Proceedings of the 10th WSEAS International Conference on SYSTEMS, Vouliagmeni, Athens, Greece, July 10-12, 2006 (pp449-454)

Water cut virtual sensor design using a Neo-fuzzy neuron and Statistical techniques in oil production

RICHARD GARCÍA-MÉNDEZ*, EDGAR CAMARGO-CONTRERAS^, FRANCKLIN RIVAS-ECHEVERRÍA*, ELIEZER COLINA-MORLES*, MARIANELLA LUZARDO-BRICEÑO**
Universidad de Los Andes, Mérida, Venezuela 5101
*Laboratorio de Sistemas Inteligentes, Facultad de Ingeniería
**Escuela de Estadística, Facultad de Ciencias Económicas y Sociales
^PDVSA-AIT. Gestión Tecnológica. Edo. Zulia, Venezuela

Abstract: - In this work a water cut virtual sensor for an oil extraction process test separator is presented. This virtual sensor is constructed using the pressure, temperature, flow and water cut variables measurement, taken using physical sensors. A previous data processing using statistical techniques is applied for determining the relation between the variables and for detecting possible atypical data. Some wells clusters are obtained from wells with similar behaviors in the considered variables and a model for each group or individual well is obtained, using a Neo-Fuzzy neuron.

Key-Words: - Virtual Sensors, Neo-fuzzy Neurons, Statistical Analysis, Oil production.

1 Introduction

The operation principle of physical variables sensors has experienced great changes throughout history. Due to the increasing use of the digital computer in supervision and industrial process control, it have arise digital computer-based sensors called "virtual sensors", which are used when is difficult to measure important variables or when the respective sensor is very expensive, too slow or inexact; in these situations the virtual sensors become necessary and nowadays they are more and more popular. Virtual sensors are programs that infer the value of a variable from the existing information about other variables. The programs can consist of a mathematical model, heuristic models or intelligent model [1].

The intelligent models that have been widely used as virtual sensors are the artificial neural networks and the fuzzy models. In this application the intelligent models are used to make complex associations, nonlinear and highly multidimensional between the existing data on other variables and the variable to be estimated [1].

Artificial neural networks consist of a system that tries to emulate the biological neural networks behavior concerning the learning and the generalization capability. With the purpose of taking advantage of artificial neural networks and the capacity of handling vague information provided by the fuzzy logic models, a new structure called Neo-fuzzy neuron has been proposed [4, 9, 12], which has demonstrated to give good results in behavior representation of complex systems.

The data concerning process variables that will be used to find relations must be representative of the process, which is verified by means of the application of certain statistical techniques and insured by means of controlled experiments design.

In the crude handling processes at the oil industry, a key variable exists that is the water cut level, which is defined as the amount of water present in the crude and is measured mainly at flow stations, where the dehydration chemical treatment for the crude begins [2].

Usually, water cut level has been measured using very expensive devices for the industry; so, at the moment the disposition of these sensors in all the flow stations is very difficult.

This work is based on the necessity to determine the water cut level without incurring in elevated costs associated to the acquisition and maintenance of a physical sensor, but making use of the existing information on other variables of the process, provided by other sensors.

2 Virtual Sensors Application and Design

Virtual Sensors implantation can be done by means of dedicated software or through specific computational tools for the execution of neural networks structures algorithms, fuzzy logic or from mathematical models estimation that approximate the processes dynamics [3].

A virtual sensor can be interpreted as a map that transforms input signals from a process into a considered variable, where the transformation is learned by the virtual sensor structure, according to the interrelations between the input variables and the variable to be estimated. Figure 1 depicts the illustration of a virtual sensor [3].

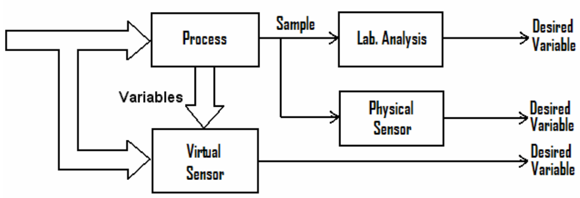

Figure 1. Virtual sensor illustration

3 Neo Fuzzy Neuron

Neo-fuzzy neuron constitutes a tool that offers great advantages for modeling complex systems by the simplicity of its structure, consisting of a single neuron, which the difference enormously of the artificial neuronal networks, where several neurons are included and can be numerous when the system to be modeled is very complex. Whereas in artificial neural networks it is necessary to change the number of layers, the number of neurons in each layer and the activation function to find the structure that obtains a good adjustment, in the neo-fuzzy neuron it is only necessary to change the number of fuzzy partitions in the input variables, allowing this way to find the most suitable structure with greater facility.

The structure of the neo-fuzzy neuron is shown in figure 2, where the synaptic weights are not constant but nonlinear functions of the inputs, represented by fuzzy logic models based on a collection of “If – Then” rules, that use an approximated reasoning in the inference process. This structure does not have an activation function, but it posses a summing point that generates the output when adding the fuzzy logic model outputs for each input [4, 9].

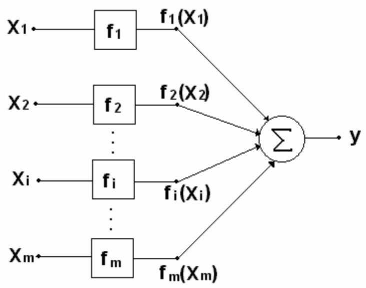

Figure 2. Neo-fuzzy Neuron

The input variables spaces are divided in several segments that will constitute the fuzzy subgroups of each variable. Each of these segments, as it is shown in figure 3, is characterized by a triangular complementary membership function.

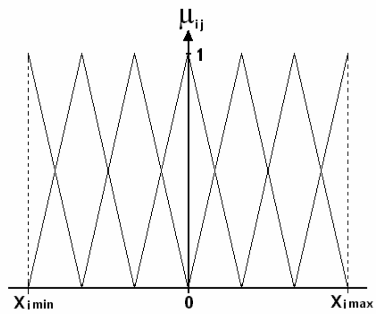

Figure 3. Membership Functions

The output synapses of each fuzzy logic model is obtained by means of an inference mechanism using fuzzyfication and defuzzyfication processes as is shown in figure 4.

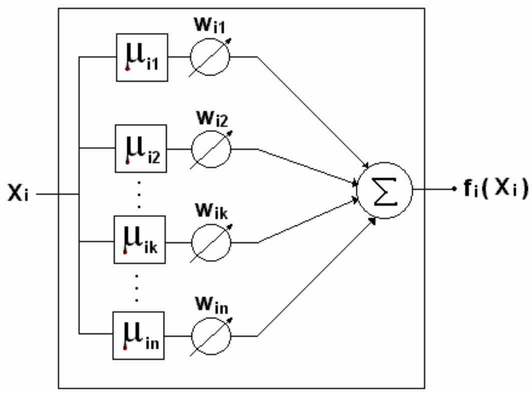

Figure 4. Neo fuzzy Neuron Synapse

The fuzzy neuron output "y" is given by the following equation:

$$y = f1(x1) + f2(x2) + ... + fm(xm) = \sum_{i=1}^{m} fi(xi)$$

3.1. Neo-fuzzy Neuron Learning

As in conventional artificial neural networks, learning in a Neo-fuzzy neuron consists about synapse modification in such a way that the errors between the desired outputs of the neuron outputs are minimized. Considering that in Neo Fuzzy neurons the synapse is represented by a fuzzy logic model with a set of "If – Then" rules, whose consequent are constant weights, and for a single signal input two rules are always activated, the constant weights of each synapse that influence the output are two and these are due to modify to obtain the desired output. This way, the learning for a Neo fuzzy neuron consists of modifying two constant weight of each synapse, corresponding to the activated rules related to a specific input, until obtaining the desired output. This is made by means of gradient descendent algorithm with the objective of minimizing an error functional.

4 Statistical Techniques

For virtual sensors construction, there is a group of statistical techniques important to take into account:

4.1. Correlation Coefficients

The correlation coefficient between two variables provides a measurement of the linear association among them. The correlation coefficient takes values in the interval [- 1.1] and is 1 or - 1 when the linear dependency between a pair of variables is exact.

4.2 Outliers Detection

Atypical data, also known as outliers, are defined as those data that have a unique characteristics combination that clearly difference them of others. Atypical data cannot be characterized as beneficial or problematic, but they must be analyzed in order to determine the type of information is providing. The problem of these data is that could not be representative of the population and could distort the inferred behavior using a model. Also it can happen that, although they are different from most of the data, belong to a valid segment of the population and, therefore, this indicates the representativeness lack of the sampled data [5].

Although there are diverse techniques for identifying outliers, the decision to exclude or not an atypical data depend on the investigator opinion, and when it is classified as atypical it must be made a study to determine the cause of this fact because this cause can be errors of registry data [6].

4.3. Cluster Analysis

A widely used technique for making observations groups with similar behaviors is the cluster analysis. The objective of the cluster analysis is grouping the observations in such a form that the observations from a same group are very similar respect to the considered variables and that observations included in different groups have a different behavior with respect to these variables [7].

4.4. Random sampling

Sampling is a statistic tool that has as objective to determine a part from a universe that must be studied in order to make inferences concerning the whole universe. Between the advantages of the sampling tool are that they allow to obtain greater exactitude and greater focus in the obtained results [8].

5 Water Cut Virtual Sensor Design

5.1 Environment conditions.

Wells extracted flow, contain crude, gas, and also presents water concentration and other impurities, which generates a multiphasic state. In such sense, a crude - gas separation process is made to the flow collected at stations, and then it is measured the water cut. This allows concluding, depending on its value, if the extraction method needs to be changed.

The selected device used for making the water cut virtual sensor design is a test separator, also known as measurement separator. The water cut measurements are made by means of installed physical equipment located at the test separator output. At the separator output there are available also pressure, temperature and flow measurements provided by another installed sensor.

The selected test separator receives crude out coming from 29 different wells, identified by numbers from 1 to 29.

5.2 Statistical analysis

Calculation of Correlations:

Table 1. Correlation Coefficients

	Water Cut	Pressure	Temperature	Flow
Water Cut	1			
Pressure	0,10982291	1		
Temperature	-0,03909244	0,19409418	1	
Flor	-0,05730942	-0,44046204	-0,06584726	1

It can be appreciated in table 1, that correlations between the variables are very low. When it was made training with diverse conventional neural structures, different parameters and configurations, was not obtained appropriate generalization results with low estimation errors.

Wells clustering:

When detecting very low correlations between water cut and pressure, temperature and flow variables, and not finding a neural model that can be adjusted to the experimental data, it was important to make wells groups with similar behaviors concerning the four considered variables. The behavior difference between wells was verified when finding very similar values of pressure, temperature and flow associated to very different water cut values, corresponding to different wells.

The wells groups with similar behaviors concerning water cut, pressure, temperature and flow variables, was obtained using the French software Spad V4.5 [16, 17]. The identified groups are the following ones:

Group I: Wells 15, 16, 20 y 24.

Group II: Wells 3, 19, 21 y 25.

Group III: Wells 1, 5, 8, 11, 14, 22, 23 y 27.

Group IV: Wells 2, 6, 10, 12, 28 y 29.

Group V: Wells 4, 7, 9, 18, 13 y 26.

Well 17 was not considered for having very few data of the considered variables.

After calculating the correlations matrices for each wells group, it were appraised that the correlation between each consider independent variable (pressure, temperature and flow) and the dependent variable (water cut) is very low; nevertheless, it was tried to find neural models that can be used for making the water cut estimation from the other variables information.

Atypical Data Detection:

With the purpose of detecting atypical values, applying the central limit theorem, it was assumed that each variable has approximately a normal distribution and it was given the standardization to each one of them. Values located outside the interval [- 3, 3], that can be considered as atypical data, were not eliminated initially for making the training of the neural structures because of the ignorance of their origin. They were considered as atypical values, and eliminated of the training files, only when those values outside the interval [- 3.3] widely affect finding some neural structure fitness for the test and training patterns; as it will be indicated more ahead, this did not happen in many cases.

5.3 Neo-Fuzzy Neuron Training

The scheme of the water cut virtual sensor using a Neo-fuzzy Neuron is presented in figure 5.

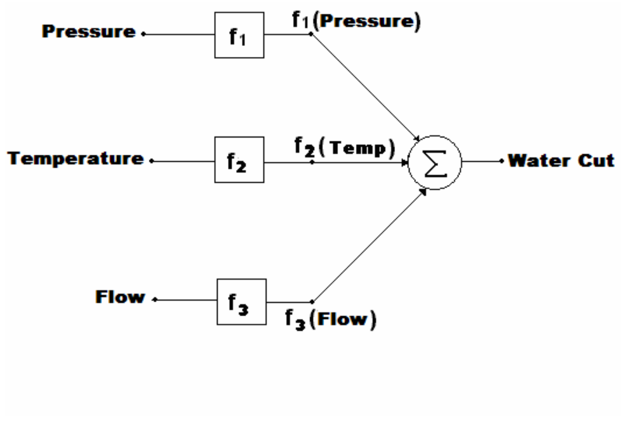

Figure 5. Water cut virtual sensor using a Neo-fuzzy Neuron scheme

Because it is not possible to obtain a single neural structure that models the complete system with all wells that are connected to the test separator, a Neo-fuzzy Neuron for each well group with similar behaviors can be trained and, according to the well that is connected to the test separator, a Neo-fuzzy neuron is activated for estimating the water cut value.

Because the Neo-fuzzy Neuron structure learning depends on the data quality used in the training stage; before separating the patterns in the test and training files, a random sampling of the patterns was made.

5.4 Obtained results

After making several experiments with Neo-fuzzy Neurons, it was found some results as the shown next.

Well Groups I:

For the wells group I, conformed by wells 15, 16, 20 y 24 it was obtained the following results:
Neo-fuzzy Neuron Structure:

Inputs Number: 3 inputs (pressure, temperature and flow)
Output Number: 1 output (water cut level)
Fuzzy sets Number:
15 sets for pressure variable
33 sets for temperature variable
31 sets for flow variable
Training Pattern Number: 126 (70 %)
Testing Pattern Number: 54 (30 %)
Training cycles Number: 250 cycles

It can be seen in Figure 6 the obtained results.

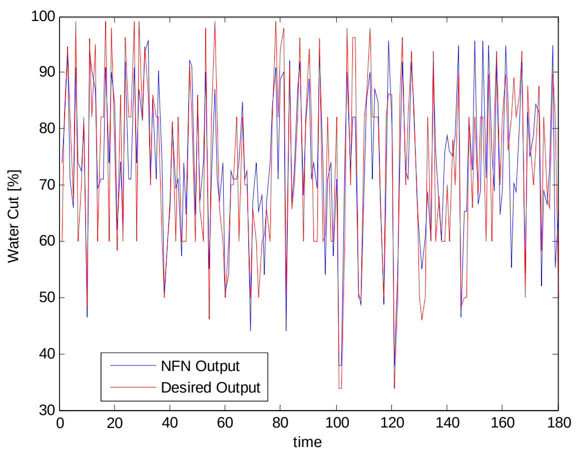

Figure 6. Neo-fuzzy Neuron Test for wells group I

The obtained Neo-fuzzy Neuron for wells group I does a good water cut estimation, having an average error for the training patterns of 9.5% and 12.5% for the testing patterns.

Wells Group IV:

For the wells group IV, conformed by wells 2, 6, 10, 12, 28 and 29 it was obtained the following results:

Neo-fuzzy Neuron Structure:

Inputs Number: 3 inputs (pressure, temperature and flow)
Output Number: 1 output (water cut level)
Fuzzy sets Number:
25 sets for pressure variable
21 sets for temperature variable
23 sets for flow variable
Training Pattern Number: 39 (67%)
Testing Pattern Number: 19 (33%)
Training cycles Number: 350 cycles

It can be seen in Figure 7 the obtained results.

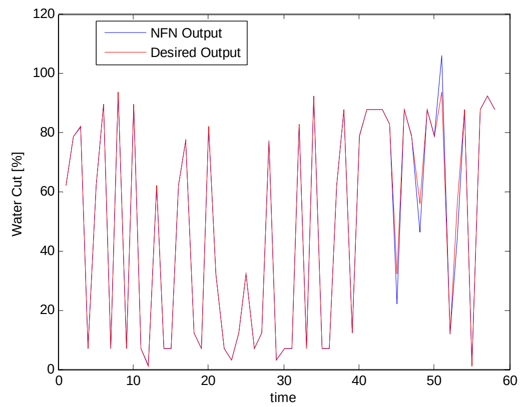

Figure 7. Neo-fuzzy Neuron Test for wells group IV

The obtained Neo-fuzzy Neuron for wells group IV gives very good water cut estimation, having an average error for the training patterns of 0.4% and 4.9% for the testing patterns.

6 Conclusion

The Neo-fuzzy neuron resulting from the concepts fusion of artificial neuronal networks and fuzzy logic turns out to be very useful for imitation complex systems, obtaining a good performance as far as the learning time required and the generalization capacity.

A necessary condition for obtaining an accurate virtual sensor using intelligent techniques is to have representative process data, with relevant information concerning the behavior, as much of the variable to be estimated and the other variables that will be used.

The statistical techniques are very useful for the selection of the data to be used for training and testing the neural structures. By means of a random data sampling, representative data used in the training patterns can be included. The standardization technique makes important contributions in the identification of atypical observations; observations with standardized values outside the interval [-3, 3] can be included in the model once it is verified that the neural structure obtains a good learning; if the structure does not obtain a good learning, these observations must be eliminated. Cluster analysis offers great advantages in the determination of systems with similar characteristics, allowing, after appreciating different observations, making groups of those similar observations, and then obtaining a model of each group separately.

The relation between the pressure, temperature, flow and water cut variables in a test separator is determined by the characteristics of the well being tested. The percentage estimation of present water in the crude, from other measurements made in the separator, requires of wells groups with similar behaviors with respect to these variables and the training of a Neo-fuzzy Neuron structure for each group of this form.

Please, follow our instructions faithfully, otherwise you have to resubmit your full paper. This will enable us to maintain uniformity in the conference proceedings as well as in the post-conference luxurious books by WSES Press. Thank you for your cooperation and contribution. We are looking forward to seeing you at the Conference.

References:

[1] Raúl R. Leal Ascencio y Enrique Herrera. *Sensores virtuales mediante redes neuronales artificiales dos estudios de caso en biotecnología*. http://iteso.mx/~rleal/archivos/conielec.pdf, 2005.

[2] Marianilca Olivares. *Diseño de sensores virtuales para estimar el nivel de corte de agua en la unidad de explotación bachaquero lago de pdvsa occidente*. Tesis de grado, Universidad de los Andes, mayo 2000.

[3] Addison Ríos; Francklin Rivas; Edgar Chacón; Jonós Montilva; Eliezer Colina. *Aspectos metodológicos para la implantación de sensores virtuales*. Notas de clase, Universidad de los Andes, Noviembre 1997.

[4] Takeshi Yamakawa. *A neo Fuzzy Neuron and Its Applications to System Identification and Prediction of Chaotic Behavior*. IEEE, 1994.

[5] M. Salvador Figueras y P. Gargallo. *Análisis exploratorio de datos*.

http://www.5campus.com/leccion/aed, 2003.

[6] Vic Barnett y Toby Lewis. *Outliers in Statistical Data*. John Wiley and Sons, 1978.

[7] Manuel Salvador Figueras. *Análisis de conglomerados o cluster*.

http://www.ciberconta.unizar.es/LECCION/cluster/INICIO.HTML, 2001.

[8] William Cochran. *Técnicas de Muestreo*. Compañía Editorial Continental, México, 1985.

[9] T. Yamakawa et al. "*A Neo fuzzy neuron and it applications to system Identification and prediction of the system behavior*". Proceedings of 2nd International Conference on Fuzzy Logic & Neural Networks". Japan. 1992.

[10] J. Aguilar J. and F. Rivas. *Introducción a las técnicas de Computación Inteligente*. Meritec. Mérida. Venezuela. 2001.

[11] Darcy Novoa, Anna Pérez, Francklin Rivas. *Fault Detection scheme using Neo-fuzzy Neurons*. IASTED International Conference on Intelligent Systems and Control Agosto 2000.Honolulu, Hawaii. USA

[12] Pérez Anna, Torres Elizabeth, Maldonado, Ronald, Rivas Francklin. "*Pisum Sativum Classification using discriminate analysis and Neural Networks*". 6th WSEAS International Conference on Neural Networks. Lisbon, Portugal 2005.

[13] Anderson T. W. *An Introduction to Multivariate Statistical Analysis*. Wiley Series in Probability and Statistics. Tercera Edición. 2003

[14] Jonson Richard and Wichern Dean. *Applied Statistical Analysis*. 4th Edition. Prentice Hall 1998.

[15] Morrison Donald F. *Multivariate Statistical Methods*. 3rd Edition. Mc Graw Hill Publishing Company 1990.

[16] SPAD. N., *Système portable pour l'Analyse des Données*. CISIA Ed., 25

Ave de l'Europe 92310 Sèvres. 1987

[17]www.bondy.ird.fr/pleins_textes/ pleins_textes_6/colloques2/36006.pdf