Drug Development and Industrial Pharmacy, 31:1019–1033, 2005

ISSN: 0363-9045 print / 1520-5762 online
DOI: 10.1080/03639040500306294

Taylor & Francis
Taylor & Francis Group

Artificial Neural Networks: Comparison of Two Programs for Modeling a Process of Nanoparticle Preparation

Névine Rizkalla and Patrice Hildgen
Faculté de Pharmacie, Université de Montréal, Montréal, Canada

ABSTRACT Artificial Neural Networks (ANNs) were used to predict nanoparticle size and micropore surface area of polylactic acid nanoparticles, prepared by a double emulsion method. Different batches were prepared while varying polymer and surfactant concentration, as well as homogenization pressure. Two commercial ANNs programs were evaluated: Neuroshell® Predictor, a black-box software adopting both neural and genetic strategies, and Neurosolutions®, allowing a step-by-step building of the network. Results were compared to those obtained by statistical method. Predictions from ANNs were more accurate than those calculated using non-linear regression. Neuroshell® Predictor allowed quantification of the relative importance of the inputs. Furthermore, by varying the network topology and parameters using Neurosolutions®, it was possible to obtain output values which were closer to experimental values. Therefore, ANNs represent a promising tool for the analysis of processes involving preparation of polymeric carriers and for prediction of their physical properties.

KEYWORDS Nanoparticles, Artificial neural networks, Genetic algorithm, Network architecture, Polynomial regression

INTRODUCTION

Polymeric nanoparticles (NP) offer many advantages regarding administration and targeting of drugs. Their preparation is a multivariate process where many factors interplay to affect the final product characteristics (Chorny et al., 2002; Das et al., 1995; Lamprecht et al., 2000; Vandervoort & Ludwig, 2001; Zambaux et al., 1998). Consequently, in order to obtain nanoparticles having some desired properties, many batches have to be prepared, which is both time and effort consuming, since levels of each variable are changed separately at a time, while keeping all other variables constant. Classical statistical methods, e.g., response surface methodology (McCarron et al., 1999; Yüksel et al., 2000) which may be combined with factorial design (Lescure et al., 1992; Seijo et al.,

Address correspondence to Patrice Hildgen, Faculté de Pharmacie, Université de Montréal, C.P. 6128 Succursale Centre-ville, Montréal (QC) H3C 3J7, Canada; Fax: 1 (514) 343-2102; E-mail: Patrice.hildgen@umontreal.ca

1990) have been applied to solve such problems. Unfortunately, they present some drawbacks related to a poor prediction (Takayama et al., 1999) and the need for a rigid experimental design. In addition, despite of the reduction of experimentation made possible using the factorial design, a certain amount of data remains essential for adequate description of complex processes.

Artificial Neural Networks

Artificial intelligence has recently gained much interest in the pharmaceutical field, to solve prediction or optimization problems, and to overcome limitations of the classical statistical methods. Artificial neural network (ANN) models were found to produce a better fitting and a higher prediction performance, compared to that of regression analysis, for the development of pharmaceutical solid dosage forms (Bourquin et al., 1997, 1998; Sathe & Venitz, 2003; Takahara et al., 1997; Türkoğlu et al., 1995). ANNs have been successfully applied to model controlled release drug delivery; Peh et al. (2000) used ANN to predict dissolution profiles from matrix-controlled release pellets. Ibrić et al. (2002) described a modeling and optimization study for dissolution from extended release tablets. Yüksel et al. (2000) applied ANN to model a preparation method of controlled release acrylic micospheres. Fan et al. (2004) optimized a novel formulation based on a PEGylated emulsion using ANN. Different controlled release matrices were studied using ANN by Reis et al. (2003).

Basic Concepts

ANNs are computational techniques which attempt to simulate some of the neurological processing ability of the brain, particularly learning and generalizing (Achanta et al., 1995). Fundamentally, they are composed of processing elements called "neurons," interconnected in various ways to form a network. They are able to quantify correlative patterns between input and output data pairs by means of iterative training. This process is known as "learning."

A model of a simple network arrangement is represented by the Multilayer Perceptron. The number of neurons in the input layer is directly determined by the number of input variables, while the number of neurons in the hidden layer(s) is adjusted to give the least error between the experimental and predicted output values, and finally the size of the output layer depends on the desired number of dependent variables. The network "learns" by adjusting the interconnection weights between layers. The answers the network is producing are repeatedly compared with the correct answers, and each time the connecting weights are adjusted slightly in the direction of the correct answers. After the training step, the ANN can predict outputs for new set of data when presented with only input values. This describes the generalization ability of the network.

Genetic Algorithm

Genetic algorithm (GA) is an optimization technique based on the concepts of biological evolution, specifically survival of the fittest. GA uses selection and recombination processes to generate new sample points with higher fitness. Once a subset of descriptors is found, these can be mapped to the property of interest using non-linear computational neural networks.

Some studies have combined both the ANN and the GA to model pharmaceutical processes (Agatonovic-Kustrin & Alany, 2001; Türkoğlu et al., 1999).

In a previous work, we studied the effect of various formulation variables on the physical properties of polymeric NP, containing a model DNA, prepared by a double emulsion technique (Rizkalla et al., in press). Three factors were found to be of primordial importance: namely the concentration of the biodegradable polymer in the organic phase, surfactant concentration in the aqueous phase, and the homogenizing pressure (HP) used to prepare the secondary emulsion.

The purpose of the present study is to model the aforementioned process by using an ANN methodology to predict NP size and micropore surface area (MPSA): two properties with relevant importance with respect to drug release from polymeric carriers. Nanoparticles were prepared from polylactic acid (PLA). The formulation factors mentioned above are selected as inputs in the modeling study. The performance of two ANNs commercial softwares: Namely NeuroShell® Predictor and NeuroSolutions® are evaluated. Finally, predicted output values from the ANNs were compared with the results obtained using a polynomial regression (PR) analysis.

MATERIALS AND METHODS

Materials

Poly (lactic acid) (PLA) M_w 60000 was currently synthesized in the laboratory. DNA, as sodium salt, and tris-EDTA buffer pH 8, were purchased from Sigma Chemical Company Inc., (St. Louis, MO, USA). Polyvinyl alcohol (PVA), 87–89% hydrolyzed, M_w 9000–10000 was supplied from Aldrich Chemical Company Inc. (Milwaukee, WI, USA). Dichloromethane (analytical grade) was obtained from Anachemia Chemicals (Rouses Point, NY, USA).

Preparation of Nanoparticles

NP were prepared by a double emulsion solvent evaporation method described in an earlier work (Rizkalla et al., in press). Briefly, an aqueous solution of DNA, as a model drug, was emulsified in an organic phase containing 5%, 7.5%, or 10% w/v of PLA in dichloromethane to form the primary emulsion. This was further emulsified in an aqueous phase consisting of a solution of 0.1%, 0.5%, or 1% w/v of PVA, using a high pressure homogenizer, set at 5000, 10000, 15000, or 20000 psi. The double emulsion was then magnetically stirred under reduced pressure to allow for solvent evaporation. The obtained suspension was centrifuged at 17,000 rpm for 30 min at 4°C. Finally, nanoparticles were freeze dried for 72 hrs, then kept at −20°C till further use.

Measurement of NP Size

Mean NP size was determined using photon correlation spectroscopy (Nanosizer N4 Plus, Coulter Electronics, Hialeah, Florida). Samples were adjusted to give a particle count rate between 5.10^4 and 1.10^6 counts/sec. Mean particle diameter was calculated according to a size distribution processor mode using the following conditions: fluid refractive index 1.33, temperature 20°C, viscosity 0.93 cp, angle of measurement 90°, and measuring time 60 sec.

Determination of NP Microporosity

Surface area of micropores was determined by nitrogen adsorption using a surface area and pore size analyzer (Autosorb-1, Quantachrome instruments, Miami, Florida). Briefly, 50–100 mg of nanoparticles were placed in glass sample holders and outgassed at 25°C for 30 min before analysis. Then, sample and reference tubes were immersed in liquid nitrogen at −196°C and the adsorption isotherm was obtained from the volume of nitrogen (cc/g) adsorbed onto the surface of nanoparticles as a function of relative pressure. Micropore surface area was determined using the Dubinin-Radushkevich (DR) method (Quantachrome, 2003).

Modeling Study Using ANN, GA, and Polynomial Regression (PR) Analysis

Data were modeled using two PC software packages, NeuroShell® Predictor, version 2.2 (Ward Systems Group, Fredrick, MD), and NeuroSolutions®, a demo version (Neurodimension Inc., Gainesville, FL).

Neuroshell® Predictor

This program uses an already built feed-forward network of the type "black-box" that adopts one of two strategies: the neural training strategy or the genetic training strategy based on a genetic algorithm. It is worthy to note that the genetic strategy here is not used for optimization, but it is adopted by the neural network in order to find the best weights.

NeuroSolutions®

On the contrary to NeuroShell® Predictor, this software enables the selection of the topology of the network, number of hidden neurons, learning algorithm, transfer functions, number of iterations, etc. It comprises two interfaces: NeuralBuilder© for the configuration of networks, and Neurosolutions for Excel© for data processing.

Choice of ANN Model

A number of architectures with variable complexities have been selected for this study, including:

1. **Multilayer Perceptron (MLP):** a feed-forward network trained by backpropagation.
2. **Generalized Feedforward Network (GFF):** a generalization of the MLP where connections can jump over one or more layers.

3. **Modular Network:** inputs are processed using several parallel MLPs, then results are recombined.
4. **Principal Component Analysis Network (PCA):** combines unsupervised and supervised learning. It works by finding a set of uncorrelated features, principal components, from the inputs.
5. **Self-Organizing Feature Maps (SOFM):** uses unsupervised learning to transform an input of arbitrary dimension into a one or two dimensional discrete map.
6. **Recurrent Network:** feedbacks the hidden layer to itself.
7. **The CANFIS (Co-Active Neuro-Fuzzy Interference System):** combines the membership function of fuzzy logic to a modular neural network.

Training of the Networks

Networks were trained using the delta rule back-propagation of errors (DBP). The name "back-propagation" refers to a process of propagating the error information backward from the output to the hidden neurons, during which connection weights are modified by the delta learning rule (Haykin, 1999). A

TABLE 1 **Experimental Design Used for the Modeling Study and NP Size Experimental Data**

Batch number	Polymer concentration (%w/v)	Pressure	PVA concentration (%w/v)	NP size (nm)
1	5	5000	0.1	295.3
2	5	5000	0.5	225.5
3	5	5000	1	141.3
4	5	10000	0.1	214.5
5	5	10000	0.5	177.0
6	5	10000	1	137.8
7	5	15000	0.1	229.0
8	5	15000	0.5	178.1
9	5	15000	1	128.2
10	5	20000	0.1	205.4
11	5	20000	0.5	176.3
12	5	20000	1	152.4
13	7.5	5000	0.1	387.0
14	7.5	5000	0.5	197.1
15	7.5	5000	1	169.8
16	7.5	10000	0.1	343.3
17	7.5	10000	0.5	197.0
18	7.5	10000	1	158.1
19	7.5	15000	0.1	290.6
20	7.5	15000	0.5	196.8
21	7.5	15000	1	153.2
22	7.5	20000	0.1	413.6
23	7.5	20000	0.5	188.5
24	7.5	20000	1	173.4
25	10	5000	0.1	545.2
26	10	5000	0.5	211.8
27	10	5000	1	192.6
28	10	10000	0.1	502.2
29	10	10000	0.5	194.4
30	10	10000	1	183.5
31	10	15000	0.1	437.9
32	10	15000	0.5	219.4
33	10	15000	1	178.9
34	10	20000	0.1	500.2
35	10	20000	0.5	214.9
36	10	20000	1	178.0

gradient descent method is used to minimize the error. Each network was trained three times using new random sets of initial weights and each cycle consisted of 1,000 iterations. The lowest Mean Square Error (*MSE*) was selected as the training endpoint for all networks.

Polynomial Regression (PR) Analysis

For comparison purposes, data were also analysed using SigmaStat for Windows©, version 3.0.1.

RESULTS AND DISCUSSION

Experimental Design and Obtained Data

Table 1 shows the experimental design used for the modeling study, i.e., levels of inputs for the three variables, as well as the observed mean size for 36 batches of PLA nanoparticles. These input–output sets were used as tutorial data to train the different networks. The relationship between the dependent and independent variables can be further understood by the three-dimensional plots, illustrated by the Fig. 1. This figure depicts the effect of PVA and HP on the mean size of nanoparticles for each PLA concentration. The three graphs clearly demonstrate an increase in mean NP size following polymer concentration as confirmed by an earlier viscosity measurement of 2.08 (±0.02), 2.51 (±0.01), and 3.00 (±0.04) cp for the three organic solutions at 5%, 7.5%, and 10% w/v of PLA/dichloromethane respectively. This increased viscosity results in a high tendency towards aggregation during solvent evaporation (Mainardes & Evangelista, 2005). On the other hand, when polymer concentration is kept constant, a decrease in the NP mean size could be observed with increasing PVA concentration, which is related to its stabilizing effect on the formed emulsion, and its prevention of

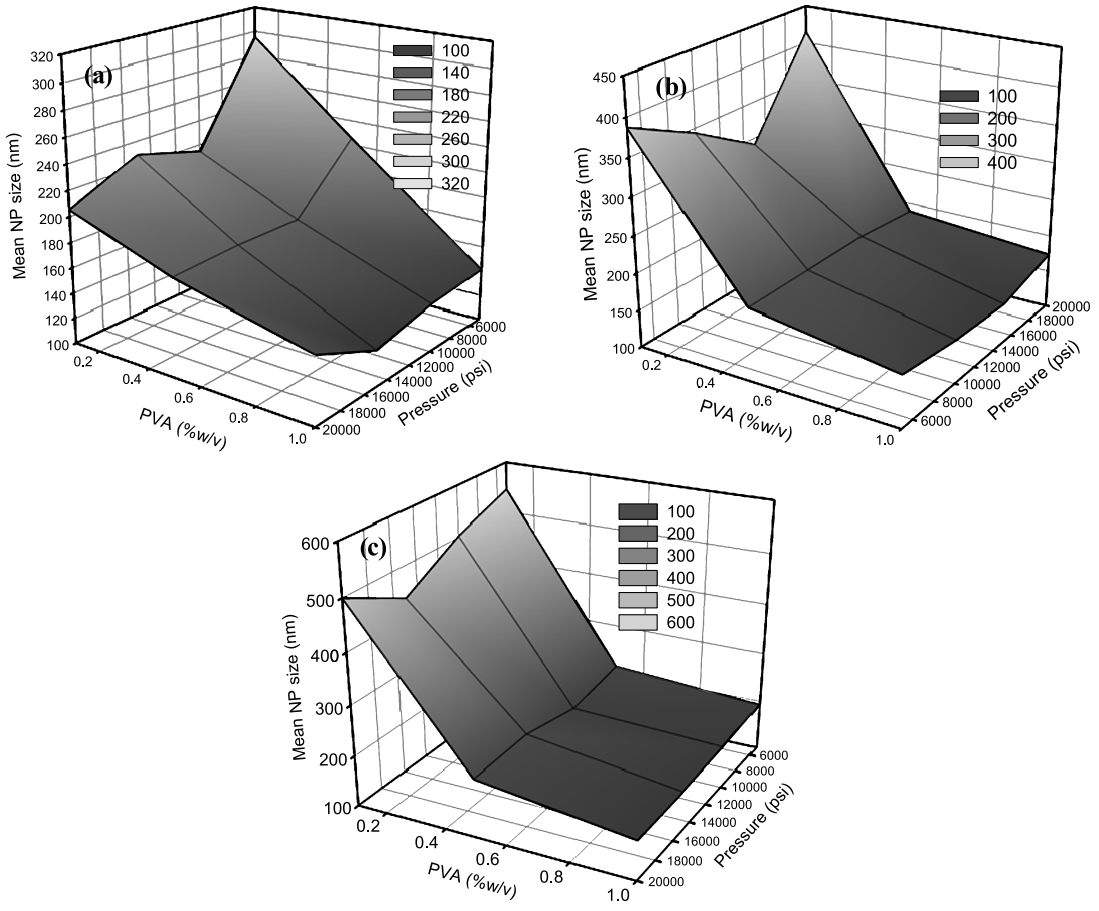

FIGURE 1 **Three-Dimensional Plots for the Effect of PVA Concentration and Homogenization Pressure on NP Size at 5%, 7.5%, and 10% w/v of PLA.**

aggregation of the newly formed droplets (Zweers et al., 2003). Graphs shown in Fig. 1(b) and 1(c) exhibit a biphasic aspect, with an abrupt drop in NP size from 400–500 nm to nearly 150–200 nm as PVA concentration increases from 0.1 to 1% w/v. Afterwards, changes in NP size occurred more gradually, till the highest PVA concentration (1% w/v). Finally, effect of HP was merely dependent on PVA, being more pronounced at low PVA concentrations, where increases in the shear forces leads to a subdivision of the emulsion globules, but almost vanishing at high concentrations of PVA, as the latter becomes then the most determining factor on NP size.

Neuroshell® Predictor

Training and Prediction

Training results generated by both neural and genetic strategies using the entire set were compared to the observed values of NP size (Fig. 2a) and MPSA (Fig. 2b). According to these two figures, the neural strategy was more successful in learning the existing relationship between input and output sets as the corresponding curve is nearly super imposable to that of the actual values. This was further confirmed by statistical data reflecting the performance of both models during training as summarized in Table 2. For

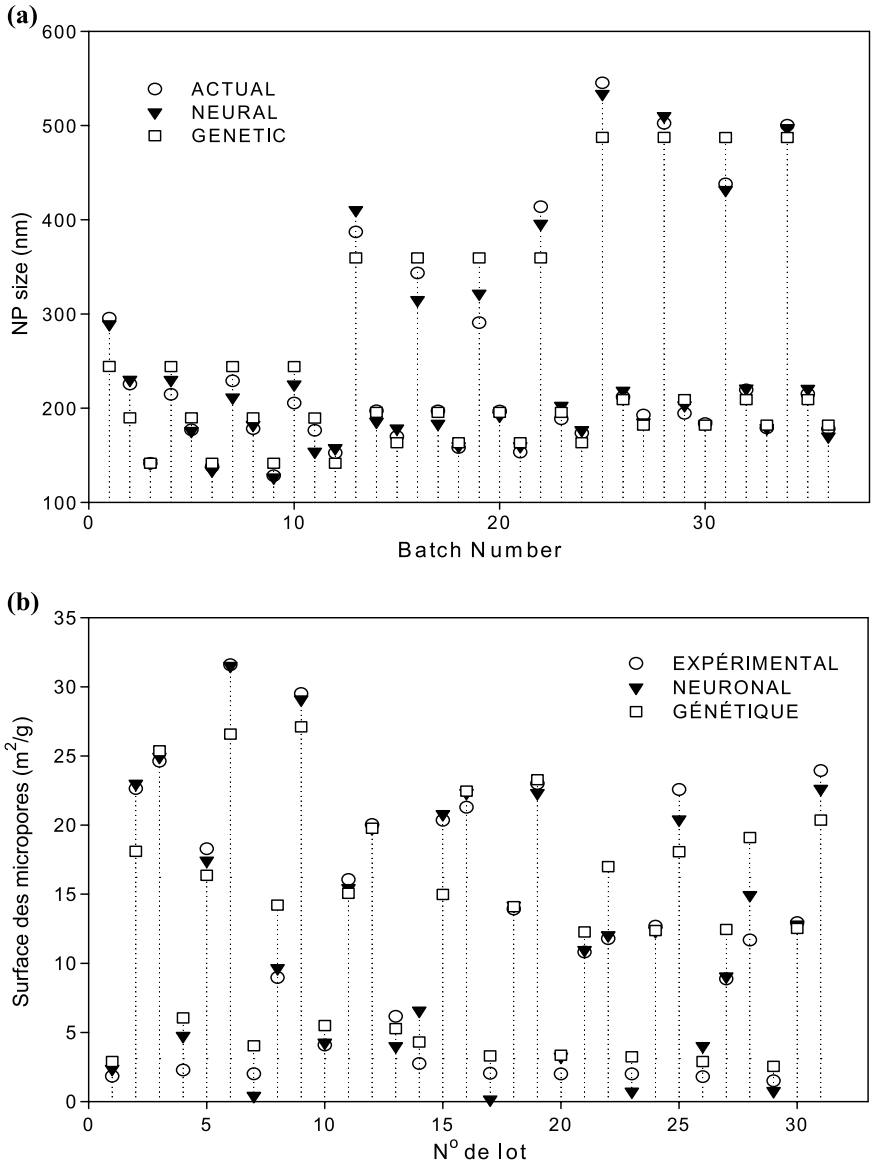

FIGURE 2 **Actual Values and Predictions of Both Neural and Genetic Strategies Using the Entire Data Set for NP Size (a) and MPSA (b).**

TABLE 2 Statistical Parameters for the Learning Performance of Neural and Genetic Models of NeuroShell® Predictor

	Neural model[a]	Genetic model[b]
NP size		
R^2	0.987772	0.912239
Average error	9.297591	22.84510
Correlation coefficient	0.993867	0.955136
MSE[c]	149.4141	1072.372
RMSE[d]	12.22351	32.74709
MPSA		
R^2	0.972990	0.849139
Average error	1.091385	2.646498
Correlation coefficient	0.986407	0.926279
MSE	2.198061	12.28131
RMSE	1.482586	3.504770

[a]No. of hidden neurons: 31.
[b]No. of generations: 204.
[c]Mean squared error.
[d]Root mean squared error.

example, concerning NP size data, a lower mean square error *MSE* (149.4) and a higher R^2 (0.987772) could be observed for the neural model, compared to the corresponding parameters of the genetic strategy (1072.4 and 0.912239, respectively). The *MSE* is defined as:

$$MSE = \frac{(V_{act} - V_{pred})^2}{N} \quad (1)$$

Errors are squared to penalize the larger errors and to cancel the effect of the positive and negative values of the differences.

In order to test the generalizing ability of the ANN, input–output values of 12 batches prepared at 7.5%w/v of PLA were excluded from the training set, and the network re-trained; later, the prediction capacity was tested on theses batches. Statistical parameters for training and prediction of NP size and MPSA are represented in Tables 3 and 4, respectively. While genetic strategy produced more precise predictions for NP size ($MSE_{\text{genetic}} = 779$) compared to the neural strategy ($MSE_{\text{neural}} = 1866.8$), the reverse occurred for the results of MPSA ($MSE_{\text{neural}} = 7.0$, $MSE_{\text{genetic}} = 12.9$). This observation leads to the conclusion that no general rule may govern such prediction abilities. Türkoğlu et al. (1999) have reported better prediction by the genetic model using this type of software. Although neural and genetic strategies showed variable degrees of fitting, it was not possible to improve the performance of either model since no modification in their structure was available.

Relative Importance of Inputs

Importance of input values is a relative measure of how significant each of the inputs is to the predictive model. The genetic strategy uses a genetic algorithm which is based on the survival of the fittest, to determine a scheme of weights for the inputs. GA would test different schemes till it finds the one that gives the best predictions for the training data (weights

TABLE 3 Statistical Parameters for the Generalization Capacity of Neural and Genetic Models of NeuroShell® Predictor for NP Size Data

	Neural model	Genetic model
Training parameters		
R^2	0.931298	0.934281
Average error	20.45882	21.24865
Correlation coefficient	0.965038	0.966933
MSE	981.9074	939.2708
RMSE	31.33540	30.64752
Testing parameters		
R^2	0.768583	0.903412
Average error	28.50528	17.29252
Correlation coefficient	0.881519	0.951350
MSE	1866.766	778.9807
RMSE	43.20146	27.91022

TABLE 4 Statistical Parameters for the Generalization Capacity of Neural and Genetic Models of NeuroShell® Predictor for MPSA Data

	Neural model	Genetic model
Training parameters		
R^2	0.942331	0.593145
Average error	1.390984	5.461419
Correlation coefficient	0.970737	0.779657
MSE	5.186495	36.59052
RMSE	2.277388	6.04901
Testing parameters		
R^2	0.889145	0.796128
Average error	1.967099	3.067811
Correlation coefficient	0.944313	0.900365
MSE	7.027285	12.92372
RMSE	2.650903	3.594958

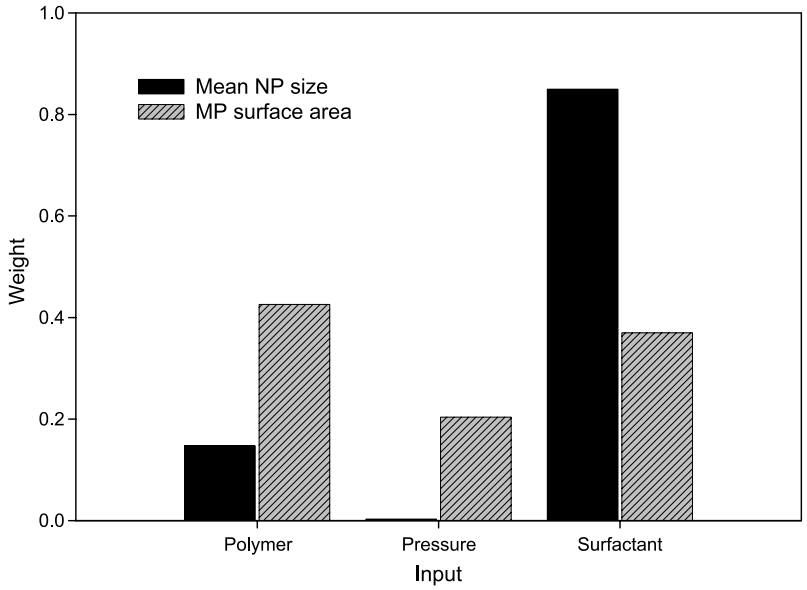

FIGURE 3 **Relative Importance of Inputs on Measured Outputs as Determined by the Genetic Algorithm.**

range from 0 to 1). Higher values are associated with more important variables (inputs). Figure 3 shows the relative importance of inputs on the NP size and MPSA. A major contribution of PVA regarding NP size can be clearly observed from the graph (weight= 0.850), while a lower influence was attributed to polymer concentration (weight=0.147). HP appeared to exert only a minimum effect (weight=0.003). On the contrary, MPSA was influenced by the three variables, with the highest contribution attributed to polymer concentration (weight=0.426), followed by surfactant concentration (weight=0.370), and finally, HP (weight=0.204). It is reported that polymer concentration affects the viscosity of the organic phase, and hence controls the evaporation rate of the organic solvent, which seems to be a major determinant of microporosity. On the other hand, the concentration of PVA is believed to have an influence on the rate of extraction (diffusion) of the organic solvent, and, consequently, affects the formation of micropores. Finally, the HP appears to have a relatively small influence on MPSA, and this role might be related to the effect of HP on NP size: As HP rises to a certain limit, NP size decreases (Fig. 1). Since smaller NP will solidify faster than larger ones, this could enable them to retain a certain microporosity,

TABLE 5 **Results of Polynomial Regression Analysis on NP Size and MP Surface Area at 10% w/v of PLA**

	Coefficients	Standard error	*t*	*p*
NP size				
y_0	675.0930	43.0386	15.6858	<0.0001
a	−0.0119	0.0072	−1.6521	0.1425
b	1156.8624	86.6257	−13.3547	<0.0001
c	0.0000	0.0000	1.4785	0.1828
d	735.4064	75.6379	9.7227	<0.0001
MPSA				
x_0	9816.7000	371.6625	26.4129	<0.0001
y_0	0.8516	0.0537	15.8607	<0.0001
a	23.3866	2.4682	9.4753	0.0007
b	4881.7841	474.8516	10.2807	0.0005
c	0.3552	0.0678	5.2375	0.0064

which would otherwise be reduced by the shrinkage of polymer chains in the semi-solid state.

Comparison with Polynomial Regression Analysis (PR)

Application of polynomial regression to data modeling resulted in two different equations for NP size and MPSA. The best fitting model for NP size was achieved using the equation

$$f = y_0 + ax + by + cx^2 + dy^2 \qquad (2)$$

where f is the NP size, x is the homogenization pressure, and y is PVA concentration, while a Gaussian model was applied for micropore surface area, represented by the equation:

$$f = a\exp(-0.5(((x - x_0)b)^2 + ((y - y_0)c)^2)) \qquad (3)$$

where f is the micropore surface, x is the homogenization pressure, and y is PVA concentration. Both equations were applied for the three different concentrations of PLA. For illustration, results of PR at 10% for both NP size and MPSA are summarized in Table 5. The t statistic is defined as the ratio of the regression coefficient to its standard error. A larger t value indicates a higher significance of the corresponding independent variable. For the size data, the model provided a significant fit ($p < 0.001$, $R^2 = 0.9828$), with a significant PVA concentration

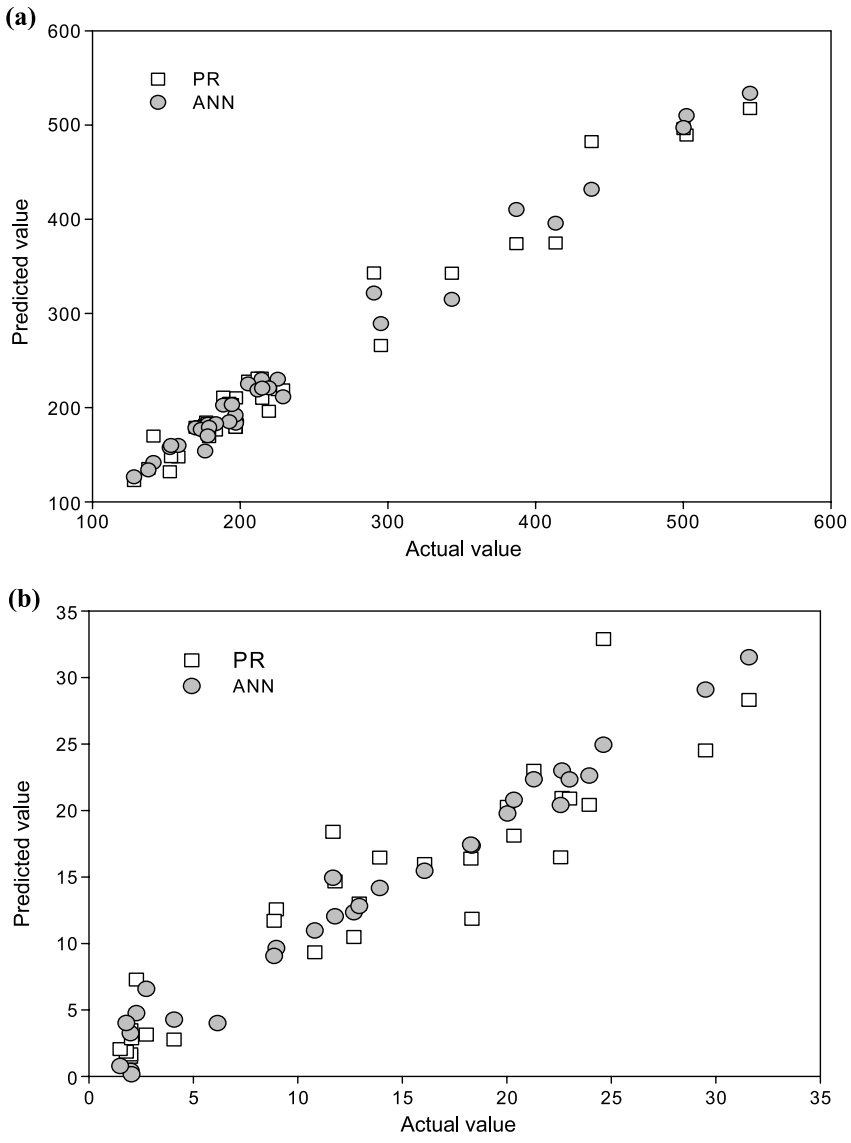

FIGURE 4 **Correlation Plots Between Both Neural and Regression Predictions and Experimental Values for (a) NP Size and (b) MPSA.**

($p < 0.0001$) but a non-significant pressure ($p = 0.1425$ and 0.1828) dependency. However, both variables appeared to significantly affect the MPSA (Table 5).

On the other hand, correlation plots shown in Fig. 4(a) and 4(b) illustrate a lower fitting obtained with PR ($R^2 = 0.96944$ and 0.86537) compared to ANN ($R^2 = 0.98772$ and 0. 97299).

Response surfaces calculated by the PR, the ANN, and the GA were plotted within the experimental limits used. Figures 5 and 6 show the three dimensional diagrams of the two response variables as a function of HP and of PVA concentration at a constant polymer concentration (10% w/v). In each figure, (a) represents the experimental data, (b) the response surface calculated by the polynomial regression, while (c) and (d) show the response as predicted by the ANN and the GA models, respectively. As it can be observed from the two graphs, the PR exhibited relatively plain surfaces for the two responses [Figs. 5(b) and 6(b)]. On the other hand, non-linear relationships were successfully mapped by the ANN [Figs. 5(c) and 6(c)]. This behavior has been reported in the literature (Takahara et al., 1997), and it is related to the capacity of the ANN to identify complex relationships between causal factors and measured responses, when they cannot be expressed by any known regression equation. Predictions of the genetic algorithm, Figs. 5(d) and 6(d), were also clearly lacking non-linearity features.

NeuroSolutions®: A Comparison Between Different Models

Using NeuroSolutions®, we focused on the effect of network topology, number of hidden neurons, and properties of training data on the performance of the network. While using Neuroshell® predictor, a separate network was built for each output; a major advantage of this program was the possibility of

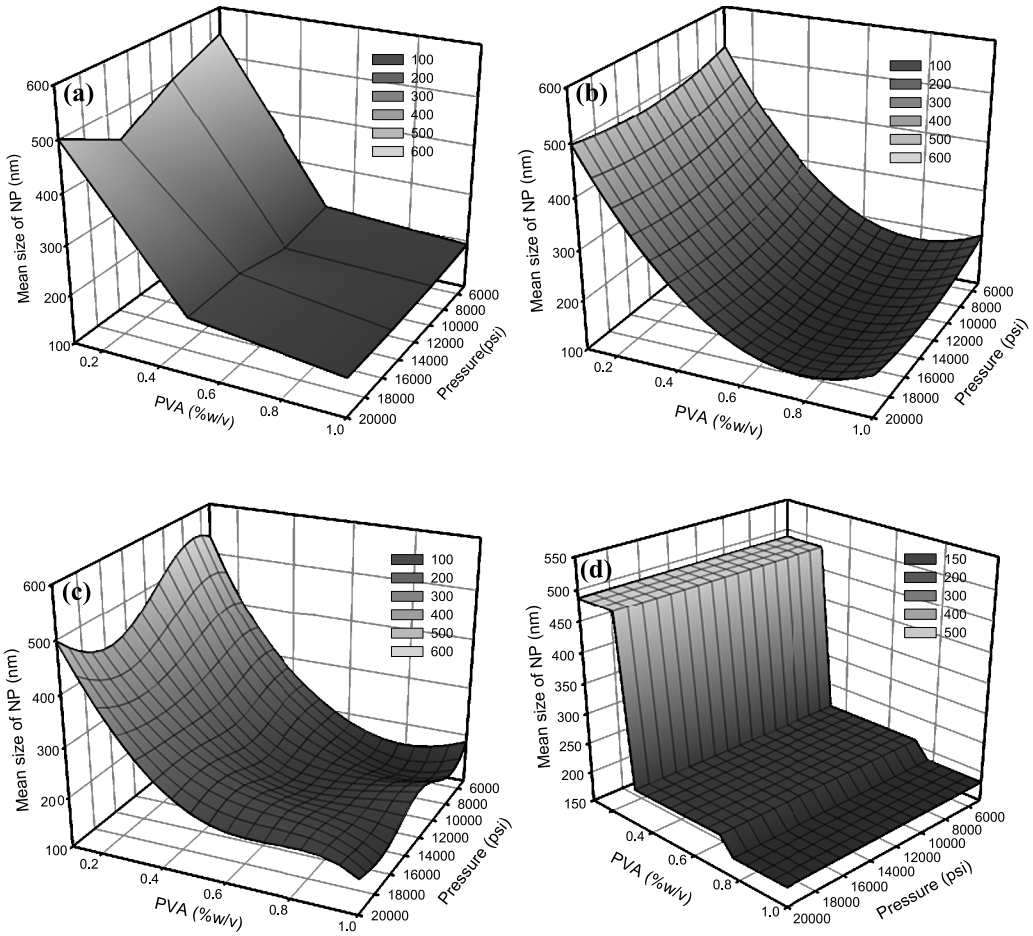

FIGURE 5 **Response Surface Plots Showing the Effect of PVA and HP on NP Size: (a) Experimental Data, (b) PR Model, (c) Neural Model, and (d) Genetic Model.**

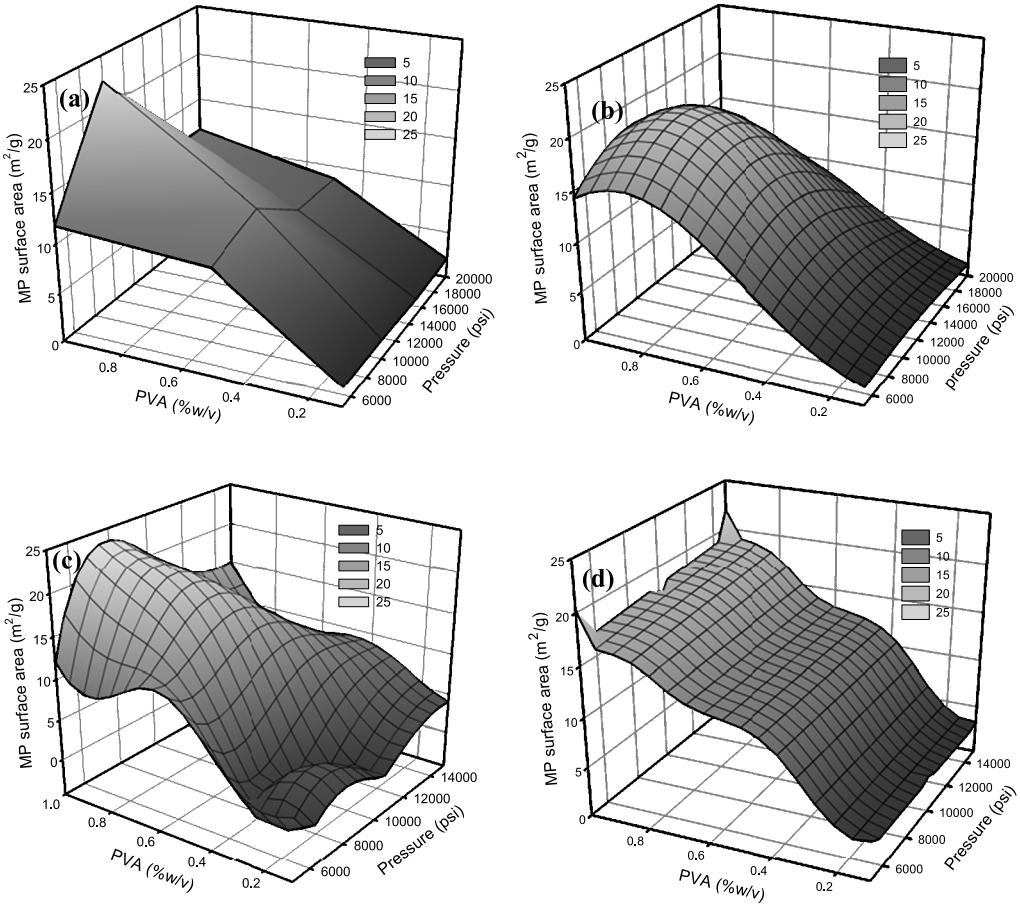

FIGURE 6 **Response Surface Plots Showing the Effect of PVA and HP on MPSA at 10% w/v of PLA: (a) Experimental Data, (b) PR Model, (c) Neural Model, and (d) Genetic Model.**

training one network for the two response variables (NP size and MPSA). It is worthy to note that regression analysis has also used two separate equations that could be considered as two independent models.

Network Topology

Hidden and output layer neurons were selected to have a hyperbolic tangent transfer function. The reason is that it is a very flexible, non-linear function. It is also continuous and differentiable. Momentum learning was applied. Step size and momentum were kept as default settings.

Different topologies demonstrated variable correlations between predicted and actual values, with R^2 values ranking in the order: CANFIS (0.9976)>PCA (0.9904)>SOFM (0.9888)>MLP (0.9876)>GFF (0.9707)>Modular (0.9041)>Recurrent (0.8766) (correlation plots not shown). In addition, the first three networks presented values of R^2 higher than that of Neuroshell® Predictor, while MLP presented a nearly equal value. The excellent learning and predicting ability of the CANFIS network might be the result of integration of a fuzzy model with a modular neural network. Fuzzy logic is based upon establishing a degree of membership criterion between input and output variables, which is wider than the classical canonical value of zero or one (Jamshidi, 2003). In other words, it is a methodology which establishes complex relationships on a continuous, rather than discrete (all or none) basis. The neurofuzzy model here seems to emphasize the synergy between the fuzzy logic and the neural network computing.

Both PCA and SOFM are based upon feature extraction of data. The unsupervised segment of the PCA finds a set of uncorrelated features from the input data, and the supervised segment of the network performs the nonlinear classification of these components using a MLP. This network converges fast. On the other hand, SOFM networks, being an alternative to classical PCA, do not use input and output data for training; instead, they map the entire training data set

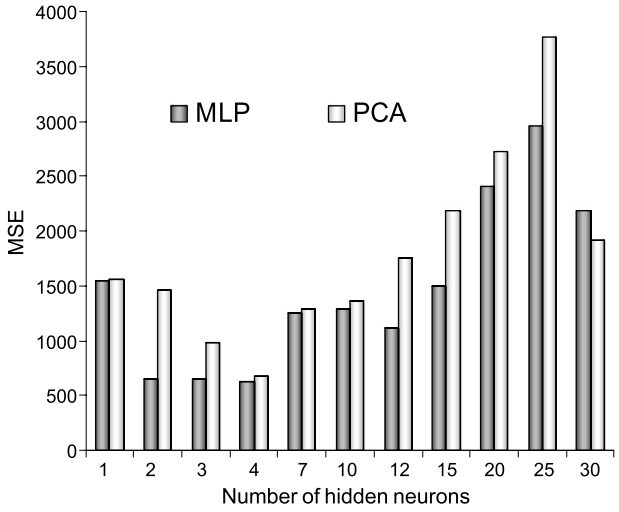

FIGURE 7 **Effect of the Number of Hidden Neurons on the Performance of MLP and PCA Networks.**

at once, which is considered an excellent method at finding relationships among variables (Agatonovic-Kustrin & Beresford, 2000). Hence, these two models provide a fast and efficient learning.

The least performance was observed with the recurrent model. Despite being considered as a powerful network, it is difficult in training and shows instability; consequently, more training cycles have to be applied in order to achieve an acceptable learning level.

Number of Hidden Neurons

Effect of the number of hidden neurons on the network performance has been studied on two networks with distinctly different architectures: namely MLP and PCA. Number of hidden neurons has been increased gradually, and the *MSE* was recorded each time. When only learning was measured, size data for 36 batches were used for training. Number of hidden neurons was increased gradually up to 30 neurons. A decrease in *MSE* was observed when hidden neurons were increased up to 15, which corresponded to the least error. Addition of hidden neurons beyond this value hindered the learning process (data not shown), as this would slow down the training process (Hussain et al., 1991). However, when prediction capacity was tested, training was performed with NP size data of 24 sets, representing batches at 5% and 10% PLA and then both networks were tested on the remaining 12 batches made with 7.5% w/v of PLA. The number of hidden neurons was increased gradually starting from one neuron, and a new network was trained each time. *MSE* obtained during the testing (prediction) was followed as a function of the number of hidden neurons as illustrated in Fig. 7. Although the two topologies demonstrated some differences in *MSE*, the same tendency could be easily observed from the graph. *MSE* decreased with increasing neurons from 1 to 4 then increased again until 25 neurons, before showing a slight decrease at 30 neurons. The increase in *MSE* is consistent with the over-fitting which reduces the generalization ability of the network making it simple to memorize the training sets.

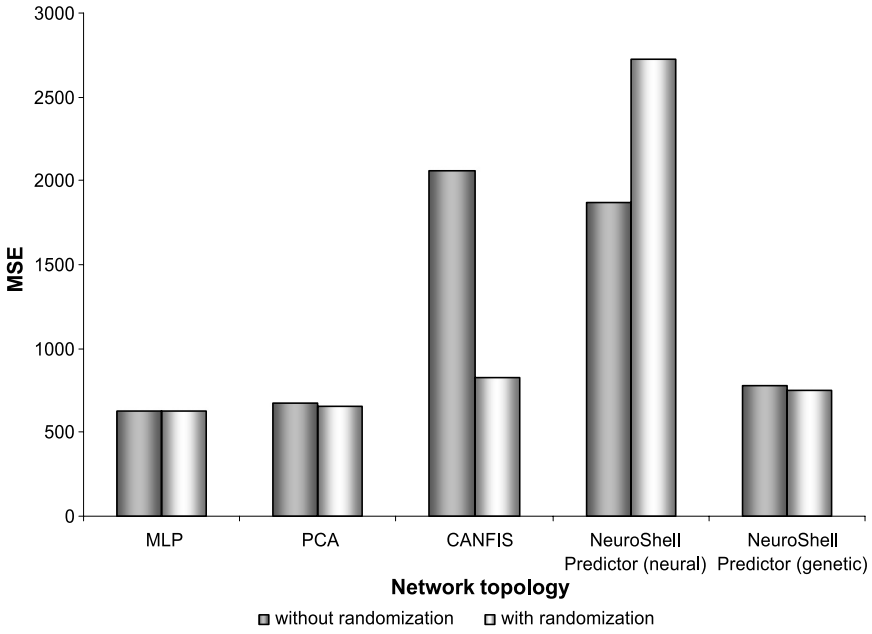

FIGURE 8 **Comparison Between ANN Performance Before and After Randomization for Different Network Topologies.**

These results suggest that the optimum number of hidden neurones could be exclusively found by a trial-and-error approach, even if some rules have been suggested for its determination (Carpenter & Hoffman, 1995). In addition, an optimum number for training might not be suitable to achieve an acceptable generalization level.

Randomization

This pretreatment of input sets constitutes a challenge to the learning and generalization abilities of the ANN, since training data will not be presented to the network according to a rigid statistical plan (Bourquin et al., 1998). Figure 8 shows *MSE* for three network models, MLP, PCA, and the CANFIS network before and after randomization. For comparison purposes, *MSE* of neural network from NeuroShell® Predictor was also included. Networks demonstrated variable responses towards randomization. No significant difference was observed before and after randomization for the MLP and PCA, in addition to the genetic model of NeuroShell® Predictor, which indicates no sensitivity of these models regarding arrangement of training sets. On the contrary, the CANFIS network showed an improvement in the data fitting, with a *MSE* decreasing from 2060 to 827, which means a better generalization capacity (Bourquin et al., 1997). Concerning the neural strategy, it exhibited a dramatic increase in the *MSE* from 1867 to 2724, demonstrating a less flexibility of the model towards the arrangement of the training data fed to the network.

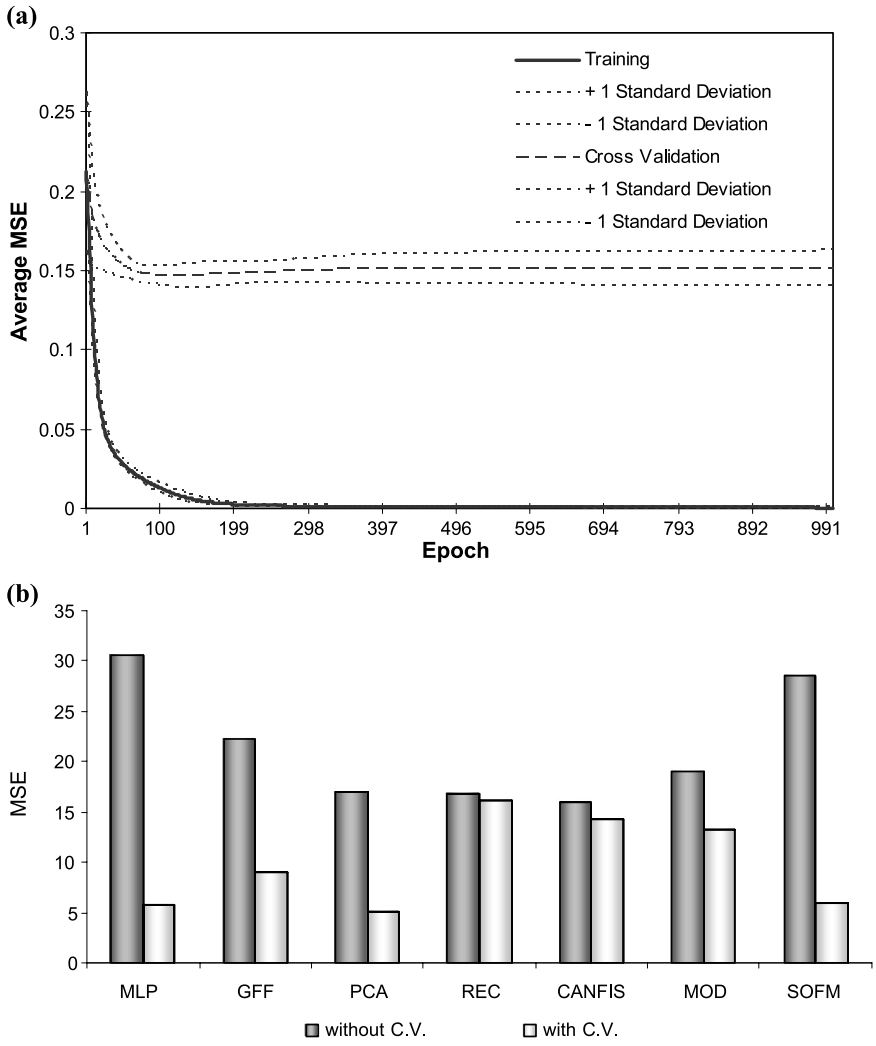

FIGURE 9 **(a) Average *MSE* with Standard Deviation Boundaries for Three Training Cycles with Cross-validation. (b) Effect of Cross-validation on the Performance of ANNs.**

Cross Validation (CV)

Input–output sets may be fed to the ANN in different ways. In the case of cross validation, a predetermined set is eliminated from the training data and is used to validate the network, i.e., to test its generalizing ability during training as it is shown in Fig. 9(a). The differences between the network output and target values of the validation data set are monitored, and the training is stopped when these differences reach a minimum. In this study, data representing values of MPSA were classified into training, cross-validation, and testing data. A significant improvement in prediction ability after cross-validation was observed for all models, except for the CANFIS and the recurrent networks, which show only a slight difference between the two values, as seen in Fig. 9(b).

This technique is known to be of high efficacy, and it is considered to be the basis of the "leave-one-out" method (Sun et al., 2003). During this process, the ANN is trained on all except one of the samples in the data set, and then validated with the sample which was omitted from the data set during training. This technique enables the use of the entire data set while allowing in the same time for an efficient on-training validation.

CONCLUSION

Artificial neural networks offered a successful tool for nanoparticle preparation analysis and modeling. Genetic algorithm represented a fast and reliable method to determine the relative importance of inputs. Predictions from ANNs were closer to experimental values than those obtained using polynomial regression analysis.

Results obtained using NeuroSolutions® confirmed that it is more advantageous to use a flexible, rather than a "black-box" program, as this would enable the free selection of different network parameters in an appropriate way to each problem. Also, pre-processing of the training data has proved to play an important role in modeling applications by neural computing.

REFERENCES

Achanta, A. S., Kowalski, J. G., & Rhodes, C. T. (1995). Artificial neural networks: implications for pharmaceutical sciences. *Drug Development and Industrial Pharmacy*, *21*, 119–155.

Agatonovic-Kustrin, S., & Alany, R. G. (2001). Role of genetic algorithms and artificial neural networks in predicting the phase behavior of colloidal delivery systems. *Pharmaceutical Research*, *18*, 1049–1055.

Agatonovic-Kustrin, S., & Beresford, R. (2000). Basic concepts of artificial neural network (ANN) modeling and its application in pharmaceutical research. *Journal of Pharmaceutical and Biomedical Analysis*, *22*, 717–727.

Bourquin, J., Schmidli, H., van Hoogevest, P., & Leuenberger, H. (1997). Basic concepts of artificial neural networks (ANN) modeling in the application to pharmaceutical development. *Pharmaceutical Development and Technology*, *2*, 95–109.

Bourquin, J., Schmidli, H., Van Hoogevest, P., & Leuenberger, H. (1998). Comparison of artificial neural networks (ANN) with classical modelling techniques using different experimental designs and data from a galenical study on a solid dosage form. *European Journal of Pharmaceutical Sciences*, *6*, 287–300.

Carpenter, W. C., & Hoffman, M. E. (1995). Understanding neural network approximations and polynomial approximations help neural network performance. *AI Expert*, *10*, 31–33.

Chorny, M., Fishbein, I., Danenberg, H. D., & Golomb, G. (2002). Lipophilic drug loaded nanospheres prepared by nanoprecipitation: effect of formulation variables on size, drug recovery and release kinetics. *Journal of Controlled Release*, *83*, 389–400.

Das, S. K., Tucker, I. G., & Ganguly, N. (1995). Evaluation of poly(isobutylcyanoacrylate) nanoparticles for mucoadhesive ocular drug delivery. I. Effect of formulation variables on physicochemical characteristics of nanoparticles. *Pharmaceutical Research*, *12*, 534–540.

Fan, T., Takayama, K., Hattori, Y., & Maitani, Y. (2004). Formulation optimization of paclitaxel carried by PEGylated emulsions based on artificial neural network. *Pharmaceutical Research*, *21*, 1692–1697.

Haykin, S. (1999). *Neural Networks a Comprehensive Foundation* (2nd ed.). Upper Saddle River, NJ: Prentice-Hall, Inc..

Hussain, A. S., Xuanqjang, Y., & Johnson, R. D. (1991). Application of neural computing in pharmaceutical product development. *Pharmaceutical Research*, *8*, 1248–1252.

Ibrić, S., Jovanović, M., Djurić, Z., Parojčić, J., & Solomun, L. (2002). The application of generalized regression neural network in the modeling and optimization of aspirin extended release tablets with Eudragit RS PO as matrix substance. *Journal of Controlled Release*, *82*, 213–222.

Jamshidi, M. (2003). Tools for intelligent control: fuzzy controllers, neural networks and genetic algorithms. *Philosophical Transactions of the Royal Society of London, A*, *361*, 1781–1808.

Lamprecht, A., Ubrich, N., Hombreiro Pérez, M., Lehr, C.-M., Hoffman, M., & Maincent, P. (2000). Influences of process parameters on nanoparticle preparation performed by a double emulsion pressure homogenization technique. *International Journal of Pharmaceutics*, *196*, 177–182.

Lescure, F., Zimmer, C., Roy, R., & Couvreur, P. (1992). Optimization of poly(alkylcyanoacrylate) nanoparticles: influence of sulphur dioxide and pH on nanoparticles characteristics. *Journal of Colloid and Interface Science*, *154*, 77–86.

Mainardes, R. M., & Evangelista, R. C. (2005). PLGA nanoparticles containing praziquantel: effect of formulation variables. *International Journal of Pharmaceutics*, *290*, 137–144.

McCarron, P. A., Woolfson, A. D., & Keating, S. M. (1999). Response surface methodology as a predictive tool for determining the effects of preparation conditions on the physicochemical properties of poly(isobutylcyanoacrylate) nanoparticles. *International Journal of Pharmaceutics*, *193*, 37–47.

Peh, K. K., Lim, C. P., Queck, S. S., & Khoh, K. H. (2000). Use of artificial neural networks to predict drug dissolution profiles and evaluation of network performance using similarity factor. *Pharmaceutical Research*, *17*, 1384–1388.

Quantachrome (2003). *Autosorb-1 Operating Manual, Version 1.27*. Quantachrome Instruments, Boysiton Beach, Florida, USA.

Reis, M. A. A., Sinisterra, R. D., & Belchior, J. C. (2003). An alternative approach based on artificial neural networks to study controlled drug release. *Journal of Pharmaceutical Sciences*, *93*, 418–430.

Rizkalla, N., Range, C., Lacasse, F.-X., & Hildgen, P. Effect of various formulation parameters on the properties of polymeric nanoparticles prepared by multiple emulsion method, in press.

Sathe, P. M., & Venitz, J. (2003). Comparison of neural network and multiple linear regression as dissolution predictors. *Drug Development and Industrial Pharmacy*, *29*, 349–355.

Seijo, B., Fattal, E., Roblot-Treupel, L., & Couvreur, P. (1990). Design of nanoparticles of less than 50 nm diameter: preparation, characterization and drug loading. *International Journal of Pharmaceutics*, *62*, 1–7.

Sun, Y., Peng, Y., Chen, Y., & Shukla, A. J. (2003). Application of artificial neural networks in the design of controlled release delivery systems. *Advanced Drug Delivery Reviews*, *55*, 1201–1215.

Takahara, J., Takayama, K., & Nagai, T. (1997). Multi-objective simultaneous optimization technique based on an artificial neural network in sustained release formulations. *Journal of Controlled Release*, *49*, 11–20.

Takayama, K., Fujikawa, M., & Nagai, T. (1999). Artificial neural network as a novel method to optimize pharmaceutical formulations. *Pharmaceutical Research*, *16*, 1–6.

Türkoğlu, M., Ozarslan, R., & Sakr, A. (1995). Artificial neural network analysis of a direct compression tabletting study. *European Journal of Pharmaceutics and Biopharmaceutics*, *41*, 315–322.

Türkoğlu, M., Aydin, I., Murray, M., & Sakr, A. (1999). Modeling of a roller-compaction process using neural networks and genetic algorithms. *European Journal of Pharmaceutics and Biopharmaceutics*, *48*, 239–245.

Vandervoort, J., & Ludwig, A. (2001). Preparation factors affecting the properties of polylactide nanoparticles: a factorial design study. *Pharmazie*, *56*, 484–488.

Yüksel, N., Türkoğlu, M., & Baykara, T. (2000). Modelling of the solvent evaporation method for the preparation of controlled release acrylic microspheres using neural networks. *Journal of Microencapsulation*, *17*, 541–551.

Zambaux, M. F., Bonneaux, F., Gref, R., Maincent, P., Dellacherie, E., Alonso, M. J., Labrude, P., & Vigneron, C. (1998). Influence of experimental parameters on the characteristics of poly(lactic acid) nanoparticles prepared by a double emulsion method. *Journal of Controlled Release*, *50*, 31–40.

Zweers, M. L. T., Grijpma, D. W., Engbers, G. H. M., & Feijen, J. (2003). The preparation of monodisperse biodegradable polyester nanoparticles with a controlled size. *Journal of Biomedical Materials Research*, *66B*, 559–566.