Inferential Estimation of Molecular Weights of Polybutadiene Rubber by Neural Networks

G. Padmavathi,[1] M. G. Mandan,[1] S. P. Mitra,[1] K. K. Chaudhuri[1] Shital J. Patel,[2] S. P. Dabke[2]

[1]*Indian Petrochemicals Corporation Limited, Baroda - 391 346, India*
[2]*Chemical Engineering Department, Faculty of Technology, M. S. University of Baroda, Baroda - 390 002, India*

Received 5 January 2004; accepted 18 August 2004
DOI 10.1002/app.21606

ABSTRACT: Molecular weight distribution, which is characterized by its averages like number average (Mn) and weight average (Mw), is one of the important properties of polybutadiene rubber (PBR), and it is difficult to measure. The objective of this work is to develop models to predict Mn and Mw from readily available process variables. Neural networks that are capable of mapping highly complex and non-linear dependencies have been adapted to develop models for the Mn and Mw of PBR. The molecular weight distribution and its averages of PBR samples collected over a wide range of operating conditions were measured by the conventional Gel Permeable Chromatograph (GPC) method. Neural networks were trained with relevant data to predict Mn and Mw from process variables. The trained networks were found to generalize well when tested with new data. © 2005 Wiley Periodicals, Inc. J Appl Polym Sci 96: 1611–1618, 2005

Key words: molecular weight distribution; number average molecular weight; weight average molecular weight; neural networks; polybutadiene rubber

INTRODUCTION

Molecular weight distribution is one of the important properties of polymers as it influences the physical, mechanical, and rheological properties of the polymers. Generally, molecular weight distribution of a polymer is characterized by molecular weight averages, like number average (Mn), weight average (Mw), z- average (Mz), and their ratio Mw/Mn known as poly dispersity. Small differences in low or high molecular weight fractions can result in significant differences in end use properties. The molecular weight distribution of a polymer is widely used in quality control, monitoring product stability and reclaiming off spec material, and development of new grades.[1] In the plant operating conditions are set in such a way that the polymer of desired molecular weight characteristics is produced. Hence, a model that estimates the value of these properties as a function of the process variables would be useful for controlling these molecular weight properties of the polymer.

PBR is produced by continuous polymerization of 1–3 butadiene in a series of three stirred tank reactors. Mooney viscosity, solution viscosity, and molecular weight distribution are important quality properties and are difficult to measure. Mooney viscosity indirectly represents molecular weight and is measured once in two hours off line in the plant laboratory. Fluid viscosity is measured once a day in the laboratory. Molecular weight is not measured as frequently as viscosity due to the tedious, complex, and time consuming GPC method. The resulting measurement delay and infrequent feedback makes automatic process control impossible. This can be overcome by development of a model that can infer the quality of the product properties from easily available process variables. Changes in some of the variables indicate changes in product quality. Hence, this work has been undertaken with an objective to develop an inferential model to predict Mn and Mw of PBR from major process variables by using the neural network approach. The trained neural networks can be used to generate the estimates of product quality corresponding to process variables instantaneously. Then the estimates can be used for process monitoring and control.

PROCESS

PBR is produced by polymerization of 1–3 butadiene in three continuous stirred tank reactors in series. The monomer, the solvent, and the catalyst system comprising of the catalyst, the cocatalyst, and the chain modifier are mixed thoroughly before entering the first reactor. The reactors are provided with jackets for temperature control and an agitator of a double helical ribbon type for mixing of high viscosity polymer solution. The reaction temperature is approximately 60–70°C, and about 88% of the monomer is converted into the polymer at the outlet of the last reactor. In the first reactor, a large amount of polymerization heat is generated, and the heat is removed by a combination of the sensible heat

---

Correspondence to: K. K. Chaudhuri (kumar.chaudhari@ipcl.co.in).

Journal of Applied Polymer Science, Vol. 96, 1611–1618 (2005)

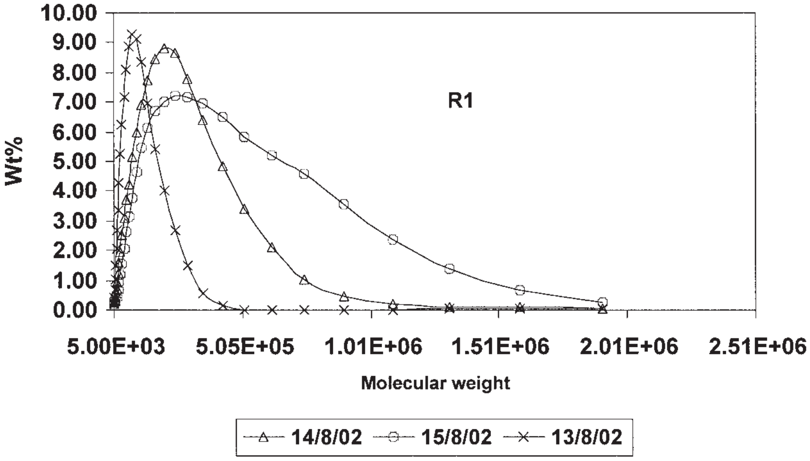

Figure 1 Molecular weight distribution of samples from R1.

required to heat up the reactor feed (monomer and solvent) and of the heat required to vaporize the refrigerant in the jacket. In the second and third reactors, a comparatively smaller amount of the polymerization heat is generated, and the heat is removed by a combination of the sensible heat required to heat up the incremental solvent to be used in the second and third reactors. The reactants enter from the bottom of the reactors and come out from the top. After the polymer is formed, the conversion and the physical properties, such as Mooney viscosity, are analyzed and they are controlled so that they fall within each process control standards. The Mooney viscosity is controlled by the chain modifier level and traces of moisture in the solvent. The polymer solution produced is mixed with a chain modifier and an antioxidant and held in the blend tank, and the final control of physical properties in the polymerization step is performed by the blending operation.

EXPERIMENTAL

PBR samples were collected from the first reactor outlet (R1) and the third reactor outlet (R3) once in two hours for analysis. The two hourly samples are mixed together to get composite samples for R1 and R3, and a portion of it is used for analysis of molecular weight distribution by the conventional GPC method. The number of samples analyzed for molecular weight distribution from R1 is 23 and from R3 is 72.

The molecular weight distribution of PBR samples was measured by GPC using a Waters Model GPC 150C at a temperature of 85°C with toluene as the solvent. The method involves dissolution of the polymer sample in the solvent to make a dilute polymer solution. Around 1 mg of PBR sample was dissolved in toluene. The dissolution time is optimized to 24 hours after trying various times starting from 4 to 48 hours. The dilute polymer solution is pumped across a packed column with a defined distribution of porosity. The packing material used in the column is a highly cross-linked porous styrene divinyl benzene gel.[2]

The polymer is separated based on the hydrodynamic volume of the polymer fractions. The molecular weight of each fraction will have different elution times corresponding to their residence time in the pores. Molecules too large to penetrate any of the pores are totally excluded from the column and elute first. Slightly smaller molecules penetrate some of the pores, are retained on the column, and elute somewhat later. Molecules small enough to penetrate all of the pores are retained on the column longest and elute last. The elution profile represents the molecular weight distribution. The amount of eluted polymer is determined by measuring the concentration by means of the refractive index.

Once all the fractions of the sample are separated, the molecular weight of the fractions are determined from the calibration curve obtained by using narrow distribution polystyrene standards. Universal calibration[3] procedure was followed to calculate molecular weight distribution of PBR samples. Figures 1–3 show the molecular weight distribution of PBR samples from R1 and R3.

Neural networks

Artificial neural networks (ANN) are renowned for their utility as universal approximators of complex non-linear relationships between process variables and product quality properties.[4] The success of the network depends on the selection of process variables, the quality of the data, and the type of network used. In the present work, the most widely used feed forward network with one hidden layer is used.[5] Figure 4 shows the typical structure of an artificial neural network. The neurons are represented by a circle. Each neuron in the hidden layer (upper layer) receives weighted inputs plus bias from each neuron in the input layer or layer below it:

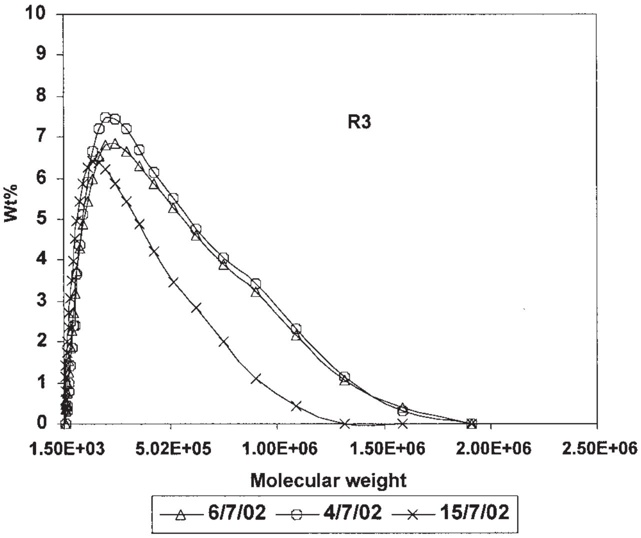

Figure 2 Molecular weight distribution of samples from R3.

$$f_{1i} = \sum_{j=1}^{S} W_{ij}p_j + b_i \quad (1)$$

where p is the input matrix, w is the initial weight matrix, b is the bias, and S is the number of variables in the input matrix. The output of the layer, a_i, which is input to the next layer, is calculated by tan-sigmoid transfer function as:

$$a_{1i} = 2/(1 + \exp(-2^* f_{1i})) - 1 \quad (2)$$

$$f_{2k} = \sum_{i=1}^{N} lw_{ik}a_{1i} + b_k \quad (3)$$

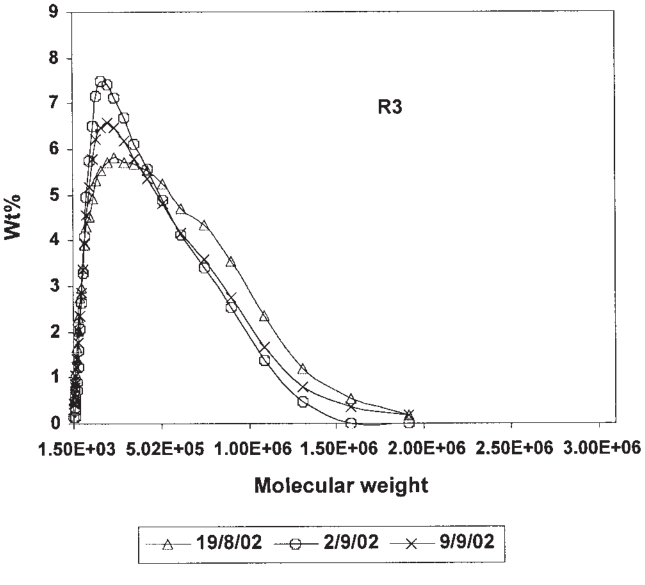

Figure 3 Molecular weight distribution of samples from R3.

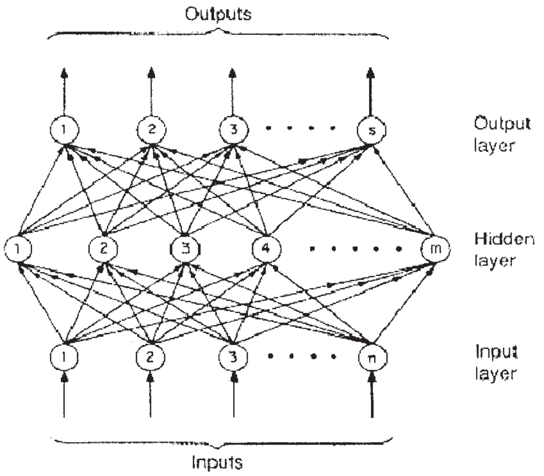

Figure 4 Typical structure of an artificial neural network.

where lw is the layer weight matrix, b_k is the bias, N is the number of neurons in the hidden layer, k is the number of targets in the output layer; and the output of the output layer a_2 is by linear transfer function, which is:

$$a_{2k} = f_{2k} \quad (4)$$

The values of weights and bias are set during the network training process. Initially, the weights are set randomly. The neural networks are trained by input and target data sets. The data set is divided randomly into a training set and testing sets. The training set is used to train the network to adjust the weights. The training of the network involves adjusting the values of the weights and bias so that the error between target values and the predicted values is minimum. The Levenberg–Marquardt back propagation minimization algorithm is used to adjust the weights. The mean square error is calculated by:

$$MSE = \left(\sum_{1}^{n}\sum_{1}^{k}(t_{pred.} - t_{act.})^2\right)/nk \quad (5)$$

where n is the number of data points, and k is the number of targets.

RESULTS AND DISCUSSION

The important process variables that control the product quality include flow rates of monomer and solvent, purity of monomer and solvent, catalyst ratios, reaction temperature, traces of moisture, and shear rate. In the plant the average of all the process variables over a day is compiled. The process variable data corresponding to the samples collected for molecular weight measurement was compiled as the input matrix. The Mn and Mw of samples from R1 and R3 are

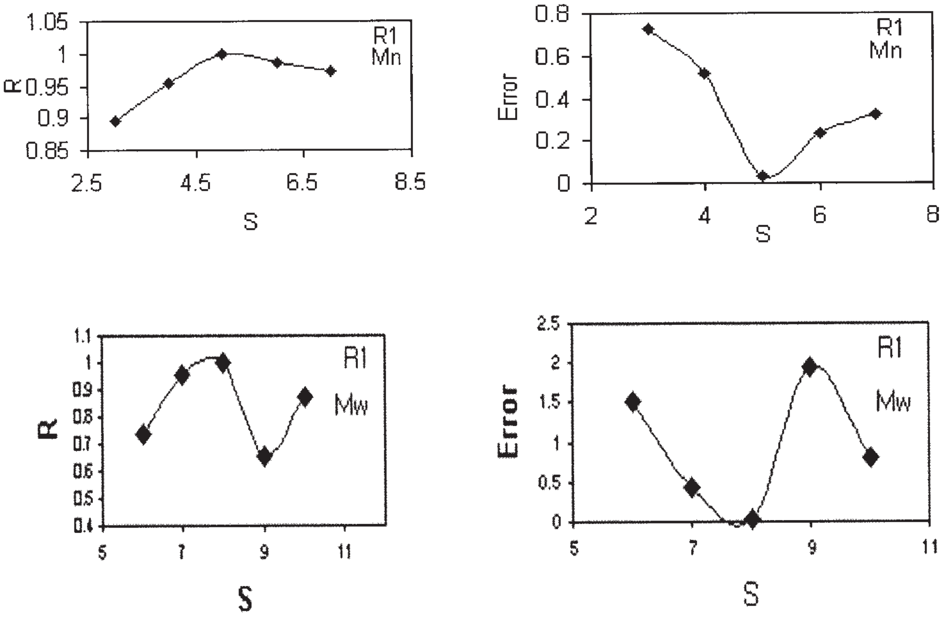

Figure 5 Optimization of state for Mn and Mw of R1.

targets, and their values were determined from the molecular weight distribution measured by GPC. The total number of data points available for training and testing a network for Mn and Mw of samples from R1 is 23. The total number of data available from R3 is 72.

Multiple linear regression has been performed, and no relation could be established between the process variables and molecular weights. The value of correlation coefficient, R, obtained was very much lower than one and mean square error was high. Multiple regression has been performed by taking the logarithm of the process variables to map the non-linearity between the process variables and the molecular weights. Though there was improvement in R value and error, it was not significant enough.

Neural network development

The data was examined thoroughly and outliers removed. The input and target are normalized to enhance the predictive performance of the network. Data normalization is very important for process monitoring and soft sensor development networks.

For training, the data has been divided into training and testing subsets by picking randomly. A three layer—input layer, hidden layer, and output layer—net-

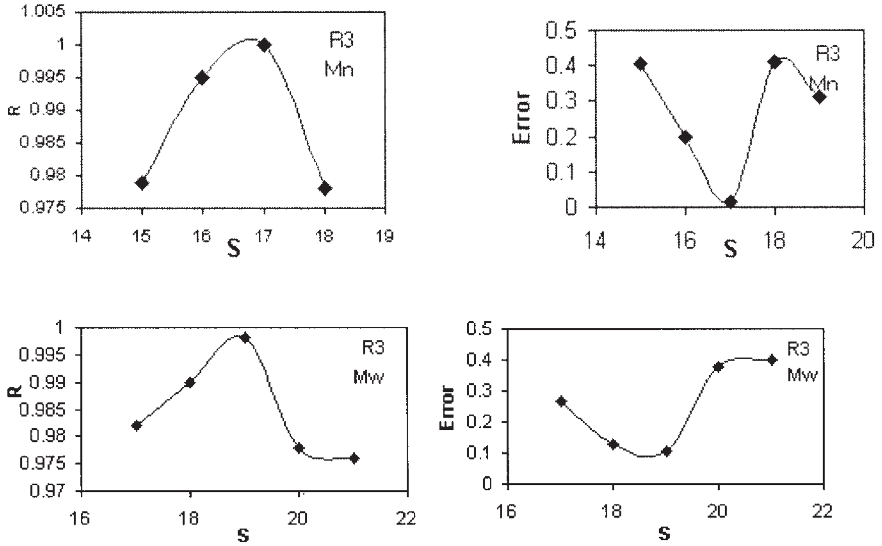

Figure 6 Optimization of state for Mn and Mw of R3.

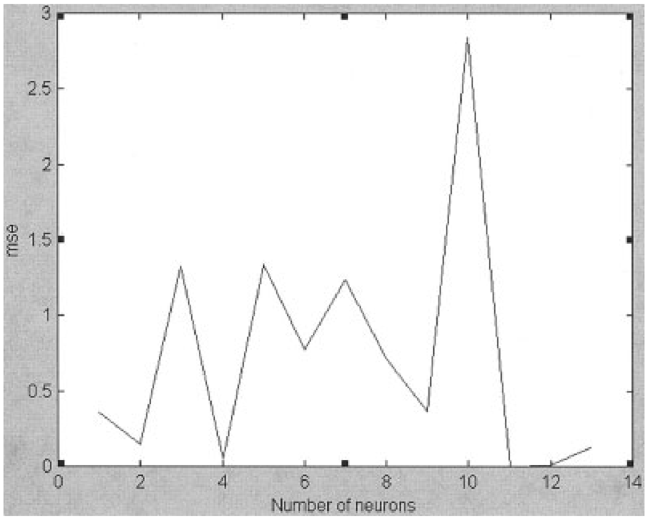

Figure 7 Mean square error vs. number of neurons for Mn of R1. (MSE = mean square error.)

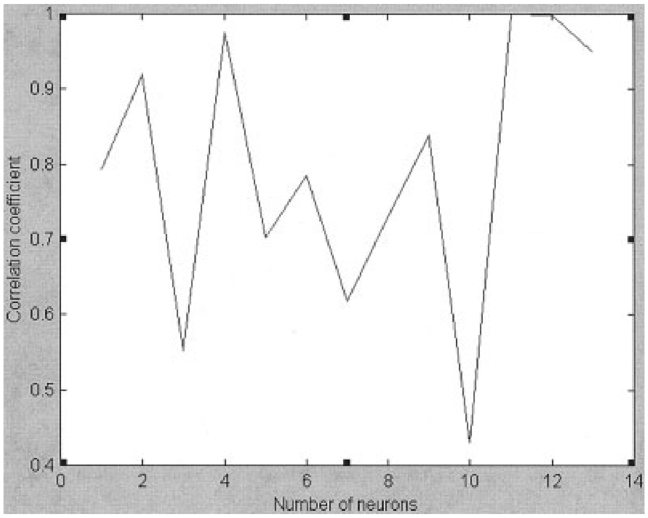

Figure 8 Correlation coefficient vs. number of neurons for Mn of R1.

work has been created with tan-sigmoid transfer function in the hidden layer and linear transfer function in the output layer for training the data. The Levenberg–Marquardt algorithm is used for training the network.

Before training a feed forward network, the weights and biases were initialized by random generation of small weights. The sequence of numbers generated is determined by the state of the random number generator. For a given network with a specified number of neurons, the value of state has been varied over a wide range. The final state has been fixed when the network performed best with respect to mean square error and correlation coefficient (Fig. 5 and 6).

Finding the number of neurons is an important task in neural network development. A number of networks were trained for the task using different numbers of neurons until the error criterion was satisfied. The error criterion is that the network has to perform its best with respect to mean square error, correlation coefficient, and test errors. The number of neurons in the hidden layer was found by trial and error. Figures 7 and 8 show the number of neurons against mean square error and correlation coefficient for prediction of Mn of PBR from R1.

Since the number of data points is limited, splitting the data into training and testing sets that are statistically significant is very difficult. For dealing with small databases, a sliding method has been reported.[6] With the sliding method, the network is trained on all patterns but one and tested on the pattern that was omitted. The same process is repeated n times (n being the number of available training patterns), each time leaving out one and calculating the error. A similar procedure was adapted in the present case for R1. For R3 the data has been divided randomly into training (95%) and test (5%).

The structure of the network was changed by increasing the number of hidden layers to two with the tansigmoid transfer function. The network was tested with different training functions, and it was found that the network performed its best when the Levenberg–Marquardt method was used. The application of the Levenberg– Marquardt algorithm to neural network training was described in literature.[7] The log-sigmoid transfer function has also been tested, and it was found that the network's performance was better with the tansigmoid transfer function. The summary of the structure of the four networks developed for modeling of Mn and Mw of R1 and R3 are presented in Table I. The first number in the number of neurons row indicates the number of input variables; the second and third numbers, if any, indicate the number of neurons in hidden layers; and the last number indicates the number of output variables in the output matrix. Figures 9–12 show the comparison of

TABLE I
Structure of the Final Networks for Mn and Mw

	Mn (R1)	Mw (R1)	Mn (R3)	Mw (R3)
State	5	8	17	19
No. of neurons	7-11-1	7-5-5-1	7-9-1	7-7-1
Transfer function (hidden layer)	tan-sigmoid	tan-sigmoid	tan-sigmoid	tan-sigmoid
Training method	Levenberg–Marquardt	Levenberg–Marquardt	Levenberg–Marquardt	Levenberg–Marquardt
Correlation coefficient	0.999	1	1	0.998

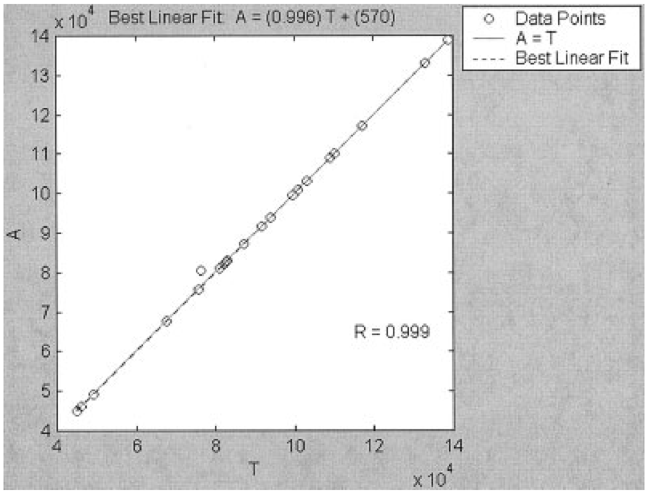

Figure 9 Comparison between experimental and predicted Mn of R1.

experimental and predicted Mn and Mw of PBR from R1 and R3, respectively.

Figures 13 and 14 show the comparison of the model predicted Mn and Mw from R1 and R3, respectively, with those determined by GPC for all the training and testing samples. The results obtained with the networks are excellent. However, in commercial reactors, the process conditions are often time varying, and the neural networks developed under these conditions lead to local models that define functions that are applicable in the neighborhood of the current state of the plant as they are developed using only those states of the plant that were close to the current one. The range of the model can be updated whenever Mn and Mw are measured in the lab for the new samples. Until then, the Mn and Mw can be estimated using these trained neural networks.

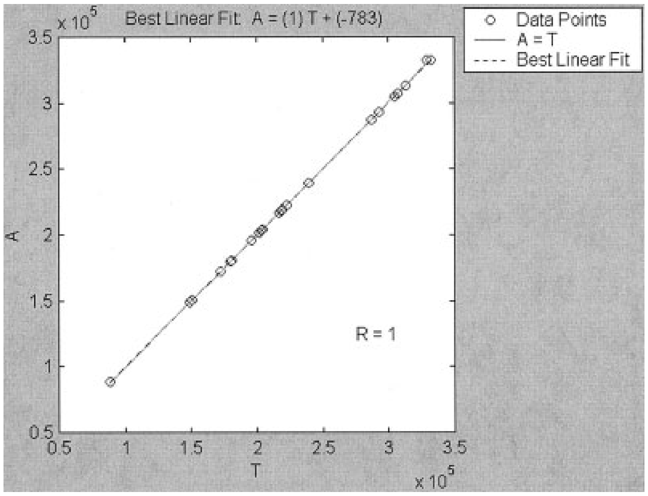

Figure 10 Comparison between experimental and predicted Mw of R1.

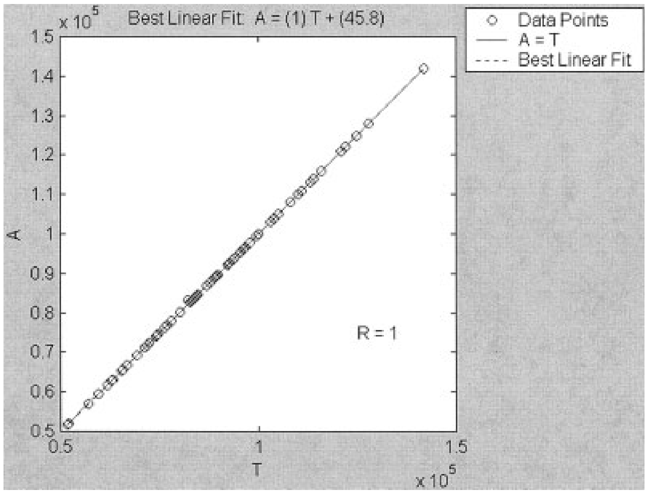

Figure 11 Comparison between experimental and predicted Mn of R3.

Parametric sensitivity

The parametric sensitivity of the networks developed for Mn and Mw of R1 and R3 has been tested by changing each variable over a certain range keeping the other variables constant so that the predicted Mn and Mw are in the range covered in the network development. The value of each variable used for simulation is between the minimum and maximum range of that variable used in a commercial process.

Effect of monomer purity

The monomer purity is in the range of 0.991 and 0.999 for most of the operating data collected corresponding to the time of the samples collected for the molecular weight analysis. Degree of purification of the monomer affects molecular weight. Its control is very important for poly-

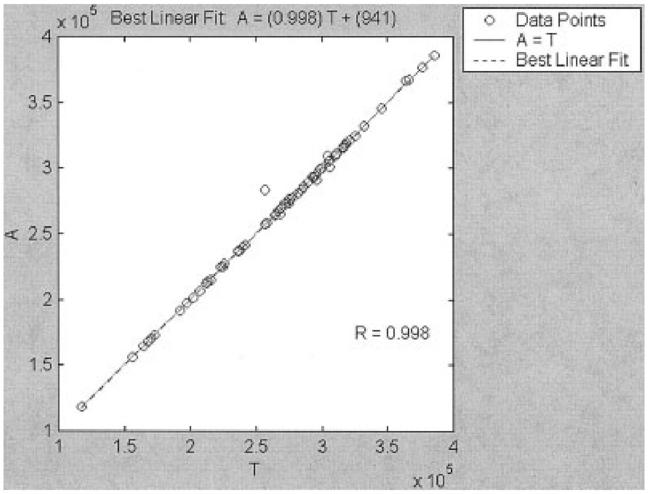

Figure 12 Comparison between experimental and predicted Mw of R3.

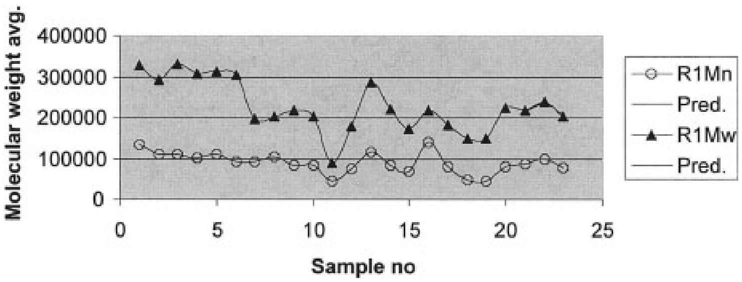

Figure 13 Measured and predicted Mn and Mw vs. sample number.

merization. It is required to maintain the monomer's purity and impurities in ppm order. The simulations carried out with the trained networks did not indicate a significant effect of monomer purity on both Mn and Mw of R1. The monomer purity is very high initially in the first reactor. This may be the reason that the Mn and Mw are almost constant with change in purity, whereas both Mn and Mw of R3 showed an increasing trend with increasing purity. In the last reactor there is a chance of building up of impurities with the addition of incremental solvent, etc.

The results reflected this observation of increasing molecular weight with purity. However, controlling of the physical properties of the polymer by a change in the conditions of the purification system requires much time until the solvent and monomer in the purification system and dry tank are completely replaced by new ones. Hence, this method is not suitable for rapid action.

Effect of flow rates of solvent and monomer

The reaction temperature in butadiene polymerization is controlled by controlling the temperature of the solvent charged, solvent to monomer ratio (S/M), and the reactor jacket liquid level. A lower S/M gives a higher monomer concentration and, hence, a higher reaction velocity. A low S/M causes disadvantages, including need for the removal of reaction heat in the reactor and too high a load on the agitator. In addition, a higher reaction rate gives a higher viscosity of the polymer liquid and makes the removal of reaction heat difficult. As molecular weight is directly proportional to viscosity, an increase of viscosity indicates an increase of molecular weight. The model predicted trend shows the increase of both Mn and Mw of R1 and R3 with monomer flow rate and tends to decrease after reaching a certain level. Hence, increment solvent shall be added to the second and third reactors to control them efficiently.

The model predicted Mn and Mw increased with an increase in solvent flow rate. At constant monomer flow rate, S/M increases with an increase in solvent flow rate, which reduces the severity of the reaction.

Effect of temperature

The predicted molecular weight decreased with an increase in polymerization temperature. Generally, in normal operation, the temperature is maintained constant to maintain the product quality and stable operation. It is an effective means for control of solution viscosity.

Effect of catalyst concentration

In most of the Ziegler–Natta catalyst systems used for production of PBR, the molecular weight of the polymer is inversely proportional to catalyst concentration. A similar trend of decreasing molecular weight with increasing catalyst level is observed for both R1 and R3.

Effect of catalyst ratios

The model predicted molecular weight increased with an increase in cocatalyst to catalyst ratio up to a certain value. Generally, this ratio is maintained at an optimum value in normal operation.

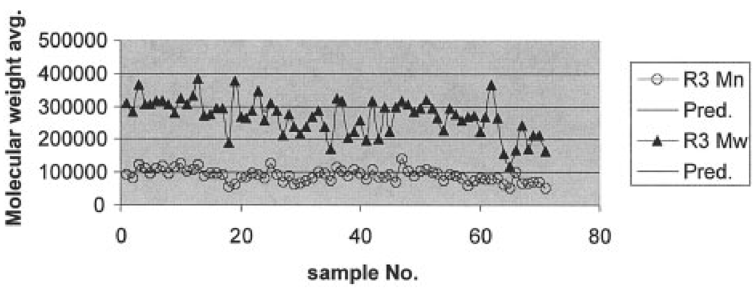

Figure 14 Measured and predicted Mn and Mw vs. sample number.

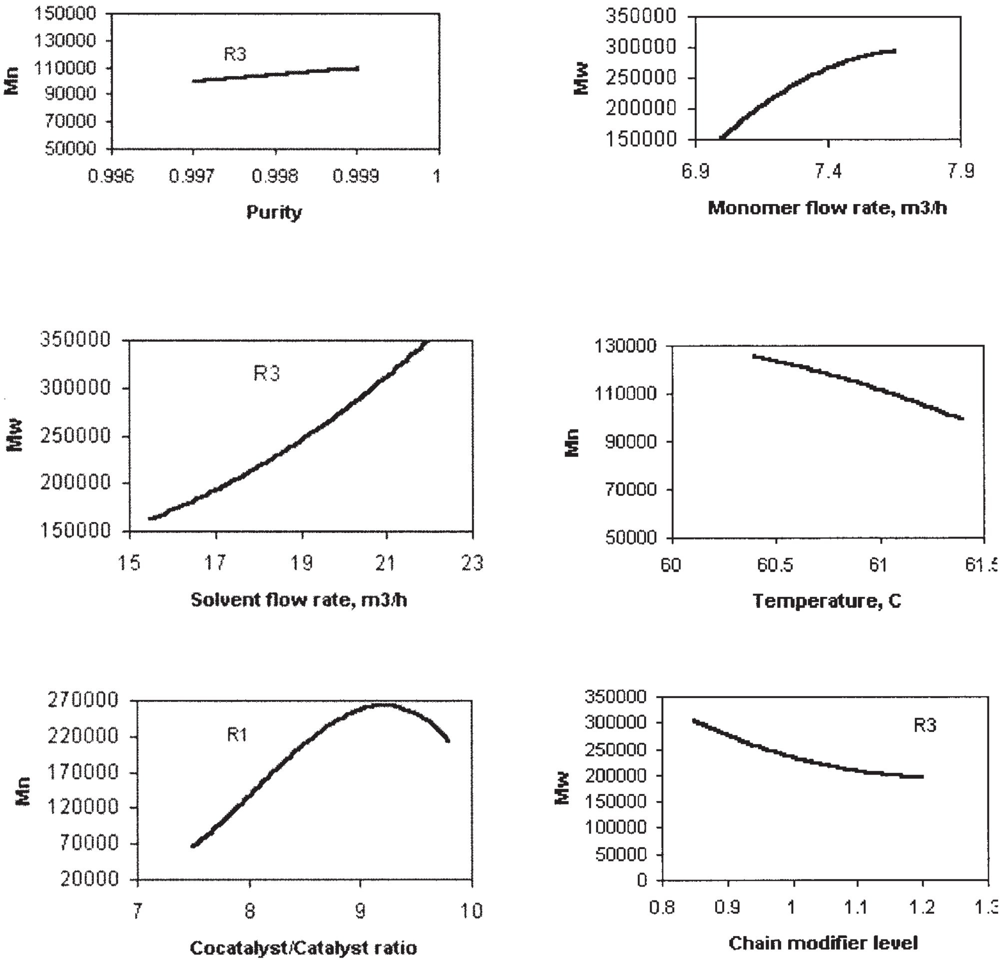

Figure 15 Parametric sensitivity of the process variables.

Effect of chain modifier

The amount of chain modifier has to be increased to decrease molecular weight. The model predicted results show a similar trend. The chain modifier terminates the growing chains and promotes fresh chains in such a way that the average molecular weight of the polymer remains within a close range.

Figure 15 shows the predicted trends of the critical process variables on the molecular weight of PBR. The sensitivity of the process variables has been tested to make sure that the networks have learned the trends properly. In real operation, a single parameter alone is not varied. The process variables are varied in such a way that the plant operation is stable and ratios of solvent to monomer, cocatalyst to catalyst, and chain modifier to cocatalyst remain in the range required to get the desired product quality.

CONCLUSION

Molecular weight distribution and their averages of around one hundred PBR samples collected over a wide range of operating conditions were measured using conventional GPC and the universal calibration method. The neural network approach has been used for inferential estimation of number and weight average molecular weights of polybutadiene rubber from the first and last reactor outlet. The trained networks can be used to estimate Mn and Mw of PBR from both R1 and R3 instantaneously at the same rate as secondary variables and, hence, corrective action can be taken in case of off spec product.

References

1. Indian Rubber Institute. Rubber Engineering; Tata McGraw-Hill Publishing Co. Ltd.: New Delhi, India, (1998); Chapter 2, pp 48–113.
2. Grubisic, Z.; Rempp, P.; Benoit, H. J Polym Sci, Lett Ed 1967, 5, 753–759.
3. GPC Operation Manual, 150C Waters Associates Inc., 1981.
4. Bulsari, A. B. Neural Networks for Chemical Engineers, Computer-Aided Chemical Engineering; 6, Elsevier: Amsterdam, 1995.
5. MATLAB - Neural Network Toolbox, Release 11.1, The Math Works, Inc., Natick, MA.
6. Tsaptsinos, D. In Neural Networks For Chemical Engineers; Bulsari, A. B., Ed.; Elsevier: Amsterdam, 1995; pp 33–75.
7. Hagan, M. T.; Menhaj, M. IEEE Transactions on Neural Networks 1994, 5(6), 983–993.