

Selection of a Mixture of Initiators for Batch Polymerization Using Neural Networks

Fabiano A. N. Fernandes

Universidade Federal do Ceará, Departamento de Engenharia Química, Campus do Pici, Bloco 709, 60455-760 Fortaleza-CE, Brazil

Received 2 July 2004; accepted 22 March 2005

DOI 10.1002/app.22421

Published online in Wiley InterScience (www.interscience.wiley.com).

ABSTRACT: The productivity in batch processes is related to reduction of the time required to complete each batch. An increase in productivity can be achieved by running the polymerization isothermally using a mixture of initiators with different decomposition rates. The amount of each initiator in the mixture can be optimized to increase productivity while not exceeding the maximum heat release that the cooling system is capable of compensating for. In

this work, neural networks are used to search for the optimum operating condition of the reactor. The procedure to find the best neural network topology is presented as well as its application. © 2005 Wiley Periodicals, Inc. *J Appl Polym Sci* 98: 2088–2093, 2005

Key words: styrene; polymerization; initiator; modeling; neural networks

INTRODUCTION

Productivity of batch processes is directly related to reduction of the time required to complete each batch. Industrial-scale reactors are designed to support a maximum rate of heat release by exothermic polymerization, which normally corresponds to autoacceleration of the polymerization rate.

Nevertheless, the average rate of heat release during batch time is significantly smaller than the maximum cooling capacity of the system, which means that the cooling system is underutilized for most of the polymerization.¹ The amount of heat that could still be released is represented by the gray region in Figure 1.²

This potential heat can come from an increase in polymerization rate in the beginning of the batch, which can be achieved using an initiator with short decomposition. Two other initiators with medium and long decomposition times are used to spread the polymerization rate and heat release over the batch time.

The formulation of an initiator mixture can be stated as an optimization problem, in which the decision variables are the amount of each initiator and the operating temperature. The constraints to be satisfied include the final desired quality of the polymer (molecular weight and polydispersity), maximum cooling capacity, and desired productivity.

Neural networks were tested to evaluate whether they could be applied to this kind of optimization

problem. In this paper optimization, using a neural network, was applied to search for the optimum initiator mixture used to promote bulk polymerization of styrene.

STYRENE POLYMERIZATION

Polystyrene can be produced in batch, solution, suspension, and emulsion polymerization. In bulk polymerization of styrene, the reaction occurs in a vessel where product is drained at the end of the reaction. The raw materials (styrene and initiators) are charged to the vessel at one prior time to reacting, and the reaction is allowed to proceed for the necessary amount of time to achieve the desired polymer properties including desired polymerization amount (conversion). When processing is complete, the product is drained from the reactor vessel. The reactor is enclosed partially in a jacket through which a heat transfer fluid, such as water, is passed for heating and cooling the contents of the vessel. Typically, batch polymerization of styrene runs at temperatures from 45 to 100 °C using initiators and also above 100 °C, when thermal initiation begins to play a major role in the formation of free radicals.³

Polymerization involves a high heat of reaction, increasing heat of reaction, and increasing solution viscosity as polymerization progresses and a corresponding decrease in heat transfer coefficient of the reaction material. Because of these problems, controlling the temperature of polymerization is very important because of the strong dependency of free-radical reaction kinetics on reaction temperature, directly af-

Correspondence to: F. A. N. Fernandes (fabiano@efftech.eng.br).

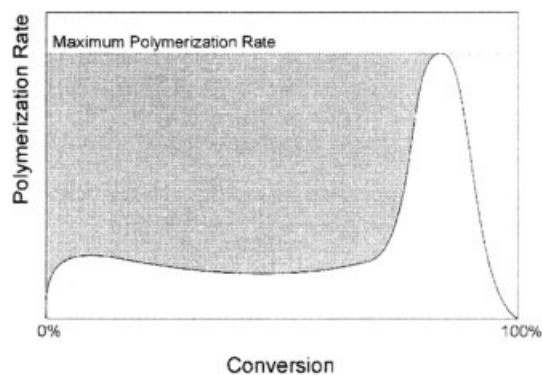


Figure 1 Polymerization rate.

fecting polymer properties such as molecular weight distribution. The styrene polymerization kinetic mechanism is well known and displays conventional kinetics up to high conversions.⁴ The propagation is very fast, if compared to the initiation step, with the presence of a mild gel effect, leading to high molecular weights. These features leave us with some challenges to be studied regarding polystyrene production, such as the use of bifunctional or tetrafunctional initiators and the optimization of initiator cocktails.

Mathematical model

To formulate an optimal initiator mixture for bulk polymerization of styrene, a representative mathematical model for the system is needed to generate the data points that will be used to train the neural network. The mathematical model used for bulk polymerization of styrene consists of a system of ordinary differential equations that represents the mass and energy balances, along with suitable equations for the kinetics of the diffusion-controlled reactions in free radical polymerization and equations for the moments of the live and dead polymer chains.

The kinetic model, used in the simulations, considers the classic steps for free-radical mechanism: initiator decomposition, initiation, propagation, transfer to monomer, and termination. The kinetic constants are shown in Table I.

Three initiators were selected and used in the initiator mixture: Vazo 52, Vazo 64, and Vazo 88, all from DuPont. These initiators are substituted azonitrile compounds that thermally decompose to generate two free radicals per molecule. Vazo 52 is a 2,2'-azobis(2,4-dimethylpentanenitrile), Vazo 64 is a 2,2'-azobis(2-methylpropanenitrile), also known as AIBN, and Vazo 88 is a 1,1'-azobis(cyclohexanecarbonitrile). The chemical structures for these initiators are shown in Figure 2 and the decomposition rate constants for these initiators are shown in Table II.

The concentration of monomer (styrene) and initiators in the batch reactor is given by

$$\frac{d[I_j]}{dt} = -k_{i,j} \cdot [I_j] \quad (1)$$

$j = \text{Vazo 52, Vazo 64 and Vazo 88}$

$$\frac{d[M]}{dt} = -k_p \cdot [M] \cdot Y_0 - k_{fm} \cdot [M] \cdot Y_0 - k_{dt} \cdot [M]^3. \quad (2)$$

The moments for live and dead polymers are given by

$$\frac{dY_0}{dt} = 2 \cdot f \cdot k_i \cdot [I] + 2 \cdot k_{dt} \cdot [M]^3 - k_{tc} \cdot Y_0^2 \quad (3)$$

$$\begin{aligned} \frac{dY_1}{dt} = & 2 \cdot f \cdot k_i \cdot [I] + 2 \cdot k_{dt} \cdot [M]^3 + k_p \cdot [M] \cdot Y_0 \\ & + k_{fm} \cdot [M] \cdot (Y_0 - Y_1) - k_{tc} \cdot Y_0 \cdot Y_1 \quad (4) \end{aligned}$$

$$\begin{aligned} \frac{dY_2}{dt} = & 2 \cdot f \cdot k_i \cdot [I] + 2 \cdot k_{dt} \cdot [M]^3 + k_p \cdot [M] \cdot (2 \cdot Y_1 \\ & + Y_0) + k_{fm} \cdot [M] \cdot (Y_0 - Y_2) - k_{tc} \cdot Y_0 \cdot Y_2 \quad (5) \end{aligned}$$

$$\frac{dQ_0}{dt} = 0.5 \cdot k_{tc} \cdot Y_0^2 + k_{fm} \cdot [M] \cdot Y_0 \quad (6)$$

$$\frac{dQ_1}{dt} = k_{tc} \cdot Y_0 \cdot Y_1 + k_{fm} \cdot [M] \cdot Y_1 \quad (7)$$

$$\frac{dQ_2}{dt} = k_{tc} \cdot (Y_0 \cdot Y_2 - Y_1^2) + k_{fm} \cdot [M] \cdot Y_2. \quad (8)$$

TABLE I
Kinetic and Other Constants for Styrene Bulk Polymerization

Styrene thermal initiation	$k_{dt} = 1.314 \cdot 10^7 \text{ L} \cdot \text{mol}/\text{min}$ $E_{adt} = 114701 \text{ J}/\text{mol}$
Propagation	$k_p = 1.302 \cdot 10^9 \text{ L} \cdot \text{mol}/\text{min}$ $E_{ap} = 32486 \text{ J}/\text{mol}$
Termination by combination	$k_{tc} = 4.920 \cdot 10^{11} \text{ L} \cdot \text{mol}/\text{min}$ $E_{atc} = 14534 \text{ J}/\text{mol}$
Transfer to styrene	$k_{fm} = 6.579 \cdot 10^8 \text{ L} \cdot \text{mol}/\text{min}$ $E_{afm} = 56215 \text{ J}/\text{mol}$
Heat of reaction	$\Delta H = -683.6 \text{ J}/\text{g}$
Free volume	$\alpha_{STY} = 1.0 \times 10^{-3}$ $T_{gSTY} = 184.8 \text{ K}$ $\alpha_{PS} = 4.8 \times 10^{-4}$ $T_{gPS} = 378 \text{ K}$
Gel effect correlation	$A = 0.348$ $m = 0.50$

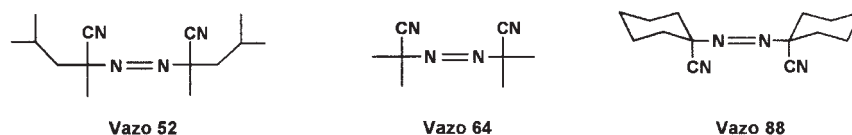


Figure 2 Chemical structures of Vazo 52, Vazo 64, and Vazo 88 initiators.

During polymerization, the propagation and termination rates become diffusion-controlled and the respective constants are given by

$$k_{tc} = k_{tc}^0 \cdot \left(\frac{\bar{M}_{w,crit}}{\bar{M}_w} \right)^m \cdot \exp \left[-A \cdot \left(\frac{1}{V_f} - \frac{1}{V_{f,crit}} \right) \right]. \quad (9)$$

The free volume (V_f) is given by

$$V_f = 0,025 + \alpha_{STY} \cdot (T - T_{gSTY}) \cdot \phi_{STY} + \alpha_{PS} \cdot (T - T_{gPS}) \cdot \phi_{PS}. \quad (10)$$

The heat release rate, which must be compensated for by the reactor cooling system to maintain the temperature constant, is calculated by

$$Q = R_p \cdot (-\Delta H) \cdot V. \quad (11)$$

The maximum heat release rate was recorded for each simulation to be used as a constraint with neural network optimization.

SELECTION OF INITIATORS VIA NEURAL NETWORK

Neural network training

Neural networks have attracted great interest as predictive models, as well as for pattern recognition. The potential for using neural networks in industrial applications is enormous. Neural networks have the ability of learning the behavior of the process and the relationships between variables, without needing a model of the phenomenological laws that rule the system. The success in obtaining a reliable and robust network depends strongly on the choice of process variables involved, as well as the available sets of data and the domain used for training purposes.⁵

TABLE II
Decomposition Constants for Vazo Initiator

Vazo 52 decomposition	$k_i = 6.25 \times 10^{10} \text{ min}^{-1}$ $Ea_i = 80578 \text{ J/mol}$
Vazo 64 decomposition	$k_i = 4.22 \times 10^{10} \text{ min}^{-1}$ $Ea_i = 84319 \text{ J/mol}$
Vazo 88 decomposition	$k_i = 3.74 \times 10^{10} \text{ min}^{-1}$ $Ea_i = 93561 \text{ J/mol}$

In general, the network consists of processing neurons and information flow channels between the neurons called interconnect. Each processing neuron calculates the weighted sum of all interconnected signals from the previous layer plus a bias term and then generates an output through its activation sigmoid functions.

In this work, the back propagation algorithm was used, as it is the most extensively adopted algorithm, which performs well. The available data were split into two sets. One set was used to train the network and the other to test its prediction capability. The activation sigmoid function used in the neural network was

$$y = \frac{1}{1 + \exp(-\Sigma x)}.$$

A random selected bias was used, and weights were updated by a Hessian approach. Special attention was given to not overtrain the network.

The data used to train the neural networks were obtained running a mathematical model for bulk polymerization of styrene. Table III presents the ranges for temperature and initiator concentration that were used in the simulations. The operating conditions (initiators concentration and temperature) were selected randomly between the range presented in Table III. Random selected data rather than factorial designed data were used since random data provides better training for this kind of neural networks.^{6,7} A total of 394 operating conditions were simulated, 298 being used to train the neural network and 96 for testing it.

Neural network selection

In view of selecting the best NN topology, NNs with one, two, and three hidden layers were tested and the

TABLE III
Ranges Used for Each Operating Condition

Variable	Range	
	Minimum	Maximum
Initiator concentration (mol/L)	0.008	0.1
Temperature (K)	333	383

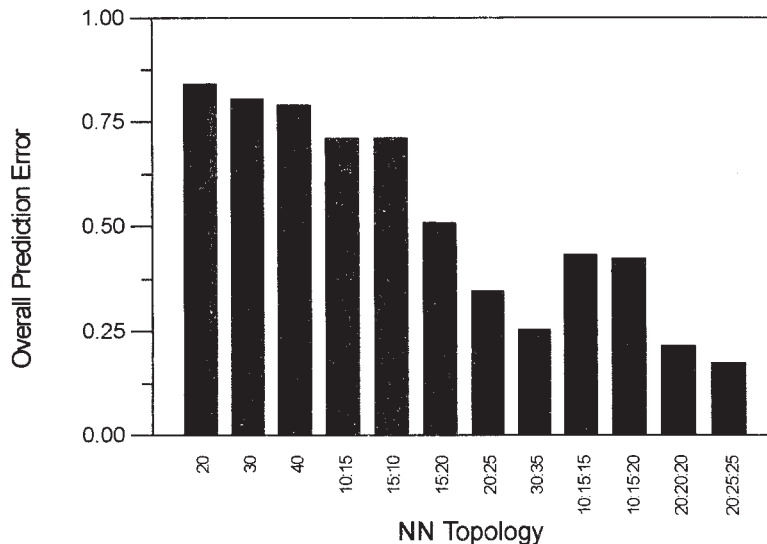


Figure 3 Overall prediction errors for several neural network topologies.

prediction errors were compared. Figure 3 presents the results for the NN topologies, where a $x:y$ topology means a neural network with two hidden layers with x neurons in the first layer and y neurons in the second layer. Lower errors indicate potentially good topologies. The overall prediction error (OPE) is calculated as

$$OPE = \frac{\sum e}{n} = \frac{\sum |(simulation\ data) - (NN\ prediction)|}{n}$$

The first step in training NNs is to compare different topologies, training them over 60,000 iterations. Training of these NNs shows the following:

- NNs with one hidden layer output very poor predictions;

- NNs with two hidden layers, if the second hidden layer has at least five more neurons than the first hidden layer, will generate better predictions than NNs with the second hidden layer with the same number or with fewer neurons as the first hidden layer;
- the same observation is valid for NNs with three hidden layers regarding the second hidden layer, while the third hidden layer may have the same number of neurons as the second hidden layer.
- Figure 4 presents the overall prediction error for five NN topologies as functions of the number of training iterations. Figure 4 shows that 40,000 iterations can be used as a decision point to decide which topology may be used in optimization of the system.
- The final decision of which NN should be used must be done by comparing the mean prediction

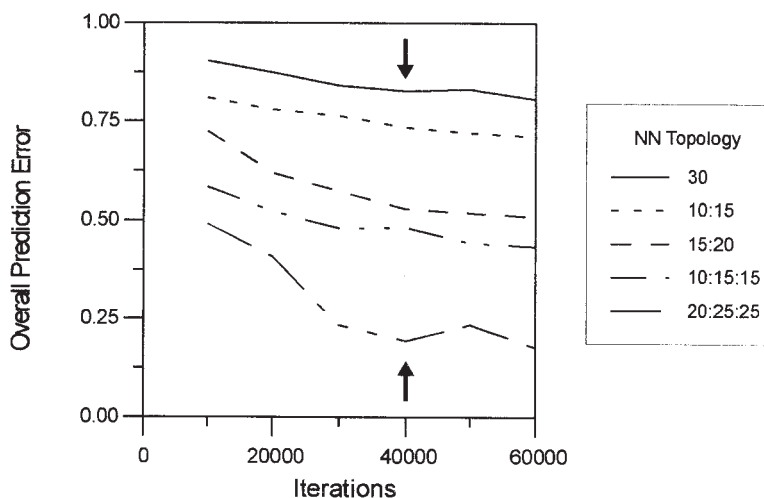


Figure 4 Overall prediction errors as a function of training iterations.

TABLE IV
Prediction Errors

Topology	Overall error	Mean prediction errors (%)			
		Vazo 52	Vazo 64	Vazo 88	Temperature
30:35	0.255	10.42	7.63	10.54	0.41
20:20:20	0.216	8.18	6.65	5.66	0.33
20:25:25	0.176	6.60	4.55	5.40	0.29

error (MPE) of each operating condition for the best topologies (Table IV). The MPE is calculated as

$$\text{MPE} = \frac{\sum e}{n}$$

$$e = \left| \frac{(\text{simulation data}) - (\text{NN prediction})}{(\text{simulation data})} \right| \times 100 [\%].$$

Training has shown that a NN with 20:25:25 topology gave the best results and was used in the optimization of the initiator mixture. To use in the optimization, this NN was further trained over 250,000 iterations, to a point where the mean error dropped to 3.8% (Table V).

Selection of the initiator mixture

Once trained, the NN can be directly used to output the optimal operating condition of batch polymerization of styrene for a given molecular weight (MW), polydispersity, productivity, and maximum heat release. The advantage of using NNs in this optimization phase is that no optimization algorithm is needed, since the output of the NN will be optimum.

The results that were obtained were promising, and the prediction errors using NN were small. Table VI shows two typical examples of the prediction that can be done for this problem, comparing the NN prediction to simulation results.

TABLE V
Predictions Errors for the Final NN

Topology	Overall Error	Mean prediction errors (%)			
		Vazo 52	Vazo 64	Vazo 88	Temperature
20:25:25	0.091	3.80	2.62	2.05	0.17

Optimization of the initiator mixture

Besides their use in the selection of initiator mixtures, trained neural networks can be used to optimize reactor productivity and polymer quality as well. Productivity can be improved by using the NN to search for new operating conditions that, for example, can increase the productivity while keeping all other polymer characteristics constant. Figure 5 shows an example of the increase in productivity that can be obtained using the neural network.

To optimize productivity, a known case (operating condition 1) was used as the starting point for optimization. A search procedure was created to find the optimum point, which consisted of increasing the productivity variable (NN input variable) while maintaining constant the values of the other input variables. Upon each increase, the trained NN outputted new operating conditions for the reactor to achieve that specified productivity. The increase in the value of productivity continued till an invalid value for the operating conditions was outputted by the NN, marking the end of the search for optimum productivity. The invalid value can be an impossible operating condition (such as a negative concentration) or a condition outside the training range.

In Figure 5, operating condition 1 marks the known operating condition, operating condition 2 marks the result found after increasing the productivity by 0.001 kg/L · min, and operating condition 3 marks the final result found, after the NN failed to output valid results. Operating condition 3 was tested with the mathematical model and the result has confirmed it as the optimum result.

When optimizing the operating conditions of a reactor using neural networks, the results should be

TABLE VI
Selection of Initiator Mixtures for Styrene Bulk Polymerization

Desired polymer and constraints	Predicted operational conditions	Prediction error (%)
$M_w = 106,150$ g/mol	Vazo 52 = 0.093 mol/L	1.70
Polydispersity = 18.4	Vazo 64 = 0.095 mol/L	0.02
Maximum heat = 53.2 kJ/L · min	Vazo 88 = 0.070 mol/L	2.37
Productivity = 0.007 kg/L · min	Temperature = 334.7 K	0.29
$M_w = 75,850$ g/mol	Vazo 52 = 0.063	1.11
Polydispersity = 11.2	Vazo 64 = 0.070 mol/L	1.60
Maximum heat = 39.1 kJ/L · min	Vazo 88 = 0.082 mol/L	1.43
Productivity = 0.012 kg/L · min	Temperature = 356.1 K	0.64

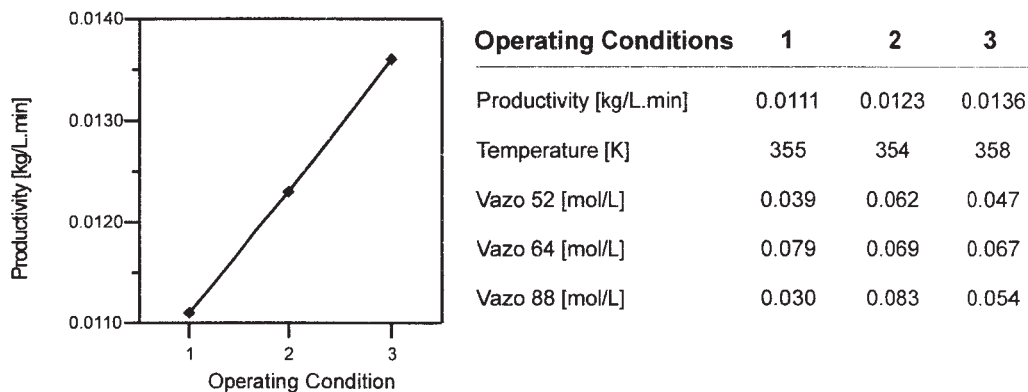


Figure 5 Productivity optimization.

carefully scrutinized, since the NN will always output a prediction but without guaranteeing its feasibility. Neural networks are not good with extrapolations, so if the prediction falls beyond the training range for a particular variable, this result may not be correct and it should be checked further with the reactor model.

CONCLUSIONS

This work presents a procedure using neural networks that can be employed to select a mixture of initiators and the operating conditions for batch polymerization of styrene to produce a polymer with a given molecular weight, polydispersity, and at the same time an enhanced productivity, as well as not exceeding the maximum heat load of the reactor jacket system. Predictions using this procedure were good, outputting estimates with less than 5% error.

Optimizations, to improve reactor productivity and polymer quality, can easily be made once the neural network has been trained. The time spent optimizing the reactor's operating conditions is low and the results are good. Special care must be taken to analyze whether the predictions lie within the range for which the NN was trained since neural networks are not particularly good with extrapolations.

References

1. Pinto, J. M.; Giudici, R. *Chem Eng Sci* 2001, 56, 1021.
2. Penlidis, A. *Polymer Reaction Engineering: Course Notes*: Waterloo, Canada, 2001.
3. Ellis, M. E. US Pat. 1999, 5986011.
4. Gao, J.; Penlidis, A. *J Macromol Sci* 1998, C38, 651.
5. Nascimento, C. A. O.; Giudici, R.; Guardani, R. *Comp Chem Eng* 2000, 24, 2303.
6. Fernandes, F. A. N., PhD Thesis, Universidade Estadual de Campinas: Brazil, 2002.
7. Fernandes, F. A. N.; Lona, L. M. F. *Pol Reac Eng* 2002, 10, 181.