Comparison of a Radial Neural Network model with a Genetic Algorithm for estimating the Particle Size Distribution of a Latex

G. Stegmayer[1], L. Clementi[2] and J. Vega[1,2]

[1] CIDISI-CONICET, Lavaise 610, 3000, Santa Fe, Argentina
gstegmayer@santafe-conicet.gov.ar
[2] INTEC-CONICET, Güemes 3450, 3000, Santa Fe, Argentina
jvega@santafe-conicet.gov.ar

Abstract. This paper presents the comparison of an Artificial Neural Network (ANN)-based model and a Genetic Algorithm (GA) as independent tools for estimating the particle size distribution (PSD) of a polymer latex, which is an important physical characteristic that determines some end-use properties of the material. The PSD of a dilute latex is estimated from dynamic light scattering (DLS) measurements, taken at several angles. To this effect, an ANN-based model and GAs are used as a tool for solving the involved ill-conditioned inverse problem. Both models are trained with a large set of measurements simulated from typical asymmetric PSDs, represented by unimodal and bimodal normal-logarithmic distributions of variable geometric mean diameters and variances. The proposed approaches are evaluated on the basis of both simulated and experimental examples.

1 Introduction

Polymers play a major role in the current production of materials, both mass consumer commodities (such as engineering plastics, rubber, etc.), and more special products such as some latexes used as adhesives, paints, coatings, reagents for medical diagnosis, etc. [1]. The production of polymers with pre-specified quality characteristics is an important scientific and technological challenge, which combines expertise in the optimization of production processes, and the characterization of the obtained products. Regarding characterization [2], it is intended to determine some quality properties by means of analytical specific techniques and physical, chemical or mechanical tests on the properties of the final material.

Product characterization involves standard procedures for signals analysis and data treatments. Usually, it is necessary to solve ill-conditioned inverse problems (i.e. inverse problems are typical of most measurement systems where only indirect measurements are available of the desired properties), combined with theoretical principles of the employed analytical techniques [3]. The resolution of such problems involves the use of numerical techniques of digital filtering, smoothing and functions regularization, to partially mitigate the inevitable noise

XXXIV Conferencia Latinoamericana de Informática (CLEI 2008), pp. 82-91, 2008.

measurement presence in the signals and systematic mistakes originated during the modelling of the associated analytical technique, which limits the accuracy and resolution of the obtained solutions.

The quality of some polymer colloids (or latexes) are normally associated to their particle size distributions (PSD), which determines some end-use properties (e.g., rheological, mechanical, and physical properties) of the material when used as an adhesive, a coating, or an ink. For example, the PSD can define the behavior of adhesives and paints, and the chemical stability of latexes; and it can influence the physico-chemical mechanisms involved in emulsion polymerization [4]. Unfortunately, there is no analytical instrumentation capable of directly measuring a PSD. For this reason indirect measurements are needed, where the measured physical variables are related to the PSD through theoretical models. Some optical techniques, such as dynamic light scattering (DLS), can estimate a latex PSD from measurements of the light scattered by particles in a dilute dispersion, when they are lightened with a monochromatic light (typically, a laser). The resolution of the resulting inverse problem is usually approached using standard regularization techniques [5], but the obtained solutions have a low resolution (i.e., inability to differentiate among particles of similar size).

Some artificial intelligence techniques have been recently proposed for investigation of the optimal operating conditions required to obtain polymers with prespecified end-use properties, either on the basis of ANN alone [6] or combined with genetic algorithms [8]. In the area of polymer characterization, the application of ANNs and GAs for the resolution of inverse problem is rather scarce. For example, NNs have been used for pattern recognition in high performance liquid chromatography [9]. For particle sizing, they have also been used to estimate: a) the radius and refractive index of homogeneous spherical particles, based on a reduced number of light scattering measurements taken at multiple angles [10], b) the PSD of an aerosol, from measurements of laser light diffraction [11], and c) the radius, aspect ratio, and orientation of cylindrical and spherical particles, from light scattering measurements at multiple angles [12]. Regarding AGs, several authors have worked in the estimation of PSDs. The work of [13] has estimated several simulated unimodal and bimodal PSDs from static light scattering measurements, but using a particular range of diameters that does not include most of commercial latex. In [14] simulated polystyrene in water PSD is estimated using light scattering but only monodispersed cases were analyzed.

Those techniques have proved effective in mitigating the effect of noise on measurements, and to achieve better solutions than those obtained through classical inversion procedures. Moreover, as far as the authors are aware, no NN-based nor GA-based method has yet been published for estimating the complete PSD of unimodal and bimodal distributions.

This paper compares the performance of ANNs and GAs when used for solving the inverse problem of estimating the latex PSD from DLS measurements. The organization of this work is the following: Section 2 introduces some fundamentals concepts of DLS measurement technique; Sections 3 and 4 present the proposed ANN and GA inverse models; Section 5 shows some simulation and experimental results for model validation; and finally, Section 6 summarizes the main conclusions of the work.

2 DLS measurements fundamentals

DLS is an optical techniques widely used for measuring mean diameters and PSD of polymer latexes with particles in the sub-micrometer range [15]. The instruments employed for DLS basically consist of: i) a monochromatic laser light that falls onto a dilute latex sample; and ii) a photometer placed at a given detection angle, θ_r, with respect to the incident light, that collects the light scattered by the particles over a small solid angle. The PSD is calculated by solving an ill-conditioned inverse problem, on the basis of a mathematical model describing the light scattering phenomena (e.g, the Mie theory [16], [17]) and, unfortunately, only a rather poor PSD resolution can be expected.

For each θ_r $(r = 1, \cdots R)$, the measurements model can be described through the following first order Fredholm equations [18], [19]:

$$g_{\theta_r}^{(1)}(\tau) = \int_0^\infty e^{-\frac{\Gamma_0(\theta_r)}{D}} C_I(\theta_r, D) f(D) dD; \quad r = 1, \cdots, R \tag{1}$$

where $f(D)$ is the unknown PSD, represented by the number of particles with diameter D; $C_I(\theta_r, D)$ is the light intensity scattered by a particle with diameter D at θ_r calculated through the Mie theory, and $\Gamma_0(\theta_r)$ depends on several experimental conditions [18]. In general, the estimation problem consists in finding the (unknown) $f(D)$ by inverting equation 1. Such inverse problem is normally ill-conditioned; i.e., small errors in the measurement (for example, small perturbations due to measurement noise) can originate large changes in the $f(D)$ estimate. Moreover, the difficulty of the inverse problem increases as the distribution becomes narrower.

To reduce the dimension of the resulting inverse problem, we propose to replace equation 1 by the mean diameter calculated with DLS measurements at each θ_r. That diameter – which we will call $D_{DLS}(\theta_r)$ – can accurately be evaluated in most commercial equipment.

Call $f(D_i)$ the discrete number PSD, where f represents the number of particles contained in the diameter interval $[D_i, D_{i+1}]$, with $i = 1, \cdots, N$. All the D_i values are spaced at regular intervals ΔD along the diameter range $[D_{min}, D_{max}]$; thus, $D_i = D_{min} + (i-1)\Delta D$, with $\Delta D = (D_{max} - D_{min})/(N-1)$. Now, for a given PSD, $D_{DLS}(\theta_r)$ can be calculated through:

$$D_{DLS}(\theta_r) = \frac{\sum_{i=1}^{N} C_I(\theta_r, D_i) f(D_i)}{\sum_{i=1}^{N} \frac{C_I(\theta_r, D_i) f(D_i)}{D_i}}; \quad r = 1, \cdots, R \tag{2}$$

The estimation problem consists in finding the PSD ordinates $f(D_i)$, by inverting equation 2.

While DLS is reliable and fast for evaluating average particle diameters, it exhibits serious limitations for estimating the PSD due to the extreme ill-conditioning of equation 2, that makes it impossible to exactly obtain the PSD by numerical methods. Regularization methods aim at improving the numerical inversion by including adjustable parameters, a priori knowledge of the solution, or some smoothness conditions [5]. While a strong regularization produces an excessively smoothened and wide PSD, a weak regularization normally originates oscillatory PSD estimates. Thus, a trade-off solution must be selected. In general, the estimation of a narrow PSD is more difficult than the estimation of a wide PSD.

3 Training data

To simplify the problem, the discrete axis D_i $(i = 1, \cdots, N)$ and the angles θ_r $(r = 1, \cdots, R)$ are assumed fixed, and only the PSD ordinates (f) and the measurement ordinates (D_{DLS}) are presented to the model. Each (a-priori known) discrete PSD lies in the diameter range [50-1400] nm, with ΔD=5 nm. Measurements were taken in the θ range [30-140] degrees, with $\Delta\theta$= 10 degrees. Each input variable $D_{DLS}(\theta_r)$ is represented by $R = 12$ discrete points.

The PSDs used for building the model were restricted to be only bimodals, with a fixed log-normal shape for each mode, i.e.:

$$f(D) = C\frac{e^{\frac{[\ln(D/D_{med1})]^2}{2\sigma_1^2}}}{\sigma_1 D} + (1 - C)\frac{e^{\frac{[\ln(D/D_{med2})]^2}{2\sigma_2^2}}}{\sigma_2 D} \qquad (3)$$

where D_{med1} and D_{med2} are the geometric means of each mode; σ_1 and σ_2 are their corresponding standard deviations; and C is the number-fraction of particles present in each mode. For simulations, C was varied between 0.1 and 0.9.

For generating the learning set, each mean diameter D_{medi}, $i = 1, 2$, was varied in the range [50-1000] nm, at intervals of 10 nm; and the standard deviations σ_i were varied in the range [0.01-0.10] nm, at intervals of 0.01 nm. A total of K = 3620 learning patterns were obtained.

4 The inverse ANN-based model

The inverse neural model consists of a general regression neural network (GRNN) [20]. A GRNN is a normalized radial basis function network in which there is a hidden unit (k) centered at every learning case [21]. In a GRNN the number of neurons equals the total number of input/target vectors (K) selected for building the model. The hidden-to-output weights (w_{ik}) are just the target values. Then, the output is simply a weighted average of those target values that are close to the given input case. Strictly, a GRNN model is directly built on the basis of the learning cases, and therefore no specific training algorithm is required. The

GRNN can also be considered as a one-pass learning algorithm, with a highly parallel structure. Even with sparse data in a multidimensional measurement space, the algorithm provides smooth transitions from one observed value to another [20].

When using GRNN models, the selection of an appropriate smoothing (spread) factor is required, to be applied to each of the radial units, which indicates how quickly the activation function decreases as the distance increases from the neuron centroid. With a small spread, the neuron becomes very selective. With larger spread, distant points have a greater influence and the function approximation will be smoother.

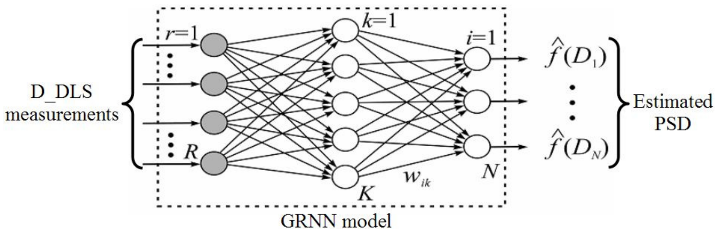

Fig. 1. Inverse GRNN model proposed for estimating the PSD of a latex.

Figure 1 shows a schematic representation of the inverse radial neural model proposed for the estimation of the latex PSD. This model is created using a set of K discrete PSDs (equation 3), and their corresponding measurements obtained according to equation 2.

5 The inverse GA-based model

GAs have proven adequate for solving complex optimization problems, particularly in the absence of linearity. The GA algorithm involves the implementation of three operators, namely Selection, Crossover, and Mutation, that apply onto a set of vectors representing the individuals in a population of possible solutions to a problem.

Each individual in a population is characterized by its "Fitness", i.e. how efficiently it solves the problem being considered. Individuals having higher fitness may have more probability of being selected as a parent to obtain the new population after evolution. Once the parents have been selected, Crossover and Mutation operators produce a new population. Each produced population is called a generation [22] and after a given number of generations, an optimal solution should be obtained.

Historically applications of GAs have been primarily restricted to those optimization problems which could easily be represented in binary form, but traditional GAs which employ binary string representations of the solution space are generally not convenient for solving continuous optimization problems [14]. In our particular case a GA has been used to optimize the parameters C, D_{med1}, D_{med2}, σ_1, and σ_2 of a bimodal normal-logarithmic distribution (equation 3). The GA algorithm aims at minimizing the following function:

$$J = \frac{\sqrt{\sum_{i=1}^{N}\left(1 - \frac{\overline{D^*}_{DLS,i}}{\overline{D}_{DLS,i}}\right)^2}}{N} \tag{4}$$

The function J of equation 4 represents the mean square error between the DLS average diameter for the individual being studied ($\overline{D^*}_{DLS}$) and the DLS average diameter for the "real" distribution ($\overline{D}_{DLS}$).

For each case studied, a random initial population is generated in the range: $100 < D_{med} < 1300$, $0,01 < \sigma < 0,3$ and $0 < C < 1$.. The population size was chosen as 400 and the crossover/mutation fraction was fixed in 70%/30%.

The fitness (F) for every individual was evaluated through equation 5:

$$F_i = \frac{[Mean(J) - J_i + Min(J)]}{\sum[Mean(J) - J_i + Min(J)]} \tag{5}$$

where J_i are the functional values obtained through equation 4 for the individual i. Thus, individuals who best fit the DLS Average Diameter will obtain lowest J and therefore greater F.

To perform the crossover, a two point crossover was adopted [22]. This operator randomly chooses two points along the parents string. Subsequently, parts of the parent between the two points chosen are exchanged to produce two new individual. A uniform mutation was implemented in which each of the parameters of each individual in the population is randomly changed with 5% of occurrence probability [14]. The best 4 individuals were selected to pass to the new generation without modification assuring that the best solutions do not get lost during the evolutionary process. The processes of selection, crossover and mutation are then applied to each generation until convergence of the J functional is met. Ten processes of estimation were implemented for each case studied adopting the best solution in the sense of minimizing J.

6 GRNN and GA models validation

Two kinds of validations were implemented. First, the models were tested with simulated (or synthetic) examples, since in these cases the solutions are a priori known, and therefore the models performance can be clearly evaluated. Then, the model was tested through an experimental example that involves a polystyrene (PS) latex of narrow PSD and known nominal diameter. In this case, the true PSD is unknown; but the best approximation is given by an independent PSD measurement as obtained from transmission electron microscopy (TEM) [23].

6.1 Validation with simulated data

A bimodal PSDs of a PS latex was simulated, $f_{sim}(D)$, which follows a bimodal log-normal distribution, with the following parameters: $D_{med1} = 200$ nm, $\sigma_1 = 0.15$, $D_{med2} = 400$ nm, $\sigma_1 = 0.075$ and $C = 0.85$ (85% weight for $D_{med1} = 200$). For this distribution, two cases were considered. In the first case, only D_{DLS} measurements without noise corresponding to this distributions were considered for input to the models. In the second case, D_{DLS} measurements having 0.1% added random noise were used as models inputs because it represents a more realistic case. In both cases, the estimated results should have been the same.

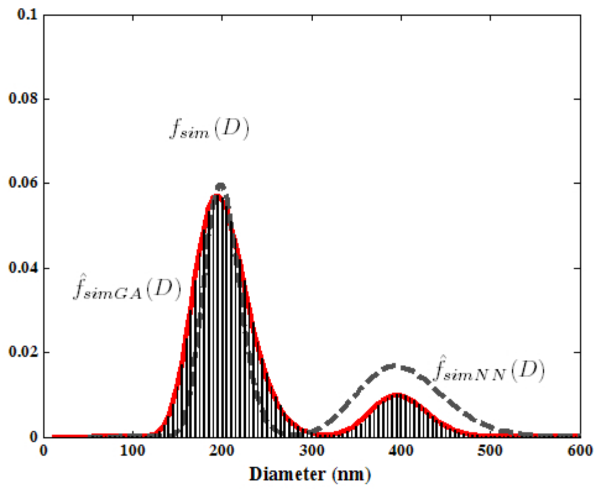

Fig. 2. Comparison of a simulated measurement $f_{sim}(D)$ with the GRNN estimate $\hat{f}_{simNN}(D)$ and the GA estimate $\hat{f}_{simGA}(D)$.

The selected bimodal PSD $f_{sim}(D)$ is represented in Figure 2 (in bars). The corresponding estimates are also represented in the figure. In the case of the GA estimate, $\hat{f}_{simGA}(D)$, its estimation is quite close to the simulated measurements. In the case of the GRNN estimate, $\hat{f}_{simNN}(D)$, its estimation has more error due to the influence that the training patterns have on the GRNN model.

In both cases, the estimates exhibits only positive values, which is practically impossible to be obtained when traditional regularization routines are used to solve the ill-conditioned inverse problem.

6.2 Validation with experimental data

A commercial latex standard of PS (from Duke Scientific) of nominal diameter 111 nm was measured through the following independent techniques: 1) DLS and

2) TEM. For the light scattering measurements, a Brookhaven instrument was used. The TEM measurement was obtained after counting about 500 particles, in a Hitachi H-7000 equipment [23].

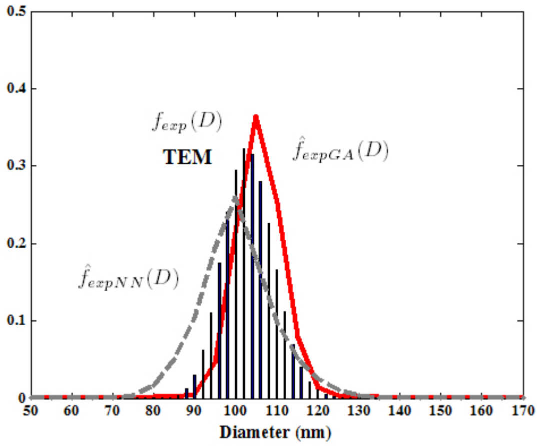

Fig. 3. Comparison of the TEM experimental measurement $f_{exp}(D)$ with the GRNN estimate $\hat{f}_{expNN}(D)$ and the GA estimate $\hat{f}_{expGA}(D)$.

The PSD obtained from TEM, $f_{exp}(D)$ is shown in Figure 3 as a histogram, and it is considered as a good approximation of the true (but unknown) PSD. The DLS measurements were fed into the trained GRNN and GA models; and the resulting estimated PSDs are indicated as $\hat{f}_{expNN}(D)$ and $\hat{f}_{expGA}(D)$, respectively. Despite the differences in the estimation, the average diameters of both estimated PSDs are near the measured value (102 nm): $\hat{f}_{expNN}(D)$ estimated an average diameter of 100 nm and $\hat{f}_{expGA}(D)$ estimated an average diameter of 105 nm.

7 Conclusions

Two models for estimating the particle size distribution of polymer latexes from DLS measurements were presented and compared. The proposed models are a general regression neural network (GRNN) and a genetic algorithm (GA), that was built on the basis of simulated log-normal PSDs, with particles in a relatively broad diameter range [50-1400] nm. The building of both models is straightforward and fast, because no training nor validation procedure is required. The proposed approaches were successfully evaluated on the basis of both simulated and experimental examples. It was observed that the resulting GAs were able to accurately recuperate PSDs of log-normal distributions with better approximation results than the GRNN model. This is due to the fact that the kind of patterns used for NN training have a strong influence on the neural model, and this could even prevent the model from being able to respond if inputs are very different from the ones the network has seen during training.

From a practical point of view, the GRNN and GA models constitute both a fast and robust tool, which additionally prove adequate for the resolution of the involved ill-conditioning non-linear inverse problem. With respect to the standard inversion techniques, these models present the advantages of not requiring any diameter range nor numerical inversion method. Also, they have proven to be insensitive to the standard noise measurements.

As future work, we are considering the combination of both models into a hybrid model or single procedure, using both strategies for PSD estimation, to be able to obtain a more accurate estimate of unknown PSDs, and to be able to compare the obtained results among all three cases: GA model, GRNN model and hybrid model. Moreover, we will consider the use of a hypothesis test to perform a more accurate comparison among all the models. We are also working on obtaining more experimental measurements of bimodal latex standards to further validate the proposed models.

References

1. Meyer, T., Keurentjes, J.: Polymer Reaction Engineering, an Integrated Approach. In: Handbook of Polymer Reaction Engineering. Chap. 1, pp. 1-15, Eds. Wiley-VCH (2005)
2. Schoenmakers, P., Aarnoutse, P: Chemical Analysis for Polymer Engineers. In: Handbook of Polymer Reaction Engineering. Chap. 20, pp. 1015-1046, Eds. Wiley-VCH (2005)
3. Kirsch, A.: An Introduction to the Mathematical Problem of Inverse Problems, Springer-Verlag, New York (1996)
4. Gilbert, R.: Emulsion Polymerization. A Mechanistic Approach, Academic Press, London (1995)
5. Tikhonov, A., Arsenin, V.: Solutions of Ill-posed Problems, Wiley, Washington (1977)
6. Fernandes, F.A., Lona, L.M.: Application of neural networks for the definition of the operating conditions of fluidized bed polymeriza- tion reactors. Polym. React. Eng. 10, PP. 181192 (2002)
7. Hanai, T, Ohki, T., Honda, H., Kobayashi, T.: Analysis of initial conditions for polymerization reaction using fuzzy neural network and genetic algorithm. Comput. Chem. Eng., 27, pp. 1011-1019 (2003)
8. Hanai, T., Ohki, T., Honda, T., Kobayashi, T., *Comput. Chem. Eng.*(2003), 27, 1011
9. Zhao, R. et al.: Application of an Artificial Neural Network in Chromatography - Retention Behavior Prediction and Pattern Recognition. Chem. & Intell. Lab. Syst., 45, pp. 163-170 (1999)
10. Ulanowski, A. et al.: Application of neural networks to the inverse light scattering problem for spheres. Appl. Optics, 37(18), pp. 4027-4033 (1998)
11. Guardani, R., Nascimento, C., Onimaru, R.: Use of Neural Networks in the Analysis of Particle Size Distribution by Laser Diffraction: Tests with Different particle Systems. Powder Tech., 126, pp. 42-50 (2002)

12. Berdnik, V., Loiko, V: Sizing of Spheroidal and Cylindrical Particles in a Binary Mixture by Measurement of Scattered Light Intensity: Application of Neural Networks. J. Quantit. Spectr. & Radiat. Transfer, 91, pp. 1-10 (2005)
13. Ye, M. et. al: Inversion of particle-size distribution from angular light-scattering data with genetic algorithms. Applied Optics 38, pp. 2677-2685 (1999)
14. Hodgson, R.J.W.: Genetic Algorithm Approach to the Determination of Particle Size Distributions from Static Light-Scattering Data; Journal of Colloid and Interface Science 240, pp. 412-418 (2001)
15. Chu, B: Laser Light Scattering, Academic Press, New York (1991)
16. Scheffold, F. et al.: PCS Particle Sizing in Turbid Suspensions: Scope and Limitations; In: Particle Sizing and Characterization, Eds. T. Provder and J. Texter (2004)
17. Glatter, O. et al.: Interpretation of Elastic Light-Scattering Data in Real Space. J. of Col. and Int. Sci. 105, pp. 577-586 (1985)
18. Vega, J. et al.: Latex Particle Size Distribution by Dynamic Light Scattering. A Novel Data Processing for Multi-Angle Measurements. J. Coll. and Int. Sci., 261, pp. 74-81 (2003)
19. Vega, J. et al.: A Method for Solving an Inverse Problem with Unknown Parameters from Two Sets of Relative Measurements. Lat. Amer. Appl. Res., 35, pp. 149-154 (2005)
20. Specht, D: A generalized regression neural network, IEEE Trans. Neural Networks, 2, pp. 568-576 (1991)
21. Wasserman, P.: Advanced methods in neural computing, Van Nostrand Reinhold, New York (1993)
22. Goldberg, D.E.: Genetic Algorithms in Search, Optimization, and Machine Learning. Addison-Wesley, New York (1989)
23. Elizalde, O., Leal, G., Leiza, J: Particle Size Distribution Measurements of Polymeric Dispersions: A Comparative Study, Particle & Particle Systems Characterization, 17(6), pp. 236-243 (2000)