Applying Soft Computing Methods to Fluorescence Modeling of the Polydimethylsiloxane/Silica Composites Containing Lanthanum

Silvia Curteanu,[1] Alexandra Nistor,[1,2] Neculai Curteanu,[3] Anton Airinei,[2] Maria Cazacu[2]

[1]*Department of Chemical Engineering, "Gh. Asachi" Technical University, Iasi, Romania*
[2]*"Petru Poni" Institute of Macromolecular Chemistry, Iasi, Romania*
[3]*Institute for Computer Science, Romanian Academy, Iasi, Romania*

Received 3 October 2009; accepted 26 January 2010
DOI 10.1002/app.32167
Published online 3 May 2010 in Wiley InterScience (www.interscience.wiley.com).

ABSTRACT: Lanthanum, as a fluorescent element, was encapsulated in complexing silica xerogel by using a one-pot procedure: silica formation from tetraethyl-orthosilicate and 3-aminopropyltriethoxysilane, its chemical modification by reacting with 2,4-dihydroxybenzaldehyde, and lanthanum complexation through the formed hydroxy-azomethine groups occurred in the same reaction system. A polydimethylsiloxane-α,ω-diol with average number molecular weight of 20,000 was added in the sol–gel system to facilitate obtaining the flexible, free-standing films. Using different molar ratios among the reactants, a series of experiments was performed and the fluorescence of the resulted compounds was evaluated, since the photophysical properties of the prepared compounds strongly depend on the composition. The acquired data have been used for fluorescence modeling based on soft computing instruments. Neural network models, individually or aggregated in stacks, were developed to estimate the fluorescence intensity depending on the reaction mixture composition: tetraethyl-orthosilicate, 3-aminopropyltriethoxysilane, 2,4-dihydroxybenzaldehyde, lanthanum acetate, and polydimethylsiloxane-α,ω-diol amounts. A procedure based on genetic algorithms was used to design simple neural networks. Two main goals are pointed out in this article: developing a general methodology for modeling the complex processes with simple or stacked neural networks, and demonstrating the improvement of the modeling performance by combining different neural networks. © 2010 Wiley Periodicals, Inc. J Appl Polym Sci 117: 3160–3169, 2010

Key words: composites; stacked neural networks; polysiloxanes; complexed lanthanum ions; fluorescence modeling

INTRODUCTION

Because of their high paramagnetism and fluorescence, the rare earth-based compounds are of high interest for many applications, such as catalysis, optoelectronic, or life sciences.[1] To use them in certain domains, these compounds are encapsulated in a matrix, usually an organic polymer. The existence of wide ranges of rare earth ions and host polymers in which these compounds can be incorporated allows to obtain products having absorption and emission bands covering the selected ranges of the electromagnetic spectrum.[2,3] However, a quenching effect occurs when a conventional doping procedure is used. This is due to the fact that only weak interactions (such as hydrogen bonding, van der Waals force, or weak static effect) do exist between polymer and inorganic parts, while the surrounding hydroxyl groups have high-energy vibration. A possibility to change this ratio of forces consists in the complexation of the rare earth ions by ligands that are covalently bound to the polymeric matrix.[4] Silica is commonly used as a rigid matrix for ligand immobilization. Silica xerogels doped with rare earths are of great interest for technological applications in the field of optical devices.[5] The immobilization of chelating groups on the silica gel surface can occur either by chemical modification of the silica or through simple physical adsorption processes.[4]

The sol–gel process based on hydrolysis/polycondensation reactions of the proper silicon alkoxydes is a convenient and versatile method for preparing transparent optical materials in mild conditions.[2,3] Furthermore, by changing the sol–gel reaction conditions, the control of the microstructure is allowed for the external shape or homogeneity. The optical properties are strongly modified when crystallization of the silica network or clustering of the rare earth-based compounds occurs.[5]

In a previous article,[6] complexed La as fluorescence center has been encapsulated in an inorganic matrix either as precomplexed form,[7] or by in situ

Correspondence to: S. Curteanu (silvia_curteanu@yahoo.com).

Contract grant sponsor: Romanian PN II type Grant; contract grant number: PC 71006/2007.

complexation[8] in a silica generating sol–gel system. In the last case, the complexing groups were attached to the silica network interpenetrated or not with a polydimethylsiloxane network. The fluorescence study revealed that, although the complexed La as such is fluorescent, the intensity of this effect depends in a complex manner on the matrix type in which it is incorporated.

Artificial neural networks (ANNs) have been increasingly used in the chemical process industry, especially for dealing with some complex nonlinear processes whose understanding is limited.[9] Polymerization processes are typical examples for this type of applications.

In many cases, neural networks used as individual models do not lead to satisfactory results. One way to improve the performance of such models is represented by grouping them into stacks, homogeneous or heterogeneous, based on the idea that individual models can capture different aspects of the process and the stack can gather them.

Stacked generalization was proposed as a method of using multiple models to provide improved accuracy or confidence intervals.[10] Wolpert[10] suggests the use of a number of ANNs, on different subsets of the data, to obtain models that are slightly different. These models are then recalled over the remainder of the data and this information is used then in the stacking process.

Several significant issues related to stacked neural networks and some examples of using these types of models, selected from the polymerization process field, are presented in the introduction of this article.

A series of articles belonging to Zhang et al.[11–16] are representative for the use of stacked neural networks in the polymerization process modeling. For instance, neural networks aggregated in stacks were used in two approaches[11]: (1) a model of the polymerization process is developed for the prediction of the initial reaction conditions, and the amount of impurities and reactor wall fouling are then calculated by comparing the predicted values with the nominal initial conditions; (2) a neural model is constructed for the dynamic behavior of the polymerization process and the predicted trajectories are then compared with the online measurements of conversion and coolant temperatures.

A different technique for aggregating multiple models using bootstrap resamples of the training data is proposed by Breiman and is known as Baggings.[12] A bootstrap resampled data set represents a subspace of the original training data set. Neural networks trained on different bootstrap resampled data sets would be more dissimilar than those trained on the same training data. When trained a bootstrap resampled data set, different neural networks will perform differently within distinct regions of the input space. Although these neural networks are correlated since they intend to model the same relationship, the independent elements among these models can be discovered through principal component analysis (PCR).[13] It is proposed in Ref. 13 that the appropriate weights for combining individual neural networks to be obtained through PCR. Bootstrap-aggregated neural networks are proposed to build inferential estimators for online evaluation of polymer quality. A technique for obtaining robust nonlinear models by aggregating multiple neural networks (data are resampled using bootstrap procedure to form several different pairs of training and testing data sets) was successfully applied to the building of software sensors for a batch polymerization reactor.[14] Additional examples of polymerization processes modeled with neural networks can be found in others approaches of Zhang et al.[15,16] Polymer melt index were successfully estimated using an aggregated neural network.[15] A batch-to-batch optimal control strategy based on the linearization of stacked neural network model is proposed in Ref. 16.

Unlike the "bagging" and "boosting" approaches, which only combine classifiers of the same type, the stacking approach can combine several different types of classifiers through a metaclassifier to maximize the generalization accuracy.[17]

An important article for the use of stacked neural networks in modeling the polymerization processes belongs to Tian et al.[18] Stacked recurrent neural networks are built to characterize the gel effect, which is one of the most difficult parts of free radical polymerization modeling. Comparative studies with the use of a single neural network show that stacked networks provide superior performance and improved robustness.

Different methods of combining neural networks, particularly their outputs, have been developed: minimizing the squared prediction error using cross validation,[10] imposing proper constraints on the weights,[19] using the confidence of an individual model as the weight for that model,[20] applying the technique of principal component regression[21] or weighted averaging of the individual networks outputs.[18]

Our group has a number of achievements in the field of modeling and optimization of polymerization processes using the tools of artificial intelligence. Individual neural networks were used to model different parameters such as conversion, polymer molecular weight, mass reaction viscosity, polydispersity index,[22–25] copolymer composition,[26] polymerization yield, swelling degree,[27,28] or the kinetics of the polymerization reaction.[29] Alternatively, stacked neural networks represented a better solution for the modeling of complex nonlinear processes with an insufficient number of experimental

data.[30,31] The influence of the reaction conditions (functional group content in silica, $CuCl_2$ solution concentration, pH, and time) upon the absorbance was evaluated based on simple neural network types (multilayer perceptron, generalized feed-forward neural networks, and Jordan Elman networks).[31] Another variant of modeling applied in Ref. 30 is based on using individual neural networks for each set of experimental data, corresponding to each sample, then aggregating them into a stack. It is a particular type of combination where the output of the stack if composed from the outputs of the individual neural networks. It was proved that stacked neural networks designed as a software assembly represent models with an accurate capacity of generalization.

In this article, we encapsulated a rare earth, lanthanum, in silica by in situ complexation with chelating groups attached to the silica, according to the procedure already described in Ref. 6. Hydroxy-azomethine groups, known to form readily complexes with transition metal ions, are generated with this aim in the same system. Polydimethylsiloxane-α,ω-diol (PDMS) was added in different proportions in the sol–gel generating silica system to induce flexibility to the prepared materials. Because of the high complexity of the process, the influence of the reaction mixture composition on the fluorescence intensity was modeled by using the artificial intelligence instruments based on the acquired experimental data series. Neural networks are developed using a procedure based on genetic algorithms to obtain the optimal structure of the model. Then, simple neural networks are combined in stacks, proving the effectiveness of these models compared with the neural networks used individually.

MATERIALS AND METHODS

Materials

Lanthanum acetate $La(CH_3COO)_3$ ($LaAc_3$), purchased from Merck was used as received, $M = 316.06$.

Tetraethyl-orthosilicate (TEOS) purchased from Fluka (purity > 98%, b.p. = 163–167°C, $d_4^{20} = 0.933$), was used as received.

3-Aminopropyltriethoxysilane (APT) (Fluka) $M = 221.37$, b.p. = 213–216, $d_4^{20} = 0.949$.

2,4-Dihydroxybenzaldehyde (AR) was prepared and purified according to the procedure described in literature[32] (yield: 33%, m.p. 135–137°C).

A polydimethylsiloxane-α,ω-diol (PDMS), having an average number molecular mass of about 20,000 (determined by GPC), has been synthesized by equilibrium cationic ring-opening polymerization of octamethylcyclotetrasiloxane (D_4) in the presence of Purolite CT 175 as a catalyst and water as a chain transfer agent.[33]

Dibutyltindilaurate (DBTDL) was received from Merck-Schuchardt, $d_4^{20} = 1.055$ and used as received.

Solvent: the tetrahydrofuran supplied by Fluka was used as such.

Methods

Equipments

Fourier transform infrared (FTIR) spectra were obtained on a Bruker Vertex 70 FTIR analyzer. Depending on the state of the sample, the analyses were performed either in transmittance mode in KBr pellets or on the films in reflectance mode (ATR).

The luminescence spectra were recorded at room temperature with a Perkin Elmer LS55 spectrometer with a solid sample holder. Emission spectra were obtained at an excitation wavelength of 365 nm. The excitation and emission slits were set at 2.5 and 12 nm, respectively.

Procedure

The preparation procedure is reported in our previous article.[6] PDMS, TEOS, and APT were mixed together in the amounts presented in Table I. A homogeneous, colorless transparent mixture is formed. The corresponding amount of AR solved in THF for 0.5 mol/mL was added as silica modifier and the stirring continued 1 h at room temperature to provide the hydroxyl-azomethine complexing groups. $La(CH_3COO)_3$ solved in water for 0.5 mol/L acidified with a drop CH_3COOH was added as La supplier to form silica containing complexed La. The mixing continued 1 h at room temperature. Then, DBTDL, as a catalyst, was added and, after stirring ½ h, the mixture was poured on Teflon foil and left to evaporate solvent. The films were maintained at room temperature for 48 h and another 24 h in vacuum at 50°C.

TEOS was used as precursor for silica. APT was added to provide aminopropyl functions on the silica. This function was reacted with 2,4-dihydroxybenzaldehyde resulting in 2,4-dihydroxyazomethine moiety. The azomethine and OH groups placed in favorable positions close chelate ring with La ions added in the reaction mixture as $LaAc_3$. To obtain silica having complexed La as a flexible free-standing film, PDMS having a molecular weight of about 20,000 was added to the sol–gel-complexing system in the final step of the reaction. In the presence of water (both resulted from the azomethine formation

TABLE I
PDMS/Silica Composites Doped with In Situ Complexed La: Reactant Amounts and Fluorescence Data

Code	TEOS (mmol)	APT (mmol)	AR (mmol)	$LaAc_3$ (mmol)	PDMS (mmol)	λ_{em}	I_{em}
RL1	22	0.64	0.64	0.25	0.01	459;515sh	9.4
RL2	22	0.64	0.64	0.50	0.0125	449;518sh	1363
RL3	22	0.64	0.64	0.75	0.0250	457;519sh	1043
RL4	22	0.64	0.64	0.125	0.050	460	17
RL5	22	0.64	0.64	0.25	0.075	464	5.5
RL6	22	0.64	0.64	0.25	0.10	459	23.6
RL7	11	0.64	0.64	0.25	0.10	460;539sh	11
RL8	44	0.64	0.64	0.25	0.10	435;564	11
RL9	66	0.64	0.64	0.25	0.10	444;550	1828
RL10	22	1.28	1.28	0.50	0.01	452;520sh	705
RL11	22	1.28	1.28	0.50	0.025	465	6
RL12	22	1.28	1.28	0.50	0.075	468	7.4
RL13	22	1.28	1.28	0.50	–	451;516sh	1073
L1	22	0.64	–	–	–	442	2193
L2	22	0.64	0.64	–	–	454;520sh	1233
L3	22	0.64	0.64	0.25	–	455;528	1215
L4	22	0.64	0.64	0.25	0.25	460	13.3
L5	22	0.64	0.64	0.50	0.25	466	5.8
L6	22	0.64	0.64	0.75	0.25	454	19.2
L7	22	0.64	0.64	0.125	0.25	458	19.2
L8	22	0.64	0.64	0.25	0.50	465	26.7
L9	22	0.64	0.64	0.25	0.75	462	20
L10	22	0.64	0.64	0.25	0.125	468	53.3
L11	11	0.64	0.64	0.25	0.25	457	10
L12	44	0.64	0.64	0.25	0.25	460	16.7
L13	66	0.64	0.64	0.25	0.25	466	11.7

reaction and added with $LaAc_3$) and DBTDL, the hydrolysis followed by condensation of the alkoxysilanes (TEOS and APT) occur. Concomitantly, Si—OH groups formed by hydrolysis of silanes react with that belonging to polydimethylsiloxane-α,ω-diol. Thus, a polydimethylsiloxane network is formed, which is connected to the silica network. The possible reaction sequences are showed in Scheme 1.

The formation of this complex structure was proved by FTIR spectra where characteristic bands are found for:

- azomethine-La complex: 1644 cm^{-1} (C=N), 1613 cm^{-1} (phenyl cycle);
- silica: the large band between 1220 and 1019 cm^{-1} enveloping from the most crosslinked to linear Si—O—Si bands, as well as the sharp band at 460 cm^{-1} assigned to the silica Si—O—Si bending vibration.
- polydimethylsiloxane: 2965 cm^{-1} (C—H from $Si—CH_3$), 1264 cm^{-1} (δ $Si—CH_3$ sym.), 798 cm^{-1} (CH_3 rocking asym. Si—C), 865 cm^{-1} assigned to Si—O—Si bond formed by condensation between silica-Si—OH and PDMS-Si—OH groups.[34]

To find the conditions in which a maximum fluorescence can be reached, a series of experiments have been performed in different conditions. The resulted materials were investigated by fluorescence spectrophotometry to evaluate values of the wavelength and emission spectra intensity, I_{em}. The acquired data concerning to the reaction conditions and fluorescence intensity of the obtained materials (Table I) were used for modeling, based on soft computing instruments (neural networks and genetic algorithms).

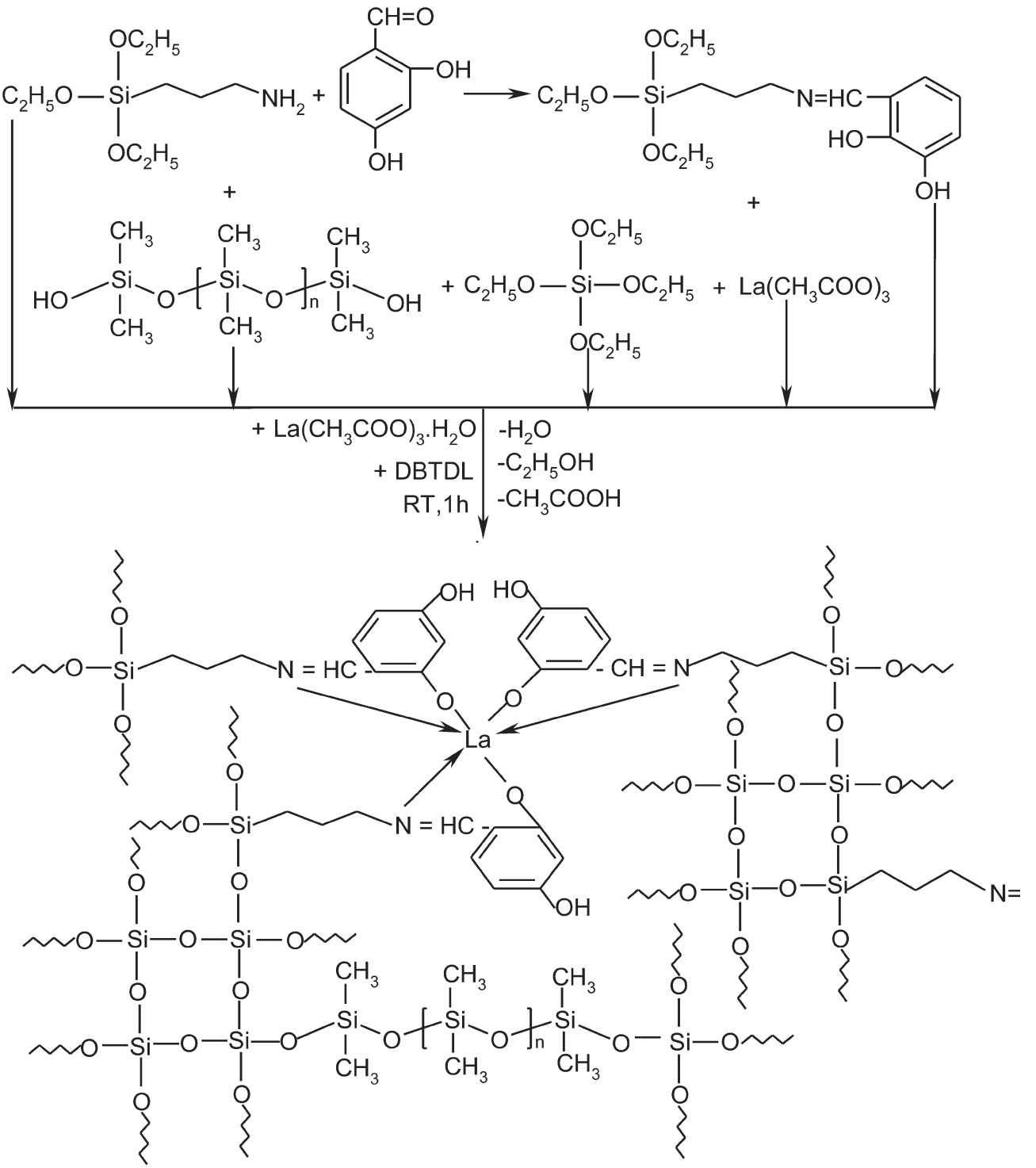
The investigations were performed after the samples were kept in the laboratory environment for about 2 months. The tests revealed that, after this time, the mass of the samples is stabilized.

Neural network modeling

Neural networks have gained significant popularity because of their capacity to create nonlinear mappings between input and output variables and to the possibility of applying them to various processes.

Two elements are very important for the application of a modeling method based on neural networks[35]: (1) Understanding the theory underlying the modeling method, as well as knowledge of advanced methods for improving the modeling results. (2) Understanding the process of applying the method to real data, including model testing and result interpretation.

In general, there is no assurance that any individual neural model has extracted all the relevant information from the data set. The idea of combining



Scheme 1 The possible reactions sequence.

neural network models is based on the premise that different neural networks can capture different aspects of the process behavior and that aggregating this information should reduce uncertainty and provide more accurate predictions. It is always possible that a stacked network could outperform a single best-trained network for the following reasons: (1) Optimal network architecture cannot always be guaranteed. (2) The optimization of the network parameters is a problem with many local minima. Even for a given architecture, final network parameters can differ between one run of the algorithm and another one. (3) Different activation functions and learning algorithms can also lead to different generalization characteristics, and no one activation function or learning algorithm is the best in all cases. (4)

Convergence criteria used for network training can lead to very different solutions, for a given network architecture.[18]

Different types of modeling are applied in this article, based on individual and stacked neural networks, to find the best model for the approached process.

The experimental data from Table I were used for developing and testing a model, which evaluates the emission spectra intensity as function of reaction conditions. TEOS, APT, AR, $LaAc_3$, and PDMS amounts are considered as input parameters. The products, formed either as film or powder, were investigated by emission spectroscopy. The values of the emission spectra intensity, I_{em}, were considered as output parameters (Table I).

Briefly, neural networks are computational architectures that combine simple units in an arrangement that can then exhibit complex behavior. One of the most familiar architectures is the feed-forward multilayer neural network or feed-forward multilayer perceptron (MLP).[35] The number of input neurons (which constitute the input layer) is equal to the number of input variables, 5 for our case study, and the output layer contains a single neuron, corresponding to the single output variable. The number of intermediate layers and the number of neurons in them are to be determined. Consequently, the modeling goal is the development of a network $5 : x : 1$ or $5 : x : y : 1$ (with one or two hidden layers) which should satisfy the imposed performance criteria.

The experimental data available in Table I are split into training and validation data sets (15% validation data) because it is more important to evaluate the performance of the neural network on unseen data than on training data. In this way, we can appreciate the most important feature of a neural model—the generalization capability.

The most common practice for developing the neural network architecture is the repetitive trial and error method, which is done by testing several topologies and comparing the prediction errors. Smaller errors indicate potentially good topologies, i.e., neural network topologies with chance to train well and to output good results. Generally, this method is time-consuming and produces uncertain results because it is based on the human expert's past experience and intuition.

There is yet another class of methods used for choosing the network topology, known as evolutionary methods. One of the most popular methods within this class is the genetic algorithm (GA)—a widely used optimization method inspired from evolution theory, in which the fittest species survive and propagate, while the less successful tend to disappear.

In a previous article,[36] we proposed a genetic algorithm based method for detecting the optimal topology for a neural network that should approximate as well as possible the test data. The representation of solutions in chromosomes must simultaneously takes into account two problems: including the information on network topology (number of hidden layers, number of neurons in these layers), and including actually the connection weights and biases of the neurons, with the purpose of verifying the network training errors. All this information is encoded by real numbers, this is why we use the real encoding for the chromosome genes.

Applying here this method, a neural network with two hidden layers, having 24 and 8 neurons, respectively, in these intermediate layers, MLP($5 : 24 : 8 : 1$), is obtained based on the mean squared error (MSE) minimization. The performances registered in the training phase for this network were: MSE = 0.00106 and r (correlation between training data and neural model predictions) = 0.9889.

The chosen representation has both advantages and disadvantages[36]:

- Advantage: the simplicity of the approach, as the genetic algorithm also accomplishes the finding of the optimum topology and the training of the neural network (determining the connection weights that allow approximating the test data).
- Disadvantage: a long training time because of the large number of chromosome genes (both information regarding the topology, and the connection weights and the biases of the neurons).

The activation function was a sigmoid bipolar and the fitness function is equivalent in the present approach to calculating the MSE for the test problem for the neural network represented by a certain chromosome. The mean squared error was computed using the following formula:

$$\mathrm{MSE} = \frac{\sum_{j=1}^{P}\sum_{i=1}^{N}(d_{ij} - y_{ij})^2}{N \cdot P} \quad (1)$$

where P is the number of output processing elements (in this case, $P = 1$), N is the number of samples in the data set, y_{ij} is the network output for exemplar i at processing element j, and d_{ij} is the desired output for exemplar i at processing element j.

A possibility to improve the performance of the model is represented by the use of stacked neural networks. Combining multiple neural networks can generate more accurate predictions than using any one of the individual networks alone.

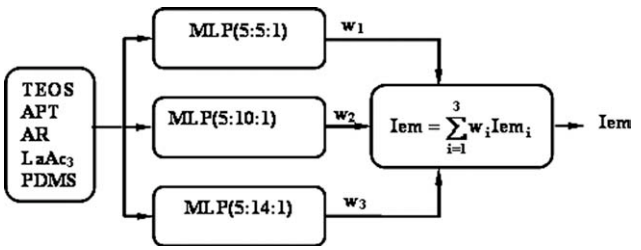


Figure 1 Stacked neural networks.

Applying the method of successive attempts, different types of feed-forward networks have been generated, varying the number of hidden layers and the number of neurons in these layers. The selected networks, in number of 3, did not record very good performances, but they benefited from the simplicity of topology. MLP (5 : 5 : 1), MLP (5 : 10 : 1), and MLP (5 : 14 : 1), with MSE around 0.006 and r approximately 0.970, have been chosen to form a stack.

The method used to combine the parallel models was the weighted summation of the individual outputs. Consequently, the performance of the stack is influenced by the aggregated individual models and their corresponding weights. Figure 1 presents the stacked neural networks for our case study.

In this article, we compare the predictions of individual and stacked neural networks, especially in the validation phase, to select a neural model with a high generalization capability.

Individual networks have been developed and tested in the framework offered by the specialized software NeuroSolutions. Subsequently, for designing the stacked neural networks, an original program was written in C#, which imports the individual models from NeuroSolutions, assemblies them in the stack by summing the outputs weightily, allows changing the weights of the single networks, and, in the same time, performs the training and the validation of the stack model. In this way, the outcomes of the program are the predictions on the training and validation data, along with the performance recorded in these phases.

RESULTS AND DISCUSSION

A route consisting in one-pot process by heating all components together was applied to prepare a series of polydimethylsiloxane/silica composites having complexed lanthanum. The reaction conditions chosen for different experiments were given in Table I as well as optical properties of the synthesized compounds.

The emission spectra of the PDMS/silica composites reveal a broad band located at 450–460 nm and a shoulder at about 520 nm with lower intensity. Different from other siliceous organic/inorganic hybrids, the position of the emission practically does not depend on the excitation wavelength in our case.[37] As a rule, it was found that the emission intensity strongly decreases when the PDMS content in composites increases. Thus, regardless the composition of the network, a PDMS content between 0.1 and 0.75 mol leads to quenching the fluorescence in the system (Table I). The increase of TEOS content (RL9) results in the intense emission, even if the PDMS content is higher. At higher contents of TEOS, a more compact network can be formed and the coordination capability of lanthanum increases, leading to a very enhanced emission. The sample RL13, which does not contain PDMS, has a higher emission by comparison with sample RL10, the both samples having the same spectral pattern. The increase of the content of APT and AR leads to the decrease of the emission intensity and, for high amounts of PDMS, the fluorescence is quenched. When the concentration of lanthanum is high enough (RL2, RL3), an increasing fluorescence intensity was observed at room temperature. This enhanced luminescence efficiency can be ascribed to the greater lanthanum coordination in the network, which determines the ligand rigidity and decreases the nonradiative energy loss of the excited state emission.

The fluorescence of the synthesized compounds was correlated with the reaction conditions by using a neural network methodology. Two issues were pursued through this modeling: (1) to model the actual process with errors as small as possible, such that the predictions to be used for inexperienced reaction conditions; (2) to develop a general modeling methodology based on neural networks, which includes opportunities for improvement through the stack combinations and which can be applied to other complex, nonlinear processes.

Individual and stacked neural networks were applied to the training and validation data sets to compare their performance and, finally, to choose the most appropriate model for the studied process.

Model estimation is performed in the training phase when the network is presented with a set of input–output pairs (desired output) and, for each input, the network calculates its output (actual output). The minimization of MSE means that the actual output to be as close as possible to desired output. This could be done by varying the network parameters, i.e., weights and biases. In this work we applied the well-known backpropagation algorithm where the minimization of MSE is done by gradient descent method and the weights are updated using the generalized delta rule.

The relative error, calculated as:

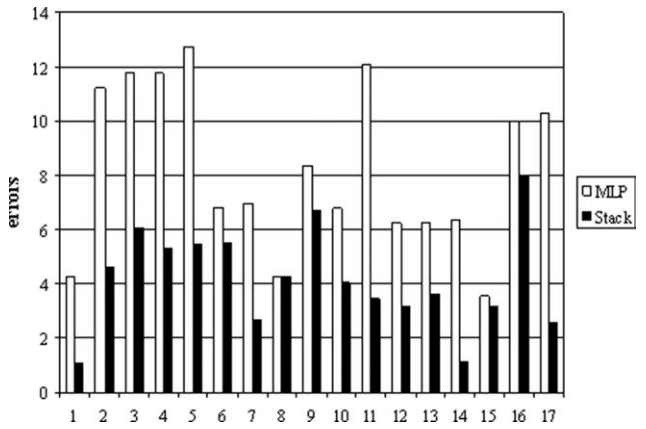


Figure 2 Errors of the individual and stacked neural network in the training phase.

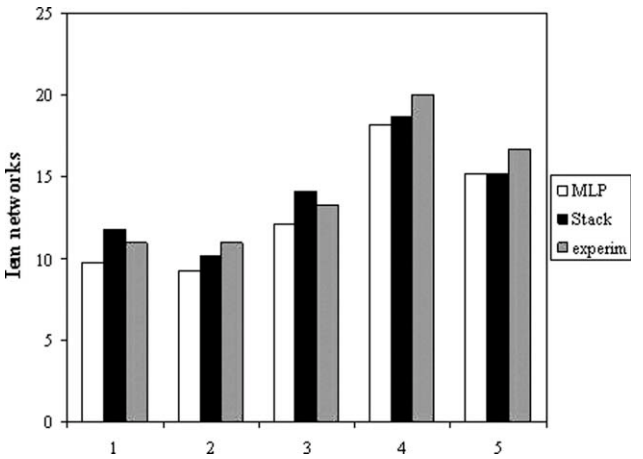


Figure 4 Predictions of the individual and stacked neural networks in the validation phase compared with the experimental data.

$$E_r = \frac{\mathrm{Iem}_{\mathrm{exp}} - \mathrm{Iem}_{\mathrm{net}}}{\mathrm{Iem}_{\mathrm{exp}}} \times 100 \qquad (2)$$

was used to evaluate the performance of the neural models (individual and stack) in the training and validation phases. In relation (2), indexes exp and net denote experimental and network values, respectively.

Different stacks were developed with the three networks, MLP (5 : 5 : 1), MLP (5 : 10 : 1), and MLP (5 : 14 : 1), the weight of each contributing to the stack being varied. To find the optimal weights, the first step was to systematically generate weights between 0 and 100% with a step of 10%, and to record the average relative error and correlation on the training and validation data sets. The best combination corresponds to the weights of 20, 30, and 50%. These percentages represent the notations w_i in Figure 1, which are the weights applied to the outputs of the three networks aggregated in stack.

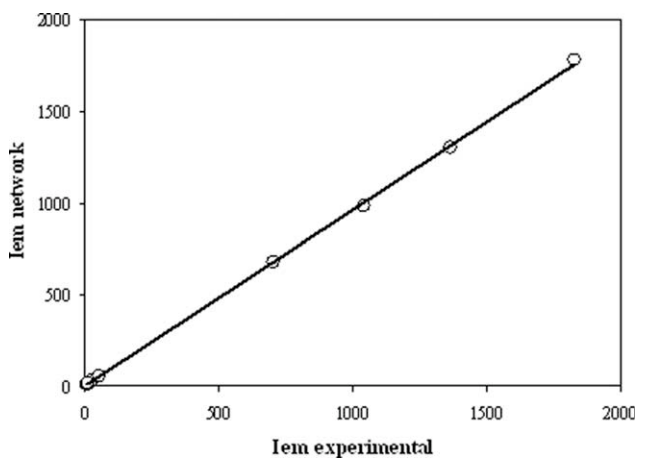


Figure 3 Predictions of the stacked neural networks in the training phase compared with the experimental data.

For the training phase, the comparison between relative errors of the MLP (5 : 24 : 8 : 1) (the network used individually) and the stacked neural network is given in Figure 2. It is evident that the stack has learned better the behavior of the process; the smaller errors mean that the predictions of the network are close to the experimental data. Figure 3 presents a comparison between experimental data and stacked neural network predictions on training data for I_{em}, proving that the model learned well the system property.

In the second stage, the neural model is used to predict output of the new input data ("unseen" data, not used in the training). The validation phase defines the generalization capability of the model. The predictions of the two selected neural models and the experimental data are compared in Figure 4. Individual neural network in Figure 4 is MLP (5 : 24 : 8 : 1).

Figures 2 and 4 takes into account the best single network-based model and the stack formed by 20% MLP (5 : 5 : 1), 30% MLP (5 : 10 : 1), and 50% MLP (5 : 14 : 1). Table II presents the performance in the training and validation phases, reflected through relative and maximum errors, for different stacked neural networks formed by the above three networks, but with other weighted outputs. It is evident that the combination 20, 30, and 50% is the best one, leading to small errors.

Figure 4 emphasizes the fact that the predictions of the stack are in better agreement with experimental data than those of the individual MLP. Therefore, combining some neural networks (preferably different neural networks) can provide a practical approach to developing a better overall model for

TABLE II
Performance of Different Stacked Neural Networks

Weights (%)			Training		Validation	
MLP(5 : 5 : 1)	MLP(5 : 10 : 1)	MLP(5 : 14 : 1)	Average error	Maximum error	Average error	Maximum error
100	0	0	6.32090	9.202	15.24210	34.40710
0	100	0	5.95300	9.0231	14.29800	29.29800
0	0	100	5.01140	8.901	12.30210	25.00300
10	50	40	4.56210	9.002	10.34600	17.02450
50	10	40	5.87560	10.2402	11.20260	21.00350
70	20	10	6.52100	10.5761	12.50910	23.69030
10	70	20	4.72010	8.8331	10.39200	19.03450
20	10	70	5.40040	9.197	8.92300	14.58210
20	30	50	4.15770	8.1	7.39030	9.87010

prediction, rather than searching for a single best-trained model.

Consequently, the neural network methodology provides useful and credible information for experimental practice, which can be a guide for future experiments or can substitute experiments that are time and material consuming.

CONCLUSIONS

A series of polydimethylsiloxane/silica composites containing complexed lanthanum ions was prepared by using different molar ratios between reactants. The study of fluorescence revealed that its intensity depends in a complex manner on the composition. Besides lanthanum as an active element, the fluorescence was also emphasized in the structures where lanthanum missed (i.e., functionalized silica). Instead, by the incorporation of PDMS as a mean to prepare flexible film, the fluorescence decreases being quenched.

The experimental data reported in this article have been used for the synthesis modeling of the polydimethylsiloxane/silica composites, namely the fluorescence property was estimated depending on the content of TEOS, APT, AR, $LaAc_3$, and PDMS in the reaction mixture.

For such a complex process, it was proven to be suitable a methodology based on neural networks, consisting of the following stages:

- Individual neural networks have been designed using a genetic algorithm technique for which the basic elements of the network topology were encoded within chromosomes. Development of an optimal architecture of the network was devised simultaneously with the network training and validation.
- Neural networks have been organized as stacks by the connection in parallel of the individual models and the weighted summation of the model component outputs. The corresponding weights of the individual model outputs have been chosen accordingly to the performance of the neural network stack model.
- Comparisons between the individual model and stack-model predictions, on one side, and experimental data, on the other side, have been made with the aim of choosing the best-performance model.

Two main conclusions can be formulated from the results of this article. Genetic algorithms represent a good and flexible method, easy to apply, with high chances to provide an optimal configuration of the neural networks. Second, stacking neural models is a possibility to improve the performance of the neural network-based model.

The proposed computing methodology is general and can be applied also to other polymerization processes whose physical and chemical governing laws are not well known. The obtained predictions provided useful information for the industrial and laboratory practice, proving to be able to effectively substitute material, energy, and highly time-consuming experiments.

References

1. Muller, G.; Maupin, C. L.; Riehl, J. P.; Birkedal, H.; Piguet, C.; Bünzli, J. C. G. Eur J Inorg Chem 2003, 22, 4065.
2. Rao, M. S.; Dubenko, I. S.; Roy, S.; Ali, N.; Dave, B. C. J Am Chem Soc 2001, 123, 1511.
3. Lai, D. C.; Dunn, B.; Zink, J. I. Inorg Chem 1996, 35, 2152.
4. Wang, Q. M.; Yan, B. J Organometal Chem 2006, 691, 545.
5. Rocca, F.; Monti, F.; Kuzmin, A.; Dalmaso, A.; Pasqualini, D. J Synchrotron Rad 1999, 6, 737.
6. Cazacu, M.; Vlad, A.; Airinei, A.; Alexandru, M. Polym Int 2008, 57, 1067.
7. Telipan, G.; Ignat, M.; Vlad, A.; Cazacu, M. J Optoelectr Adv Mater 2008, 10, 3409.
8. Marcu, M.; Cazacu, M.; Vlad, A.; Racles, C. Appl Organometal Chem 2003, 17, 693.
9. Mujtaba, I. M.; Hussain, M. A., Eds. Application of Neural Networks and Other Learning Technologies in Process Engineering; Imperial College Press: London, 2001.
10. Wolpert, D. H. Neural Networks 1992, 5, 241.

11. Zhang, J.; Morris, A. J.; Martin, E. B.; Kiparissides, C. Comp Chem Eng 1999, 23, 301.
12. Zhang, J. Bagging Predictors, Technical Report No. 421, Department of Statistics, University of California at Berkeley, USA.
13. Zhang, J. Neural Networks 1999, 12, 927.
14. Zhang, J. Neurocomputing 1999, 25, 93.
15. Zhang, J.; Jin, Q.; Xu, Y. Chem Eng Tech 2006, 29, 442.
16. Zhang, J. Chem Eng Sci 2008, 63, 1273.
17. Wang, S. Q.; Yang, J.; Chou, K. C. J Theory Biol 2006, 242, 941.
18. Tian, Y.; Zhang, J.; Morris, J. Ind Eng Chem Res 2001, 40, 4525.
19. Breiman, L. Learning 1996, 24, 49.
20. Freund, Y.; Schapire, R. In Machine Learning Proceedings of the 13th International Conference; Morgan Kaufmann: San Francisco, CA, 1996; p 148.
21. Zhang, J.; Martin, E. B.; Morris, A. J.; Kiparissides, C. Comput Chem Eng 1997, 21, 1025.
22. Curteanu, S. Cent Eur J Chem 2004, 2, 113.
23. Curteanu, S.; Petrila, C. Int J Quant Chem 2006, 106, 1445.
24. Curteanu, S.; Leon, F. Polym Plast Tech Eng 2006, 45, 1013.
25. Curteanu, S.; Leon, F.; Galea, D. J Appl Polym Sci 2006, 100, 3680.
26. Curteanu, S.; Cazacu, M. J Macromol Sci Part A: Pure Appl Chem 2007, 45, 23.
27. Curteanu, S.; Dumitrescu, A.; Mihailescu, C.; Simionescu, B. C. Polym Plast Tech Eng 2008, 47, 1061.
28. Curteanu, S.; Dumitrescu, A.; Mihailescu, C.; Simionescu, B. C. J Macromol Sci Part A: Pure Appl Chem 2008, 46, 368.
29. Nistor, A.; Piuleac, C. G.; Cazacu, M.; Curteanu, S. Mater Plast 2008, 45, 67.
30. Piuleac, C. G.; Rodrigo, M.; Cañizares, P.; Curteanu, S.; Sáez, C. Environ Model Software 2010, 25, 74.
31. Piuleac, C. G.; Curteanu, S.; Cazacu, M. Rev Roum Chim, to appear.
32. Bisagni, M.; Bun-Hoiet, N. P.; Roger, R. J Chem Soc 1955, 3688.
33. Cazacu, M.; Marcu, M. Macromol Rep A 1995, 32, 1019.
34. Tellez, L.; Rubio, J.; Morales, E.; Oteo, J. L. J Mater Sci 2003, 38, 1773.
35. Reich, Y.; Barai, S. V. Eng Appl Artif Intell 2000, 13, 685.
36. Curteanu, S.; Leon, F. Int J Quant Chem 2007, 108, 617.
37. Sa Ferreira, R. A.; Ferreira, A. L.; Carlos, L. D. J Non-Cryst Solids 2006, 352, 1225.