19th EURO Working Group on Transportation Meeting, EWGT2016, 5-7 September 2016, Istanbul, Turkey

Determining an Optimal Area to Locate a Biorefinery under Economic and Environmental Criteria

Adrian Serrano-Hernandez\textsuperscript{a,}\textsuperscript{*}, Javier Faulina, Jesús M. Pintora and Javier Belloso\textsuperscript{a}

\textsuperscript{a}Integrated Group of Logistics and Transportation, Institute of Smart Cities. Public University of Navarre, Pamplona 31006, Spain

Abstract

Facilities location is a strategic decision which has to be carefully considered because it could involve the failure or success of a business. For that reason, anything that helps decision makers to facilitate their location decision processes is of their utmost interest. The aim of this paper is, therefore, providing a methodology that could be useful for the decision makers by giving them not only an optimal point but also a whole region where they can focus on their attention. Knowing that biofuels are settling as a new alternative energy source which has been spreading around the world to reduce greenhouse gas emissions and oil dependence, this methodology is tested in the real case of locating a biorefinery in Navarre, Spain. Moreover, A Mixed Integer Linear Programming (MILP) model has been developed to generate optimal region vertices as well as some other supply chain characteristics, including, among others, which crops are going to be harvested, when they are going to be collected, and their storage levels. Additionally, two criteria were implemented in MILP model to create two optimal regions: one considering an economic criterion and other one minimizing environmental impact. As a result, two regions were drawn in the Navarrese territory that point out where a biorefinery should be located and how the supply chain should be designed.

© 2017 The Authors. Published by Elsevier B.V.

Peer-review under responsibility of the Scientific Committee of EWGT2016.

Keywords: Facility Location Problem; Biorefinery, Mixed Integer Linear Programming, Biofuels
1. Introduction

The location of a specific facility may be the difference between bankruptcy and success of its associated business. Decisions about location are said to be strategic because they require a large amount of resources which will have a long projection over time. Thus, locating a facility is a crucial decision companies usually have to face, at least, once in their life. For that reason, a lot of attention has been paid to develop several tools (Decision Support Tools, DST) that help decision makers to support their decisions in general, and their facility location decisions in particular. Such is the case that a whole branch of Operations Research deals with this kind of problems, which is known as Facility Location Problem (FLP).

Factors affecting locations decisions are unlimited, however, researchers have tried to identify the most important ones in order to measure and implement them in their DST. For instance, Chan (2001) considered two great factors classes: tangible factors (easy to quantify), which include transport costs, staff costs, energy costs, land availability, taxes…; and intangible factors (difficult to quantify), such as government stability, competitors, costumers/workers preferences, pollution and others. Then, some qualitative and/or quantitative analyses have been performed to obtain an optimal location given the aforementioned factor (Daskin, 2013). However, many problems can arise (administrative or legal issues, underestimated cost, unconsidered negative factors, etc.) when obtaining just a point to locate a facility once you have considered all the factors. For that reason, it is useful not only providing an optimal point to set up a firm but also a whole area to be considered by decision makers, or even several areas following. By doing so, the DST became a complete tool that helps the decision makers to locate the facility, providing information of two types: positive (the set of points being candidates for the location) or negative (set of points which should not locate the facility).

As a way to illustrate the previous methodology, this paper aims to determine optimal regions to place a biorefinery. A biorefinery can be defined as a complex facility that uses biomass as feedstock for the sustainable production of a range of different products (mainly biofuels, but also chemical commodities and electricity) which requires the integration of a huge variety of technologies (Cherubini et al, 2009). The biofuels have been reaching more and more interest because they search a worthwhile substitution for fossil fuels in transportation sector. Firstly, the fact that biofuels are usable in current vehicle setup make simpler their adoption and growth (Al-Mulali, 2015). Secondly, the use of biofuels allows a reduction of the dependence of many Western countries from oil production and extraction (Kallas and Gil, 2015). Finally, biofuels are considered a way to reduce level of CO2 emissions and increase energy security (Börjesson et al, 2014). For these reasons, among others, policy makers are promoting the use of biofuels in transportation (Cansino et al, 2012) which leads, therefore, to the consideration of building facilities capable to generate such biofuels (biorefineries) over the world. In this case, we will consider the Spanish Northern region of Navarre as a place where locate a potential biorefinery (see Fig. 1).

![Fig. 1. Location of Navarre in Europe and Spain.](image-url)
2. Literature Review

Concerning location problems, the places where a facility can be settled up are frequently limited to a finite set of candidate locations (Revelle and Eiselt, 2005). In a broad sense, there exist three vast classes of FLP depending on modelling approaches: coverage, center of gravity and p-median (Daskin, 2013). Usually, the customer service depends on the distance between the customers and the facility in such a way that, in order to maximize their utility, customers are assigned to the nearest facility considering an upper bound in distance, and that is called coverage (Farahani et al, 2012). However, in other cases, this given upper bound is not exogenously taken (i.e. is a variable) and the model tries to minimize it in order to cover all demands. So, distances between any customer and the facility are, therefore, minimized (that is to say, that facility is a center of gravity) (Garfinkel et al, 1977). Finally, we can consider the numerical demand of potential customer in order to open a facility close enough to those big customers. In other words, facilities are located having the purpose of obtaining the minimum weighted distance (facility is at the median point, p-median) (Klose and Drezner, 2005).

Apart from classical models, many variations of them have appeared in order to cover new challenges. On one hand, we have paid our interest in the randomness of some variables: demand with some extreme factors (Murali et al., 2012), or agricultural production (Serrano et al., 2015). Other models we could have considered are: multiobjective FLP (Amim and Zhang, 2013), location integrated with routing (Location Routing Problem (Escobar et al., 2015)), and a miscellaneous group of cases due to Owen and Daskin (1998), Şahin and Süral (2007), Drezner and Drezner (2007) and Melo et al, (2009). On the other hand, some works related to FLP solution techniques have been gaining importance along the time. Thus, many algorithms have been developed to face new approaches in FLP modelling such as combining simulation and optimization (Munoz-Villamizar et al, 2014), genetic algorithms (Korac et al, 2012) or Pareto-based metaheuristics in multiobjective approaches (Rahmati et al. 2012); even though exact methods are widely used (Tragantalerngsak et al, 2000). Furthermore, Mixed Integer Linear Programming (MILP) formulation is required to consider binary decisions such as location and assigning (Diabat et al, 2013).

Considering jointly the biorefinery and the location literature, only a few works have analyzed biorefinery locations from analytical point of view. Most of them are supported by Geographical Information Systems such as Yu et al. (2014), using Net Present Value as function objective to be maximized (Marvin et al, 2012) as well as multiobjective programming where You et al. (2012), for instance, considered economic and environmental criteria. Thus, the contribution of this paper is twofold. On one hand, this article makes a contribution in the scarce field of biorefinery location, solving a real optimization problem. Therefore, the fierce competitive environment for the feedstock in the local markets is taken into account. A high detailed data has been used (genuine biomass production and availability, biomass prices, transportation costs and storage costs) in order cover all the relevant issues that affect to location decisions. On the other hand, a novel methodology based on optimal regions is implemented. That methodology provides helpful graphical solution to the decision makers to focus on a limited, but important, area to make their location decisions.

3. Methodology

Locating a biorefinery is not straightforward at all and many variables should be considered. Moreover, the Facility Location Problem usually considers a set of points as potential locations (Zanjirani, and Hekmatfar, 2009), an element which is not particularly helpful for us since we want to derive optimum areas. A Mixed Integer Linear Programming (MILP) model is then developed to generate the vertices of the optimal region. Here, two criteria were selected to be included independently in the objective function: economic and environmental criteria. The former included purchase, transport and stock costs, the latter consisted of the minimization of distance driven. We take for granted that this is not an ‘Environmental Impact Minimization Problem’ since we are not considering many other factors which are traditional in this problem (road gradient, payload… (Demir et al, 2014) but it is an approximation to that model.

Location will not be the only variable we are going to optimize due to the fact that it has a critical influence in the supply chain management of the biorefinery. For that reason, additionally data we considered are related with
feedstock production and consumption, prices, transportation costs, depreciation, warehouses, harvest timing… in order to optimize many supply chain characteristics.

3.1. Characterizing the biorefinery environment

As a first restriction, the biorefinery must be located inside Navarrese boundaries (see Figure 1), which give us a solution space of 10,000 km². Secondly, biomass availability was carefully computed in order to avoid significant food scarcity which could have an impact on prices. With this regard, a genuine number of rape, alfalfa, rice, oat, corn, wheat, and barley as well as agroindustrial wastes are spreading over Navarre and closer regions to be processed by the biorefinery (see Fig. 2.a). Actually, biorefinery size is determined following this criterion, reaching 150,000 net biomass tons per year, that is, once humidity and depreciation have been removed from purchased biomass.

3.2. Optimal area framework

Having the purpose of determining optimum areas, we deployed a grid over Navarre in which every point would be a ‘potential location’ (see Fig. 2.b). The distance between each point is about 20 kilometers but this resolution could be arbitrarily modified: the more distance between points we consider, the bigger optimum area we have, and vice versa. Concerning the optimal region size, it would be useless a quite huge optimal region and, similarly, a very small one would have no sense, because the decision maker would have no choices. Then, we will solve the problem with all the points obtaining ‘the best location’. Later, the previous point will be removed from the sample and the problem will be solved again obtaining another ‘best location’ different from the first one. By doing so four times, a squared will be drawn. A fifth point would be required in order to derive the direction of the optimum area.
3.3. The model with economic criterion

A Mixed Integer Linear Programming (MILP) model is built to determine, as mentioned, not only location of the biorefinery but also its supply chain characteristics. In this sense, Table 1 summarizes the sets that will be used, noting that individual crops productions will be aggregated within those located in the same municipality resulting 222 crops. Observe, too, that we consider 13 different kinds of products (rape, alfalfa, rice, oats, corn, wheat, and barley as well as agroindustrial wastes) that have their particular harvest time through the year: obviously, products are not available during the whole year. Meanwhile, Table 2 and Table 3 show the decision variables and parameters selected to be incorporated into the MILP model.

Table 1. Set definitions.

<table>
<thead>
<tr>
<th>Set</th>
<th>Description</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>Set of crops</td>
<td>$i = 1, 2...222$</td>
</tr>
<tr>
<td>J</td>
<td>Set of points</td>
<td>$j = 1, 2...323$</td>
</tr>
<tr>
<td>P</td>
<td>Set of products</td>
<td>$p = 1, 2...13$</td>
</tr>
<tr>
<td>T</td>
<td>Set of months</td>
<td>$t = 1, 2...12$</td>
</tr>
</tbody>
</table>

Table 2. Decision variables.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$X_j$</td>
<td>1 if the Biorefinery is built in potential location $j$, 0 otherwise</td>
</tr>
<tr>
<td>$Y_{i,t,j}$</td>
<td>1 if crop $i$ is used at time $t$ to serve potential location $j$, 0 otherwise</td>
</tr>
<tr>
<td>$Q_{p,i,t,j}$</td>
<td>Tons of product $p$ bought in crop $i$ at time $t$ to serve potential location $j$</td>
</tr>
<tr>
<td>$BS_{p,t,j}$</td>
<td>Stock of product $p$ at time $t$ in potential location $j$</td>
</tr>
</tbody>
</table>

Table 3. Parameter description.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
<th>Unit</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$h_p$</td>
<td>humidity of product $p$</td>
<td>%</td>
<td>0.12-0.50</td>
</tr>
<tr>
<td>consumption$_p$</td>
<td>biorefinery monthly consumption</td>
<td>Tn</td>
<td>12500</td>
</tr>
<tr>
<td>onseason$_{p,t}$</td>
<td>1 if product $p$ is available at $t$</td>
<td>-</td>
<td>0 or 1</td>
</tr>
<tr>
<td>dist$_{i,j}$</td>
<td>distance from $i$ to $j$</td>
<td>Km</td>
<td>0-200</td>
</tr>
<tr>
<td>seasondur$_p$</td>
<td>season duration of product $p$</td>
<td>Months</td>
<td>2-8</td>
</tr>
<tr>
<td>price$_p$</td>
<td>price of product $p$</td>
<td>€</td>
<td>60-90</td>
</tr>
<tr>
<td>production$_{p,i}$</td>
<td>total production of $p$ in $i$</td>
<td>Tn</td>
<td>0-20,000</td>
</tr>
<tr>
<td>$\alpha_{p,i}$</td>
<td>use factor of product $p$ in $i$</td>
<td>%</td>
<td>0.15-0.45</td>
</tr>
<tr>
<td>$FCost$</td>
<td>transportation fix cost</td>
<td>€/Tn</td>
<td>8.23</td>
</tr>
<tr>
<td>$VCost$</td>
<td>transportation variable cost</td>
<td>€/Tn/km</td>
<td>0.094</td>
</tr>
<tr>
<td>$s$</td>
<td>stock cost</td>
<td>€/Tn/month</td>
<td>0.945</td>
</tr>
<tr>
<td>$\gamma$</td>
<td>losses on stock</td>
<td>%</td>
<td>1</td>
</tr>
<tr>
<td>$\delta$</td>
<td>losses on transportation</td>
<td>%</td>
<td>0.5</td>
</tr>
</tbody>
</table>
Being the model formulation as follows:

\[
\text{min } \text{TotalCosts} = \text{BiomassCosts} + \text{TransportCosts} + \text{StorageCosts} \tag{1}
\]

\[
\text{BiomassCost} = \sum_{p} \sum_{i} \sum_{j} \sum_{t} Q_{p,i,t,j} \cdot \text{price}_p \tag{1.1}
\]

\[
\text{TransportCost} = \sum_{i} \sum_{p} \sum_{j} \sum_{t} (\text{FCost} + \text{VCost} \cdot \text{dist}_{i,j}) \cdot Q_{p,i,t,j} \tag{1.2}
\]

\[
\text{StorageCost} = \sum_{p} \sum_{j} \sum_{t} B_{s,p,t,j} \tag{1.3}
\]

Subject to:

\[
\sum_{j} X_{j} = 1 \tag{2}
\]

\[
\sum_{i} Q_{p,i,t,j} \cdot (1-\delta) + B_{s,p,t-1,j} \cdot (1-\gamma) = \frac{\text{consumption}_{p,i,t}}{1-h_p} + B_{s,p,t,j} \forall p, \forall j, \forall t \tag{3}
\]

\[
Q_{p,i,t,j} \leq AB_{p,i,t} \cdot Y_{j,i,t} \forall i, \forall p, \forall j, \forall t \tag{4}
\]

\[
AB_{p,i,t} = \text{production}_{p,i} \cdot \alpha_{p,i} \cdot \frac{\text{onseason}_{p,t}}{\text{seasondur}_{p}} \tag{5}
\]

\[
\sum_{p} \text{consumption}_{p,j,t} = 12,500 \cdot X_{j} \forall j, \forall t \tag{6}
\]

\[
Y_{j,i,t} \leq X_{j} \forall i, \forall j, \forall t \tag{7}
\]

\[
Y_{j,i,t} \in \{0,1\} \tag{8}
\]

\[
X_{j} \in \{0,1\} \tag{9}
\]

Where, equation (1) is the objective function to be minimized, representing the total costs of the process, which has three sources of costs: the cost of buying biomass (1.1), transportation cost (1.2) and storage costs (1.3). Equations (2)-(9) are restrictions in which constrain (2) establishes that just one biorefinery can be built. Constrain (3) defines storage flows taking into account potential losses and product humidity. Constrain (4) limits biomass that can be purchasable (a percentage \(\alpha_{p,i}\) of total production) which is defined in equation (5) according to physical characteristics in which a particular product can be harvested. Constraint (6) determines biorefinery size, a continuous intake of 12,500 monthly net tons which makes 150,000 annual net tons.

Finally, in order to solve this MILP, the problem was coded in GAMS software language in which CPLEX solver was called. This procedure was run in a standard personal computer, Intel ® Core ™ 2 Quad CPU Q6600 @ 2.40 GHz, and 3.42 GB RAM.
3.4. The model with environmental criterion

It is possible to have an environmental analysis of the problem, assuming some hypothesis about distances and their relationship with pollutant emissions. Thus, in this problem we have focused on externalities derived from biomass transportation, mainly greenhouse gas (GHG) emissions. Since emissions are a function of fuel consumption, the factors affecting fuel consumption will affect emissions as well (Demir et al, 2014). There exist plenty of factors we could consider such as distance, payload, road gradient, speed, driver behaviour… However, due to data limitation we have just selected distance as an environmental impact source. Thus, a new objective function is used instead of using the previous one that minimizes total distance:

$$\min \, Distance = \sum_i \sum_j \sum_t Y_{i,j,t} \cdot dist_{i,j}$$

After the consideration of this objective function, keeping the same constraints depicted in the previous section, we have somehow designed an environmental model easy to solve. We are going to solve it with the same software and computers than it was explained in the previous section.

4. Results

Table 4 shows numerical results corresponding to the vertices of the optimal region with economic criterion whereas Table 5 displays the environmental criterion ones. Absolute results are not shown because of confidential issues; however, change percentages versus the best value (shadow) in each category (i.e. distance, purchase costs, transport costs, storage costs and total costs) are reported. Note that in economic criterion approach, Point 47 gets first position accounting for a 222.77% increasing the distance with respect to the first position in environmental criterion (Point 65). On the environmental criterion side, Point 65 achieves the first position accounting for an 8.72% increasing of total cost to first position in the economic criterion (Point 47). Finally, the reader should take in mind that for each solution displayed in Table 4 and 5, a particular configuration of supply chain is adopted. This implies that, even though two points are in both set of solutions, their purchases, transport and storage polices may be totally different. This is the case, for instance, in point 65 that is the third option in economic approach whereas first in environmental.

Fig. 3 shows graphical representation of the results in which red area represents optimal region with environmental criterion and blue area represents optimal region with economic criterion among all the candidate points (blue dots).

<table>
<thead>
<tr>
<th>#</th>
<th>Distance (km)</th>
<th>Costs (£)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Purchases</td>
<td>Transport</td>
</tr>
<tr>
<td>1</td>
<td>Point 47</td>
<td>222.77%</td>
</tr>
<tr>
<td>2</td>
<td>Point 66</td>
<td>195.67%</td>
</tr>
<tr>
<td>3</td>
<td>Point 65</td>
<td>224.25%</td>
</tr>
<tr>
<td>4</td>
<td>Point 46</td>
<td>201.83%</td>
</tr>
<tr>
<td>5</td>
<td>Point 29</td>
<td>198.23%</td>
</tr>
</tbody>
</table>
Table 5. Results using environmental criterion

<table>
<thead>
<tr>
<th></th>
<th>ENVIRONMENTAL CRITERION</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Distance (km)</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>Point 65</td>
</tr>
<tr>
<td>2</td>
<td>Point 66</td>
</tr>
<tr>
<td>3</td>
<td>Point 85</td>
</tr>
<tr>
<td>4</td>
<td>Point 86</td>
</tr>
<tr>
<td>5</td>
<td>Point 67</td>
</tr>
</tbody>
</table>

Fig. 3. Graphical representation of optimal regions

Optimal region: Environmental Criterion

Optimal region: Economic Criterion
5. Conclusions

Decisions regarding locations are strategic from the business point of view, determining many other decisions at the tactical and operational levels. Thus, providing a good place to set up a facility is critical and demanding. For that reason, researchers have developed a whole branch concerning this issue which is called the Facility Location Problem.

In this paper, it is proposed a framework to help decision makers to make easier their decisions about location by providing, not only a point to set up a plant, but also a whole region where the decision makers can focus on. We illustrated this analysis in the real case of locating a biorefinery in Navarre under economic and environmental criteria. Then, a Mixed Integer Programming model was formulated and solved drawing the regions we were looking for, one for each criterion. In the process, the MILP formulation gave us some key supply chain elements: we were able to identify key products, key crops and key months. Analyzing these results, we can isolate and study the optimal management of purchases, transport and stock policies.

As a result, decision makers can focus on a little but important area to locate their facilities where other subjective factors can be added, those that could not be taken into account in the first step because of their complexity or intangibility (Chan, 2001).

In our example, space where decision maker would focus on is about 0.03% of the total area. Moreover, they can select the criterion they prefer to optimize supply chain: economic or environmental. Noting that, there exist some points along the common line that shares both regions; what implies that decision makers could select a point along this line without caring about economic/environmental criteria. This conclusion is quite a convenient one because by selecting one of those points, decisions about supply chain structure (which will determine economic/environmental criteria) can be put off.

Acknowledgments

This work has been partially supported by the Government of Navarre (Biorefinery Navarra project IIM14196.R11) the Spanish Ministry of Economy and Competitiveness (TRA2013-48180-C3-P and TRA2015-71883-REDT), and the Ibero-American Program for Science and Technology for Development (CYTED2014-515RT0489). Likewise, we want to acknowledge the support received by the CAN Foundation in Navarre, Spain (Grant CAN2015-70473).

References

Chan, Y. 2001 Location theory and decision analysis. Cincinnati: South Western College.


