Systems & Control Transactions

Research Article - Peer Reviewed Conference Proceeding
ESCAPE 35 - European Symposium on Computer Aided Process Engineering
Ghent, Belgium. 6-9 July 2025
Jan F.M. Van Impe, Grégoire Léonard, Satyajeet S. Bhonsale, Monika E. Polańska, Filip Logist (Eds.)

PSE PRESS

Computer-Aided Design and Optimization of Lycopene Production Process from Tomato Waste

Nereyda Vanessa Hernández-Camacho [a], Fernando Israel Gómez-Castro [a*], Mariano Martín [b], Ehecatl Antonio del Rio-Chanona [c] and Oscar Daniel Lara-Montaño [d]

[a] Universidad de Guanajuato, Campus Guanajuato, División de Ciencias Naturales y Exactas, Departamento de Ingeniería Química, Guanajuato, Guanajuato, Mexico.
[b] Universidad de Salamanca, Departamento de Ingeniería Química, Salamanca, Spain.
[c] Imperial College London, Centre for Process Systems Engineering (CPSE), Department of Chemical Engineering, London, United Kingdom.
[d] Universidad Autónoma de Querétaro, Facultad de Ingeniería, Juriquilla, Querétaro, Mexico.
* Corresponding Author: fgomez@ugto.mx

ABSTRACT

The extraction of lycopene from tomato waste has been largely evaluated at an experimental level, leading to the creation of polynomial models or response surfaces that allow the representation of the extraction behavior. However, these studies are based on laboratory level and an extraction process has not yet been scaled up. This study evaluates the design and optimization of the lycopene extraction process from tomato waste. The proposed model is solved through a link between Python and Aspen Plus, performing the optimization a genetic algorithm (GA) in Pymoo. The minimum value of TAC is 211,692.2 USD/yr, corresponding to a production of 2.29 g/h of lycopene, starting from 1000 kg/h of tomato waste. This work represents a first approach to the design of a commercial-scale lycopene production process.

Keywords: tomato waste, lycopene, solvent extraction, stochastic optimization.

INTRODUCTION

Approximately 13.2% of the world's food is lost before being harvested [1]. Thus, food waste management has high relevance. Particularly in Mexico, the agricultural sector contributes with 3.4% of the national GDP, where the fruits and vegetables sector stand out with 45% of the sector's exportations. Tomatoes contributes to exportation with 8.41%. Mexico is the main supplier of this product worldwide, with a 19% share of the exported volume in the period 2003-2017 [2]. However, not all tomatoes are used due to their short shelf life, causing approximately 30% of the crop to be wasted [3]. In addition, waste from tomato sauce production can also be valorized. Tomato waste can be used to isolate high-value compounds such as carotenoids, polyphenols, vitamins, fibers, flavonoids, among others. Among these, lycopene is considered a well-known carotenoid and the most abundant pigment in tomatoes, responsible for giving them their color. It is used as a raw material in the cosmetic, pharmaceutical and food industries [4]. Lycopene from tomato waste is traditionally recovered by solvent extraction, but research on this topic has only been carried out through laboratory-scale experimentation [5, 6]. An important aspect in lycopene extraction is the type of solvent used in the extraction, as this can have economic, environmental or health drawbacks, depending on the application for which it is used. Some studies work with traditional solvents [6,7], others use enzymes that help with greater extraction [5,8], and others work with supercritical fluids [9]. Various alternatives have been proposed, but all of them have been developed at laboratory level, without scaling up.

This work addresses the systematic design and optimization of an industrial-scale process for obtaining lycopene from tomato waste. A comparison is made between a base case and the optimized design using a conventional solvent mixture: acetone-n-hexane. A stochastic single-objective optimization strategy is employed to minimize to total annual cost (TAC) of the proposed process.

METHODOLOGY

Process Design

Tomato waste is used as raw material to produce lycopene. This waste is passed through a tray-type dryer, reducing the water content to 10%. The waste is then subjected to a mill until it reaches an average particle size of 78 mm. Lycopene is then extracted from tomato waste using an acetone-n-hexane solvent mixture in a stirred tank reactor. The lycopene rich stream passes to a flash tank to separate large volumes of solvents and concentrate the lycopene. Finally, the lycopene stream is purified in a distillation column to achieve 95% purity, food grade. The process is summarized in Figure 1.

Mathematical modeling

The mathematical model of the process consisted of the design equations for the drying, grinding and extraction stages, which have been coded in Python, linked to the simulation in Aspen Plus for the separation and purification stages (flash tank and distillation column).

For the drying stage the main equations are given by eqs. (1), (2) and (3) [10]:

$$\frac{q_T}{\dot{m}_s} = c_{ps}(T_{sb} - T_{sa}) + X_a c_{pL}(T_v - T_{sa}) + (X_a - X_b)\lambda + X_b c_{pL}(T_{sb} - T_v) + (X_a - X_b)c_{pv}(T_{va} - T_v) \quad (1)$$

Where q_T it is the necessary heat (W), $\dot{m}_s$ is the solid biomass flowrate (g/s), T_{sa} is feed temperature (°C), T_v is vaporization temperature (°C), T_{sb} is final temperature of the solids (°C). T_{va} is final stream temperature (°C), λ is latent heat of vaporization (J/g°C). c_{ps}, c_{pL}, c_{pv} are the specific heats of solid, liquid and vapor (J/g°C).

$$q_T = \dot{m}_s c_{sb}(T_{hb} - T_{ha}) \quad (2)$$

In equation (2), c_{sb} is the gas heat capacity for inlet moisture (J/g°C), T_{hb} is hot gas inlet temperature (°C), T_{hb} is hot gas outlet temperature (°C).

$$q_T = UA\overline{\Delta T} \quad (3)$$

In equation (3), U is the overall heat transfer coefficient (W/m²°C), A is the area and $\overline{\Delta T}$ is the average temperature difference (°C).

For the grinding stage the main equations is [10]:

$$\frac{P}{\dot{m}} = 0.3162 W_i \left(\frac{1}{\sqrt{D_{pb}}} - \frac{1}{\sqrt{D_{pa}}}\right) \quad (4)$$

Where P is the power (kW), $\dot{m}$ is the waste stream rate (t/h), D_{pb} is the particle size at outlet (mm),D_{pa} is the particle size at inlet (mm) and W_i is work index (kW/t) [11].

For the extraction stage, equation 5 was used:

$$Y = 0.009621 - 0.0028125A + 0.0010875B - 0.0014875C + 0.00525D + 0.00045AB + 0.00265AC - 0.0022875AD - 0.000775BC + 0.0000375BD - 0.0009625CD \quad (5)$$

Where Y is the amount of lycopene (mg/g) extracted from the biomass. A, B, C and D are the normalized temperature, residence time, acetone/n-hexane ratio, and volume of solvent mixture, respectively. This extraction model was obtained from experimental data [6]. The normalization of the variables is given by:

$$L_{Ni} = (L_i - L_{Lower})/\,(L_{Upper} - L_{Lower}) \quad (6)$$

Where L_{Ni} is the value of normalized variable L_i, L_{Upper} is de upper limit of variable i and L_{Lower} is the lower limit of variable i.

For the solvent separation and lycopene purification stage, the flash tank and the distillation column were modeled in Aspen Plus. These devices were linked to Python to run the process continuously.

Optimization variables

For this problem, seven variables were considered,

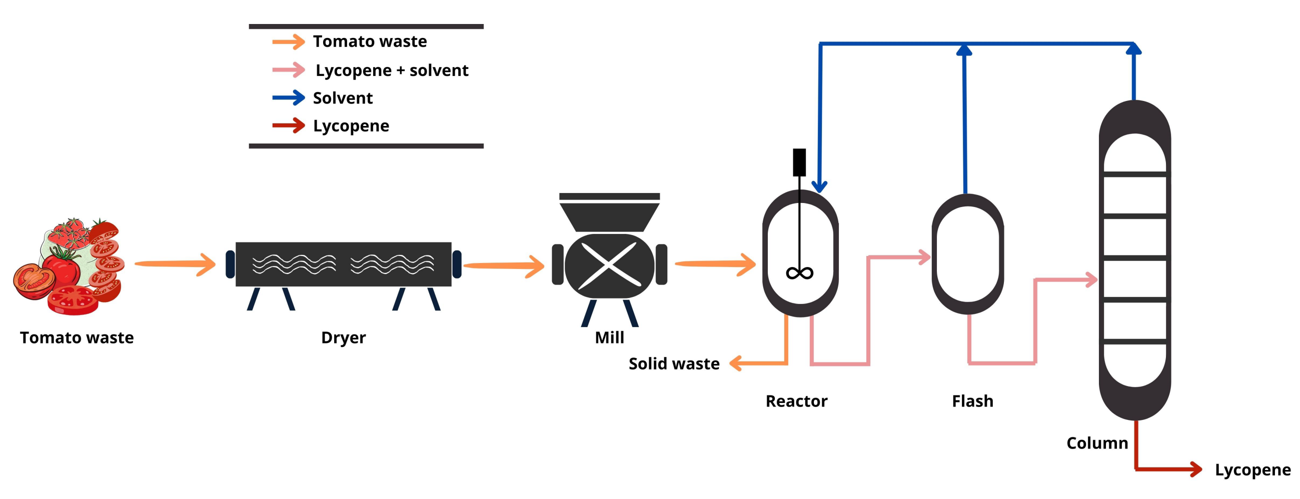

Figure 1: Diagram of the lycopene production process.

the first four correspond to the variables of the lycopene extraction model and the other three variables correspond to the distillation column for purification. These variables and their limits are presented in Table 1.

Table 1: Optimization variables for lycopene production.

Variable	Symbols	Units	Limits
Temperature	T_{ext}	°C	[30, 50]
Residence time	t	min	[1, 60]
Acetone/n-hexane ratio	A:H	%	[25, 75]
Volume of solvent mixture	V_{solv}	L	[10, 30]
Number of stages	N_t	-	[5, 20]
Feed stage	N_F	-	[1, N_t]
Reflux ratio		-	[0.1, 2]

The feed stage has a dynamic upper limit, Herrera-Velázquez et.al. worked with these kinds of limits [12], as it depends on the number of total stages selected. This strategy prevents the feed stage from taking a value greater than the total number of stages in the column.

Process optimization

The mathematical model implemented in Python and linked with Aspen Plus used the Pymoo environment [13]. Having a model that includes the link with Aspen Plus, where the exact equations to be considered are not explicitly available, made stochastic optimization an adequate approach, where the use of Pymoo can be advantageous [14]. The optimization was performed using the genetic algorithm (GA) method considering a population of 100 individuals with 21 generations, implying a total of 2100 evaluations.

Objective function

The objective function is the total annual cost (TAC). To calculate the equipment cost the Guthrie method [15] and the equations presented by Turton et al. [16] were used. Equipment cost has been adjusted to 2023 with a CEPCI of 797.9. Additionally, five years of payback time are assumed. Cost calculations are represented by equations (7), (8) and (9).

$$CT_{equip} = \sum_{i=1}^{7} C_{i,equip} \quad (7)$$

$$CT_{oper} = \sum_{i=1}^{n} C_{i,cooling} + \sum_{i=1}^{n} C_{i,steam} + \sum_{i=1}^{n} C_{i,elect} + \sum_{i=1}^{n} C_{i,solvent} \quad (8)$$

$$TAC = CT_{oper} + \left(\frac{CT_{equip}\left(\frac{CEPCI\ 2023}{CEPCI\ 2001}\right)}{payback\ time} \right) \quad (9)$$

Where, $C_{i,equip}$ is the cost of each of equipment (USD), considering in addition to the five main equipment, the condenser and reboiler of the distillation column. CT_{equip} is the total cost of equipment (USD), $C_{i,cooling}$ is the cooling utility cost (USD/yr), $C_{i,stream}$ is the steam utility cost (USD/yr). $C_{i,elect}$ is the cost of electricity (USD/yr), $C_{i,solvent}$ is the cost of solvent used in the process (USD/yr), CT_{oper} is operating cost (USD/yr), and TAC is the total annual cost (USD/yr).

A flowchart of the optimization algorithm is shown in Figure 2.

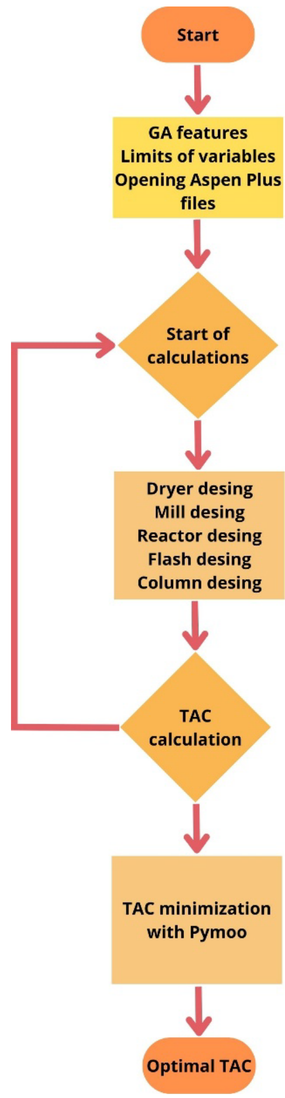

Figure 2. Flowchart of the optimization process.

Base case

The base case contemplates the following values of the optimization variables: a temperature of 40°C, a residence time of 30.5 min, an acetone/n-hexane ratio of 50%, a solvent volume of 50 L, 6 stages in the distillation column with feed in the 3rd stage and a reflux ratio of 0.9, for a lycopene purity of 95%mol. The values of the variables associated to the extraction stage (A, B, C and D) correspond to the central points for the boundaries shown in Table 1. On the other hand, the values for the design variables of the distillation column have been obtained from short-cut methods. This case is compared with the design obtained from optimization.

RESULTS

TAC optimization

The Python code linked to Aspen Plus for TAC minimization was run using Pymoo with GA, a population of 100 individuals and 21 generations. The algorithm took a total of 81.33 minutes to execute the allotted 2100 evaluations. Figure 3 shows how the algorithm progressed in seeking to minimize TAC over the iterations. A greater dispersion in the results is observed in the first iterations, the red line marks the best solution found at that time and continues to be compared, showing that out of 2100 option evaluated, 1957 converged and at that time the change between the solution found stabilizes and finds the optimum.

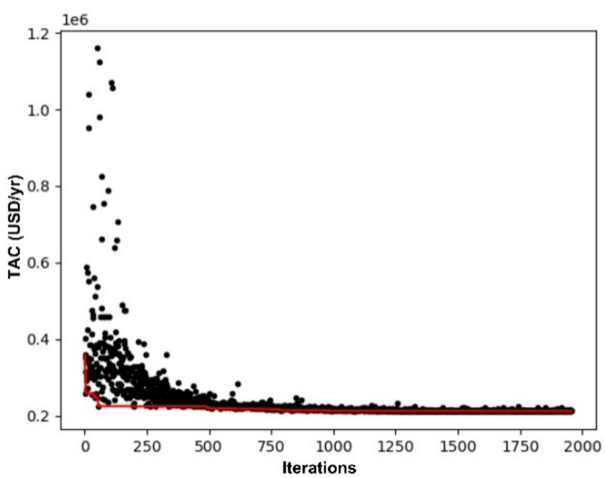

Figure 3. Iterations vs TAC.

Comparison between solutions

Table 2 shows a comparison between the variables and the TAC value between the base case and the optimized solution. The results show that the optimal solution requires a higher temperature as well as a longer residence time and reflux ratio compared to the base case. The favorable solvent ratio is 25% acetone and 75% hexane, and a volume of 10 liters of solvent is sufficient. For the case of distillation column, it requires one less stage than the base case and the feed is kept on the third stage. The optimal solution shows a reduction of the TAC by 147,575.7 USD/yr, equivalent to a reduction of 41.07% of the total annual cost. Table 3 shows more details of the lycopene production process for the optimal solution.

Table 2: Comparison of the base case and the optimal solution.

Variable	Units	Base case	Optimal Solution
Temperature	°C	40	49.89
Residence time	min	30.5	51.39
Acetone/n-hexano ratio	%	50	25
Volume of solvent mixture	L	50	10
Number of stages	-	6	5
Feed stage	-	3	3
Reflux ratio	-	0.9	1.29
TAC	**USD/yr**	**359,267.9**	**211,692.2**

The study allowed obtaining relevant details of the process and the characteristics of the equipment that allow the TAC to be 211,692.2 USD/yr for a production of 2.29 g/h of lycopene, starting from 1000 kg/h of tomato waste. For this desing, the dryer, the mill and the flash tank are equipment that have no optimization variables, so these equipment behave in a similar way during the optimization process. The reactor and the distillation column are the equipment that present the greatest changes in the operating variables and dimensions.

The lycopene/acetone-n-hexane mixture is easy to separate and therefore the thermal loads in the flash tank, the condenser and the reboiler of the distillation column have low values. The solvent/lycopene ratio is something that can affect the maximum purity that can be achieved in the column, since if there is too little lycopene compared to the solvent, more energy is required. Here, it is important to highligh that the reboiler temperature must not be higher than 80°C, since the lycopene can start to degrade. Therefore, it is important to take care of the operating conditions of the distillation column to avoid damaging the product.

CONCLUSIONS

The results indicate that it is possible to develop and optimize a lycopene production process from tomato waste. The design was obtained using Python linked to Aspen Plus, then optimized with a genetic algorithm to minimize the TAC. The results show that the optimized processs achieved a TAC of 211,692.2 USD/yr, reducing the initial TAC of 359,267.9 USD/yr by 41.07%. This study was carried out for a specific experimental model of lycopene extraction, but there are other mathematical extraction models in the literature that are good to be able to compare in the future to explore other alternatives, especially in terms of solvents, to be used to improve the lycopene extraction process. Lycopene extraction models have only reached the formulation of an extraction polynomial or response surfaces, but considering the design on a large scale, it allows to visualize which aspects are important to improve for scaling-up such route for the conversion of wastes into a valuable product.

Table 3: Details of the optimal solution.

Equipment	Variable	Units	Result
Dryer	Area	m^2	1.8
	Cost	USD	19,853.49
Mill	Volume	m^3	43.78
	Power	kW	51.84
	Cost	USD	258,332.00
Reactor	Volume	m^3	13.12
	Residence time	min	51.39
	Temperature	°C	49.89
	Cost	USD	258,332.00
Flash	Volume	m^3	10.60
	Pressure	atm	1.0
	Temperature	°C	70
	Qh	kW	14.67
	Cost	USD	240,400.52
Column	Pressure	Atm	0.1
	Volume	m^3	0.3
	H	m	1.82
	Stages	-	5
	Feed stage	-	3
	Reflux ratio	-	1.29
	Cost	USD	75,849.65
Condenser	Area	m^2	1.0
	Qc	kW	0.000118
	Cost	USD	31,867.17
Reboiler	Area	m^2	1.0
	Qh	kW	0.000117
	Cost	USD	31,867.17
	Operating cost	USD	28,393.79
	TAC	USD/yr	211,692.2
	Tomato waste	Kg/h	1000.00
	Lycopene production	g/h	2.29

ACKNOWLEDGEMENT

The authors thank Universidad de Guanajuato and Universidad de Salamanca for supporting this work, and JCyL SA085P24 funding. N.V. Hernández-Camacho acknowledge SECIHTI for the scholarship granted to develop her doctoral studies.

REFERENCES

1. United Nations. The Sustainable Development Goals Report: Special Edition (2023). https://unstats.un.org/sdgs/report/2023/The-Sustainable-Development-Goals-Report-2023.pdf (accessed September 20, 2024).
2. Montaño Méndez IE, Valenzuela Patrón IN, Villavicencio López KV. Competitividad del tomate rojo de México en el mercado internacional: análisis 2003-2017. *Revista Mexicana de Ciencias Agrícolas* 12 (7): 1185-1197 (2021). https://doi.org/10.29312/remexca.v12i7.2531
3. Méndez-Carmona JY, Ascacio-Valdez JA, Alvarez-Perez OB, Hernández-Almaraz AY, Ramirez-Guzman N, Sepúlveda L, Aguilar-González MA, Ventura-Sobrevilla JM, Aguilar CN. Tomato waste as a bioresource for lycopene extraction using emerging technologies. *Food Biosci* 49: 101966 (2022) https://doi.org/10.1016/j.fbio.2022.101966
4. Kuvendziev S, Lisichkov K, Marinkovski M, Stojchevski M, Dimitrovski D, Andonovikj V. Valorization of tomato processing by-products: Predictive modeling and optimization for ultrasound-assisted lycopene extraction. *Ultrason Sonochem* 110: 107055 (2024) https://doi.org/10.1016/j.ultsonch.2024.107055
5. Catalkaya G, Kahveci D. Optimization of enzyme assisted extraction of lycopene from industrial tomato waste. *Sep Purif Technol* 219:55-63 (2019) https://doi.org/10.1016/j.seppur.2019.03.006
6. Poojary MM, Passamonti P. Optimization of extraction of high purity all-trans-lycopene from tomato pulp waste. *Food Chem* 188:84-91 (2015) http://dx.doi.org/10.1016/j.foodchem.2015.04.133
7. Kaur D, Wani AA, Oberoi DPS, Sogi DS. Effect of extraction conditions on lycopene extraction from tomato processing waste skin using response surface methodology. *Food Chem* 108:711-718 (2008) https://doi.org/10.1016/j.foodchem.2007.11.002
8. Dolatabadi Z, Rad AHE, Farzaneh V, Feizabad SHA, Estiri SH, Bakhshabadi H. Modeling of the lycopene extraction from tomato pulps. *Food Chem* 190:968-973 (2016) http://dx.doi.org/10.1016/j.foodchem.2015.06.069
9. Huang W, Li Z, Niu H, Li D, Zhang J. Optimization of

operating parameters for supercritical carbon dioxide extraction for lycopene by response surface methodology. *J Food Eng* 89: 298-302 (2008) http://dx.doi.org/10.1016/j.jfoodeng.2008.05.006

10. McCabe WL, Smith JC, Harriott P. Operaciones Unitarias en Ingeniería Química. Mc Graw Hill (2007)
11. Williams O, Eastwick C, Kingman S, Giddings D, Lormor S, Lester E. Investigation into the applicability of Bond Work Index (BWI) and Hardgrove Grindability Index (HGI) tests for several biomasses compared to Colombian LA Loma coal. *Fuel* 158: 379-387 (2015) http://dx.doi.org/10.1016/j.fuel.2015.05.027
12. Herrera-Velázquez JJ, Cabrera-Ruiz J, Alcántara-Avila J, Hernández-Hernández EJ, Perez-Gudiño MA, Luna-Perez R, Hernández S. Optimization of the bioethanol dehydration process using $MgCl_2$ as mass separating agent. *Chem Eng Process* 201:109809 (2024) https://doi.org/10.1016/j.cep.2024.109809
13. Blank J, Deb K. Pymoo: Multi-Objective Optimization in Python. *IEEE Access* 8: 89497-89509 (2020) https://doi.org/10.1109/ACCESS.2020.2990567
14. Moriwaki M, Herrera-Velázquez JJ, Cabrera Ruiz J, Matsuda K, Alcántara-Avila R. Synthesis of hybrid membrane distillation processes with optimal structures for ethanol dehydration. *Comput Chem Eng* 178: 108385 https://doi.org/10.1016/j.compchemeng.2023.108385
15. Guthrie KM. Data and techniques for preliminary capital cost estimating. *Chem Eng* 114-142 (1969)
16. Turton R, Shaeiwitz J, Bhattacharyya D, Whiting W. Analysis, Synthesis, and Design of Chemical Processes. Pearson Education (2018)