THE APPLICATION OF NEURAL NETWORKS TO PRODUCTION PROCESS CONTROL

A dissertation submitted to
Kent State University Graduate School of Management
in partial fulfilment of the requirements
for the degree of Doctor of Philosophy

by

James H. Hamburg

March, 1996

UMI Number: 9623042

Dissertation written by

James H. Hamburg

B.S.E.E, University of Wisconsin-Milwaukee, 1972

P.E., Electrical, State of Ohio, 1979

M.S.E.E., The University of Akron, 1983

M.B.A., Baldwin-Wallace College, 1985

Ph.D., Kent State University, 1996

Approved by

_______________ Co-Chair, Doctoral Dissertation Committee

_______________ Co-Chair, Doctoral Dissertation Committee

_______________ Members, Doctoral Dissertation Committee

____

____

Accepted by

_______________ Doctoral Director, Graduate School of Management

_______________ Dean, Graduate School of Management

ACKNOWLEDGMENTS

I would like to take this opportunity to acknowledge my indebtedness to all of my committee members. I would like to thank Dr. Gregory Madey for his assistance in providing access to his personal library for my literature review. Dr. Michael Hu must be acknowledged for his support in the experimental design formulation. Special thanks must be extended to Dr. G. J. Weinroth for his valuable insight into simulation, creative ideas, and critiques of my dissertation. Above all, I want to thank Dr. D. E. Booth for his endless support, guidance, and help in the preparation of this dissertation.

I also would like to take this opportunity to thank Dr. S. Prasad for his insights into the joint estimation procedure. Additionally, Dr. J. Denton needs to be acknowledged for his assistance and use of his neural network algorithm.

Finally, I would like to express my love and appreciation to my wife. This dissertation would never have been possible without her invaluable assistance and support.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

The goal of this dissertation was to develop a neural network that can be used to control real-time production processes. In production and nuclear processes, the monitoring of product quality and the safeguarding of materials are crucial due to today's business climate and world events. Complexities in dealing with these two aspects arise because of observation interdependency and the influence of process disturbances, i. e. outliers. Unfortunately, the independent and identically distributed assumption of traditional control chart methods isn't always applicable. Therefore, the importance of this research in achieving its goal lies in its ability to apply the neural network algorithm with simulation approach enhancement to various process control applications especially those which are autocorrelated.

In achieving this goal certain objectives were met. The first objective was to select an appropriate neural network architecture for this research. The second objective was to investigate the neural network's ability to interpret control chart data. The third objective was to determine the neural network's ability to detect process disturbances, i. e. outliers. The fourth objective was to compare the neural network method against other process control methods based on the same process data sets. The results of meeting these objectives indicate the strong potential of implementing the neural network algorithm with the simulation approach enhancement to real-time on-line applications.

Nuclear industry applications were first examined in the research because it is of critical importance to monitor and safeguard nuclear materials. In Scientific American's January 1996 issue, the cover story indicated that approximately two hundred instances of smuggling nuclear materials out of the former Soviet Union have occurred. The actual data sets chosen for this research came from a study at the AGNS Barnwell Nuclear Fuels plant and a report done by the Energy Research and Development Administration. The neural network algorithm with the simulation approach enhancement was applied to these data sets. The results were compared against the results of other control methods (Joint Estimation, Data Bounding, and Polynomial Smoothing) based on these data sets. The neural network algorithm with the simulation approach enhancement was found to be equivalent or better than current methods in the detection of outliers and the recognition of terminal points. Thus, it is believed that the neural network approach is a strong candidate for use in these applications.

Production process applications were next reviewed. In production processes the monitoring of product quality is crucial due to today's business climate. The actual data sets chosen for this section of the research came from a process of dyeing woolen yarn, a process in which pump bore holes are cut into automobile diesel engine blocks, a continuous sheet-like process, and a transmission parts manufacturing process. The neural network method with the simulation approach enhancement was applied to these data sets and the results were compared against the results of the other methods also applied to these data sets. Again, the neural network algorithm with the simulation approach enhancement was found to be equivalent or better than current methods in the detection of outliers and the recognition of terminal points. Thus, again, it is believed that the neural network approach is a good alternative for use in these applications.

The dissertation presents the problem statement, the background for the study, its significance, and a thorough literature review. Next, the simulation computer program, the data sets used in this research, and the neural network algorithm are described. The applications and the evaluations of the other control methods used in this research are discussed. Finally, the importance of the research and possible future research endeavors are presented.

This addition to process control should provide the control industry and operating managers involved in production processes or nuclear material processes a useful tool to confront the events occurring in the world today.

CHAPTER 1: INTRODUCTION

Chapter Overview

In this chapter, the problem statement is presented along with a basic overview of the research. The general topic of neural network techniques is described and the significance of the topic is discussed. The primary focus of the literature review is explained. Finally, the dissertation structure is established.

Problem Statement

In recent years, neural networks have emerged as an important tool in the growing area of artificial intelligence. They represent a promising paradigm in computational technology that enhances many applications. As a result of a model of the learning process being inspired by the human brain, neural networks can handle complex tasks efficiently and effectively. Thus, the ground work is in place to continue research into the theoretical development and the application of this method to real world problems.

This dissertation investigates the benefits of using a neural network to analyze data to determine when abnormalities (outliers) occur due to a production process change or disruption. Currently, production process changes or disruptions are identified and corrected by process control techniques. In this dissertation, process control is defined as maintaining the performance of a production process at its ultimate capability level. The process capability level is the producing of a uniform quality product (DataMyte Corporation, 1989). At least six techniques are currently used in industry for process control purposes. These techniques are time series analysis, Proportional-Plus-Integral-Plus-Derivative (PID), Statistical Process Control (SPC), Batch Control, Cascade Control, and Statistical Quality Control (SQC).

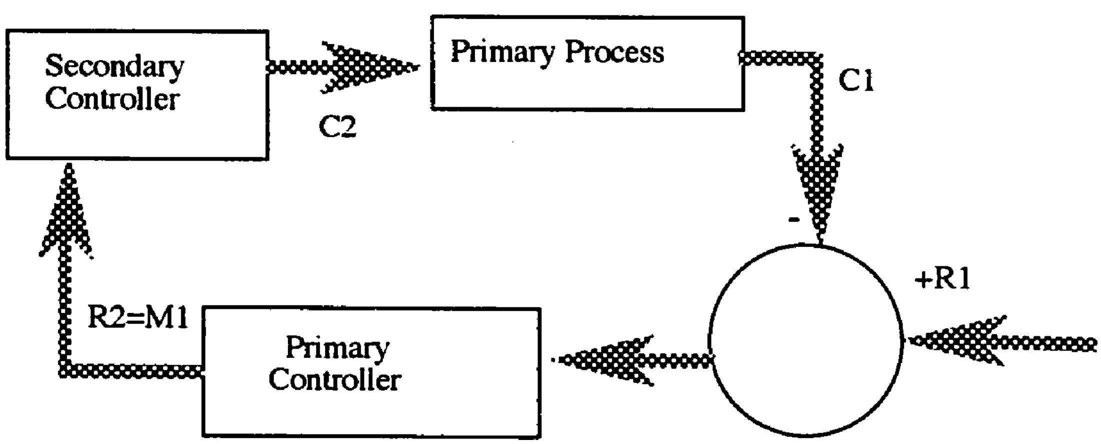
Time series based control methods have been applied to advance the abilities of SPC for controlling a process and detecting out-of-control situations (outliers). Alwan and Roberts (1988), Booth (1984; 1986) and Prasad (1990; 1993) have used time series models in their research on SPC.

The Proportional-Plus-Integral-Plus-Derivative (PID) controller is important because it permits the interaction of the proportional, integral, and derivative control modes to maintain a process variable (Shinskey, 1988). The initial settings for the three control modes are determined by the particular process and any further adjustment of these settings is done by an operator based on his/her own knowledge of the process. Thus, the accuracy of this control technique can be question because of the introduction of human error.

Statistical Process Control (SPC) is defined as the use of statistical methods, especially control charts, for analyzing and controlling a production process (DataMyte Corporation, 1989). This method is further described in chapter two.

Batch control is defined as the use of a controller to maintain a discontinuous process characterized as having a zero load. Zero load is the period between the completion of one batch and the start of another batch. If a process variable overshoots its set point (the controlling value for the variable, example: 95 degrees or 6 pounds of limestone) during a batch, the overshoot is permanent for that batch until the next batch is started. To correct for this occurrence, Proportional-Plus-Derivative control is currently being used (Shinskey, 1988).

Cascade Control is a method in which the output of one controller is fed into a second controller to manipulate its setpoint and can be represented by the following block diagram. The first advantage of using this technique is that the disruptions in the secondary controller are corrected before influencing the primary process. Another advantage is that a secondary controller improves the speed of the primary controller's response. A final advantage is that it provides the means for high performance control in the face of random disturbances and undue phase shift (90 degree lag) in the secondary controller (Shinskey, 1988).



Statistical Quality Control (SQC) is the application of statistical methods for measuring and improving the quality of processes. Statistical Process Control (SPC) is the application of statistical methods for analyzing and controlling variations in the process. Therefore, SPC is one procedure employed by SQC to maintain the quality control of the production process (DataMyte Corporation, 1989).

Neural networks have the potential to provide improvements in production process control. These improvements are based on neural network properties such as adaptive learning, real-time operation, the elimination of a required operating expert, and a quick response to a process disruption. Thus, neural networks have the potential of providing a highly accurate method for production process control.

Background

This research examines several applications of neural networks in process control, including important control applications in the nuclear industry. Also, several neural network methods will be discussed.

The nuclear industry was selected as a prime example of the process control industry because of the critical importance in controlling nuclear processes (reactors, material disposal, etc.), providing safeguards, and monitoring nuclear material inventories. In 1982, Goldman et. al. reviewed the origins and development of nuclear material safeguards along with the role of statistics. They noted that the formation of the International Atomic Energy Agency (IAEA) in 1957 was a major breakthrough concerning nuclear material safeguards. The IAEA was given the mission to independently verify the quantity of material unaccounted for over a specified period and establish the limits of accuracy for all nuclear facilities. Also, Goldman et. al. discussed four common accounting problems which included shipper-receiver differences, temporarily distributed material balances, spatially distributed material balances, and data verification. Speed and Culpin (1986) concurred with Goldman et.

al. in their review. They further examined various statistical methods which included the likelihood ratio, CUMUF test, page test, power-one type test and robust test, and found problems with them all in dealing with the diversion of nuclear material for unauthorized use. Speed and Culpin also noted that none of these methods took into account the cost of a false alarm concerning a diversion. Further, they stated that Shewhart control charts traditionally used in statistical quality control and time series techniques proved to be of limited or no real value. Thus, the results of the examinations on these techniques indicated that a problem exists currently in meeting the needs of the nuclear industry for safeguarding nuclear material.

In the recent article by B. Hileman (1994) the international crisis concerning the monitoring and safeguarding of nuclear material in relationship to the break-up of the Soviet Union and the lack of cooperation by North Korea is discussed. Hileman pointed out that the National Academy of Sciences (NAS) warns in a current report that a "clear and present danger to national and international security is posed." The experts agree safeguarding and monitoring of nuclear material could be done fairly well in the U.S. However, within the former Soviet Union and North Korea they believe it is questionable. The article further presents long and short term proposed disposal methods. The short term approaches include bilateral monitoring, former Russian states stopping the production of weapon grade materials, burying, burning, making glass logs, or U.S. purchasing Russian nuclear waste. The long term goal is to develop reactors for the disposal of nuclear materials. There are four types of reactors under consideration for this disposal. They are the advanced light-water reactors, the modular high-temperature gas-cooled reactors, the advanced liquid-metal reactors, and the accelerator-based conversion reactors. Presently, the U.S. and Russia have developed certain bilateral arrangements on nuclear material.

More recently, articles have appeared in the magazines Time and Newsweek concerning the theft and selling of nuclear material. In the Time article, Nelan (1994) described examples of stolen nuclear material from Russia being sold in Germany. The reason given for these thefts is that the nuclear material safeguards in Russia are being undermined by personnel having their pay reduced, being paid late, or not being paid at all. The fear expressed in the article is that the buyers of this nuclear material are either terrorist groups, or more likely, countries like Iran, Iraq, Libya, Pakistan, or North Korea. However, the real problem is that Russian officials will not admit their current safeguard methods are seriously flawed. As a result the National Academy of Sciences has developed recommendations to deal with the build up of Plutonium in the U.S. and Russia. Despite this concern for safeguarding nuclear material, the Russians are still skeptical of the West as they continue to view their Plutonium as a national treasure.

New process control techniques are being examined by industry, and neural networks may have the potential to be extremely important to the nuclear industry. As a result, this dissertation will examine neural networks to determine if they provide such a new technique that the nuclear industry and others will be able to use.

An example of the neural network approach is the GRG2 paradigm which in previous research has demonstrated successful performances (Denton, 1991; Osyk, 1991; and Kang, 1991). In that research, it was found that the GRG2 technique provided better scalability, faster results, and better quality solutions than other neural network methods (Subramanian, and Hung, 1991). Then too, it was found in previous research that the GRG2 technique was superior to an alternate technique, Back Propagation, not only in amount of time to train a network to arrive at a higher quality solution, but also in scalability relative to problem size (Hung and Denton, 1991). The exceptional performance of the GRG2 technique was also studied by comparing three benchmarks (parity, double spiral, and binary encoder). Parity was used to determine whether an n-vector of binary digits had an odd or even sum. The result of using this benchmark was that the GRG2 technique was faster and produced a better quality solution with smaller variability than Back Propagation. The double spiral benchmark consisted of data points on two interleaved spirals going around the origin three times. The result of using this benchmark was that the GRG2 technique was able to solve a problem that the Back Propagation network failed to solve because the objective function value fluctuated around 14.99 for several hundred iterations. The binary encoder benchmark occurs when a set of input patterns are mapped to a set of identical output patterns by a small set of hidden units. The result of this benchmark was that the GRG2 technique was able to solve a problem that Back Propagation was not able to solve because the objective function value did not improve beyond the 17.88 threshold for several hundred iterations (Subramanian and Hung, 1991). Thus, this study does demonstrate GRG2's significant importance to neural network techniques. The GRG2 technique is discussed further in Chapter two.

Studying the feasibility of using neural networks for process control is important inasmuch as they have the potential of providing better methods for process control. Neural networks have certain advantages that other techniques do not. The network learns about the current process on its own as well as making adjustments in its learning to account for process changes. An overall mathematical or scientific theory is not important for the neural network to function. The type of product has no impact on the neural network performance because the network is directly controlling the process, and thereby indirectly maintaining the product quality. The network lessens the possibility of human error because there is no necessary intervention or interaction by a person. It anticipates the abnormalities (outliers) in order to correct for them within the process because it has the ability to learn and adjust to process changes without outside interference. It removes the need for trial and error in establishing the optimum control for the process because it learns the information from the process itself (Blaesi and Jensen, 1992). In general, neural networks have been seen to be able to solve the more difficult problems in less time than other conventional methods (Francett,1989).

However, there are problems in using neural networks. For example, neural networks sometimes need a large amount and a wide range of data to develop an accurate model (Blaesi and Jensen, 1992). Nevertheless, the justification for the study of neural networks in process control is that there are the potentials for better quality control, for higher productivity, for improved environmental conditions, for increased safety capabilities, for achieving the status required by governmental regulations, and for decreasing human intervention. Hopefully the goal, to be able to obtain faster and more complete solutions with neural networks responding to a process change, will be achieved.

Neural Network Technique

At this point, it would be useful to provide a general overview of neural networks. Initially, neural networks were inspired by biological systems, (See Figure

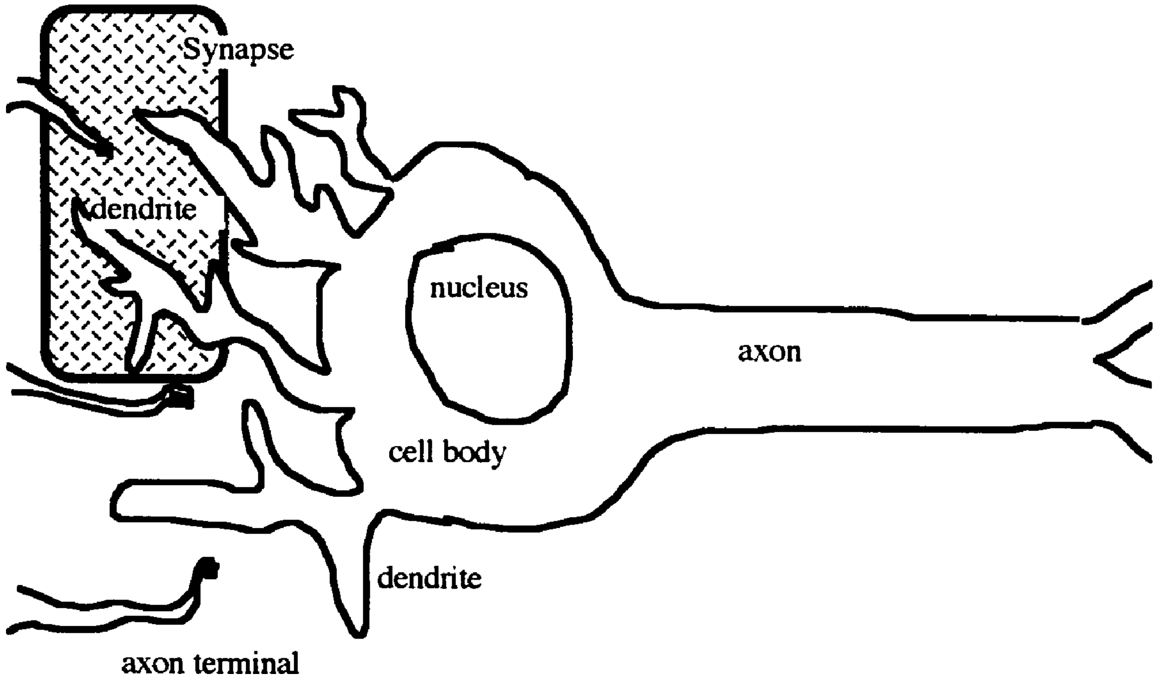


Figure 1: Biological System

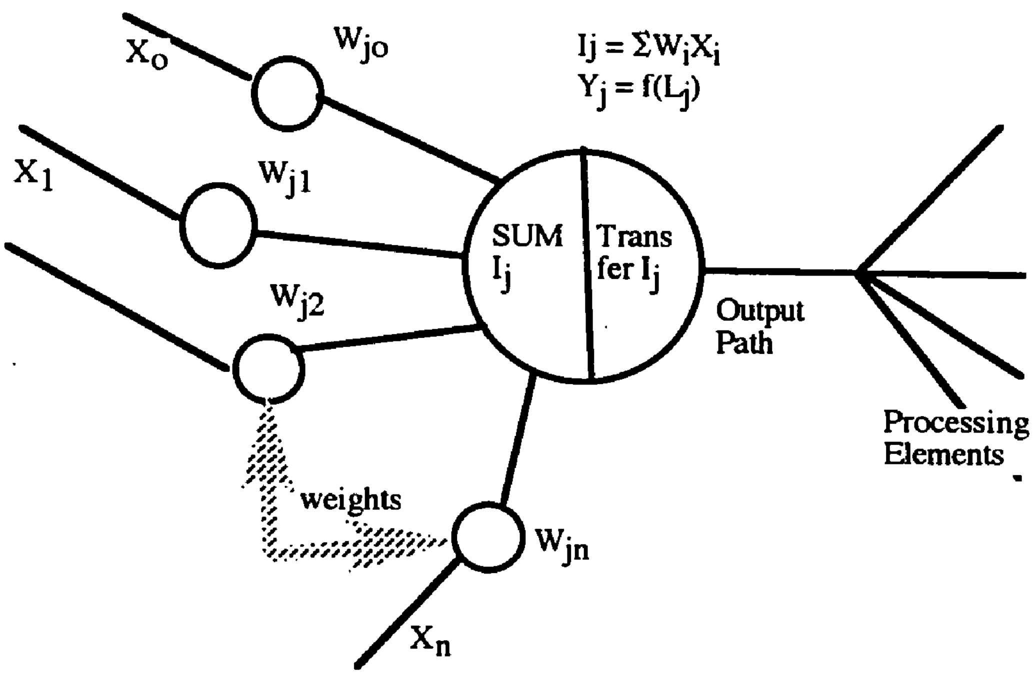


Figure 2: Neural Computing Network

1) in particular the human brain, which processes large amounts of information. The human brain is made up of at least ten billion neurons which are interconnected in a dense manner. Each neuron can be considered a microprocessing unit with a center (the nucleus). This nucleus receives and combines signals from various passive input channels which are called dendrites. If the combined signal to the nucleus is strong enough, the output channels, which are called axons, will be activated and information will be transferred. This transfer of information is done chemically and can be measured by electrical side effects. The axon from each neuron is connected to dendrites of other neurons by a junction called a synapse. At this junction the transfer of information relies on the amount of chemical released by the axon and received by the dendrites. As a result the synapse combined with the neuron form the basic memory mechanism of the brain (NeuralWare, Inc., 1991).

In the study of neural networks, the component analogous to a neuron is the processing element (PE) (See Figure 2). The processing elements of the network are the nodes that have various input paths and at least one output path. Input paths (X_n) are analogous to the dendrites. Each input path is multiplied by a weighted value. The weighted value is analogous to the synaptic strength. When all modified input paths reach the node, they are summed. This combination is multiplied by a transfer function which could be a sigmoid or hyperbolic tangent. The sigmoid function has this form.

$$f(x) = 1/(1 + e^{-x})$$

The hyperbolic tangent function has this form.

$$f(x) = (e^{x} - e^{-x})/(e^{x} + e^{-x})$$

The multiplied combination of the input paths by the transfer function is then directly sent to the output path. This output path (y_j) is analogous to the axon which can be either connected to other nodes as an input via the weighted values, or designated as the final output (NeuralWare Inc., 1991).

In recent years neural networks have become a popular technique for various applications. They date back to 1943 when McCulloch and Pitts issued a paper titled, "A Logical Calculus of Ideas in Nervous Activity" (Harston and Maren, 1992; NeuralWare, Inc., 1991). A more in depth historical review is provided in chapter two.

Of all the various neural network methods, the connectionist type, multilayer feedforward network (MFN), is the most widely chosen. This network (See Figure 3) consists of an input layer, an output layer, and at least one hidden layer. Each of these layers are composed of nodes. The various nodes are connected together by arcs. Information enters the input layer nodes. The input layer nodes multiply a transfer function to the information and then send the resultants via the arcs. These arcs have a weighted value which modifies the resultants before presenting it to the hidden layer node(s). The hidden layer node(s) provides the network with the ability to learn complex information from the input (Touretzky and Pomerleau, 1989; Qin, 1992). There are three basic types of hidden layer nodes. The first type receives a value from a single input. The second type receives values from a pair of inputs in the same direction. The third type receives values from a pair of inputs in the opposite direction (Malkoff, 1990). When the arcs reach the hidden layer nodes, they are summed. This summation is then adjusted by a transfer function and this value is sent out via another set of arc(s). The arc(s) also is modified by a weighted value and is connected to either

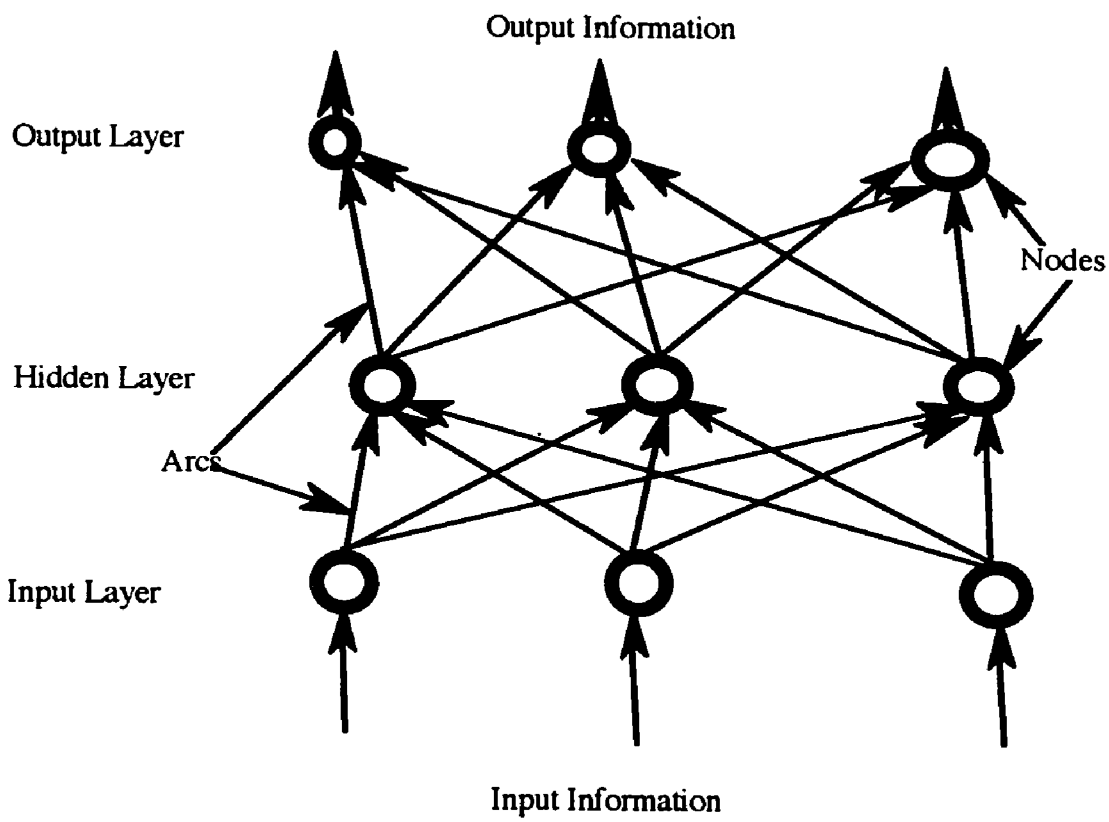


Figure 3: The MFN Neural Network

an additional set of hidden layer node(s) or the output layer node(s). The number of hidden layers is governed by the needs of the application. Finally, when the signals via the arc(s) reach the output layer node(s), they are summed. The summed signal is the final output information (Hinton, 1990; Qin, 1992). Other neural network methods are considered in chapter two.

The neural network technique has two phases. These phases are training and testing. In the training phase, data is presented to the network to achieve a particular outcome. The network learns to achieve the necessary outcome by adjusting the

RESEARCH OBJECTIVES
Short Term Select an appropriate neural network architecture Investigate its ability to interpret control chart data Determine its ability to detect process disturbances (outliers) Compare neural networks to time series and control chart methods based on the same process data sets
Long Term Demonstrate neural network's capabilities to control an unknown real-time process Develop a strategy for implementing the research goal

Table 1: Research Objectives

weighted values associated with the arcs. These adjustments are accomplished as the network sends the information from the input layer through the hidden layer to the output layer. When the network has been trained, other data sets are then presented to test the network's performance. This testing of the neural network determines whether the network can function to solve particular applications. This topic will be discussed further in chapters two and three.

This research investigates the learning ability of neural networks to anticipate production process disturbances (outliers). The study utilizes data sets from various industries, but in particular, the nuclear industry for application examples and an algorithm developed by Denton (1993) as the neural network paradigm. Thus, the focus of this dissertation is not only to determine whether or not neural networks can be implemented as a technique for production process control but also whether they can be used as an additional element to a nuclear material safeguard system.

Significance

The significance of this research can be seen in its ability to take process control "one step further into the future." It attempts to improve the quality control of both the product and process, allowing for real-time operation, and hopefully also reducing control cost. The goal of this research was to develop a neural network that can be used to control real-time processes. To achieve this goal, certain objectives were met (See Table 1).

By meeting the research goal and objectives, the process control industry is provided with a much needed tool for the future. A good example to demonstrate the potential of this research is contained in the article by Malkoff (1990). Malkoff developed a neural network algorithm called STOCHASM for real-time signal detection and classification. The algorithm along with sophisticated sensor technology has correctly classified signals embedded in noise and subject to considerable uncertainty within 500 milliseconds of signal detection. This feature can also enhance other potential investigations of applying neural networks in industrial applications. Hopefully, the neural network technique will enable the process control industry to free people from many of the mundane tasks in process control and allow them to concentrate on other more important tasks. In addition, the product is improved because the process is controlled more tightly and the reactions to process variations are more immediate. In turn, sales potential will increase due to providing a high quality product to the customer. Also, cost reduction is realized as industries increase productivity, while significantly reducing scrap, equipment down time, maintenance, and even manpower. This dissertation has the potential of providing a truly significant contribution in the application of neural networks to control real-time processes. Finally, it expands the knowledge base of the academic community.

Focus of the Literature Review

An extensive review of the current literature is highly important to lay the groundwork for developing the dissertation's direction. The literature review includes the topics of time series analysis, neural networks, and process control. In planning the literature review, applications were chosen to demonstrate how these various topics

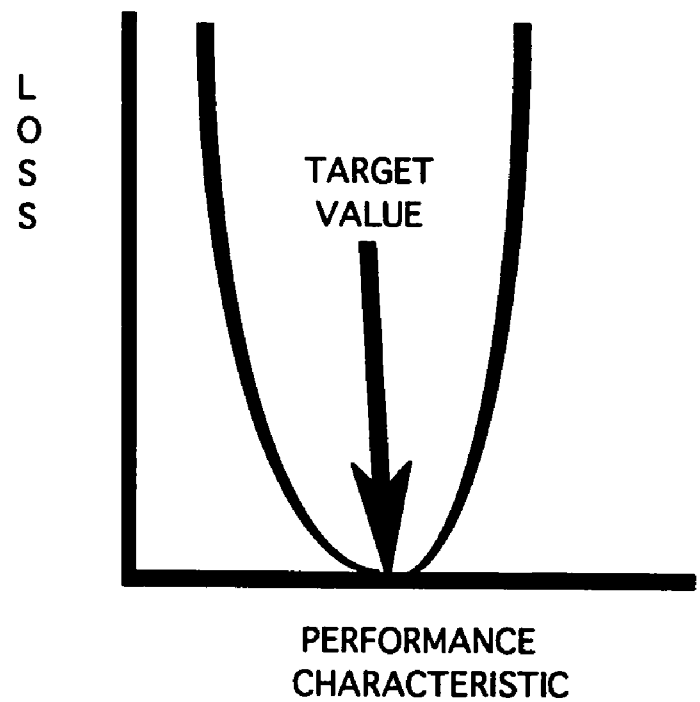


Figure 4: The Taguchi Loss Curve

have been implemented for process control. The most recent literature indicates an increased use of neural networks in process control. The reason for this interest in neural networks for process control is the desire for a highly reliable, predictable, accurate, fault resistant, and responsive method. The process control industry believes that neural networks have the potential for providing a method to meet these relevant issues and a basis for a new generation of flexible electronic equipment.

Furthermore, the literature contains many examples of attempts to use neural networks for process control. These examples include robotics (Xu, Scherrer, and Schweitzer, 1990; Lee and Kil, 1990), automated inspection systems (Hueter, 1993), process fault detection and diagnosis (Jokinen, 1991), multi-target tracking (Kuczewski, 1990), PVC pipe extruder (Smith and Dagli, 1991), and temperature control (Owens, 1992). In one particular example, Pugh (1991) compared the performance of neural networks to the traditional X-Bar control charts for various values of process shift. The data for Pugh's investigation was generated by a Monte Carlo simulation. Pugh trained the neural network for various shift conditions. One of these shift conditions was a parabolically distributed shift similar to the Taguchi Loss Curve (See Figure 4, Dehnad, 1989; and Tunner, 1990). The probability of the shift was in proportion to the loss imparted by the shift. As a result of the study Pugh formulated several conclusions. First, the training of a neural network with multiple shifts decreases error and training time. Second, a neural network trained with the shift contoured according to the Taguchi Loss Curve offered a slight performance improvement over the traditional X-Bar chart. Third, neural networks can be developed to approach the performance of standard X-Bar control charts in terms of Type I error. Finally, they can be created to out perform these charts in terms of Type

II error (Pugh, 1991).

Dissertation Summary

In the introduction, an overview of the dissertation is presented which incorporates the problem statement, the background for the dissertation, its significance, and the focus of the literature review. In the second chapter, an extensive review of the current literature is provided in order to probe the aspects of neural networks, time series analysis, and process control. The literature review deals with items such as definitions, techniques, applications, and a comparison of topics. The third chapter presents an analysis of the various data sets used in the research, an overview of a simulation computer program, and a review of Denton's algorithm along with the methodological design to be used to perform the study. The fourth chapter summarizes and evaluates the results from various methods based on nuclear material data sets. Chapter five summarizes and evaluates the results from various methods based on production process data sets. Also, in chapters four and five a comparison of the results from the neural network technique and other various techniques is discussed. The final chapter discusses the importance of the research results and possible future research endeavors.

CHAPTER 2: LITERATURE REVIEW

Chapter Overview

This section of the dissertation presents an in depth examination of current literature covering the topics of time series analysis, neural networks, production process control, and their relationships. Each topic is explored with respect to pertinent definitions, methods, and applications. Further, the topic of outliers is reviewed along with their impact on time series methods. A summary of the historical development of neural networks is included to provide an understanding of this topic. The areas of neural networks and time series are compared in terms of their ability to learn and help in decision making. Process control is examined to consider the use of time series and neural network methods including some representative applications. During this investigation, it was determined whether or not neural networks were better suited for process control than standard control chart or time series methods, for neural networks do possess certain advantages over these methods. The advantages are their ability to adapt to process changes and to learn as the process is operating. Finally, the nuclear industry is examined along with the relevant issues and applications. This industry was selected as a typical process control industry in order to perform analysis on data sets by neural networks and determine if neural networks could be used as an important safeguards method.

Time Series Methods

Definition

Time series analysis provides a method or model that can be used for analyzing historical data on particular variables in order to extrapolate them into the future (Bowerman and O'Connell, 1979). It is the basic strategy for many forcasting techniques that rest on the assumption that the data pattern identified will continue into the future. Bowerman and O'Connell (1979) in their book defined a time series as "a chronological sequence of observations on a particular variable." As a result, a time series does its forecasting without taking into account the relationships between the inputs and the outputs (Makridakis, Wheelwright, and McCree, 1983). Makridakis, Wheelwright, and McCree (1983) provided two good arguments for treating time series forecasting as a black box. "First, the system might not be understood, and even if it was understood; it may be extremely difficult to measure the relationships assumed to govern its behavior. Second, the main concern may be only to predict what will happen and not to know why it happens" (Suh, 1991). Thus, one difficulty with time series analysis is that the method or model which best fits the past data pattern might not necessarily provide the best forecast. The reason for this flawed forecast is that the data pattern did not continue into the future.

Time series approaches can be classified into two methods, 1) univariate (extrapolation or projection) where forecasts are based only on past observations of a given series and 2) multivariate (or causal) where forecasts are based partly on observations of other explanatory variables (Chatfield, 1988). The univariate method is useful when conditions are expected to remain the same. However, this method does not perform well when forecasting the impact of changes in management policies. On the other hand, the multivariate method is advantageous to the business world because it allows management to evaluate the impact of various alternative policies. Nevertheless, the method does have several flaws. First, it is difficult to use. Second, it requires historical data on all the variables including the variable to be forecasted in the model. Finally, the method depends on the forecaster's ability to accurately predict future values of the independent variables. Regardless of the disadvantages mentioned, the multivariate method is used quite often to generate predictions (Bowerman and O'Connell, 1979). The paradox is that the multivariate method is not as good as the univariate method in theory or in practice (Chatfield, 1988).

For many years, time series forecasting was based on the decomposition of trend, cyclic variation, seasonal variation, and irregular fluctuation. Decomposition of trend demonstrates the upward or downward movement of a variable over a period of time. Cyclic variation represents the recurring up and down movements around trend levels. Seasonal variation is a periodic pattern that will be completed within a calendar year and then repeats yearly. Irregular fluctuation consists of the erratic movements in a time series that follow no recognizable or regular pattern. These factors can occur individually, in any combination, or all together in a time series analysis. Consequently, there is no one time series method or model that is best suited for every situation. However, the selection of the appropriate time series technique should be based on the available data and the study objectives (Bowerman and O'Connell, 1979).

In 1991, Hwarng and Hubele (page 884) applied Back Propagation to classify the six above mentioned control patterns. What they found was that a neural network classifier with binary input and output worked well enough to serve as a supplement to traditional control charts. Smith (1993) extended the research in this area by considering a single model of a simultaneous X-bar and R chart along with location and variance shifts. He developed an experiment to train a neural network using a reasonable subgroup size of 10 to recognize a shift in a controlled process and to determine whether the mean or variance had changed. As a result, Smith was able to demonstrate the compatibility between neural networks and Shewhart X-bar and R control charts for large shifts in mean or variance. Further, the research indicated that neural networks out-performed the control charts for small shifts. Smith did note that one significant benefit of a neural network was its ability to model multiple control strategies simultaneously.

Techniques

In general, there are various time series methods and/or models from which to chose. Each method depending on the situation has its strengths and weaknesses in performing the forecast. Therefore, some prominent techniques are considered in this dissertation. These techniques include the Box-Jenkins method and its subcategories, Auto-Regressive Moving Average (ARMA), and Auto-Regressive Integrated Moving Average (ARIMA). A discussion of each of these techniques is provided. Also, the topic of outliers and the importance of their effects on time series analysis is given. Finally, two applications are presented to demonstrate the use of time series techniques for process control.

Box-Jenkins Technique

The Box-Jenkins technique gained popularity around 1970, both for its theoretical developments and its suitability for real-world applications (Andersen and Weiss,1984). These applications have included astronomy, economics, accounting, and many others. The Box-Jenkins technique is probably the most user dependent approach because the researcher is required to make decisions frequently throughout the process (Andersen and Weiss,1984).

The Box-Jenkins technique consists of four steps. The first step involves using the observations of a time series to identify a tentative model to be used in forecasting future values. During this identification step, a determination of whether the times series is stationary or nonstationary is made. A stationary time series fluctuates around a constant mean μ while a nonstationary time series has no constant mean. Next, either autocorrelation or partial autocorrelation is selected to measure the relationships between any two time series observations separated by a lag of K time units. An autocorrelation function of a stationary time series tends to die down with increasing lag K time units while for a nonstationary time series it tends to die down very slowly. The second step involves estimating the unknown parameters of the tentatively identified model. These estimates are often least squares estimates. The third step, the diagnostic check, involves testing the adequacy of the tentatively identified model, and if need be, suggesting improvements to the model. It has been found that an effective way to measure the overall adequacy of the identified model is to examine the quantity which determines the first K autocorrelations of the residuals. This quantity is known as the Box-Pierce Chi-Square statistic and is represented by the following formula (Bowerman and O'Connell, 1979).

$$Q = (n - d) \sum_{\ell-1}^{k} [(r_\ell) (\epsilon)]^2$$

where: n = number of observations in the original time series

d = degree of differencing that was used to transform the original time series into a stationary time series.

$(r_\ell)(\epsilon)$ = the sample autocorrelation of the residuals at lag ℓ

ϵ are the residuals

The final step is to use the developed model to forecast future values of the time series. Statistical confidence intervals can be formulated for the future values of the time series analysis because of the theoretical statistical basis behind the general Box-Jenkins technique (Bowerman and O'Connell, 1979). This time series technique has been automated into a forecasting expert system called AUTOBOX (Sharda and Patil, 1990).

Even though the Box-Jenkins method is more likely to produce more accurate forecasts than regression or exponential smoothing; it does have disadvantages. These disadvantages encompass the requirement for a large amount of data, a great investment in time, and dependency on other resources in order to build an accurate forecasting model. These disadvantages are particularly important if a desired model is to forecast a time series with seasonal variations (Bowerman and O'Connell, 1979).

ARMA (Auto-Regressive Moving Average)

The ARMA model is one of the most widely used forecasting techniques for stationary time series. This model contains both the Moving Average (MA) and the Auto-Regressive (AR) techniques. The Moving Average (MA) technique is simply the average of the latest N Data points which can be generalized for order q and denoted by MA(q). MA(q) is represented by:

$$Z_t = \mu + \epsilon_t - \theta_1\epsilon_{t-1} - \theta_2\epsilon_{t-2} \cdots\cdots - \theta_q\epsilon_{t-q} = \theta(B)\epsilon_t + \mu$$

$$Z_t - \mu = \theta(B)\epsilon_t$$

where: $\theta(B)$ is a function of the back-shift operator B

'backwards' operator B is defined by the relation $B.Z_t = Z_{t-1}$

ϵ_t are the residuals

μ is the mean

(Bowerman and O'Connell, 1979; Anderson and Weiss, 1984; and Newbold and Bos, 1994).

On the other hand, the Auto-Regressive (AR) technique can be generalized for order p and denoted by AR(p). AR(p) can be represented by:

$$Z_t = \mu + \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \cdots\cdots + \phi_p Z_{t-p} + \epsilon_t$$

$$Z_t - \phi_1 Z_{t-1} - \phi_2 Z_{t-2} - \ldots\ldots - \phi_p Z_{t-p} = \mu + \epsilon_t$$

$\phi(B)Z_t = \mu + \epsilon_t$

where: $\phi(B)$ is a function of the back-shift operator B

'backwards' operator B is defined by the relation $B.Z_t = Z_{t-1}$

μ is the mean

ϵ_t are the residuals

The term "Auto-Regressive" is used because Z_t, the current value of the time series, is "regressed," or expressed as a function of previous values of the same time series (Bowerman and O'Connell, 1979; Anderson and Weiss, 1984; and Newbold and Bos, 1994). By combining the moving average technique and the auto-regressive technique into the ARMA model, the result will require a relatively small total number of parameters and provide for an excellent representation of an actual stationary time series (Newbold and Bos, 1994).

According to McDonald (1989), the traditional approach to ARMA models has implicitly assumed that the underlying residuals are normally distributed. Consequently, minimizing the sum of the squares of the estimated residuals yields estimators which are asymptotically equivalent to the maximum likelihood estimators (McDonald, 1989).

The complete ARMA (p,q) process is represented by:

$$X_t - \phi_1 X_{t-1} - \phi_2 X_{t-2} - \cdots - \phi_p X_{t-p} = \epsilon_t + \theta_1 \epsilon_{t-1} + \cdots + \theta_q e_{t-q} + \mu$$

or

$$\Phi(B)X_t = \Theta(B)\epsilon_t + \mu$$

where: $\Phi(B)$ and $\Theta(B)$ are functions of the back-shift operator B

'backwards' operator B is defined by the relation $B.Z_t = Z_{t-1}$

ϵ_t are residuals

μ is the mean

A special case for a stationary time series would translate into the AR component when q is equal to zero. In order for the AR component to be considered stationary, the AR process must have Φ be less than one and fluctuate around constant mean (μ). The resulting equation corresponds to the AR(p) process:

$$\Phi(B)X_t = \epsilon_t + \mu$$

where the solutions to $\phi(Y) = 0$ all lie outside the unit circle

(Bowerman and O'Connell, 1979; Anderson and Weiss, 1984; and Newbold and Bos, 1994).

Further, the autocorrelation function (ACF) and the partial autocorrelation function (PACF) for this process can be expressed. The autocorrelation function for variations of this process takes the form:

$$\Phi(B)\ \rho(k) = 0 \quad \text{for } k > q$$

where: $\phi(B)$ is a function of the back-shift operator B

'backwards' operator B is defined by the relation $B.Z_t = Z_{t-1}$

$\rho(k)$ is the autocorrelations of a process and

denoted by $\rho(k) = \text{corr}(Z_t, Z_{t-k})$

and is more complex for $\rho(k)$, when $k \leq q$ (Anderson and Weiss, 1984; and Newbold and Bos,1994). The estimate of $\rho(k)$ is designated "the sample autocorrelation at lag K" and is denoted by the symbol r_k.

r_k takes the form:

$$r_k = \frac{\sum_{t=a}^{n-k} (Z_t - \bar{Z})(Z_{t+k} - \bar{Z})}{\sum_{t=a}^{n} (Z_t - \bar{Z})^2}$$

where $\bar{Z}$ is the average of all Z_t observations, i.e.

$$\bar{Z} = \frac{\sum_{t=a}^{n} Z_t}{n - a + 1}$$

Finally, the autocorrelation function of a stationary time series tends either to die down with increasing lag k or to cut off after a particular lag k=q (Bowerman and O'Connell, 1979; and Newbold and Bos, 1994). On the other hand, the partial autocorrelation function highlights the differences between any two time series observations separated by a lag of k time units. The partial autocorrelation can be denoted as follows:

$$\phi(B)\, \rho\,(kk) = 0 \quad \text{for } k<q$$

where: $\phi(B)$ is a function of the back-shift operator B

'backwards' operator B is defined by the relation $B.Z_t = Z_{t-1}$

ρ(kk) is the partial autocorrelation for a process

The estimate r(kk) is designated "the sample partial autocorrelation at lag K," and is denoted by the symbol r_{kk}.

r_{kk} takes the form:

$$r_{kk} = \begin{cases} r_1 & \text{if } k = 1 \\ \dfrac{r_k - \sum_{j=1}^{k-1} r_{k-1,j}\, r_{k-j}}{1 - \sum_{j=1}^{k-1} r_{k-1,j}\, r_j} & \text{if } k = 2,3,\cdots \end{cases}$$

where $r_{kj} = r_{k-1,j} - r_{kk} r_{k-1,k-j}$ for $j = 1,2,\cdots\cdots,k-1$

The partial autocorrelation function of a stationary time series tends to either die down with increasing lag k or to cut off after a particular lag k=q (Bowerman and O'Connell, 1979; and Newbold and Bos, 1994).

A comparison of the MA, AR, and ARMA time series models versus the ACF and PACF functions are presented in Table 2 (Anderson and Weiss, 1984).

	ACF	PACF
AR(p)	declines exponentially	cuts off after p lags
MA(q)	cuts off after q lags	declines exponentially
ARMA(p,q)	declines exponentially	declines exponentially

Table 2: Autocorrelation Functions vs Processes

The advantage of ARMA is that a model can be constructed with very few parameters. Its disadvantage is that too much emphasis is placed on outliers due to the technique of least squares used by ARMA both for identification and estimation (Booth, 1984 and McDonald, 1989). This disadvantage indicates that another method which de-emphasizes outliers would be desirable.

ARIMA (Auto-Regressive Integrated Moving Average)

The ARMA model has been used successfully to represent the behavior of a stationary time series in a wide variety of applications. However, for many applications this usage is neither realistic nor supported by empirical observations. The stationary model is only appropriate for a simple transformation of a business or economic time series. Thus, in various levels of business and the economy, the time series elements are not stationary and require a way to deal with periodic changes. The ARIMA model is often selected over other models in this case because of the addition of the integrated process (Newbold and Bos, 1994). There are other reasons for its selection which include being based on a solid foundation of mathematical theory, yielding optimal univariate forecasts based on the Box-Jenkins technique, and handling a wide variety of forecasting conditions (Pankratz, 1983).

The ARIMA model of order (p,d,q) can be denoted by ARIMA (p, d, q) where p is the number of auto-regressive terms, d is the degree of differencing, and q is the number of moving average terms in the model. The model further can be represented with a non-zero mean μ as follows:

$$\phi(B)\ [\ (1-B)^d X_t-\mu] = \theta(B)a_t$$

where: $\phi(B)$ and $\theta(B)$ are functions of the back-shift operator B

'backwards' operator B is defined by the relation $B.Z_t = Z_{t-1}$

a_t is white noise

$(1-B)^d X_t$ is a stationary series

In general this model has proven to be very successful in business and economic applications (Newbold and Bos, 1994).

Even though this type of time series model is often selected for use, it has its weaknesses (Pack and Downing, 1983). The model can be very complex and require a large amount of computing time. The model requires a minimum of points to produce an accurate forecast. If the time series is too short, the forecast may not be meaningful. Thus, the model is better suited for long range forecasts. Finally, the model's overall performance can be affected by outliers.

Outliers

Outliers have a massive impact on identification, estimation, and forecasting in time series analysis. Outliers are observations in a set of data which appear to be inconsistent with the remainder of the data set. Booth (1984) and Booth et. al. (1990) showed that knowing the outlier type aids in determining assignable causes in SPC. Further, Acar and Booth (1987) suggested the implementation of methods that detect outliers in real time SPC. First, the methods find the outliers (the points with potential assignable causes). Second, they identify the outlier type which helps to locate the production process problem. Therefore, by this investigation a problem with using standard control charts in SPC has been recognized.

When dealing with outliers in time series analysis, the first action taken is to identify their locations and types. Fox (1972) made the first contribution in defining types of outliers. His type I outlier is the isolated independent error which is independent of other observations. His type II outlier is the inherent type of anomalous observation which affects succeeding observations (Barnett and Lewis, 1979). Denby and Martin (1979) followed Fox's (1972) terminology by relating their innovational outlier to the type I outlier and additive outlier to the type II outlier. They defined the identification process as an underlying stochastic process that generates an outlying observation. To resolve the problems outliers cause, Denby and Martin developed the Generalized M-Estimation (GM) method to better estimate an AR(1) time series with outliers. Hillmer, Bell, and Tiao (1983), Tsay (1988), and Tiao (1990) expanded the discussion on outliers and their classification (Chen and Liu, 1993).

In current time series thought, there are four possible types of outliers in SPC. They are an innovational outlier (IO), an additive outlier (AO), a level shift outlier (LS), and a temporary change outlier (TC). A TC outlier produces an initial effect in a time series analysis at a particular point in time which then dies out gradually over time. An LS outlier produces an abrupt and permanent step change to the time series analysis. An AO outlier causes an immediate and a singular effect on the observed time series (Chen and Liu, 1993). A time series is affected by an AO outlier in two ways. The first way is through the carry-over effect. The second is through the bias in the estimates which could imply that incorrect parameter values were used in calculating the time series (Ledolter, 1989). Finally, the IO outlier is more complex because it affects both stationary and nonstationary time series. For a stationary time series, an IO outlier will produce a temporary effect which is connected to a continuing process problem. For a nonstationary time series, an IO outlier may produce various effects depending on the ARIMA (p, d, q) model. They include: "1) an initial effect at the time of the intervention and a level shift from the second period of the intervention, 2) an initial effect at the time of intervention that will gradually converge to a permanent level shift, 3) a seasonal shift if the time series follows a pure seasonal model, and 4) an annual trend change if the time series follows a multiplicative seasonal model" (Chen and Liu, 1993).

Chen and Liu (1993) have developed a procedure to identify and determine the effects of outliers. This procedure is important because it detects multiple types of outliers and gives reliable parameter estimates in their presence. This procedure has three stages. Stage one develops the initial parameter estimation and outlier detection. There are four steps in this stage. First, the maximum likelihood estimates of the model parameters are computed based on an adjusted time series to obtain residuals. Second, the determination of outliers and their types are made. Third, if no outliers are found, then the process moves to step four. If outliers are found, they are removed from the residuals and observations. The process then returns to step two. Fourth, if no outliers are found in the first pass, then the process is stopped. Otherwise the process proceeds to stage two when no outliers are detected after repeated passes (Chen and Liu, 1993).

Stage two is the joint estimation of outlier effects and model parameters.There are also four steps in this stage. First, the multiple regression model is used to jointly estimate the effects of possible outliers of various types which have been identified.

Second, the statistics (τ) of the estimators are computed and compared to a critical value (c). When the $\min_j |\tau_j|$ < or = c where j = 1,.....,m, the outlier at that particular time is deleted and step one is repeated. If the $\min_j |\tau_j|$ is greater than c, then the process moves to step three. Third, only the outlier effects are removed that are significantly based on the iterations of steps one and two. Fourth, the maximum likelihood estimates of the model parameters are computed based on the results of step three. If the relative change of the residual error from the previous estimator is greater than a predetermined tolerance, the process would go back to step one for additional iterations. Otherwise, the process goes to step one of stage three (Chen and Liu, 1993).

The final stage is the detection of outliers based on the final parameter estimates. There are two steps in this stage. First, the residuals are computed by filtering the original time series based on the parameter estimates. Second, these residuals are iterated through stages one and two with modifications that (a) the parameter estimates found in stage one are fixed to those obtained at stage two step four, and (b) steps three and four are omitted in stage two. The estimated outlier effects of the last iteration at stage two step one are the final estimates of the effects of the detected outliers. This entire procedure does simplify the computations involved in joint detection of multiple outliers because it employs detection of a single outlier in an inner-loop iteration (Chen and Liu, 1993).

Applications

There are many applications which use time series forecasting to predict future events based on past observations. These applications include weather forecasting, financial market forecasting, product sales forecasting, inventory level forecasting, and many more. An example demonstrating the usage of time series forecasting for process control is in power system loading. The Puget Sound Power and Light Company used a load series based on an hourly profile that was dominated by A.M. and P.M. load peaks. They applied ARMA time series method to forecast peaks for power demand in the future. As a result, the company was better able to manage its power distribution system (Connor and Atlas, 1991).

Another example was proposed by Alwan and Roberts (1988). They suggest that data correlation in the SPC process be modeled by an ARIMA time series. When the ARIMA model is complete, two charts, Common-Cause Control (CCC) chart and Special-Cause Control (SCC) chart, will be created and used to monitor the process. The CCC chart is a plot of fitted values or forecasted values which provides the current and predicted status of the process. With no control limits, the chart essentially signals when there are systematic variations in the process. Wardell et. al. (1992) extended the work of Alwan and Roberts by deriving limits for the CCC chart for the specific case of ARMA(1,1) model and restricted it to the steady state variance case. Their control limits for the CCC chart ($\mathbf{CL_F}$) are:

$$\mathbf{CL_F} = \boldsymbol{\mu} \pm \mathbf{L_F}\,\boldsymbol{\sigma}_F$$

where: $\boldsymbol{\sigma}_F$ is the steady state variance

$\mathbf{L_F}$ is the constant multiplier of the standard deviation of the forecast

$\boldsymbol{\mu}$ is the mean value

The results showed that by adding the limits to the CCC chart it improved the detection of shifts in the process mean and indicated a process change before other methods were able to do so.

The SCC chart is comparable to a conventional control chart. However, residuals are plotted instead of actual observations. The residuals will show a departure from statistical control when any process disturbances occur. Wardell et. al. (1992) derived control limits for the SCC chart based on the mean of the residuals being zero, which results in the center line of the SCC chart being zero. Thus, the control limits ($\mathbf{CL}_R$) for the SCC chart are:

$$\mathbf{CL}_R = \pm \mathbf{L}_R \boldsymbol{\sigma}_R$$

where: $\boldsymbol{\sigma}_R$ is the standard deviation of residuals

$\mathbf{L}_R$ is a multiplier and usually assumed to be 3

By establishing these limits, Wardell et. al. were able to determine that the SCC chart had the ability to forecast the quality of a future process. Wardell et. al. (1994) developed the run length distributions for the SCC chart to perform further research against standard SPC charts. Even though the results were promising, Wardell et. al. still felt that the traditional control chart should be considered for certain cases.

However, the Alwan and Roberts approach does have some appealing aspects. The first aspect is that the process is correlated. This correlation allows for forecasts of future process quality. The second aspect is that since the SSC chart is based on the assumption that residuals are random, the assumptions for the traditional SPC are met. Therefore, the traditional tools of SPC can be utilized. The third aspect is that the SSC chart can be used to detect any assignable cause relating to a process disturbance. The fourth aspect is that the level of sophistication required by the user is minimal because of the availability of user friendly software packages to fit the time series models and the straight forwardness of the methodology used to obtain the charts. The final aspect is that this method can be applied to any type of time series method.

In a recent article, Prasad et al. (1995) examined a new robust time series method for Statistical Process Control (SPC). This new method called Joint Estimation was developed by Chen and Liu (1993). It has the capability of detecting four types of outliers: Innovational Outliers (IO), Additive Outliers (AO), Level Shift (LS), and Temporary Change (TC). These outlier types will assist in determining assignable causes for process variations. To demonstrate this method's capabilities for statistical process control, Prasad et al. (1993) compared the results of the Joint Estimation method against the Shewhart control chart method, and the M-Type Iterative method based on the data set from monitoring the acidity in dye liquor. Grant and Leavenworth (1980, p. 93) provided the data set which presented the average pH values across five Hussong kettles over time. The evaluation was done in terms of the accuracy of modeling the underlying process and the ability to detect and identify various assignable causes. The results of the study showed that the Joint Estimation method was able to identify all the assigned abnormalities while the other two methods weren't. This technique is discussed further in chapters four and five.

Neural Networks

Definition

Today, neural network technology has become one of the most widely investigated topics in information systems. The fascination of this field is in its ability to learn from being exposed to information or data, and then utilizing the data and information to make decisions similar to those of a human brain. At this point, several basic definitions of neural network technology will provide insight into this new field. Klimasauskas (1991) says, "Neural Networks (Artificial Neural Systems) are an Information Processing Technology inspired by studies of the brain and nervous system." Kohonen (1987) writes, "The artificial neural networks are massively paralleled interconnected networks of simple (usually adaptive) elements and their hierarchical organizations which are intended to interact with the objects of the real world in the same way as the biological nervous systems do." Hecht-Nielsen (1989) adds, "A neural network is a computing system made up of a number of simple, highly interconnected processing elements, which processes information by its dynamic state response to external inputs." Maren (Maren, Harston, and Pap, 1990) continues, " Neural networks are computational systems, either hardware or software, which mimic the computational abilities of biological systems by using large numbers of simple, interconnected artificial neurons."

When these definitions are analyzed, they all consider neural networks to be like a biological nervous system stimulated by information to make decisions. This unique capability of processing information provides specific advantages. The first advantage is adaptive learning. Neural networks have the ability to learn how to do tasks based on the data given by training or initial experience. Neural networks learn how to perform certain tasks by being trained with illustrative examples. This learning capability has a long term potential for allowing neural networks to be used in environments where data are in flux. The second advantage is self-organization. Neural networks can create their own organization or representation of the information they receive during learning and operation. Neural networks are able to self-organize information received because of their adaptive learning capabilities. Because of self-organization leading to generalization, neural networks can respond when novel data situations are presented. The third advantage is fault tolerance via redundant information coding. Fault tolerance has two distinct aspects. Neural networks have the ability to recognize patterns which are noisy, distorted, or even incomplete. Also, neural networks can continue to perform even with a partial destruction of the network. This fault tolerant feature exists because neural networks have redundant information coding. The fourth advantage is real-time operation. Neural network computations may be carried out in parallel. To achieve this capability, special hardware devices have been and continue to be designed and manufactured. This concept of neural networks is a very competitive alternative for low-level, real-time pattern recognition and classification. The final advantage of neural networks is their ease of insertion into existing technology. Because of the near-term availability of neural networks on specialized chips and their potential to offer improved performance on discrete tasks, they will facilitate modular upgrades to existing systems. Finally, they offer incremental design improvements for systems under development (Maren, Harston, and Pap, 1990).

Along with these advantages, there are other major reasons for using neural networks. A neural network can be developed with less programming time and less checkout time than it takes to formulate a theoretical model. A neural network does not have to know the exact operating conditions of a process. If a process unit (such as a variable exceeding the tolerance level) changes, the neural network can learn and adjust to the change without human interference. Thus, the neural network requires a fraction of the computational time needed for a theoretical model to solve the problem (Blaesi and Jensen, 1992).

However, neural networks do have disadvantages. Sartori (1992) and his colleagues felt that neural networks had a tendency to confuse complexity in relationships with randomness. Further, they felt that neural networks had difficulty distinguishing between what is a signal and what is noise. On the other hand, Blaesi and Jensen (1992) felt that neural networks needed too much data for a good model. Also, for a good model to be developed, they concluded that a wide range of data points is also needed. Therefore, the researcher must be aware of these disadvantages when constructing his neural network.

History

Scientists and researchers have had an endless interest in the human brain which has inspired work in the area of neural networks. These networks are a composite of many simple processing elements that operate in parallel in a similar manner to the processing found in the human nervous system. The historical progression of neural networks can be seen in Figure 5.

The study of neural networks began with a paper written by McCulloch and Pitts in 1943 titled, "A Logical Calculus of Ideas Imminent in Nervous Activity." Based on this paper two prominent fields received their beginnings, artificial intelligence and expert systems. Nevertheless, Rosenblatt recognized the limitations of McCulloch-Pitts' models and referred to them as "logical contrivances" (Kohonen,1987; Maren, Harston, and Pap, 1990; and NeuralWare Inc., 1991).

In 1957, Rosenblatt developed the first major project in the field of neural networks with his development of the Perceptron network. The first Perceptron network was powerful and capable of limited learning. The learning in this model was accomplished by comparing the product of the input and the weight vectors to the threshold value which sets the criterion for the output value. If the product value exceeds the threshold value, then the output is assigned a value of one. If not, the value of the output is zero. The Perceptron further possessed a great deal of plasticity, a capability of making limited generalizations, and the ability to properly categorize patterns. The Perceptron had limitations because of it being a developmental device. The Perceptron method could not be described in terms of the Automata Theory. The Automata Theory is a design of mathematical models describing methods of information transformation in digital systems. In fact, the Automata Theory is a part of the Computability Theory which deals with problems in computer science, activity of the nervous systems (neural networks), and so on (Mikolajczak, 1991). Also, the Perceptron method had no capability to represent the basic Exclusive OR (XOR) function. Nevertheless, Rosenblatt's work with the Perceptron stimulated further research in the area of neural networks (Maren, Harston, and Pap, 1990; NeuralWare, Inc., 1991).

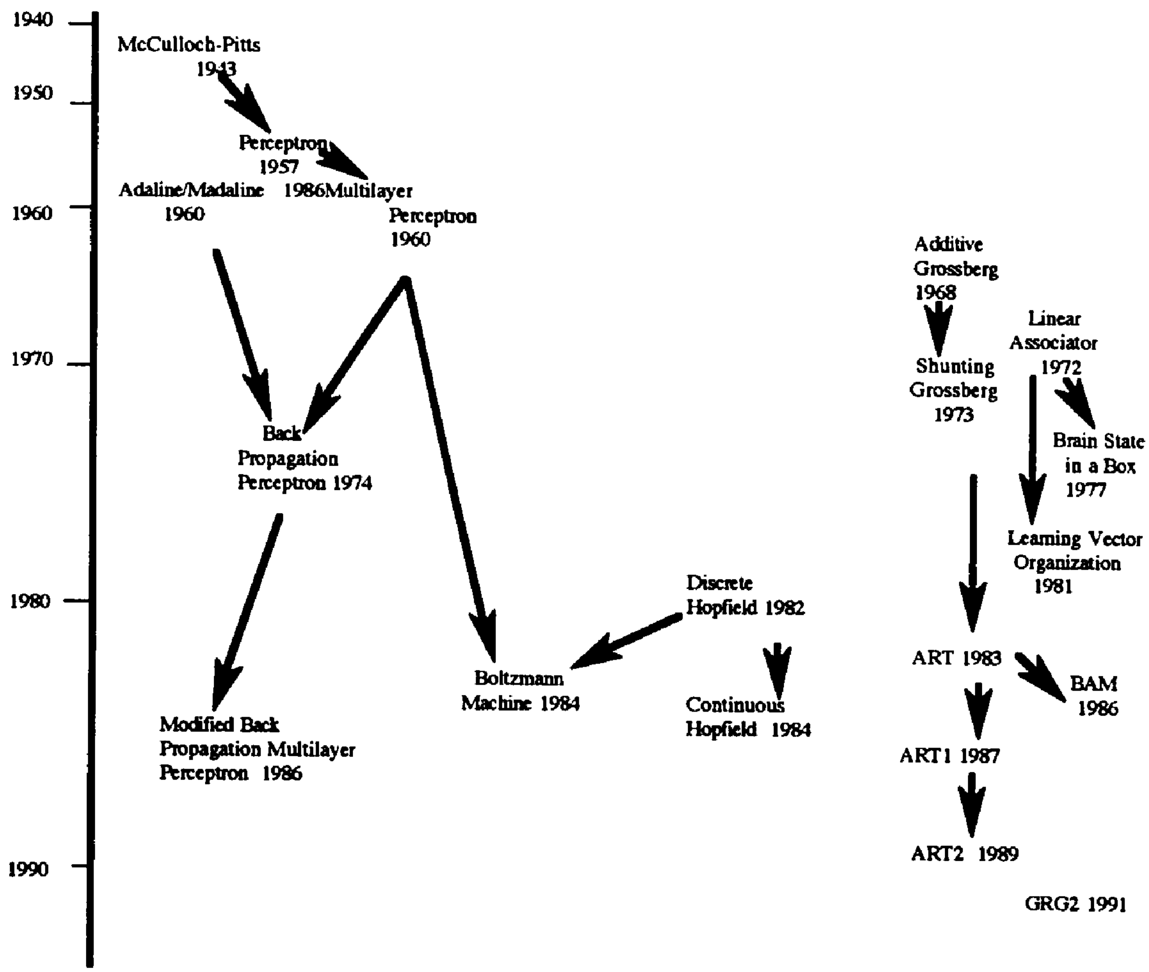


Figure 5: The History of Neural Networks

In 1959, Widrow developed an adaptive linear model called Adaline (Adaptive Linear Neuron). The model used the Widrow-Hoff learning control mechanism to sample the inputs, the output, and the desired output. "Widrow-Hoff learning performs a gradient descent algorithm in weight space and is guaranteed to converge to the unique set of weights which give the minimum mean square error between the desired and the actual outputs for an example set." (NeuralWare, Inc., 1991) From the results, an error signal was computed and the trainable weights adjusted to reduce the error. This process converged to a stable condition in approximately five times as many learning trials as there were weights. Widrow used Adaline to develop adaptive filters that eliminated echoes on phone lines. This application was the first time neural networks were applied commercially. Other applications of the Adaline model were in character recognition, speech recognition, weather prediction, and adaptive control (NeuralWare, Inc., 1991).

However, the last limitation of the Perceptron led to Minsky and Papert's paper Perceptrons: An Introduction to Computational Geometry (1969) which outlined the limitations of Perceptron based neural networks. In this paper, they concluded that the Perceptron cannot handle inputs that are visually nonlinear. Therefore, the Perceptron neural network could not perform an Exclusive OR (XOR) operation. Minsky and Papert pointed out that the extension to a multi-layer perceptron might be able to solve the Exclusive OR (XOR) problem. However, there was no mechanism in place at that time for resolving how to adjust the weights when the output was incorrect. The results of their work were important because at the time, there existed a large number of functions that were not linearly separable. Therefore, the Perceptron algorithm could not represent those functions. Thus, their released findings had a damaging effect on neural network research, and it declined severely. These two highly influential people in the field of Artificial Intelligence effectively interrupted the research on neural networks for almost twenty years. Nevertheless, research did continue in a limited manner (Maren, Harston, and Pap, 1990; NeuralWare, Inc., 1991).

One of the researchers who did continue the work in neural networks was Anderson (1972). He worked on developing a linear model called the Linear Associator. This model is based on the Hebbian Principle that the connections between neuron-like elements are strengthened every time they are activated. The Hebbian Principle produces a desired output for input stimuli for an example set provided they are orthogonal to each other (NeuralWare, Inc., 1991). Another of these researchers was Kohonen (1972). He did fundamental work in adaptive learning which formulated rules for weights to be modified in a manner only dependent on the previous weight value, and the post and presynaptic values. Another of his contributions is the Principle of Competitive Learning, in which processing elements compete to respond to an input stimulus and the winner adapts itself to respond more strongly to that stimulus. In general, Competitive Learning consists of the lateral interactions of neighboring groups of neurons on the brain. Kohonen's concept originated from a study of self-organizing maps which studied how information was received at sensory organs and mapped topologically onto one and two dimensional areas of the brain. The Competitive Learning method is also unsupervised. This method is discussed later in this chapter. A third researcher, during this period, was Grossberg (1973). His field of concentration was in the use of neurological data to build neural computing models. One of his contributions is the class of networks called Adaptive Resonance Theory (ART) which reflects his commitment to physiological modeling because of certain properties within these networks. The first of these properties is "one shot" learning as opposed to a lengthy training time found in other methods. The second property is the fact that learning and recall only depend on local information. The third property is the fact that noisy input data does not corrupt the network. The final property is the fact that the network is self-organizing (NeuralWare, Inc., 1991).

Another of Grossberg's contributions was his proof of the Cohen-Grossberg Theorem which concerns stability during recall of a general class of networks. This stability theorem is formulated in terms of the Lyapunov function which guarantees the convergences of a network to an equilibrium in response to any external stimulus. This concept is also supported by the Hopfield theorem (NeuralWare, Inc., 1991). Finally, Hopfield (1982) was able to free the field of neural networks from the stigma placed on it by Minsky and Papert in 1969. He did so in 1982 by presenting a paper entitled "Hopfield Model" to the National Academy of Sciences. In this paper, Hopfield presented a computing system consisting of interconnected processing elements that seek an energy minimum. Due to Hopfield's standing in the scientific community, as well as his personal charisma and enthusiasm, he was able to relegitimatize and rekindle the interest in neural computing.

Currently there are many researchers working in the area of neural networks. One of the most significant advancements in neural computing has been the development of multi-layer systems which can learn and categorize a case of complex class categories. Rumelhart and his Parallel Distributed Processing group (1982, 1987) used this concept in the development of the Back Propagation model. The Back

Propagation model is a multilayer network consisting of one input layer, one output layer, and at least one hidden layer. The layers are fully connected with arrows indicating the flow of information. This model will be discussed later in this chapter. Other researchers deserving of mention include Hinton and Sejnowski (1984, 1986) for their development of the Boltzmann model which modified the Hopfield network. Kosko (1986, 1987) developed the Bi-Directional Associative Memory (BAM) network which is a variation of the Grossberg network. Finally, Hung and Denton (1991) using a General-Purpose Nonlinear Optimizer (GRG2) developed an interface for implementing neural network training. Hung and Subramanian (1991) indicated the superiority of this method over Back Propagation.

Techniques

In general, neural networks with relatively complex architectures tend to be more powerful in learning, but take more time to train. A simple neural network is shown in Figure 6 (Rumelhart et. al., 1987; NeuralWare, Inc., 1991). The architecture of this network consists of a number of layers among which information is processed as indicated by the arrows. The input layer receives the information and adjusts it with the use of a linear or nonlinear transfer function. The output layer sends out the processed results. In between these two layers, there is at least one hidden layer which allows the network to learn complex data. The actual number of hidden layers will depend on the application complexity. Each layer is connected with its succeeding layer by a set of weights (Rumelhart et. al., 1987; NeuralWare, Inc., 1991). Although there are several mathematical models of neural networks, only three will be examined in this dissertation. They are Adaptive Resonance Theory (ART), Back Propagation, and General Purpose Nonlinear Optimizer (GRG2).

ART (Adaptive Resonance Theory)

ART is an unsupervised learning model and was developed in order to analyze how brain networks can autonomously learn about a changing world in a rapid but stable fashion and in real-time. These models originated from the analysis of simpler types of models in use today. Amari and Takeuchi (1978), Bienenstock, Cooper and Munro (1982), Kohonen (1984), and Carpenter and Grossberg (1990) further enhanced the growth of these models (Zornetzer, Davis, and Lau, 1990).

Grossberg (1976) formally introduced ART to demonstrate how to embed a Competitive Learning model into a self-regulating control structure. This structure has the ability to learn and recognize an arbitrary sequence of input patterns in a continuously stable and efficient manner. Because of this ability, ART differs significantly from other alternative learning schemes in several ways (See Table 3).

The importance of these differences is that ART has the capability of learning without supervision. It has a match phase mode which enables quick and stable learning to take place while buffering the system's memory against external noise. In summary, the ART model updates its learned codes with new information either from new approximate matches or from selected nodes initiating a learning of new recognition categories and defends its learned codes from being deleted by an influx of new input events (Zornetzer, Davis, and Lau, 1990; and NeuralWare, Inc., 1991).

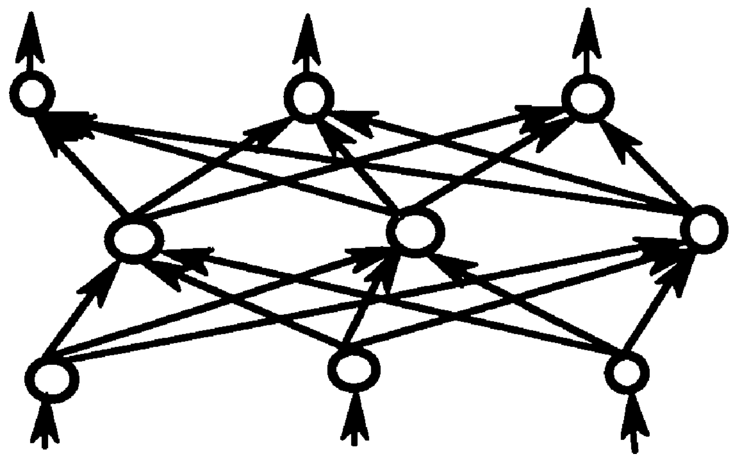


Figure 6: Simple Neural Network

Grossberg and Carpenter expanded ART into the network models ART1 and ART2. ART1 was designed to respond to an arbitrary sequence of binary input patterns. On the other hand, ART2 was an extension of ART1 in that it classified with binary input patterns. However, it could also classify with an arbitrary sequence of analog input patterns. Since ART1 was formulated first, was the simpler of the two, and demonstrated many of the important aspects of ART2, this dissertation will confine its discussion to ART1.

The ART1 model consists of two interconnected layers of nodes, labeled F1 and F2 (See Figure 7). As a stream of input patterns arrive at the F1 layer of nodes, these patterns are normalized and stored in short term memory (STM). Then, the patterns are multiplied by adaptive weights. These adaptive weights are the long term memory (LTM) traces between F1 nodes and F2 nodes. The F2 layer of nodes is designed with the capability of selecting the node which receives the largest input (winner-take-all). Each node of the F2 layer sums its input patterns and the F2 node receiving the largest total input is chosen the winner and allowed to fire. When this process is complete, the matching process begins between "the bottom up input pattern" and "the top down learned expectation," to identify a mismatch. When a mismatch is recognized at a F2 node, the node will be deactivated for the remaining of the matching cycle and another F2 node is selected. By doing this process, an examplar is created and put into LTM for the model to remember what input pattern it just "saw" (Zornetzer, Davis, and Lau, 1990; NeuralWare, 1991; and Kang, 1991).

Carpenter, Grossberg, and Reynolds (1991) have introduced a new ART architecture called Predictive ART or ARTMAP. ARTMAP is a supervised real-time learning and classification method for nonstationary data. A classifier system operates in the context of an environment that sends messages to the system and provides it with reinforcement based on the behavior it displays. ARTMAP learns to classify arbitrarily many ordered vectors into categories based on predictive success. A Predictive Art system includes two ART modules connected by an Inter-ART associative memory. It is similar to the Back Propagation network in that both are supervised learning systems. However, the two systems differ in many respects (See Table 4).

With supervised learning, one input vector is associated with another input vector on each training trial. Then during a test trial, a new input vector is presented that has never been seen before, and in turn, the network predicts an output vector. The system's performance is evaluated by comparing the predicted output vector with the correct answer.

Back Propagation

Back Propagation is the most widely used learning paradigm by which a multi-layer network is developed for pattern recognition or classification and utilizes the supervision technique. This algorithm allows for weights to be adjusted by back propagating errors from the output layer to the input layer (Kang, 1991; and NeuralWare, Inc., 1991). It uses a gradient descent method to search for optimal parameter settings for the network (Rumelhart, McClelland, and the PDP Research Group, 1986). The gradient descent method that Back Propagation employs is called the Delta Rule which always points along the direction of the steepest rise of a curve. The Delta Rule minimizes the mean square error with respect to the connection weights in the network and does it in the most efficient manner (Caudill, 1990). The Delta Rule will be discussed and analyzed later in this chapter. As for a typical Back Propagation network (See Figure 8), it is comprised of an input layer, and output layer, and at least one hidden layer. This architecture makes the network a multilayer feedforward network similar to the Perceptron network developed by Rosenblatt. The difference is that there are no hidden layers found in Rosenblatt's model. These additional hidden layer(s) allow the model to overcome the problem of linear separability that existed with the Perceptron network when performing more complicated mapping functions.

Back Propagation's wide acceptance has been credited to its simplicity of use, flexibility in dealing with application data, and its ability to learn complicated multi-dimensional mapping. This ability to learn complicated multi-dimensional mapping, in the words of Werbos, allows Back Propagation to go "Beyond Regression" (Hecht-

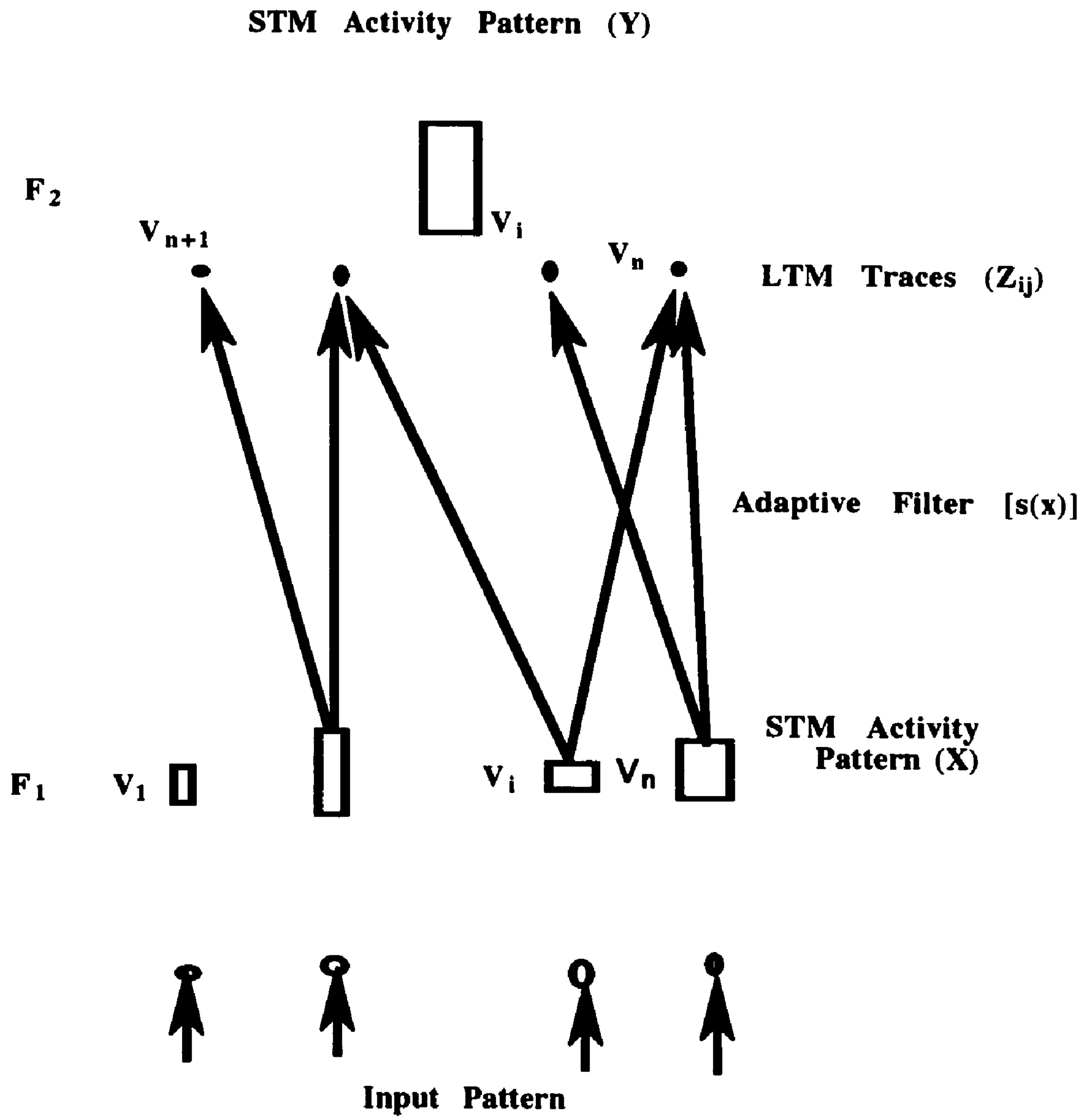


Figure 7: ART1 Model

ART	Alternative Learning Properties
Real-Time (on Line) Learning	Lab-Time (off-line) Learning
Nonstationary World	Stationary World
Self-Organizing (unsupervised)	Expert Supplies Correct Answer (supervised)
Fast Learning	Slow Learning or Oscillation Catastrophe
Learn in Match Phase	Learn in Mismatch Phase
Fast Adaptive Search for Best Match	Search Tree
All Properties Scale to Arbitrarily Large System Capacities	Key Properties Deteriorate as System Capacity is Increased
Rapid Direct Access to Codes of Familiar Events	Recognition Time Increases with Code Complexity

Table 3: Comparision of ART vs. Alternative Learning Schemes

Properties	Predictive ART	Back Propagation
Supervised	yes	yes
Self-organizing	yes	no
Real-time	yes	no
Self-stablizing	yes	no
Learning	fast or slow match	slow mismatach

Table 4: ARTMAP vs Back Propagation

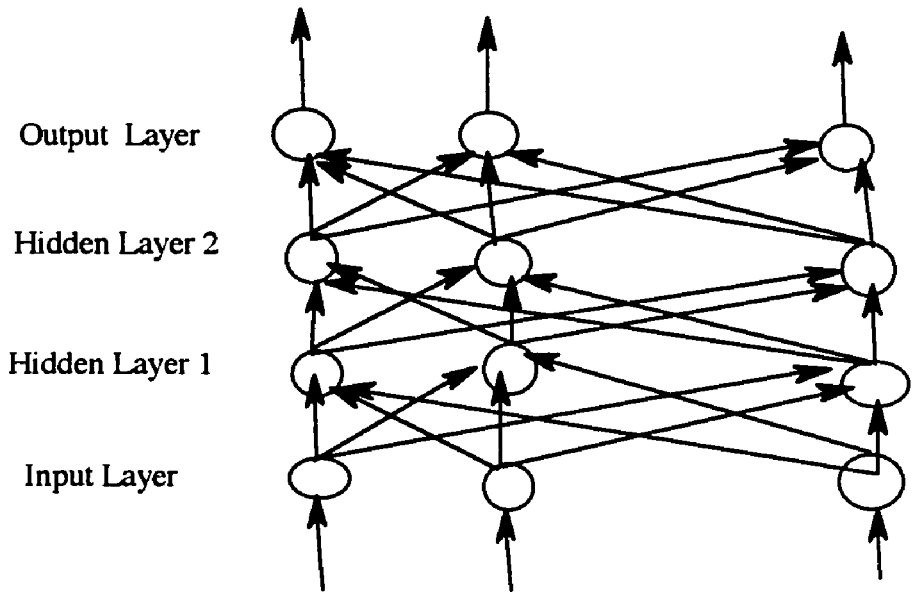


Figure 8: Typical Back Propagation Network

Nielsen, 1989). Even though Back Propagation does suffer from several limitations which will be discussed later in this chapter, it is still a widely used network to solve problems in pattern recognition, classification, and signal processing applications.

Rumelhart and his Parallel Distributed Processing (PDP) group are credited with developing Back Propagation into a usable technique and advocating it to a wide audience. Back Propagation was originally introduced by Bryson and Ho in 1969 (Hecht-Nielsen, 1989). Then it was independently rediscovered by Werbos in 1974, by Parker in the mid 1980's, and by Rumelhart, Williams, and the PDP group in 1985. The PDP group became aware of Parker's efforts after reviewing Parker's 1985 report (Hecht-Nielsen, 1989). Werbos' effort was not acknowledged until mid 1987 when Parker discovered it. Bryson and Ho's work was uncovered in 1986 by Le Cun (Hecht-Nielsen, 1989). Consequently, still older works might emerge concerning

Back Propagation (Hecht-Nielsen, 1989).

The principal algorithm used by Back Propagation for adjusting the weights of a network is the generalized Delta Rule. "The Delta Rule is a type of learning algorithm where weighted values are modified to reduce the difference between the desired output and the actual output of a processing element" (NeuralWare, Inc., 1991). This reduction of difference is accomplished by moving the weight vector from its current position on the bowl and then following the negative gradient of the bowl to a new position closer to the minimum error (See Figure 9). The minimum error corresponds to the ideal weight vector for a particular input pattern. Consequently, the closer the minimum error is to zero, the better the weight factor is (Caudill, 1990). Thus, the Delta Rule, which follows the gradient of a curve, is called a gradient-descent learning algorithm (Caudill, 1988). This learning algorithm always takes the most efficient route from the current position of the weight vector to the ideal position, based on the current input pattern. Therefore, the Delta Rule not only minimizes the mean squared error but does so in the most efficient possible manner (Caudill, 1990).

Back Propagation involves two phases. During the first phase, the input patterns are presented and propagated forward through the network to compute the output patterns. These computed output patterns are compared to the desired output patterns resulting in an error value (E). The second phase involves a backward pass through the network (similar to the forward pass) during which the error signal is passed to each unit in the network and appropriate weighted value changes are made (Rumelhart, McClelland, and the PDP Research Group, 1986). These weighted value changes are calculated using:

$$W_{new} - W_{old} = \frac{\beta E X}{|X|^2}$$

Where: X - input vectors
W - weight vectors
β - learning constant
E - error value

An important factor to note about this equation is that only weighted values are modified (Caudill, 1990). Another important factor is that the learning constant β , which is a measure of the speed of convergence of the weight vector to the minimum error position, should be in the range of 0 to 1 (Caudill, 1990).

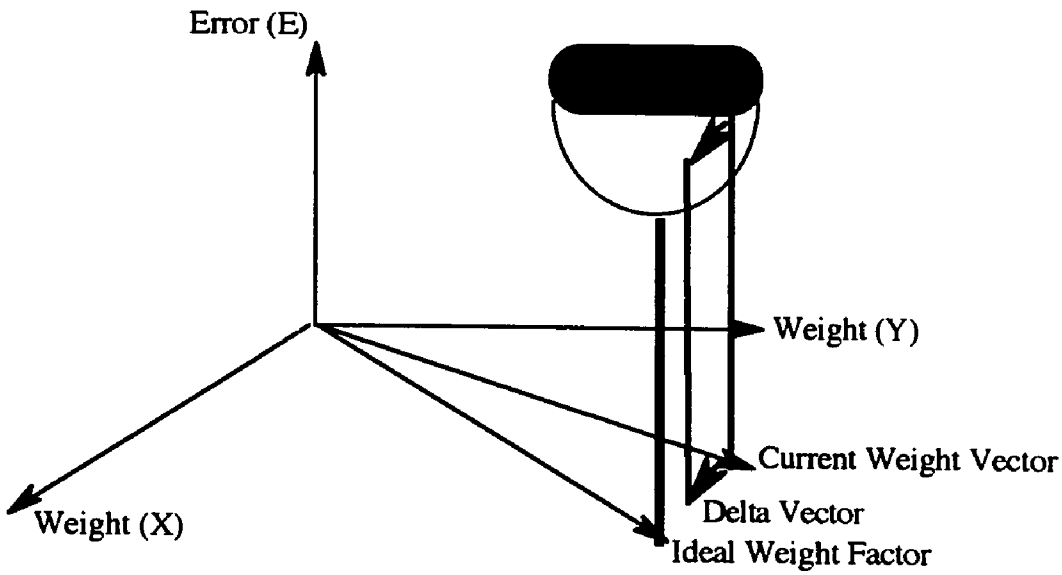


Figure 9: Delta Rule

Back Propagation commonly uses two non-linear transfer functions. They are the sigmoid and the hyperbolic tangent. The sigmoid transfer function scale has a range of 0 to +1 and is defined as $1/(1+\epsilon^{-x})$. On the other hand, the hyperbolic transfer function scale has a range of -1 to +1 and is defined as $(\epsilon^{x}-\epsilon^{-x})/(\epsilon^{x}+\epsilon^{-x})$. The Back Propagation method utilizes these functions to transfer its inputs to its output and their derivatives to scale the error. As a result, these functions allow multi-layer networks to deal with non-linear cases (NeuralWare, Inc., 1991).

As stated previously, Back Propagation, despite its popularity, does have limitations. It suffers from long learning time, poor local minima (a point on the error surface where the network will oscillate and never find a solution), and network paralysis. Then too, the error surface of multilayer Perceptrons (Perceptrons which have intermediate layers called hidden layers between the input and output layers), unlike that of single Perceptrons, have flat areas, shallow troughs, high points, and low points in high dimensional space. Thus, the Back Propagation network has the potential to get trapped in any one of the various error surfaces, because it may contain a local minima. However, researchers who have been able, through lengthy experiments to discover local minima, nevertheless experience no serious problems (Cheung, Lustig, Kornhauser, 1990).

The most critical problem when implementing Back Propagation is its slow training process time. Researchers have been attempting to improve the training process. However, they are experiencing little success. Some of the suggested ideas for improvement include the use of Newton's Method (White, 1990), Fourier Analysis (Jones, 1992), Power Series (Chen and Manry, 1990), Parallel Forward Propagation (Abe, 1990), and Higher Order Derivatives (Parker, 1987).

GRG2 (General-Purpose Nonlinear Optimizer)

Even though Back Propagation is currently the most widely used feedforward neural network algorithm, researchers continue to seek improvements because of the limitations that were previously mentioned. One of these improvements is the use of GRG2 for the training system base. GRG2 is a widely distributed nonlinear optimization software package that offers better scalability for large problems, faster results, and a better quality of solutions. V. Subramanian and M. Hung (1991) in their research verified GRG2's superiority to Back Propagation.

GRG2 uses a generalized reduced gradient algorithm (which refers to the gradient vector for a subset of variables, which are to change without violating the constraints in solving a problem). It is a Fortran based program that uses supplied subroutines to evaluate the objective function for a particular gradient subroutine. If a gradient subroutine is not specified, then the program will use either the method of forward differences or central differences in order to estimate the gradient. Denton and Hung (1991) developed a method of interfacing GRG2 for the training of neural networks.

Applications

Due to the current interest level in neural network research, a multitude of successful applications have emerged. This development has been brought about by more powerful multilayer networks, improved learning techniques, and enhanced computer capabilities. From these elements, neural network applications have provided this literature review with a wide range of subjects to discuss. These subjects will include classification, process control, robotics, signal processing, and business. In this research, the application of neural networks would fall under classification and process control. The remainder of this section will present many of the current applications. A summarization of this material is presented in Table 5.

Classification

Classification is introduced first because of the variety of ways it can be accomplished. Early classification methods that were based on mathematical or statistical methods were not sufficiently tolerant to random noise. Thus, these methods tend to exhibit poor performance. The recent renaissance of neural network research has brought about many new classifiers. They even allow for learning the implied relationships between the data while making no assumptions about the distribution of the data. The study by Xu, Krzyzak, and Suen (1991) used neural networks to investigate the combining of multiple classifiers in conjunction with an associative switch. As the authors stated, "When an unlabeled pattern is input to each individual classifier, it also goes to the neural network for associatively calling out a code which controls the switch to decide whether the result of each classifier could pass through as a final result." The concept was applied to a problem of combining multiple classifiers for recognizing totally unconstrained hand written numerals. The results indicated that the associative switch could produce improvements in pattern recognition. Another article by Wong and Vieth (1990) integrated neural networks into statistical classification for the purpose of overcoming the problems of noise effects of the data sets' performances on early classification systems. They developed their artificial neural network using a medical data set. The data set contained patient records having twelve attributes with two to four unique values where each was diagnosed to have one of four diseases. The authors used seventy-five percent of the randomly selected records for the training set and twenty-five percent for the testing set. Then the authors ran both Back Propagation and their network for ten trials and compared the average of the results. It was shown that the new network was far superior in diagnosing the correct disease group 95% of the time versus 80% of the time and the incorrect group 4% versus 20%.

In a paper by Subramanian, Hung, and Hu (1993), the researchers investigated the capabilities of neural networks for classification. A comparison was done of the performance difference between neural networks and two discriminant analysis models (linear and quadratic). The study was based on four cases: combinations of two and three groups, and two and three variables. Certain factors were considered in the study such as sample size, proportion of group memberships, and degrees of overlap. The analysis of these factors was accomplished by using the analysis of variance (ANOVA) program. The results showed that even under the best conditions for discriminant analysis models, neural networks are quite competitive. Further, the authors concluded that neural networks are more impressive because they are less sensitive to changes in sample size, number of groups, number of variables, proportions of group memberships, and degrees of overlap among groups.

In another paper, Patuwo, Hu and Hung (1993) investigated using neural networks for two-group classification. The main priority of the research was to study two important issues in building neural networks. The issues were network architecture and size of training samples. The researchers developed two experiments for examining two-group classification. In the first experiment, the objective was to determine the appropriate architecture of neural network classifiers.

APPLICATION	EXAMPLES	REFERENCES
Classification	General	- L.Xu, A. Krzyzak, and C. Y. Suen (1991) - A. K. C. Wong and J. O. Vieth (1990) - V. Subramanian, M. Hung, and M. Hu (1993) - E. Patuwo, M. Hu, and M. Hung (1993)
	Pattern Recognition	- Z. Lo and D. B. Bavarian (1991) - N. A. Thacker and J. E. W. Mayhew (1990) - I. N. M. Papadakis (1991) - L. Gupta and A. M. Upadhye (1991)
	Speech Recognition	- M. E. Hoff (1988) - Sejnowski and Rosenberg (1987)
Robotics	Kinematic	- S. Lee and R. M. Kil (1990)
Signal Processing	Ultrasonic Signals	- K. Shahani, L. Udpa, and S. S. Udpa (1991)
Business	Financial	- H. White (1988) - P. K. Coats and L. F. Fant (1992)
	Bankruptcy	- K. Tam (1991)
	Stock Selection	- F. S. Wong, P. Z. Wang, T. H. Goh, and - B. K. Quek (1992)
	International Conflict	- P. A. Schrodt (1991)

Table 5: Neural Network Application Summary

Two-group, two-variable classification problems were used. The networks had two input nodes and one output node. The number of hidden nodes was varied (3,5, and 7). As a result of the first experiment, it was found that the appropriate architectural choice for a neural network or the choice of the sample size depended upon whether to maximize the classification rate of the training samples or the generalizability of the neural network. Therefore, depending on the researcher's objective, the researcher could chose to maximize the classification rate in a training sample by using large networks or generalize the network results by using small networks with large training samples. In the second experiment, neural network models are compared against classical models such as discriminant analysis (linear and quadratic) and nonparametric methods (K-nearest neighbor and linear programming). Based on the information from the first experiment, the DISCRIM procedure from Statistical Analysis System (SAS) Institute was used to solve the three types of two-group classification problems. The results from experiment two indicated that neural networks compare favorably to these other methods in terms of classification rates in the training samples but not in the test samples.

A major subdivision of classification is pattern recognition. Pattern recognition is the mapping of a large set of data into a smaller set of pre-specified classes using some measurement criterion. The process of pattern classification is composed of two stages, a feature extraction stage and a decision-making stage. In the first stage, the large dimensional pattern is transformed into a set of measurements. In the second stage, some parameters are adapted to external signals or vectors to improve the overall performance of the classification. Then too, as in most pattern recognition applications, the underlying statistical distributions of data or their functional forms are known (Lo and Bavarian, 1991).

With the advent of neural networks, research in pattern classification has increased. Lo and Bavarian (1991) developed a Neural Piecewise Linear Classifier based on Kohonen's Learning Vector Quantization and Self-Organizing feature map. A data set extracted from ship images was used to evaluate their classifier. The results demonstrated that the classifier had a strong self-organizing property and performed better than the two other classifying methods used. Meanwhile, Thacker and Mayhew (1990), recognizing that real world patterns are complicated by noise and imperfect data, developed a method called CLAM (Contextual Layered Associative Memory) in an attempt to extend conventional one-layered models to permit multi-layering. Their network was comprised of a fixed number of layers (which are used to classify accumulated classifications from previous layers), a flexible number of nodes (which allows for the classification network to grow from a few seed nodes during training), and flexible connectivity (which permits connections to be maintained according to the demands of the training set). According to the authors, this network was not a cure-all to the problems of context-sensitive classification, but a step towards making neural networks available for classification.

Then too, Papadakis (1991) developed a new pattern classifier based on the Hamming network. The Hamming network is composed of three layers. The first layer is the input layer with N processing elements. The second layer is the category layer with M processing elements. The third layer is the output layer. This neural network is a minimum error classifier for binary vectors which is defined using Hamming distance. The distance is measure between two binary vectors which is defined by the number of bits in the input that do not match corresponding bits in the categories (NeuralWare, Inc., 1991). The new classifier called PACNET (Pattern Classification Neural Network) can be used either as an optimum minimum error classifier or a pattern classifier with predetermined noise tolerances. As a result, the scheme avoids all the disadvantages of the Hopfield model when used as a pattern classifier. Gupta and Upadhye (1991) developed a neural network to address the ineffectiveness of classifiers to compensate for time variations resulting from random partial occlusion. They were able to compensate for the above by adding a non-linear alignment stage to the neural network output. The resulting classifier is capable of tolerating high degrees of random noise and occlusion in shapes. Their experimental results verified the effectiveness of the neural network classifier in classifying shapes with time variations.

Another area under classification is speech recognition. This is a difficult problem because of the amount of information the system must retain in order to establish the rules of a specific language. Also, the problem is aggravated by the dynamic time varying feature of speech. Neural networks have been proposed to be used in this area in order to take advantage of their capabilities in adaptivity and flexibility. Hoff (1988) developed a word recognizer for processing speech that allows voice-controlled data entry. This method can handle a two hundred word vocabulary, respond in less than half of a second, and provide better than ninety-nine percent accuracy while being trained to recognize a specific speaker. Another application was developed by Sejnowski and Rosenberg (1987) called NETTALK. NETTALK is based on a multi-layered network and trained using Back Propagation. Its main function is to learn how to pronounce an English text. Using this method, they were able to achieve a performance rate of ninety-seven percent for one thousand commonly used words and an eighty percent rate for twenty thousand words from a dictionary.

Robotics

In the area of robotics, Lee and Kil (1990) developed a new method of accomplishing robot kinematic control based on a Bi-Directional Mapping Neural Network (BMNN). BMNN is composed of a multi-layer feedforward network with hidden units having sinusoidal activation functions and a feedback network forming a recurrent loop around the feedforward network. The feedback network iteratively generates joint angle updates based on a Lyapunov function to modify the current joint angle in such a way that the output of the forward network converges to the desired Cartesian position and orientation. BMNN has certain advantages over conventional approaches. The advantages are accurate computations of robot forward and inverse kinematic solutions for simple training, the ability of handling "one-too-many" inverse mappings that are required for redundant arm kinematic solutions, and the automatic generation of arm trajectories (Lee and Kil, 1990).

Signal Processing

Neural networks are playing an increasingly larger role in the classification of ultrasonic signals. The reason for this increased role lies in the fact that the method does not require any explicit a priori statistical information. Shahani, Udpa and Udpa (1991) incorporated neural networks in the classification of Ultrasonic NDE

(Nondestructive Evaluation) Signals. The analysis of the signals represents the interaction between defects and inhomogeneities in a specimen and an ultrasonic wave launched into material under inspection. The authors selected two neural networks (the multi-layer Perceptron and the Time Delay Neural Network) to compare for classifying signals. The results showed that the Time Delay Neural Network offered fewer misclassifications than the multi-layer Perceptron did for the same architecture. Therefore, the classification results obtained were better than using traditional pattern classifiers.

Business

Neural networks provide opportunities to handle a wide range of tasks in the business world. In the financial area, White (1988) used neural network modeling to search and decode nonlinear regularities in assessing price movements. He successfully applied his technique on the IBM daily common stock returns. Coats and Fant (1992) used neural network modeling to forecast financial distress. They compared the neural network approach against multiple discriminant analysis (MDA) for ninety-four firms. The results showed that the neural network method correctly forecast ninety-one percent of the distressed as distressed, and ninety-six percent of the healthy firms as healthy. MDA only identified seventy-two percent and eighty-nine percent respectively. Tam (1991) compared neural network models to statistical and machine learning techniques in order to predict bank bankruptcy. The evaluation was based on four dimensions (robustness, predictive accuracy, adaptability, and explanatory capability). The statistical and machine learning techniques included

MDA, Factor Logistics, k Nearest Neighbor, and the ID3 algorithm. MDA is a linear model and was applied in this study using the Fisher type of maximizing the ratio of between-groups and within-groups variance. Factor logistic is a factor analysis and was applied in this study using the West model. The West model identified the important factors influencing banking and provided factor scores for each observation used. k Nearest Neighbor (kNN) is a nonparametric method and was applied in this study to assign an observation to a group to which the majority of its k Nearest Neighbors belonged. The ID3 algorithm is a rule-based classification tree and it employs a non-backtracking splitting procedure that recursively partitions a set of examples into disjointed subsets which intends to maximize the entropy of the split subsets. The empirical results of Tam's research showed that neural networks offer a better predictive accuracy than the techniques just mentioned.

Wong, Wang, Goh, and Quek (1992) outlined an Intelligent Stock Selection (ISS) system which extends the neural network method to handle fuzzy, probabilistic, and Boolean information. Fuzzy information is data which has ambiguities, inconsistencies, incompleteness, and noise. Boolean information is data which primarily indicates whether associated statements are true or false. Probabilistic information is data which is based on a probability in the range 0 to 1. The ISS system was designed around an integrated network architecture, based on a building block called a neural gate. A neural gate is similar to the processing cell used in neural networks but designed for a Fuzzy Net to process fuzzy rules, noisy information, probabilistic and Boolean data. The ISS model was tested on a database containing about eight hundred stocks over a three year period. The initial results from the study were satisfactory.

Schrodt (1991) implemented a neural network to predict international conflict outcomes and compared it against discriminant analysis (a linear model), multinomial logit analysis (a nonlinear maximum likelihood model), and the ID3 algorithm (rule-based classification tree). In split-sample tests using Butterworth's International Conflict data set (Schrodt, 1991), the neural network out-performed both discriminant analysis and ID3 in terms of accuracy. It was roughly comparable in accuracy to the multinomial logit analysis. The neural network was less successful than discriminant analysis and multinomial logit analysis at predicting nonmodal values of the dependent variable.

As a result of neural networks' current popularity, many manufacturing applications also have been initiated. Therefore, it is important to understand the learning process of neural networks and its relationship to time series methods. Many of the manufacturing applications that will be discussed later in this chapter are able to be accomplished due to the fact that neural networks have the ability to learn.

Learning Process of Neural Networks

The potential of neural networks for time-series forecasting has been explored previously in the literature. Werbos (1974) first indicated that neural networks trained by the Back Propagation method could outperform the Box-Jenkins method, when forecasting over a long period of time. Lapedes and Farber (1987) demonstrated that neural networks using the Back Propagation technique could be used to predict future sample points in a chaotic time-series better than conventional time series methods. Sharda and Patil (1990) indicated that the performance of the neural network approach was comparable to that of the Box-Jenkins method for one period ahead forecasts. Tang, Almeida, and Fishwick (1990) also stated that neural networks could be a promising alternative approach to time series forecasting. A logical reason for this optimism is the ability of neural networks to learn from actual process data and be able to be more sensitive to turning point fluctuations.

Caudill (1988) states that learning is fundamental to the neural network approach. Learning implies that the processing element somehow changes its input/output behavior in response to changes in the environment. Neural networks learn by changing the weights on the inputs. Thus, the process for precisely determining how to change the weights in response to a given input and output pair is learning.

The learning technique in neural networks can be separated into two categories, supervised or unsupervised. Supervised learning means the network has some omniscient input present during training that always knows what the correct answer should be. A selected output, which represents the desired response, is developed for each input and the actual output is compared with the corresponding selected output. The resolved difference is fed back into the network to modify the weights and reduce the difference. On the other hand, unsupervised learning means that the network has no knowledge of the correct result and therefore cannot know what the correct response should be. Thus, this method does not need a selected output which allows it to be similar to the biological learning system.

As was discussed previously, Caudill (1988) said that learning was an attribute of neural networks which somehow changes its input/output behavior in response to the environment. The primary components of neural networks are processing elements and the interconnections. The processing elements, the neural network equivalent of biological neurons, are generally simple devices that receive a number of input signals and, based on those inputs, either generate a single output signal or do not. The output signal is sent to other processing elements as inputs via the interconnections.

Then too, a neural network is neither sequential nor even necessarily deterministic, but has many processing elements working in parallel. It has no separate memory array for storing data and its processing units are not highly complex. There are no precoded instructions for the neural network to execute. Instead, neural networks achieve their overall state when they reach some equilibrium condition after adjusting the weights of the interconnections and taking the weighted sum of all its inputs. As Grossberg (1991) states, "Neural network models are not merely a more convenient way to realize well-known computational ideas. They embody new computational theories."

At present, neural networks are being investigated to be used for time series forecasting and process control. In doing this investigation, certain problems have been encountered. The first problem is that sequential training pairs are statistically dependent. This means that if the training pairs are presented to the network in sequential order, then any trends in the data can result in slow convergence as the network tries to model the current trend rather than the underlying model of the time series data. To avoid this situation, it has been suggest that the order in which the training pairs are presented to the network should be random rather than sequential. Second, the assumptions that the network parameters are held constant after training implies that the training set fully describes the relationship to be learned by the neural network and that the relationship is static. Neither of these assumptions can be guaranteed for applications dealing with time series data. Finally, due to the limited size of the training sequence, it is possible that the dynamics of the time series are not fully described by the training sequence (Khotanzad and Fowler, 1991).

Therefore, based on the study by Khotanzad and Fowler (1991), they suggest that neural networks should be trained with randomly ordered data rather than sequential data. Also, the learning rate parameter should be increased as the RMS prediction error decreases. Finally, the network should continue to update its parameters as data outside of the training set is made available.

Process Control

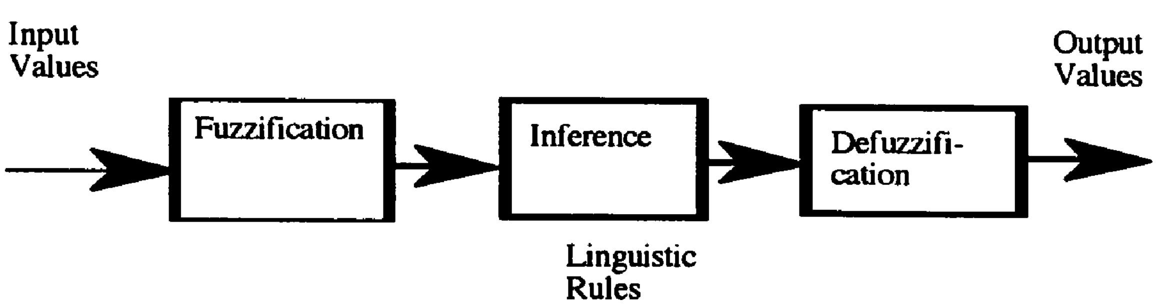
Definition

Process Control and Process Capabilities depend on how a researcher views the subject. Process Control is the total control of the process or operation from the beginning to the end of manufacturing a product. Process Capabilities, on the other hand, refer to the proportion of quality variation that may be expected from a production process operating at particular conditions which include line speeds, temperature and pressure set points, and others. Thus, the importance of controlling the process is to increase productivity while improving quality (DataMyte Corporation, 1989).

The methods that are available for controlling a process include Statistical Process Control (SPC), Time Series, Neural Networks, Fuzzy Logic, and others. What follows is a basic explanation of some of these methods with special emphasis on time series and neural networks.

Traditional SPC was introduced by Shewhart (1931) to examine and control an industrial process. The principal goal of SPC is to identify extraneous disturbances to a process, typically a manufacturing process, in order to maintain a predictable output (Wardell, Moskowitz, and Plante, 1991). The underlying assumption of SPC is that observations are independent and normally distributed about a mean. This variation to non-random disturbances is categorized into two types, expected and acceptable and unexpected and unacceptable. Consequently, SPC uses control charts to present a visual display of the process characteristics to the operator. Thus, the operator will decide from reviewing the charts the status of the process and whether a process correction is necessary. Even though, different SPC charts are available, the result from analyzing the SPC charts is only an indication of the process status. This indication doesn't provide the operator with what is causing the problem, where the problem is, how to correct the problem, or whether it is a major or a minor problem.

Fuzzy Logic (FL) was introduced by Zadeh (1965) to model subjective, indeterminate concepts with intermediate values. The following block diagram is a representation of a FL controller.



There are two advantages of using FL. First, FL reacts much faster to over and under shoots, or a process disturbance than a standard PID control method. Second, it handles problems where insufficient mathematical models or sensory data exist. The disadvantages of FL are that it does not lend itself to autotuning and cannot condition signals from sensors for use in the inference block of the FL diagram. As to its applications, FL has been used in automatic transmissions, anti-braking systems, temperature controllers, vision systems, motion controllers, and PLC's (Programmable Logic Controllers). Recently, FL has been incorporated with Neural Networks (Back Propagation) to design more accurate control systems (Berg, 1993).

Techniques

Times Series

Box and Jenkins (1976) presented a times series approach to standard Statistical Process Control (SPC) methods. They indicated that these methods could continue to be used as a screening mechanism for the process while the time series analysis of the systematic non-random variations could help determine the underlying common causes that influence the process operation. This time series analysis would show what systematic variables need adjusting to provide continuous improvement of the process.

Alwan and Roberts (1988) proposed a possible solution to the problem of data correlation in SPC environments. Their proposal was to model the process with a time series technique (ARIMA). After the modeling was completed, the process was monitored by the Common-Cause Control (CCC) chart and the Special-Cause Control (SCC) chart. The CCC chart is composed of fitted values which provide a view of the current level of the process and its change over time with no control limits. The SCC chart is composed of residuals which screen the operation of the process for outliers. These time series techniques contain many benefits. They include taking advantage of the fact that a process is correlated to make forecasts of future quality, utilizing the assumption that residuals are random, implementing traditional SPC tools, using the SCC chart to detect any assignable cause, and being able to deal with any type of time series. Even with these advantages, the research to investigate the performance of time series techniques versus traditional SPC procedures has been limited.

Prasad (1990) and Booth (1984) discussed new approaches of applying the time series analysis to process control for early detection of material losses in critical inventory control. In particular, Prasad presented a dynamic form of time series analysis to identify both the innovative and additive outliers which led to determining whether or not a process disruption will continue or is a one time event. This method was a major advancement in SPC and made it easier to recognize process problems.

Recently, Prasad, Booth, Hu, and Deligonul (1995) compared the performances of two time series methods (Joint Estimation and M-type iterative) and Shewhart control charts for Statistical Process Control (SPC). The Joint Estimation method was recently developed by Chen and Liu (1993). This method is a three-stage process which identifies all four types of outliers in SPC applications. In applying these methods, a data set from Grant and Leavenworth (1980, p. 93) was used. The three methods were used to locate problems in the pH dye liquor data set. The results demonstrated that the Joint Estimation method was able to identify all the assignable causes in the data set. The authors did recommend the implementation of this method for monitoring production processes, but did recognize that a high degree of operator training is required. Therefore, the time series methods which have been presented will aid those in process control to decide which SPC methodology is best suited to their manufacturing needs.

Neural Networks

Smith and Dagli (1991) examined controlling an industrial process with supervised, feedforward neural networks. Process monitoring and control consists of numerous inputs and outputs which are normally nondeterminative and do not adhere to known probability distributions. By training neural networks with the Back Propagation method, Smith and Dagli determined that significant benefits were achieved over traditional statistical methods. These advantages encompassed the ability to discern complex relationships and trends rather than assuming distributions, the ability to integrate in real-time large amounts of continuous data, and the ability to handle noisy or incomplete data.

Psaltis, Sideris, and Yamamura (1987) used neural networks to control a given process (plant) (See Figure 10). They applied several learning techniques to train the neural controller. These techniques were indirect learning, general learning, and specialized learning. From their simulation, they concluded that individually generalized or specialized training can provide the desired results in many cases. Further, some prior general learning can help specialized learning in yielding better results.

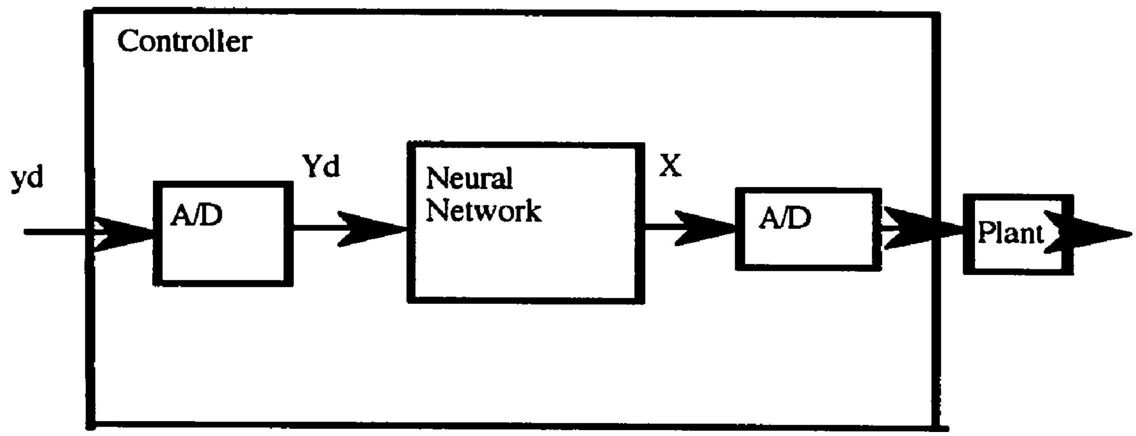


Figure 10: The Operation Mode Block Diagram

Hwarng and Hubele (1991) recognized the difficulties of implementing SPC for automatic identification of special disturbances in a process. Therefore, they proposed a control chart pattern recognition methodology based on the Back Propagation technique. This method provided the ability for real-time Statistical Process Control and the evaluation of observations to determine whether a pattern (such as a trend, cycle, stratification, etc.) exists. It was found that neural networks had major advantages over traditional approaches for control chart pattern recognition. First, the flexibility of training allows neural networks to fit into today's dynamic manufacturing environment. Second, the ability for high-speed computations, as a result of their highly interconnected processing elements, enables neural networks to be implemented in real-time applications.

Pugh (1991) performed a comparison of a neural network to several traditional statistical control chart techniques. He discovered that the Back Propagation neural network could be developed to approximately equal the performance of standard X-Bar control charts for Type I error. However, it exceeded the performance of these charts for Type II errors. Finally, he noticed that a network trained with the shift contoured according to the Taguchi cost curve could offer a slight improvement over a traditional X-Bar chart.

Owens (1992) in his article indicated that a single-hidden-layer Back Propagation network, which allows arbitrarily complex nonlinear interactions, often gives more accurate predictions of input-output relationships than traditional statistical methods. The network can be trained to simulate the dynamics of the time-dependent relationships between the input variables and the output responses of a complex process. Thus, the network made it easier for non-linear mapping between the process input variables and the corresponding product output responses. Finally, it required no prior knowledge about the functional relationship between input variables and output responses.

Manufacturing Applications -- General

In manufacturing today, there are many examples of control applications utilizing neural networks. R. Escobedo developed a neural network based engineering retrieval system for Boeing. Boeing engineers submit their designs to the retrieval system and within a few minutes, they are told whether or not the part they designed exists somewhere in Boeing's vast inventory (Babb, 1993). Georgia Tech electrical engineers created a high-speed analog integrated circuit (IC) amplifier that "learns" to control its own performance through the use of neural network technology. They are trying to automate the process by an on-chip device that will generate inputs, measure the outputs, and make the adjustments. Once in service, the IC amplifier could periodically retrain itself depending on its time in service and other operational changes (Control, 1993).

With new manufacturing techniques being required to meet competition, Chu and Tsai (1994) have developed an approach for cellular manufacturing (CM). The first stage of designing an effective CM is called a cell formation (CF). In the CF process either similarly designed parts or parts requiring similar processing are grouped together with their machinery to form machine cells. The authors approached the CF problem by applying the Fuzzy mathematical programming method. The benefit of this method is that it allows for and takes into account the ambiguity, incomplete information, and uncertainty inherent in most real world situations. In their study, they investigated three different models as well as fuzzy programming. From this study, they concluded that the fuzzy programming approach not only provided decision makers with better and more flexible ways of representing CF problems, but also led to an improved overall performance.

Present day machine vision technology applied to SPC is generally limited to detecting the existence of a bad characteristic in a product. With the implementation of neural networks, machine vision has been able to develop into intelligent product inspection systems. Hueter (1993) described applying an automated vision inspection system which incorporated neural network technology for an apple quality sorting system. The system sorted the apples according to color, uniformity for four color gradations, and for four surface defects while taking into account that the stem and glare are not defects. The results from the system realized significant labor savings and gains in quality consistency. Beck, McDonald, and Brzakovic (1989) presented a self-training visual inspection system using a neural network classifier. The system consisted of a control unit, a signal processing unit, and a connectionist classifier (neural network) (See Figure 11).

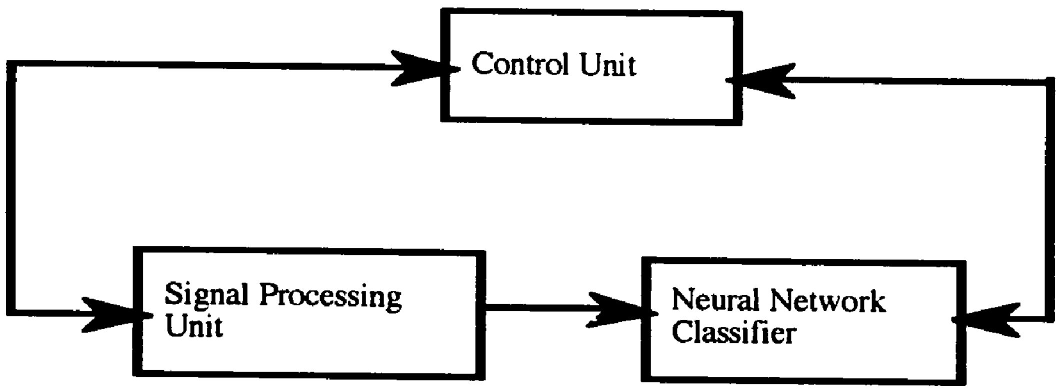


FIGURE 11: Components of the Inspection System

The control unit's task was to generate flaws of known characteristics and superimpose them on digitized images. The signal processing unit had the task of generating a pattern vector containing flawed features that adequately distinguished the class of each flaw. The classifier had to learn to detect and classify flaws in digitized images of surfaces with known characteristics. The learning was accomplished with the use of Back Propagation. To evaluate the performance of the system, the authors developed three diagnostic indicators to explain the behavior of the inspection system within the context of two different inspection tasks. The three diagnostic indicators included the ability of the classifier to learn to classify a set of pattern vectors, to indicate whether or not the classifier is functioning within a reliable operating zone, and the first layer weights. As for the two different inspection tasks, they included the detection and classification of flaws on a printed surface of know characteristics and the detection and classification of flaws on surfaces characterized by a uniform intensity distribution. As a result of testing the inspection system, the reliability was found weak depending on the inspected surface.

Today, researchers are investigating the applying of neural network technology for flight management (FMS) and vehicle management (VMS) systems. Reibling and Olinger (1990) developed a mathematical model for the basis of their neural network architecture for planning the best direction of movement. The model was based on the electric field theory. The difficulty in computing the best path for aircraft trajectory or vehicle location is that real-time solutions are required which current technology cannot accomplish. As a result of their efforts, they determined that the paths obtained by their model are reasonable, but that considerations should be given to determining if another partial differential equation may provide better results.

FMS is essentially a global optimization problem constrained by the dynamics of the systems being trained. Kuczewski (1987) in his article took the Grossberg-Mingolla Boundary Contour System (BCS) and modified it to perform temporal associations and interpolations between target reports properly correlated in time and space. As a result, a two dimensional Interpolative Probability Field (IPF) capable of performing the multi-target tracking task was developed. This method was applied to aircraft for improving landing capabilities. The advantages of IPF are in its parallel implementation, its ease of adaptation, and its achievement of the track-independent calculation time. VMS utilizes neural networks to interpret visual images. A group of graduate students and professors at Carnegie Mellon University created a VMS which they labeled ALVINN (Autonomous Land Vehicle In a Neural Network) (Business Week, 1992). They were able to train ALVINN to recognize various images such as trees, parked cars, and pavement. Even with these capabilities, it was difficult for ALVINN to stay on the road for more than a mile or two.

Finally, robotics has become very important for increasing productivity and quality in manufacturing processes. Xu, Scherrer, and Schweitzer (1990) established a goal of not only placing a robotic gripper of three fingers on an arbitrary object but also being able to grasp the object. They accomplished their goal with the use of the Hopfield neural network and a simple mechanical gripper. The advantage of this approach was its learning capacity. The results obtained showed that this scheme behaved in a promising fashion.

Manufacturing Applications --Process

An important topic in Process Control is sensor technology and its relationship to intelligent diagnostics. Sensor data is easily obtained from many manufacturing operations and its immediate but thorough analysis is crucial. Neural networks provide an ideal platform for this function. First, they have the ability to work simultaneously with large amounts of multivariate data. Second, they have the aptitude for melding data and detecting important patterns or trends. Finally, they have the ability to be integrated into larger control systems (Smith and Dagli, 1991). Smith and Dagli presented an example of applying SPC to a plastic pipe extruding operation. They compared regression against the Back Propagation neural network method for four studies using ten test observations each. The results confirmed their assumption that neural networks can offer alternatives to traditional analysis without suffering in quality. Denton (1994) did a similar study comparing regression and neural networks for four different cases. The cases included a standard regression model, a regression model with an outlier, a regression model with multicollinearity, and a regression model with a specification error. In each case, the neural network performed the same regression function. The results revealed that the neural network out performed regression at times but yielded none of the rich statistical information that regression can furnish.

Owens (1992) utilized the Back Propagation neural network to automatically generate an optimal control strategy for a text book example of a simple heat exchanger. The exact analytical solution to the model for the heat exchanger is presented below.

$$Y = Y_o \text{ Exp}(-Ft) + (t + u^2/F)\{1 - \text{Exp}(-Ft)\}$$

where:
Y_o = Output temperature at t = 0
t = Time
u = Heater voltage
F = Flow rate
Y = Output temperature

The network was structured with eight inputs, four hidden units, and one output. The control strategy for varying the voltage implemented by the neural network was performed in steps. First it increased the voltage to its maximum value. Second, it held the voltage at this value until the output temperature reached its goal. Third, it allowed the voltage to oscillate about the steady-state value for several time intervals.

Finally, it reached the output temperature and achieved steady-state. This strategy produced a smooth and rapid controlled output temperature curve.

Jokinen (1991) compared neural models with respect to noise using a process fault detection and diagnosis problem. The data that was used was obtained from a simulated chemical process. The process was composed of a reactor and a distillation column. The author performed a comparison between Back Propagation and the Dynamically Capacity Allocating (DCA) network by adding a noise pattern to each. DCA models an open learning system that estimates the functional form between variables along with the model parameter. This network is particularly suitable for a continuous learning system because it is capable of spatially selective updating. As a result of the author's simulation, it was found that the continuously learning DCA network model was superior in performance to Back Propagation.

Manufacturing Applications --Nuclear Industry

Background

The nuclear industry originated in the early 1950's when prototype units demonstrated that nuclear power could be generated (See Figure 12). From these modest beginnings, the government in 1954, through the Atomic Energy Act, provided access to nuclear technology to industry for commercialization. Even though the industry was given this access, there was still concern for financial liability in case of an accident. Thus, the Price Anderson Act was passed in 1957 to protect owners and operators of nuclear power plants from such liability.

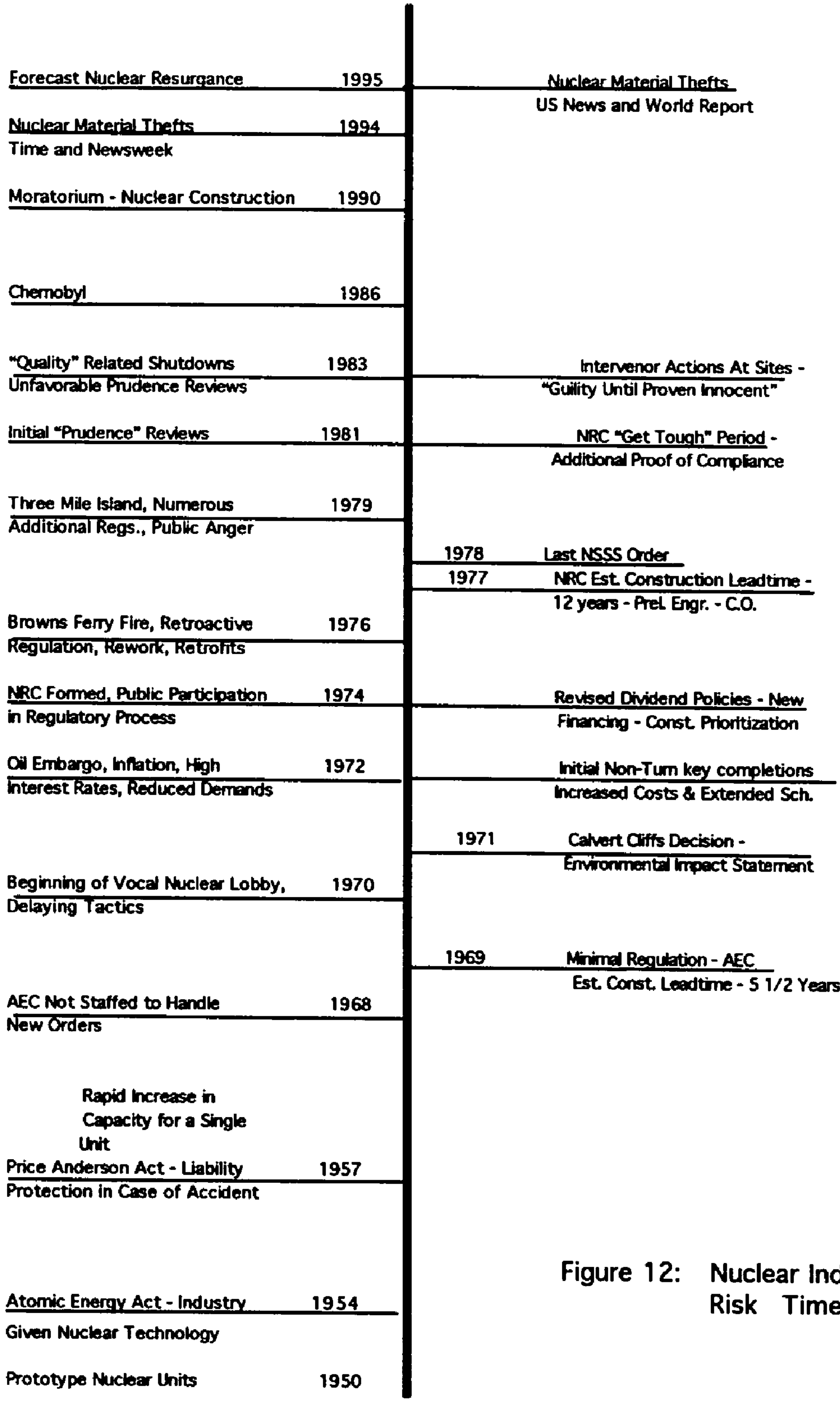


Figure 12: Nuclear Industry Owner Risk Timeline.

During the mid to late 1960's, the nuclear industry realized its maximum growth. It was able to meet this growth by performing turn key contracts and utilizing utility management techniques learned from the fossil-fuel industry. Due to this rapid growth, the nuclear industry lost sight of its responsibility to monitor plant constructions and to control the campaign for larger nameplate capacity nuclear units, which in turn, would expose them to increasing financial risk.

In the late 1960's and through the 1970's, citizens' groups started becoming very active in forcing the federal and state governments to start addressing the risks of constructing new nuclear power plants. In 1971, the Calvert Cliffs decision made the National Environmental Protection Agency (NEPA) a part of the Atomic Energy Commission (AEC) and required an Environmental Impact Statement before granting a construction permit to any company. Because of this act and continued public outcry, the nuclear industry faced a decline of approvals for permits and delays in construction.

In 1973 - 1974, the nuclear industry encountered financial difficulties because of the Arab oil embargo which raised fuel prices and the rate of inflation. As a result of the embargo it made the cost of borrowing money impossible. The impact of this event was that construction of new plants were either delayed or canceled. In turn, the nuclear industry changed its approach towards contracting, staffing, and monitoring, because it became clear that the ultimate accountability for quality, safety, environment, and compliance with regulation was theirs. Then, in 1974, the Nuclear Regulatory Commission (NRC) was formed to be directly involved in the management process of nuclear power plants. Further, it encouraged public participation in the nuclear regulatory process.

Although this regulatory process appeared to be functioning properly, when the Three Mile Island (TMI) accident occurred in 1979, it demonstrated a lack of thoroughness and questioned the effectiveness of the NRC. In response to this situation, the NRC issued a wave of new regulations covering all aspects of the nuclear industry. Because of these new regulations, the nuclear industry adopted a new approach of dedicated project management with increasing owner involvement and control.

In 1986, the worst nuclear accident occurred at Chernobyl, Ukraine. Even though scientists explained what happened, the general public concern about the risks connected with nuclear power could not be silenced. By 1989, the relative cost of nuclear power, public uproar, and the political climate led to the placement of a moratorium on the construction of additional nuclear power plants until 1994 (Allen and Oregon, 1989; Sweet, 1990).

At present, General Electric and Westinghouse have initiated studies into developing a safe and environmentally correct nuclear power plant.

Relevant Issues

When studying the nuclear industry, it is easy to understand the issues confronting it. In addition to satisfying the political debates, the private interest groups, the governmental regulations, the public concerns, and the financial costs, the nuclear industry is faced with providing a nuclear unit that meets additional criteria (See Table 6).

Requirements to be Met:
* safe working atmosphere * environmentally correct * accountable for the nuclear material * radiation safe * preventative maintenance program * staffed properly * quality power provider * accident free * reliable and efficient operation * constructed and operated economically * proper safeguards established

Table 6: Nuclear Unit Criteria

Along with the above mentioned issues, the nuclear industry in order to assure complete confidence must also deal with three other important side issues. They are the proliferation of nuclear material into weapons, theft of nuclear material, and waste storage. The proliferation of nuclear material into weapons and nuclear waste storage issues (Hileman, 1994) were discussed earlier in this chapter. In Hileman's article, it was pointed out that due to the break-up of the Soviet Union here is a clear and present danger with respect to these issues. The theft of nuclear materials issue has been brought to the forefront by articles in Time, Newsweek, and U.S.News and World Report. In the Time article discussed earlier in this chapter, Nelan (1994) points out that nuclear material thefts are also occurring because of the break-up of the Soviet Union. In the U.S. News and World Report article, Zimmermann and Cooperman

(1995) point out that the latest thefts of nuclear material from the Soviet Union have been arranged by organized-crime groups. This situation prompted FBI Director Louis Freeh to state that any sophisticated crime group with a network in international distribution and connections with the right buyers is a grave and immediate threat. As a result there is a great concern that current nuclear material safeguards are not adequate and nuclear material could fall into the wrong hands. Due to today's environmental concerns, the public perceives these three issues as being very critical and directly connected to the future of nuclear power.

One of the major factors contributing to these relevant issues is a general lack of understanding by our society of the concept of acceptable risk. Bromley (Rogers, 1991) has stated that if all areas of risk were eliminated completely, the potential for future advances would be greatly reduced. The most society can hope for is that the nuclear industry will put into place methods that will produce safeguards to minimize risk and provide a nuclear unit meeting the previously mentioned criteria.

Application

Human inventory is the current method for determining the removal, loss, or tampering of nuclear material. However, other methods have been examined in order to improve the monitoring of nuclear material. One of these methods is time series analysis which does not take into account all the elements governing nuclear material accountability such as receiving the raw material, monitoring material within a facility, and human intervention. Thus, in this dissertation, neural networks will be used to take into account all the elements governing nuclear material and to provide a method for overall accountability of nuclear material. The results will hopefully not only underscore the capabilities of neural networks as a superior method of forecasting, but also provide a breakthrough for the nuclear industry to demonstrate to the public a significant reduction of risk for the use of nuclear power in the future.

Another application of neural networks in the nuclear industry was done by Jouse and Williams (1991). Jouse and Williams looked at applying Drive Reinforcement Theory and a modified Barto-Sutton algorithm to a startup of a pressurized water reactor. The Drive Reinforcement Theory is a real-time learning mechanism for the synthesis of networks that interacts with its environment. On the other hand, the Barto-Sutton algorithm has neural-like elements which incorporate two types of associative memories and a reactor simulator as an additional layer. The reactor simulator was used to coordinate the pre- and post-synaptic activations across the reactor system. The training of the network was based on Stochastic initial conditions to meet a set of complex time dependent and only piecewise continuous objective functions. The results of this approach led to a network architecture that could learn to control a reactor simulator during a complex task (startup).

Today, the nuclear industry is applying Statistical Process Control methods and time series methods to assist in nuclear material management and safeguards. Goldman, Picard, and Shipley (1982) considered safeguards as the primary way of protecting special nuclear materials (SNM). Safeguarding of SNM in a broad perspective is a combination of two elements. The first element is the monitoring of the movement and the containing of SNM to prevent loss. The second element is the accounting procedures implemented to keep track of the quantities and the locations of SNM. Goldman et. al. did review specific issues in the relationship of safeguards concerning SNM. These issues included shipper-receiver differences, near-real-time accounting, spatially distributed materials balance, and data verification. The shipper-receiver differences occur between what the receiver receives and what the shipper sends. The difference is analyzed against a statistical value called the "limit of error" (LE). The usage of LE's has created controversies causing alternative methods to be considered. The near-real-time accounting issue deals with material losses in process equipment clean out and in process inventory. The accounting procedure depends on taking timely measurements of nuclear materials and is illustrated by the following equation.

$$MB_t = I_{t-1} - I_t + T_t$$

where

I_t	inventory measured at the end of day t
T_t	measured difference of materials transferred into and out of the balance area during day t
MB_t	normally distributed about zero when all material is properly handled (material balance)

This material balance is also discussed in the book by Avenhaus (1977). The spatially distributed materials balance deals with material being transferred between areas of a facility. This issue requires that a calculation of a materials balance for each area be done and the two materials balances are statistically analyzed. The last issue is data verification where measurements taken by an operator are selectively verified by an inspector. The approach used here is based on material unaccounted for (MUF) by the operator and is statistically analyzed. The authors hoped to spark interest in looking at different safeguard methods. Speed and Culpin (1986) also examined the role of statistical methods in nuclear material accounting regarding safeguards. In their article, Speed and Culpin reviewed several statistical tests which included CUMUF test, Page test for MUF residuals, Power-One type test for CUMUF (Cumulative MUF), Robust test, and the likelihood ratio test. As a result of their analysis, Speed and Culpin found that problems existed in the usage of standard control charts, time series methods, and statistical methods for the tracking and accounting of nuclear material. With no other current state-of-the-art methods, statistical methods are playing a vital role because measurements are not always exact and assurances are not always certain. Speed and Culpin agreed with Goldman, Picard, and Shipley on the two principal roles that statistics can play. Further, they pointed out three levels of safeguards. They are within an individual plant, the nationary regulatory body, and the International Atomic Energy Agency. Also, Speed and Culpin reviewed the material balance equation of Goldman et al. (1982) and agreed with their approach. Goldman et. al. and Speed and Culpin further agreed that the material balance equation was being affected by errors. These errors included systematic errors, non-measurement errors, and inventory errors. Systematic errors arise from plugged probes, solid buildup in tanks, miscalculation of measurement devices, and so on. Non-measurement errors arise from operators misreading, mistranscribing, miscalculating, and so on. Inventory errors deal with the diversion of nuclear materials over time which cause the errors. Finally, Speed and Culpin discussed anomalies and alarms. The authors examined the probabilities for the detection of false alarms in relationship to the diversion of nuclear material. What they concluded was that assigning cost, such as the cost of a false alarm versus the cost of not identifying a diversion, was the only way to deal with trade-offs and establishing process control limits for the desired sensitivity of any safeguard method.

Even with these methods, the nuclear industry senses a need for a more highly reliable, predictable, accurate, fault resistant, and responsive method. With this need in mind and the successful applications presented, the logical direction for the nuclear industry is to continue the implementation of neural networks. Thus, utilizing the nuclear industry data that is available for this dissertation, the proper direction for further integration of neural networks is in the process control of nuclear materials management.

CHAPTER 3: RESEARCH METHODOLOGY

Chapter Overview

The previous two chapters provided the foundation to proceed with the research methodology. The research methodology delineates the procedure to examine the performance of the neural network approach and meet the research objectives. The research methodology consisted of first selecting data sets that were representative of nuclear material processes and other production processes. Second, the neural network algorithm was selected based on performance and ability to handle all the data sets used in the research. Third, to enhance the neural network algorithm, a computer simulation program was developed to ensure that outliers were present in the training data set. Fourth, the various data sets were analyzed using the neural network algorithm with the simulation enhancement to deal with outliers and terminal points. Finally, the results from this approach were compared against results from other proposed methods based on the same data sets. As a result, the research methodology meets the objectives of this research which included selecting an appropriate neural network architecture, investigating the neural network's ability to interpret control chart data, determining the neural network's ability to detect process disturbances (outliers), and comparing the neural network approach against time series or control chart methods based on the same process data sets. Thus, the critical aspect of this research methodology is that process control by a neural network based algorithm with the simulation enhancement has not been applied to the monitoring of nuclear material previously nor to more than one type of production process. Therefore, this research considers these aspects.

The chapter begins with a discussion of the research methodology's main elements which includes the various data sets analyzed, the simulation computer program, Denton's neural network algorithm, and the neural network architecture. The analysis of various data sets will include a discussion on data verification and validation. Next, the simulation computer program is reviewed. This computer program provides a technique to ensure that outliers are incorporated into the training data set for the neural network for there is no guarantee that an actual data set will have outliers included. Then Denton's neural network algorithm is examined with various selected parameters described. The neural network's architecture for this study is explained using the input layer, the hidden layer (or layers), the output layer, the weights, and the arcs. Finally, a discussion establishing the basis for comparing various methods currently being used with the neural network method for process control is presented.

Main Elements

Review of the Various Data Sets

The data sets utilized in this dissertation are presented in the following tables. In Table 7, the data presented was obtained from a process of dyeing woolen yarn. In this process, it is desirable to control the acidity of the dye liquor. The data was collected during a normal work week for a period of approximately five weeks. Generally, each day, one of three determinations of the acidity was taken. This data was then used to develop the X-Bar and R values. An X-Bar control chart was plotted with the center line at 4.22 and the control limits at 4.05 and 4.39. From this chart it can be seen that the process had experienced a fluctuation (Grant and Leavenworth, 1980).

In Table 8, the data shown on inventory differences was taken from the Energy Research and Development Administration's 1977 "Report on Strategic Special Nuclear Material Inventory Differences." The inventory differences were for Plutonium, enriched Uranium, Pu-238, and U-233 from 1949 to the first half of 1977. The data was acquired from four nuclear facilities [Los Alamos Scientific Laboratory (LASL), Oak Ridge National Laboratory (ORNL), Richland Hanford, and Savannah River]. The analysis of this data was done on the basis of no inventory difference given a value of zero and inventory differences of less than 50 grams given a value of -0.5 (Chernick, Downing, and Pike, 1982). In the original article Chernick et. al. (1982) wanted to investigate the effects of outliers on time series data by applying the influence function for the autocorrelations, p(k), of a stationary time series. This influence function was defined by Hampel (1974). Chernick et. al. defined their function as:

$$I(H, p(k),(z_t,z_{t+k}) = z_t z_{t+k} - p(k)(z_t^2+[z_{t+k}]^2)/2$$

Where:
- H = bivariate distribution function
- p(k) = autocorrelation at lag k for a stationary time series
- z = point of interest
- t = period of time

Date	X-Bar	R	Date	X-Bar	R
Jan.			Feb.		
30	4.17	0.14	14a	4.25	0.11
31a	4.15	0.30	14b	4.26	0.26
31b	4.08	0.20	15a	4.10	0.18
Feb.					
1a	4.07	0.09	15b	4.14	0.23
1b	4.13	0.10	16a	4.20	0.52
1c	4.22	0.24	16b	4.24	0.17
2a	4.33	0.65	19a	4.21	0.46
2b	4.33	0.17	19b	4.11	0.20
5a	4.54	0.58	20a	4.07	0.40
5b	4.50	0.22	20b	4.22	0.12
6a	4.54	0.22	21a	4.11	1.34
6b	4.61	0.18	21b	3.72	0.96
7a	4.36	0.44	22a	4.18	0.35
7b	4.61	0.20	22b	4.29	0.31
8a	4.37	0.23	23a	4.17	0.20
8b	4.54	0.23	23b	4.14	0.13
8c	4.29	0.32	26a	4.32	0.26
9a	4.35	0.62	26b	4.26	0.08
9b	4.31	0.28	27a	4.16	0.51
12a	4.32	0.20	27b	4.25	0.25
12b	4.36	0.40	28a	4.28	0.09
13a	4.27	0.40	28b	4.26	0.15
			Mar.		
13b	4.28	0.38	1	4.14	0.11

Table 7: Acidity of Dye Liquor

Facility	LASL		Richland	ORNL			Savannah
Material	U-233	PU-238	U-233	U-233	PU-238	U-233	PU-238
Data Set	1	2	3	4	5	6	7
Year							
1949	0.1	-1.1	1.0	-0.05	-0.05		
1950	0.1	-2.1	0.3	-0.05	-0.1		
1951	-0.2	-7.8	-2.9	0.0	-0.05		
1952	0.0	-3.7	1.0	-0.1	-0.05	-0.05	
1953	-5.7	-1.7	0.3	-0.1	-0.3	-0.05	
1954	-8.1	-1.8	-24.6	- 0.4	-0.6	-0.05	
1955	-0.4	-1.1	-23.7	0.3	-0.1	0.1	
1956	0.9	-1.5	-11.6	-0.05	0.5	-0.1	
1957	-3.6	-1.0	-55.7	-0.1	-0.1	0.1	
1958	-9.3	-1.4	-86.4	-0.1	-0.4	-0.8	
1959	-2.1	0.4	-90.1	0.0	-0.3	-1.4	-0.05
1960	-4.9	-2.6	-143.7	-0.05	0.1	-0.05	-0.05
1961	-3.1	-1.9	-169.2	-0.05	-0.1	-0.1	-0.05
1962	-6.8	-1.2	-106.8	0.0	-0.7	-0.4	-0.1
1963	-1.5	-0.4	-66.9	-0.05	-0.9	-0.1	-0.5
1964	-3.5	-2.1	-94.9	-0.05	-0.05	-0.9	-1.9
1965	-3.4	0.9	-118.8	-0.05	0.9	-0.6	0.8
1966	-3.4	1.0	17.2	-0.05	-0.3	0.3	1.3
1967	-1.0	-0.7	-1.5	-0.3	-0.2	-0.4	-1.0
1968	3.2	-6.1	-32.0	-0.6	0.3	0.1	-2.5
1969	0.2	-14.3	62.9	-0.4	-0.7	-0.5	-1.7
1970	1.9	1.3	-9.0	-0.5	-0.6	-0.5	-19.6
1971	-0.1	5.7	0.9	-0.4	-0.3	-0.05	-2.9
1972	6.1	-15.7	-40.6	-0.2	-2.0	-0.05	1.6
1973	-1.3	5.2	-49.1	-0.2	-0.1	-0.1	2.1
1974	-4.8	-79.6	21.2	-0.2	-0.3	-0.05	-2.1
1975	-5.2	4.8	-3.4	-0.1	-0.1	-0.05	-1.3
1976	-0.05	22.7	6.7	-0.05	0.2	-0.05	-2.0
1977	0.2	-2.6	-3.0	-0.05	-0.05	-3.3	-2.5

Table 8: Inventory Differences

Item #	Data	Item #	Data	Item #	Data
1	0.5	29	-0.17	57	-0.87
2	0.3	30	-3.11	58	-0.51
3	0.1	31	2.74	59	-1.04
4	0.15	32	0.37	60	-1.55
5	-1.05	33	-0.86	61	0.35
6	-0.5	34	-0.18	62	1.03
7	0.5	35	-0.17	63	0.69
8	-0.5	36	-1.55	64	0.17
9	1.0	37	-1.38	65	-0.17
10	1.91	38	-0.69	66	-0.69
11	0.69	39	-0.69	67	2.42
12	0.0	40	-3.1	68	1.73
13	0.2	41	-1.05	69	5.5
14	0.5	42	-0.33	70	1.74
15	-11.55	43	-0.52	71	0.0
16	-3.45	44	-1.2	72	0.34
17	-2.6	45	-1.04	73	0.0
18	0.2	46	-3.45	74	-0.69
19	0.15	47	-1.38	75	6.21
20	0.05	48	-0.68	76	-5.51
21	0.5	49	-1.21	77	0.17
22	0.15	50	0.17	78	0.0
23	-1.39	51	-0.52	79	0.17
24	-1.76	52	0.0	80	0.17
25	-0.1	53	-0.34	81	0.35
26	0.1	54	0.17	82	0.34
27	0.56	55	-1.72	83	0.41
28	4.48	56	-1.9	84	-2.13

Table 9: Nuclear Material Loss

The authors applied their technique to three different applications. The data sets from one of the three different applications are presented in Table 8. For the data sets in Table 8, Chernick et. al. used a critical cut-off value of one. As a result, the authors were able to establish outliers for the data sets in Table 8 as well as for the other sets of data in their study.

In Table 9 the data presented represents minirun three which is one of the five miniruns obtained from the experiments at the AGNS Barnwell Nuclear Fuels plant (Cobb, Dayem, Baker, Ellis, Ehinger, and Crawford, 1981). The Los Alamos Safeguards System Group conducted these experiments to demonstrate near-real-time accounting and control techniques for a nuclear facility. The group performed their study on a Plutonium Purification Process. The Plutonium Purification Process has various cycles that constitute its operation. The experiments dealt with the second and the third Plutonium cycles that used natural uranium solutions. A total of five miniruns were completed with each being seven days in duration except for minirun five which was five days in duration. The authors used various evaluating techniques on the minirun data which included material balance accounting, in-process inventory determinations, and a decision analysis method (Shipley, 1978). The decision analysis method was implemented utilizing a Los Alamos developed computer program on the AGNS minicomputer (PDP 11/35). The results of the miniruns indicated that near-real-time accounting can detect losses (outliers) of nuclear materials from the process area of a large nuclear facility (Cobb, Dayem, Baker, Ellis, Ehinger, and Crawford, 1981). Therefore, these data sets provide this research with authentic test data for determining if neural networks can function accurately in a nuclear facility.

In addition, three other production process data sets will be analyzed. These data sets deal with monitoring the quality of production processes. They include the diameters of injection pump bore holes of automobile diesel engine blocks (Quesenberry, 1986), the sheet-like process (Johnson, et. al. 1974), and transmission parts from a manufacturing process (Quesenberry, 1990). Also, various other studies have used these data sets which include Sebastian (1994) and Prasad (in press) to identify process disturbances (outliers) and relate them to a particular production process problem. As a result, it has been found that various methods, especially the neural network approach with the simulation computer algorithm are suitable for on-line real-time process control.

Each of the data sets used in this study were verified and validated. Verification refers to a set of activities that ensure correct implementation of a specific function or data. It addresses the adequacy of the data sets, particularly during the evaluation and correlation of the network design with the neural network requirements (Touchton and Rausch, 1992). Validation refers to a separate set of tests directed toward justifying that the data sets satisfy the research requirements. These tests can also ensure the functionality and operability of the neural network (Touchton and Rausch, 1992). As a result, these various data sets did meet the criteria for verification and validation.

Denton's Neural Network Algorithm

Today, neural network technology has become an important topic in process control. The fascination with this field is based on the neural network's ability to learn from being exposed to information or data and then to utilize the information or data to make decisions in a manner that is similar to a human brain. Hecht-Nielsen (1989) provided a definition, that supports this idea, "A neural network is a computing system made up of a number of simple, highly interconnected processing elements, which processes information by its dynamic state response to external inputs."

This unique capability of a neural network for processing information provides specific advantages. These advantages include adaptive learning, self-organization, fault tolerance via redundant coding, real-time operating ability and ease of insertion into existing technologies (Maren et. al., 1990). Beside these advantages, there are other major reasons for using neural networks in process control. The algorithm learns a process on its own. The approach works with any product. It anticipates process disturbances. Finally, it solves difficult problems in less time. Thus, the neural network's ability to respond to process changes satisfies the goal of obtaining faster and more complete solutions to process changes.

The neural network algorithm used in this study was developed by Denton (1993). The program formulates a set of network weights using standard back propagation training or nonlinear optimization algorithms and line searches in conjunction with backward error propagation. Backward error propagation refers to the method of assigning errors to hidden network nodes. It is not to be confused with the training method known as back propagation. Backward error propagation does require that the network structures be feedforward and because of this requirement, the output of any node cannot become part of the input of that node.

The algorithm does require various parameter values to be established. However, if they are not, there are default values for the unspecified parameters. The parameters with their default values are presented in Table 10. The actual specification of parameters is accomplished within the definition file. The definition file specifies training parameters, instructions for node connections, and/or output control information. It further specified the neural network architecture, (In this case, it is two input nodes, one output node, and five hidden nodes.), the tolerance for training and testing phases, and the number of test data sets. Table 11 presents an example of a definition file.

For the algorithm to use the definition file, the problem must be assigned a name and then the file will be labeled "NAME.def." Besides the definition file two other files are required. These are the training and the testing files. The training file contains the data that consists of input data points and the desired output for each point. This file is labeled "NAME.trn." The testing file contains the data that consists of the input data points and the output for each data point set to zero. This file is labeled "NAME.txx" where xx is a number between 01 and 99.

When the neural network computer program is initiated, the computer program reads the definition file and checks for a training file, the number of test files, and parameters with designated values. One of the more important parameters is called the seed parameter. The seed parameter is an unique value for each run and unless fixed in the definition file, a different default seed parameter value will be used for each run. The seed parameter selection is important in the successful training of a neural network and its repeatability. The computer program provided various outputs such as the final weights of the training phase and the results of the training and testing phases. The results of the neural network computer program applied to the data sets previously mentioned in this chapter can be seen in the various Tables in chapters 4 and 5.

Table 10: Neural Network Algorithm Parameters

Parameter	Default Value
1. Number of Input Nodes	Must Specify
2. Number of Output Nodes	Must Specify
3. Number of Hidden Nodes	Must Specify
4. Traineps- Success Criteria During Training	0.05
5. Testeps- Success Criteria During Testing	0.05
6. Initwt- Specifies Source of Initial Weights	Randomly Generated with a Uniform Distribution Center Zero
7. Seed- Initializes a Random Number Generator for Weight Generation	System Clock
8. Test- Number of Test Sets	Must Specify
9. Input Transformation	None
10. Output Transformation	None
11. Termobj- Termination Criterion on Objective Function Value	Must Specify
12. Maxiter- Termination on Number of Iterations	Largest Representable Integer
13. Maxsuc- Termination on Number of Successes	Set to the Largest Possible Integer Representation
14. Search Direction	lbfgs Method
15. Store- Number of Vectors Store	3
16. Gradzero- Termination When Max. Gradient Component is < a Floating Point Number	0.001
17. Improvezero- Termination Occurs if Objective Change < Improvezero for Improvenum Iterations	0.001

Table 10: (continued)

Parameter	Default Value
18. Wrange- Range Around Zero for Initial Weights to be Generated	5
19. Bound- Symmetric Bound for Weight Values	+10 +/- 10e
20. Node Connection	Standard

Table 11: Definition File

Parameter	Value
1. Input Nodes	2
2. Output Nodes	1
3. Hidden Nodes	5
4. Traineps	0.1
5. Testeps	0.01
6. Termobj	0.0001
7. Tests	01
8. Connect	Standard

The Simulation Approach to Training the Neural Network

The simulation computer program begins by reading two files. One file contains the total number of data points, the multiplying factor of the robust standard deviation based on the desired control limits for the process and the selected control method's sensitivity (e.g. 2.5 (robust standard deviation)), and the indication of the target value (zero or the robust process mean). The robust process mean and robust standard deviation are described below. The target value or center line represents the average level or value for the production process from which the control limits are calculated (e.g. target value +/- 2.5(robust standard deviation)). The second file is the complete data set. After reading these two files, the computer program then splits the data set into two groups, A and B. A contains the first 80% of the observations and B contains the remaining 20%. Other splits are of course possible. A robust process mean and robust standard deviation were computed from the A set as described in Booth, 1986. These two values are used as the basis for developing our simulated training data set.

After calculating the robust process mean and the robust standard deviation, the simulated data set can be developed by the computer program. The simulated data set is created by using the sine function (i.e. to generate both + and - outliers). The sine function provides a means to provide outlying data points and balance them above and below the target value as depicted in Figure 13. First, the maximum plus and minus variation about the robust process mean or target value is established. These values are calculated by multiplying the robust standard deviation by four. This multiplying factor of four was selected because a data point located within three standard deviations is considered an in-control point. Thus, by using a value of four we insure that there will be simulated outliers to train on in our new simulated data set. Now, the simulation program can start to create the simulated training data set. The simulated data set is divided into sixty individual time periods. The simulated data set could be divided into any number of individual time periods depending on the application. We felt that sixty individual periods were reasonable for a sine function with our data sets. For the first nine time periods, the data points for the simulated data set are equal to the robust process mean. Next, for the periods ten through thirty, the data points for the simulated data set are calculated using the upper half of the sine function from zero to pi. The result from calculating the sine function for each period of time is multiplied by the maximum plus limit of four times the robust standard deviation. The resultant value is added to the robust process mean to generate + outliers. For the periods thirty-one through thirty-nine, the data points for the simulated data set are again equal to the robust process mean. Finally, for the periods forty through sixty, the lower half of the sine function from pi to two pi is used to calculate a value for each period of time. This resultant value is multiplied by the minimum negative limit of four times the robust standard deviation and then added to the robust process mean to get the final data points (the lower outliers) for the simulated data set.

With the simulated data set complete and if the target value is selected at zero, the robust process mean will be subtracted from each simulated data point forming a new simulated data set. However, if the target value is to be the robust process mean, the formulated simulated data set will not be altered (See Figure 13). The sine function was selected in order to assure that outliers would be in the training data set. We are currently developing methods to completely simulate different outlier types using this approach. Hopefully then the simulation computer program will have the potential of incorporating the equations for the outlier types as described by Chen et. al., 1992 and Thome, 1995, and in turn, training the neural network to identify any particular type of outlier. This will allow the analyst to be sure that he/she can detect any expected process problem type. Finally, the training data set for our program is organized from the simulated data set by three column vectors. The first two columns are the input data for training and the third column vector is the output data for training. The first column contains the simulated data set. The second column is the simulated data set points but shifted forward one observation to account for a time period change. Thus, if we call the simulated data X, we have the situation as shown in Table 12. In Table 12, column three is the output data for training which contains a zero for a point inside the established control limits and a one for a point outside the established control limits (an outlier). Thus, the training data set will contain outliers. The test data column vectors are formed in the same way using the test data sets (A Union B) except that column three initially contains all zeros. As the algorithm discovers outliers those observations are changed to ones and thus the outliers are identified. The reason that the form of the first two vectors in Table 12 was used is that the vast majority of real data we have worked with are AR(1) and therefore a one period lag is appropriate. We started the subscripts at t rather than zero to indicate that we can start this procedure anywhere in the series.

Table 12: Column Vectors

Column 1	Column 2	Column 3
X_t	X_{t+1}	
X_{t+1}	X_{t+2}	
X_{t+2}	X_{t+3}	
X_{t+3}	X_{t+4}	
etc.	etc.	

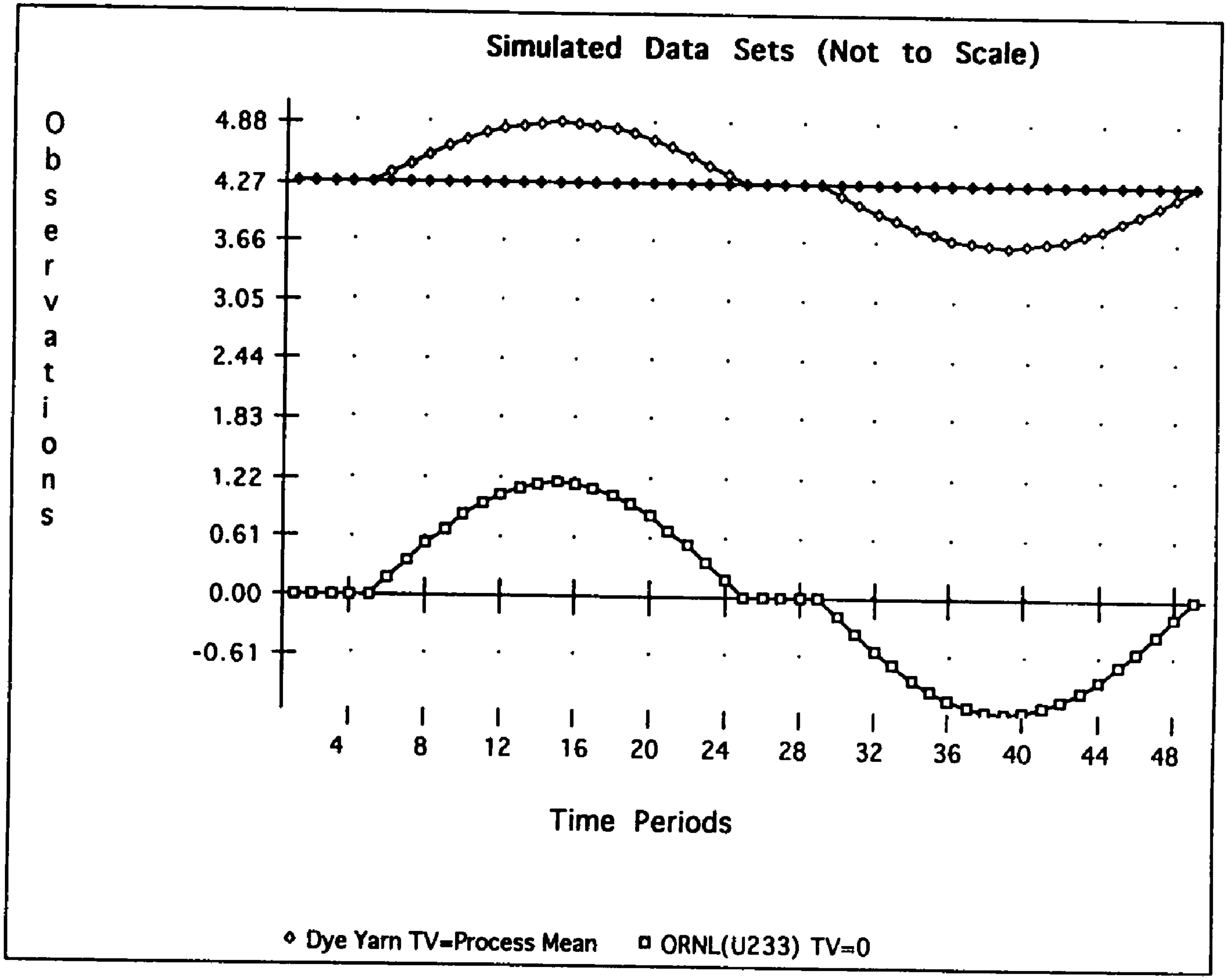
Now, at this point we have three data sets per application, simulated, A, and B. We train on the simulated data set and test on the union of A and B. Notice that we are not using the A set as the training data set (we only use its robust process mean and robust standard deviation to develop the simulated training data set). Thus, we are not training and testing on the same data set.

Using this approach, we are able to develop a complete training set of data containing outliers, even if all the process data available is in control.

Neural Network Architecture

For this research, a particular neural network was selected to satisfy the two criteria of time shifts and proper number of hidden nodes. Thus, the architecture that

Figure 13



Simulated Data Sets (Not to Scale)
Observations
4.88
4.27
3.66
3.05
2.44
1.83
1.22
0.61
0.00
-0.61
4
8
12
16
20
24
28
32
36
40
44
48
Time Periods
Dye Yarn TV=Process Mean
ORNL(U233) TV=0

was implemented for this research consists of two input nodes and one output node. Also, it used one hidden layer with five hidden nodes. The two nodes in the input layer were chosen because the vast majority of the data sets used in this research were of an AR(1) type and therefore, a one period lag was appropriate. Thus, X_t is applied to Node 1 and X_{t+1} is applied to Node 2 (See Figure 14). The choice for the number of hidden nodes is based on the current literature which states that the most widely selected number of hidden nodes is 2n+1 with n being the number of input variables (Patuwo, Hu, and Hung, 1993). The sensitivity of this factor can be measured by adjusting the number of hidden nodes (Patuwo, Hu, and Hung, 1993). With this type of neural network architecture (See Figure 14) a thorough analysis was accomplished and firm conclusions drawn.

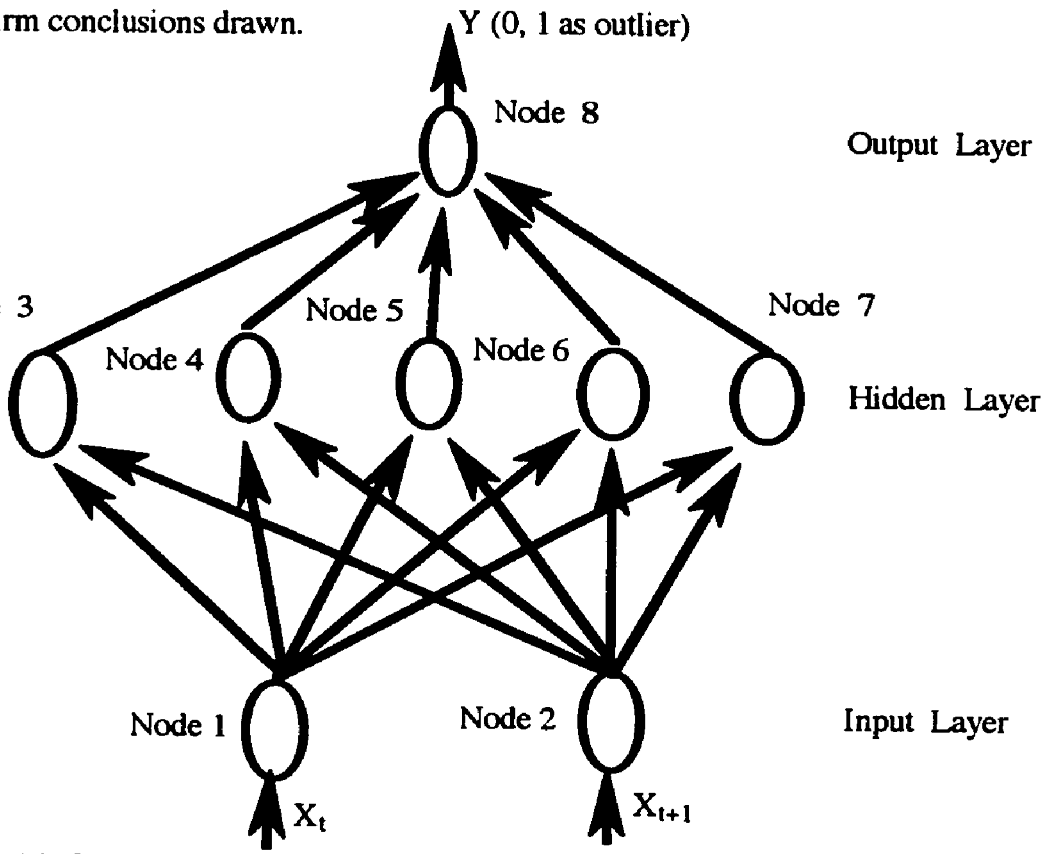


Figure 14: Research Neural Network Architecture

In performing the training phase, the neural network is presented a data set with various process disturbance points to achieve a particular outcome. This training is important because it sets up the neural network to look at streams of data coming from an actual process and identifying process disturbances early enough to prevent a problem from occurring and in turn reducing false alarms. The data set presented to the neural network is defined as the training data set and is composed of column vectors with the first two columns being the input data and the third column being the output data (See Previous Section for Details). Each output data point is a zero if that data point is within the control limits and a one if that data point is outside the control limits. However, a problem could arise with using actual data sets in two ways. First, the actual data set may not have enough data points for both the training and the testing of a neural network. Second, the actual data set may possibly not have any process disturbance points (i.e. outliers) for the proper training of the neural network. Therefore, a method that would provide a training data set based on actual data from a production process and one that is sure to contain all needed examples of process problems would be of great benefit to the statistical process control industry for use in applying neural networks. In response to this idea, this research has developed such an approach based on simulation, which was previously described. The basic idea behind the simulation was that control process data sets don't always contain outliers (i.e. out-of-control points). However, the purpose of the training phase was to train the neural network to detect outliers. Thus, this research used a simulation based on real control process data to generate an appropriate training data set containing outliers.

The neural network computer program provides various outputs such as the final weights of the training phase and the results of the training and testing phases.

The results of the neural network computer program applied to the data sets in this research can be seen in the various Tables (See Chapters four and five). Again, it must be stressed that due to the importance of not overlooking an outlier, as in the case of production processes, management must be willing to tolerate a certain frequency of false alarms. Thus, a questionable outlier can serve as an early warning signal that the production process might be deteriorating. Management is then able to be more aware of a potentially abnormal situation, and the operator can examine the critical points in the process, look for leaks, and monitor the process more closely.

Comparison Criteria for Research Results

In chapters four and five, the research results from applying various methods on various data sets are tabulated. Then, certain items were determined for each method based on the following notation:

TOI = total outliers identified

T+OI = total positive outliers identified (nuclear data sets)

NA+I = number of actual outliers

NFA = number of false alarms

N+OM = number of outliers missed

The comparison of the various methods was based on the tabulation of the above items but more importantly the analysis was done on the detection of outliers while potentially reducing false alarms. Also, the various methods were compared as to their capability to handle terminal points. The research compared primarily the neural network approach with the simulation enhancement against other approaches developed by the Kent State University research group. Previous publications showed that these methods are as good or better than those currently in use (Prasad et. al., 1990, 1994, 1995, and Sebastian et. al., 1994, 1994, 1995). Also, the capability of the various methods to detect the final data point of a data set when it is a process disturbance point (outlier) is important because the final data point is the latest piece of information available from the production process. It further demonstrates their use for controlling a real-time on-line process.

CHAPTER 4: RESEARCH RESULTS FROM ANALYZING THE DETECTION OF NUCLEAR MATERIAL LOSSES

Chapter Overview

A series of repeated nuclear material balances forms a time series of often autocorrelated observations. Outliers, deviations from an in-control production process or time series pattern, indicate an out-of-control situation relative to the process norm. In this chapter, various methods, especially neural networks, were examined with respect to their use to detect nuclear material diversions or losses more rapidly and accurately than currently used methods. The neural network technique was enhanced with the use of a simulation computer program for creating the training data set. This simulation approach provides the opportunity of including outliers of various types in a data set for training the neural network because an actual process data set used for training possibly may not have outliers. Also in this chapter, the various methods are compared on their ability to identify outliers and handle terminal points. These various methods were tested on data sets of nuclear balances with known removals and the results were tabulated and described. Based on these results, the researcher believes the algorithms used will assist the nuclear industry in process control, provide a new approach to nuclear material safeguards, and also provide a new approach to training neural networks for process control applications.

Nuclear Material Safeguards and Related Research Methods

In Nuclear Materials Accounting the objective is to accurately determine if there have been material diversions or losses (e.g., leakage or theft) as soon as possible to minimize the threat to the environment and the public. Even though some progress has been made in this area, the control measures against leakage, theft, or other loss of radioactive material have continued to be of concern. The international crisis concerning the monitoring and safeguarding of nuclear material in the former Soviet Union has been highlighted by several instances of nuclear material from the former Soviet Union being found in Germany (Hileman, 1994 and Nelan, 1994). Furthermore, the National Academy of Sciences (NAS) has warned that a "clear and present danger to national and international security is posed" (Nelan, 1994). In addition, the leakage of radioactive material into the environment has the potential to adversely affect the well-being of the world's population. Therefore, any method that could detect losses or diversions quickly and accurately would indeed provide a significant contribution to nuclear safeguards as well as public safety.

This research demonstrates that nuclear material loss or diversion can be detected through the detection of statistical outliers (Youden et. al., 1975; Beckman, 1983 and Miller, 1984). Outliers may be thought of as observations in a set of data which appear to be inconsistent with the remainder of the data set. Traditional statistical process control (SPC) charts which attempt to detect outliers as do other current methods are generally based on two assumptions. First, the system's observations in a time series are independent (Anderson, 1987) and identically distributed (IID) about the process mean at any time, t. This independence suggests that the data set has no defined pattern. Second, the underlying distribution is normal when the process is in statistical control (Wardell et. al., 1992). In reality, the IID assumption of conventional control chart methods does not always hold. Therefore, traditional SPC procedures may be ineffective and inappropriate for monitoring and/or controlling the production process in these situations (Wardell et. al., 1992).

In nuclear material accounting, the material balance equations (Chernick et. al., 1982 and Sebastian et. al., 1994) are used to monitor and control nuclear material within a facility. The nuclear industry does often use other equations instead of the material balance equations which are called the material unaccounted for (MUF) equations. The MUF equations are equal to the material balance equations multiplied by a negative one. The positive MUF equation equates to a loss or removal of material. Thus, the results based on these equations could be a negative material balance, a positive material balance, or zero (which indicates no material loss). A negative material balance can occur because of leakage, theft, accounting errors, operator errors, etc. A positive material balance results from a hidden gain because of a systematic measurement error or an unaccounted for measurement for an abnormal amount of material in the process that later reversed itself (e.g. caking). Sebastian et. al. (1994) provides an in-depth discussion of this topic.

In order for a material balance to be documented, hundreds of measurements and estimates are often necessary. This large amount of data and related sources of error create difficulty in measuring the dependency among observations. Outliers can distort these parameters and affect the standard error value and thus, inaccurate conclusions may result. The advantage of the methods demonstrated in this research is that they detect the effect of these outlying ("out-of-control") observations.

Booth (1984 and 1986) details the relationship between time series outlier types and nuclear material accounting. An additive outlier (AO) affects a single observation. It is an abrupt removal or one time loss or theft. An innovative outlier (IO) affects sequential observations. This type can indicate leakage or systematic theft over a period of time. Thus, the early identification of an outlier and its type will minimize any loss or theft because the organization using this process would have the necessary information to deal with the problem effectively and quickly.

Prasad et. al. (1995) tested the joint estimation procedure (Chen et. al., 1993) on nuclear material inventory data sets and achieved satisfactory results. This procedure was used because of its ability to handle deviations from any type of time series model of a process. Also, it is a robust and discriminating process which identifies outlier types, using standard hypothesis testing - not only additive (an abrupt removal or one time loss) and innovative (a continuing loss of short duration) but also temporary change (TC) (a continuing loss of longer duration) and level shift (LS) (a long term protracted removal or continuing loss throughout the remaining series) outlier types can be observed.

Sebastian et. al. (1994) recently tested polynomial smoothing algorithms and data bounding algorithms for nuclear material inventory data sets again with satisfactory results. A polynomial smoothing algorithm combines a commonly used smoother with an "outlier detection procedure for process control." A data bounding algorithm is essentially a means of adjusting data points that lie beyond certain limits, thus smoothing significant peaks and troughs. These methods have been adapted to outlier detection in safeguards development. Thus with the proper fine-tuning of the sensitivity constants for these algorithms, the process engineer can choose the algorithm which is best suited for the application.

The problem of Nuclear Materials Accounting procedures and the estimations of associated parameter values constitutes an unique application of outlier detection to a crucial societal problem. Previously, the joint estimation procedure, data bounding, polynomial smoothing, and neural networks have been successfully tested on conventional industrial process control data sets with known process disturbances (outliers) (Prasad, 1990; Sebastian, 1994). In this research, an enhancement to the neural network approach for use in nuclear material accounting was discussed, and in fact, for any process control problem. This enhancement is based on a new simulation algorithm which was discussed in chapter three. This simulation algorithm creates a training data set from an actual data set of any size. The importance of this algorithm is that outliers can be incorporated into a training data set for a neural network while an actual process data set used for training may not include outliers. Thus, the simulation algorithm provides a crucial enhancement to the neural network method for the early and more accurate detection of outliers (i.e., process disturbances) while potentially reducing the possibility of false alarms. This procedure in conjunction with the neural network approach has also been successfully tested on production process data sets of various sizes with known process disturbances (Hamburg, 1995). Based on the results, the research shows that these methods, especially the neural network with the simulation enhancement feature, can be applied to nuclear material processes.

Nuclear Material Safeguard Data Sets with Know Diversions

To accomplish a more meaningful comparison, nuclear material balance data sets with known outliers (i.e. removals) were used. These data sets were taken from actual nuclear inventories known to have protracted removals of nuclear material. To illustrate the capabilities of the neural network approach, two particular cases of nuclear material removal have been selected. Each case will be discussed briefly. A discussion of the testing of various methods on the data set from each case will be presented in the next section.

In 1980, at the AGNS Barnwell Nuclear Fuels Plant (Dayem et. al., 1984) a series of experiments were performed in order to improve the detection of nuclear material losses or diversions in the different unit process accounting areas (UPAA). The material balance data were recorded periodically on a near-real-time basis with the aid of a computerized nuclear materials control and accounting system.

The experiments involved the physical removal of nuclear material in different forms from a Plutonium Purification Process (PPP). Also, location of removals, time, quantity, concentration, and type of diversion were noted. These experiments dealt with the second and the third Plutonium cycles that use natural uranium solutions. The original researchers called these experiments "miniruns."

In 1977, the Energy Research and Development Administration issued a report titled "Report on Strategic Special Nuclear Material Inventory Differences" (Chernick et. al., 1982). In this report, data was presented on inventory differences for plutonium, enriched uranium, U-233, and Pu-238 from 1949 to 1976 from various sites. These sites included Los Alamos Scientific Laboratory (LASL), Oak Ridge

National Laboratory (ORNL), Richland Hanford, and Savannah River. The Nuclear Regulatory Commission (NRC) has the responsibility for checking on nuclear material stored at their facilities. Thus, when the nuclear material inventory at a facility has a change, the NRC must investigate to determine the type of change. A negative change corresponds to possible losses or thefts while a positive change corresponds to a gain. This pattern is the reverse of MUF direction. This data was initially used by Chernick et. al. (1982) in their study of nuclear material safeguards.

This research first considered miniruns 3C and 5B of the AGNS study. Table 13 presents the location and timing of removals for the two minirun data sets. The total protracted removal across the extended time period by volume of solution and the corresponding uranium content is recorded in column 5.

Next, the research considered four of the seven nuclear material inventory differences (data sets) from Chernick et. al.(1982). These data sets were chosen at random and included U-233 from LASL, U-233 from Richland, U-233 from ORNL, and Pu-238 from Savannah River. Even though the data of the material balances presented is on an annual basis, it is still useful for testing the neural network algorithm and comparing it to other algorithms. However, due to the time period, a host of assignable causes are possible and consequently timely corrective actions are precluded except for after the most recent observation. In the article, Chernick et. al.(1982) used the Influence Function Matrix approach to detect outliers in the nuclear material balance.

Due to the importance of not overlooking an outlier, as in the case of production processes, management must be willing to tolerate a certain frequency of false alarms. Thus, a questionable outlier can serve as an early warning signal that the production process might be deteriorating. Management is then able to be more aware of a potentially abnormal situation, and the operator can examine the critical points in the process, look for leaks, and monitor the process more closely.

Experiment	Location	Start Time/Date	End Time/Day	Total Removal Vol. (L)	U. (Kg)	OBS.
Minirun 3C		0815/07/18	1204/07/21			
	Pulse Col. (2AP)	1645/07/18	0845/07/19	96	4.224	8-24
	Pulse Col. (3AP)	1645/07/18	1645/07/19	96	4.253	8-32
	Surge Tk. (1BP)	1615/07/18	0850/07/21	96	5.722	56-72
	"U" settled on	bottom of acid	concentrators	= 120Kg		
Minirun 5B		1600/11/18	1630/11/21			
	Surge Tk. (1BP)	0000/11/19	1200/11/19	174	10.2	9-20
	Surge Tk. (1BP)	1445/11/19	0115/11/20	177.2	11.1	23-34

Table 13: Location and Timing of the Removals

Testing and Comparing Various Methods for Outlier Detection

AGNS Barnwell Nuclear Fuels Plant Minirun Experiments with Know Diversions

Minirun 3C. Tables 13 and 14 show that there are three protracted removals from three locations in this experiment, two of which are in part simultaneous: 0.4186 Kg/hr during observations 8-24; 0.1701 Kg/hr during observations 25-32; and 0.3366 Kg/hr during observations 56-72. The 120 Kg that settled out in the concentrators averages to 1.58 Kg/hr. Over the entire period, the separated solids (120 Kg) and the amount removed (14.199 Kg) or 134.2 Kg should be an average for the material unaccounted for (MUF) of 1.77 Kg/hr over the 76 observations.

Neural network (1.1 SD, as described in chapter three), data bounding and polynomial smoothing detected the very first observation (no. 8) in the initial protracted removal period. The joint estimation procedure and neural network (2.5 SD) missed the first observation, but did detect no. 10, the negative observation. The joint estimation procedure detected the first positive protracted removal at no. 14 while the neural network (2.5 SD) detected it at no. 61. The neural network (1.1 SD) had the highest percentage of actual positive outliers identified and also was able to detect the terminal point. CUSUM missed the initial protracted removal period completely but did very well on the final protracted removal period.

Ideally, any comparison of results among different methods should be established upon equal attributes. However, the proposed methods do not have easily comparable attributes. Each method requires a different number of sensitivity constants. Even though the different methods are not exactly comparable, the comparisons done give a reasonable idea of their relative performances.

Minirun 5B. This run has two protracted removal periods with only a two hour interruption separating them, as can be seen in Tables 13 and 15. Both protracted removals are from the surge tank (1BP): 10.2 Kg (0.850 Kg/hr) during observations 9-20, and 11.1 Kg (0.925 Kg/hr) during observations 23-34. The protracted removal periods had proportionally less positive MUF's than the nonremoval periods, making it very difficult to detect true losses while minimizing false alarms.

The neural network algorithms (2.5 SD) and (0.9 SD) immediately detected a negative outlier at observation no. 1 while data bounding and polynomial smoothing detected a negative outlier at observation no. 9. None of these methods detected a positive outlier until observations no. 17 and no. 18. The joint estimation procedure only detected negative outliers at observations no. 19 and no. 20 (Prasad et. al., pending). This result could have occurred because the sensitivity constants may not be comparable to the other methods. All five methods failed to detect any loss in the second removal period (nos. 23-34) until observation no. 33. Again, CUSUM was not useful since it detected 56 consecutive outliers from observations nos. 3-58, well before and past the removal periods.

Chernick et. al. (1982) Nuclear Material Inventory Difference Data Sets without Known Diversions

In examining the comparison results presented in Tables 16 through 19, it can be seen that the neural network approach did detect the same outliers as the influence function matrix approach. However, the neural network approach in Tables 16 - 18 also detected additional outliers, and this difference in number could be related to different sensitivity constants and quantifiable attributes used by each method. Even with this possible difference between methods, there are certain observations that can be made. In Table 16, the two additional points detected by the neural network approach could be indicating additional outlying points that are an indication of a problem in the material balance at LASL. In Table 17, the influence function matrix approach indicates period 1964 (-94.9) as an outlier and 1962 (-106.8) as a point within the control limits. However, the neural network approach detected the points in just the opposite way. The most significant observation to be made concerns Table 18. Here the authors state, "Year 77 is worth noting because it illustrates the outlier being detected at the time of occurrence." Therefore, with the neural network approach's ability to detect the last observation (1977) as an outlier, it demonstrated its capabilities for use in real-time on-line process control. In Table 19, the outlier detected was the only outlier in the data set but it was nearly seven times larger in absolute value than the next most extreme observation.

Analysis. The outliers in these data sets were difficult to identify because each removal per period was minute and generally did not exceed the normal in-control fluctuations. In fact, the removals were within the (0.2 SD) of the target value. The percentages of positive material unaccounted for (MUF) observations during the removal and nonremoval periods are similar, making it difficult to detect losses. In minirun 3C (Table 13), for example, little difference can be seen between removal and nonremoval periods. Then too, in minirun 5B, there is a greater frequency of positive MUF's during the removal periods.

Even though the five methods adapt to process change as though it were a new process, the joint estimation procedure does this to a greater extent. Thus, the joint estimation procedure will detect only the first outlier during a protracted removal period. Then, when the process reverts back to its normal state, it may also detect that change as an outlier. Finally, the joint estimation procedure did have a lower incidence of false alarms because of the method's approach.

Precipitated Solids. In minirun 3C, outlier detection was further hampered by 120 Kg of uranium solids settling to the bottom of the concentrators. If this process was in a steady state condition or at equilibrium, the precipitation of solids would have been no problem. However, new material was constantly entering the process. Since these solids apparently were not considered in the nuclear inventories, both positive and negative MUF's would be recorded. However, it is assumed that the original researchers conducted the experiments properly and removed the solids from the concentrators before the beginning of each run. Nevertheless, there is no way of knowing at what rate the particles settled out. If the rate was constant, then the uranium precipitated out of solution at a rate of 1.58 Kg/hr for one period in minirun 3C. This precipitation exceeded the rate of the protracted removals which ranged from 0.170 Kg/hr to 0.735 Kg/hr which certainly complicates outlier detection.

Doubt continues to exist as to whether an identified outlier represents a loss (removal) or simply solids that have precipitated out of the solution. Thus, a number of the false alarms could be expected. However, the methods should be able to adapt to the settled precipitation of solids if the rate of precipitation is a relatively constant part of the process.

Again, all of the methods had difficulties in detecting all of the known outliers. However, in comparison to other methods, neural networks did quite well, though care must be exercised because of the different sensitivity constants that each method uses. Finally, the joint estimation procedure is more adaptable to detecting the first outlier of an adverse process change because it is adapting to what is essentially a new process.

Discussion

A drawback of some methods is in the fact that neither neural network, data bounding, nor polynomial smoothing as discussed in this research can identify the outlier type directly. On the other hand, the joint estimation procedure, utilizing complex statistical mathematics and mathematically modeling the process, can identify the type of outlier. However, the researcher believes that the neural network method could be trained to identify the type of outlier by the use of a simulated training data set incorporating the mathematical formula for each type of outlier. Researchers at Kent State University are currently developing such an approach.

In process control, especially in dealing with nuclear material, the researcher believes that both primary and secondary control procedures should be in place. The primary control procedure could be either a neural network, joint estimation, or polynomial smoothing which would have the main responsibility of controlling the process. The secondary control procedure could be one of the procedures just mentioned, CUSUM, traditional Control Charts, or other nuclear material accounting procedures. The secondary control method would function either as a backup in case of a failure of the primary control procedure or a verification of the primary control procedure. The various methods should be examined for use in a particular process prior to complete implementation.

The neural network approach, as the primary control method, can be trained, possibly with the use of simulation, to adjust to an out-of-control situation, especially in the case of consecutive outliers, as though it represents a new process or process adjustment. However, when the primary method is the joint estimation approach, it will usually only detect the beginning of a protracted removal or leakage situation. When the leakage stops and the process is brought back into control there is another process change back to the original condition. Therefore, the next point may appear as an outlier to the primary control method. Since the outliers during the adverse process change period may have gone undetected if the joint estimation approach is monitoring the new out-of-control process, the secondary control method is indispensable to keep beta (probability of a Type II error) low.

Real-Time On-Line Process Control Capability. The results from the methods presented in this research may lead the reader to conclude that the methods may only apply to retrospective control. However, these algorithms are capable of handling end points. Further, they can also apply to real-time on-line process control. Tables 14, 15, and 18 demonstrate that any outlier identified would have been detected by the neural network method if it were a final point. All identified outliers with assigned causes, except for possibly two exceptions, could be considered final points and thus a test could be conducted.

If an outlier is found by both the primary and secondary control methods, the practitioner can feel more assured that the process disturbance point is indeed an outlier and not a false alarm. Thus using the simulated data set for training a neural network to accurately pinpoint the outlier would eliminate some of the false alarms.

Any of the previous methods may confirm the indications of the neural network or vice versa to minimize alpha (the probability of a Type I error) when a false alarm could have a major economical impact. Similarly, a second method, whether it be a neural network or not, could be used as a means of outlier detection to minimize beta (probability of Type II error) if the economical impact of overlooking an outlier is crucial. By using a secondary confirmation method to detect outliers, the process control operator has at hand a large amount of information on a pending process disturbance. Therefore, future research will be required to verify the aforementioned statements.

Summary

The methods presented in this chapter are effective across diverse applications. The neural network approach is particularly useful in the case of autocorrelated data. It can handle severely abnormal distribution situations which are due to the presence of many outliers. Further, the neural network approach is enhanced by the simulation computer program in the detection of outliers and the reduction of false alarms.

Neural networks have certain advantages that the other techniques do not. First, the network learns about the current process on its own; as well as making adjustments in its learning to account for process changes. Second, an overall knowledge of mathematical or scientific theory is not important for daily application of the neural network. Third, the type of product has no impact on the neural network performance because the network is directly controlling the process, and thereby indirectly maintaining the product quality. Fourth, the network lessens the possibility of human error because there is no necessary intervention or interaction by a person if a feedback loop is constructed. Fifth, it anticipates the abnormalities in order to correct for them within the process because it has the ability to learn and adjust to process changes without outside interference. Sixth, it removes the need for trial and error in establishing the optimum control for the process because it learns the information from the process itself (Blaesi et. al., 1983). In general, neural networks have been seen to be able to solve the more difficult problems in less time than other conventional methods (Francett, 1989).

The simulation computer program provides important features in developing a training data set for a neural network. This computer program provides a technique to ensure that outliers are incorporated into the training data set for training the neural network because there is no guarantee that an actual data set will contain outliers. It further provides means of incorporating actual formulas for particular outliers (Chen et. al., 1992 and Thome, 1995). Overall, the simulation computer program provides a researcher with total flexibility in developing a training data set for a neural network based on the researcher's needs.

It can not be overly stressed that the most important feature of these methods which have been presented is the ability to identify outliers well within the usual three standard deviations from a target value. They are a favorable addition to those methods already available to the process control industry involved in nuclear material safeguards.

Table 14: Minirun 3C-UPAA PPP(2AP, 3AP, 1BP)
Detection of Outliers among Uranium Inventory Differences (MUF) - AGNS Barnwell Nuclear Fuels Plant

* = detection of outlier
*+ = detection of outlier beyond TV ± 3(MSD) for data bounding or polynomial smoothing
total number of data points = 76
min. value = -40 Kg, max. value = 15.1 Kg, target value(TV) = 0
total no. of observations during removal periods = 42
total no. of observations during nonremoval periods = 34
no. of + observations during removal periods = 26(61.9%)
no. of + observations during nonremoval periods = 21(61.8%)
amount of uranium that settled out = 120 Kg or 1.6 Kg/hr

Data bounding and Polynomial Smoothing:
modified process average = 1.145
modified standard deviation = ± 3.571
with %drop = 5%(4)
median = 1.65

Neural Networks:
Weighted Process Mean = 1.616
S. D. = ± 3.855

Time Period	MUF	DB 1.1SD ref. 206	Poly Smooth 1.1SD ref. 206	Joint Estim. ref. 206	CUSUM ref. 206	NN 2.5SD	NN 1.1SD	Assigned Cause
4	-6.5	*	*				*	none/unknown
8	5.5	*	*				*	protracted
10	-8.0	*	*	*		*	*	removal
13	1.8							0.4186 Kg/hr
14	4.7			*			*	OBS. 8-24 17 periods
15	4.4							" "
16	6.2							" "
17	2.8							" "
18	5.2						*	" "
19	6.7	*	*				*	" "
20	3.1							" "
22	1.1							" "
23	3.1							" "
24	1.3							" "
25	5.3	*	*				*	protracted
26	0.3							removal
29	-7.8	*	*	*			*	0.1701 Kg/hr
30	0.3							OBS. 25-32
31	1.9							8 periods
37	5.4	*	*				*	none/unknown
39	7.2	*	*	*			*	" "

Table 14: (continued)

Time Period	MUF	DB 1.1SD	Poly Smooth 1.1SD	Joint Estim.	CUSUM	NN 2.5SD	NN 1.1SD	Assigned Cause
42	6.6	*	*				*	none/unknown
45	7.0			*			*	" "
56	0.45							protracted
58	-5.0	*					*	removal
59	3.0							0.3366 Kg/hr OBS. 56-72
60	1.8							17 periods
61	15.1	*+	*+	*	*	*	*	" "
62	-5.2	*	*				*	" "
63	3.6				*			" "
64	2.4				*			" "
65	3.6				*			" "
66	-1.0				*			" "
67	1.4				*			" "
68	1.6				*			" "
69	-2.3				*			" "
70	0.0				*			" "
71	0.7				*			" "
72	-40.0	*+	*+	*	*	*	*	" "
74	-5.2			*			*	none/unknown

Total (TOI)	13	12	8	11	3	17
Total (T+OI)	7	7	4	7	1	10
Actual(NA+I)	4(15.4%)	4(15.4%)	2(7.7%)	7(26.9%)	1(3.8%)	6(23.1%)
False (N+FA)	3(14.3%)	3(14.3%)	2(9.5%)	0	0	4(19%)
Missed(N+OM)	22	22	24	19	25	20

Table 14: Minirun 5B-UPAA PPP(1BP)
Detection of Outliers among Uranium Inventory Differences (MUF) - AGNS Barnwell Nuclear Fuels Plant

* = detection of outlier
*+ = detection of outlier beyond TV $\pm$ 3(MSD) for data bounding or polynomial smoothing
total number of data points = 68
min. value = -2.8Kg, max. value = 1.3Kg, target value(TV) = 0
total no. of observations during removal periods = 24
total no. of observations during nonremoval periods = 44
no. of + observations during removal periods = 9(37.5%)
no. of + observations during nonremoval periods = 26(59.1%)

Data bounding and Polynomial Smoothing:
modified process average = 0.04
modified standard deviation = $\pm$0.613
with %drop = 2%(2)
median = 0.075

Neural Networks:
Weighted Process Mean =0.036
S. D. = $\pm$ 0.482

Time Period	MUF	DB 0.9SD ref. 206	Poly Smooth 0.9SD ref. 206	Joint Estim. ref. 206	CUSUM ref. 206	NN 2.5SD	NN 0.9SD	Assigned Cause
1	-1.5					*	*	none/unknown
2	-1.7					*	*	" "
3	-2.8	*+	*+		*	*	*	" "
4	-1.8	*			*	*	*	" "
9	-1.1	*	*		*	*	*	protracted
13	0.05				*			removal
14	0.18				*			0.850 Kg/hr
15	0.1				*			OBS. 9-20
16	0.3				*			" "
17	0.75	*			*		*	" "
18	0.95	*	*		*		*	" "
19	-0.35			*	*			" "
20	-1.3	*	*	*	*	*	*	" "
30	0.15				*			protracted
31	0.2				*			removal
33	0.8	*	*	*	*		*	0.925Kg/hr OBS. 23-34
43	1.0	*			*		*	none/unknown
48	0.85	*			*		*	" "
49	1.3	*			*	*	*	" "
58	0.5						*	" "

	DB	Poly Smooth	Joint Estim.	CUSUM	NN 2.5SD	NN 0.9SD
Total (TOI)	10	5	3	56	7	13
Total (T+OI)	6	2	1	28	1	7
Actual (NA+I)	3(33.3%)	2(22.2%)	1(11%)	9(100%)	0	3(33.3%)
False(N+FA)	3(11.5%)	0	0	19(73.1%)	1(3.8 %)	4(15.4%)
Missed (N+OM)	6	7	8	0	9	6

Table 16: DETECTION OF OUTLIERS IN LASL U-233 MATERIAL BALANCES

* = detection of outlier
total number of data points = 29
min. vallue = -9.3, max. value = 6.1, target value = process mean
Weighted Process Mean = -2.008
Standard Deviation = 2.817

Period	MB	Influence Function Matrix ref. 50	Neural Network 1.5 SD
1952	0.0		
1953	-5.7		
1954	-8.1	*	*
1955	-0.4		
1956	0.9		
1957	-3.6		
1958	-9.3	*	*
1959	-2.1		
1960	-4.9		
1961	-3.1		
1962	-6.8		*
1963	-1.5		
1967	-1.0		
1968	3.2		*
1969	0.2		
1970	1.9		
1971	-0.1		
1972	6.1	*	*
1973	-1.3		

Total (TOI)		3	5

Table 17: DETECTION OF OUTLIERS IN RICHLAND U-233 MATERIAL BALANCES

* = Detection of outliers
total number of data points = 29
min. value = -169.2, max. value = 62.9, target value = process mean
Weighted Process Mean = -35.495
Standard Deviation = 36.620

Period	MB	Influence Function Matrix ref. 50	Neural Network 1.5 SD
1958	-86.4		
1959	-90.1		
1960	-143.7	*	*
1961	-169.2	*	*
1962	-106.8		*
1963	-66.9		
1964	-94.9	*	
1965	-118.8	*	*
1966	17.2		
1967	-1.5		
1968	-32.0		
1969	62.9		*
1970	-9.0		
Total (TOI)		4	5

Table 18: DETECTION OF OUTLIERS IN ORNL U-233 MATERIAL BALANCES

* = detection of outlier
total number of data points = 26
min. value = -3.3, max. value = 0.3, target value = 0
Weighted Process Mean = -0.221
Standard Deviation = 0.297

Period	MB	Influence Function Matrix ref. 50	Neural Network 1.5 SD
1957	0.1		
1958	-0.8		*
1959	-1.4	*	*
1960	-0.05		
1963	-0.1		
1964	-0.9		*
1965	-0.6		*
1966	0.3		
1976	-0.05		
1977	-3.3	*	*

Total (TOI) 2 5

Table 19: DETECTION OF OUTLIERS IN SAVANNAH PU-238 MATERIAL BALANCES

* detection of outlier
total number of data points = 19
min. value = -19.6, max. value = 2.1, target value = 0
Weighted Process Mean = -0.576
Standard Deviation = 2.076

Period	MB	Influence Function Matrix ref. 50	Neural Network 1.5 SD
1969	-1.7		
1970	-19.6	*	*
1971	-2.9		

Total (TOI)		1	1

CHAPTER 5: RESEARCH RESULTS FROM ANALYZING THE QUALITY MONITORING OF PRODUCTION PROCESSES

Chapter Overview

In production processes, the monitoring of product quality is critical especially in light of today's business climate. Complexities in dealing with product quality can arise because of observation interdependency and the influence of process disturbances, i.e. outliers. Unfortunately, the independent and identically distributed assumption of the traditional control chart isn't always applicable. Thus, in this chapter various methods are tested on autocorrelated data sets from actual production processes with greater emphasis placed on the neural network method. The neural network technique is enhanced with the use of a simulation computer program for creating training data sets. This simulation approach provides the opportunity of including outliers of various types in a data set for the training of the neural network. The research concluded that the neural network method along with the simulation approach provides the production process control industry with certain advantages, such as permitting dependency among observations and determining the effects of outliers, in dealing with the quality of a production process. Based on results reported herein, this research found that a neural network approach to statistical process control is both viable and often an improvement over currently used methods.

Production Process Control and Related Research Methods

For use in production processes, the monitoring methods of product quality must be able to detect process disturbances (outliers) in a timely manner. The most often used monitoring method is the traditional statistical process control (SPC) chart. However, the assumptions for SPC charts are not always appropriate for many real industrial production processes (Booth, 1984 and Alwan et. al., 1988). Thus, the control chart method can be unreliable (Prasad et. al., 1995).

A key area in quality control monitoring that has been investigated extensively for various production processes is the detecting of significant shifts in the process mean. These shifts in the process mean are often adverse process changes. Currently, the statistical methods being used for monitoring these occurrences include Shewhart Control Charts and CUSUM charts. In the language of these charts, points that are outside three standard deviations of the process mean are usually considered to be process disturbances (outliers). When this situation exists, the process is said to be "out-of-control" and assignable causes of the difficulty are sought. Otherwise, we say the process is in-control, i.e. operating within its control limits. However, as previously mentioned in chapter one, these techniques were also found to have problems.

This research examined various techniques that are more robust, (i.e. outlier resistant) for use in the detection of outliers in production processes. This robustness allows the methods to be used to detect outlying observations, which is essentially what the process control analyst wants to do. These techniques have been compared based on data sets with and without assignable causes for process disturbances. The main emphasis of this research is on the neural network technique and its capabilities. Overall, the results from this analysis show the ability of these methods and especially neural networks to not only detect outliers in a timely manner but also handle terminal points.

Booth (1984 and 1986) detailed the relationship between time series outlier types and production process changes. In that work it was pointed out that an additive outlier (AO) indicated a one-time process disturbance (e.g. a one-time bad production run). Further, Booth pointed out that an innovative outlier (IO) indicates a continuing production problem (e.g. machine wear). In particular, Booth et. al. (1990) demonstrated that knowing the outlier type aids in determining the assignable causes for production disturbances. Acar and Booth (1987) discussed various methods for implementation in real-time SPC.

Recently, Sebastian et. al. (1994) and Prasad et. al. (1995) tested their respective methods on the same production process data sets and achieved satisfactory results. In particular, Prasad et. al. (1995) showed that the typing of outliers directly aids in determining assignable causes. In this research, the neural network algorithm with the simulation approach enhancement was also applied to the same production process data sets and the following discusses the results.

Illustrations of Testing and Comparing Various Methods on SPC Applications

Process of Dyeing Woolen Yarn

The data set was taken from Grant and Leavenworth (1980, p. 94). They described the production process in the following manner. "In the dyeing of woolen yarns, it is desirable to control the acidity of dye liquor. Unless the dye liquor is sufficiently acid, the penetration of color is unsatisfactory; on the other hand, too - acid liquor affects durability of the production made yarn. Acidity is conveniently measured as pH. A low pH corresponds to high acidity, and vice versa. In any dyeing operation there is a band of pH values within which the best results as to both color penetration and durability are obtained.

The acidity of the dyeing solution depends not only on the constituents put into the dye liquor but also on the characteristics of the wool being dyed. From time to time it is necessary to use wools from sources that have different characteristics. Although blends of wools from various sources are made, successive blends will differ somewhat from one another."

In this process, it is desirable to control the acidity of the dye liquor. The sample data was collected from each of five Hussong kettles by an inspector twice a day for a period of approximately five weeks. The X-Bar and the R value were calculated for each sample. An X-Bar Control Chart was plotted having a center line of 4.22 and control limits at 4.05 and 4.39. See Table 19 for a view of the sequence of observations discussed here.

As for the assignable causes, the first process disturbance began on February 1 (OBS. no. 4) when a new blend of entirely different wools was introduced to the process and immediately the acidity dropped. Then on February 5 (OBS. no. 9) after the old surplus stock had been used the acidity fell below the control limit and continued out of control until corrective measures were taken on February 8 (OBS. no. 15). Afterwards, the process was under satisfactory control until February 21(OBS.

no. 35) when it had a brief departure. This brief departure was caused by two batches of improperly neutralized carbonized stock.

Analysis. The neural network approach detected the very first observations (no. 3 & 4) in which a new blend of entirely different wools was introduced into the process (See Table 20). Then at observation (no. 9), all methods detected the process disturbance due to the old surplus stock being used and affecting the acidity. The polynomial smoothing method did miss observation (no. 10) while the neural network detected point (no. 32) as being a process disturbance point without a known cause. The joint estimation method only detected three process disturbances but they were the leading point or return point of a process disturbance. Thus, the Joint Estimation procedure indicates process changes and enhances the ability to determine process disturbance types. However, the results from these methods are not entirely comparable due to their different number of sensitivity constants. Even with this difference, the neural network approach does provide as good, and even in some cases, a better analysis of the data set.

The Diameters of Injector Pump Bore Holes of Automobile Diesel Engine Blocks

This data set is taken from a production process in which pump bore holes are cut into automobile diesel engine blocks (Quesenberry, 1990). The data consists of deviations from a target value (measured in 10^{-5} inches) and collected from a run of forty observations after a restart of the process following a shut down period. Thus, the author felt that a method which provided an estimate for the process mean and standard deviation was required.

The production process is a metal machinery process and is prone to have process disturbances (outliers). These process disturbance points can be caused by metal chips from the machining process which fall against the cutting blade and cause problems before they burn or fall off. If these process disturbances are used in computing automatic adjustments for the process, there is a potential for serious over adjustment to occur.

The original author used the sequential uniform residuals (SUR) approach to screen for outliers in the data set. The approach was able to identify three outliers at observation no.'s 18,19, and 32. These observations had unknown assignable causes even though it was indicated that metal chips could be the cause. The results of this approach are compared against results from other approaches and the comparison is presented in Table 21.

In examining Table 21, the presentation of the actual number of outliers and the number of false alarms is absent because the assignable causes are unknown. Even with this condition, the results do provide some insight into the capabilities of the methods. In this case, the Polynomial Smoothing approach was able to detect four outliers while the Data Bounding and the Joint Estimation approaches each detected five outliers. The neural network approach detected the most outliers, six. However, the most significant point to make deals with the last observation. It can be seen that the neural network approach was the only one that was able to detect the last point as an outlier. Prasad et. al. (1995) indicated that the last outlier falls at the end of the series and the identification of the type of such an outlier could be questionable. Sebastian (1994) indicated that the Polynomial Smoothing approach needed a special technique to detect the last point while Data Bounding relied on the + or - 3(MSD) level to detect the last point. Therefore, the researcher feels sure that the neural network approach does offer a potentially powerful tool to the process control industry.

The Sheet-Like Process

This data set is taken from a continuous sheet-like process in which "it is desired to control the weight at 1.25 per unit area" (Johnson et. al., 1974). The data consists of deviations from the target value and were identified as an AR(1) model with $\Phi = 0.65$.. The estimated standard deviation was equal to 0.043. The process deviations were not given assignable causes by the original authors. They did say that "this example indicates that in a control scheme it is important to consider the correlation structure when designing the scheme."

In examining Table 22, the presentation of the actual number of outliers and the number of false alarms is absent because the assignable causes are unknown. It can be seen that the neural network approach did detect twice as many outliers as the Joint Estimation approach and this difference could be attributed to different sensitivity constants. Here again, the neural network approach does provide a good alternative to other methods for the process control industry.

Detection of Outliers in a Transmission Parts Manufacturing Process

This data set is composed of diameter measurements from forty-five consecutively produced automatic transmission parts (Quesenberry,1990). In this process, tool wear is often the cause for producing outlying measurements. The transmission parts were made immediately following an adjustment to the tool. In this case, the out-of-control points were not assigned a cause. Therefore, the actual number of outliers and the number of false alarms are not recorded.

The original author used the sequential uniform residual (SUR) approach to screen for outliers in the data set. The approach was able to identify outlying measurements (outliers) at observations no.'s 19 and 43. The results of this approach are compared to other methods in Table 23.

In reviewing the results of Table 23, it can be seen that both the neural network approach and the joint estimation approach detected outlying measurements (outliers) at observation no.'s 19 and 43. These two observations were also detected by the SUR approach. The point to be made is that the neural network approach detected observation no. 2 which had a value approximately equal to observation no. 43. Therefore, the neural network approach may be able to give an early signal of a deterioration in the process.

Discussion

Some of the methods, for example, data bounding, polynomial smoothing or the sequential uniform residual approach, discussed in this chapter, have a weakness in not being able to directly identify outlier types. On the other hand the joint estimation procedure, utilizing complex statistical mathematics, and mathematically modeling of the process, can identify the outlier type. However, the researcher believes that the neural network method could be trained to identify the outlier type by the use of a simulated training data set incorporating the mathematical formulas (Chen et. al., 1992 and Thome, 1995) for each outlier type. Such possibilities are currently being investigated at Kent State University.

Neural Networks, Polynomial Smoothing, Data Bounding, and Joint Estimation are capable of handling end points (i.e. the latest piece of information available from the production process) and thus can be applied to real-time on-line process control. An example of this capability is demonstrated in Table 21, where the neural network approach was able to detect the last point in the data set as being a process disturbance point (outlier). Thus, the results do show, especially for the neural network approach, that these methods can be used to control real-time on-line processes.

Summary

The methods presented in this chapter are effective across diverse applications. As we have seen the neural network approach is particularly useful in the case of autocorrelated data (Chemick et.al., 1982, Sebastian, 1994, and Prasad, 1995). The term autocorrelation is discussed in chapter two. Neural networks have certain advantages that the other techniques do not. First, the network learns about the current process on its own, as well as making adjustments in its learning to account for process changes. Second, an overall theory is not important for the neural network to function. Third, the type of product has no impact on the neural network performance because the network is directly controlling the process, and thereby indirectly maintaining the product quality. Fourth, the network lessens the possibility of human error because there is no necessary intervention or interaction by a person. Fifth, the network anticipates the abnormalities in order to correct for them within the process because it has the ability to learn and adjust to process changes without outside interference assuming the establishment of a feedback loop. Sixth, it removes the need for trial and error in establishing the optimum control for the process because it learns the information from the process itself (Blaesi et. al., 1992).

In summary, the neural network approach with the simulation computer program was shown to be superior or at least equivalent to other methods in the detection of outliers and reducing the probability of false alarms. Consequently, this research will be able to provide the process control industry with a tool to enhance the reduction of quality problems and erroneous out-of-control situations. It can not be overly stressed that the most important feature of the neural network approach along with the simulation enhancement is the ability to identify outliers often well within the usual three standard deviations from a target value. In other words, this method shows a change in the production process without having to go to the usual three standard deviation limit. Thus they are a favorable addition to those methods already available to the process control industry.

Table 20: Detection of Outliers among Sample Averages of pH of Dye Liquor from Five Hussong Kettles

* = detection of outlier
*+ = detection of outlier beyond TV± for data bounding or polynomial smoothing
total number points of data points = 46
min. value = 3.72, max. value 4.63, target value = 4.22

Data bounding and Polynomial Smoothing:
modified process average = 4.253
modified standard deviation = + / - 0.099
with % drop = 20%
median = 4.255

Neural Networks:
Weighted Process Mean= 4.268
Standard Deviation+/- 0.156
Joint Estimation:
Standard Deviation = + / - 0.168

Period	X-Bar pH	Grant Leavenworth pages 94&95	DB 1.1SD ref. 206	Poly Smooth 1.1 SD ref. 206	Joint Estim. 2.5SD ref. 206	Neural Network 1.1SD	Assigned Cause
1	4.17						
2	4.15						
3	4.08					*	Different Wools
4	4.07					*	Different Wools
5	4.13						
6	4.22						
7	4.33						
8	4.33						
9	4.54	*	*	*	*	*	Old Surplus Stock Used Up
10	4.50	*	*			*	" "
11	4.54	*	*	*+		*	" "
12	4.61	*	*	*+		*	" "
13	4.63	*	*	*+		*	" "
14	4.61	*	*	*+		*	" "
15	4.37						corrective
16	4.54	*	*	*+	*	*	measures
17	4.29						taken
18	4.35						
19	4.35						
20	4.32						
21	4.36						

Table 20: (continued)

32	4.07					*	none/unknown
33	4.22						
34	4.11						
35	3.72	*	*	*+	*	*	improper neutralization
36	4.18						
37	4.29						
38	4.17						
39	4.14						

Total (TOI)	8	8	7	3	11
Actual (NA+I)	7	7	6	2	7
False (NFA)	0	0	0	0	1
Missed (N+OM)	0	0	1	5	0

Table 21: DETECTION OF OUTLIERS FOR BORE HOLE PRODUCTION PROCESS

* = detection of outliers
*+ = detection of outlier beyond TV +/-3(MSD) for Data Bounding or Polynomial Smoothing
total number of data points = 40
min. value = -87, max. value = 184, target value = 0

Data Bounding & Polynomial Smoothing
modified process average = -17.75
modified Standard Deviation = +/-15.65
with % drop = 10%(5)
median = -19.5

Neural Network
Weighted Process Mean = -14.868
S.D. = +/- 17.050

R = outlier detected with the aid of the Linear Regression End Point Manipulation Procedure

Period	OBS	SUR ref. 205	DB 1.5SD ref. 205	Poly Sm 1.5SD ref. 205	Jt Est 3SD ref. 205	Neural Network 2.5SD	Assigned Cause
1	-37		*				unknown
18	144	*	*+	*+	*	*	"
19	184	*	*+	*+	*	*	"
20	30						
28	-44					*	"
31	-2						
32	-87	*	*+	*+	*	*	"
33	-26				*		"
34	-21						
35	-28						
36	-45					*	"
40	-62		*+	*+R	*?	*	"
Total (TOI)		3	5	4	5	6	

Table 21: DETECTION OF PROCESS DEVIATIONS FOR A CONTINUOUS SHEET-LIKE PROCESS

* = detection of outlier
min. value = -0.16, max. value = 0.10, target value = 0

Neural Network
Process Weighted Mean = 0.009
Standard Deviation = + / - 0.030

Period	Weight/ Unit Area	Joint Est. 3SD (ref. 191)	Neural Network 2.55SD	Assigned Cause
1	0.06			
2	0.09		*	unknown
3	0.07			
6	-0.06			
7	-0.08		*	"
8	-0.02			
28	-0.02			
29	-0.13	*	*	"
30	-0.10		*	"
31	-0.04			
63	0.03			
64	-0.02			
65	0.07	*		"
66	0.01			
67	0.03			
68	0.03			
69	0.10	*	*	"
70	0.08		*	"
71	0.06			
82	-0.06			
83	-0.16	*	*	"
84	-0.14		*	"
85	-0.03			

Total (TOI) 4 8

Table 23: DETECTION OF OUTLYING MEASUREMENTS FOR TRANSMISSION PARTS PROCESS

* = detection of outliers
min. value = -0.05, max. value = 0.010, target value = 0

Neural Network
Process Weighted Mean = 0.000
Standard Deviation = + / - 0.003

Period	OBS	SUR ref. 191	Joint Est. 3SD ref. 191	Neural Net.3SD	Assigned Cause
1	0.004				
2	0.009			*	unknown
3	0.005				
4	0.001				
17	0.0				
18	-0.002				
19	-0.015	*	*	*	unknown
20	0.0				
21	-0.002				
41	-0.003				
42	-0.003				
43	0.010	*	*	*	unknown
44	0.0				
45	0.002				
Total (TOI)		2	2	3	

CHAPTER 6: DISCUSSION AND CONCLUSIONS

Chapter Overview

This chapter discusses the results, the research findings, and the importance of the study. Major findings of the study and their strategic implication are presented first. Next, the future research opportunities are discussed. Finally, a summary of the dissertation is presented.

Importance of the Research

Based on the results from this study, it has been demonstrated that the neural network algorithm along with the simulation computer program will provide the process control industry with important tools for controlling a real-time on-line production process. As stated in chapter one, the significance of this research is its ability to take the process control industry "one step further into the future." The current literature supports this idea and indicates that there is both a trend and an increased interest in applying the neural network approach along with any enhancement (i.e. simulation approach to create training data sets) to improve production process control. The reason for this support is that the process control industry is searching for an approach, which will be highly reliable, predictable, accurate, fault resistant, and responsive, to provide the next generation of flexible electronic equipment. As a result, this research has the potential of providing the tool that the process control industry needs for the future. Thus, the goal and objectives of this research, as stated in chapter one, have and should continue to support this direction.

As for chapter two, it provided an in-depth and thorough analysis of the current and past literature on topics such as time series, neural networks, and production process control. The review investigated various aspects of each topic which included definitions, different approaches, and applications. In examining the topic of time series, the most important issue discussed was outliers and their massive impact on identification, estimation, and forecasting. However, in this research the main emphasis was placed on neural networks. Here, a comprehensive history of neural networks was presented. As a result of the current interest in the neural network approach, a multitude of successful applications have emerged. These applications include classification, robotics, signal processing, business, and process control. Table 5 provides a summary of neural network applications. In the final section of this chapter, the process control topic was discussed. The current methods of control available to the process control industry were reviewed. Next, various applications were examined with special emphasis placed on the nuclear industry. This industry has been placed in the forefront because of recent articles in both Time and Newsweek. These articles basically state that a potential international crises exists concerning the monitoring and safeguarding of nuclear material in relationship to the break-up of the Soviet Union. Therefore, the nuclear industry is seeking a tool to alleviate this situation and assist in addressing the relevant issues (See Table 6) of the nuclear industry.

Next, in chapter three, the major aspects of the research methodology were articulated. In the first section, the data sets used for the study were discussed. The data sets were taken from the nuclear industry and various other production processes. When the data sets were applied, the researcher noticed that some of the data sets had assignable causes for process disturbances while others didn't. Further, the amount of data points and the process for data collection were different among the data sets. Thus, these data sets from actual processes provided significant testing for various methods' ability to detect outliers. In the second section, the neural network algorithm chosen for this study was reviewed. The algorithm used was developed by Denton (1993). Denton's computer algorithm reads in data and calculates a set of network weights using standard back propagation training or nonlinear optimization algorithms and line searches in conjunction with backward propagation. The program does require certain parameters to be established in a definition file. These parameters include the number of input nodes, output nodes, and hidden nodes, training and testing tolerances, the weight seed value, etc. The weight seed value is critical in successfully training the neural network and for the repeatability of the neural network to analyze correctly data sets from the same process after it has been trained. In the third section, the simulation computer program used in this study was described. This simulation approach provided the means of creating a training data set that had the flexibility for the detecting of outliers, the future typing of outliers, and the handling of terminal points. Furthermore, the amount of actual data needed from a real-time process is reduced. Consequently, this simulation approach provides researchers using neural network algorithms a very powerful tool to enhance the potential for success. In the fourth section, the neural network architecture used in this study was described. The network consists of two input nodes, one output node, and five hidden nodes. The two input nodes allow for the time series affect to be taken into account by having the data presented to node two the same data as node one receives but shifted one time period. The number of hidden nodes is based on a formula of 2n+1 (Patuwo et. al., 1993) and provides for network sensitivity. The final section of chapter three discussed the criteria used to compare various methods and especially the neural network method for detecting process disturbances (outliers) for various production and nuclear process data sets. The most crucial analysis in this study was based on two points. The first point was the correct detection of outliers as early as possible in order to reduce false alarms. The second point was the detection of an outlier when the outlier was a final point which would indicate the neural network's capability of controlling a real-time on-line production process.

In chapter four, the study demonstrated that various methods, in particular neural networks, could detect nuclear material diversions or losses more rapidly and accurately than currently used methods. Also, these various methods, especially neural networks, were described. The neural network algorithm used for this study was developed by Denton (1993). The algorithm formulates a set of network weights using standard back propagation training or nonlinear optimization algorithm and line searches in conjunction with backward error propagation. The algorithm was enhanced by the use of the simulation approach to create a training data set for a neural network. The simulation approach allowed for reducing the amount of actual data required, reducing the uncertainty of the training data set not having outliers, incorporating outlier detection, and providing for the future typing of outliers.

In performing the analysis on the two selected cases (AGNS Barnwell Nuclear Fuels Plant, 1984, and Chernick et. al., 1982, data sets), all the methods investigated had problems detecting the outliers because of the very small amounts of nuclear material being removed. Also, the precipitation of solids out of the solution provided a factor that affected the results for outlier detection and possibly created a number of false alarms. Thus, all the various methods had difficulty in detecting all of the known outliers because of the different sensitivity constants that each method uses. However, the key issue was to demonstrate the capability to control a real-time on-line production process and the neural network approach was able to show this capability of detecting a final point of a data set. Tables 13, 14, and 17 provide the supporting information. Therefore, these various methods, especially the neural network approach with the simulation enhancement, do provide the process control industry with valuable tools for dealing with nuclear material safeguards.

Finally, in chapter five, the study examined the monitoring of product quality for various production processes. This study analyzed four data sets using various methods with special emphasis again on the neural network approach and the simulation computer algorithm to detect process disturbances (outliers). Again, the neural network algorithm used for this analysis was Denton's algorithm and the simulation computer program was also the same one previously described. Once more, the results are not entirely comparable because of their different number of sensitivity constants.

The resulting analysis of the data sets showed that the neural network approach with the simulation enhancement was demonstrated to be superior or at least equivalent to the other methods in the detection of outliers and reducing the probability of false alarms. Further, the results in Table 20 established that the neural network approach with the simulation enhancement has the capability to control a real-time on-line production process. Therefore, this study was able to provide the process control industry with a tool to enhance the reduction of quality problems and erroneous out-of-control situations.

Future Research Applications

This research determined that neural networks can successfully identify process disturbances (outliers) early enough in various applications to possibly correct a situation before it becomes a problem. Therefore, to finalize the objectives of this dissertation, future research efforts will be required to be investigated in the following areas.

First, a manufacturing company needs to be found that would be willing to share information about each of its processes. This information would include the number of input variables for each process (i.e. temperatures, pressures, line speeds, etc..), the number of output variables for each process (i.e. quality parameters, number pieces produced, etc..), the problem data for each process, and the assigned causes for process disturbances. Also, the researcher would need to be able to observe the processes operating in order to develop a total understanding of each of the processes. The researcher would then need to convince the company to actually apply the neural network approach with the simulation enhancement in order to actually control one the processes in a real-time on-line experiment. This element would require presenting a case that would contain: the strategic benefits to the company, the results from phase four, a cost/benefit analysis to demonstrate possibly savings, and the improvements to product quality and overall production. Thus, the concept for the research goal would be implemented to control a real-time on-line production process where the results could be verified by a comparison against the instrumentation taking continuous measurements.

Another future research direction is the incorporation of outlier typing into the simulation computer program. By accurately identifying the outlier types immediately after the first outlier is detected, this identification would make the neural network approach more effective in reducing false alarms, preventing process disturbances to occur, and improving product quality. To accomplish this outlier typing, the researcher would need to incorporate formulas for the different types of outliers into the simulation computer algorithm. The formulas could be similar to the ones in the papers by Chen et. al. (1992) and Thome (1995). As a result, the researcher would be able to be assured that he/she could detect and type any process problem.

Next, the simulation computer program could be enhanced through future research by determining the most appropriate functional curve to be used. The functional curve used in this research was the trigonometric sine function. By properly matching the functional curve to the process, the researcher possibly could detect more accurately the process disturbances (outliers) and better assign type designations to each outlier detected. With this capability, the process would then incur less false alarms and more early indications of process problems.

Further testing of neural network algorithms with the simulation enhancement needs to be done to ensure that this concept will have broad use and acceptance. To accomplish this goal, two more data sets could be candidates (automobile emissions data and a computer transaction disk retrieval data set) for testing the concept's capabilities. These data sets provide for evaluating this research concept on processes outside the production process industry.

Finally, future research is required in determining whether the concept of the primary and secondary control schemes could be used as a means of outlier detection to minimize both the probability of a Type I error and/or a Type II error from occurring. Integrating the neural network approach with another method, such as fuzzy logic, joint estimation, or an expert system, could provide the process control industry with an even more powerful set of tools, while providing management with more information to prevent production disruptions. Finally, this concept would also provide for a safety back-up in case the primary or the secondary control method were to fail.

Summary

The neural network approach with the simulation enhancement has been demonstrated to be effective, flexible, and adaptive to a wide variety of process data sets. This concept has proven to be a useful tool for: the detection of process disturbance points (outliers), providing an earlier warning system of possible process disturbances, and reducing false alarms than some of the other methods discussed. Further this concept has been able to detect outliers that occur as end points in production process data and thus demonstrates its capability to control a real-time on-line production process.

The neural network approach does have some unique advantages which enhanced the success of this research. These advantages include adaptive learning, real-time self-organization, and solving difficult problems in less time. Thus, the neural network approach is able to respond to process changes faster and obtain more complete solutions to process problems.

The most important element of this research is the various applications to which the neural network algorithm with the simulation computer program enhancement have been applied. The second element is that the simulation computer program ensures that outliers are incorporated into the training data set. The third element is that the simulation computer program provides for the incorporation of formulas for a particular outlier type. Finally, a researcher is provided with maximum flexibility in the development of a training data set for any neural network algorithm. Therefore, future research efforts with other methods requiring a training data set should be enhanced by using this simulation computer program.

Based on the results of this research and recent successes of other researchers in the neural network area, the potential for the implementation of the neural network approach with the simulation enhancement looks very promising for a variety of real-time on-line applications. Therefore, this concept provides the control industry and operating managers involved in production process control and product quality a tool to confront the global competition in today's business climate.

BIBLIOGRAPHY

Abe, Shigeo, "Learning by parallel forward propagation," International Joint Conference on Neural Networks, San Diego, California, Vol. 2, IEEE Neural Networks Council, Michigan, (June 17-21, 1990), pp. 99-104.

Acar, William and David E. Booth, "Easier quality control: Combining problem classification with time series analysis," Production and Inventory Management, Vol. 28, No. 3, 1987, pp. 53-58.

Ahmad, Subutai and Gerald Tesauro, "Scaling and generalization in neural networks: A case study," Neural Networks, Vol. 1, Supplement 1, 1988, pp. 3-9.

Allen, Robert B., "Several studies on natural language and back-propagation," International Joint Conference on Neural Networks, Seattle, Vol. 1, IEEE Service Center, New Jersey, July 8-12, 1991, pp. 335-340.

Allen, T. J., "Project management risk, strategy and structure in the nuclear industry: A historical prospective," Project Management Journal, Vol. 20, No. 4, (December 1989), pp. 21-30.

Alwan, L. C. and H. V. Roberts, "Time Series Modeling for Statistical Process Control," Journal of Business Economic Statistics, Vol. 6, No. 1, 1988, pp.87-95.

Andersen, Allan and Andrew Weiss, "Forecasting: The Box-Jenkins approach," in The forecasting accuracy of major time series methods, , edited by S. Makridakis, A. Anderson et. al., John Wiley and Sons, New York, 1984, pp. 167-200.

Anderson, Dana Z., Neural information processing systems, American Institute of Physics, New York, 1988.

Anderson, R. L., Practical statistics for analytical chemists, Van Nostrand Reinhold, New York, 1987.

Andes, David, Bernard Widrow, Michael Lehr, and Eric Wan, "MRIII: A robust algorithm for training analog neural networks," International Joint Conference on Neural Networks, Washington D.C., Vol. 1, Lawrence Erlbaum Associates, New Jersey, (January 15-19, 1990), pp. 533-536.

Avenhaus, Rudolf, Material accountability: Theory, verification, and application, John Wiley and Sons, New York, 1977.

Babb, Michael, "Neural nets: No hype, please," Control Engineering, (June 1993), p. 49.

Bailey, A. W., "Automatic evolution of neural net architectures," International Joint Conference on Neural Networks, Washington D.C., Vol. 1, Lawrence Erlbaum Associates, New Jersey, (January 15-19, 1990), pp. 589-592.

Baillie, R. T. and T. Bollerslev, "Predictions in dynamic models with time-dependent conditional variances," Journal of Econometerics, Vol. 52, No.1/2, (1992), pp. 91-113.

Barnett, Vic and Toby Lewis, Outliers in Statistical Data, Second Edition, John Wiley and Sons, New York, 1978.

Bartos, Frank J., "Fuzzy logic widens its appeal to industrial controls," Control Engineering, (June 1993), pp. 64-67.

Beck, H., D. McDonald, and D. Brzakovic, "A self-training visual inspection system with a neural network classifier," International Joint Conference on Neural Networks, Washington, D.C., Vol. 1, 1989, pp. 307-311.

Beckman, R. J, and R. D. Cook, "Outliers," Technometrics, Vol. 25, No. 2, 1983, pp. 119-149.

Belfore, L. A., II, B.W. Johnson, and J. H. Aylor, "Modeling of fault-tolerance in neural networks," IJCNN, Vol. 1, (January 15-19, 1990), pp. 325-328.

vom Berg, Haubold, "Fuzzy logic: A clear choice for temperature control," I&CS, Vol. 66, No. 6, (June 1993), pp. 39-41.

Blaesi, J. and B. Jensen, "Can neural networks compete with process calculations," Intech, Vol. 39, No. 12, (December 1992), pp. 34-37.

de Bollivier, M., P. Gallinari, and S. Thiria, "Cooperation of neural nets for robust classification," International Joint Conference on Neural Networks, San Diego, California, Vol. II, (June 17-21, 1990), pp. 113-120.

Booth, D. E., Some applications of robust statistical methods to analytical chemistry, Unpublished Ph.D. Dissertation, The University of North Carolina at Chapel Hill, 1984.

Booth, D. E., "A method for the early identification of loss from a nuclear materials inventory," The Journal of Chemical Information and Computer Sciences, Vol. 26, 1986, pp. 23-28.

Booth, D. E.," Regression methods and problem banks", Tools for Teaching, Module 626, COMAP, INC., 1986.

Booth, D. E., William Acar, Thomas Isenhour, and Sharif Ahkam, "Robust time series models and statistical process control," Journal of the Industrial Mathematics Society, Vol. 40, Part 1, 1990, pp. 73-97.

Borson, D., "A different kind of nuclear waste," Business and Society Review, (N.D.), pp. 26-29.

Bowen,W. M. and C. A. Bennett, Statistical methods for nuclear material management, U.S. Government Printing Office, Washington D.C., 1988.

Bowerman, B. L. and R. T. O'Connell, Time Series and Forecasting, Duxbury Press, North Scituate, Massachusetts, 1979.

Box, George E. P. and Gwilym M. Jenkins, Time series analysis forecasting and control, Revised edition, Holden-Day, San Francisco, 1976.

Bretz, E. A., "Can standardization revive the nuclear-power industry?," Electrical World, Vol. 205, No. 1, (January 1991), pp. 54 & 55.

Carpenter, Gail A., Stephen Grossberg, and John H. Reynolds, "ARTMAP: Supervised real-time learning and classification of nonstationary data by a self-organizing neural network, Neural Network, Vol. 4, 1991, pp. 565-588.

Caudill, M., "Neural networks primer: Part I," AI Expert Neural Network Primer, NeuralWare, Inc., Pittsburgh, PA, (December 1987), pp. 46-52.

Caudill, M., "Neural networks primer: Part II," AI Expert Neural Network Primer, NeuralWare, Inc., Pittsburgh, PA, (February 1988), pp. 55-61.

Caudill, M., "Neural networks primer: Part III," AI Expert Neural Network Primer, NeuralWare, Inc., Pittsburgh, PA, (June 1988), pp, 53-59.

Caudill, M., "Neural networks primer: Part IV," AI Expert Neural Network Primer, NeuralWare, Inc., Pittsburgh, PA, (August 1988), pp. 61-67.

Caudill, M., "Neural networks primer: Part V," AI Expert Neural Network Primer, NeuralWare, Inc., Pittsburgh, PA, (November 1988), pp. 57-64.

Caudill, M., "Neural networks primer: Part VI," AI Expert Neural Network Primer, NeuralWare, Inc., Pittsburgh, PA, (February 1989), pp. 61-67.

Caudill, M., "Neural networks primer: Part VII," AI Expert Neural Network Primer, NeuralWare, Inc., Pittsburgh, PA, (1990), pp. 51-57.

Caudill, M., "Neural networks primer: Part VIII, AI Expert Neural Network Primer, NeuralWare, Inc., Pittsburgh, PA, (1990), pp. 58-64.

Chapman, D., "The eternity problem: Nuclear power waste storage," Contemporary Policy Issues, Vol. 8, (July 1990), pp. 80-93.

Chang, C., G. C. Tiao, and C. Chem, "Estimation of time series parameters in the presence of outliers," Technometrics, Vol. 30, No. 2, 1988, pp. 193-204.

Chatfield, Chris, "The future of time-series forecasting," International Journal of Forecasting, Vol. 4, 1988, pp. 411-419.

Chen, C. and G. C. Tiao, "Random level shift time series models, ARIMA approximation, and level shift detection," Journal of Business and Economic Statistics, Vol. 8, 1990, pp. 170-186.

Chen,Chung, Lon-Mu Liu, and Gregory Hudak, "Modeling and forecasting time series in the presence of outliers and missing data," Scientific Computing Assoaciates, Working Paper No. 128, February, 1992.

Chen, Chung and Lon- Mu Liu, "Joint estimation of model parameters and outlier effects in time series," Journal of the American Statistical Association, Vol. 88, 1993, pp. 284-298.

Chen, David S. and Ramesh C. Jain, A robust backpropagation learning algorithm for function approximation, pp. 1-12. Unpublished Manuscript, University of Michigan AI Lab.

Chen, Mu-Song and Michael T. Manry, "Back-propagation representation theorem using power series," San Diego, California, Vol. 1, IEEE Neural Networks Council, Michigan, (June 17-21, 1990), pp. 643-648.

Cheng, Bing, and D. M. Titterington, "Neural networks: A review from a statistical perspective," Statistical Science, Vol. 9, No. 1, (February, 1994), pp. 2-54.

Cherkassky, Vladimir and Nikolaos Vassilas, "Performance of back propagation networks," International Joint Conference on Neural Networks, San Diego, California, Vol. 1, IEEE Neural Networks Council, Michigan, (June 21-24, 1987), pp. 77-84.

Chernick, Michael R., Darryl J Downing, and David H. Pike, "Detecting outliers in time series data," Journal of the American Statistical Association, Vol. 77, No. 380, (December 1982), pp. 743-747.

Chester, Daniel, "New trends in neural networks," Scientific Computing and Automation, Vol. 8, No. 6, (May 1992), pp. 43-48.

Cheung, Raymond K. M., Irving Lustig, and Alain L. Kornhauser, "Relative effectiveness of training set patterns for back propagation, International Joint Conference on Neural Networks, Washington, D.C., Vol. 2, 1989, pp. 673-678.

Chin, David, "A more efficient global optimization algorithm based on Styblinski and Tang (1990)," submitted as letter to editor in Neural Networks.

Chu, Chao-Hsien and Chang-Chun Tsai, "Manufacturing cell formation in a fuzzy environment," pp. 1432-1434 (submitted for publication).

Co, Tomas B., "Induction of neural networks for parallel binary operations," International Joint Conference on Neural Networks, Washington D.C., Vol. 2, IEEE TAB Neural Network Committee, California, 1989, pp. 43-47.

Coats, Pamela K. and L. Franklin Fant, "A neural network approach to forecasting financial distress," The Journal of Business Forecasting, Vol. 10, No. 4, (Winter 1991-1992), pp. 9-12.

Cobb, D. D., H. A. Dayem, A. L. Baker, J.H. Ellis, M. H. Ehinger, and J. M. Crawford, "Demonstration of near-real-time accounting: The AGNS 1980 miniruns," Nuclear Materials Management, (Spring 1981), pp. 34-43.

Connor, Jerome and Les Atlas, "Recurrent neural networks and time series prediction," International Joint Conference on Neural Networks, Seattle, Vol. 1, IEEE Service Center, New Jersey, (July 8-12, 1991), pp. 301-306.

Culioli, Jean-Christophe and Vladimir Protopopescu, "Neural networks models for linear programming," International Joint Conference on Neural Networks, Washington D.C., Vol. 1, Lawrence Erlbaum Associates, New Jersey, (January 15-19, 1990), pp. 293-296.

Dagli, C. H., S. R. T. Kumara, and Y. C. Shin, (Ed.), Intelligent engineering systems through artificial neural networks, ASME Press, New York, 1991.

Dahl, Edward Denning, "Accelerated learning using the general delta rule," International Joint Conference on Neural Networks, San Diego, California, Vol. 2, IEEE Neural Networks Council, Michigan, (June 21-24, 1987), pp. 523-530.

DataMyte Corporation, DataMyte handbook fourth edition: A practical guide to computer data collection for statistical process control, Minnetonka, Minnesota, 1989.

Davis, Lawrence, "Mapping classifier systems into neural networks," Advances in neural information processing systems 1, Morgan Kaufmann Publishers, Inc., San Mateo, California, 1989, pp. 49-56.

Dayem, H. A.; D. D.Cobb; E. A. Hakkila, and C. A. Ostenak, Demonstrations of near-real-time accounting: the AGNS 1980-1981 miniruns, Los Alamos Laboratory, Los Alamos, 1984.

Dehnad, Khosrow, Quality control, robust design, and the Taguchi Method, Wadsworth & Brooks/Cole, Pacific Grove, CA, 1989.

Denby, Lorraine and R. Douglas Martin, "Robust estimation of the first-order autoregressive parameter, Journal of the American Statistical Association, Vol. 74, No. 365, (March 1979), p. 140.

Denton, James W., A robust and efficient training algorithm for feedforward neural networks, Ph.D. Dissertation, Kent State University, OCLC No. 26045963, (December, 1991).

Denton, James W., Instruction manual optinet neural network training software, University of West Virginia, Morgantown, West Virginia, 1993.

Denton, James W., "How good are neural networks for causal forecasting?," Journal of Business Forecasting Methods and Systems, Vol. 14, No. 2, (Summer,1995), pp. 17-20.

Denton, James W., Ming S. Hung and Barbara A. Osyk, "A neural network approach to the classification problem," Expert Systems with Applications, Vol. 1, 1990, pp. 417-424.

Downing, D. J.; D. H. Pike, and G. W. Morrison, "Analysis of MUF data listing ARIMA models," Nuclear Material Management, Vol. 7, No. 4, 1978, pp. 80-86.

Doyle, John, "Serial, Parallel and neural computers," Futures, Vol. 23, No. 5, (July-August 1991), pp. 577-593.

Dray, Darrell L., "Process control: Back to the basics," Instruments and Controls Society, Vol. 67, No. 1, (January 1994), pp. 67-73.

Flam, F., "Neural nets: A new way to catch elusive particles?," Science, Vol. 256, (May 29, 1992), pp. 1282 & 1283.

Flick, Thomas E., Lee K. Jones, Richard G. Priest, and Charles Herman, "Pattern classification using projection pursuit," Pattern Recognition, Vol. 23, No. 12, 1990, pp. 1367-1376.

Fox, A. J., "Outliers in time series," Journal of the Royal Statistical Society, Series B, Vol. 34, 1972, pp. 350-363.

Francett, B., "Neural nets arrive," Computer Decisions, (January 1989), pp. 58-62.

Fuchs, A. and H. Haken, "Pattern recognition and associative memory as dynamical processes in nonlinear systems," IEEE International Conference on Neural Networks, San Diego, California, (July 24-27, 1988), pp. 217-224.

Fukaya, M., M. Kitagawa, and Y. Okabe, "Two-level neural networks: Learning by interaction with environment," Proceedings of the 1988 Connectionist Models Summer School, Carnegie Mellon University, Morgan Kaufmann Publishers, San Mateo, California, 1989, pp. 531-539.

Fusaro, Dave, "Tomorrow's control technology," Control, Vol. V, No. 9, (September 1992), pp. 24-28.

"Georgia tech engineers embed neural networks in amplifier," Control, Vol. VI, No. 1, (January 1993), p. 12.

Gevins, Alan S., "How the human brain thinks and acts: Results of latest research using neural networks to analyze brain signals," International Joint Conference on Neural Networks, San Diego, California, Vol. 4, IEEE Neural Networks Council, Michigan, (June 21-24, 1987), pp. 41-47.

Goldsman, David and Barry L. Nelson, "Methods for selecting the best system," 1991 Winter Simulation Proceedings, The Society for Computer Simulation, International, San Diego, California, (December 1991), pp. 177-186.

Goldman, Aaron S., Richard R. Picard, and James P. Shipley, "Statistical methods for nuclear materials safeguards: An overview," Technometrics, Vol. 24, No. 4, (November 1982), pp. 267-275.

Grant, Eugene I. and Richard S. Leavenworth, Statistical quality control, Fifth edition, McGraw-Hill Book Company, New York, 1980.

Grossberg, S., "Burning the candle from both ends," IEEE Expert, Vol. 6, No. 1, (February 1991), pp. 3-6.

Guiver, John and C. C. Klimasauskas, "Applying neural networks part IV: Improving performance," PC AI, (July-August 1991), pp. 34-41.

Gupta, L. and A. M. Upadhye, " A neural network approach to the partial shape classification problem," International Joint Conference on Neural Networks, Seattle, Vol. 1, IEEE Service Center, New Jersey, 1991, pp. 215-220.

Hagiwara, Masafumi, "Novel back propagation algorithm for reduction of hidden units and acceleration of convergence using artificial selection," International Joint Conference on Neural Networks, San Diego, California, Vol. 1, IEEE Neural Networks Council, Michigan, (June 17-21, 1990), pp. 625-630.

Hamburg, J. H., The application of neural networks to production process control, Presentation at the 27th Central Reporting Regional Meeting of the American Chemical Society, 1995.

Hammond, R. P., "Nuclear power risks," American Scientist, Vol. 62, (March-April 1974), pp. 155-160.

Handelman, David A., Stephen H. Lane, and Jack J. Gelfand, "Integration of knowledge-based system and neural network techniques for autonomous learning machines," International Joint Conference on Neural Networks, Washington, D.C., Vol. 1, 1989, pp. 683-688.

Hecht-Nielsen, Robert, "Theory of the backpropagation neural network," International Joint Conference on Neural Networks, Washington, D.C., Vol. 1, 1989, pp. 593-605.

Hileman, B., "U.S. and Russia face urgent decisions on weapons plutonium," Chemical and Engineering News, Vol. 72, No. 24, (June 13, 1994), pp. 12-25.

Hillman, David, "Knowledge systems based on cascading neural nets," AI Expert, Vol. 6, No. 12, (December 1991), pp. 46-53.

Hillmer, S. C., W. R. Bell, and G. C. Tiao, "Modeling considerations in the seasonal adjustment of economic time series," Applied Time Series Analysis of Economic Data, Washington, D. C., U. S. Bureau of the Census, 1983, pp. 74-100.

Hirschy, J. R., "Forecasting: A time for the crystal ball," Corporate Controller, Vol. 4, No.2, (November-December 1991), pp. 58-59 & 61.

Hoff, M. E., "Speaker dependent isolated word speech recognition," DARPA Neural Network Study, Final Report, Lincoln Laboratory, Massachusetts Institute of Technology, Lexington, Massachusetts, Part IV, (October, 1987-February, 1988), pp. 55-57.

Hoptroff, R. G., M. J. Bramson, and T. J. Hall, "Forecasting economic turning points with neural nets," International Joint Conference on Neural Networks, Seattle, Vol. 1, IEEE Service Center, New Jersey, (July 8-12, 1991), pp. 347-352.

Hoversten, P. and K. Johnson, "Atoms can enrich lives or end them," USA Today, (December 2, 1992).

Hueter, Geoffrey J., "Neural networks automate inspections," Quality, Vol. 32, (January 1993), pp. 41- 44.

Hughes, Ronald G., "Simulation for design, test and evaluation, and training: Reconciling the differences," 1991 Winter Simulation Conference Proceedings, The Society for Computer Simulation, International, San Diego, California, (December 1991), pp. 231-236.

Hung, Ming S. and James W. Denton, "Training neural networks with the GRG2 nonlinear optimizer," European Journal of Operational Research, Vol. 69, 1993, pp. 83-91.

Huntley, David, "Neural nets: An approach to the forecasting of time series," Social Science Computer Review, Vol. 9, No. 1, (1991), pp. 27-38.

Hwarng, H. Brian and Norma F. Hubele, "Back-propagation pattern recognizers for x-bar control charts," Computers and Industrial Engineering, Vol. 24, 1993, pp. 219-235.

Hwarng, H. Brian and Norma F. Hubele, " X-bar chart pattern recognition using neural nets," ASQC Quality Congress Transactions (Milwaukee: American Society for Quality Control), pp. 884-889.

Hwarng, H. Brain and Norma F. Hubele, "X-bar control chart pattern identification through efficient off-line neural network training," IIE Transactions, Vol. 25, 1993, pp. 27-39.

International Atomic Energy Agency, The Agency's Programme and Budget for 1991 and 1992, Austria, International Atomic Energy Agency, (July 1990).

International Atomic Energy Agency, The Agency's Budget for 1992, Austria, International Atomic Energy Agency, (August 1991).

Johnson, R. and M. Bagshaw, "The effect of series correlation on the performance of CUSUM test," Technometrics, Vol. 16, 1974, pp. 103-113.

Jokinen, Petri A., "Comparison of neural network models for process fault detection and diagnosis problems," International Joint Conference on Neural Networks, Seattle, Vol. 1, IEEE Service Center, New Jersey, (July 8-12, 1991), pp. 239-244.

Jones, Lee K., "Constructive approximation for neural networks by sigmoidal functions," Proceedings of the IEEE, Vol. 78, No. 10, (October 1990), pp. 1586-1589.

Jones, Lee K., "A simple lemma on greedy approximation in hilbert space and convergence rates for projection pursuit regression and neural network training," The Annals of Statistics, Vol. 20, No. 1, 1992, pp. 608-613.

Jones, Lee K., Correction and addition to 'Constructive approximation for neural networks by sigmoidal function,' " Proceedings of the IEEE, Vol. 79, No. 2, (February 1991), p. 243.

Jones, Lee K., "Good weights and hyperbolic kernels for neural networks, projection pursuit, and pattern classification: Fourier strategies for extracting information from high-dimensional data," Proceedings of the IEEE, Vol. 80, No. 7, (July 23, 1992), pp. 1-44.

Jouse, Wayne C. and John G. Williams, "The control of nuclear reactor start-up using drive reinforcement theory," Intelligent engineering systems through artificial neural networks, Proceedings of the Artificial Neural Networks in Engineering (ANNIE '91), (November 10-13, 1991), ASME Press, New York, pp. 537-544.

Kang, Suh, An investigation of the use of feedforward neural networks for forecasting, Kent State University Ph.D. Dissertation, OCLC No. 26088347, (August, 1991).

Kao, Y. and Y. B. Moon, "A unified group technology implementation using the backpropagation learning rule of neural networks," Computers and Industrial Engineering, Vol. 20, No. 4, (1991), pp. 425-437.

Keats, J. Bert and Douglas C. Montgomery, Statistical process control in manufacturing, Marcel Dekker, Inc., New York, 1991.

Khotanzad, Alireza and Mark Fowler, "Neural network based time series forecasting," Intelligent engineering systems through artificial neural networks, Proceedings of the Artificial Neural Networks in Engineering (ANNIE '91), (November 10-13, 1991), ASME Press, New York, pp. 813-818.

Klimasauskas, C .C., "Applying neural networks part 1: An overview of the series" PC AI, (January-February 1991), pp. 30-33.

Klimasauskas, C. C., "Applying neural networks part 2: A walk through the application process," PC AI, (March-April 1991), pp. 27-34.

Klimasauskas, C. C., "Applying neural networks part III: Training a neural network," PC AI, (May-June 1991), pp. 20-24.

Klimasauskas, C. C., "Applying neural networks part V: Integrating a trained network into an application," PC AI, (September-October), pp. 36-41, 60-62.

Klimasauskas, C. C., "Training a neural network," PC AI, (May-June 1991), pp. 20-24.

Klimasauskas, Jane (Ed.), Advanced technology for developers, High-Tech Communications, Sewickley, PA, Vol.1, (May 1992).

Klopf, A. Harry and James S. Morgan, "The role of time in natural intelligence: Implications for neural network and artificial intelligence research," International Joint Conference on Neural Networks, Washington D.C., Vol. 2, IEEE TAB Neural Network Committee, California, 1989, pp. 97-100.

Kohonen, T., "State of the art in neural computing," IEEE International Conference on Neural Networks, Vol. I, (June 21-24, 1987), pp. 79-90.

Kosko, B., Neural Networks and Fuzzy Systems, Prentice-Hall Inc., Englewood Cliffs, New Jersey, 1992.

Koutsougeras, C. and C. A. Papachristou, "Training of a neural network for pattern classification," IEEE International Conference on Neural Networks, San Diego, California, Vol. 1, (July 24-27, 1988), pp. 247-254.

Kuczewski, Robert M., "Neural network approaches to multi-target tracking," IEEE International Conference on Neural Networks, Vol. 4, (June 21-24, 1987), pp. 619-633.

Landenberger, Keith and David E. Booth, "Robust neural networks for statistical process control," DSI Proceeding, March 1, 1991, pp. 1-19.

Lapedes, A. and Farber, R., "Nonlinear signal processing using neural networks: Prediction and system modeling," Technical Report LA-UR-87-2662, Theoretical Division, Los Alamos National Laboratory, Los Alamos, N.M., 1987.

Lapedes, A. and Farber, R., "How neural net work," Neural Information Processing Systems, (Anderson, D. Z., ed.), American Institute of Physics, New York, 1988.

Law, Averill M. and W. David Kelton, Simulation modeling and analysis, McGraw-Hill Book Company, New York, 1982.

Lawton, Asa W., David C. Chance and John Y. Cheung, "An initial look at real-time neural network models of classical conditioning," Intelligent engineering systems through artificial neural networks, Proceedings of the Artificial Neural Networks in Engineering (ANNIE '91), (November 10-13, 1991), ASME Press, New York, pp. 973-979.

Ledolter, J., "The effect of additive outliers on the forecasts from ARIMA models," International Journal of Forecasting, Vol. 5, No. 2, (1989), pp. 231-240.

Lee, B. and B. F. Ingram, "Simulation estimation of time-series models," Journal of Econometrics, Vol. 47, (1991), pp. 197-205.

Lee, Sukhan and Rhee M. Kil, "Robot kinematic control based on bidirectional mapping neural network," International Joint Conference on Neural Networks, San Diego, California, Vol. 3, IEEE Neural Networks Council, Michigan, (June 17-21, 1990), pp. 327-335.

Lehar, S. and John Weaver, "A developmental approach to neural network design," International Joint Conference on Neural Networks, San Diego, California, Vol. 2, IEEE Neural Networks Council, Michigan, (June 21-24, 1987), pp. 97-104.

Lester, Bruce P., "A neural network simulation language based on multi-pascal," International Joint Conference on Neural Networks, San Diego, California, Vol. 3, IEEE Neural Networks Council, Michigan, (June 21-24, 1987), pp. 347-354.

Levin, Esther, Raanan Gewirtzman, and Gideon F. Inbar, "Neural network architecture for adaptive system modeling and control," International Joint Conference on Neural Networks, Washington, D.C., Vol. 2, 1989, pp. 311-316.

Lippmann, R. P., "An introduction to computing with neural nets," IEEE ASSP Magazine, (April 1987), pp. 4-22

Lo, Zhen-Ping and B. Bavarian, " A neural piecewise linear classifier for pattern classification," International Joint Conference on Neural Networks, Seattle, Vol. 1, IEEE Service Center, New Jersey, July 8-12, 1991, pp. 263-266.

Makridakis, S., Wheelwright, S. C. and McCree, V. E., Forecasting: Methods and Applications, Second Edition, John Wiley and Sons, New York, 1983.

Makridakis, S. et. al., The Forecasting Accuracy of Major Time Series Methods, John Wiley and Sons, New York, 1984.

Malkoff, Donald B., "A neural network for real-time signal processing," Advances in neural information processing systems 2, Morgan Kaufmann Publishers, Inc., San Mateo, California, 1990, pp. 248-255.

Mannaert, Herwig and Andre Oosterlinck, "A self-organizing system for objects analysis and identification, International Joint Conference on Neural Networks, San Diego, California, Vol. 2, IEEE Neural Networks Council, Michigan, (June 17-21, 1990), pp. 819-824.

Maren, A. C. Harston, and R. Pap, Handbook of Neural Computing Applications, Academic Press, Inc., San Diego, California, 1990.

Marshall, E. "Nuclear fuel accounts in bad shape," Science, Vol. 211, 1981, pp. 147-150.

McCullough, Claire L., "An anticipatory fuzzy logic controller utilizing neural net prediction," Simulation, Vol 58, No. 5, (May 1992), pp. 327-332.

McDonald, J. B., "Partially adaptive estimation of ARMA time series model," International Journal of Forecasting, Vol. 5, No. 2, 1989, pp. 210-230.

McWilliams, Gary, "The father of AI says his child has gone astray," Business Week, No. 3256, (March 2, 1992), p. 105.

Mikolajczak, B. (Ed.), Algebraic and structural automata theory, Elsevier Science Publishers B.V., Amsterdam, Netherlands, 1991.

Miller, David P. et. al. (panelists), "AI: What simulationists really need to know," 1991 Winter Simulation Conference Proceedings, The Society for Computer Simulation, International, San Diego, California, (December 1991), pp. 204-209.

Miller, J. C. and J. M. Miller, Statistics for Analytical Chemistry, Halsted, New York, 1984.

Minai, Ali A. and Ronald D. Williams, "Back-propagation heuristics: A study of the extended delta-bar-delta algorithm," International Joint Conference on Neural Networks, San Diego, California, Vol. 1, IEEE Neural Networks Council, Michigan, (June 17-21, 1990), pp. 595-600.

Montgomery, D. C., Introduction to Statistical Quality Control, 2nd ed., John Wiley & Sons, New York, 1991.

Murre, J. M. J., R. H. Phaf, and G. Wolters, "Calm networks," International Joint Conference on Neural Networks, Washington, D.C., Vol. 1, 1989, pp. 649-656.

Namphol, Aran, Mohammed Arozullah, and Steven Chin, "Higher Order data compression with neural networks," International Joint Conference on Neural Networks, IEEE, Seattle, Washington, (July 1991), pp. 55-59.

Nelan, B. W., "Formula for terror," Time, (August 29, 1994), Vol. 144, No. 9, pp. 46-51.

Neter, John, William Wasserman, and Michael H. Kuther, Applied linear statistical models, Second Edition, Irwin, Homewood, IL, 1985.

NeuralWare, Inc., Neural Computing, Pittsburgh, PA, 1991, Author.

Newbold, Paul, and T. Bos, Introductory business and economic forecasting, South-Western Publishing, Cincinnati, OH, 1994.

Osyk, Barbara, The use of neural networks for fault analysis in a computer integrated manufacturing system, Ph.D Dissertation, Kent State University, OCLC #26286644, (December, 1991).

Owens, Aaron J., "Artificial neural networks for process modeling and process control," ASA ,1992, pp. 1-6.

Oyster, J. Michael and Josef Skrzypek, "Computing shape with neural networks: A proposal," International Joint Conference on Neural Networks, San Diego, California, Vol. 4, IEEE Neural Networks Council, Michigan, (June 21-24, 1987), pp. 335-344.

Pack, D. J. and Downing, D. J., "Why didn't Box-Jenkins win (again)?," 3rd International Symposium on Forecasting, 1983.

Padgett, Mary Lou and Thaddeus A. Roppel, "Neural networks and simulation: Modeling for applications," Simulation, Vol 58. No. 5, (May 1992), pp. 295-305.

Paik, Eugene, David Gungner, and Josef Skrzypek, "UCLA SFINX - A neural network simulation environment," International Joint Conference on Neural Networks, San Diego, California, Vol. 3, IEEE Neural Networks Council, Michigan, (June 21-24, 1987), pp. 367-375.

Pankratz, A., Forecasting with univariate Box-Jenkins models, John Wiley and Sons, New York, 1983.

Parker, David B., "Optimal algorithms for adaptive networks: Second order back propagation, second order direct propagation, and second order hebbian learning," International Joint Conference on Neural Networks, San Diego, California, Vol. 2, IEEE Neural Networks Council, Michigan, (June 21-24, 1987), pp. 593-600.

Patuwo, Eddy, Michael Y. Hu, and Ming S. Hung, "Two-group classification using neural networks, Decision Sciences. Vol. 24, No. 4, (July-August 1993), pp. 825-845.

Picard, R. R., "Sequential analysis of material balances," Journal Institute Nuclear Materials Management. Vol. 15, No. 2, 1987, pp. 38-42.

Pike D. H.; G. W. Morrison, and D. J. Downing, Time series analysis techniques applicable to nuclear material accountability data. Oak Ridge Natiional Laboratory, 1978.

Pineda, Fernando J., "Time dependent adaptive neural networks," Advances in neural information processing systems 2. Morgan Kaufmann Publishers, Inc., San Mateo, California, 1990, pp. 710-718.

Port, Otis, "Sure, it can drive, but how is it at changing tires?," Business Week. No. 3256, (March 2, 1992), pp. 98-99.

Powell, Bill, Anne Underwood, Seema Nayyar, and Charles Fleming, " Eyes on the future," Newsweek. Vol. CXXI, No. 22, (May 31, 1993), pp. 39-41.

Prasad, Sameer, Robust montoring method for early detection of deviations in a time series process with applicatins in Operations Management., Ph.D. Dissertation, Kent State University, OCLC #23006757, 1990.

Prasad, Sameer, David Booth, Michael Y. Hu, and Seyda Deligonul, Monitoring the quality of a chemical production process using the Joint Estimation Procedure." Journal of Chemical Information and Computer Science. Vol. 35, 1995, pp. 53-58.

Prasad, Sameer, David Booth, Michael Y. Hu, and Seyda Deligonul, "The detection of nuclear material losses," Decision Sciences. Vol. 26, No. 2, 1995, pp. 265-281.

Psaltis, Demetri, Athanasios Sideris, and Alan Yamamura, "Neural Controllers", International Joint Conference on Neural Networks. San Diego, California, Vol. 4, IEEE Neural Networks Council, Michigan, (June 21-24, 1987), pp. 551-558.

Pugh, G. Allen, "Synthetic neural networks for process control," Computers and Industrial Engineering. Vol. 17, No. 1-4, 1989, pp. 24-26.

Pugh, G. Allen, "A comparison of neural networks to SPC charts," Computers and Industrial Engineering. Vol. 21, No. 1-4, 1991, pp. 253-255.

Qin, Si-Zhao, Integration and comparison of neural networks and conventional modeling techniques. Dissertation, 1992.

Quesenberry, C. P., "Screening outliers in normal process control data with uniform residuals, " Journal of Quality Technology. Vol. 18, No. 4, 1986, pp. 226-233.

Quesenberry, C. P., "Screening outliers in process control regression data with uniform residuals," Journal of Quality Technology, Vol. 22, No. 2, 1990, pp. 87-94.

Reibling, Lyle A. and Michael D. Olinger, "A neural network implementation of parallel search for multiple paths," International Joint Conference on Neural Networks, Washington D.C., Vol. 2, Lawrence Erlbaum Associates, New Jersey, (January 15-19, 1990), pp. 675-678.

Reece, Peter, "Perceptrons and neural nets," AI Expert, (January 1987), pp. 50-55.

Reinschmidt, Kenneth F., "Neural networks: Next step for simulation and control," Power Engineering, Vol. 95, No. 11, (November 1991), pp. 41-45.

Ripley, B. D., "Neural networks and related methods for classification," Journal Royal Statistical Society, Vol. 56, No. 3, 1994, pp. 409-456.

Roberts, H. V., Beyond statistical process control: statistical process monitoring and improvement. Presentation at the meeting of the Chicago chapter of the American Statistical Association, 1989.

Roberts, Markus, "Twelve neural network cliches AI Expert, (August 1988), pp. 40-46.

Rogers, K. C., "Safety is the key to a nuclear future," Public Utilities Fortnightly, (April 1991), pp. 54-56.

Rosenfeld, E., "Neurocomputing - A new industry," IEEE International Conference on Neural Networks, Vol. 4, (June 21-24, 1987), pp. 831-838.

Rousseeuw, Peter J. and Annick M. Leroy, Robust regression and outlier detection, John Wiley and Sons, New York, 1987.

Rumelhart, D. E. and J. L. McClelland, Parallel Distributed Processing, 3rd Ed., The MIT Press, Cambridge, Massachusetts, 1987.

Saint Donat, Jean, Naveen Bhat and Thomas J. McAvoy, "Optimizing neural net based predictive control," Proceedings of the 1990 American Control Conference, American Automatic Control Council, San Diego, California, 1990, pp. 2466-2471.

Salam, Fathi M. A. and Shi Bai, "A feedback neural network with supervised learning," International Joint Conference on Neural Networks, San Diego, California, Vol. 2, IEEE Neural Networks Council, Michigan, (June 17-21, 1990), pp. 263-268.

Sartori, David E., "Neural networks, statistics and experimental designs," Scientific Computing and Automation," Vol. 8, No. 11, (October 1992), pp. 4-6.

Sastri, T., B. Fores, and J. Valdes, "Detecting points of change in time series," Computers Operation Research, Vol. 16, No. 3, (1989), pp. 271-293.

Schrodt, Philip A., "Prediction of interstate conflict outcomes using a neural network," Social Science Computer Review, Vol. 9, No. 3, 1991, pp. 359-380.

Sebastian, Paul R., Polynomial smoothing and data bounding applied to industrial process control and nuclear material safeguards, Ph.D. Dissertation, Kent State University, OCLC #32172826, (August, 1994).

Sebastian, Paul R., David E. Booth, and Michael Y. Hu, "Using polynomial smoothing and data bounding for the detection of adverse process changes in a chemical process," Journal of Chemical Information and Computer Science, Vol. 34, No. 4, 1994, pp. 881-889.

Sebastian, Paul. R., David E. Booth, and Michael Y. Hu, "Using polynomial smoothing and data bounding for the detection of nuclear material diversions and losses." Journal of Chemical Information and Computer Science, Vol. 35, No.3, 1995, pp. 1-8.

Sejnowski, T. J. and Rosenberg, C. R., "Parallel networks that learn to pronounce English text," Complex Systems, Vol.1, 1987, pp. 145-168.

Shadmehr, Reza and David Z. D'Argenio, "A comparison of a neural network based estimator and two statistical estimators in a sparse and noisy data environment," International Joint Conference on Neural Networks, Washington D.C., Lawrence Erlbaum Associates, Inc., Hillsdale, New Jersey, (January 1990), pp. 289-292.

Sharda, R. and Patil, R. B., "Neural networks as forecasting experts: An empirical test," IEEE INNS International Joint Conference on Neural Networks, II, Washington D.C., (January 1990), pp. 491-494.

Shinskey, F. G., Process control systems: Application, design, and adjustment, McGraw-Hill Book Company, New York, 1988.

Sillin, J. C., "Spent nuclear fuel - What is it and where to put it?," Public Utilities Fortnightly, (November 1990), pp. 26- 27 & 52.

Smith, Andrew B., "A parallel PDP network simulator," International Joint Conference on Neural Networks, San Diego, California, Vol. 3, IEEE Neural Networks Council, Michigan, (June 21-24, 1987), pp. 377-384.

Smith, A. E., "Control chart interpretation using neural computing," International Journal Production Research, Vol. 32, No. 2, 1994, pp. 309-320.

Smith, D. J., "Nuclear suppliers prepare for the next generation of orders," Power Engineering, Vol. 94, No. 3, (March 1990), pp. 18-25.

Soule, L. M., "Basic control modes," Chemical Engineering, (October 20, 1969), pp. 115-119.

Spall, James C. and John A. Cristion, "Direct adaptive control of nonlinear systems using neural networks and stochastic approximation," IEEE Transactions on Neural Networks, Vol. 4, No. 1, 1993, pp. 1-35.

Spall, James C., "Multivariate stochastic approximation using a simultaneous perturbation gradient approximation," IEEE Transactions on Automatic Control, Vol. 37, No. 3, (March 1992), pp. 332-341.

Speed, T. P. and D. Culpin, "The role of statistics in nuclear materials accounting: Issues and problems," J. R. Statis. Soc. A., Vol. 149, part 4, 1986, pp. 281-313.

Stornetta, W. Scott and B. A. Huberman, "An improved three-layer, back propagation algorithm," International Joint Conference on Neural Networks, San Diego, California, Vol. 2, IEEE Neural Networks Council, Michigan, (June 21-24, 1987), pp. 637-643.

Stranahan, Susan Q., "Reactors redux," Audubon, (January-February 1992), pp. 21-26.

Studt, Tim, "Neural networks: Computer toolbox for the '90's," Research and Development Magazine, (September 1991), pp. 36-42.

Subramanian, Venkat and Ming S. Hung, "A GRG2-based system for training neural networks: Design and computational experience", ORSA Journal on Computing, Vol. 5, No. 4, (Fall, 1993), pp. 386-394.

Subramanian, Venkat and Ming S. Hung, "Application of GRG2 for training neural networks," International Joint Conference on Neural Networks, Seattle, Vol. 2, IEEE Service Center, New Jersey, (July 8-12, 1991), , pp. 207-213.

Subramanian, Venkat, Ming S. Hung, and Michael Hu, " An experimental evaluation of neural networks for classification," Computers and Operations Research, Vol. 20, No. 7, 1993, pp. 769-782.

Sun, G. Z., H. H. Chen and Y. C. Lee, "Learning decision trees using parallel sequential induction network," Abstracts of the First Annual INNS Meeting, Boston, MA, Vol. 1, (September 6-10, 1988), p. 54.

Swartz, Evan I., "Smart programs go to work," Business Week, No. 3256, (March 2, 1992), pp. 97-105.

Sweet, C., "Does nuclear power have a future?," Energy Policy, Vol. 18, No. 5, (June 1990), pp. 406-422.

Swift, A. L. and G. J. Janacek, "Forecasting non-normal time series," Journal of Forecasting, Vol. 10, No. 5, 1991, pp. 501-520.

Szu, Harold and Joseph Landa (Ed.), "Future perspectives on neural network research," International Joint Conference on Neural Networks, San Diego, California, Vol. 1, IEEE Neural Networks Council, Michigan, (June 21-24, 1987), pp. 159-168.

Tam, K., "Neural network models and the prediction of bankruptcy," Omega International Journal of Management Science, Vol. 19, No. 5, (November 5, 1991), pp. 429-445.

Thacker, Neil A. and John E. W. Mayhew, "Designing a layered network for context sensitive pattern classification," Neural Network, Vol. 3, 1990, pp. 291-299.

Thesen, Anne and Laurel E. Travis, "Introduction to simulation," 1991 Winter Simulation Conference Proceedings, The Society for Computer Simulation, International, San Diego, California, (December 1991), pp. 14-21.

Thome, H., "A Box-Jenkins approach to modeling outliers in time series analysis," Sociological Methods and Research, Vol. 23, No. 4, 1995, pp. 442-478.

Touchton, Robert A, "Verification and validation (V&V) issues and experience for real-time expert systems in a nuclear power plant," International Control Engineering Exposition and Conference, (February 24-26, 1992), pp. 14-20.

Touretzky, David S. and Dean A. Pomerleau, "What's hidden in the hidden layers?," Byte, (August 1989).

Tsay, R. S., Outliers, level shifts, and variance changes in time series," Journal of Forecasting, Vol. 7, 1988, pp. 1-20.

Tucker, G. K. and D. M. Wills, A simplified technique of control system engineering, Honeywell Inc., Fort Washington, PA, 1958, 1960, 1962.

Tunner, J., "Is an out-of-spec product really out-of-spec?" Quality Progress, Vol. 23, No. 12, 1990, pp. 57-59.

Ungar, L. and Foster, B., "Neural nets to establish weights for combining forecasts," ORSA/TIMS Joint National Meeting, Nashville, (May, 1991).

Vasilopoulous, A. V. and A. P. Stamboulis, "Modification of control chart limits in the presence of data correlation," Journal of Quality Technology, Vol. 10. No. 10, 1978, pp. 20-30.

Vidal, Jacques J., Joseph C. Pemberton, and James M. Goodwin, "Implementing neural nets with programmable logic," International Joint Conference on Neural Networks, San Diego, California, Vol. 3, IEEE Neural Networks Council, Michigan, (June 21-24, 1987), pp. 539-545.

Wang, Gou-Jen and Denny K. Miu, "Unsupervising adaption neural-network control," International Joint Conference on Neural Networks, San Diego, California, Vol. 3, IEEE Neural Networks Council, Michigan, (June 17-21, 1990), pp. 421-428.

Wang, Jun and B. Malakooti, "A feedforward neural network for multiple criteria decision making," Computers Operations Research, Vol. 19, No. 2, 1992, pp. 151-167.

Wardell, Don G., Herbert Moskowitz, and Robert D. Plante, "Control charts in the presence of data correlation," Management Science, Vol. 38, No. 8, (August 1992), pp. 1084-1105.

Wardell, Don G., Herbert Moskowitz, and Robert D. Plante, "Run-length distributions of special cause charts for correlated processes," Technometrics, Vol. 36, No. 1, (February, 1994), pp. 3-27.

Wei, William W. S., Time series analysis: Univariate and multivariate methods, Addison-Wesley Publishing Company, Inc., New York, 1990.

Weiss, S. M. and C. A. Kulikowski, Computer Systems That Learn, Morgan Publishers, Inc., San Mateo, California, 1991.

Werbos, P. J., Beyond regression: New tools for prediction and analysis in the behavioral sciences, Ph.D. Thesis, Harvard University, 1974.

Wheeler, D. J. and D. S. Chambers, Understanding statistical process control, Second Edition, SPC Press, Knoxville, Tennessee, 1992.

Wheeler, D. J., Understanding variation: The key to managing chaos, SPC Press, Inc., Knoxville, Tennessee, 1993.

White, Halbert, "Economic prediction using neural networks: The case of IBM daily stock returns," International Joint Conference on Neural Networks, San Diego, California, Vol. 2, IEEE Neural Networks Council, San Diego, (July 24-27, 1988), pp. 451-472.

White, Ray H., "The learning rate in back-propagation systems: An application of Newton's method," International Joint Conference on Neural Networks, San Diego, California, Vol. 1, IEEE Neural Networks Council, Michigan, (June 17-21, 1990), pp. 679-684.

Williams, P. and P. N. Woessner, "The real threat of nuclear smuggling," Scientific American, Vol. 274, No. 1, (January, 1996), pp. 40-44.

Williams, R. H. and H. A. Feiveson, "Diversion-resistance criteria for future nuclear power," Energy Policy, Vol. 18, No. 6, (July-August 1990), pp. 543-549.

Williams, Tom, "Fuzzy logic to make rapid inroads in the next five years," Computer Design, Vol. 32, (July 1993), pp. 43,46.

Williams, T., (ed.), "Special report: Bringing fuzzy logic and neural computing together," Computer Design, Vol. 33, No. 8, (July 1994), pp. 69-84.

Wong, Andrew K. C. and John O. Vieth, "Neural networks in statistical classification," International Joint Conference on Neural Networks, Washington D.C., Vol. 1, Lawrence Erlbaum Associates, New Jersey, (January 15-19, 1990), pp.553-556.

Wong, R. S., P. Z. Wang, T. H. Goh, and B. K. Quek, "Fuzzy neural systems for stock selection," Financial Analysts Journal, Vol. 48, No. 1, (January-February 1992), pp. 47-52, 74.

Xu, Guanghan, HansKaspar Scherrer and G. Schweitzer, " Application of neural networks on robot grippers," International Joint Conference on Neural Networks, San Diego, California, Vol. 3, IEEE Neural Networks Council, Michigan, (June 17-21, 1990), pp. 337-342.

Xu, Lei, Adam Krzyzak, and Ching Y. Suen, "Associative switch for combining multiple classifiers," International Joint Conference on Neural Networks, Seattle, Vol. 1, IEEE Service Center, New Jersey, (July 8-12, 1991), pp. 43-48.

Yang, Hedong and Clark C. Guest, "Performance of back propagation for rotation invariant pattern recognition," International Joint Conference on Neural Networks, Seattle, Vol. 1, IEEE Service Center, New Jersey, July 8-12, 1991, pp. 365-370.

Yao, Yong and Qing Yang, " Programming neural networks: A dynamic-static model, International Joint Conference on Neural Networks, Washington D.C., Vol. 1, Lawrence Erlbaum Associates, New Jersey, (January 15-19, 1990), pp. 345-348.

Youden, W. J. and E. H. Steiner, Statistical manual of the Association of Official Analytical Chemists, AOAC; Washington D.C., 1975, Vol. 30.

Zink, W. T., T. P. Vogl, and J. K. Mangis, "Neural networks as classifiers of noisy patterns: An experimental comparison with Bayesian classifiers," INNS First Annual Meeting, Boston, Vol. 1, (September 6-10, 1988), p. 61.

Zornetzer, Steven F., Joel L. Davis, and Clifford Lau, An introduction to neural and electronic networks, Academic Press, Inc., San Diego, California, 1990.

PUBLICATIONS AND PRESENTATIONS BASED UPON THIS DISSERTATION

1) Hamburg, James H., David E. Booth, and G. Jay Weinroth, "A neural network approach to the detection of nuclear material losses," Journal of Chemical Information and Computer Science, In Press, 1996.

2) Hamburg, James H., David E. Booth, and G. Jay Weinroth, "A neural network approach to monitoring the quality of production processes," Submitted to Decision Sciences, on September 25, 1995.

3) Hamburg, James H., "The application of neural networks to production process control," Presentation at the 27th Central Regional Meeting of the American Chemical Society, Akron, May, 1995.

4) Hamburg, James H., David E. Booth, and G. Jay Weinroth, "A neural network approach to the detection of nuclear material losses," Proceedings at OAI Neural Network Symposium, Ohio University, August, 1995.

5) Hamburg, James H., David E. Booth, and G. Jay Weinroth, T. Isenhour, "A neural network approach to the detection of nuclear material losses," Presentation at the 28th Central Regional Meeting of the American Chemical Society, Dayton, 1996.

6) Hamburg, James H., David E. Booth, and G. Jay Weinroth, "A neural network approach to the detection of nuclear material losses," Invited Symposium Presentation, Duquesne University, February, 1996.

APPENDIX : SIMULATION COMPUTER PROGRAM

```
C
C        DEVELOPMENT OF A DATA SET USING SIMULATION
C
C
C       MAIN PROGRAM
C
C       ESTABLISH COMMON STATEMENT FOR PROGRAM
C
        COMMON ND,DP(100),DPRO(100),DPR(100),NDP,QQ,
     1TEMP,MUI,TEMPI,DPRA(100),TEMPII,SD,MUA,ZZ,
     2W(100),SWY,SW,PWY,MUN,MUC,MUF,TP,BP,SIMDS(1
     300),CPI,CPII,NA,CA(100,100),MXHL,MXLL,DPms(100
     4,100)
        REAL*8 A,B,C,D,E,F,G,H,M,O,P,Q,R,S,T,U,V,W,X,Y,Z
C
C       READ THE FILE THAT CONTAINS THE NUMBER REPRESENTING
C       THE TOTAL NUMBER OF DATA POINTS IN THE DATA SET =ND
C       THE TARGET VALUE IS EITHER ZERO OR THE PROCESS MEAN=MTV
C       CONTROL LIMIT MULTIPLIER = ZSD
C
C
        READ(11,*) ND,MTV,ZSD
  350 FORMAT(2X,2I5,5X,F5.2)
        WRITE(6,500)
  500 FORMAT('1',20X,'TOTAL NUMBER OF DATA POINTS')
        WRITE(6,350) ND,MTV,ZSD
C
C       READ THE FILE THAT CONTAINS THE ACTUAL DATA SET
C
        READ(12,*) (DP(I),I=1,ND)
        WRITE(6,505)
  505 FORMAT('1',20X,'ACTUAL DATA SET')
C
C       WRITE ACTUAL DATA SET TO A FILE
C
        WRITE(6,290) (DP(IS),IS=1,ND)
  290 FORMAT(10X,F20.3)
C
C       CALCULATE 80% LEVEL FOR THE TOTAL NUMBER OF
C       ACTUAL DATA POINTS
C
        NDP=ND*0.8
C
C       DEVELOP A NEW DATA SET WITH 80% LEVEL OF POINTS
C       FROM THE ACTUAL DATA SET STARTING WITH THE FIRST
C       POINT
C
        DO 10 J=1,NDP
        DPRO(J)=DP(J)
   10 DPR(J)=DP(J)
        WRITE(6,510)
  510 FORMAT('1',20X,'80% DATA SET')
C
C       WRITE NEW FORMED DATA SET TO FILE
C
        WRITE(6,290) (DPR(JS),JS=1,NDP)
C
C       CALCULATE THE MEDIAN STANDARD DEVIATION "D"
C       AND THE WEIGHTED MEAN "MU"
C
```

```
C
C     FIRST STEP:SORT DATA SET DPR FROM LOW TO HIGH
C
      RNDP=NDP
      DO 20 K=1,NDP-1
      DO 30 L=K+1,NDP
      IF(DPR(K) .LE. DPR(L)) GO TO 30
      TEMP=DPR(K)
      DPR(K)=DPR(L)
      DPR(L)=TEMP
   30 CONTINUE
   20 CONTINUE
      WRITE(6,515)
  515 FORMAT('1',20X,'INITIAL SORTED DATA SET')
C
C     WRITE SORTED DATA SET TO FILE
C
      WRITE(6,290)  (DPR(KS),KS=1,NDP)
C
C     DETERMINE THE INITIAL WEIGHTED MEAN
C
      IF ( (NDP/2)  .NE.  (RNDP/2.0) )  GO TO 40
      WRITE(6,700)   rndp,ndp
  700 FORMAT('1','COMPARED NDP TO RNDP',5x,f10.3,5x,i5)
      CUI=(DPR(NDP/2)+DPR((NDP/2)+1))/2
      GO TO 50
   40 CUI=DPR((RNDP/2.0)+0.5)
C
C      WRITE OUT INITIAL WEIGHTED MEAN
C
   50 WRITE(6,520)
  520 FORMAT('1',20X,'INITIAL WEIGHTED MEAN')
      WRITE(6,290)  CUI
C
C     DETERMINE THE DIFFERENCE BETWEEN CREATED DATA SET
C     POINTS & THE INITIAL WEIGHTED MEAN & DEVELOP A NEW
C     DATA SET WITH THE ABSOLUTE VALUE OF THE DIFFERENCES
C
      DO 60 IJ=1,NDP
      TEMPI=DPR(IJ)-cUI
      DPRA(IJ)=ABS(TEMPI)
   60 CONTINUE
      WRITE(6,525)
  525 FORMAT('1',20X,'ABSOLUTE VALUE DATA SET')
C
C     WRITE NEWLY CREATED DATA SET TO FILE
C
      WRITE(6,290)  (DPRA(LS),LS=1,NDP)
C
C     SORT DATA SET DPRA FRON LOW TO HIGH
C
      DO 70 IK=1,NDP-1
      DO 80 IL=IK+1,NDP
      IF(DPRA(IK)  .LE.  DPRA(IL)) GO TO 80
      TEMPII=DPRA(IK)
      DPRA(IK)=DPRA(IL)
      DPRA(IL)=TEMPII
   80 CONTINUE
   70 CONTINUE
      WRITE(6,555)
```

```
  555 FORMAT('1',20X,'SORTED ABS VALUE DATA SET')
      WRITE(6,290) (DPRA(LIS),LIS=1,NDP)
C
C     DETERMINE THE MEDIAN OF THE ABS VALUES FOR DPRA
C
      IF((NDP/2) .NE. (RNDP/2.0)) GO TO 90
      CUA=(DPRA(NDP/2)+DPRA((NDP/2)+1))/2.0
      GO TO 100
   90 CUA=DPRA((RNDP/2.0)+0.5)
C
C      DETERMINE THE STANDARD DEVIATION
C
  100 SD=CUA/0.6745
      WRITE(6,530)
  530 FORMAT('1',20X,'ABS WEIGHTED MEAN & STANDARD DEV.')
C
C     WRITE STANDARD DEVIATION & MU VALUES TO A FILE
C
      WRITE(6,300) CUA,SD
  300 FORMAT(10X,F20.3,10X,F20.3)
C
C     DETERMINE THE WEIGHTED MEAN FOR THE DATA SET
C
  140 DO 110 IN=1,NDP
      ZZ=(DPR(IN)-CUI)/SD
      IF(ABS(ZZ) .LE. 1.5) GO TO 120
      W(IN)=1.5/ABS(ZZ)
      GO TO 110
  120 W(IN)=1.0
  110 CONTINUE
C
C     WEIGTED VALUE ARRAY
C
      WRITE(6,565)
  565 FORMAT('1',20X,'WEIGHT VALUE ARRAY')
      WRITE(6,290) (W(LIN),LIN=1,NDP)
      SWY=0
      SW=0
      DO 130 JI=1,NDP
      PWY=DPR(JI)*W(JI)
      SWY=SWY+PWY
      SW=SW+W(JI)
  130 CONTINUE
      write(6,710)
  710 format('1',20x,'the yis * Ws sum and the ws sum')
      write(6,300) sw,swy
      CUN=SWY/SW
      WRITE(6,560)
  560 FORMAT('1',20X,'NEW CALCULATED WEIGHTED MEAN')
C
C      WRITE NEW CALCULATED MU TO FILE
C
      WRITE(6,290) CUN
      CUC=CUN-CUI
      IF(ABS(CUC) .LE. 0.001) GO TO 150
      CUI=CUN
      GO TO 140
C
C     FINAL WEIGHTED MEAN FOR THE DATA SET
C
```

```
  150 CUF=CUN
      WRITE(6,535)
  535 FORMAT('1',20X,'80% NO. VAL.,STD. DEV.,& FIN. WEIG. MEAN')
C
C     WRITE NUMBER OF DATA POINTS,STANDARD DEVAITION,
C     & FINAL WEIGHTED MEAN TO FILE
C
      WRITE(6,310) NDP,SD,CUF
  310 FORMAT(5x,I5,5x,F15.3,5x,F15.3)
C
C     ESTABLISH HIGH AND LOW POINTS
C
      PHI=3.1415926536
      TP=(4.0*SD)
      BP=(4.0*SD)
C
C     TOP AND BOTTOM LIMITS
C
      WRITE(6,570)
  570 FORMAT('1',20X,'TOP AND BOTTOM LIMITS')
      WRITE(6,300) TP,BP
C
C     CONSTRUCT THE SIMULATED DATA SET
C
      DO 160 JL=1,9
  160 SIMDS(JL)=CUF
      DO 170 JN=10,30
      AP=JN-10
      AI=AP/20.0
      AC=PHI*AI
      ACPI=TP*SIN(AC)
      CPI=CUF+ACPI
      IF(CPI .EQ. CUF) GO TO 180
      WRITE(6,735) AI,AC,CPI
  735 FORMAT(2x,F10.3,2x,F20.8,2x,F15.3)
      SIMDS(JN)=CPI
      GO TO 170
  180 SIMDS(JN)=CUF
  170 CONTINUE
      DO 190 NN=31,39
  190 SIMDS(NN)=CUF
      DO 200 NI=40,60
      API=NI-20
      AII=API/20.0
      ACI=PHI*AII
      ACPII=BP*sin(ACI)
      CPII=CUF+ACPII
      WRITE(6,725) AII,ACI,CPII
  725 FORMAT(2x,F10.3,2x,F20.8,2x,F15.3)
      IF(CPII .EQ. CUF) GO TO 210
      SIMDS(NI)=CPII
      GO TO 200
  210 SIMDS(NI)=CUF
  200 CONTINUE
      WRITE(6,540)
  540 FORMAT('1',20X,'SIMULATED DATA SET')
      IF(MTV .NE. 0) GO TO 509
      DO 810 JMM=1,60
  810 SIMDS(JMM)=SIMDS(JMM)-CUF
C
```

```
C     WRITE SIMULATED DATA SET TO FILE
C
  509 WRITE(6,290)  (SIMDS(NJ),NJ=1,60)
C
C     DEVELOP A TWO COL ARRAY WITH COL. TWO ONE VALUE
C     AHEAD OF THE FIRST COL. WHICH TAKES INTO ACCOUNT
C     TIME
C
      NA=0
      DO 220 NL=1,2
      DO 230 NNL=1,59
      CA(NNL,NL)=SIMDS(NNL+NA)
  230 CONTINUE
      NA=1
  220 CONTINUE
      WRITE(6,545)
  545 FORMAT('1',20X,'SIMULATED TWO COL. DATA SET')
C
C     WRITE NEW CREATED TWO COL. DATA SET
C
      DO 320 IMD=1,59
      WRITE(6,340)  (CA(IMD,JMD),JMD=1,2)
  320 CONTINUE
  340 FORMAT(10X,F20.3,10X,F20.3)
C
C     CALCULATED CONTROL LIMITS
C
      IF(MTV .NE. 0) GO TO 800
      CUF=0.0
  800 CXHL=CUF+(ZSD*SD)
      CXLL=CUF-(ZSD*SD)
C
C     DETERMINE IF VALUES IN THE 2ND COL. ARE WITHIN
C     THE CONTROL LIMITS & LABEL VALUES WITH A 1.0 OR 0.0
C
      DO 240 NJI=1,59
      IF(CA(NJI,2) .GT. CXHL) GO TO 250
      IF(CA(NJI,2) .LT. CXLL) GO TO 250
      CA(NJI,3)=0.0
      GO TO 240
  250 CA(NJI,3)=1.0
  240 CONTINUE
      WRITE(7,550)
  550 FORMAT('1',20X,'FINAL SIM 3 COL. DATA SET')
C
C     WRITE FILE DATA SET ARRAY TO A NEW FILE
C
      DO 260 NC=1,59
      WRITE(7,280)  (CA(NC,NR),NR=1,3)
  260 CONTINUE
  280 FORMAT(2x,f20.3,1x,f20.3,1x,f20.3)
c
c     form the master data set col. array
c
      NNA=0
      DO 585 III=1,2
      DO 590 JJJ=1,ND-1
      DPms(JJJ,III)=DP(JJJ+NNA)
  590 CONTINUE
      NNA=1
```

```
  585 CONTINUE
c
c      write two col. master data set array
c
       write(6,720)
  720 format('1',20x,'master data set-two col.')
      do 730 njj=1,nd-1
      write(6,340) (dpms(njj,jnj),jnj=1,2)
  730 continue
      DO 595 LLL=1,ND-1
  595 DPms(LLL,3)=0.0
      WRITE(8,575)
  575 FORMAT('1',20X,'TEST DATA SET')
      DO 600 KKK=1,ND-1
  600 WRITE(8,580) (DPms(KKK,NIL),NIL=1,3)
  580 FORMAT(2x,f20.3,1x,f20.3,1x,f20.3)
      STOP
      END
```