OPTIMAL ALLOCATION OF FACTS DEVICES: CLASSICAL VERSUS METAHEURISTC APPROACHES

Y. Del Valle*, J. Perkel*, G. K. Venayagamoorthy**, and R. G. Harley*

* *Department of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA 30332 USA. Email: yamille.delvalle@gatech.edu*
** *Real-Time Power and Intelligent Systems Laboratory, Department of Electrical and Computer Engineering, Missouri University of Science and Technology, Rolla, MO 65409 USA*

Abstract: The application of optimization techniques to allocate and size FACTS devices in a power system is still under study. Currently, there is no widely accepted method since many researchers claim their methods to be better than others. This paper compares the effectiveness of classical and metaheuristic optimization approaches in a simple but realistic case study of optimal allocation of FACTS (Flexible AC Transmission System) devices, considering steady state and economic criteria. Concepts and details about the optimization process that tend to be overlooked in the literature are discussed together with some considerations about statistical analyses in the case of metaheuristic approaches.

Keywords: FACTS devices, classical optimization, Benders' decomposition, branch and bound, evolutionary computation techniques, genetic algorithm, particle swarm optimization, bacterial foraging algorithm.

1. INTRODUCTION

The topic of optimal allocation of FACTS (Flexible AC Transmission System) devices is still in a relatively early stage of investigation. Currently, there is no widely accepted method in the academic circles since many researchers claim their methods to be "better" than others. Considering the present state-of-the-art in this area, the comparison of different methods, particularly between classical and metaheuristic approaches, has been difficult to evaluate because each study focuses on different problem formulations, system sizes and conditions.

This paper provides a common background for comparing the performance of classical and metaheuristic optimization algorithms. A simple but realistic case study of optimal STATCOM allocation (a type of FACTS devices), considering steady state and economic criteria, is used to assess the performance of two classical methods: Bender's decomposition and Branch-and-Bound (B&B), and four metaheuristic approaches: Genetic Algorithm (GA), Particle Swarm Optimization (PSO), Enhanced-PSO, and Bacterial Foraging Algorithm (BFA).

It is important to note that the focus of this paper is not to find a solution to the particular problem, but rather to illustrate and comment on important details about the optimization process that tend to be overlooked in the literature and are unknown to most readers: the discussion about local versus global optimality (for a given objective function), understanding convexity assumptions (that do not apply only to the objective function), and the importance of the algorithm's convergence into feasible regions. Moreover, considering metaheuristic approaches, a statistical analysis is required to evaluate their performance since the use of the typical average and standard deviations must be validated to be meaningful.

The following sections of this paper provide: an optimization background (section 2), concepts and issues that should not be disregarded (section 3), a problem description (section 4), optimization algorithms (section 5), simulation results (section 6), and concluding remarks (section 7).

2. BACKGROUND

The optimization techniques used to solve the optimal allocation of FACTS devices can be decomposed into two primary groups: classical approaches and metaheuristic algorithms, consisting of mainly evolutionary computation techniques (ECTs). A third group of alternative methods, such as modal analysis, may also be considered. However, these methods are primarily based on technical feasibility rather than on finding optimal solutions.

2.1 Classical Optimization Techniques

In the literature, two classes of classical optimization methodologies have been applied to this problem: (i) Mixed Integer Linear Programming (MILP) [1]-[3] and (ii) Mixed Integer Non-Linear Programming (MINLP) [4]-[6].

The MILP formulation, as the name indicates, requires the relationships between all variables to be linear. Thus, this approach can be only used together with DC power flow. The main algorithms for solving the MILP problem are Bender's Decomposition [1], Branch and Bound (B&B), and Gomory cuts [2], [3]. The concluding remarks of the MILP approach indicate that the optimization process is performed in an efficient manner, but the DC power flow represents a limitation for the type problems that can be addressed.

The MINLP formulation allows for the use of non-linear objective function and constraints, thus, AC power flow can be used in this case. The algorithm most widely utilized for solving the MINLP problem is Bender's Decomposition [4]-[6]. Unfortunately, it has been reported that the size and non-convexity of the problem, which depend on the system parameters, are critical issues that may cause convergence problems.

2.2 Metaheuristic Techniques

Computational intelligence based techniques, such as Genetic Algorithm (GA) [4], [5], [7]-[10], Particle Swarm Optimization (PSO) [11]-[13], Simulated Annealing (SA) [7], [14], Tabu Search (TS) [13], [14], and Evolutionary Programming (EP) [15], [16], are alternative methods for solving complex optimization problems.

Candidate solutions play the role of individuals in a population and the cost function determines the environment where the solutions exist. Evolution of the population then takes place and, after the repeated application of biological or social operators, the optimal solution is reached. In general ECTs perform well in MINLP problems. However the scalability of these methods requires further investigation.

3. OPTIMIZATION: CONCEPTS AND ISSUES

3.1 Optimization of offline problems

A common misperception is the belief that the problem of optimal allocation of FACTS devices is not challenging from the optimization perspective because it is an offline problem. Some presume that the solution is as simple as arranging a number of computers in parallel and letting them run, for as long as it takes, until all possible solutions are found and the best one selected among them.

The fact is that, even when this approach is theoretically possible to perform for any system, in practice the number of calculations required to find the solutions to the problem can grow extremely fast as the size of the system increases and the objective function becomes more sophisticated (evaluation of transient performance is computationally intensive). If it is also required to satisfy the N-1 or N-2 contingency criteria, or add stochastic components and uncertainties to the system, the number of cases to evaluate simply becomes uncountable.

Therefore, the study of optimization algorithms applied to system planning problems, such as the problem of FACTS allocation is not trivial.

3.2 Convexity assumptions

The concept of convexity is mostly analyzed in the case of the objective function: if the function is strictly convex a unique optimal solution is guaranteed (Fig. 1.a).

This characteristic is most desirable but it rarely occurs in power system problems. Most of the time the plot of the objective function resembles the function in Fig. 1.b. As a result, gradient descent algorithms are prone to getting trapped in local valleys (local minima). In these cases, special mechanisms, such as injecting randomness to the search, must be considered.

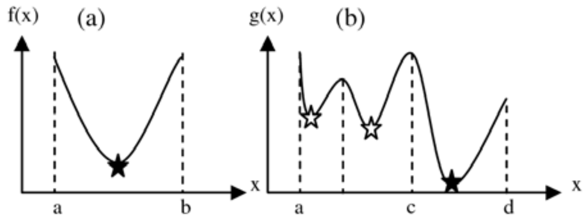

Fig. 1: Global versus local minima

The convexity assumption also applies to the feasible region. For example, in the case of linear programming problems, the optimum can be found (either by simplex method or interior point method) if the feasible region is a convex set, as shown in Fig. 2.a (as opposed to Fig, 2.b) [17].

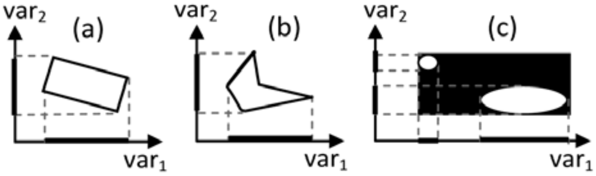

Fig. 2: Convexity of the feasible region

A worst case is presented in Fig. 2.c. where the feasible region consists of several small areas (white) scattered among the area limited by the upper and lower bounds of the decision variables, var_1 and var_2 (black area).

This type of feasible region, as shown later in the paper, is typical when technical constraints are imposed in the power system. The optimization algorithms in this case should have efficient exploration mechanisms so that feasible solutions can be found fast and therefore minimum computational effort is wasted wandering around in infeasible areas.

3.3 Global Optimality

Another aspect that tends to be overlooked in the literature is the discussion of global versus local optimality. Contrary to general opinion, this topic is not related to comparing the values of different objective functions applied to the same power system. Instead, it implies the understanding that, for a given objective function, the problem may a have a unique optimal solution, thus a local optimum is also the global optimum (Fig. 1.a) or the problem may have several local optimal points plus the global optimum. Fig. 1.b represents the latter, where the first white star represent the minimum in the interval [a, b], the second white star is the local minimum in the interval [b, c], and the black star is global minimum in the overall interval [a, d].

The previous concept may seem superfluous, however once an optimization algorithm provides a solution, normally there are no guarantees about its quality. Proof of global optimality can be obtained but only under very specific conditions as in the case of linear programming problems [17]. In the case of MINLP problems, the capability of each algorithm to find the global optimum, without getting trapped in local minima, has to be studied separately.

3.4 Statistical Analysis for metaheuristic methods

Particularly, in the case of metaheuristic algorithms, statistical analysis is required to assess their performance. It is important to note that, currently in literature, the main statistical values used to compare the performance among different metaheuristic optimization algorithms are the mean value of the objective function and its standard deviation.

Intuitively, most people, with some statistical background, understand the average value as the expected outcome of a specific trial and the standard deviation as a measure of the variability of this outcome, nevertheless this true meaning can be concluded if and only if the data comes from a Gaussian distribution.

In the literature, results for a normality test, such as the Anderson-Darling, the Shapiro-Wilk or similar tests, are not typically reported, therefore conclusions about the performance of the optimization methods may be questionable. An example when the data do not distribute normal will be shown to illustrate the importance of this issue. In that case, other statistical tools like Weibull analysis must be used to derive conclusions that are statistically significant.

4. PROBLEM DESCRIPTION

The problem to be addressed consists of finding the optimal placement (bus number) and power rating (MVA) of multiple STATCOM units in a 45 bus system, part of the Brazilian power network (Fig. 3) [18].

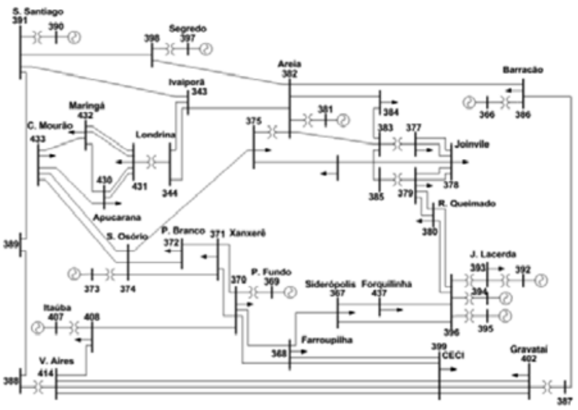

Fig. 3: 45 Bus section of the Brazilian power system

The main objective is to minimize the bus voltage deviations throughout the power system at minimum cost. The reasons for selecting the objective criteria and specific power system are: (i) the power system is not large; therefore an exhaustive manual search can be performed to find the global optimum, (ii) the problem has a reduced, scattered and non convex feasible region, and (iii) only a steady state criterion is considered to avoid possible discrepancies when transient analysis is also included [19].

4.1 Objective Function

Two goals are considered: (i) to minimize voltage deviations in the system and (ii) to minimize the cost. Thus, two metrics J_1 and J_2 are defined as in (1) and (3).

$$J_1 = \sqrt{\sum_{1}^{N} (V_k - 1)^2} \tag{1}$$

Where:

J_1 is the voltage deviation metric,
V_k is the p.u. value of the voltage at bus k, and
N is the total number of buses.

The total cost function, C_{total}, consists of two components: a fixed cost per unit that is installed in the system and a variable cost that is a linear function of each unit size:

$$C_{total}(M) = C_f \cdot M + C_v \cdot \sum_{p=1}^{M} \eta_p \quad (2)$$

Where:

M is the number of units to be allocated,
C_f is the fixed cost per unit,
C_v is the cost per MVA, and
η_p is the size in MVA of unit p.

Since $C_f >> C_v$, it is convenient to normalize each term of the cost function prior to its inclusion in the objective function (3).

$$J_2 = \frac{C_f \cdot M}{C_f \cdot M_{\max}} + \frac{C_v \cdot \sum_{p=1}^{M} \eta_p}{C_v \cdot M_{\max} \cdot \eta_{\max}} = \frac{M}{M_{\max}} + \frac{\sum_{p=1}^{M} \eta_p}{Max\_MVA} \quad (3)$$

Where:

J_2 is the cost metric,
M_{max} is the maximum number of STATCOM units to be allocated, and
η_{max} is the maximum size in MVA of each STATCOM unit.

The multi-objective optimization problem can now be defined using the weighted sum of both metrics J_1 and J_2 to create the overall objective function J shown in (4).

$$J = \omega_1 \cdot J_1 + \omega_2 \cdot J_2 \quad (4)$$

The weight for each metric is adjusted to reflect the relative importance of each goal. In this case, considering the maximum magnitudes of J_1 and J_2, it is decided to assign values of $\omega_1 = 1$ and $\omega_2 = 0.5$, such that both metrics have equal importance.

4.2 Decision Variables

The decision variables are the location of the STATCOM units and their sizes. These variables can be arranged in a vector as:

$$x_i = [\lambda_1 \quad \eta_1 \quad \ldots \quad \lambda_M \quad \eta_M] \quad (5)$$

Where:

λ_p, $p=1...M$, is the location (bus number) of STATCOM unit p.

All components of the decision vector are integer numbers, thus $x_i \in Z^{2M}$.

4.3 Constraints

There are several constraints in this problem regarding the characteristics of the power system and the desired voltage profile. Each constraint represents a limit in the search space, which in this particular case corresponds to:

- Generator buses are omitted from the search process since they have voltage regulators to regulate the voltage.
- Bus numbers are limited to {1, 2,…, N}.
- Only one unit can be connected at each bus.
- The number of units: 1 M 5.
- The size of each unit: 0 η_p 250 MVA.
- The desired voltage profile requires N additional restrictions defined as:

$$0.95 \leq V_k \leq 1.05 \ , \quad \forall k \in \{1,2,...,N\} \quad (6)$$

Each solution that does not satisfy the above constraints is considered infeasible.

5. OPTIMIZATION ALGORITHMS

For the optimal allocation of multiple FACTS units in a 45 bus system, six algorithms are fully developed and compared: Bender's decomposition, B&B, GA, PSO, Enhanced-PSO, and BFA.

5.1 Benders' Decomposition

This method separates two sets of decisions that are made into two consecutive stages. In the first stage of the decision making, some of the constraints are delayed to reduce the complexity of the original (master) problem. In the second stage, some of the parameters that influence the decision, whose values were originally uncertain, are known and fixed after the first decision vector is found. Thus the secondary problem is reduced in complexity and in the number of variables [20], [21].

In the case of the STATCOM allocation problem, the master problem considers the decision vector in (5) that can be naturally separated into one sub-vector for selecting optimal locations and another sub-vector for choosing optimal sizes.

The separation of the constraints can be stated as follows:

- First stage: sizes of the STATCOM units become delayed constraints, thus the reactive power limits for these devices are relaxed in the solution of the power flow. The voltage reference is set to 1.0 p.u. for each STATCOM controller. The objective function corresponds to the voltage deviation metric defined in (1).

- Second stage: with the locations of the devices determined, the set of constraints is limited to those related with the maximum size of each unit. The objective function includes the voltage deviation metric and the cost metric as in (4).

5.2 Branch and Bound

B&B is a classical approach to search for an optimal solution by evaluating only a subset of the total possible solutions. The main steps in the algorithm are [17], [20], [21]:

- Branching: the set of feasible solutions is partitioned into simpler subsets. At each iteration, one of the promising subsets is chosen and an effort is made to find the best feasible solution within it.
- Bounding: the algorithm proceeds to find upper and lower bounds for the optimal objective value. There is only one upper bound u at each stage, which corresponds to the lowest among the objective values of all the feasible solutions that have appeared so far.
- Pruning: if at certain a stage, one of the subsets has a lower bound which is greater than the current upper bound, then the algorithm prunes (discards) that set.

Branching, bounding and pruning are repeated until the optimal solution is found.

For this particular problem, the objective function is defined as in (4). The branching strategy corresponds to the "depth-first search": for each subset of feasible locations, branching is performed by dividing progressively the STATCOM size intervals into smaller sub-intervals. The bounding and pruning strategies help to narrow the search by discarding as many sub-intervals as possible until the optimal value, for the particular subset of feasible locations, is found. In the next stage another subset of feasible locations is chosen, and the process is repeated until the set of all feasible locations is covered.

5.3 Genetic Algorithm (GA)

GA is an ECT that patterns itself after Charles Darwin's "survival of the fittest" concept. Each chromosome represents a possible solution to the problem. Through selection of parents, crossover between members of the current population, and mutation of the offspring, the population evolves and, after a number of generations, it approaches an optimal solution [22], [23].

For this particular problem, the chromosomes are defined as the decision vector in (5) and the fitness of each chromosome is evaluated through the use of the objective function in (4).

After the fitness of the entire population has been assessed, a subgroup of chromosomes is selected to become the parents for the next generation. For this particular case, elitism and "roulette wheel" are used as the selection mechanisms. Once the two parents are chosen, crossover between them produces two offspring. For each offspring, there is a chance that any number of its genes may be mutated; the mutation probability applies to each gene independently resulting in anywhere from zero to all genes being mutated.

The previous generation is replaced by the new generation and the entire process is repeated until the maximum number of generations is reached.

The parameters used in this study are shown in Table I [24].

Table 1: GA Parameters

Parameter	Optimal value
Percentage of elite members	10%
Crossover probability	85%
Mutation probability	5%

5.4 Particle Swarm Optimization (PSO)

The PSO algorithm considers that each particle represents a potential solution to the problem, thus the particles are defined as the decision vector in (5). The quality of the solution, that allows the best position for each particle and the swarm to be determined, is assessed using the fitness function defined in (4).

At each iteration, t, the position of each particle is determined by [25], [26]:

$$\vec{x}_i(t) = \vec{x}_i(t-1) + \vec{v}_i(t) \qquad (7)$$

The velocity of each particle is determined by both the individual and group experiences:

$$\begin{aligned} v_i(t) = {} & w_i \cdot v_i(t-1) + \ldots \\ & c_1 \cdot r_1 \cdot (p_i - x_i(t-1)) + \ldots \\ & c_2 \cdot r_2 \cdot (p_g - x_i(t-1)) \end{aligned} \qquad (8)$$

Where:

w_I is a positive number between 0 and 1,
c_1 and c_2 are the cognitive and social acceleration constants respectively,
r_1 and r_2 are random numbers with uniform distribution in the range of [0, 1].
p_i is the individual best position found by the corresponding particle, and
p_g is the global best position found by the entire swarm.

To avoid the divergence of the swarm, a maximum velocity for each dimension of the problem hyperspace is defined (v_{max}).

Additionally, since integer variables are included in the optimization problem, the Integer-PSO version is used, where the particle's position is rounded off to the nearest integer [26].

The PSO parameters used in this study are presented in Table II [24].

Table 2: PSO Parameters

Parameter	Optimal value
Inertia constant (w_i)	Linear decrease (0.9 to 0.1)
Individual acceleration constant (c_1)	2.5
Social acceleration constant (c_2)	1.5
V_{max} for bus location	9
V_{max} for STATCOM size	50

5.5 Enhanced-PSO

For this particular application, the canonical PSO algorithm described in the previous section is enhanced to facilitate the search through the problem hyperspace [24].

The additional logic in each individual is defined by the following rules:

- If the corresponding particle's best position, *pbest,* and the swarm's best position, *gbest,* are both feasible solutions then the velocity update is performed according to (8).
- If the particle has not found a feasible solution yet, then it is better to rely on the social knowledge and the velocity update equation is replaced by:

$$v_i(t) = w_i \cdot v_i(t-1) + c \cdot rand \cdot (p_g - x_i(t-1)) \qquad (9)$$

Where:
c is a single acceleration constant: $c = c_1 + c_2$,
rand is a random number with uniform distribution in the range of [0, 1].

- If none of the particles have found a feasible solution (*gbest* and *pbest* values are both infeasible) then the velocity of each particle is updated using a random value of the maximum velocity as shown in (10).

$$v_i(t) = [r_1 \cdot v_{\max}(1) \quad r_2 \cdot v_{\max}(2) \quad r_3 \cdot v_{\max}(3) \quad r_4 \cdot v_{\max}(4)] \qquad (10)$$

Where:
r_h is a random number with uniform distribution in the range of [0, 1] and
$v_{max}(h)$ is the maximum velocity in the h^{th} dimension of the problem hyperspace.

5.6 Bacterial Foraging Algorithm (BFA)

BFA is based on the movement patterns of *E. coli* in the intestines. Each individual, in this case a bacterium, represents a possible solution to the problem as in (5).

The algorithm considers four successive steps [26], [28]:

- Chemotaxis: the bacteria move towards better nutrient concentrations. For the N_c chemotactic steps the direction of movement is given by:

$$\theta^i(j+1,k,l) = \theta(j,k,l) + C(i) \cdot \varphi(j) \qquad (11)$$

Where :
$C(i)$ is the step size,
j is the number of chemotactic step, k is the reproduction step, and l is the index for the elimination event.
$\varphi(j)$ is the unit length of random direction taken at each step.

The bacterium continues to move in the same direction (given that the fitness function value improves) and stops when the number of repetitions reaches a maximum of N_s.

- Swarming: All the bacteria have a cell-to-cell attraction via attractant and a cell-to-cell repulsion via repellant, with respect to other bacteria. Thus the movement of each bacterium towards better nutrient concentrations can be represented by:

$$J_o(i,j,k,l) + J_{cc}(\theta, P) \qquad (12)$$

Where:
$J_o(i,j,k,l)$ is the fitness function and
J_{cc} is the term that defines the attraction-repulsion to other bacteria [24].

The fitness function, J_o, corresponds to the objective function in (4) plus a penalty function defined as the number of buses in the system that violate the voltage profile constraint in (6).

- Reproduction: after chemotaxis, the population of bacteria is allowed to reproduce. S_r $(S_r=S/2)$ bacteria having the worst objective function value die and the remaining part split into two keeping the population size constant.

- Elimination-Dispersal: each bacterium is eliminated with a probability of p_{ed}.

The BFA parameters used in this study are [24]:

Table 3: BFA Parameters

Parameter	Optimal value
Number of bacteria (S)	20
Number of chemotactic cycles (N_c)	30
Number of swim steps (N_s)	3
Number of reproductions (N_{re})	3
Number of elimination-dispersal loops (N_{ed})	2
Probability of elimination (P_{ed})	0.5
Maximum distance ($C(i)$)	4
Attraction coefficients $d_{attract}$ and $w_{attract}$	0.1
Repellent coefficient d_{repel} and w_{repel}	0.05

6. SIMULATION RESULTS

6.1 Exhaustive search

An exhaustive search is performed on the problem of optimally allocating M STATCOMS to the power system in Fig. 3 by running a powerflow solution for each case in order to determine the global optimum. The solution indicates that the minimum number of devices needed to satisfy the constraints in Section IV-C is two and the computational effort corresponds to 37,196,250 power flows.

The feasible region of the problem is reduced, scattered and non convex. It is not possible to plot the entire feasible region since the dimensions are greater than three, however for illustrative purposes Fig. 4 shows the best scenario considering all possible bus locations and maximum STATCOM size of 250 MVA for each unit.

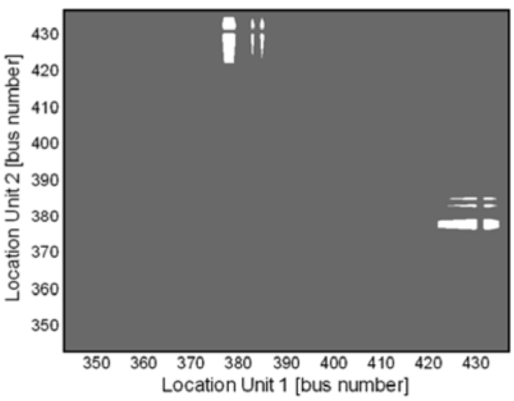

Fig. 4: Feasible region (white) vs. problem hyperspace

Fig. 5 shows the percentage of the feasible region with respect to the total number of cases [24].

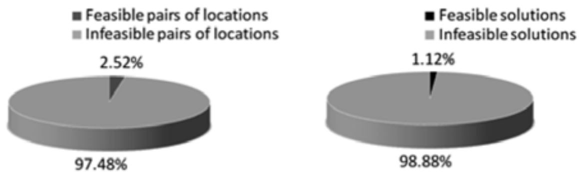

Fig. 5 (a): Percentage of feasible locations over total possible combinations. (b): Percentage of feasible solutions over total problem hyperspace

The global optimal solution is to place one STATCOM unit of 75 MVA at bus 378 and the second unit of 92 MVA at bus 433. The effect of the two STATCOM units is shown in Fig.6.

After the devices are optimally placed, all bus voltages are in the desired range of ±5% voltage deviation. Additionally, the voltage deviation metric J_1 improves by 26.5 % from an original value of 0.2482 to 0.1824.

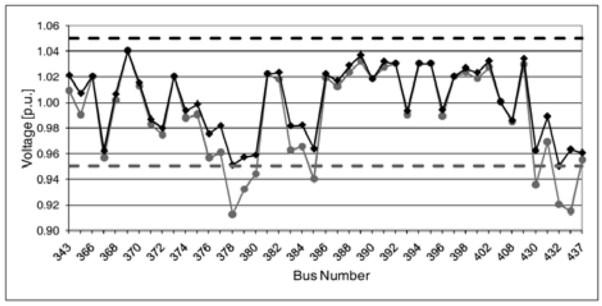

Fig. 6: Voltage profile without (-•-) and with STATCOM units (-♦-)

6.2 Metaheuristic approaches

Convergence into feasible regions: In order to evaluate the performance of the metaheuristic optimization algorithms (GA, PSO, Enhanced-PSO, and BFA), 150 trials are carried out for each algorithm. At each trial, the number of power flow evaluations (PF) is recorded until the first feasible solution is found. If no feasible solution is found, then the algorithm stops when the number of power flow evaluations reaches a maximum of 2000 [24]. Additionally, a performance indicator called Success Rate is calculated to determine the percentage of time that the algorithm is able to converge into feasible regions.

The Anderson-Darling normality test is performed to determine if the datasets for each algorithm are normally distributed. The results of this analysis show that with better than 99.5% certainty, the data are not normally distributed. Thus, other statistical distributions must be used.

The Weibull distribution is an appropriate alternative to analyze data of this type. This distribution is used extensively to study extreme-value data. In this case, the number of power flows to the *first* feasible solution [29] corresponds to such extreme-value data.

A two-parameter Weibull distribution is fitted to each dataset in the least-squares sense. In each case, the correlation is greater than 0.95, indicating that the choice of Weibull is suitable. Fig. 7 shows the resulting probability plots for each technique and Table IV shows the corresponding statistical parameters.

Table 4: Statistical values two-parameter Weibull Distribution

	GA	PSO	Enhanced PSO	BFA
Minimum PF	67	28	22	24
Maximum PF	>2000	>2000	379	1834
Success Rate	30	20.7	100	100
Scale (α)	4329	8650	147	326
Shape (β)	1.1	0.8	2.5	1.2

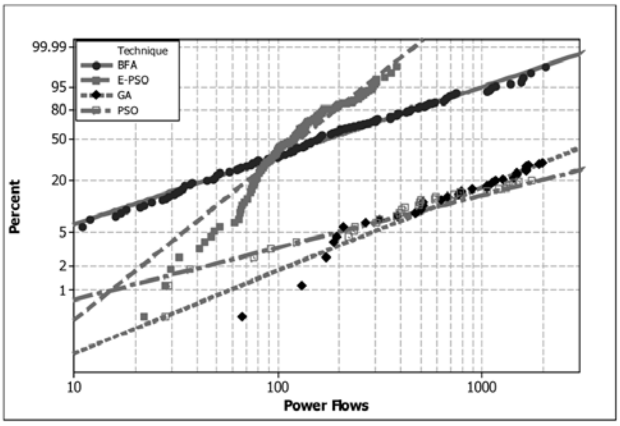

Fig. 7: Weibull plots of all algorithms

Table IV indicates that, based on the ranges for the number of power flow evaluations, the Enhanced-PSO is faster in finding feasible solutions compared to the other algorithms. Moreover its Success Rate is 100% versus 20.7% for canonical PSO and 30% in the case of GA.

Additionally, the Weibull parameters, α and β, carry important physical meanings. The scale parameter, α, corresponds to the characteristic time to find the first feasible solution. This is defined as the number of power flows needed to obtain a feasible solution in 63.2% of the trials. The shape parameter, β, represents the slope produced by data when plotted on a Weibull plot (Fig. 7). More interestingly, the shape parameter provides insight into how the algorithms are able to seek out feasible solutions:

- $\beta < 1$: Less likely to find feasible solutions as the number of power flows increases.
- $\beta = 1$: Same likelihood of finding feasible solution regardless of the number of power flows that are performed.
- $\beta > 1$: As the number of power flows increases so does the likelihood of locating a feasible solution.
-

Applying these concepts, the resulting characteristic time to find a feasible solution is 147 and 326 power flows for Enhanced-PSO and BFA, respectively. The canonical PSO and GA are only able to find feasible solutions in at most 30% of the trials while the rest of the values are censored. This leads to characteristic times of 4329 and 8650 for GA and PSO, respectively.

In addition, the Enhanced-PSO is the only algorithm with a shape parameter grater than one, which means that this algorithm offers the most efficient means of locating feasible regions.

The probability of obtaining a feasible solution in any number of power flows (or less) for each of the techniques, can also be read off from Fig. 7. Equally, the probability may be specified and then the maximum number of power flows required to find feasible regions may be read off.

Global Optimality: For further comparison of the performance of Enhanced-PSO and BFA algorithms, their capabilities for finding the global optimal solution are investigated. Thus, statistical values are calculated over a set of 50 trials, with 20 particles and 100 iterations, for each algorithm. In this case, the Anderson-Darling normality test gives *p*-values greater than 0.05, indicating that the data have a Normal distribution for both cases. Table V provides the additional indicators to evaluate the accuracy in finding the optimal solutions.

TABLE 5: Statistical analysis for optimal solutions

Parameter	Enhanced PSO	BFA
Minimum J value	0.91745	0.92441
Maximum J value	1.08390	1.36422
Average J value	0.98791	1.14765
Standard deviation J value	0.04167	0.09654

The accuracy in finding the optimal solution is higher in the case of the Enhanced-PSO algorithm with a standard deviation of 0.0417 as compared to 0.0965 of BFA, which is more than two times larger. In terms of the maximum and average values of the objective function value, they indicate a clear advantage of the Enhanced-PSO over BFA. Furthermore, the Enhanced-PSO algorithm finds the global optimum for this problem.

Figure 8 shows the degree of sub-optimality in the solutions provided by the Enhanced–PSO algorithm. It is possible to note that 70% of the time, the solutions found have 10% or less difference with respect to the global optimum. Additionally, the difference does not exceed the value of 20% in any of the trials.

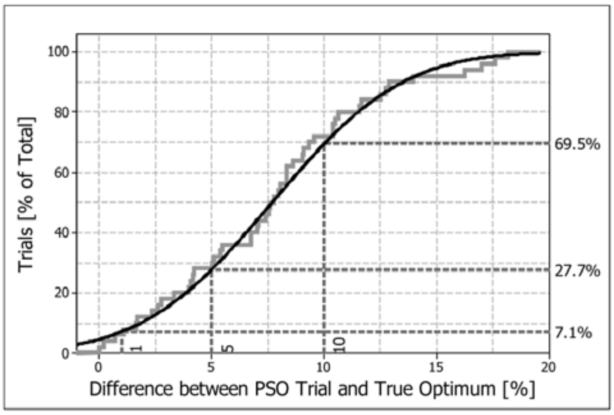

Fig. 8: Degree of sub-optimality for Enhanced-PSO algorithm

Scalability: As the previous section demonstrates, the evolutionary computation techniques are able to find solutions in a small fraction of the number of power flows required for an exhaustive search. One aspect of concern about their use is their capability to effectively solve optimization problems when the size of the power system is increased.

The Enhanced-PSO algorithm is applied to illustrate how the algorithm performs when the power system is changed from the Brazilian 45 bus system to the IEEE 118 bus network [30].

Figure 9 shows the capability of the algorithm to converge into feasible regions using boxplots (the box represents the middle 50% of data). The maximum number of iterations is 100 for both cases and the number of particles are 20 and 50 for the 45 bus and 118 bus system respectively. This figure provides evidence that the performance of the algorithm is not substantially affected by the size of the system. In both cases, feasible solutions are found in fewer than 17% of the maximum allowed power flow computations and the inter-quartile ranges (difference between the first and third quartile, that spans the middle 50% of the data) are fairly similar (4.5% and 5% for 45 bus and 118 bus systems respectively).

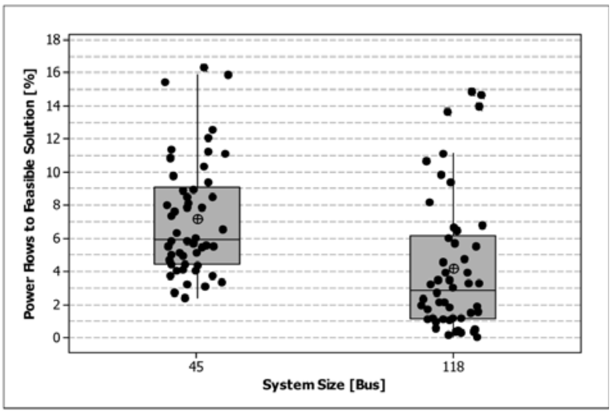

Fig. 9: Convergence into feasible regions

Due to the large number of power flow computations, an exhaustive search is not performed for this case. To give an idea of the computational time involved, for one power flow computation taking 125 msec, then the total time to run exhaustive search in the 45 bus system is 54 days, however for the 118 bus system the total time is 475 days (almost one year and four months). For this reason, the quality of the optimal solutions is assessed only by statistical analysis over 50 trials.

Table VI, shows the results for the 45 bus and 118 bus systems. The maximum, average and standard deviation values are presented as percentages with respect to the minimum objective function value.

Table 6: Optimal solutions – 45 and118 bus system

Parameter	45 Bus	118 Bus
Minimum *J* value	0.91745	0.8734
Maximum *J* value [%]	118.1	109.2
Average *J* value [%]	107.7	100.7
Standard deviation [%]	4.54	1.34

Comparing the percentages in both columns, it is possible to conclude that there are not significant differences in the performance when the size of the power system is increased.

6.3 Classical versus metaheuristic approaches

Table VII summarizes the overall performance data for the classical and best metaheuristic algorithms. The parameters considered for evaluating the performance of each method are the ability of the corresponding algorithm to find the global optimal solution and its computational effort.

Table 7: Algorithms' performance – 45 bus system

Parameter	Benders	B&B	Enhanced PSO
Bus 1, Size 1 (MVA)	(378, 75)	(378, 67)	(378, 75)
Bus 2, Size 2 (MVA)	(433, 92)	(430, 150)	(433, 92)
J value	0.9174	1.0170	0.9174
Voltage deviation (J_1)	0.1824	0.1819	0.1824
Time (sec)	18,611	846	666
Power flows	63,095	2,155	2,000

Considering the ability of the algorithms to find the global optimal solution, Enhanced-PSO and Bender's decomposition, are able to find the best solution. On the other hand, the B&B algorithm gets trapped in a local minimum.

Concerning the computational effort, the number of fitness function evaluations for Benders' decomposition is 31.5 times larger than the Enhanced-PSO, with the resultant increase in computational time. Nevertheless, both algorithms require only a fraction of the total computational effort required by the exhaustive search, 0.17% and 0.005% for Benders, decomposition and Enhanced-PSO, respectively.

7. CONCLUSIONS

This paper compares several optimization algorithms applied to the problem of optimal allocation of FACTs devices in the power system: classical approaches such as Bender's decomposition and Branch and Bound (B&B) algorithms, and metaheuristic techniques such as Genetic Algorithm (GA), Particle Swarm Optimization (PSO), Enhanced-PSO and Bacterial foraging Algorithm (BFA).

Emphasis is placed on aspects of the optimization process that tend to be overlooked in the literature:

Convergence into feasible regions: for this type of application the feasible region is reduced, scattered and non-convex, therefore special consideration has to be given to the exploratory capabilities of the optimization algorithms. Enhanced-PSO algorithm is introduced to show the importance of this aspect: with simple rules to enhance the initial exploration of the problem hyperspace, this algorithm is capable of finding feasible solutions in 100% of the cases and twice as fast as compared to its closest competitor, the BFA algorithm, and 30 times faster than the canonical PSO.

Statistical analysis of algorithm's performance: performance of metaheuristic techniques is mostly analyzed using parameters, such as average value and standard deviation, which assume the data to be normally distributed. This paper shows that there are cases when the normality assumption does not hold. Weibull analysis is presented as an example of how statistical tools correctly applied in those cases can lead to interesting conclusions about the underlying search mechanism of metaheuristic algorithms.

Global Optimality: until now there is no proof that metaheuristic algorithms provide global optimality. This paper analyzes this aspect using a simple but realistic case study of optimal STATCOM allocation considering steady state and economic criteria. An exhaustive search is carried out on a 45 bus system to find the global optimum of the problem, and then statistical results are obtained for different optimization algorithms. The algorithm with the best performance is capable of finding the global optimum at least once over 50 trials, and in at least 70% of the time, the degree of sub-optimality is less than 10%.

Scalability of metaheuristic algorithms: scalability is investigated using as an example the Enhanced-PSO algorithm. Results are obtained for the IEEE 118 bus system and compared with the 45 bus system. Considering both, the convergence to feasible solutions and the degree of sub-optimality of the optimal solutions found, there is evidence that indicates that the performance of the algorithm is not affected by the size of the system.

Classical versus metaheuristic approaches: the classical approaches, Bender's decomposition and B&B are compared with the Enhanced-PSO considering the capability of the algorithms in finding the global optimal solution and their computational effort. Bender's decomposition and Enhanced PSO are capable of finding the global optimum however B&B becomes trapped in a local optimal point. For all algorithms, the number of power flow computations is small (compared to the exhaustive search), but a comparison between them favors the metaheuristic algorithm since Bender's decomposition takes 30.5% more computational effort than the Enhanced –PSO.

A final concluding remark is that, at present, there is no optimization method that universally outperforms all others. The selection of an algorithm is problem dependent, and this paper makes a particular effort in showing different aspects that should be considered while choosing an optimization method for solving the problem of allocating FACTS devices.

ACKNOWLEDGEMENTS

This work has been partly supported by NSF Grants ECS #0524183 and ECCS #0348221.

REFERENCES

[1] D. Wanhong, T. Lie, "Optimal compensation of variable series capacitors for improved economic dispatch in power systems," Proc. of the IEEE International Conference on Energy Management and Power Delivery (EMPD) 1995,Vol. 2, Nov. 1995, pp. 732-737.

[2] F. Lima, F. Galiana, I. Kockar and, J. Munoz, "Phase shifter placement in large-scale systems via mixed integer linear programming," *IEEE Transactions on Power Systems*, Vol.18, no.3, Aug.2003, pp.1029-1034.

[3] A. Sharma, S. Chanana, S. Parida, "Combined optimal location of FACTS controllers and loadability enhancement in competitive electricity markets using MILP," Proc. of the IEEE Power Engineering Society General Meeting, 2005. Vol. 1, Jun. 2005, pp. 670–677.

[4] N. Yorino, E. El-Araby, H. Sasaki, S. Harada, "A new formulation for FACTS allocation for security enhancement against voltage collapse," *IEEE Transactions on Power Systems*, Vol. 18, no. 1, Feb. 2003, pp. 3-10.

[5] E. El-Araby, N. Yorino, H. Sasaki, "A comprehensive approach for FACTS devices optimal allocation to mitigate voltage collapse," Proc. of the IEEE Transmission and Distribution Conference and Exhibition 2002: Asia Pacific. Vol. 1, Oct. 2002, pp. 62-67.

[6] E. De Oliveira, J. Marangon Lima, K. De Almeida, "Allocation of FACTS devices in hydrothermal systems," *IEEE Transactions on Power Systems*, Vol. 15, no. 1, Feb. 2000, pp. 276-282.

[7] S. Gerbex, R. Cherkaoui, A. Germond, "Optimal location of FACTS devices to enhance power system security," Proc. of the IEEE Power Tech Conference Proceedings, 2003 Bologna, Vol. 3, Jun. 2003, pp. 7-14.

[8] S. Gerbex, R. Cherkaoui, A. Germond, "Optimal location of multi-type FACTS devices in a power system by means of genetic algorithms," *IEEE Transactions on Power Systems*, Vol. 16, no. 3, Aug. 2001, pp.537–544.

[9] L. Ippolito, P. Siano, "Selection of optimal number and location of thyristor-controlled phase shifters using genetic based algorithms," Proc. of the IEE Generation, Transmission and Distribution, Vol. 151, no. 5, Sept. 2004, pp. 630–637.

[10] D. Radu, Y. Besanger, "A multi-objective genetic algorithm approach to optimal allocation of multi-type FACTS devices for power systems security," Proc. of the IEEE Power Engineering Society General Meeting, 2006, Jun. 2006, pp. 8-16.

[11] H. Shaheen, G. Rashed, S. Cheng, "Optimal location and parameters setting of unified power flow controller based on evolutionary optimization techniques," Proc. of the IEEE Power Engineering Society General Meeting, 2007. Jun. 2007, pp. 1-8.

[12] M. Saravanan, S. Slochanal, P. Venkatesh, P. Abraham, "Application of PSO technique for optimal location of FACTS devices considering system loadability and cost of installation," Proc. of the IEEE 7th International Power Engineering Conference (IPEC) 2005. Vol. 2, Dec. 2005, pp. 716-721.

[13] H. Mori, Y. Maeda, "A hybrid method of EPSO and TS for FACTS optimal allocation in power systems," Proc. of the IEEE International Conference on Systems, Man and Cybernetics (ICSMC '06), 2006. Vol. 3, Oct. 2006, pp. 1831-1836.

[14] W. Ongsakul, P. Jirapong, "Optimal placement of multi-type FACTS devices by hybrid TS/SA approach," Proc. of the IEEE International Symposium on Circuits and Systems (ISCAS), Vol. 3, May 2003, pp. III-375 - III-378.

[15] J. Hao, L. Shi, Ch. Chen, "Optimising location of unified power flow controllers by means of improved evolutionary programming," Proc. of the IEE Generation, Transmission and Distribution, Vol. 151, no. 6, Nov. 2004, pp.705–712.

[16] W. Ongsakul, P. Jirapong, "Optimal allocation of FACTS devices to enhance total transfer capability using evolutionary programming," Proc. of the IEEE International Symposium on Circuits and Systems (ISCAS), Vol. 5, May 2005, pp. 4175-4178.

[17] D. Bertsimas and J. N. Tsitsiklis, *Introduction to linear optimization*, Athena Scientific Series in Optimization and Neural Computation, 1997.

[18] S Mohagheghi, Y. del Valle, G.K. Venayagamoorthy and R.G. Harley, "A Proportional-Integrator Type Adaptive Critic Design Based Neurocontroller for a Static Compensator in a Multimachine Power System". *IEEE Transactions on Industrial Electronics.* Vol. 54, No.1, February 2007, pp. 86-96.

[19] B. Kumar, S. Singh, S. Srivastava, "Placement of FACTS controllers using modal controllability indices to damp out power system oscillations," IET Generation, Transmission & Distribution, Vol. 1, no. 2, March 2007, pp. 209-217.

[20] G.L. Nemhauser and L.A. Wolsey, *Integer and combinatorial optimization*, Wiley-Interscience, New York, 1999.

[21] K. G. Murty, *Operations research: deterministic optimization models*, Prentice Hall, 1995.

[22] D. Goldberg, *Genetic algorithms in search, optimization and machine learning*. Addison-Wesley Pub. Co. 1989.

[23] L Schmitt, "Theory of genetic algorithms," Theoretical Computer Science (259), 2001, pp. 1-61.

[24] Y. del Valle, M. Digman, A. Gray, J. Perkel, G. Venayagamoorthy, and R. Harley, "Enhanced Particle Swarm Optimizer for Power System applications", Proc. of the IEEE Swarm Intelligence Symposium, 2008 (SIS 2008). St. Louis, Missouri, September 21-23, 2008.

[25] R. Eberhart, Y.Shi and J. Kennedy, *Swarm Intelligence.* San Francisco, CA: Morgan Kaufmann, 2001.

[26] Y. del Valle, G.K. Venayagamoorthy, S. Mohagheghi,, J. C. Hernandez, and R.G. Harley, "Particle Swarm Optimization: Basic Concepts, Variants and Applications in Power System". *IEEE Transactions on Evolutionary Computation*, Vol. 12, No. 2, April 2008, pp: 171-195.

[27] T.K. Das, G.K. Venayagamoorthy, "Bio-inspired Algorithms for the Design of Multiple Optimal Power System Stabilizers: SPPSO and BFA", Conference Record of the 2006 IEEE Industry Applications Conference, 41st IAS Annual Meeting. Vol. 2, 8-12 Oct. 2006, pp:635 – 641.

[28] K. M. Passino, "Biomimicry of bacterial foraging for distributed optimization and control," Control System Magazine, IEEE, vol. 22, Issue 3, pp. 52-67, June 2002.

[29] R. B. Abernethy, *The New Weibull Handbook*, North Palm Beach, FL, Robert B. Abernethy, 1993.

[30] The IEEE 118 Bus Test System. [Online]. Available: http://www.ee.washington.edu/research/pstca/.