Decomposition Algorithm for Geographic Information System Based Mixed-Integer Linear Programming Models: Application to Sewage Sludge Amendment

ARTICLE *in* INDUSTRIAL & ENGINEERING CHEMISTRY RESEARCH · NOVEMBER 2013

Impact Factor: 2.59 · DOI: 10.1021/ie402471w

4 AUTHORS, INCLUDING:

Andrei Kostin

Research and Development Centre of Rosneft,…

Laureano Jiménez

Universitat Rovira i Virgili

I&EC research
Industrial & Engineering Chemistry Research
Article
pubs.acs.org/IECR

Decomposition Algorithm for Geographic Information System Based Mixed-Integer Linear Programming Models: Application to Sewage Sludge Amendment

P. Vaskan, G. Guillén-Gosálbez,* A. Kostin, and L. Jiménez

Departament d'Enginyeria Química (EQ), Escola Tècnica Superior d'Enginyeria Química (ETSEQ), Universitat Rovira i Virgili (URV), Campus Sescelades, Avinguda Països Catalans, 26, 43007 Tarragona, Spain

ABSTRACT: We present a decomposition strategy for mixed-integer linear programming (MILP) models that are formulated on the basis of geographic information system (GIS) data. Our algorithm relies on decomposing the MILP into two levels, a master problem and a slave problem between which we iterate until a termination criterion is satisfied. The former is constructed using a K-clustering statistical aggregation method that reduces the computational burden of the model. The solution of this level is used to guide the search in the slave model. A case study that addresses the optimal design of sewage sludge amendment in Catalonia (NE of Spain) is introduced to illustrate the capabilities of the proposed approach.

1. INTRODUCTION

Geographic informational systems (GIS) were initially developed as a tool for storing and displaying all forms of geographically referenced information. In the recent past, however, there has been a growing interest on the application of GIS in the solution of various social and economic problems. Particularly, GIS has been used in the context of spatial decision analysis for the assessment of potential locations for different types of systems considering various inputs simultaneously, with a recent growing interest placed on its application to environmental problems. As an example, Nadalet al.,[1] Poggio and Vrščaj[2] investigated the use of GIS for human health assessment, whereas Johnson et al.[3] Schriever and Liess[4] applied GIS in the assessment of the ecological exposure and environmental risk of several systems.

GIS can be combined with multicriteria decision analysis (GIS-MCDA) to address problems in which different (typically conflictive) criteria must be accounted for in the analysis. Malczewski[5] investigated the use of GIS-based tools in land-use suitability analysis, whereas Passuello et al.[6] applied GIS and MCDA to the management of sewage sludge.

The capabilities of GIS and spatial analysis can be further enhanced through its integration with optimization tools. Grabaum and Meyer[7] investigated the use of GIS to support decision making in planning problems. Wang et al.[8] developed a GIS model to identify the best location for future land uses in the Lake Erhai basin in China. Mapa et al.[9] combined GIS and mathematical modeling for the solution of location-allocation problems arising in the management of education facilities. Jung et al.[10] integrated GIS and optimization tools for the effective control of parcel delivery services. Marcoulaki et al.[11] developed an integrated framework based on stochastic optimization and GIS for the design of pipeline systems.

One problem in which the combined use of GIS and mathematical programming holds good promise is the treatment of sewage sludge in agricultural areas. The production of sewage sludge (SS) has grown rapidly during the last years, mainly due to the increase of the world population. Despite recent advances, the question on how to treat the SS still remains open. One effective method for this is to reuse it as a fertilizer in the agricultural sector, an alternative encouraged by the European Community, which promotes the recycling of organic matter and nutrients to soils.[12] Identifying the best agricultural areas for SS amendment is challenging because this strategy shows not only benefits to both soil and crops but also disadvantages due to the potential contamination of the fields.

GIS tools for land classification are well suited for this problem, as they allow identifying the best regions for SS amendment from information available in spatial databases.[6] These tools are mainly descriptive; that is, they provide valuable information about the system, but no guidelines on how to solve the underlying problem. In this general context, there is a strong motivation for developing systematic tools that integrate GIS and optimization to facilitate decision-support in this area.

Vaskan et al.[13] investigated the combined use of GIS and MILP (mixed-integer linear programming) for identifying optimal agricultural areas for sewage sludge amendment in the area of Catalonia. The combined use of GIS and optimization tools led to complex MILP models due to the spatially explicit nature of the problems addressed. In these MILPs, the decision variables are defined for every pixel of the GIS map, thereby giving rise to mathematical models with a very large number of variables and constraints. In our previous paper,[13] we overcame this limitation by considering a GIS map with low resolution. Although this strategy simplifies the calculations, it offers no guarantee of convergence to the global optimum of the original problem (i.e., the one defined for the original map with high resolution).

In work, we propose a rigorous decomposition algorithm for the efficient solution of GIS-based MILPs that exploits their

Received: July 30, 2013
Revised: October 10, 2013
Accepted: October 18, 2013
Published: October 18, 2013

particular structure. This strategy allows handling models based on GIS maps with high resolution. Our approach is based on decomposing the problem into two hierarchical levels between which the algorithm iterates until a termination criterion is satisfied. We illustrate the capabilities of our strategy via its application to the optimal location of agricultural areas for sewage sludge amendment. Numerical results show that our approach achieves reductions of orders of magnitude in CPU time (as compared to the full space GIS-based MILP) while still yielding near optimal solutions.

The article is organized as follows. Section 2 formally states the problem. Section 3 introduces a rigorous decomposition algorithm to tackle GIS-based MILP problems. Some numerical results are presented and discussed in section 4, and the conclusions of the work are finally drawn in section 5.

2. PROBLEM STATEMENT

We consider as a test bed to illustrate the capabilities of our approach the optimal allocation of agricultural areas for sewage sludge (SS) amendment. We next formally state the problem of interest before describing the MILP derived to solve it. To this end, we consider a superstructure like the one depicted in Figure 1. Given are a set of wastewater treatment plants (WWTP) and a set of agricultural areas that can receive the sludge sent by the plants. Each field is described by coordinates expressed in meters, and it features an acceptable capacity (CAPar(j)) in tons per year of sludge, and an application cost (in euro $ton^{-1}year^{-1}$). In addition, every plant is characterized by a pair of Cartesian coordinates and a total production of SS per year (CAPpl(i)). The goal of the analysis is to determine the optimal distribution of SS production among a set of agricultural areas so that the total cost is minimized.

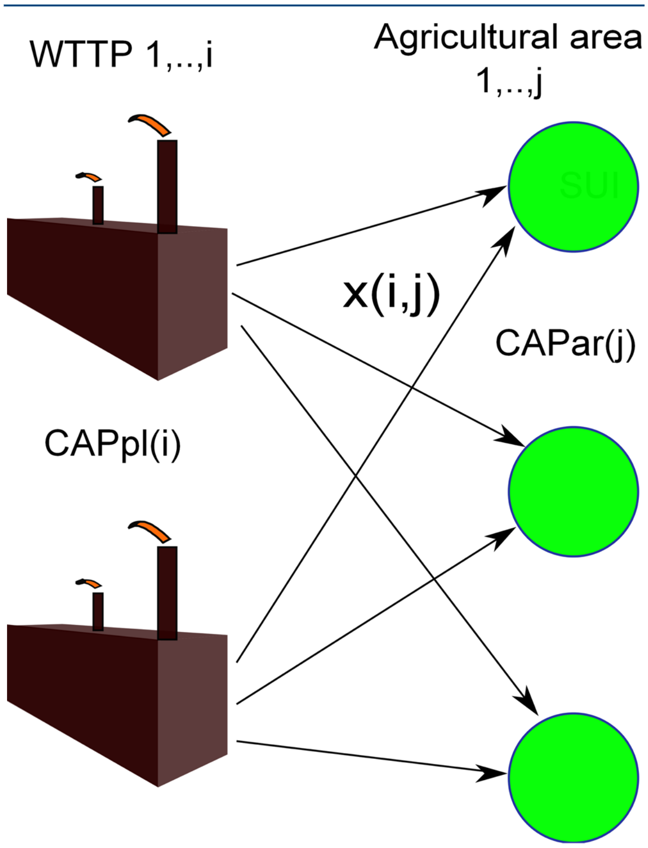

Figure 1. Superstructure of the supply chain problem.

3. MODEL FORMULATION

The MILP used in this work, which is taken from Vaskan et al.,[13] is based on the superstructure showed in Figure 1. The MILP seeks to determine the optimal flows to be established between the wastewater treatment plants and the agricultural fields considering the cost as unique criterion. This is a major difference with respect to the original bicriteria model that optimized the cost along with the environmental impact. For the sake of completeness of this work, we next describe the equations of the MILP. Further details can be found in our previous publication.

3.1. MILP Model.

3.1.1. Capacity Limitations. We define the continuous variable $x(i,j)$, which denotes the amount of SS sent from plant i to agricultural soil j. The total amount of sludge sent from a plant to the fields is equal to the plant capacity (represented by the parameter CAPpl), as shown in the following equation:

$$\sum_j x(i,j) = \mathrm{CAPpl}(i) \qquad \forall\, i \tag{1}$$

$$\sum_j x(i,j) \le \mathrm{CAPar}(j) \qquad \forall\, j \tag{2}$$

Furthermore, the amount of SS sent to an agricultural field must not exceed its capacity (parameter CAPar(j)) The amount of SS sent from a plant to a field must lie within lower and upper bounds if a transportation link is established between them, and should be zero otherwise:

$$\underline{X(i,j)}\, z(i,j) \le x(i,j) \le \overline{X(i,j)}\, z(i,j) \qquad \forall\, j, i \tag{3}$$

In this equation $z(i,j)$ is a binary variable that represents the existence of a transportation link between plant i and field j. $z(i,j)$ equals 1 if a transportation link is established, and 0 otherwise. In the same equation, $\underline{X(i,j)}$ and $\overline{X(i,j)}$ denote the minimum and maximum allowable flows of SS, respectively, that can be transported between i and j. The total amount of SS sent to field j must lie within lower and upper limits, provided the field is used for SS amendment:

$$\underline{Y(j)}\, y(j) \le \sum_i x(i,j) \le \overline{Y(j)} y(j) \qquad \forall\, j \tag{4}$$

In this equation, $y(j)$ is a binary variable that takes the value of 1 if area j is used and 0 otherwise, and $\underline{Y(j)}$ and $\overline{Y(j)}$ are lower and upper bounds, respectively, on the total amount of SS disposed on field j.

3.1.2. Objective Functions. The model minimizes the total cost (TC), which is obtained as follows:

$$\mathrm{TC} = \mathrm{TRC} + \mathrm{AC} \tag{5}$$

Here TRC represents the transportation cost from the SS plants to the fields (TRC) and AC is the application cost associated with SS in the fields. The transportation cost is given by

$$\mathrm{TRC} = \sum_i \sum_j \mathrm{tc}\, x(i,j)\, \lambda(i,j) \tag{6}$$

where $\lambda(i,j)$ represents the distance between plant i and field j in kilometers, tc is the cost of transporting 1 ton of SS per km of distance (euro ton^{-1} km^{-1}), and $x(i,j)$ denotes the amount of SS transported from i to j expressed in ton per year (ton year^{-1}).

The application cots are calculated from the amount of SS disposed as follows:

$$\mathrm{AC} = \sum_i \sum_j \mathrm{ac}(j)\, x(i, j) \tag{7}$$

where ac(j) represents the application cost of 1 ton of SS per year (euro ton^{-1} year^{-1}).

The overall MILP can be expressed in compact form as follows:

$$\begin{aligned} &\min \mathrm{TC}\ (M) \\ &\text{s.t.} \quad h(x, y) = 0 \\ &\qquad g(x, y) \leq 0 \\ &\qquad x \subset \mathbb{R},\ y \subset 0, \mathbf{1} \end{aligned}$$

where x are continuous variables and y binary ones. Functions $h(x,y)$ are equality constraints that model mass balances, whereas $g(x,y)$ are inequality constraints that define capacity limitations. This MILP tends to be large because decision variables need to be defined for every pixel of the GIS map. GIS maps typically show thousands of pixels (or even hundreds of thousands). Hence, they might lead to optimization models showing a large computational burden. In the section that follows, we introduce a method to expedite this type of GIS-based MILPs.

4. BILEVEL DECOMPOSITION

In the MILP model presented above, the number of potential areas for SS amendment depends on the number of pixels in the GIS map. In other words, we consider the option of sending the SS to as many different locations as pixels contained in the GIS map. As an example, for a GIS map with 13 984 pixels, we would define an MILP containing 55 940 continuous variables, 69 920 binary variables, and 153 832 equations. The model size is hence quite sensitive to the number of pixels, which can grow rapidly as we increase the map resolution. More precisely, the total number of binary variables (BV) can be expressed as follows:

$$\mathrm{BV} = |I| + |IJ| \tag{8}$$

where $|I|$ is the cardinality of the set of plants and $|J|$ is the cardinality of the set of fields.

To expedite the solution of this GIS-based MILP, we propose an algorithm that decomposes the model into two hierarchical levels, a master and a slave level, between which we iterate until a stopping criterion is reached. The scheme of the bilevel algorithm is shown in Figure 2. The master MILP contains the same equations of the original MILP, but it is defined for a smaller number of (aggregated) pixels. This MILP identifies the aggregated regions where SS should be sent and provides in turn a lower bound on the cost (LB problem). In the lower lever, we disaggregate the aggregated pixels and remove those regions discarded by the master MILP. This slave MILP provides an upper bound on the cost (UB problem). After the slave MILP is solved, an integer cut is added to the master MILP to remove those solutions explored so far in previous iterations. The master and slave MILPs are then solved iteratively until a termination criterion is reached. We describe in detail the two levels of the algorithm in the ensuing sections.

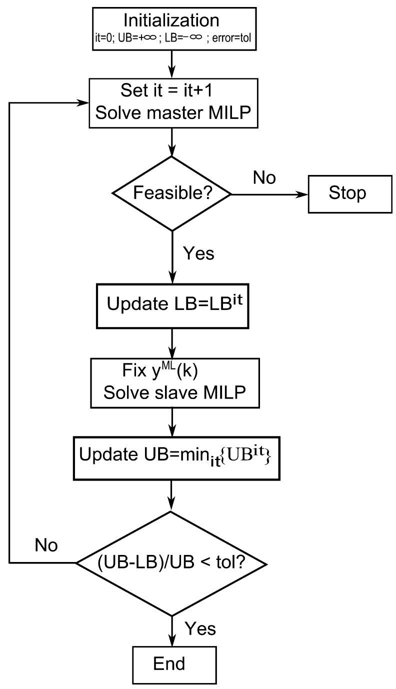

Figure 2. Flow chart for the bilevel decomposition algorithm.

4.1. Master MILP: k-Means Clustering Method.

As already mentioned, the master MILP is constructed by aggregating pixels in the original model. To this end, we use a k-means clustering method. The k-means clustering is a partitioning method that aggregates data into clusters such that observations within each cluster are as close to each other as possible and as far from observations in other clusters as possible. In the context of our application, each observation corresponds to a pixel with a given location in the space of coordinates. This makes such a clustering aggregation very useful for spatially explicit problems.

Each cluster is defined by the centroid or center and its member objects (pixels). The goal is to determine the centroid with the minimum sum of distances from all objects in that cluster. k-means uses an iterative algorithm that minimizes the sum of distances from each object to its cluster centroid, over all the clusters. This algorithm moves objects between clusters until the sum cannot be decreased any further. The result is a set of clusters that are as compact and as well-separated as possible.

To clarify this technique we consider a simple example with 15 fields, each one defined by given coordinates. These fields are aggregated into three clusters with minimum total sum of distances between centroids and fields. After applying the k-means strategy, we identify three clusters containing different numbers of pixels (Figure 3). Further details on this method can be found in Hartigan and Wong[14] and Kanungo et al.,[15] and implementation details are available in Matlab.[16]

After performing the aggregation, we slightly modify the original MILP to accommodate the new aggregated clusters. Let $JK(k)$ be the set of pixels j contained in the aggregated cluster k. To this end, we use the following equations:

$$\mathrm{ac}^{\mathrm{ML}}(k) = \min_{j \subset JK} \mathrm{ac}(j) \tag{9}$$

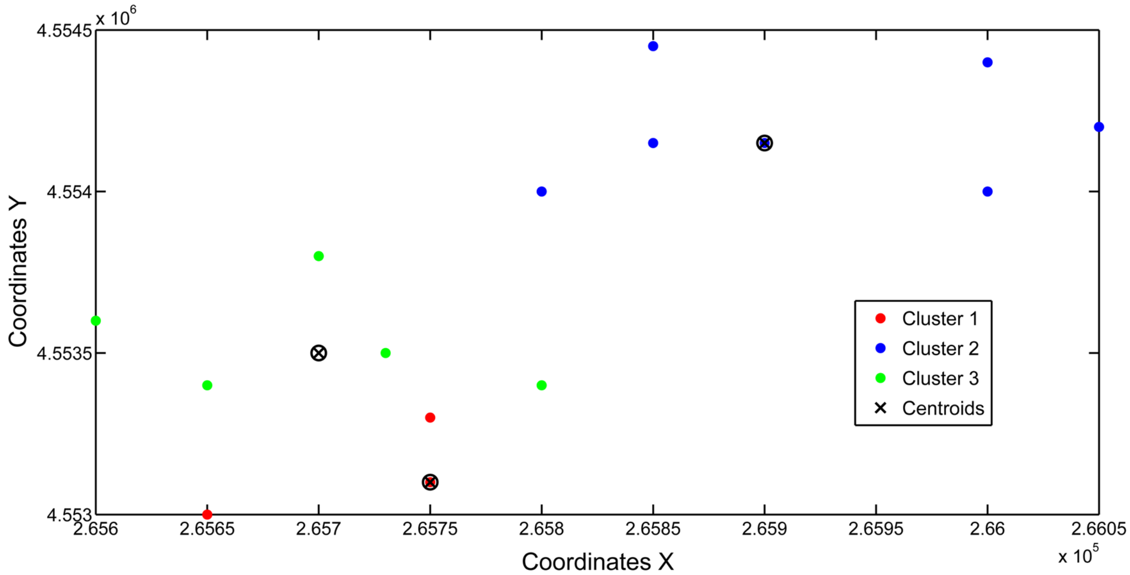

Figure 3. k-means aggregation example.

$$\lambda^{ML}(i, k) = \min_{j \subset JK} \lambda(i, j) \tag{10}$$

$$\mathrm{CAPar}^{ML}(k) = \sum_{j \subset JK} \mathrm{CAPar}(j) \tag{11}$$

Hence, the values of the application cost ($\mathrm{ac}^{ML}(k)$) and distance ($\lambda^{ML}(i,k)$) of an aggregated pixel (i.e., k clusters) in the master MILP correspond to the minimum values among the pixels contained in the cluster (note that in this notation ML stands for master level). Furthermore, the capacity of the aggregated pixel ($\mathrm{CAPar}^{ML}(i,k)$) is given by the sum of capacities of all of the pixels of the cluster. Because of the manner in which it is constructed, the master MILP is guaranteed to provide a rigorous lower bound on the total cost. The master MILP identifies in a systematic and rigorous manner the aggregated pixels that will receive the sludge from the treatment plants. As will be shown in the next section, this information is used to reduce the number of variables and constraints of the slave MILP.

Note that the master MILP is defined for the k aggregated clusters, rather than for the j fields. Apart from this, the master MILP is identical to the MILP model described above. It therefore includes the following equations:

$$\sum_k x^{ML}(i, k) = \mathrm{CAPpl}(i) \qquad \forall\ i \tag{12}$$

$$\sum_i x^{ML}(i, k) \le \mathrm{CAPar}^{ML}(k) \qquad \forall\ k \tag{13}$$

$$\underline{X^{ML}(i, k)}\, z^{ML}(i, k) \le x^{ML}(i, k) \le \overline{X^{ML}(i, k)}\, z(i, k) \quad \forall\ k, i \tag{14}$$

$$\underline{Y^{ML}(k)}\, y^{ML}(k) \le \sum_i x^{ML}(i, k) \le \overline{Y^{ML}(k)}\, y^{ML}(k) \quad \forall\ k \tag{15}$$

The model minimizes the total cost (TC^{ML}), which is obtained as follows:

$$\mathrm{TC}^{ML} = \mathrm{TRC}^{ML} + \mathrm{AC}^{ML} \tag{16}$$

$$\mathrm{TRC}^{ML} = \sum_i \sum_k \mathrm{tc}\, x^{ML}(i, k)\, \lambda^{ML}(i, k) \tag{17}$$

$$\mathrm{AC}^{ML} = \sum_i \sum_k \mathrm{ac}^{ML}(k)\, x^{ML}(i, k) \tag{18}$$

4.2. Slave MILP. As already mentioned, the master problem identifies the aggregated pixels where SS should be disposed. In the slave problem, we disaggregate this information assuming that the pixels belonging to the active clusters of the master MILP (i.e., those for which $y^{ML}(k)$ equals 1 in the master MILP) can be utilized for SS amendment. In contrast, if a pixel defined in the slave MILP does not belong to any of the aggregated pixels selected in the master problem, then it is removed from the slave model. The slave MILP contains therefore the same equations of the original MILP, but fewer constraints and variables, because pixels that are not selected in the master MILP are omitted in the formulation. Let JAK^{it} be the set of pixels j contained in the aggregated clusters k that are active in the solution of the mater MILP (those for which $y^{ML}(k)$ takes a value of one) in iteration it of the algorithm. With this notation, the slave MILP includes the following constraints:

$$\sum_{j \subset \mathrm{JAK}} x(i, j) = \mathrm{CAPpl}(i) \qquad \forall\ i \tag{19}$$

$$\sum_{j \subset \mathrm{JAK}} x(i, j) \le \mathrm{CAPar}(j) \qquad \forall\ j \tag{20}$$

$$\underline{X(i, j)}\, z(i, j) \le x(i, j) \le \overline{X(i, j)}\, z(i, j) \qquad \forall\ j \subset \mathrm{JAK}, i \tag{21}$$

$$\underline{Y(j)}\, y^{ML}(j) \le \sum_i x(i, j) \le \overline{Y(j)}\, y^{ML}(j) \qquad \forall\ j \subset \mathrm{JAK} \tag{22}$$

The model minimizes the total cost (TC), which is obtained as follows:

$$\mathrm{TC} = \mathrm{TRC} + \mathrm{AC} \tag{23}$$

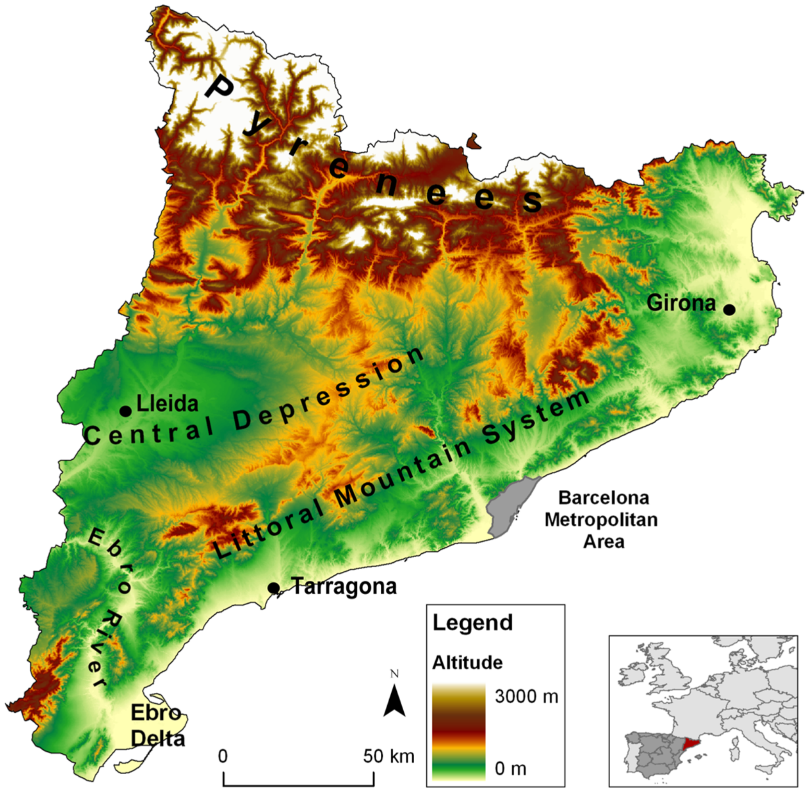

Figure 4. Map of Catalonia.

$$\mathrm{TRC} = \sum_{i} \sum_{j \subset \mathrm{JAK}} \mathrm{tc}x(i, j)\, \lambda(i, j) \tag{24}$$

$$\mathrm{AC} = \sum_{i} \sum_{j \subset \mathrm{JAK}} \mathrm{ac}(j)\, x(i, j) \tag{25}$$

After calculating the master MILP, we next solve the slave MILP minimizing the total cost and fixing the values of $y(j)$ to the values provided by the master problem (thereby disaggregating the information obtained therein). Note that the slave subproblem provides an upper bound on the cost. This is because the slave MILP contains the same equations of the original MILP, but it is solved in a reduced domain with a smaller feasible space. Finally, after solving the slave MILP, we derive an integer cut to exclude solutions identified so far in previous iterations.

Hence, the master model works with data from the k-means clustering aggregation, whereas the slave model works with disaggregated clusters obtained from the solution of the master model. Both problems (slave and master) are solved iteratively until a termination criterion is satisfied.

In summary, the steps of the bilevel decomposition algorithm are the following:

1. Aggregate the data into the desired number of clusters using the k-means clustering aggregation.
2. Set iteration count it = 0, upper bound UB = +∞, lower bound LB = −∞, and tolerance error = tol.
3. Set it = it + 1. Solve the MILP master problem (LP):
   If problem (LP) is infeasible, then stop.
   Otherwise, set the current lower bound to $\mathrm{LB} = \mathrm{LB}^{\mathrm{it}}$
4. Disaggregate the pixels of the master MILP and fix variables $y^{\mathrm{ML}}(k)$ (eq 14) obtained from step 2, in the slave problem and solve it.
5. Update the upper bound (UB) to $\mathrm{UB} = \min_{\mathrm{it}}\{\mathrm{UB}^{\mathrm{it}}\}$ where $\mathrm{UB}^{\mathrm{it}}$ represents the objective function value associated with the optimal solution in iteration it.
6. Check the convergence criteria:
   a. If $(\mathrm{UB} - \mathrm{LB})/\mathrm{UB} < \mathrm{tol}$, then stop. The solution corresponding to UB (i.e., the solution of the slave model in the iteration with minimum cost) satisfies the termination criterion (i.e., it can be regarded as optimal within the predefined optimality gap).
   b. Otherwise, go to step 3.

5. CASE STUDY: CATALONIA

We apply our method to a case study based on Catalonia. Catalonia is a province of Spain located in the Northeastern part of the Iberian Peninsula. The total area of Catalonia is 32 114 km^2, with an agricultural area available for cultivation of near 5000 km^2. The relief is very different from the mountains on the north, to the flat at the center and the coast (Figure 4). The Mediterranean climate and precipitation levels favors the existence of agricultural sectors. The Catalonian agriculture is mainly based on the production of wine, wheat, rice, barley, olive, grapes, fruits, nuts, and vegetables. The total population of Catalonia is near 7 350 000 people. It is divided into four provinces: Barcelona, Tarragona, Girona, and Lleida, with a population of 5 416 447, 788 895, 731 864, and 426 872 people, respectively. The production of sewage sludge (SS) has been growing in the recent past and near 83% of the total production of SS was applied on agricultural soils.[6]

6. RESULTS AND DISCUSSION

We illustrate the capabilities of our approach through its application to a case study based on Catalonia. The input data for the MILP were taken from ref 6. We consider three different levels of aggregation in the problem (all of them for the same agricultural area of 505 176 ha): 126 294, 31 517, and 13 984 pixels, each one with a surface of 4, 16, and 36 ha, respectively. The MILP model and the bilevel algorithm were both implemented in GAMS and solved with CPLEX on an AMD Athlon 2.99 GHz, 3.49 GB of RAM machine. The optimality gap set for CPLEX was 3%, whereas the bilevel algorithm was executed considering a tolerance (relative error between the lower and upper bounds) of 3%.

We should make a remark concerning the use of models with a large number of pixels. In general, it is desirable to include as many pixels as possible in the analysis, because the decisions involved might be rather sensitive to the scale of the map. Moreover, the main characteristics of the map areas are in some cases very sensitive to the scale, which motivates the need to define a large number of pixels for a better assessment of the performance of each alternative. Hence, back to our example, it is more convenient for an adequate analysis to consider 4 ha for every pixel (126 294 pixels in total) rather then 36 ha (13 984 pixels in total). Unfortunately, this leads to more complex problems.

Particularly, we solved a set of problems of increasing complexity involving a different number of cities in Catalonia. We consider first the location of WWTPs in four main cities: Barcelona, Girona, Tarragona, and Lleida, and then solve the same problem considering additional locations (i.e., Terrasa, Vic, Amposta, and Montblanc).

Numerical results for different levels of complexity are presented in Tables 1−3. The goal is to illustrate the performance of the algorithm as compared to the full-space method. The objective in these problems was to minimize the cost as single objective function. In all of the cases, the optimality gap set for the bilevel algorithm (i.e., 3%) was reached in one single iteration.

Table 1. Computational Results for 13 984 Pixels with GAP 3% (CPLEX)

	binary variables	continuous variables	equations	time (s)	cost (euro)
		Four Plants			
full space	69 920	55 940	153 832	5.58	6 425 371
bilevel (UB)	12 980	10 386	28 562	0.81	6 436 821
LB	6 995	5 598	15 395	0.42	6 257 572
		Five Plants			
full space	83 904	69 924	181 801	7.66	5 381 347
bilevel (UB)	15 540	12 952	33 677	1.2	5 393 824
LB	8 394	6 997	18 194	0.55	5 214 386
		Six Plants			
full space	97 888	83 908	209 770	9.64	5 138 030
bilevel (UB)	18 102	15 518	38 798	1.23	5 149 251
LB	9 793	8 396	20 993	0.69	4 971 599
		Seven Plants			
full space	111 872	97 892	237 739	10.56	5 055 427
bilevel (UB)	20 824	18 223	44 260	1.64	5 070 094
LB	11 192	9 795	23 792	0.8	4 889 444
		Eight Plants			
full space	125 856	111 876	265 708	11.2	5 000 648
bilevel (UB)	23 652	21 026	49 942	1.8	5 015 416
LB	12 591	11 194	26 591	0.88	4 828 968

Table 2. Computational Results for 31 517 Pixels with GAP 3% (CPLEX)

	binary variables	continuous variables	equations	time (s)	cost (euro)
		Four Plants			
full space	157 585	126 072	346 695	14.44	6 392 357
bilevel (UB)	28 795	23 038	63 355	1.91	6 399 849
LB	15 760	12 610	34 678	1.14	6 268 917
		Five Plants			
Full space	189 102	157 589	409 730	18.19	5 350 280
bilevel (UB)	43 278	36 067	93 776	2.94	5 358 057
LB	18 912	15 762	40 983	1.2	5 227 879
		Six Plants			
full space	220 619	189 106	472 765	22.03	5 107 400
bilevel (UB)	40 152	34 418	86 048	3.5	5 115 116
LB	22 064	18 914	47 288	1.63	4 985 148
		Seven Plants			
full space	252 136	220 623	535 800	24.97	5 021 803
bilevel (UB)	46 248	40 469	98 286	4.36	5 031 313
LB	25 216	22 066	53 593	3.38	4 898 018
		Eight Plants			
full space	283 653	252 140	598 835	29.55	4 965 311
bilevel (UB)	52 533	46 698	110 913	4.59	4 978 131
LB	28 368	25 218	59 898	1.86	4 839 344

We start by generating results for the lowest resolution (i.e., 13 984 pixels). As observed in Table 1, the proposed approach shows better numerical performance than the full-space method. First, the computational time is less for our bilevel strategy because in every level of the algorithm we have less number of equations and variables than in the full space problem. Second, the value of the objective function obtained from the bilevel strategy is very close to the value generated by the full space problem. Note that although we fixed an optimality gap of 3% for CPLEX, we obtain indeed the global optimum in all the runs (i.e., we solved again fixing a 0% gap, and we got the same results). On average, our bilevel algorithm reduces the CPU time by a factor of almost 1 when compared to the full space approach.

We next increase the map resolution and repeat the calculations (Table 2). As seen, we get very similar results as in the previous case. The CPU times of both methods increase but are still within low limits. In addition, the bilevel method still outperforms the full-space one, achieving almost 1 order of reduction in the CPU time compared to the full space method.

Finally, Table 3 shows the results corresponding to the maximum map resolution. As seen, the full space MILP gets intractable when we increase further the number of pixels (i.e., 126 294 pixels), which leads to a prohibitive computational burden. As observed, the full space method must solve an MILP with 631 470 binary variables and 505 192 continuous variables, which turns out to be intractable. In contrast, our bilevel strategy keeps the model size tractable and can thus handle large problems in reasonable CPU times (CPU time around 40−50 s), while still providing near optimal (i.e., optimality gap of 3%) solutions. Hence, our approach allows for

Table 3. Computational Results for 126 294 Pixels with GAP 3% (CPLEX)

	binary variables	continuous variables	equations	time (s)	cost (euro)
		Four Plants			
full space	631 470	505 192	1 389 266	out of memory	
bilevel (UB)	114 280	91 426	251 422	10.97	6 391 374
LB	63 150	50 522	138 936	5.06	6 326 472
		Five Plants			
full space	757 764	631 486	1 641 858	out of memory	
bilevel (UB)	136 728	113 942	296 251	15.16	5 349 945
LB	75 780	63 152	164 197	6.89	5 284 988
		Six Plants			
full space	884 058	757 780	1 894 450	out of memory	
bilevel (UB)	158 879	136 184	340 463	18.27	5 109 052
LB	88 410	75 782	189 458	8.02	5 044 526
		Seven Plants			
full space	1 010 352	884 074	2 147 042	out of memory	
bilevel (UB)	183 064	160 183	389 020	20.36	5 025 745
LB	101 040	88 412	214 719	9.14	4 959 139
		Eight Plants			
full space	1 136 646	1 010 368	2 399 634	out of memory	
bilevel (UB)	206 550	183 602	436 060	32.08	4 970 541
LB	113 670	101 042	239 980	14.28	4 901 357

the solution of MILPs based on maps with higher resolution while keeping the CPU time low.

As seen also in Tables 1−3, the complexity of the model, and therefore the CPU time spent in its solution, increases with the number of plants as well as the number of pixels, being the second factor the most critical one.

Finally, we should note that for this particular problem the objective function does not improve significantly as we increase the map resolution (approximately 1%, which corresponds to 50 000 euros in absolute values). Note, however, that in general it is not possible to predict an exact interval within which the objective function will fall when we increase the number of pixels. Hence, because this difference might be much larger in other problems, it is always recommended to use the highest map resolution available.

7. CONCLUSIONS

This work has proposed a decomposition method for MILPs that are formulated on the basis of GIS maps. Our approach is based on a bilevel decomposition strategy that makes use of a clustering aggregation algorithm. In the first level, we solve a lower bounding problem to identify the aggregated pixels that will receive the sludge from the treatment plants. In the upper bounding problem, this information is disaggregated to obtain a rigorous upper bound on the cost.

We applied our method to a case study based on sewage sludge amendment in Catalonia. Numerical examples showed that our tool provides near optimal solutions in a fraction of the CPU time required by the full space model. Our method thus solves in an efficient manner large-scale MILPs based on GIS maps with a high resolution. The strategy presented herein is general enough to be applied to similar MILPs used in spatial-decision analysis that address problems arising in chemical and process industries.

■ AUTHOR INFORMATION

Corresponding Author

*G. Guillén-Gosálbez: e-mail, gonzalo.guillen@urv.cat.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The authors acknowledge the financial support of the following institutions: Spanish Ministry of Education and Science (CTQ2009-14420-C02, CTQ2012-37039-C02, DPI2012-37154-C02-02 and ENE2011-28269-CO3-03), Spanish Ministry of External Affairs (projects PHB 2008-0090-PC) and the Catalonian Government (FI programs).

■ APPENDIX

Indices

i wastewater treatment plants
j agricultural soils
k aggregated soils (i.e., clusters)

JAK^{it} Sets

I i:i is a wastewater treatment plants
J j:j is the agricultural soil contained in a pixel
JK k: set of pixels j contained in the aggregated cluster k
JAK^{it} it: set of pixels j contained in the aggregated clusters k that are active in the solution of the mater MILP (those for which $y^{ML}(k)$ takes a value of one) in iteration it of the algorithm

Parameters

CAPpl(i) capacity of plant i
CAPar(j) capacity of field/pixel j
$\underline{X(i,j)}$ minimum allowable flows of sewage sludge from plant i to field j
$\overline{X(i,j)}$ maximum allowable flows of sewage sludge from plant i to field j
$\underline{Y(j)}$ lower bound on the total amount of SS disposed on field j
$\overline{Y(j)}$ upper bound on the total amount of SS disposed on field j
tc unitary transportation cost from plants to fields
$\lambda(i,j)$ distance between plant i and field j

ac(j) application cost of SS in field/pixel j

$ac^{ML}(k)$ application cost defined for the aggregated pixel k (master level)

$\lambda^{ML}(i,k)$ distance between plant i and cluster/aggregated pixel k (master level)

$CAPar^{ML}(k)$ capacity of the aggregated pixel k (master level)

$\underline{X^{ML}(i, k)}$ minimum allowable flows of sewage sludge from plant i to cluster k (master level)

$\overline{X^{ML}(i, k)}$ maximum allowable flows of sewage sludge from plant i to cluster k (master level)

$\underline{Y^{ML}(k)}$ lower bound on the total amount of SS disposed on cluster k (master level)

$\overline{Y^{ML}(k)}$ upper bound on the total amount of SS disposed on cluster k (master level)

Variables

$x(i,j)$ amount of SS sent from plant i to agricultural soil j

$z(i,j)$ binary variable that represents the existence of a transportation link between plant i and field j

$y(j)$ binary variable that represents the use of field j

$x^{ML}(i,k)$ amount of SS sent from plant i to cluster k (master level)

$z^{ML}(i,k)$ binary variable that represents the existence of a transportation link between plant i and cluster k (master level)

$y^{ML}(k)$ binary variable that represents the use of a cluster k (master level)

TRC total transportation cost

AC application cots of SS

TC total cost

TRC^{ML} total transportation cost (master level)

AC^{ML} application cots of SS (master level)

TC^{ML} total cost (master level)

REFERENCES

(1) Nadal, M.; Kumar, V.; Schuhmacher, M.; Domingo, J. L. Definition and GIS-based characterization of an integral risk index applied to a chemical/petrochemical area. *Chemosphere* **2006**, *64*, 1526−1535.

(2) Poggio, L.; Vrščaj, B. A gis-based human health risk assessment for urban green space planningAn example from Grugliasco (Italy). *Science of the total environment* **2009**, *407*, 5961−5970.

(3) Johnson, M. S.; Korcz, M.; von Stackelberg, K.; Hope, B. K. *Decision Support Systems for Risk-Based Management of Contaminated Sites*; Springer, 2009; pp 1−19.

(4) Schriever, C. A.; Liess, M. Mapping ecological risk of agricultural pesticide runoff. *Sci. Total Environ.* **2007**, *384*, 264−279.

(5) Malczewski, J. GIS-based land-use suitability analysis: a critical overview. *Progress in planning* **2004**, *62*, 3−65.

(6) Passuello, A.; Cadiach, O.; Perez, Y.; Schuhmacher, M. A spatial multicriteria decision making tool to define the best agricultural areas for sewage sludge amendment. *Environment international* **2012**, *38*, 1−9.

(7) Grabaum, R.; Meyer, B. C. Multicriteria optimization of landscapes using GIS-based functional assessments. *Landscape and Urban Planning* **1998**, *43*, 21−34.

(8) Wang, X.; Yu, S.; Huang, G. Land allocation based on integrated GIS-optimization modeling at a watershed level. *Landscape and Urban Planning* **2004**, *66*, 61−74.

(9) Mapa, S. M. S.; Lima, R. d. S.; Mendes, J. F. *Combining Geographica Information Systemas (GIS) and mathematical modeling to location-allocation problems in education facilities management*. 2007,

(10) Jung, H.; Lee, K.; Chun, W. Integration of GIS, GPS, and optimization technologies for the effective control of parcel delivery service. *Computers & Industrial Engineering* **2006**, *51*, 154−162.

(11) Marcoulaki, E. C.; Papazoglou, I. A.; Pixopoulou, N. Integrated framework for the design of pipeline systems using stochastic optimization and GIS tools. *Chem. Eng. Res. Des.* **2012**,

(12) Passuello, A.; Schuhmacher, M.; Mari, M.; Cadiach, O.; Nadal, M. A Spatial Multicriteria Decision Analysis to Manage Sewage Sludge Application on Agricultural Soils. *Environmental Modeling for Sustainable Regional Development: System Approaches and Advanced Methods* **2010**, 221.

(13) Vaskan, P.; Passuello, A.; Guillén-Gosálbez, G.; Schuhmacher, M.; Jiménez, L. Combined use of GIS and mixed-integer linear programming for identifying optimal agricultural areas for sewage sludge amendment: A case study of Catalonia. *Environmental Modelling & Software* **2013**,

(14) Hartigan, J. A.; Wong, M. A. Algorithm AS 136: A k-means clustering algorithm. *Journal of the Royal Statistical Society. Series C (Applied Statistics)* **1979**, *28*, 100−108.

(15) Kanungo, T.; Mount, D. M.; Netanyahu, N. S.; Piatko, C. D.; Silverman, R.; Wu, A. Y. An efficient k-means clustering algorithm: Analysis and implementation. *Pattern Analysis and Machine Intelligence, IEEE Transactions on* **2002**, *24*, 881−892.

(16) MatLab, http://www.mathworks.es/. 2013,