Computers and Chemical Engineering 152 (2021) 107384

Contents lists available at ScienceDirect

Computers and Chemical Engineering

journal homepage: www.elsevier.com/locate/compchemeng

Simultaneous optimization of the design of the product, process, and supply chain for formulated product

Manuel Taifouris[a], Mariano Martín[a,*], Alberto Martínez[b], Nats Esquejo[c]

[a] Department of Chemical Engineering, University of Salamanca. Plz. Caídos 1-5. 37008, Spain
[b] Procter and Gamble. Brussels Innovation Center, Temselaan 100, 1853 Strombeek-Bever, Belgium
[c] Procter and Gamble. Newcastle Technical Center, Whitley Rd, Longbenton, Newcastle Upon Tyne, Tyne And Wear NE12 9SR, England

ARTICLE INFO

Article history:
Received 17 October 2020
Revised 27 April 2021
Accepted 28 May 2021
Available online 31 May 2021

Keywords:
Supply chain
Process and product design
Integrated problem
Formulated products

ABSTRACT

In this work, an integrated framework and a solution procedure are developed for the design of formulated products. The framework considers the design of the manufacturing process, the products, and supply chains simultaneously. The problem is a multi-period MINLP. A solution procedure consisting of two stages is developed. In stage 1, the model is initialized with the data provided by the market analysis and the optimal formulation for each product and for each location, individually. In stage 2, the model is optimized within the feasible region using the information from stage 1, and the amounts of raw material purchased from each supplier, product manufactured at each location, and product purchased by each customer are determined. The algorithm is used to evaluate the design of powder detergents in Europe. Together with the location of the facilities, product composition, suppliers and price policies are selected.

1. Introduction

The market for consumer products is more competitive than ever as a result of globalization and increasingly demanding customers (Litster and Bogle, 2019). Customers do not only demand that a product meets their needs and likes (Tijskens and Schouten, 2014) but also personalized products, low lead time, quality improvements, sustainable and healthy products, and traceability. Therefore, for a new product to be competitive, it is not only necessary to adjust production costs, but also to comply with what is described above. The use of mathematical optimization techniques for the integrated design of processes and products is a powerful tool to screen among a large number of feasible products by considering only those that can be competitive (Taifouris et al., 2020b) and with a limited environmental impact (Martín and Martínez, 2013), so that resources and time are focused on a detailed design of the most promising products. This allows savings in development costs and reduces the time to launch new products. As a result, the process community has focused its attention on the integrated design of processes and products, with an increase in publications in recent years (Bernardo and Saraiva, 2005; Gani, 2004; Ng and Gani, 2018; Zaman et al., 2018). However, the supply chain has been studied in a subsequent stage. Integrating the design of the supply chain within the process and product design problem may result in more efficient products since it allows the centralization of production and promotes economies of scale (Martín and Martínez, 2018; Tsay et al., 1999). It also facilitates personalization in product design by considering the availability of raw materials and customer demands locally, increasing the flexibility of the entire process by achieving coordination between the three levels (product, process, and supply chain) (Marsillac and Roh, 2014). This coordination allows reducing lead times and improving customer service (Ellram et al., 2007). The integration of product, process, and supply chain design is addressed with the concept of three-dimensional concurrent engineering (3DCE) and, even though it is known since the beginning of 2000 (FINE, 2009), its practical application has been quite limited due to the complexity of the mathematical problem (Caniato et al., 2012; Ellram et al., 2008).

Formulated products are a type of complex products manufactured by mixing the ingredients so that the final product possesses specific physicochemical properties aligned with the needs and likes of its target customer (Audet et al., 2004). Its design depends on its formulation. As a result, the entire supply chain from the suppliers to the product that features the customer expectations, including the processing of the mixture across the manufacturing facility, has an impact on the selection of the process con-

* Corresponding author.
E-mail address: mariano.m3@usal.es (M. Martín).

https://doi.org/10.1016/j.compchemeng.2021.107384

ditions and the different ingredients, (Martín and Martínez, 2013). In formulated products, each different formula constitutes a new product, which facilitates its personalized design for different types of customers. Therefore, the integration of product, process, and supply chain design is especially important in these products. Currently, there is a limited number of scientific publications focused on the simultaneous design of processes and products, in the case of formulated products (Almeida-Rivera et al., 2007; Martín and Martínez, 2013; Zhang et al., 2017), some of them reach integrated supplier selection within the optimization model (Taifouris et al., 2020a), but do not integrate a complete supply chain. The integrated problem for the design of the product, process, and supply chain design is very difficult to address. On the one hand, pooling problems have bilinear products associated with mass balances (Misener et al., 2010), while the models to estimate the raw material prices are usually nonlinear (Audet et al., 2004). Besides, the models to estimate product quality, performance and process constraints are typically highly nonlinear equations, including bilinearities too (Martín and Martínez, 2013). On the other hand, supply chains often have decision variables such as the location of the factories, the customers, as well as the suppliers of raw materials (Yue and You, 2014). These variables are binary variables, transforming the integrated problem into a large non-convex mixed-integer nonlinear programming (MINLP) model. Therefore, models that optimize supply chains usually have the production process and product design fixed, focusing on the analysis of the logistics. In these cases, the models to be solved are usually mixed integer-linear models (MILP) (Allaoui et al., 2018; Liu et al., 2020; Zahiri et al., 2018). Beyond a certain model size, commercial solvers cannot solve this type of problem and the development of specific algorithms for each problem is usually necessary. These algorithms can be genetic algorithms (Lin et al., 2009), algorithms based on linear relaxations (Thanh et al., 2012, 2010; Yue and You, 2014),or on decompositions, such as the Benders (Sahinidis and Grossmann, 1991) and Lagrange (Trespalacios and Grossmann, 2016) decompositions. However, no cases of application of these algorithms have been found in the integrated design of products, processes, and supply chains for the design of formulated products.

In this work, an optimization framework has been developed to address the integrated process-product-supply chain design problem. It has been applied for the optimized design of a specific formulated product, detergent powder at a continental scale. The mathematical model to be optimized is a large non-convex MINLP whose solution is approached by developing a decomposition algorithm. The framework and the algorithm are general and flexible aiming at its use beyond the case study for large-scale problems. The rest of the paper is organized as follows. In Section 2, the mathematical optimization model is developed including the description of the problem, the supply chain integration into the reference work, and the development of the algorithm. In Section 3, the model is applied to a case study in Europe. In Section 4, the results are shown and in Section 5, the conclusions are discussed.

2. Mathematical model

2.1. Description of the problem

The problem we address in this work is the simultaneous product, process, and supply chain design for a specific type of formulated product, detergent powder. The objective is to develop a mathematical model and a decomposition algorithm that allow us to optimize not only the formulation of the product and the operating conditions of the plant but also the location of these plants, the selection of the best ingredients and suppliers, as well as the amount sold to each customer.

We use the work of Taifouris et al. (2020a) as a starting point and refer the reader to that work regarding the details on the modelling of the process constraints, the product performance, and the price policies for the ingredients. The final product consists of a formulated material within the family of consumer products, detergent powder. This is obtained by mixing up to 14 different ingredients, classified into 8 groups. The ingredients, their abbreviations, and their carbon footprints (HC) are shown in Table 1.

In each group, the ingredients differ by their price and their environmental impact. Several suppliers are considered for each ingredient depending on their nature whether they are organic, inorganic, or enzymes. The distances between suppliers and factories are calculated and evaluated from an economic point of view. The price of some of the ingredients (i.e. builders, bleaches, fillers, and enzymes) is considered fixed throughout the year by a contract, selecting the best price policy (linear, exponential, constant elasticity and fixed) based on the amount used of them in the production, while the price of the other ingredients (i.e. surfactant, antifoaming agents, and polymers) may vary throughout the year, due to market fluctuations.

During the production process, which mainly consists of ingredients mixing, slurry drying, and the addition of additives, the different units must operate under certain conditions so that the production of the final product is feasible and its performance meets the consumer needs. Therefore, it is necessary to introduce a series of constraints. Due to confidentiality issues, the ones selected based on open literature were the particle size and cake strength of the final product. Three different types of products are produced with different prices and performances. In this way, a larger market spectrum can be addressed. To characterize the three products, the performance of each one is added in the model as a constraint, setting the maximum and minimum values for each product. In addition, the model is multi-period because the contracts with the suppliers are multiannual. We consider two years as the time horizon.

The objective of this work is to add the dimension of the supply chain to the problem formulated by Martin and Martínez (2013) and Taifouris et al. (2020a), to provide information on the ingredients and detergent composition, the selection of suppliers, price policies, and the location of the production facilities from an economic point of view. This type of model is characterized by its large size and mathematical complexity, having a large number of nonlinear equations and bilinear products. This makes it quite difficult to solve this type of problem directly using a commercial solver. The size and complexity of the problem, especially for supply chains at the continental scale, require the development of a methodology that allows solving the problem considering the entire problem of the detergent production that is, the product, process, and supply chain simultaneously.

2.2. Supply chain design

For the sake of the length of the manuscript, we refer readers to previous work for the models regarding the process and product constraints. In this work, we only indicate how the supply chain is added as well as the development of the algorithm to address its solution. The basic model was presented in Taifouris et al. (2020a). Building on that formulation, see also the **supplementary material**, the dimension of the supply chain is added. This requires the addition of a new dimension to many variables, the location (loc). The modified variables are shown in Table 2. In addition, it is necessary to add new variables and parameters.

In addition, the variable $cv_{ye,j,cli,loc}$ is defined to represent the amount of the product 'j' sold to the customer 'cli' by the factory located at location 'loc' in the year 'ye'. This variable is included in the objective function since, together with the price of the prod-

Table 1
List of considered ingredients and their associated environmental impact.

Group	Ingredient	Abbreviation	HC (tCO_2/t_k)
Surfactant	Linear alkyl aryl sulfonates	LAS	4.20
	Alcohol ethoxylates and alkyl amides	AE	3.70
Builder	Polyphosphates	STPP	1.01
	Zeolite	ZEO	1.76
Bleach	Sodium perborate tetrahydrate	S. PERBO	0.40
	Sodium percarbonate.	S. PERCA	0.40
Fillers	Sodium sulfate	S.SU	0.30
	Xylene sulphonate	X.SU	0.03
Antifoaming agents	-	ANTI	1.76
Enzymes	Protease	PRO	3.69
	Lipase	LIP	3.69
Polymers	Sodium polyacrylate	S. POLY	0.02
	Polyethylene glycol	POLYGLY	0.17
Water	Water	WAT	0.00

Table 2
List of variables and parameters with location dimension.

$x_{ye,i,l,loc}$	$costeing_{ye,i,sup,po,loc}$	$MassprodT_{j,loc}$
$y_{ye,l,\ j,loc}$	$CtransU_{sup,loc}$	$CostP_{ye,i,sup,po,loc}$
$z_{ye,i,\ j,loc}$	$Particle_{j,loc}$	$ccp_{ye,i,sup,po,loc}$
$p_{ye,l,\ k,loc}$	$Cakest_{j,loc}$	$CTT_{sup,loc}$
$PQ_{j,k,loc}$	$MassIng_{ye,i,sup,po,loc}$	$division_{ye,i,sup,po,loc}$
$cv_{ye,j,cli,loc}$	$Massprod_{ye,j,loc}$	
$distance_{sup,loc}$	$Ctrans_{sup,loc}$	
$Performance_{j,loc}$	$division_{ye,i,sup,po,loc}$	

ucts, it represents the income obtained by each factory. The total income is calculated by Eq. (1).

$$Income = \sum_{ye,j,cli,loc} cv_{ye,j,cli,loc} \cdot priceproduct_j \tag{1}$$

Where $priceproduct_j$ are the prices of the three products considered and their values are in the supplementary material. It is considered that the price of the products only depends on the quality perceived by the clients. The relationship between price and location is not considered directly since it can be indirectly related to purchasing power and perceived quality. Both depend on the location and set the demand for different products with different prices, affecting the price of the final product chosen.

The introduction of this new variable implies the introduction of two new inequalities and one equality constraint so as to be able to meet the demand Eqs. (2), ((3), and (4)).

$$\sum_{loc} cv_{ye,j,cli,loc} \le DemandMax_{ye,cli,j} \quad \forall ye,\ j,\ cli \tag{2}$$

$$\sum_{loc} cv_{ye,j,cli,loc} \ge DemandMin_{ye,cli,j} \quad \forall ye,\ j,\ cli \tag{3}$$

$$\sum_{cli} cv_{ye,j,cli,loc} = Massproduc_{ye,j,loc} \quad \forall ye,\ j,\ loc \tag{4}$$

Where $DemandMax_{ye,cli,j}$ and $DemandMin_{ye,cli,j}$ are the maximum and minimum demand of the product 'j' by the client 'cli' in the year 'ye', respectively. $Massproduct_{ye,j,loc}$ is the amount of the product 'j' that is produced in the location 'loc' in the year 'ye' and is defined by Eq. (5).

$$Massproduc_{ye,j,loc} = \sum_{l} y_{ye,l,j,loc} + \sum_{i} z_{ye,i,j,loc} \quad \forall ye,\ j,\ loc \tag{5}$$

Where $z_{ye,I,j,loc}$ and $y_{ye,l,j,loc}$ are the flows of material, see also the supplementary material for further details.

The equation that determines the transportation costs from the suppliers to the factory has already been described in the work by Taifouris et al. (2020a). In this work, a similar strategy is followed to compute the transportation cost for the product sold to customers. First, the unit transport cost is calculated. For that, the cost of transportation from the production plant at location 'loc' to the customer 'cli' is determined. The price is calculated considering consumption of 25 L/100 km and a diesel cost of 1 €/L. Therefore, the price per kilometer is 0.25 €/km and the transport cost is calculated by Eq. (6).

$$CtransCli_{loc,cli} = distC_{loc,cli} \cdot 0.25€/km \quad \forall loc,\ cli \tag{6}$$

Where $distC_{loc,cli}$ is the distance between the location where the product is produced and the customer who buys them. The unit cost is given by Eq. (7).

$$CtransCliU_{loc,cli} = \frac{CtransCli_{loc,cli}}{LoadingCapacity} \quad \forall loc,\ cli \tag{7}$$

Where the loading capacity of each truck is 7 t and the total cost of transportation is calculated by Eq. (8).

$$TotalCostTc = \sum_{ye,j,cli,loc} CtransCliU_{loc,cli} \cdot cv_{ye,j,cli,loc} \tag{8}$$

Border taxes, in the case of application, are included in the initial cost of raw materials ($c0_{i,sup}$, see supplementary material), and in the price of the products ($priceproduct_j$).

Finally, it is necessary to add the fixed cost (FixCost) to penalize the installation of a large number of small facilities. If this cost were not added, the mathematical optimization model would choose to build as many plants near the suppliers and customers as needed with the aim of reducing transportation costs. Fixcost is accounted for using a binary variable ($bi1_{loc}$) and added to the objective function. Therefore, the new objective function is given by Eq. (9).

$$\begin{aligned} profit = & \sum_{ye,j,cli,loc} cv_{ye,j,cli,loc} \cdot priceproduct_j \\ & - \sum_{ye,i,sup,po,loc} CostP_{ye,i,sup,po,loc} \cdot ccp_{ye,i,sup,po,loc} \\ & - \sum_{l,loc} Cpool_l \cdot \sum_{ye,j} y_{ye,l,j,loc} - TotalCostTC \\ & - \sum_{ye,i,sup,po,loc} CtransS \cdot ccp_{ye,i,sup,po,loc} \\ & - \sum_{loc} bi1_{loc} \cdot FixCost_{loc} \end{aligned} \tag{9}$$

Another difference between the formulation of this work and that of Taifouris et al. (2020a) is that it is necessary to calculate the distances between locations, suppliers, and customers. We use the coordinates to determine the distances between these three variables, using the Eqs. (10)–(14). The coordinates can be consulted in the supplementary information.

$$\Lambda lat = |lat1 - lat2| \cdot \frac{pi}{180} \tag{10}$$

$$\Lambda long = |long1 - long2| \cdot \frac{pi}{180} \tag{11}$$

$$a = \sin^2\left(\frac{\Lambda lat}{2}\right) + cos\left(lat1 \cdot \frac{pi}{180}\right) \cdot cos\left(lat2 \cdot \frac{pi}{180}\right) \cdot sin^2\left(\frac{\Lambda long}{2}\right) \tag{12}$$

$$c = 2 \cdot \arcsin\left(\sqrt{a}, \sqrt{(1-a)}\right) \tag{13}$$

$$D = R \cdot c \tag{14}$$

Where ‘lat’ is the latitude and ‘long’ is the longitude of the position of the supplier, location, or customer. D is $distC_{loc,cli}$(km) or $distance_{sup,loc}$(km) (depending on the coordinates used) and R is the radius of the Earth (6371 km). Note that in Taifouris et al. (2020a), the variable "distance" was a vector that only depended on the supplier, since one location alone was considered. However, in this work, this variable becomes a matrix that depends on the supplier and the location.

2.3. Development of the algorithm linearization-solution

With a new dimension added (locations) and the fact that the number of locations to be considered is expected to be large, the size of the formulated optimization problem becomes another complexity. In addition, the binary variables associated with the fixed cost ($bi1_{loc}$) transform the previous problem into a mixed-integer nonlinear programming problem (MINLP). As a result, the model cannot be solved directly using a commercial solver (i.e. DICOPT or BARON). Therefore, an algorithm is developed to address the problem. It consists of two stages, such as presented in Fig. 1.

In the first stage, the MINLP is linearized and transformed into a $MILP_{P1}$ using the results of the NLPs of each location ($NLPs_{P1}$). The MILP is solved using a commercial solver and its results are used to fix the decision variables (suppliers, ingredients, and pricing policies). This information is saved in a binary variable, $bi2_{ye,\ i,\ sup,\ po,\ loc}$ and is used to transform the initial problem (MINLP) into an NLP (the ‘$bi1_{loc}$’ binaries are also fixed by the $MILP_{P1}$), passing to the second stage of the algorithm. In the second stage, a final solution is obtained. Note this new binary variable $bi2_{ye,\ i,\ sup,\ po,\ loc}$ was not in the original MINLP. They are generated in the linearized program ($MILP_{P1}$) to save the decision variables and send them to the global model (NLP_{P2}).

2.3.1. Stage 1: linearization of the global problem (MINLP to MILP)

The first stage consists of a linearization of the MINLP. The MINLP contains both nonlinear equations (i.e. price policies) and bilinear products (i.e. product performance equations, mass balances, etc.). To linearize the nonlinear equations, we apply piecewise linear approximations. The nonlinear equations correspond to the equations used to calculate the discount applied to the price of ingredients (see supplementary material and Taifouris et al., (2020a)). These equations strongly depend on the upper and lower limits of raw material availability, so linearization cannot be applied using the same points in all ingredients since the linearization would not be adequate. Therefore, analyzing the nonlinear profiles, for each ingredient different linearization points are chosen, adjusting to the shape of the curves. To approximate the bilinear products, McCormick envelopes are used. Regarding McCormick envelopes, the proper selection of the upper and lower limits of the involved variables determines the accuracy of the approximation of the bilinear products. To compute these limits, the mathematical models developed in previous work are used. They correspond to the optimization considering a single location (each one of the NLP within NLP_{P1} in Fig. 1), a single customer, and multiple suppliers for each type of ingredient. If we take into account the consideration that the optimal composition depends on the demand, the suppliers available for that factory, and local legislation, we can apply the NLPs developed in previous works to each of the locations considered in this work to obtain a first estimate of the optimal composition per location.

Since only one location is considered in each model NLP_{P1}, the number of equations is not larger than in previous works and therefore they can be solved to global optimality using a commercial solver. In addition, each model is independent of any other. Thus, they can be solved in parallel, so the total computational time corresponds to the solution time of a single NLP. After solving them, we can use this information to calculate the limits of the McCormick envelopes, considering that the correct value will be in a margin of 20% with respect to the value obtained from the NLPs, the limits are calculated by Eqs. (15)–(16). This value can be readjusted if needed.

$$PQ^L_{j,k,loc} = 0.9 \cdot PQ_{NLP1-68\ j,k} \quad \forall j, k, loc \tag{15}$$

$$PQ^U_{j,k,loc} = 1.1 \cdot PQ_{NLP1-68\ j,k} \quad \forall j, k, loc \tag{16}$$

Where $PQ^U{}_{j,k,loc}$ and $PQ^L{}_{j,k,loc}$ are the upper and lower limits of the McCormick envelopes, respectively. In order to initialize these NLP models, we need representative values of the demand per location to fix a maximum and minimum value of production. In the first iteration, we do not know the fraction of the total production that may be assigned to each location and therefore, the maximum and minimum total demand is fixed to each location. In this way, each NLP_{P1} has the same demand. This is updated from the second iteration onwards. When the McCormick limits are established and the $MILP_{P1}$ is solved more realistic values for the demand are selected for each location, readjusting the optimal composition of the product in each NLP_{P1} (in each location) and improving the estimation of the McCormick limits. Subsequently, this information is sent back to the $MILP_{P1}$ and new results are obtained, better than previous results. When the results of two iterations are within 1%, the iterative process stops, and the results are given as valid.

2.3.2. Stage2: Solution of the global model without linearization (NLP_{P2})

Once the $MILP_{P1}$ is solved, its results are used to set suppliers, ingredients, locations to build factories, and pricing policies in the original problem. In order to extract this information, it is necessary to add Eq. (17) to the $MILP_{P1}$ model.

$$ccp_{ye,i,\sup,po,loc} - bi2_{ye,i,\sup,po,loc} \cdot Lsub_{i,\sup} \le 0 \quad \forall ye, i, \sup, po, loc \tag{17}$$

Therefore, if ‘ccp’ is different from zero, ‘bi2’ will be 1. As ‘ccp’ is only different from zero when the best suppliers, ingredients, locations, and policies are selected, ‘bi2’ will be 1 for that combination of suppliers, ingredients, locations, and policies. Due to the need for the use of binary variables, we add the Eq. (18) in the $MILP_{P1}$ to use the same price policy and supplier all years, and therefore, it is no longer necessary to include penalty functions to ensure that the same supplier and price policy are selected over the years. That was a particular reformulation needed in the

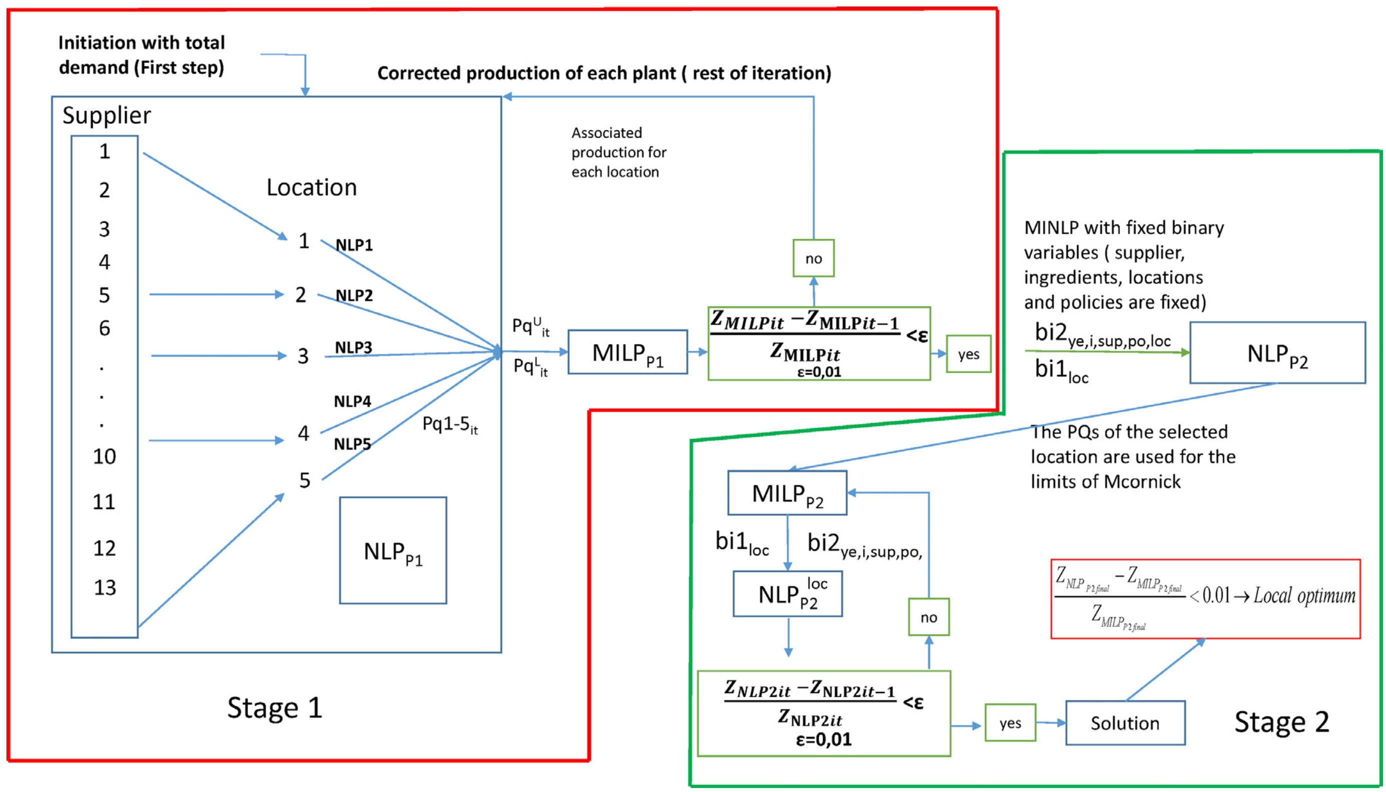

Fig. 1. Example of application of the algorithm linearization-solution.

previous work to avoid converting the optimization problem in an MINLP (Taifouris et al., 2020a).

$$bi2_{1,i,\sup,po,loc} = bi2_{2,i,\sup,po,loc} \quad \forall i, \sup, po, loc \tag{18}$$

Besides, Eq. (19) is added to select only one pricing policy.

$$\sum_{po} bi2_{ye,i,\sup,po,loc} \le 1 \quad \forall ye, i, \sup, loc \tag{19}$$

Eq. (20) is added to select locations, using 'bi1'.

$$bi2_{ye,i,sup,po,loc} - bi1_{loc} \le 0 \quad \forall ye, i, sup, po, loc \tag{20}$$

It is necessary to remember that $bi1_{loc}$ is introduced in the model to determine the fixed costs, as indicated in previous sections (see the objective function in section 2.2). The binary variables can be used differently. If 'bi2' is fixed, 'ccp' can only be different from zero for the case in which 'bi2' is one. Therefore, the values of 'bi2' of $MILP_{P1}$ are added in the NLP_{P2} using the Eq. (21).

$$ccp_{ye,i,\sup,po,loc} - bi2_{ye,i,\sup,po,loc} \cdot \underset{i,\sup}{Lsup} \le 0 \quad \forall ye, i, \sup, po, loc \tag{21}$$

The fixed value for the binary 'bi1' is introduced in the objective function of the NLP_{P2} to determine the fixed costs. Note both 'bi1' and 'bi2' are variables in the MILP $_{P1}$ and parameters in the NLP_{P2}. Since all binary variables are set by the $MILP_{P1}$, the initial MINLP is transformed into an NLP_{P2} that let us find a solution in the feasible region of the problem. However, it should be noted that the global optimum cannot be guaranteed, but it is possible to solve the large integrated process product and supply chain design problems that cannot be addressed using a commercial solver directly.

Next, once the NLP_{P2} has been solved, the results corresponding to the optimal formulation of the detergents, PQ are used as the next limits in the MILP $_{P2}$ model, transforming the problem in an iterative process that finishes when the difference between the result in the objective function of the NLP_{P2} and the previous iteration is less than an ε value (see the Fig. 1). The only difference between $MILP_{P1}$ and $MILP_{P2}$ is that the McCormick envelope limits are set by NLP_{P1} in the case of $MILP_{P1}$, while in the case of $MILP_{P2}$ are set by the results of NLP_{P2}.

Table 3
Demands of the costumers.

Demand							
	Product(t/year)				Product(t/year)		
Costumer	1	2	3	Costumer	1	2	3
1	100	20	5	8	50	25	5
2	10	100	15	9	100	10	5
3	5	100	10	10	150	15	5
4	2.5	50	15	11	50	5	5
5	5	100	10	12	25	100	10
6	100	50	5	13	10	25	100
7	15	50	15	14	100	25	5

3. Case of study

The aim of this work is to optimize the design of the process, product, and supply chain simultaneously. We apply the methodology shown in section 2 to the case of Europe. We consider 68 locations to build the factories and 29 possible suppliers corresponding to the major chemical companies in the continent. The demand considered for each customer is shown in Table 3. Even though the different models integrated by the algorithm make it possible to establish different demands for each year, the same demand is established in order to facilitate the analysis of results and analyze

Fig. 2. Locations, suppliers, and customers.

the coherence of the results provided by the algorithm. Suppliers, locations, and customers can be seen in Fig. 2. The customers correspond to real locations of major retailers across Europe.

4. Results

In this section, the results of applying the algorithm developed in this work to the case study described in Section 3 are commented. On the one hand, the computational performance of the algorithm is discussed, indicating the computation time of each stage of the algorithm, the number of iterations necessary to solve the proposed problem and its comparison with commercial solvers. On the other hand, the results referring to the optimal selection of suppliers, locations, formulation, and customers are analyzed. The algorithm is executed in an Intel Core i7-7700 computer at 3.6 GHz (4,2GHz as turbo frequency), 65W of TDP, 4 core with 8 threads, and 32 Gb of RAM (1200MHz).

4.1. Computational performance of the algorithm

In Table 4, the total CPU time of each stage and the number of iterations of each iterative process within each stage are shown, and this information is also compared with that of applying two commercial solvers (BARON and DICOPT).

For the case study presented in this work, only the application of the algorithm has given any results. BARON was used to solve NLP_{P1}, CPLEX for the MILP's, and CONOPT for NLP_{P2}. In stage 1, 2 iterations were necessary to converge. The total computation time of stage 1 is 1714 seconds. Most of this time is used to solve NLP_{P1}s that consume 500s of solution time for each iteration. With regards to stage 2, the second iteration provided a result similar to the one obtained in the first iteration. The iterative process does not improve the solution. This may be due to the solution provided by the NLP solver, CONOPT. Changing the limits of the McCormick envelopes with the information from the NLP_{P2} does not significantly improve the solution in this specific case. Therefore, the total computation time of stage 2 is 358 s. The upper bound, provided by the relaxed problem (MILP) is 10860 while the lower bound (NLP_{P2}) is 10834. The difference is only **0.24%**, giving as optimal results those obtained in the first iteration of stage 2. Note that we cannot ensure the global optimum. Global solvers (i.e BARON), due to the complexity and size of the case study presented, do not provide us with any results and even a local solver such as DICOPT does not provide us with a valid result either. Therefore, the result provided by the algorithm is taken as the only optimal value that can be calculated and it is evaluated if the result is logical with the location of the suppliers, the demands of the clients, the distribution of the production of the factories, and the selection of policies.

4.2. Results of the case study

While the $MILP_{P1}$ sets the decision variables, such as, which suppliers, ingredients, locations to build the factories and pricing policies to choose for each ingredient, the NLP_{P2} sets the amounts of raw material purchased, the amount of product produced, the amount sold to each customer, and the formulation of detergents. Table 5 shows the purchased amount of the selected ingredients and Table 6 shows the total amount of each product produced by each location. Fig. 3 shows the locations chosen as optimal and the suppliers that supply the ingredients to the locations.

As it can be seen in Table 5, the selected ingredients are the cheapest (see supplementary information) since environmental impact limits have not been established so that the results are easier to interpret and verify that the algorithm is working properly. Each location centralizes the purchases of its ingredients in a single supplier by type of ingredient (organic, inorganic or enzymes) since they have enough raw material available. The centralization of purchases allows obtaining the maximum possible discount in

Fig. 3. Selected locations and suppliers.

Table 4
Computational results.

Stage	Problem size		Iteration number	Z	ε	Time per iteration(s)	Total time(s)
1	NLP_{P1}	Eqs= 12573 Vars=16876	1	-	-	500	1714
			2	-	-	500	
	$MILP_{P1}$	Eqs=3568820 Vars= 3019881(413)	1	10846	-	706	
			2	10860	1E-03	8	
2	NLP_{P2}	Eqs= 864950 Vars= 882247	1	10829	-	170	358
			2	10834	4E-04	172	
	$MILP_{P2}$	Eqs=3568820 Vars= 3019881(413)	1	10861	-	9	
			2	10860	-	7	
Commercial Solvers							
BARON	Eqs= 1236069 Vars=974037(68)		-	No solution	-	-	72000
DICOPT			-	Infeasible	-	-	6548

the prices of the ingredients. In addition, the closest suppliers for each of the locations are chosen, reducing transport costs.

The optimal composition of each product in each selected location can be seen in Table 7. The optimal composition of the 3 products is very similar in the 3 selected locations. Therefore, it is possible to refer to products 1, 2 and, 3 without loss of generality or specifying the factory that produces it. On the one hand, product 1 requires a larger fraction of enzymes in its composition (2.5 times higher) to meet its expected performance. On the other hand, product 2 has the same fraction of the enzyme as product 3, but its percentage of bleach is higher. It can be seen (in Eq. (21). of the supplementary information) that the bleaches have a very important effect on the washing performance and allow product 2 to have an intermediate performance between products 1 and 3 without having to increase its concentration in enzymes since these have a very high cost.

Regarding production, if the demands shown in Table 3 and the locations of the customers (see Fig. 2) are analyzed, it can be seen that the countries of central Europe demand a larger amount of the high-quality and high-price product (product 1), while Spain, Portugal and Italy demand lower-quality and lower-price products (product 2). Finally, the countries of Eastern Europe demand the lowest priced products (product 3). For this reason, the three factories installed have a different distribution of products in their production (see Table 6). The most manufactured product in each of them depends on the demand of customers close to them. Thus, the plant located in Spain shows a higher production of product 2, since the closest customers have a higher demand for this product. Similarly, it occurs with the factory located in France and Poland. Noted that the plant located in France is the one that produces the most, not only because it has more customers around it, but also because they are the ones that consume the most. Therefore, it is

Table 5
Amount purchased of each ingredient by the selected supplier.

Ingredient	Amount (t) Supplier	Policy	A	B	C
LAS	4	-	120.72	0.00	0.00
	11	-	0.00	286.50	0.00
	28	-	0.00	0.00	90.69
ZEO	3	2	423.02	0.00	0.00
	10	3	0.00	1000	0.00
	24	3	0.00	0.00	306.16
S. PERBO	10	2	0.00	197.66	0.00
S. PERCA	3	3	123.50	0.00	0.00
	24	3	0.00	0.00	32.13
S.SU	3	2	95.91	0.00	0.00
	10	3	0.00	374.40	0.00
	24	3	0.00	0.00	53.00
ANTI	4	-	0.81	0.00	0.00
	11	-	0.00	1.91	0.00
	28	-	0.00	0.00	0.53
PRO	6	3	9.12	0.00	0.00
	9	2	0.00	37.83	0.00
	25	3	0.00	0.00	7.73
S. POLY	4	-	0.73	0.00	0.00
	11	-	0.00	0.71	0.00
	28	-	0.00	0.00	0.36
POLYGLY	4	-	0.08	0.00	0.00
	11	-	0.00	1.20	0.00
	28	-	0.00	0.00	0.17
WAT	29	-	31.09	9.55	39.23

Table 6
Total produced amount of each product by each location.

Product	Production(t) A	B	C
1	75	1200	170
2	600	650	100
3	130	60	260

the one that produces the highest benefit. Thus, there is a centralization of production, being factory B the most prominent case. This allows reducing both the costs of raw materials (economy of scale) and the costs of distributing products. Regarding the price policies, we can see in Table 5 that all the ingredients are purchased using policies 2 and 3. Because both policies provide a very similar reduction in prices for the purchased amounts of ingredients, it is not possible to select the policy clearly between the two. However, policies 1 and 4 are not chosen in any case. In fact, if the raw material were purchased using policy 1 instead of policy 2 or 3, the cost of the raw material would be between 18-26% higher.

5. Conclusion

An algorithm has been developed for the integrated design of the process, product, and supply chain for formulated products. It consists of two stages. Stage 1 initializes the problem, providing the information for bounds and sets the decision variables. Stage 2 is responsible for product design and selection of the amount purchased for each ingredient from each supplier selected in stage 1, the amount sold to each customer and the amount manufactured in each of the previously set locations.

From the results of this work, it can be concluded that for the integrated design of products, processes, and supply chains, it is necessary to adapt the problem in order to be solved. If the model is complex enough, commercial solvers cannot find a solution. After the development of an algorithm adapted to the particular case of detergent production in Europe, it was found that the algorithm could find an optimal solution (even though a global optimum could not be ensured) in less than an hour. After analyzing the results, the solution found was logical with the problem posed. It was decided to build 3 different factories, one focused on each type of product and each type of market to reduce the transportation costs of finished products. Purchases were centralized to achieve the largest possible discount and policies 1 and 4 were discarded. Further work may include multi-objective optimization.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

CRediT authorship contribution statement

Manuel Taifouris: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Writing - original draft, Writing - review & editing, Visualization. **Mariano Martín:** Conceptualization, Methodology, Investigation, Writing - original draft,

Table 7
Optimal composition of the products in each location.

Location	Optimal formulation%($t_k/t_p \bullet 100$) A								
Ingredients/Products	**Surfactant**	**Builder**	Bleach	**Filler**	**Antifoam**	**Enzyme**	**Polymer**		**Water**
	1	4	6	7	9	10	12	13	Water
1	15.00%	60.00%	5.00%	16.87%	0.10%	2.43%	0.00%	0.10%	0.50%
2	15.00%	50.00%	18.79%	10.00%	0.10%	1.00%	0.10%	0.10%	5.01%
3	15.00%	60.00%	5.41%	17.89%	0.10%	1.00%	0.10%	0.10%	0.50%
Location	**B**								
Ingredients/Products	**Surfactant**	**Builder**	**Bleach**	**Filler**	**Antifoam**	**Enzyme**	**Polymer**		**Water**
	1	4	6	7	9	10	12	13	Water
1	15.00%	55.51%	5.00%	21.23%	0.10%	2.56%	0.00%	0.10%	0.50%
2	15.00%	45.82%	20.68%	16.80%	0.10%	1.00%	0.10%	0.00%	0.50%
3	15.00%	60.00%	5.41%	17.89%	0.10%	1.00%	0.10%	0.00%	0.50%
Location	**C**								
Ingredients/Products	**Surfactant**	**Builder**	**Bleach**	**Filler**	**Antifoam**	**Enzyme**	**Polymer**		**Water**
	1	4	6	7	9	10	12	13	Water
1	15.00%	60.00%	5.00%	10.00%	0.10%	2.43%	0.00%	0.10%	7.37%
2	25.00%	48.16%	10.63%	10.00%	0.10%	1.00%	0.10%	0.00%	5.01%
3	15.46%	60.00%	5.00%	10.00%	0.10%	1.00%	0.10%	0.00%	8.34%

Writing - review & editing, Visualization, Supervision, Project administration, Funding acquisition. **Alberto Martínez:** Conceptualization, Writing - review & editing, Supervision, Project administration. **Nats Esquejo:** Conceptualization, Writing - review & editing, Supervision, Project administration, Funding acquisition.

Acknowledgement

The authors appreciate P&G and PSEM3 USAL for funding the research and MT acknowledges the JCyL for a PhD fellowship.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.compchemeng.2021.107384.

References

Allaoui, H., Guo, Y., Choudhary, A., Bloemhof, J., 2018. Sustainable agro-food supply chain design using two-stage hybrid multi-objective decision-making approach. Comput. Operat. Res. 89, 369–384. doi:10.1016/j.cor.2016.10.012.

Almeida-Rivera, C., Jain, P., Bruin, S., Bongers, P., 2007. Integrated product and process design approach for rationalization of food products. Comput. Aided Chem. Eng. 24, 449–454. doi:10.1016/S1570-7946(07)80098-8.

Audet, C., Brimberg, J., Hansen, P., le Digabel, S., Mladenović, N., 2004. Pooling problem: Alternate formulations and solution methods. Manag. Sci. 50, 761–776. doi:10.1287/mnsc.1030.0207.

Bernardo, F.P., Saraiva, P.M., 2005. Integrated process and product design optimization: a cosmetic emulsion application. Comput. Aided Chem. Eng. 20, 1507–1512. doi:10.1016/S1570-7946(05)80093-8.

Caniato, F., Caridi, M., Crippa, L., Moretto, A., 2012. Environmental sustainability in fashion supply chains: An exploratory case based research. Int. J. Prod. Econ. 135, 659–670. doi:10.1016/j.ijpe.2011.06.001.

Ellram, L.M., Tate, W., Carter, C.R., 2008. Applying 3DCE to environmentally responsible manufacturing practices. J. Cleaner Prod. 16, 1620–1631. doi:10.1016/j.jclepro.2008.04.017.

Ellram, L.M., Tate, W.L., Carter, C.R., 2007. Product-process-supply chain: an integrative approach to three-dimensional concurrent engineering. Int. J. Phys. Distribut. Logistics Manag. 37, 305–330. doi:10.1108/09600030710752523.

FINE, C.H., 2009. Clockspeed-based strategies for supply chain design 1. Product. Operat. Manag. 9, 213–221. doi:10.1111/j.1937-5956.2000.tb00134.x.

Gani, R., 2004. Chemical product design: Challenges and opportunities. Comput. Chem. Eng. 28, 2441–2457. doi:10.1016/j.compchemeng.2004.08.010.

Lin, L., Gen, M., Wang, X., 2009. Integrated multistage logistics network design by using hybrid evolutionary algorithm. Comput. Ind. Eng. 56, 854–873. doi:10.1016/j.cie.2008.09.037.

Litster, J., Bogle, I.D.L., 2019. Smart process manufacturing for formulated products. Engineering doi:10.1016/j.eng.2019.02.014.

Liu, S., Papageorgiou, L.G., Shah, N., 2020. Optimal design of low-cost supply chain networks on the benefits of new product formulations. Comput. Ind. Eng. 139, 106189. doi:10.1016/j.cie.2019.106189.

Marsillac, E., Roh, J.J., 2014. Connecting product design, process and supply chain decisions to strengthen global supply chain capabilities. Int. J. Prod. Econ. 147, 317–329. doi:10.1016/j.ijpe.2013.04.011.

Martín, M., Martínez, A., 2018. On the effect of price policies in the design of formulated products. Comput. Chem. Eng. 109, 299–310. doi:10.1016/j.compchemeng.2017.11.019.

Martín, M., Martínez, A., 2013. A methodology for simultaneous process and product design in the formulated consumer products industry: The case study of the detergent business. Chem. Eng. Res. Des. 91, 795–809. doi:10.1016/j.cherd.2012.08.012.

Misener, R., Gounaris, C.E., Floudas, C.A., 2010. Mathematical modeling and global optimization of large-scale extended pooling problems with the (EPA) complex emissions constraints. Comput. Chem. Eng. 34, 1432–1456. doi:10.1016/j.compchemeng.2010.02.014.

Ng, K.M., Gani, R., 2018. Chemical product design: advances in research and teaching. Comput. Aided Chem. Eng. Elsevier B.V. 21–32. doi:10.1016/B978-0-444-64241-7.50003-3.

Sahinidis, N.v., Grossmann, I.E., 1991. Convergence properties of generalized benders decomposition. Comput. Chem. Eng. 15, 481–491. doi:10.1016/0098-1354(91)85027-R.

Taifouris, M., Martín, M., Martínez, A., Esquejo, N., 2020a. On the effect of the selection of suppliers on the design of formulated products. Comput. Chem. Eng. 141, 106980. doi:10.1016/j.compchemeng.2020.106980.

Taifouris, M., Martín, M., Martínez, A., Esquejo, N., 2020b. Challenges in the design of formulated products: multiscale process and product design. Curr. Opin. Chem. Eng. doi:10.1016/j.coche.2019.10.001.

Thanh, P.N., Bostel, N., Péton, O., 2012. A DC programming heuristic applied to the logistics network design problem. Int. J. Prod. Econ. 135, 94–105. doi:10.1016/j.ijpe.2010.09.025.

Thanh, P.N., Péton, O., Bostel, N., 2010. A linear relaxation-based heuristic approach for logistics network design. Comput. Ind. Eng. 59, 964–975. doi:10.1016/j.cie.2010.09.007.

Tijskens, L.M.M., Schouten, R.E., 2014. Modeling Quality Attributes and Quality Related Product Properties. In: Postharvest Handling: A Systems Approach. Elsevier Inc., pp. 411–448. doi:10.1016/B978-0-12-408137-6.00015-6.

Trespalacios, F., Grossmann, I.E., 2016. Lagrangean relaxation of the hull-reformulation of linear generalized disjunctive programs and its use in disjunctive branch and bound. Eur. J. Operat. Res. 253, 314–327. doi:10.1016/j.ejor.2016.02.048.

Tsay, A.A., Nahmias, S., Agrawal, N., 1999. Modeling Supply Chain Contracts: A Review. Springer, Boston, MA, pp. 299–336. doi:10.1007/978-1-4615-4949-9_10.

Yue, D., You, F., 2014. Game-theoretic modeling and optimization of multi-echelon supply chain design and operation under Stackelberg game and market equilibrium. Comput. Chem. Eng. 71, 347–361. doi:10.1016/j.compchemeng.2014.08.010.

Zahiri, B., Jula, P., Tavakkoli-Moghaddam, R., 2018. Design of a pharmaceutical supply chain network under uncertainty considering perishability and substitutability of products. Inf. Sci. 423, 257–283. doi:10.1016/j.ins.2017.09.046.

Zaman, U.K.uz, Rivette, M., Siadat, A., Mousavi, S.M., 2018. Integrated product-process design: Material and manufacturing process selection for additive manufacturing using multi-criteria decision making. Rob. Comput. Integr. Manuf. 51, 169–180. doi:10.1016/j.rcim.2017.12.005.

Zhang, L., Fung, K.Y., Zhang, X., Fung, H.K., Ng, K.M., 2017. An integrated framework for designing formulated products. Comput. Chem. Eng. 107, 61–76. doi:10.1016/j.compchemeng.2017.05.014.