International Journal of Computers, Communications & Control
Vol. II (2007), No. 2, pp. 143-148

Self-Organizing Maps for Analysis of Expandable Polystyrene Batch Process

Mikko Heikkinen, Ville Nurminen, Yrjö Hiltunen

Abstract: Self-organizing maps (SOM) have been successfully applied in many fields of research. In this paper, we demonstrate the use of SOM-based method for the analysis of Expandable PolyStyrene (EPS) batch process. To this end, a data set of EPS-batch process was used for training a SOM. Reference vectors of the SOM were then classified by K-means algorithm into six clusters, which represent product types of the process. This SOM could also be used for estimating the optimal amounts of the stabilisation agent. The results of a validation data set showed a good agreement between the actual and estimated amounts of the stabilisation agent. Based on this model a Web application was made for test use at the plant. The results indicate that the SOM method can also be efficiently applied to the analysis of the batch process.
Keywords: Neural networks, self-organizing maps, process control, batch process

1 Introduction

Batch processes are typically based on predefined process recipes. If process circumstances, chemicals and recipes are constant, the product should basically be always the same. A batch process is also commonly used for producing Expandable PolyStyrene (EPS). However, in practice this polymerisation reaction is a very sensitive process and numerous variables affect it, which makes the process difficult to control. The EPS production has to be able to follow fast the aims and quality requirements of the market, which causes additional demands on the process control.

Archived process data is an important resource for the knowledge management of the process and it can be used for the optimization and improvement of productivity. Recent applications have demonstrated that artificial neural networks can provide an efficient and highly automated method for modelling industrial data [1], [2]. In particular, studies, which use standardized protocols, are most likely to benefit from automated ANN analysis [1], [2]. Self-organizing maps [1], [3]-[5] have also been successfully applied in many areas of research and are thus a tool for process optimization. The SOM method offers an efficient means of handling complex multidimensional data, which is typically the situation in industrial applications. In addition, the SOM method is robust for missing values of data.

Here, we apply self-organizing maps to the analysis of an EPS-batch process. The optimal amounts of the stabilisation agent can be estimated using the SOM model. In the study we have also included some features of the supervised approach in the designed unsupervised method.

2 Methods

2.1 The process and the data

The studied process was a typical suspension polymerisation batch process, which is commonly used for producing EPS (Expandable PolyStyrene). The polymerisation stage is executed in a pressure-temperature range below the boiling point of styrene-water suspension system. After the polymerisation stage the process continues into the impregnation stage, where the blowing agent is impregnated into the beads. The impregnation stage assumed to be negligible in the means of bead size distribution.

The biggest challenge in the suspension polymerisation process is to achieve the required bead size distribution. It is common knowledge that the basic variables in the term of the bead size are the mixing

properties and the amount and quality of the suspension stabilizers. However the suspension polymerisation of styrene is a very sensitive process and numerous variables affect it. Most of these variables cannot be measured or followed by in a reasonable way. For example to analyse all impurities from all raw materials is too heavy a task for any industrial laboratory. Some variables are quite easily measurable, but have not been traced due to the assumption that they would not have a significant contribution to the process. To be able to model the process the studied system required elimination of the variables, which were assumed to be inessential.

The data for the model were divided into three groups: recipe, results and process parameters. Process parameters, such as actual reactor temperature, were measured and stored automatically from each batch every one minute. Process parameter data was not used for modelling in the first part of this study and the target is to add it later to the model. Table I shows the recipe and result variables.

The data contained 15 production campaigns and about 450 batches in 4 reactors. The data had to be divided into two separate groups due to the process changes in the stabilisation system. These changes limited the size of useful data to 251 batches, which is a low amount for accurate modelling. Fortunately more production campaigns will be executed almost every month and more data will be available for further studies.

2.2 Computational methods

Self-organizing maps

Self-organizing maps (SOMs) are an artificial neural network methodology, which can transform an n-dimensional input vector into a one- or two-dimensional discrete map. The input vectors, which have common features, are projected to the same area of the map e.g. (in this case described as ŞneuronsŤ). Each neuron is associated with an n-dimensional reference vector, which provides a link between the output and input spaces. During learning, the input data vector is mapped onto a particular neuron (best matching unit, BMU) based on the minimal n-dimensional distance between the input vector and the reference vectors of the neurons. Then the reference vectors of the activated neurons are updated. When the trained map is applied, the best matching units are calculated using these reference vectors. In this unsupervised methodology, the SOM can be constructed without previous a priori knowledge [1].

The data were coded into 11 inputs for the SOM. All input values were variance scaled. The SOM having 676 neurons in a 26x26 hexagonal arrangement was constructed. The linear initialization and

Table 1: Recipe and result variables of the particular StyroChem suspension polymerisation process.

Recipe Variables	*Unit*	*Result Variables*	*Unit*
Amount of stabilisation agent	% from MS	Under sized	%
Mixing speed	rpm	Product A	%
Polymerisation temperature	°C	Product B	%
Reactivity	min	Product C	%
Amount of styrene	kg	Product D	%
Additional stabilisation	kg	Product E	%
Product type	-	Product F	%
Batch number in campaign	-	Over sized	%
Reactor	-	Mean particle size	mm
		Delta	mm
		Narrowness	-

batch training algorithms were used in the training of the map. A Gaussian function was used as the neighbourhood function. The map was taught with 10 epochs and the initial neighbourhood had the value of 6. The SOM Toolbox [7] was used in the analysis under a Matlab-software platform (Mathworks, Natick, MA, USA).

The data set of an EPS-batch process (n= 251 batches) was divided into two subsets. The first subset (the first 190 batches) was the training set, which was used for training the map. The other subset (the last 61 batches) was the test set. Variables of each batch and amounts of the stabilisation agent were used as an input for the SOM in the training phase. However, the amounts of the stabilisation agent were missing values in the test phase, i.e. best matching neurons for the test set were sought using only those other batch variables specified above. Estimated values for the amounts of the stabilisation agent were obtained from the reference vectors of neurons.

K-means method

The K-means algorithm was applied to the clustering of the map. The K-means method is a well-known non-hierarchical cluster algorithm [8]. The basic version begins by randomly picking K cluster centers, assigning each point to the cluster whose mean is closest in a Euclidean distances sense, then computing the mean vectors of the points assigned to each cluster, and using these as new centers in an iterative approach.

3 Results and discussion

The map was obtained by training a self-organizing network with the training set of an EPS-batch process. The map and the six clusters calculated by the K-means method are shown in Figure 1. These clusters represent different types of the products. The brief descriptions of the clusters are also illustrated in Figure 1.

The method was validated by using the test set, i.e. the last 61 data vectors were in test. The results, when the SOM was trained to estimate the amounts of the stabilisation agent, are illustrated in Fig. 2. The correlation coefficient between the actual amount of the stabilisation agent and the estimated one was 0.851 (Fig 2 a). In Fig. 2 b) the actual and estimated amounts of the stabilisation agent are shown as a function of the batch number.

Table 2: Variables used for modelling.

Variables
Under sized
Product A
Product B
Product C
Product D
Product E
Mean particle size
Batch number in campaign
Amount of stabilisation agent in previous batch
Amount of stabilisation agent

A Web application based on this SOM model was made for estimating the optimal amount of stabilisation agent in practice at the plant. Figure 3 illustrates the interface of this application. The same

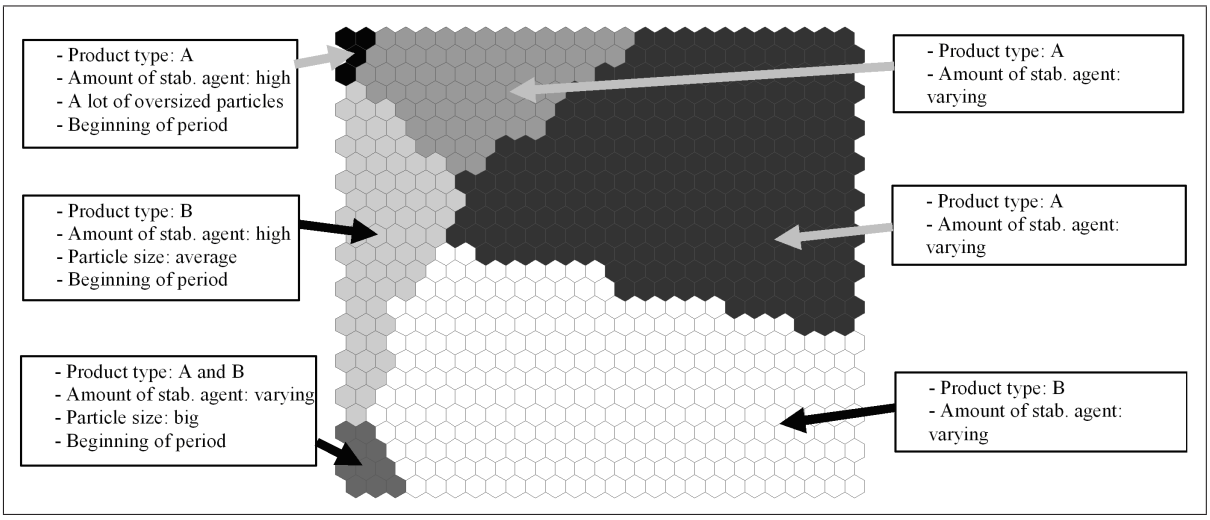

Figure 1: SOM using the data of an EPS-batch process. The background colours visualize the six clusters of the map. Short descriptions for each cluster are also shown.

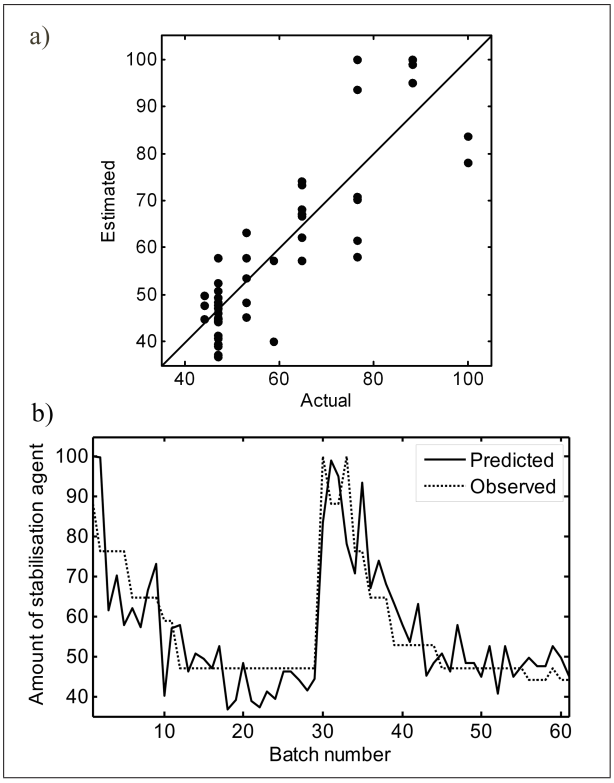

Figure 2: a) The correlation between the actual amount of the stabilisation agent and the one estimated by the SOM analysis. The results are from the test set (the last 61 batches) and the value for the correlation coefficient is 0.851. b) The actual and estimated amounts of the stabilisation agent as a function of the batch number.

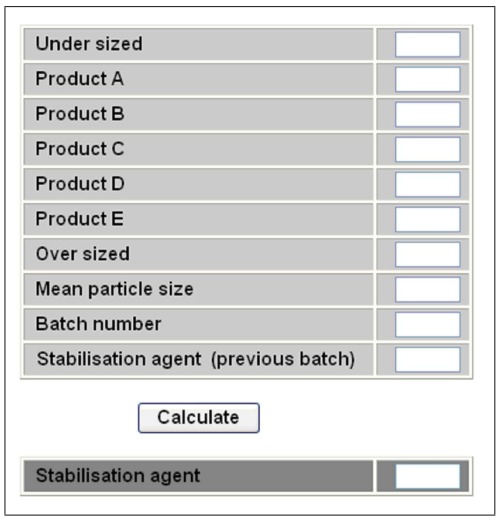

Figure 3: The interface of the Web application.

variables, which are shown in Table 2, have been used in this application. A user gives as many as possible out of the ten upmost variables and the application calculates the lowest one. The first seven variables define the target amounts of the product.

Our earlier results showed that the SOM method could be successfully applied to process state monitoring and optimization of NOx emissions in the case of a continuous process [6]. The results presented here illustrate also the advantages of using SOM method in the analysis of a batch process. Figure 3 shows that the method can be used for estimation of optimal parameters of the process. It seems to reduce varieties of the process and so helps to get better products. Because our SOM method is also non-sensitive for the presence of missing values, it is feasible in the analysis of industrial data. Furthermore, SOM analysis does not require extensive knowledge of neural nets and it can easily be included in any kind of software. An attractive property of SOM is also that it can be retrained, if new product types of the process are to be analysed.

4 Conclusion

The SOM analysis provides an efficient and automated method for data analysis in the process industry. The present study shows that this kind of data-driven approach is a fruitful way of analysing a batch process.

5 Acknowledgement

This research was supported by StyroChem Ltd.

References

[1] T. Kohonen, *Self-organizing Maps*, Springer-Verlag, Berlin Heidelberg New York, 2001.

[2] S. Haykin, *Neural Networks: A Comprehensive Foundation*, Upper Saddle River, NJ: Prentice Hall, 1999.

[3] J. Kaartinen, Y. Hiltunen, P. T. Kovanen, M. Ala-Korpela, Classification of Human Blood Plasma Lipid Abnormalities by 1H Magnetic Resonance Spectroscopy and Self-Organizing Maps, *NMR Biomed*, Vol. 11, pp. 168-176, 1998.

[4] M. T. Hyvönen, Y. Hiltunen, W. El-Deredy, T. Ojala, J. Vaara, P. T. Kovanen, M. Ala-Korpela, Application of Self-Organizing Maps in Conformational Analysis of Lipids, *Journal of the American Chemical Society*, Vol. 123, pp. 810-816, 2001.

[5] M. Heikkinen, M. Kolehmainen, Y. Hiltunen, Classification of process phases using Self-Organizing Maps and SammonŠs mapping for investigating activated sludge treatment plant in a pulp mill, *Proceedings of the Fourth European Symposium on Intelligent Technologies and their implementation on Smart Adaptive Systems*, pp. 281-297, 2004.

[6] M. Heikkinen, A. Kettunen, E. Niemitalo, R. Kuivalainen, Y. Hiltunen, SOM-based method for process state monitoring and optimization in fluidized bed energy plant, ICANN 2005, *Lecture Notes in Computer Science 3696*, Eds. W. Duch, J. Kacprzyk, E. Oja, S. Zadrozny, Springer-Verlag Berlin Heidelberg, pp. 409-414, 2005.

[7] Homepage of SOM toolbox, Helsinki University of Technology, Laboratory of Computer and Information Science (CIS), http://www.cis.hut.fi/projects/somtoolbox/.

[8] J. MacQueen, Some methods for classification and analysis of multivariate observations, *In Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, Vol. I: Statistics, University of California Press, Berkeley and Los Angeles, pp. 281-297, 1967.

Mikko Heikkinen
University of Kuopio, Department of Environmental Sciences
P.O. Box 1627, FIN - 70211 Kuopio, Finland
E-mail: Mikko.Heikkinen@uku.fi

Ville Nurminen
StyroChem Ltd
P.O. Box 360, FIN - 06101 Porvoo, Finland
E-mail: Ville.Nurminen@styrochem.com

Yrjö Hiltunen
University of Kuopio, Department of Environmental Sciences
P.O. Box 1627, FIN - 70211 Kuopio, Finland
E-mail: Yrjo.Hiltunen@uku.fi

Received: March 6, 2007