An Evolutionary Algorithm Approach to Generate Distinct Sets of Non-Dominated Solutions for Wicked Problems

Marcio H. Giacomoni

Assistant Professor

Civil and Environmental Engineering

February 16th 2017

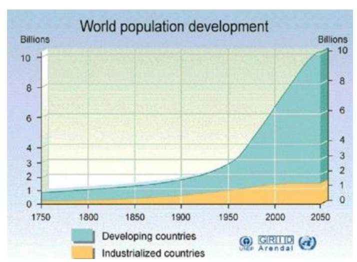
Zechman, E.M., M. Giacomoni, and M. Shafiee (2013) "An Evolutionary Algorithm Approach to Generate Distinct Sets of Non-Dominated Solutions for Wicked Problems" *Engineering Applications of Artificial Intelligence* 26(5), 1442-1457.

UTSA Water Resources Systems Analysis Lab:

PROBLEMS

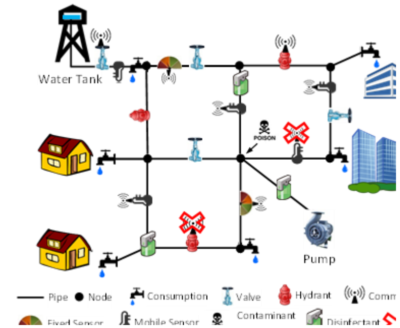
Sustainability of the Built and Natural Environments

Drought Management and Water Conservation



Stormwater Management and Green Infrastructure

Resilience and Security of Cyber-Physical Systems



TOOLS

Simulation Models

Hydrologic and Hydraulic
Water Networks
Land Use Change
Population Growth

Optimization Algorithms

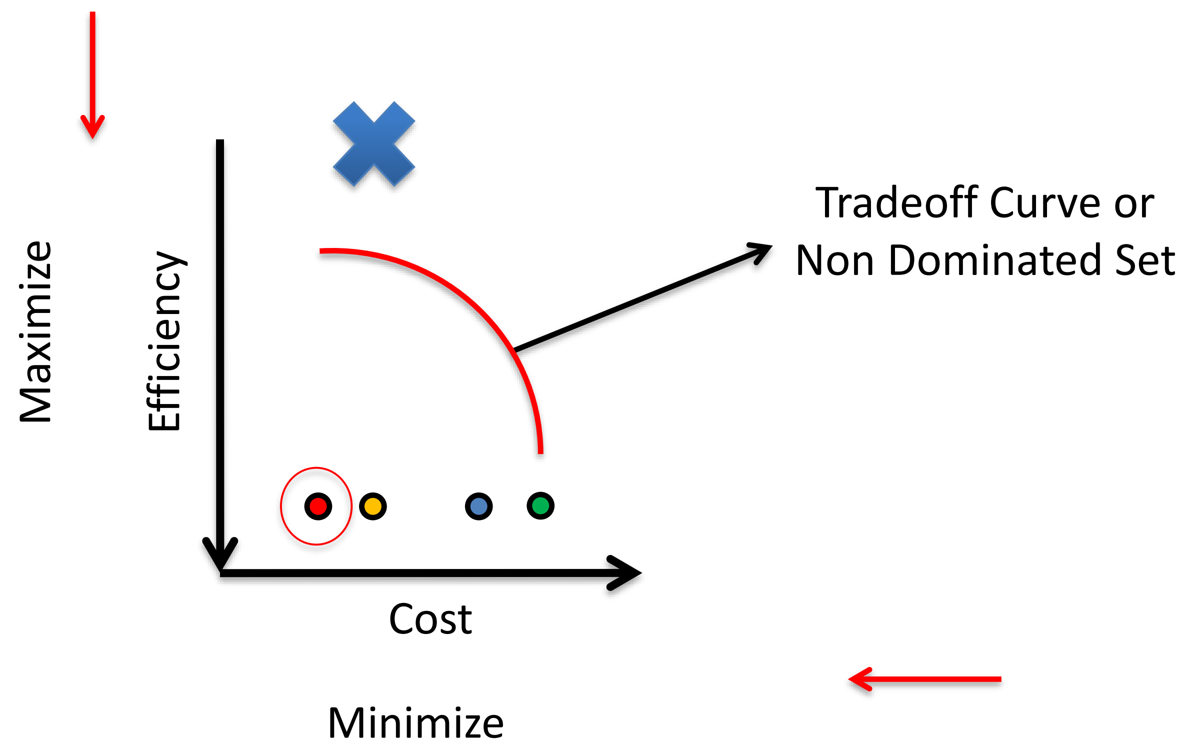
Evolutionary Computation
Single/multi-objective problems
Method for Generating Alternatives

Data Collection and Analysis

GIS and Remote Sensing
Monitoring

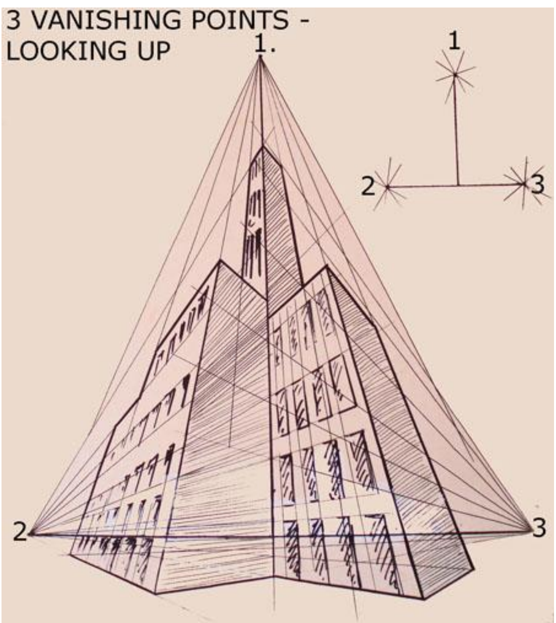
Engineering Problems

- Many engineering problems have multiple objectives:
  - Pareto front should be identified to represents the trade-off among conflicting objectives



Engineering Problems

- Real world problems are ill-posed:
  - Multiple perspectives (social, environmental, political)
  - Some objectives are difficult to model mathematically

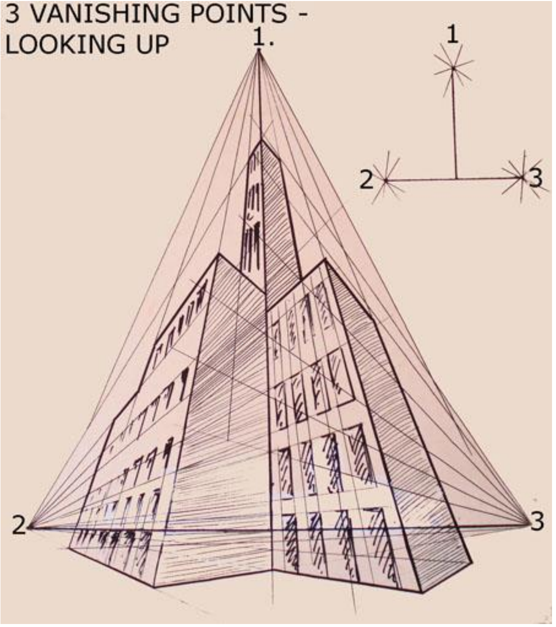


It's White and Gold!!!

No it's NOT!!! It's Blue and Black!!!

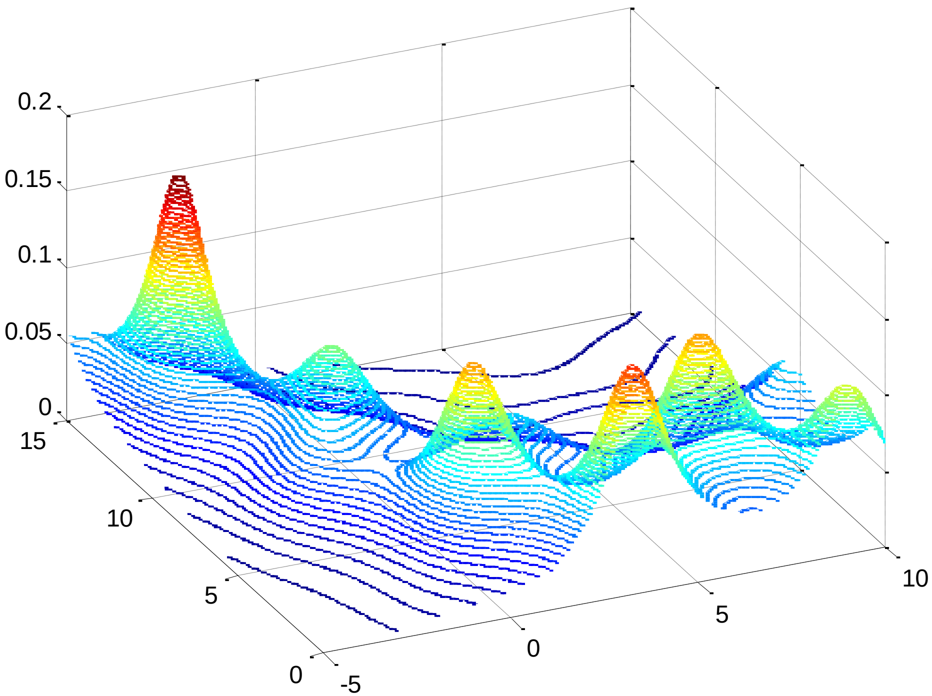
Engineering Problems

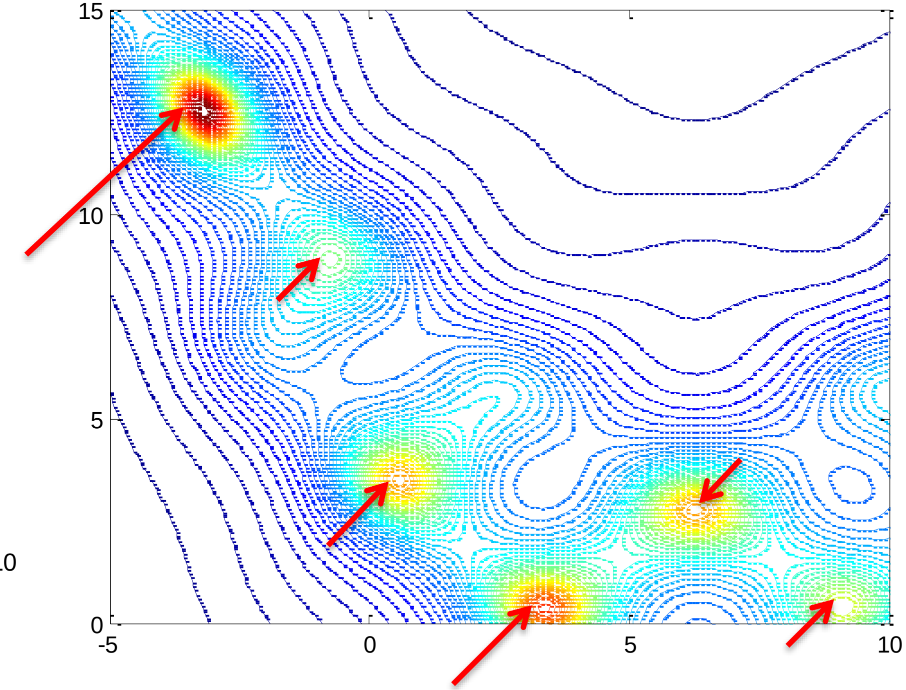
- Real world problems are ill-posed:
  - Multiple perspectives (social, environmental, political)
  - Some objectives are difficult to model mathematically



Engineering Problems

- The fitness landscapes for realistic problems, are often non-linear, complex, and multi-modal.





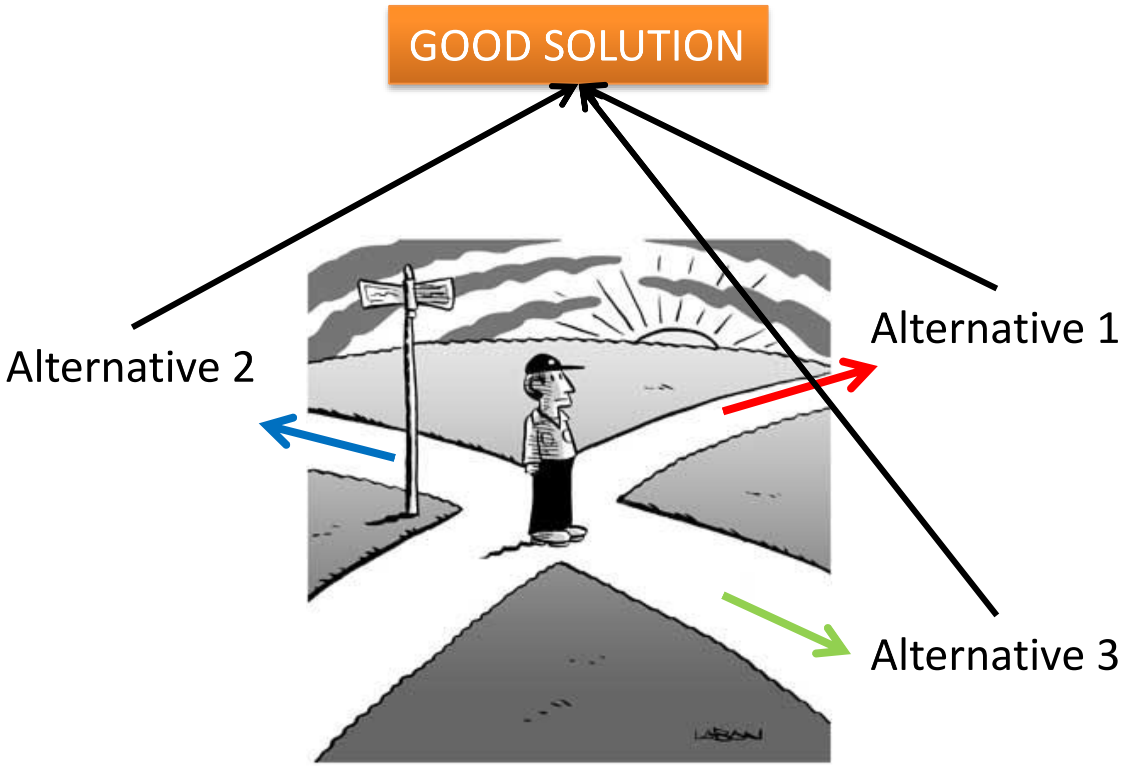
$$Maximize\ z = f(\boldsymbol{x}, \boldsymbol{y}) = \frac{1}{[a(y - bx^2 + cx - d)^2 + e(1 - f)\cos(x)\cos(y) + \log(x^2 + x^2 + 1) + e]}$$

$$x \times y \in [-5,10] \times [0,15] \qquad a = 1\,;\mathrm{b} = \frac{5.1}{4\pi^2}; c = \frac{5}{\pi}; d = 6\,; e = 10\,; f = \frac{1}{8\pi}$$

http://www.complexity.org.au/ci_louise/vol05/munteanu/munteanu.html

Modeling to Generate Alternatives

- Decision making can be aided through identification of alternative solutions



Modeling to Generate Alternatives

Original Problem:
Maximize $Z_k = f_k(X) \; \forall \; k = 1, \ldots, K$. (K – number of objectives)
Subject to $g_i(X) \leq b_i \quad \forall \; i = 1, \ldots, M$. (M – number of constraints).

Optimal Solution:
X^*, *with objective values of* Z_k^*

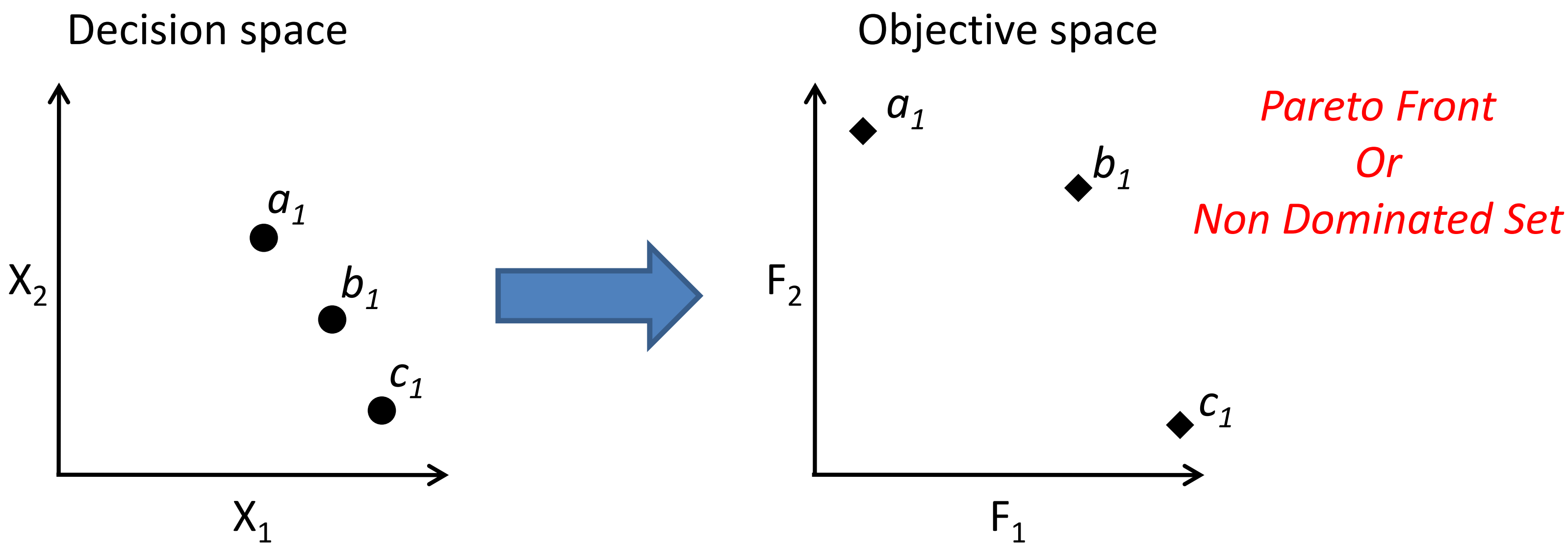
New Optimization Problem:
Maximize $D = \Sigma_j \mid x_j - x_j^* \mid$.
Subject to $g_i(X) \leq b_i \quad \forall \; i = 1, \ldots, M$.
$f_k(X) \geq T(Z_k^*)$.

Brill, E. D. Jr. (1979). The use of optimization models in public-sector planning. Management Science, 25(5), 413-422.

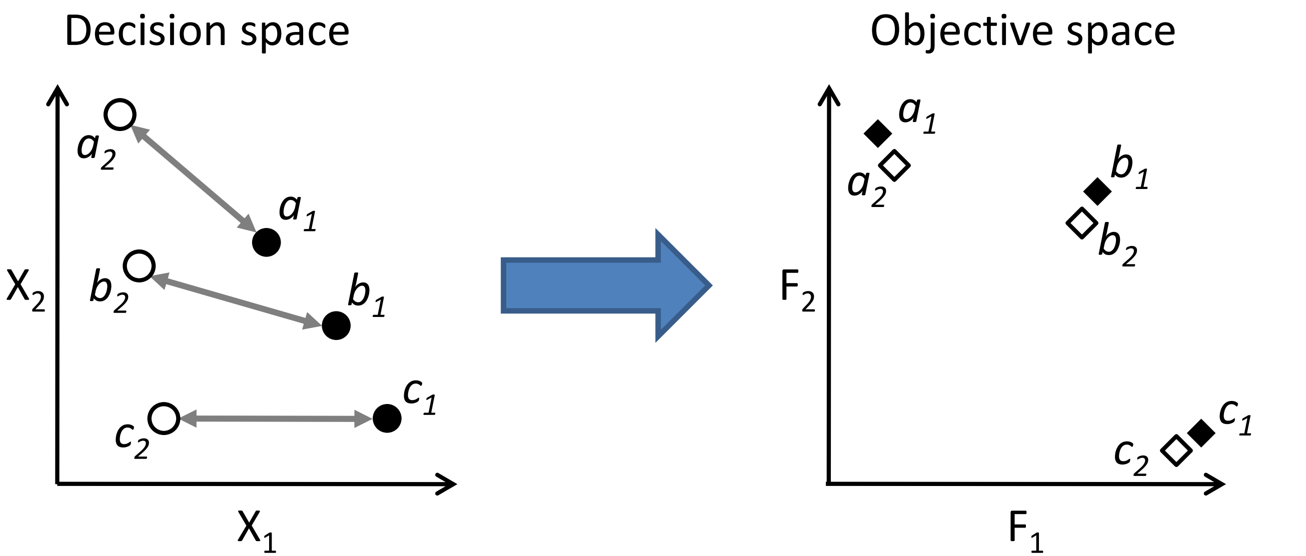
Research objective

- Develop an algorithm to identify a set of alternative Pareto fronts that are made up of solutions that map to similar regions of the objective space while mapping to maximally different regions of the decision space.

Multi-objective problems



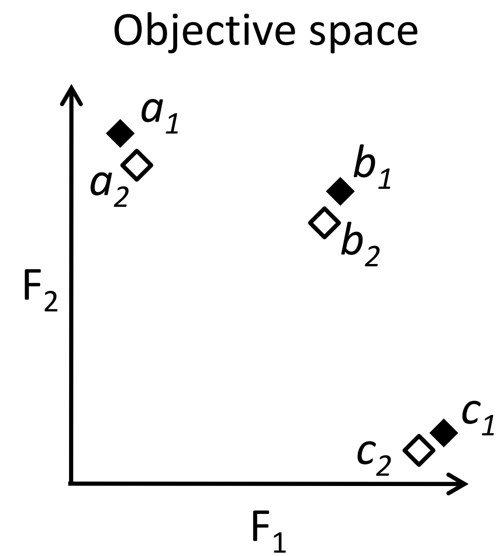
Multi-objective multi-modal problems



Two Sets of non-dominated solutions:
Pareto Front 1 - (a1, b1, and c1)
Pareto Front 2 - (a2, b2, and c2)

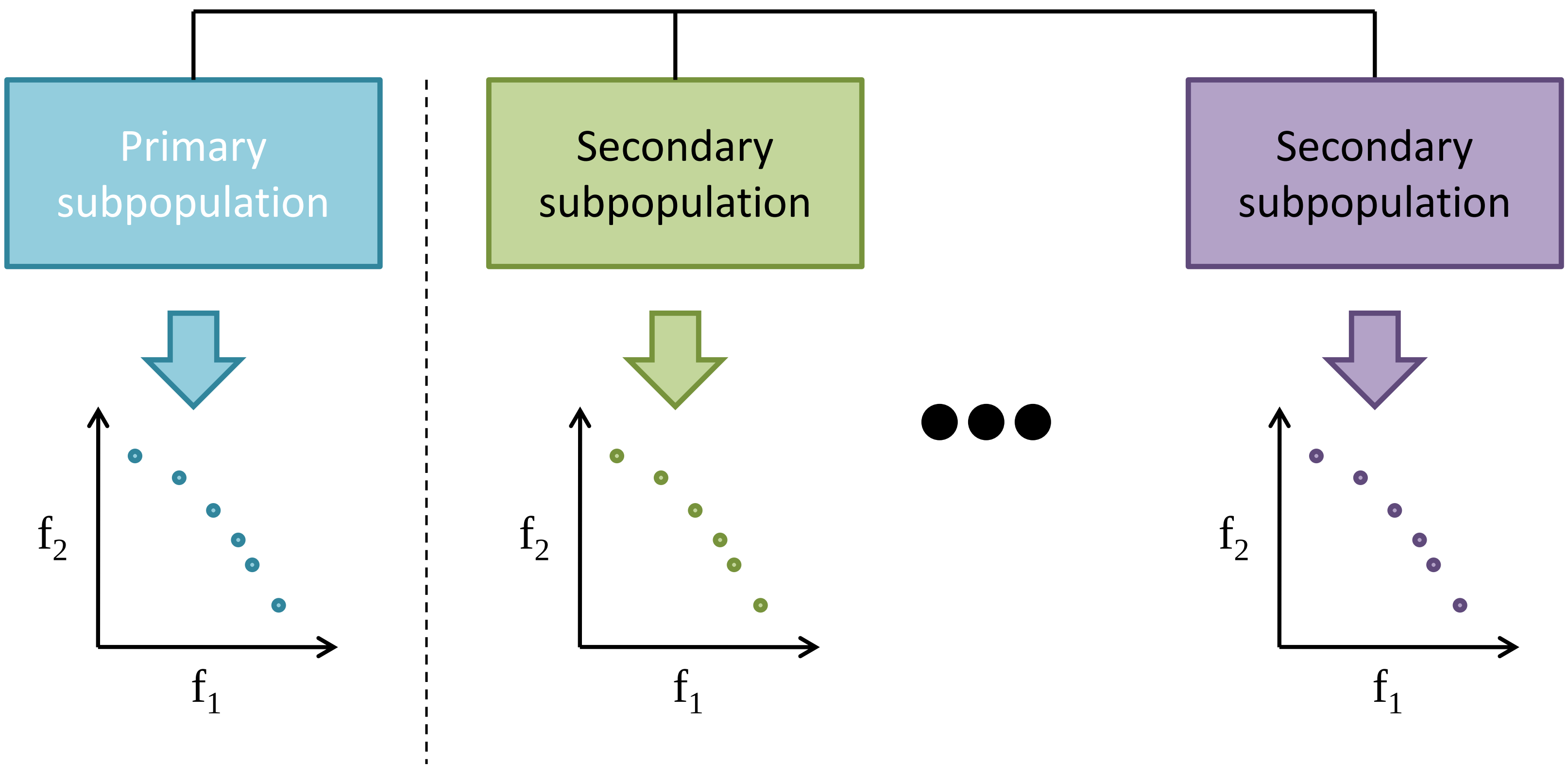
Multi-objective Evolutionary Algorithms to Generate Alternative Non-dominated Sets

- Use a set of populations to converge to alternative sets of non-dominated solutions
  - Each subpopulation will evolve one Pareto front that is different in decision space from other subpopulations
  - First subpopulation executes a conventional MOEA to find a typical Pareto front
  - Secondary subpopulations find alternative Pareto fronts



Multi-objective Niching Co-evolutionary Algorithm (MNCA)

- Multiple sub-populations co-evolve to distinct sets of non-dominated solutions



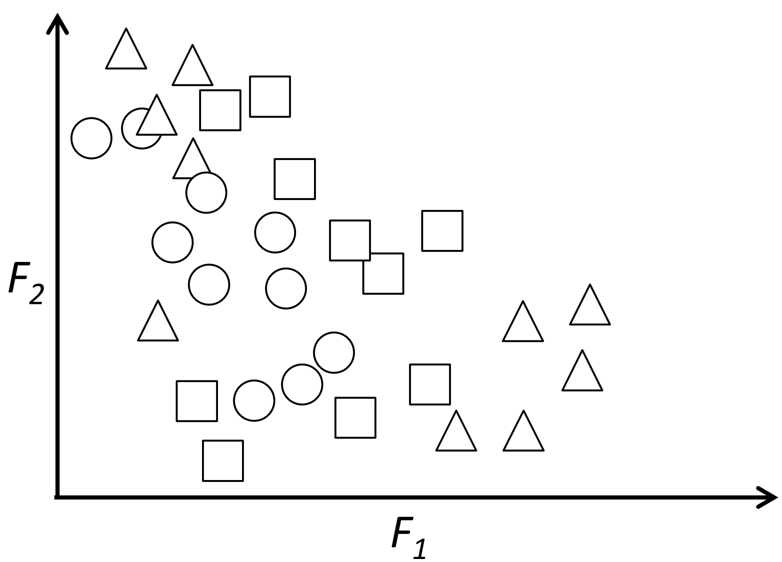
Zechman, E.M., M. Giacomoni, and E. Shafiee (2013) "An Evolutionary Algorithm Approach to Generate Distinct Sets of Non-Dominated Solutions for Wicked Problems" Engineering Applications of Artificial Intelligence 26(5), pp. 1442-1457

Multi-objective Niching Co-Evolutionary Algorithm (MNCA)

1. Group all solutions into clusters based on proximity in objective space
   - K-means clustering
2. Distance calculation
   - The distance of one solution is calculated in decision space to solutions that fall in the same cluster but in different subpopulations
3. Target Front
   - A target front is created, based on the first front of non-dominated solutions from first subpopulation
4. Feasibility Assignment
   - Label solutions in secondary subpopulations as feasible if they dominate any point in the target front
5. Selection:
   - First Subpopulation: NSGA-II operator
   - Secondary Subpopulations: Crowding Distance and Binary Tournament
     - Infeasible: rank and NSGA-II Crowding distance
     - Feasible: Crowding distance using four solutions

MNCA Algorithmic steps

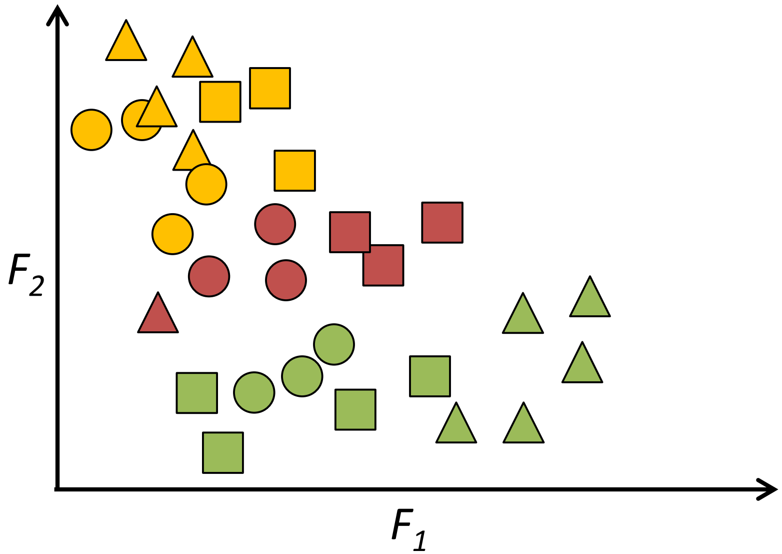
1. Group all solutions into clusters based on proximity in objective space
   – K-means clustering



Objective space

MNCA Algorithmic steps

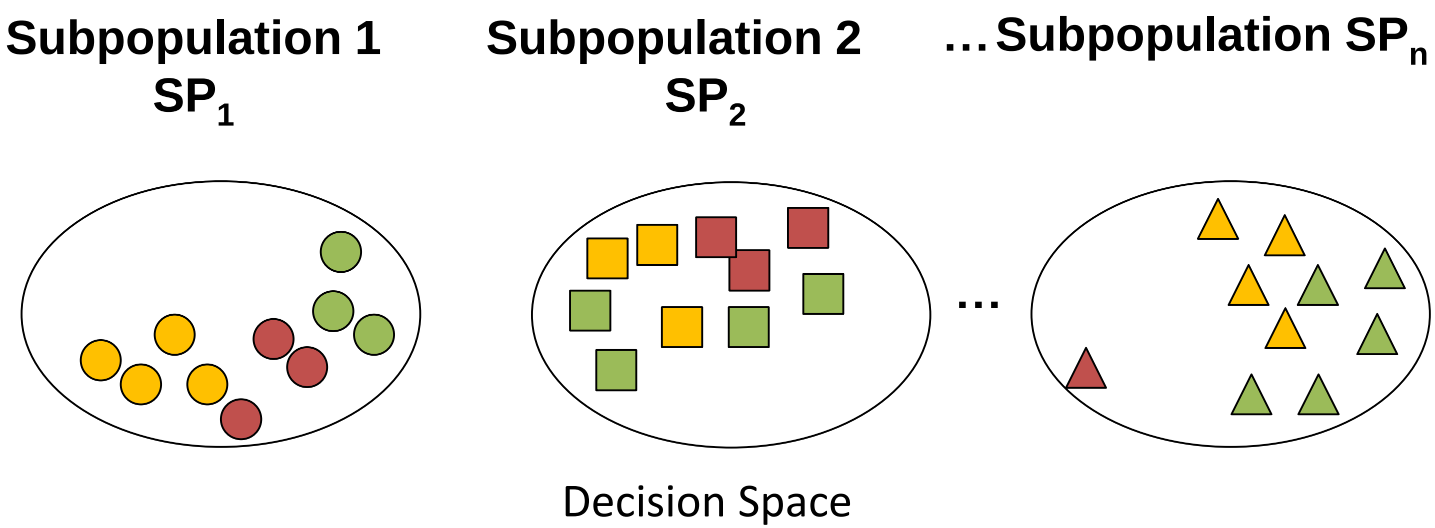
1. Group all solutions into clusters based on proximity in objective space
   - K-means clustering



Objective space

MCA Algorithmic steps

1. Group all solutions into clusters based on proximity in objective space
   - K-means clustering

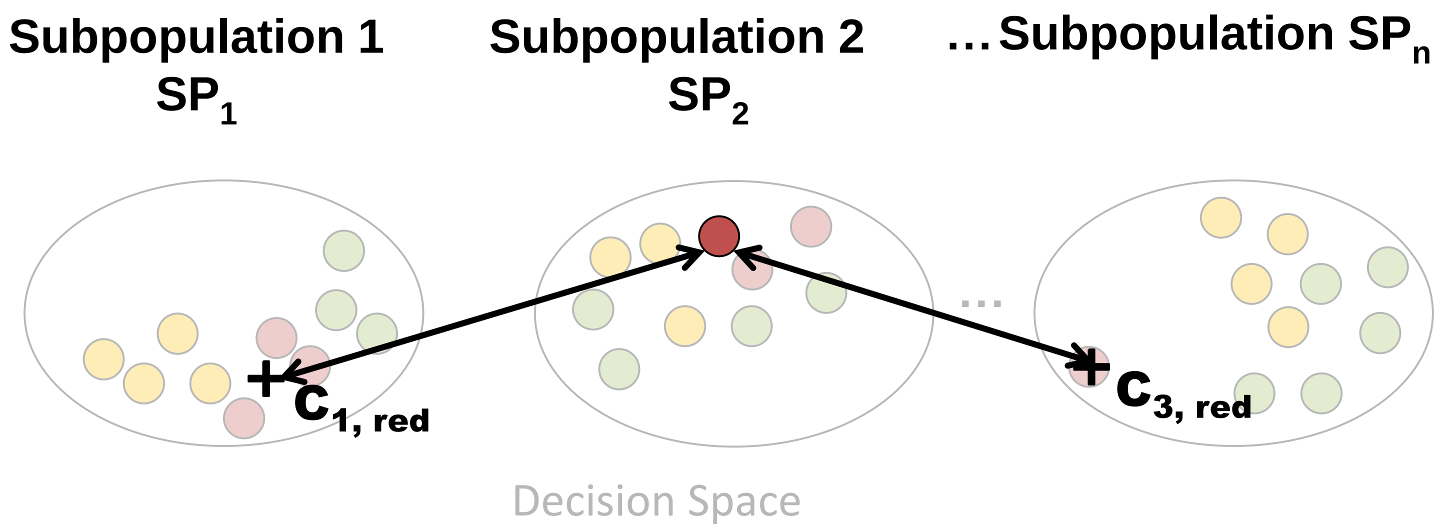


Colors represent different clusters that are formed based on similarities in objective space

MCA Algorithmic steps

1. Group all solutions into clusters based on proximity in objective space
   – K-means clustering
2. Distance calculation
   – The distance of one solution is calculated in decision space to solutions that fall in the same cluster but in different subpopulations

Distance calculation



Distance is calculated for each solution to centroid of same cluster in other subpopulations

Solution $Red_{2,i}$ Distance = minimum(Distance to $C_{1,\ red}$; Distance to $C_{3,\ red}$)

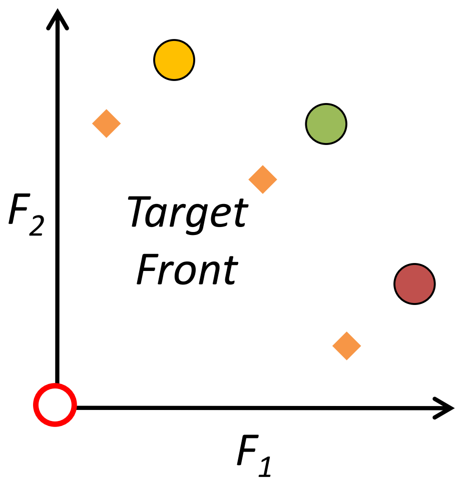
MNCA Algorithmic steps

1. Group all solutions into clusters based on proximity in objective space
   - K-means clustering
2. Distance calculation
   - The distance of one solution is calculated in decision space to solutions that fall in the same cluster but in different subpopulations
3. Target Front
   - A target front is created, based on the first front of non-dominated solutions from first subpopulation

Target front

$$Z_i' = T(Zi - WP_i) + WP_i$$

Z_i' is the a point on the target front
Z_i is the value of the i^{th} objective
T is the target reduction (i.e., 80%)
WP_i is the worst point for the i^{th} objective



Objective space

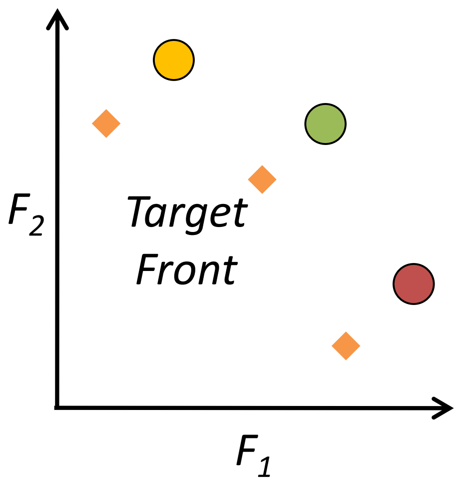
MNCA Algorithmic steps

1. Group all solutions into clusters based on proximity in objective space
   - K-means clustering
2. Distance calculation
   - The distance of one solution is calculated in decision space to solutions that fall in the same cluster but in different subpopulations
3. Target Front
   - A target front is created, based on the first front of non-dominated solutions from first subpopulation
4. Feasibility Assignment
   - Label solutions in secondary subpopulations as feasible if they dominate any point in the target front

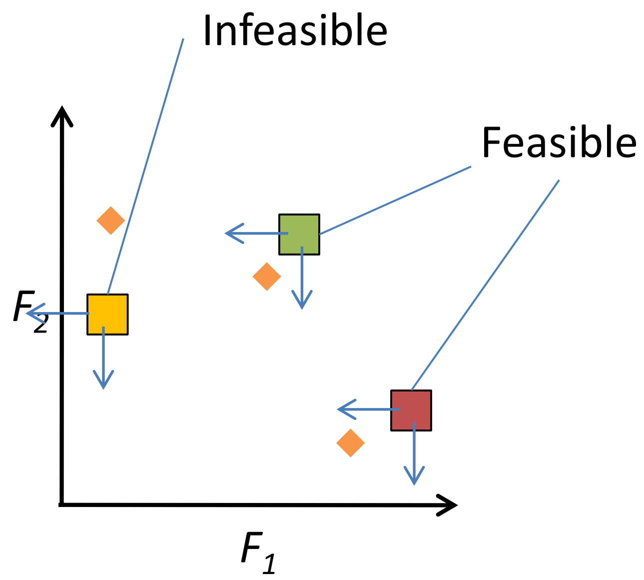
Target front

$$Z_i' = T(Zi - WP_i) + WP_i$$

Z_i' is the a point on the target front
Z_i is the value of the i^{th} objective
T is the target reduction (i.e., 80%)
WP_i is the worst point for the i^{th} objective



Objective space



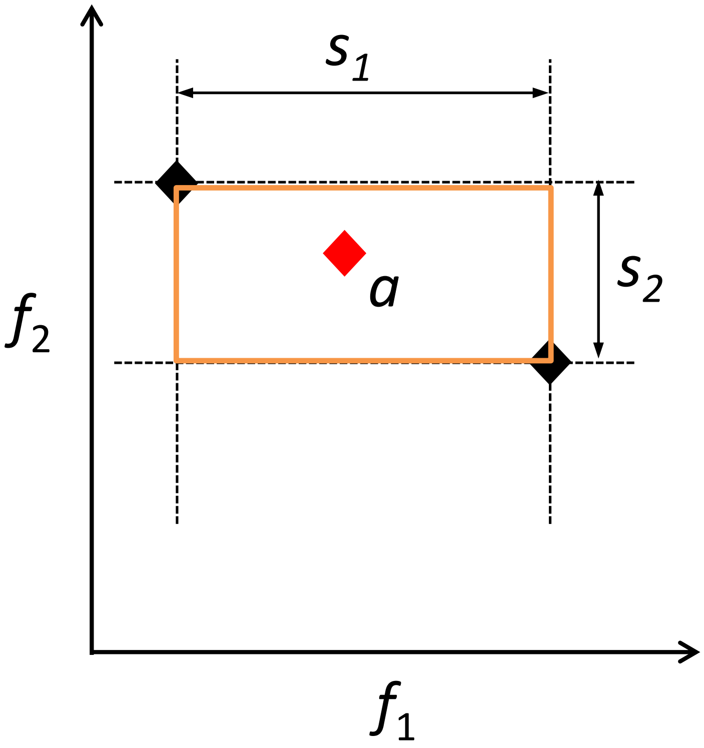
Objective space

MNCA Algorithmic steps

1. Group all solutions into clusters based on proximity in objective space
   – K-means clustering
2. Distance calculation
   – The distance of one solution is calculated in decision space to solutions that fall in the same cluster but in different subpopulations
3. Target Front
   – A target front is created, based on the first front of non-dominated solutions from first subpopulation
4. Feasibility Assignment
   – Label solutions in secondary subpopulations as feasible if they dominate any point in the target front
5. Selection:
   – First Subpopulation: NSGA-II operator
   – Secondary Subpopulations: Crowding Distance and Binary Tournament
     - Infeasible: rank and NSGA-II Crowding distance
     - Feasible: Crowding distance using four solutions

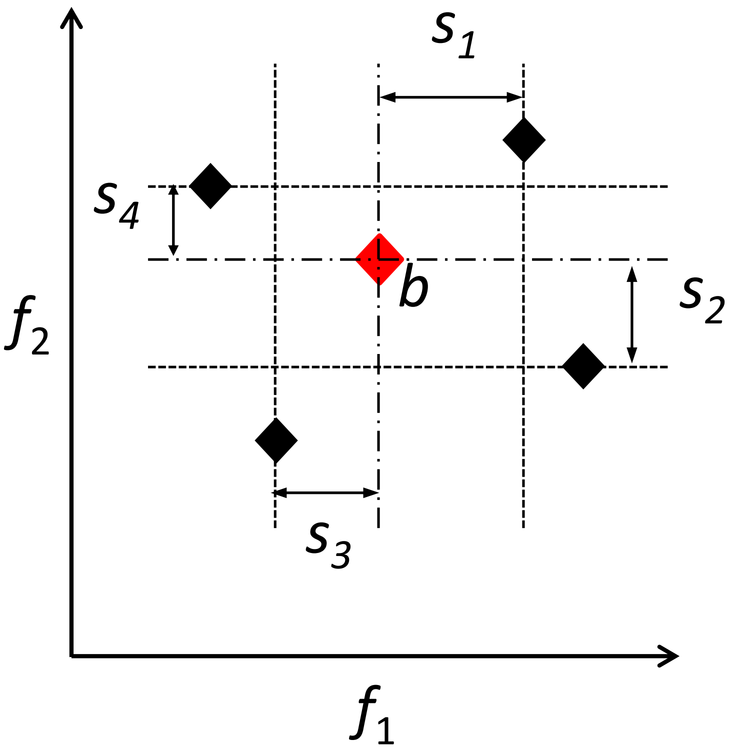
Crowding Distance

Infeasible Solutions



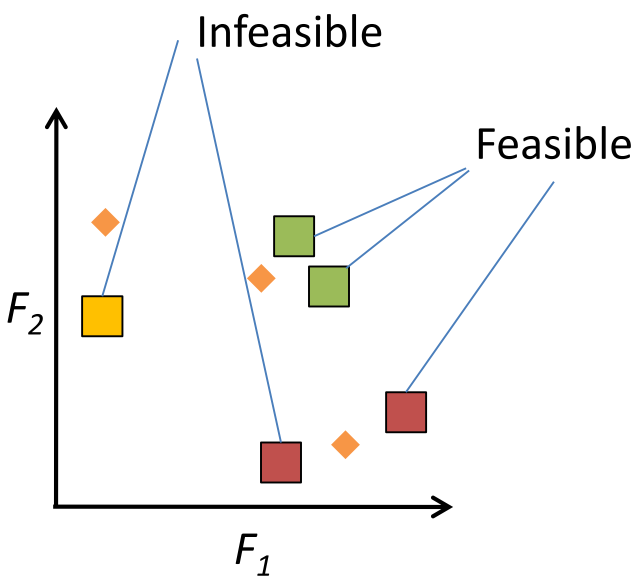
$$cd_b = \sum_{i=1}^{2} s_i$$ (NSGA-II)

Feasible Solutions



$$cd_a = \sum_{i=1}^{4} s_i$$

Binary Tournament

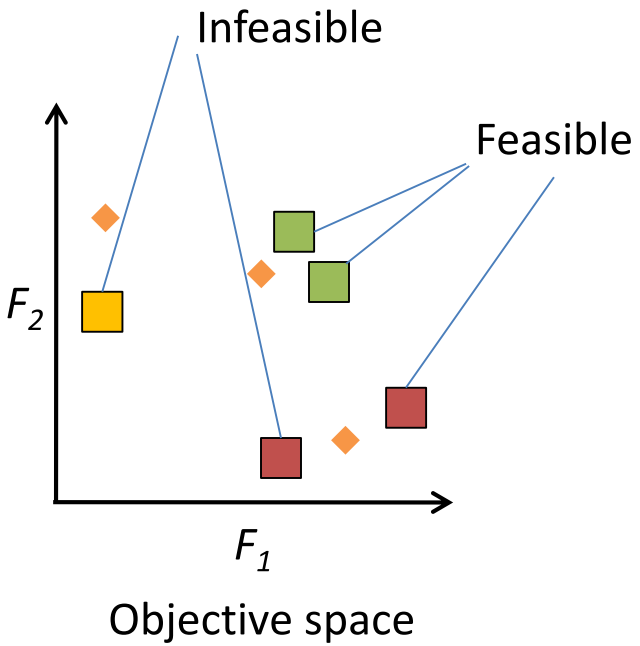


Objective space

Infeasible X Feasible

Feasible WINS!!!

Binary Tournament



Feasible Feasible

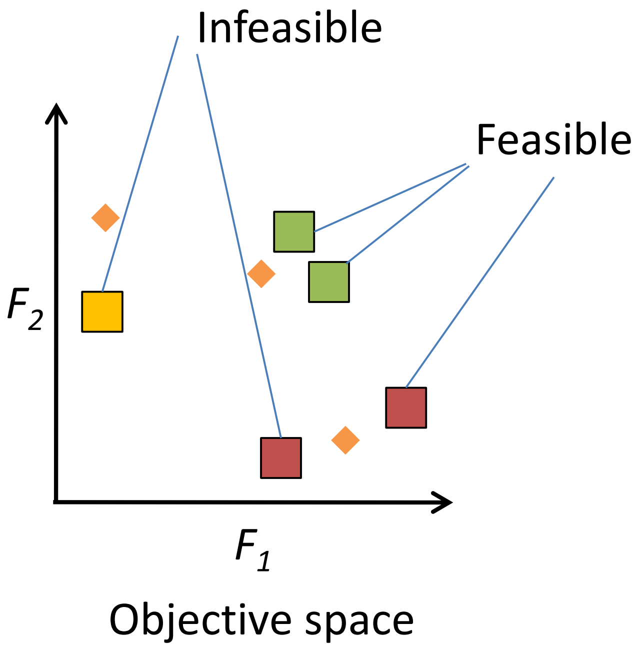
Same Cluster

X

Higher Distance WINS!!!

Solutions within the same region of objective space:
Maximize Distance increasing diversity.

Binary Tournament



Feasible Feasible

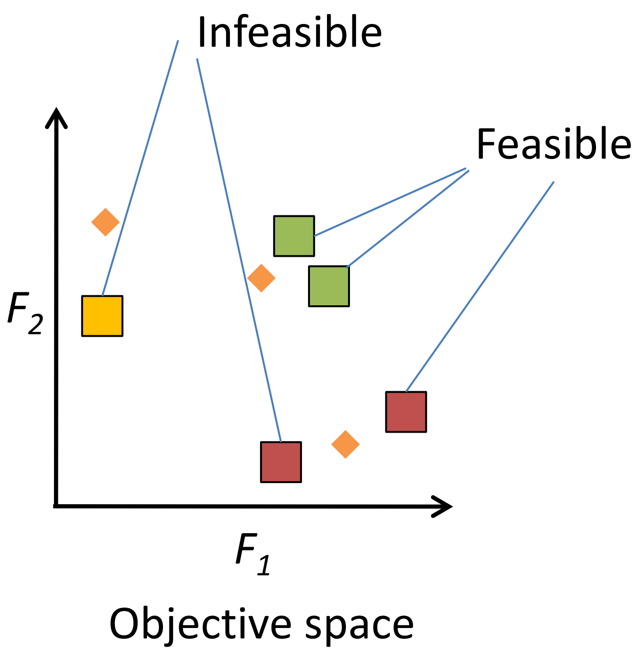
Different Clusters

X

Higher Modified Crowding Distance WINS!!!

Solutions are from different parts of the objective space: pressure to improve the coverage across the Pareto front.

Binary Tournament



Infeasible X Infeasible

Lower Front Rank or Crowding Distance WINS!!!

Niching-CMA

- Covariance Matrix Adaptation Evolution Strategy Niching Technique.
- Group solutions in niches based on their proximity in both decision and objective space.
- Outperformed other multi-objective and diversity-enhancing methodologies.

O. Shir and T. Beck, "Niching with derandomized evolution strategies in artificial and real-world landscapes," Natural Computing, vol. 8, pp. 171–196, 2009, 10.1007/s11047-007-9065-5.

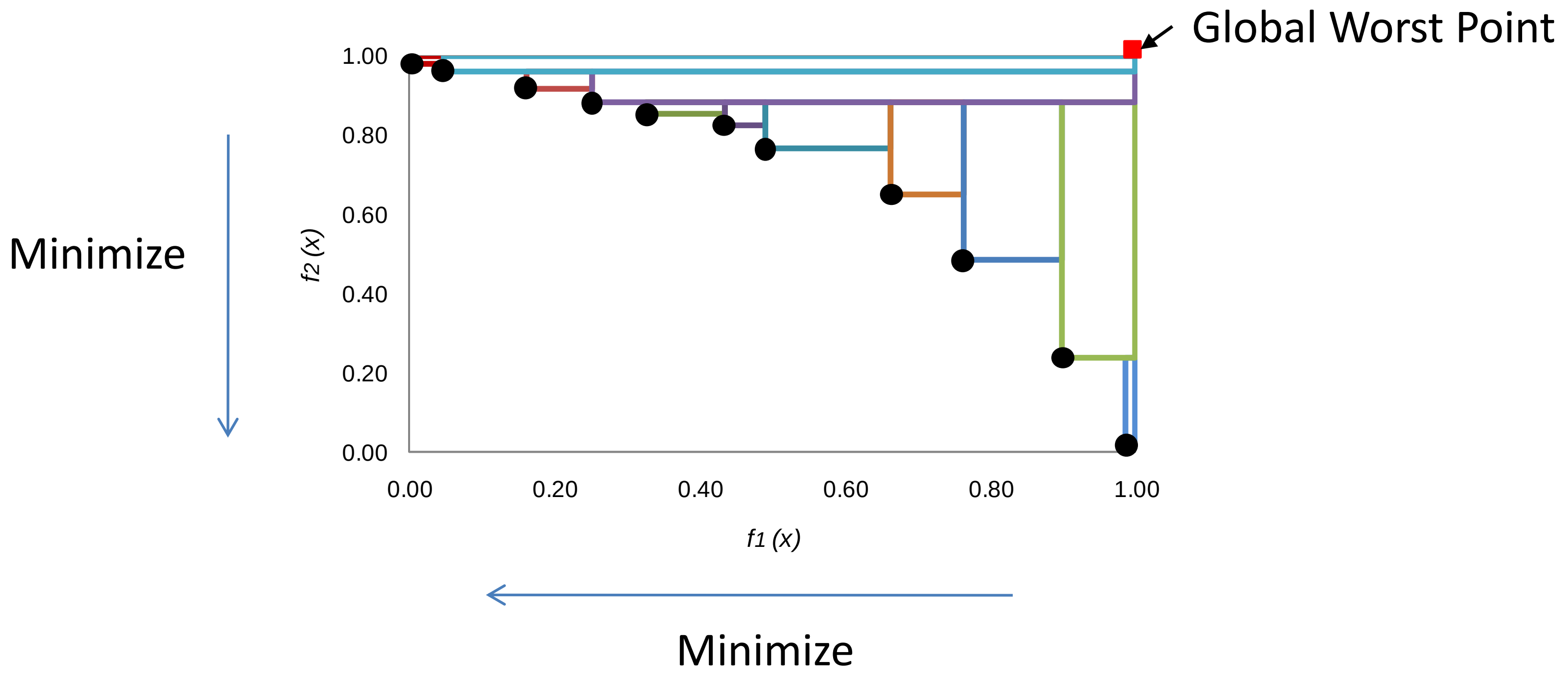
Algorithmic Settings

Parameter	Setting
Population Size	50
Subpopulations	2
Generations	1000
Mutation	1%
Clusters	5
Target	90% (95%)[1]

[1] 95% was used for the function Two-on-One

Hypervolume

- Represents the size of the space covered by the non-dominated set
- Used as an indicator to measure the quality of a non-dominated set



Decision Space Diversity

- Assess diversity in decision space based on the distance between all pair of individuals in a population
- Average of the distances between pairs of solutions and is normalized by the diameter of the decision space

$$diversity = \frac{1}{R \times pop(pop-1)} \sum_{i=1}^{pop} \sum_{j=i+1}^{pop} d(X_i, X_j) \quad (7)$$

where $d(X_i, X_j)$ is the Euclidean distance in decision space between the ith and jth solutions in a subpopulation. The equation for R is given as

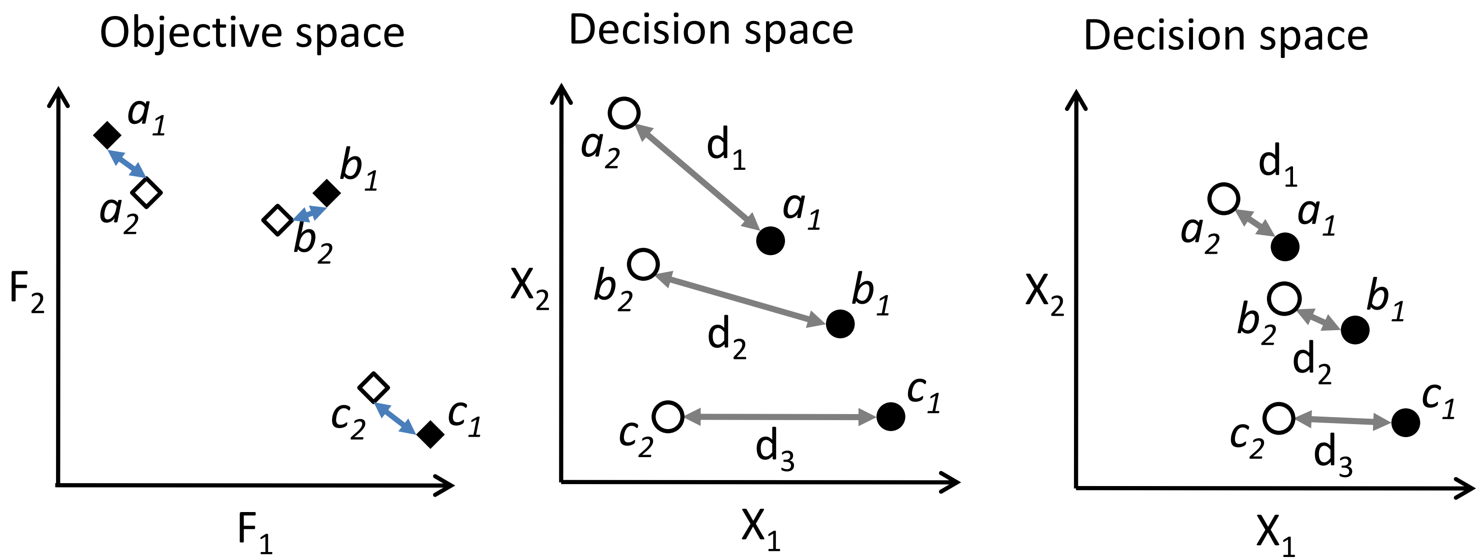
$$R = \sqrt{\sum_{i=1}^{N} (x_{i,max} - x_{i,min})^2} \quad (8)$$

O. Shir and T. Beck, "Niching with derandomized evolution strategies in artificial and real-world landscapes," Natural Computing, vol. 8, pp. 171–196, 2009, 10.1007/s11047-007-9065-5.

Paired Solution Diversity

- New metric to assess set of solutions that are distant in decision space though similar in objective space.
- Pair solution of one subpopulation with solution of another subpopulation that is nearest in objective space, and for each pair, calculating the Euclidean distance in the decision space.

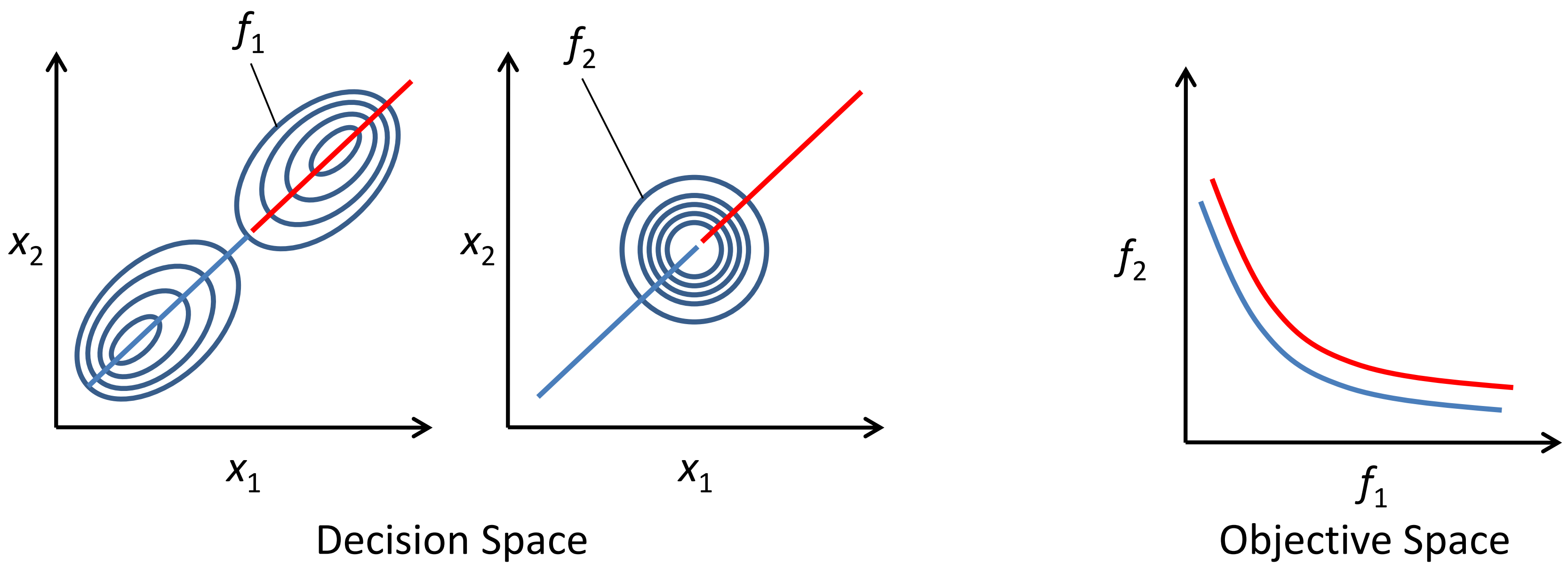
Paired Solution Diversity



$$diversity_{ps} = \frac{d_1 + d_2 + d_3}{3}$$

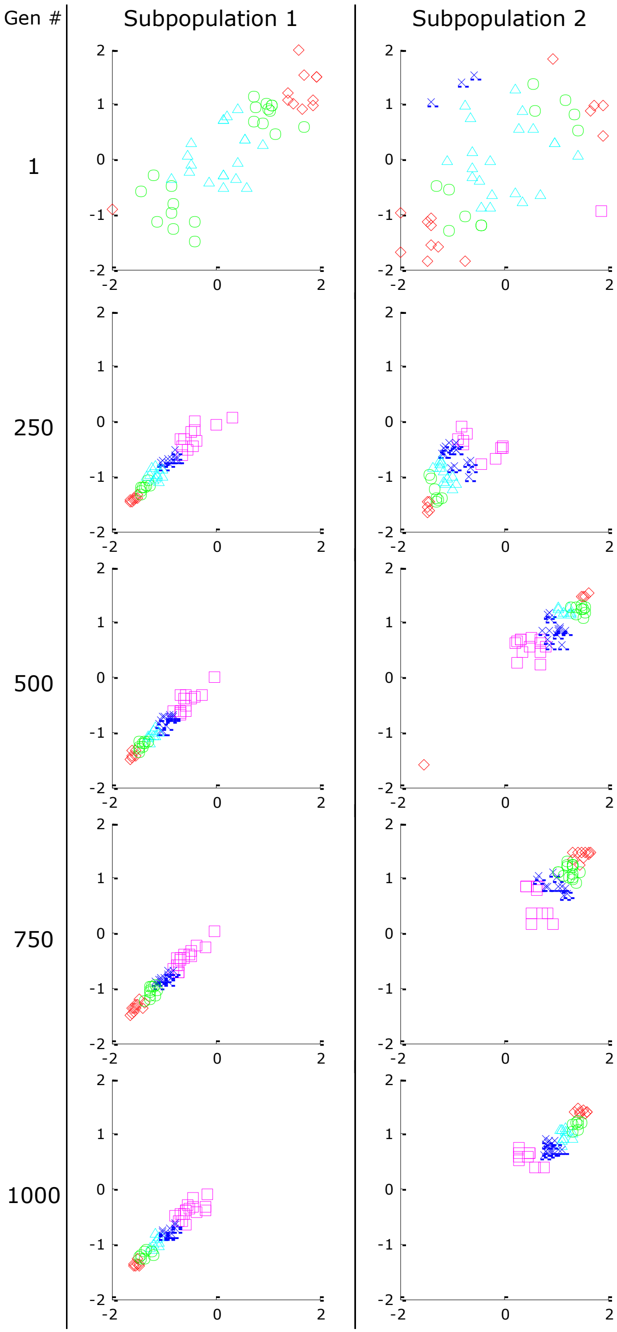
Test Function Two-on-One

- Two-on-One is a bi-modal problem composed of a fourth degree polynomial with two optima and a second-degree sphere function.

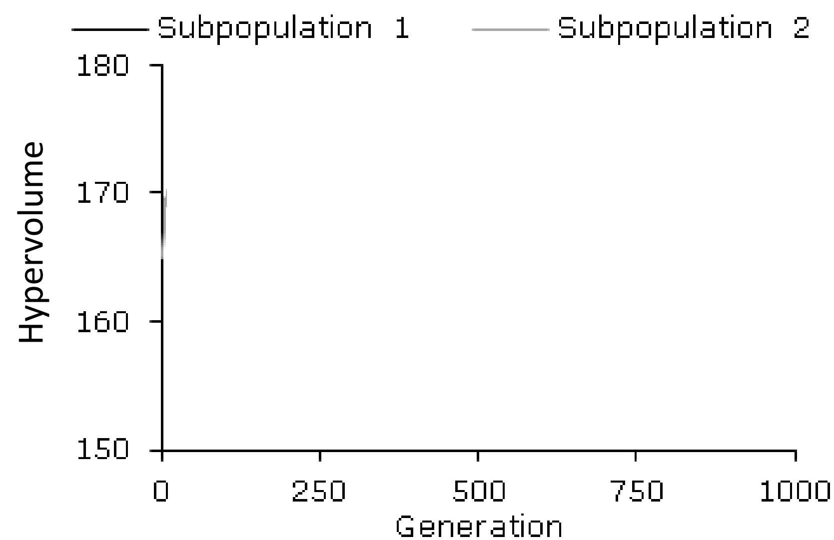


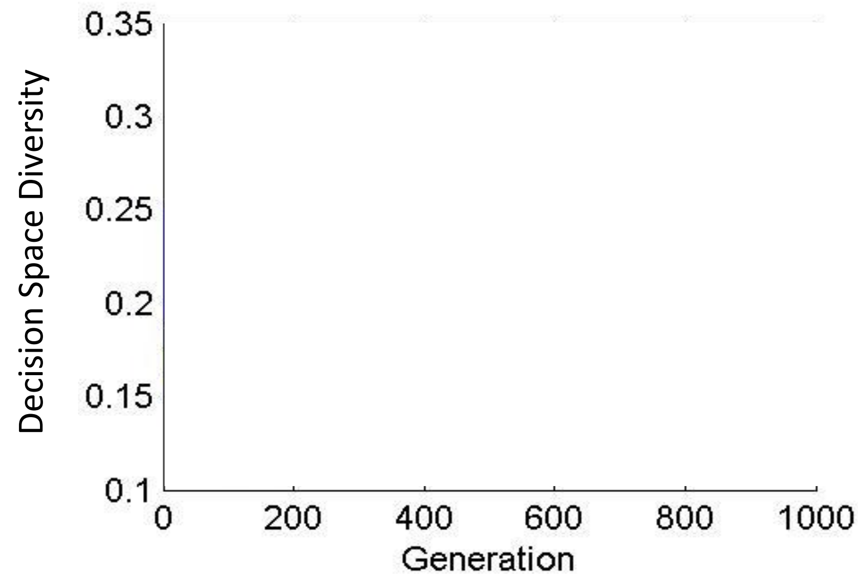
M. Preuss, B. Naujoks, and G. Rudolph, "Pareto Set and EMOA Behavior for Simple Multimodal Multiobjective Functions," in Parallel Problem Solving from Nature - PPSN IX, 2006, pp. 513–522.

Decision Space

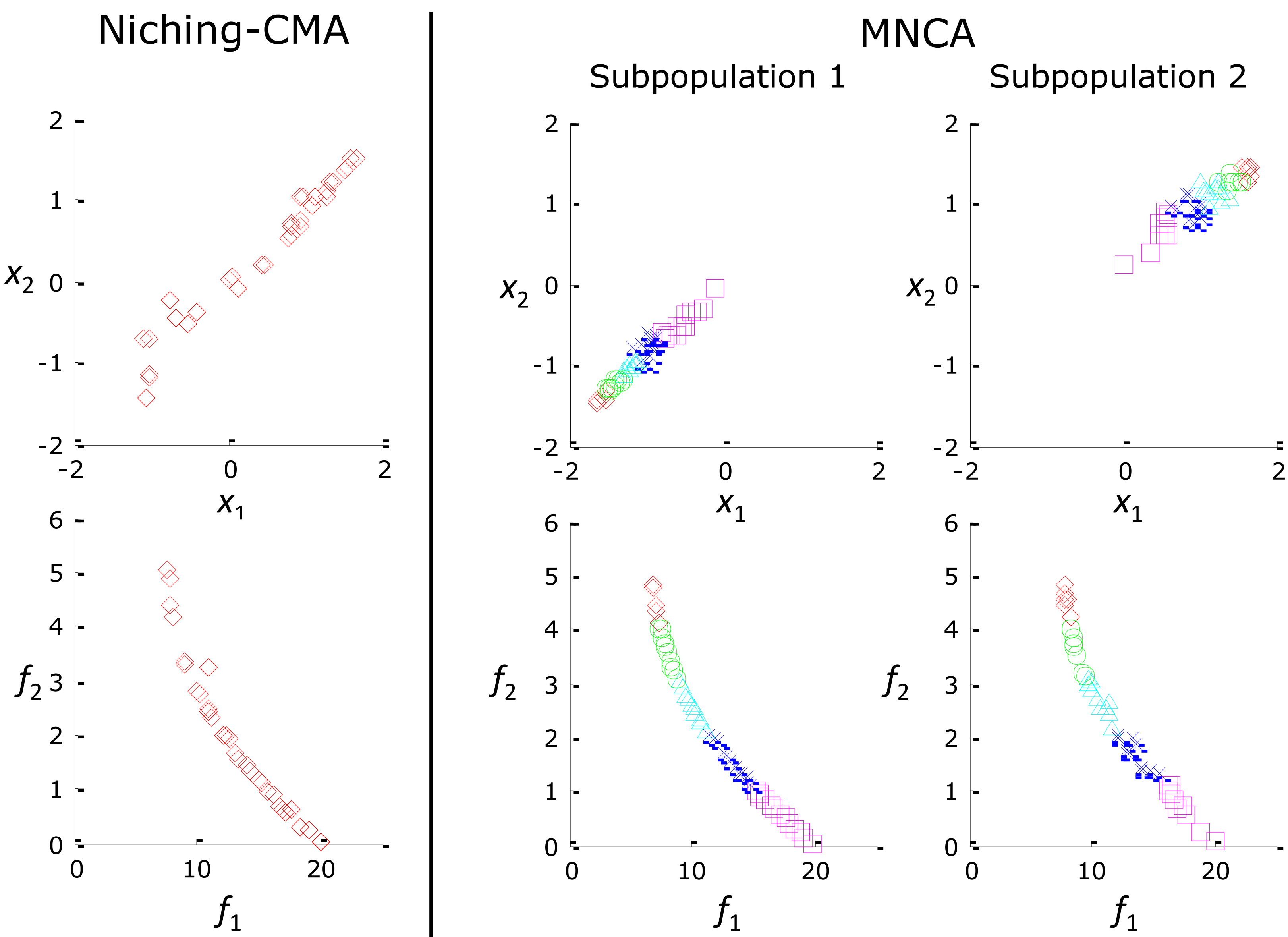


Function Two-on-One



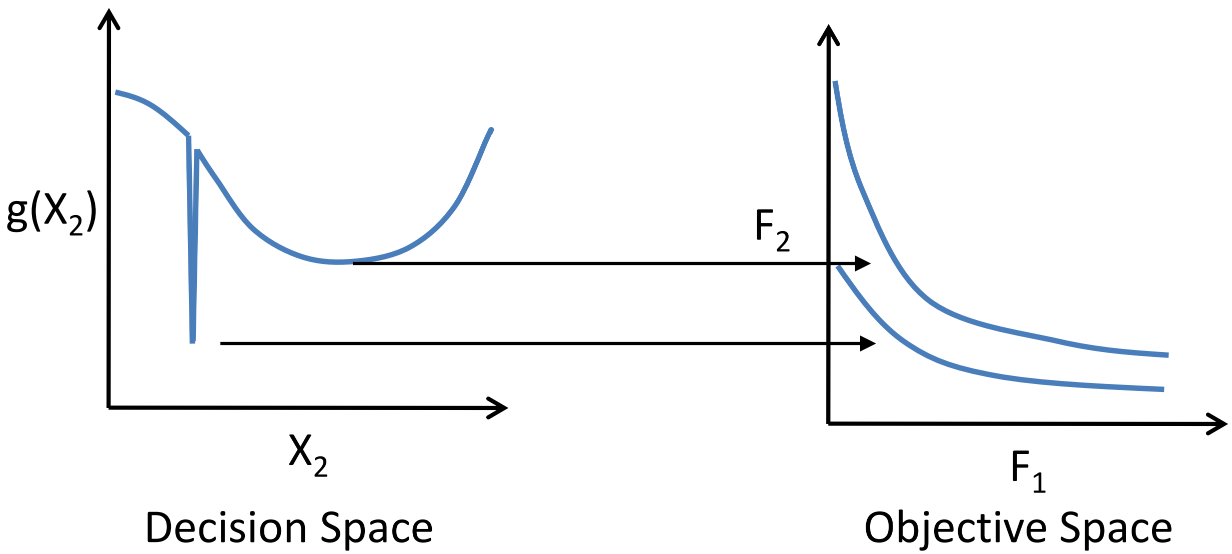


Function Two-on-One



Test Function Deb99

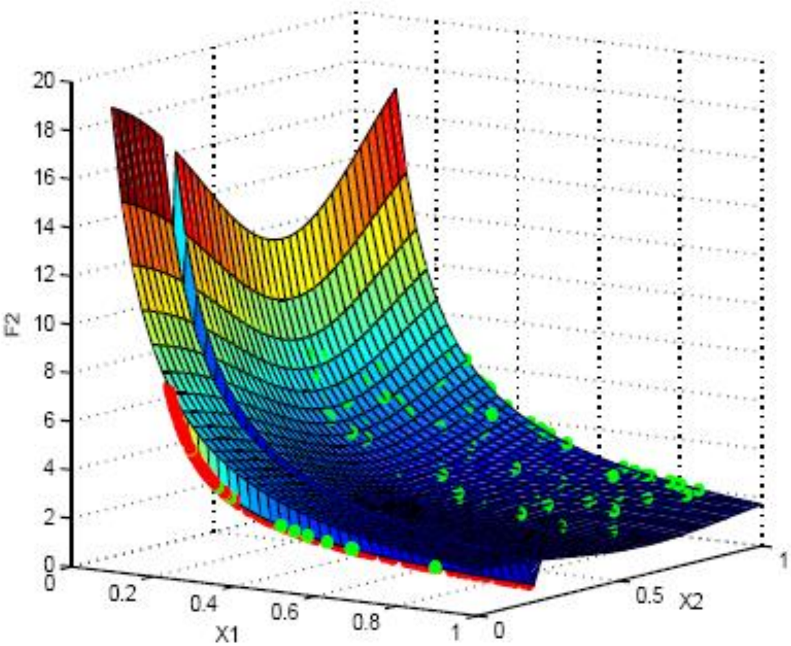
- Deb99 is a deceptive multi-objective problem that has a global optima difficult to identify and a local optima located in a long flat valley that is easy to find.

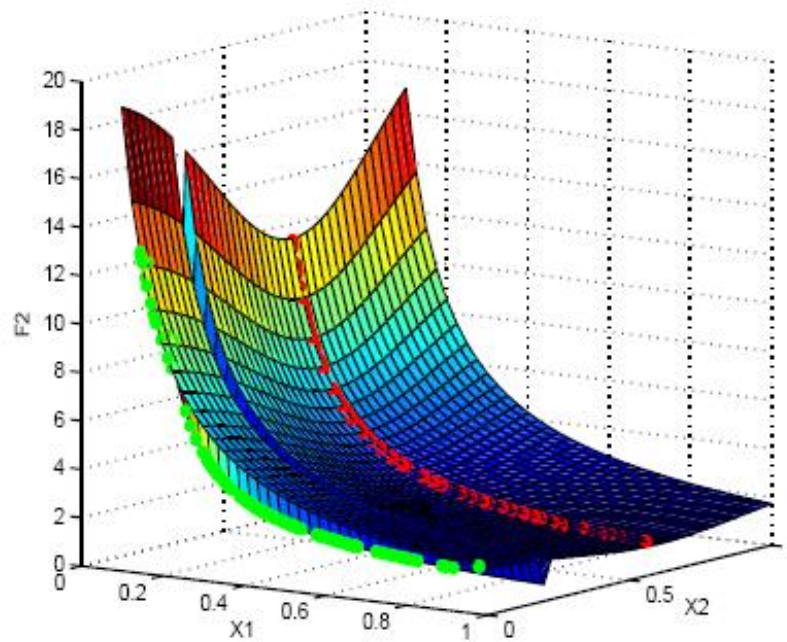


M. Preuss, B. Naujoks, and G. Rudolph, "Pareto Set and EMOA Behavior for Simple Multimodal Multiobjective Functions," in Parallel Problem Solving from Nature - PPSN IX, 2006, pp. 513–522.

Deb99 Test Function

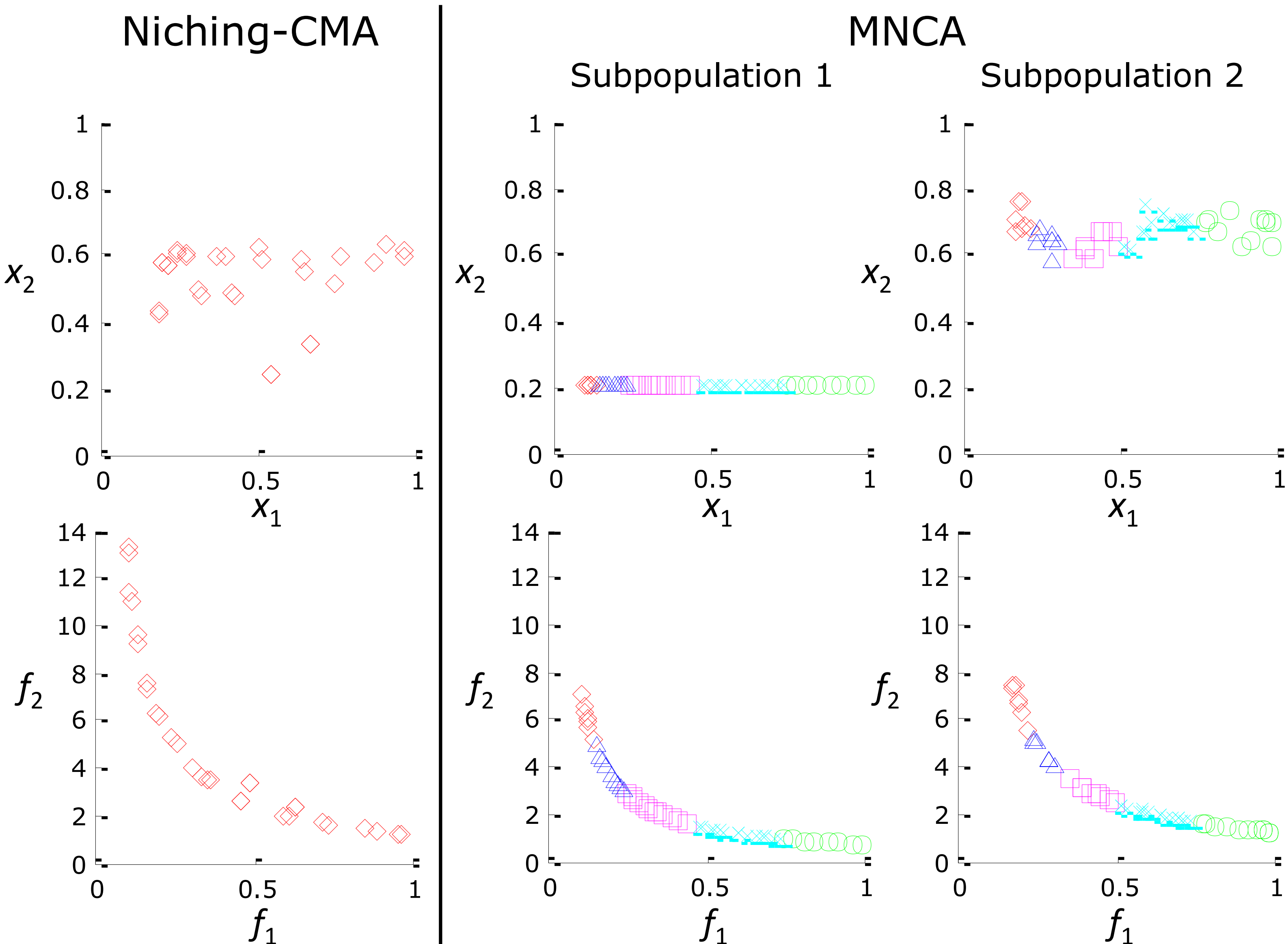
Deb99	Minimize $f_1 = x_1$; Minimize $f_2 = \frac{g(x_2)}{x_1}$	$0 \leq x_i \leq 1$
	$g(x_2) = 2 - exp\left\{-\left(\frac{x_2-0.2}{0.004}\right)^2\right\} - 0.8exp\left\{-\left(\frac{x_2-0.6}{0.4}\right)^2\right\}$	$i = 1, 2$





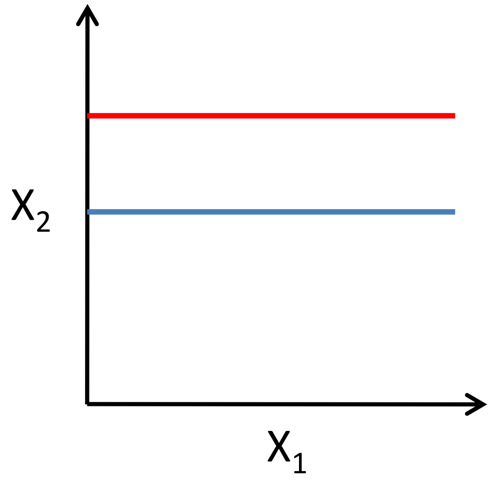
First subpopulation
Second subpopulation

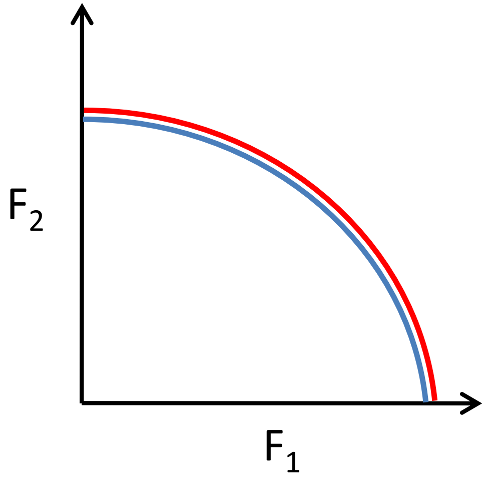
Function Deb99



Lamé Supersphere

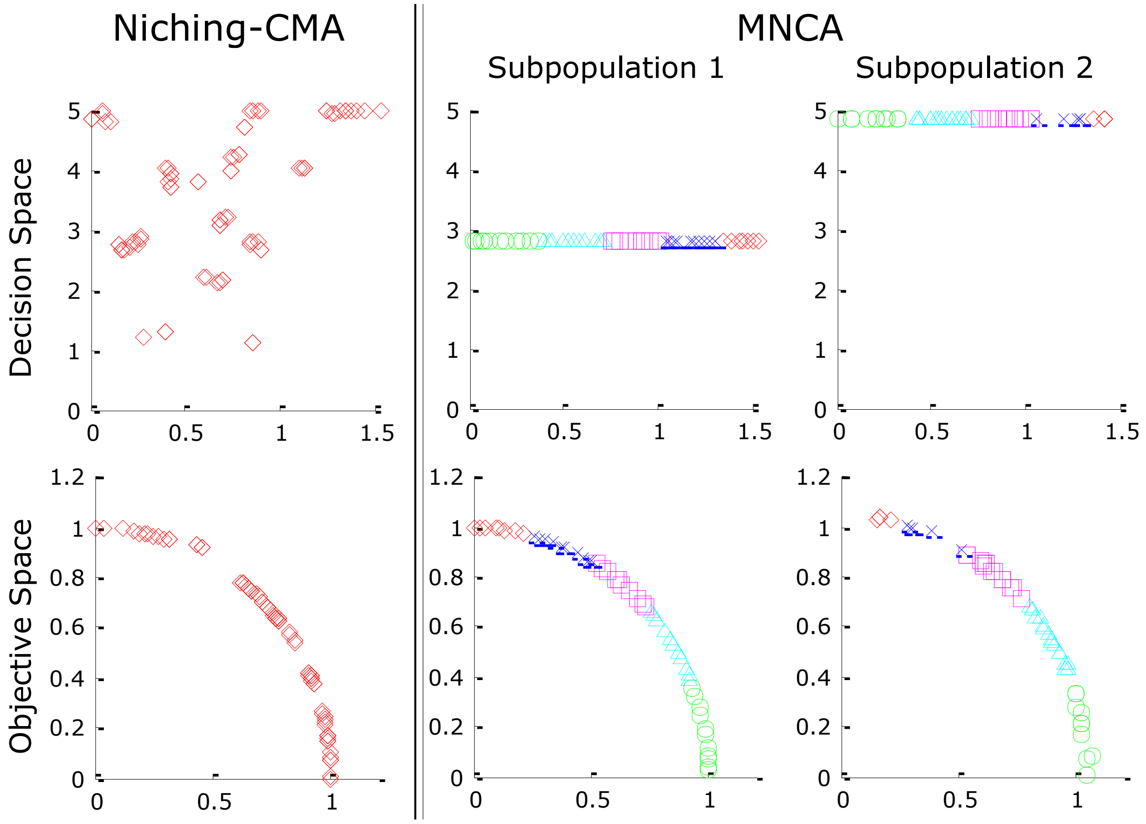
- Lamé Supersphere is a multi-modal problem global with spherical geometry in objective space and equidistant parallel lines in decision space.





M. Preuss, B. Naujoks, and G. Rudolph, "Pareto Set and EMOA Behavior for Simple Multimodal Multiobjective Functions," in Parallel Problem Solving from Nature - PPSN IX, 2006, pp. 513–522.

Function Lame Supersphere



Hypervolume

Test Function	Niching CMA	MNCA	
		Subpopulation 1	Subpopulation 2
Two-on-One	169.6 ± 1.9	**173.6 ± 0.1**	165.6 ± 3.2
Deb99	7.95 ± 0.69	**9.12 ± 0.01**	7.45 ± 0.35
Lamé Superspheres	3.12 ± 0.13	**3.19 ± 0.02**	2.95 ± 0.14

Decision Space Diversity

Test Function	Niching CMA	MNCA
Two-on-One	0.231 ± 0.031	0.298 ± 0.032
Deb99	0.285 ± 0.032	**0.359 ± 0.015**
Lamé Superspheres	**0.329 ± 0.039**	0.112 ± 0.007

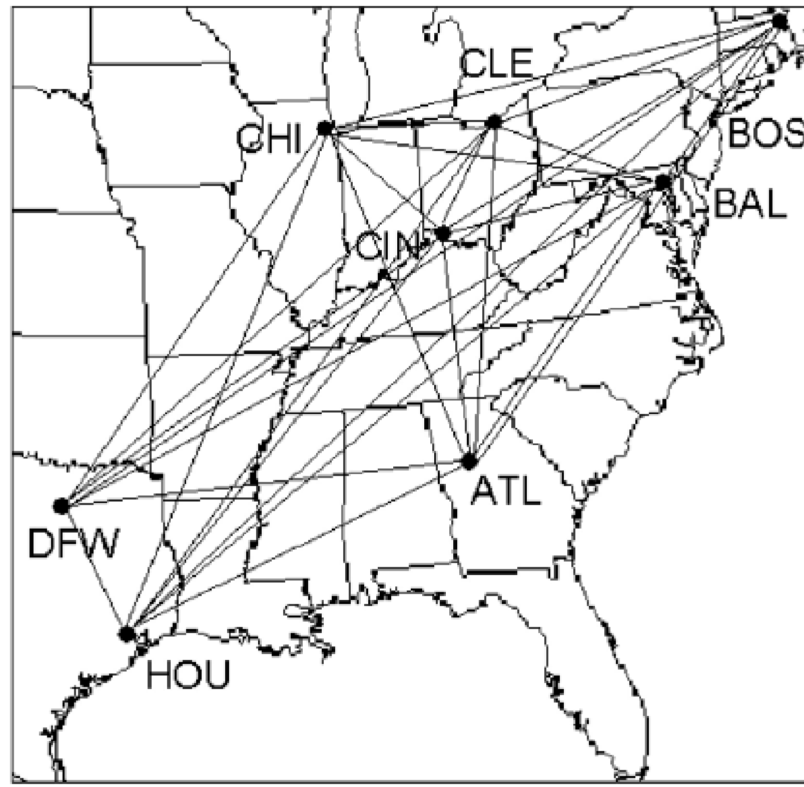
Paired Solution Diversity

Test Function	MCA	MNCA
Two-on-One	2.158 ± 0.239	**2.64 ± 0.49**
Deb99	0.23 ± 0.06	**0.47 ± 0.08**
Lamé Superspheres	0.79 ± 0.05	**1.59 ± 0.04**

Realistic Planning Problem: Airline Routing

- Determine a set of 28 routes to connect 8 cities to maintain objectives:
  - minimize f_1 : Cost
  - minimize f_2 : Number of stops

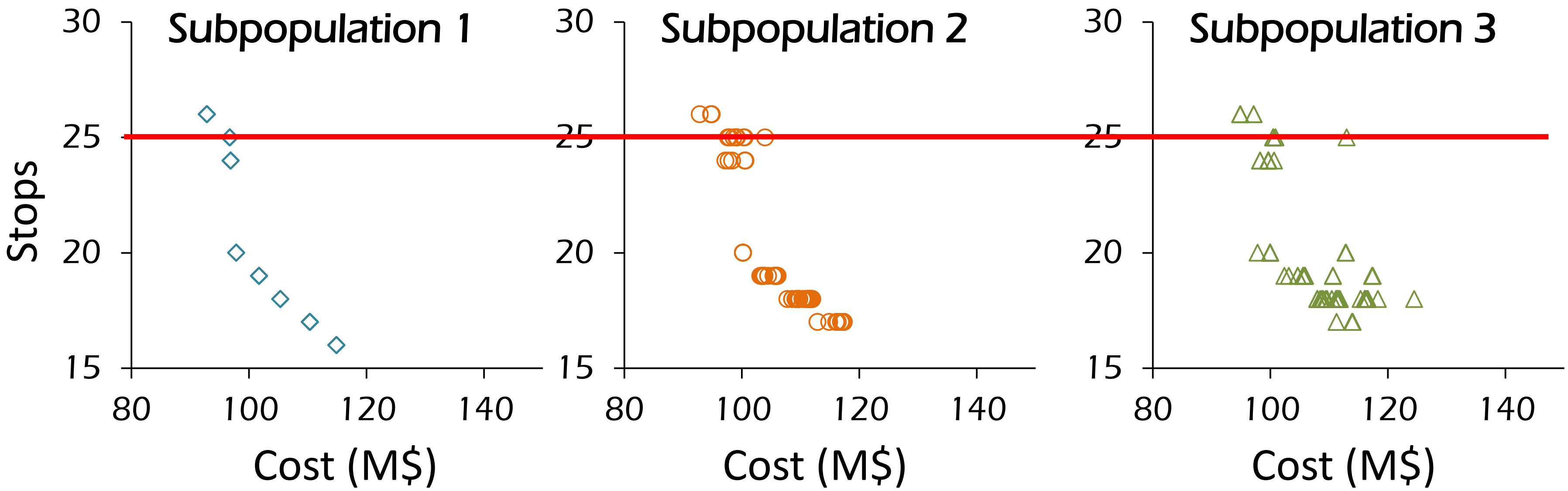
- Given:
  - Number of requests at each city
  - Cost of initial set-up
  - Cost of each flight
  - Unit cost for each passenger
  - Potential type of connections:
    - Directly
    - Indirectly- through multiple-leg connections



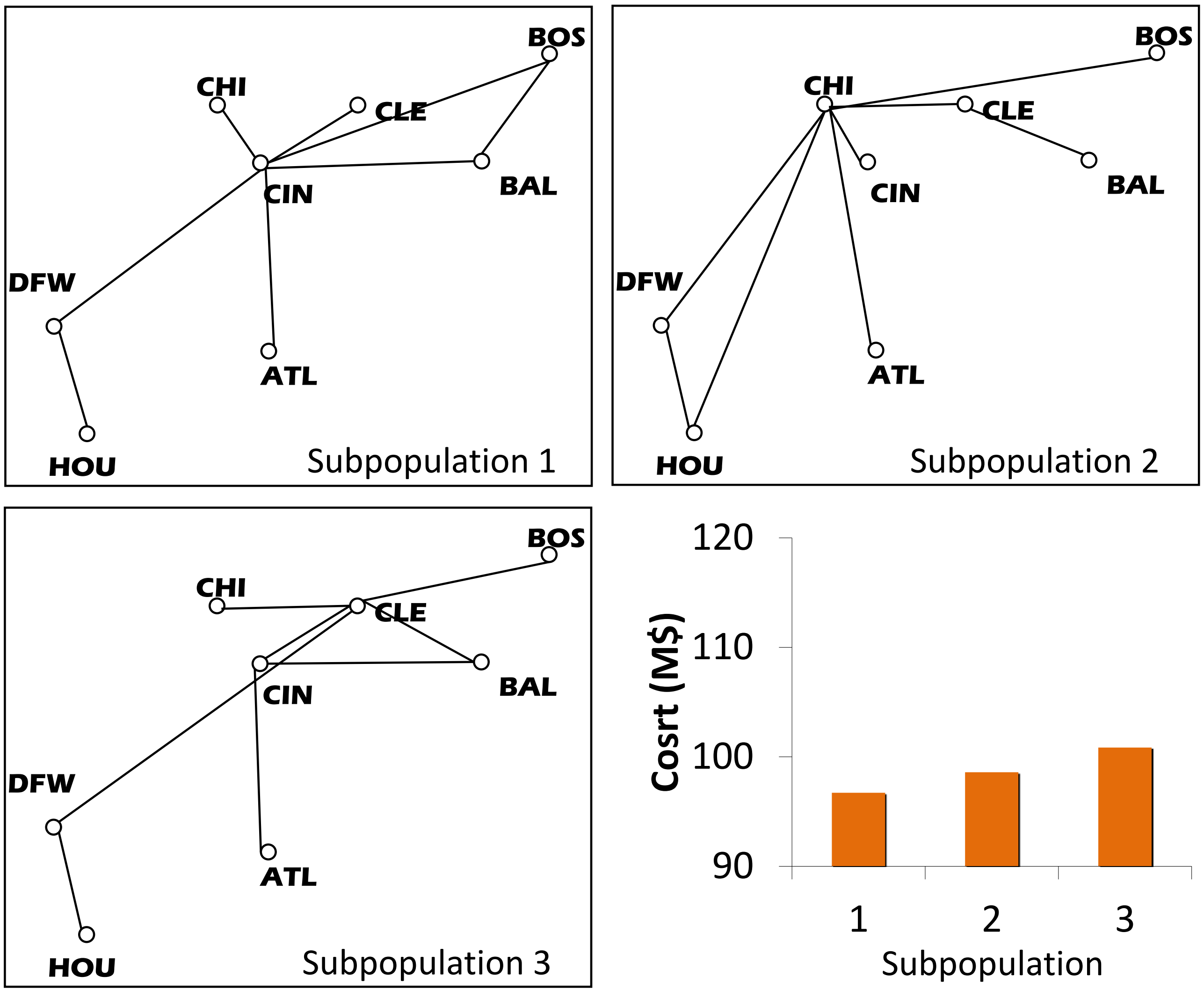
Settings for MNCA

Parameter	Setting
Population Size	100
Subpopulations	3
Generations	100
Mutation	1%
Clusters	5
Relaxation coefficient	90%

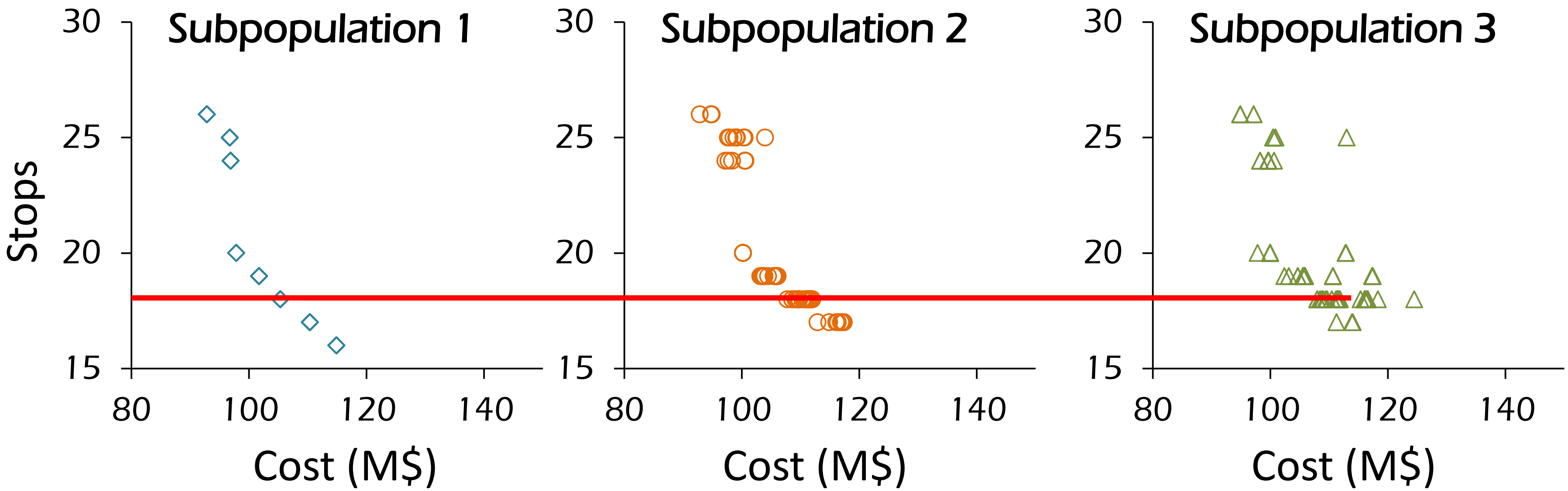
Distinct sets of non-dominated solutions



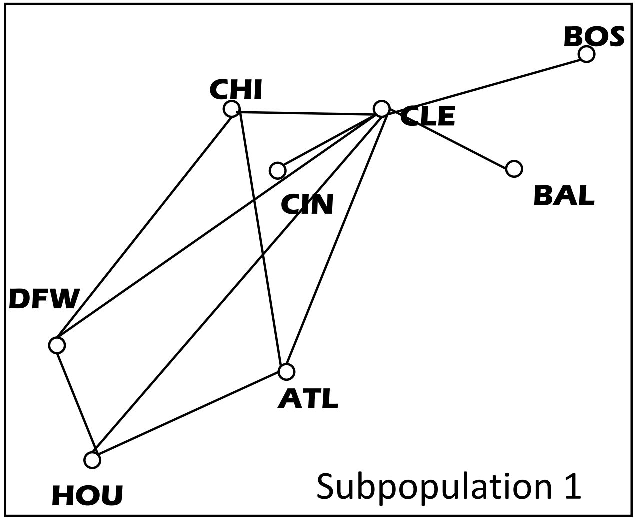
Twenty five-stop alternative solutions

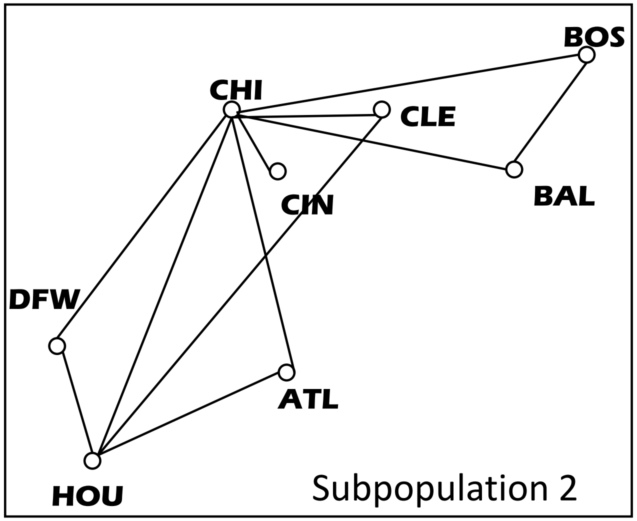


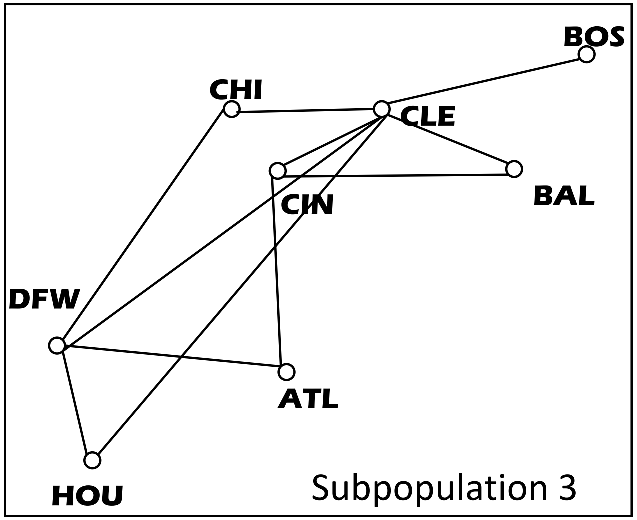
Distinct sets of non-dominated solutions

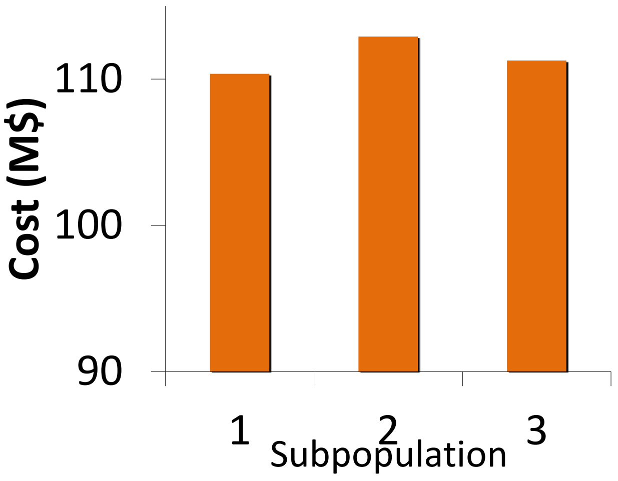


Eighteen-stop alternative solutions

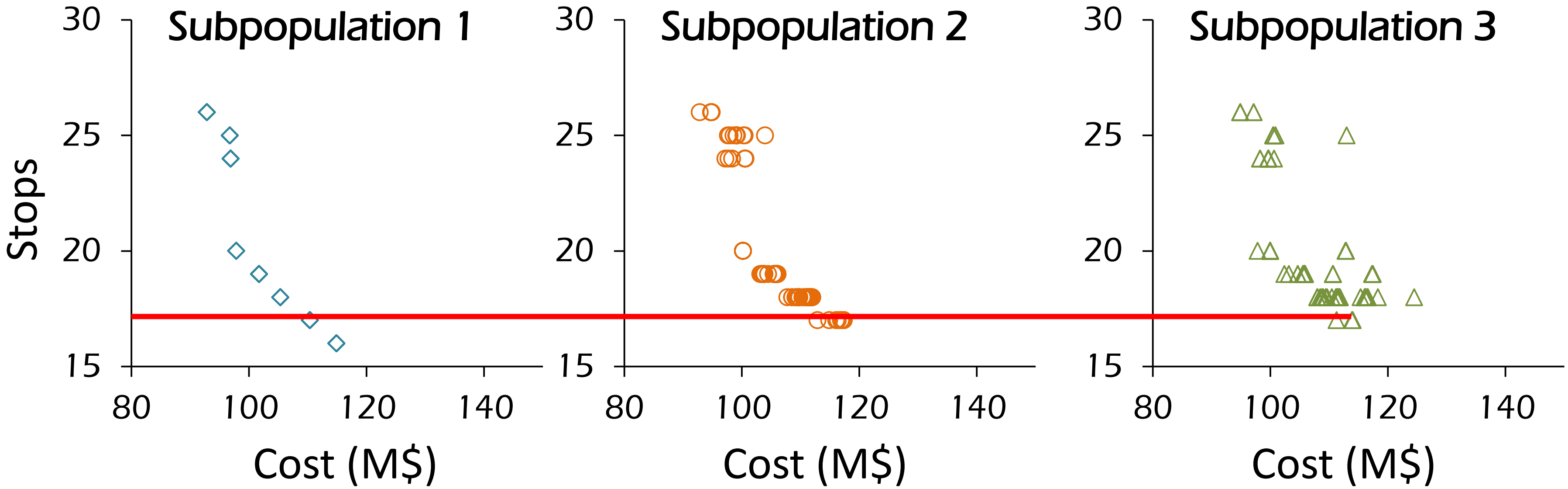




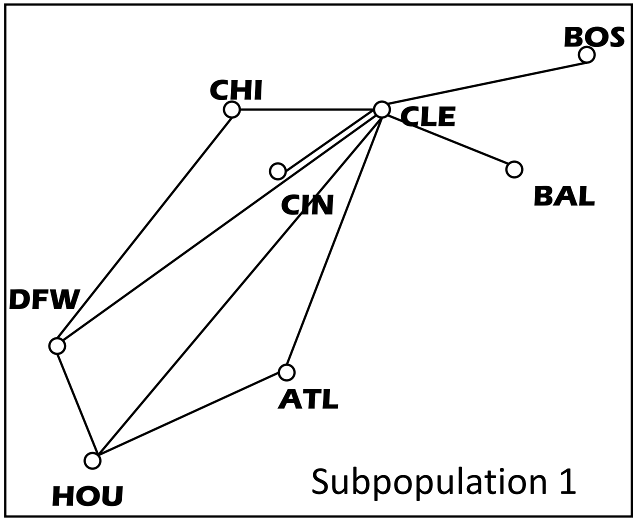


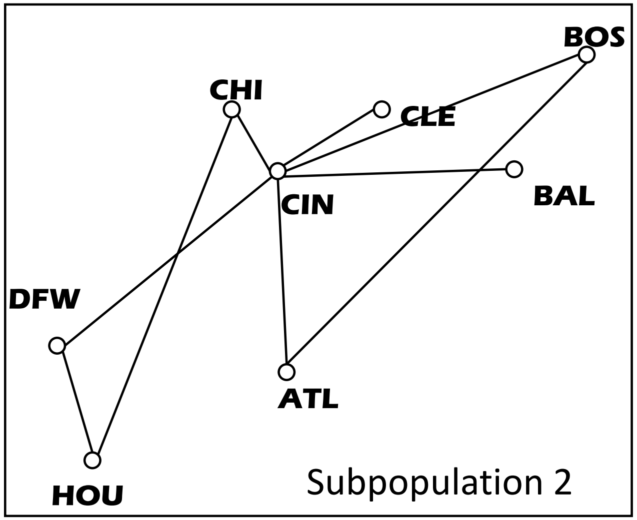


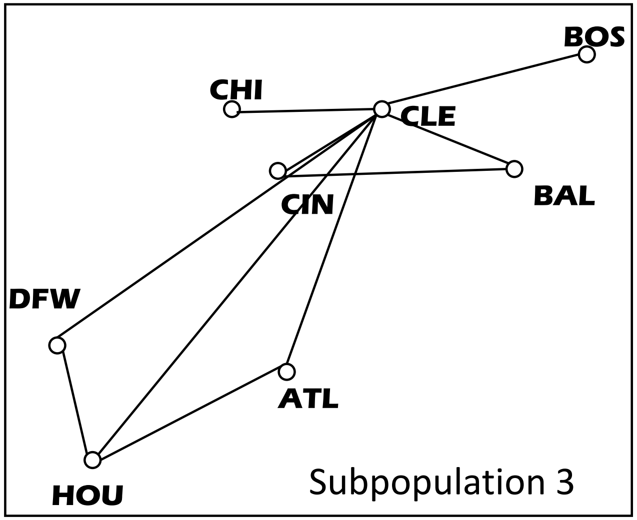
Distinct sets of non-dominated solutions

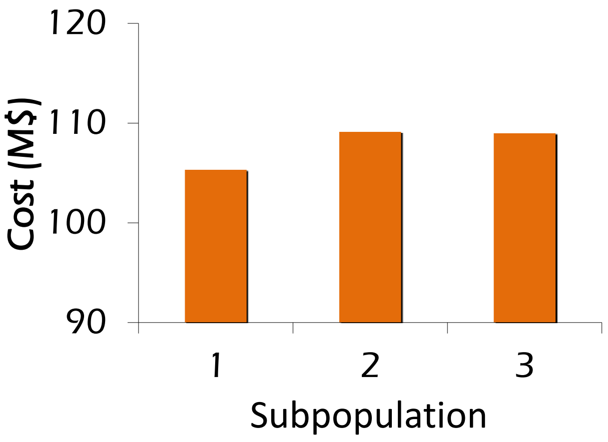


Seventeen-stop alternative solutions

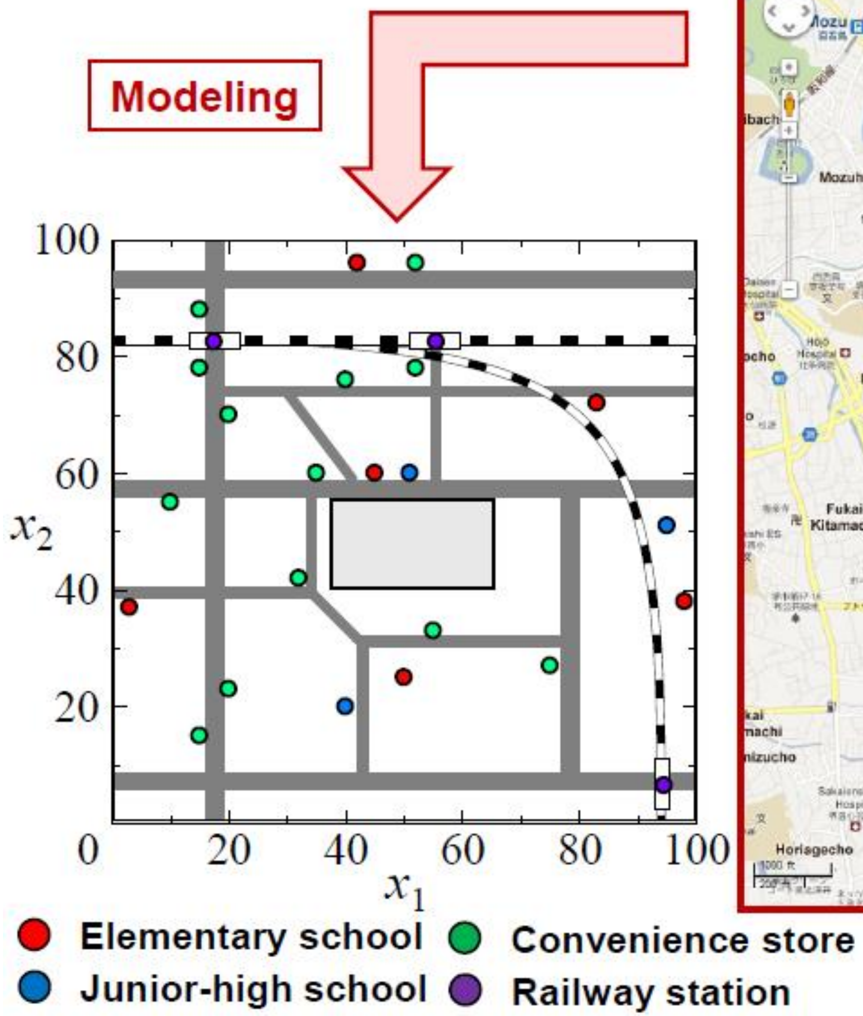


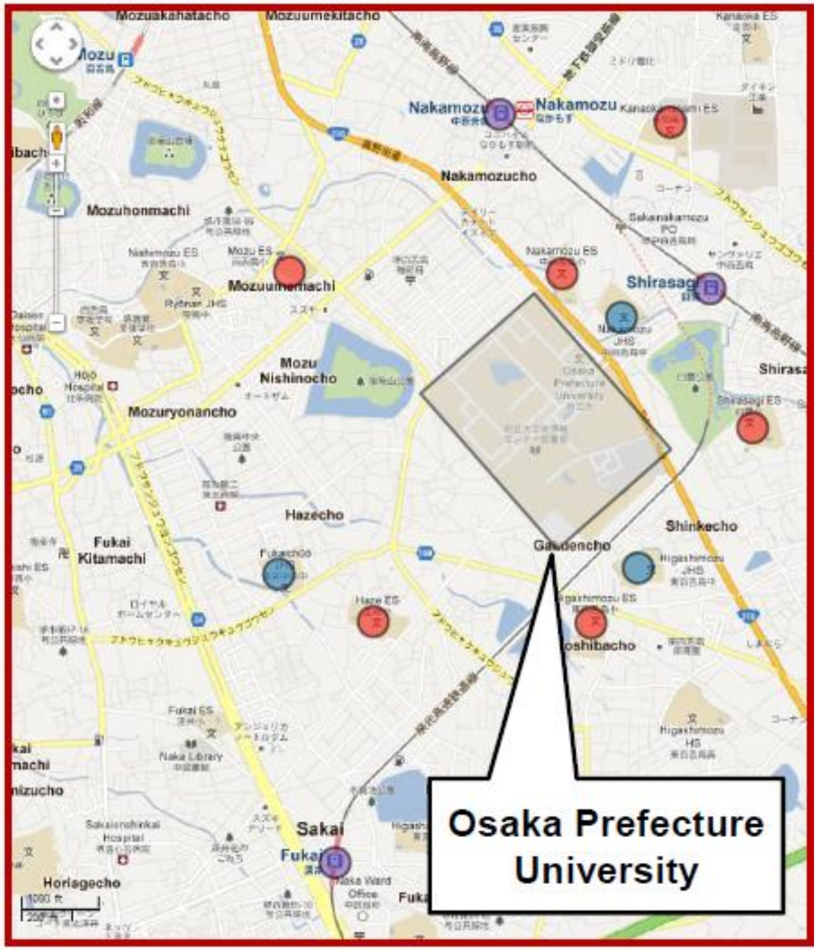






Where should Osaka Prefecture University be located?

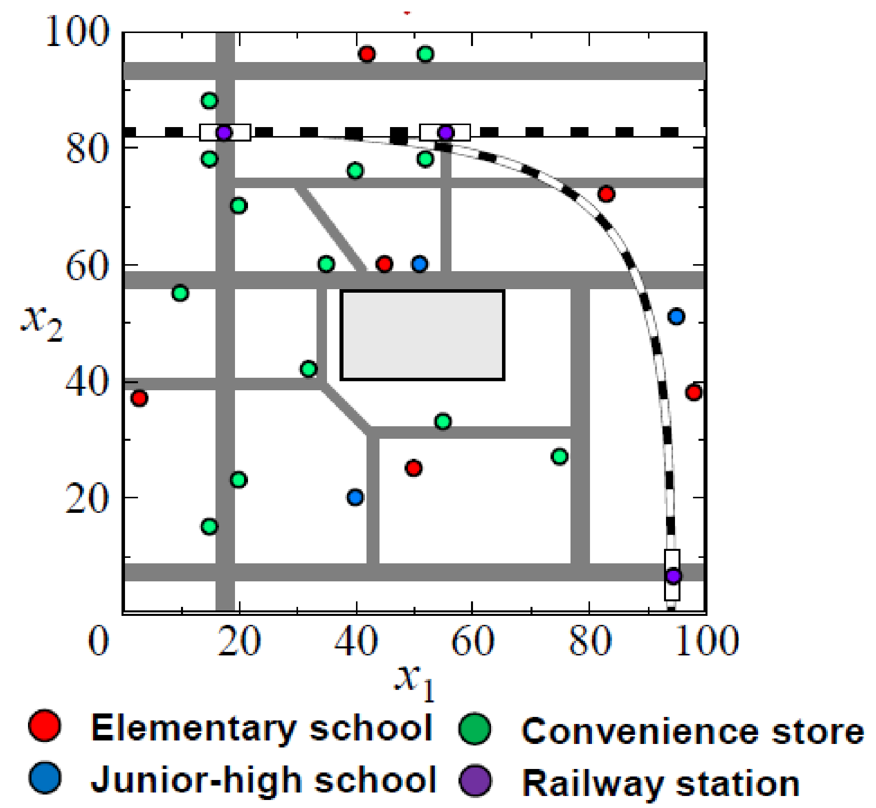




OFU – Problem Formulation

2 decision variables and 4 objective values

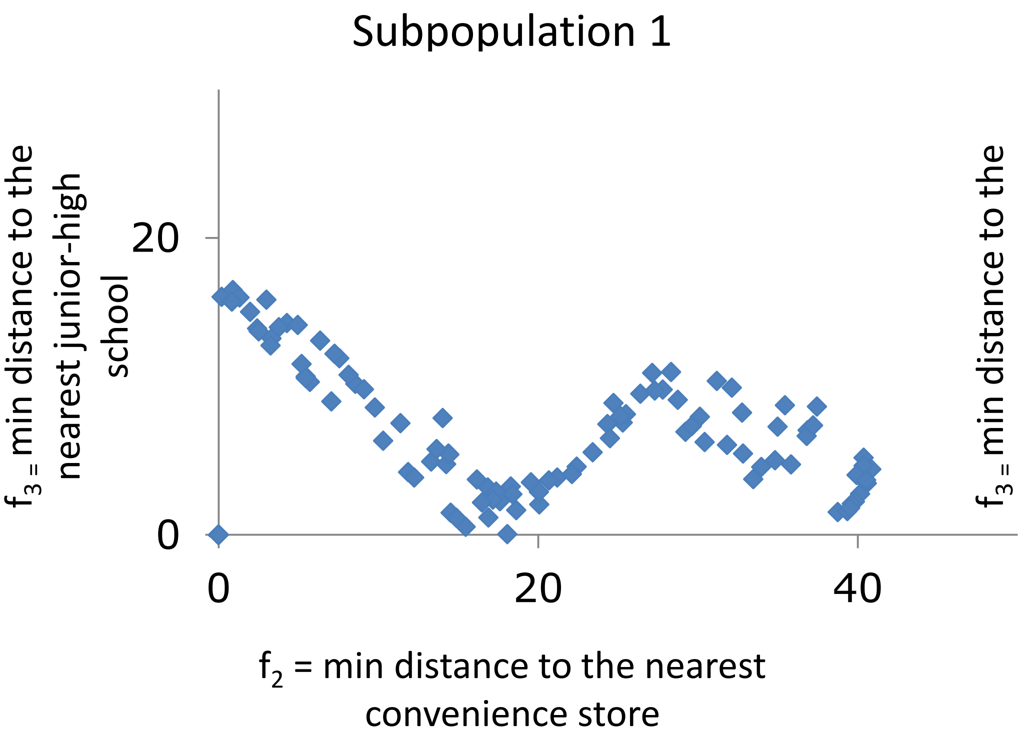
$f_1(x_1,x_2)$ = min distance to the nearest elementary school
$f_2(x_1,x_2)$ = min distance to the nearest convenience store
$f_3(x_1,x_2)$ = min distance to the nearest junior-high school
$f_4(x_1,x_2)$ = min distance to the nearest railway station

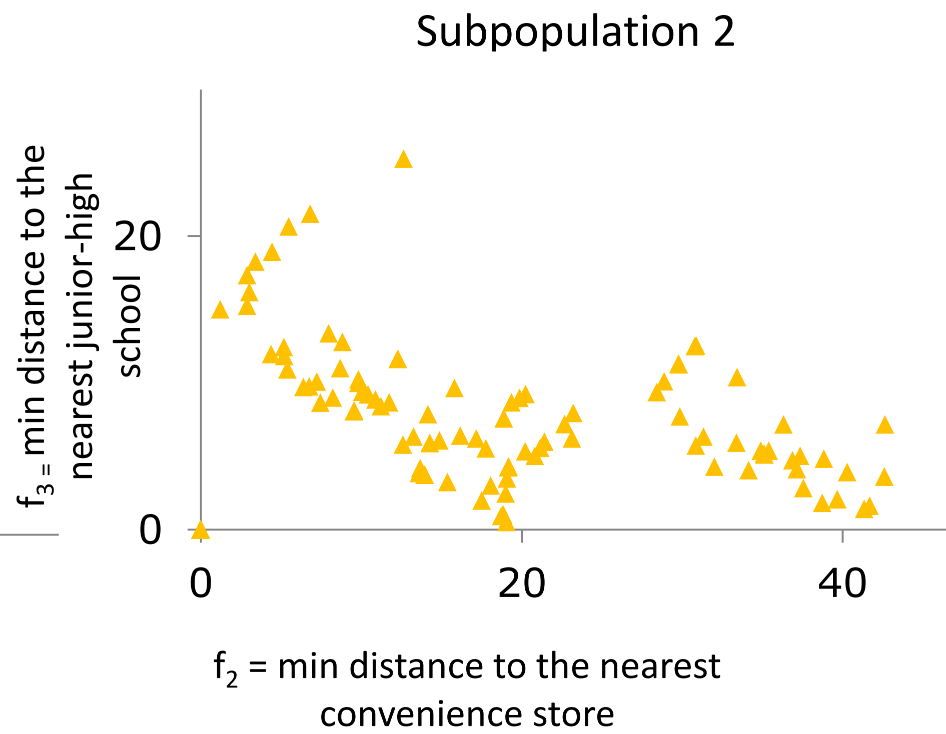


OFU – Algorithm settings

Parameter	Setting
Population Size	100
Subpopulations	2
Generations	100
Mutation	1%
Clusters	5
Target	95%

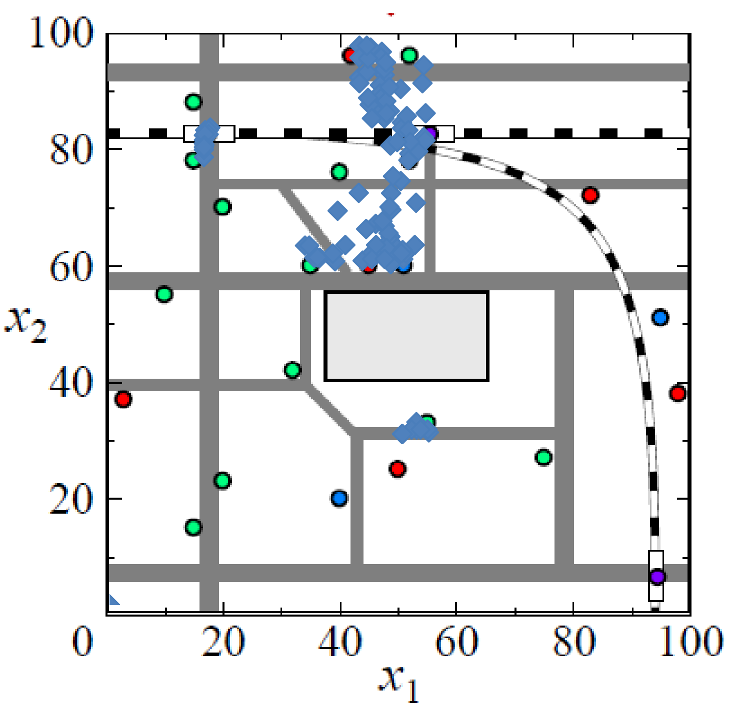
Results - Objective Space



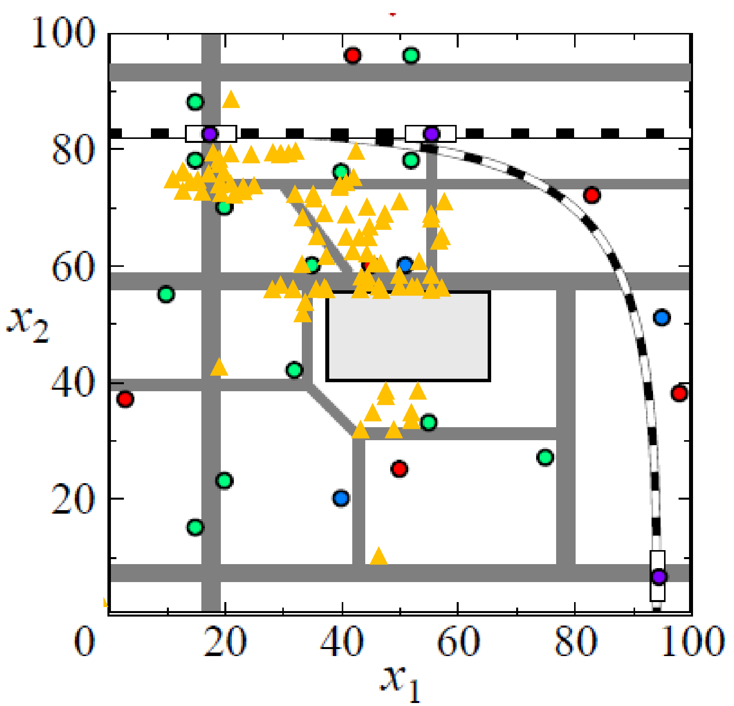


Results - Decision Space

Subpopulation 1



Subpopulation 2

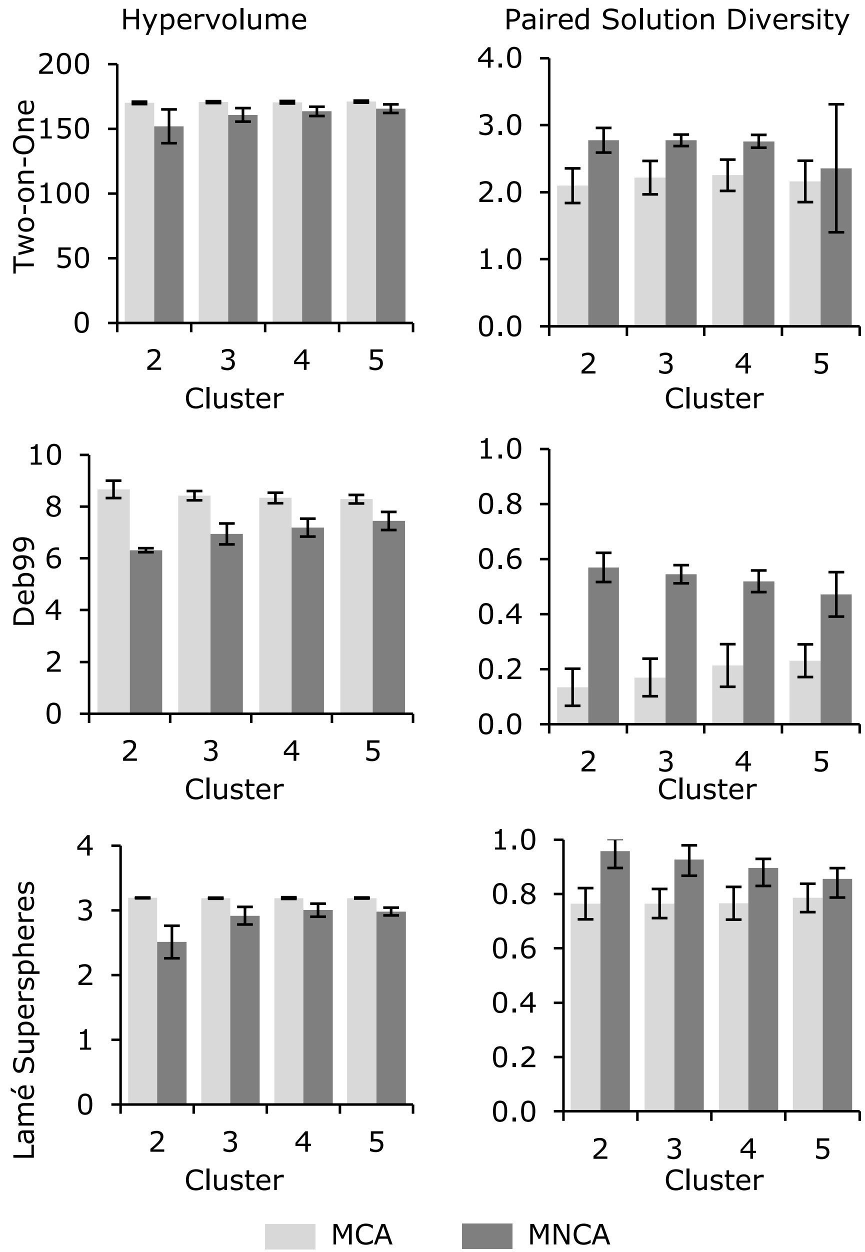


Conclusions

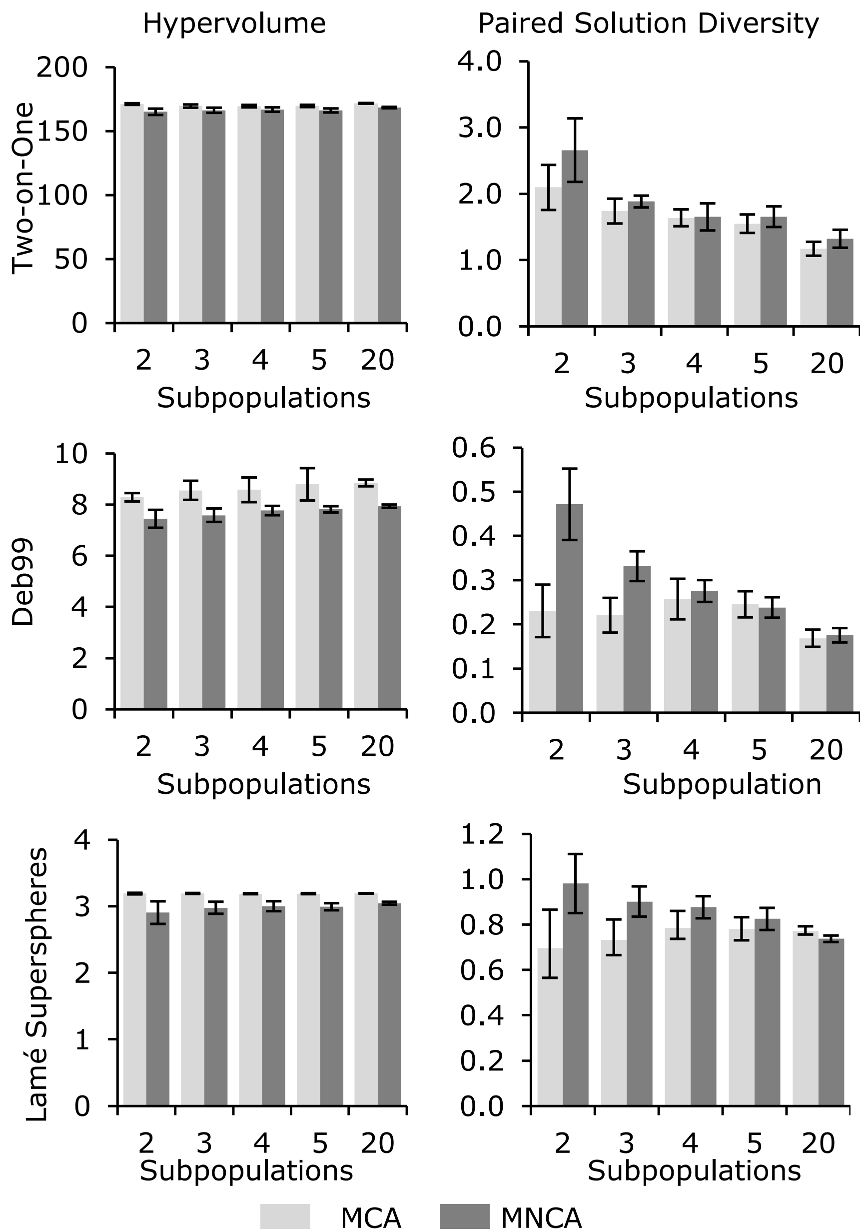
- Results demonstrate the ability of a new algorithm, MNCA, to identify sets of non-dominated solutions for test problems.
- Outperforms state-of-the-art (Niching-CMA) in diversity algorithms for multi-objective problems.
- A new metric to assess diversity among subpopulations was developed.
- MNCA maximizes diversity in decision space while identifying complete alternative Pareto fronts.

Thank you for your Attention!

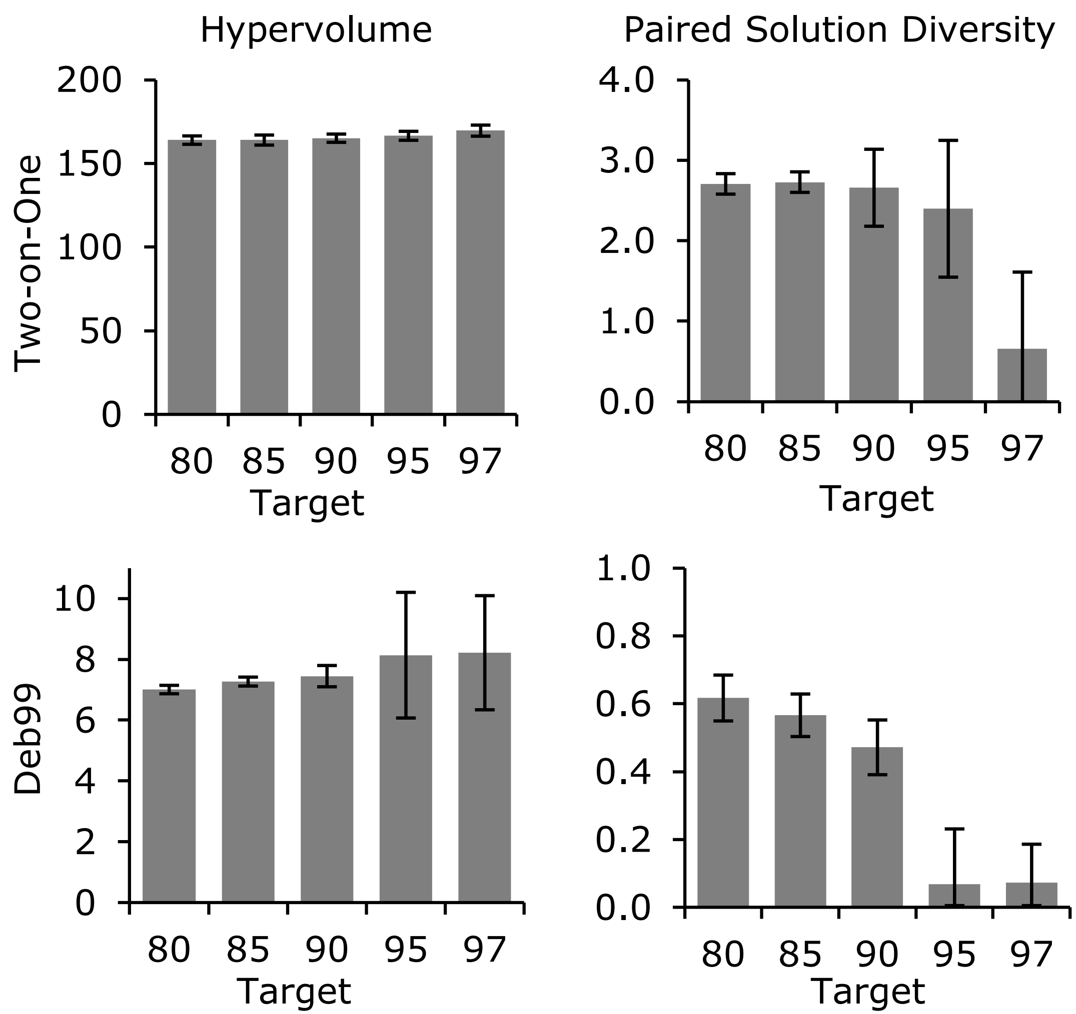
Number of Clusters



Number of Subpopulations



Targets



Hypervolume
Paired Solution Diversity
Two-on-One
Deb99
200
150
100
50
0
4.0
3.0
2.0
1.0
0.0
10
8
6
4
2
0
1.0
0.8
0.6
0.4
0.2
0.0
80 85 90 95 97
Target