X-Bar and R Control Chart Interpretation Using Neural Computing

Alice E. Smith
Assistant Professor
Department of Industrial Engineering
University of Pittsburgh
1031 Benedum Hall
Pittsburgh, Pennsylvania 15261
412-624-5045
412-624-9831 (fax)

Abstract. This paper formulates Shewhart mean (X-bar) and range (R) control charts for diagnosis and interpretation by artificial neural networks. Neural networks are trained to discriminate between samples from probability distributions considered within control limits and those which have shifted in both location and variance. Neural networks are also trained to recognize samples and predict future points from processes which exhibit long term or cyclical drift. The advantages and disadvantages of neural control charts compared to traditional statistical process control are discussed.

Keywords. control charts, neural networks, statistical quality control, artificial intelligence.

Revised for *International Journal of Production Research*

March, 1993

X-Bar and R Control Chart Interpretation Using Neural Computing

Abstract. This paper formulates Shewhart mean (X-bar) and range (R) control charts for diagnosis and interpretation by artificial neural networks. Neural networks are trained to discriminate between samples from probability distributions considered within control limits and those which have shifted in both location and variance. Neural networks are also trained to recognize samples and predict future points from processes which exhibit long term or cyclical drift. The advantages and disadvantages of neural control charts compared to traditional statistical process control are discussed.

1. Introduction

Control charts are commonly used in production environments to analyze process parameters to determine if a controlled process is within or out of control, i.e. to distinguish between assignable and common, also called chance, causes. Some manufacturing processes which benefit from control chart tracking are filtration, extraction, fermentation, distillation, refining, reaction, pressing, metal cutting, heat treatment, welding, casting, forging, extrusion, injection molding, spraying, and soldering (Miller and Walker 1988). Although computationally simple, control charts are sometimes complex to use correctly because the sample points come from non-specified probabilistic distributions and usually require interpretation by a skilled user.

Artificial neural networks trained by supervised techniques have been documented as good alternatives for pattern classification and prediction. These skills can be put to use for the interpretation of process control from input of control chart samples. Through learning of varying input conditions matched with control status, a network can decide on control status when faced with new inputs. Besides exceeding control boundaries, control chart points can provide information about the long term condition of the process through symptomatic shapes, runs and drifts. If these can be correctly identified from a small sample by neural networks, then the process can be investigated expediently. Neural networks can also forecast future control chart point(s), thus contributing to the diagnosis of process condition in borderline conditions.

This paper briefly discusses Shewhart control charts and their uses, then covers previous applications of artificial intelligence to statistical process control. The research focuses first on the data representation required to successfully train a backpropagation neural network to recognize control chart patterns. Number of inputs and preprocessing which enhance network performance are discussed. The ability of networks to discriminate within control from out of control situations based on small, probabilistic samples is presented. A second focus is the ability of the network to recognize sequences of noisy data points as belonging to patterns which reflect long term undesirable drift in the process, and the third area of research is the prediction of future subgroup means based on immediate past subgroup means.

2. Overview of Control Charts

W. A. Shewhart first proposed the use of control charts in 1931, which commonly bear his name as Shewhart control charts (Shewhart 1931). A typical Shewhart control chart is shown in Figure 1. Through the ensuing years many different formulations became known, however most manufacturers use versions of the early control charts which track sample mean (X-bar charts) and sample range (R charts) as checks on the process state and the process variability. One recent study shows 71% of responding firms use X-bar control charts and 64% use range charts (usually in tandem), far greater percentages than the nearest competitors (Saniga and Shirland 1977). This paper focuses on these two since they are the most common control charts.

INSERT FIGURE 1 HERE

A difficulty with control charts is the determination of whether a process is actually within control or not. Since sample points are subject to noise due to the process itself, and measurement and human imperfections, they form a non-specified probabilistic distribution. Another difficulty is the selection of appropriate heuristic rules to interpret the position and shape of control chart points. The eight most accepted rules, with two examples being 14 points in a row alternating up and down and 6 points in a row steadily increasing or decreasing, are discussed in (Nelson 1984). The process randomness and the choice of interpretive heuristics affect the performance of a control chart. Errors are generally classified as Type I (α) Errors, which are false alarms, and Type II (β) Errors, which are missed disturbances. There may be different relative penalties for Type I and Type II Errors.

3. Artificial Neural Networks for Process Control

3.1. Overview of Neural Networks

Artificial neural networks, a branch of artificial intelligence, are massively parallel computing mechanisms emulating the biological brain, which store intelligence in their many interconnecting weights. These variable weights connect nodes (neurons) both in parallel and in sequence. The entire mechanism hierarchically processes vector input through the network of nodes and weights, arriving at a vector output.

Neural networks have been noted as being particularly advantageous for modeling systems which contain noisy, fuzzy and uncertain elements. They learn models by iterating through a large number of exemplar vectors. Relationships can be auto-associative (relating an input with itself), or hetero-associative (relating an input with another output). Learning can take place through internal grouping (self organizing or competitive learning) or through paired training sets (supervised learning).

For modeling control data, a supervised approach is preferable since calibrated training data is usually available and it is advantageous to pre-specify the desired output. The most well known of supervised techniques is backpropagation, which adjusts initially randomized weights during training according to the steepest gradient along the error surface (Werbos 1974, Rumelhart et al. 1986). Weights are adjusted in proportion to their contribution to the output by recycling the squared error signal back through the layers of weights. Typical backpropagation neural networks, which are more properly termed multi-layered perceptrons trained by backpropagation, are fully connected, feed forward only, and use a sigmoidal transfer function at the nodes to evaluate weighted input sums. An input layer, an output layer and at least one hidden layer are required to model non-linear systems (Funahashi 1989, Hornik et al. 1989), however it has been suggested that for analog input, a two hidden layer network is superior (Lapedes and Farber 1988, Smith and Dagli 1991 B). Figure 2 shows a typical two hidden layer backpropagation neural network for process monitoring.

INSERT FIGURE 2 HERE

3.2. Neural Networks in Process Control

In process control, work has been done to neurally relate input parameters to product variables in both an associative and a predictive model. Association tasks usually are diagnostics for use during manufacturing. Burke identified tool wear states during machining using a competitive learning network, a task normally done by human operators (Burke 1989). A similar subject was pursued by Guillot and Ouafi who used a supervised network to recognize tool breakage for use in untended machining (Guillot and Ouafi 1991). A third neural model for monitoring during machining used the frequency of the vibration signals to classify if machine deterioration was taking place (Knapp and Wang, in press). The plastics industry was studied with a backpropagation correlation model of injection molding process parameters and product defects for diagnostics and corrective action (Wu et al. 1991). The input vectors were the quality defects while the output vectors contained recovery instructions.

Predictive models attempt to estimate product parameters based on process conditions before product manufacture. Andersen et al. used backpropagation networks to relate input parameters (arc current, arc voltage, travel speed and wire speed) predictively to quality measures of a weld (bead width, penetration, reinforcement height and cross section area) (Andersen et al. 1990). Okafor pursued a similar approach for estimating surface roughness and bore tolerance in milling using input variables of cutting force components, acoustic emission and spindle vibration in a moving window size of five (Okafor et al. 1990). Smith and Dagli related many input variables of a plastic extrusion process to prediction of the lot quality with backpropagation (Smith and Dagli 1991 A). Smith used a similar approach for injection molding of brake linings to predict product quality and its variability (Smith 1993).

These works suggest that all processes can be diagnosed and modeled successfully by neural networks, although results may not necessarily be superior to statistical or other analytical techniques. There is sustained interest however, for reasons other than superiority of performance. One reason is the ability to learn relationships through the data itself rather assuming probability distributions or explicitly coding an empirical model. A learned relationship eliminates any error due to erroneous parametric or analytical assumptions. The second reason is that training can handle multiple, related or non-related, inputs and outputs simultaneously. This has important advantages for control charts as demonstrated later in this paper where a single neural network simultaneously serves as both X-bar and R charts. The third reason is that a neural network can dynamically adjust to changing line conditions by continuous training, or sporadic retraining. This means a manufacturer can improve the neural network control chart model by accumulation of additional training data, then performing incremental training or occasional new batch training. The fourth reason is that hardwired neural networks are expected to be readily available in the near future, which will facilitate compact, cost effective, real time control. A neural network chip alleviates the need for continuous monitoring by PC or other computing platform.

3.3. Past Research on Control Charts and Intelligent Computing

Some earlier work has been done to relate intelligent computing to manufacturing control charts. Most of these have taken the form of using expert, or knowledge-based, systems to select proper control methodologies and advise on the analysis of the selected methodologies (Alexander and Jagannathan 1986, Dybeck 1987, Scott and ElGomayel 1987, Evans and Lindsay 1988, Hosni and Elshennawy 1988, Eid and Losier 1990, Willborn 1990, Dagli and Smith 1991, Smith and Yazici 1992). These systems do a good job where the problem involves analysis of relatively few alternatives and the analysis depends mostly or wholly on qualitative information. Thus, selection of the proper control chart or advice on corrective actions given a certain diagnosis are appropriate venues for the expert system approach. However for an expert system to handle control chart data gathering and analysis, a scheme to codify the pattern analysis must be devised and incorporated as part of the system logic. This is a formidable barrier to successful use of expert systems for control chart interpretation.

A simple use of neural networks for control charts is to simply act as barriers for signaling a point within control limits or beyond control limits. This was done by (Yazici and Smith 1992) for a plastic extrusion process. A neural network acted as X-bar control chart monitor for four process variables simultaneously. The neural network was able to detect all instances of exceeding 3σ control chart limits. While the neural network did not offer detection advantages over analytical or human monitoring, the ability to simultaneously monitor multiple variables was unique, which would become particularly advantageous when using hardware neural network chips on the line for this purpose.

A few papers have used neural computing to detect location (usually mean) and variance (usually range) shifts. Pugh published two papers on using backpropagation networks to learn when mean shifts had occurred for a sample size of five (Pugh 1989, Pugh 1991). He found the neural networks produced average run length results about equal to a standard X-bar control chart with 2σ control limits, and improved significantly on Type II errors over X-bar charts. Guo and Dooley looked at positive shifts in both mean and variance using backpropagation neural networks compared to cumulative sum and moving sum charts (Guo and Dooley 1992). They found their best network reduced errors in classification about 40% from control chart heuristics. Hwarng and Hubele used a backpropagation classifier on six control patterns - trend, cycle, stratification, systematic, mixture and sudden shift (Hwarng and Hubele 1991). They found a neural classifier with binary input and output performed well enough to serve as a supplement to traditional control charts.

This paper extends the above research by considering a single model of a simultaneous X-bar and R chart, and investigating both location and variance shifts. Results are compared to standard Shewhart diagnostic rules. Furthermore, four typical control patterns are learned, using various sequential sample sizes and noise, for both classification and prediction, out to five time increments.

4. Interpretation of Process Shifts

The object of the initial research in this paper was to train neural networks using a rational subgroup size of 10 to recognize whether a controlled process had shifted, and if so, whether the mean or variance had changed. This was a *simultaneous* X-bar and R model. All samples were assumed to conform to rational subgrouping. Training sets of 1500 samples were created with the following components:

A.	500 Normal with $\mu = 0$ and $\sigma = 1$	(In Control)
	500 Normal with $\mu = 0$ and $\sigma = 2$	(Small Variance Shift)
	500 Normal with $\mu = 1$ and $\sigma = 1$	(Small Mean Shift)
B.	500 Normal with $\mu = 0$ and $\sigma = 1$	(In Control)
	500 Normal with $\mu = 0$ and $\sigma = 3$	(Large Variance Shift)
	500 Normal with $\mu = 3$ and $\sigma = 1$	(Large Mean Shift)

Figure 3 shows a training sample of 10 with the in control process distribution, and two altered ones (small mean and variance changes). All data was normalized to between 0.1 and 0.9 via the following procedure:

$$n_i = 0.1 + 0.8\left(\frac{x_i - x_{min}}{x_{max} - x_{min}}\right)$$

where n_i = the normalized value, x_i = the raw data value, x_{min}= the minimum raw data value in all the samples, and x_{max} = the maximum raw data value in all the samples. The normalization was required for proper handling by the unipolar sigmoidal transfer function which asymptotically approaches 0 and 1 in its extremes.

INSERT FIGURE 3 HERE

Training took place on randomly selected sets of 75% of the 1500 data and testing took place on the remaining 25% of the data. Efforts to train a network to recognize parameter shifts using raw data alone (the 10 sample points) were unsuccessful. These networks dealt with the variability of the inputs by training to produce identical outputs regardless of input. Statistical characterization of the raw data was needed to assist the network learn to discriminate between close distributions. Therefore, along with the 10 data points of each sample, the subgroup mean, standard deviation and range were calculated.

Successful networks were trained on either the raw data and the calculated statistics (a total of 13 inputs) or on the calculated statistics only (a total of 3 inputs). A single output was required for all networks with the following normalized coding: 0.5 for the in control distribution, 0.9 if the distribution had a shifted variance, and 0.1 if the distribution had a shifted mean. An alternative output scheme would be multiple binary coded neurons, as is done for shape detection in Section 5. Table 1 shows the performance statistics of two hidden layer networks on the test sets. One network architecture had 3 neurons in each hidden layer and the other had 10 neurons in each hidden layer. It has been generally recognized that a neural network must have enough hidden neurons to have the capacity to learn the model adequately, however overfitting the network with excess hidden neurons can impair generalization ability. It can be seen from Table 1 that a network found classification easier for distributions which are more radically shifted (set B), and it also assisted classifications for set A if all the data were included as inputs. For these sets, a larger network assisted classification as well, although not dramatically.

INSERT TABLE 1 HERE

It is important to compare the neural network results to those which would have been obtained using a standard Shewhart control chart strategy. This was done by calculating control chart 3σ limits for both X-bar and R charts for a subgroup size of 10, a mean of 0 and a standard deviation of 1. The limits were ± 0.949 about the mean of 0 for the X-bar chart and ± 3*0.797 about an expected range of 3.078 for the R chart using accepted control chart practices (Mendenhall and Sincich 1992). Test sets A and B were applied to these control limits, and a decision of out of control was reached if any point exceeded the 3σ limit.

The neural networks' decisions were interpreted according to the following schedule: 0 to 0.3 is a 0.1 (mean shift), 0.3 to 0.7 is 0.5 (no shift), 0.7 to 1.0 is 0.9 (variance shift). This type of interpretive schedule is typical when dealing with neural network output since it is usually imprecise. Table 2 shows the percentage of wrong decisions for the Shewhart control chart strategy and the neural networks.

INSERT TABLE 2 HERE

For a large shift in mean or variance, the neural networks performed comparably to a standard control chart, however when the shift was more subtle, and the extra information of the raw data itself was provided, the neural networks made slightly more than half the errors of the control charts, i.e. 16.8 and 11.7 percent incorrect for the two neural networks with all the inputs versus 27.5 percent incorrect for the control chart. The type of errors of each formulation was also different. In our experiments the manual control chart made primarily Type II Errors, that is missed disturbances. The neural network made both Type I and Type II Errors. This improvement of the neural networks' Type II Error rate relative to the manual control charts' Type II Error rate substantiates the results of Pugh (Pugh 1991) discussed in Section 3.3. The networks were trained with the assumption that the penalties for Type I and Type II Errors were identical, and therefore made both kinds of errors in approximately equal proportions. This could easily be remedied to reflect different relative penalties for Type I and Type II Errors by shifting the interpretive schedule for an out of control decision, that is moving the decision boundary further from the class with the lower relative error penalty. This procedure and the resulting power curves for neural networks are discussed fully in (Twomey and Smith 1993).

Most neural network classification errors were outputs very close to the decision schedule borders, i.e. near 0.3 or 0.7. Over the six networks, only 1.6% of the test set were clear misses, that is misclassified by as much as 0.5. The errors near decision borders could be analyzed more closely by human experts or with neural prediction of points, as discussed in Section 6, to achieve even better precision.

5. Interpreting Long Term and Cyclical Drifts

Besides discriminating process shifts by samples, recognizing patterns of samples as symptomatic of process problems is important. This is usually done manually by examining sequential points and looking for runs in certain areas of the control chart, or looking for an enduring pattern. These patterns can indicate tool wear, shift or operator differences, or other undesirable conditions.

Nine neural networks were trained to recognize the following shapes: flat (in control), sine wave (cyclic), sloped (trend or drift) and bimodal (stratification). Specific parameters concerning the shapes were selected randomly for each training and testing vector within pre-specified bounds as described below. This variety of training and testing shape parameters was used to facilitate generalization of the trained network to many specific instances of the shapes during operation. The following describes the generation of the sine wave, trend and bimodal shapes. The sine wave cycle frequency was fixed at approximately 65 increments, that is it would take 65 sequential observations to move completely through one cycle. The point at which the cycle began was randomly selected for each training and testing vector. The sign of the trend was randomly selected for each training and testing vector, while the absolute value of the slope was fixed at 0.1 (10% of σ). Therefore the network was trained to recognize the trend pattern as a noisy linear movement, rather than in a specific direction. The bimodal shape was selected to have random durations at each extreme for each training and testing vector.

Each shape was subjected to Normally distributed noise with $\mu = 0$ as illustrated in Figure 4. The networks differed by the number of sequential training inputs, which ranged from 5 to 20, and the standard deviation of the noise, which ranged from 0.1 to 0.3. Each network was trained on 1200 vectors (300 of each shape). Networks were tested on 300 randomly selected different vectors, also equally divided into the four shapes. A two component binary vector was used to classify patterns with the following code: flat (0 0), sine wave (1 1), trend (0 1) and bimodal (1 0). An output between 0 and 0.4 was considered a 0 and an output between 0.6 and 1.0 was considered a 1.

INSERT FIGURE 4 HERE

Table 3 summarizes the results of these nine networks for recognizing patterns and shifts. Networks improved classification when data was less noisy and additional points were available to help ascertain the pattern. All in all the networks did very well, classifying 69.3% correctly at the worst case (5 data points with maximum noise) and 99.0% correctly at the best case (20 data points with minimum noise). The ease of each shape classification changed with noise and number of input points. For low noise all misclassifications were either sine wave or bimodal. For noisier data with 5 data input points, the flat shape was the most difficult to classify correctly. For noisier data with 10 or 20 input points, the sine wave shape was the most difficult to classify correctly.

INSERT TABLE 3 HERE

6. Control Predictions To Assist Interpretation

Neural networks have been noted to perform well as predictors (White 1988, Sharda and Patil 1990, Tang et al. 1990). This ability can be taken advantage of when using neural control charts. Predicted points can be analyzed to see if they support the identified cycle, drift or trend, or if they indicate continued acceptable randomness of the process. This paper examines prediction of points for the shapes described in Section 5 with the future aim of using the predictions to confirm shape identification, especially in conditions near decision borders. For example distinguishing between a bimodal and a sine wave pattern may be assisted by predicting the next few points using the most recent observations.

Two groups of predictive networks were tested, each predicting equal numbers of the four shapes described in Section 5. Data was generated in the same manner as described in Section 5 except that the desired network response was the next sequential point(s) in the pattern rather than a binary classification. Note that this predictive task is quite difficult. First the network has to internally identify the input vector's pattern class from among the four classes (flat, trend, sine wave or bimodal), and then predict the next point(s) for that identified pattern and the specific input vector's parameters.

The first group of predictive networks predicted a single future point using 20 sequentially ordered inputs. The inputs had degrees of noise dispersion ranging from σ = 0.05 to σ = 0.2. Table 4 shows the percentage of correct predictions (those within 10% of actual), the RMS error and σ of error for the random test set. As expected, noise impeded prediction accuracy.

INSERT TABLE 4 HERE

The second group of networks predicted from one to five future points *simultaneously* based on 20 sequentially ordered inputs with σ = 0.1, i.e. the first network predicted the next sequential point, the second network predicted the next two sequential points, and so on, until the fifth network predicted the next five sequential points of the pattern. Identical training and testing sets were used on these five networks. Table 5 shows the results of these networks. To be considered correct all simultaneous points has to be predicted to within 10% of the actual value. Accuracy declined with multiple points, however RMS error and error dispersion were fairly constant indicating that individual prediction performance on more future points was not impaired. This relatively good predictive performance of the neural networks offers the potential that predictions can be used as further evidence in borderline cases for decisions concerning process state.

INSERT TABLE 5 HERE

7. Conclusions and Discussion

This paper has demonstrated that neural networks can be comparable to Shewhart X-bar and R control charts for large shifts in mean or variance, and can out perform them for small shifts. Neural networks work best when they have benefit of both raw sample data and sample statistics, although the statistics themselves are adequate to detect large shifts. A significant benefit of the neural approach is that a single network can model multiple control strategies simultaneously; in this paper, X-bar and R charts in a single network. More work on the crossover point where neural networks can become superior to traditional X-bar and R control chart interpretation is needed, as is the sensitivity of neural networks to the number and kind of input statistics calculated from the raw data and the subgroup size of the raw data itself. We are pursuing research with smaller subgroup sizes, deviations from the rational subgrouping assumptions, and multiple shift distributions in a single training set, all more characteristic of actual conditions found in industry. Another area for futher study is the comparison of neural control charts to Shewhart control charts with other heuristic rules such as those in (Nelson 1984).

For shape interpretation and prediction, networks performed best with minimal noise and maximum number of inputs. All neural networks proved capable of good quality decisions regarding pattern identification even in light of sparse and noisy data. The predictive networks could simultaneously predict out to five sample increments without much lost of accuracy. We are pursuing work on comparing the shape recognition performance of the neural networks to human judgment, heuristic run rules, and on looping back point predictions to the shape identification neural networks to ascertain their value for adding evidence to correctly classify difficult samples, i.e. those near decision borders.

Given the widespread use of control charts in both manufacturing and service industries, and the current difficulties with proper interpretation of plotted results, the promising results of a neural computing approach bear further research. Since neural networks are commonly available in software form running on PC and workstation platforms, and will soon be practical in VLSI format, they are viable options for statistical control in production environments.

Acknowledgments

The author would like to acknowledge the helpful comments and suggestions of two anonymous reviewers.

References

Alexander, S. M. and V. Jagannathan, 1986, Advisory system for control chart. *Computers & Industrial Engineering*, **10**, 171-177.

Andersen, Kristinn, George E. Cook, Gabor Karsai and Kumar Ramaswamy, 1990, Artificial neural networks applied to arc welding process modeling and control. *IEEE Transactions on Industry Applications*, **26**, 824-830.

Burke, Laura I., 1989, *Automated Identification of Tool Wear States in Machining Processes: An Application of Self-Organizing Neural Network.* Ph.D. Thesis, University of California - Berkeley.

Dagli, Cihan H. and Alice E. Smith, 1991, A prototype quality control expert system integrated with an optimization module. *Proceedings of the World Congress on Expert Systems*, 1959-1966.

Dybeck, Martin, 1987, Taking process automation one step further: SPC. *Proceedings of the Sixth Annual Control Engineering Conference*, 643-651.

Eid, M. S. and G. Losier, 1990, QCMS: a quality control management system. *Computers & Industrial Engineering*, **19**, 495-499.

Evans, James R. and William M. Lindsay, 1988, A framework for expert system development in statistical quality control. *Computers & Industrial Engineering* **14**, 335-343.

Funahashi, Ken-ichi, 1989, On the approximate realization of continuous mappings by neural networks. *Neural Networks*, **2**, 183-192.

Gibra, Isaac N., 1975, Recent developments in control chart techniques. *Journal of Quality Technology* **7**, 183-192.

Guillot, Michel and Abderrazak El Ouafi, 1991, On-line identification of tool breakage in metal cutting processes by use of neural networks. In *Intelligent Engineering Systems Through Artificial Neural Networks* (C. Dagli, S. Kumara and Y. Shin, Editors) (New York: ASME Press), 701-709.

Guo, Yuan and Kevin J. Dooley, 1992, Identification of change structure in statistical process control. *International Journal of Production Research*, **30**, 1655-1669.

Hornik, Kurt, Maxwell Stinchcombe and Halbert White, 1989, Multilayer feedforward networks are universal approximators. *Neural Networks*, **2**, 359-366.

Hosni, Yasser A. and Ahmad K. Elshennawy, 1988, Knowledge-based quality control. *Computers & Industrial Engineering*, **15**, 331-337.

Hwarng, H. Brian and Norma F. Hubele, 1991, X-bar chart pattern recognition using neural nets. *ASQC Quality Congress Transactions*, 884-889.

Knapp, Gerald M. and Hsu-Pin Wang, in press, Machine fault classification: a neural network approach. *International Journal of Production Research*.

Lapedes, Alan and Robert F. Farber, 1988, How neural nets work. In *Neural Information Processing Systems* (Dana Z. Anderson, Editor) (New York: American Institute of Physics), 442-456.

Mendenhall, William and Terry Sincich, 1992, *Statistics for Engineering and the Sciences* (San Francisco: Dellen Publishing).

Miller, Richard K. and Terri C. Walker, 1988, *Artificial Intelligence Applications in Manufacturing* (Madison, GA: SEAI Publications).

Montgomery, Douglas C., 1980, The economic design of control charts: a review and literature survey. *Journal of Quality Technology*, **12**, 75-87.

Nelson, L. S., 1984, The Shewhart control chart - tests for special causes. *Journal of Quality Technology*, **16**, 237-239.

Okafor, A. C., M. Marcus and R. Tipirneni, 1990, Multiple sensor integration via neural networks for estimating surface roughness and bore tolerance in circular end milling. *1990 Transactions of NAMRI/SME*, 128-137.

Pugh, G. Allen, 1989, Synthetic neural networks for process control. *Computers & Industrial Engineering*, **17,** 24-26.

Pugh, G. Allen, 1991, A comparison of neural networks to SPC charts. *Computers & Industrial Engineering*, **21**, 253-255.

Rumelhart, D. E., J. L. McClelland and the PDP Research Group, 1986, *Parallel Distributed Processing: Explorations in the Microstructure of Cognition, Volume 1* (Cambridge, MA: MIT Press).

Saniga, Erwin M. and Larry E. Shirland, 1977, Quality control in practice...a survey. *Quality Progress,* **10**, 30-33.

Scott, L. L. and J. I. El Gomayel, 1987, Development of a rule based system for statistical process control chart interpretation. In *Quality: Design, Planning, and Control* (R. E. DeVor and S. G. Kapoor, Editors) (New York: ASME), 73-91.

Sharda, Ramesh and Rajendra B. Patil, 1990, Neural networks as forecasting experts: an empirical test. *Proceedings of the International Joint Conference on Neural Networks*, II-491-494.

Shewhart, W. A., 1931, *Economic Control of Quality of Manufactured Product* (Princeton, NJ: Van Nostrand Reinhold).

Smith, Alice E., 1993, Predicting product quality with backpropagation: a thermoplastic injection molding case study. *International Journal of Advanced Manufacturing Technology,* **8**, in press.

Smith, Alice E. and Cihan H. Dagli, 1991 A, Controlling industrial processes through supervised, feedforward neural networks. *Computers & Industrial Engineering,* **21**, 247-251.

Smith, Alice E. and Cihan H. Dagli, 1991 B, Relating binary and continuous problem entropy to backpropagation network architecture. In *Applications of Artificial Neural Networks II* (Bellingham, WA: SPIE), 551-562.

Smith, Alice E. and Hulya Yazici, 1992, An intelligent composite system for statistical process control. *Engineering Applications of Artificial Intelligence*, **5**, 519-526.

Tang, Zaiyong, Chrys de Almeida and Paul Fishwick, 1990, Time series forecasting using neural networks vs. Box-Jenkins methodology. *Proceedings of the First Workshop on Neural Networks: Academic/Industrial/NASA/Defense*, 95-100.

Twomey, Janet M. and Alice E. Smith, 1993, Power curves for pattern classification networks. *Proceedings of the 1993 IEEE International Conference on Neural Networks,* in press.

Werbos, Paul J., 1974, *Beyond Regression: New Tools for Prediction and Analysis in the Behavioral Sciences,* unpublished Ph.D. Thesis, Harvard University.

White, Halbert, 1988, Economic prediction using neural networks: the case of the IBM daily stock returns. *Proceedings of the International Joint Conference on Neural Networks,* I-451-458.

Willborn, Walter W. O., 1990, Expert systems in support of quality management. *ASQC Quality Congress Transactions*, 758-763.

Wu, Hsien-Jung, Cheng-Shin Liou and Hsu-Heng Pi, 1991, Fault diagnosis of processing damage in injection molding via neural network approach. In *Intelligent Engineering Systems Through Artificial Neural Networks* (C. Dagli, S. Kumara and Y. Shin, Editors) (New York: ASME Press), 645-650.

Yazici, Hulya and Alice E. Smith, 1992, A composite system approach for intelligent quality control. *Proceedings of the IIE Research Conference*, 325-328.

Table 1. Network Performance on Test Sets A and B.

Training/ Test Set	Inputs	Hidden Layer Size	RMS Error	Mean Abs. Error	σ of Error
A	All	3	0.1783	0.1308	0.1213
A	All	10	0.1619	0.0777	0.1420
A	Statistics	3	0.2109	0.1597	0.1378
B	All	3	0.0438	0.0233	0.0371
B	All	10	0.0302	0.0119	0.0278
B	Statistics	3	0.0210	0.0118	0.0174

Table 2. Wrong Decisions of Control Charts and Neural Networks.

Training/ Test Set	Inputs	Hidden Layer Size	Neural Net Percent Wrong	Shewhart Control Percent Wrong
A	All	3	16.8	27.5
A	All	10	11.7	27.5
A	Statistics	3	28.5	27.5
B	All	3	0.8	0.5
B	All	10	0.5	0.5
B	Statistics	3	0.0 (None)	0.5

Table 3. Pattern Recognition Networks

# Input Points	σ of Noise	# Correct of 300	# Missed of 300			
			Flat	Slope	Sine wave	Bimodal
5	0.1	288	0	0	10	2
5	0.2	226	39	15	10	10
5	0.3	208	32	39	18	3
10	0.1	292	0	0	5	3
10	0.2	268	3	7	16	6
10	0.3	246	9	17	23	5
20	0.1	297	0	0	0	3
20	0.2	293	0	0	0	7
20	0.3	280	6	1	1	12
Over All Networks		88.8%	3.3%	2.9%	3.1%	1.9%

Table 4. Predictive Networks With Varying Data Dispersion.

σ of Noise	% Correct	RMS Error	σ of Error
0.05	82.0	0.7487	0.6244
0.10	79.7	0.7698	0.6442
0.20	72.7	0.8227	0.6400

Table 5. Predictive Networks With Varying Number of Predictions

# Predictions	% Correct	Average RMS Error	Average σ of Error
1	80.3	0.7690	0.6457
2	72.3	0.7677	0.6263
3	66.0	0.7006	0.5498
4	66.3	0.6609	0.5219
5	62.0	0.6485	0.5064

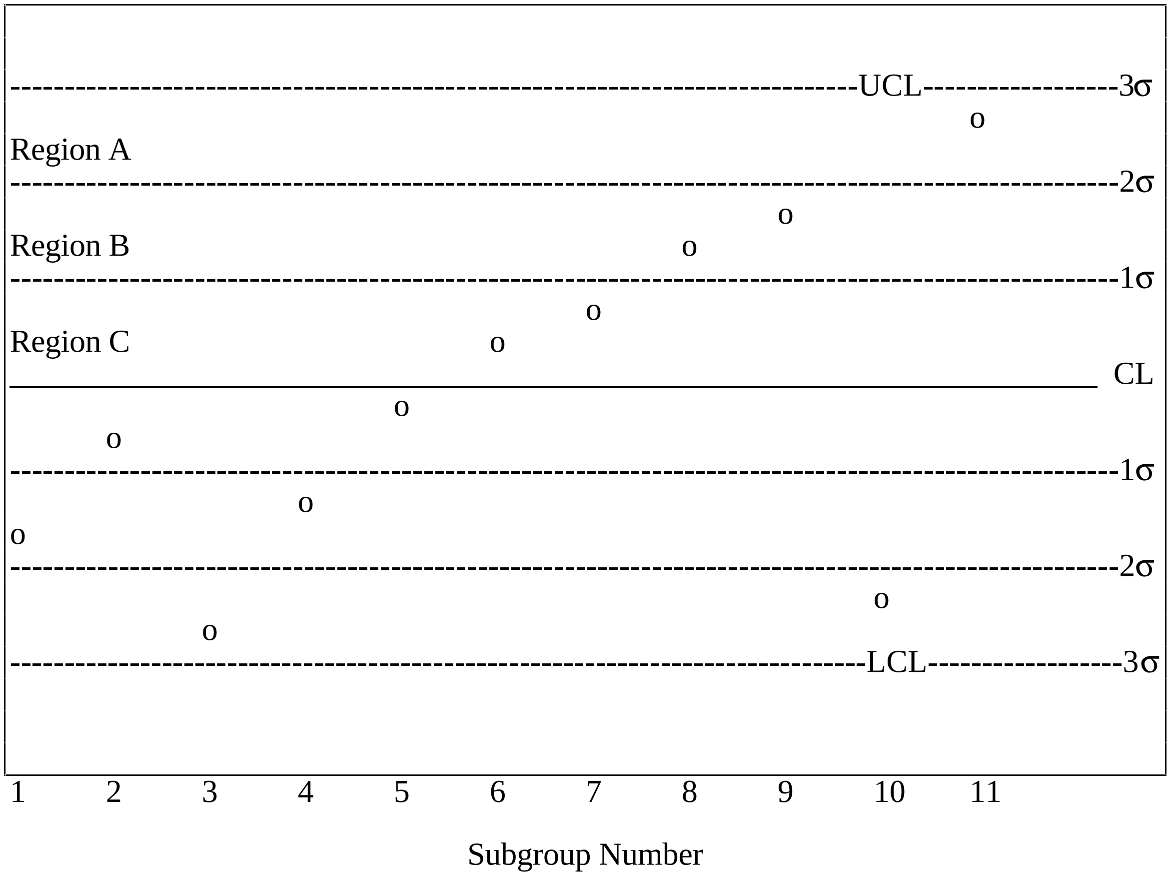

Figure 1. Typical Shewhart Control Chart.

where the output of each neuron = $\frac{1}{1+e^{-\sum wx}}$ and *wx* = weighted sum of inputs to that neuron.

Figure 2. Backpropagation Neural Network With Unipolar Sigmoidal Transfer Function.

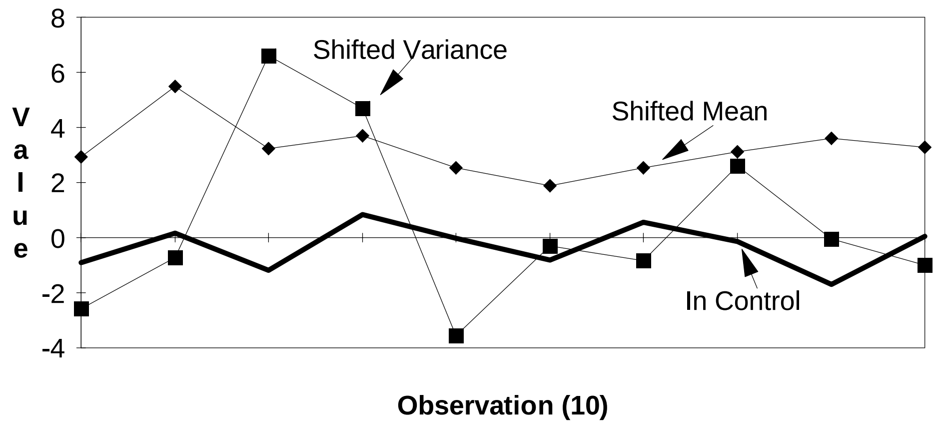

Figure 3. A Randomly Generated Sample of 10 of Non-Shifted Distribution, Small Mean Shift and Small Variance Shift.

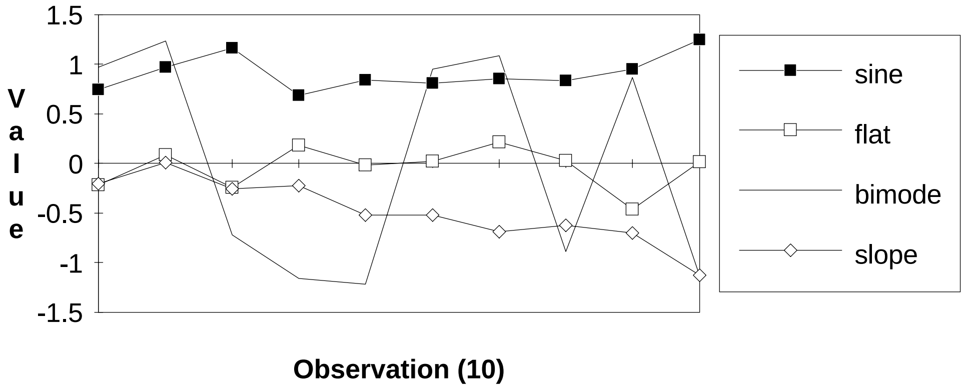

Figure 4. A Randomly Generated 10 Sample Sequential Input for Four Symptomatic Shapes With Medium Noise ($\sigma = 0.2$).

Figure Captions

Figure 1. Typical Shewhart Control Chart.

Figure 2. Backpropagation Neural Network With Unipolar Sigmoidal Transfer Function.

Figure 3. A Randomly Generated Sample of 10 of Non-Shifted Distribution, Small Mean Shift and Small Variance Shift.

Figure 4. A Randomly Generated 10 Sample Sequential Input for Four Symptomatic Shapes With Medium Noise ($\sigma = 0.2$).