सिद्धिमूलं प्रबन्धनम्
भा. प्र. सं. इन्दौर
IIM INDORE

Capacitated Facility Location-Allocation Problem for Effluent Treatment in an Industrial Cluster

Saurabh Chandra

saurabh@iimidr.ac.in

Amit Kumar Vatsa

amitv@iimidr.ac.in

Manish Sarkhel

f14manishs@iimidr.ac.in

WP/02/2019-20/OM&QT

January 2020

Capacitated facility location-allocation problem for effluent treatment in an industrial cluster

Abstract

We present a location-allocation problem for effluent treatment in a cluster of processing units. The problem involves installing effluent treatment plants of appropriate capacities at suitable locations and allocating processing units to these plants. This problem is formulated as a mixed integer non-linear programming problem with nonconvex treatment costs, for which an exact convexification strategy is proposed. An outer approximation based branch and cut approach is presented as an exact solution method to solve practical size instances. For solving larger instances, a hybrid heuristic approach based on outer approximation and mixed-integer linear programming neighborhood based search is presented.

Keywords: shared effluent management, facility location-allocation, mixed integer non-linear program, outer approximation

1. Introduction

Many industrial units such as textile, leather, paper, food processing, etc. produce toxic effluents as a part of their manufacturing processes. Worldwide countries enacted laws to restrict the dangerous contamination of surrounding land and water bodies. India for instance, through the water prevention and control of pollution act 1974, has mandated all processing units to treat their effluents before disposing it off to the environment. Wastewater processing is a challenge for the medium and small scale enterprises (MSEs) in India due to the high cost of installation and maintenance of the effluent treatment units. To solve this problem, the Government of India promotes the installation of a shared network of common effluent treatment plants (CETP) for a group of processing units for joint wastewater management.

There are multiple advantages associated with CETPs from the perspective of MSEs. First, member firms can save on effluent treatment costs through the economies of scale in operating a larger effluent treatment facility. Second, MSEs may not have access to extra land within their premises to set up their effluent treatment facility. CETPs are located at a convenient location, considering the location characteristics like terrain and spatial distribution of processing units. This can lead to a reduction in allocation/transportation costs, which are typically a substantial percentage of the overall cost. Lastly, the difference in effluent characteristics emanating from these firms serves to homogenize the effluent transported to a CETP, which helps in setting up standard CETPs at common locations instead of customized ones for every processing unit.

The design of a network of CETPs for an industrial cluster involves many steps. The first step consists of determining the potential locations for the installation of the CETPs. Next, the exact locations for CETP installation need to be identified, followed by choosing the capacities of CETPs to be installed at each site. Furthermore, the processing units to be allocated to each installed CETP needs to be determined, such that the total discharge from the allocated processing units does not exceed the installed capacity of the corresponding CETP, at each location. At each CETP location, the installation and treatment costs show economies of scale with increasing volumes of effluent. On the contrary, these costs show dis-economies of scale in terms of the level of pollutant concentration. The marginal cost of treatment of effluent goes up with the pollutant concentration, making it increasingly costly to treat the effluent.

We present a new variant of the facility location-allocation model to address the problem of designing a network of CETPs for an industrial cluster. We present a mixed integer nonlinear programming (MINLP) formulation for the problem. We observe that this MINLP is non-convex, since the treatment costs exhibit economies of scale w.r.t. the volume of effluent and dis-economies w.r.t pollutant concentration. To find an exact solution approach for the model, we first propose a convexification strategy. Next, an outer approximation based branch and cut algorithm is presented as an exact solution approach. To solve larger instances in reasonable times, a hybrid heuristic approach based on outer approximation and mixed integer linear programming (MILP) based neighborhood search is proposed. Computational experiments on multiple data instances derived from a real scenario, followed by analyzing the impact of various problem parameters on different costs and model output are presented.

In Section 2, we present a brief literature review of the location-allocation problem relevent to our context. Section 3 describes the problem formulation with a case study. Section 4 describes the solution strategies for the problem. Computational analysis of the proposed solution methods is presented in Section 5. A multi-objective analysis of different financial and social costs in the problem are discussed in Section 6. Finally, we conclude with remarks and future directions.

2. Literature Review

The literature related to this work comes from two strands of ongoing work in logistics research and optimization: facility location-allocation problems and waste management. We discuss some relevant papers in this regard below:

2.1. Facility location allocation problem

In an early presentation of location-allocation problems, Cooper (1963) discusses the numerical aspects of various classes of these problems. Both exact and heuristic methods are analyzed. Various extensions and versions of these problems are discussed in the literature since then. Zhou and Liu (2003) present stochastic programming models for capacitated facility location-allocation problems. Network simplex, stochastic simulations, and genetic algorithms are integrated into the proposed solution approach. Capacitated facility location-allocation problems are considered as belonging to the class of NP-hard problems. We refer to some of the recent literature on these types of problems.

Specific to applications in the field of supply chain management, Manzini and Gebennini (2008) present an MILP formulations aimed at designing multi-period, multi-stage, and multi-commodity location-allocation problems in logistics distribution systems. Chen and Ting (2008) present a single-source capacitated facility location problem. A hybrid algorithm combining Lagrangian heuristic and Ant Colony algorithm is proposed as an efficient solution approach. A single source facility location is a particular class of facility location problems in which each customer is served from only one facility. Harris et al. (2014) present a capacitated facility location-allocation problem with flexibility at the allocation level. Paper presents multi-Objective optimization with CO_2 emissions and financial costs. An evolutionary algorithm coupled with Lagrangian relaxation is presented as an efficient solution approach. Mogale et al. (2018) present a food grains silo location-allocation model in the context of the Indian grain distribution system. The problem is formulated as a multi-objective, multi-modal, and multi-period planning problem with dwell time. Total supply chain network cost and lead-time are taken as conflicting objectives. Further, Pareto based multi-objective algorithms are implemented as solution approaches. Baharmand et al. (2019) present a location-allocation model for locating distribution centers in a disaster zone. Trade-offs between response times and logistics costs are considered. Wu and Yang (2018) analyze the location decisions of Chinese manufacturing firms by integrating a flow capturing location model with the traffic assignment model. A genetic algorithm is used to solve the model.

In other applications, Hammad et al. (2017) address a multi-objective facility location problem, concerning the location of noise-sensitive and noise-generating facilities in urban environments. An augmented ϵ-constraint method is used to handle multiple objectives, and a Benders decomposition approach is proposed to solve large instances. Liu et al. (2019) present a bi-objective optimization model to determine the optimal temporary medical service locations and medical service allocation plan by maximizing the number of expected survivals and minimizing the total operational cost. Lin et al. (2019) propose a location-allocation model for a multi-classification yard location problem. An upper-level decision involves the selection of potential yard locations with yard size and capacity, followed by the lower level problem of determining railcar re-classification. A simulated annealing algorithm is presented as a solution approach. In an application in the related area to the current study, Gokbayrak and Kocaman (2017) propose a continuous location-allocation problem as a mixed integer quadratically constrained problem. Each facility has a fixed opening cost and has to operate under coverage distance limitations. Applications are suggested in the spatial planning of water and energy access networks. A three-stage heuristic algorithm is proposed for the problem.

2.2. Waste management

de Figueiredo and Mayerle (2008) discuss a problem of designing minimum-cost recycling network with required throughput to determine the optimal number and location of receiving centers along with incentive mechanisms to stimulate the collection of used or unrecoverable products. In a logistics network design problem in waste management, Parker et al. (2010) present an MINLP model for analyzing the economic potential and infrastructure configuration designing a pathway for hydrogen production from agricultural waste and delivery to usage locations. Kim et al. (2011) present an optimization model for the design of biomass processing network for biofuel. Zhao et al. (2016) discuss the network design problem for a regional hazardous waste management system. The model deals with the location of various waste facilities and determining transportation routes to the facilities. The problem is formulated as a multi-objective MILP model, and three multi-objective optimization approaches are implemented to find good solutions.

A class of studies deals with general waste collection mechanisms. Ramos et al. (2014) present a planning problem of recyclable waste collection systems using a multi-product, multi-depot, vehicle routing problem aimed at minimizing the objectives of distance and CO_2 emissions. A decomposition solution approach is developed as a solution approach for a real case study. Miranda et al. (2015) address the problem of designing a household waste collection system for insular areas. Mixed integer programming model integrates site selection, scheduling, and routing.

Investigation of the research related to MINLPs involving power functions, Wang et al. (2013) propose quadratic outer approximation approaches for solving fuel consumption rate functions in a berth allocation problem incorporating ship fuel consumption minimization. Wang and Meng (2012) present an efficient outer-approximation method to solve MINLP designed for achieving sailing speed optimization for container ships. Wang et al. (2015) propose a global optimization method employing linearization, outer approximation, and range reduction techniques to solve a discrete transportation network design model.

While the literature on location-allocation problems is extensive, the incorporation of nonconvex operational costs within the same is sparse. To the best of our knowledge, strategic decisions involving the installation of multiple effluent treatment facilities taking into account the nature and quantity of effluents is missing from the literature. Our model attempts to fill this void. By integrating the aspects as mentioned above, we provide a decision-making framework for the installation and configuration of common effluent treatment facilities. The data used, along with our model, captures the intricacies relevant in the Indian setting. However, the model can be tweaked and extended in other contexts.

3. Problem description and mathematical formulation

This section presents a detailed description of the CETP location-allocation problem using a case study. Tirupur is a city located in the southern state of Tamil Nadu in India. It has a major cotton textile manufacturing cluster. A large number of integrated and small-scale textile manufacturing units are located in and around the city. The small scale textile manufacturing units are primarily involved in one or multiple stages of the textile manufacturing process. One of the essential stages of textile manufacturing is dyeing and bleaching. It is done to impart color to the fabric. The dyeing/bleaching process consumes large amounts of water and produces effluents. The effluent discharged from these firms contains a lot of dissolved salts used as binding agents during the dyeing process. These salts are toxic to living organisms and need to be neutralized before disposal. So far, eighteen CETPs have been established in Tirupur to treat effluents coming out of the dyeing and bleaching units in the area. These cater to the needs of around 380 dyeing and bleaching units.

Most of the dyeing and bleaching units using the CETPs are small and medium scale enterprises. They lack the resources to set up their own individual effluent treatment plants. Hence, they cooperate with other such small-scale units to set up CETPs. The number of processing units allocated to each CETP is different. It is primarily dependent upon the pollutant concentration and the effluent volume discharged by the allocated processing units, and the distance of these units from the potential CETP locations. The potential CETP locations are determined by the geography of the area. The effluent transportation from the processing units to the assigned CETPs are carried out either through pipelines or using water tankers designed to carry the effluents.

The CETP facility location-allocation problem deals with identifying CETPs of appropriate capacities to be installed at most suitable locations and allocating each processing unit to exactly one installed CETP. The objective is to minimize the cost of CETP installation, transportation, and treatment cost of the effluent. We consider discrete capacity options for CETP installation at each location. The installation cost of a CETP depends on its capacity. The effluent treatment cost is a function of effluent volume and pollutant concentration in the effluent.

We now present the all notations and symbols used in the model are described in the following subsection, followed by the mathematical model of the problem.

3.1. Notations

Indices

i	potential CETP location
j	processing unit
t	CETP capacity type

Sets

$\mathcal{I}$	set of all potential CETP locations
$\mathcal{T}$	set of all CETP capacity types

$\mathcal{F}$ set of all processing units

Decision variables

y_{ti} 1, if CETP of capacity type t is installed at location i, 0 otherwise

z_{ji} 1, if processing unit j is allocated to CETP at location i, 0 otherwise

x_i volume of effluent transported to a CETP at location i

p_i weight of pollutant transported to a CETP at location i

Parameters

C_{ti}^F fixed cost of installing a CETP of capacity t at a potential CETP location i

C_{ji}^A transportation cost when processing unit j is allocated to a CETP at location i.

CAP_t treatment capacity of CETP capacity type t

F_j^e effluent discharge from a processing unit j

F_j^p pollutant concentration of effluent from a processing unit j

K constant representing the multiplier of the effluent treatment cost function

q constant representing the exponent of volume of effluent treated in a CETP

r constant representing the exponent of pollutant weight treated in a CETP

3.2. MINLP model

The problem involves which potential CETP locations within the set I to be selected for installation and the capacity type t of CETP to be installed at each selected location i. Each installed location caters to a subset of processing units in the cluster. The overall system costs consist of the installation cost of the CETPs, the allocation cost between processing units and CETPs, and cost of treatment of effluent through the CETPs. The allocation cost comprises of the effluent transportation cost between processing units and respective CETPs. Figure 1 illustrates the problem structure using a small example.

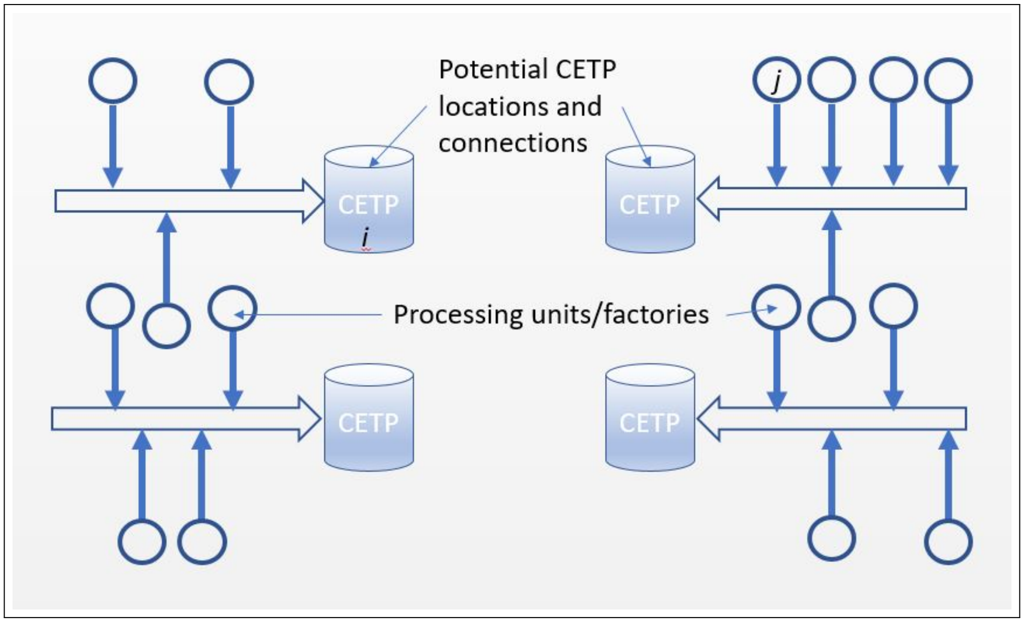

Figure 1: An illustrative example of the CETP location-allocation problem

Next, we present the mathematical formulation of the problem.

$$\text{P: Min } Z = \sum_{i\in\mathcal{I}}\sum_{t\in\mathcal{T}} C^F_{ti} y_{ti} + \sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{F}} C^A_{ji} z_{ji} + \sum_{i\in\mathcal{I}} K x_i^q p_i^r \tag{1}$$

$$s.t. \quad \sum_{i\in\mathcal{I}} z_{ji} = 1 \qquad \forall j \in \mathcal{F} \tag{2}$$

$$\sum_{t\in\mathcal{T}} y_{ti} \leq 1 \qquad \forall i \in \mathcal{I} \tag{3}$$

$$z_{ji} \leq \sum_{t\in\mathcal{T}} y_{ti} \qquad \forall i \in \mathcal{I}, j \in \mathcal{F} \tag{4}$$

$$x_i = \sum_{j\in\mathcal{F}} F^e_j z_{ji} \qquad \forall i \in \mathcal{I} \tag{5}$$

$$\sum_{j\in\mathcal{F}} F^e_j z_{ji} \leq \sum_{t\in\mathcal{T}} CAP_t y_{ti} \qquad \forall i \in \mathcal{I} \tag{6}$$

$$p_i = \sum_{j\in\mathcal{F}} F^e_j F^p_j z_{ji} \qquad \forall i \in \mathcal{I} \tag{7}$$

$$x_i, p_i \geq 0 \qquad \forall i \in \mathcal{I} \tag{8}$$

$$z_{ji} \in \{0, 1\} \qquad \forall i \in \mathcal{I}, j \in \mathcal{F} \tag{9}$$

$$y_{ti} \in \{0, 1\} \qquad \forall i \in \mathcal{I}, t \in \mathcal{T} \tag{10}$$

The objective function (1) minimizes the total cost of installation of CETPs, the allocation cost of processing units to the installed CETPs, and the treatment cost of effluent through the CETPs. The last nonlinear objective function term is derived from previous studies on CETP design (Mundle et al., 1995) and is elaborated in subsection 3.3. Constraints (2) ensure that each processing unit is allocated to exactly one CETP. Constraints (3) ensure that at most, one CETP is installed at each potential location. Constraints (4) do not allow a processing unit to be allocated to a potential CETP location unless one is installed there. Constraints (5) ensure that the total effluent transported to a CETP must be equal to the discharge from the processing units allocated to it. Constraints (6) are the capacity constraints for effluent treatment at each CETP. Finally, constraints (7) calculate the total pollutant weight from all allocated processing units transported to a CETP. Constraints (8) impose non-negativity condition of x_i and p_i variables. Finally, the constraints (9) and (10) restrict the z_{ji} and y_{ti} variables to take only binary (0,1) values.

3.3. Functional characteristics of the treatment cost

In a technical note on CETP design, Mundle et al. (1995) propose that the treatment cost in a CETP is a nonlinear function of the total volume and the concentration of the effluent. The functional form is presented in the last term of the objective function (1). Here C_i^T is the effluent treatment cost at CETP i, c_i is the concentration of the effluent, and s is the exponent of effluent volume treated at i, x_i. The treatment cost shows the economy of scale in terms of x_i and dis-economy of scale in terms of c_i. We derive the expression for treatment cost in terms of effluent volume and pollutant weight from the existing relationship between effluent volume and concentration, as concentration itself is a function of effluent volume.

$$C_i^T = Kx_i^s c_i^r = Kx_i^s \left(\frac{p_i}{x_i}\right)^r = Kx_i^{(s-r)} p_i^r = Kx_i^q p_i^r \tag{11}$$

Proposition 1. *Treatment cost function in equation* (11) *is non-convex when there exists economies of scale w.r.t. effluent volume and dis-economies of scale w.r.t. effluent concentration.*

Proof. The treatment cost function of equation (11) will have $s < 1$ when there exists economies of scale w.r.t. effluent volume, and $r > 1$ with dis-economies of scale in terms of effluent concentration (hence $s - r = q < 0$). The Hessian matrix for the above non-linear function is as follows:

$$\mathbf{H} = K.\begin{bmatrix} q(q-1)x_i^{q-2}p_i^r & qrx_i^{q-1}p_i^{r-1} \\ qrx_i^{q-1}p_i^{r-1} & r(r-1)x_i^q p_i^{r-2} \end{bmatrix} = Kx_i^{q-2}p_i^{r-2}\begin{bmatrix} q(q-1)p_i^2 & qrx_ip_i \\ qrx_ip_i & r(r-1)x_i^2 \end{bmatrix} \tag{12}$$

$$\det(\mathbf{H}) = Kx_i^q p_i^r q.r(1-q-r) \tag{13}$$

The function C_i^T is convex if and only if the following conditions are satisfied:

1. det(**H**) is non-negative: As $q < 0$, this condition is satisfied when $q + r \geq 1$ (i.e. $s \geq 1$). Clearly, this condition is not satisfied.
2. Diagonal elements of matrix **H** are non-negative:

(a) Since $q < 0$, the first diagonal element is positive.

(b) Since $r > 1$, the second diagonal element is positive.

Since the determinant of Hessian matrix is negative, the treatment cost function as described in equation (11) is non-convex. □

The values of parameters of the nonlinear treatment cost are estimated from the values derived from real data of the case study. Demonstrating economies of scale in effluent volume, s varies from 0.85 to 0.95. The values of the exponent of pollutant concentration, parameter r, varies from 1.2 to 1.4.

4. Solution strategy

The non-linear terms in the objective function (1) is a non-convex bi-variate function in effluent volume and pollutant weight. First, we implement a convexification strategy to this term. We notice that the variable x_i is specified by constraint set (5), which is an expression involving only the binary variables. Hence for convexification, we introduce a variable w_i where $w_i = x_i^2$. In terms of w_i, the treatment cost at CETP location i can now be expressed as:

$$C_i^T = Kw_i^{q/2} p_i^r \quad where, w_i = x_i^2 \tag{14}$$

Further, we replace constraint set (5) with

$$w_i = (\sum_{j \in \mathcal{F}} F_j^e z_{ji})^2 \quad \forall i \in \mathcal{I} \tag{15}$$

Since, z_{ji} are binary variables, we linearize the above constraint set using the following set of constraints:

$$w_i = \sum_{j=1}^{|\mathcal{F}|-1} \sum_{j'=j+1}^{|\mathcal{F}|} 2F_j^e F_{j'}^e \zeta_{ijj'} + \sum_{j \in \mathcal{F}} (F_j^e)^2 z_{ji} \qquad \forall i \in \mathcal{I} \tag{16}$$

$$\zeta_{ijj'} \leq z_{ji} \qquad \forall i \in \mathcal{I}, j \in \{1, 2, .., |\mathcal{F}| - 1\}, j' \in \{j + 1, .., |\mathcal{F}|\} \tag{17}$$

$$\zeta_{ijj'} \leq z_{j'i} \qquad \forall i \in \mathcal{I}, j \in \{1, 2, .., |\mathcal{F}| - 1\}, j' \in \{j + 1, .., |\mathcal{F}|\} \tag{18}$$

$$\zeta_{ijj'} \geq z_{ji} + z_{j'i} - 1 \qquad \forall i \in \mathcal{I}, j \in \{1, 2, .., |\mathcal{F}| - 1\}, j' \in \{j + 1, .., |\mathcal{F}|\} \tag{19}$$

$$\zeta_{ijj'} \in \{0, 1\} \qquad \forall i \in \mathcal{I}, j \in \{1, 2, .., |\mathcal{F}| - 1\}, j' \in \{j + 1, .., |\mathcal{F}|\} \tag{20}$$

Proposition 2. *$Kw_i^{q/2} p_i^r$ is a convex function for $r \in [1.2, 1.4]$, $s \in [0.85, 0.95]$ and $q = s - r$.*

Proof. Replacing q with $q/2$ in equation (13), it can be seen that the determinant of Hessian matrix will be non-negative when $1 - q/2 - r \leq 0$, i.e. $q/2 + r \geq 1$. This implies $(s + r)/2 \geq 1$ or, $s + r \geq 2$. Clearly, $r \in [1.2, 1.4]$ and $s \in [0.85, 0.95]$ satisfy this. □

Proposition 3. *Constraint set* (19) *will be redundant at optimality for* $r \in [1.2, 1.4]$ *and* $s \in [0.85, 0.95]$.

Proof. As shown above, the treatment cost given in equation (14) is a convex function for $r \in [1.2, 1.4]$ and $s \in [0.85, 0.95]$. Further, the treatment cost decreases with increasing w_i (since, $q = s - r < 0$). Moreover, for any given value of variables y_{ti} and z_{ji}, it can be seen from (16) that w_i increases with increasing $\zeta_{ijj'}$. Hence, only the upper bound $\zeta_{ijj'}$ will be be binding at optimality (i.e the lower bound of $\zeta_{ijj'}$ will be redundant). □

Proposition 4. $\zeta_{ijj'}$ *can be relaxed as continuous and the constraint set* (20) *is redundant.*

Proof. For a given solution y_{ti}, z_{ji}, it is clear that $\zeta_{ijj'}$ is upper bound by constraint sets (17) and (18). Further, as argued in proof of Proposition 3, only the upper bound of $\zeta_{ijj'}$ will be binding at optimality. Thus, $\zeta_{ijj'}$ variables value will be determined by the lower of the upper bounds given by constraint sets (17) and (18). Hence, $\zeta_{ijj'} := \min(z_{ji}, z_{j'i}) \leq 1$. Therefore, even if $\zeta_{ijj'}$ is relaxed as continuous and the constraint set (20) is removed, the solution will not change. □

We have transformed a non-convex MINLP to a convex MINLP. Convex problems are much easier to deal with as compared to non-convex problems. Quoting Rockafellar (1993), "In fact the great watershed in optimization isn't between linearity and nonlinearity, but convexity and nonconvexity". Additionally, some exact methods that cannot be applied to a non-convex problem, can now be applied to our transformed problem. We next present one such exact method that is widely applied to a convex MINLP.

4.1. Outer approximation cutting plane method

Duran and Grossmann (1986) present an outer approximation (OA) cutting plane algorithm for solving MINLPs, where all the functions involving discrete variables are linear, and the nonlinear functions in the continuous variables are convex in the underlying structure. The OA solution method works by decomposing the original MINLP into a subproblem, which is a nonlinear program in continuous variables and a master problem, which is a MILP in the discrete variable and a single continuous variable. The step by step algorithm for the OA method implemented as a branch and cut to solve MINLPs is shown by Duran and Grossmann (1986) and Fletcher and Leyffer (1994). Shahabi et al. (2013) demonstrate the step-by-step application of the OA algorithm to solve a robust shortest path problem.

The MINLP formulation for the CETP facility location-allocation problem (1) - (10), is separable in linear and nonlinear terms. The terms involving the binary variables (y_{ti} and z_{ji}) are linear in the formulation, and the objective function term estimating the treatment costs are a nonlinear function of the continuous variables (x_i and p_i). The original non-convex nonlinear term is convexified using the strategy reported earlier. We show the formulations for the resulting master problem (MP), and the sub-problem (SP) as per the OA method implemented to the given problem.

$$MP: \quad minimize \quad Z_{MP} = \mu + \sum_{i\in\mathcal{I}}\sum_{t\in\mathcal{T}} C^F_{ti} y_{ti} + \sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{F}} C^A_{ji} z_{ji} \tag{21}$$

Subject to

$$w_i - (\sum_{j\in\mathcal{F}} F^e_j z_{ji})^2 = 0, \quad \forall i \in \mathcal{I}, \tag{22}$$

$$\mu \geq \sum_{i\in\mathcal{I}} K(w^k_i)^{q/2}(p^k_i)^r + \sum_{i\in\mathcal{I}}(w_i - w^k_i)K\frac{q}{2}(w^k_i)^{\frac{q}{2}-1}(p^k_i)^r + \sum_{i\in\mathcal{I}}(p_i - p^k_i)K(w^k_i)^{\frac{q}{2}} r(p^k_i)^{r-1}, \tag{23}$$

Constraints (2)-(4), (6), (7), (9), (10), (16)-(18)

$$w_i, p_i \geq 0, \quad \forall i \in \mathcal{I}, \tag{24}$$

The constraint (23) is the outer approximation cutting plane, as applied to the model. Here w^k_i and p^k_i are a known pair of feasible values of the decision variables, w_i and p_i, respectively. At first, MP is solved to optimality without adding the OA cutting plane. The solution to MP gives a lower bound on the optimal solution to original problem P and gives a feasible set of values for the binary decision variables y_{ti} and z_{ji}. The known values of the binary decision variables are used to estimate the effluent volumes, pollutant weights, and treatment costs. The total cost calculated by adding the MP objective with the treatment costs gives an upper bound on the objective function. We use the derived values of continuous variables in the previous step to formulate a new cutting plane, which is then added to the MP. This process is continued iteratively, till the gap between the lower and upper bounds reaches a target level.

The above description of OA method is a classical version, in which the MP is solved to optimality at each iteration. As the number of successive cuts increases, it becomes increasingly difficult to solve the MP. We implement the OA method in a branch and cut framework. In this method, the incumbent solution in the branch-and-bound search tree is passed as known values of binary decision variables to generate a new OA cut. This is facilitated by implementing the branch-and-cut routine using the callback function provided by commercial solvers (like Gurobi), to intervene in the branch-and-bound tree search process. The callback function adds the OA cut directly at each node in the branch-and-bound tree while solving the MP. Moreover, the generated OA cuts are added to the MP as *lazy constraints*. When a cut is operationalized as a *lazy* cut, the solver does not add all the generated cuts at each node, but performs a feasibility check on the lazy cuts and adds only those which are infeasible. Our computational results suggest that this implementation is faster than the classic OA implementation. Figure 2 presents a flowchart of this implementation of the OA algorithm. The stopping criteria, as mentioned in the decision step, is the mixed integer programming gap calculated as the ratio of the difference between the upper bound on the MILP solution and lower bound from a relaxed LP solution to the upper bound.

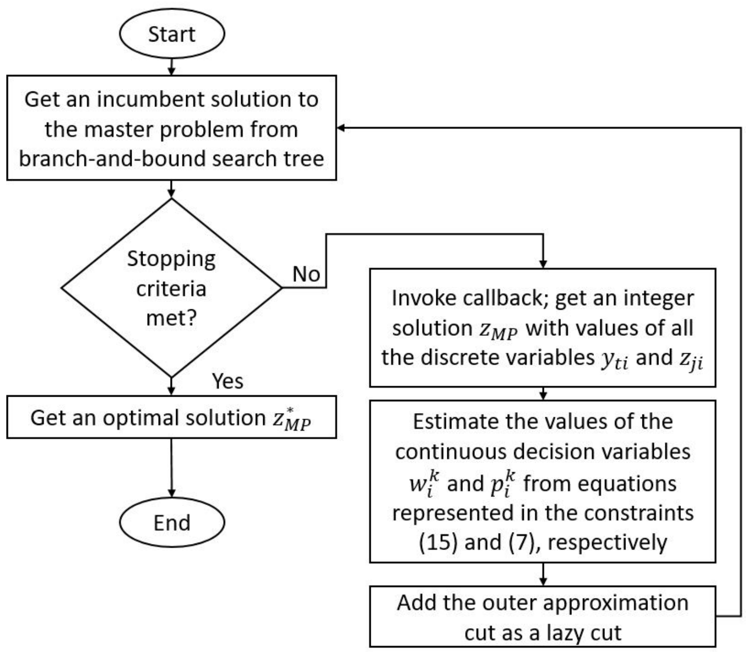

Figure 2: Flowchart of outer approximation based branch-and-cut implementation

4.2. Hybrid outer approximation and MILP based neighborhood search heuristic

The OA based branch-and-cut algorithm could solve only medium-sized instances. So, to solve large-sized instances, we present a hybrid heuristic that incorporates an MILP based neighborhood search incorporating the OA based branch-and-cut method.

The heuristic presented here works as follows. We obtain multiple starting, feasible solutions to the problem instance. We set up the instance as an MILP model in a commercial solver using the OA- branch and cut solution approach. The binary decision variables are initially fixed to the values of a starting feasible solution, and the decision variable values and the initial objective function value are recorded as the best solution. A local neighborhood search between each pair of CETP potential locations, at least one of which is currently in the solution, is carried out using different operators. On performing any operation, if the current solution is better than the previous known best solution, the later is updated to the new solution. The pairwise local search is carried out with all identified starting feasible solutions, and the best solution is recorded for each start. The best solution found across all the different starts is reported as the final heuristic solution. The heuristic consists of multiple procedures and sub-procedures, which are explained below.

Algorithm 1 implements the OA- based branch and cut algorithm to solve a problem instance and gives as output the optimal objective and the optimal values of the decision variables.

Algorithm 1 MINLP solution implementation

procedure FUNCTION –MINLPSOLVE()
    solve the MINLP using the OA branch and cut method (ref: figure 2)
    $Obj \leftarrow$ optimal objective function value
    $Y' \leftarrow$ set of optimal solution values of the decision variables y_{ti} for $t \in \mathcal{T}$, $i \in \mathcal{I}$
    $Z' \leftarrow$ set of optimal solution values of the decision variables z_{ji} for $j \in \mathcal{F}$, $i \in \mathcal{I}$
    return Obj, Y', Z'
end procedure

While conducting the pairwise search, in some steps, we do not solve the sub-problem using a commercial solver but use simple logic to evaluate a potential improvement in solution. In such cases, we derive a known set of values for the decision variables and estimate its objective using the *ObjEval* function as described in algorithm 2.

Algorithm 2 Function evaluating the objective for know decision variable values

procedure FUNCTION –OBJEVAL($\bar{Y}, \bar{Z}$)
    $\bar{Y} \leftarrow \{\bar{y}_{ti} \mid t \in \mathcal{T}, i \in \mathcal{I}\}$
    $\bar{Z} \leftarrow \{\bar{z}_{ji} \mid j \in \mathcal{J}, i \in \mathcal{I}\}$
    for $i \in \mathcal{I}$ **do**
        $\bar{x}_i \leftarrow \sum_{j \in \mathcal{F}} F^e_j \bar{z}_{ji}$
        $\bar{p}_i \leftarrow \sum_{j \in \mathcal{F}} F^e_j F^p_j \bar{z}_{ji}$
    end for
    $Obj \leftarrow \sum_{i \in \mathcal{I}} \sum_{t \in \mathcal{T}} C^F_{ti} \bar{y}_{ti} + \sum_{i \in \mathcal{I}} \sum_{j \in \mathcal{F}} C^{FC}_{ji} \bar{z}_{ji} + \sum_{i \in \mathcal{I}} K \bar{x}^q_i \bar{p}^r_i$
    return Obj
end procedure

Two avoid the MIP based neighborhood search heuristic getting stuck in local optima, we employ a multi-start method, where diverse starting solutions are used. Two extreme starting feasible solution sets are obtained through procedures explained as algorithms 3 and 4. The StartSol1 method generates a starting feasible solution with only one large CETP installed at a heuristically determined best location. Initially, all the decision variables are initialized to zero. For each potential CETP location, we estimate the weighted average distance, w^{dist}_i, as a weighted average of distances of a CETP from all the processing units, weights being the pollutant discharge through each processing unit. The location having the minimum weighted distance is identified as i^b. The largest CETP is installed at this location and all the processing units are allocated to the CETP installed at this location. The procedure yields a Y_1 and Z_1 as solution vectors for one of the starting feasible solutions.

Another starting feasible solution represented as StartSol2 represents another extreme with the smallest possible CETPs installed at as many locations as possible. Initially, all the decision variables are initialized to zero. A de-

Algorithm 3 Starting feasible solution with one large facility

procedure StartSol1
  $Y_1 \leftarrow \{\bar{y}_{ti} = 0 \,|\, t \in \mathcal{T}, i \in \mathcal{I}\}$
  $Z_1 \leftarrow \{\bar{z}_{ji} = 0 \,|\, j \in \mathcal{J}, i \in \mathcal{I}\}$
  $i^b \leftarrow \phi$
  $w_i^{dist} \leftarrow \frac{\sum_{j \in \mathcal{F}} D_{ji} F_j^e F_j^p}{\sum_{j \in \mathcal{F}} F_j^e F_j^p}$
  $w_{min} \leftarrow \min_{i \in \mathcal{I}} w_i^{dist}$
  $i^b \leftarrow \{i \,|\, w_i^{dist} = w_{min}\}$
  $t_L \leftarrow$ capacity of the largest CETP type
  $\bar{y}_{t_L i^b} \leftarrow 1$
  for $j \in \mathcal{F}$ **do**
    $\bar{z}_{ji^b} \leftarrow 1$
  end for

scending list of locations as per the weighted distance is generated as I_{ranked}. CAP_{min} is a lower limit on the size of the smallest CETP and is estimated as the ratio of the total flow through all the processing units and the number of potential locations. In steps 6 - 11, the smallest CETP type having a capacity just greater than or equal to CAP_{min} is identified. In the next step, a pollutant level weighted distance of each processing unit from a potential location is evaluated as w_{ji}. We start installing a CETP to each potential location as per the sequence in the I_{ranked} set and allocated as many processing units as possible, given the capacity constraint of the CETP, as per the increasing value of w_{ji} associated with the processing units. Once all the processing units are allocated, the algorithm terminates. Finally, we report Y_2 and Z_2 as solution vectors for the second starting feasible solution.

Algorithm 7 details the steps followed in the MILP based neighborhood search heuristic. A complete pairwise neighborhood search between each pair of potential CETP locations is conducted for each starting feasible solution represented by the index k. The best solution is initialized as the decision variable solution vectors, Y^{best} and Z^{best} taking null values and best objective value Obj^{best} taking a very large value represented as M. Next, a MILP instance is created by declaring the variable vectors, Y, and Z, and freezing the values of these variables to the known values from the starting feasible solution sets, Y_k and Z_k. For each pair of different potential CETP locations, i_1, and i_2, a local search is carried out through a process of multiple operators. First, the CETP types, t_1 and t_2 currently installed at the two locations is identified. In step 13, we check if only one of the locations has a CETP installed, we perform two operations. First, we swap the locations by installing a CETP at the empty location and check for improvements in objective value. The $Swap$ operator, as explained in algorithm 5 interchanges the values of known decision variable values corresponding to the locations i_1 and i_2 in the vectors Y_k and Z_k, represented as $\bar{y}_{ti_1}$, $\bar{y}_{ti_2}$, $\bar{z}_{ji_1}$, and $\bar{z}_{ji_2}$, across the two locations. The updated vectors Y_k and Z_k are evaluated by the operator $CheckUpdate$ described as algorithm 6. This function evaluates the objective value for the known solution vectors and compares it with the best-known solution found so far. If there is an improvement in the solution, the best solution vectors Y^{best} and Z^{best}, and objective value, Obj^{best} are updated with the newfound improved solution. Steps 17 - 22 implement the reshuffle operator, where

Algorithm 4 Starting feasible solution with multiple small CETPs

procedure StartSol2
  $Y_2 \leftarrow \{\bar{y}_{ti} = 0 \mid t \in \mathcal{T}, i \in \mathcal{I}\}$
  $Z_2 \leftarrow \{\bar{z}_{ji} = 0 \mid j \in \mathcal{J}, i \in \mathcal{I}\}$
  $I_{ranked} \leftarrow$ sorted list of location i as per decreasing values of w_i^{dist}
  $CAP_{min} \leftarrow \frac{\sum_{j \in \mathcal{F}} F_j^e}{n(\mathcal{I})}$
  for $t \in \mathcal{T}$ **do**
    if $CAP_t \geq CAP_{min}$ **then**
      $t_s \leftarrow t$
      $CAP_{min} \leftarrow CAP_t$
      quit for loop
    end if
  end for
  for $j \in \mathcal{F}$ **do**
    for $i \in \mathcal{I}$ **do**
      $w_{ji} \leftarrow F_j^e F_j^p D_{ji}$
    end for
  end for
  for $i \in \mathcal{I}_{ranked}$ **do**
    $\bar{y}_{t_S i} \leftarrow 1$
    $F_i^w \leftarrow$ sorted list of the processing units j as per the weights w_{ji}
    for $j \in F_i^w$ **do**
      $\bar{z}_{ji} \leftarrow 1$ till the capacity CAP_{t_S} is not violated for the given i
    end for
  end for

Algorithm 5 Swapping the fixed values of binary decision variables between a pair of locations

procedure Function Swap(Y_k, Z_k, i_1, i_2)
  swap the values of $\bar{y}_{ti_1}$ and $\bar{y}_{ti_2} \forall t \in \mathcal{T}$
  swap the values of $\bar{z}_{ji_1}$ and $\bar{z}_{ji_2} \forall j \in \mathcal{F}$
  update the vectors Y_k and Z_k
end procedure=0

Algorithm 6 Checking and updating the best solution

procedure Function CheckUpdate($Obj^{best}, Y^{best}, Z^{best}, Y_k, Z_k, Obj^1$)
  if $Obj^1 \leq Obj^{best}$ **then**
    $Obj^{best} \leftarrow Obj^1$
    $Y^{best} \leftarrow Y_k$
    $Z^{best} \leftarrow Z_k$
  end if
end procedure

a sub-MILP is solved using the *MINLPsolve* function by unfixing the decision variables corresponding to the current location pair. The *CheckUpdate* function checks and updates any improvement in solution.

In the second case, where both the locations in the current pairing are currently installed, as identified in step 24, three operators are employed. First, the *Swap* operator is implemented as previously explained. Second, steps 28 - 38 implements the integration operator, which tries to combine a single CETP to handle all the processing units assigned to both the locations at one of the locations. Next, steps 39 - 47 describe a reshuffling operator where a sub-MILP is solved on unfixed decision variables across the location pair. Finally, the best improvement identified so far is reported.

5. Computational analysis

The two solution approaches discussed in the Section 4 have been implemented in the academic version of the commercial solver Gurobi 8.1.1. with Python 2.7.10 programming language. All computational tests are performed using a DELL Precision T5610 with Intel Xeon CPU E5-2620 v2 @ 2.10 GHz – 6 cores CPUs and 32.0 GB RAM.

5.1. Test instances

The computational analysis is done using real data estimated from the Tirupur industrial cluster. A group of 380 processing units is engaged in similar production activities and release the same type of effluent as discharge. There are eighteen potential CETP locations, based on the terrain of the region and relative distances between different processing units. We have considered 3, 6, and 9 capacity types of CETP for analysis. To compare the solution results from the exact and heuristic solution approaches presented in section 4, we create smaller data sets from the original data described here. Table 1 presents the details of all the data instances used for experimentation. The columns 2 - 4 list the number of entities in terms of the number of processing units, the number of potential CETP locations, and the number of CETP capacity types considered, respectively.

The data related to the distances of individual processing units from each potential CETP location, effluent discharge rates in kl/day, and level of pollutant discharge in g/kl are presented as a minimum, maximum, and average values in table 2. Other costs and parameters considered for the model are shown in table 3.

Algorithm 7 MILP based neighborhood search heuristic

procedure MILPHEUR
  for $k \in \{1, 2\}$ **do**
    $Y^{best} \leftarrow \phi$
    $Z^{best} \leftarrow \phi$
    $Obj^{best} \leftarrow M$
    $Y \leftarrow$ vector of decision variables y_{ti} for $t \in \mathcal{T}, i \in \mathcal{I}$
    $Z \leftarrow$ vector of decision variables z_{ji} for $j \in \mathcal{F}, i \in \mathcal{I}$
    freeze variables in Y to the derived values of Y_k
    freeze variables in Z to the derived values of Z_k
    for $i_1, i_2 \in \mathcal{I} | i_1 \neq i_2$ **do**
      $t_1 \leftarrow$ CETP type currently installed at i_1
      $t_2 \leftarrow$ CETP type currently installed at i_2
      if $\sum_{t\in\mathcal{T}} \bar{y}_{ti_1} + \sum_{t\in\mathcal{T}} \bar{y}_{ti_2} = 1$ **then**
        Swap(Y_k, Z_k, i_1, i_2)
        $Obj^1 \leftarrow ObjEval(Y_k, Z_k)$
        CheckUpdate($Obj^{best}, Y^{best}, Z^{best}, Y_k, Z_k, Obj^1$)
        $t_m \leftarrow max(t_1, t_2)$
        for $t \in \{1, ..., t_m\}$ **do**
          unfreeze the decision variables y_{ti_1}, y_{ti_2}, z_{ji_1}, and z_{ji_2}
        end for
        $Obj^1 \leftarrow MINLPsolve().Obj$
        CheckUpdate($Obj^{best}, Y^{best}, Z^{best}, Y_k, Z_k, Obj^1$)
      end if
      if $\sum_{t\in\mathcal{T}} \bar{y}_{ti_1} + \sum_{t\in\mathcal{T}} \bar{y}_{ti_2} = 2$ **then**
        Swap(Y_k, Z_k, i_1, i_2)
        $Obj^1 \leftarrow ObjEval(Y_k, Z_k)$
        CheckUpdate($Obj^{best}, Y^{best}, Z^{best}, Y_k, Z_k, Obj^1$)
        $flow_{12} \leftarrow \sum_{j\in\mathcal{F}} F^e_j \bar{z}_{ji_1} + \sum_{j\in\mathcal{F}} F^e_j \bar{z}_{ji_2}$
        $t_C \leftarrow$ smallest CETP type with capacity greater than or equal to $flow_{12}$
        for $i, i' \in i_1, i_2 | i \neq i'$ **do**
          $\bar{y}_{ti} \leftarrow 1$
          $\sum_{t\in\mathcal{T}} \bar{y}_{ti'} \leftarrow 0$
          for $j \in \mathcal{F}$ **do**
            $\bar{z}_{ji} \leftarrow \bar{z}_{ji} + \bar{z}_{ji'}$
          end for
          $Obj^1 \leftarrow ObjEval(Y_k, Z_k)$
          CheckUpdate($Obj^{best}, Y^{best}, Z^{best}, Y_k, Z_k, Obj^1$)
        end for
        $\mathcal{T}^r \leftarrow \{t \,|\, t \in [1, ..t_C - 1],\ t \notin \{t_1, t_2\}\}$
        if $\mathcal{T}^r \geq 1$ **then**
          for $t \in \mathcal{T}^r$ **do**
            unfreeze the decision variables y_{ti_1}, y_{ti_2}, z_{ji_1}, and z_{ji_2}
          end for
          $Obj^1 \leftarrow MINLPsolve().Obj$
          CheckUpdate($Obj^{best}, Y^{best}, Z^{best}, Y_k, Z_k, Obj^1$)
        end if
      end if
      Report $Obj^{best}, Y^{best}, Z^{best}$
    end for
    Report $Obj^{best}, Y^{best}, Z^{best}$ across all starts (k) as final heuristic solution

Table 1: Data instances used in computational analysis

	Problem entities		
Instances	#proc.units	#cetp.locs	#Cetp.types
1	20	2	3
2	20	2	6
3	20	2	9
4	30	3	3
5	30	3	6
6	30	3	9
7	40	3	3
8	40	3	6
9	40	3	9
10	50	4	3
11	50	4	6
12	50	4	9
13	60	5	3
14	60	5	6
15	60	5	9
16	80	5	3
17	80	5	6
18	80	5	9
19	100	6	3
20	100	6	6
21	100	6	9
22	150	8	3
23	150	8	6
24	150	8	9
25	200	13	3
26	200	13	6
27	200	13	9
28	380	18	3
29	380	18	6
30	380	18	9

Table 2: Data value ranges related to the processing units

	Minimum	Maximum	Average
Distance from processing units to potential CETP locations in km	0.10	23.40	9.00
Effluent flow from the processing units in kl/day	166.96	570.70	316.01
Pollutant discharge from the processing units in g/kl	309.56	492.08	439.08

Table 3: Parameter values considered

Parameter	**Value**
K	0.68
s	0.9
r	1.3

To estimate the CETP capacity types, we have collected primary and secondary data related to different CETP capacity options applicable to the scenario, as presented in table 4. We use this data to develop an approximate relationship between CETP size in terms of designed capacity (kl/day), cost of installation, and area required for installation. For each data instance, we calculate the total effluent volume from all the processing units and consider CETP capacity types corresponding to that volume. For example, if the total effluent volume from all the processing units is 120,000 kl/day, we consider three types with capacities 40,000, 80,000, and 120,000 for a data instance with three types considered and similarly for instances with six and nine CETP types.

Table 4: Installation costs and area required for multiple CETP types

CETP types	Plant Size (kl/day)	Capital Costs (in 100,000 INR)	Land Area (m2)
1	2000	37.88	7950.5595
2	3750	29.81	10634.275
3	5400	77.02	12588.613
4	8000	83.9	15099.214
5	11000	139.53	17496.179
6	12000	97.98	18214.907
7	41725	574.04	32421.444
8	47000	307.84	34257.307
9	54000	391.16	36530.068

5.2. Computational results

The computational results from the proposed solution approaches implemented on the data sets are presented in table 5. First column lists the data instances described previously in table 1 in terms of problem elements and size. For each instance, we present the results of both the OA-based branch and cut algorithm as well as the MILP based neighborhood heuristic in the same row for easy comparison. Column-2 shows the number of lazy cuts added during the MILP branch and bound process. The branch and bound search for each instance was constrained to run for a maximum of four hours (14,400 seconds). The time shown in column-3 includes problem buildup, branch and bound root relaxation time and branch and bound search time till either optimal solution or the time limit of four hours is reached. So, in some instances the total reported time exceeds four hours, although the branch and bound search time was always lesser then the specified number. The best objective value obtained from the branch and cut approach is reported in the column-4. For two of the instances, 23^{rd} and 28^{th}, we could not get any solution. The best LP lower bound obtained from the MILP branch and bound solution is reported next with the optimality gap reported in column-5. Finally, we report the solution obtained from the proposed heuristic approach. The best objective value obtained, running time till completion and optimality gap from the LP lower bound in column-4 are presented consecutively.

Table 5: Computational results

Instances	OA-based branch and cut					MILP based neighborhood search heuristic		
	#lazycuts	Run.time (sec)	Obj.best (mil. INR)	LP.lb	Opt.gap (%)	Obj.Best (mil. INR)	Run.time (sec)	Opt.gap (%)
1	1	3.40	47.64	47.64	0.00	47.64	4.27	0.0
2	1	5.55	47.63	47.63	0.00	47.63	3.56	0.0
3	1	1.26	47.52	47.52	0.00	47.52	3.73	0.0
4	2	11.90	68.39	68.39	0.00	68.39	5.92	0.0
5	2	15.23	67.34	67.34	0.00	67.34	6.29	0.0
6	1	9.20	67.32	67.32	0.00	69.66	6.66	3.4
7	3	25.41	90.73	90.73	0.00	90.79	7.98	0.1
8	3	31.56	89.05	89.05	0.00	89.05	8.89	0.0
9	0	19.19	88.89	88.89	0.00	88.89	10.00	0.0
10	6	73.66	109.03	109.03	0.00	111.23	14.26	2.0
11	12	110.71	108.68	108.68	0.00	108.68	15.03	0.0
12	4	58.80	108.35	108.35	0.00	109.32	14.16	0.9
13	13	246.04	130.28	130.28	0.00	130.57	23.24	0.2
14	7	348.63	129.34	129.34	0.00	129.35	19.93	0.0
15	7	185.64	129.21	129.21	0.00	129.51	23.02	0.2
16	9	354.98	169.79	169.79	0.00	169.90	36.68	0.1
17	24	500.09	169.22	169.22	0.00	169.23	55.17	0.0
18	15	271.57	169.17	169.17	0.00	170.41	46.51	0.7
19	18	2,875.11	196.70	196.70	0.00	197.40	56.68	0.4
20	1	3,234.91	190.81	190.81	0.00	195.43	80.94	2.4
21	12	2,594.23	190.93	190.93	0.00	198.42	81.48	3.8
22	36	32,073.11	277.00	276.93	0.02	289.61	209.44	4.4
23	-	28,589.15	-	-	-	278.73	269.15	-
24	8	17,659.12	279.14	257.61	7.71	280.24	387.84	8.1
25	20	28,869.61	374.97	339.44	9.48	358.44	1,432.53	5.3
26	8	43,310.90	377.54	318.48	15.64	352.10	1,319.95	9.5
27	13	46,433.96	401.28	320.05	20.24	357.01	1,183.47	10.4
28	2	36,960.00	-	-	-	654.92	13,451.97	-
29	3	45,965.00	688.81	510.73	25.85	615.85	9,552.97	17.1
30	2	57,866.29	645.08	464.16	28.05	584.54	18,854.18	20.6

Computational analysis reveals that the OA-based branch and cut algorithm presents itself as a reliable solution approach for the proposed MINLP problem. For instances with less than 150 processing units, the solution method reported optimal or near-optimal results, although the solution time increases rapidly from a few seconds for the smaller instances to almost an hour. For instances with more than or equal to 150 processing units, the running time exceeded the specified time limit, and the achieved optimality gap widened with problem size. The heuristic approach yields quick solutions to smaller instances 1 - 22, with an average optimality gap of 0.8% ranging from 0 to 4.4%. Even for the larger instances it reports, on an average, 7% lower than the best-found integer solution from the exact method in substantially lesser solution times in comparison to the exact OA-based method.

6. Trade-offs between financial and social costs of a CETP system

The problem consists of multiple objectives that may conflict with each other and are expected to vary widely with changes in parameters. We would like to understand the dynamics of trade-offs among various financial costs inherent in the CETP system. The processing units collaborating to install a group of CETPs would mainly focus on financial costs. Although the purpose of the CETP system is to minimize environmental damage, large-scale operations have some negative consequences on the local environment and society. One such cost associated with a CETP system is the land area usage for CETP installation. Firms do bear the actual monetary cost of this resource when installing and using a CETP system, but the societal costs of land area usage have a larger negative impact on the region. So, we analyze the social cost of land area usage along with the installation, transportation, and treatment costs of the CETP system as a multi-objective nonlinear integer programming problem. It consists of four different objectives- CETP installation cost (Obj^I), CETP treatment cost (Obj^T), CETP transportation cost (Obj^A), and the total area used in CETP installation (Obj^{Ar}). The four objectives can be expressed as shown in equations 25 - 28. Here A_t refers to the area usage in square meter (sq.m) in installing and operating a CETP of type t. We are not attaching any additional per-unit cost parameter to the social cost, as subjective estimations are varying widely in terms of estimated impact on the local environment and society in the long term. The social costs are estimated in its original unit of square meter (sq.m) for the total area used. We use the smallest data instance 3 (refer to table 1) with 20 processing units, 2 CETP potential locations, and 9 CETP types for our multi-objective analysis.

$$Obj^I = \sum_{i \in \mathcal{I}} \sum_{t \in \mathcal{T}} C^F_{ti} y_{ti} \tag{25}$$

$$Obj^T = \sum_{i \in \mathcal{I}} K x^q_i p^r_i \tag{26}$$

$$Obj^A = \sum_{i \in \mathcal{I}} \sum_{j \in \mathcal{F}} C^A_{ji} z_{ji} \tag{27}$$

$$Obj^{Ar} = \sum_{i \in \mathcal{I}} \sum_{t \in \mathcal{T}} A_t y_{ti} \tag{28}$$

We present the trade-offs between the four objectives. We do a pairwise comparison of each pair of objectives mentioned above using the ξ-constraint method. In this, we first find the minimum and maximum feasible values of all objectives by running the whole model independently w.r.t. each objective at a time. Table 6 lists the maximum and minimum feasible values of each objective.

Table 6: Maximum and minimum feasible values of different objectives

	Minimum	Maximum
Obj^I	7,635,924	15,271,848
Obj^C	33,681,941	46,381,944
Obj^T	5,149,277	5,426,897
Obj^A	13,820	27,641

Installation cost and transportation cost are inversely related to each other. For the given parameters, we get a Pareto frontier with three Pareto optimal points, as shown in figure 3. There is a sudden increase in transportation cost, below a threshold reduction in installation cost, as fewer CETPs are needed for a lesser installation cost.

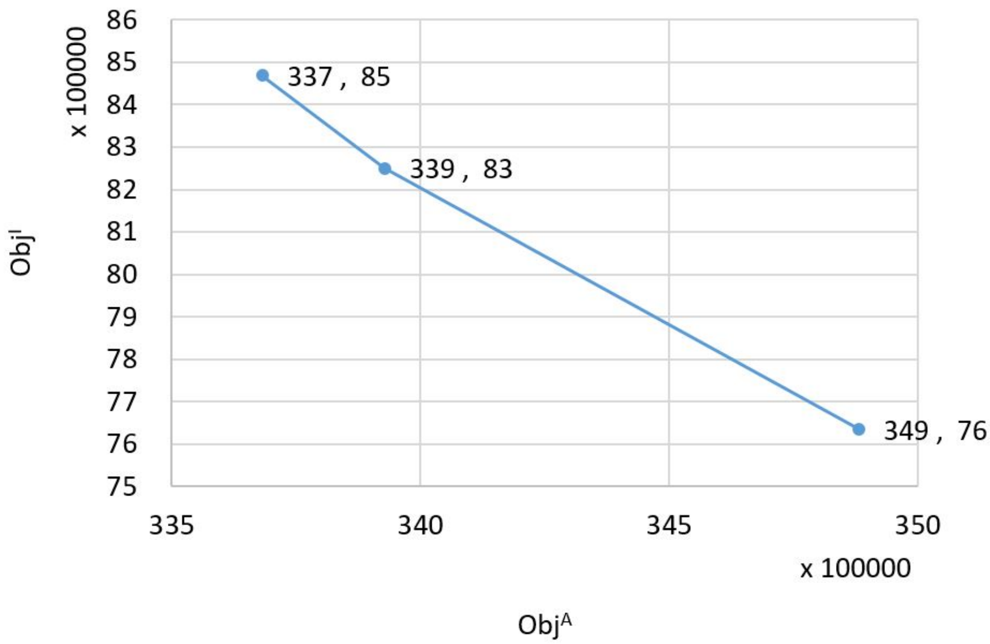

Figure 3: Installation cost vs transportation cost

We can see an inverse relationship between treatment cost and transportation costs. A total of 10 Pareto points can be derived. Small changes in transportation cost leads to corresponding small changes in the treatment cost. Lower transportation cost implies more number of CETPs, which leads to higher treatment costs, as we are unable to take advantage of the economy of scale in treatment. Figure 4 shows the interrelationship between the two objectives.

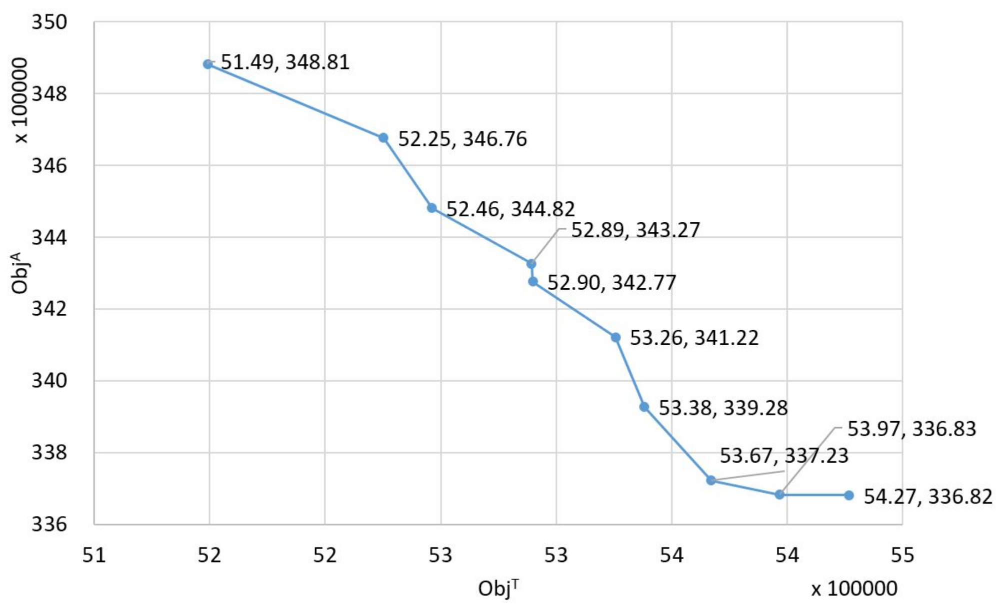

Figure 4: transportation cost vs treatment cost

There exists an inverse relationship between transportation cost and area used. Three Pareto optimal points are identified for the data. One can see a sudden jump in the area used on reducing transportation costs, as more CETPs are required. Figure 5 presents the trade-off between the two objectives.

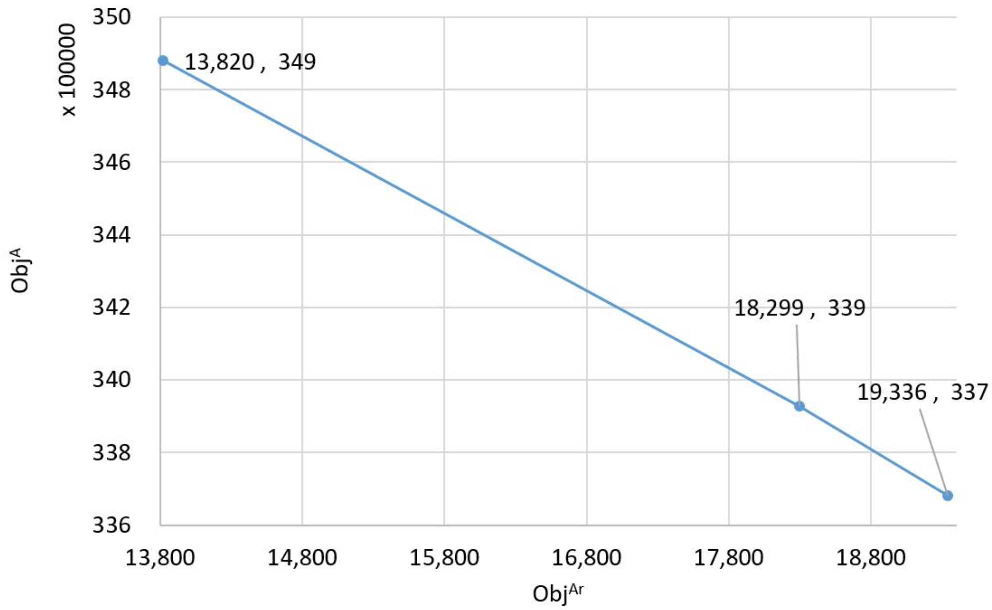

Figure 5: transportation cost vs Area

The three objectives of installation cost, treatment cost, and the area used are directly related to each other, and we get a single Pareto optimal point for all three objectives together, as given in table 7.

Table 7: Pareto optimal point for installation cost, treatment cost and area used

Obj^T	Obj^I	Obj^{Ar}
5,149,272	7,635,924	13,820

7. Conclusions and future work

The implementation of a shared ecosystem of effluent treatment facilities serving a cluster of processing units of similar industries is a positive trend for small and medium enterprises. A strategic problem in this scenario is designing an efficient and effective CETP system, which includes selecting appropriate types of CETPs in terms of treatment capacity, deciding locations for CETP installations, and allocating processing units to each installed CETP. The problem is modeled as a capacitated facility location-allocation problem with nonconvex treatment costs. An MINLP mathematical formulation is presented for the problem. Thus, the model extends the capacitated facility location model in two ways- by considering multiple discrete options of facilities, in terms of capacity, cost, and other physical characteristics, to be installed at a location and nonconvex treatment costs. As capacitated facility location problems are NP-hard, finding optimal solutions to this important strategic problem can be computationally challenging.

A case study from an industrial cluster is presented, and data is estimated from a real planning scenario. Further, an exact convexification strategy is proposed for the nonconvex treatment costs. We present two solution approaches for the MINLP model. First, an outer approximation cutting plane algorithm is implemented in a branch and cut framework. To solve large instances, we incorporate the the branch and cut method into a MILP based neighborhood search heuristic. The computational results with the two approaches suggest that branch and cut method can solve medium sized instances to optimality in a reasonable time. The heuristic approach yields good quality solutions in comparatively smaller computational time for larger instances. We analyze the trade-offs between various cost components and the social cost of area usage through multi-objective analysis.

Future research can explore interconnected CETPs and incorporate uncertainties in effluent discharge from processing units. Further, seasonality in effluent from processing units can be considered to develop a multi-period decision making model.

References

H. Baharmand, T. Comes, and M. Lauras. Bi-objective multi-layer location–allocation model for the immediate aftermath of sudden-onset disasters. *Transportation Research Part E: Logistics and Transportation Review*, 127:86–110, 2019.

C.-H. Chen and C.-J. Ting. Combining lagrangian heuristic and ant colony system to solve the single source capacitated facility location problem. *Transportation research part E: logistics and transportation review*, 44(6):1099–1122, 2008.

L. Cooper. Location-allocation problems. *Operations research*, 11(3):331–343, 1963.

J. N. de Figueiredo and S. F. Mayerle. Designing minimum-cost recycling collection networks with required throughput. *Transportation Research Part E: Logistics and Transportation Review*, 44(5):731–752, 2008.

M. A. Duran and I. E. Grossmann. An outer-approximation algorithm for a class of mixed-integer nonlinear programs. *Mathematical programming*, 36(3):307–339, 1986.

R. Fletcher and S. Leyffer. Solving mixed integer nonlinear programs by outer approximation. *Mathematical programming*, 66 (1-3):327–349, 1994.

K. Gokbayrak and A. S. Kocaman. A distance-limited continuous location-allocation problem for spatial planning of decentralized systems. *Computers & Operations Research*, 88:15–29, 2017.

A. W. Hammad, A. Akbarnezhad, and D. Rey. Sustainable urban facility location: Minimising noise pollution and network congestion. *Transportation Research Part E: Logistics and Transportation Review*, 107:38–59, 2017.

I. Harris, C. L. Mumford, and M. M. Naim. A hybrid multi-objective approach to capacitated facility location with flexible store allocation for green logistics modeling. *Transportation Research Part E: Logistics and Transportation Review*, 66:1–22, 2014.

J. Kim, M. J. Realff, J. H. Lee, C. Whittaker, and L. Furtner. Design of biomass processing network for biofuel production using an milp model. *Biomass and bioenergy*, 35(2):853–871, 2011.

B. Lin, S. Liu, R. Lin, J. Wang, M. Sun, X. Wang, C. Liu, J. Wu, and J. Xiao. The location-allocation model for multi-classification-yard location problem. *Transportation Research Part E: Logistics and Transportation Review*, 122:283–308, 2019.

Y. Liu, N. Cui, and J. Zhang. Integrated temporary facility location and casualty allocation planning for post-disaster humanitarian medical service. *Transportation Research Part E: Logistics and Transportation Review*, 128:1–16, 2019.

R. Manzini and E. Gebennini. Optimization models for the dynamic facility location and allocation problem. *International Journal of Production Research*, 46(8):2061–2086, 2008.

P. A. Miranda, C. A. Blazquez, R. Vergara, and S. Weitzler. A novel methodology for designing a household waste collection system for insular zones. *Transportation Research Part E: Logistics and Transportation Review*, 77:227–247, 2015.

D. Mogale, M. Kumar, S. K. Kumar, and M. K. Tiwari. Grain silo location-allocation problem with dwell time for optimization of food grain supply chain network. *Transportation Research Part E: Logistics and Transportation Review*, 111:40–69, 2018.

S. Mundle, U. Shankar, and S. Mehta. Incentives and regulation for pollution abatement with an application to waste water treatment. 1995.

N. Parker, Y. Fan, and J. Ogden. From waste to hydrogen: an optimal design of energy production and distribution network. *Transportation Research Part E: Logistics and Transportation Review*, 46(4):534–545, 2010.

T. R. P. Ramos, M. I. Gomes, and A. P. Barbosa-Póvoa. Economic and environmental concerns in planning recyclable waste collection systems. *Transportation Research Part E: Logistics and Transportation Review*, 62:34–54, 2014.

R. T. Rockafellar. Lagrange multipliers and optimality. *SIAM review*, 35(2):183–238, 1993.

M. Shahabi, A. Unnikrishnan, and S. D. Boyles. An outer approximation algorithm for the robust shortest path problem. *Transportation Research Part E: Logistics and Transportation Review*, 58:52–66, 2013.

D. Z. Wang, H. Liu, and W. Szeto. A novel discrete network design problem formulation and its global optimization solution algorithm. *Transportation Research Part E: Logistics and Transportation Review*, 79:213–230, 2015.

S. Wang and Q. Meng. Sailing speed optimization for container ships in a liner shipping network. *Transportation Research Part E: Logistics and Transportation Review*, 48(3):701–714, 2012.

S. Wang, Q. Meng, and Z. Liu. A note on "berth allocation considering fuel consumption and vessel emissions". *Transportation Research Part E: Logistics and Transportation Review*, 49(1):48–54, 2013.

S. Wu and Z. Yang. Locating manufacturing industries by flow-capturing location model–case of chinese steel industry. *Transportation Research Part E: Logistics and Transportation Review*, 112:1–11, 2018.

J. Zhao, L. Huang, D.-H. Lee, and Q. Peng. Improved approaches to the network design problem in regional hazardous waste management systems. *Transportation research part E: logistics and transportation review*, 88:52–75, 2016.

J. Zhou and B. Liu. New stochastic models for capacitated location-allocation problem. *Computers & industrial engineering*, 45 (1):111–125, 2003.