PREDICTION OF RHEOMETRIC PROPERTIES OF COMPOUNDS BY USING ARTIFICIAL NEURAL NETWORKS

GUSTAVO ARIEL SCHWARTZ*

UNIVERSIDAD DE BUENOS AIRES, FACULTAD DE CIENCIAS EXACTAS Y NATURALES, DPTO. DE FÍSICA, CIUDAD UNIVERSITARIA, PABELLÓN 1, BUENOS AIRES (1428), ARGENTINA

ABSTRACT

The ability of an Artificial Neural Network (ANN) to evaluate the variability of rheometric properties of rubber compounds from their formulation is presented. Because of the complexity and non-linearity of mixing processes; an exact mathematical treatment of the problem is extremely difficult, or even impossible. The use of artificial neural networks (ANNs) might be very useful to analyze these processes, since they have the ability to map nonlinear relationships without prior information about process or system models. In this work a three-layer ANN is used and the optimum parameters are determined. The results are compared with theoretical and experimental published data. The dependence of the rheometric properties as a function of compound components is also analyzed. Finally, the sensibility matrix concept is introduced. The sensibility matrix allows us to calculate the minimum expected variability, for a given compound, due to the weight tolerances of its components.

INTRODUCTION

The rheometric properties of a compound are affected by many factors, such as the variability of the raw material, weighing errors, process variations, etc. It's very difficult to know exactly how these factors affect the final properties of a given compound.[1,2] Due to the characteristics of the compounds and the complexity of the elaboration process, it is virtually impossible to predict the rheometric properties of the compound in terms of first principles. The use of artificial neural networks (ANNs) can help us to find complicated dependencies among input variables (formulation of the compound) and output ones (properties). The ANNs have the ability to map nonlinear relationships among variables, without prior information about the process.[3,4] Recently, ANNs have been used successfully to predict the copolymer composition as a function of reaction conditions and conversion.[5]

We can consider the mixing process as a "grey box," and train a neural network using the available experimental data. The ANNs are mathematical models that have the ability to learn the correlation between input and output values. The training of the network consists of introducing a set of correlated inputs and outputs, called examples. From these examples, the network goal is modeling the relationship between the input and output variables, by adjusting the node connection weights. When the system converges to a stable solution we can enter a new formulation and get its rheometric properties.

In this work, the implementation of a neural network is analyzed in order to predict the rheometric properties of new compounds from their formulation. Network characteristics are analyzed in order to guarantee the stability and convergence of the solutions, and some applications and possible extensions of this kind of treatment are discussed.

IMPLEMENTATION

In this work, the analog version of a commercial ready-made ANN called NeuroShell was used. In order to simplify the model, only a subset of components and a subset of properties were taken into

* Tel: 54-11-4576-3300 (x255); fax: +(54 11) 4576-3357; e-mail: schwartz@df.uba.ar

TABLE I

COMPOSITION OF THE SELECTED COMPOUNDS[a]

Code	NR	CB	Oil	Sulfur	Acc	PA	M_L	M_H	t 50
Comp 01	10.000	7.366	0.946	10.000	6.061	6.667	5.308	8.788	5.162
Comp 02	10.000	7.366	0.708	10.000	6.061	6.667	6.812	8.426	4.824
Comp 03	10.000	6.273	1.138	7.250	3.939	8.333	4.773	5.024	4.868
Comp 04	6.000	7.205	10.000	6.000	6.667	8.333	4.222	4.243	6.191
Comp 05	10.000	6.150	1.672	7.300	3.939	10.000	4.463	4.716	4.912
Comp 06	10.000	6.110	1.937	7.300	3.939	1.667	3.414	4.486	4.191
Comp 07	10.000	7.666	1.302	5.625	1.212	6.667	6.759	5.358	4.838
Comp 08	8.000	4.605	2.284	4.350	6.061	6.667	3.264	3.037	6.743
Comp 09	3.500	7.268	3.011	5.500	6.970	10.000	6.841	7.275	6.257
Comp 10	5.000	8.146	0.722	7.500	5.455	0.000	7.657	10.000	5.809
Comp 11	0.000	10.000	4.244	7.500	4.000	10.000	10.000	9.550	10.000
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

[a] All values are normalized in the range 0–10 (see text).

account. Six relevant variables (components) of compounds formulations were considered in this analysis: weight percentage of carbon black (CB) and oil, phr (weight per hundred rubber) of natural rubber (NR), sulfur, accelerator (Acc) and process aid (PA). Others ingredients in the formulation were considered not important. Three output variables were taken into account: low torque (M_L), high torque (M_H) and time to 50% cure (t 50). Twenty-four compounds were selected. The values were normalized in the range 0–10. Table I shows a partial list of the selected compounds. The values of rheometric properties are the average of more than one-hundred samples of each compound.

NEURAL NETWORKS ARCHITECTURE

The ANNs have evolved from the intent of modeling the cerebral activity. The basic processing element of a neural network is the artificial neuron or simply node. The basic artificial neuron can be modeled as a non-lineal device of multiples inputs, each one with a weight w_{ij}, and a nonlinear function $f(x)$, usually sigmoid. This simple model of an artificial neuron sums up the n heavy inputs and passes the result through the nonlinear function $f(x)$ in accordance with the equation[3]

$$y_j = f\left[\sum_{i=1}^{n} \omega_{ij} x_i + \theta_j\right] \tag{1}$$

where θ_j is a threshold or external offset, x_i is the input i, w_{ij} the corresponding weights and y_j represent the output. Figure 1 shows a simplified outline of a basic artificial neuron.

The neurons usually operate in parallel and they are configured in regular arrangements. They are usually organized in layers and two types of feedback connections are allowed: within the layer and toward adjacent layers. Figure 2 shows the general structure of a three-layer network as the one used in this work. The network is feedforward propagation and consists of an input layer, a hidden layer and an output layer. The number of nodes of the input layer is generally similar to the number of input variables (6 in this case), and the same happens with the output layer (3 nodes). The number of nodes of the hidden layer is one of the most important considerations to take into account when solving problems using multilayer networks of direct propagation. In general, more variables and more examples make necessary more nodes; however, an excessive number of nodes could generate an over-trained network that adjusts all the examples perfectly but it doesn't supply accurate results.

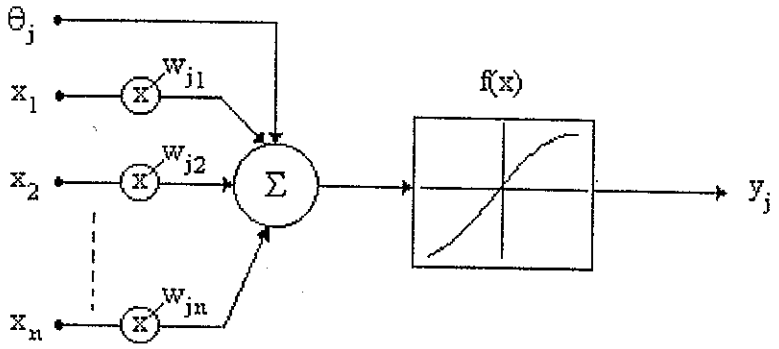

FIG. 1. — Simplified functional model of an artificial basic neuron cell.

Training the network to learn consists of presenting it with a set of correlated inputs and outputs, called examples. The system learns by adjusting the weights w_{ij} of the node connections, in such a way to minimize the difference between the calculated (y_k) and the given (d_k) output. This means minimizing the mean square error given by Equation (2):

$$E = \sum_{l-1}^{r} \sum_{k=1}^{m} [d_k^l - y_k^l]^2 \tag{2}$$

The most extended method of minimizing the error function is the back-propagation algorithm, which is a generalization of the steepest method. The adjustment of each individual weight is as follows[3]

$$\Delta\omega_{ij} = -\eta \frac{\partial E}{\partial \omega_{ij}} \tag{3}$$

where η is the learning constant which influences the convergence speed and the effectiveness of the learning process. In general, the optimum value of η depends on the problem that is being analyzed, and only small values of η guarantee stable solutions.

NETWORK OPTIMIZATION

The learning parameter η should be chosen small to provide minimization of the total error function E. However, for a small η the learning process becomes very slow. On the other hand,

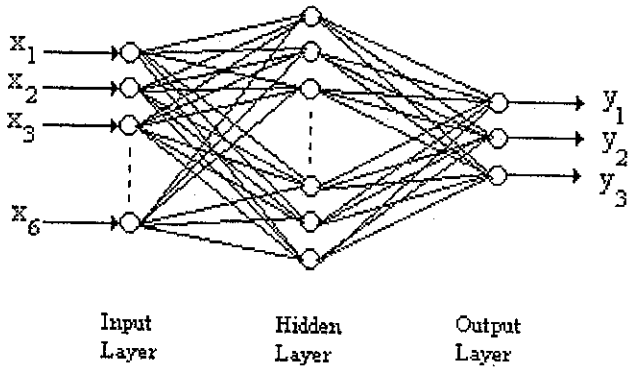

FIG. 2. — The general structure of a three-layer back-propagation network.

TABLE II

PARAMETERS FOR ANNs OPTIMUM PERFORMANCE

α	n	η	ε
.9	12	.035	.0015

large values of η correspond to fast learning, but lead to parasitic oscillations which prevent the algorithm from converging to the desired solution. Moreover, if the error function contains many local minima, the network might get trapped in some local minimum, or get stuck on a very flat plateau. One simple way to improve the standard back-propagation learning algorithm is to smooth the weight changes by over-relaxation,[3] i.e., by adding the momentum term (α) that increase the rate learning without magnifying the parasitic oscillations.

The network training finishes when all the errors E_j are below a previously established error, ε. Like happens with the quantity of hidden nodes, too small of errors can fit the data but could give a wrong generalization.

Regrettably there's no way to know previously which are the parameters that optimize the network for a given problem, and these should be determined by trial and error. Also, since the initial weights are assigned at random, the solutions are not always the same. A right selection of the parameters will minimize the difference among the successive runs. These approaches and a prediction error smaller than 10%, were the boundary conditions imposed to the parameters. A systematic trial and error test was carried out and the optimum parameters were determined. Table II shows the results.

RESULTS AND DISCUSSION

PROPERTIES OF NEW COMPOUNDS

Once the optimum parameters were determined the network is ready to be used. In order to ensure that the network predicts correctly the rheometric properties of the new compounds, the 24 compounds were divided into two subsets. One for training (training set) and the other for test (test set). The network training was carried out with the data of the first subset. During the training process, the neurons learn the relationship between the output variables and the input ones by adjusting the connection weights. After the ANN has been trained, it can be used to predict the compounds properties of the test set. Table III shows the results obtained from 19 elements for the training set and 5 for the test set.

TABLE III

CALCULATED VS REAL VALUES FOR FIVE COMPOUNDS OF THE TEST SET

Real values			Calculated values			Relative errors, %			
M_L	M_H	$t\,50$	M_L	M_H	$t\,50$	M_L	M_H	$t\,50$	Averages
7.788	9.616	6.563	6.168	9.618	6.584	20.804	0.010	0.365	7.060
5.103	5.383	6.687	4.278	4.806	7.595	16.140	10.720	13.696	13.519
3.732	3.466	9.175	3.234	3.544	8.746	13.364	2.246	4.732	6.781
7.821	8.303	8.514	5.904	5.880	10.593	24.517	29.176	24.421	26.038
6.516	5.698	7.986	6.810	5.752	7.617	4.501	0.939	4.590	3.344
					Averages	15.865	8.618	9.561	11.348

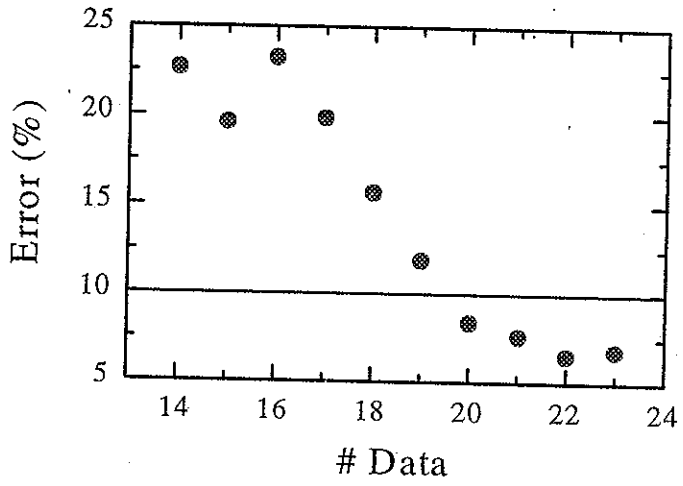

FIG. 3. — Average relative error as a function of the number of data in the training group.

The greater and more representative the training set, the smaller will be the error for the new compounds properties calculation. This fact can be verified by passing elements from the test group to the training one, re-training the network and re-calculating the values for the new test group. Figure 3 shows the dependence of the average relative error as a function of the number of data in the training group.

EXPERIMENTAL DESIGN

There are several works in the literature about the compound optimization using experimental design.[6–9] This kind of method, although it has been used with relative success, has shown some important limitations. (a) It is expensive and not always possible to implement all conditions that experimental design requires. (b) The number of experiences to carry out increases as d^N, where

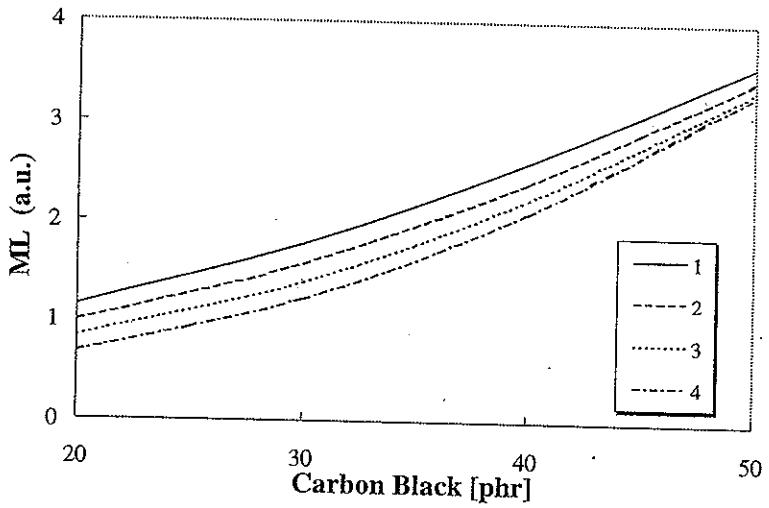

FIG. 4(a). — Low torque as a function of carbon black load for different amounts of sulfur (100% NR, oil 20, Acc 1, PA 0.2).

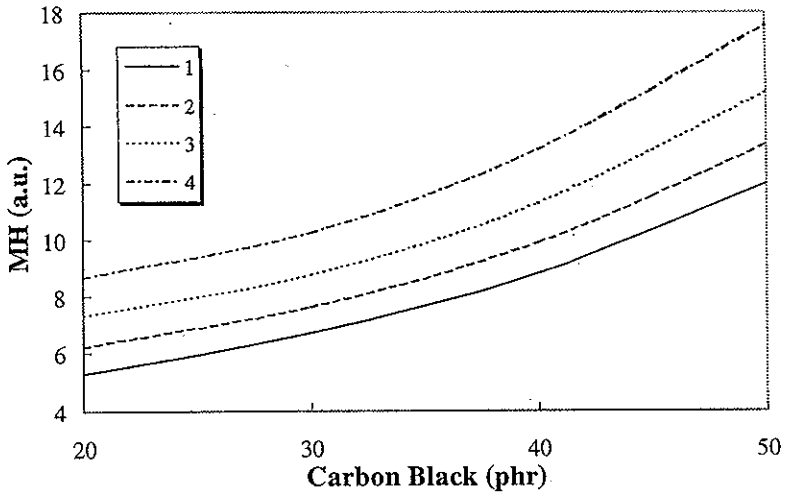

FIG. 4(b). — High torque as a function of carbon black load for different amounts of sulfur (100% NR, oil 20, Acc 1, PA 0.2).

N is the number of variables and d the quantity of values that each variable can take. And (c), for designs with a reasonable quantity of experiences, only lineal or quadratic effects can be seen.

The use of neural networks can attenuate many of these limitations. It is possible to use a well-trained network to generate, by numerical simulation, any experimental design independently of the required number of experiences. Moreover, it doesn't matter if a given point of the experiment design was measured or not, the network will predict the value from the available information.

In this way, for instance, we can analyze the dependence of M_L as a function of carbon black load, for different quantities of sulfur. Figure 4(a) shows the obtained results.

The dependence of the low torque (M_L) with carbon black load obeys to the generalization made by Smallwood[10] and other authors from the Einstein viscosity equation[11]; in this way we

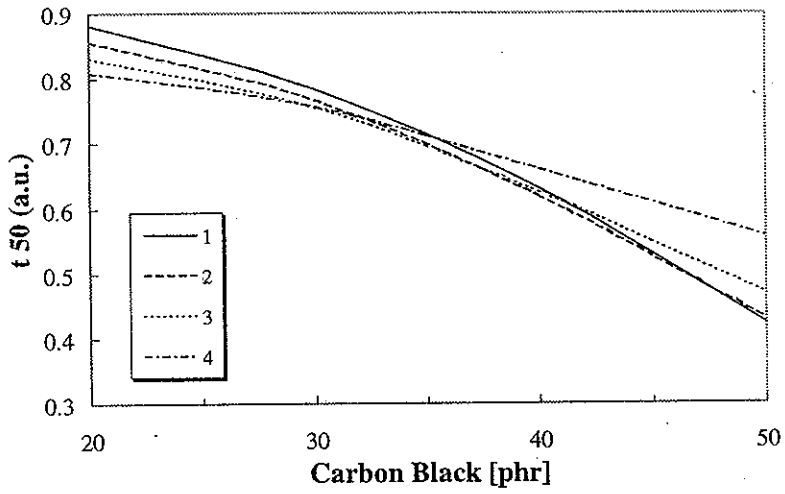

FIG. 4(c). — Time to 50% cure ($t50$) as a function of a carbon black load for different amounts of sulfur (100% NR, oil 20, Acc 1, PA 0.2).

have that

$$M_L = M_{L,\,\text{Rubber}}(1 + a_1 c + a_2 c^2 + \cdots) \tag{4}$$

where c is the filler volume fraction. Note that M_L decreases as sulfur increases, this fact probably responds to the effect of "hide variables" that were not included in the model.

Figure 4(b) shows the known dependence of the M_H as function of CB load. It can be seen that, for a given CB load, M_H increases with the sulfur content. This is a consequence of cross-links density increment. Finally, Figure 4(c) shows how the cure time to 50% (t 50) diminishes appreciably with CB load independently of sulfur quantity (at least for small quantities) in good agreement with well known theoretical results.[12,13]

SENSIBILITY TO INPUT VARIATIONS

One interesting and useful analysis we can do with the ANN is to calculate the sensibility of a given compound to the weight variation of some of its components. We define this sensibility as

$$\xi_{ij}^{*} = \left.\frac{\partial y_i}{\partial x_j}\right|_{x_k = cte(k \neq j)} \tag{5}$$

where y_i is the desired output variable and x_j is the input one, ξ_{ij} is then specialized in the x_i of the compound that is being analyzed.

Multiplying by x_i and dividing by y_i, we normalize ξ_{ij} obtaining in this way a relative quantity that can be compared with any other, independently of the x_i and y_i values.

$$\xi_{ij} = \left.\frac{x_j}{y_i}\,\frac{\partial y_i}{\partial x_j}\right|_{x_k = cte(k \neq j)} \tag{6}$$

Generalizing, we can calculate a matrix ξ whose elements correspond to the sensibility of each output variable to each input one. Specializing ξ in the x_i of a given compound, the sensibility matrix that characterizes that compound will be obtained. In this way, weight tolerances could be defined taking into account the most sensitive variables.

Each one of the ξ_{ij} can be calculated taking small variations around x_j and entering these data to the network. As a result we will obtain the corresponding y_i values. From these values it is easy to calculate the ξ_{ij}. Table IV shows the sensibility matrix for a typical compound.

Multiplying the matrix elements by the relative weight tolerance of each component and adding each line, we obtain the expected variability, due to weight errors, for each compound rheometric property. Table V shows the comparison of the values obtained in this way with the variability observed experimentally. The calculated values are slightly inferior because they don't take into account variability of either the process or the raw material. These values set an inferior limit to the expected variability for this compound.

TABLE IV
SENSIBILITY MATRIX FOR A TYPICAL COMPOUND

ξ_{ij}	NR	CB	Oil	Sulfur	Acc	PA
M_L	0.03	2.14	−0.2	−0.08	0.12	0.02
M_H	−0.27	1.46	−0.25	0.40	0.15	−0.02
t 50	−0.25	−0.83	0.03	0.05	0.01	0.04

TABLE V

CALCULATED VS REAL VALUES FOR THE VARIABILITY OF A TYPICAL COMPOUND

Δ, %	Calculated[a]	Real[a]
M_L	4.74	7.23
M_H	4.01	4.98
t 50	2.04	2.63

[a] Values are normalized in the range 0–10.

CONCLUSIONS

The implementation of an artificial neural network to predict the rheometric properties of a compound from their formulation was developed in this work. The calculated values are in very good agreement with those obtained experimentally. It was shown that besides the prediction of the rheometric properties, the network can be used to analyze the dependence and sensibility of the different properties with the parameters of the formulation.

Even with the limitations imposed in this first approach to the problem, the obtained results are encouraging. The network may be still improved, optimizing the learning algorithms and developing a validation algorithm of the input data. The natural extension of this treatment will be to predict the physical properties from the formulation of the compound. Finally, it is necessary to do a more in depth analysis that includes information about the mixing process and allows separating the different types of rubbers, carbon blacks, etc. that have been treated here without distinction.

ACKNOWLEDGMENTS

I want to thank to L. De Bernardez for many contributions as well as helpful discussions.

REFERENCES

[1] G. Kraus, "Reinforcement of Elastomers," John Wiley & Sons, New York, 1965.

[2] "Goodyear Tech Book for the Rubber Process Industries," Goodyear Tire & Rubber Company, Akron, Ohio, 1980.

[3] A. Cichocki and R. Unbehauen, "Neural Networks for Optimization and Signal Processing," John Wiley & Sons, New York, 1996.

[4] J. Hertz, A. Krogh, and R. G. Palmer, "Introduction to the Theory of Neural Computation," Addison Wesley, Redwood City, CA, 1990.

[5] H. Ni and D. Hunkeler, *Polymer* **38,** 667 (1997).

[6] F. E. Welsh and R. J. Emerson, *Tire Technol. Int.* 96 (1997).

[7] I. A. Amraee and A. R. Teimorian, *Tire Technol. Int.* 91 (1997).

[8] G. C. Derringer, RUBBER CHEM. TECHNOL. **61,** 377 (1988).

[9] G. E. P. Box, W. G. Hunter, and J. S. Hunter, "Statistics for Experimenters," John Wiley & Sons, New York, 1978.

[10] H. Smallwood, *J. Appl. Phys.* **15,** 758 (1944); RUBBER CHEM. TECH. **21,** 667 (1948).

[11] A. Einstein, *Ann. Physik* **19,** 289 (1906); **34,** 1591 (1911).

[12] M. L. Studebaker, RUBBER CHEM. TECHNOL. **30,** 1400 (1957).

[13] M. L. Studebaker and L. G. Nabors, *Rubber Age* **80,** 837 (1957).

[*Received May 1, 1999; revised May 9, 2000*]