MONITORING OF DISTILLATION COLUMN OPERATION THROUGH SELF-ORGANIZING MAPS

Y.S. Ng and R. Srinivasan*

Laboratory for Intelligent Applications in Chemical Engineering,
Department of Chemical and Biomolecular Engineering,
National University of Singapore, 10 Kent Ridge Crescent, Singapore 119260
**Corresponding Author: Email: chergs@nus.edu.sg*

Abstract: Monitoring of chemical processes is becoming increasingly difficult as a result of growing complexity and larger scale of operation. In this paper, Kohonen self-organizing maps (SOM) is used to monitor the operation of a lab-scale distillation column and to identify process states. SOM projects high-dimensional data to a lower two dimensional grid maps while preserving the metric relations of the original data. The results from this paper show that using this property of SOM, process monitoring can be performed effectively through observing time series trajectory of process operations on SOM while fulfilling the objective of state identification at the same time. Occurrence of a fault will result in the deviation from the normal operating trajectory. Root cause identification can also be performed through simple visualization of component planes.

Keywords: monitoring, self-organizing systems, fault diagnosis, Neural-networks

1. INTRODUCTION

Incidents occurred in the past have shown the catastrophic consequences of improper monitoring of chemical processes; the resulting effect is not merely the loss of capital from shareholders but also injuries and loss of precious human lives. There is an increasing need for monitoring technique so that abnormal situations can be detected in the shortest amount of time. Over the past few decades, a lot of monitoring techniques have been introduced e.g. enhanced trend analysis, principal-component analysis, artificial neural-networks, etc; some have been implemented in industry successfully (Venkatasubramanian, *et al*., 2003), but what is still lacking in the chemical plant is a powerful visualization technique that help plant personnel to visualize the evolution of a process and detect abnormal situations efficiently. These situations are worsening with the growing scale and complexity of chemical processes; at any instant, the number of variables that plant personnel need to track can easily range in the thousands. Due to the above mentioned reasons, monitoring through conventional control charts has become less effective and there is a need for new and effective methods to help plant personnel to monitor the progression of chemical processes.

In this paper, Self-Organizing Maps (SOM) has been used for the purpose of monitoring of chemical processes. SOM is a powerful projection method which is capable of projecting high dimensional data onto a lower two dimensional grid while preserving the metric relationships of the original data. The effort to perform process monitoring is reduced many fold by SOM as compared to conventional methods since visualization is performed on a much lower two dimensional grid. The organization of this paper is as follows: A brief description of the underlying concepts behind SOM is presented in Section 2 Section 3 presents the review of SOM-based process analysis while Section 4 presents the method of using SOM for process monitoring and state identification. Monitoring and fault detection through SOM for a lab-scale distillation column is presented in Section 5.

2. THE SELF ORGANIZING-MAP

SOM was first proposed by Kohonen (Kohonen, 1981) and has been widely used since then. SOM algorithm employs nonparametric regression technique which involves the fitting of discrete, ordered reference vectors, to the distribution of input feature vectors. The dimensions of the prototype vectors are equal to the dimensions of input vectors. The hexagonal lattice type has been employed as it is more effective for visualization. Consider an input vector, $\vec{x}$, given by

$$\vec{x} = \{x_1, x_2, x_3, ..., x_n\} \qquad (1)$$

$\vec{x}$ will be compared with all reference vectors, say $\vec{m}_i$,

$$\vec{m}_i = \{m_1, m_2, m_3, ..., m_n\} \qquad (2)$$

Each $\vec{m}_i$ also represents a node on a hexagonal output grid. The map $\vec{m}_i$ which gives the smallest Euclidean distance with $\vec{x}$ is defined as the best-matching unit (BMU), represented here as c,

$$c = \arg\min_i \{\| \vec{x} - \vec{m}_i \|\} \qquad (3)$$

During the training of the SOM, the reference weight vector $\vec{m}_i$ will be updated and the topological neighbors around the BMU are updated as well by moving them towards the training sample $\vec{x}$. The SOM learning rule is given by

$$\vec{m}_i(t+1) = \vec{m}_i(t) + a(t) h_{ci}(t)[\vec{x}(t) - \vec{m}_i(t)] \qquad (4)$$

where $h_{ci}(t)$ is the neighborhood function and $a(t)$ is the learning rate factor. It is necessary that h_{ci} reduces to zero as the training proceeds to guarantee convergence. A large value of a is employed initially and usually decreases monotonically with time (Kohonen, 2000). During the training phase, SOM will fold onto the pattern formed by the training data and neighboring units are pulled nearer together because of the neighborhood relations. Thus neighboring units are in a sense, similar to each other. The SOM is usually visualized using the U-matrix (Ultsch, 1990), where distances of each map unit to its corresponding neighbors are calculated and displayed through gray or color-scales. A similar technique can also be applied to visualize the individual components (or variables) so that correlations among components can be investigated and abnormality can be detected at the variable level. The combinatory use of U-matrix and component planes can so fulfill the objectives of monitoring of chemical processes on both a plant wide and variable level.

3. SOM-BASED PROCESS ANALYSIS

SOM has been widely used as a visualization tool for unsupervised learning. (Vesanto, 2002). It has several powerful features which make it popular for the task of process monitoring. SOM implements an ordered dimensionality reduction through the mapping of the input feature vectors while preserving the most crucial topological and metric relationships of the original data, hence producing a similarity map of the input feature vectors The self-organizing maps is readily explainable, simple and easy to visualize (Vesanto, 1999). SOM has found successes in diverse fields of applications. Li Xiao, *et al.* (2003) used component plane presentation of SOM for microarray data analysis of yeast cells and human breast tumors, permitting the visualization of transcriptional changes of tumor sample at a genome scale. López-Rubio, *et al*. (2003) proposed a self-organizing neural model that performs principal components analysis. Kolehmainen, *et al.*(2003) used SOM and Sammon's mapping to identify the growth phases of brewer's yeast. In their work, the exhaust gas from the fermenter was measured with ion mobility spectrometry and a new growth phase for the fermentation process has been identified in addition to the four phases reported in the literature. While the previous applications of SOM are mainly on data clustering, some recent applications of SOM on process monitoring and fault diagnosis are presented in the next subsection.

Deventer, *et al.* (1996) demonstrated how disturbances in a plant for froth flotation of minerals can be visualized with SOM. They track the changes in operating conditions through an on-line computer system utilizing features extracted from froth images and visualize the degree of dispersion of the various input feature vectors through SOM. Chan, *et al*. (2001) presented a modified version of Kohonen SOM called constrained Kohonen networks (CKN) to overcome the problem of redundant sensors by constraining the weight vectors in the parity space. Srinivasan and Gopal (2002) showed how SOM can be used to extract operating information from operating data from a fluidized catalytic cracking unit. Jämsä-Jounela, *et al.* (2003) presented a SOM based fault diagnosis system for a smelter based on heuristic rules. SOM has been used to determine the coefficient for oxygen enrichment and detection of aggregations in various part of the plant after principal component analysis. Abonyi, *et al*., (2003) applied SOM to a polyethylene process for product quality estimation. They developed multiple local linear model for the Philips polyethylene process through piecewise linear regression with SOM. More than 4000 scientific publication on SOM have been written to date (Kohonen, 2001) and the majority of them deal with the classification of input feature vectors and data mining. A comprehensive reference of SOM research has been compiled by Kaski, *et al*., (1998). In this work, we propose a methodology for process monitoring and state identification through the use of SOM.

4. PROPOSED SOM-BASED MONITORING METHODOLOGY

The proposed methodology can be summarized as follows. Measurements of the normal operating process will first be collected and range normalized before being projected onto the U-matrix of SOM. The clusters formed on the U-matrix are interpreted and different states, which correspond to different operating phases is identified. Different clusters will be separated by large distances, modeled in the SOM's U-matrix by a border of darker color, predefined by the user. During the on-line monitoring phase, the trajectory being formed by the real time data will follow the trajectory of the normal operating data and significant deviation from the normal operating trajectory can be regarded as a fault. Duration of operation can be read from the size of the hexagon on the U-matrix, in which the size of the hexagon is directly proportional to the duration of operation. State identification of the process is done through the identification of the location of the projected operating data (data hit) on SOM. The proposed methodology is illustrated next through a case study.

5. MONITORING OF DISTILLATION-COLUMN OPERATION

The proposed method has been tested on an Armfield lab-scale distillation column. The schematic of the distillation column is shown in Figure 1. The distillation column is of 2 meters height and 20cm width and has 10 trays, where the feed enters at tray 4. The system is well integrated with a control console for controllers setting and data recording. Cold startup of the distillation column with ethanol-water at 30% v/v as bottoms is performed following the standard operating procedure (SOP) as shown in Table 1. The feed passes through a heat exchanger before being fed to the column. 19 variables --- all tray temperatures, reboiler and condenser temperature, reflux ratio, top and bottom column temperature, feed pump power, reboiler heat duty, cooling water inlet and outlet temperature - are measured at 10-second interval. The startup normally takes two hours and different faults such as sensor fault, failure to open pump, too high a reflux ratio etc., can be introduced at different states of operation.

The constructed SOM for a normal startup of distillation column is shown in Figure 2. The fully trained SOM consists of 30 x 8 map units and its corresponding bar-plane is shown in Figure 3. Each unit of the bar-plane is a representation of the magnitude of the 19 variables of the distillation column after normalization. Two observations can be seen from Figure 2. Firstly, the startup process has been observed to follow a trajectory on the U-matrix. Each label represents the corresponding operating time (x 10 seconds). Secondly, it has been observed that any significant deviation from the normal operating trajectory is an indication of a faulty operation.

By comparing Figure 2 and Figure 3, even without any prior knowledge on the process, one can deduce that during the normal run, most of the variables are increasing over time as represented by the trajectory on the U-matrix as indicated in Figure 2 Different clusters which correspond to different operating phases can be defined for each U-matrix or component plane constructed. As can be seen from the U-matrix on Figure 2, four clusters of operations can be identified for the startup of distillation column, each representing different phases of the start-up of distillation column. The border of each cluster is separated by map units of darker color

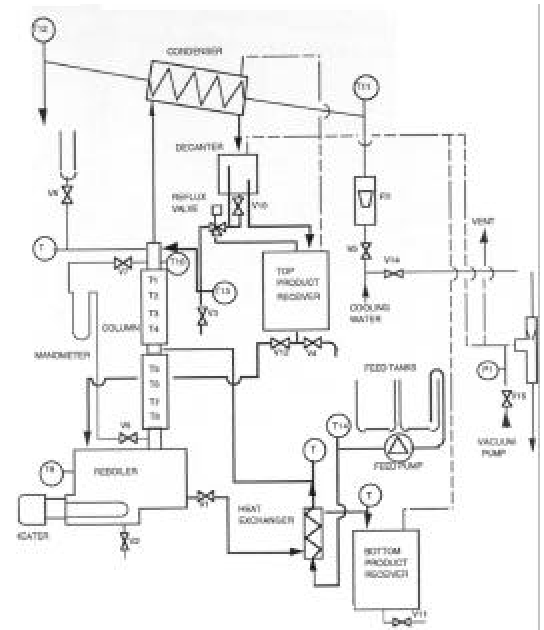

Figure 1. Schematic of the lab scale distillation column.

Table 1 Standard operating procedures (SOP) for the startup of distillation column

	Distillation column startup SOP
1.	Set all controllers to manual
2.	Fill reboiler with liquid bottom product
3.	Open reflux valve and operate the column on full reflux
4.	Establish cooling water flow to condenser
5.	Start the reboiler heating coil power
6.	Wait for all of the temperatures to stabilize
7.	Start feed pump
8.	Activate reflux control and set reflux ratio
9.	Open bottom valve to collect product
10.	Wait for all the temperatures to stabilize

which can be interpreted as a boundary marker between different clusters. The color code on the right tells the exact distance among neighboring units, which were trained using Gaussian neighborhood function during the batch training phase. Operating trajectory from the starting of the heating process till time 3630s can be regarded as a separate state from the rest of the data points – reboiler heating phase. Time series 3630s to 3820s is the boiling phase where evaporation of the bottom liquid takes place and there is an abrupt rise of temperature in the distillation tower. Time series 3840s to 4800s correspond to the full reflux heating phase before the feed was introduced and thus causing a switch in operating phase as shown from time trajectory from 4930s to 6340s, which is representing the steady state operating phase of the tower.

The use of SOM to monitor faulty operations is shown in Figure 4 and Figure 5. Scenario 1: Figure 4

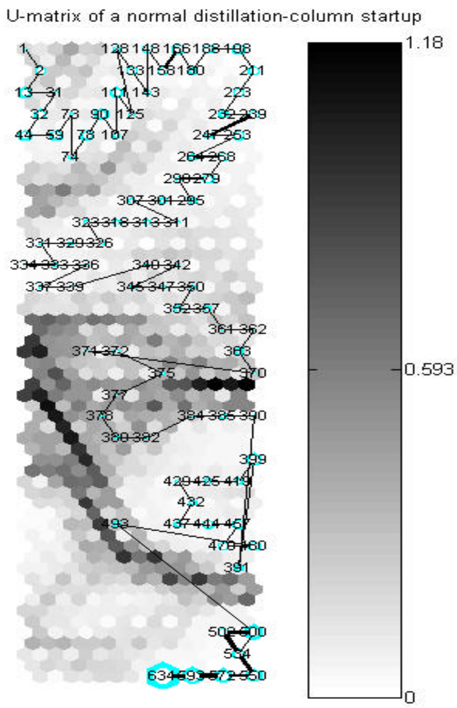

Figure 2. Trajectory of a normal run on U-matrix. Labels indicate the corresponding operating time (x10 seconds).

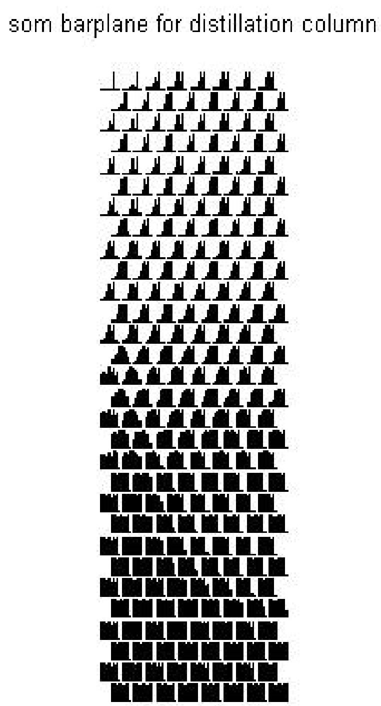

Figure 3. Bar-plane for the trained self-organizing maps.

shows the trajectory of a faulty run. Two trajectories can be observed from Figure 4, the light solid trajectory is the trajectory of a normal run whereas the dark solid line is the trajectory of the faulty run. The first fault occurred from 2650s (shown as 265 in Figure 4), where operator failed to turn on the feed pump and thus causing deviation from normal run. The situation is later rectified by the operator and the recovery can be seen on the U-matrix where the

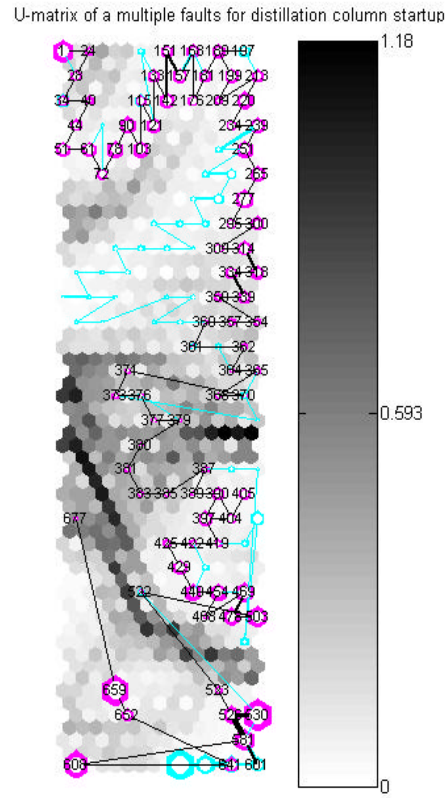

Figure 4. Trajectory of Scenario 1 (Dark solid line represents the faulty run while light solid line represents the normal run).

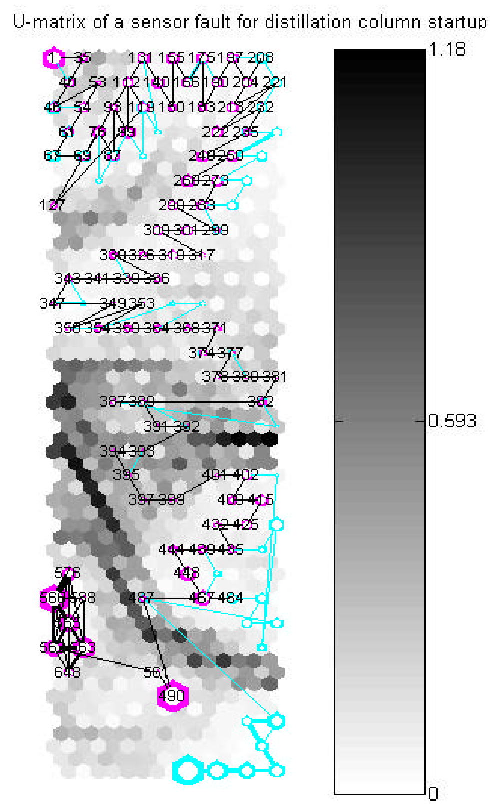

Figure 5. Trajectory of Scenario 2 (Dark solid line represents the faulty run while light solid line represents the normal run).

trajectory returns to follow the trajectory of a normal run at 3570s. A second fault occurred at time 6520s, when a feed flow of 1.5 times the normal value was introduced. The corresponding trajectory was found to deviate from normal operating trajectory from 6520s to 6770s. Secnario 2: The second scenario consists of high reflux ratio and temperature sensor fault as shown in Figure 5. At 4900s, the operator accidentally set the reflux ratio to 2 times higher as compared to the normal run and the column is operated in this condition for a long period of time (as seen from the size of the data hits), before a temperature sensor break down at 5600s.

While U-matrix is a simple 2-dimensional plot for operators to perform process monitoring, it does not contain sufficient information for diagnosing the fault. The component planes of the corresponding run have to be investigated in order to locate the fault. For the second scenario (Figure 5), variable 6 can be identified to be the main cause of deviation as it moves into a cluster of lower value (see Figure 6). The corresponding operating chart for variable 6 is given in Figure 7. Similar analysis can also be applied toother scenarios for fault identification (not shown in this paper).

The knowledge inferred from the normal operating run in Figure 2 can also be applied to other cases such as the one in Figure 4 and Figure 5. The operating state of a different run can be directly obtained from the location of the data point on the self-organizing maps. As an illustrative example, for the operation of distillation column shown in Figure 4, even though there are faults in the system, the process of state identification can still be done fairly

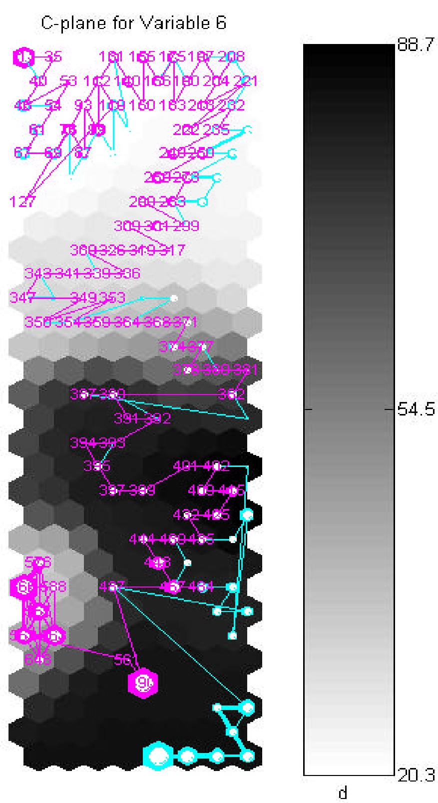

Figure 6. Component plane of variable 6 (5th tray temperature sensor).

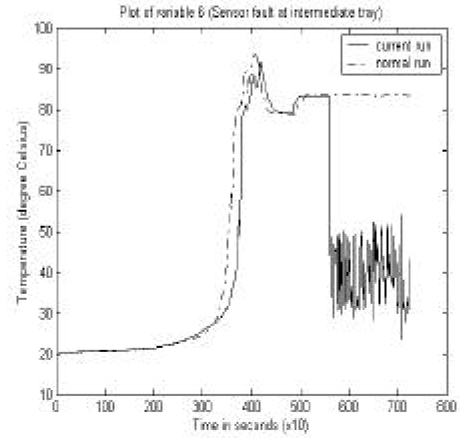

Figure 7. Time variations of variable 6 (5th tray temperature sensor).

successfully. Reboiler heating phase can be observed from time 0 – 3650s. Evaporation of bottoms has been observed to take place from 3680s – 3870s while full reflux operation prior to feed is from 3890s – 5220s, and lastly the steady state operation can be seen to take place from 5230s – 6770s. Similar interpretations can be applied to other case studies as well.

6. CONCLUSIONSAND DISCUSSION

This paper illustrates how Kohonen SOM can be used for monitoring of process dynamic transitions, in this case the operation of a distillation column. High dimensional data has been projected onto a lower two dimensional grid maps, consequently the task of process monitoring has been simplified as compared to conventional control charts approach. One only needs to keep track of the trajectory in order to deduce whether the operation is normal or abnormal. A faulty run will result in deviation from normal trajectory, and component planes can then be analyzed to locate faulty variables for root cause identification. This paper has also shown that SOM can be used for state-identification. The exact phase of different run at any instant can be inferred from the location of the data hit on the self-organizing maps. Our current research is targeted at automating the SOM-based approach to improve the efficiency of monitoring. Two methods that use the results from SOM for automated abnormality detection and diagnosis have been formulated. Firstly, statistical analysis has been used to generate variable specific residual by using SOM as local model during the transition. Secondly, a pattern recognition approach is being explored for diagnosis and fault isolation purposes. By using a fault database and machine learning approaches, faults can be identified in a short amount of time. A robust formulation of normal trajectories is also being investigated to accommodate multiple normal operating conditions.

REFERENCES

Abonyi, J., Nemeth, S., Vincze., Arva, P., (2003). Process analysis and product quality estimation by self-organizing maps with an application to polyethylene production. Computers in Industry 52. 221-234.

Chan, C.W., Jin, H., Cheung, K.C., Zhang, H.Y., (2001). Fault detection of systems with redundant sensors using constrained Kohonen networks. Auomatica 37. 1671-1676.

Deventer, J.S.J.V., Moolman, D.W., Aldrich, C., (1996). Visualization of plant disturbances using self-organizing maps, Computers Chem. Engng Vol.20, pp.S1095-S1100.

Jämsä-Jounela, S.L., Vermasvuori, M., Endén, P., Haavisto,S., (2003). A process monitoring system based on the Kohonen self-organizing maps. Control Engineering Pract ice 11. 83-92.

Kaski, S., Kangas, J., Kohonen, T., (1998). Bibliography of self-organizing map (SOM) papers: 1981-1997, Neural Computing Surveys, 1:102-350.

Kohonen, T., (1981). Automatic formation of topological maps of patterns in a self-organizing system, Proc. 2SCIA, Scand. Conf. on Image Aanlysis, pp 214-220, Helsinki, Finland.

Kohonen, T., (2000). Self-Organizing Maps, Springer Series in Information Sciences, Springer, Berlin, Germany.

Kohonen, T., (2001). Seiffert, U., Jain, L. C., Editors, Self-Organizing Neural Networks: Recent Advances and Applications, Phisica-Verlag Heidelberg, Springer-Verlag, Germany.

Kolehmainen, M., Rönkkö, P., Raatikainen, O., (2003). Monitoring of yeast fermentation by ion mobility spectrometry measurement and data visualization with self-organizing maps. Analytica Chimica Acta 484. 93-100.

López-Rubio,E., Muñoz-Pérez, J., Gómez-Ruiz, J.A., (2003). A principle component analysis self-organizing map. Neural Networks. In-press.

Srinivasan, R., Gopal, S., (2002). Extracting information from high-dimensional operations data using visualization techniques, AIChe meeting, Indianapolis, #271c.

Ultsch,A., Siemon, H.P., (1990). Kohonen's self organizing feature maps for exploratory data analysis, Proceedings of International Neural Network Conference (INNC'90), Kluwer academic Publishers, Dordrecht, pp. 305-308.

Venkatasubramanian, V., Rengaswamy, R., Yin, K. & Kavuri, S.N., (2003). A review of fault detection and diagnosis Part I: Quantitative model-based methods, Computers and Chemical Engineering 27, 293-311.

Vesanto, J., (1999). SOM-based data visualization methods, Intelligent Data Analysis 3, 111-126.

Vesanto, J., (2002). Data Exploration Process Based on the Self-Organizing Map, phD thesis, Helsinki University of Technology, Acta Polytechnica Scandincavica: Mathematics and Computing Series No. 115, Department of Computer Science and Engineering, FIN-02015 HUT, Finland.

Vesanto, J., Himberg., Alhoniemi, E., Parhankangas, J., (2000). SOM Toolbox for Matlab 5. SOM toolbox team, Helsinki University of Technology, Finland.

Xiao, L., Wang, K., Teng, Y., Zhang, J., (2003). Component plane presentation integrated self-organizing map for microarray data analysis. FEBS Letters 538. 117-124.