Int. J. of Computers, Communications & Control, ISSN 1841-9836, E-ISSN 1841-9844
Vol. V (2010), No. 2, pp. 205-215

On the Use of the FuzzyARTMAP Neural Network for Pattern Recognition in Statistical Process Control using a Factorial Design

J. A. Vazquez-Lopez, I. Lopez-Juarez, M. Peña-Cabrera

J. A. Vazquez-Lopez
Instituto Tecnologico de Celaya
CIATEC, A.C. Posgrado Interinstitucional en Ciencia y Tecnologia (PICyT)
Industrial Engineering Department
Celaya, Gto. MEXICO
E-mail: antoniov@itc.mx

I. Lopez-Juarez
Centro de Investigacion y de Estudios Avanzados del IPN - Unidad Saltillo
Robotics and Advanced Manufacturing Department
Carretera Saltillo-Monterrey Km 13. Ramos Arizpe, Coahuila. MEXICO
E-mail: ismael.lopez@cinvestav.edu.mx

M. Peña-Cabrera
National Autonomous University of Mexico
Engineering Computing Science and Automation Department
Institute of Research on Applied Mathematics and Systems
Mexico, City. MEXICO
E-mail: mario@leibniz.iimas.unam.mx

Abstract: Time-series statistical pattern recognition is of prime importance in statistics, especially in quality control techniques for manufacturing processes. A frequent problem in this application is the complexity when trying to determine the behaviour (pattern) from sample data. There have been identified standard patterns which are commonly present when using the X chart; its detection depends on human judgement supported by norms and graphical criteria. In the last few years, it has been demonstrated that Artificial Neural Networks (ANN's) are useful to predict the type of time-series pattern instead of the use of rules. However, the ANN control parameters have to be fixed to values that maximize its performance. This research proposes an experimental design methodology to determine the most appropriate values for the control parameters of the FuzzyARTMAP ANN such as: learning rate (β) and network vigilance (ρ_a, ρ_b, ρ_{ab}) in order to increment the neural network efficiency during unnatural pattern recognition.
Keywords: Statistical Process Control, Control Charts, Artificial Neural Network (ANN), FuzzyARTMAP, and Factorial Design.

1 Introduction

To preserve product quality an accurate knowledge of the production process is necessary. This requires the automation of quality control systems and the use of control charts as introduced by Dr. Walter A Shewart to observe the behaviour of the manufacturing process.

Control charting is the key point in Statistical Process Control (SPC) implementation. The correct application of these Control Charts requires satisfying statistical assumptions such as the independence of the random variable and symmetry in its probability distribution [1]. If these assumptions are met then the use of Control Charts is correctly applied since the Upper and Lower limits are established as $\pm 3\sigma$ from the global mean of the X random variable. In figure 1, the probability distribution of X is shown under both circumstances with symmetry and without symmetry.

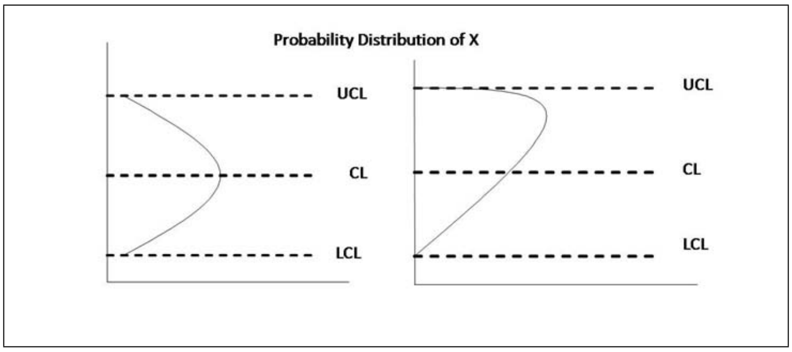

Figure 1: X Symmetry importance to establish the control limits

The power of an X chart is based on its capacity to differentiate special and natural causes of variation; however, important disadvantages on the use of this form of quality control exist because traditional control chart with control limit can only indicate when to seek a disturbance and not where and what to look for, generating then hurried and sometimes mistaken diagnostics [2].

Trying to know where and what has happened in the manufacture process may be possible by using pre-established rules in combination with human judgment. These rules are commonly referred to as: points outside the control limits, run of consecutive points, non-random patterns and points near the control limits [1]. The efficiency of the use of these rules has been investigated and it has been found that is not enough to recognise the type of statistical pattern ([3], [4], [5], [6], [7], [8], [9], [10], [11], [12]), which would give the correct answer to the questions of where and what to look for. This is why researchers suggest the use of the neural networks as an alternative approach to identify variation data in statistical patterns [4].

A neural network is a soft computing system [25], which consists of a number of elements (nodes) strongly interconnected that have the ability to process information as a result of a process of dynamic work of those nodes and connections to external points to the network [13]. Neural networks are efficient in recognizing data variation [2], [5], [14], especially in asymmetric probability distributions [2]. Among the existing neural networks, the Fuzzy ARTMAP Network is widely recognized due to its on-line and fast learning capability for pattern recognition tasks [15], [16].

For many processes of manufacture, the parameters can be obtained with neural network, with exception in many complex biotechnological processes. In these cases, [23] proposed the use of grey-box models which combine a priori knowledge expressed in terms of a white-box model, with a black-box model such as neural network, and [24] developed a Matlab®Toolbox for the construction of grey-box neural network models.

This paper is organised as follows: In the next section an introduction to the ART theory is given first followed by the standardisation and coding algorithm for the input data in the neural network using the Monte Carlo method. Section 3 describes the statistical pattern generation. Training and testing results are provided in Section 4. Section 5 describes the factorial design for the selection of the network parameters as well as the experimental results. An analysis for the pattern variation findings is given in Section 6 while Conclusions and Future work are given in Section 7.

2 Adaptive resonance theory

The Adaptive Resonance Theory (ART) [17] was developed by Stephen Grossberg and Gail Carpenter at Boston University to solve the called stability-plasticity dilemma. That is, the system is sensitive to novelty capable of distinguishing between familiar and unfamiliar events (plastic) and still remains stable. Different model variations have been developed to date based on the original ART-1 algorithm for binary input patterns [18], ART 2-A for analogue and binary input patterns [19], and ART 3 based on chemical transmitters. Supervised learning is possible through ARTMAP [20] that uses two ART modules and its variants, Fuzzy ARTMAP [16], Gaussian ARTMAP [21] and ART-EMAP even though there are many other variants adapted for specific applications [22]. In the next section a brief explanation of the mechanics of ART-1 and Fuzzy ARTMAP is given.

2.1 ART-1

The ART-1 architecture consists of two parts: attentional subsystem and orienting subsystem as illustrated in figure 2. The attentional subsystem is made up of two layers of nodes F_1 and F_2. In an ART network, information in the form of processing-element output reverberates back and forth between layers. If a stable resonance takes place learning or adaptation can occur. On the other hand, the orienting subsystem is in charge of resetting the attentional subsystem when an unfamiliar event occurs.

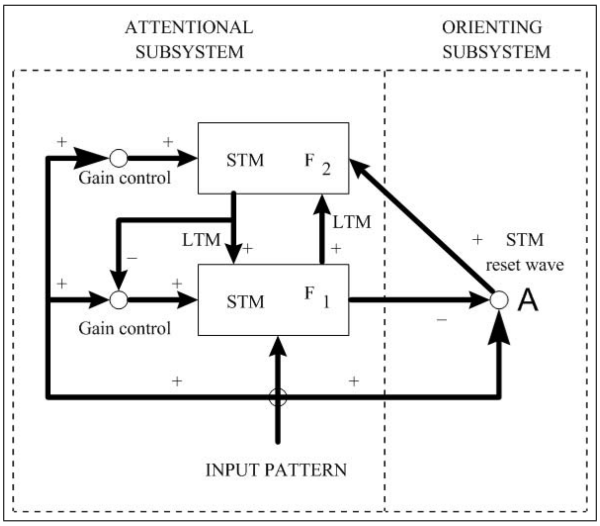

Figure 2: Basic ART Architecture

A resonant state can be attained in one of two ways. If the network has learned previously to recognise an input vector, then a resonant state will be achieved quickly when that input vector is presented. During resonance, the adaptation process will reinforce the memory of the stored pattern. If the input vector is not immediately recognised, the network will rapidly search through its stored patterns looking for a match. If no match is found, the network will enter a resonant state whereupon the new pattern will be stored for the first time. Thus, the network responds quickly to previously learned data, yet remains able to learn when novel data is presented, hence solving the so-called stability-plasticity dilemma. The activity of a node in the F_1 or F_2 layer is called short-term memory (STM) whereas the adaptive weights are called long-term memory (LTM). Gain controls handle the discrete presentation of the input signals. A vigilance parameter measures how much mismatch is tolerated between the input data and the stored patterns, which can be used to control the category coarseness control of the classifier.

2.2 Fuzzy ARTMAP

In the Fuzzy ARTMAP (FAM) network there are two modules ART_a and ART_b and an inter-ART module $Map-field$ that controls the learning of an associative map from ART_a recognition categories to ART_b categories. This is illustrated in figure 3.

The Map field module also controls the match tracking of ART_a vigilance parameter. A mismatch between Map field and ART_a category activated by input I_a and ART_b category activated by input I_b

increases ART_a vigilance by the minimum amount needed for the system to search for, and if necessary, learn a new ART_a category whose prediction matches the ART_b category. The search initiated by the inter-ART reset can shift attention to a novel cluster of features that can be incorporated through learning into a new ART_a recognition category, which can then be linked to a new ART prediction via associative learning at the $Map-field$.

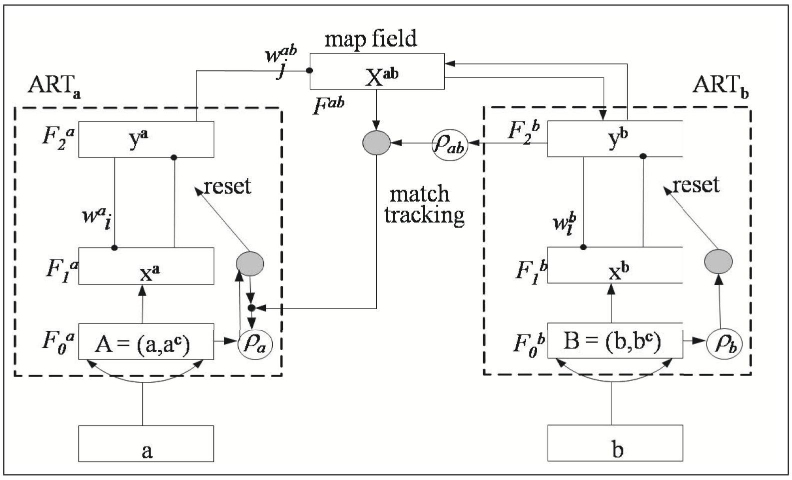

Figure 3: FuzzyARTMAP Architecture

A vigilance parameter measures the difference allowed between the input data and the stored pattern. Therefore this parameter is determinant to affect the selectivity or granularity of the network prediction. For learning, the FuzzyARTMAP has 4 important factors: Vigilance in the input module (ρ_a), vigilance in the output module (ρ_b), vigilance in the Map field (ρ_{ab}) and learning rate (β). These were the considered factors in this research.

2.3 Standardization and codification

The use of the neural network requires two important mathematical considerations, which are the standardization and the codification of the input data [2]. The training and testing data needs to be pre-processed in these two stages. The standardization means that the data have to be linearly transformed from data with mean (μ) and standard deviation (σ) into data with μ = 0 and σ = 1 using equation 1, as a result, the sample data is within the interval (-3.9, +3.9).

$$Y_t = \left(\frac{x_t - \mu}{\sigma}\right) \tag{1}$$

where:
x_t = sample value at sampling time t.
Y_t = standardized value from x_t.
μ = process mean.
σ = process standard deviation.

The x_t data are generated by a process simulator of Monte Carlo, according to equation 2

$$x_t = \mu + n_t + d_t \tag{2}$$

where:
μ = process mean.
n_t = common cause variation at sampling time t.

d_t = special disturbance at time t ($d_t = 0$ when the pattern is natural).

Shift.

$$d_t = ud \tag{3}$$

where:
u = parameter to determine the position of shifting (0: before shifting; 1: after shifting).
d = displacement of mean in terms of σ.

Trend slope.

$$d_t = st \tag{4}$$

where:
s = trend slope in terms of σ.
t = sampling time.

On the other hand, with the codification of Y_t the variation interval of [0,1] is obtained, which is a requirement for the neural network operation that reduces the effects of common causes of variation (noise) [2]. The codification of x_t considered the interval [-7.625, 7.625], whose range is greater to the expected Y_t range.

3 Pattern data generation

A specific value x_t of sample data is obtained from the sum of three mathematical considerations:

- Global and historical effect (μ).
- Natural variation effect (n_t).
- Disturbance variation effect (d_t).

Mathematically, equation 2 expresses this situation. In terms of industrial quality, these effects can be thought of as the global and historical mean obtained from experience (i), thought of as data variation which is unavoidable and it is always present (ii); and finally, the data variation due to disturbances which is associated to special causes that may cause the process to be out of statistical control (iii).

When a sample data has only influence on natural causes of variation, then $n_t > 0$ and $d_t = 0$, and the pattern data will be natural. On the other hand, if $d_t > 0$, then the pattern data will be unnatural, and it means that a cause of special variation has occurred in time t. It must be noticed that $0 < n_t < d_t$ for any type of special pattern data. If the d_t value is very similar to n_t then neural network output can be misleading between a special pattern and a natural one.

3.1 Natural Pattern

The data used for this pattern were generated using the Monte Carlo simulator using equation 2 with μ = 0 and $\sigma = 1$. An example of this type of pattern is shown in figure 4. The graph data comes from a time-series of X that did not consider any trend or shift in the global mean and with data distribution randomly assigned.

3.2 Shift Pattern

Data used for either downward shift or upward shift shows two data set separated by an abrupt change as shown in figure 5. This occurs because the reference mean also changes. This can be positive or negative and its magnitude depends on the special variation cause in the manufacturing process.

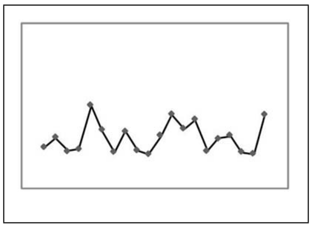

Figure 4: Natural Pattern

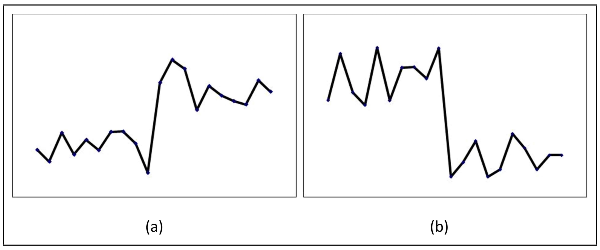

Figure 5: Shift Patterns (a)Upwards (b)Downwards

3.3 Trend Pattern

The type of pattern can be distinguished at a glance due to its upward or downward trend. In terms of mathematically what happens is that for any X_t in the data series (where t is not the last data), there will be points in time $t+1, t+2, t+3, \ldots, t+n$ of higher magnitude (upward trend) or lower magnitude (downward trend). Through linear regression is always possible to find out the slope magnitude which is also the magnitude of the effect that caused the special variation. This type of pattern can be observed in figure 6.

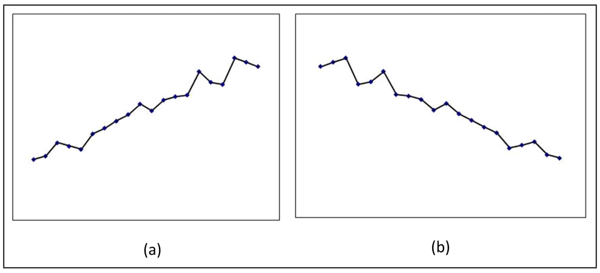

Figure 6: Trend Patterns (a)Upwards (b)Downwards

4 Training and testing

Five pattern types with diverse effects of special variation as mentioned above were studied. Table 1 shows the corresponding information indicating the used values for d_t and the output binary code used during the training and testing phase. The number of patterns for the input vector was 51 for training and 1,350 for testing.

5 Factorial design

The experimentation required 120 tests based on 8 runs and 3 replicates. The information for the experimental design is shown in table 2. An analysis of variance from multiple experiments revealed the

Table 1: d_t values considered for each pattern type and output vector code

Pattern Type	d_t	**Code**
Natural	0.0	1 0 0 0 0
Upward shift	$+0.5, +1.5, +2.5, +3.5$	0 1 0 0 0
Downward shift	$-0.5, -1.5, -2.5, -3.5$	0 0 1 0 0
Upward trend	$+0.1, +0.2, +0.3, +0.4$	0 0 0 1 0
Downward trend	$-0.1, -0.2, -0.3, -0.4$	0 0 0 0 1

significant factors of operation for the neuronal network. The results are given in section 5.1 and in all cases, there were no violations to the normality and independence of the residuals $e_{i,j}$.

Table 2: Experimental Design

CONCEPT		**DESCRIPTION**	
Type of experiment		**2^k, fraction 1/16**	
Number of factor		**7**	
Resolution		**III**	
Number of replicates		**3**	
Generating process of Alias		**D = AB, E = AC, F = BC and G = ABC**	
Independent variable (response)		**Efficiency of the neural network (η) = correct predictions/total vectors.**	
Significance level		**0.1**	
FACTORS		**LEVELS**	
DESCRIPTION		**LOW**	**HIGH**
A	ρ_{a1}	**0.2**	**0.8**
B	ρ_{ab1}	**0.2**	**0.8**
C	β_1	**0.2**	**1.0**
D	$\rho_{b1,2}$	**0.2**	**0.8**
E	ρ_{ab2}	**0.2**	**0.8**
F	β_2	**0.2**	**1.0**
G	ρ_{a2}	**0.2**	**0.8**

5.1 Experimental results

The results lead to the identification of the factors that most influence the prediction efficiency of the neuronal network (characterization). Results also showed the determination of the best combination for the factor levels (relative optimization) that generated higher efficiencies. It was observed that at the level of significance of the experimental test, all the factors influence in the efficiency of the neural network. The relative optimal levels are shown in table 3.

The efficiency of the neural network (η) considering the factors and levels indicated in table 3 are given in table 4 and figure 7. The experimental test validation was carried out with 17,000 data samples created by simulation.

Table 3: ANN Factors and optimal levels

Factors		Optimal Levels
A	(ρ_{a1})	0.8
B	(ρ_{ab1})	0.2
C	(β_1)	1.0
D	($\rho_{b1,2}$)	0.8
E	(ρ_{ab2})	0.2
F	(β_2)	0.2
G	(ρ_{a2})	0.6

Table 4: Neural network efficiency (η)

Pattern Data	η
Natural	30 %
Upward shift	93.5 %
Downward shift	96.4 %
Upward trend	99.8 %
Downward trend	96.4 %

6 Analysis of pattern variation

There is a direct relationship between the value d_t assumed by each special pattern data (table 1), the value of the standard deviation (σ) of the sample data and the efficiency of the neuronal network (η). table 5, shows this situation. In all cases, while the absolute value of d_t increases the corresponding σ and η also increases. A polynomial regression analysis was carried out (σ and d_t are the independent variables and η the dependent variable). The coefficient of correlation was high and typically above 89% (see table 6). This fact indicates that as d_t approaches n_t and σ is low then the neural network may predict a wrong pattern. On the opposite, with higher values of d_t and σ the neural network prediction efficiency increases.

Table 5: Relationship between d_t, standard deviation (σ) and the neural network efficiency (η).

Pattern Type	**Parameters**	**Values**			
Upward shift	η	76.0%	86.8%	92.7%	93.5%
Downward shift	η	42.4%	77.6%	88.2%	96.4%
	d_t	±0.5	±1.5	±2.5	±3.5
	σ	1.009	1.237	1.620	2.089
Upward trend	η	22.0%	25.0%	73.8%	99.8%
Downward trend	η	79.1%	82.8%	95.8%	96.4%
	d_t	±0.1	±0.2	±0.3	±0.4
	σ	0.539	0.670	0.862	1.080

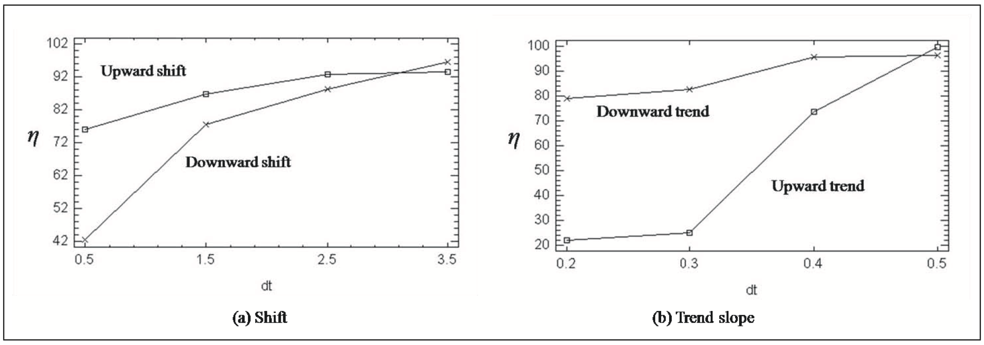

Figure 7: Neural network efficiency (η) vs d_t

Table 6: Equations for the multiple linear regressions

Pattern Type	Multiple Regression	R^2
Upward shift	Efficiency = $107.7 + 21.4(d_t) - 42.5(\sigma)$	99.8
Downward shift	Efficiency = $132.2 - 61.0(d_t) - 119.7(\sigma)$	99.3
Upward trend	Efficiency = $-121.5 - 294.1(d_t) + 317.5(\sigma)$	96.0
Downward trend	Efficiency = $77.3 - 92.3(d_t) - 15.0(\sigma)$	89.21

7 Conclusions

This investigation confirms the obtained results from previous studies with respect to the efficiency of the neural network in the recognition of statistical pattern data. It is also demonstrated, that as the effect of the special cause approaches close to zero, the efficiency decreases because the standard deviation of these data is smaller or equal to 1, i.e., the standard deviation of a data set from a natural pattern. Another result from this investigation is the definition of the most appropriate parameter values for the FuzzyARTMAP that facilitated the use of this neuronal network. It is important to mention that in this application the efficiency of the network for the control chart pattern recognition depends greatly on the d_t and σ values.

Future work has been envisaged to look at the dt and σ values in the confusion zone and their relationship with the sampling window size in order to analyse the neural network behaviour in this zone.

Acknowledgements

The authors wish to thank CIATEC, A.C., and the Posgrado Interinstitucional en Ciencia y Tecnologia (PICyT) for the facilities provided to J. A. Vazquez-Lopez during his doctoral studies.

Bibliography

[1] D.C. Montgomery. *Introduction to Statistical Quality Control*. Third edition. John Wiley & Sons, New York, 1991.

[2] R.S. Guh. Real-time pattern recognition in statistical process control: a hybrid neural network/decision tree-based approach. *IMechE, Part B: J. Engineering Manufacture*. Vol. 219, No. 3. pp. 283-298, 2005.

[3] T. L. Lucy-Bouler. Application to forecasting of neural network recognition of shifts and trends in quality control data. *Proceedings of the World Congress on Neural Networks (WCNN'93)*. Vol. 1, pp. 631-633, UK, 1993.

[4] R.S. Guh. Robustness of the neural network based control chart pattern recognition system to non-normality. *Int. J. Qual. Reliability Mgmt*. Vol. 19(1), pp. 97-112, 2002.

[5] C.W. Zobel, D.F. Cook, Q.J. Nottingham. An augmented neural network classification approach to detecting mean shifts in correlated manufacturing process parameters. *Int. Journal of Production Research*. Vol. 42 (4), pp. 741-758, 2004.

[6] F. Zorriassantine, J.D.T. Tannock. A review of neural networks for statistical process control. *Journal of Intelligent Manufacturing*. Vol. 9, pp. 209-224, 1998.

[7] C.S. Cheng. A multi-layer neural network model for detecting changes in the process mean. *Computers Ind. Engng*. Vol. 28(1), pp. 51-61, 1995.

[8] H.B. Hwarng, and C. W. Chong. Detecting process nonrandomness through a fast and cumulative learning ART-based pattern recognizer. *Int. J. Prod. Res.*, Vol. 33(7), pp. 1817-1833, 1995.

[9] C.S. Cheng. A neural network approach for the analysis of control chart patterns. *Int. J. Prod. Res.*, Vol. 35(3), pp. 667-697, 1997.

[10] E.S. Ho and S. I. Chang. An integrated neural network approach for simultaneous monitoring of process mean and variance shifts - a comparative study. *Int. J. Prod. Res.*, Vol. 37(8), pp. 1881-1901, 1999.

[11] R.S. Guh and J.D.T. Tannock. Recognition of control chart concurrent patterns using a neural network approach. *Int. J. Prod. Res.*, Vol. 37(8), pp. 1743-1765, 1999.

[12] M.A. Wani, and D.T. Pham. Efficient control chart pattern recognition through synergistic and distributed artificial neural networks. *Proc. Instn Mech. Engrs, Part B: J. Engineering Manufacture*. Vol. 213(B2), pp. 157-169, 1999.

[13] D.F Cook, C.W. Zobel and M. L. Wolfe Environmental statistical process control using an augmented neural network classification approach. *European Journal of Operational Research*. Vol. 174, pp. 1631-1642, 2006.

[14] J.A. Swift. Development of a Knowledge-Based Expert System for Control Chart Pattem Recognition and Analysis. *Doctoral Thesis*. Oklahoma State University, 1987.

[15] Gerson Tontini. Pattern Identification in Statistical Process Control Using Fuzzy Neural Networks. *IEEE Management Sciences*. Department Regional University of Blumenau (FURB), 1996.

[16] G.A Carpenter, S. Grossberg, N. Markuzon, J.H. Reynolds, and D.B. Rosen. Fuzzy ARTMAP: A neural network architecture for incremental learning of analog multidimensional maps. *IEEE Transactions on Neural Networks*. Vol. 3 (5), pp. 698-713, 1992.

[17] G. A. Carpenter, S. Grossberg. Adaptive Resonance Theory (ART). *The Hanbook of Brain Theory and Neural Networks*. Edited by M. A. Arbib, The MIT Press. pp. 79-82, 1995.

[18] G. A. Carpenter, S. Grossberg. A Massively Parallel Architecture for a Self-Organizing Neural Pattern Recognition Machine. *Computer Vision, Graphics, and Image Processing*. Academic Press, Inc. pp. 54-115, 1987.

[19] G. A. Carpenter, S. Grossberg, D. B. Rosen. ART 2-A: An Adaptive Resonance Algorithm for Rapid Category Learning and Recognition. *Neural Networks*. Vol. 4, pp. 493-504, 1991.

[20] G. A. Carpenter, S. Grossberg, J. H. Reynolds. ARTMAP: Supervised Real-Time Learning and Classification of Nonstationary Data by Self-Organizing Neural Network. *Neural Networks*. pp. 565-588, 1991.

[21] J. R. Williamson. Gaussian ARTMAP: A Neural Network for Fast Incremental Learning of Noisy Multidimensional Maps. *Neural Networks*. Vol. 9, No. 5, pp. 881-897, 1996.

[22] G. A. Carpenter, W. D. Ross. ART-EMAP: A Neural Network Architecture for Object Recognition by Evidence Accumulation. *IEEE Trans. on Neural Networks*. Vol. 6, No. 4, pp. 805-818, 1995.

[23] D. Psichogios, L. Ungar. A Hybrid Neural Network-First Principles Approach to Process Modeling. *Computers & Chemical Engineering*. Vol. 38(10), pp. 1499-1511, 1992.

[24] G. Acuña, and E. Pinto. Development of a Matlab®Toolbox for the Design of Grey-Box Neural Models. *Int. Journal of Computers, Communications & Control*. Vol. I, No. 2, pp. 7-14, 2006.

[25] Imre J. Rudas, János Fodor. Intelligent Systems. *Int. Journal of Computers, Communications & Control*, 3(S):132-138, 2008.

Jose Antonio Vazquez-Lopez was born in Celaya, Guanajuato, Mexico in 1972. He graduated with a BEng and MSc in Industrial Engineering from Instituto Tecnologico de Celaya in 1999 and 2001 respectively. He obtained his PhD from the Posgrado Interinstitucional en Ciencia y Tecnologia (PICyT) in 2009. Currently, he is a Lecturer at The Instituto Tecnologico de Celaya. His main research interest is in statistical industrial engineering, where he has also served as industrial consultant for several companies.

Ismael Lopez-Juarez was born in Tequixquiac, Mexico in 1968. He obtained a BEng from The National Autonomous University of Mexico (UNAM) in 1991. He obtained an MSc in Instrument Design and Application at The University of Manchester in 1996 and a PhD in Intelligent Robotics at The Nottingham Trent University in 2000, both in th U.K. He was the founder and leader of the Mechatronics and Intelligent Manufacturing Systems Research Group at CIATEQ, A.C. during 2000-2006. Currently, he is a Principal Researcher at the Robotics and Advanced Manufacturing Systems Research Group at CINVESTAV-Saltillo. Dr. Lopez-Juarez has published over one hundred papers in refereed international journals and conference proceedings in the area of Instrumentation, Self-adaptive Industrial Robotics and Neural Networks.

Mario Peña-Cabrera was born in Queretaro, Mexico in 1954. He graduated with a BEng in Mechanical and Electrical Engineering with specialization in control, communications and electronics from The National Autonomous University of Mexico (UNAM). He holds a MEng in Electronics from the same University and a MEng from The McMaster University, Canada. He obtained his PhD degree in Science and Technology from PICyT, SEP-CONACYT, Mexico. His areas of interest are in robot vision, automation and digital control, areas in which he has been working for over 30 years. Currently, He is the Head of the Computing Engineering Science and Automation Department at The Institute of Research on Applied Mathematics and Systems in The National Autonomous University of Mexico (UNAM).