Journal of Scientific & Industrial Research
Vol. 65, July 2006, pp. 565-573

# Artificial neural network for solving paper industry problems: A review

Kumar Rajesh* and A K Ray
Department of Paper Technology, Indian Institute of Technology Roorkee, Saharanpur Campus, Saharanpur 247 001

*Received 01 July 2005; accepted 29 March 2006*

The review focuses on applications of artificial neural networks (ANNs) in various subsystems of pulp and paper making processes. Some applications are pertinent to modeling of pulping, bleaching and energy conservation process while others are related to control of refining, washing of brown stock, combustion of black liquor, lime kiln operation, grade changes and quality control. A general algorithm is also developed to demonstrate how to solve the complex problem using ANN for paper mill subsystems.

**Keywords:** Artificial intelligence, Fault diagnosis, Neural networks, Paper mill simulation, Prediction of paper properties

**IPC Code**: G06N3/02; D2/F

## Introduction

Most systems in paper industry are generally nonlinear and often too complicated to be accurately described with physical models. Neural networks (NNs) are powerful tools that can solve a variety of nonlinear modeling problems. Artificial neural network (ANN) can assimilate operating data from an industrial process and learn about the complex relationships existing within the process, even when the input-output information is noisy and imprecise. There are many methods (Table 1) available for solving problems through ANN; Back-propagation network (BPN) is more widely applicable.

## Principle of the Back-propagation Network

BPN consists of multiple layers of neurons. For a three-layer NN system, there is an input layer, a middle layer (sometimes referred to as a hidden layer), and an output layer. The network is constructed in such a way that each layer is fully connected to the next layer. In BPN, a randomized set of weights on the interconnections is used to present the first pattern to the network. There are various steps to learn the network. Feeding into layer "a" is the input vector I, layer "a" has L nodes, $a_i$(i=1 to L); layer "b" has m nodes, $b_j$(j=1 to m); and layer "c" has n nodes (Fig. 1). Layer parameters are as follows: Inputs to input layer, $I_1$, $I_J$, $I_L$; Outputs from input layer, $a_1$, $a_i$, $a_L$; Outputs from hidden layer, $b_1$, $b_j$, $b_m$; Outputs from output layer, $C_1$, $C_k$, $C_n$; Targets, $T_a$, $T_b$, $T_c$; Weight between input and hidden layer, V; and Weight between hidden layer and output layer, W. The size of the weight is adjusted with the help of learning rate, η, and to speed up the training, momentum rate, α, is used. If η is low, there is possibility of failure in convergence, if η is large, the search path will oscillate. In most cases, η is set to 0.9 and α is assumed as 0.7.

### Algorithm

***Step 1***

Randomly choose the weights $v_{ij}$ and $w_{jk}$ (between 0 to 1) the internal threshold values must be assigned (input layer threshold=0; all hidden and output layer thresholds must be equal to one).

***Step 2***

Introduce the input into NN and calculate the output from first layer (input layer) using

$$X_i=I_i+T_{1i},$$

Output from first layer $a_i=f(x_i)=1/(1+e^{-X_i})$

***Step 3***

Knowing the output from the first layer, calculate the output from second layer (hidden layer), $b_j$.

$$b_j=f(\Sigma^{l}_{i=1}(v_{ij}*a_i)+T_{2j})$$

*Author for correspondence
E-mail: rajtisotra@yahoo.com

Table 1—Comparison between neural networks

| S No. | Methods | Remarks |
|---|---|---|
| 1 | Perceptron neural network | It is a single layer network, it is only able to classify linearly separable patterns, |
| 2 | MLFF (multi-layer feed-forward) Back-propagation network | It is based on gradient descent learning(supervised learning). The multi-layer network is a powerful extension of perceptron network, multi-layer networks can be used as function approximators. It is one of the easiest networks to understand because its learning and update procedures are intuitively appealing and based upon a relatively simple concept. |
| 3 | Adaptive resonance theory (ART) | It is based on competitive learning(unsupervised learning). The ART networks are designed to achieve learning stability while maintaining sensitivity to novel inputs. The general complexity of the network is a limitation. |
| 4 | Learning vector quantization (LVQ) | LVQ network uses both supervised and unsupervised learning to recognize clusters. LVQ networks can be trained to recognize classes made up of multiple unconnected regions. |

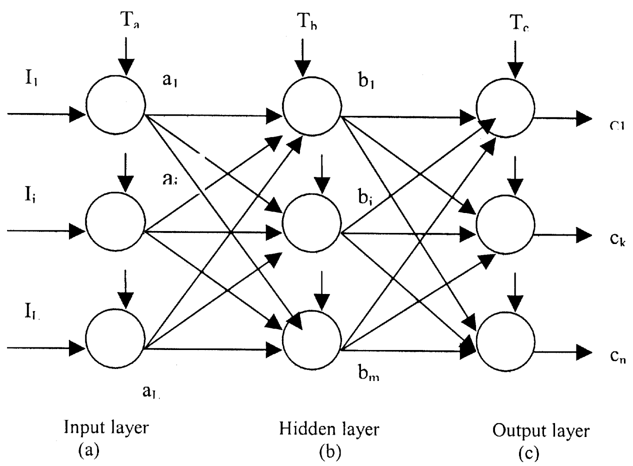


Fig. 1—Multi layer feed-forward neural network

***Step 4***

Knowing the output from second layer, calculate results from output layer, $c_k$.

$$c_k=f(\Sigma^m_{j=1}(w_{jk}*b_j)+T_{3k})$$

***Step 5***

Calculate total mean square error, E

$$E=\Sigma^p_{p=1}\Sigma^n_{k=1}(d_k-c_k)^2$$

where p is the training pattern presented to the input layer, n is the number of nodes on the output layer.

***Step 6***

Calculate gradient descent term for k'th node in the output layer as

$$\delta^p_{3k}=(d-c)\,df/dx$$
$$f(x)=1/(1+e^{-x})$$
$$df/dx=e^{-x}/(1+e^{-x})^2$$

***Step 7***

Calculate gradient descent for jth node on the hidden layer as

$$\delta_{2j}=(\Sigma\delta_{3k}w_{jk})\,df/dx_j$$
$$X_j=(\Sigma v_{ij}\,a_i+T_{2j})$$

where k denotes a node in the output layer.

***Step 8***

Knowing gradient decent term for hidden and output layer, calculate weight changes (between input & hidden layer nodes, hidden layer & output layer nodes) as

$$\delta v_{ij\,new}=\eta\delta_{2j}a_i+\alpha\delta v_{ij}$$
$$\delta v_{jk,new}=\eta\delta_{3k}b_j+\alpha\delta w_{jk}$$
*for* $v_{ij,new}$.
$$\Delta v_{11},new=\eta\delta_{21}a_1+\alpha\delta v_{11}.$$
$$\Delta v_{12},new=\eta\delta_{22}a_1+\alpha\delta v_{12}.$$
$$\Delta v_{13},new=\eta\delta_{23}a_1+\alpha\delta v_{13}.$$
*and for* $\delta w_{jk,new}$
$$\Delta w_{11},new=\eta\delta_{31}b_1+\alpha\delta w_{11}.$$
$$\Delta w_{12},new=\eta\delta_{32}b_1+\alpha\delta w_{12}.$$
$$\Delta w_{13},new=\eta\delta_{33}b_1+\alpha\delta w_{13}.$$

***Step 9***

After knowing the weight changes, update the weights according to the equations

$$W_{jk,new}=w_{jk}+\Delta w_{jk,new}$$
$$V_{ij,new}=v_{ij}+\Delta v_{ij,new}$$

where, f( ) is sigmoid function; $T_{2j}=1$ & $T_{3k}=1$ are bias functions added to weighted input; j=1 to m; k=1 to n.

Steps 1-9 are to be repeated for all training patterns until the network predictions are within some defined tolerance of acceptability.

## Application of ANN to Pulp and Paper Mill

Combination of a neural network system, a rule based system and a conventional computational system, provides a tool to handle pulp and paper mill problems with simplicity and effectiveness. Paper making process consists of a large number of subsystems (unit operations and processes). Some of the applications related to modeling and control of paper mill subsystems are depicted below.

### ANN to Modeling and Simulation

***Modeling Kappa Number in a Continuous Kamyr Digester***

Kamyr digester, employed for converting raw materials (wood, bamboo, bagasse etc.) into pulp in a continuous manner, necessitates sophisticated control of pulp parameters. In kraft pulping, pulp quality is generally measured by kappa number ($\kappa$), which is related to lignin content remaining in the pulp. Inferential models for $\kappa$ are developed using partial least square regression (PLS) and NN. Dayal *et al*[1] investigated above technique to build empirical models for $\kappa$ using historical data from an industrial continuous Kamyr digester. A static model for $\kappa$ based on the 21 input process parameters has been developed.

***Modeling Pulp Bleaching Process***

Objective of bleach plant is to obtain brightness (D) of pulp at a desired level while minimizing the use of bleaching chemicals like $Cl_2$, $ClO_2$ etc. There are various sequences of bleaching like CEH, CEDD, DEDED etc. mainly used by Indian paper mill. An ON-LINE measurement of D-stage brightness is desired for process control. However, because of sensor limitations, D is measured OFF-LINE by an hourly manual test. To ensure adequate pulp D, operators occasionally control to a higher than required D with a corresponding increased use of bleaching chemicals. The purpose of the model is to provide a real time indication of D so that operators can reduce chemical use. A bleach plant model has been developed using ANNIE[2] (artificial neural network integrated environment), in which data preparation, network topology selection, network training, evaluation, modification, code generation and simulation are integrated (Fig. 2).

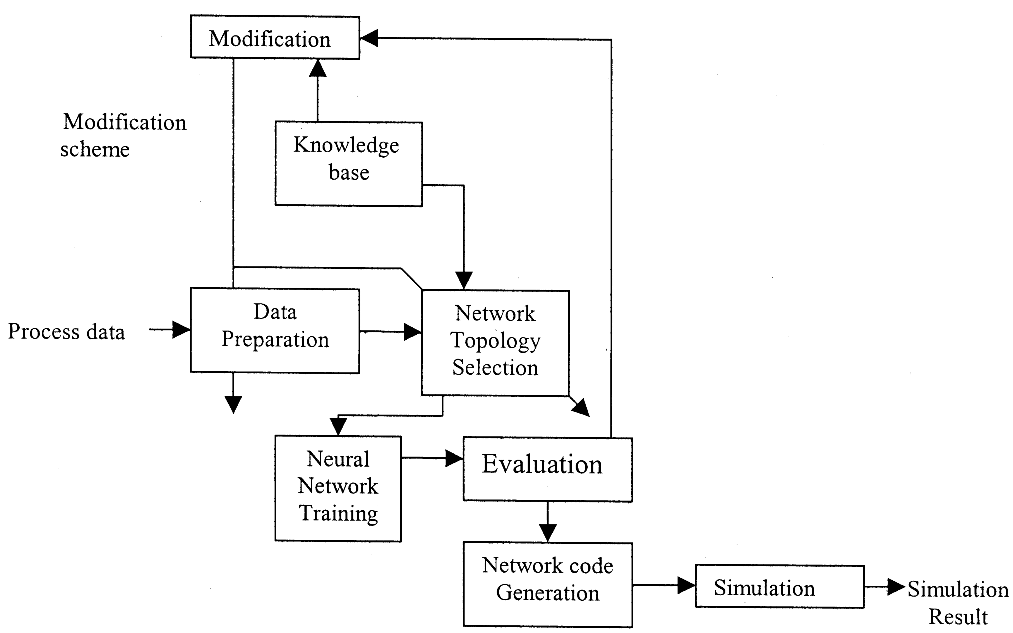


Fig. 2—Software system organization of artificial neural network integrated environment

***Modeling a Scrubber in a Paper Machine Ventilation System for Energy Conservation***

Paper machines require a large quantity of hot dry air for the drying process. This air is used for the machine hood and pocket ventilation. After drying process, mixture of hot and evaporated water, which has a specific enthalpy several times greater than the air delivered to drying hood, is conducted through the system of heat exchangers called recuperators. The last stage of system of heat exchanger usually represents a scrubber. Using feed-forward NNs[3], outlet water temperature of the scrubber process was estimated. Networks are trained with experimental data carried out on two pilot scrubbers. NNs consisted of seven inputs and one output. The inputs were mass flow rate, temperature, inlet humidity of the supply air, mass flow rate of the water, inlet water temperature, height of the scrubber and water pressure in the nozzle. The output was the outlet water temperature from the scrubber.

**Applications to Control**

***Control of Refiner***

NN model can predict pulp quality at 10 sec intervals and trains and tunes models in real time. It can also be used for estimating stock consistency using fiber properties, stock linear velocity, stock temperature, and dilution flows. Kooi & Khorasani[4] proposed a BPN as a controller to replace self-tuning regulator for a closed loop control of energy in the woodchip refining processes. Both static and dynamic NN controllers perform very well in emulating self-tuning regulator for processes having a non-minimum phase properties. Dynamic NN controller provides a satisfactory control compared to the self-tuning regulator.

***Control of Brown Stock and Bleach Washer***

Brown Stock Washer is one of the subsystems of a pulp mill to separate clean brown pulp from black liquor obtained from a digester through a multi-stage countercurrent washing. The main variables affecting the process is input consistency in the vat, rpm of washer, input flow rate, temperature, vacuum and vat level etc. Rudd[5,6] used NN to determine real time values for soda loss, washer mat consistency and washer mat unit density in brown stock washer, the former is used as soft sensor. The target is to control the black liquor solids carried out by the pulp mat to the bleach plant. Results show reduction (25%) in standard deviation of black liquor solids using 8 days trial. The controller also maintained larger disturbances in an automatic mode for the input variables. These values are used to control the washing operation. Data from a process domain is collected along with known results to develop training sets used to train configured NNs (Fig. 3).

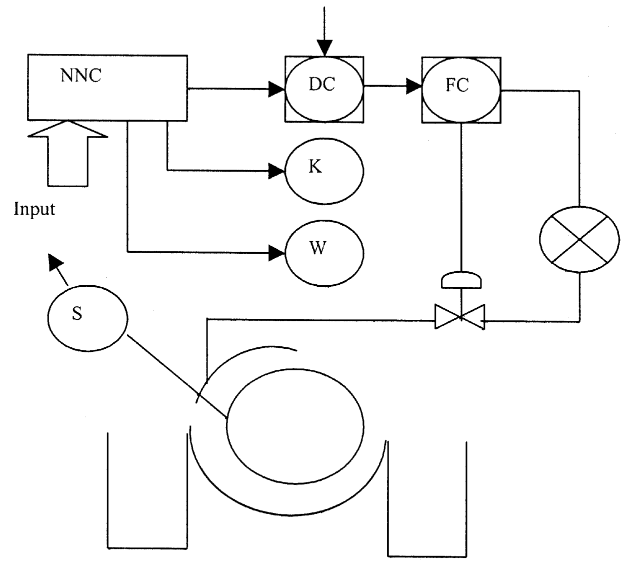


Fig. 3—Neural network based control strategy

The NN based control strategy consists of dilution factor controller (DC), consistency (K), weight (w), speed(s) and flow controller(FC). NN controller (NNC) gets a signal for a set point or target value. The output signal from NNC goes to dilution factor controller, which in turn sets out a signal to the flow controller in a cascading mode. The output signal from DC goes to control the flow of water to the filter mat. Similar configuration can be drawn for single stage bleach plant washer.

***Control of Black Liquor Combustion***

Black liquor obtained from Brown stock washer undergoes concentration (up to 55-65%) in a multiple effect evaporator system. This liquor is then led to spray in the recovery furnace for combustion. The inorganic components ($Na_2CO_3$, $Na_2S$ and very small amount of NaOH) remain in the same form or the other and sink through the char bed (principally carbon), which is formed, by decomposition, and pyrolysis of organic constituents (lignin, hemicelluloses etc.) present in black liquor. The shape and size of the char bed are the important parameters for control of combustion in recovery furnace. Ozaki *et al*[7] described a system where the bed shape

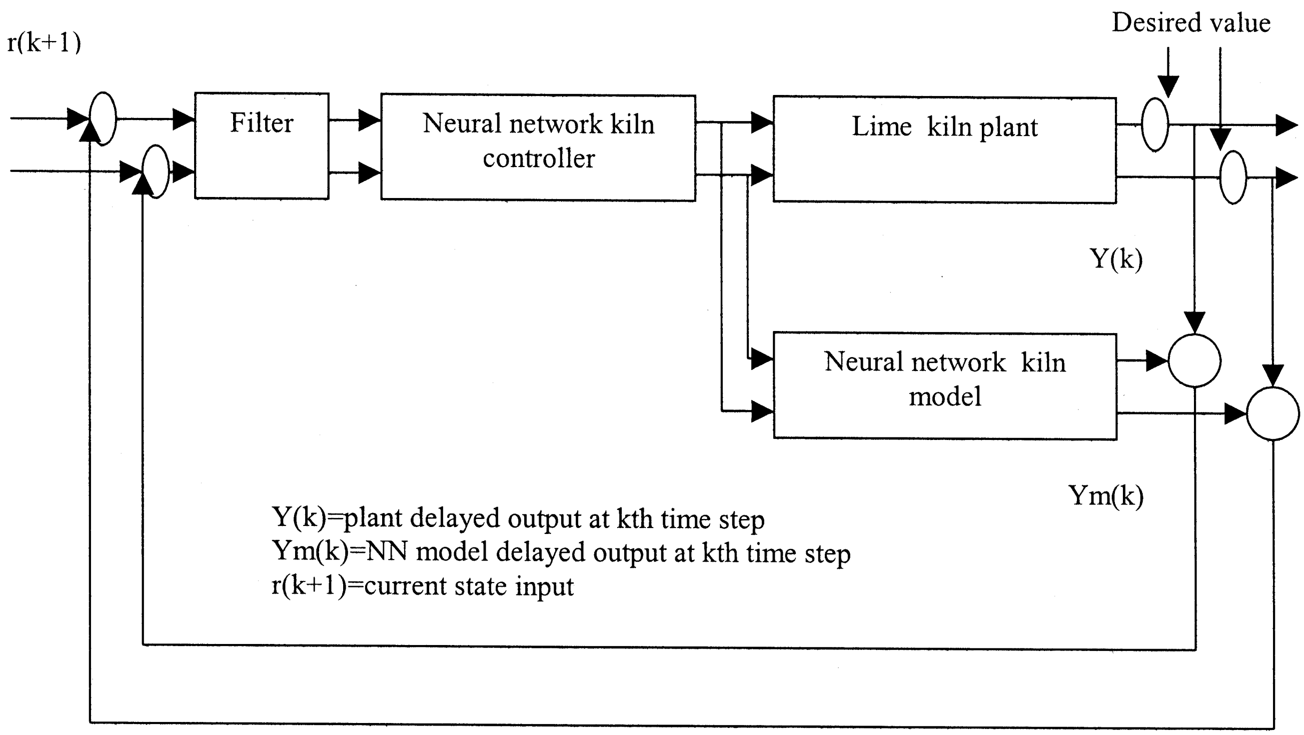


Fig. 4—Neural network based control strategy for lime kiln

from the image analysis system is first classified by three layers back propagation NN in to three classes as ideal shape, wide based and high, and high. This is followed by a fuzzy logic controller, which changes the airflows in to the recovery boiler by using the classifier information.

***Control of Lime Kiln***

Limekiln is used primarily to produce lime required to convert green liquor to white liquor by causticization process using three to four stage cocurrent causticizing reactor (back mix flow reactor-CSTR). White liquor is the feed to digester to get pulp from various fibrous raw materials. The control of limekiln is essential for getting good quality lime in terms of size and reactivity. Control itself includes both burning zone control and combustion control. Internal model control strategy[8] (IMC) for limekiln control uses multilayer feed forward BPN with 8 inputs, 2 outputs, and 2 hidden layers with 20 and 10 nodes (Fig. 4).

***Application to the Control of Grade Changes***

In paper mill, grade changes usually occur due to market demand. During these changes, many drives and valves that define the process parameters require control. Objective is to change the grade as quickly as possible from one acceptable quality to another while avoiding breaks. Grade change weakens stability of the process and therefore increases risk of breaks. While a paper machine is running, paper produced between grades that do not conform to quality standards is generally rejected and recycled in to the process. When the quality level required for a new grade has been reached, reeler begins again to accumulate paper (Fig. 5). A grade change in a paper mill demonstrates following sequences: i) Normal production; ii) Preparation for a grade change; iii) Execution of grade change; and iv) Termination of grade change. Main activities of the preparation stage are: i) Functional mapping of production line; ii) Operational plan; iii) Qualitative validation and planning; iv) Scanning of history and operator-experience database; and v) Simulation tests.

The execution stage of the grade change consists of the pre-change operations that must be carried out during a normal run and a transient grade change. In the paper machine, for example change in basis weight, filler content and speed cause a shift in the load level of the drying section. Since the dynamics of the drying section are slower than the dynamics of the flow and speed up processes, the changes in the steam

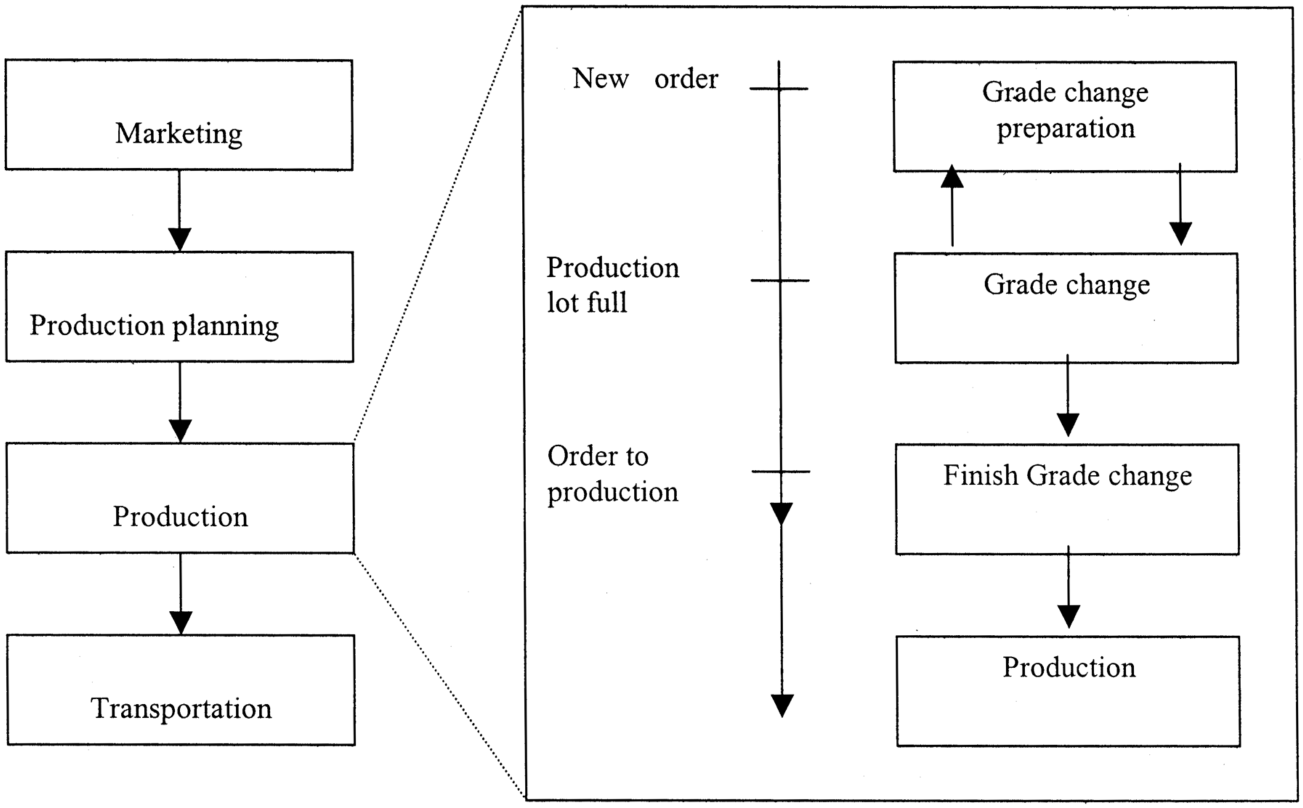


Fig. 5—General view of grade change in a paper mill.

pressure must be done in good time before the grade changes. The selection of appropriate methods for the transient state control is influenced by requirement set for the operation. To obtain the additional benefits, more efficient algorithms and calculation tools are required. It must, however, be noticed that more sophisticated methods also need more work both at the planning and the implementation stages. Benefits increase in the following order: single input-single output (SISO) PID control, multi-variable feed forward and feedback control, and model predictive control. Use of NN technology is beneficial if the process model is very nonlinear or its structure is unknown.

## Application to Quality Control

### *Quality Control of Paper Machine*

Expert systems and NN have been used in quality control system in paper mills. The primary objectives of a paper quality expert system are: i) Secure quality of paper; ii) Minimize variations between shifts; iii) Reduce production costs; iv) Support operators; v) Provide a flexible simulation tool; vi) Use existing knowledge; and vii) Train new staff members. The basic functions of one system are to collect real time process and quality data, evaluate the measured quality against customer specifications, recommend necessary corrective actions and simulate fulfilling of these actions (Fig. 6). The quality model represents the process knowledge as a quality matrix. Rules come from mill experts and known relationships between control and quality variables. Several workstations can operate with the same knowledge. One could be in the paper machine and the other could be in the paper laboratory. Fig. 7 shows the operation environment. Q-OPT is a NN kernel used in the quality control of paper machines. Fig. 8 shows a general structure that shows the main functions as real time predictions of output values using on-line measurements, optimization, sensitivity analysis and statistical analysis.

## Applications to Prediction of Paper Properties

Curl is usually a paper defect and paper can be rejected if care is not taken during processing to eliminate curling. Curl can only be measured reliably OFF-LINE after manufacture, making it difficult to control. Edward & Murray[9] predicted curl from

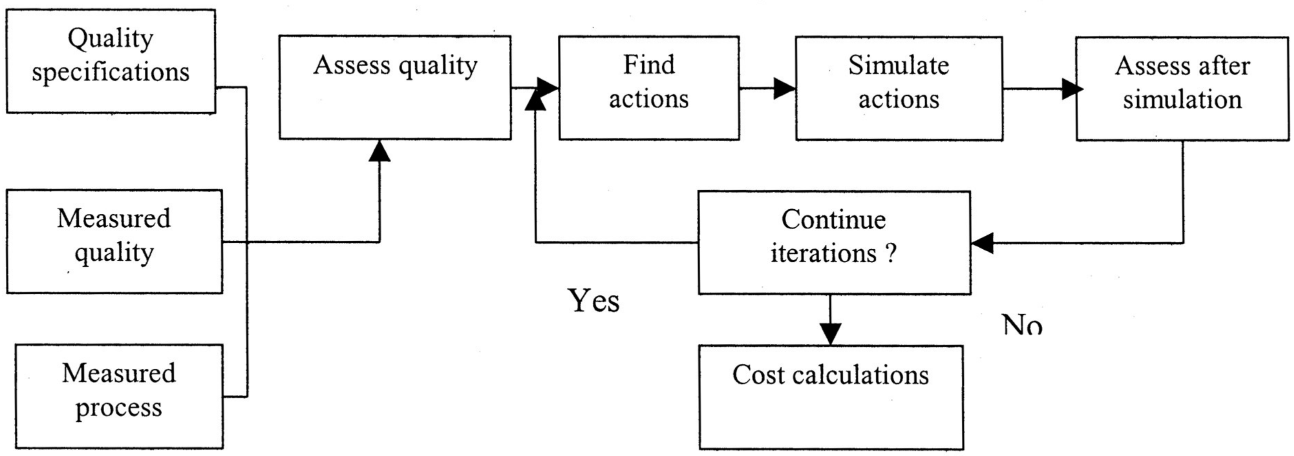


Fig. 6—Main function of the paper quality expert system

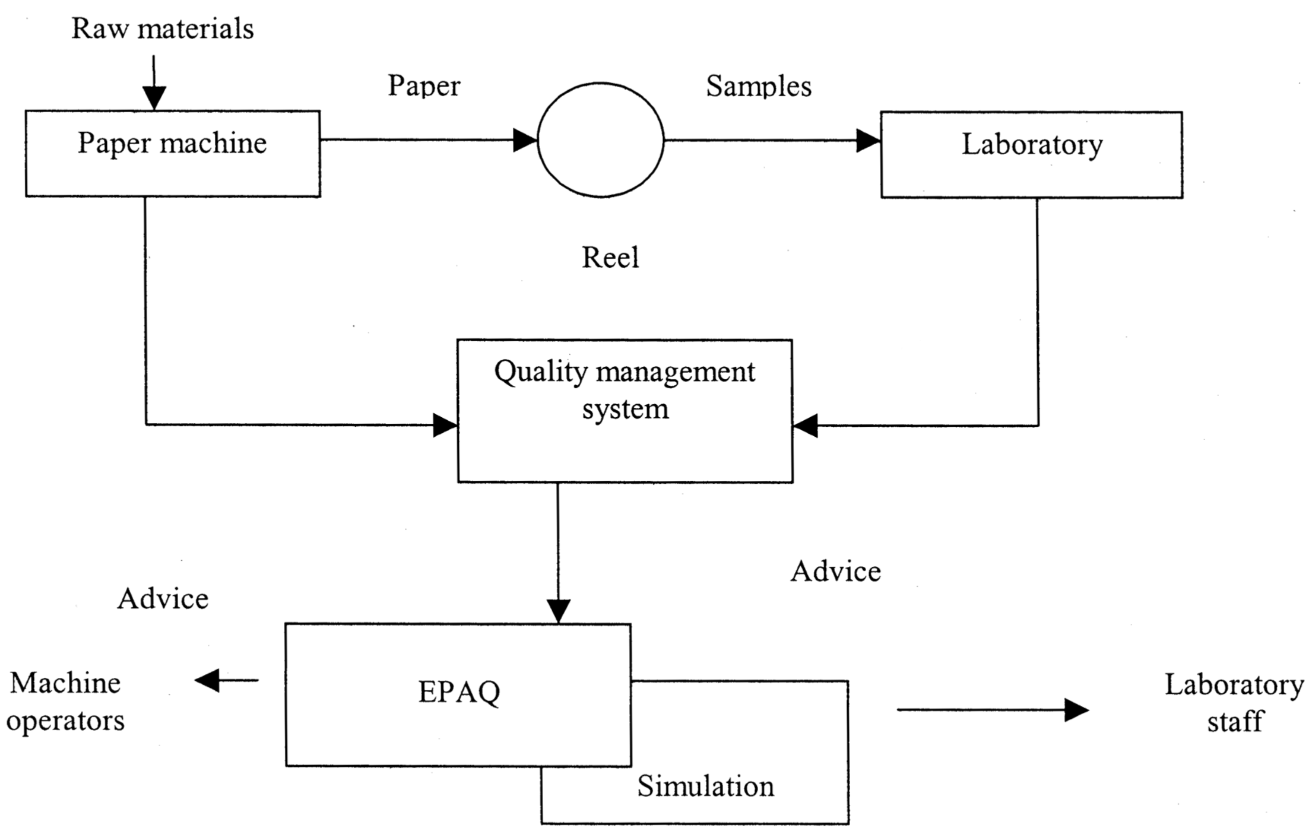


Fig. 7—Operation environment of the paper quality expert system

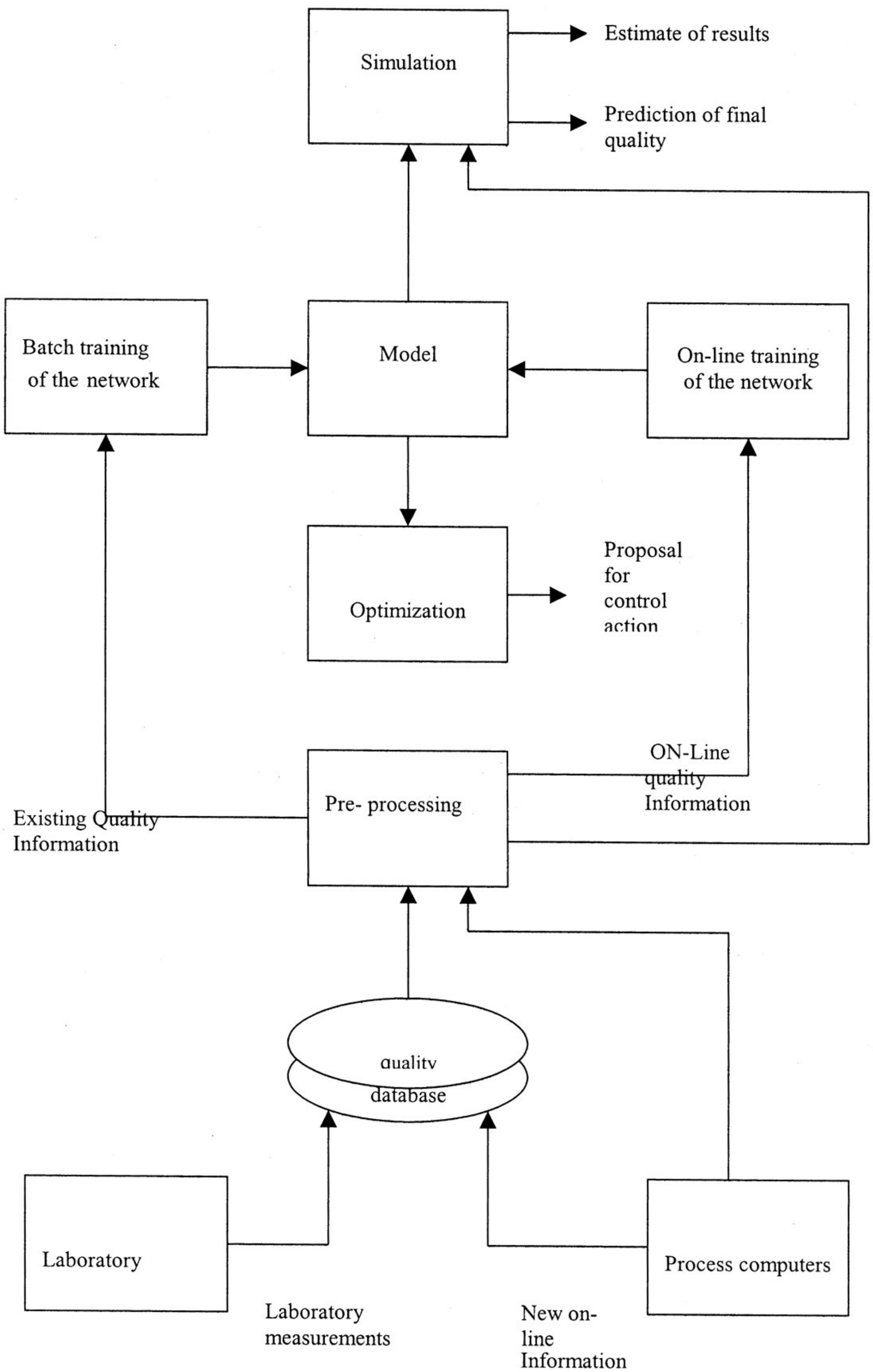


Fig. 8—Structure of a neural network system for quality control of paper mills

parameters defining the current characteristics of a reel and the plant machinery using NN techniques.

## Conclusions

A review of feed-forward back-propagation network is highlighted with supervisory strategy. Applications of neural network for modeling, simulation and fault diagnosis have been exemplified with the subsystem used in a paper mill notably modeling of kappa number in Kamyr digester, pulp bleaching process, and control of brown stock washer and bleach washer, incineration of black liquor, lime kiln, control of grade changes, quality control and finally prediction of curl a typical paper property. Advantages of using ANN based control systems are cited to demonstrate how to achieve a robust control of the above paper mill subsystems, which are difficult to control, by conventional controllers.

## References

1 Dayal B S, Macgregor J F, Taylor P A, Kildaw R & Marcikic S, Application of feed-forward neural networks and partial least squares regression for modeling kappa number in a continuous kamyr digester, *Pulp & Pap Canada*, **95** (1994) 26-32.

2 Zhu J, Xia Q, Rao M & Zurcher J, A neural network for modeling pulp process, *Pulp & Pap Canada*, **98** (1997) 35-38.

3 Milosavljevic N & Heikkila P, Modeling a scrubber using feed-forward neural networks, *TAPPI J*, **82** (1999) 197-201.

4 Kooi S B L & Khorasani K, Control of the woodchip refiner using neural networks, *TAPPI J*, (1992) 156-162.

5 Rudd J B, Using a neural network controller for advanced control applications, *Proc Int Engg Conf*, Nashville, TN, 30 Sept-3 October 1991, 101-113.

6 Rudd J B, Using a neural network systems for advanced process control, *TAPPI J,* **74** (1991) 153-159.

7 Ozaki N, Yamazu S & Tateoka K, *Proc Int Conf Control Syst* (SPCI, Stockholm) 1994, 307.

8 Ribeiro B, Dourado A & Costa E, *Proc Int. Conf Artificial Neural Networks and Genetic Algorithms* (Springer Verlag, Innsbruck) 1993, 117.

9 Edward P J & Murray A F, The application of neural networks to the papermaking industry, *IEEE Trans Neural Networks*, **10** (1999) 1456-1464.