Available online at www.sciencedirect.com

ScienceDirect

Procedia CIRP 41 (2016) 448 – 453

www.elsevier.com/locate/procedia

48th CIRP Conference on MANUFACTURING SYSTEMS - CIRP CMS 2015

Design performance analysis of a Self-Organizing Map for statistical monitoring of distribution-free data streams

Massimiliano Di Mauro[a], Matteo F. Maggioni[a], Marco Grasso[a,b,*], Bianca Maria Colosimo[a]

[a]*Dipartimento di Meccanica, Politecnico di Milano, Via La Masa, Milano, 20156, Italy*
[b]*Laboratorio MUSP, Via Tirotti 9, Piacenza, 29122, Italy*

* Corresponding author. Tel.: +39-02-239-98500; fax: +39-02-239-98202. *E-mail address:* marco.grasso@musp.it

Abstract

In industrial applications, the continuously growing development of multi-sensor approaches, together with the trend of creating data-rich environments, are straining the effectiveness of the traditional Statistical Process Control (SPC) tools. Industrial data streams frequently violate the statistical assumptions on which SPC tools are based, presenting non-normal or even mixture distributions, strong autocorrelation and complex noise patterns.
To tackle these challenges, novel nonparametric approaches are required. Machine learning techniques are suitable to deal with distributional assumption violations and to cope with complex data patterns. Recent studies showed that those methods can be used in quality control problems by exploiting only in-control data for training (such a learning paradigm is also known as "one-class-classification").
In recent studies, the use of distribution-free multivariate SPC methods was proposed, based on unsupervised statistical learning tools, pointing out the difficulty of defining suitable control regions for non-normal data. In this paper, a Self-Organizing Map (SOM) based monitoring approach is presented. The SOM is an automatic data-analysis method, widely applied in recent works to clustering and data exploration problems. A very interesting feature of this method consists of its capability of providing a computationally efficient way to estimate a data-adaptive control region, even in the presence of high dimensional problems. Nevertheless, very few authors adopted the SOM in an SPC monitoring strategy. The aim of this work is to exploit the SOM network architecture, and proposing a network design approach that suites the SPC needs. A comparison study is presented, in which the process monitoring performances are compared against literature benchmark methods. The comparison framework is based on both simulated data and real data from a roll grinding application.

Peer-review under responsibility of the scientific committee of 48th CIRP Conference on MANUFACTURING SYSTEMS - CIRP CMS 2015

Keywords: Statistical process control; Multivariate distributions; Machine learning; Self-organizing map

1. Introduction

In the mechanical industry there is an increasing interest in the quality monitoring of manufacturing processes based on signals acquired from one or more sensors during the process itself. These machines, equipped with several sources of information, represent a real data-rich environment with enormous streams of data that are available and exploitable for many purposes (e.g., monitoring, diagnostics, predictive maintenance, etc.).

In industrial applications a common practice is to extract synthetic features from one or multiple sensors to characterize the stability of the ongoing process. This practice is known in the literature as index-based monitoring, which consists in designing and implementing control charts to monitor performance of the manufacturing process based on extracted indexes. Despite of the various advantages provided by in-process monitoring tools [1-3], the conventional SPC assumptions relative to the underlying data distribution may not be appropriate for the design of signal-based control charts.

On the other hand, nonparametric and machine learning techniques have proved to be suitable to deal with distributional assumption violations and to cope with complex data patterns. One interesting feature of this category of methods consists of their distribution-free properties, and hence, they provide flexible solutions to extend the application field of signal-based

Peer-review under responsibility of the scientific committee of 48th CIRP Conference on MANUFACTURING SYSTEMS - CIRP CMS 2015
doi:10.1016/j.procir.2015.12.054

SPC. Recent studies showed that those methods can be used in quality control problems by exploiting only in-control data for training, by allowing the implementation of traditional classification techniques in the SPC frame in which no information on the nature of possible departures from the natural condition is available. This kind of approach goes under the category of methods that is known as "one class classification" or "novelty detection" [4].

In this work, we present the Self-Organizing Map (SOM), which is a Machine Learning tool originally designed and applied for unsupervised clustering and data exploration problems. A very interesting feature of this method consists of its capability of providing a computationally efficient way to estimate a data-adaptive control region making it a promising tool to deal with one-class classification problems.

As very few authors adopted the SOM in an SPC monitoring strategy, the aim of this work is to exploit the SOM network architecture and compare it on real industrial data against other benchmark techniques (e.g., the fuzzy-ART-based scheme proposed by Pacella and Semeraro [5]).

The performances of the method are evaluated using Monte Carlo simulations in the presence of mixture distributions (a.k.a., as multimode data, [6]) and a real dataset acquired during roll grinding operations.

The paper is organized as follows: Section 2 presents the industrial case study to motivate the need for nonparametric methods; Section 3 presents the framework of the Self-Organizing Map monitoring approach; Section 4 provides a performance comparison analysis based on a real case study in transverse roll grinding; Section 5 concludes the paper.

2. Distribution-free data streams: a real case study

The growing sensor technology together with increasing computational capabilities have enabled the development of industrial quality monitoring tools based on the real-time analysis of different signals acquired during production process. The signal-based SPC framework is based on monitoring variables, which usually represent heterogeneous quantities that come from one or multiple sensors. For this reason an information synthesis step aimed at extracting a reduced set of variables from raw signals is usually performed. These processed variables allow characterizing the ongoing process and detecting possible shifts from an in-control state.

The pre-processing step of raw signals is usually done by time-domain, frequency-domain or more complex kinds of analysis. The assumption of multi-normality is frequently violated in practice and data transformation to normality may be a very difficult task [7].

In addition to non-normality, discrete manufacturing processes may exhibit a "multimode" nature that yields clustered data clouds within the space spanned by the monitored variables, under in-control conditions. This leads to a challenging violation of common distributional assumptions (e.g. Fig 1). It is worth to notice that, when multimode (clustered) data refer to in-control conditions, one may be interested in estimating a control region that globally adapts to

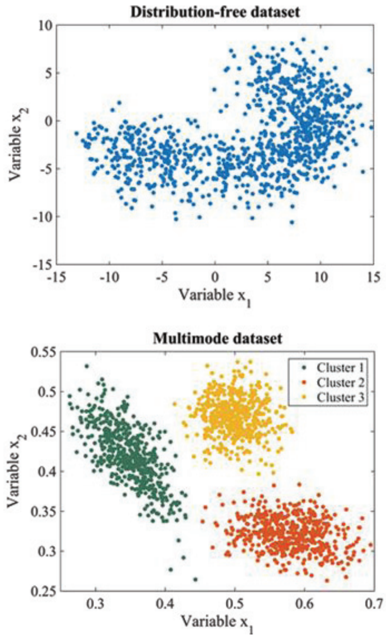

Fig. 1. Examples of datasets where traditional SPC assumptions are violated.

the clustered pattern. Because of this, the one-class-classification paradigm is still applicable, but the applied method must cope with the multimode nature of training data.

An industrial example in which clustered signal data are acquired and distributional assumptions are frequently violated is the case study of a transverse roll grinding operation, where the information coming from two accelerometer sensors installed on the machine are synthesized and used to monitor the stability of the cutting process.

The product of the presented case-study consists of large cylindrical rolls for metal sheets milling operations. It is well known that process vibrations are one of the most critical issues in grinding processes, which may cause undesired undulations on the manufactured surface [8]. The waves generated on the workpiece surface, called chatter marks, are created by the relative vibration between the grinding wheel and the workpiece, resulting in a depth-of-cut modification after one workpiece revolution. The phase shift between the surface waves and the current relative vibration makes the process unstable when the chattering condition is reached.

Fig. 2 (left panel) shows the surface of the grounded roll in absence of chatter-marks, whereas Fig. 2 (right panel) shows

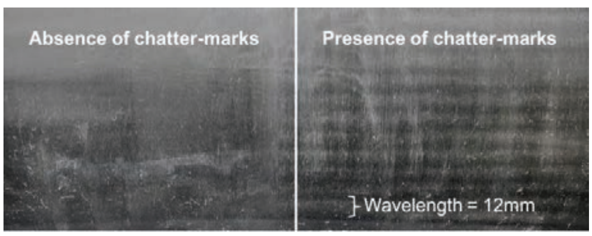

Fig. 2. Surface quality of the grinded roll.

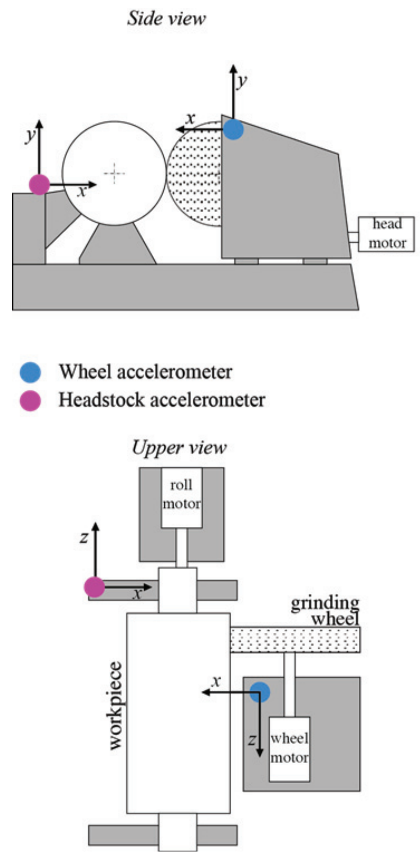

Fig. 3. Qualitative scheme of the machine tool used for the experiments.

the wavelength of the chatter-marks distributed on the surface of the roll. An experimental campaign was performed to collect real data during cylindrical grinding processes under stable and unstable cutting conditions. The workpiece used for the experiments was a special alloyed steel roll for hot rolling, having an initial diameter of 500 mm and an axial length of 1700 mm. The machine was equipped with a resin bonded wheel with a diameter of 790 mm, and a width of 70 mm.
A qualitative scheme of the machine tool used for the experiments is shown in Fig. 3.

Three tri-axial accelerometer sensors were mounted on the wheel head, tailstock and on the headstock, (as shown in Fig. 3). The acceleration signals along the x-axis were acquired with a sampling rate of 2 kHz and segmented into sliding time windows of 1 second in duration. The vibration signal within the ith time window were processed online to compute the root mean square indexes. The result of this pre-processing step is a trivariate quality characteristic $\{x_i \in R^2,\ i = 1, 2, \ldots\}$, where $x_i = rms_{signal\ i}$.

We are focusing on this specific case study because this kind of process involves more grinding cycles, each one consisting of multiples runs performed with different cutting parameters that may vary within given ranges. This situation yields to a multimode process in which the in-control distribution of the monitored indices is characterized by sequentially changing distributions, which correspond to different combinations of cutting parameters.

Fig. 7 shows the multimode distribution of $x_i = \left[rms_{signal1,i},\ rms_{signal2,i}\ rms_{signal3,i} \right]$ under in control conditions for different combinations of cutting parameters. Clustered data represent the natural pattern that characterizes the IC condition as a consequence of the multimode behavior of the process. The natural process variability leads to switching between one mode and another causing shift, which should not be signaled by an appropriately designed control chart. On the other hand, the out-of-control state characterized by the chattering condition, needs to be quickly detected and suppressed [8] in order to avoid undesired undulations of both the workpiece and the wheel, and to prevent the execution of extra grinding cycles to cope with those chatter marks.

The case study of the transverse roll grinding regards one of the most challenging violations of the traditional statistical process control distributional assumptions, which motivates the use of more sophisticated techniques. For this reason we discuss the applicability of the SOM based control chart approach in this real industrial scenario, comparing its performances against the different methods mentioned in this study.

3. Process monitoring via Self-Organizing map

3.1. The Self-Organizing map

The SOM has been used for visualization of correlation patterns, clustering data, monitoring of operation state, and as a novelty detection tool [9].

Thanks to its ability to automatically detect features in the dataset, it has a clear advantage compared to other Machine Learning techniques based on supervised learning, which require target values to be known.

For these reason, the SOM not only has been widely applied to the visualization of high-dimensional data [9] but has also been successfully used in various engineering applications [10] covering areas like pattern recognition, image analysis, process monitoring and control, and fault diagnosis.

The SOM algorithm performs a topology preserving mapping of the input data from its high-dimensional data space onto a two-dimensional grid. By doing this, the relative distances between data points are preserved and a roughly approximation of the probability density function of the data can be estimated.
The detection of out-of-control departures from natural process conditions can be implemented based on the so-called quantization error [9]. The larger is the quantization error, the larger is the expected departure from the in-control state. The methodology is briefly described in the following sub-sections.

3.2. Learning framework

During the iterative training procedure, the SOM creates a topology preserving mapping from high-dimensional space onto map units so that the relative distance between data points are preserved. The SOM consists of neurons organized on an array, and the size of the neuron grid can be changed according to the requirements. Each neuron is characterized by an n-dimensional weight (a.k.a. codebook) vector,

$\mathbf{w}_i = [w_{i1},\ w_{i2},\ldots,w_{in}]$, where n is the size of the input data space, such that each neuron is connected to the adjacent ones by a neighborhood relation, which dictates the topology (or structure) of the map.

In each training step, one sample vector, X, is drawn randomly from the input data set and the distance between it and all the weight vectors of the SOM is calculated by using the Euclidian distance measure. The neuron whose weight vector is closest to the input vector, X, is referred to as the "best matching unit" (BMU). After the BMU is identified, the weight vectors of the BMU are updated, and its topological neighbors are moved closer to the input vector in the input data space. The neurons can be interpreted as n-dimensional points that tend to occupy the areas where the density of training data is higher. If training data are clustered, the neuron topology will reflect such a multimode nature. The final topology will represent the spreading of training data regardless of the actual distribution, which makes the SOM a suitable technique to design nonparametric monitoring tools.

By exploiting the architecture of the SOM, we found a correlation between the number of codebooks (i.e., the neuron weights), the topology of the map and the false alarm rate (type I error). This means that the selection of the number of neurons has a great impact on the performances of the monitoring system. In Fig. 4, the approach proposed for the selection of the number of codebooks is schematically outlined. The dataset is divided into a "training" dataset and "tuning" dataset. The former is used to train the SOM, the latter to test the Type I error performances on a different set of data. Starting from the input dataset and an initial number of codebooks, a SOM network is trained. Then, the Euclidean distance between the input data and the BMUs is calculated and its 100(1-α)% percentile is calculated by means of kernel smoothing density estimation [11]. The corresponding threshold is tested on the tuning data: if the estimated threshold does not reach the desired α target value (denoted by α*), then the number of codebooks is increased until the false alarm rate matches the desired target.

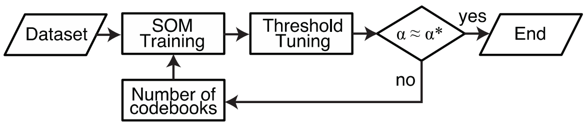

Fig. 4. Learning framework.

In Fig. 5, the average run length (ARL) is plotted against the possible combinations of codebooks that constitute the input of the SOM network. For the specific input dataset, the target ARL could be reached with 4 different combinations of codebooks. By selecting this map configuration it is possible to assure the desired false alarm rate along the monitoring phase.

3.3. Control chart-based monitoring: a simulation study

When the SOM network is trained with the selected codebooks map and the cutoff threshold has been estimated, it is possible to monitor if the new observations belong to the in-control region by calculating the Euclidean distance of each of

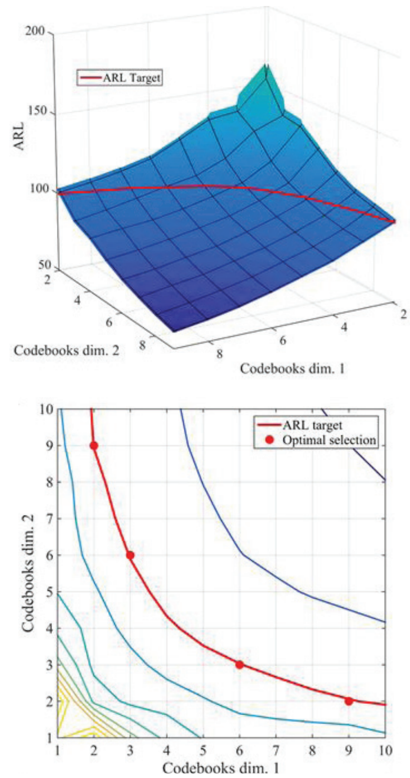

Fig. 5. Selection of the number of codebooks

them from its BMU and compare it to the estimated threshold. If the Euclidean distance of the new observation is higher than the threshold value, this observation will be marked as out-of-control, otherwise it will be considered as an effect of the natural variability of the process.

To motivate the usage of the SOM for monitoring free-form multivariate distributions, a simple simulation study is presented. A non-normal bivariate distribution is generated according to equation (1).

$$\begin{aligned} x_1 &= \delta_1 \cdot r \cdot sin(A) + \beta \cdot s \\ x_2 &= \delta_2 \cdot r \cdot cos(A) + \beta \cdot s \\ A &= c_1 \cdot \pi + \alpha \cdot c_2 \cdot \pi \\ & \beta \sim \mathcal{N}(0,1)\ \alpha \sim \mathcal{U}(0,1) \end{aligned} \tag{1}$$

Three scenarios have been simulated introducing a distortion in the distribution geometry by acting on the δ coefficients. Scenario 1 acts on the coefficient δ_1 while keeping δ_2 unchanged; scenario 2 does the opposite; scenario 3 modifies both the coefficients δ_1 and δ_2. Five different values of $\delta = [0.9, 0.7, 0.5, 0.3, 0.1]$ have been selected to prove the effectiveness of the reviewed methods in detecting deviations from the in-control state of δ = 1.

To compare the performances in terms of average run length, two competitors have been selected as representative of consolidated approaches for MSPC. The first one is the

Hotelling's T^2 control chart with empirical limits, and the second one is the one-class classification variant of the Fuzzy ART, presented and discussed in [5]. For the simulation study, 1000 replicates of each shift and scenario condition have been considered; the results of the simulation are shown in Table. 1.

Table 1. Simulated data: comparison in terms of ARL.

	ARL scenario 1 & 2			ARL scenario 3		
Shift	Fuzzy ART	T2	SOM	Fuzzy ART	T2	SOM
0	98,98	99,59	99,36	100,33	99,02	99,87
1	159,74	122,84	104,06	317,61	181,53	89,92
2	314,33	111,08	41,09	5437,56	713,07	14,99
3	384,26	75,85	12,91	9902,87	3361,31	3,68
4	399,1	49,95	5,73	9988,33	7608,6	1,72
5	462,84	32,45	3,61	10000	8977,52	1,33

By examining the results, it was possible to see that each considered method yields the same results in scenario 1 and 2, which is why they are unified in the results table. In both cases, the SOM approach proved to be the best performer, by an earlier detection of the distribution modification (lower ARL). While SOM and T^2 increase their performances when greater values of shifts are considered, the Fuzzy ART shows an increasing weakness in detecting out-of-control conditions. This is mainly due to the control region generated by this methodology and the fact that the out-of-control condition consists of a displacement of the data towards the inside of the banana-shaped distribution. Fig. 6 shows the control region of each compared method, where it is possible to notice that the SOM control region fits the in-control data distribution better than its competitors.

When, in Scenario 3, the shift occurs on both the δ1 and δ2 coefficients, the data distribution changes dramatically and only the SOM approach is capable of detecting this modification. The reason why the competitors fail in the identifications is due to the fact that the out-of-control distribution falls entirely in their control region, which is not properly fitting the in-control distribution.

As a result of this simulation study, the SOM seems to be a flexible tool for distribution-free process monitoring. Future studies may be aimed at comparing the SOM methodology with other machine learning method based on the one-class-classification paradigm [6].

4. Main results

Signal-based SPC usually follows an information synthesis step aimed at extracting a reduced set of variables from raw signals. These variables allow characterizing the ongoing process and detecting possible shifts from an in-control state. In the case of a transverse roll grinding operation, where three accelerometer sensors are used to monitor the stability of the process, the RMS of each signal is monitored. Transverse roll grinding involves consecutive cycles composed by different process runs, each one executed with different cutting parameters, which yields to a transition from one operating mode (i.e., one set of cutting parameters) to the following one.

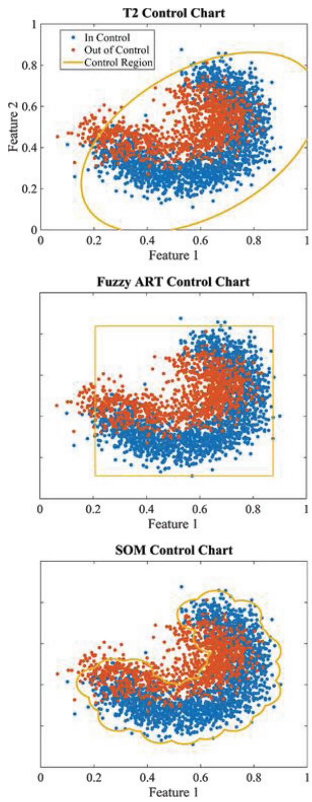

Fig. 6. Comparison of different control regions.

This causes a shift in the monitored time series that should not be signaled as an alarm, since it corresponds to a natural transition between two consecutive in-control states.

In Fig. 7, in addition to non-normality, the probability density function shows a multimodal behavior, which is a challenging violation of common distributional assumptions that motivates the usage of clustering tools such as the SOM.

Here, the presented SOM-based control chart is compared against its competitors on the traverse roll grinding case. The three methods are evaluated in terms of percentage of out-of-control identified (Tab. 2) and in terms of number of samples

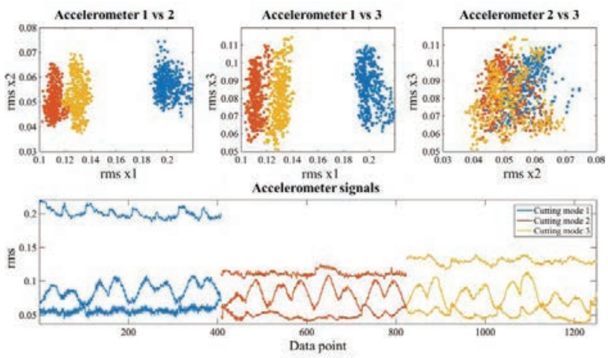

Fig. 7. Stable cutting, multimode real industrial dataset.

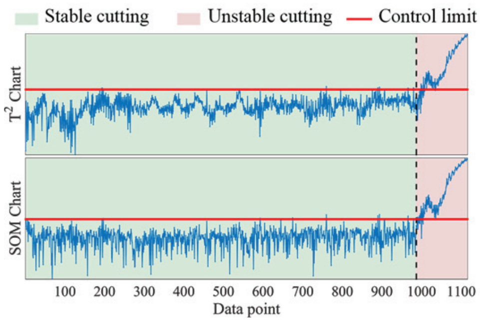

Fig. 8. Comparison of T^2 and SOM control charts.

before alarm (Tab. 3). All the approaches were trained with 1250 samples of stable cutting conditions and tested against 350 samples of unstable cutting conditions. The compared methods should identify the change in the cutting condition when the process switches from a stable to an instable mode.

Table 2. Real industrial data: comparison in terms of percentage of out of control identified.

	Real industrial data: % of out of control identified		
Cutting mode	Fuzzy ART	T2	SOM
Unstable (chatter)	60.30%	60.30%	68.03%

From the analysis of the results it turns out that even in the real industrial application the SOM shows a higher value of out-of-control correctly identified, thanks to its ability of better fitting the real data distribution. Furthermore, the proposed approach better performs in quickly identifying a deviation from the normal process conditions, which means that it is a much more reactive monitoring approach compared to the Hotelling's T^2 chart and the Fuzzy ART methodology.

Table 3. Real industrial data: comparison in terms of samples before alarm.

	Number of samples before alarm		
Cutting mode	Fuzzy ART	T2	SOM
chatter	105	106	9

In Fig. 8 it is possible to see the T^2 and the SOM-based statistics with their empirical control limits set with type-I error of 1%. As mentioned before, the SOM chart signals an alarm before the T^2 does, which can be better understood by observing the quicker growth of the SOM statistics in the red region of the chart where the process is unstable.

5. Conclusions

In this paper, the Self-Organizing Map has been applied as a tool for recognizing out-of-control process conditions in the state of a multivariate manufacturing operation, which has a multimode and distribution free behavior. In particular, a SOM-based monitoring system has been proposed for multivariate manufacturing quality monitoring and compared to other benchmark approaches. The SOM-based process monitoring approach is capable of providing an assessment of the current process state, which is achieved by calculating the Euclidean distance between the RMS value of the acquired accelerometers and their best matching templates. By monitoring the stability of the calculated Euclidean distance with an estimated control threshold, it is possible to identify changes in the cutting conditions which lead to a reduction of the workpiece quality. In this study, an algorithm is proposed to select the input parameters when designing the SOM-based control chart to cope with real-world applications.

The SOM is used as a one-class-classifier, hence it does not require previous information about abnormal pattern appearances but needs only normal operation datasets for the training. This feature makes its application more flexible and easier to be used in real industrial application compared to other supervised monitoring approaches. The results with simulated data demonstrate that the SOM chart is more sensitive to process shifts than the benchmark MSPC control charts like the Hotelling's T^2 and the Fuzzy ART when monitoring multivariate processes which have a multimode and distribution free behavior. In comparison with the other reviewed monitoring scheme, the SOM chart showed the ability of creating data adaptive control regions by performing a better fit of the real probability distribution function. According to the simulations and the real industrial case analysis the SOM-based chart can be considered an effective and promising monitoring tool for unsupervised quality monitoring of industrial process, which present multimode and distribution free behavior. Future studies will be aimed at comparing this methodology with other one-class-classification schemes for distribution-free process monitoring.

References

[1] Jin J., Shi, J. (1999). Feature-Preserving Data Compression of Stamping Tonnage Information Using Wavelets, Technometrics, 41:4, 327 – 339

[2] Jin, J., Shi, J. (2001). Automatic Feature Extraction of Waveform Signals for In-process Diagnostic Performance Improvement, Journal of Intelligent Manufacturing, 12, 257 – 268

[3] Maggioni M., Marzorati E., Grasso M., Colosimo B.M., Parenti P. (2014). In-Process Quality Characterization of Grinding Processes: a Sensor-Fusion Based Approach, ASME 12th Biennial Conference on Engineering Systems Design, ESDA 2014, June 25-27, 2014, Copenhagen, Denmark

[4] Tax D. M. J. (2001). One-Class Classification; Concept-Learning In The Absence Of Counter-Examples, Ph.D. thesis, Delft University of Technology

[5] Pacella M., Semeraro Q., (2005). Understanding ART-Based Neural Algorithms As Statistical Tools For Manufacturing Process Quality Control, Engineering Applications Of Artificial Intelligence, 18, 645-662

[6] Grasso, M., Colosimo, B. M., Semeraro, Q., Pacella, M. (2014). A Comparison Study of Distribution-Free Multivariate SPC Methods for Multimode Data. Quality and Reliability Engineering International, DOI: 10.1002/qre.1708

[7] Montgomery D. C. (2008). Introduction to Statistical Quality Control, John Wiley & Sons, 6th Ed

[8] Parenti P., Leonesio M., Cassinari A., Bianchi G., Monno M. (2013). Model-Based Identification of Chatter Marks during Cylindrical Grinding, International Conference on Advanced Manufacturing Engineering and Technologies, NEWTECH 2013, October 27-30, 2013 Stockholm, Sweden

[9] Kohonen, T. (2013). Essentials of the self-organizing map. Neural Networks, 37, 52-65.

[10] Kohonen, T., Oja, E., Simula, O., Visa, A., & Kangas, J. (1996). Engineering applications of the self-organizing map. Proceedings of the IEEE, 84(10), 1358-1384.

[11] Bowman, A. W., Azzalini A. (1997). Applied Smoothing Techniques for Data Analysis, Oxford University Press, New York