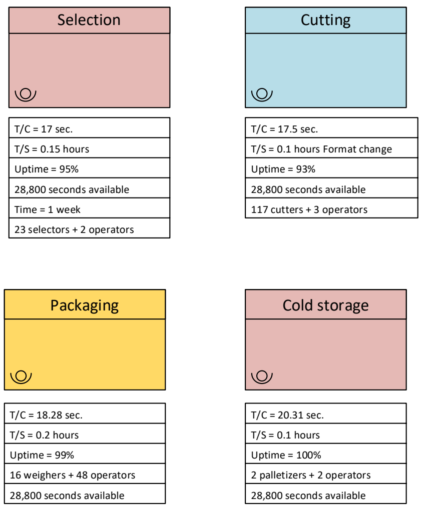


Fig.9 Information on each operation.

Between each operation in Figure 10, a triangle is observed, representing the amount of inventory flowing between them.

The amount of inventory is expressed in the days required for its execution. There is also a line at the bottom of the graph, which represents the time of added value and the time of unrealized value. Through this analysis, it was determined that for the production of a box of products, we have a production time of 6.57 days and an operating time of 94.31 seconds. Two trucks are also observed. The truck on the left side represents the reception of raw material that is sent by our suppliers, these shipments are made daily. The truck on the right side represents the shipments made by the company; these shipments are made every Monday. Table III shows the hourly salary of the personnel in the different areas of the company. We can see that most of the personnel receive the minimum agricultural salary.

Table III
Hourly wage per each operation

Capacidad		Equivalent ratio of one hour's pay to the value of one day's minimum wage.
Selection	Selection	16.28%
	Selection supplier	16.28%
	Selection weight control	16.28%
Cutting	Cutting supplier	16.28%
	Punnet cutter	18.60%
	Weighing of cut raw material	16.28%
Packing	Punnet coder	16.28%
	Punnet filling	18.60%
	Weighing of punnet	18.60%
	Punnet closing	18.60%
	Bulk bag filling	18.60%
	coder/weigher	16.28%
	USA Packaging - Selection/Weighing	23.26%
Cold storage	Box labeling	16.28%
	Stacking and palletizing	23.26%
	Storage	18.60%
Maintenance	Cleaning	16.28%
	Maintenance	20.93%

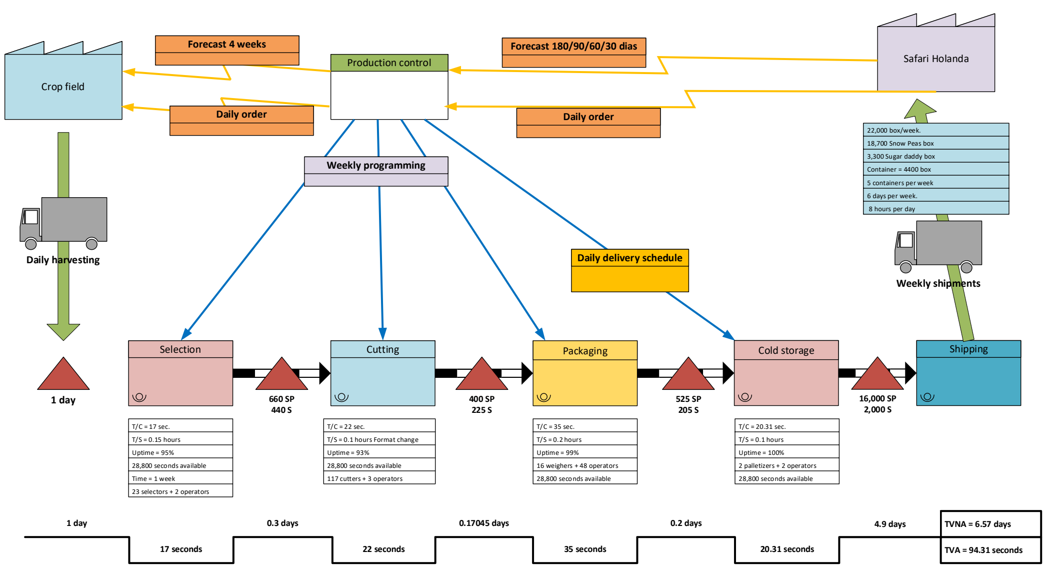


Fig. 10 VSM of the current situation.

To find the production cost ratio is calculated by multiplying each worker's salary by 8 hours and dividing by the raw material input, which gives us an average cost indicator of 1.22 soles per kilogram.

B. *Improved situation*

In this section, we will present the proposals that arise from what was mentioned in the previous point, where we investigated the use of tools to reduce or eliminate these causes so that the identified problem is partially or solved. Figure 11 shows the causes and the tool used to control the problem.

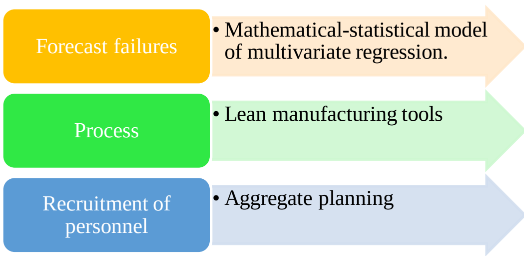


Fig.11 Improvement Proposals.

B.1 Phase 1: Forecasting method

The following questions were analyzed: Where should we cultivate? and why do yields per hectare vary from one field to another, if the same engineers are in charge of one field to another? Due to the aforementioned, the factors that impact the average yield in kilograms per hectare were evaluated. To validate the correct accuracy of the forecast, the 25 field lots that the company owns around Humay were analyzed. Among the main factors we have:

X_{Altura}**:** this is the height of the soil to be planted, concerning sea level. It is measured in meters above sea level.

$X_{Temperatura}$**:** the average temperature from planting to harvest, which is measured daily. It is measured in degrees Celsius.

$X_{Humedad\ Relativa}$**:** the relative humidity of the environment is measured from planting to harvest, and then averaged. It is measured in RH (%).

$X_{Variedad\ de\ Semilla}$**:** this is the type of seed to be used. All these seeds are certified seeds.

$X_{Densidad\ poblacional}$**:** is the number of plants planted in one hectare (usually there is a 5cm separation between plants, and a 1.10 separation between rows. However, out of one hectare of land, one hectare is not planted, because there are transit areas). Plants/ha are measured

$X_{Estación}$**:** is the season in which the sowing of the pea seed begins.

$X_{Tipo\ de\ Arveja}$**:** indicates whether the pea is Snow Peas or Sugar Snap.

$X_{Nutrientes}$**:** is the amount of N, P2O5, and K2O placed on one hectare of peas sown under normal conditions (no pests). It is measured in kg/ha

$X_{Tipo\ de\ suelo}$**:** Characteristic of the soil where it is sown. Peas adapt to a variety of soils except for very clayey soils. It prefers loam, clay loam, fertile, deep, and well-drained.

X_{Acidez}**:** the pea plant is moderately tolerant to acidity. It is requested in PH

$X_{Profundidad\ de\ siembra}$**:** Indicates the depth at which the seed of the pea plant was planted (cm).

$Y_{Tn/ha}$ **:** Yield per hectare measured in tn/ha. It is the response to the behavior of the previous factors.

To perform the multiple linear regression analysis, harvest data were collected from 75 fields (Humay, Huancano, Hualla, Auquis, and Quitasol), each of them analyzed with the probable factors that affect the yield per hectare of the pea plant. Subsequently, the behavior of the data is validated, analyzing that they do not present outliers, since this can distort the quality of prediction of the forecasting method, the outliers were calculated by the interquartile difference.

Subsequently, the regression model was calculated. The linear regression equation found uses the 11 variables and determines the model's coefficient of determination, R^2.

$$Tn/ha = -6841 + 1.03X_{Altura} + 974X_{Temperatura} + 91.7_{XHR(\%)} - 78X_{Variedad\_de\_Semilla} + 0.00025X_{densidad\_poblacional} - 1610X_{Estacion} - 1731X_{Tipo\_de\_Arveja} - 27.2X_{nutrientes} - 4444X_{Tipo\_de\_suelo} - 630X_{Acidez} + 2443X_{Profundidad\_de\_siembra}$$

Analysis of Variance

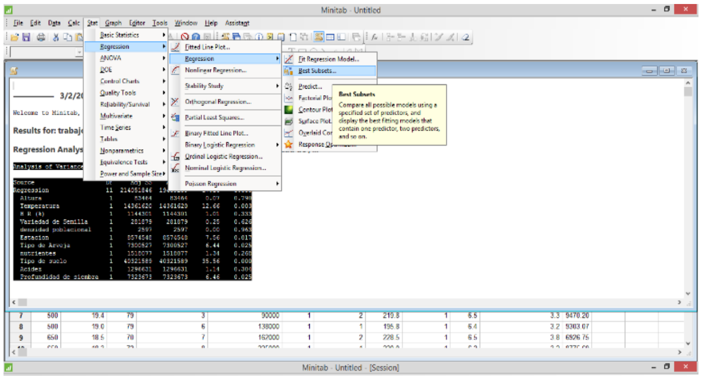
Source	DF	Adj SS	Adj MS	F-Value	P-Value
Regression	11	214051846	19459259	17.16	0.000
Altura	1	83464	83464	0.07	0.790
Temperatura	1	14361620	14361620	12.66	0.003
H R (%)	1	1144301	1144301	1.01	0.333
Variedad de Semilla	1	281879	281879	0.25	0.626
densidad poblacional	1	2597	2597	0.00	0.963
Estacion	1	8574548	8574548	7.56	0.017
Tipo de Arveja	1	7300527	7300527	6.44	0.025
nutrientes	1	1518077	1518077	1.34	0.268
Tipo de suelo	1	40321589	40321589	35.56	0.000
Acidez	1	1296631	1296631	1.14	0.304
Profundidad de siembra	1	7323673	7323673	6.46	0.025
Error	13	14742452	1134035		
Total	24	228794298			

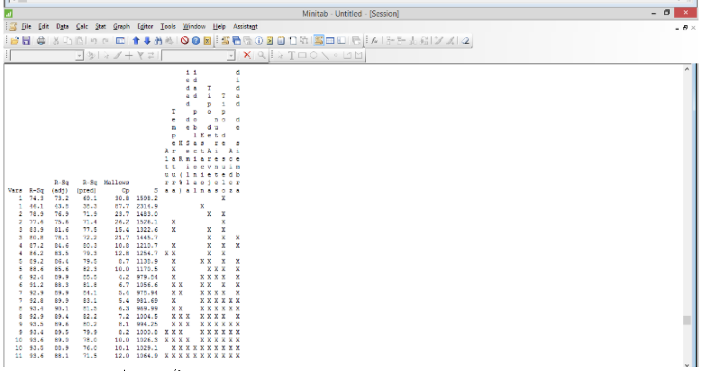
Model Summary

S	R-sq	R-sq(adj)	R-sq(pred)
1064.91	93.56%	88.10%	71.51%

In the report we observe that the P-value is less than α, therefore the null hypothesis Ho is rejected. Concluding that there is at least one factor that influences the model. However, with this analysis, we cannot assure that the model obtained is an appropriate model to predict the yield per hectare. Therefore, we need to carry out a more exhaustive analysis, looking for the correlation between the predictor variables. For this purpose, we performed the analysis of the best subsets of variables that predict the behavior of the variable tons/hectare. Figure 12 shows the analysis of the best subsets with the help of Minitab.

By performing the subset analysis, we can choose the 6, 7 or 8-variable model, depending on our statistical target value, either the coefficient of determination or Mallows' Cp. If we use the adjusted coefficient of determination to discriminate in a model, we observe that there is no significant difference between the 7-variable model and the 8-variable model. It is even better to select the 6-variable model since this model can explain 89.9% of the variability of the data (we should choose the highest value of R^2_{adj}).





Response is Tn/ha

Vars	R-Sq	R-Sq (adj)	R-Sq (pred)	Mallows Cp	S	Altura	Temperatura	HR (%)	Variedad de Semilla	densidad poblacional	Estacion	Tipo de Arveja	nutrientes	Tipo de suelo	Acidez	Profundidad de siembra
1	74.3	73.2	69.1	30.8	1598.2									X		
1	46.1	43.8	38.3	87.7	2314.9						X					
2	78.9	76.9	71.9	23.7	1483.0							X		X		
2	77.6	75.6	71.4	26.2	1526.1		X							X		
3	83.9	81.6	77.5	15.4	1322.6		X					X		X		
3	80.8	78.1	72.2	21.7	1445.7							X		X		X
4	87.2	84.6	80.3	10.8	1210.7		X					X		X		X
4	86.2	83.5	79.3	12.8	1254.7	X	X					X		X		
5	89.2	86.4	79.5	8.7	1138.9		X				X	X		X		X
5	88.6	85.6	82.3	10.0	1170.5		X					X	X	X		X
6	92.4	89.9	85.5	4.2	979.84		X				X	X	X	X		X
6	91.2	88.3	81.8	6.7	1056.6		X	X			X	X		X		X
7	92.9	89.9	84.1	5.4	978.94		X	X			X	X	X	X		X
7	92.8	89.9	83.1	5.4	981.69		X				X	X	X	X	X	X
8	93.4	90.1	81.5	6.3	969.99		X	X			X	X	X	X	X	X
8	92.9	89.4	82.2	7.2	1004.5		X	X	X		X	X	X	X		X
9	93.5	89.6	80.2	8.1	994.25		X	X	X		X	X	X	X	X	X
9	93.4	89.5	79.9	8.2	1000.8	X	X	X			X	X	X	X	X	X
10	93.6	89.0	78.0	10.0	1026.3	X	X	X	X		X	X	X	X	X	X
10	93.5	88.9	76.0	10.1	1029.1		X	X	X	X	X	X	X	X	X	X
11	93.6	88.1	71.5	12.0	1064.9	X	X	X	X	X	X	X	X	X	X	X

Fig.12 Analysis of best subsets of predictor variables

We do not use the R2 because the more explanatory variables we incorporate into the model, the higher the coefficient of determination, since the sum of squares of the regression decreases as the number of explanatory variables increases. Therefore, when we want to carry out a comparative analysis between several models, we use adjusted R^2, see equation 2.

$$\bar{R}^2 = 1 - \frac{n-1}{n-k}(1 - R^2) \qquad (2)$$

Likewise, if we discriminate using the Mallows Cp factor, we should look for the smallest value of the coefficient. Mallows' Cp helps to choose between several regression models. It allows to achieve of an important balance between the number of predictors in the model. So, we should look for the model where the Mallows Cp is small and very close to the number of predictors in the model, see equation 3.

$$C_p = \frac{SCE_p}{S^2} - N + 2P \qquad (3)$$

Where:

$$SCE_p = \sum_{i=1}^{N} (Y_i - Y_{pi})^2$$

It is the error of the sum of squares for the model with p regressors.
Ypi = predicted value of the i-th observation of Y of the P regressors.
S^2 = residual mean square after regression on the full set of K regressors and can be estimated by mean square error MSE.
N = sample size.

Therefore, the 6-variable model is chosen, which has a Mallow's Cp of 4.2. The two criteria converge in that it is convenient to use the 6-variable model. The regression model is detailed in equation 4.

$$y = \beta_0 + \beta_1 Temperatura + \beta_2 estación + \beta_3 tipo\ de\ arveja + \beta_4 nutrientes + \beta_5 tipo\ de\ suelo + \beta_6 profundidad \qquad (4)$$

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Regression	6	211512618	35252103	36.72	0.000
Temperatura	1	23742573	23742573	24.73	0.000
Estacion	1	8749161	8749161	9.11	0.007
Tipo de Arveja	1	8879475	8879475	9.25	0.007
nutrientes	1	7364430	7364430	7.67	0.013
Tipo de suelo	1	95110443	95110443	99.06	0.000
Profundidad de siembra	1	12843352	12843352	13.38	0.002
Error	18	17281680	960093		
Total	24	228794298			

Model Summary

S	R-sq	R-sq(adj)	R-sq(pred)
979.844	92.45%	89.93%	85.45%

With the model obtained, not only was it possible to determine a function to better forecast the raw material to be harvested in the field. The model also helped the field engineer to make decisions to influence the variables that affect the yield of the fields and improve production per hectare. After one year of study, it was observed that the fields yielded approximately 10 tons per hectare, thus improving the initial harvest by 30%. These data also had an impact on the quantity to be processed in the agroindustrial plant. The forecast regression is shown in equation 5.

$$Tn/ha = 738 + 906\,Temperatura - 1395\,Estacion - 1488\,Tipo\ de\ Arveja - 46.5\,nutrientes - 4817\,Tipo\ de\ suelo + 2670\,Profundidad\ de\ siembra \qquad (5)$$

B.2 Phase 2: Lean Manufacturing Tools

We start the proposal by asking ourselves: How do we increase productivity? What adds value to my production? We analyze each operation in detail.

***Process*:** When performing the value stream mapping shown in Figure 10, we observed that there are redundant operations, that is, operations that do not add value. The heart of this research lies in linking the cutting area with the packaging area.

It does not add value for a cutting worker to receive raw material, cut it, place it in a bale, then it transferred by the cutting supply personnel, weighed, then the bale is emptied onto a table, after which the filling personnel places the cut product on the punnets.

It adds value if a cutting worker receives raw material, cuts it and places the cut product on the punnets.

Technology as support: We observed that there are basic problems in two areas:

Selection: the productivity of the sorters is held back by removing white flowers that come with the snow peas, this is a critical process, given that if the flowers are not removed, the product could rot or decrease its shelf life (the flower perishes faster than the snow peas). For this reason, the selection belt was modified, incorporating a flower sucker, in addition to placing a tunnel with ultraviolet light to reduce the bacterial load and to remove the flowers in their entirety, which increased the flow of selected raw material per hour. Figure 13 shows the design of the selection belt.

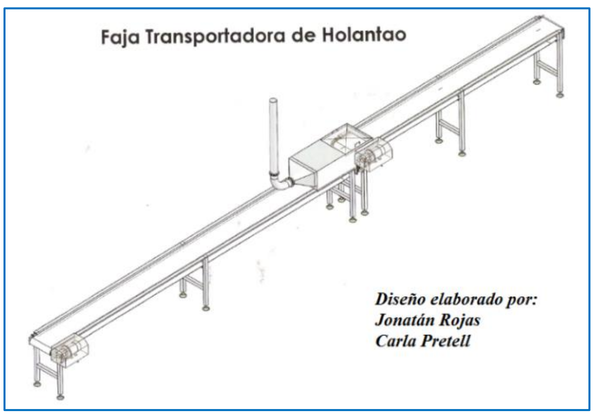


Fig.13 Improved sorting belt. With UV tunnel + flower sucker.

Cutting: After explaining that the cutting and packing areas must be joined, we also need to reconfigure the distribution within the plant. That is, if the cutting personnel now pack in punnets, how, where, and when will these punnets be weighed and sealed? The analysis determined that the supply of people transporting the trays with the cut product would produce shrinkage, or the punnet filling presentation would be disarranged so that the increased productivity in cutting was diminished by the subsequent operations. Given this, a cutting table was designed with a small belt in the center of the table, where the punnets would be placed so that at the end of the table they would be weighed, and at the end of the line they would be sealed and boxed. The design of the table with a conveyor belt can be seen in Figure 14.

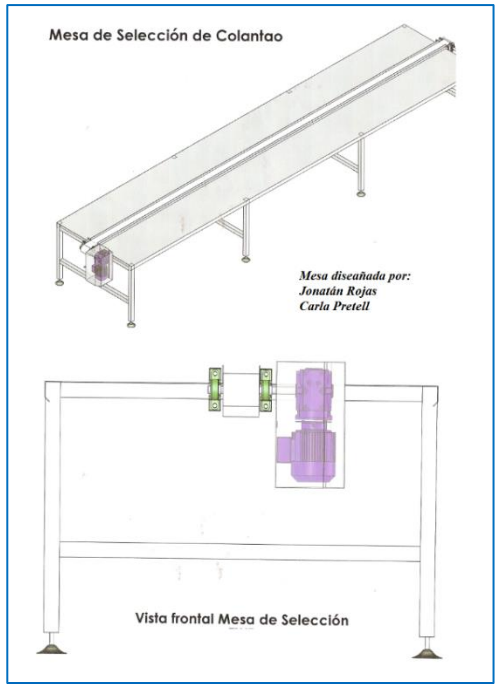


Fig.14 Modified cutting table with belt.

Table IV shows the new production activities, and it can also be seen that the cutting and packaging area is united, through continuous training and the search for improvements in the work stations, the productivity of each work station is improved. Table IV shows that workers with a 4-hour workday would already receive the minimum wage.

Table IV

New hourly wage based on process improvements

Capacity		Kg/hour-operator	Equivalent ratio of one hour's pay to the value of one day's minimum wage.
Selection	Selection	100	24.19%
	Selection supplier	750	24.19%
	Selection weight control	750	27.91%
Cutting + Packing	Cutting supplier	800	19.77%
	Punnet cutter	14	23.26%
	Weighing of cut raw material	150	18.60%
	Punnet coder	50	23.26%
	Punnet filling	150	23.26%
	Weighing of punnet	16	23.26%
	Punnet closing	150	18.60%
Cold storage	Box labeling	900	18.60%
	Stacking and palletizing	300	23.26%
	Storage	800	37.21%
Maintenance	Cleaning	4 personas	18.60%
	Maintenance	2 personas	32.56%

Finally, a cost analysis is performed and the new production cost indicator is validated to be US$ 0.38 per kilogram.

Figure 15 shows the new process plant layout that maximizes production flow.

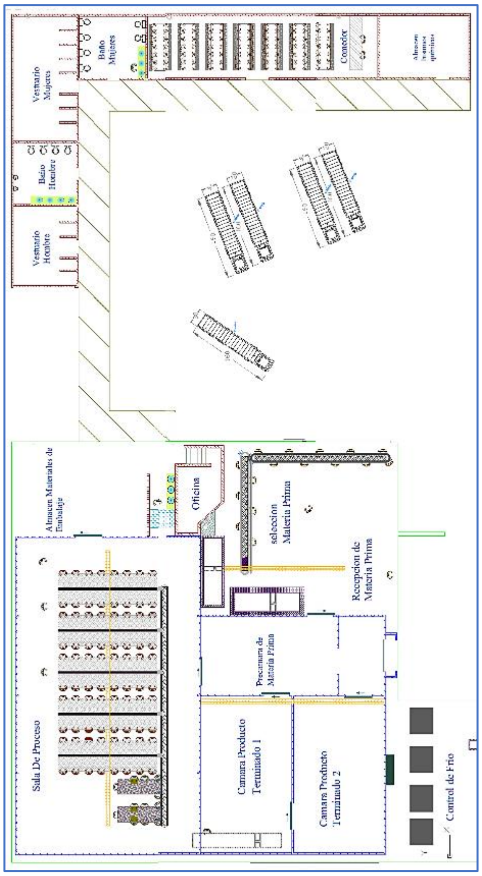


Fig.15 New layout of the production plant.

B.3 Phase 3: Aggregate planning - S&OP

Finally, having as input data a correct forecasting model, and having determined an optimal capacity of the production plant, the next step is to find the number of people in each period of the production campaign. For which we will apply mathematical modeling tools - linear programming to determine the labor force per work period, see table V and VI.

Table V

Forecast of the amount of raw material harvested.

Month	**Week**	SP	SS	**Total**
JUN	23	0	0	0
	24	0	0	0
	25	337.500	0	337.500
	26	1196.415	0	1196.415
JUL	27	2797.862	0	2797.862
	28	6867.923	0	6867.923
	29	10657.058	117	10774.058
	30	12553.718	292.5	12846.218
	31	13253.373	765	14018.373
AUG	32	13073.234	1912.5	14985.734
	33	11892.708	2529	14421.708
	34	10166.963	3865.5	14032.463
	35	10361.358	3811.5	14172.858

Table VI

Forecast of the amount of raw material harvested

Month	**Week**	SP	SS	**Total**
SEP	36	9913.748	4949.235	14862.983
	37	10894.545	4187.588	15082.133
	38	11267.438	3776.175	15043.613
	39	11275.308	4445.438	15720.746
OCT	40	10612.881	4423.320	15036.201
	41	10503.716	5168.115	15671.831
	42	10165.748	4789.058	14954.805
	43	8584.659	6230.588	14815.247
	44	8640.000	5624.235	14264.235
NOV	45	8572.500	6255.000	14827.500
	46	8212.500	4635.000	12847.500
	47	6750.000	3105.000	9855.000
	48	4410.000	1485.000	5895.000
DEC	49	1935.000	675.000	2610.000
	50	810.000	270.000	1080.000
	51	450.000	0	450.000
	52	180.000	0	180.000

The linear model in LINGO software is detailed below:

```
Minimize total cost = Normal time production cost + Overtime
production cost + Personnel hiring cost + Personnel firing
cost + Inventory cost.
Subject to:
month_i
Production and inventory:
Initial inventory + production x_i = Final inventory + demand
for the month _i
Number of persons to work during the month:
Persons mes_i - hires_i + layoffs_i = Persons mes_i-1
```

```
MODEL:
SETS:
PERIODO:DEMANDA,OPERARIOS,OPERARIOS_ADICIONALES,OPERARIOS_D
ESPEDIDOS,INVENTARIO_FINAL;
ENDSETS
DATA:
PERIODO=JUNIO JULIO AGOSTO SETIEMBRE OCTUBRE NOVIEMBRE
DICIEMBRE;
DEMANDA,COSTOTN,CCA,CCD,COSTOINV,COSTOTE,PRODUCCION,PORCENT
AJE=
@OLE('D\Tablas.xlsx');
@OLE('D\Tablas.xlsx')=
OPERARIOS,OPERARIOS_ADICIONALES,OPERARIOS_DESPEDIDOS,INVENTAR
IO_FINAL;
OPERARIO_AL_INICIO=212;
ENDDATA
!FUNCIÓN OBJETIVO;
MIN=@SUM(PERIODO(I):COSTOTN*OPERARIOS(I)+CCA*OPERARIOS_ADIC
IONALES(I)+CCD*OPERARIOS_DESPEDIDOS(I);


!ST;
!Inventario final, desde Junio a Diciembre;
@FOR(PERIODO(I)|I#GE#2#AND#I#LE#9:INVENTARIO_FINAL(I)=INVEN
TARIO_FINAL(I-1)+PRODUCCION*OPERARIOS(I)-DEMANDA(I));
!Condición de diciembre;
@FOR(PERIODO(I)|I#EQ#12:INVENTARIO_FINAL(I)=INVENTARIO_FINAL_DI
CIEMBRE);
!Personal de Junio a Diciembre;
@FOR(PERIODO(I)|I#GE#2#AND#I#LE#9:OPERARIOS(I)-
OPERARIOS_ADICIONALES(I)+OPERARIOS_DESPEDIDOS(I)=OPERARIOS(
I-1));
END [21]
```

The personnel required after applying the linear modeling is shown in Figure 16. It should be noted that the model aims to minimize total production costs, hiring costs, and firing costs.

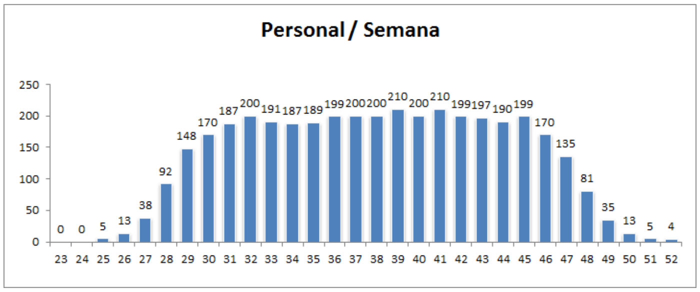


Fig.16 Personnel capacity level in a production plant.

IV. RESULTS

It is initially observed in the case of the deficiency in production planning, which started with a bad forecast and therefore ended with bad planning inside the production plant.

With the improvement proposals, the level of confidence in the forecast developed in the field was raised, and this model was applied in the company, bringing with it an improvement in both the level of prediction and in the field yield, that is to say, knowing the variables that affect the yield, making it possible to try to control them and therefore improve the yield of tons harvested per hectare planted. The confidence level increased to 90%, and field yields have increased to 12 tons per hectare.

The application of lean manufacturing tools also reduced the average process cost from approximately US$1.2 to US$0.38 per kilogram, and impacted production capacity, with an increase in production capacity from 5.6 tons to 14.5 tons per day with the same labor force. This improvement is considered the most important since it not only benefits the company but also the worker since his income is higher than the minimum wage.

Finally, with the application of the aggregate plan, we optimized on a workforce basis, i.e. only at the worker level, just the optimum number of workers to develop appropriate competencies within the company.

The application of the 3 phases of improvement is the ones that make possible the reduction of the production cost to a third of its initial cost.

V. CONCLUSIONS

This research was carried out over approximately two years. The improvement carried out with lean manufacturing tools was the most complex, since it consisted of changing the worker's mentality in quality and production improvement tools. Sessions were developed with them not only for training but also for talks entitled What prevents me from being better? How should we work? Are we quality agents? These sessions were developed with all members of the organization. Later when this improvement was executed, there were uncertainties in its implementation, for example, a worrying case was that the production improvement was not reflected in a reduction of costs, then doing the validations of the production runs it was determined that the improvement and impact are gradual since it is a process that is coupled to the learning curve.

The improvements made are small but had a significant impact on the company, reducing the cost by more than 60% was a revolution, which was shared with all members of the organization: board of directors, engineers, and operators.

REFERENCES

[1] Rojas, J. & Silva, W. (2014). Mejora en el abastecimiento y recolección de materia prima, aplicando heurísticas de Ruteo de Vehículos. Twelfth LACCEI Latin American and Caribbean Conference for Engineering and Technology
[2] INEI. (2019). PERÚ: Producto Bruto Interno por Años, según Actividades Económicas 2007 - 2019. https://www.inei.gob.pe/estadisticas/indice-tematico/pbi-de-los-departamentos-segun-actividades-economicas-9110/
[3] MINAGRI. (2010). Lineas de cultivos emergentes. https://www.minagri.gob.pe/portal/download/pdf/sectoragrario/agricola/lineasdecultivosemergentes/SAUCO.pdf.
[4] PROMPERU. (2010). PERÚ - Productos Naturales. Lima: Comisión de Promoción del Perú para la Exportación y el Turismo.
[5] MIDAGRI (2022). Estadística Agropecuaria - Herramientas Tecnológicas. https://siea.midagri.gob.pe/portal/
[6] Tavera, M., Davila, S. & Prado, A. (2009). "Asistencia técnica para los planes operativos (POS)" - PROYECTO UE-PERU/PENX. Código Ref. MINCETUR – POB 1.3.6.1.
[7] Nolasco, J. (2004). Evaluación de diferentes densidades de siembra de haba (Vicia faba L.) Como cultivo trampa para trips (Triphs sp.) EN EL CULTIVO DE ARVEJA CHINA (Pisum sativum L.) En la aldea Xeabaj, Santa Apolonia, Chimaltenango". Universidad de San Carlos de Guatemala.
[8] Gujarati, D. & Porter, D. (2010). Econometría. Mc Graw Hill.
[9] Yanchang Zhao (2014). R and Data Mining: Examples and Case Studies. Published by Elsevier.
[10] Ezequiel Uriel (2013). Regresión lineal múltiple: estimación y propiedades. Universidad de Valencia.
[11] Montgomery, D., and Runger, G. (2018). Applied statistics and probability for engineers. Arizona State University. John Wiley & Sons.
[12] Rahani, A. (2012). Production Flow Analysis through Value Stream Mapping: A Lean Manufacturing Process Case Study. Procedia Engineering, 41, 1727 – 1734.
[13] Womack, J. & Jones, D. (2012). Lean Thinking (Primera ed.). Barcelona, España: Grupo Planeta.
[14] Liker, J. (2004). The Toyota way. New York: McGraw-Hill.
[15] Hirano, H. (2002). Manual para la implantación del JIT: una guía completa para la fabricación "just-in-time". Portland, Oregón: Productivity Press.
[16] Domínguez Machuca, J. & Álvarez Gil, M. (1995). Dirección de operaciones. Madrid: McGraw-Hill.
[17] Orlicky, J., Plossl, G., & Orlicky, J. (2011). Orlicky's material requirements planning. New York: McGraw-Hill.
[18] Ivanov, D., Tsipoulanidis, A., & Schonberger, J. (2017). Global Supply Chain and Operations Management A Decision-Oriented Introduction to the Creation of Value. Springer International Publishing.
[19] Krajewski, L., & Malhotra, M. (2022). Operations Management: Processes and Supply Chains. Pearson Education Limited 2022.
[20] Jacobs, F., & Chase, R. (2018). Operations and Supply Chain Management. McGraw-Hill Education.
[21] Modelo propuesto por Christian Santos Cornejo Sánchez, Profesor de área de investigación de operaciones de la Pontificia Universidad Católica del Perú