UNIVERSITY OF ALBERTA

# Application of Artificial Intelligence to Improve Pulp Mill Operations

By

© James Zurcher B.Sc. (Mech. Eng.) P.Eng

A Thesis submitted to the Faculty of Graduate Studies and Research in Partial Fulfilment of the requirements for the degree of Master of Science

IN

PROCESS CONTROL

DEPARTMENT OF CHEMICAL ENGINEERING

Edmonton, Alberta
Spring 1995

ISBN 0-612-01668-4

UNIVERSITY OF ALBERTA

RELEASE FORM

NAME OF AUTHOR: James Zurcher
TITLE OF THESIS: Application of Artificial Intelligence to Improve Pulp Mill Operations
DEGREE: M.Sc.
YEAR THIS DEGREE GRANTED 1995



James Zurcher

Box 1356
Grande Prairie
Alberta
T8V 4Z1

Dec 21/94

**UNIVERSITY OF ALBERTA**

**FACULTY OF GRADUATE STUDIES AND RESEARCH**

THE UNDERSIGNED CERTIFY THEY HAVE READ, AND RECOMMEND TO THE FACULTY OF GRADUATE STUDIES AND RESEARCH FOR ACCEPTANCE, A THESIS ENTITLED **APPLICATION OF ARTIFICIAL INTELLIGENCE TO IMPROVE PULP MILL OPERATIONS** SUBMITTED BY **MR. JIM ZURCHER** IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF **MASTER OF SCIENCE** IN **PROCESS CONTROL**.

Dr. Ming Rao

Dr. S. Shah

Dr. S. AbouRizk

Date: Dec. 21, 94

This thesis is dedicated to my mother and father, Jean and Bill Zurcher, whose continued support and encouragement made this possible.

# Abstract

For most of the last f[illegible] years Artificial Intelligence (AI) has been confined to academic study. In the past decade, however, it has experienced an explosive growth in industry. So pervasive has been its recent growth that it prompted Tom Peters, world-renowned industry watcher, to espouse "that any senior manager in any business of almost any size who isn't at least learning about artificial intelligence (AI), and sticking a tentative toe or two into AI's waters, is simply out of step, dangerously so." [1]

AI encompasses a varied array of technologies including virtual reality, speech recognition, natural language, intelligent robotics, expert systems and neural networks. The research covered by this thesis involves learning about the expert systems and neural network technologies, theories and tools, and then applying this knowledge to solve appropriate problems within the Weyerhaeuser pulp mill operations at Grande Prairie, Alberta.

Applying new technology, however, involves more than simply finding an elegant solution to a problem. There is a human side to technology implementation that must also be managed in order for any new technology to be effectively deployed and maintained. The knowledge gained in this area is discussed in this thesis.


---

[1] Allen, Mary Kay and Helferich, Omar Keith, Putting Expert Systems to Work in Logistics, Council of Logistics Management, Oak Brook, Il, 1990, p. 158

# ACKNOWLEDGEMENTS

I would like to thank Dr. Ming Rao for his generous guidance and support on this project and on my reintroduction to the academic world.

I would like to thank NSERC for their financial support

Financial support for this project by Weyerhaeuser Canada and the Canadian Alberta Forest Partnership is greatly appreciated.

I would also like to thank the members of my committee, namely the chairman, Dr. U. Sundararaj, and the committee members Dr. S. Shah and Dr. S. AbouRizk.

# Table of Contents

# Bibliography

131

## List of Tables

# List of Figures

# Chapter 1

# Introduction

The Pulp and Paper Industry has always been known for applying only sufficient levels of technology to solve the problems at hand. This has resulted in lower levels of technology being used than in other chemical industries. This reduced application is caused by the mistaken perception that the process is simple and inexpensive and therefore does not warrant the added complexity and risk that state of the art technology entails. This has been especially true in the real time information and process control technologies. Recently, several factors have combined to create a need for sophisticated technology to be applied against the industry's unique problems.

The drivers for this change include increased global competition, increased environmental awareness, improved technology to measure environmental impact and an increase in cost of the forest resources. The environmental drivers have served to illustrate to the industry how complex the pulp and paper process truly is. The increased cost of the diminishing forest reserves, the competitive pressures of globalization and escalating capital expenditures for the physical plant have dispelled the perception of an inexpensive process.

Even a process as simple as repulping used newspapers becomes very complex and expensive given the need to control to a high quality target while maintaining excellent environmental performance. In order to maintain costs and quality control the pulp and paper industry has begun to use technology successfully applied in other chemical industries, such as oil and gas. Two of the most exciting state of the art technologies increasingly applied are model-based control and expert systems.

Although the need for advanced technology is now recognised, there remains little data on how and where to implement advanced solutions. This thesis will review several applications of Artificial Intelligence to pulp mill operations. The purpose of these examples is to illustrate a variety of industry problems that can be addressed with advanced information technology, problems that could not be solved using traditional methods. As the examples are reviewed, the features that make them successful implementations will be discussed. This will allow the reader to reapply the methodology, if not the application.

The thesis comprises three examples. Chapter 2 deals with an expert system designed to help control pulp inventories at a customer's plant. In addition to requiring an expert system, it was determined that basic control theory could be applied to this problem. This is a marked departure from the accepted approach currently used by the industry.

The third chapter deals with the development of models to analyse plant operation. These models will eventually be incorporated into a control strategy, but that is beyond the scope of this paper. Three approaches to modelling are reviewed: parametric models (regression analysis); neural network models; and, adaptive logic models which are a hybrid of the first two methods. Each of these methods is illustrated with two examples. One model relates pulp machine throughput to the position of the basis weight valve. The second model is of pulp strength which is governed by a large number of processing parameters.

Chapter 4 deals with the issue of how to integrate the use of advanced technology into the manufacturing organizations that they support. The major issues are how to get the required organizational resources to develop and maintain the systems, and how to ensure that systems are used effectively and remain healthy contributors to the bottom line over the years.

The thesis looks at these three examples, illustrates their development, describes how they solve real industrial problems and then talks about the organizational and human factor issues that need to be addressed when implementing these types of projects in an industrial operation.

# Chapter 2

# Expert Systems

## 2.0 Overview

Expert Systems technology is one of the most popular and successful areas in the broad field of Artificial Intelligence (AI). The concept of computer programs that emulate human thought was at the root of the first AI efforts in the early 40's. Unfortunately, to date, mimicking human thought has proven to be too ambitious a target. Rather than attempting to recreate human thought processes, expert systems are designed to capture human knowledge and expertise. This is a worthwhile and more achievable goal. As a result there is a proliferation of expert systems and expert system development tools on the market[2].

Expert systems fulfil a real need. In general, expert systems are used where data exists in conjunction with a set of rules that are used to interpret the data and arrive at some conclusion. This is typical of problems found in the chemical industry. Industrial

---

[2] Allen, Mary Kay and Helferich, Omar Keith, Putting Expert Systems to Work in Logistics, Council of Logistics Management, Oak Brook, Il, 1990, p. xiv

applications of expert systems solve problems that:

- require analysis of large amounts of data
- are governed by complex rules
- utilize knowledge that is required infrequently
- utilize knowledge that is no longer available due to inaccessibility of experts due to vacations, transfers, retirement, and other such causes.

The later part of this chapter contains a discussion about the solution to a problem that exhibits these characteristics.

With the rule-based expert system, knowledge is represented as a series of rules as typified by IF-THEN type of logic. By simplifying human knowledge to this level, computer programs can be designed to generate the same conclusions that a human expert would provide, given the same problem description. As well, expert systems usually incorporate the capability to explain the process used to arrive at that conclusion to the user. Expert systems consist of these major components: the knowledge base, the database, the inference engine, and the user interface which perform the aforementioned functions respectively.

**Knowledge base.** The knowledge base is the knowledge of a human expert collected into a series of rules. A simple example could be:

- IF TANK LEVEL LESS THAN 1 METRE

  THEN TANK LEVEL IS LOW

- IF TANK LEVEL IS LOW

  THEN STOP PUMP

**Data base.** While the knowledge base represents the static knowledge of the expert, the data base contains the dynamic data of the specific problem being solved. In the above example, to determine if the pump should be stopped, not only are the rules (knowledge) required, but the actual level of the tank (data) is needed before the decision can be made. The data base describes the current situation, and the knowledge base provides the rules to be used when interpreting the data.

**Inference engine.** The inference engine is the heart of an expert system. Given a set of data, it parses through the rules to arrive at some conclusion. This is referred to as forward chaining in that the inference engine starts with the data and goes "forward" through the rules to arrive at a conclusion. Backward chaining is the process of starting with the conclusion and working backwards through the rules to determine whether the conclusion is true or not[3]. In the above example, given the tank level we could forward chain through the rules and produce the conclusion that the pump should (or should not) be stopped. Or we could assume that the pump should be

---

[3]**Zinn, Walter and Marmorstein, Howard,** ***Comparing Two Alternative Methods of Determining Safety Stock Levels: The Demand and the Forecast Systems,*** **Journal of Business Logistics, Volume 11 #1, Nov 1990, p. 40**

stopped and backward chain through the rules and data to determine if that were a valid assumption.

**User interface.** The user interface performs two independent sets of functions for the expert system. During the reasoning stage, it asks the user for any needed data that is not contained in the data base. Once a conclusion is reached, most user interfaces allow the user to query the system and find out what rules and data contributed to arriving at the conclusion.

Expert systems generally are written using a special development environment called an expert system shell. The shell contains the inference engine and the user interface. The knowledge base is contained in a separate file. While the knowledge base can be modified by the user, the shell itself usually cannot be. In the examples that follow, MetaCOOP, a shell developed in the Intelligence Engineering Laboratory at the University of Alberta was used. This shell was selected over other shells because its object oriented approach to knowledge representation provided a superior development environment. As well, because the source code was available, it could be modified to suit the needs of this project. These changes are described later in this chapter.

In the examples that follow, an expert system was developed in conjunction with a spreadsheet to provide knowledge, numerical analysis, and interfaces that solved inventory management problems for Weyerhaeuser Canada. The primarily source of knowledge was the logistics manager at the Grande Prairie Alberta pulp mill.

## 2.1 Continuous Replenishment

Continuous Replenishment is a process designed to deliver to the customer just the right quantity and quality of material at just the time that it is required. In its purest form, this is known as Just-In-Time delivery.

Ideally, the supplier would be able to deliver each item as it is required continuously throughout the day. This is possible in cases where the customer and supplier are in close proximity and delivery time is negligible.

When delivery time is significant, Just-In-Time delivery can still be achieved if the delivery time of the product from the supplier to the customer is constant and if the customer's usage can be accurately predicted for a period of time greater than the delivery time. Although examples of these situations exist, the normal situation is that usage is not fully predictable and variations in delivery time occur. This necessitates the accumulation of inventory at, or close to, the customer. This excess inventory is known as "safety stock". The amount of inventory that must be maintained is dependent on how well the usage can be predicted and how little deviation in delivery times occur[4]. The ideal safety stock level is the minimum level that adequately protects the customer from major deviations in delivery or usage. An expert system can be used to help minimize this inventory.

---

[4] Zinn, Walter and Marmorstein, Howard, *Comparing Two Alternative Methods of Determining Safety Stock Levels: The Demand and the Forecast Systems*, Journal of Business Logistics, Volume 11 #1, Nov 1990, p. 61

The current estimate is that maintaining pulp inventory represents a yearly cost of up to 25% of its value. This is based on the cost of foregoing investment opportunities, interest costs, quality degradation, extra handling, storage costs, and so forth. Because of this cost, an economic decision must be made whether to protect against all deviations. Protecting against deviations that may occur only once a decade might result in a very high safety stock level being maintained. The cost of maintaining this inventory may well exceed the cost of the potential supply disruption. The measure of a customer's sensitivity to disruption of supply is known as confidence level. A high confidence level indicates a high sensitivity.

The Distribution Department at the Weyerhaeuser pulp plant in Grande Prairie has developed a system to maintain minimum safety stock at their customers' plants. Even during periods of low pulp prices, the savings available to a customer by this inventory minimization is substantial.

The Continuous Replenishment Process (CRP) is composed of three subsystems:

**Safety Inventory Calculator**. This is used to determine the minimum inventory required by the customer to compensate for variances in usage, order processing and transit time; and in some circumstances, damaged or deteriorated supply. This subsystem is usually used when initially setting up a customer. It is also used to illustrate the advantages of CRP to the customer by showing how much inventory can be reduced. In the latter example, the system is used by sales personnel who may have a lower level of understanding of the calculations involved.

**Inventory Replenishment Planning**. This develops a pulp shipment plan based on the customer's pulp usage, required safety inventory levels, the customer's receiving constraints, production variations, and shipment scheduling. Once this plan is developed, it must be continually updated to account for changes in usage, transit time, etc. Experience has shown that forecasted usage is highly inaccurate and this results in a need to continually fine tune the shipment plan. This subsystem is generally run by people knowledgeable in CRP.

**Shipment Tracking**. This subsystem tracks shipments from the time they leave the plant until they are received and unloaded by the customer. This subsystem was not automated in this project.

Each subsystem, while seemingly straight forward, involves substantial troubleshooting as each component is complex and subject to errors. For example, when the systems are initially being developed, there is often very little data available. A common mistake is to perform the safety inventory calculation on a set of data that is too small or does not have a normal distribution in which case a safety inventory can not be calculated. This type of problem is not easy to detect yet it will result in improper inventories being calculated.

Generally, there are two elements in each subsystem: numerical algorithms and expertise knowledge. The numerical algorithms can be readily handled by a spreadsheet. The expertise knowledge, however, currently exists only in a handful of experts in the plant. Neither element by itself is sufficient for the system to operate. The knowledge is of little use without the numerical methods and, because the numerical methods are prone to inaccuracies, their use requires an expert.

The objective of this project was to capture the expertise knowledge and to develop the required numerical methods. These were then integrated in a computer system so that the CRP was not dependent on experts to function. It was felt by leadership within the plant that this was an ideal opportunity to introduce advanced computer technology into the pulp mill.

The CRP was bound by several user constraints. The system had to:

- be easy to use;
- link to automatic data retrieval systems where possible (to minimize manual entries); and,
- use an Excel interface as that is the current plant standard.

Data is stored either in Excel spreadsheets or an Oracle database, both of which are corporate standards. The system runs on an IBM compatible PC.

Ongoing maintenance of the system was considered and the system developed with tools appropriate to the tasks. Using Excel to capture heuristic knowledge would have been an inappropriate application of technology, resulting in ongoing maintenance problems. Similarly an expert system development tool was not suitable for significant numerical calculations.

In the sections that follow, we will discuss the two implemented subsystems: the expert system to aid in the calculation of safety inventory and the inventory replenishment tool. Both of these subsystems were only developed to a framework level. The initial goal of the project was to illustrate the technology to plant and divisional resources. As such, many more rules are required to completely flesh out the system.

## 2.2 Development of a Safety Inventory Expert System

The standard demand formula[5] for calculating safety stock is well known and given by:

$$SS_d = Z * \sqrt{(t * S_d^2) + (d^2 * S_t^2)} \qquad (2.1)$$

where

$SS_d$ = Safety Stock under the demand system

$Z$ = Confidence Limit

$t$ = Average Lead-time ( Order Processing Time + Delivery Time )

$S_d$ = Standard Deviation of Demand

$d$ = Average demand

$S_t$ = Standard Deviation of Lead-time

The reader is refered to Zinn for an explanation of the above formula. The calculation of safety inventories used at the Weyerhaeuser plant involves applying a simplified version of this formula[6] given by:

$$SS = Z * CV_{demand} * d * \sqrt{T} \qquad (2.2)$$

where

---

[5] Zinn, Walter and Marmorstein, Howard, *Comparing Two Alternative Methods of Determining Safety Stock Levels: The Demand and the Forecast Systems*, Journal of Business Logistics, Volume 11 #1, Nov 1990, p. 97

[6] Jacob, H.M. *Current Best Approach for Raw Material Inventory Tracking*, Confidential Procter & Gamble document, 1990

$SS$ = Safety Stock Level

$CV_{demand}$ = Coefficient of Variance of Demand

$T$ = Maximum lead-time plus all processing time. This includes the time between each system review, order processing time and transit time.

The simplified formula results because the choice has been made to use the maximum value (occurring during normal conditions) for the leadtime factor. Choosing this, as opposed to using the average lead-time, allows the standard deviation of lead-time ($S_t$) to be assumed to be 0. As a result, the last term of the demand formula is eliminated. Using this assumption, a simple change of form of the equation (using the definition of the coefficient of variance) is easily achieved resulting in the simplified formula shown in equation 2.2.

After the data is input by the user, the calculated output must be validated by an expert before the calculated safety inventory is accepted. The objective of this part of the project was to automate the calculation and validation processes.

The data flow diagram is shown in Figure 2.1.

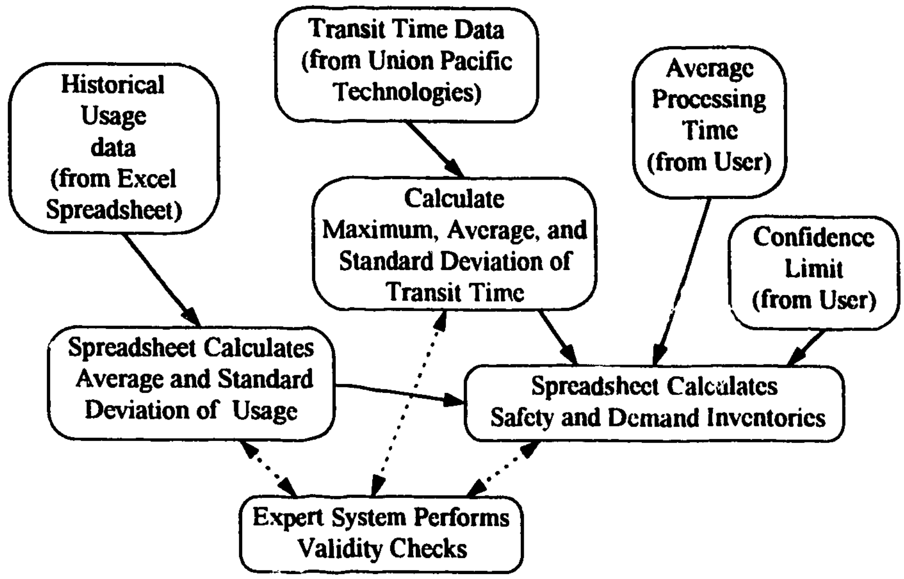


**Figure 2.1: Safety Inventory Calculation Data Flow Diagram**

Figure 2.1 illustrates the data flow once it has been prepared for inclusion into an Excel spreadsheet.

The three sources of data are:

**Historical Usage** which is contained in mainframe files. An ASCII report is generated, imported into Excel and parsed into a usable format.

**Transit Time** data which is obtained from Union Pacific Technologies (UPT) in the form of an ASCII report. This is parsed and loaded into an Oracle database. SQL routines in Excel are then utilized to extract data from the database

**Average Processing Time** (time to prepare all paperwork before shipment) and the confidence limit (how many standard deviations of safety are desired) which are entered by the user into the safety inventory calculation spreadsheet.

The safety inventory calculations are contained in an Excel spreadsheet part of which is shown in Figure 2.2. The spreadsheet is colour coded to quickly distinguish between the fields the user must enter, the fields based on calculations, and the fields containing knowledge base data.

| | Transit Time | Usage |
|---|---|---|
| AVERAGE | 9.97 | 202.64 |
| STDEV | 1.45 | 43.49 |
| CV | 0.15 | 0.21 |
| Count | 408 | 408 |
| | | |
| Confidence Limit= | 3 | |
| Avg Processing time | 17 | days |
| Safety Inventory= | 678 | tonnes |
| Demand Inventory= | 1111 | tonnes |
| D/SI err Ratio | 64% | |
| | | |
| Calculated from UPT & Usage Data | | |
| User Entered Data | | |
| Caculated Results | | |
| Knowledge Base Messages | | |

Check Results

Help on Results

| TRANSIT TIME | USAGE |
|---|---|
| 9 | 195 |
| 9 | 165 |
| 9 | 246 |
| 10 | 196 |
| 11 | 199 |
| 10 | 119 |
| 11 | 218 |
| 11 | 206 |
| 10 | 267 |
| 9 | 261 |
| 11 | 221 |
| 11 | 246 |
| 7 | 178 |
| 10 | 272 |
| 10 | 137 |

**Figure 2.2: Safety Inventory Spreadsheet**

In use a macro is executed automatically when the spreadsheet is opened. This macro parses the ASCII reports for transit time and historical usage, extracts the required data, and enters it into the appropriate cells in the spreadsheet.

Once the macro has completed filling in the spreadsheet, the user enters the confidence limit and the average processing time. The spreadsheet automatically calculates the required safety inventory, the demand inventory and the ratio of demand to safety inventory. The user may then click on the "Check Results" button to invoke the MetaCOOP based expert system to perform a validity check on the input data and the results.

As the goal of the project was simply to demonstrate the capabilities of expert systems, the knowledge base is quite limited. Based on the success of the demonstration it will be expanded in the future. The current knowledge base contains the following reasonableness checks which represent some of the knowledge used in the manual system:

- Ensure that a sufficient historical data is used. Currently the threshold is set to 150 data points.
- Ensure that the Standard calculation and the Demand calculation give similar answers. The two calculations are considered similar if they are within 25% of each other.
- Ensure that the variability of the historical data is not too large. Variability is too great if the coefficient of variance is greater than .33.

The above rules were those specified by the logistics manager (although she does plan to implement more robust statistic measures in the future).

To meet the requirement of no extraneous output, the expert system operates without sending any messages to the user directly. All I/O is performed through the spreadsheet. All the user sees is two icons appearing for approximately 5 seconds while the expert system runs in the background. When the expert system analysis is complete, the spreadsheet is updated with messages that state whether the data is normally distributed or not, whether the data count is OK and whether the ratio of demand to safety inventory is within limits. If the user is unsure of any messages that are displayed, he or she can click on the "Help on Results" button which activates a hypertext help system. This help system explains the rules, messages, and assumptions used in the calculation of safety inventory. The hypertext help system was written using Guide version 3.1[7]. Guide was chosen as previous experience with this package had shown it to be a full-featured and easy to use hypertext development system.

Once the appropriate safety inventory level is calculated the user is ready to build an inventory replenishment plan. In practice this would first be done in conjunction with the safety inventory calculation. It would then be recalculated on a regular basis as usage, transit times, or the variance of either of these change.

[7] OWL International, Inc. 2800 156th Avenue SE, Bellevue WA, 98007 USA, (206)747-3203

## 2.3 Development of an Inventory Replenishment Expert System

The calculation of a shipping plan that maintains a customer's inventory at a target level is a highly manual process that is difficult to computerize. Unlike the safety calculations, inventory replenishment is not based on a well known set of calculations.

As part of this project, several commercial inventory management programs were investigated. Although the algorithms are considered proprietary, observations showed that they search for an optimal solution through a highly iterative methodology. These programs are expensive, slow, and require powerful workstations. Manual systems handle inventory management as illustrated in Figure 2.3. A key feature is the lack of inventory level feedback to the shipment plan.

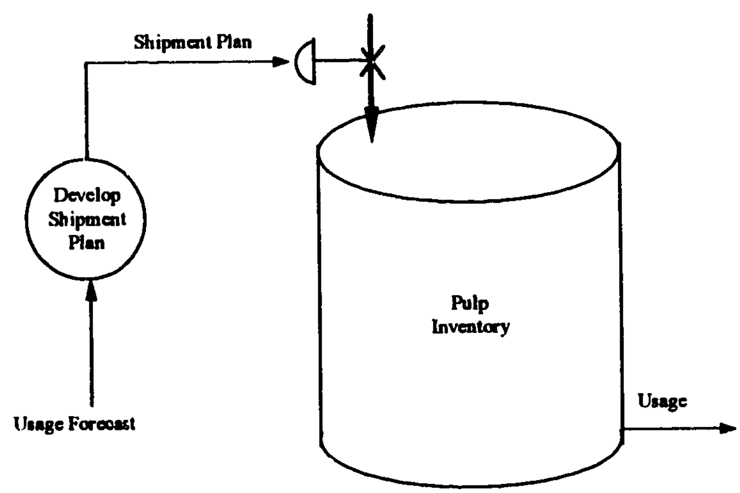


**Figure 2.3: Tank Analogy to Inventory Management as Used by Industry Today**

The first step in building a system for Grande Prairie was to develop an algorithmic approach that reduced the iterative nature of inventory management and captured the more robust aspects of their inventory replenishment planning process. The basic pulp shipment process is shown in Figure 2.4.

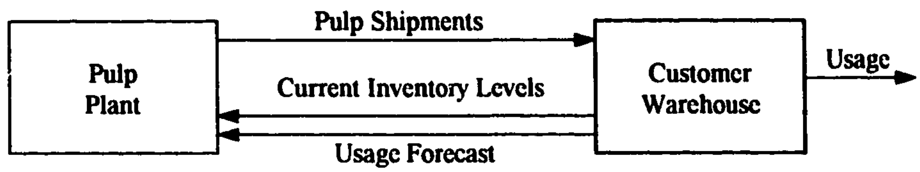


**Figure 2.4: Pulp and Information Flow**

This system is analogous to the simple tank system shown in Figure 2.5 and the control procedures are similar.

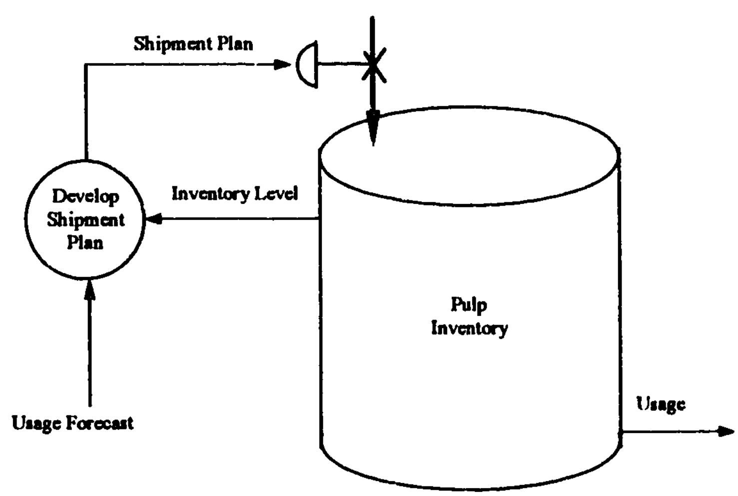


**Figure 2.5: Tank Analogy to Inventory Management as Used by Weyerhaeuser**

A simulation was built by the Logistics Manager that modelled a customer's pulp usage. This model incorporated random normal variations in transit time and usage. It was very successful in emulating a typical customer. This model was utilized when analyzing various approaches to inventory management.

Because usage is completely independent from inventory, the warehouse can be represented as a pure integrating process. Inventory replenishment can be a simple application of control theory learned in undergraduate engineering.

The three major feedback control schemes, P, PI, and PID were applied to this (simulated) "process" and their outputs compared. Of particular interest was their ability to maintain inventories around a setpoint (in this case 500 tonnes) with minimal variance. The minimum, maximum, average and standard deviation of the inventory levels are compared in Table 2.1 and the ongoing inventories are shown in Figures 2.6, 2.7 and 2.8. As well, an inferential control scheme was applied and is shown in Figure 2.9. These figures are included to show the effect of the discrete nature of this process on the control schemes. The large fluctuations (at steady state) of inventory level are caused by the long periods between shipments (2-3 days).

The usage also fluctuates in the first 13 days. This is because of the simulation used. The simulation models days 1-13 as having already occurred and therefore models fluctuations typically seen at a customer's plant. Usage for days 14-36 are usage forecasts. Typically these forecasts are simply an average daily usage (as shown here).

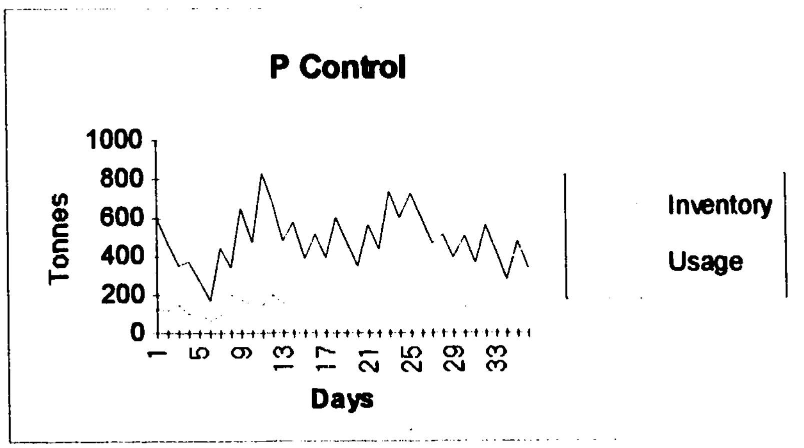


**Figure 2.6: Proportional Control**

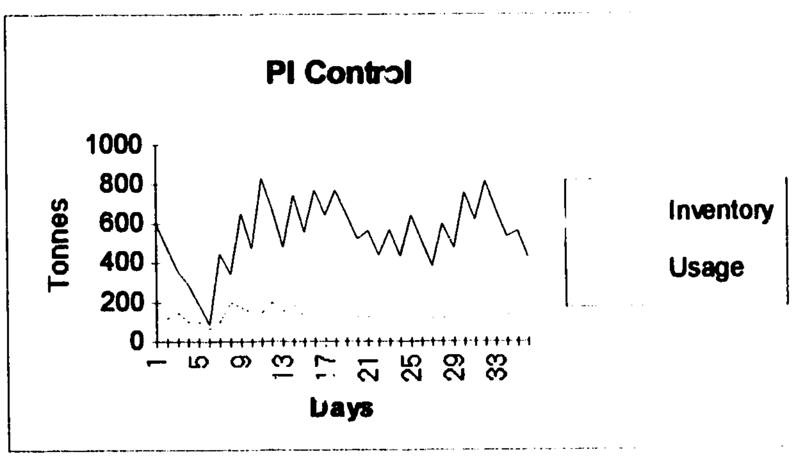


**Figure 2.7: Proportional Integral Control**

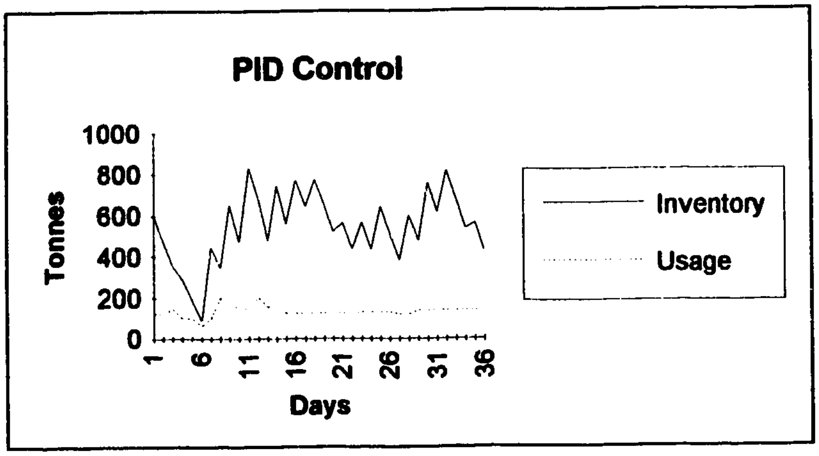


**Figure 2.8: Proportional Integral Derivative Control**

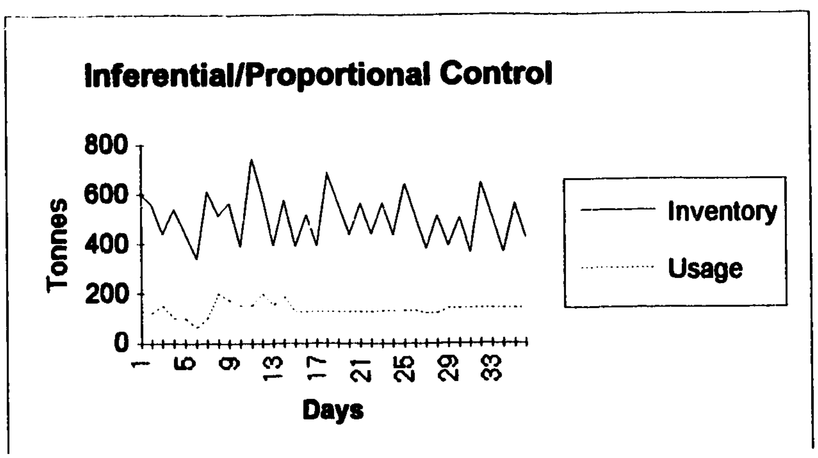


**Figure 2.9: Proportional and Inferential Control**

| Control Type | P | PI | PID | P/Inferential |
|---|---|---|---|---|
| Maximum Inv | 829 | 829 | 758 | 745 |
| Minimum Inv | 173 | 89 | 5 | 341 |
| Average Inv | 501 | 544 | 437 | 511 |
| Std Dev | 141 | 174 | 195 | 103 |

**Table 2.1: Control comparison**

The three feedback control schemes were all able to control the inventory when usage variation was low. However, step changes in usage typically resulted in wide fluctuations in inventory levels. This is due to:

- the highly discrete nature of the shipping process.
- the irregular shipping schedule, i.e. Mon, Wed, & Fri rather than every day.
- constraints at the user site on the number of cars that can be unloaded each day.

Of the three causes, the last was probably the most influential. Through several simulations with different tuning parameters, it was observed that the controller output could often exceed the customer's ability to unload pulp. This often resulted in system instability. With proper tuning, of the three feedback schemes, pure proportional control showed the least oscillatory behaviour as measured by range and standard deviation of the inventory levels.

Although the proportional scheme provides adequate control, the customer's usage forecast can be incorporated into the control scheme to provide more robust inventory management. In some manual implementations of inventory management, only the usage forecast is used to develop shipment plans. In these systems, adjustments are made on an irregular frequency to accommodate inaccuracies in the forecast. Grande Prairie's initial manual system was of this type, although inventory adjustments were made on a frequent (usually weekly) basis. A combination of inferential and proportional control approximately replicates the current manual system used in Grande Prairie.

The optimum choice of control scheme for inventory control represents an intriguing opportunity for further research. For this thesis however, the proportional/inferential control scheme provides adequate performance and has the advantage of capturing the essence of the current manual system. The following example illustrates the control algorithm as built for the inventory replenishment spreadsheet.

To calculate the number of tonnes that need to be shipped on any given day, the following steps are executed.

1. Calculate today's inventory. This is yesterday's inventory plus what was unloaded yesterday less what was used yesterday.

2. Calculate forecast usage. This is the sum of the projected usage between now and the next time a shipment can occur. (Note, the spreadsheet accommodates "odd" shipping schedules. For example, shipments usually occur on Monday, Wednesday, and Friday, but during shutdowns, rail problems, etc., it is able to accommodate any shipment schedule.)

3. Calculate the inventory shortfall (excess). This is the safety inventory less today's inventory. This is the deviation from the target (safety) inventory levels. Multiplied by a gain this would be the response from a proportional controller.

4. Calculate the number of tonnes needed. This is the forecast usage plus the inventory shortfall. The forecast usage represents the inferential portion of the control scheme. The number of tonnes needed is what is required in today's shipment.

The above control scheme provides a satisfactory, though non-optimal solution to inventory management. An expert that reviews the shipment plan from this spreadsheet can usually improve the plan. The expert typically would change slightly the target inventory to accommodate transportation outages, step changes in usage, and other factors. To capture the expert's knowledge, an expert system was built that evaluates the shipment plan, applies the heuristic knowledge of the expert, changes the target inventory by some offset, and then reruns the shipment plan. This may be repeated several times.

The data flow diagram for the project is shown in Figure 2.10.

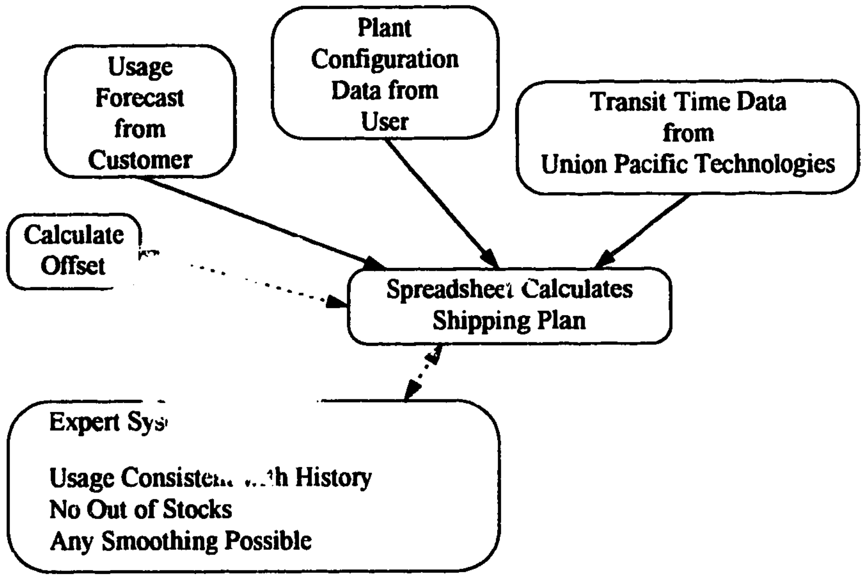


**Figure 2.10: Inventory Replenishment Calculation Data Flow Diagram**

Once the data is entered into the Excel spreadsheet the shipment plan is calculated. The expert system is then run to calculate the proper change to the target inventory (in a similar manner to that for the safety inventory).

As in the safety inventory subsystem, Excel is used for the user interface. Colour coding readily identifies user entry and calculated cells. The spreadsheet has a button that executes the expert system. Once the user has entered all of the data into the spreadsheet, they push the expert system button and the plan is updated. The average length of time to calculate a shipment plan is 15 seconds, as compared to the 1 to 2 hours required by the manual method.

As can be seen from the above, the customer desired a very simple system. They did not want to interact with an expert system if at all possible. Significant effort was required to limit the expert system interface to simply a button on an Excel spreadsheet. As stated earlier, the reasons for using MetaCOOP were that its object oriented approach provided a superior development environment and the source code could be modified to provide the minimal user interaction that had been specified by the customer.

## 2.4 MetaCOOP Enhancements

MetaCOOP is an object oriented expert system shell that was developed by the Intelligence Engineering Lab at the University of Alberta. The original program was a full featured windowing program run on a Sun Workstation. This had been ported to the PC environment to produce a non-windowing version (PCMetaCOOP). As part of this project, this non-windowing version was modified to be Microsoft Windows compliant with support for Dynamic Data Exchange (DDE). The modified version of MetaCOOP (WMetaCOOP) met the following user/developer considerations:

**Operates under Microsoft Windows as a Windows program.** This was required so that the system would be consistent with other programs operated by the user.

**Automatically executes a knowledge base (reasons) when run.** The user did not wish to learn another package and therefore running the expert system had to be transparent to the user.

**Provides no extraneous output to the user.** As above, the user was not interested in the details of how a result was obtained. The expert system had to be transparent.

**Supports DDE (both send and receive).** As the user was experienced with Microsoft Excel this was chosen as the interface. DDE was chosen as the most practical method to interface with the expert system.

**The source code to the base MetaCOOP program contains minimal changes.** Other students in the Intelligence Engineering Lab planned enhancements for Meta-COOP. By minimizing the changes made at this time, their work was eased.

**Provides an improved developer interface.** After the development of this project was complete, plant personnel would have to maintain the knowledge base. The existing developer interface was very difficult to use and required improvement for a production environment.

The MetaCOOP program has three main functions: compiling, reasoning, and executing the user's methods. From an I/O standpoint, the compiling function consists solely of output while the reasoning and methods functions incorporates both input and output. From a user interface standpoint the data flow within the program is as shown below:

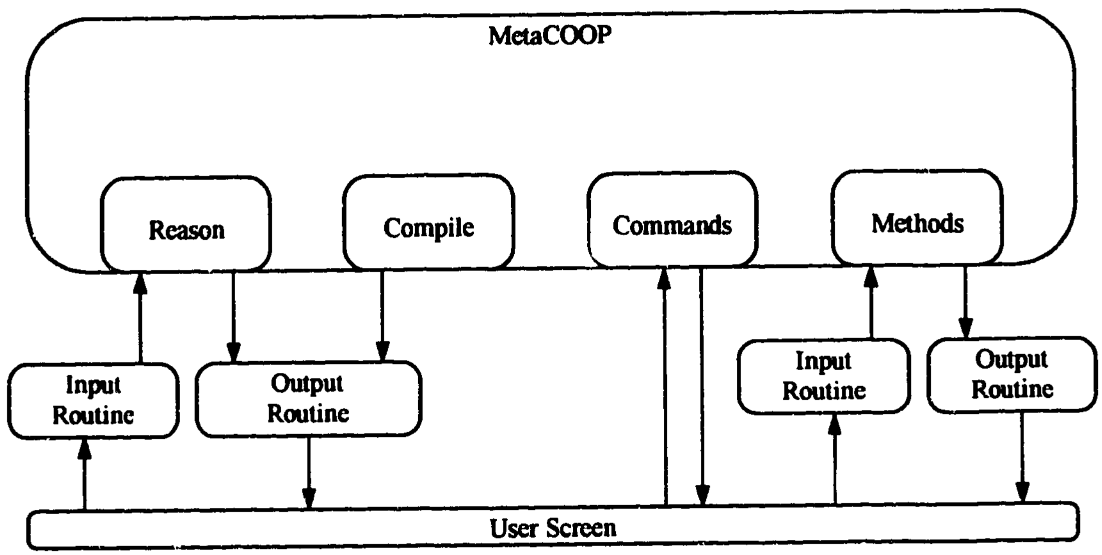


**Figure 2.11: PCMetaCOOP User Interface Data Flow**

The compile and reason functions use a common output routine while all other I/O is through dedicated routines. A Windows wrapper was designed that intercepts the output from MetaCOOP, decides what function is active at the time and then sends that output to one of three windows (user, reason, or compile windows). Input to the reasoning function is trapped and redirected to a dialogue box. Input to the method is not supported. This implemented the required Windows interface without major modifications to the main program. When the main MetaCOOP program is updated, less than 20 lines of code need to be changed in order to restore Windows functionality.

One of the advantages of the new Windows interface is that it provides a 250 line display which the user can scroll to view messages. One of the problems with PCMetaCOOP is that compile messages quickly scroll off the screen. This is no longer a problem with WMetaCOOP.

The data flow with the windows wrapper is shown below

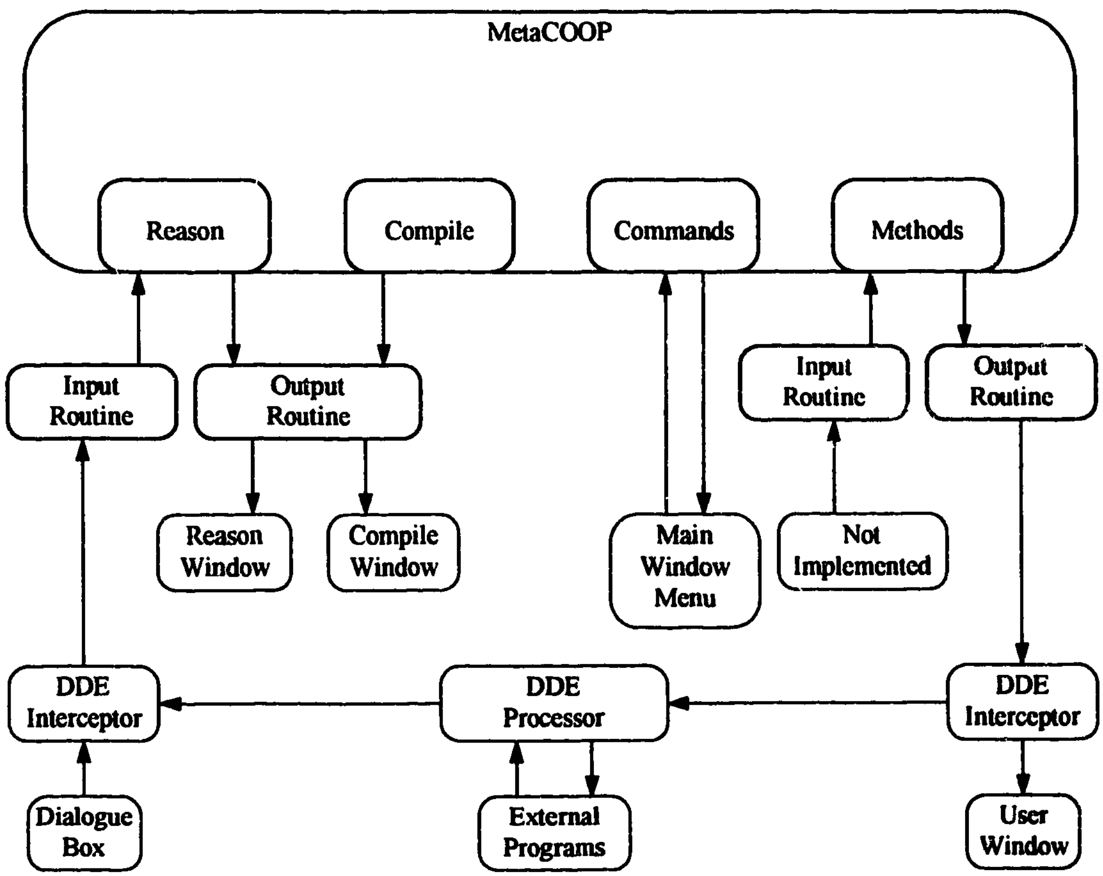


**Figure 2.12: WMetaCOOP User Interface Data Flow**

As can be seen, the input and output routines trap DDE requests and redirect those to the DDE processor sub-function. The interception is based on finding the string "DDE" in the output message or input prompt. Although not the most elegant method, it was the only way to meet the constraint of minimal changes to the base MetaCOOP program.

So that the user does not become burdened with extraneous messages, all but the user window can be disabled. Again, this was a marked change from the original development environment which output numerous messages for every rule that fired. For a new user it was nearly impossible to determine which were extraneous messages and which were important. WMetaCOOP can be configured to discard the extraneous messages and only display those that are required by the user. From a user perspective, WMetaCOOP was found to be much easier to use.

The PCMetaCOOP program is designed for developer use only. It expects a user to load the desired file, compile it, and give it the reason command to execute the knowledge base. This process is not suitable for an end user. WMetaCOOP allows a file name to be entered on the command line. If the file is a source code (as indicated by the extension FRA) then that file is loaded and WMetaCOOP waits for developer commands. If the file is a compiled knowledge base (as indicated by the extension CBK) WMetaCOOP starts up, loads the file, performs the reason command, and then exits. This allows WMetaCOOP to run in a background mode without user intervention.

PCMetaCOOP was written in standard Borland C. The enhancements were written in Borland C++ version 3.1. Some complexity was added by interfacing C to C++, however the advantages of object oriented programming outweighed the difficulties.

The change of PCMetaCOOP to WMetaCOOP was a challenging but rewarding programming endeavour. It required much learning, not only in how expert systems are developed, but also in how to write C++ based Windows compliant programs. A great deal of planning was required to determine the optimum methods for enhancing what existed with PCMetaCOOP to provide a package that users with lower skills in expert systems could use.

All design requirements of the upgrade were met.

## 2.5 Project Development

The development of this project followed the five stages of expert system development[8]:

1. **Identification Stage.** The characteristic of the problem was identified with the assistance of the department manager. In this stage, the various functions of the expert were studied and the goals of the project were developed. It was here that it was decided to provide a small example solution to prove the technology to the business leaders prior to developing a full scale implementation.

2. **Conceptualization Stage.** The concepts and relationships were developed. This was especially important in the Inventory Replenishment project. In this stage the actual data flows were investigated. Although Figures 2.1 and 2.10 appear simple, it required much analysis to arrive at that data flow as being representative of the manual system.

3. **Formalization Stage.** The data structures and knowledge base were defined and data flow diagrams prepared prior to any programming.

4. **Implementation Stage.** Implementation is basically complete for the Safety Inventory, partially complete for the Inventory Replenishment and has not yet begun for the Shipment Tracking portions of the project.

[8] Rao, Ming and Qiu Hainming, Process Control Engineering, Gordon and Breach Langhorne PA., 1993, p. 370

5. **Evaluation Stage.** Testing and modification has been conducted on simulated data.

As part of the development process, requirements for the user interface were defined. The plant currently uses Microsoft Excel, and wished to continue to use that as the primary interface. We were not familiar with any expert system shells interfaced to Excel in the manner that we required. As we had the source code for MetaCOOP we modified it to meet the project's specifications.

## 2.6 Summary of CRP Project

The CRP project accomplished the following goals.

1. The implementation of expert systems in pulp operations were proven to be a justifiable expense. It has eliminated a barrier to approaching CRP as part of a team assignment. The technician who operates the CRP out of Grande Prairie is very knowledgeable. Despite the high level of multiskilling inherent in Grande Prairie's team system, other members of her team are not sufficiently skilled to operate CRP without the expert system. Management at the plant was just as impressed by the expert system's ability to remove this barrier, as they were by its capacity to help manage the CRP.

2. The application of standard control algorithms to inventory management was shown to be a valid approach. This approach has greatly eased the process of calculating shipments.

3. The MetaCOOP program was extensively modified to improve its user interface. The specifications used in modifying MetaCOOP have been incorporated in later versions of the program. As a side benefit I learned much about Windows and C++ programming.

There remains much work to be done to fully complete this project, but the basic frame work is in place. The objective of proving the benefit of AI to typical pulp mill problems was achieved. Progress on the project has been delayed due to the death of the manager who was championing the process across the pulp division.

## Future Research

A major opportunity for future research is the application of modern control theory to the field of inventory replenishment. This in conjunction with a complete rule base would deliver an extremely capable inventory management system.

# Chapter 3

# Modelling with Neural Networks

## 3.0 Overview

Models are important to the efficient operation of complex chemical processes, such as those found in a pulp mill. They are used in model-based control systems and decision support systems. They are also important to various empirical process optimization studies that are used to develop a better understanding of pulp processes[9].

This chapter discusses process models that predict or simulate pulp and paper process behaviour over a finite time span. The purpose of discussing these models is to illustrate how a network based model can overcome some of the problems that are encountered in specific situations when using traditional models. These case studies also illustrate that network models are not a panacea, because as will be shown, the models are not perfect and their predictions require significant analysis. Although the models may be imperfect, the reader should gain an appreciation for the capabilities of two more tools that can aid them in their modelling efforts.

---

[9] Seborg, Dale E.; Edgar, Thomas F.; Mellichamp, Duncan A., Process Dynamics and Control, John Wiley and Sons, New York, 1989, p. 164

The chapter is broken into three parts:

Part 1 introduces the concept of modelling in general (as it is practised today). Regression analysis is the most common modelling tool in use today. The popularity of regression modelling is increasing as many computer programs are now available to do most of the work in building this type of model.

The discussion of regression analysis is followed by a review of neural networks, specifically, the back propagation network (BPN). Neural networks provide the ability to model processes that were heretofore difficult or impossible to model using parametric models such as regression analysis. One of the major barriers to their acceptance has been the "black box" stigma associated with neural networks. By showing how neural networks can solve heretofore impossible modelling challenges, and provide substantial process insight, some of the mystique will be removed from this methodology.

Part 2 describes applications of network modelling using problems from the Weyerhaeuser pulp mill at Grande Prairie. An overview of the pulp mill operation lays the groundwork for understanding the two examples of modelling that will be presented. In the first example, models that relate basis weight valve setting with production rate are illustrated. Both regression analysis and neural networks models are developed so a comparison can be drawn between the results of the two methods. The second example is a neural network that relates finished product pulp strength to various mill operating parameters. This illustrates an instance where a neural network successfully modelled a process that had proven impossible to model using traditional techniques.

Part 3 revisits the basis weight model and uses a new tool called Adaptive Logic Networks. ALNs are a hybrid of network and parametric models. Some of the short comings of neural networks are discussed in this section. ALNs represent a significant new addition to the modeler's toolbox.

## 3.1 Modelling Theory Overview

The simulation of a process or activity in order to predict future responses is applied to almost every field of human endeavour from chemical processes to psychological studies. Simulation aids in better understanding the process, and, in the case of chemical processes, permits the user to effect short and/or long term optimizations. These optimizations can be achieved by implementing[10]:

- model-based controls.
- model-based decision making.
- model-based process studies.

**Model-based control** systems are those automated process control systems which incorporate models in order to provide improved process control with minimal human intervention. The most common type of such system is inferential control where the control system adjusts control elements to achieve the desired process output based on the model's prediction of the response of the process to changing inputs or raw material feed. This is contrasted to the most common non-model control - feedback control - which adjusts inputs based on measurements of the process output. Feedback control requires deviations in the output to occur before remedial action is initiated, while the model-based inferential control takes proactive action, based on the

[10] Seborg, Dale E.; Edgar, Thomas F.; Mellichamp, Duncan A., Process Dynamics and Control, John Wiley and Sons, New York, 1989, p. 10

model's prediction that some corrective action is required to respond to a fluctuation in the process inputs.

**Model-based decision making** provides decision support data to operators in situations where manual control is required. A common application of this type is called a "soft sensor". The name derives from the fact that the software (model) simulates sensor output to the user. An actual hardware sensor may not be viable for a variety of reasons, such as cost, reliability or technology limitation. This situation is very common as there are many circumstances where current sensor technology is inadequate and process control relies on manual testing. Examples are K# testing, pulp strength tests, viscosity testing, etc. Part of the research discussed in this paper relates to a model designed to provide pulp strength information that reduces operations requirement for a manual test which requires one day to perform. A twenty four hour sampling time is obviously not conducive to adequate control.

**Model-based process studies** use models to improve our knowledge of a process. Many processes are well understood when operating in a very narrow range of "normal" conditions. For large processes, the expense and complexity of changing operating conditions for extensive testing outside this narrow range is prohibitive. In these circumstances, a model can be developed to simulate the process and process studies are directed against it. Such process studies are also useful where a process operates correctly but is not fully understood. In this case, a model is constructed that incorporates known and assumed relationships. If the model correctly simulates the observed process behaviour then the relationships are assumed to be correct and knowledge has been gained about the process. Because of the empirical nature of the model the relationships can only be assumed true for those ranges over which the model has been tested. A common mistake of the modeler is to extrapolate (or interpolate) the use of the model for assumed relationships outside the test range. This can result in significant process upsets and/or questions about the validity of the model.

There is a difference between models that are based on theoretical knowledge (mass/energy balances, etc.) and those that are empirical in nature. Although the first two categories of models (model-based control and decision making) can either be theoretical or empirical, models in the last category are typically empirical. A theoretical model is preferred over an empirical model as it will generally be more accurate. In many cases however, a perfect theoretical model cannot be built due to inefficiencies in the process, imperfectly understood chemical reactions, or insufficient measurements. In these cases, the modeler must resort to an empirical model. The remainder of this chapter will deal with empirical models and how best to represent mathematically a data set that is obtained from some process.

For models in the third category (model-based process studies), there will be a focus on obtaining a relatively simple mathematical relationship that can be analyzed to provide information on the process itself. This will require modelling techniques that provide readily understandable mathematical forms. For models in the first two categories (model-based decision making and model-based control), we will focus on obtaining mathematical forms that can be implemented in a computer. Extracting an understandable mathematical form will be of less importance. The most important aspect of these models will be to accurately simulate process behaviour for some range of input data.

Three types of empirical modelling tools will be discussed. Regression models and those based on adaptive logic networks provide readily understood mathematical relationships. Neural network modelling techniques provide models that can be easily implemented in computer systems and in some cases are better able to emulate the behaviour of a process than other modelling techniques. For each modelling tool, an example will be given that illustrates the strength or weakness of the technique.

### 3.1.1 Regression Analysis

A model provides a relationship between independent variables (inputs) and dependent variables (outputs) in some mathematical form. Regression analysis is a technique often used to find the best relationship[11]. The modeler must chose the mathematical form, although the simple linear form is most common. If the linear form is chosen, then the regression model is referred to as linear regression. Multiple regression and non-linear regression models are used as well. Regression models provide very good performance and can provide a very accurate model if the problem is relatively simple

[11] Wittnik, Dick R., The Application of Regression Analysis, Allyn and Bacon, Inc. Needham Heights, Ma., 1988, p. 2

The simplicity and accuracy of the approach make regression analysis to be the preferred choice for process models. In many cases however, this modelling technique fails to provide an adequate model.

The failure of the technique can be due to correlation between the "independent" variables, more variables than can be easily handled by the modeller, or difficulty in obtaining an assumption about the base mathematical form of the relationship being modelled. The latter problem is due to the trial and error nature of regression analysis. The typical procedure is to assume the simplest reasonable mathematical form and then use increasingly complex forms until a model of acceptable accuracy is obtained.

The measure of accuracy normally chosen is the square of the error (summed over the useful range of the model). This choice eliminates the sign from the error measurement and penalizes a model for having large deviations from the actual data. A model with a few large erroneous values will generally be considered worse than a model with many small errors, as the modeler can then be assured that no single modelled value deviates significantly from actual value. Although use of the least square of the error is typical, the choice of performance measurement is always problem dependent and in some cases modeler dependent. For example mean absolute error (the sum of the absolute values as opposed to the squares) puts less penalty on large errors. This measure is used in control systems where the squared error measure "can lead to unacceptably large accelerations and jerks"[12]. In this thesis the square of the error is used as the performance measurement.

### 3.1.2 Neural Networks

Neural network research has its roots in the 1940's era research into how people learn. Several works of this era stand out in the field of Neurocomputing. In 1943 Warren McCulloch and Walter Pitts[13] showed that any arithmetic or logical function could be computed by a simple neural network[14]. Other research of the time proposed the study of computers that simulated brain structure. The most notable was found in the papers and books by Norbert Wiener and John von Neuman[15]. Probably the most significant contribution was by Donald Hebb[16] who, in 1949, proposed that learning occurred at the neuron level and went on to describe a learning law for the synapses of the neuron. This was important as it provided an example of a learning method that researchers could implement when developing computer programs that could "learn". These and other works in the late 40's and early 50's led to the 1957 development of

---

[12] Hecht-Nielson,R., Neuralcomputing, Addison-Wesley, New York, 1989, p. 114

[13] McCulloch, W.S. and Pitts, W., *A logical calculus of the ideas immanent in nervous activity*, Bulletin of Math. Bio., 5, 1943, pp. 115-133

[14] Rarisini, T. and Zoppoli, R., *Neural Networks for Feedback Feedforward Nonlinear Control Systems*, IEEE Transactions on Neural Networks, IEEE Neural Networks Council, May 1994, p. 447

[15] Hecht-Nielson,R., Neuralcomputing, p. 15

[16] Hebb,D., The Organization of Behaviour, Wiley, New York, 1949

the Mark I Perceptron by Frank Rosenblatt, Charles Wightman, and others. The results of this research became part of the then new field of artificial intelligence.

As time progressed, a great deal of "hype" was generated by the neural network community. "For example, there were widely publicized predictions that artificial brains were just a few years away from development, and other incredible statements."[17]. Because of the wide divergence between promise and reality, neural network research declined and was mostly ignored, especially by agencies responsible for funding research. Although research in neural networks was already declining, Marvin Minsky and Seymor Papert are usually credited with causing this dormant period in neural network research[18]. In the last decade, the capability of the networks has begun to catch up with the previous hype and the field is once again flourishing.

---

[17] Hecht-Nielson,R., Neuralcomputing, p. 16

[18] Kempka, Anthony A., *Activating Neural Networks: Part I*, AI Expert, Miller Freeman, San Francisco, June 1994, p. 33

### 3.1.3 Neural Network Training

The key feature of neural networks is their ability to "learn" from data that is presented to them. The most common type of network used in modelling is a feedforward model called the Back Propagation Network (BPN). Building a BPN consists of the following steps:

- chose a topology i.e., number of input, output and hidden nodes, and number of hidden layers
- initialize the weights to random values
- input a data vector (exemplar) to the network and calculate the network's output
- compare the network output to the desired output and determine the error
- propagate this error back through the network (from output layer to input layer), adjusting the weights to reduce the error. This step is the source of the name for this network.
- continue steps 3 - 5 until a minimum or satisfactory level of error is found

This type of network requires that a training set of data be available, i.e., a set of data for which the desired output is known. This style of learning is called supervised learning. There are networks that can learn in an unsupervised mode. However they are more suited to classification problems and are not typically used in modelling.

### 3.1.4 Back Propagation Network Learning

The advantage of neural nctworks is their ability to learn functions based on a set of exemplars. Learning occurs when the weights within the network are changed to reduce the error between the network output and the measured data. Initially BPN's were trained using a gradient descent methodology called the Delta Learning Rule which was first used by Widrow and Hoff[19].

Gradient descent is best understood by visualizing an error surface composed of the total error of the network for each set of weights in the network (because network error is a function of the network weights). There is at least one point on this surface which represents the set of weights for which the error is a minimum. Assuming an initial random set of weights, gradient descent is the changing of the weights in the network to step down this surface until a minimum is found. The size of the step taken is given by the learning rate n.

### 3.1.5 Modified Learning Rules

The gradient descent based Delta Learning Rule has several undesirable features. For small values of n, finding the minimum on the error surface can involve a significant amount of time. As well, the rule has no mechanism to reject local minima. Modification of learning rules to speed learning and more effectively handle local minima have been major research areas for neural networks in the past.

---

[19] Rao, Valluru B. and Rao, Hayagriva V., C++ Neural Networks and Fuzzy Logic, Management Information Source, Inc., New York, 1993, p. 96

Of the many learning rules developed over the years, two modifications to the Delta Learning Rule have been particularly successful. These are the Delta Bar Delta (DBD) rule and the Extended Delta Bar Delta (EDBD) rule.

## Delta Bar Delta Learning

The Delta Bar Delta rule was developed by Jacobs[20] to improve the speed at which a minimum is located. It uses the heuristic that if weight changes are of the same sign for several iterations then the learning rate can be increased. However, if the weight changes alternate signs for several iterations, then a minimum is being straddled and the learning rate should be reduced. Increases in learning rate are constant while decreases are geometric. This helps to keep the learning rate from becoming too large too fast. Decreasing geometrically ensures that the learning rates do not become negative[21].

---

[20] Jacobs, R.A., *Increased Rates of Convergence Through Learning Rate Adaption*, Neural Networks, Volume 1, 1988, pp. 295-307

[21] NeuralWare Inc., Neurocomputing, a Technology Handbook for Professional II/+ and NeuroWorks Explorer, NeuralWare Inc., Pittsburg Pa., 1993, p. NC-131

## Extended Delta Bar Delta Learning

This modification to Delta learning was developed by Minai and Williams[22] to incorporate heuristics not found in the Delta Bar Delta learning rule. The major change is the addition of a momentum term. This momentum term is generated by adding an amount proportional to the last weight change to the current weight change. If the Delta rule continually causes a weight change in one direction, the momentum term will gradually increase the weight change and the learning will increase in speed.

The analogy of a ball picking up speed (or momentum) as it rolls down a hill is very appropriate. This means for long gradual gradients in the error surface, the learning will generally be shorter than DBD learning. Of course, the "down" side of momentum is when the minimum is found, the weight changes will continue due to the momentum effect and the weight will overshoot the minimum. This is not necessarily bad. If the minimum is only a local minimum, then the weights may "roll up" out of the valley and on to the next minimum.

Decreasing learning rates occur in a similar manner to DBD learning. Increased learning rates are a geometrically decreasing function of the weighted gradient components. This means that changes in learning will be greatest in areas of high curvature. A cap is placed on the learning rate and momentum.

--- ,linai, A.A. and Williams,R.D., *Acceleration of Back-Propagation through Learning Rate and Moment[illegible]daption*, Inernational Joint Conference on Neural Networks, Volume 1, January 1990, pp. 6[illegible]-679

## Network Configuration with DBD and EDBD learning

Both DBD and EDBD modify the basic Backpropagation approach to the nodes in the neural network. With Backpropagation, the learning rates are established for the entire network. With DBD, each node has it's own learning rate. EDBD extends this so that each node also has its own momentum term. As well, EDBD incorporates a memory. The weights for the current iteration (and the associated error) are saved. If on the next iteration, a larger error is generated, the previous weight values are restored and the learning rate and momentum reduced.

This gives a brief overview of neural networks. In practice many commercial packages are available so that minimal familiarity with the learning rules is sufficient. This allows the user to focus on simulating a process rather than writing network software. Even though commercial packages are available it is important to understand the basic principles underlying these packages. Without this understanding, choosing the proper network configuration becomes a purely trial and error process which can be very time consuming.

## 3.2 Pulp Mill Applications

Modelling is an important tool for process engineers. To illustrate the types of modelling encountered in general, and the application of neural networks in particular, two case studies are investigated:

1. a feedforward production rate control scheme, and

2. a pulp strength study.

In order to assist the reader to better understand the models, a brief review of a pulp mill operation is included.

Pulp mill operations are often described in terms of major functional blocks as shown in Figure 3.1.

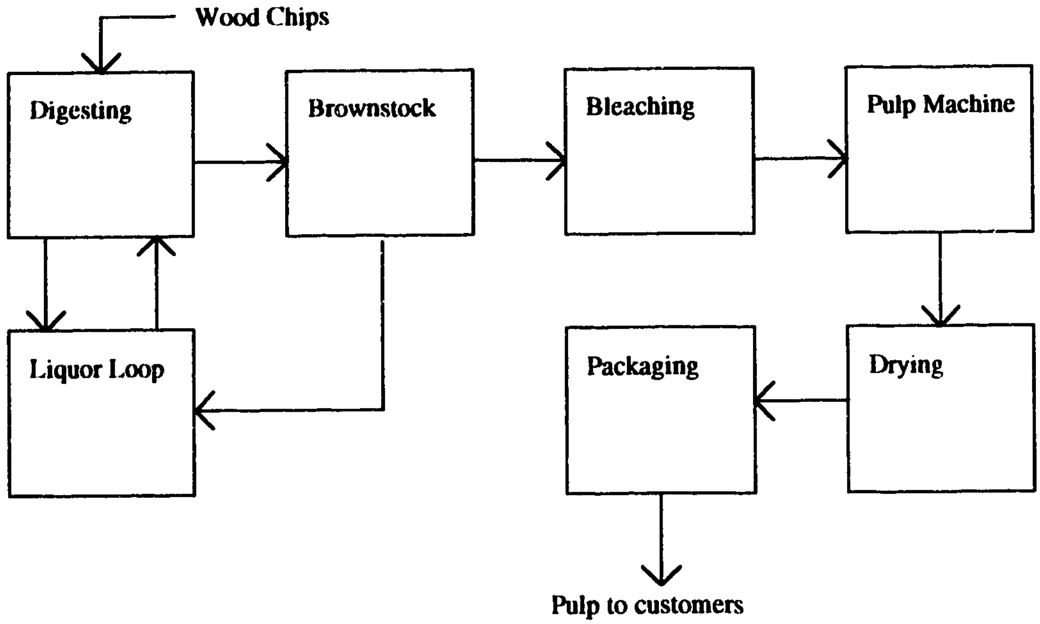


**Figure 3.1: Pulp Mill Operation**

## Wood Chips:

Wood is made up of roughly four materials (each of equal volume): cellulose fibre, lignin, and water, and air. The objective of the Kraft process is to separate the cellulose fibres from each other and from the other three constituents with minimal damage to the cellulose fibres. "Kraft" refers to the type of chemical processing used. The word is derived from the German word for strong, so named because the resulting pulp is stronger than that produced by other processes. Such other processes as Thermal Mechanical Pulping (TMP) or Chemi-Thermal Mechanical Pulping (CTMP) separate the cellulose fibres from each other by grinding the wood chips into a fibrous state. This causes major fibre damage in the process. Kraft pulping retains fibre integrity and strength.

Cellulose fibre is a component of all plantlife. The most common example of raw cellulose is cotton batting which is very similar in appearance to fully bleached woodpulp. Lignin is a complex organic binder which holds the cellulose fibres together. An example of lignin is sap.

## Digesting

Digesting combines woodchips and a caustic (Sodium Hydroxide) "white liquor" at a high temperature and pressure for a specific time period to extract the lignin from the cellulose fibre. The industry refers to this as cooking. The correct combination of chemical strength, temperature, pressure and time will provide a very strong pulp of low lignin content. If some component of the combination is too low, the wood will not "cook" completely. This will result in excess lignin being left in the cellulose causing increased usage of expensive chemicals in the bleaching stage. If some component is too high, the cellulose fibres will be degraded, and a weaker pulp will result. One test for how well the pulp is cooked is called the $K^{\#}$ test, so named because it consists of chemically preparing the pulp sample and then titrating the sample with a potassium compound. The amount of potassium used provides a numeric indication of cooking level, the higher the number, the less cooking has occurred.

## Brownstock

Digesting produces a mixture of cellulose fibre and spent (black) liquor. Brownstock is the process where the spent liquor is separated from the cellulose. This is done through various devices that hold the fibre against a screen while the liquor is forced out of the pulp by forcing wash liquor through the pulp mat. Throughout the process, as high strength liquor is removed from the pulp, it is washed with lower concentration liquor until the final stage where nearly pure water is used.

## Liquor Loop

The spent cooking liquor recovered from brownstock contains a significant amount of organics (lignin). The main process in the liquor loop is the recovery boiler where the combustion of the organics yields a chemical residue that can be converted back into white liquor. The steam created in the boiler is used to generate power and heat for the rest of the process. The liquor loop makes the Kraft process very economical in terms of chemical and energy usage.

## Bleaching

The pulp that leaves brownstock is brown in colour, about the same as brown wrapping paper. The goal of bleaching is to whiten the pulp so that it can be used in writing papers. This whitening is accomplished by a variety of chemicals. Although Kraft identifies a common liquor loop and digesting process, there are many bleaching processes in use today. Chemicals used in these processes are $Cl_2$, $ClO_2$, $H_2O_2$, $O_2$, $O_3$, and others. Bleaching processes are rapidly changing at the current time due to environmental concerns. For example, chlorine ($Cl_2$), the most common bleaching chemical 15 years ago, is now absent from most bleach plants because of the tendency of the chlorine to bind with harmless dioxins to form chlorinated dioxins, some of which are extremely toxic. Chlorine dioxide ($ClO_2$) does not produce chlorinated dioxins to the same degree and is currently the bleaching chemical of choice for many plants. The two main bleaching chemicals at the Grande Prairie plant are chlorine dioxide and hydrogen peroxide ($H_2O_2$).

## Pulp Machine and Drying

Once the pulp leaves bleaching it is chemically ready to be used by the customer - usually a paper machine of some sort. As the pulp that leaves the bleach plant is in a slurry of 88% water, it must be concentrated prior to shipping. This is the job of the pulp machine. The pulp machine forms a sheet on a moving screen. This sheet is then pressed to remove as much water as possible. At this point the sheet is about 52% water. The remainder is removed by thermal drying until a final target of 10% water by weight is achieved.

### Packaging

Like all processes previously mentioned, the pulp machine is continuous, so the packaging department receives an endless sheet 226 inches wide travelling at 600 fpm. This is converted into packages of pulp approximately 32"x32"x16" which are loaded into boxcars and shipped to the customer. Each package contains about 600 pounds of pulp. Customers typically purchase pulp in boxcar loads, which at current market prices, costs in excess of $50,000/boxcar. In order to ensure that the pulp purchased is of specified quality, several performance tests are performed on the finished product. These include moisture content, dirt content, brightness, viscosity and strength.

## 3.2.1 Basis Weight Valve Model

One of the greatest difficulties facing process control engineers is adequately handling those processes with long deadtimes, i.e., the time from when a change is made at the input of a process until a resulting change is seen at the output. The pulp machine is an area where deadtime can be as much as 15 minutes and hence presents a significant challenge to process control.

The most common type of control is feedback control, which consists of taking corrective action when an error is observed at the output. With long deadtime processes this becomes impractical. For example, if the control system is designed to observe the outputs of the pulp machine and take corrective actions once an error is observed, the error would continue for 15 minutes before the corrective action could be observed at the output. A mechanism that responds to errors in the inputs, so that the output is maintained in control is much preferred. This is referred to as model reference feedforward control. The modelling tools described earlier provide a foundation for developing such a control scheme.

To illustrate how these models provide such foundation, an example relating to pulp machine throughput will be used. Throughput (measured in tonnes per day at the dryer output) is controlled by the stock flow rate through the basis weight valve to the headbox. Stock consistency prior to the basis weight valve is controlled to a nominal 3% consistency although this varies significantly. The stock flows through the headbox and on to the rest of the drying process, where machine throughput is measured by on-line sensors. Measured values prior to the headbox are consistency and valve position. For the problem we will be discussing, changes in dilution water flow do not have a major impact on machine production. The process is shown in Figure 3.2.

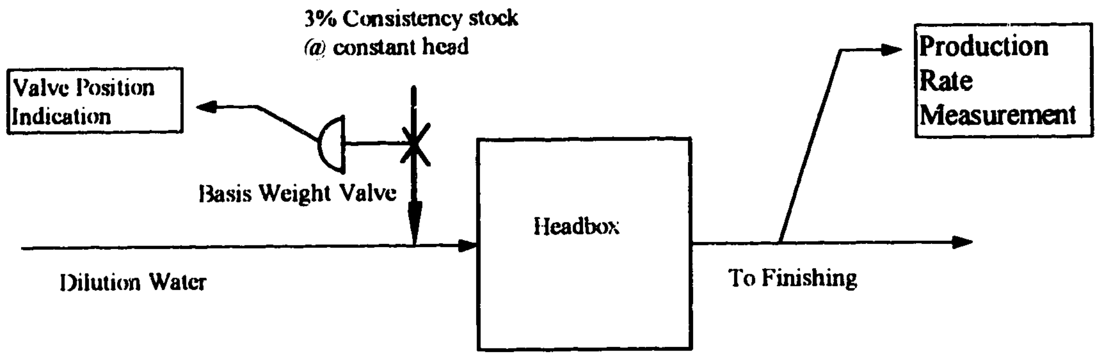


**Figure 3.2: Basis Weight Valve Diagram**

During operation, when a change in throughput is required, the operator manually opens (or closes) the basis weight valve slightly, while at the same time increasing (or decreasing) dilution water flow. This increases (or decreases) stock flow to the headbox while maintaining headbox consistency. The change in stock flow causes a change in machine throughput. To automate this process requires that given a target production rate, a basis weight valve setting can be determined. In order to determine this, a model is required that provides a relationship between consistency, valve position, and throughput. Once this model is available it can be incorporated into a feedforward scheme to reduce the impact of consistency variations.

One approach would be to use the valve curve to obtain the relationship directly. Unfortunately, the curve for this valve/positioner combination is not readily available. However the problem is simple enough that it provides a good test for the modelling techniques. Two approaches will be examined: regression analysis and a neural network model.

## Modelling with Regression Analysis

Because flow through a valve is usually governed by the square of the open area, it is assumed that flow onto the fourdrinier will have the form

throughput = a*cons(b*pos+c*pos$^2$)

a, b,and c are constants to be determined, cons is consistency, and pos is the basis weight valve position as read from the positioner. Note that the linear term for pos is included on the chance that the positioner or valve construction modifies the behaviour such that a linear relation results. Using regression techniques on basis weight and consistency data from the Grande Prairie operation for 1993, the values for the coefficients were found to be

a = .057

b = negligible

c = 1.83

This indicates that there is no actual linear component to the relationship although given the small range of valve position and slight curvature the relationship can be considered linear.

The scatter plot of predicted values versus actual values is shown in Figure 3.3.

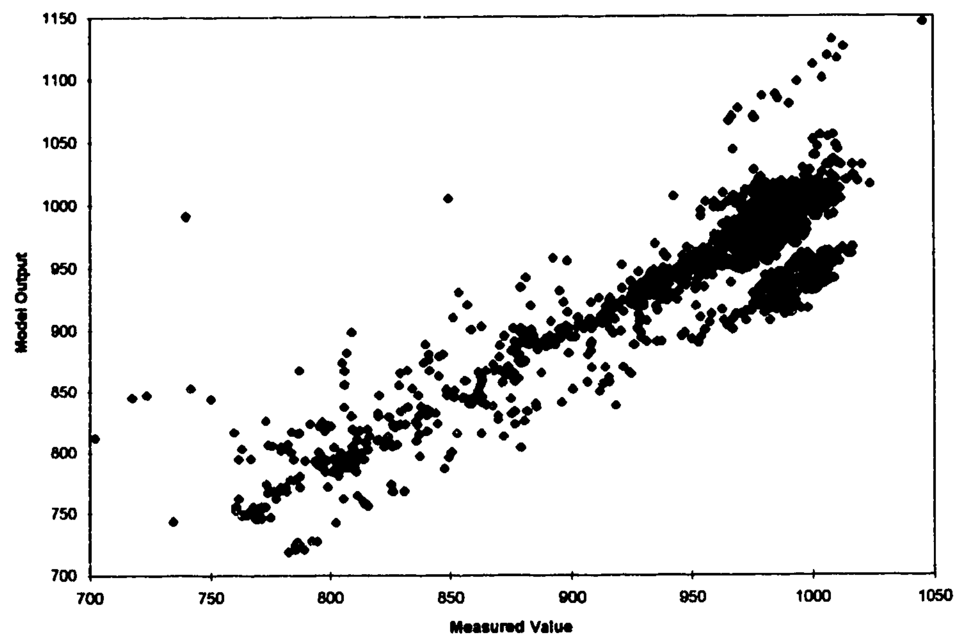


**Figure 3.3: Regression Model**

As can be seen, there is significant "noise" in the readings. The plot also indicates a "shadow" pattern that sits under the main pattern, indicating that there may be another factor in the flow. Possible factors would include changes in pressure (which was assumed to be constant), or modifications to the positioner/valve mechanism during the period of data collection.

At this point it would seem to be proper to perform some further analysis to determine the overall accuracy of this model. However, as can be seen from Figure 3.5 the regression model is too far from accurate to be of any value. This illustrates that given the wrong mathematical form, regression analysis can generate very poor "best" fits. As two other models will be developed in later sections, the trial and error process of finding the proper form for this model was not attempted.

## Modelling with Neural Networks

A neural network was built to compare the results of this type of model to the standard regression analysis. The network was built using 4 layers, 2 input nodes, 20 hidden nodes in the first hidden layer, 5 nodes in the second hidden layer and 1 linear output node. The learning mechanism was extended delta bar delta learning. The network was trained for 99,000 passes through the training data used in the regression analysis example.

This network topology was chosen as it was similar to that developed for the pulp strength model described later. The first network topology attempted was successful and no attempt was made to minimize the size of the network because, as will be seen later, a better model was developed using an alternate modelling method. As the purpose here was to illustrate the capability of networks to model a function it was felt that further work on refining the model would be superfluous.

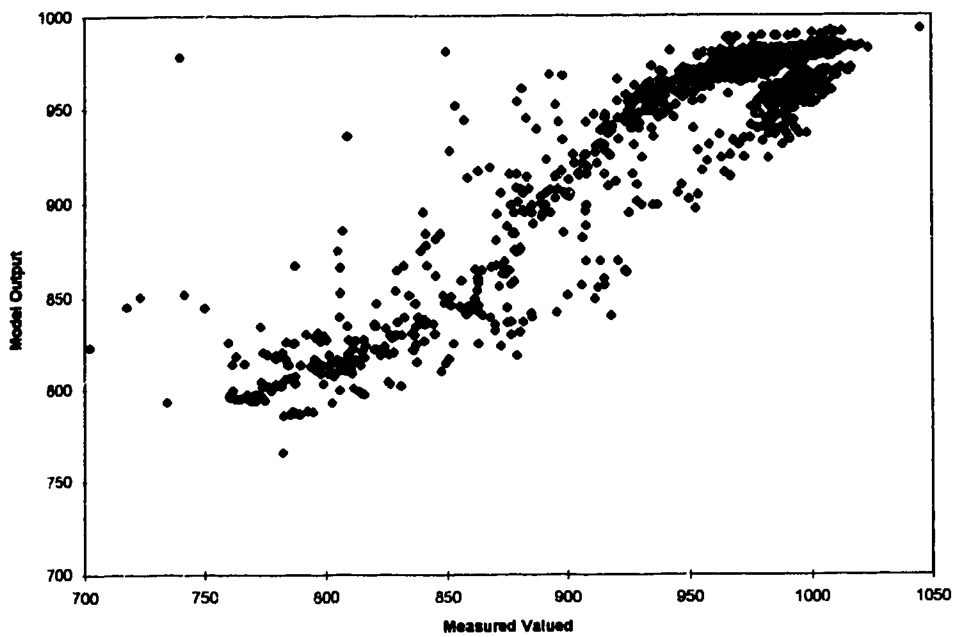


**Figure 3.4: Neural Network Model**

A comparison of the results of the regression analysis (shown in Figure 3.3) and the neural network model (shown in Figure 3.4)show the "squashing" of the neural network output at the upper and lower extremes of the data range. Methods to evercome this squashing effect will be discussed later. Despite this squashing, both models *appear* to yield a fair simulation of the output data. However, as previously discussed, Figure 3.5 shows that the regression model is not as robust as the neural network model.

## Data Analysis

If the analysis were left at this stage some significant factors could be overlooked. In truth, as will be illustrated, neither model can be verified as correct. To properly analyze the models, first, the input data must be scrutinized.

A visual examination of the distribution of the data points in the above figures show that the bulk of the data occurs around a throughput of 975 which is a typical operating rate for this plant. There is no indication in these figures of the distribution of the input data, i.e., consistency and valve position. As we will see, this distribution of data is important to understanding the capabilities of the models.

Table 3.1 shows this data distribution with each number in the table representing the number of samples at those conditions.

| | Pos | 47 | 48 | 49 | 50 | 51 | 52 | 53 | 54 | 55 | 56 | 57 | Total |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Cons** | | | | | | | | | | | | | |
| **2.40** | | | | 4 | | | | | | | | | 4 |
| **2.80** | | | | | | | | | 1 | | | | 2 |
| **2.90** | | | | | | | | 1 | 1 | | | | 2 |
| **3.00** | | | | 1 | | | 2 | 2 | | | | 1 | 6 |
| **3.10** | | 1 | 1 | 5 | 2 | 2 | 4 | 21 | 15 | 43 | | | 94 |
| **3.20** | | 84 | 135 | 290 | 255 | 264 | 461 | **1417** | **2122** | **1870** | 67 | | 6965 |
| **3.30** | | | | 2 | 1 | 1 | [illegible] | 1 | 13 | 19 | 39 | 19 | 96 |
| **3.40** | | | 1 | | | | | 2 | 1 | 1 | 3 | 6 | 14 |
| **3.50** | | 1 | | | | | | | | | | | 1 |
| **3.90** | | | | | | | | | | | | | |
| **Total** | | 86 | 137 | 302 | 258 | 267 | 469 | 1444 | 2153 | 1933 | 109 | 26 | [illegible]78 |

**Table 3.1: Data Distribution Valve Position vs. Consistency**

As can be seen, 75% of the data occurs when the range for position is from 53 to 55 and consistency is 3.2%. 97% of the input data occurs for a consistency of 3.2. Obviously, there is insufficient data to draw firm conclusions about model performance for other consistencies with any reasonable degree of confidence[23]. A comparison of model predicted output and actual values for a consistency of 3.2 is shown in Figure 3.5.

[23] Eberhardt, Keith R., *Survey Sampling Methods*, Handbook of Statistical Methods for Engineers and Scientists, Harrison M. Wadsworth (Editor), McGraw-Hill, New York, 1989, pp. 9.1-9.18

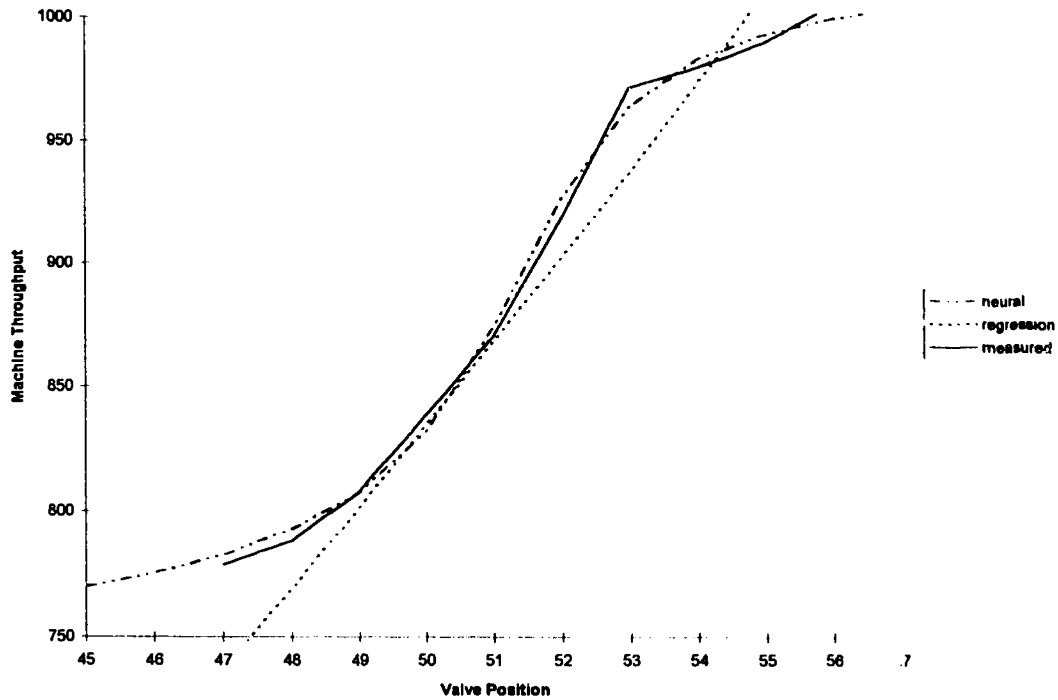


**Figure 3.5: Model Comparisons**

The actual output data clearly shows a break at a valve position of about 52. It appears that the "normal" operating range of 53-55 follows a different set of rules than the rest of the range. The neural network identifies this difference and models it correctly. The regression model is restricted in its form and therefore does not model the change.

The problem that the modeler faces at this point is that the pattern of the output data must be explained as the measured curve is counterintuitive. There are other factors (such as inconsistent operating strategies, maintenance changes, data collection problems etc.) that could cause this pattern and they should be investigated prior to using the neural network model. Clearly, the regression model is unsuitable as is. In this example, further tests can be conducted to verify the reasons for the pattern. However, in more complex processes, this may not be possible. In such cases, the neural network model may be used as is or as a basis for further research.

Which method of modelling is better? The regression models should be better if they are based on good a priori knowledge. In the above example, the knowledge used to build the regression model, was the same knowledge that was used to state that the input data was counterintuitive. Although the neural network does an excellent job of predicting the output data, because the output data is suspect, then so is the model. If good a priori knowledge exists, then a regression model should be possible. If this knowledge does not exist, then the neural network can give excellent simulation, but there is no easy method to verify that the model is doing anything other than replicating some set of previously seen output data. If replication is all that is required, then neural networks without further analysis are excellent tools. However, if confidence that the network is learning a function is required, then further testing is needed.

If the requirements of a problem are such that a neural network is feasible, then there is an added advantage in their use - the ease of on-line training. In cases where the proper output will be known at some time in the future, it can be backpropagated through the network, updating the weight vectors. This is very advantageous in those processes that change their characteristics over time, as happens in this basis weight valve problem. Over time, the positioning mechanism will degrade and periodically be repaired. As this occurs, the relationship between indicated valve position and throughput will change slightly. A relationship modelled using regression would have to be recalculated, whereas, a neural network could be configured to learn on-line and automatically adjust for any changes in the process.

### 3.2.2 Pulp Strength Model

The relatively simple basis weight problem described above demonstrates the complexities of modelling even a simple process with minimal variables. Within our pulp mill operations, there are very complex processes with multiple variables and even larger deadtimes. Modelling such processes with traditional tools such as regression analysis has proven to be ineffective. It is in these situations that neural network models will excel. One such process in the Grande Prairie operation is that which involves controlling the strength of our finished product.

Wood pulp is the raw material for many consumer products, such as tissues, paper towels, newspapers, sausage skins, cigarette paper, magazine stock, and diapers. Each product is distinguished by certain properties. For example, we expect that tissues will be soft and strong, paper towels will be absorbent when wet, and newspapers will be strong and stiff. One of the tools the papermaker uses in creating these properties is the choice of raw materials.

Woodpulp from northern latitudes tends to have long, stiff fibres which result in strong but hard paper, while pulp from southern areas tends to be softer and weaker. The papermaker creates a "pulp recipe" to obtain the desired base material characteristics for their product. The pulp produced at the Weyerhaeuser pulpmill in Grande Prairie is one of the premier high strength Northern Softwood Kraft (NSK) pulps.

Because strength is an important feature of the pulp produced there, operations personnel want to have a complete understanding of how all process variables affect the pulp strength. Unfortunately, the process is so complex, and so many factors influence pulp strength, that there is not a complete understanding of these factors.

Pulp strength (also known as tensile strength) is affected by several factors associated with the individual cellulose fibres, such as length, strength, damage, straightness, and inter-fibre bonding. Each of these fibre properties is impacted by the chemical processing that occurs throughout the process. Experts in the pulping process have determined 60 process parameters that can possibly impact pulp strength[24]. For twenty years, attempts have been made to build a model that correlated a subset of these parameters with the actual measured pulp strength. A subset was chosen because a model with 60 variables is difficult to build with traditional tools, and many of these variables are correlated. The process has been to try to eliminate the correlated variables and build a model with the half dozen or so variables that were left. No model using this process was satisfactory.

---

[24] Eberhardt, Keith R., *Survey Sampling Methods*, Handbook of Statistical Methods for Engineers and Scientists, Harrison M. Wadsworth (Editor), McGraw-Hill, New York, 1989, pp. 9.1-9.18

The basis weight valve model showed a comparison of regression analysis with a neural network model. Typically, a parametric model is preferred as it provides a good basic understanding of the process with a minimum of effort. In many cases however, regression or other techniques cannot provide an accurate model. One example is where there are many variables, some of which are correlated with each other to some degree. Pulp strength is one of these cases.

Because of the structure of neural networks, correlated variables do not pose a problem. In fact, as every input node is connected to every first layer hidden node, the network's assumption is that every variable is correlated. The size of the model also poses little problem to the modeler. The major impact model size creates is the need for larger amounts of data and longer computing time while training.

The value for a pulp strength model is two fold. First and foremost, the effect of the 60 variables on pulp strength currently is not well understood. With a model that accurately predicts pulp strength, investigations can be made to determine what the key variables are and how to better control them. Currently, if the pulp strength gets out of control, it is not clear which variables should be the immediate focus areas. The only alternative is to try to bring all 60 back into specifications. This is typically a difficult task. If operators knew which were the top ten variables to control pulp strength, response to strength variations would be quicker.

Second, pulp strength is a finished product test and is used to determine which customer gets which pulp. Each customer has specific product and product wrapping specifications. Typically, high strength pulp wrapped for one customer cannot be sent to another customer. Obviously low strength pulp wrapped for a high strength customer would not be suitable for any of the customers. The problem is that the pulp strength test is a manual test that requires up to a full day to complete. By the time that test results reveal that pulp strength has decreased unexpectedly, the chances are high that the pulp is already en route to the customer. This means that the shipment must either be recalled or diverted to another customer. In either case, a substantial amount of rework on the part of the shipping office and loss of face with the customer is incurred.

A pulp strength model that accurately predicts tensile properties of the finished product would allow improved decision making relative to both operation of the plant and wrapping and shipping procedures. A neural network was constructed to model pulp strength.

As stated previously, experts had identified 60 possible candidate variables. The mill information system retains hourly averages of this data and it was used in the model. Each hour has 60 input variables and one output variable. This represents one input vector or exemplar. The number of exemplars needed to train a network is dependent upon the number of weights that must be modified. The first step is to determine the number of weights in this network.

This network is configured as 60 inputs nodes, 1 hidden layer with 20 nodes, another hidden layer with 5 nodes and 1 node in the output layer. Currently there are inconsistent approaches to choosing network topology.

This choice of topology was based on the rule of thumb as taught by NeuralWare trainers (and others) to have a roughly pyramidal shaped network. As well, this form of network was similar to a network that performed reasonably well learning the Mexican hat function.

Prior to this project, a brief study was conducted on how a network's topology impacted it's ability to model functions. In this study the 2-dimensional Mexican hat function was used. This function is characterised by many inflection points. Although not an exhaustive study, about 20 different topologies were tried and it was found that those that incorporated a single hidden layer had difficulty learning more than 2 inflection points. Those that had two hidden layers were much better at learning the complexities of that function.

Results of the neural networks versus desired Mexican hat function values exhibited the same "squashing" as evident in the basis weight valve problem shown in Figure 3.4. As can be seen from that figure, the network predictions tend to "flatten out" at the extremes. This appears to be a common problem with the real valued (as opposed to classification oriented) neural networks. The problem is due to the use of a logistic function (which appropriately is referred to as a squashing function). Several methods are used to combat this:

- **Data scaling** of input and output values is often used to reduce the range of the data and therefore force it to the centre (and more linear) portion of the s shaped logistic.

- **Greatly increased learning times** are sometimes used to decrease squashing but has the disadvantage of often incurring overfitting problems.

- **Choice of logistic function** can be changed. The standard sigmoid function of $f(x)=1/(1-e^{-x})$ is often replaced with the hyperbolic tangent. The sigmoid varies from 0 to 1 while the hyperbolic tangent varies from -1 to +1. This seems to provide better performance.

- **Data elimination** can improve the performance. If the data set is normally distributed then most of the data will be near the mean value. There will be little outlying data for the neural network to train with. If there is more than enough data in the training set to train the network, one can remove some data points near the mean. This will force the neural network to spend a larger percentage of time learning the outliers. The goal should be to have a uniform or random distribution of data points as opposed to a normal distribution. That is, the data histogram should look like a box rather than a bell.

For the pulp strength model the last two methods were used to overcome the lack of performance at extreme values which was evident in early attempts at the model. As well the network topology was the result of the tried and true approach of trial and error. Over 100 networks were attempted over a period of two years with varying degrees of success. Another issue in choosing topology is ensuring that sufficient training exemplars exist.

For this network there are 60*20=1200 connections between the input layer and the first hidden layer, 20*5=100 connections between the hidden layers, and 5*1=5 connections between the second hidden layer and the output layer. In total there would be a total of 1,305 connections. The large number of hidden nodes (25) was chosen because, based on the difficulty others had encountered building pulp strength models, it seemed probable that the relationship of tensile to the input variables would be complex. The large number of hidden nodes should help in capturing that complexity. The next step was to determine how many exemplars are available for training.

A good rule of thumb is that two thirds of the available data should be used for training and the remainder for testing. The training set should be twice as large as the test set. Rounding to even numbers gives us a training set of 1,000 exemplars and a test set of 400. Although the modeler should obtain as much data as possible, there should be at least 1 exemplar per network connection. Fewer exemplars than this increases the risk that there will be insufficient data to properly train the network. Clearly our 1,305 connections are too many for the 1,000 exemplars that we have.

Fortunately, in this problem, we are fairly sure that the full 60 variables are not required. Which of the 60 are not significant is not known. If only 48 variables were used then there would only be 1,065 exemplars are needed, which is close to the 1,000 that are available. At this stage it was decided to build a model with 48 inputs using 1,000 exemplars. There was a risk that the model could not be trained but if that were to happen, more variables could be eliminated or a different network configuration chosen. In fact, this network was found to be trainable, although the final network configuration was modified slightly from this initial configuration in order to obtain a more accurate model.

The standard deviation, minimum, and maximum of the 60 variables that were identified as potential factors were calculated. During 1993, pulp strength had a standard deviation of 5%. Variables that had very little deviation, or extremely large deviations compared to this were eliminated. Although this is a fairly arbitrary method, the resulting variables that were dropped did not appear to be influential on pulp strength. The resulting 48 variables were used for the network model. The variables evaluated are listed in Table 3.2 . (Note: Table 3.2 has been removed in order to protect confidential information.)

The actual networks were built using a commercial tool (Professional II+ by Neural Ware[25]). This allowed quick development of the models and allowed investigations into the impact of different network configurations. After a capable model was produced several slightly different topologies were tried so as to fine tune the result. The best success was obtained with a network configured with 48 inputs, 1 output and 10 hidden nodes in the each of the two hidden layers. This only slightly reduced the

[25] NeuralWare Inc., Building IV, Suite 227 Penn Center West, Pittsburgh, Pennsylvania USA, 15276, (412) 787-8222

total number of hidden nodes, 20 as opposed to the original guess of 25 and so retained the networks ability to capture complex functions. However, moving some of the nodes from the first hidden layer to the second, reduced the total number of connections in the network to 680, thereby increasing the ratio of exemplars to connections. The topology of the network is illustrated in Figure 3.6.

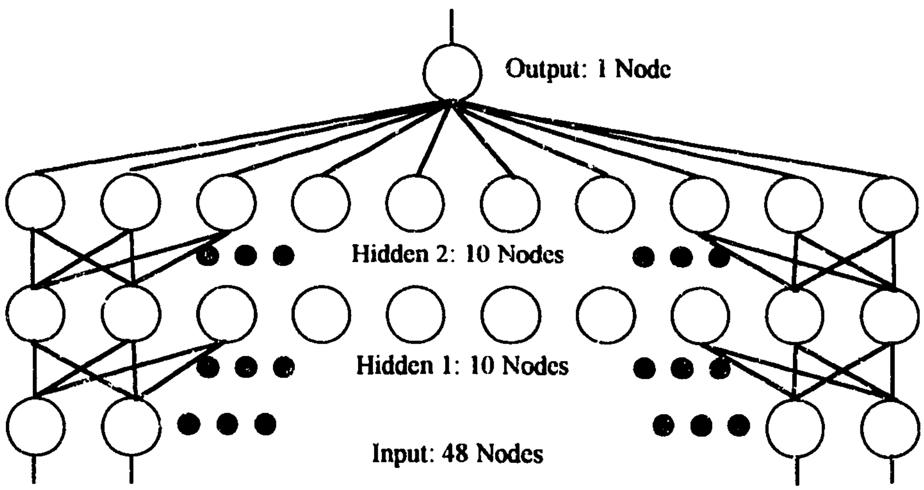


**Figure 3.6: Neural Network for Pulp Strength Model**

The plot of network prediction versus actual values is shown in Figure 3.7. As can be seen, good agreement is obtained by the network although significant noise exists. This network has been used to predict excursions of pulp strength from specification.

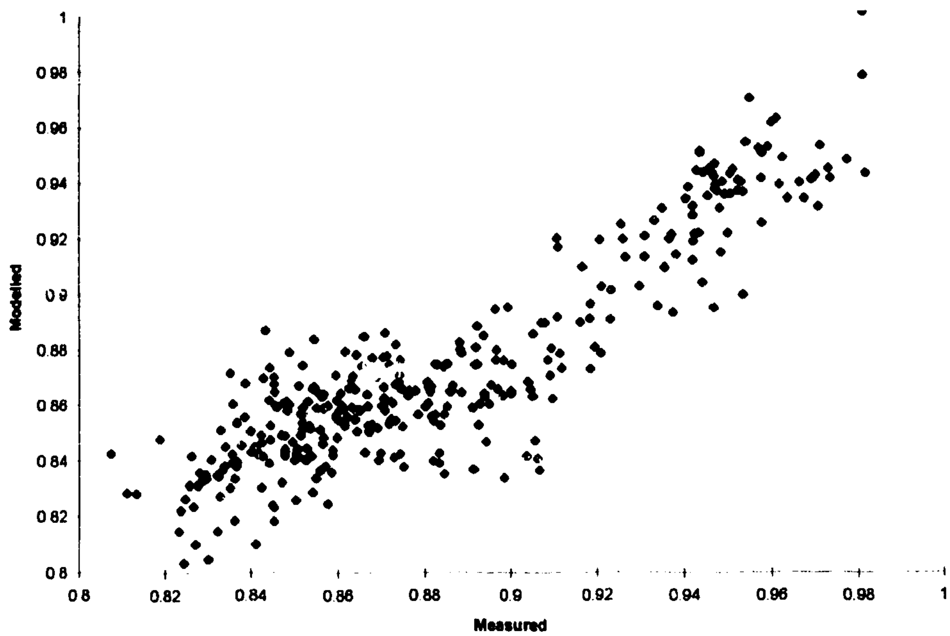


**Figure 3.7: NN Prediction versus Actual Pulp Strength (6pt mvg avg)**

The more formal methods of model quality evaluation and/or discrimination are not considered here. An analysis using the regression coefficients indicates that the model explains 72% of pulp strength variation when evaluated on any individual sample. Most of the remaining variability is due to the large variability inherent in the actual pulp strength test. To measure trends, the six point moving average is used. This represents a 24 hour average. In this case the model correlation is 90%.

This model more than satisfies the original objective of indicating trends in pulp strength. In use it has successfully provided that type of information. As an added benefit, the model was analyzed to determine key contributors to pulp strength variability. These key contributors and their relative impacts are shown in Table 3.2.

In order to maintain confidentiality of Weyerhaeuser data, in all but one case, the variable names have been replaced with the name of the area in which they reside. Brownstock pitch dispersant is named so that it can be used as an illustration of the type of analysis that can be carried out on the other variables.

| # | PIS | CV% | Rank | Impact | Description |
|---|---|---|---|---|---|
| | | | | | |
| 21 | 491 | 7.45 | 01 | 23.7 | Brownstock pitch dispersant |
| 26 | 900 | 8.52 | 02 | 23.0 | Machine variable 1 |
| 48 | 1734 | 27.7 | 03 | 20.7 | Brownstock variable 2 |
| 31 | 47 | 9.80 | 04 | 20.5 | Digesting variable 1 |
| 39 | 1558 | 18.54 | 05 | 19.4 | Bleaching variable 1 |
| 18 | 306 | 6.72 | 06 | 19.3 | Bleaching variable 2 |
| 49 | 474 | 29.44 | 07 | 19.0 | Brownstock variable 3 |
| 17 | 1618 | 6.62 | 08 | 18.9 | Brownstock variable 4 |
| 47 | 272 | 25.97 | 09 | 18.2 | Bleaching variable 3 |
| 10 | 568 | 5.40 | 10 | 18.1 | Recovery variable 1 |
| 13 | 248 | 6.15 | 11 | 18.1 | Digesting variable 2 |
| 50 | 1823 | 45.71 | 12 | 17.5 | Recovery variable 2 |
| 6 | 1552 | 3.88 | 13 | 17.4 | Bleaching variable 4 |
| 29 | 266 | 9.24 | 14 | 17.4 | Brownstock variable 5 |
| 8 | 180 | 3.99 | 15 | 17.3 | Bleaching variable 5 |
| 9 | 297 | 4.71 | 16 | 17.2 | Bleaching variable 6 |
| 27 | 201 | 8.57 | 17 | 17.0 | Brownstock variable 6 |
| 28 | 848 | 9.08 | 18 | 16.9 | Machine variable 2 |
| 43 | 304 | 20.86 | 19 | 17.0 | Bleaching variable 7 |
| 25 | 504 | 8.31 | 20 | 16.8 | Digesting variable 3 |
| 4 | 246 | 2.40 | 21 | 16.8 | Recovery variable 3 |
| 22 | 846 | 8.07 | 22 | 16.8 | Machine variable 3 |
| 24 | 922 | 8.11 | 23 | 16.7 | Machine variable 4 |
| 20 | 476 | 7.23 | 24 | 16.4 | Brownstock variable 7 |
| 45 | 1733 | 25.66 | 25 | 16.3 | Brownstock variable 8 |
| 1 | N/A | N/A | 26 | 16.3 | Machine variable 5 |
| 5 | 303 | 3.42 | 27 | 16.2 | Bleaching variable 8 |
| 41 | 290 | 19.00 | 28 | 16.2 | Bleaching variable 9 |
| 11 | 247 | 5.72 | 29 | 16.0 | Recovery variable 4 |
| 35 | 1039 | 11.59 | 30 | 16.0 | Bleaching variable 10 |

* Indicates value not used in neural network

**Table 3.2: Variables Used in Pulp Strength Model**

| I# | PIS | CV% | Rank | Impact | Description |
|---|---|---|---|---|---|
| | | | | | |
| 7 | 921 | 3.92 | 31 | 15.9 | Machine variable 6 |
| 40 | 72 | 18.59 | 32 | 15.8 | Digesting variable 3 |
| 12 | 451 | 5.89 | 33 | 15.8 | Bleaching variable 11 |
| 33 | 252 | 10.82 | 34 | 15.6 | Digesting variable 4 |
| 44 | 383 | 22.06 | 35 | 15.6 | Brownstock variable 9 |
| 14 | 300 | 6.33 | 36 | 15.5 | Bleaching variable 12 |
| 38 | 1551 | 14.3 | 37 | 15.2 | Bleaching variable 13 |
| 36 | 293 | 12.26 | 38 | 15.1 | Bleaching variable 14 |
| 16 | 40 | 6.59 | 39 | 15.0 | Digesting variable 5 |
| 42 | 350 | 19.71 | 40 | 14.3 | Digesting variable 6 |
| 15 | 648 | 6.46 | 41 | 14.3 | Recovery variable 5 |
| 32 | 1959 | 9.83 | 42 | 14.0 | Bleaching variable 15 |
| 46 | 271 | 25.75 | 43 | 10.8 | Bleaching variable 16 |
| 23 | 341 | 8.11 | 44 | 9.5 | Bleaching variable 17 |
| 37 | 1575 | 13.30 | 45 | 8.9 | Bleaching variable 18 |
| 34 | N/A | 10.98 | 46 | 8.5 | Machine variable 7 |
| 19 | 39 | 6.78 | 47 | 6.1 | Brownstock variable 10 |
| 30 | N/A | 9.47 | 48 | 1.4 | Machine variable 8 |
| 51* | 1410 | 51.29 | | | Water treatment variable 1 |
| 52* | 1409 | 65.31 | | | Water treatment variable 2 |
| 53* | 342 | 105.93 | | | Digesting variable 7 |
| 54* | 1412 | 106.87 | | | Water treatment variable 3 |
| 55* | 1500 | 195.47 | | | Water treatment variable 4 |
| 56* | 228 | 234.94 | | | Bleaching variable 19 |
| 57* | 1162 | 247.78 | | | Machine variable 9 |
| 58* | 1822 | 353.30 | | | Recovery variable 6 |
| 59* | 1501 | 465.44 | | | Water treatment variable 5 |
| 60* | 1335 | 471.30 | | | Machine variable 10 |
| 2* | 309 | .54 | | | Bleaching variable 20 |
| 3* | N/A | .59 | | | Machine variable 11 |
| Out | N/A | 5.43 | | | Pulp strength |

* Indicates value not used in neural network

**Table 3.2: Variables Used in Pulp Strength Model**

The first column of the table represents the input node number in the network that the variable is entered. The coefficient of variance (CV) column is shown as a percent.

The impact column represents the results of a simulation to determine how much pulp strength changes when subjected to a change in the input variables by a set amount. A vector was built that contained the average value for each input. A vector was then built for each of the other 48 variables in which all variables were their average value except for one. For that variable, input was changed by 10% of its CV. The change was calculated based on CV to ensure that the relative impacts of the different variables were the same.

Each vector was presented to the network and the output recorded. The change in pulp strength (relative to its CV) is the value shown in the impact column. The rank column is simply the order of sensitivity of pulp strength to perturbations in the input variable. The other columns contain descriptions of the actual variables used and are blank to maintain confidentiality of Weyerhaeuser data.

According to the model, the greatest impact on pulp strength is caused by pitch dispersant. Unfortunately, experts assert that this cannot be true, i.e., adding more pitch dispersant will not make a stronger pulp. However, pitch dispersant is a measure of woodchip properties. In other words, as the raw wood changes, operators make compensating changes in pitch dispersant. The raw wood is known to be a significant contributor to pulp strength, but there are no on-line measurements of wood properties, and even off-line measurements are inconclusive. Therefore, pitch dispersant is likely signalling wood property changes to the model.

Another possibility is that the experts are wrong. Pitch dispersant is felt to be a weaker contributor to pulp strength so it is not modelled. It may be that pitch dispersant has a greater impact than previously thought. This highlights possible problems with a priori knowledge. If a priori knowledge is incorrect, greater problems can occur than if it did not exist at all. As well this illustrates a problem with traditional modelling techniques. Typically the modeler is forced to minimize the number of variables that are input to the model due to limitations in the modelling technique. This may result in important relationships being overlooked. The ability of neural networks to model problems with large dimensionality results in fewer variables being overlooked.

Scanning the list of variables shown in Table 3.2 one can see that there are many rankings that are expected, and many that are surprising. The next phase of modelling will be to verify these results using both network and traditional approaches. In many cases, designed experiments will be required. The next phase of modelling is on hold waiting for a new high speed data acquisition system to be installed in the spring of 1995 and research into chip quality measurements that will be performed by the Laser Institute in 1995.

In summary, neural networks provide a simple approach to modelling that can determine base relationships where other traditional techniques are less successful. As with any technology, the user must be conversant with the tools that they use, or inappropriate results may be obtained. The availability of commercial packages greatly lessens the burden on today's modeler. However, this availability does require increased diligence on the part of the modeler to ensure that he or she understand exactly what the tools are doing and how to take the greatest advantage of them.

The use of a commercial tool is only part of developing a [illegible]ful model. Whatever the modelling tool, proper data preparation, data analysis and model verification is required. These steps are usually the most difficult portion of the modelling exercise.

Previously we compared neural networks to regression analysis. Given the ability of neural networks, it may appear that models built using this tool are capable of working wonders; and, in many cases they can. However, they have areas of weakness as well. Neural networks are excellent at modelling processes with high dimensionality. Neural networks are also excellent at rejecting noisy signals. However, studies I conducted as part of a Computing Science course showed that neural networks have difficulty simulating very dynamic functions such as the Mexican Hat function. A network based tool that does not suffer from this limitation is the Adaptive Logic Network.

## 3.3 Adaptive Logic Networks

### 3.3.1 Introduction

In this section a model of the basis weight valve problem will be developed through an Adaptive Logic Network (ALN) using development tools provided by Dendronic Decision Limited[26]. ALNs have several advantages that make them useful for a wide range of problems[27]. ALNs have the features of both regression and neural network models. Because of this, they can be used in place of either of these modelling techniques.

ALN's are comprised of two parts. The first part is a tree of logic functions. Each node in this tree has either an AND or an OR function. Like a neural network, the tree is organized in layers and all nodes in the same layer have the same function. The function alternates each layer. For example, if the top layer uses an AND function, then the next layer will use an OR function. All of the nodes in the logic tree have other nodes below them, either more logic nodes, or terminal nodes.

Terminal nodes have no nodes beneath them and represent the second part of the ALN. Each terminal node develops a single hyper-planar surface. In two dimensions, this is a linear segment. In dimensions higher than three, the surface developed is a

---

[26] Dendronic Decisions Limited, 3824-108 St, Edmonton, Alberta, T6J 1B4, (403) 438-8285

[27] W. W. Armstrong, et al, *Learning and Generalization in Adaptive Logic Networks*, Artificial Neural Networks, T. Kohonen et al (Editors), Elsevier Science Publishers B. V.

hyperplane. The number of dimensions is determined by the number of input variables. If a problem has four input variables, then four dimensional linear segments are required and each terminal node would have four inputs.

During training, the linear segments are developed in terms of slope and intercept. Given an input, each terminal node (linear segment) provides an output value. Using a best fit scheme, the linear segment is built so that it provides accurate values for some range of the input data. The binary tree is responsible for learning which linear segment to use for which range of data.

A simple tree adapted for modelling continuous processes is shown in Figure 3.8. In this example, a three-dimensional surface is being modelled, therefore, each terminal node has three inputs to define the linear segment.

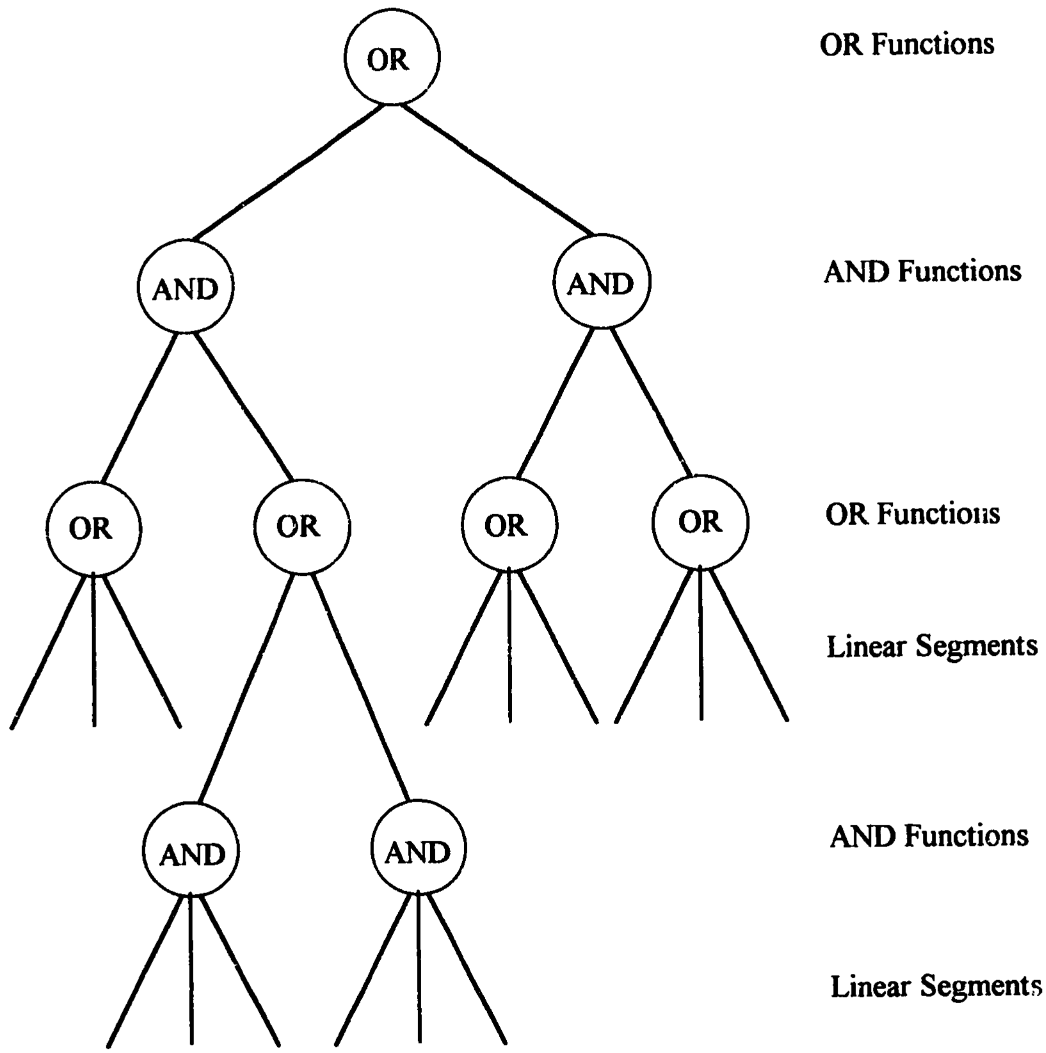


**Figure 3.8: Simplified ALN Structure**

This example has 5 planar surfaces. Terminal nodes are attached to the lowest and next lowest layers at random. The terminal nodes are split over two layers so that the binary tree can develop a relationship with planar segments based on both AND and OR logic functions.

When illustrating neural network structures, much of the detail is not shown so that the drawing is easy enough to understand. The simplicity of ALNs is demonstrated by the fact that the above example is a full featured and robust configuration. Adequate models could be built using fewer processing elements than shown above.

### 3.3.2 Developing models with ALNs

In general terms, ALN's operate much like BPN's. Training and test data preparation is similar. Rather than specifying the number of hidden layers, nodes, and learning algorithm, the ALN user specifies the number of layers, the logical function to use in the topmost layer, and the fan-in of the network. Fan-in refers to the number of connections to each non-terminal node. For example, in Figure 3.8 the fan-in is 2 as each non-terminal node has 2 nodes connected to it. The number of layers and the fan-in control the number of linear pieces used in modelling the function.

Because of randomness built into the network, it is not possible to determine the exact number of linear pieces that will develop, but an order of magnitude can be ascertained. This is a limitation of the current beta version being used and will be fixed in the final release version of the software. The reader is referred to the references for further details on the actual structure and operation of the network

Once the basic structure of the ALN is built, the modeler can input the existing knowledge of the process. This knowledge takes the form of two types of data: monotonicity and slope.

Inputs can be monotonic increasing, monotonic decreasing, or free. Monotonic increasing means that the output variable always increases when the input increases. If the output of a process is known to be monotonically increasing with some specific input, then the ALN should always provide an increase in the output when presented with an increase in the input.

By specifying that the input is monotonic increasing, the ALN will ensure that no planar segment is developed with a negative slope. A positive (or zero) slope will be enforced, even if the input data indicates a negative slope is required in some region. Usually this aids in smoothing small perturbations in the input data. However, it also can uncover inconsistencies in the input data.

Similar to the monotonicity, the maximum slope of a relationship can be specified. This helps reject noise or other inconsistencies in the input data.

### 3.3.3 ALN model of Basis Weight Valve Position

To help understand the operation of an ALN and to compare its performance to that of regression analysis and neural networks, an ALN was developed to model the basis weight problem discussed earlier in the chapter. The data used is the same as discussed in section 3.2.1 . The model was built using basis weight valve position and throughput data when the consistency was 3.2%.

Figures 3.9 and 3.10 show the results of the models.

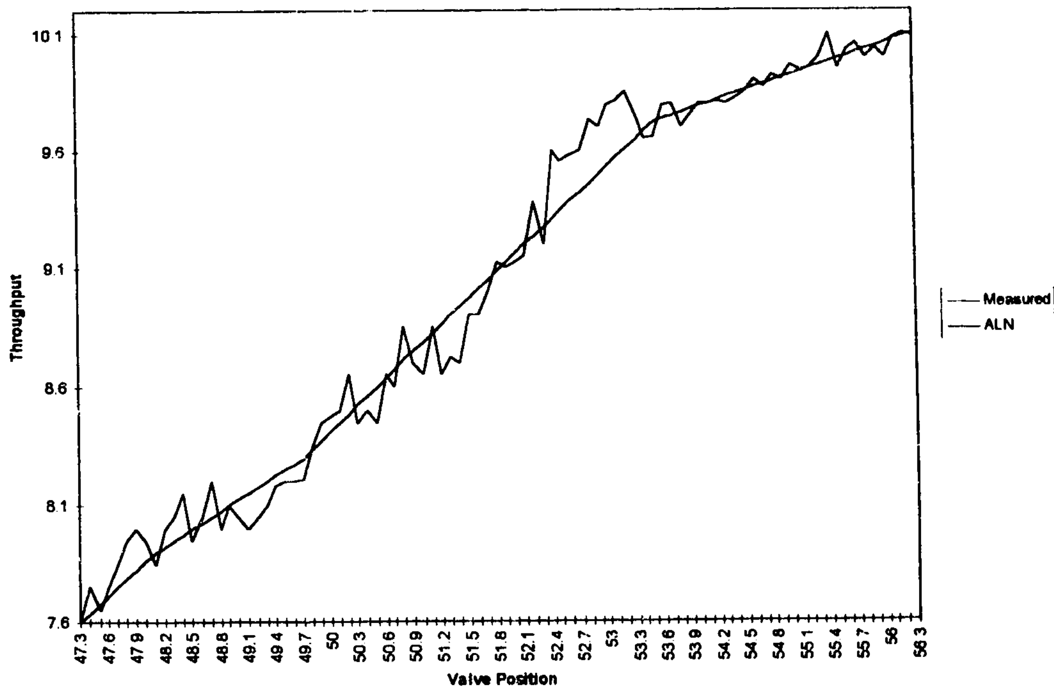


**Figure 3.9: ALN Model of Machine Rate**

The first figure shows the comparison of actual measurements to that predicted by an ALN comprising 4 layers with a fan-in of 2. The topmost node has an OR function. This configuration should provide $2^{(4-1)}=8$ linear segments which is much larger than required. Because we knew what the curve looked like, (from the neural network study) it was determined that three linear segments would be sufficient to model this problem. The above configuration was the smallest configuration that provided an acceptable model

It was discovered that this was a bug in the beta software which will be fixed in the release version. Typically, if the user can determine how many linear segments are required then the ALN is constructed to provide that. If the modeler cannot determine the number of segments needed, then a sufficiently large number is used.

The second figure shows the comparison of actual measurements versus that predicted by a neural network with 2 hidden layers of 3 nodes each using the extended delta bar delta learning rule.

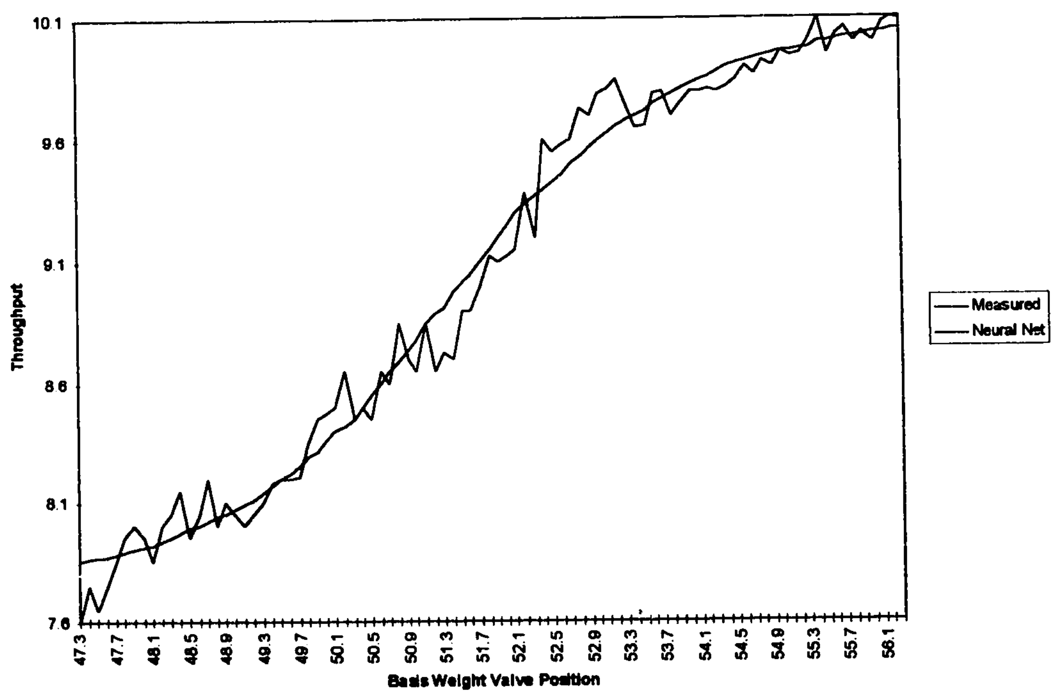


**Figure 3.10: Neural Network Model of Machine Rate**

As can be seen, both methods give good results. Both models took about the same time to train (about 1 minute from 100 data points). It should be possible to compare the two models quantitatively by comparing the RMS. error of each model. Unfortunately, the beta version of the ALN software does not provide this value. This was another recommendation from this project that will be implemented in the final production version. From a use standpoint, both methods were similar.

The major difference was the use of a priori knowledge and the analysis possible with the ALN. The neural network provides an excellent prediction of throughput based on basis weight valve position. However, there is little capability to extract a simple mathematical relationship from the neural networks complex model. The ALN model also provides an excellent simulation of the process, and the linear approach can be clearly seen. The neural network shows a sigmoidal type relationship, whereas the ALN provides basically two linear segments to model most of the operating range. The ALN clearly shows a break point in the relationship at a valve position of about 53.4. This break point is difficult to determine with the neural network.

In using the ALN, the concept of the Simplest Accurate Model (SAM) was employed (i.e. what is the simplest model that emulates the process with a desired accuracy). Because of the complexity of neural networks, SAM has no meaning when using them.

The ALN model was used to determine the fit using 1, 2, 3, and even more linear segments. For the basis weight study, the 2 segment model was the simplest accurate model. Because of the simplicity of the model, it can be very easily implemented in a process control situation. Not only will the implementation be easy, but the draw on computing resources will be minimal. Where the neural network will require 17 multiplication and 16 addition operations, the linear model requires 1 logical test, 1 multiplication, and 1 addition. In other words, the linear model should be over 15 times faster than the neural network implementation. This speed difference is important in situations where very quick response is required or where there are significant numbers of models employed. An example of an ALN that uses this high speed capability is in active suspensions for military vehicles traversing rough open terrain.

In summary, neural networks and ALN's provide powerful tools for modelling complex processes. ALN's have an advantage where a priori knowledge exists, fast response is required or further analysis of the model is desired. Where these requirements do not exist, either method is adequate. Neural networks are particularly good where noisy systems or highly dynamic processes must be modelled. No matter which method is chosen, data preparation and model evaluation [illegible] important steps in building an acceptable model. A comparison of each of the three techniques is shown in Table 3.3

| | Simple | Many Variables | Can use a priori knowledge | must use a priori knowledge |
|---|---|---|---|---|
| Regression | Yes | No | Yes | Yes |
| Neural Networks | No | Yes | No | No |
| Adaptive Logic Networks | Yes | Yes | Yes | No |

**Table 3.3: Comparison of Modelling Techniques**

## 3.4 Summary of Modelling with Network Models

Whether trying to find the best fit for a straight line through a two-dimensional set of points, or mapping a complex surface to a multi-dimensional data set, regression analysis has long been the mainstay of the process engineer. This is even truer today when computers have eliminated most of the drudgery from the process. Although suited to any mathematical form, regression is most commonly used to fit straight lines to nearly linear data sets.

Linear models form the most common model in industry. The ubiquitous PID controller provides a linear relationship between the dependent and independent variables. When non-linearities in the process cannot be ignored, the simple expedient of gain scheduling is often employed. This is simply an extension of the linear approach (i.e., assume that the modelled function is made up of two or more linear pieces).

Why are linear approximations popular? Certainly, from an historical perspective, previous generations of modelers did not have the computing capacity required to develop and implement large numbers of non-linear models. Two reasons for the continued popularity of linear models are that by design they are easier to understand than non-linear models, and they typically provide adequate accuracy. It has only been recently that the influx of modern computer power has made these restrictions obsolete and non-linear modelling methods are beginning to find acceptance in the modelling community. The indication of this acceptance is evidenced in the increased implementation of neural network based models. Neural networks are now a major field, not only for the academics, but for the modelling community.

The Backpropagation Network (BPN) is the most common type of neural network used in modelling. It has the advantage of being relatively easy to program, simple to use, and can achieve a good fit to a known data set. Naturally, if this data set is an accurate representation of some larger data set, then the network will provide a satisfactory model of that data set as well. Although this point seems obvious and hardly worth stating, ignoring it is one of the major stumbling blocks of neural networks.

This real problem is one created by the exuberance of the neural network community. Several papers have described neural networks that are able to extrapolate or interpolate into areas for which there is little or no training data. In other words, they have contradicted the above "obvious" comment. Unfortunately, these examples are often interpreted by inexperienced modelers to imply that all neural networks can interpolate or extrapolate with similar accuracy.

The examples of good interpolation and extrapolation with neural networks are due to fortuitous circumstance as opposed to any systemic capability. The only systemic impact is that, because neural networks are much more powerful than traditional methods, they are capable of generating a wide range of results in areas where there are no training points. In some cases, the output is what is desired. However, it just as easily can be an undesired result. The training mechanisms provide no guidance to the network where there is a scarcity of training data. It has been shown that even well trained neural networks can generate a very large (and erroneous) change in output when interpolating between two inputs.

One of the advantages of a linear regression model (or even a non-linear regression model) is that the modeler is usually aware of the limitations of the model. The fit of the model output to actual data is closely scrutinized. This scrutiny is possible because the complexity of problem that can be solved by regression is usually restricted.

Neural networks, on the other hand, can model very complex processes. Visualizing a best straight line or parabola to a set of points is easy compared to visualizing a 10-dimensional hyperplane. Usually, the modeler can "see" how well a regression model simulates the measured data while they often have little idea of what the neural network output looks like. Typically, all the neuromodeler has is a measure of total RMS. error.

Another criticism of neural networks is their "black box" methodology. Regression models take the approach of finding the simplest mathematical model that can fit a set of known data points. Neural networks, on the other hand, take the opposite approach. They provide an extremely complex mathematical model and fit it to the data points. Once again, it is easy to understand the regression model that comprises a few linear or quadratic terms. A simple neural network will have over a hundred exponential terms and is impossible to decipher directly.

The greatest difficulty with neural networks is their inability to make use of a priori knowledge. In other words, not only do they allow users that have no understanding of the underlying process to build models that appear satisfactory, but, if some knowledge of the process exists, there is no method to incorporate this knowledge into the model. If the modeler has no knowledge of the process, this is of little concern. However, if the knowledge exists, it is a very frustrating limitation.

So, the question is, if neural networks cannot use a priori knowledge, provide very complex models, and are black boxes, why are they so popular? The reason is that when treated with the proper care, they are very powerful. Neural networks fill a niche in the modelling world between the simple regression models and the more sophisticated models. They are able to deal with large data sets with correlated variables and large dimensionality.

Another reason is that up to now, there have been no better techniques to share that niche with neural networks. Recent developments in adaptive logic networks have positioned them to take a more dominant role in modelling. ALNs have all the advantages of neural networks while remaining simple, fast, and deterministic.

In this chapter it has been shown that advanced modelling techniques are extensions of traditional techniques. These extensions allow added capability that is needed when building empirical models of real chemical processes. A simple example, the basis weight valve position, was used to highlight the modelling approach used by each technique. This example was chosen because it was a real problem, but simple enough that process complexities would not hamper the illustrative nature of the example. The use of the model in the mill environment was of secondary importance.

The tensile problem, on the other hand, is neither easy to model nor does it lend itself to illustrating different approaches between modern and traditional techniques. For twenty years, experts in modelling have attempted to model this problem and using standard approaches have failed. The ability to provide a working model with neural networks shows the capability of these types of models to succeed where other methods fail.

ALNs are another new technique that hold great promise. The technique is so new that the example in this thesis was built using a beta version of the software. Further developments in the next several years likely will result in this becoming a key modelling tool for many engineers.

Besides showing that these new artificial intelligence based models are very useful to the pulp industry, this chapter has shown that the underlying mathematics of these approaches is sound. With the advent of powerful computers it has become too easy for people to accept with blind faith that any output from a commercial computer program must be correct. People with little understanding of what these packages are meant to do often misuse them. In the end, Leonardo da Vinci says it best:

> Those who ... practice without knowledge are like the sailor who gets into a ship without rudder or compass and who never can be certain

whither he is going. Practice must always be founded on sound theory ... without this nothing can be done well[28].

[28] da Vinci, Leonardo, The Notebooks of Leonardo da Vinci, Jean Paul Richter (Editor), The General Publishing Company, Ltd. Toronto, 1970, Volume I p. 18

# Chapter 4

# Organizational Aspects of Implementing Advanced Technologies

## 4.0 Overview

As can be seen from the foregoing, using advanced technology is a practical way to solve many problems in today's pulp mill environment. However, providing the technology is only one piece of the puzzle. To operate and manage a large chemical plant requires the proper delivery of the technologies used by the plant effectively. Proper delivery of technology is a broad subject and cannot be covered in depth here. However, any discussion of the applications of advanced technology is incomplete without a review of these issues.

In this context, delivery encompasses a much broader definition than simply "giving" the results of some development to the organization. Delivery includes:

- the ongoing development of the delivering organization.
- integration of the technology with the needs of the customer.
- education systems to keep non-technical people abreast of the capability of the technology to solve important problems.
- systems to ensure that the technology is contributing to bottom line results, both in the long term and the short term.

The proper delivery or implementation of technology is difficult. This is especially so with technologies such as AI (expert systems, neural nets and ALNs) that are new and not well understood by the organization. The challenge is to build acceptance both at the individual level and the organizational level. Industry has experienced some key learnings in those areas over the past several years.

## 4.1 New technologies in the Control Room

AI, once only a topic of discussion in the halls of academia, is now being put to practical use in the control rooms of industry. The implementation process for this and other new technologies has not been without its difficulties. The purpose of this portion of the thesis is to review some of the key learnings that we have gained over several years of implementing small advisory systems in the Grande Prairie plant. Similar considerations apply to other computer technologies as well. The ability to implement these technologies is founded on a clear understanding of where to apply the technology (in this case, advisory or expert systems); understanding of the barriers to acceptance at the operator level; and awareness of prerequisites for success.

### 4.1.1 Problems Where Advisory Systems Shine

To receive acceptance, advisory (or expert) systems must be applied to the proper types of problems. Generally such problems are those that no two operators would typically solve the same way. Five types of problems in this category are:

- Complex problems with more data than can be easily assimilated by one person.
- Problems that are not easily discernible until too late to take effective action.
- Problems that occur infrequently.
- Problems for which solutions are non-intuitive.
- Problems where there is a lack of knowledge or training in the operator community

**Complex problems with more data than can be easily assimilated by one person.**

These types of problems require the analysis of large amounts of data and this is rapidly becoming the most frequent use of advisory systems due to the impact of distributed control systems (DCS). The DCS has allowed a significant decrease in the operator to control loop ratio. Although the typical control room of today has banks of screens showing details of the plant's operation for each operator, there is much more data to be monitored than can be displayed at one time. Alternate tracking methods are used for those process variables that cannot be continuously monitored by the operator. The most common method (and a very rudimentary advisory system) is to install alarms that alert the operator when predefined process conditions occur.

Many operations find that simple alarming is of limited use as the process variables go into alarm condition too often, either because many variables go into alarm during a single upset, variables vary around a trigger point setting off the alarm regularly, or because a variable's alarm state is determined by other variables in the process. As a result, operators often turn off or ignore the alarms, with the risk of missing an alarm that requires response. Smart alarming is required and in many cases artificial intelligence is the ideal solution.

**Problems that are not easily discernible until too late to take effective action.**

Many processes exhibit non-linear behaviour - that is, if a problem is detected early, it can be easily fixed. But, if the problem is not noticed in time, it becomes difficult or impossible to fix. To make matters worse, these types of problems are often complex problems as well. Operators often require significant amounts of data to recognize a problem, especially in the early stages. In these cases an on-line expert system can monitor the significant variables, and alert an operator in the early stages of an upset.

In the Grande Prairie plant, digester level control was a problem of this type. The level sensor provided reliable data most of the time. Often, however, it would provide a false indication of level. Only by looking at several other variables could one determine if the reading was valid. In this situation responding quickly to an impending level upset was trivial, while if the upset became large enough, it could take days to properly line out the process again. We used an expert system to monitor all of the variables on-line to provide a more reliable level indication than the level sensor alone.

**Problems that occur infrequently**

In these situations, the appropriate response is generally well understood, but because of the long time between occurrences, the operators often need to relearn the response. Traditionally they "go to the manual" . However, on-line systems - in the form of expert systems or hypertext documents - are becoming more popular.

Occasionally, an infrequent problem has many of the same symptoms as a frequent problem. Here an advisory system can be used to quickly identify the real problem without the human prejudice of expecting the problem to be the frequent one.

**Problems for which solutions are non-intuitive**

There are many examples in industry where the correct response to a process change is not obvious to a casual observer. Many years of experience are needed for an operator to learn these types of processes as they are complex, use large amounts of data or just respond contrary to what one would expect.

**Problems where there is a lack of knowledge of training in the operator community**

The last example is most common in areas that are experiencing high levels of turnover or where processes are extremely complex. Here the operator is not fully conversant with the proper response to the process and must consult others. In these cases, an on-line tool is useful in providing assistance. This also encompasses the occasions when the resident expert retires or quits. When this occurs, an expert system that has captured the knowledge can be a valuable tool while that expertise is redeveloped.

### 4.1.2 Barriers to Acceptance

Once an appropriate project is identified, there are some potential barriers to acceptance by the operating community that must be addressed. Five potential barriers are:

- Fear of losing jobs.
- Fear of leading edge technology.
- Fear of computers.
- Fear of reduced stature for experts.
- Lack of visible results.

If worked properly and in a timely manner, most of these barriers can be avoided. If they are ignored or considered too late, these barriers can destroy an otherwise good application.

**Fear of losing jobs.**

The fear of layoffs caused by technology upgrades is real. In our plant, although implementing new technology has resulted in a reduced work force, we have reduced at a pace that matches attrition, so there have been no layoffs. Even with this record, we need to address job security each time a new technology installation occurs.

**Fear of leading edge technology**

Although artificial intelligence has existed in academia for many years, it is still a relative newcomer on the industrial scene. Justifying an artificial intelligence application is still much more difficult than justifying a new pump. A pump is a well understood technology while artificial intelligence is still a new concept in many industries. At our plant we have made a concerted effort to educate everyone (and especially plant leadership) on the benefits of the technologies that we use. By educating proactively we find fewer issues to deal with when we implement a new project.

**Fear of computers**

This fear is mostly in the computer illiterate operators. For these operators we have a training program in basic computer operation. The key to this has been the development of such simple applications as a basic messaging system or systems that help them manage their personal work affairs, such as vacation or shift schedules . By increasing their use of computers for general items, they more easily adapt to more complex process and business computer systems.

**Lack of visible results**

Many advisory systems provide business benefit by avoiding or speeding up the resolution of problems. In these cases the benefit is not what you see, but what you don't see: long downtimes, poor quality, and so forth. In these cases careful documentation is required to show the before and after results. Otherwise, it may appear that the new technology has not done anything.

**Fear of reduced stature for experts**

For many experts, much of the prestige of their jobs comes from being called on to help others. Placing their knowledge into a computer can reduce their feelings of self worth and may cause problems in the future. It is important to help them understand how capturing their knowledge on a computer is beneficial to them. In our plant we had one expert on chemical addition that was regularly called for his advice. He was very proud of this role and it was not until he realized how many of the calls were occurring at 2-3 o'clock in the morning that he asked us to build an application that captured his knowledge (and let him get a full night's sleep!).

### 4.1.3 Prerequisites for Success

There are several prerequisites for a successful application. An advisory system must:

- Solve a "real" problem.
- Provide an agreed upon solution.
- Be easy to use with a familiar format.
- Give the correct results.
- Be significantly better than existing systems.

**Solve a real problem**

Typically we have found that operators will not waste their time using a system that does not benefit them. Although this may seem obvious, we have relearned it on several occasions. The most typical example is when a manager comes up with what they think is a great idea and requests an advisory system be built. In some cases the operators have not been properly consulted and when the solution is implemented we find that they don't believe it adds benefit, and the "great idea" sits unused. Before building an application, it is important that the end user understands and agrees to its purpose. If the end user is not the originator of the request, then the originator must also understand the need for user acceptance.

**Provide an agreed upon solution**

By their very nature, advisory systems tend to provide standardized solutions. All users must concur that the advice provided by the system is the proper advice. If a group of 10 operators each use a different operating strategy, then an advisory system that provides one of those strategies as its advice will only be used by one operator.

The advisory system can rarely be used to enforce a standard solution. The operators must accept the standard solution before the advisory system can be effectively used. If there is contention about the correct operating strategy then this needs to be resolved before the complexity of a computer system is added to the problem.

**Easy to use with a familiar format**

For many operators the concept of advisory systems is new. If the system also has a new and strange interface, its use will be limited. The user interface must be as consistent as possible with the existing tools that the operators use.

**Give the correct results**

This again seems fairly obvious, but it is surprising how many applications are developed that do not provide complete accuracy. For a human expert, occasional errors are a sign of being human. Computers do not have the advantage of being human. Any errors in any piece of advice are an indication to the operator that the entire system is flawed. Either the system must provide completely accurate advice or the limitations of the system must be well understood and accepted by the operators. One of the methods we have used is to enrol the operators in the verification of the advisory system. As new learnings are made or errors are encountered, the operators incorporate the new information into the system. This not only improves the system but it also keeps the operators as part of the development loop and keeps them from becoming too reliant on the computer system.

**Better than existing systems**

If the existing system to get advice (e.g. ask another operator) is satisfactory, then no new system will be successful. Advisory systems need to be easier, quicker and/or more accurate than existing methods.

### 4.1.4 Summary

New technologies (such as advisory systems) are quickly gaining more acceptance in industry to solve a variety of problems. Process information managers can assist in this process when developing such technologies by following a few basic rules: focus on the needs of the users and apply the correct technology; ensure operators are involved in the identification of problems to be solved and the development of the system to solve them; and, ensure that users are satisfied with the solutions that the system will provide.

## 4.2 Technology Delivery at the Organizational Level

Successful implementation of new technologies on an application by application basis alone is not sufficient to ensure effectiveness on an organizational level. This requires a much more systemic approach to the selection and deployment issues. A business that produces a product or provides a service uses many different technologies in pursuit of its objectives. These technologies can include financial systems, human resource technologies, computer technologies (both mainframe and PC) and others. How these "support" technologies are delivered can be critical to the success of an enterprise. Clearly a business that has control of all technologies used in its process, has a much better chance of survival than one that does not. This portion of the thesis will describe one method for ensuring that technologies support the business objectives and continue to add value.

### What is a Technology?

The dictionary defines technology as the "science of industrial and mechanical arts"[29]. This is certainly an all encompassing description, and one that is satisfactory for our purposes. Often, technology is narrowly defined so as to include only those things that have some associated hardware, such as computer technology. When discussing support requirements, this narrow definition minimizes the needs of the "soft" technologies such as human resourcing, financial analysis, etc. Conceptually, there is

[29] The Penguin Concise English Dictionary, Revised Edition, 1969, p. 744

little difference between a "hard" technology and a "soft" technology and the techniques discussed here do not just explain how to organize a computer department, but can be applied to almost any business endeavour.

**What is the problem?**

Why should business be concerned about the proper implementation of technology? The proper implementation of technology is key to the success of a business. Although there are examples of businesses that are successful without a systematic support plan for technology, there are many more examples of businesses that fail due to lack of adequate support systems. With proper organization and direction, those technologies that support improved profitability can be nurtured to contribute the optimum amount. Equally important, those that do not contribute fully, can be eliminated or contracted out, reducing costs to the organization.

In many plants today, a wide range in quality of technology implementations exist. Typically there will be several implementations of different technologies that cover the spectrum of too little, too much, and just the right amount of support structure.

For example, in our plant, we have people that are unable to adequately use the interoffice electronic mail system. This results in increased costs as paper systems (mail runners, typists, etc.) must continue to be supported in addition to the electronic system. Therefore two systems are being supported (with a real cost) where one would suffice if it were properly implemented. At the other end of the spectrum, we have a person who used to write their own editors for their PC. Not that there was a real deficiency with the standard editor, but rather, programming is something this person enjoys doing, so he wrote his own packages, editors, spreadsheets, etc. Here we have effort being spent developing technology that does not move the business forward. Finally, there are technology implementations that, with minimal support, have greatly contributed to business results for over a decade. All three examples of technology implementation have their associated costs, but each delivers radically different results to the organization.

## What does business need

How can a business ensure that all the technology being used is contributing adequately and is not under or over implemented? The answer is, plan for the future and evaluate the present. As in everything else, if the objective is well understood and communicated, then the chances of success are increased dramatically. Continual evaluation of progress against the objectives will almost always guarantee a successful and durable implementation (as well as timely elimination).

Another question that business is often faced with is what to contract out to third parties and what to keep as part of the core organization. If the technology objectives are well understood, then the decision is easily made. When the objective is not well understood, incorrect decisions often result. To help develop objectives and evaluations a model of the ideal technology implementation was developed. This model is called the Technology Delivery Model or alternatively the VALUE model.

### 4.2.1 The VALUE Model

The VALUE model was the result of several managers at the Procter & Gamble (now Weyerhaeuser) Grande Prairie Pulp Mill trying to understand why some technology installations were very successful, while other applications were implemented, contributed at a high level for six months, and then their contribution dropped off while they continued to draw heavily on the organization's resources (in the form of support and maintenance). The successful implementations were examined and two common features were discovered. Further investigation revealed that these same features were missing from those technologies that were poorly implemented. The two features were the existence of well understood objectives, and monitoring or feedback mechanisms.

A technically competent support staff was found to be common to both types of implementation so it was determined that this was important to success, but it did not preclude misuse of the technology. An assumption had been that technically competent support staff led to overuse of the technology (the editor programmer) but our research showed that if the objectives were understood, communicated and monitored, there was little chance of inappropriate use of technology. We were also able to determine that without quality technical input, objectives were likely to be poorly developed.

The two features - objectives and monitoring - were further broken down into four tasks. Setting objectives was split into the tasks of Visioning, Applying, and Learning. Monitoring was found to be better described by Enabling as the task was more encompassing than just checking whether objectives were being met. Finally a fifth task was added as it was found that in the successful implementations, the support group rarely used the technology directiy, but rather supported various users in the plant. Therefore, in order to make the model complete, a Utilization step was added. These five terms:

Visioning

Applying

Learning

Utilization

Enabling

classify the tasks necessary to successfully implement a technology. In practice, each task class is executed iteritively and concurrently, so there is no logic to the order (other than its value as an acronym).

## The Objective

Visioning, Applying, and Learning are the processes used to develop and communicate a clear objective with and to the organization. The objective is described in terms of what decisions the users make and how the technology is used to help them accomplish that. If the organization has a human resource or work system plan then that is usually a key part or influence on the technology vision.

## Visioning

The knowledge gained in the learning phase (described below) is used to develop a vision of how the technology benefits the business. This vision is not described in terms of the technology itself, but in terms of what the users are doing and the results they achieve. What is achieved is of p.ime importance, how it is done is secondary.

One successful vision (for our process information system) was written in the form of a trip report describing the plant five years in the future. The trip report described how the operators interacted with each other, for which decisions they were responsible, and how they made these decisions. Attached to this was a brief outline of what technologies were required to fulfil this vision.

Typically the vision is built with substantial discussion between the support groups and the customers. Once a preliminary vision is developed it is shared throughout the organization for feedback and modified as required before it is implemented.

## Application

Once the vision is developed and understood, projects may be initiated to turn the vision into reality. More often however, the vision is used as a guide when implementing projects justified by other means. For example, only 1 project was initiated as a result of the previously mentioned information system vision. However, when one of the computers needed to be replaced due to high maintenance costs, the vision was used to determine which replacement option would be consistent with future needs. This highlights another key feature of the vision, it is a living document that sets direction as opposed to a project implementation timeline.

## Learning

For each successful technology, one or more people understood what the state of the art technology was and what its capabilities were. Rarely was the state of the art implemented. But, there was a good understanding of what was possible. This knowledge was gained through reading literature, attending selected conferences, and talking to peers in other industries and sectors.

The technologists also had a very good understanding of what their customers required. The customer/supplier relationship between the technologist and customer was more of a partnership than customer/supplier interface. For example, in the four person real time information system team, 2 people had transferred to the group after extensive experience in operations, one transferred from maintenance, and only the department head had formal computer science training. Successful technology groups do not just rely on customer input, but have developed an in-depth understanding of the customer's business, either through previous experience, rotational assignments, or other methods.

**Utilization**

Utilization is the actual task of using the technology. In an information system, it would describe all of the applications that a user sees, decisions that the user makes, training that the user requires, and so on. This can be a very detailed document and can be integrated into other programs, such as defining the skills required for promotion. Routine reporting of benefits derived from using the technology can also be generated in this phase.

This is often the most visible part of a technology, and in poorly implemented technologies, this is the only part. Utilization can consume vast resources and, without the other parts of the model, there is little chance of the utilization contributing effectively to business results. Unfortunately, the vast resources consumed by utilization are often hidden as the users often have other responsibilities. Two hundred operators who use an information system 10% of the time represent twenty dedicated people. This is never as noticeable as the 4 support personnel dedicated to the system (and usually located in a very visible location).

**Enabling**

The last phase is comparing the current contribution of the technology to the objectives. If outages are detected then corrective measures are initiated. This can range from the use of training programs to the modification of the objectives to deciding to abandon the technology altogether. The results of this comparison are broadly shared with the organization. This achieves two objectives. First it maintains the organization's focus on the objectives for the technology. Second, it highlights how valuable (hopefully) the technology is to business success.

## 4.2.2 Application Examples

This model has been successfully applied to several technologies. The two best examples are the plant's real time information system and control engineering system.

**Real time information system**

The plant's real time information system was implemented in 1982 with very clear and well communicated objectives. Over the years the objectives and benefits of this technology have been shared with the organization every 6 to 12 months. Since 1982 many changes have occurred in the objectives.

In 1982, the information system's objective was to provide real time information to its primary customer - the operator. As Distributed Control Systems (DCS) were implemented, the objective of the information system moved to an historical focus with the primary customer now being the process engineer. By maintaining a continual emphasis on th ctive of the technology (even though it was changing), the change to the current re. evolutionary. Without continual monitoring, it is very likely that the real mation system would have remained unchanged until it was discovered at some late date that it was obsolete and not being used. Through all these changes the technology has continued to contribute to business results, as shown by the records of the annual assessments.

## Control Engineering

As the DCS was implemented in the plant it was felt that there was a need for engineers who could leverage this new technology to provide better quality pulp at a reduced cost. A vision was developed with the focus on what decisions people needed to make. This vision clearly laid out the objectives of the technology. Based on these objectives, it was determined that contracting out the control engineering role was inappropriate. Two new control engineers were hired and systems set up to support them. Again, as that group has evolved, the objectives have changed, but because the evolution has been managed, there is no question about value of the team to the plant objectives

**Contractors**

Often the question of whether to contract out parts of the business arises. In many cases this is a very good strategy. A first decision point is to determine if your organization has the skills and desire to manage the technology using the VALUE model. If utilization is all that can be supported, then contracting out to an organization that can provide the visioning and evaluation systems is appropriate. Clearly the need is to find a third party that has the capability to support all steps of the VALUE model and not just the utilization aspect.

If the business has the skills to support the VALUE model, the next test is whether the technology is critical to business results. For example, most businesses need computer systems, but for many industrial plants, this technology is not considered critical to the core business. The same business might decide that process engineering is more critical in the development of a proprietary competitive advantage. In this example, computer systems would be open to outsourcing, while process engineering would not.

In some cases, how a technology is implemented can determine whether to outsource. In our real time information system, the strong operation background of the support team cannot be replicated by a third party.

## 4.3 Summary

Technologies can have a major influence on whether a business succeeds or fails. A well managed business requires that all its parts be well managed as well. This means well thought out and communicated objectives, and continual evaluation of progress against objectives. A clear vision that is a living document is a key method to share objectives and set business direction. The VALUE model has been found to be an excellent tool for ensuring this occurs.

When these elements exist at the organizational level the support organizations are able to develop into cohesive groups that can truly impact the bottom line. This impact will be a long term effect rather than a short term gain that is usually seen in organizations that do not foster this type of development.

Although organizational support is important, the individuals and teams implementing applications must also use some methodology to ensure success of their applications. Following the guidelines outlined in section 4.1 has proven crucial to implementing successful applications at the Grande Prairie pulp mill.

By maintaining a focus on success at both the organizational level and the team level, a business can successfully utilize technology to achieve its corporate goals.

# Chapter 5

# Conclusions

Increased global competition, increased environmental awareness, improved technology to measure environmental impact and an increase in cost of the forest resources have created a need for improved tools to operate the increasing complex pulp and paper plants of Canada. Artificial intelligence provides some of these tools. This thesis reviewed several applications of artificial intelligence to pulp mill operations. These examples illustrate a variety of problems that can be addressed with artificial intelligence, problems that could not be solved using traditional methods.

The first application discussed involved the use of an expert system to aid customer service representatives in determining safety inventory levels and inventory replenishment plans. It was determined that basic control theory could be applied to this problem. This is a marked departure from the accepted approach currently used by the industry.

The second chapter of the paper dealt with the development of models to analyze plant operation. Three approaches to modelling were investigated, parametric models (regression analysis), neural network models and adaptive logic models. The two examples show how AI based networking techniques can overcome problems that traditional modelling was incapable of solving.

The final chapter consisted of two papers that dealt with the issue of how to integrate the use of advanced technology into the manufacturing organizations they support. The major issues addressed were how to get the required organizational resources to develop and maintain the systems, and how to ensure the systems are used effectively and remain healthy contributors to the bottom line over the years.

The paper looked at three examples, illustrated their development, described how they solve real industrial problems and then talked about the organizational and human factor issues that need to be addressed when implementing these types of projects in a plant.

## Future Research

Areas for future research include:

- research into the application of modern control theory to inventory management.
- research into the relationship between problem complexity and network topology.
- research into the application of Adaptive Logic Networks to chemical processes.
- research into the impact of processing parameters on pulp strength

These areas of research would provide important contributions to the study of the application of artificial intelligence to the pulp and paper industry.

# Bibliography


Alexander, Igor (Editor), Neural Computing Architectures, The Design of Brain-like Machines, MIT Press, Cambridge Ma, 1989

Allen, Mary Kay and Helferich, Omar Keith, Putting Expert Systems to Work in Logistics, Council of Logistics Management, Oak Brook, Il, 1990

Antonelli, P.L. and Bradbury, R.H., *Volterra-Hamilton Models in the ecology and Evolution of Colonial Organisms*, Course notes (to be published) for Mathematical Modelling Course 472 University of Alberta, Spring 1994

Antonelli, Peter L. (Editor), Mathematical Essays on Growth and the Emergence of Form, University of Alberta Press, Edmonton, 1985

Antsaklis, P.J., *Control Theory Approach*. Mathematical Approaches to Neural Networks, J.G.Taylor (Editor), North-Holland, New York, 1993

Barecca, Cliff, *Variables with Possible Impacts on Tensile*, Confidential Weyerhaeuser document, 1991

Barr, Avron and Feigenbaum, Edward A. (Editors), The Handbook of Artificial Intelligence Volumne II, , Addison-Wesley Publishing Company Inc., Reading Ma., 1982

Barr, Avron; Cohen, Paul R.; Feigenbaum, Edward A. (Editors), The Handbook of Artificial Intelligence Volumne IV, Addison-Wesley Publishing Company Inc., Reading Ma., 1989

Barr, Dean S. and Mani, Ganesh, *Using Neural Nets to Manage Investments*, AI Expert, Miller Freeman, San Francisco, February 1994

Burrus, Daniel, Techno Trends How to Use Technology to Go Beyond Your Competition , HarperBusiness, 1994

Cox, David, *Using C++ for Back Propagation*, AI Expert, Miller Freeman, San Francisco, July 1994

da Vinci, Leonardo, The Notebooks of Leonardo da Vinci, Jean Paul Richter (Editor), The General Publishing Company, Ltd. Toronto, 1970, Volume I

Dayal, B.S., *Feedforward Neural Networks for Process Modelling and Control. Master Thesis*, McMaster University, 1992

Dendronic Decisions Limited, 3824-108 St, Edmonton, Alberta, T6J 1B4, (403) 438-8285

Eberhardt, Keith R., *Survey Sampling Methods*. Handbook of Statistical Methods for Engineers and Scientists, Harrison M. Wadsworth (Editor), McGraw-Hill, New York, 1989

Fogarty, Donald W.; Blackstone, John H.; Hoffmann, Thomas R., Production & Inventory Management, South-Western Publishing Co., Cincinnati Oh., 1991

Freeman, J.A., Simulating Neural Networks with Mathematica, Addison-Wesley, New York, 1994

Hebb,D., The Organization of Behaviour, Wiley, New York, 1949

Hecht-Nielson,R., Neuralcomputing, Addison-Wesley, New York, 1989

Jackson, Peter, Introduction to Expert Systems (2d. ed.), Addison-Wesley Publishing Company Inc. Reading Ma. 1990

Jacob, H.M. *Current Best Approach for Raw Material Inventory Tracking*, Confidential Procter & Gamble document, 1990

Jacobs, R.A., *Increased Rates of Convergence Through Learning Rate Adaption*, Neural Networks, Volume 1. 1988

Kempka, Anthony A., *Activating Neural Networks: Part I*, AI Expert, Miller Freeman, San Francisco, June 1994, P33

Kernighan, Brian W., The C Programming Language (2d. ed.), Prentice-Hall, Toronto, 1988

Kleinbaum, David G.; Kupper, Lawrence L.; Muller, Keith E., Applied Regression Analysis and Other Multivariate Methods (2d. ed.), PWS-Kent Publishing Company, Boston, 1992

Kurtz, Max., Handbook of Applied Mathematics for Engineers and Scientists, McGraw-Hill Inc., New York, 1991

Masters, Timothy, Practical Neural Network Recipies in C++, Academic Press Inc, New York, 1993

McCulloch, W.S. and Pitts, W., *A logical calculus of the ideas immanent in nervous activity*, Bulletin of Math. Bio., 5, 1943

Middleton, Richard H. and Goodwin, Graham C., Digital Control and Estimation A Unified Approach, Prentice Hall, Englewood Cliffs, N.J., 1990

Minai, A.A. and Williams, R.D., *Acceleration of Back-Propagation through Learning Rate and Momentum Adaption*. Inernational Joint Conference on Neural Networks.

Volume 1. January 1990

Morrison, Foster, The Art of Modeling Dynamic Systems, John Wiley & Sons, Inc., New York, 1991

Mosteller, Frederick; Rourke, Robert E.K.; Thomas, George B Jr., Probability with Statistical Applications, Addison-Wesley Publishing Company, Reading Ma, 1961

Murray III, William H. and Pappas, Chris H., Windows Programming: An Introduction, Osborne McGraw-Hill, Berkely, 1990

Murray, Don, *Tuning Neural Networks with Genetic Algoritms*, AI Expert, Miller Freeman, San Francisco, June 1994

NeuralWare Inc. Building IV, Suite 227 Penn Center West, Pittsburgh, Pennsylvania USA, 15276, (412) 787-8222

NeuralWare Inc., Neurocomputing, a Technology Handbook for Professional II/+ and NeuroWorks Explorer. NeuralWare Inc., Pittsburg Pa., 1993

Ogata, Katsuhiko, Modern Control Engineering (2d. ed.), Prentice-Hall, Englewood Cliffs, New Jersey, 1990

OWL International, Inc. 2800 156th Avenue SE, Bellevue WA, 98007 USA, (206)747-3203

Rao, Ming and Qiu Hainming, Process Control Engineering, Gordon and Breach, Langhorne PA., 1993

Rao, Valluru B. and Rao, Hayagriva V., C++ Neural Networks and Fuzzy Logic, Management Information Source, Inc., New York, 1993

Rarisini, T. and Zoppoli, R., *Neural Networks for Feedback Feedforward Nonlinear Control Systems*, IEEE Transactions on Neural Networks, IEEE Neural Networks Council, May 1994

Rich, Elaine and Knight, Kevin, Artificial Intelligence, McGraw-Hill Inc, New York, 1991

Seborg, Dale E.; Edgar, Thomas F.; Mellichamp, Duncan A., Process Dynamics and Control. John Wiley and Sons, New York, 1989

Stroustrup Bjarne, The C++ Programming Language (2d. ed.), Addison-Wesley Publishing Company, Reading Ma., 1991

Tushman, Michael L. and Moore William L. (Editors), Readings in the Management of Innovation, Harper Business, 1988

Vondrak, Ivo, *Object Oriented Neural Networks*, AI Expert, Miller Freeman, San Francisco, June 1994

W. W. Armstrong, et al, *Learning and Generalization in Adaptive Logic Networks* Artificial Neural Networks, T. Kohonen et al (Editors), Elsevier Science Publishers B. V.

West, Mike and Harrison, Jeff, Bayesian Forecasting and Dynamic Models, Spring-Verlag, New York, 1989

Window on Technology Volume IV Artificial Intelligence Revisited, International Artificial Intelligence, Toronto

Wittink, Dick R. The Application of Regression Analysis, Allyn and Bacon Inc. Boston, 1988

Xiao, Q.; Rao, M.; Shen, X.; Ying, Y.; Zurcher, U.J., *Systematic Modeling and Decoupling Design and Control of a Pressurized Headbox,* Canadian Electrical and Computer Engineering Conference, Vancouver, Sept. 1993

Zauderer, Erich, Partial Differential Equations of Applied Mathematics, John Wiley & Sons, New York, 1989

Zinn, Walter and Marmorstein, Howard, *Comparing Two Alternative Methods of Determining Safety Stock Levels: The Demand and the Forecast Systems*, Journal of Business Logistics, Volume 11 #1, Nov. 1990

Zurcher, U.J., *Adaptive Logic Networks, A New Approach to Process Modelling,* CPPA Western Branch Meeting, Fall 1994

Zurcher, U.J., *Putting Artificial Intelligence to Work,* IEEE Conference on Control Applications, Vancouver, Sept. 1993