2011 International Nuclear Atlantic Conference - INAC 2011
Belo Horizonte,MG, Brazil, October 24-28, 2011
Associação Brasileira de Energia Nuclear - ABEN
ISBN: 978-85-99141-04-5

ANALYSIS OF INPUT VARIABLES OF AN ARTIFICIAL NEURAL NETWORK USING BIVARIATE CORRELATION AND CANONICAL CORRELATION

Valter Magalhães Costa[1] and Iraci Martinez Pereira[1]

[1] Instituto de Pesquisas Energéticas e Nucleares (IPEN / CNEN - SP)
Av. Professor Lineu Prestes 2242
05508-000 São Paulo, SP
valter.costa@usp.br

ABSTRACT

The monitoring of variables and diagnosis of sensor fault in nuclear power plants or processes industries is very important because a previous diagnosis allows the correction of the fault and, like this, to prevent the production stopped, improving operator's security and it's not provoking economics losses. The objective of this work is to build a set, using bivariate correlation and canonical correlation, which will be the set of input variables of an artificial neural network to monitor the greater number of variables. This methodology was applied to the IEA-R1 Research Reactor at IPEN. Initially, for the input set of neural network we selected the variables: nuclear power, primary circuit flow rate, control/safety rod position and difference in pressure in the core of the reactor, because almost whole of monitoring variables have relation with the variables early described or its effect can be result of the interaction of two or more. The nuclear power is related to the increasing and decreasing of temperatures as well as the amount radiation due fission of the uranium; the rods are controls of power and influence in the amount of radiation and increasing and decreasing of temperatures; the primary circuit flow rate has the function of energy transport by removing the nucleus heat. An artificial neural network was trained and the results were satisfactory since the IEA-R1 Data Acquisition System reactor monitors 64 variables and, with a set of 9 input variables resulting from the correlation analysis, it was possible to monitor 51 variables.

1. INTRODUCTION

The growing improvement technology verified in the last decades requires of the machines and, in general, in large processes systems (chemical industries, nuclear power plants, automotive industries, etc) bigger capacity and velocity of operation increasing the production and decreasing the customary hard work. However, the complexity of large processes systems, in addition to the various components of hardware and software and ever increasing demand for safety and reliability, has led to the development of efficient system of monitoring and diagnosis because is necessary to keep the equipment in operation without no-programmed stop provoking economic losses and increasing the costs to repair some damaged equipment. [1]

The features of a nuclear reactor that belong to the design of the control system are based in mathematical analysis and are such complex, because are no-linears and multi-variables, such that is impossible to write mathematical equations to describe it whole and the simplified equations contain very hypotheses on the behavior of the reactor that represent only an approximation of their real behavior [2] [3] [4].

Any power system should be controlled and the expert systems develop an important role in the support of the operation of nuclear systems. The operator has to be skillful to identify unusual condition such as a plant accident scenario, equipment failure or an external disturbance to the system and the expert system has show the necessary information to your action because accidents have to be limited in such way that radioactive materials not escape to the environment [5] [6] [7] [8].

In situations that exist a lot of data or lot of variables to be monitor and there is no a structured knowledge about them is very helpful the use of Artificial Neural Networks (ANN) in the monitoring and control of this processes. In applications of ANN a fundamental problem is to find a suitable representation of the multivariate data because the adequate selection of the inputs variables is important to be chosen the smallest number of variables that contain the necessary information to the neural network. Normally, we turned to the knowledge of the expert to the selection of these variables and, however, it would be interesting a method of automatic selection of this variables such independent of the prior knowledge of the expert [9] [10].

One of the central problems in the field of neural networks is the model selection such that have been intensively investigated and a vast array of methods and results have been suggested to perform this task [11]. Anders and Korn (1998) examine how model selection in neural networks can be guided by statistical procedures such as hypothesis tests, information criteria and cross validation [12]. Gomm, Evans and Williams **(**1996**)** present investigations into the application of neural-network techniques to identify the nonlinear dynamics of a real, laboratory-scale, process unit, exhibiting non-linearity and typical disturbance characteristics, and subsequently, the on-line performance of a predictive controller incorporating the neural-network model [13]. Witczak, Korbicz, Mrugalski, Patton (2005) focused on the problem of designing a robust fault detection scheme with group method of data handling (GMDH) neural networks and showed how to determine the structure and parameters of the neural network as well as how to estimate the modeling uncertainty of the resulting neural model. It is shown that the algorithms proposed to tackle the above tasks make it possible to obtain a suitably accurate mathematical description of the system [14]. Gou and Fyfe (2003) review a recent neural implementation of Canonical Correlation Analysis and show, using ideas suggested by Ridge Regression, how to make the algorithm robust [15].

The aim of this work is the selection of the inputs variables of ANN. For this, will be used the Correlation Coefficient of Pearson and the Multiple Correlation Coefficient to examine the linear relations between two set of variables and for the calculation of the Multiple Correlation Coefficient we used the method of the Canonical Correlation Analysis.

With this methodologies want, in the whole of all variables monitor of a nuclear power plant, to build a set, not necessary minimum, which will be set input variables of ANN and, like way, to monitor the biggest number of variables. The variables selected will be value based in the efficiency of the monitoring of sensors of the reactor IEA-R1 by ANN and will compare the results obtained by ANN and by Multiple Correlation Coefficient.

2. LINEAR CORRELATION

2.1. Linear Regression Equation

Some kinds of relations encountered in everyday can be expressed conveniently and accurately in terms of linear equations that, in general, can be written as $Y = a + bX$, where Y is the dependent variable, X is independent variable and the constants a and b determines the position of the line on the graph. Let, now, X and Y variables with n observations each, where (X_i, Y_i) , $1 \le i \le n$ can be plotted in a scatter diagram in the graphic below.

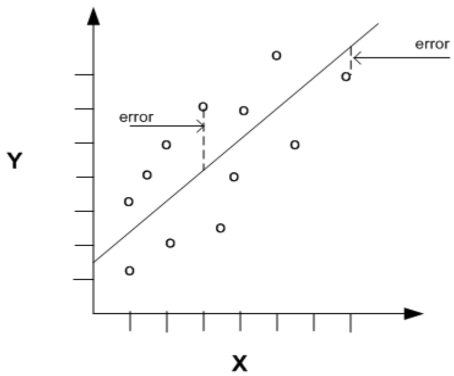

Figure 1: Dispersion graph.

The goal of linear regression is to adjust the values of slope and intercept to find the line that best predicts Y from X. More precisely, the goal of regression is to minimize the sum of the <u>squares</u> of the <u>vertical</u> distances of the points from the line, that is, for any X_i have (X_i, Y_i) e $(X_i, \hat{Y}_i)$ where Y_i (real value or observed value) and $\hat{Y}_i$ (predict value by straight line), labeling $\hat{Y} = a + bX$ of regression line, the aim is located the better line that can be adjusted to the points, that is, the parameters a e b so that the error $|Y_i - \hat{Y}_i|$ is minimized. [18]

The criterion for best fit is named "least-square" and in this fitting a straight line to a set of data by this criterion, the values of a and b in the equation are chosen so as to minimize the sum of the squared discrepancies $\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2$. Replacing $\hat{Y}_i$ for $\hat{Y}_i = a + bX_i$ and labeling de G this difference, yields:

$$G = \sum_{i=1}^{n}[Y_i - (a + bX_i)]^2 = \sum_{i=1}^{n}(Y_i - a - bX_i)^2 \qquad (1)$$

Squaring the expression in parentheses on the rights, differentiating with respect to a and b and setting each of these expressions equal to zero and rearranging, we have:

ISBN: 978-85-99141-04-5

$$b = \frac{n\sum_{i=1}^{n} X_i Y_i - \sum_{i=1}^{n} X_i \sum_{i=1}^{n} Y_i}{n\sum_{i=1}^{n} X_i^2 - (\sum_{i=1}^{n} X_i)^2} = \frac{\sum_{i=1}^{n} X_i Y_i - \frac{\sum_{i=1}^{n} X_i \sum_{i=1}^{n} Y_i}{n}}{\sum_{i=1}^{n} X_i^2 - \frac{(\sum_{i=1}^{n} X_2)^2}{n}} \quad \text{and} \quad a = \bar{Y} + b\,\bar{X}$$

To simplify notation, we define:

$$S_{xx} = \sum (X - \bar{X})^2 = \frac{n\sum X^2 - (\sum X)^2}{n} = \sum X^2 - \frac{(\sum X)^2}{n} \qquad S_{yy} = \sum (Y_i - \bar{Y})^2 = \frac{n\sum Y^2 - (\sum Y)^2}{n} = \sum Y^2 - \frac{(\sum Y)^2}{n}$$

$$S_{xy} = \sum (X_i - \bar{X})(Y_i - \bar{Y}) = \frac{n\sum XY - \sum X \sum Y}{n} = \sum XY - \frac{\sum X \sum Y}{n}$$

And using this notation, we have:

$$b = \frac{S_{xy}}{S_{xx}} \qquad (2)$$

2.2. Correlation Coefficient of Pearson

The correlation coefficient measures the strength of relationship between X and Y and it is related to linear regression as follow:

Considering the identity $Y_i - \bar{Y} = (Y_i - \hat{Y}_i) + (\hat{Y}_i - \bar{Y})$ and squaring both sides of the equation we have:

$$(Y_i - \bar{Y})^2 = (Y_i - \hat{Y}_i)^2 + (\hat{Y}_i - \bar{Y})^2 + 2(Y_i - \hat{Y}_i)(\hat{Y}_i - \bar{Y})$$

Adding the last term over the n sample observations and replacing $\hat{Y}_i$ by $\bar{Y}_i + b(X_i - \bar{X})$, we have:

$$\sum (Y_i - \bar{Y})^2 = \sum (Y_i - \hat{Y}_i)^2 + \sum (\hat{Y}_i - \bar{Y})^2 \quad (*)$$

Each of the terms in this equation must be greater than or equal to zero, therefore, obviously:

$$\sum (Y_i - \bar{Y})^2 \geq \sum (Y_i - \hat{Y}_i)^2 \quad \text{e} \quad \sum (Y_i - \bar{Y})^2 \geq \sum (\hat{Y}_i - \bar{Y})^2$$

Now, dividing both side of (*) by $\sum (Y_i - \bar{Y})^2$, we obtain:

$$r^2 = 1 - \frac{\sum (Y_i - \hat{Y}_i)^2}{\sum (Y_i - \bar{Y})^2} = \frac{\sum (\hat{Y}_i - \bar{Y})^2}{\sum (Y_i - \bar{Y})^2} \Rightarrow r = \pm\sqrt{1 - \frac{\sum (Y_i - \hat{Y}_i)^2}{\sum (Y_i - \bar{Y})^2}} = \pm\sqrt{\frac{\sum (\hat{Y}_i - \bar{Y})^2}{\sum (Y_i - \bar{Y})^2}} \qquad (3)$$

Considering the left side of the equality sign (right side of the implication sign) in the equation (by [16] [17]):

If $\sum (Y_i - \hat{Y}_i)^2 = 0$ (if there is no error of estimation), than the left side of the equation would be exactly 1 and, in this case, would be a perfect linear relationship between X and Y. On the other hand, if there were no linear relationship between X and Y than we have $\hat{Y} = \bar{Y}$ and the left side of the equality sign would equal zero.

Considering now the member of the right, if there were no linear relationship between X and Y and Y=$\bar{Y}$, than the numerator of this term would be $\sum (Y_i - \bar{Y})^2$ and the expression on the right would exactly 1, indicating a perfect linear relationship between X and Y. Each side of the expression defines the square of Pearson correlation coefficient, r. Thus, the equations of r and r^2 have the properties:

1. r is zero if there is no linear relationship between X and Y. In this case $\sum (Y_i - \hat{Y}_i)^2$ has its maximum value, $\sum (Y_i - \bar{Y})^2$ and $\sum (\hat{Y}_i - \bar{Y})^2$ has minimum value zero.
2. The possible values of r range from -1 a +1 where r = 1 or r = -1 indicate perfect linear relationship between X e Y. that is, $\sum (Y_i - \hat{Y}_i)^2 = 0$ and $\sum (Y_i - \bar{Y})^2 = \sum (\hat{Y}_i - \bar{Y})^2$. Thus, $r^2 = 1$ and than, r = 1 or r = -1.

2.3. Alternative Formula to the Correlation Coefficient

As $\hat{Y}_i = a + bX_i$, we can write $\sum (Y_i - \hat{Y}_i)^2 = \sum (Y_i - a - bX_i)^2$ and replacing a for $\bar{Y} - b\bar{X}$ yields:

$$\sum (Y_i - \hat{Y}_i)^2 = \sum (Y_i - \bar{Y} + b\bar{X} - bX_i)^2 = \sum_{i=1}^{n} [(Y_i - \bar{Y}) - b(X_i - \bar{X})]^2$$

Squaring the expression of the last equality:

$$\sum_{i=1}^{n} [(Y_i - \bar{Y})^2 + b^2 \sum (X_i - \bar{X})^2 - 2b \sum (X_i - \bar{X})(Y_i - \bar{Y})$$

We have now:

$$\sum (Y_i - \hat{Y}_i)^2 = S_{yy} + b^2 S_{xx} - 2bS_{xy}$$

Like $b = \dfrac{S_{xy}}{S_{xx}}$, yields:

$$\sum (Y_i - \hat{Y}_i)^2 = S_{yy} + \frac{S_{xy}{}^2 S_{xx}}{S_{xx}^2} - 2\frac{S_{xy}^2}{S_{xx}} = S_{yy} - \frac{S_{xy}^2}{S_{xx}}$$

Using now this last equality and the fact that $S_{yy} = \sum (Y_i - \bar{Y})^2$ we may substitute for $\sum (Y_i - \hat{Y}_i)^2$ e S_{yy} to obtain:

$$r = \pm\sqrt{1 - \frac{S_{yy} - \frac{S_{xy}^2}{S_{xx}}}{S_{yy}}} = \pm\sqrt{\frac{S_{xy}^2}{S_{xx}S_{yy}}} = \frac{S_{xy}}{\sqrt{S_{xx}S_{yy}}} \Rightarrow r = \sum_{i=1}^{n}\left(\frac{x_i - \bar{x}}{\sqrt{S_{xx}}}\right)\left(\frac{y_i - \bar{y}}{\sqrt{S_{yy}}}\right) \quad (4)$$

3. MULTIPLE CORRELATION ANALYSIS

3.1. Canonical Correlation Analysis (CCA)

Let us consider that X has p variables and Y has q variables in your sets, the CCA looks for linear combinations of the p and q variables as follow:

$$Z_x = u_1 X_1 + \ldots + u_p X_p \qquad\qquad Z_y = v_1 Y_1 + \ldots + v_q Y_q$$

where the canonical variates Z_x and Z_y have maximum correlation, i.e. the weight vectors u and v are chosen such that corr(Z_x, Z_y), the Pearson correlation coefficient between Z_x and Z_y is maximized. [16] [17] [18]
It can be showed that corr(Z_x, Z_y) is:

$$r_{z_x z_y} = \frac{\sum z_x z_y}{\sqrt{\sum z_x^2 \sum z_y^2}} = \frac{u' S_{xy} v}{\sqrt{(u' S_{xx} u)(v' S_{yy} v)}} \quad (5)$$

where

$$u = \begin{pmatrix} u_1 \\ . \\ . \\ . \\ u_p \end{pmatrix} \quad v = \begin{pmatrix} v_1 \\ . \\ . \\ . \\ v_p \end{pmatrix} \quad S_{xx} = \begin{pmatrix} S_{11} & S_{12} & . & . & . & S_{1p} \\ . & . & & & & . \\ . & . & & & & . \\ . & . & & & & . \\ S_{p1} & S_{p2} & & & & S_{pp} \end{pmatrix}$$

$$S_{yy} = \begin{pmatrix} S_{11} & S_{12} & . & . & . & S_{1p} \\ . & . & & & & . \\ . & . & & & & . \\ . & . & & & & . \\ S_{q1} & S_{q2} & . & . & . & S_{pp} \end{pmatrix} \quad S_{xy} = \begin{pmatrix} S_{x_1y_1} & S_{x_1y_2} & . & . & . & S_{x_1y_q} \\ . & . & & & & . \\ . & . & & & & . \\ . & . & & & & . \\ S_{x_py_1} & S_{x_py_2} & . & . & . & S_{x_py_q} \end{pmatrix}$$

In matrices Sxx, S_{yy}, S_{xy}, are computation, respectively, by:

$$S_{jk} = \sum_{i=1}^{n}(X_{ij} - \bar{X}_j)(X_{ik} - \bar{X}_k);\ j,k = 1,\ldots,p \qquad S_{jk} = \sum_{i=1}^{n}(Y_{ij} - \bar{Y}_j)(Y_{ik} - \bar{Y}_k);\ j,k = 1,\ldots q.$$

$$S_{x_j y_k} = \sum_{i=1}^{n}(X_{ij} - \bar{X}_j)(Y_{ik} - \bar{Y}_k) \quad j = 1,\ldots p \ \text{ e } \ k = 1,\ldots,q.$$

It can be showed which the square of the correlation coefficient between Z_x and Z_y is given by:

$$r^2 = R_{xx}^{-1}R_{xy}R_{yy}^{-1}R_{yx} \Rightarrow r = \sqrt{R_{xx}^{-1}R_{xy}R_{yy}^{-1}R_{yx}} \qquad (6)$$

Where,

$$R_{xx} = \begin{pmatrix} 1 & r_{x_1x_2} & r_{x_1x_3} & \cdots & r_{x_1x_p} \\ r_{x_2x_1} & 1 & & & r_{x_2x_p} \\ & & . & & \\ & & & . & \\ r_{x_px_1} & r_{x_px_2} & & & 1 \end{pmatrix} \quad R_{yy} = \begin{pmatrix} 1 & r_{y_1y_2} & r_{y_1y_3} & \cdots & r_{y_1yq} \\ r_{y_2y_1} & 1 & & & r_{y_2y_p} \\ & & . & & \\ & & & . & \\ r_{y_py_1} & r_{y_py_2} & & & 1 \end{pmatrix}$$

$$R_{xy} = \begin{pmatrix} r_{X_1Y_1} & r_{X_1Y_2} & r_{X_1Y_3} & \cdots & r_{X_1Y_q} \\ r_{X_2Y_1} & r_{X_1Y_2} & & & r_{X_2Y_q} \\ & & . & & \\ & & & . & \\ r_{X_pY_1} & r_{X_pY_2} & & & r_{X_pY_q} \end{pmatrix}$$

such that the elements of the matrix represent the correlation between subscripted two variables. [16] [19] [20]
In particular, if

$$Y = (Y_1) \quad \text{and} \quad X = \begin{pmatrix} X_1 \\ X_2 \\ . \\ . \\ . \\ X_P \end{pmatrix}$$

them, $Z_y = v_1Y_1$ (with v_1=1) and $Z_x = u_1X_1 + ... + u_pX_p$

Thus, we have: $R_{xy} = \begin{pmatrix} r_{X_1Y_1} \\ r_{X_2Y_1} \\ . \\ . \\ . \\ r_{X_pY_1} \end{pmatrix}$ $\quad R_{yx} = \begin{pmatrix} r_{Y_1X_1} & r_{Y_1X_1} & \cdots & r_{Y_1X_p} \end{pmatrix}$

And so,

$$(1)\begin{pmatrix} r_{Y_1X_1} & r_{Y_1X_1} & \cdots & r_{Y_1X_p} \end{pmatrix} \begin{pmatrix} 1 & r_{x_1x_2} & r_{x_1x_3} & \cdots & r_{x_1x_p} \\ r_{x_2x_1} & 1 & & & r_{x_2x_p} \\ & & . & & \\ & & & . & \\ r_{x_px_1} & r_{x_px_2} & & & 1 \end{pmatrix}^{-1} \begin{pmatrix} r_{X_1Y_1} \\ r_{X_2Y_1} \\ . \\ . \\ . \\ r_{X_pY_1} \end{pmatrix} = r^2 I = r^2$$

4. ARTIFICIAL NEURAL NETWORKS

A neuron is an information-processing unit that is fundamental to the operation of a neural network and is identified by three basic elements:

1. A set of synapses or synaptic weight.
2. An adder for summing the input signals, weighted by respective synapses of the neuron.
3. An activation function for limiting the amplitude of the output of a neuron.

The neuronal model includes the bias, denoted by b_k. The bias has the effect of increasing or lowering the net input of the activation function.

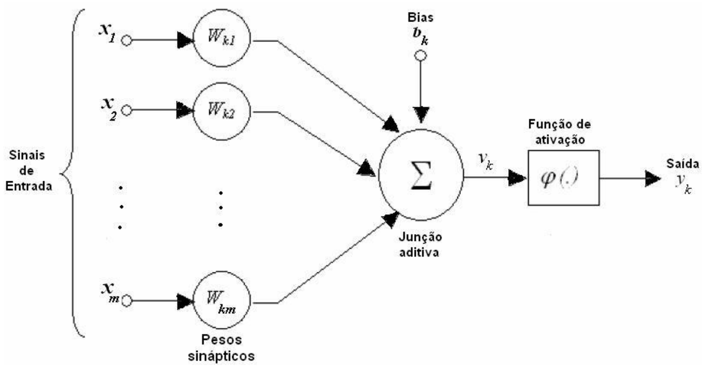

Figura 2. Modelo de neurônio base para projetos de ANN
Fonte: Haykin (1999)

In mathematical terms, a neuron may be described by following pair of equations:

$$u_k = \sum_{j=1}^{m} w_{kj} x_j \qquad y_k = \varphi(u_k + b_k) \qquad (7)$$

where $x_1, x_2, ..., x_m$ are the input signals; $w_{k1}, w_{k2,}..., w_{km}$ are the synaptic weights; u_k is the linear combiner output due to the input signals; b_k is the bias; φ is the activation function; y_k is the output signal of the neuron.

The choice of the network architecture is important aspect because the manner in which the neurons of a neural network are structured is intimately linked with the learning algorithm used to train the network. Like way, in this paper we consider only classes of one-hidden layer feedforward networks with activation function sigmoid and the learning algorithm is the Backpropagation. The Backpropagation algorithm is a learning algorithm supervisioned that, briefly, is: The inputs are multiplied by the adaptive weights; the output is obtained by passing the sum of the weighted inputs through an activation function (sigmoid function). Initially set to small random values, the weights are adjusted after each presentation of a new input pattern. The adaptation rule is given by: $\Delta w_{ji}(n) = -\eta \frac{\partial E(n)}{\partial w_{ji}(n)}$, where η is the learning rate, $E(n) = \frac{1}{2}\sum_{j \in C} e_j^2(n)$, e is the error vector, that is, the difference between the actual and desired outputs. The learning task of interest is to use the algorithm to approximation of functions [10] [21] [22] [23] [24].

4.1. Neural Network Validation

To evaluate the approximation or performance of the neural network we used the value of energy error $E_{média} = \frac{1}{N}\sum_{i=1}^{N} E(n)$ and to estimate the modeling quality we used the correlation test $r^2 = 1 - \frac{\sum (y_i - \hat{y}_i)^2}{\sum (y_i - \bar{y})^2}$, where y and $\hat{y}$ are the target and network output, respectively, $\bar{y}$ is the average of the target y over the test set. [25]

The value r is a fraction between 0.0 and 1.0, and has no units. The value 0.0 means that knowing X does not help you predict Y or there is no linear relationship between X and Y, and the best-fit line is a horizontal line going through the mean of all Y values. When r equals 1.0, all points lie exactly on a straight line with no scatter (to see figure below).

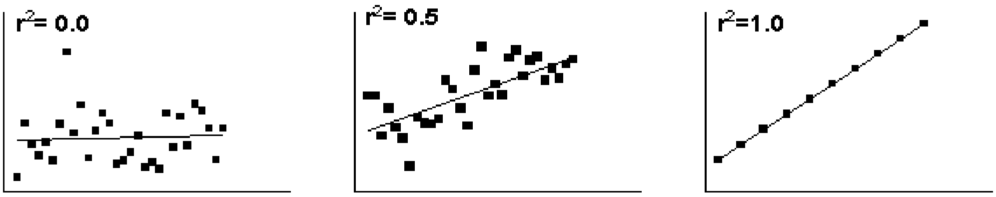

Figure 3. Scatter graphs

5. PHYSICS CONSIDERATIONS OF REACTORS

The nuclear reactors possess the mechanism to control the nuclear fission and must be understood as source of heat that replaces the combustion in the gas plant or steam plant. The liberation of energy by nuclear fission coming from kinetic energy of the fission fragments, neutrons and alpha and beta particles derived from radioactive decay such that the heat generator by three sources correspond about 90%, or more, of the energy generated and the 10% correspond to the gamma rays produced in the fission [6] [7] [19].

A nuclear power reactor is a thermal machine and is thermodynamic coupled to the power system. It's service and economic value depend strongly on thermal performance of the fuel and coolant in the core and, besides, is the rate as that the heat is removed that limit the level power. In nuclear plants the heath flow from fuel element to the water coolant and should be removed continuously [7] [26].

The power is determined by rate of fission and the quantity of neutrons is related to production of heat and, in consequence of the law of Energy Conservation, the heat produced in the elements fuel induce the increase of the temperature of the water coolant. Apart from, the source principals of the radiation are resulted of the nuclear fission. The control of reactor by rod of control/safety with high cross-sections and has the function of to regulate the power [27].

6. WORK DEVELOPMENT

6.1. Data Base and Neural Network Preparation

The operation data of the IEA-R1 Research Reactor at IPEN used refer at the measurement of IEA-R1 Data Acquisition System (SAD) at 2009 between 19 and 22 in January. A concise list of the variables follows below [28] [29]:

Table 1: Variables Monitored

1	T1	**17**	T20-T7	**33**	PPtcB	**49**	Z4
2	T2	**18**	T22	**34**	PStcA	**50**	R1M3
3	T3	**19**	F1M3	**35**	PStcB	**51**	R2M3
4	T4	**20**	F23	**36**	NBPtcA	**52**	R3M3
5	T4-T3	**21**	F2M3	**37**	NBPtcB	**53**	R4M3
6	T6	**22**	F3M3	**38**	N1	**54**	R5M3
7	T7	**23**	C1	**39**	N2	**55**	R6M3
8	T8	**24**	C2	**40**	N3	**56**	R7M3
9	T9	**25**	L1	**41**	N4	**57**	R8M3
10	T9-T8	**26**	F1	**42**	N5	**58**	R9M3
11	T13	**27**	F2	**43**	N6	**59**	R10M3
12	T15	**28**	F3	**44**	N7	**60**	R11M3
13	T16	**29**	F4	**45**	N8	**61**	R12M3
14	T16-T15	**30**	delta P	**46**	Z1	**62**	R13M3
15	T19	**31**	Preator	**47**	Z2	**63**	R14M3
16	T20	**32**	PPtcA	**48**	Z3	**64**	NOBRES

The variables of 1 to 18 refer to Temperatures, 19 to 22 refer to flow of Primary Circuit, 23 and 24 refer to Conductivity, 25 refer to poll Level, 26 to 29 refer to flow of Secondary Circuit, 30 refer to Difference in pressure in the core of the reactor, 31 to 37 refer to Thermal Balance, 38 to 45 refer to Power, 46 to 49 refer to Rod of Control, 50 to 58 refer to Monitor Area (Radiation) and 59 to 64 refer to Ducts Monitor (Radiation).

In order, the neural networks development in this work belong to the class of multilayer perceptron with one hidden layer, the nodes of the hidden layer use the activation function sigmoid and the number of nodes in the output layer is one because, in all cases, we have only one predict variable and the number of nodes in the input layer is equal to the number of variables of input of the neural network.

The neural networks trained with the Backpropagation algorithm was implemented using only the data sets that were divided into a training set (70% of the data set) and a test set (30% of the data set) and to determine the end of the training procedure we used maximum number of training epochs (Iterations) that, in this case, is equal to 1000 epochs. All calculations were performed with Matlab 7.4 using a NN Toolbox.

The variables Power, Rate of flow of primary circuit, Rod of control/security and Difference in pressure in the core of the reactor (ΔP), because, for hypothesis, almost whole of monitoring variables have relation with the variables early described or it's effect can be result of the interaction of two or more. The Power is related to the increasing and decreasing of temperatures as well as the amount radiation due fission of the uranium; the Rods are controls of power and, for consequence, influence in the amount of radiation and increasing and decreasing of temperatures; the Rate of flow of primary circuit has function of the transport of energy by removing of heat of the nucleus and the pool is responsible as many for moderating as armour-plate.

Like this, labeling B= {Power, Rate of flow of Primary Circuit, Rod of Control/Security and ΔP} we computed the correlation between B and all another variables monitoring by SAD (coefficient of multiple correlation) and we tested the efficiency in approximation by ANN. The results obtained to the variables of B where to each variable monitored we used the others variables of set as input of ANN:

Table 2. Approximation by ANN using B

Variable	Error Energy	Goodness of adjust
19	1.297.635	0.0479
30	5,75E-04	0.6529
39	0.0608	0.9965
46	880.056	0.9755
47	0.1351	0.9999
48	0.0598	1.0000
49	0.2819	0.9999

By Canonical Correlation Analysis theory we can add variables at any set, whose correlation with the monitor variable is more than any bivariate correlation between the monitor variable and those belong to set, to improve the correlation coefficient and, like way, to try improve the quality of linear prediction. In this case, as to 19 as to 46 the results was unsatisfactory. However, addition to B the variable 26, where correlation with 19 is 0.99 (corr (26,19) = 0.9920), we obtain 0.0018 to error energy and 1 to goodness of adjust. Analyzing 46, of the point view of the correlation, their bigger correlations are with 47, 48 49 (to see table below)

Table 3. Bivariate correlation between entre "19 and B" and "46 and B"

19	**30**	**39**	**46**	**47**	**48**	**49**
Correlation	-0.9495	0.1011	0.1514	0.1463	0.1457	0.1456

46	**19**	**30**	**39**	**47**	**48**	**49**
Correlation	0.1514	-0.2382	0.8143	0.8937	0.8947	0.8946

As the aim is find a set of variable which make approximation, by ANN, of the greatest number of variables we have two alternatives: 1) To add at B the variable 26; and 2) To add at B the variable 26 and exclude the variable 46.

Anyway, we excluded the variable 49 of both situations 1) and 2) due, firstly, to high correlation among 47, 48 e 49 and second because are variables that have the same function (table 4).

Table 4. Cross-correlation between 47, 48 e 49

Variable	47	48	49
47	1.0000	0.9999	0.9999
48	0.9999	1.0000	1.0000
49	0.9999	1.0000	1.0000

Labeling B' the set obtained in 1) and B'' the set obtained in 2) we computed the correlation coefficient and we tested the efficiency in approximation by ANN of variables of SAD and as B' as B'' was unsatisfactory the approximation by neural network of, 5, 20, 21, 22, 23, 27, 28, 29 31, 35, 38, 54, 55, 56, 59, 60, 61, 62 e 63.

Observing the table 5 integrated to the correlation map that provide the level of linear association between the variables unsatisfactory approximated by grade colors where the yellow means high correlation and red means low correlation we have that 20, 21 e 22 have high correlation among themselves. Of the some way, 28 and 35 have high correlation with 20 and the pair (21, 22) has high correlation with 28 and 35.

By adding 20, 21 and 22 to B'', the approximation by ANN of each variable has been satisfactory using the others as input (table 6). Analyzing this new set (C) using it as input variables we have an increasing in the quantity of variables satisfactory approximated by ANN and by linear fit.

Table 5. Cross-correlation between variables and Correlation Map

variables

Correlation	20	21	22	28	35
20	1.00	0.44	0.90	0.79	-0.92
21	0.44	1.00	0.02	0.40	-0.07
22	0.90	0.02	1.00	0.69	-0.99
28	0.79	0.40	0.69	1.00	-0.73
35	-0.92	-0.07	-0.99	-0.73	1.00

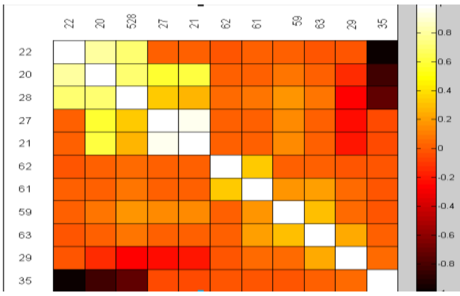

Table 6: Approximation by ANN using C

Variable	Error energy	Goodness of adjust
19	0.0053	1
20	0.3910	0.9999
21	0.0664	1
22	0.0019	1
26	2,79E-04	1
30	5,58E-04	0.7242
39	0.1697	0.9900
47	319.347	0.9908

7. RESULTS

The study of the physic behavior of reactors that allowed the selection of a base (label B) and have been monitored 39 variables using B as input set. Likewise, have been monitored all variables which belong to B, except the variables 19 and 46 and, to add the variable 26, that have high correlation with 19 have been satisfactory of the point view of monitoring by ANN.

To use, only, the measure of the goodness of adjust to assess the performance of the ANN it is not reasonable, because we can have a high goodness of adjust without low error energy. For example, the variable 60 has linear correlation with B' equal to 0.9671, good goodness of adjust, however high error energy. The situations observed in the implementation of the method have been:

- Lower error energy and lower Goodness of the adjust;
- Lower error energy and high Goodness of the adjust;
- High error energy and high Goodness of the adjust.

The set C (19, 20, 21, 22, 26, 30, 39, 47 e 48) allowed, considering the measure of performance of ANN, the approximation by ANN of 51 variables. The variables unsatisfactory approximated were: 5, 23, 29, 38, 46, 54, 55, 56, 59, 60, 61, 62 and 63.

8. CONCLUSIONS

A analysis of the variables monitored of the reactor IEA-R1 do IPEN both from a phenomelogy study and Canonical Correlation Analysis theory point view showed to be a efficient method because has allowed to relate variables that were previously unconnected and, in this way, it become possible to obtain a input set of a ANN such as we monitored 51 variables.

To add variables to the initial set to increase the degree of linear correlation between the predicted variable and the input set showed to be a interesting tool to select the input variables of the neural network once that, in all variables monitored, the high correlation is related to the performance of the ANN because the high correlations presented have led to reduction of the error energy.

The measures "error energy" and "goodness of adjust" have to be used simultaneously and the ideal condition is "Lower error energy and high Goodness of the adjust".

REFERENCES

[1] Bueno, E.I., "Utilização de Redes Neurais Artificiais na Monitoração e Detecção de Falhas em sensores do Reator IEA-R1", Dissertação apresentada ao IPEN como parte dos requisitos para obtenção do Grau de Mestre em Ciências na Área de Tecnologica Nuclear-Reatores, 2006.

[2] Patton, R. J.; Chen, J.; Nielsen, S. B., **"**Model-Based Methods for Fault Diagnosis: Some Guidelines**",** *IEEE Transactions Of The Institute Of Measurement And Control*, **Vol. 17**, No. 2, 1995.

[3] Ogata, K. *Modern Control Engineering* 4th ed., Prentice Hall, N.J., USA, 2001.

[4] Y, T., Su-xia; H., Chong, L.; Fu-yu, Z. "An improved implicit multiple model predictive control used for movable nuclear power plant"**,** *Nuclear Engineering and Design*. **v 240** ,p 3582–3585, 2010.

[5] Glasstone, S.; Sesonske, A. *Nuclear reactor engineering* 4th ed. v II, Chapman and Hall, New York, NY , 1994.

[6] Murray, R. L. *Nuclear Energy* 4th ed., Pergamon Press, Elmsford, NY, 1993.

[7] Glasstone, S. "Principles of Nuclear Reactor Engineering", Princeton, N.J.:Van Nostrand . 1956.

[8] Lugger, G. F. *Artificial Intelligence: structure and strategies for complex problem solution* 4th ed. Addison-Wesley, Harlow, Angleterre; Toronto, 2002.

[9] Braga, A. P.; Carvalho, A. P. L. F.; Ludemir, T. B., *Fundamentos De Redes Neurais Artificiais*, Rio De Janeiro: 11º Escola De Computação/Universidade Federal do Rio de Janeiro, 1998.

[10] Haykin, S. *Neural Networks: A Comprehensive Foundation*, 2th ed. Prentice Hall, New Jersey, USA, 1999.

[11] Finnof, W.; Hergert, R.; Zimmermann, G. “Improving Model Selection by Nonconvergent Methods**”** *Neural Networks*, **v 6**, pp 771-783, 1933.
[12] Anders, U.; Korn, O, “Model selection in neural networks**”** *Neural Network*, **v 12**, pp 309-323, 1999.

[13] Gomm, J.B.; Evans, J.T. ; Williams, D, “Development and performance of a neural network predictive controller”, *Control Eng. Practice,* **v 5**, No. 1, p. 49-59, 1997.
[14] Witczak, M., Korbicz, J. Z., Mrugalski, M.., Patton, R.J., “A GMDH neural network-based approach to robust fault diagnosis: Application to the DAMADICS benchmark problem”, *Control Engineering Practice,* **v. 14**, pp 671–683, 2006.
[15] Gou, Z.; Fyfe, C. “A canonical correlation neural network for multicollinearity and functional data” *Neural Network*, **v 17**, pp 285-293, 2004.
[16] “Lindeman, Ricchard’’, Introduction to Bivariate and Multivariate Analysis**,** Scott, Foresman and Company, Glenview, Ill, 1985
[17] Retherford, R. D., Choe, M, K., *Statistical Models for Causal Analysis*, Wiley-Interscience, Honolulu, Hawaii, 1993.
[18] Hsieh, W. W., “Nonlinear canonical correlation analysis by neural networks”, *Neural Network***, v 13**, pp 1095-1105, 2000.
[19] Hotelling, H., “Analysis of a complex of statistical variables into principal components” *Journal of Educational Psychology*, **v 24,** pp 417-441, 1933.
[20] Hotelling, H., “The most predictable criterion”, *Journal of Educational Psychology*, **v 26**, pp 139-142, 1935
[21] Hornik, K., “Multilayer Feedforward Networks are Universal Approximators”, *Neural Networks,* **Vol. 2**, pp. 359-366, 1986
[22] Leshno, M.;Ya, V. I.; Pinkus A.;Schocken, S., “Multilayer Feedforward Networks With a Nonpolynomial Activation Function Can Approximate Any Function*”, Neural Networks*, **Vol. 6**, pp. 861-867, 1993
[23] Stinchcombe, M.B., “Neural network approximation of continuous functionals and continuous functions on compactifications”, *Neural Networks***, v. 12** , pp 467–477, 1999.
[24] Funahashi, K. I. “On the Approximate Realization of Continuous Mappings by Neural Networks”, *Neural Networks*, **v. 2**, pp. 183-192, 1989
[25] Minin, A.; Velikova, M.; Lang, B.; Daniels, H., “Comparison of universal approximators incorporating partial monotonicity by structure” *Neural Networks*, **v 23**, pp 471-475, 2010.
[26] Tong, L. S., “Heat transfer in water-cooled nuclear reactors”, *Nuclear Engineering and Design*, **v 6**, pp 301-324, 1967
[27] Szymendera, L., “A Study on the temperature effect coupling in reactor radiation attenuation analysis”, *Nuclear Engineering and Design*, **v 7,** pp 201-222, 1968.
[28] Tanomaru, N.; Hiromoto, M. Y. C., *Manual De Instalação e Operação do SAD IEA-R1*, 1998.
[29] Monteiro, P. R. B.; Gonçalves, I. M. P. M., *Sistema De Aquisição De Dados Do Reator IEA-R**1***, INAC 2002 – *International Nuclear Atlantic Conference Proceedings (Cdroom)*, Outubro 2002.