SIMULATION OF PULPING PROCESS USING NEURAL NETWORKS AND HYBRID MODEL

Polowski, N.V.; Aguiar, H.C.; Maciel Filho,R.

Laboratory of Optimization Design and Advanced Process Control. Faculty of Chemical Engineering- State University of Campinas ,P.O. 6066 – ZIP: 13081- 970 – Campus – SP- Brazil

ABSTRACT

A Neural Network Model is developed for a Kraft digester which is a fundamental stage in the pulp production. A deterministic model is used to describe the main features of the process and will provide the data to verify the performance of the artificial neural network modeling. The paper shows that the artificial neural network is a good way to represent the process since it gives equivalent results when compared to deterministic model, and it is easier and cheaper to be developed. *Copyright © 2002 IFAC*

keywords: neural networks, hybrid model, pulp, optimization

1. INTRODUCTION

Pulp and paper industry is one of the most important productive activities and, in fact, it is used as a way of measuring the development of the country. As it is the case for a number of processes, there is a large incentive for the development of cheaper and more environmentally, more friendly operation, but still enhancing product quality. This usually requires the use of a set of tools, as advanced control strategies and real time monitoring and sometimes optimization, which needs a suitable mathematical representation of the process.

Even in a net integrated environment, as real time optimization, in order to find out which are the best operating conditions and policy, a reliable and easy to use mathematical model is required. A deterministic representation, which is developed from the mass and energy balance equations, is a convenient way to obtain a model of the process. On the other hand, artificial neural networks model is an alternative approach which may be very general and easy to be developed since many industrial process have data acquisition system normally used to process monitoring. The objective of this paper is to show the potential of the neural network model to represent the kraft digester and to describe the most important steps to build up such kind of a model aimind the development of hybrid model.

The approach based on neural network is particularly interesting when the phenomena taking place in the process are very complex so that deterministic description will require knowledge and information on heat and mass transfer coefficients, besides kinetics, to allow a reasonable process description. This is the ease for many chemical processes but particulary for pulping process. In fact in such process there is a lack of data for kinetic and especially transport data related to diffusion through the wood chips by the chemicals. In this case, industrial process data may be not enough and laboratory data are usually expensive and time consuming to be obtained. This is typically the case for the wood digester. Wood is a complex structure of fibres glued together by lignin. In order to make pulp, it is necessary to separate them, which can be done mechanically, chemically or by a combination of the two processes. The focus here is on chemical pulp process, more specifically in kraft pulps since they are important industrial process. The kraft process consists of a chemical reaction

between the lignin and a cooking liquor, composed mainly of sodium hydroxide and sodium sulfide. However, the liquor can also attack other wood components, such as carbohydrates (cellulose, hemicellulose and others), degrading them, which is not desirable since they are the final product. A high quality pulp should have, among other properties, fibres in good conditions, for that they will contribute significantly for the resistance of the final paper.

Although the composition of the wood used in the process will partially determine the properties of the paper, these latter are also altered during the manufacturing process. According to the operating conditions, the fibres can be more or less degraded, bleached or protected. It will depend on the way in which the reactions develop, attacking mostly the lignin, or the fibres, which is carried out in the digester (Aguiar and Maciel Filho, 1997).

2. KRAFT DIGESTER

The continuous digester is a large pressurized vessel, divided into zones (Figure 1). The first zone is responsible for the impregnation of the wood with liquor. It is very important that the wood receives liquor entirely and uniformly, so that, at the end of the cooking, the pulp is uniform and the amount of rejects is minimum. A heating zone follows the impregnation zone, leading the process to the suitable temperature for the lignin reaction which predominates over the carbohydrates reactions. The chips stay in the cooking zone for a period, and then proceed to the washing and discharge zones. The liquor is added and extracted from the digester in different areas, in order to keep the reaction happening properly, and its flow patterns vary from zone to zone (Assumpção et al.,1988).

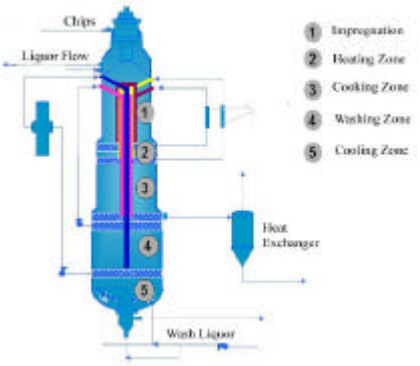

Figure 01 Continuous Kraft Digester

3. FIRST PRINCIPLES MODEL

The kraft pulping of wood is essentially based on the simultaneous occurrence of two phenomena: the first, of a physical nature, involves the mass transfer of the cooking chemicals from the bulk liquor surrounding the wood chips to the liquid filling the pores of the wood structure; and the second is related to the chemical reactions between these reactants, inside the chips, and the main wood components (i.e., lignin and carbohydrates). Because of them, it is of great importance for the understanding of the pulping process to quantify their behavior (Egas et. al., 2002).

Theoretically, first principles model should reproduce the modeled process in any situation, even in extreme conditions. For that reason, their application in simulators and control instruments are spread in various classes of processes. Many cooking models, with different approaches, can be found in literature. One of the earliest kinetic model was developed by Vroom (1957), who used an Arrhenius type expression for the reaction rate temperature dependence to derive the "H-factor". The H factor, which combines the cooking temperature and cooking time into a variable, is at the heart of many control schemes used for kraft pulping. H-factor models that are highly empirical in nature are very useful for batch pulping, but they are not suitable for continuous flow digesters such as the kamyr digester. The major disadvantage of using the H-factor is that it does not include mass transfer and the variation of chemical concentration during the process.

There are several different approaches to the chemistry of the process in the literature. Most of them use the same general form (Saltin, 1992), although some simplifications and modifications are usually applied (Aguiar, Maciel Filho, 1997):

$$-dW/dt=(k_1[OH^-]^a+k_2[OH^-]^b[HS^-]^c)(W-W_\infty) \quad (1)$$

where: W is the wood component concentration; k1and k2 are rate constants that are correlated as a function of temperature using the Arrhenius equation $\left(k = Ae^{-\Delta E/RT}\right)$; W_∞ is the wood component after infinite time; a,b and c are the kinetic orders; $[OH^-]$ and $[HS^-]$ are the OH^- and HS^- concentrations, respectively, and t is time.

Some approaches consider a three step reaction, with three differential equation of the same type above, one for each phase (Gustafson et al., 1983). Gustafson's model predicts the effect of

inadequate diffusion on pulping properties but cannot predict the consequences of incomplete penetration on the pulping performance and final pulp properties. Another approach divide the lignin into high and low reactivity lignin, according to the kind of ethers they contain (Saltin, 1992; Mirams and Nguyen, 1994). The model developed by Mirams and Nguyen (1994), which is based on batch cooking process includes mass transfer and intrinsic reaction kinetics based on parallel reactions of lignin, hemicellulose and cellulose.

This study considers the simultaneous reactions of two lignins, one reacting faster and the other suffering a slow reaction. There is also a third type of lignin, in a very low concentration, that does not react with cooking liquor. It is assumed the presence of two types of cellulose and hemicellulose, which do not consume the sulfide ion. Unlike most models, this one considers a simple kinetic equation for the extractives. Wood influence is accounted for in the model through density and porosity (Mirams and Nguyen, 1994; Aguiar and Maciel Filho, 1998). A diffusion term is included to describe the transport of cooking chemicals, assuming the same diffusion coefficient for both, OH^- and HS^-.

The deterministic model equations can be written as follows:

- Lignin

$$\frac{dL_1}{dt} = k_{L_1} \exp\left(\frac{E_{L_1}}{RT}\right)[OH]^a [HS]^b L_1 \quad (2)$$

$$\frac{dL_2}{dt} = k_{L_2} \exp\left(\frac{E_{L_2}}{RT}\right)[OH]^a [HS]^b L_2 \quad (3)$$

- Hemicellulose

$$\frac{dH_1}{dt} = k_{H_1} \exp\left(\frac{E_{H_1}}{RT}\right)[OH]^a H_1 \quad (4)$$

$$\frac{dH_2}{dt} = k_{H_2} \exp\left(\frac{E_{H_2}}{RT}\right)[OH]^a H_2 \quad (5)$$

- Cellulose

$$\frac{dC_1}{dt} = k_{C_1} \exp\left(\frac{E_{C_1}}{RT}\right)[OH]^a C_1 \quad (6)$$

$$\frac{dC_2}{dt} = k_{C_2} \exp\left(\frac{E_{C_2}}{RT}\right)[OH]^a C_2 \quad (7)$$

- Extractives

$$\frac{dE}{dt} = -0.6E \quad (8)$$

- Hydroxide

$$\frac{\partial OH}{\partial t} = \frac{1}{X^2}\frac{\partial}{\partial x}\left(D\frac{\partial OH}{\partial x}\right) + \quad (9)$$
$$+\frac{r}{0.1e}\left[3.2\times10^{-3}\frac{dL}{dt} + 6.1\times10^{-3}\left(\frac{dH}{dt}+\frac{dC}{dt}\right) + 7.0\times10^{-3}\frac{dE}{dt}\right]$$

$$\frac{\partial HS}{\partial t} = \frac{1}{X^2}\frac{\partial}{\partial x}\left(D\frac{\partial HS}{\partial x}\right) \quad (10)$$

The orthogonal collocation technique was used to solve the partial differential equations (the equations involving diffusion), and the complete system was solved by the LSODE integrator. One inconvenience when generating deterministic models is that parameter estimation methods require acquisition of data not readily available from mill DCS. For instance the concentrations of substances that are not usually measured at the mill. In order to overcome such difficulties, laboratory data is used to adjust mill process parameters (Aguiar, 1997).

4. NEURAL NETWORKS

The most common network used for process modeling is the feedforward neural net. Such net contains at least one input layer, one output layer and one hidden layer. The connections between neurons (or nodes) in this type of network are transmitted forward, from each neuron in a layer to all neurons in the following layer, as shown in Figure 2. Each process variable value is given to one node in the input layer. The neurons on the first hidden layer receive a weighted summation of the signals from input nodes, added to a bias term. Inasmuch as the weights are specific for each connection, and the bias terms are particular for each receiving node, all the neurons will receive a distinct value. The summed values are altered by a transfer function which transforms the signal to a value in a limited range (Figure 3). The sigmoidal function, which transfers numbers to the [0,1] range, is usually used, but other functions can be applied, such as sine, tangent, etc. The transformed value will be the output of the node and will be transmitted to the next layer in the same manner, and from one layer to the other until the output layer releases the net output.

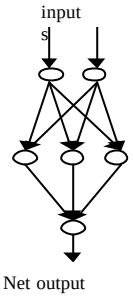

Figure 2 Feedfroward neural network information transference.

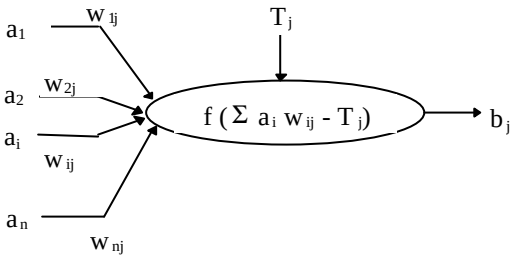

Figure 3 A single neuron

Many backpropagation algorithms can be found in literature. The algorithm used in this work is the Generalized Delta Rule (GDR) algorithm, a gradient descent method that minimizes the sum of squares of the residual (Savkovic-Stevanovic, 1994). The amount and quality of training sets are essential for the success of the artificial neural network simulations. The number of hidden layers and the neurons in each hidden layer can be tested in order to obtain better results. Also the learning rate, a coefficient equivalent to a step size and the momentum term, which keeps the direction of descent from changing too rapidly from step to step can be altered in order to improve the net efficiency. These parameters affect not only the accuracy of outputs, but also the training time, which can be considerably long in some cases especially for on-line trainny (Aguiar and Maciel Filho, 1997).

This ability to recognize and to reproduce cause-effect relationships through training, and filtering noise or irrelevant data, and yet the possibility of working with multiple input and output vectors (MIMO), make neural networks efficient in representing even nonlinear systems. That justifies the choice of neural nets to model the relationship between process operation and its product characteristics in complex processes. Furthermore, the models based on this method may be developed by using process variables that are familiar to mill personnel and easily accessed as inputs, in order to predict outputs that are usually difficult or expensive to obtain. Used in this fashion, the neural network can work as a process predictor. In this paper the neural network used to predict kappa number from process conditions will be described. The usual measurement of pulp properties and delignification extent is the kappa number, which indicates the amount of lignin left in the pulp.

Nevertheless, neural nets have their limitations. The training step is very time consuming, especially with industrial noisy data. They can not extrapolate far beyond their training range, and also do not allow interpretation of the phenomena they are simulating. An alternative approach to pick up advantage of both deterministic and Artificial Neural Network approach is to develop a hybrid model as follow.

5. HYBRID MODEL

Although mathematical models can describe relatively well some behavior of the process, they make a number of assumptions that turn them into approximated models, and often are not able to reproduce properly the nonlinearities of the systems. On the other hand, neural networks represent non linear patterns properly, but do not allow extrapolation. Table 1 describes the 9 inputs variables used on Hybrid Model..

Table 1 - Selected input variables

Variables	Unities
Total effective alkali	g/l NaOH
Temperature (top)	^{o}C
H Factor (top)	–
Temperature (CD1)	^{o}C
H Factor (CD1)	–
Effective alcali (CD1)	g/l NaOH
Temperature (CD2)	^{o}C
H Factor (CD2)	–
Effective alcali (CD2)	g/l NaOH

This work proposes a hybrid model, based on the known part of a first principle model, combined to a neural network to account for the phenomena that is not comprised in the deterministic equations, in order to overcome limitations of both methods.

Hybrid models may be developed in many different combinations. Knowledge of the process may be used to impose constraints (as inequalities) to the neural model. The network can also be used as default, but the deterministic model should be used when the variables are out of training range. Also, the mathematical model may be used in situations when data is not available. A common approach is having the first principles model as the basis, with the neural net calculating unknown parameters. Still, the network can be used to learn the deviation between the mathematical model output and the aimed output. In fact, this may be easier than learning the complex relationships between input variables and process results. Other possible approach is using the deterministic model as reinforcement for the function relationship between inputs and outputs.

The choice of how to combine both parts depends on the precedence of data or first principle system knowledge, and it is particular to each case. The approach considered in this work uses the mathematical model as a reinforcement for the relationship between inputs and outputs. The mathematical model receives information from process variables and generates the kappa number according to the kinetic equations. The result is then fed into the net along with other process variables, and the net yields a final kappa number. The scheme is described in Figure 4.

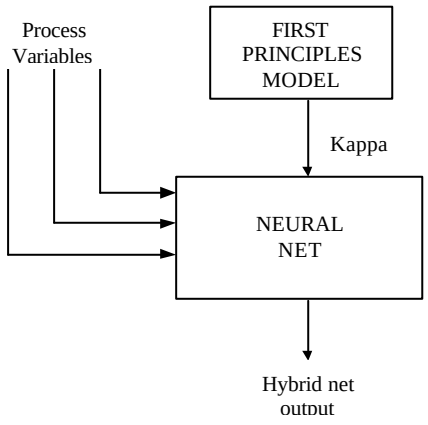

Figure 4 Hybrid model scheme

The mathematical model receives information from process variables and generates the kappa number according to the kinetic equations. The result is then fed into the net yields a final kappa number (Figure 4). Training is carried out as described for a pure network training section. However the mathematical model calculates an extra input variable, which is the calculated kappa number, for each training set. A hybrid network using the neural net to estimate the unknown parameters is other development not considered in this work.

6. RESULTS

Many adjustments in network configuration and analysis of the input variables set must be done in order to obtain the best network configuration for the modeled process. The net was initially set to be trained with one hidden layer and all 22 available variables. Table 2 describes all the 22 inputs variables. However, it did not converge to the predetermined minimum error. Changes in net configuration were done gradually until best results were obtained. The parameters to be optimized in a neural network are: number of layers and neurons in each layer, momentum term, offset, learning rate, error tolerance, transfer function and random initial weights algorithm.

Table 2 - Initial input variables

Variable	Unities
Production rate	RPM/h
Total white liquor flow	m^3/h
Wood density	kg/m^3
Total Aakali charge	%
Total effective alkali	g/l NaOH
Sulfidity	%
Temperature (top, CD1, CD2)	^oC
H Factor (top, CD1, CD2)	-
Aakali charge (top, CD1, CD2)	%
Effective alkali (top, CD1, CD2)	g/l NaOH
White liquor flow (top, CD1, CD2)	m^3/h

The predicted variable in the model was the kappa number, due to its importance to the next sections of the process and its significance for mill personnel. Regression and graphical analysis of the data were carried out in order to eliminate the variables that do not affect kappa number significantly. As a result, only 9 variables were used to feed the input neurons in the network. Figure 5 illustrates the improvement obtained for kappa number prediction with previous data analysis.

The number of industrial profiles was not the desired number of training sets for a neural network but it reflects the potential to be used even with relatively few data. In any case, it is expected that when more

data is gathered, network prediction will improve considerably. The hybrid network was trained in the same way. The tendency dictated by the mathematical model result is responsible for a slight improvement in network prediction, as indicated in Figure 6. The training time, when measured as number of computational iterations, was shorter for the hybrid network than for the pure neural net. This is a relevant information when on-line applications either for advanced control or process optimization are desired.

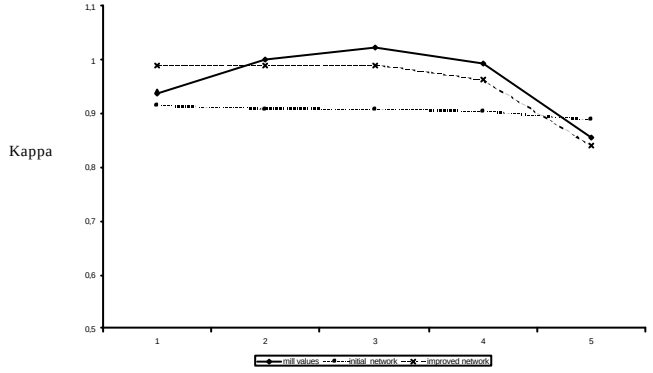

Figure 5 Kappa prediction

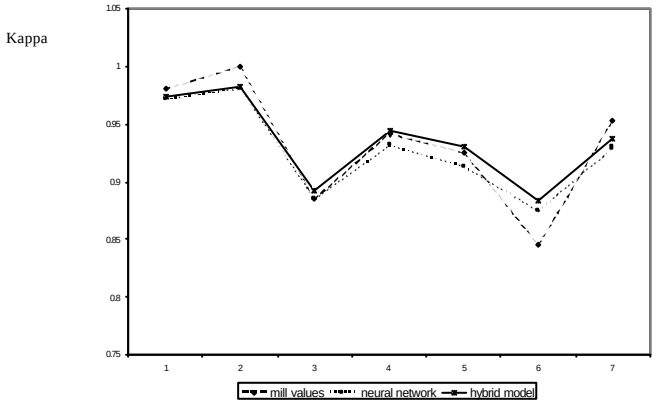

Figure 6 Comparison between hybrid model and neural model predictions

7. CONCLUDING REMARKS

Both, neural network and hybrid network models produced similar results, showing to be adequate tools to model an industrial cooking process, even with a reduced number of training sets. The simple approach used to combine first principles knowledge and neural network features produced improved results when compared to the pure net. However, it should be seen as the starting point for a more complex model, where process knowledge and pattern learning capacity can be better combined in order to improve prediction capability and to provide a more complete approach of the process (when all variables and relevant phenomena influencing results are accounted for).

The expected benefit is that simulation program assists in decision making on process policies according to the desired product characteristics, and enables mill personnel to correct the process before undesired pulp grades are produced. Furthermore the software can be used to design specific pulp characteristics that are requested by customers, and can be achieved by specific parameters controlled in the process. It can also be a useful tool for training operators and engineers. The software was written in Fortran 77 (public domain) with an easy to use interface.

9. REFERENCES

Aguiar, H.C. and Maciel Filho, R., (1997). Modelling and Optimization of Pulp and Paper Processes Using Neural Networks, ESCAPE 8.

Assumpção, R.M.V. et al., (1998). Celulose e Papel – Tecnologia de fabricação da pasta celulósica, chapter VI, Vol. 1, IPT.

Egas, A. P. V., et al., (2002). Experimental Methodology for Heterogeneous Studies in Pulping of Wood, *Ind. Eng. Chem. Process Des. Dev.41.2529-2534*.

Gustafson, R.R. et al., (1983). Theoretical Model of Kraft Pulping Process, *Ind. Eng. Chem. Process Des. Dev.* 22(1):87.

Jiménez, G., Gustafson, R.R. and McKean, W.T., (1989). "Modeling Incomplete Penetration of Kraft Pulping Liquor", *Journal of pulp and paper science* 15(3):J110.

Mirams, S. and Nguyen, K. L., (1994). Application of Predictive Kinetic Models in Eucalypt Kraft Pulping, *International Pan Pacific Conference, Tappi Proceedings*, 73-82.

Saltin, J.F., (1992). A Predictive Dynamic Model for Continuous Digesters, *Pulping Conference, Tappi Proceedings*, 261-268.

Savkovic-Stevanovic, J, (1994). *Computers and Chem. Engng* 18(11/12):1149

Vroom, K. E., (1957). The 'H' Factor: A Means of Expressing Cooking Times and Temperature as a Single Variable, *Pulp Pap. Mag. Can.*, 58, n.3, 228.