Abstract

In the next decades, a significant increase is expected in the amounts of CCA-treated wood waste that annually need to be properly disposed. This waste should be recycled only after its remediation, so planning and optimisation of the remediation units location is of major importance.

A location model for CCA-treated wood waste was implemented using Geographic Information Systems (ArcGIS 8.2), with geographic information, namely land use information and the results of a questionnaire sent to Portuguese wood preservation industries. Two different clustering methods (Self-Organizing Maps and K-means) were tested in different conditions to solve the multisource Weber problem using SOMToolbox for MATLAB.

The solutions obtained with the data and with both clustering methods could be used to decide on the location of these plants. SOM provided more robust and reproducible results than K-means, with the disadvantage of longer computing times. The main advantage of K-means, compared to SOM, is the reduced computing time (considering an average of all the runs, the K-means computing time is half the SOM computing time) together with the fact that it allows to obtain the best solutions in the majority of the cases, in spite of bigger variances and more geographical dispersion.

Keywords: CCA-treated wood waste; Location models; Self-Organizing Maps (SOM); K-means

1. Introduction

There is a growing concern about the environmental impacts and the difficulty to dispose treated wood products at the end of their service life. In the next decades, a significant increase is expected in the amounts of treated wood removed from service, e.g. in USA, about 15 million m$^3$ in 2010 and about 18 million m$^3$ in 2020 (McQueen et al., 1998), in Denmark, until 2010, about 200,000 m$^3$ (Ottosen and Christensen, 2006), and in Portugal, we estimate that 90,000 m$^3$ of treated waste wood will need to be properly disposed in 2020.

Typical waste disposal options for treated wood waste, such as incineration and landfill disposal, are becoming more costly because of increasingly strict regulatory requirements and recycling of treated wood waste in wood composites has been progressively increasing. Several remediation methods to extract the preservatives from the wood have been developed (Claussen, 2000; Illman et al., 2000; Ribeiro et al., 2000; Clausen and Kenealy, 2004; El-Fatah et al., 2004; Humar et al., 2004; Oskoui, 2004; Christensen et al., 2006). Due to the transport costs associated with remediation plants, such plants should be located in an optimized location to receive as much treated wood waste as economically feasible.

Location decisions are frequently made at all levels of human organization and are often strategic in nature, involving large sums of capital resources and with long term economic
effects. Location models provide a systematic means by which the decision-maker can explore the various alternatives in order to identify an optimal strategy. There are several types of location problems, with different objectives and types of constraints. All of them are usually very complex and are sub-optimally solved using heuristics. Besides that, location of waste management facilities represents one of the most complicated location decisions faced by environmental policies (Llurdes et al., 2003; Najm and El-Fadel, 2004).

The main objective of this study was the development of a location model to optimise the location of remediation plants for CCA-treated wood waste for further recycling, minimizing costs and considering environmental criteria. Other objectives were: (i) make an estimation of the magnitude of CCA-treated wood waste to be disposed in Portugal; (ii) formulation of the location model; (iii) simulation of different conditions and scenarios with discussion of the results; (iv) testing and comparison of different methods to solve the location problem; (v) production of a final map with the “optimal” locations for the remediation infrastructures; and (vi) identification of further developments.

2. CCA-treated wood waste remediation

Incorporation of CCA waste wood in composite materials such as wood cement, structural composites and particleboard has been identified as the recycling option with the highest potential (Solo-Gabriele et al., 1998; Cooper, 2003). Increased use of waste wood by the composite industry using fiber and small particles appears to be a solution. Treated wood can be recycled in composite materials as it is or after removing contaminants through a remediation process (Cooper, 2003). Different methods for wood preservative removal have been tested with high removal efficiency potential. An extensive review of these methods can be found in Gomes (2004) and in Helsen and Van den Bulck (2005).

In this study, we focus the electrodialytic process, with which our team obtained removal efficiencies of 93% of Cu; 95% of Cr and 99% of As from CCA treated sawdust (Ribeiro et al., 2000; Moreira et al., 2005), remediated wood chips (Velizarova et al., 2002, 2004; Ribeiro et al., 2007) at laboratory scales. At pilot scales, the best remediation efficiencies were 88% of Cu, 82% of Cr and at least 96% of As (Villum-Hansen et al., 2003; Christensen et al., 2004; Pedersen et al., 2005; Christensen et al., 2006).  

3. The Portuguese situation

Limited information was available concerning wood preservation industry and use of preserved wood products, as well as on disposal practices for CCA-treated wood waste in Portugal. For that reason, a questionnaire for the preservation industry was developed, inquiring about the treatment process used by each plant, the quantity of treated wood produced and the type of products manufactured. It was sent to all Portuguese wood preservation industries and the response rate obtained was 45%.

Of the respondent industries, the total treated wood production in 2003 was 75,282 m³. The most common products produced by the treating plants were vineyard stakes (26% of the production) and fence posts (20%) (Fig. 1).

The inventory of CCA-treated wood waste in Portugal was based in the results of the questionnaire and in the treated wood production, exports and imports data available in the National Statistics Institute (INE). This quantification was obtained through a mass balance approach [see Eq. (1)] modified from Solo-Gabriele et al. (1998). The estimation of CCA-treated wood waste production was done for a time span of twenty years, until 2022. The approach used can be represented by the following equation:

\[
\frac{dU}{dt} = P - E + I - D
\]

where \( U \) is the mass of CCA-treated wood in use at any time \( t \); \( dU/\text{df} \) is the rate of change of \( U \) with respect to time (mass/time); \( P \) is the mass of CCA-treated wood produced per unit time; \( E \) is the mass of CCA-treated wood exported per unit time; \( I \) is the mass of CCA-treated wood imported per unit time and \( D \) is the mass of CCA-treated wood disposed per unit time. We assume that the CCA-treated wood production in Portugal is null since the put in force of the Commission Directive 2003/2/EC, of January 6, in June 2004.

The service life of treated wood products is different, according to their use. Lumbers, timbers and fence posts used in residential applications, for example, generally have short service lives (Solo-Gabriele et al., 1998). Cooper (2003) refers that, in USA, one study showed 67% of decks replaced were less than 10 years old and the average service life of decks was only 9 years. Other products, like utility poles and railway sleepers, generally have longer service lives (40—60 years). In Portugal, according to our results, the treated wood is used mainly for agricultural purposes and the treated wood use pattern is not the same that in the USA. So, the following average service lives were assumed: (a) for poles 40 years; (b) for sawn
wood, fence posts and other products 20 years; and (c) for playgrounds and garden furniture 10 years.

The results presented in Fig. 2 show a peak in the annual production of CCA-treated wood waste in 2022, where it will reach approximately 58,000 Mg. This estimation does not take into account changes in style and tastes, nor degradation and uncontrolled combustion of treated wood waste.

4. Model implementation

4.1. Dataset

Location problems are usually formalized by considering a number of demand nodes, each with a given location and weight (amount of resources demanded), a number of facility locations that must be found, and a cost function that takes into account the relation between the demand nodes and the facility locations. In this case, the demand nodes will be producers of waste wood, the facilities will be the remediation units, and the cost function will be the Euclidean distance between the demand nodes and the facilities, considering the amount of wood generated at each demand node.

Taking into account the results of the questionnaire to the preservation industries and land use information, a Geographic Information System (GIS) model was used to identify the places where treated wood was in use and potentially would become a waste.

We have also considered a simpler approach where the demand nodes were the centroids of each census tract of Portugal (NUTS 5) and the demand was represented by the number of inhabitants, since consumer lumber makes up the majority of current and future expected production of CCA treated wood — about 82% of expected waste wood generated in 2010 (Cooper, 2003).

4.2. Formulation of the location model for the remediation units

The problem we are trying to solve consists in selecting an optimal configuration of CCA-treated wood waste remediation units, to minimize transportation costs, considering where this waste is produced.

The problem is subject to the following constraints: (i) demand is concentrated in a set of discrete locations (demand nodes) that have fixed total demands; (ii) every demand node shall be allocated to the closest remediation unit; (iii) there is no interaction between remediation units; and (iv) there are no spatial constraints (forbidden areas, barriers) to the location of the units. It can be located anywhere on the planar coordinate system.

The following assumptions were also made: (i) although the facilities in study are hazardous waste treatment units, they were not considered as obnoxious facilities, like others in literature [see for example Giannikos (1998); Nema and Gupta (1999); Lahdelma et al. (2002); Llurdes et al. (2003)], therefore we assumed that the facilities were “desirable”, since a model to include complex sociologic phenomena like NIMBY (Not In My Back Yard) would be extremely difficult to implement; (ii) we assume for all remediation units the electrodialytic process, that has proven to be efficient in the removal of Cr, Cu and As in pilot scale [see more details in Villumsen (2003) and in Christensen et al. (2006)], although the optimization model is not dependent of the remediation technology chosen; (iii) capital or fixed costs were not considered and we assume that these costs are equal to all remediation units; (iv) the transportation costs between the demand nodes and the remediation units are proportional to the Euclidean distance; and (v) we did not consider the capacity of the remediation units (uncapacitated facility location problem).

This problem is called location-allocation problem or multisource Weber problem (Drezner et al., 2002) and can be formulated as:

\[
\min_{x_i, y_i, z_{ij}} \sum_{j=1}^{p} \sum_{i=1}^{n} z_{ij} \cdot w_j \cdot d_j(x_i, y_i)
\]

subject to \(\sum_{i=1}^{p} z_{ij} = 1; z_{ij} \in [0, 1]; i = 1, 2, \ldots, p; j = 1, 2, \ldots, n\); where \(p\) facilities must be located to satisfy the demand of \(n\) demand nodes, \(x_i, y_i\) denote the coordinates of the \(i\)th facility, \(d_j\) the Euclidean distance from \((x_i, y_i) \) to the \(j\)th user, \(w_j\) the demand (or weight) of the \(j\)th user and \(z_{ij}\) the fraction of this demand which is satisfied from the \(i\)th facility. There is always an optimal solution with all \(z_{ij} \in \{0, 1\}\), i.e., each demand node is satisfied from a single facility. The objective function is neither convex nor concave and may have a large number of local minima (Drezner et al., 2002). This problem was proven to be NP-hard (Non-deterministic Polynomial-time hard) in Megiddo and Supowit (1984). NP-hard problems constitute a complexity class of decision problems that are intrinsically harder than those that can be solved by a non-deterministic Turing machine in polynomial time. No algorithm has been found to solve these problems in a computationally efficient manner, which means that they remain difficult problems to solve optimally. The most well know instance of an NP-hard problem is the TSP (Travelling Salesman Problem),

Fig. 2. Estimated production of CCA-treated wood waste based on expected service life and in the different uses of treated wood products.
where a salesman has to visit at least one time a number of different cities, but try to minimize the distance travelled.

According to the classification scheme proposed by Hamacher and Nickel (1998) the location model can be classified as: \( N/P + \text{alloc}/\Sigma \). \( N \) in this classification means that \( n \) points have to be located. \( P \) represents the Euclidean plane, while alloc in means that we are dealing with an allocation problem where demand nodes are allocated to the facilities. The distance function (\( d_2 \)) is the Euclidean distance. Finally, \( \Sigma \) means that the objective function considered is the classical Weber objective function or minimum objective function (minimizing the sum of weighted distances).

4.3. Methods used

The location-allocation problem solving process can be viewed as properly clustering \( N \) demand nodes into \( P \) groups served by the associated \( P \) remediation units, to achieve the minimum objective.

The K-means algorithm (MacQueen, 1967) is one of the best known classic statistical clustering methods. It generates a specific number of disjoint, non-hierarchical clusters. The K-means method is numerical, unsupervised, non-deterministic and iterative, and is basically a stochastic hill-climbing optimisation technique.

On the other hand, the works of Lozano et al. (1998) and Hsieh and Tien (2004) show that Self-Organizing Maps (SOM) or Kohonen maps may be a potential heuristic method meeting the requirements of both quality of solution and speed of computation. The SOM provides a visual representation of a vector quantification algorithm that places a number of vectors into a high-dimensional input data space in such a way that they approximate the original data patterns in an ordered manner. The SOM provides a visual representation of the data itself. The results for Test 1 are given in Table 1 and 2. When using SOM, we have to provide an output space grid upon which the neurons are set, i.e. the neurons are organized into a dimensional and fully laterally connected topology. In all grids we used a rectangular topology. Training a SOM usually involves two "training phases", in which a specific number of disjoint, non-hierarchical clusters. The K-means method is numerical, unsupervised, non-deterministic and iterative, and is basically a stochastic hill-climbing optimisation technique.

4.4. Experimental conditions

In order to determine if the increasing amounts of CCA-treated wood waste estimated have an effect on the location of the remediation units, we made experiments for three different years: 2010, 2015 and 2020 for the demand nodes that represent waste wood quantities. Regarding the population dataset, only the most recent data available from the 2001 census were used. We considered three different possibilities for the number of remediation units (5, 10 and 15 remediation units) in order to test different SOM grids and to ensure good territory coverage.

Two different tests (Test 1 and Test 2) were made, with different experimental conditions that are summarized in Tables 1 and 2. When using SOM, we have to provide an output space grid upon which the neurons are set, i.e. the neurons are organized into a dimensional and fully laterally connected topology. In all grids we used a rectangular topology. Training a SOM usually involves two "training phases", in which a specific number of disjoint, non-hierarchical clusters. The K-means method is numerical, unsupervised, non-deterministic and iterative, and is basically a stochastic hill-climbing optimisation technique.

In all experiments we used the SOM Toolbox for MATLAB developed at Helsinki University of Technology (Vesanto et al., 2000). All runs of the location model were executed in a 2.00 GHz CPU Pentium 4 PC running Windows XP, using MATLAB R12.

5. Results and discussion

5.1. Results from Test 1

The objective of our location problem was to minimize the average weighted distance to the remediation unit. This distance corresponds to the mean quantization error that is the difference between each data vector and its best matching unit (BMU) or, in other words, the neurons as idealized prototypes of the data and the data itself. The results for Test 1 are given in Table 1 and show that the increase in the number of epochs does not affect considerably the average distance to a remediation unit. An epoch is defined as the processing of all the input patterns once, thus each input pattern will be processed as many times as the number of epochs. In fact, increases of 100 times the number of epochs do not influence greatly the mean quantization error or average distance to a remediation unit. For this reason, in Test 2 we used only 5 training epochs in the first learning phase and 10 training epochs in the second.
The location of the remediation units obtained is very similar in the three years considered (2010, 2015 and 2020). The results obtained with the population data as demand nodes were similar to the ones achieved using the location and quantity of CCA-treated wood waste, particularly in the cases where we consider 10 and 15 remediation units. Geographically, the solutions obtained provide a good coverage of the country and of the demand nodes.

5.2. Results from Test 2

The average distance to a remediation plant in each of the experiments in Test 2 is given in Table 3. The results are presented side by side according to the experimental conditions and to ease the comparison between the two clustering methods tested (SOM and K-means). The minimum average distance for the same experimental conditions (demand nodes and number of remediation units) is marked in bold. We can observe that K-means obtains average distances slightly lower than SOM in 5 experiments (a1, a3, b1, b2 and c1). The best solution is always obtained with K-means (also marked in bold).

The results show that the remediation units locations are alike in the three years considered (2010, 2015 and 2020), as we have also observed in Test 1. Another important remark is the consistency and the reproducibility of the results of the SOM experiments, observed in Test 2. K-means, however, was more unstable. The variance in SOM solutions was appreciably lower, as we can see comparing the standard deviation of SOM and K-means average distance in Table 3 and the remediation units location in Fig. 3.

5.3. Discussion

The two tests have proved the SOM robustness to solve this problem and, consequently, its applicability. The location model was verified and tested, as advocated by Jakeman et al. (2006), considering quantitative criteria and checking...
if the interactions and outcomes of the model are feasible and defensible, given the objectives and the prior knowledge.

In Test 2, the comparison of the two clustering methods tested allows us to conclude that both have advantages and disadvantages. One of the main drawbacks in using K-means is that it is quite sensitive to local minima, and to a certain degree we verified that in our tests. There are several ways of dealing with this problem, the most common of which is to re-initialize the algorithm several times with different seeds (starting values for the K-means clustering procedure) and then choose the best solution. The main advantage of this method, compared to SOM, is the reduced computing time allied to the fact that it allows to obtain best solutions in most cases. SOM provided more robust and reproducible results, with the disadvantage of longer computing times (Fig. 4).

The location of the remediation units obtained is very similar in the three years considered (2010, 2015 and 2020) in both tests. This can be explained by the fact that the location of the demand nodes is the same in those years; the only difference is the amount of waste in each demand node. In fact this constitutes one limitation of the model, since we cannot predict and incorporate the changes in land use with the passage of time.

Comparing the grids used in the two tests, in the case of 10 and 15 remediation units, that is 5 × 1 vs. 10 × 1 and 5 × 3 vs. 15 × 1, respectively, we can see that the results are very similar, particularly for 10 remediation units. It was expected that the two-dimensional grids would force proximity between the remediation units, which can be seen, to a certain level, in CC9 results and was verified in Gomes et al. (2004). The results show that the 15 × 1 grid produces better results than the 5 × 3 grid, i.e., lower average distances to remediation units (about 1 km less). This can be explained by the fact that the “tension” exerted in each unit by the neighbouring units is much higher in the case of the matrix configuration and limits the plasticity of the SOM to adapt to the particular distribution of the dataset (Baçao et al., 2004).

6. Conclusions and further developments

This paper represents a first approach to the emerging issue of CCA-treated wood waste management in Portugal. One of the most important conclusions is the applicability of the tested methods to solve this location problem. The solutions obtained with our data and with both clustering methods could be used to decide on the location of these units. The tested methods could also be used in other similar problems.

SOM provided more robust and reproducible results than K-means, with the disadvantage of longer computing times. The main advantage of K-means, compared to SOM, is the

<table>
<thead>
<tr>
<th>Year</th>
<th>SOM</th>
<th>K-means</th>
</tr>
</thead>
<tbody>
<tr>
<td>2010</td>
<td>a1</td>
<td>43.9 43.8 0.021</td>
</tr>
<tr>
<td></td>
<td>a2</td>
<td>25.1 24.9 0.091</td>
</tr>
<tr>
<td></td>
<td>a3</td>
<td>19.3 18.7 0.164</td>
</tr>
<tr>
<td>2015</td>
<td>b1</td>
<td>43.8 43.8 0.040</td>
</tr>
<tr>
<td></td>
<td>b2</td>
<td>24.9 25.0 0.022</td>
</tr>
<tr>
<td></td>
<td>b3</td>
<td>19.2 18.5 0.220</td>
</tr>
<tr>
<td>2020</td>
<td>c1</td>
<td>44.01 44.0 0.041</td>
</tr>
<tr>
<td></td>
<td>c2</td>
<td>24.9 24.9 0.050</td>
</tr>
<tr>
<td></td>
<td>c3</td>
<td>19.0 18.7 0.347</td>
</tr>
</tbody>
</table>
reduced computing time together with the fact that it allows to obtain best solutions (minimum average distance to a remediation unit) in the majority of the cases, in spite of bigger variances and further geographical dispersion.

Another important conclusion is that the population data can be a good approximation to this kind of location problems where the demand is related to consumption and/or waste generation. In both tests, the solutions obtained with population data were in close proximity to the solutions obtained with CCA-treated wood waste demand nodes, particularly when the number of remediation units is larger. This can be important when information about the location or weighted demand of demand nodes is not available. Population data can then be used as an approximation to the problem.

The approach used can easily be adapted to a more complex (and realistic) formulation of our problem. On one hand, additional sources of preserved wood waste can be added. On the other hand, both algorithms used can apply distances derived from a distance matrix with actual road distances between possible locations, thus using more realistic input data.

As further developments, a more realistic formulation of the problem would be the use of a network model and not a continuous location model, considering the distances on that network instead of the Euclidean distances. In this case, the remediation units would be located in the nodes of the network and we would be dealing with the classic p-median problem that involves the location of facilities in such a manner that the total weighted distance of all users to their closest facility is minimized.

Furthermore, we could consider a capacitated location model, defining constraints about the capacity of the remediation units to treat the amounts of CCA-treated wood waste located in the demand nodes. The location of the wood recycling industries that will use the remediated CCA-treated wood as raw material could also be included in the model formulation, allowing to consider other cost minimization objectives.

Finally, the model developed in this study must be regarded as a decision aid tool that can help decision makers to better understand a situation, rather than as a black box that finds the optimum solution. The model can merely offer a number of solutions that may be considered satisfactory; however, the ultimate decision (and responsibility) always lies with the decision maker(s), which should be involved in the model development, as sustained by Jakeman et al. (2006).

Acknowledgments

Project POCTI/32927/AGR/2000 approved by FCT and POCTI, with FEDER funds, partly supported this study.

References

Impacts of Preservative-Treated Wood Conference, 8–11 February 2004, Florida Center for Environmental Solutions, Orlando, Florida.


