A Method of Transfer Functions and Block Diagrams to Study the Contribution of Variables in Artificial Neural Network Process Models

Edwin G. Boza Condorena and Aline Carvalho da Costa

***Abstract*—In this paper is proposed the use of transfer functions and block diagram algebra to describe cause and effect relationships in artificial neural network process models. Explicit formulae are derived for feedforward neural networks with an arbitrary number of inputs, outputs and hidden layers. This novel approach provides a general interpretation of the information existing in neuron interconnections. Numerical applications for the particular case of transfer functions expressed in terms of relative contributions of weights, showed that the method was applicable in the evaluation of the effect of inputs on a given output, with good results.**

***Index Terms*— Variables contribution; Backpropagation; Artificial neural networks; Process modeling.**

I. INTRODUCTION

Knowledge of the relative importance of input variables in a neural network model provides two basic advantages [1] - [4]: a) The information can be used to build the optimal neural network model via the selection of inputs, which can improve the generalization capability of the model and allow for faster training of the neural network, with economic savings if measurement of the variables is expensive; b) It provides a better understanding of the process model since the irrelevant variables are identified.

Another point to be considered is that when a great number of input variables are available, but the size of the training set is limited, the likelihood of overfitting is increased [5] [6]. In this case, the selection of variables that have the strongest influence on the output is of extreme importance.

Many researchers have tried to express the importance of input variables in neural network models and proposed the use of various methods [3][4][7] -[10].

In the present paper is described a method that uses transfer functions and block diagrams to express input-output dependencies in ANNs, then it is shown that by using this method the concept of analyzing weights to estimate the importance of the inputs can be extended to topological net structures with multiple inputs, multiple outputs and multiple hidden layers.

Manuscript received July 19, 2011; revised August 07, 2011. This work was supported in part by the CAPES and FAPESP.

E.G. Boza-Condorena is the corresponding author.Phone: 19-8119-1822; e-mail: ebozac2003@ yahoo.es.
He is with the School of Chemical Engineering, State University of Campinas, 13083-970, Campinas, SP, Brazil
A.C. da Costa is with the School of Chemical Engineering, State University of Campinas, 13083-970, Campinas, SP, Brazil. (e-mail: accosta@feq.unicamp.br).

II. THE TRANSFER FUNCTION METHOD

A. *Basic concepts*

Transfer functions

A transfer function **G** is a mathematical statement that relates an input, **x,** to an output, **y,** of a system [11][12] .

$$G = \frac{y}{x} \qquad (1)$$

The transfer function **G** transforms the input **x** (cause) into an output **y** (effect) as shown in (2).

$$y = G * x \qquad (2)$$

Block diagrams

The block diagram can be used to describe cause and effect relationships throughout a dynamic system [13]. Figure 1 shows the block diagram for (2).

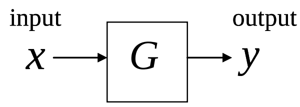


Fig. 1. Transfer function

The transfer functions in block diagrams are represented by blocks. The lines that interconnect blocks represent signals that flow interconnecting the system elements. Other components are the pickoff point and the summing point. The four components of a block diagram are shown in Fig. 2. A signal from input x_1 flows into the pickoff point (***P***) and two signals flow out of the point. In the summing point (***S***) two signal flows are added, the result is one signal that flows out of the point.

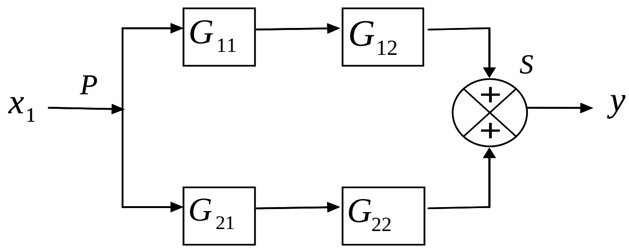


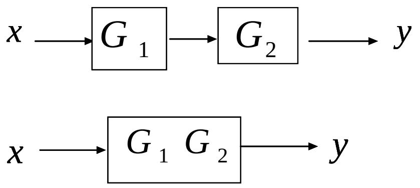
Fig. 2. Blok diagram for transfer functions between input $\mathbf{x_1}$ and output **y**.

Basic rules of block diagram algebra

The cause and effect relationship between input $\mathbf{x_1}$ and output **y** from the block diagram in Fig. 2 is obtained using

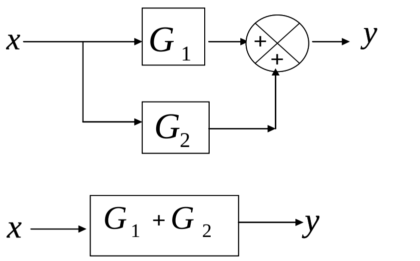
ISBN: 978-988-19251-7-6
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

two basic rules of block diagram algebra, which are: blocks in series combine with each other by multiplication and bocks in parallel combine with each other by algebraic addition. Fig. 3 and Fig. 4 show these two basic rules and the mathematical representation.



$$y = G_1 * G_2 * x = G_2 * G_1 * x$$

Fig. 3. Rule 1: blocks in series



$$y = (G_1 + G_2)x = G_1 * x + G_2 * x$$

Fig. 4. Rule 2: blocks in parallel

Equation (3) shows the result of applying the basic rules of block diagram algebra first to the blocks in series and then for the resulting blocks in parallel of Fig. 2. Equation (4) is the result of applying the associative property and Equation (5) is the mathematical statement of the transfer function between input $\mathbf{x_1}$ and output $\mathbf{y}$. Fig. 5 shows the reduced block diagram obtained from Fig. 2.

$$y = G_{11} * G_{12} * x_1 + G_{21} * G_{22} * x_1 \qquad (3)$$

$$y = (G_{11} * G_{12} + G_{21} * G_{22}) * x_1 \qquad (4)$$

$$\frac{y}{x_1} = G_1 = (G_{11} * G_{12} + G_{21} * G_{22}) \qquad (5)$$

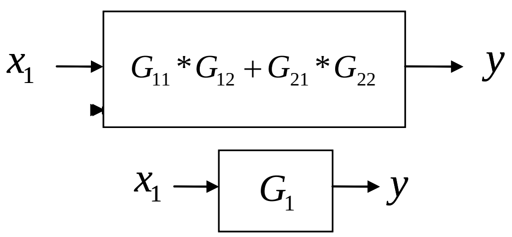


Fig. 5. Reduced block diagram

Fig. 5 shows that the effect of the input $\mathbf{x_1}$ on the output $\mathbf{y}$ is expressed by a transfer function $\mathbf{G_1}$.

B. Generalized equations for the effect of inputs on the outputs (E_{OI})

Notation used

Figure 6 shows the notation used for transfer function in neural networks structure.

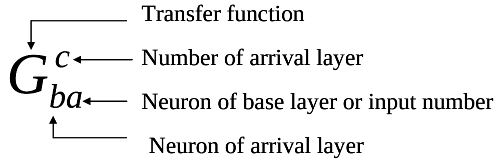


Fig. 6. Notation used for transfer function the effect of the input on the output (E_{OI})

Fig. 7 shows a portion of a neural network structure with one input, two neurons in the hidden layer and one neuron in the output layer. It also presents the transfer function associated to each connection between neurons of adjacent layers.

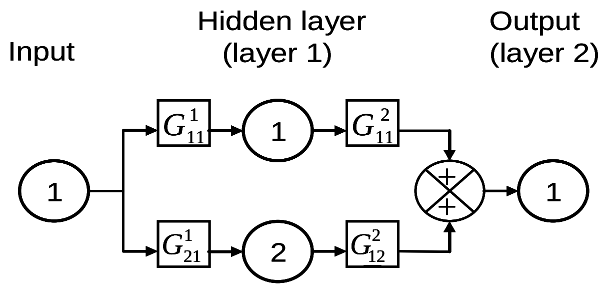


Fig. 7. Block diagram to show transfer functions in the neural network structure.

From Fig. 7, we obtain the effect of the input on the output (E_{OI}), using the basic rules of block diagram algebra.

$$E_{OI} = G_{11}^1 * G_{11}^2 + G_{21}^1 * G_{12}^2 \qquad (6)$$

Fig. 8 shows a neural network with two layers and a topology 3: 2: 2.

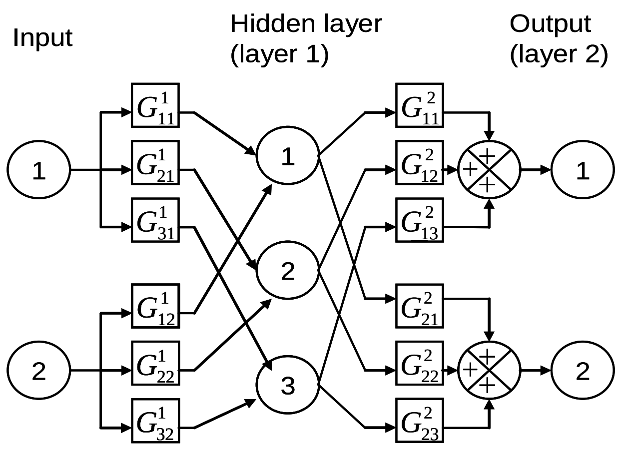


Fig. 8. Transfer functions in neural network with two inputs, three neurons in the hidden layer and two neurons in the output layer

From Fig. 8, the transfer functions for the effects of each input on the output are the following:

$$E_{11} = G_{11}^1 * G_{11}^2 + G_{21}^1 * G_{12}^2 + G_{31}^1 * G_{13}^2 \qquad (7)$$

ISBN: 978-988-19251-7-6
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

$$E_{12} = G_{12}^1 * G_{11}^2 + G_{22}^1 * G_{12}^2 + G_{32}^1 * G_{13}^2 \quad (8)$$

$$E_{21} = G_{11}^1 * G_{21}^2 + G_{21}^1 * G_{22}^2 + G_{31}^1 * G_{23}^2 \quad (9)$$

$$E_{22} = G_{12}^1 * G_{21}^2 + G_{22}^1 * G_{22}^2 + G_{32}^1 * G_{23}^2 \quad (10)$$

From (7) to (10) we can infer that the general equation for effects from input **I** on the output **O**, when the feedforward neural network has two layers, with ***n*1** neurons in the hidden layer is,

$$E_{OI} = \sum_{j=1}^{n1} G_{jI}^1 * G_{Oj}^2 \quad (11)$$

With a similar procedure were obtained general equations for more complex topologies.

The general equation for effects from input **I** on the output **O**, when the feedforward neural network has three layers, with ***n*1** neurons in hidden layer 1 and ***n*2** neurons in layer 2 is:

$$E_{OI} = \sum_{i=1}^{n1} \sum_{j=1}^{n2} G_{iI}^1 G_{ji}^2 G_{Oj}^3 \quad (12)$$

The general equation for the effect of input **I** on the output **O**, when the neural network has **r** layers, with **n(1)** neurons in hidden layer **1**, **n(2)** neurons in the hidden layer **2**, **n (r-1)** neurons in the **r-1** hidden layer, is:

$$E_{OI} = \sum_{i(1)=1}^{n(1)} \cdots \sum_{i(r-2)=1}^{n(r-2)} \sum_{i(r-1)=1}^{n(r-1)} G_{i(1)I}^1 \ldots G_{i(r-1)i(r-2)}^{r-1} G_{Oi(r-1)}^r \quad (13)$$

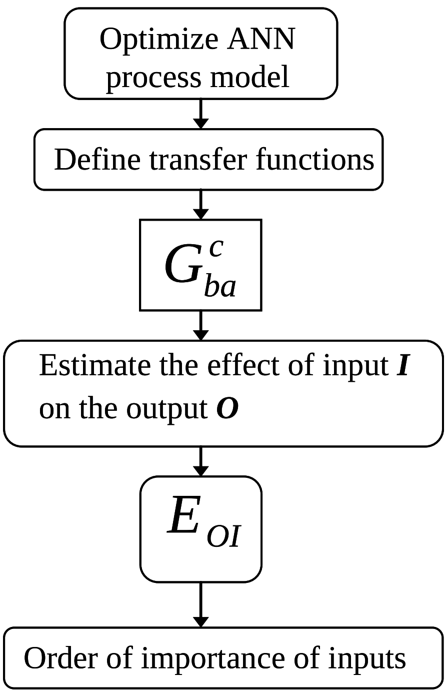


Fig. 9. Algorithm to apply the method of transfer functions

III. NUMERICAL APPLICATIONS

A. ***Particular case****: Relative contributions expressed in terms of absolute values of the weights as transfer function.*

The weights associated to each connection between neurons of adjacent layers were used in transfer functions. Fig. 10 shows the notation used.

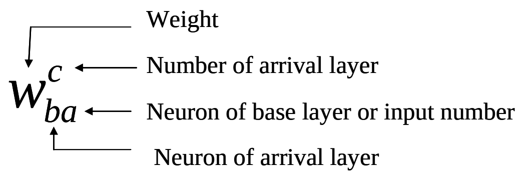


Fig. 10. Notation used for weights in neural networks

Equation (14) shows the transfer function expressed as a relative contribution of the absolute values of the weights.

$$G_{ba}^c = \frac{\left| w_{ba}^c \right|}{\sum_{a=1}^{na} \left| w_{ba}^c \right|} \quad (14)$$

Where:

G is a transfer function, **c** is the number of the arrival layer, **b** is a neuron of the arrival layer, **a** is a neuron of the base layer or an input number, **na** is the number of neurons in the base layer or the inputs, $\left| w_{ba}^c \right|$ is an absolute value of the connection weight, $\sum_{a=1}^{na} \left| w_{ba}^c \right|$ is the sum of the weights that correspond to the connections between the neuron **b** of the arrival layer and each one of the **na** neurons of the base layer.

The relative contribution of an input **x** to an output **y**, when compared to the relative contributions of the other inputs to the same output, expresses the relative effect of input **x** on output **y**, and hence its importance in the set of inputs, as a factor to produce a response **y**.

Expressing (11) – (13) in terms of relative contributions, we obtain (15) – (17) to estimate the effects of the input variables on the output variables (E_{OI}):

From (11) the general equation in terms of relative contributions as transfer functions for the effects of input ***I*** on the output ***O***, when the feedforward neural network has 2 layers, with ***n*1** neurons in the hidden layer (layer 1), and ***ni*** inputs, is,

$$E_{OI} = \sum_{j=1}^{n1} \frac{\left| w_{jI}^1 \right|}{\sum_{i=1}^{ni} \left| w_{ji}^1 \right|} * \frac{\left| w_{Oj}^2 \right|}{\sum_{j=1}^{n1} \left| w_{Oj}^2 \right|} \quad (15)$$

From (12) the general equation in terms of relative contributions as transfer functions for the effects of input **I** on the output **O**, when the feedforward neural network has

3 layers, with ***n*1** neurons in the hidden layer 1, ***n*2** neurons in the hidden layer 2, and ***ni*** inputs, can be written:

$$E_{OI} = \sum_{i=1}^{n1}\sum_{j=1}^{n2} \frac{\left|w_{iI}^{1}\right|}{\sum_{h=1}^{ni}\left|w_{ih}^{1}\right|} * \frac{\left|w_{ji}^{2}\right|}{\sum_{i=1}^{n1}\left|w_{ji}^{1}\right|} * \frac{\left|w_{Oj}^{3}\right|}{\sum_{j=1}^{n2}\left|w_{Oj}^{3}\right|} \qquad (16)$$

Finally, from (13), the general equation in terms of relative contributions as transfer functions for the effect of input **I** on the output **O**, when the neural network has **r** layers, ***ni*** inputs, **n(1)** neurons in layer 1, **n(2)** neurons in layer 2, … **n(r-1)** neurons in layer **r-1** is:

$$E_{OI} = \sum_{i(1)=1}^{n(1)} \cdots \sum_{i(r-2)=1}^{n(r-2)} \sum_{i(r-1)=1}^{n(r-1)} \frac{\left|w_{i(1)I}^{1}\right|}{\sum_{g=1}^{ni}\left|w_{i(1)g}^{1}\right|} * \cdots$$

$$\cdots * \frac{\left|w_{i(r-1)i(r-2)}^{r-1}\right|}{\sum_{i(r-2)=1}^{n(r-2)}\left|w_{i(r-1)i(r-2)}^{r-1}\right|} * \frac{\left|w_{Oi(r-1)}^{r}\right|}{\sum_{i(r-1)=1}^{n(r-1)}\left|w_{Oi(r-1)}^{r}\right|} \qquad (17)$$

Example 1: Neural network with 5: 2: 1 topology, with five inputs, two hidden neurons and one output neuron

In order to test the ability of the proposed method to determine the order of importance of the influence of inputs on outputs, a neural network was trained to describe the process of Isar et al. (2006) [14]. The connection weights of the neural network are shown in Table 1.

Table 1
Example 1: ANN weights associated to each connection between neurons

Hidden layer			Output layer	
Weights (W_{ij})	Destination		Weights (W_{ij})	Destination
	Neuron 1	Neuron 2		Neuron 1
Source inputs	j =1	j =2	Source	*O* =1
I = 1	4.2452	0.2094	j = 1	0.0436
I = 2	3.2627	0.0638	j = 2	-1.4816
I = 3	14.7474	0.2515		
I = 4	9.5223	0.0907		
I = 5	0.9688	-0.2702		
$\sum_{i=1}^{k=5}\lvert w_{ji}^{1}\rvert$	32.7464	0.8856	$\sum_{j=1}^{n=2}\lvert w_{Oj}^{2}\rvert$	1.5252

The effects of the inputs on the output, E_{OI} , are obtained using (15) . In this case: ni = 5 (number of inputs), $n1$= 2 (number of neurons in the hidden layer), O = 1 (number of outputs).

Fig. 2 shows the order of importance of the effects of all the inputs on the output using the transfer function method (Equation 15), the Garson method [7] and the method that uses raw connection weight values [10]. In order to compare the obtained results with the reference values, all the data set was normalized in the range [0.1, 0.9].

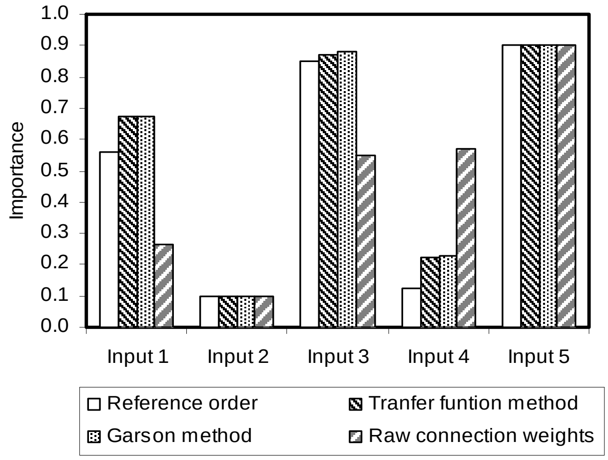


Fig. 11. Comparison of the importance of the inputs' effects using different methods.

The reference order was established using the coefficients of the mathematical model (expressed in terms of coded factors) obtained using experimental design methodology by Isar et al. (2006) [14].

In this example, the transfer function method and the Garson method [7] predicted the same order of importance of the inputs as the reference mathematical model [14]. The use of the raw connection weight values proposed by Olden et al. (2004) [10] predicted the order of importance of the inputs erroneously. Two factors that probably contributed to the wrong result were the big difference in magnitude between the highest and the lowest weight values in Table 1 and the great dispersion of weight values around the mean (variance = 23.33). The use of relative contributions expressed in terms of absolute values of connection weights, as transfer functions, as proposed in this work, led to better results in this case.

Example 2: Neural network with 11:5:5:1 topology, to model a batch fermentation process for bioethanol production from sugarcane bagasse

Introduction

The second case study is a process from the work of Andrade et al. (2009) [15]. These authors proposed a mathematical model to describe a batch ethanol fermentation process, and used Plackett-Burman designs [16] to estimate the order of importance of the effects of the kinetic parameters of the mathematical model on the concentrations of biomass (*X*) and ethanol (*P)*, as well as on the substrate consumption time (t).

In this numerical application, the order of importance of the kinetic parameters on one of the outputs of the process (the substrate consumption time) was determined using the transfer function method, and the results were compared with those obtained by Andrade et al. (2007) [15] using Placket-Burman (PB) designs. Both results were compared with the reference order, obtained by applying the 'perturb' method to the mathematical model proposed by Andrade et al. (2007) [15]. The reference values for the kinetic parameter were taken from Andrade et al. (2007) [15].

ISBN: 978-988-19251-7-6
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

The ‘perturb’ method was applied to build the profiles of the effects of the following inputs on the response: fermentation time for total substrate consumption:

1) μ_{max} = maximum specific growth rate (h^{-1})
2) X_{max} = biomass concentration when cell growth ceases (kg/ m^3)
3) P_m =product concentration when cell growth ceases (kg/ m^3)
4) Y_x = limit cellular yield (kg/kg)
5) Y_{px} = yield of product based on cell growth (kg/kg)
6) K_s = substrate saturation parameter (m^3/kg)
7) K_i = substrate inhibition coefficient (m^3/kg)
8) m_p= ethanol production associated with growth (kg/[kg.h])
9) m_x = maintenance parameter (kg/[kg.h])
10) m = parameter used to describe cellular inhibition
11) n = parameter used to describe the inhibition by product.

The changes used in the 'perturb' method consisted of variations of -20%, -10%, +10% and +20%, which were produced in the selected input variable around the reference values of the kinetic parameters, while keeping all the other inputs constant. The results are shown in Fig. 13.

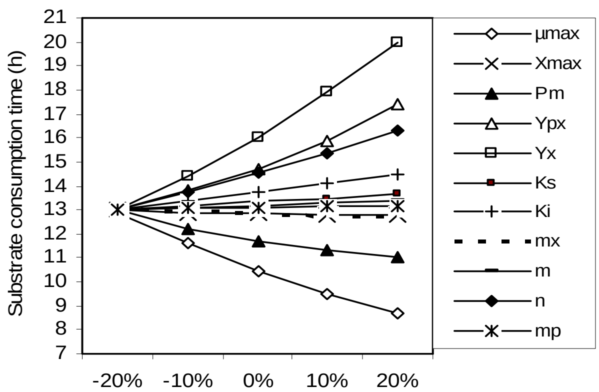


Fig. 12. Profiles of the variation of the output variable according to the changes in the input variables

Fig. 12 shows that there are six inputs with greater influence on the output: Y_x, Y_{px}, μ_{max}, *n*, P_m, and K_i.The most important input affecting the output (fermentation time for total substrate consumption) was Y_x, the limit in cell yield (kg/kg).

In order to quantitatively express the information shown in Fig. 13, the S (*index*) was defined, which measures the cumulative sum of changes in the slope value. The slope is the ratio of the output change to the difference between two consecutive scaled values of inputs, and is calculated by the following equation:

$$S(index) = \sum_{i=1}^{ni} \left| O_{i+1} - O_i \right| / \left| sI_{i+1} - sI_i \right| \tag{18}$$

Where: s*I* = scaled value of input, O = output value,
i = interval number between two consecutive scaled values of inputs.

ANN model topology

With the purpose of applying the transfer functions method, a feedforward neural network with 11:5:5:1 topology, with 11 inputs, 5 neurons in the first hidden layer, 5 neurons in the second hidden layer and 1 output, was selected to model the relationships between the above mentioned factors and the response: fermentation time for total substrate consumption.

The number of neurons in the hidden layers was determined using the cross-validation technique, in order to avoid model overfitting and to achieve good generalization from the training dataset. This technique splits the data sample into a training dataset and a validation dataset. The performance of the trained neural network was evaluated by the ability to predict the elements of the validation dataset, which was expressed in terms of the mean square error (*mse*).

$$mse = \frac{1}{N} \sum_{k=1}^{N} \left(t_k - ev_k^O \right)^2 \tag{19}$$

In (19), *k* is the number of data points in the validation dataset, which varies between 1 and *N*, t_k is the *k*th *target value,* and ev_k^O is the *k*th net estimated output value.

Comparison of the results for the importance of the effects of input variables on the outputs

The relative importance of all the input variables on a given output using the weight values for this particular ANN structure was obtained using (12).

When the transfer functions are fractional contributions expressed in terms of absolute values of weights,. (12) adopts the form of (16). In this case: *ni* = 11 (number of inputs); *n*1 = 5 (number of neurons in the first hidden layer); *n*2 = 5 (number of neurons in the second hidden layer); O = 1 (output); and *I* = 1,2,3,…,11 (output number).

The orders of importance of the six inputs with greater influence on the output, obtained using the transfer function method and the Placket-Burman design [15], are shown in Fig. 13.

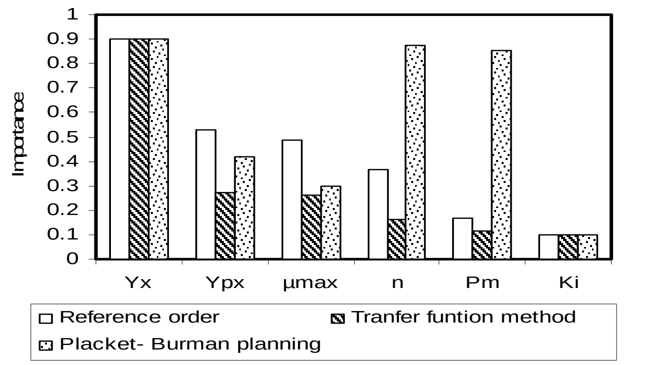


Fig. 13 Comparison of the importance of the inputs’ effects on the fermentation time for total substrate consumption, using different methods.

Fig. 13 shows that the order of importance obtained using the transfer function method with the absolute weight values was the same as that obtained by applying the perturb method to the phenomenological mathematical model of the process, but that the Plackett-Burman design methodology led to a different order of importance of the kinetic parameters for the substrate consumption time

ISBN: 978-988-19251-7-6
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

IV. Conclusion

The transfer function method provided a general framework to study the information contained in the weights of ANN based models, which was useful into estimating the importance of inputs on outputs, and facilitated the derivation of applicable equations for models based on neural networks with two or more layers of neurons, extending the possibilities of analyzing these cases with respect to the weight based methods found in the literature.

Transfer functions between adjacent neurons in ANNs could be used to relate inputs and outputs. In numerical applications, good results were obtained when using relative contributions expressed in terms of absolute values of weights as transfer functions, in order to estimate the effects of inputs on outputs.

A good neural network model of a process gives secure information about the relative importance of the input variables, which highlights the importance of the availability of good models to describe the dynamic behavior of chemical and biotechnological processes.

The results showed that the proposed method, which uses transfer functions and the rules of block diagram algebra to estimate the importance of the effects of inputs on outputs in ANN based models, can give better results than the application of Plackett Burman designs.

References

[1] A. Engelbrecht, and I. Cloete, "Feature extraction from feedforward neural networks using sensitivity analysis," In *Proceedings of the International Conference on Systems, Signals, Control, Computers*, Durban, South Africa,1998, pp. 221-225.

[2] *Z. Reitermanová*, "Feedforward Neural Networks – Architecture Optimization and Knowledge Extraction," *In* WDS'08 Proceedings of Contributed Papers, Part I: Mathematics and Computer Sciences, *Prague, Czech Republic, 2008, pp. 159–164*.

[3] Z. Huang, H. Chen, C. J. Hsu, W. H. Chen and S. Wu, "Credit rating analysis with support vector machines and neural networks: a market comparative study," *Decis Support Syst*., vol. 37, no. 4 , pp. 543–558, Sept. 2004.

[4] J. Tikka, "Simultaneous input variable and basis function selection for RBF networks," *Neurocomputing*, vol. 72 no. 10-12, pp. 2649-2658, Jun. 2009.

[5] N. Chapados, and Y. Bengio, "Input decay: Simple and effective soft variable selection," In *Proceedings of the 2001 IEEE/INNS International Joint Conference on Neural Networks*, Washington D.C., 2001, pp.1233 – 1237.

[6] T. Similä and J. Tikka, "Combined input variable selection and model complexity control for nonlinear regression," *Pattern Recognit Lett*, vol. 30, no. 3, pp. 231-236, Feb. 2009.

[7] G. D. Garson, "Interpreting neural-network connection weights," *AI Expert*, vol. 6, no. 7, pp. 47-51, Apr. 1991.

[8] P. Hajela and Z. P. Szewczyk, "Neurocomputing strategies in structural design on analyzing weights of feedforward neural networks," *Struct Multidisc Optim*., vol. 8, no. 4, pp. 236-241, Dec. 1994.

[9] Y. Dimopoulos, P. Bourret and S. Lek, "Use of some sensitivity criteria for choosing networks with good generalization ability," *Neural Process Lett*, vol. 2, no. 6, pp. 1-4. 1995.

[10] J. D. Olden, M. K. Joy and R. G. Death, "An accurate comparison of methods for quantifying variable importance in artificial neural networks using simulated data," *Ecol Modell*, vol. 178, no. 389-397, Nov. 2004.

[11] W. Luyben, *Process modeling, simulation and control for chemical engineers*. New York, NY: McGraw-Hill, 1990.

[12] D. E. Seborg, T.F. Edgar, and D.A Mellichamp, *Process Dynamics and Control*. Hoboken, New Jersey, NJ: John Wiley & Sons, Inc., 2003.

[13] C. Mei, "On Teaching the Simplification of Block Diagrams," *Int J Eng Ed*, vol. 18, no. 6, pp. 697-703, My. 2002.

[14] J. Isar, A. Lata, S. Saurabh and S.A. Rajendra, "Statistical method for enhancing the production of succinic acid from *Escherichia coli* under anaerobic conditions," *Bioresour Technol*, vol. 97, pp. 1443–1448, Sept. 2006.

[15] R. R. Andrade, "Procedure for development of a robust mathematical model for alcoholic fermentation process," M.S. thesis, State University of Campinas, Sao Paulo, Brazil, 2007.

[16] R.L. Plackett and J.P. Burman, "The design optimum multifactorial experiments," *Biometrika*, vol. 33, no.4, pp. 305-325, 1946.