https://dx.doi.org/10.22161/ijaers.68.30 *ISSN: 2349-6495(P) | 2456-1908(O)*

Bioethanol Production and Statistical Modeling from Fruit Residual Biomass Potential

Rhuan C. M. Ribeiro[1], Matheus S. Viana[2], Glauber T. Marques[3], Paulo C. Santos Júnior[4], Pedro S. S. Campos[5], Otavio A. Chase[6], J. Felipe Almeida[7]

[1–7] Federal Rural University of Amazonia (UFRA), Cyberspatial Institute (ICIBE), Belém-PA, Brazil
Email: {rhuan.ribeiro; matheus.viana; glauber.marques; paulo.junior; pedro.campos; otavio.chase; felipe.almeida}@ufra.edu.br

Abstract *— This work aims to describe an experimental procedure for the synthesis of bioethanol by the alcoholic fermentation of organic matter, from the use of discarded fruits. Based on the procedures performed on the statistical analysis of factorial experiments was used to verify the influence of the independent variables: the amount of must and fermentation time, in relation to yield response. The alcoholic fermentation was obtained from the pulp of apples (Malus communis) and tangerines (Citrus reticulata), as well as by microorganism (Saccharomyces cerevisiae). As a result, the maximum yield value was around 17.5% v.v^{-1}, which gives fruit residues a high potential for use in bioethanol production. The statistical evaluation was used to optimize the input condition and the value of 19.06% v.v^{-1} has been estimated. Thus, this text presents a model of economic viability and its environmental importance due to the use of organic waste.*

Keywords *— **Biofuels, Experimental production, Residual biomass, Statistical analysis.***

I. INTRODUCTION

Nowadays, the world's energy resources are expanding, due to the increasing demands of the capacity to transform the productive factors combined with the continuous change in consumption habits and in the face of the increase of the world population on Planet Earth. However, specialized agencies warn of the depletion of traditional energy resources, and more than that, they warn of the danger of the continued use of traditional energy sources based on fossil sources (AKKARI, RÉCHAUCHÈRE, et al., 2018). These intensifications of production mean overburdening the environment, which is clear evidence of its contribution to global imbalances. In other words, overloading energy resources means overburdening all other resources on the planet, exerting significant pressure on the use of common goods in the economy (ZILBERMAN, 2017). Therefore, guaranteeing energy efficiency and at the same time not overloading the environment is the problem to be faced by the world in the coming years, thus evidencing the search for renewable energies that cause fewer negative externalities to the environment (MME, 2015).

The use of renewable energy, especially based on solar photovoltaic, wind and biomass sources, corresponds to a technological innovation that breaks the existing paradigm because it is a new method of production, sustainable and non-aggressive of the environment. In this context comes bioethanol, a liquid biofuel derived mainly from renewable biomass that presents some important differences in relation to petroleum-derived fuels. Among these are the high oxygen content, which accounts for about 35% by weight of bioethanol, low toxicity and high biodegradability, which in general allow for cleaner combustion and increased engine performance (BUCKERIDGE e SOUZA, 2017).

Biofuels are classified according to their process of production and can be qualified as first or second-generation biofuels. The first generation being produced from the raw material saccharide and starch, such as sugarcane, corn, and soybeans, also used as food for humans and animals, second-generation biofuels come from industrial waste, agricultural waste or urban waste (BUCKERIDGE e SOUZA, 2017; BAJPAI, 2013). Among the main advantages of second-generation bioethanol production is the absence of a threat to food production, as well as the agricultural land (BAJPAI, 2013). The synthesis may be derived from the enzymatic hydrolysis of polysaccharides contained in lignocellulosic matter, followed by the fermentation of fermentable sugars (EHTESHAMI, VIGNESH, et al., 2016). However, these three processes can be performed independently: SHF - separate hydrolysis and fermentation; combined SSF - simultaneous

https://dx.doi.org/10.22161/ijaers.68.30 *ISSN: 2349-6495(P) | 2456-1908(O)*

saccharification and fermentation; or SSCF - simultaneous saccharification and co-fermentation (BAJPAI, 2013).

The organic waste to produce biofertilizers used in agriculture, and obtaining biogas and bioethanol for energy cogeneration, is expressed by different methodologies for this purpose, in specialized literature. Azevedo et al. (2007) evaluated the production of bioethanol from the persimmon juice and observed that the factors of the initial concentration of inoculum, soluble solids and initial pH did not influence the alcoholic fermentation of the organic matter. In the study presented by Lima et al. (2015) in terms of efficiency and yield of the process of obtaining cellulosic ethanol by means of the alcoholic fermentation of the hydrolyzed liquors with the use of the yeast *Saccharomyces cerevisiae*. These authors obtained maximum yield and efficiency values of 0.445g of ethanol/g of bagasse and 87.1% for the hydrolyzed liquor with the addition of cashew juice. Ylitervo (2008) used dried oranges peels to produce bioethanol by enzymatic hydrolysis with the application of the *Mucor Indicus* fungus, hydrolyzing and converting into sugars, in the bioethanol production by two fermentable sugar process, and has been obtained a yield of 0,36g/g after 24 h.

In contrast, a large amount of biomass residues with high potential for energy generation is wasted each year, because about 35% of Brazil's agricultural production goes to waste and, consequently, the country is among the ten countries that more wasted food (GOULART, 2008). A study by the Food and Agriculture Organization (FAO) estimates that each year about 1.3 billion tons of food is wasted or lost in the world, equivalent to one-third of global production. In fact, apples alone are 3.7 trillion units annually (EMBRAPA, 2017; FAO, 2013).

All processing of fruit and vegetables produce waste which, when treated for recovery, is generally used as fertilizer or animal feed. However, with the growing worldwide concern about preservation and environmental impacts, more sustainable means of agricultural production and alternative energy generation are sought. With this, several biotechnological processes have been developed to transform agroindustry waste into products of economic value, adding usefulness to a disposal item (BUCKERIDGE e SOUZA, 2017; BRUNELLE, DUMAS e SOUTY, 2014). In this context, we have been proposed an experimental bioethanol alcoholic fermentation process and to validate this synthesis by statistical optimization, aiming the reuse of discarded fruits.

II. MATERIALS AND METHODS

The organic material was collected on July 15, 2018, where 7 kilograms of fruit were collected (Fig. 1). The fermented fruit was produced on a bench scale, and for the production, procedures were used a batch reactor with a capacity of 500 ml. Else, a reflux distillation processor was used with independent temperature controllers and water flow condensation chamber, in addition to the transfer of manually controlled internal stages. For these experiments we used: apples (*Malus Communis*) e tangerines (*Citrus reticulata*); yeast (Biological yeast); sucrose; and, distilled water.

Fig. 1: Organic material collected

2.1 The preparation of the substrate (must)

The preparation process begins with the washing of the fruits and the extraction of the juice (must) containing water, alcohol and the solid by-products was obtained with the aid of a domestic blender used to grind and homogeneously mix the matter.

Fig. 2: The fermentative must in preparation

To expand the scale for the works in the future, the microorganism *Saccharomyces cerevisiae* was used because it is the most economically viable. The planning indicated a volume of around 6 (six) liters needed for the development of the ten trials. So, in the process of must preparation, we used 5 liters fruit pulps mixing 1-liter distilled water. After this, we also add 500 g sucrose to favor the more rapid growth of the microorganisms, the number of yeasts was added in the concentration of 10 $g.L^{-1}$, totaling 60 g. This solution was mixed and transferred to the storage containers.

2.2 Fermentation

In this stage, the conversion of sugars into ethyl alcohol (ethanol) and carbon dioxide (CO_2) occurred. The

https://dx.doi.org/10.22161/ijaers.68.30 *ISSN: 2349-6495(P) | 2456-1908(O)*

process has been completed by storing the must in a coupled vessel of a three-piece vacuum chamber, subjected to an ambient temperature of 28°C, and containing a valve that was used to collect the must according to the fermentation time previously determined. The fermented samples were collected at intervals in the range of 33.1 to 134.9 hours under fermentation conditions.

2.3 Distillation

At first, the fermented mixtures were filtered to separate any solid residues from the liquid. After this, the broth was poured into the distillation boiler, where an element inside the vessel heated the substance so that the alcohol boiled and rose in the form of steam through the distillation column, all this process was done with the aid of the panel GT-6000 from Marcraft. To maximize the alcohol content of the samples the vessel temperature was controlled and remained between the boiling points of alcohol and water, 78.3 and 100°C, respectively, with a total process time of about 4 hours.

2.4 Calculation of the variable studied

The yield calculation was developed from Equation (1). Where v_{EP} represents the volume of ethanol produced and v_{MF} is the volume of the fermented must.

$$\text{Yield} = V_{EP} / V_{MF} \times 100 \quad (1)$$

2.5 Statistical methods

Systematic sequencing of factorial experiments, from the levels of the variables worked, were specified, applying a planning 2^2 with four factorial points, counting on four more axial points and two central points, totalizing 10 trials.

For the analysis of the results, we have been used the computational software's MINITAB® v.18 and STATISTICA® v.10. By means of these can be possible to obtain the multiple linear regression, based on the experimental plan (central planning), the need to adjust the model and its adequacy through analysis of variance (ANOVA), to validate the relationship between the variables. We also verified the hypotheses of normality and constant variance (homoscedasticity) of the residues, in order to admit if the proposed model satisfies them. The generation of the response surface graphs was suggested for a better visualization of the effects found and verification of the optimal point revealed in the experiments that represents a maximum value in the response. Based on the optimization of the process, the ideal conditions for input variables were established, being able to estimate the maximum yield (%) $v.v^{-1}$.

III. RESULTS AND DISCUSSION

Table 1 shows the real and coded levels of input variables obtained with the planning.

Table 1: Shows the real and coded levels of input variables obtained with the planning.

Variables	Level - 1,414	Level -1	Median	Level +1	Level + 1,414
Quantity of must (mL)	337,86	400	550	700	762,13
Fermentation time (hours)	33,1	48	84	120	134,9

Once these conditions were maintained, the Central Composite Planning (CCP) for each variable was elaborated in two levels of work plus two points (alpha), defined as points of extreme variation. So, each variable is studied whereas three previously defined levels (-1, 0, +1) and two other levels added (–α and +α), such parameters were used in order to obtain a delineation, where the variance and covariance matrix is diagonal and the estimated parameters are not correlated with each other (SMUCKER, KRZYWINSKI e ALTMAN, 2019). In Table 2, is presented the planning matrix with the coded values of two input factors and answers acquired on the accomplishment of the experiments in relation to the levels studied.

Table 2: Factorial design matrix 2^2 with coded values and experimental yield data (%) $v.v^{-1}$ of production

Trails	QM (mL) (coded)	TF (h) (coded)	R (%) $v.v^{-1}$
01	-1	-1	13,3
02	-1	+1	13,7
03	+1	-1	16,6
04	+1	+1	17
05	-1,414 (-α)	0	11,5
06	1,414 (+α)	0	17,5
07	0	-1,414 (-α)	13,3
08	0	1,414 (+α)	13,5
09 (PC)	0	0	10,2
10 (PC)	0	0	10,2

QM – Quantity of must; TF – Fermentation time; R – Yield; PC – Central Point.

According to the Policy N° 64/2008 (BRAZIL, 2008), fermented fruit is the liquid with an alcoholic strength in the range of 4,00% to 14,00% in volume and 20°C. These values are directly proportional to the yield (%) $v.v^{-1}$ of production. The highest value of alcoholic yield found in

https://dx.doi.org/10.22161/ijaers.68.30 *ISSN: 2349-6495(P) | 2456-1908(O)*

this study was 17,5% $v.v^{-1}$ (experiment 6). Therefore, is necessary to analyze the percentage concentration of alcohol obtained in the final product, which has been not considered in the present study, and will also be possible to attribute the high alcohol content to these conditions. Result higher than found by Dantas & Silva (2017) that used the alcoholic fermentation of *Spondias tuberosa* and the yeast *Saccharomyces cerevisiae* has been found alcoholic yield of 12,54% $v.v^{-1}$.

In this work, the maximum result (17,5% $v.v^{-1}$) is much higher than that found by Gomes, Lima, Rabelo, Oliveira & Silva (2010) also for fermented of *Spondias tuberosa*, being 11,6% $v.v^{-1}$ of alcohol. However, being much lower than found by Pereira, Gallina, Banczek, Maia & Rodrigues (2016) that used *Cyperus esculentus* as raw material and the enzymatic route for the hydrolysis, found at the maximum of their experiments the value of 29,08% $v.v^{-1}$ of alcohol. The variations of these results are due to the fact of the different processes, types, and concentration of the inoculum, as well as the temperature used in the experiments, among others.

When a Multiple Linear Regression Model (MRLM) is proposed to statistically evaluate a certain procedure, one of the objectives is to find the simplest and best-presented model. For a better analysis, it is necessary to consider that by adding more terms in the systematic part of the model, we increase the Quadratic Sum of Regression (QSReg), remaining to know if such increase is significant, in order to verify if the addition of the extra terms contribute to the best description of the variable studied (ALTMAN e KRZYWINSKI, 2015). However, it is necessary to request an ANOVA table corresponding to the reduced model (Table 3) in comparison with the table of the complete model, presented in Table 4.

Table 3: ANOVA short form (no additional terms)

Source	Degree of freedom (GF)	Quadratic sum (QS)	Root Mean Square (RMS)	Test F	*p*-Value
Regression	2	28,5921	14,2961	2,87	0,123
QM (linear)	1	28,4456	28,4456	5,71	0,048
TF (linear)	1	0,1466	0,1466	0,03	0,869
Error	7	34,8439	4,9777	-	-
Lake of fit	6	34,8439	5,8073	-	-
Pure error	1	0,0000	0,0000	-	-
Total	9	63,4360	-	-	-
(%) Explained variation	R^2	R^2 (Adjusted)			
	45,07	29,38			

The ANOVA of this model shows an R^2 value than expected for predictive models and the values-p found is above the set level of significance ($\alpha = 0,005$). Also, it is noted high value of the lack of adjustment, allowing the understanding that such model needs adjustments as the studied variable. Table 4 presents the model by adjusted configuration submitted to additional terms.

Table 4: ANOVA complete model (with the addition of terms)

Source	Degree of freedom (GF)	Quadratic sum (QS)	Root mean Square (RMS)	Test F	*p*-Value
Regression	4	60,0821	15,0205	22,39	0,002
QM (linear)	1	21,7510	21,7510	32,43	0,002
TF (linear)	1	15,5978	15,5978	23,25	0,005
QM (quadratic)	1	27,4400	27,4400	40,91	0,001
TF (quadratic)	1	16,5069	16,5069	24,61	0,004
Error	5	3,3539	0,6708	-	-
Lake of fit	4	3,3539	0,8385	-	-
Pure error	1	0,0000	0,0000	-	-
Total	9	63,4360	-	-	-
(%) Explained variation	R^2	R^2 (Adjusted)			
	94,71	90,48			

https://dx.doi.org/10.22161/ijaers.68.30 *ISSN: 2349-6495(P) | 2456-1908(O)*

The adjusted model can be used for predictive purposes since the descriptive values-p are below the fixed significance level and the quadratic mean of the lack of adjustment (0.8385) allows to state that the model does not present statistical evidence for to make a new adjustment. The coefficient of determination (R^2 adjusted) was satisfactory at a significance level of 95%, indicating the high significance of the model studied. The adjusted R^2 value (90.48) means that 90.48% of the total variation around the mean is explained by the regression (ALTMAN e KRZYWINSKI, 2015; CHARNET, FREIRE, et al., 2008). Therefore, it is understood that there is no need to formalize a hypothesis test to compare Quadratic Sum of Regression (QSReg), due to the perceptible improvement of the model after an additional term. The respective statistical coefficients of the empirical mathematical model are described by Equation (2), where the values in bold correspond to the statistically significant parameters.

$$\text{Yield } (\%) = \mathbf{46{,}26} - \mathbf{0{,}1072\ QM} - 0{,}2427\ TF + \mathbf{0{,}000109\ QM^2} + \mathbf{0{,}001467\ TF^2} \quad (2)$$

In Equation (2), the linear term of the fermentation time variable is not in bold, because the value-p of this term was below the measure of the significance level set (α = 0,005), indicating a negligible contribution of this term to the model.

In fact, once these formulations have been established, it is necessary to evaluate the suitability of the adjustment made. For this propose, we examined the distribution of the residual values, calculated as the difference between the predicted values, according to the current model, and the observed values, as can be seen in Fig. 3. An analysis of the normality and constant variance (homoscedasticity) assumptions of residues is also important. Thus, Fig. 4 and Fig. 5 were also proposed for better visualization of these results.

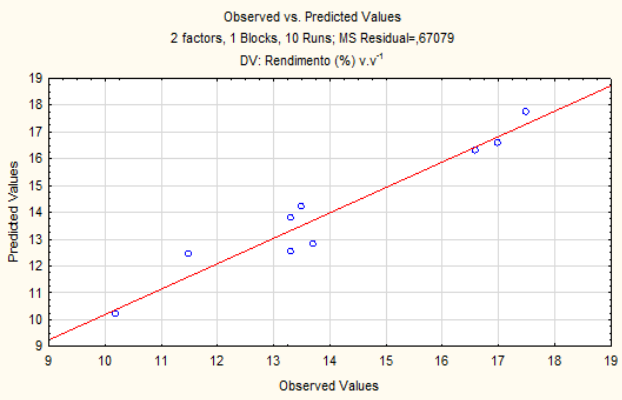

Fig. 3: Graph of predicted values vs. observed values

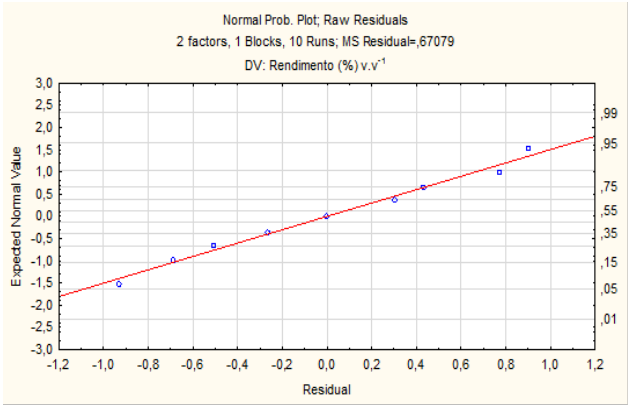

Fig. 4: Normal probabilistic residues diagram

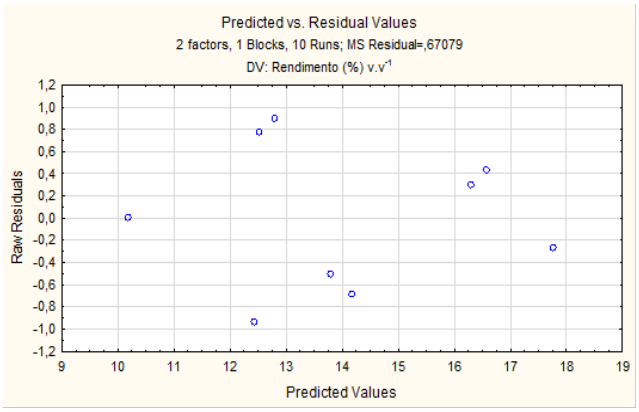

Fig. 5: Diagram of the residual's vs predicted values

By analyzing the previous graphs, we can present that the set of observed values (model residuals) are very close to following a theoretical distribution. In the visual interpretation of Fig. 3 and Fig. 4, it is observed that the points are distributed close to the straight identity. And the residues also meet the assumptions of constant variance (homoscedasticity), as seen in Fig. 5 (BIAZUS, SOUZA, et al., 2005).

Based on the results obtained, it is relevant to examine the surface and contour plot of the dependent variable as a function of the factors. The surface is satisfactorily described by Equation (2), which is shown in Fig. 6. This verification allows obtaining the regions of greatest (maximum) and minor (minimum) results of the studied response.

https://dx.doi.org/10.22161/ijaers.68.30 *ISSN: 2349-6495(P) | 2456-1908(O)*

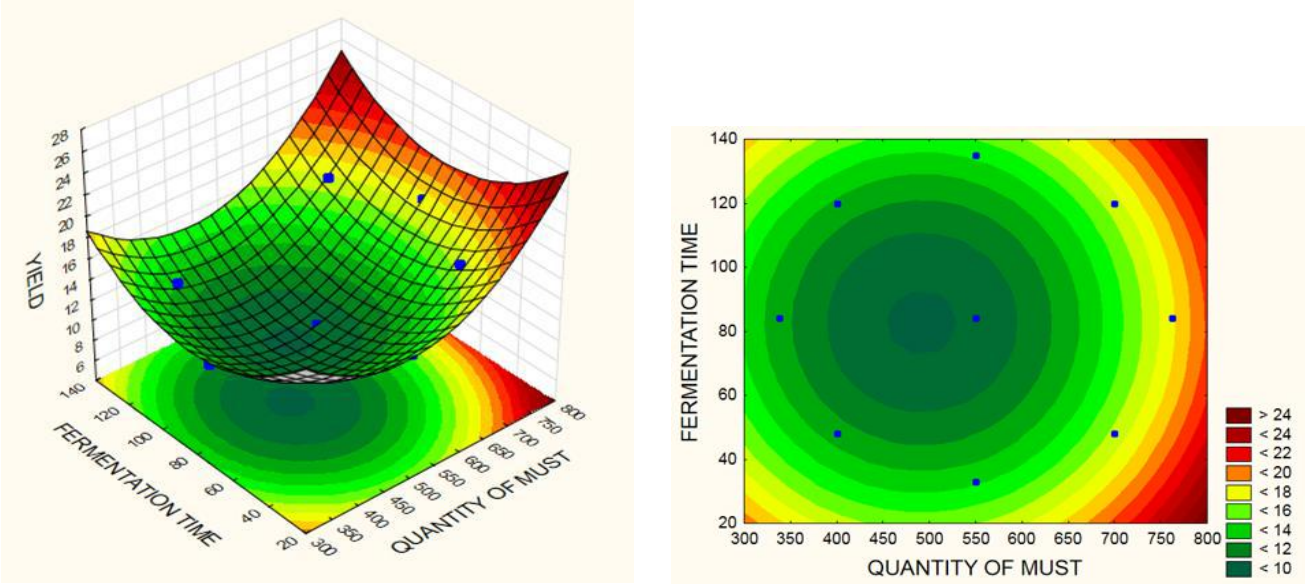

Fig. 6: Response and boundary surface: yield (%) v.v-1 (%) depending on the quantities of must and fermentation time

Fig. 6 shows the highest percentage yield value found (17,5% $v.v^{-1}$), maintaining a quantity of must at the level 1,414 (+α), which is equivalent to 762,13 mL and fermentation time ratio corresponding to 84 hours (center point), coded values of Table 2 (experiment 6) and quantified in Table 1, are the levels that maximize the bioethanol experimental production of this work.

According to the specialized literature (BOTHAST e SCHLICHER, 2005; NICHOLS, MONCEAUX, et al., 2008; QUINTERO, MONTOYA, et al., 2008), the fermentation requires a total time of 48 to 72 hours and reaches a final ethanol concentration of 10-14 % 46 $v.v^{-1}$.

In addition, it is emphasized that the samples were exposed to the flame test, which consists of setting fire to the obtained solution to verify if it starts with combustion with a certain facility. All the tests were successful and have been observed that the alcoholic concentration of the bioethanol solutions was at considerable concentration levels. Fig. 7 presents the optimized condition of the bioethanol production process using the applied methodology.

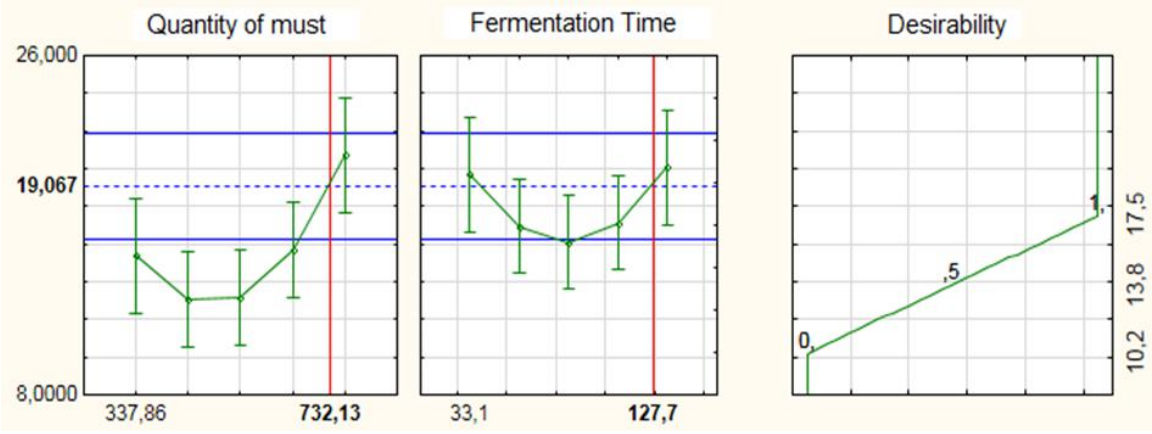

Fig. 7: Optimum yield values of must (mL) and fermentation time (h)) optimized for yield (%) $v.v^{-1}$ in the bioethanol production process

The optimal conditions determined in the optimization of this process were 732.13 (mL) for the quantity of must and 127.7 (hours) of fermentation time. The independent variables in such optimized conditions (according to Fig. 8), presented an approximate yield value of 19,07% $v.v^{-1}$.

In this sense, the conditions of these variables are within the range of experiments performed (line projected in green), demonstrating that the optimization occurred successfully.

https://dx.doi.org/10.22161/ijaers.68.30 *ISSN: 2349-6495(P) | 2456-1908(O)*

IV. FINAL CONSIDERATIONS

The use of apples (*Malus communis*) and tangerines (*Citrus reticulata*), as raw material to obtain bioethanol from the alcoholic fermentation process, presented high economic viability and significant potential from an environmental point of view for the use of organic residues. Since these fruits were collected directly from the garbage collection. The result consistent with the experiment, report that such raw materials can be used to produce bioethanol and this conclusion is based on the high 17,5% $v.v^{-1}$, which confers to the fruit residues a competitive potential considering ethanol coming from the sugar cane, however without additional expenses for agricultural inputs, such as planting, harvesting, transportation, among others. The implementation of a factorial planning in the study of the production was efficient and relatively simple to use as a strategy of optimization and analysis about this process, being able to be considered relevant to the development of researches as to the use of products or byproducts derived from industrial sources or which can be processed for clean and renewable energy generation.

ACKNOWLEDGEMENTS

The authors would like to acknowledge CNPq, CAPES and UFRA for supporting this work

REFERENCES

[1] AKKARI, M. E. (2018). et al. A meta-analysis of the greenhouse gas abatement of bioenergy factoring in land use changes. Scientific Reports, v. 8, pp. 1-8.

[2] ALTMAN, N.; krzywinski, M. (2015). Simple linear regression. Nature Methods, v. 12, n. 11, pp. 999-1000.

[3] AZEVEDO, A. W. M. et al. (2007) Viability Study of Ethanol Obtention from Kaki (Saccharomyces cereviseae) Fruit Must.

[4] BAJPAI, P. (2013). Advances in Bioethanol. [S.l.]: Springer.

[5] BIAZUS, J. P. M. et al. (2005). Optimization of Drying Process of Zea Mays Malt to Use as Alternative Source of Amylolytics Enzymes. Brazilian Archives of Biology and Technology, v. 48, pp. 185-190.

[6] BOTHAST, R. J.; SCHLICHER, M. A. (2005). Biotechnological processes for conversion of corn into ethanol. Applied Microbiology and Biotechnology, v. 67, n. 1, pp. 19-25.

[7] BRUNELLE, T.; DUMAS, P.; SOUTY, F. (2014). The Impact of Globalization on Food and Agriculture: The Case of the Diet Convergence. Journal of Environment & Development, v. 23, n. I, pp. 41–65.

[8] BUCKERIDGE, M. S.; SOUZA, A. P. D. (2017). Advances of Basic Science for Second Generation Bioethanol from Sugarcane. [S.l.]: Springer.

[9] CHARNET, R. et al. (2008). Analysis of Linear Regression Models. 2ª. ed. Campinas: Unicamp.

[10] DANTAS, C. E. A.; SILVA, J. L. A. (2017). Fermented Umbu alcohol: Production, fermentation kinetics and physical-chemical characterization. Holos, Natal, v. 2, pp. 108-121.

[11] EHTESHAMI, S. M. M. et al. (2016). Numerical investigations on ethanol electrolysis for production of purê hydrogen from renewable source. Applied Energy, v. 170, pp. 388-393.

[12] EMBRAPA. (2017). Brazilian Agricultural Research Corporation. Government of Brazil.

[13] FAO. (2013). Food wastage footprint: Impacts on natural resources. Food and Agriculture Organization. Roma, pp. 63.

[14] GOMES, E. M. S. et al. (2010). Production of alcoholic fermentation from Umbu pulp. IFAL Scientific journal, v. 1, n. 1, pp. 59-65.

[15] GOULART, R. M. M. (2008). Food Waste: A Public Health Problem. Integração, São Paulo, p. 285-288.

[16] LIMA, E. E. et al. (2015). Production of second generation ethanol from the bagasse of cashew nuts. Caatinga, Mossoró, v. 28, n. 2, pp. 26-35.

[17] BRAZIL - MINISTRY OF AGRICULTURE, LIVESTOCK AND SUPPLY. (2008). Ordinance No. 64, April 23, 2008. Official Diary of the Federative Republic of Brazil.

[18] MME. (2015). Ministry of Mines and Energy. Government of Brazil.

[19] NICHOLS, N. N. et al. (2008). Production of Ethanol from Corn and Sugarcane. In: J. W. E. Al. (Ed.). Bioenergy: ASM Press, Washington, DC.

[20] PEREIRA, S. C. L. et al. (2016). Study of the production of alcohol using Cyperus esculentus (Tiririca) as raw material. Virtual Journal of Chemistry, v. 8, n. 5.

[21] QUINTERO, J. A. et al. (2008). Fuel ethanol production from sugarcane and corn: Comparative analysis for a Colombian case. Energy, v. 33, n. 3, pp. 385-399.

[22] SMUCKER, B.; KRZYWINSKI, M.; ALTMAN, N. (2019). Two-level factorial experiments. Nature Methods, v. 16, pp. 209–212, march.

[23] YLITERVO, P. (2008). Production of ethanol and biomass from orange peel waste by mucorindicus. University College of Borås, Borås, pp. 1-61.

[24] ZILBERMAN, D. (2017). Indirect land use change: much ado about (almost) nothing. GCB Bioenergy, v. 9, pp. 485–488.