Disservice Representation Using the Gini Coefficient in Semi-desirable Facility Location Problems

Haluk Yapicioglu
Department of Industrial Engineering
Anadolu University
Eskisehir, TURKEY, 26470
hyapicio@anadolu.edu.tr

Alice E. Smith
Department of Industrial and Systems Engineering
Auburn University
Auburn, Al, 36849
smithae@auburn.edu

***Abstract*— We consider various bi-objective models for the semi-desirable facility location problem. In these problems, the disservice caused by the facility is traditionally measured by distance-related objective functions. In this paper, we modify the objective function representing the disservice using the Lorenz curve and the Gini coefficient. Both of these concepts are widely used in the economics literature to measure the discrepancy in wealth distribution within a population. The use of the Gini coefficient enables the measurement of how the disservice caused by the facility varies across different Pareto optimal solutions. We use a bi-objective particle swarm optimizer (bi-PSO) to compare how the change in the objective function representing the disservice affects the recommended location of the facility. Results suggest that some solutions identified as "Pareto optimal" by traditional formulations are dominated by other solutions when the Gini coefficient is used. Additionally, the use of the Gini coefficient causes a change in the "optimal" location of a semi desirable facility in some instances.**

Keywords-- semi-desirable facility location problem, Gini coefficient, multiobjective PSO

I. INTRODUCTION

Semi-desirable facilities are inevitable components of an industrialized society. While providing necessary services, semi-desirable facilities also create undesirable effects such as bad smells, hazardous chemicals and traffic congestion. Garbage dump sites, water treatment facilities, and airports are among the examples of semi-desirable facilities. On the one hand, decision makers would like to locate such facilities as close as possible to population centers, due to considerations of transportation cost. On the other hand, locating such facilities close to population centers may not be the best option when their undesirable effects are taken into account. Thus, in order to solve semi-desirable facility location problems, bi-objective models are widely used in the literature. In a bi-objective model, one of the objective functions represents the transportation costs from/to the facility and the other represents the disservice imposed by the facility on the population.

The research objective of this paper is to use the Gini coefficient as the objective representing the disservice imposed by the facility and compare the results with those obtained by traditional models in the literature. To do so, we use a previously devised bi-objective particle swarm optimizer, biPSO [2].

The paper is organized as follows. In Section II, we discuss previously studied semi-desirable facility location models followed by the definition of the Lorenz curve, the Gini coefficient and the implications of the use of the Gini coefficient in Section III. Section IV gives the details of the calculation of the Gini coefficient. Section V introduces the details of the bi-PSO used in this paper. Results are presented in Section VI, followed by a discussion in Section VII.

II. PREVIOUSLY STUDIED SEMI-DESIRABLE FACILITY LOCATION MODELS

There is a large literature on the semi-desirable facility location problem. An extensive survey on these models can be found in [1], [2] and [3]. Among the proposed models, we consider two where the objective function that defines obnoxious effects of the facility is the sum of undesirable effects over the set of population centers. Details of these models are provided below.

Skriver and Andersen [4] formulate the semi-desirable facility location as follows:

$$\begin{aligned} \min \quad & f_1(\mathbf{x}) = \sum_{i=1}^{n} w_{1i} d(\mathbf{x}, \mathbf{a}_i) \\ \min \quad & f_2(\mathbf{x}) = \sum_{i=1}^{n} w_{2i} d(\mathbf{x}, \mathbf{a}_i)^{-b} \end{aligned} \tag{1}$$

Where $d(\mathbf{x}, \mathbf{a}_i)$ is the Euclidean distance between the fixed population center point $\mathbf{a}_i = (a_i^x, a_i^y)$ and the facility to be located at point $\mathbf{x} = (x^x, x^y)$. Each objective is weighted by $\mathbf{w}_1$ and $\mathbf{w}_2$. $f_1(\mathbf{x})$ represents the total transportation costs. $f_2(\mathbf{x})$ represents the disservice created by the facility.

Yapicioglu et al. [2] use the weighted total distance to represent the transportation costs and define a piecewise non-increasing function to represent the facility's disservice. Their formulation is as follows:

$$\min \quad f_1(\mathbf{x}) = \sum_{i=1}^{n} w_{1i} d(\mathbf{x}, \mathbf{a}_i)$$
$$\min \quad f_2(\mathbf{x}) = \sum_{i=1}^{n} f_{2i}(\mathbf{x}, \mathbf{a}_i) \tag{2}$$

where

$$f_{2i}(\mathbf{x}, \mathbf{a}_i) = \begin{cases} M, & \text{if} \quad w_{2i} d(\mathbf{x}, \mathbf{a}_i) \le d_1 \\ M - m(w_{2i} d(\mathbf{x}, \mathbf{a}_i)), & \text{if} \quad d_1 < w_{2i} d(\mathbf{x}, \mathbf{a}_i) \le d_2 \\ 0, & \text{if} \quad d_2 < w_{2i} d(\mathbf{x}, \mathbf{a}_i) \end{cases} \tag{3}$$

M, m, d_1 and d_2 are the function parameters that are customized with respect to the problem at hand. Yapicioglu et al. [2] used this formulation both with Euclidean and Manhattan distances.

III. THE LORENZ CURVE AND THE GINI COEFFICIENT

The Lorenz curve and the Gini coefficient are concepts used to measure the equity of a good or a service among the constituents of a population. The Lorenz curve [5] is a graphical representation of a distribution of a good or service across the members of a population. The Gini coefficient is defined as the area between a 45 degree line and the line representing the cumulative equity in a Lorenz curve [6]. Translated into the facility location literature, the Gini coefficient provides a measure of the fairness of the location of a facility with respect to the fixed locations of the population centers that are receiving service from the facility. Drezner et al. [7] first used the concept of the Gini coefficient in the context of a facility location problem and proposed a solution approach to find the optimal location of a single facility on a plane. In this paper, we use the Gini coefficient to measure the cumulative disservice created by the facility located at point $\mathbf{x} = (x^x, x^y)$, over the population centers $\mathbf{a}_i = (a_i^x, a_i^y)$, i = 1, 2, …, n. Using this definition of the Gini coefficient (as explained in detail in subsequent sections), the disservice values are normalized by dividing the individual disservice values by the cumulative disservice. Thus, the Gini coefficient achieves at its minimum (0) when the facility is located *equidistant* to all population centers, no matter how bad that location is from a disservice point of view. Since we are using the Gini coefficient to model the undesirable effects, a bad location actually refers to a location that is very close to population centers.

Consider the two cases depicted in Table 1. Assume that the larger the value of the arbitrary function $f_i(\mathbf{x})$, the larger the disservice to the fixed point at i. In Case 1, the facility creates minimal disservice to population centers 1, 2, 3 and 4 whereas the disservice to population center 5 is 96 times higher than that of the others in the example. For Case 1, the Gini coefficient is calculated as 0.76 (Figure 1). In Case 2 however, the facility incurs much higher disservice to population centers 1, 2, 3 and 4 and the disservice incurred to population center 5, comparatively speaking, is lesser. In Case 2, the Gini coefficient is calculated as 0.08 (Figure 2). Note that in both cases the total disservice created by the facility is 100 units. When we use the Gini coefficient to model the disservice generated by a semi-desirable facility, we favor a more equitable distribution of overall disservice.

TABLE I. LORENZ CURVES, CASE1, CASE 2

	i	1	2	3	4	5
Case 1	$f_i(\mathbf{x})$	1	1	1	1	96
	$\sum_i f_i(\mathbf{x})$	1	2	3	4	100
Case 2	$f_i(\mathbf{x})$	15	20	20	20	25
	$\sum_i f_i(\mathbf{x})$	15	35	55	75	100

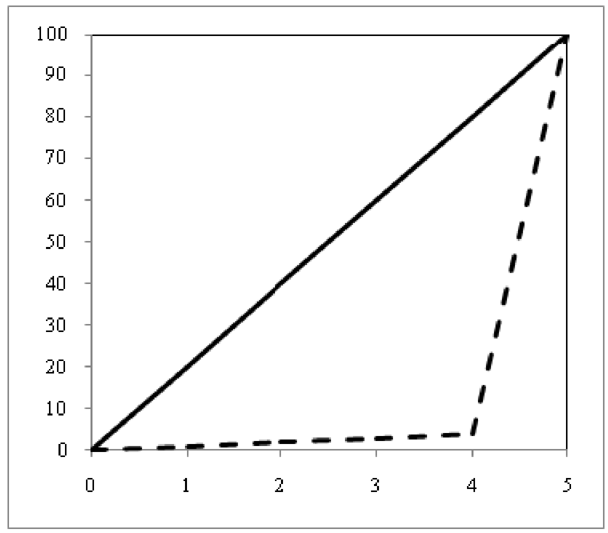

Figure 1 Lorenz Curve, Case 1

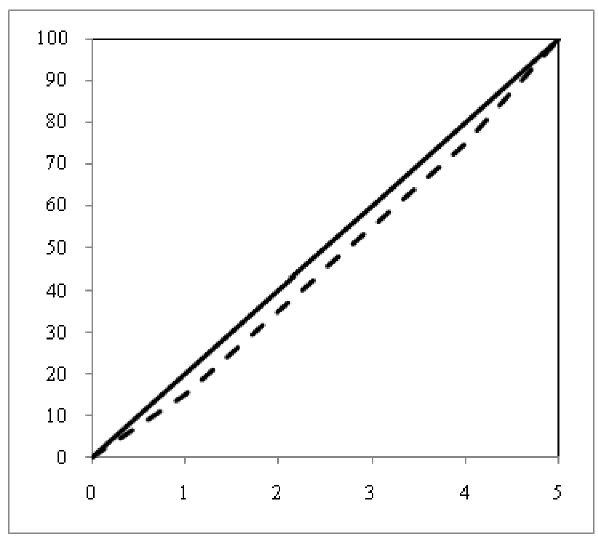

Figure 2 Lorenz Curve, Case 2

IV. CALCULATION OF THE GINI COEFFICIENT FOR THE DISSERVICE CAUSED BY A FACILITY

As stated previously, the Gini coefficient is defined over the Lorenz curve (dotted line in Figures 1 and 2). In a Lorenz curve, the x axis shows the cumulative percentage of the population, whereas the y axis is used to depict the cumulative percentage of the good and/or the service. A 45 degree straight line from (0, 0) to (1, 1) defines the most equitable distribution and is called "the line of equity" (the straight line in Figures 1 and 2).

In this paper, we work with the three models (the model from [4] and model from [2] with both Euclidean and Manhattan distances). To calculate the Gini coefficient, we use $f_2(\mathbf{x})$ in every case. For this, we calculate the disservice to each population center as defined by the $f_2(\mathbf{x})$ of the respective model and then sort these values in increasing order. Then the Gini coefficient for the facility located at $\mathbf{x}$ is given by [7]

$$G(\mathbf{x}) = \frac{\sum_{i=1}^{n}\left[\frac{2i-1}{n}-1\right] f_{2(i)}(\mathbf{x})}{f_2(\mathbf{x})} \tag{4}$$

Where $f_{2(i)}(\mathbf{x})$ is the i^{th} smallest disservice caused by the facility at $\mathbf{x}$. In all cases, transportation costs are calculated by $f_1(\mathbf{x})$.

V. Bi-objective PSO

Particle Swarm Optimization (PSO) was first proposed by Kennedy and Eberhart [8, 9, 10]. They were inspired by the collective behavior of animals, specifically the movement patterns of bird flocks. In such populations, individuals determine their travel path considering both their previous experience (cognition) and the collective behavior of the entire flock (social interaction). In this fashion, the flock synergistically locates a food source. PSO is a population based optimization method in which individuals (e.g. particles) use the principles of flocking in search of a best solution of an optimization problem. In PSO, each particle is composed of three vectors (current location, **c**, current velocity, **v**, and the best solution found by the particle, **p**) and two fitness values (*c-fitness* and *p-fitness*, that represent the particle's current and best value of the objective function, respectively). At each iteration of the PSO the current position of the particle is updated according to equations below:

$$\begin{aligned} v_{id} &= K\left(v_{id} + \varphi_1 \mathrm{U}(0,1)(p_{id} - c_{id}) + \varphi_2 \mathrm{U}(0,1)(p_{gd} - c_{id})\right) \\ c_{id} &= c_{id} + v_{id} \end{aligned} \tag{5}$$

where

- $\mathbf{c}_i$ the current position of particle i,
- $\mathbf{v}_i$ the current velocity of particle i,
- $\mathbf{p}_i$ the best solution identified by particle i,
- $\mathbf{p}_g$ the best solution in the neighborhood of particle i,
- φ_1, φ_2 learning rates governing the cognition and social interaction within the swarm, respectively,
- d is the d^{th} dimension of the corresponding vector, d = 1, 2, … D,
- U(0, 1) is a uniform random number in the interval [0, 1],

$$K = \frac{2}{\left|2-\varphi-\sqrt{\varphi^2-4\varphi}\right|} \text{ where } \varphi = \varphi_1 + \varphi_2 \text{ and } \varphi > 4$$

is the constriction coefficient controlling the velocity update to prevent explosion and provide stability as explained in [11]. Once **c** is updated, *c-fitness* is calculated, and if it is better than *p-fitness*, **p** and *p-fitness* are also updated.

After successful implementations in single-objective problems, PSO was first applied to multi-objective settings by Moore and Chapman [12], followed by Coello Coello and Lechuga [13], Hu and Eberhart [14], Fieldsend and Singh [15], and Parsopoulos et al. [16]. Yapicioglu et al. [2] also developed a bi-objective PSO (bi-PSO) where the non-dominated solutions discovered during the search are kept in an external repository called Archive. During the velocity update, a randomly selected member of the external repository serves as the swarm's best. Thus, the formulae used to update the particle modifies to:

$$\begin{aligned} v_{id} &= K\left(v_{id} + \varphi_1 \mathrm{U}(0,1)(p_{id} - x_{id}) + \varphi_2 \mathrm{U}(0,1)(p_{gd} - x_{id})\right) \\ x_{id} &= x_{id} + v_{id} \\ g &= \mathrm{int}\left(\mathrm{U}(0,1)\left|\mathrm{Pareto\_Front}\right|\right) \end{aligned} \tag{6}$$

Like the constriction PSO, once **c** is updated, *c-fitness* is also updated. Then if *c-fitness* dominates *p-fitness*, **p** *and p-fitness* are also updated and new solutions are checked against the solutions in Pareto Front for Pareto dominance. The pseudo-code of the bi-PSO is provided below:

```
Procedure bi-PSO{
    t = 0;
    Initialize Swarm(t); //Swarm at tth cycle
    Evaluate the particles of the Swarm(t);
    Initialize Archive (t); //Repository at tth cycle
      While (Not Done) {
          for each particle i    Swarm(t) {
              Select a member from Archive(t) randomly;
              Update (i, Archive(t));
              Evaluate particle i;
          }end for
          Update Archive(t);
          t = t + 1;
      } end while
} end procedure
```

In this paper, $\varphi_1 = \varphi_2 = 2.05$, the swarm size is 500, and the search terminates after 2,000 iterations. We did not set any limits on the size of the Archive as our experimentation suggested that it was not necessary. The bi-PSO used in this paper is advantageous due to its relative ease of implementation and its proven success in semi-desirable facility location models [2].

VI. Comparison of Results: Original Disservice Formulations vs. the Gini Coefficient

In this section, we consider the disservice formulations and problems from [4] and [2] and compare the results from those papers with ones using the same models but employing the Gini coefficient. Figures 3, 5 and 6 depict the differences. In these figures, Fixed Points refer to the population centers for the problem at hand (depicted with ▪), the efficient set obtained using the original formulation is referred to as Orig. Form. (depicted with ×) and the efficient set obtained using the Gini coefficient for the disservice is referred to as Gini (depicted with ♦).

Figure 3 provides the efficient sets obtained by bi-PSO for the original formulation from the Skriver and Andersen model [4] and the effect of the change in the objective function representing the obnoxious effects with the Gini coefficient. As can be verified from Figure 3, the efficient regions lying on the left side found using the original formulation from [4] are no longer identified when we employ the Gini coefficient to calculate the disservice. To investigate the cause of this we manually calculated the Gini coefficients for those points lying on the right side of Figure 3 (a total of 17 points). When we depict these solutions' objective functions, as given in Figure 4, we can verify that those points are actually provides inferior solutions for the Gini formulation in terms of Pareto dominance.

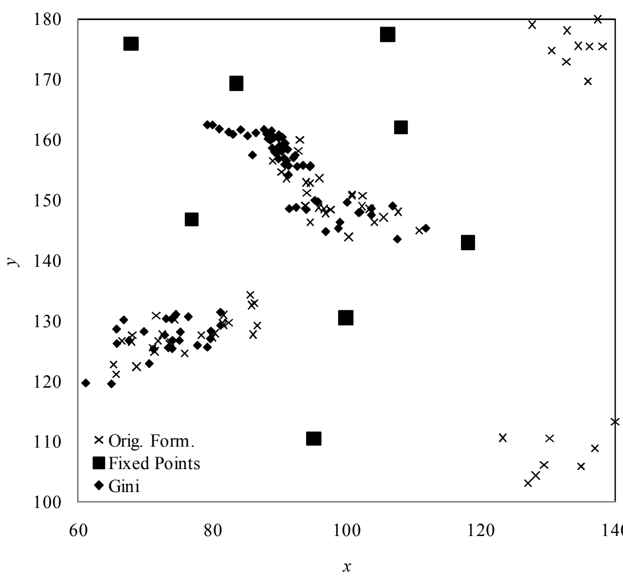

Figure 3 Fixed points (population centers), efficient set from original formulation from [4] and efficient set from the Gini formulation

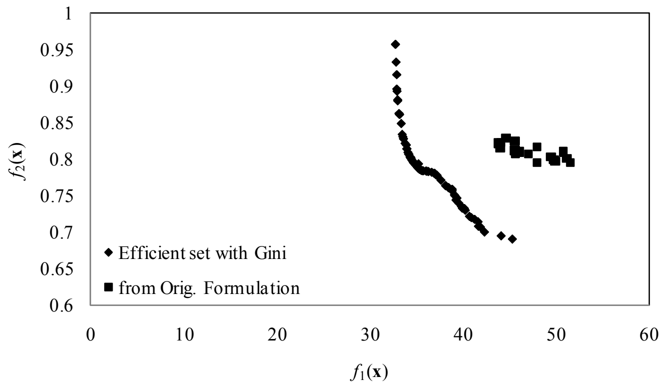

Figure 4 Pareto front and points from the right side of Figure 3 using the Gini objective functions

Using the formulation from [2], the effects of the Gini coefficient are tested using both Euclidean and Manhattan distances with function parameters M = 200, m = 1, d_1 = 10 and d_2 = 30 on a seven population center example problem. Figure 5 provides a comparison of these two different formulations for the Euclidean distance. The original formulation has an efficient set scattered across the feasible region, especially over the bottom part. The formulation with the Gini coefficient however, finds all efficient solutions within the convex hull of the fixed points.

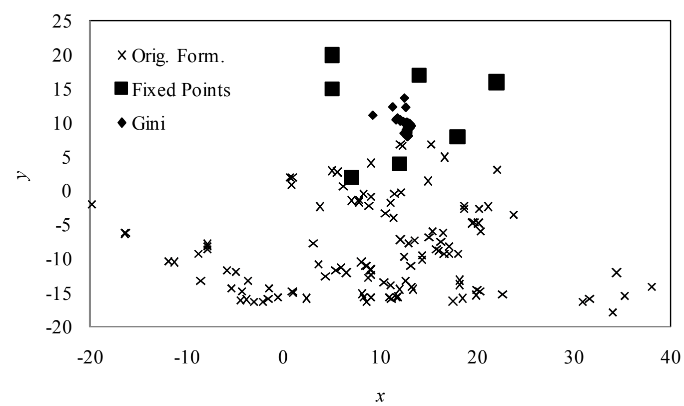

Figure 5 Fixed points (population centers), efficient set from original formulation from [2] and efficient set from Gini formulation, using Euclidean distance

In the case of Manhattan distance, a similar pattern to that of found with the use of Euclidean distance is observed. With the original formulation from [2], the efficient set is scattered across the feasible region, whereas with the Gini coefficient, the efficient set is again bounded by the convex hull of the population centers, as depicted in Figure 6.

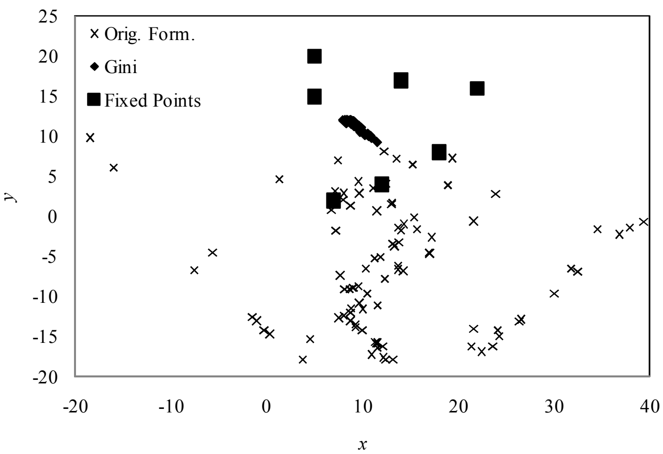

Figure 6 Fixed points (population centers), efficient set from original formulation from [2] and efficient set from Gini formulation, using Manhattan distance

VII. Discussion of the Results

As seen in the previous section, with the use of the Gini coefficient in the Skriver and Andersen model [4], two of the efficient regions identified with the original formulation no longer belong to the efficient set. With the Yapicioglu et al. model [2], the results are more pronounced. With the exception of very few points, the efficient regions obtained by the two different approaches to quantify undesirable effects do not overlap. A common observation in all test problems is that when we use the Gini coefficient to measure the undesirable effects, the suggested locations of the facility occurs at regions where the transportation costs are smaller. With the Gini coefficient, the total cost of disservice is actually higher, but, it is more equally distributed among the population centers. In other words, the justice in terms of the obnoxious effects of the facility is achieved via locating the facility at locations that are close and equidistant from the population centers. The bi-objective PSO was efficient and effective in identifying the Pareto optimum locations of the facility in all cases. It is well suited to this application in complex facility location problems. Future work could include the location of multiple semi-obnoxious facilities which should be readily handled by the PSO.

References

[1] Erkut, E. and Neuman, S. "Analytical models for locating undesirable facilities." European Journal of Operational Research, vol. 40, pp. 275-291, 1989.

[2] Yapicioglu, H., Smith, A.E. and Dozier G.V. Solving the semi-desirable facility location problem using bi-objective particle swarm." European Journal of Operational Research, vol. 177, pp. 733 – 749, 2007.

[3] Karasakal , E. and Nadirler, D. "An interactive sSolution approach for a bi-objective semi-desirable location problem," Journal of Global Optimization, vol. 42, pp. 177-199, 2008.

[4] Skriver, A.J.V. and Andersen, K.A. "The bicriterion semi-obnoxious location (BSL) problem solved by an ε-approximation," European Journal of Operational Research, vol. 146(3), pp. 517 – 528, 2003.

[5] Lorenz, M. O. "Methods for measuring the concentration of wealth." Journal of the American Statistical Association, vol. 9, pp. 209–219, 1905.

[6] Gini, C. "Measurement of inequality and incomes," The Economic Journal, vol. 31 , pp. 124–126, 1921.

[7] Drezner, A., Drezner, Z. and Guyse, J. "Equitable service by a facility: Minimizing the Gini coefficient," Computers and Operations Research, vol. 36, pp. 3240—3246, 2009.

[8] Kennedy, J. and Eberhart, R.C. "Particle swarm optimization," in: Proceedings of the 1995 IEEE International Conference on Neural Networks, pp. 1942–1948, 1995.

[9] Kennedy, J. "The particle swarm: Social adaptation of knowledge," in: Proceedings of the 1997 International Conference on Evolutionary Computation, Indianapolis, IN, 1997, pp. 303–308, 1997.

[10] Kennedy, J. and Eberhart, R.C. Swarm Intelligence, Morgan Kaufmann Publishers, San Francisco, 2001.

[11] Clerc, M. and Kennedy, J. "The particle swarm—explosion, stability, and convergence in a multidimensional complex space," IEEE Transactions on Evolutionary Computation vol. 6 pp. 58–73, 2002.

[12] Moore, J. and Chapman, R. "Application of particle swarm to multiobjective optimization", Department of Computer Science and Software Engineering, Auburn University, 1999.

[13] Coello Coello, C.A. and Lechuga, M.S. "MOPSO: A proposal for multiple objective particle swarm optimization," in: Proceedings of the 2002 Congress on Evolutionary Computation, vol. 2, pp. 1051–1056, 2002.

[14] Hu, X. and Eberhart, R. "Multiobjective optimization using dynamic neighborhood particle swarm optimization," in: Proceedings of the 2002 Congress on Evolutionary Computation, vol. 2, pp. 1677–1681, 2002.

[15] Fieldsend, J.E. and Singh, S. "A multi-objective algorithm based upon particle swarm optimisation, an efficient data structure and turbulence," in: Proceedings of the 2002 UK Workshop on Computational Intelligence, pp. 37–44, 2002.

[16] Parsopoulos, K.E., Tasoulis, D.K., and Vrahatis, M.N., "Multiobjective Optimization Using Parallel Vector Evaluated Particle Swarm Optimization," in: Proceedings of the IASTED International Conference on Artificial Intelligence and Applications (AIA 2004), vol. 2 pp. 823-828, 2004.