Tomasz Schlieter
Institute of Computational Mechanics and Engineering,
Silesian University of Technology,
ul. Konarskiego 18A, 44-100 Gliwice
email: tomasz.schlieter@polsl.pl

APPLICATION OF SELF-ORGANIZING MAPS IN VISUALIZATION OF MULTI-DIMENSIONAL PARETO FRONTS

Summary – The aim of this study is to show the application of Kohonen's self-organizing maps for the purpose of visualization. Multi-dimensional sets of solutions of test functions for optimization and real multiobjective optimization problems are presented using self-organizing maps and traditional methods for comparison.

Keywords: self-organizing map, Kohonen's neural network, Pareto front, visualization

1   Introduction

The aim of this study is to investigate the application of Kohonen's self-organizing maps for the purpose of approximating n-dimensional optimization solution space as n colored two dimensional maps to enhance the dialogue between the analyst solving an optimization task and a decision maker. Such an approach to visualization should be able to reduce the dimensionality of solution data while clearly showing the values of optimization criteria and outlining the nature of conflict between the criteria, including objectives tradeoffs and discontinuities on the front. Additionally it should handle data clustering and highlighting of selected solutions.

There are two main approaches to multi-dimensional optimization tasks, one of which is to seek a single optimal solution, having the objectives scalarized and therefore converting a multiobjective optimization problem into a single objective problem e.g. using Weighting Method or ε-Constraint Method [2]. Another approach is to seek not a single optimal solution but a set of non-dominated (Pareto optimal) solutions.   The former approach is relatively simple in implementation and do not require any additional visualization tools as the results are straightforward, while the latter provides analyst with more information about the problem, including objectives tradeoffs, but is usually harder to implement, comparison of solutions' quality being an

issue. This approach generates problems in understanding the results in case of more than two objectives, getting even worse as the number of objectives and thus dimensionality of the Pareto front increases.

Effective visualization techniques are essential as the graphical interpretation of data obtained from the optimization task is a link providing the decision maker with the information from the analyst. There are several ways to overcome difficulties in visualization of high-dimensional Pareto frontiers e.g. decision maps and interactive methods. Self-organizing maps can be utilized to project high-dimensional Pareto fronts as points on two-dimensional maps.
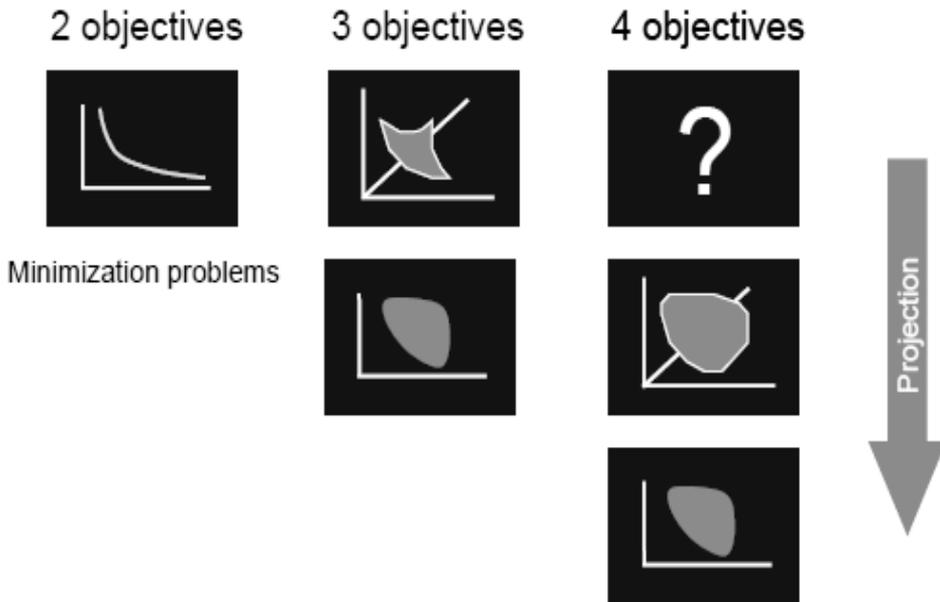

Fig. 1.   Visualization of Pareto front [5]

2   Self-organizing maps by Kohonen

Self-Organizing Map (SOM) is an artificial neural network introduced by Kohonen in [4], used for visualization of high-dimensional data sets. SOM is used to generate a similarity graph of input data. In case of multiobjective optimization results visualization SOM can be used to convert high-dimensional sets of non-dominated solutions into simple geometric relationships of their image points on a regular, usually hexagonal, 2D grids of nodes. It has numerous applications: visualization, data mining and clustering to be named in terms of working with optimization solutions as input data.

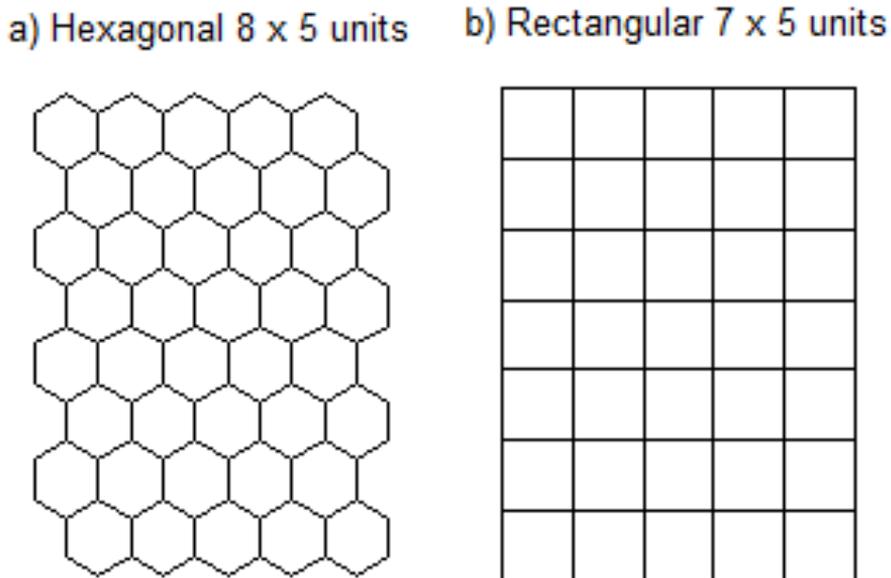a) Hexagonal 8 x 5 units   b) Rectangular 7 x 5 units

Fig. 2.   SOM 2D topology

   Each SOM node (unit) is described by its codebook vector, consisting of map weights. Size of codebook vectors is equal to a size of a single input data sample. Codebook vectors' values are initialized in the first phase of creating a network. Random or linear (fig. 3) initialization are used, the latter is proven to be more effective. In the training part of the algorithm input data is presented to the network and the best-matching unit (BMU) is chosen among all map units using a Euclidean distance as a criterion to be minimized. BMU and its neighbors are adjusted towards an input pattern, hence the topological order of data is preserved in SOM – the data located close to each other in the input space are also close to each other on the map. The Gaussian neighborhood function is used in this paper. SOM training is an iterative task, usually consisting of rough training and fine tuning. Trained SOM can be presented as a number of grids of colored nodes. In case of optimization results visualization, each grid corresponding to a single optimization criterion, and nodes corresponding to particular solutions, with the value of criterion depicted as unit's color (fig 3 a) and b)). Often a measure of similarity between data in the neighboring units is desired to be depicted, thereof another grid called U-matrix, showing the unified Euclidean distance between codebook vectors of the neighboring units (fig. 3 c)).
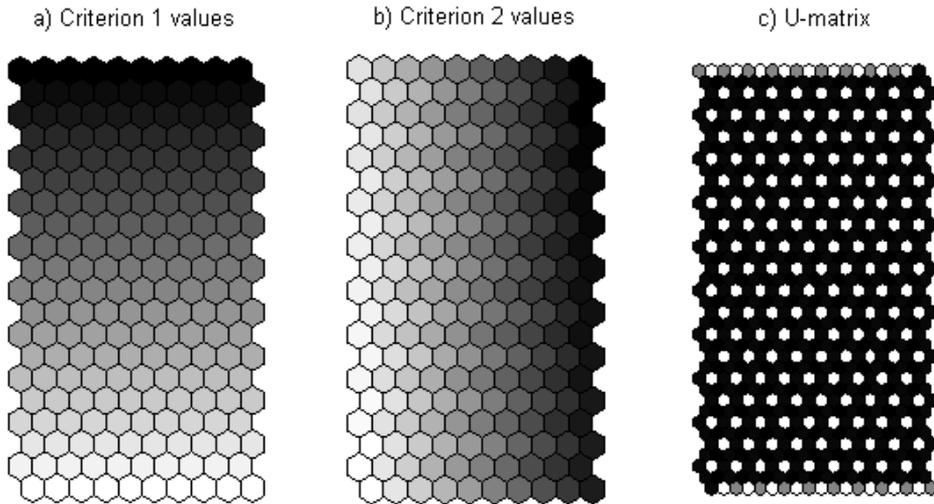
Fig. 3.   Presentation of a 20×10 units SOM for the 2 criterion optimization
          before the training, linear initialization only, no input data presented to
          the network.

For the purpose of creating and displaying SOM Toolbox Version 2.0
[7], a toolbox for MATLAB was used.

3   Test function solutions

To show the way SOM can be utilized for the purpose of visualization
of multi-dimensional Pareto fronts two objective optimization test
functions with known solutions [6] were chosen. In author's opinion two
objectives are sufficient to present the idea of SOM-based visualization
in comparison with traditional presentation of results on a single X-Y
plane. An example of four objective optimization is presented in chapter
4. The ZDT2 and ZDT3 test functions represent a non-convex front and
a front consisting of several noncontinuous, convex parts.
Formulation of ZDT2 test problem:

$$f_1 = x_1$$
$$f_2 = g(x)\left[1 - \left(x_1/g(x)\right)^2\right]$$
$$g(x) = \left[1 + \frac{9}{n-1}\right]\left(\sum_{i=2}^{n} x_i\right)$$
$$0 \leq x_i \leq 1, \quad i = 1, \ldots, n$$

(1)

Functions f_1 and f_2 are to be minimized, function g is an ancillary function used to define f_2 only.
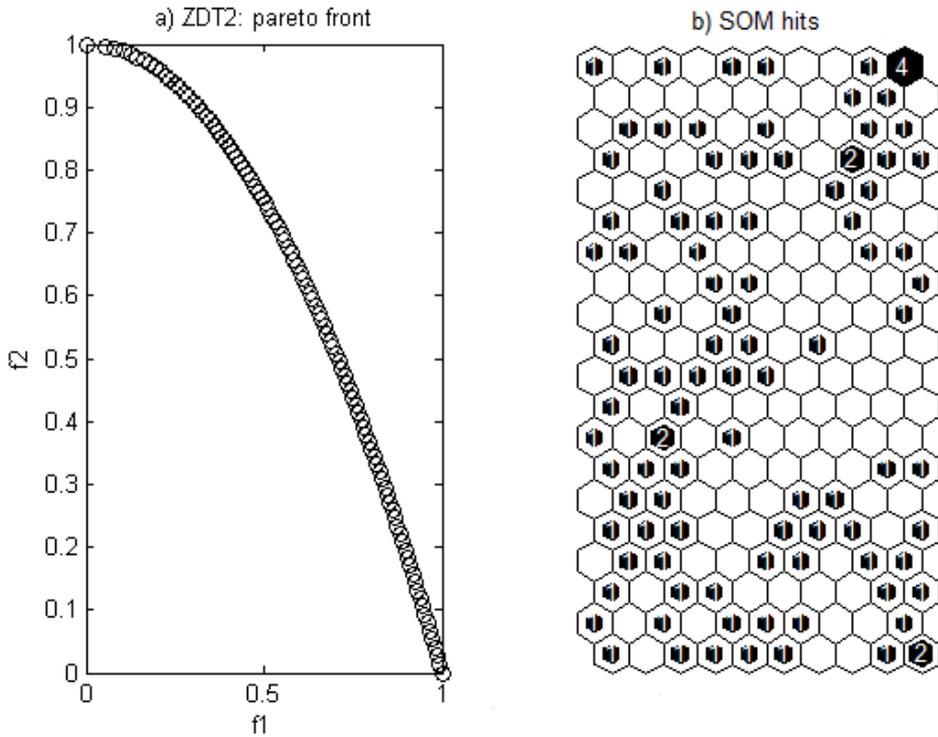

Fig. 4.   a) 100 samples of evenly distributed Pareto-optimal solutions to ZDT2 and b) BMUs location on SOM for the samples.
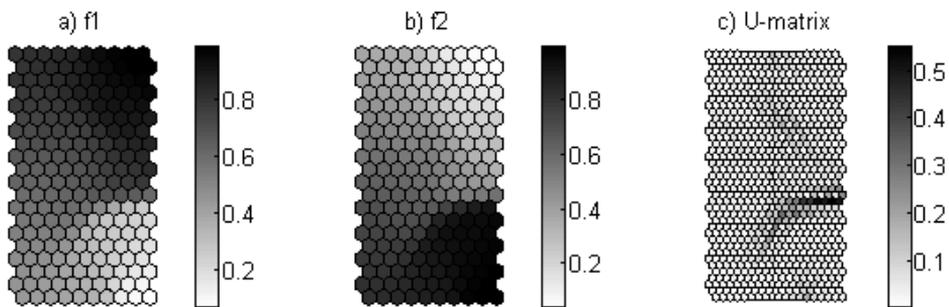

Fig. 5.   SOM presentation of the ZDT2 results on a 20×10 units map, trained using batch algorithm.

It can be noticed that the f_1 and f_2 objectives are contradictory and tradeoffs between them are similar in the entire objective space. Hence

the Pareto front is continuous and the samples are evenly distributed there is no remarkable leaps in the objective values and the U-matrix values are low for most of the units.

Formulation of ZDT3 test problem:

$$f_1(x) = x_1$$

$$f_2(x) = g(x)\left[1 - \sqrt{x_1/g(x)} - x_1/g(x)\sin(10\pi x_i)\right]$$

$$g(x) = \left[1 + \frac{9}{n-1}\right]\left(\sum_{i=2}^{n} x_i\right)$$

$$0 \le x_i \le 1, \quad i = 1,\ldots,n$$

(2)

Functions f_1 and f_2 are to be minimized, function g is an ancillary function used to define f_2 only.
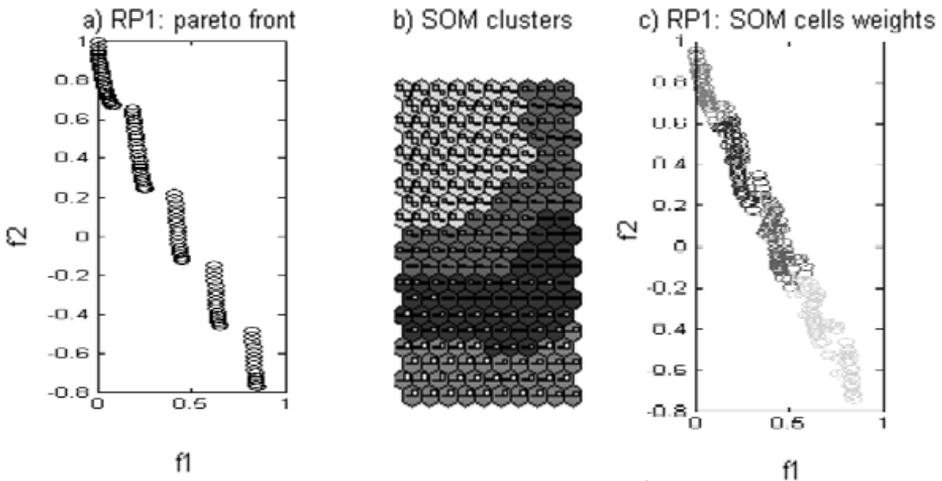

Fig. 6.   100 samples of evenly distributed Pareto-optimal solutions to ZDT3 and clustering of the SOM. Bars representing codebook vectors.

4   Real optimization problems solutions

Two solved problems were chosen to show how to deal with the visualization of real optimization problems in terms of two and four objectives. Both problems are stated, well described and solved in [1] and [3] while this study is supposed to propose an alternative method of visualization for the results presented in a traditional way by the authors. Traditional methods are also included in this chapter to be compared with the SOM approach.

Real problem 1 (RP1) concerns multiscale evolutionary optimization of Functionally Graded Materials. Example of a three-point bended beam is described in [1]. Objectives to be minimized are: f_1 – the maximum deflection of the beam and f_2 – the total cost of the structure, represented by the total volume fraction of the reinforcement. Design variables are fibre dimensions in the sections of a bended beam.
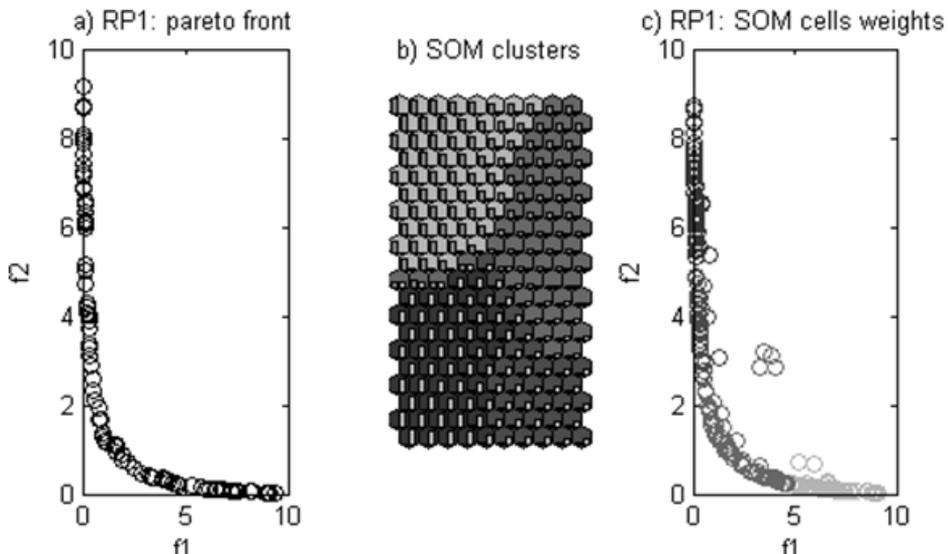

Fig. 7.   86 Pareto-optimal solutions to RP1 and clustering of the SOM. Bars representing codebook vectors.

It can be noticed that there are some cells in SOM whose weights are located in a distance from the Pareto front. These cells are surrounded by high values of U-matrix and can be located on the map easily and should not be taken into consideration by the decision maker.

In the fig. 9 BMUs of five design variables sets were highlighted. Design variables were sorted and numbered in an ascending order based on f_1 value. DV1 is a solution with minimal deflection, while DV 84 is a solution with maximum deflection among non-dominated solutions. Areas on the map where a small increase in one objective function causes a significant decrease in the contradictory objective function (both criteria to be minimized) can be noticed. Such a presentation of the results gives a decision maker a picture of severe tradeoffs existing amid the non-dominated solutions.

Real problem 2 (RP2) concerns multiobjective optimization of electrothermal microactuators by means of Immune Game Theory Multiobjective Algorithm (IMGAMO). Example of coupled mechanical, thermal and electrical analysis of the chevron-type actuator is described

in [3]. Objectives to be minimized are: f_1 – volume of the structure, f_2 – maximum value of the equivalent stress, f_4 – total heat generated by the actuator, objective to be maximized is: f_3 – deflection of the actuator. Design variables are the parameters describing shape of the actuator (fig. 10). The IMGAMO algorithm dismisses evaluation of criteria based on domination (which is troublesome for more objectives) in favor of a metaphor of playing games between players representing optimization criteria and is proven to be superior to popular multiobjective optimization algorithms (e.g. NSGA-II – nondominated sorting genetic algorithm II) in terms of working with optimization problems concerning more than 3 criteria.
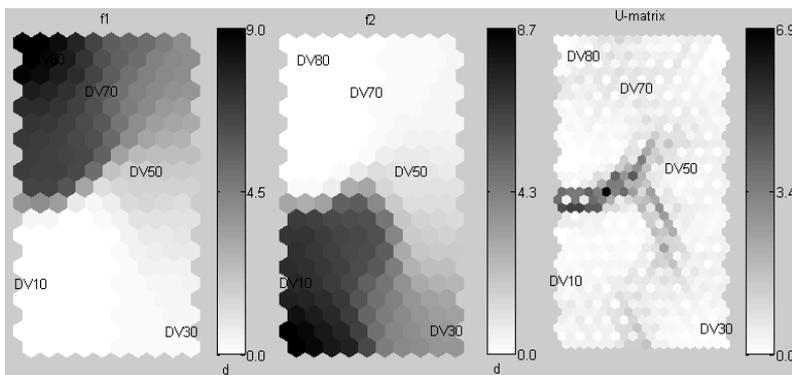

Fig. 8.   SOM presentation of the RP1 results with some of the design variables highlighted.
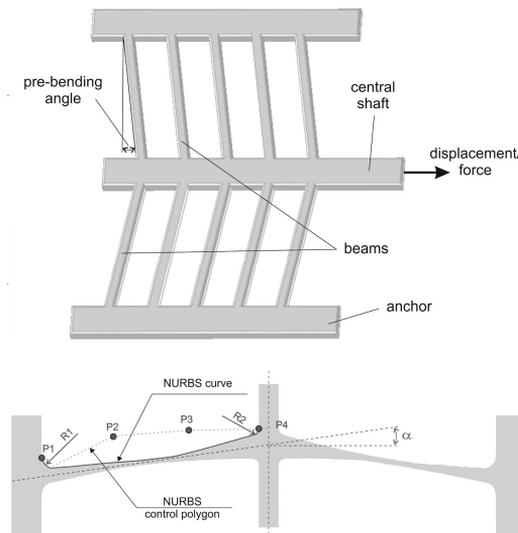

Fig. 9.   Chevron-type actuator and geometry to be optimized

Results in case of four objectives are hard to be displayed in a clear, simple and precise manner and in such a way to ensure the decision maker is provided with complete information on the non-dominated solutions including tradeoffs between objectives. Among the visualization method in this case multiple plot method, showing how the individual objective function values are correlated can be highlighted (fig. 11). This method can be also used for higher order data sets. In general for n dimensional data set $(n$-$1)^2$ plots of pair-wise relationships between the variables are drawn and histograms for each variable can be added.
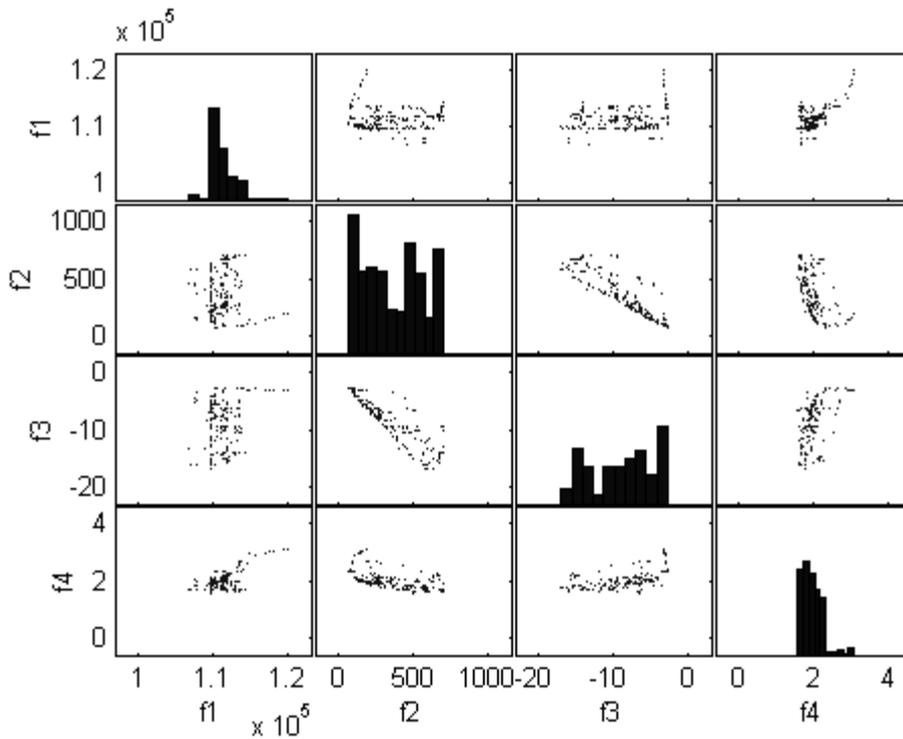

Fig. 10.   Multiple plot method for RP2 results.

Another method, dedicated to 4D data sets is using three dimensions and color to represent the fourth (fig. 12). The visualization lacks clarity and precision.
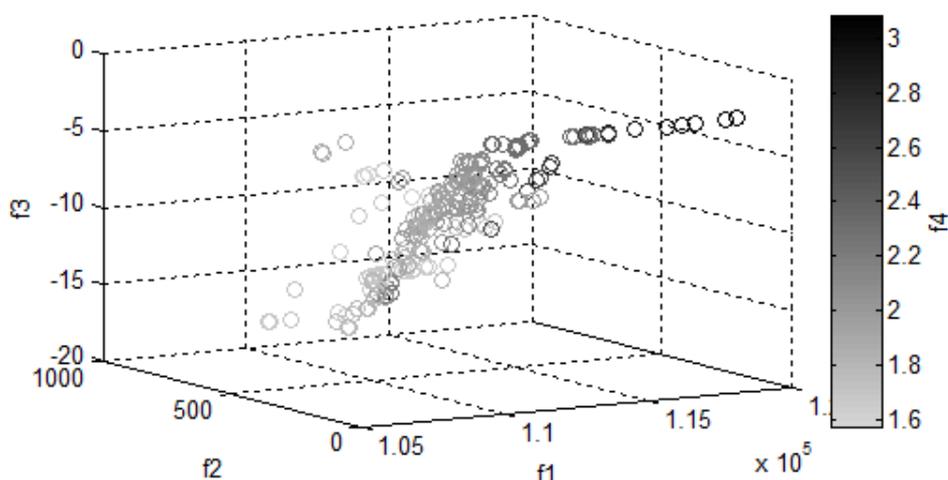
Fig. 11.   3D colored plot method for RP2 results.

5   Conclusions

Application of SOM to visualize multi-dimensional sets of Pareto-optimal solutions is an efficient method that can help the dialogue between analyst and the decision maker. It can be used to visualize tradeoffs between objectives and to group similar solutions together in a clear and simple manner. Particular solutions, corresponding to a design variables sets or single objective optimal solutions can be highlighted.

In case of large number of solutions SOM of bigger size can be utilized and more clusters can be chosen for the purpose of data mining. SOM generation and training is quick and doesn't require much computational time or memory compared to optimization tasks presented in the chapter 4.

Multiobjective optimization is a developing area and efficient methods to solve problems concerning four and more objectives are an object of author's further research.

Using self-organizing maps to visualize the 4D data overcomes the problems of large number of plots (fig. 11) and ambiguity of display (fig. 12). To extract specific information on the Pareto solutions it is possible to draw additional maps including U-matrix, clustered data or a display of chosen design variables sets.
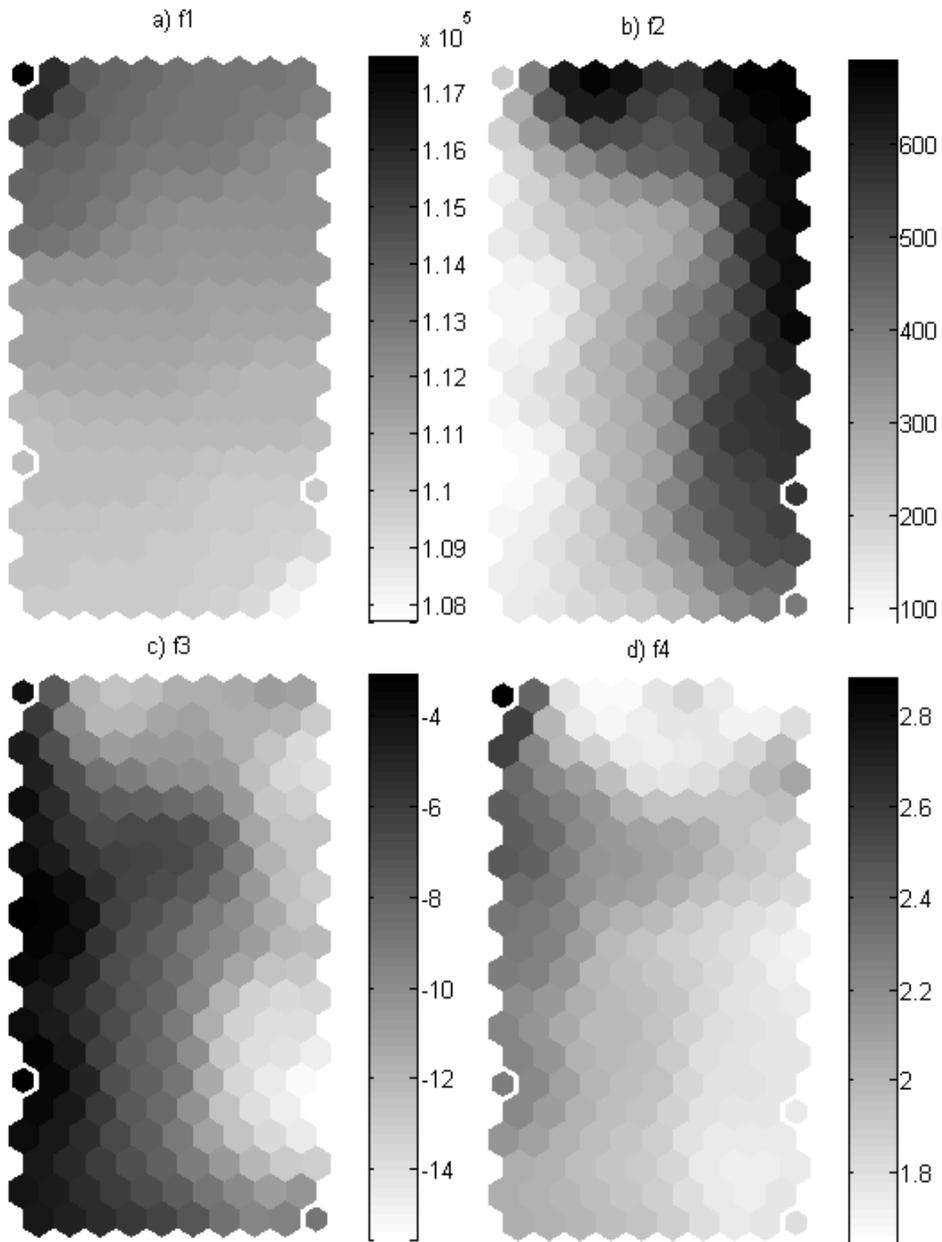
Fig. 12.   SOM presentation of the RP2 results. The BMUs for the best
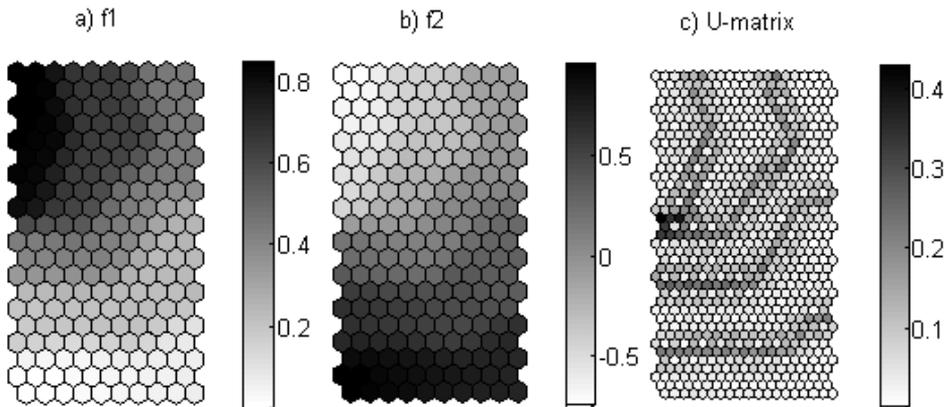solutions regarding each criterion highlighted with white border.

Fig. 13.   SOM presentation of the ZDT3 results.

In this example the SOM was additionally clustered using UPGMA (Unweighted Pair Group Method with Arithmetic Mean), a simple agglomerative (bottom-up) hierarchical clustering method. The cluster number was chosen *a priori* as four. Clusters correspond to continuous intervals of the Pareto fronts and as there are five such intervals and four clusters, two of the intervals are linked in one cluster. There are visible leaps in the objective values and increased U-matrix values. U-matrix can be divided to parts corresponding to continuous intervals of the Pareto front, by the units with increased distance, therefore the analysis of the U-matrix can be utilized to help choose an appropriate number of clusters. In terms of real optimization problems it can be useful to provide the decision maker with the clustered information and a set of design variables for a representative unit in each cluster.

References

[1]   Beluch W., Hatłas M., Multiscale Evolutionary Optimization of Functionally Graded Materials, *Advances in mechanics: theoretical, computational and interdisciplinary issues*, 2016, pp. 83-91

[2]   Branke J., Deb K., Miettinen K., Słowiński R., *Multiobjective Optimization. Interactive and Evolutionary Approaches*, Springer, Berlin, 2008

[3]   Długosz A., Jarosz P., Multiobjective optimization of electrothermal microactuators by means of Immune Game Theory Multiobjective Algorithm, *Advances in mechanics: theoretical, computational and interdisciplinary issues*, 2016, pp. 141-146

[4]   Kohonen T., *Self-Organizing Maps*, Springer, New York, 2001

[5]   Obayashi S., Sasaki D., Visualization and Data Mining of Pareto Solutions Using Self-Organizing Map, *Lecture Notes in Computer Science*, 2003, pp. 796-809

[6]   Zitzler E., Deb K., Thiele L., *Comparison of Multiobjective Evolutionary Algorithms: Empirical Results, Evolutionary Computation*, 2000, pp. 173-195

[7]   http://www.cis.hut.fi/projects/somtoolbox/about, access 18-X-2016

ZASTOSOWANIE SAMOORGANIZUJĄCYCH MAP W CELU WIZUALIZACJI WIELOWYMAROWYCH FRONTÓW PARETO

Streszczenie – Celem pracy jest przedstawienie zastosowania samoorganizujących map Kohonena do wizualizacji. Wielowymiarowe zbiory niezdominowanych rozwiązań testowych funkcji optymalizacji oraz rzeczywistych problemów optymalizacji wielokryterialnej są przedstawione z wykorzystaniem samoorganizujących map oraz tradycyjnych metod dla porównania.

Słowa kluczowe: samoorganizująca mapa, sztuczna sieć neuronowa Kohonena, front pareto, wizualizacja