A Neural Network Approach to Quality Control Charts

Thomas STÜTZLE
Department of Statistics and O.R.
University Complutense, Madrid

Abstract

In this paper Quality Control Charts without memory are compared to neural networks trained with the Backpropagation algorithm. Neural networks are used to decide whether a process is under statistical control or out of control. As only the last sample is used to decide upon the state of the production process, a comparison to Shewhart-control charts leads automatically to a comparison between statistical tests and neural networks. By using a combined control chart to control the process mean and the process variability, the kind of classifications of the kind of change is considered explicitly. Finally neural networks are used to classify the kind of change occurred, considering only the last sample.

1 Introduction to quality control charts

Control Charts are used as an aid to check whether a production process is under or out of statistial control. A process is said to be out of statistical control when some parameters like the process mean or variability have changed. In order to control the paramters of a production process at constant intervals of time, a sample of the process is taken and the realization of a quality characteristic is measured. Let $\mathbf{x_t} = (x_{t1}, x_{t2}, .., x_{tn})^T$ be the vector of these measurements. This vector can be interpreted as a realization of the random vector $\mathbf{X_t} = (X_{t1}, X_{t2}, .., X_{tn})^T$. The running of a control chart can be considered as a periodically repeated statistical test, where the null hypothesis is $\mathbf{H_0}$: *The process is under statistical control* and the alternative hypothesis is $\mathbf{H_1}$:*The process is out of statistical control.* As with every test errors of type 1 and type 2 can be commited. The possible errors are **Error of type 1**: *Some action is taken although the process is under control* and **Error of type 2**: *No action is taken although the process is out of control.* The errors of type 1 are also called false alarms. False alarms are to be avoided because an interruption of the production process and an unnecessary search for possible errors cause avoidable costs.

Often Shewhart control charts are used to control the mean and the variability of a production process. In this paper a $\bar{x}$-chart, that uses the sample mean as sample statistic, is chosen to control the process mean and a s-chart, that makes use of the sample standard deviation, is taken to control the process variability. These control charts do not consider past samples and only the most recent sample is used to test whether some parameter has changed.

To control the mean and the variability of a production process at the same time, often a $\bar{x}$-chart and a s-chart are used in parallel. This possibility was also considered here. The interesting point for a combined control chart is, that a sole increase of the process variability results in an increase of the probability that the mean chart gives an out of control signal, independently of the sample size n chosen. Otherwise a change of the process mean has no effect on the s-chart. This may lead to false classifications of the kind of process change. These false classifications are to be avoided because a correctly classified change may give an indication for possible faults and facilitate the search for them.

One possibility to avoid these false classifications would be to use past samples as inputs for some classification procedures and to use them to distinguish the different kinds of change.

2 Applications of neural networks in quality control

Recently there have been some attempts to use neural computing for special tasks in the Statistical Process Control (SPC). One approach is to compare control charts to neural networks. Pugh (1991) tried to imitate a Shewhart Control Chart for the process mean with neural networks. In Smith (1994) a neural network was trained to replace a combined $\bar{x}$-R-chart.

Another kind of application of neural networks was considered in the article of Guo, Dooley (1992). When a shift of the parameters by two control charts for the process mean and the process variability was detected, they used neural networks to determine whether the out-of-control signal was actually owed to a shift of the process mean or of the process variability. They compared the results obtained with neural networks to a quadratic discriminant function and some obvious heuristics that consider the control chart that first signals a change in the parameters. Here it is tried to improve the evaluation of a Quality Control System to give a better indication for possible errors.

Another possibility is the interpretation of control charts to detect unnatural patterns on the charts. This approach was described in the articles of Hwarng, Hubele (1993) and Pham, Oztemel (1994). Some unnatural patterns on control charts are described for example in Montgomory (1985). These unnatural patterns also indicate an out-of-control situation. In Hwarng, Hubele the Backpropagation-algorithm was used, whilst Pham, Oztemel make use of some Learning Vector Quantizer algorithm (LVQ).

In the last two possible approaches, neural networks were used as a possible classification tool for patterns on control charts. This may be the more usefull application of a nonparametric method like neural networks. Nevertheless in this paper neural networks are compared to Shewhart Control Charts especially considering the different errors of type 1 and type 2 as in the articles of Pugh and Smith different weights for this two type of errors were not considered. Here the results of the neural networks are compared to control charts for the process mean and the process varibility. Next neural networks are trained to imitate a combined $\bar{x}$-s control chart.

Basically, in this paper neural networks are therefore used to imitate a special statistical test. Usually neural networks can be used as a special classification procedure, but a statistical test might be interpreted also as some kind of classification.[1] In a statistical test the null hypothesis and the alternative hypothesis may be interpreted as two different classes. In terms of control charts these two classes can be interpreted as *the process is under statistical control* and *the process is out of statistical control.*

It is well known, that the output of neural networks when used for classification approximates under certain conditions bayesian a posteriori probabilities, see e.g. Richard, Lippman (1992). This has also some implications for statistical tests, what will be investigated in the first experiment. Actually statistical tests may be formulated by a bayesian approach where the posterior probabilities are used.

3 Comparison of Neural Networks to Shewhart Control Charts

3.1 Basic assumptions and practical considerations

3.1.1 Assumptions for the comparison

For the comparisons a sample size of $n = 6$ was chosen. For the generation of the samples by Monte Carlo simulation it was supposed that $X_{t1}, X_{t2}, .., X_{tn}$ are independent, identically $N(\mu_t, \sigma_t^2)$ distributed. If the process is under statistical control the parameters are chosen $\mu_0 = 0$ and $\sigma_0 = 1$. For the different control charts a significance level of $\alpha = 0,01$ is chosen. This results in an overall level of error of type 1 of $\alpha = 0.0199$ for the combined $\bar{x}$-s control chart.

3.1.2 Partitioning of the samples into sets

Three disjunctive sets of 1500 samples each were generated. The first set was used as the training set. As the different weights of errors of type 1 and type 2 had to be considered, a second set of samples was used to determine the interpretation limits for the network outputs in a proper way, as explained in section 3.1.4. The last set was used as a test set, to obtain the final results for the comparison.

3.1.3 Preprocessing of the samples

It was also one objective of this investigation to investigate the effect of the different possibilities for the preprocessing of the samples. In the presentation of the results differences caused by the preprocessing will be discussed only if they lead to obviously different results.

One possibility is, not to preprocess the sample vector $\mathbf{x} = (x_1, x_2, .., x_n)^T$ and take it as input for the neural net.

Additionally sample statistics like the sample mean and the sample standard deviation can be calculated and used as additional inputs for the neural net. This should not be necessary, since the neural networks should be able to imitate a control chart without additional sample statistics.

Another possibility is to order every sample in increasing order, so that $x_{(1)} < x_{(2)} < .. < x_{(n)}$. Such a preprocessing has the advantage that the jth input neuron always gets the j-largest sample value as input. With this kind of preprocessing e.g. the sample range, that has a direct relation to the process variability, can easily be calculated as a linear function.

A further possibility is to consider only the upper half of the mean control chart and to 'reflect' the sample realizations falling beneath the middle line of the $\bar{x}$-chart with respect to this middle line. This can easily be done, since the $\bar{x}$-chart is symmetrical with respect to the middle line. So only one control limit has to be considered.

3.1.4 Classification considering errors of type 1

One possibility for clasification with neural networks is to choose one output neuron for each class. This corresponds to a vectorial codification of the objective outputs. In this case the jth class may be represented by a basis vector whose jth component is 1.

To interpret the outputs, one way is to assign the input to the class that is represented by the output neuron with the highest output value or to assign a class only if a specified limit is exceeded. For the comparison between neural

[1] Or every classification might be interpreted as some special kind of statistical test.

networks and quality control charts it was considered essentially, that also with the neural networks a rate of error of type 1 comparable to control charts can be observed. Therefore here another possibility was used to determine proper limits for the interpretation of the outputs.

An independent set of samples was used to determine these interpretation limits. The number of errors of type 1 and type 2 made by the corresponding control chart were compared to the number of errors of type 1 and type 2 made by the neural network for different values of the output limits. Then the interpretation limits for the outputs were chosen so, that approximately the same number of errors of type 1 as with the control charts were made. If there were several possible values, also the errors of type 2 were considered to choose adequate limits. The values for the output limits were examined over a fixed range of values with a stepsize of 0.01. The final results for the comparison between neural networks and control charts are then based on the evaluation of the test set with the fixed limits.

Another way for the representation of the outputs is to take one output neuron and to assign each class a certain value as the objective output. This can be considered as a scalar codification of the outputs. Here interpretation limits have to be assigned to divide the output into different intervals. No noticeable difference between the two types of representation of the objective outputs were found in the experiments.

3.1.5 Training of neural networks

For the comparisons to the different control charts feedforward neural nets, that were trained with the Backpropagation algorithm, were used. As activation functions for the hidden and the output neurons a sigmoid function was used. To determine the time when to stop the training, a validation set was used, as it is proposed in Hecht-Nielson (1990).

3.2 Discussion of the experimental results

3.2.1 Approximation of posterior probabilities and statistical tests

The approximation of the posterior probabilities of neural networks and the resulting relation to statistical test is examined here, considering a simple test with $H_0: \mu_0 = 0$ *and* $H_1: \mu_1 = 1$. Here σ_0 is assumed to be 1.[2]

For this test, the posterior probabilities for the two parameters can be calculated explicitely and can be compared to the outputs of a neural network. It can be checked also, whether the outputs sum 1 as they should, if posterior probabilities are approximated. This test can be formulated as a Neyman-Pearson test or as a bayesian test. The both test formulations can be shown to be equivalent. For the Neyman-Pearson test a level $\alpha = 0.01$ is chosen. For the bayesian test prior probabilities of the parameters have to be given or estimated. Here it is assumed that $P(\mu_0) = P(\mu_1) = 0.5$, so for each set 750 samples with μ_0 and 750 samples with μ_1 were generated. With these prior probabilities and the assumptions made here, the posterior probability for the parameters can be calculated and for the parameter μ_1 a value for its posterior probability can be determined, from that one should choose parameter μ_1 to get a bayesian test equivalent to a Neyman-Pearson test with a fixed level α. In this example, this probability is approximately $P(\mu_1|\mathbf{x}) = 0.937$. The interpretation limit of the output of the neural nets trained for this test should be therefore close to this value.

For the neural networks two output neurons were used to represent the two different parameters or classes. Nets with and without a hidden layer were trained and then the errors of type 1 and type 2 for the test set were compared to the performance of the corresponding statistical test. The statistical test corresponds to the application of an one-sided control chart with the upper control limit of $\tau_{0.99} \cdot 1/\sqrt{6} = 0.9497$.

In table 1 the interpretation limits for the node that represents parameter μ_1, the maximum and the minimum of the summed outputs, that should be near 1, and the mean absolute deviation of the output of the neuron that represents parameter μ_0 from the correct posterior probability were calculated. Here e.g. 6–3–2 refers to a neural net with 6 input, 3 hidden and 2 output neurons.

Architecture	errors of type 1/type 2	interpretation limits	max. of summed outputs	min. of summed outputs	mean absolute deviation from posterior prob.
QRK	5 / 354	/	/	/	/
6–2	6 / 354	0.92	1.00002	0.99998	0.020986
6–3–2	6 / 343	0.91	1.01231	0.98492	0.030631
6–5–2	6 / 385	0.93	1.0264	0.97846	0.026766
6–7–2	7 / 341	0.90	1.00354	0.99674	0.027840

Table 1 Results for the approximations of posterior probabilities

With respect to the approximation to the posterior probabilities, the net with architecture 6–2 gives the best performance. This can be best seen taking account of the mean absolute deviation of the posterior probabilities from the network outputs. Also the maximum and the minimum of the sum of the outputs are very close to 1. Compared to the net 6–2, the nets that use hidden units give quite large values for the maximum and the minimum. The limits

[2] This test can be seen also as a typical classification problem, see e.g. Duda, Hart (1973).

for the interpretation of the output are for all networks close to the theoretical value of 0.937. This should bethis way, as with a good approximation of the posterior probabilities, the interpretation limits should also be close to this theoretical value. Also the number of errors of type 1 and type 2 should be close to the number of errors for the statistical test, as with the posterior probabilities a bayesian test, equivalent to a Neyman-Pearson test, can be formulated and this bayesian test depends on the posterior probabilities. The experimental results show, that neural networks can be trained to imitate some special statistical tests also with respect to a certain rate of error of type 1 that has to be considered. So e.g. the net 6–2 gives nearly the same results as the corresponding statistial test. Only one error of type 1 more is made. The nets with additional hidden units give similar results, so e.g. net 6–3–2 gives even less errors of type 2 as net 6–2.

3.2.2 Comparison to a $\bar{x}$-chart

In this section a more general comparison to a $\bar{x}$-chart is carried out. The samples were generated so, that approximately for one third of the samples the mean was $\mu_0 = 0$, for approximately one third a negative shift, uniformly distributed in $[-3\sigma_0; -0.5\sigma_0]$, and for the last third a positive shift, uniformly distributed in $[0.5\sigma_0; 3\sigma_0]$, was supposed.[3] For the comparisons only the number of errors of type 1 and type 2 were considered. Another possibility would have been to estimate the average run length for some selected shifts. One objective of the experiments was also to investigate the effect of different kinds of preprocessing for the samples. The preprocessing and the determination of correct outputs can be chosen in such a form, that theoretically no hidden layer is necessary to imitate a $\bar{x}$-chart. This is the case when the samples are 'reflected', as described in section 3.1.3 or when an increase and a decrease of the mean are considered explicitly as classes. As there were only slight differences between the different kinds of output representation and preprocessing, these will not be discussed here.

In the test set, where 501 samples did not undergo a change of the mean, 5 errors of type 1 and 203 errors of type 2 were made with the $\bar{x}$-chart. The results of the neural networks have to be compared with these numbers of errors.

In table 2 some results of neural networks for different types of preprocessing of the data are given. Here 'reflected' refers to the use of only one half of the control chart, ordered refers to the ordering of the samples and with codification the presentation of the objective output is meant. The different number of output nodes is due to the representation of the objective outputs.

Architecture	errors of type 1	errors of type 2	remarks on preprocessing an output representation
6–2	5	206	reflected, ordered data, vectorial codification
6–2	5	200	not reflected, ordered data, vectorial codification
6–1	5	196	not reflected, ordered data, scalar codification
6–3	5	207	not reflected, not ordered, vectorial codification

Table 2 Results for the neural networks for the comparison to a shewhart control chart for the process mean

Generally it can be said, that neural networks are able to imitate a $\bar{x}$-chart and may obtain quite the same results with respect to errors of type 1 and type 2. This can be seen comparing the results for the neural networks given in table 2 to the results of the test set for a $\bar{x}$-chart. Accidentally all the neural nets give 5 errors of type 1 with little differences in the number of errors of type 2. The kind of preprocessing of the samples had only very little influence on the results. This seems reasonable, since the kinds of preprocessing considered here do not influence on the decisions made by the $\bar{x}$-chart either.

Another outcome is, that no hidden layer is necessary to obtain results, that compare to a $\bar{x}$-chart. But this is only the case if no hidden units are necessary from theoretical considerations.[4] It is also not necessary to use additional inputs like the sample mean.

Finally it should be noticed that the results presented here could be expected to be this way since the problem considered here is basically a linear problem. This is so because the sample mean, that is a sufficient statistic for the parameter μ of a normal distribution, is also linear function, that should be easily approximated by neural networks.

3.2.3 Comparison to a s chart

For the comparison between a standard deviation chart and neural networks only an increase of the process variability was considered. For the generation of the samples it was supposed, that approximately half of the samples were subject to a shift of the standard deviation, that was taken as uniformely distributed in $[1.2\sigma_0; 3\sigma_0]$. For the test set, where 742 samples did not undergo a change of the standard deviation, 7 errors of type 1 and 325 errors of type 2 were made. In the following table some results obtained with neural networks are summarized.

[3] Here a special interpretation of the prior class probabilities was supposed. This interpretation corresponds to three classes, that have equal prior probabilities. Another possibility would have been to choose for one half of the samples a shift that can be positive or negative and for the other half no shift at all.

[4] Hidden units are only necessary when the samples are not 'reflected' and only the classes shift/no-shift are distinguished. For a justification of this result, the problem may be compared to the well known XOR-Problem.

Architecture	errors of type 1	errors of type 2	remarks
6–2	8	297	reflected, ordered data
6–9–2	9	278	reflected, ordered data
6–15–2	9	398	reflected, unordered data
6–13–5–2	11	397	reflected, unordered data

Table 3 Results for the neural networks for the comparison to a shewhart control chart for the process variability

The most salient result is, that neural networks can give comparable results to a s-chart, if the ordered samples are taken as inputs. Then even a network without hidden layer suffices. If the unordered samples are taken as inputs, a considerably higher number of errors of type 2 is made. This means, that here the kind of preprocessing of the samples has a noticeable influence on the results.

It should be observed here, that the sample standard deviation actually is a nonlinear function of the sample. If no ordering of the sample is made, hidden units are necessary to obtain reasonable results. The worse performance of neural nets, that take unordered samples as inputs, shows, that neural networks were not able to approximate good enough the decision rule of a s-chart. One reason may be, that not enough training examples were used as inputs for the neural net. Another reason may be, that the architecture of the neural net is not complex enough to approximate the desired function. Here no further experiments were made to investigate the reasons of this worse performance. If the sample standard deviation or the sample range are used as additional inputs to the unordered samples, similar results to a s-chart are obtained. This seems to be natural as the sample standard deviation is also the statistic used by the s-chart and the neural network only has to 'learn' to interpret this statistic correctly.

With the ordered samples, similar results to a s-chart could be obtained. It makes no sense to talk about an overall error rate, if the two different kinds of errors are not considered. We will limit us here to state, that neural nets and the s-chart give approximately the same result, if the samples are ordered. It might be interesting to investigate the slightly different results in more detail.

The reason for the good performance of neural nets with ordered inputs may be explained. The greatest and the smallest sample value suffice to calculate the sample range, that can also be used to control the process variability. Here the input for the neural network is the vector whose components are the order statistics of the sample, so that the sample range could be calculated as a linear function of these inputs. If no hidden unit is used, basically a monotone transformation of the weighted sum of the order statistics is calculated in the output neurons. If one examines the weights for these networks, it can be seen that the smallest and the greatest value have also the greatest weights associated with them.

3.2.4 Comparison to a combined $\bar{x}$-s chart

If one uses a combined $\bar{x}$-s-chart, there are three classes that are to be distinguished. These classes may be interpreted as *The process is under statistical control*, *A change of the process mean occurred* and *A change of the process variability occurred.* A further class like *Both parameters changed* might be considered.

As a consequence of the results obtained in the last two sections, the ordered samples were used as inputs. Here the samples were generated so, that approximately one third of the samples had no shift, approximately the second third had a shift of the parameter μ, uniformly distributed in $[\mu_0 - 3\sigma_0; \mu_0 - 0.5\sigma_0] \cup [\mu_0 + 0.5\sigma_0; \mu_0 + 3\sigma_0]$, and the last third had a shift of the standard deviation, uniformely distributed in $[1.2\sigma_0; 3\sigma_0]$. In table 4 the evaluation of the test set with a combined control chart is given and explained below.[5]

position of sample points	samples without shift	samples with shift of μ	samples with shift of σ
on both charts inside c.l.	495	99	175
outside c.l. on $\bar{x}$-chart	5	387	28
outside c.l. on s-chart	6	1	232
outside c.l. on both charts	0	2	70
sum	506	489	505

Table 4 Results of the test data set for a combined $\bar{x}$-s-chart. Errors of type 1: 11 Errors of type 2: 274 False classifications: 31

With a $\bar{x}$-s-chart a shift of a parameter is indicated, if at least one chart gives an out of control signal. E.g. a shift of μ may be indicated on the $\bar{x}$-chart, on the s-chart or on both charts at the same time. No error is only made if a shift of σ is indicated on the s-chart, a shift of μ is indicated on the $\bar{x}$-chart and if no shift occurred, both charts do not indicate an out of control situation.

The control charts made $11(= 5 + 6)$ errors of type 1 and $274(= 99 + 175)$ errors of type 2. Another kind of error occurs when the kind of shift is classified in a wrong way. It is striking here, that a lot of times (70) a shift of σ is indicated on both charts. This is mainly due to the possible values for the shifts of the standard deviation up to $3\sigma_0$. To determine the number of false classifications it should be considered that the $\bar{x}$-chart reacts to an increase of the process variability with a higher probability of false alarms whereas the s-chart is not sensible to changes of μ. If both charts give an out of control situation, this should be therefore interpreted actually as a shift of σ. So the total number of false classifications amounts to $31(= 28 + 1 + 2)$.

[5] In the table c.l. means control limit.

Next an exemplar result of a neural net, that was trained to distinguish between the three classes is given. This result is representative of others that were obtained with other architectures of neural networks that were trained for this problem.

network 6–5–3	samples without shift	samples with shift of μ	samples with shift of σ
no change shown	491	120	194
shift of μ shown	8	367	14
shift of σ shown	7	2	297
shift of μ and σ shown	0	0	0
sum	506	489	505

Table 5 Results of the test data for a neural net, architecture 6–5–3. Errors of type 1: 15 Errors of type 2: 314
False classifications: 16

It can be seen, that for the neural network the number of errors of type 1 and of type 2 is higher than for the combined control chart. Only the number of false classifications $(16 = 14+2)$ is smaller.[6] In general neural nets could not reach the lower numbers of errors of type 1 and type 2 of the $\bar{x}$-s-chart. They only made less false classifications. Nevertheless the false classifications are mainly of that kind, that a shift of σ is interpreted as a shift of μ. It should also be noticed, that in this example it never occurred, that a shift is indicated on both output neurons that indicate shifts of μ or σ.

It should be stated here, that the main problem is, that neural nets do not reach the number of errors of type 1 or type 2 of the control chart. Although they offer an advantage with respect to a lower number of false classifications, this should be no reason to use neural nets to replace a combined control chart, as the essential thing are the lower numbers of errors of type 1 or type 2. The false classifications could also be handeled by a further interpretation of the control charts considering past samples. Nevertheless neural nets could offer an advantage in a combination with control charts as it is discussed in the next section. One reason for neural networks not to reach this low numbers of errors of type 1 or type 2 may be, that they also minimize explicitely false classifications because of the objective function they use, whereas with the application of a combined $\bar{x}$-s-chart the classification of the kind of change is not considered.

3.2.5 Classification of the kind of change

It is interesting, that the number of false classifications with neural networks is smaller than with the combined control charts. This leads to the idea, that neural networks may only be used as a classification procedure to classify the kind of change, once that such a change is indicated by a combined control chart. To investigate whether with such a combined Shewhart-chart/neural-net scheme the number of false classifications can be reduced, a neural network was trained to distinguish between shifts of the mean and shifts of the standard deviation. A test set of 2000 samples was then evaluated by a combined $\bar{x}$-s-chart. If the combined $\bar{x}$-s-chart indicated an out of control situation, the kind of shift was classified by a neural network. The number of false classifications given by the neural network was then compared to the number of false classifications made by the combined $\bar{x}$-s-chart.

The samples were generated so, that only shifts of μ, uniformely distributed in $[\mu_0 + 0.5\sigma_0; \mu_0 + 2\sigma_0]$, or shifts of σ, uniformely distributed in $[1.2\sigma_0; 2\sigma_0]$, were considered.

The evaluation of the test data with a combined $\bar{x}$-s-chart is given in table 6.

Results for the test data for the combined $\bar{x}$-s-chart	samples with shift of μ	samples with shift of σ
on both charts inside c.l.	706	688
only outside c.l. on the $\bar{x}$ chart	256	64
only outside c.l. on the s-chart	9	240
Outside c.l. on both charts	7	30
sum	978	1022

Table 6 Results of the test data for a combined control chart. False classifications: 80

With the combined control chart $80(= 64 + 9 + 7)$ false classifications would be made. In table 7 now the results of the test of the kind of classification with the neural network is given. Here only the 606 samples where an out of control situation was detected are considered.

Classification of shift with neural nets	neural net shows shift of σ	neural net shows shift of μ
on both charts outside c.l., shift of σ	30	0
on both charts outside c.l.,shift of μ	2	5
outside c.l. on the $\bar{x}$ chart, shift of σ	38	26
outside c.l. on the $\bar{x}$ chart, shift of μ	4	252
outside c.l. on the s chart, shift of σ	238	2
outside c.l. on the s chart, shift of μ	5	4

[6] This holds in general for all other neural nets, that were trained for this problem.

Table 7 Classification of the kind of change with neural nets. False classifications: 39

With neural nets $39(= 2 + 26 + 4 + 2 + 5)$ false classifications are made. This means, that the number of false classifications has approximately been halfed. This seems to indicate, that by using neural nets for only this special classification task, an improvement could be achieved. Anyway these result should be carefully interpreted because in this experiments only the last sample was considered.

4 Summary and final comments

In this paper it was essentially investigated whether neural networks can imitate a certain model, like it is assumed for the control charts. This problem can be interpreted as a special kind of statistical test.

Generally it can be pointed out, that neural networks can be trained to imitate the control charts considered here. This holds for the $\bar{x}$-chart as well as for the s-chart. Neural nets can be trained to consider the different kinds of error of type 1 and type 2. If both kind of changes are considered, the results of neural networks with respect to errors of type 1 and type 2 are worse than those of a combined $\bar{x}$-s-chart. Only for the number of false classifications an improvement could be found. So it seems to be more reasonable to use neural networks for special tasks such as the classification of the kind of shift. A correct classification of the kind of change of the process parameters is important as it may facilitate the search for possible errors.

The main drawback of the problems treated here is, that past samples are not considered. Past samples are used e.g. by the CUSUM-charts that are able to detect small shifts of the parameters faster than Shewhart charts. There are also additional rules for the interpretation of the Shewhart charts that give them an artificial memory. Here it would be one possibility to investigate whether neural networks are also able to imitate or even to give an improvement over these charts. Considering the patterns of the past samples, neural networks can also be used to detect special patterns on control charts like mixtures or a gradual increase/decrease of the mean or the process variability. These special patterns indicate also changes of the parameters. Another possibility is, that different patterns may indicate certain errors occurred during the running of a machine. As the use of computers increases also in the Statistical Process Control, adequate algorithms or classification procedures may be used to obtain a better evaluation of the production process. This evaluation has to be made in a real time manner. One possibility here is the use of neural networks for this task, but still further investigation is needed.

References

Duda, R. and Hart, P. *Pattern Classification and Scene Analysis,* John Wiley & Sons, New York, 1973.

Guo, Y. and Dooley, K.J. *Identification of Change Structure in Statistical Process Control,* International Journal of Production Research, Vol.30, 1992, 1655–1669.

Hecht-Nielson, R. *Neurocomputing*, Addison-Wesley Publishing Company, 1990.

Hwarng, H.B. and Hubele, N.F. *Back Propagation Pattern Recognizers for X Control Charts: Methodology and Performance,* Computers & Industrial Engineering, Vol. 24, 1993, 219–235.

Montgomery, D.C. *Introduction to Statistical Quality Control,* John Wiley & Sons, 1985.

Pham, D.T. and Oztemel, E. *Control Chart Pattern Recognition Using Learning Vector Quantization Networks* International Journal of Production Research, Vol.32, 1994, 721–729.

Pugh, G.A. *A Comparison of neural Networks to SPC Charts,* Computers & Industrial Engineering, Vol.21, 1991, 253–255

Richard, M.D. and Lippman, M.P. *Neural Network Classifiers Estimate Bayesian a posteriori Probabilities,* Neural Computation, Vol.3, 1991, 461–483.

Smith, A.E. *X-bar and R control Chart Interpretation Using Neural Computing,* International Journal of Production Research, Vol.32, 1994, 309–320.

Stützle, T. *Vergleich Neuronaler Netze zu Qualitätsregelkarten ohne Gedächtnis,* M.S. Thesis at the University of Karlsruhe, Faculty of Economics, Karlsruhe, 1994.