Research Article

Hybrid modelling and kinetic estimation for polystyrene batch reactor using Artificial Neutral Network (ANN) approach

Mohammad Anwar Hosen,* Mohd Azlan Hussain and Farouq S Mjalli
Chemical Engineering Department, University of Malaya, Kuala Lumpur, Malaysia

Received 7 October 2009; Revised 9 February 2010; Accepted 9 February 2010

ABSTRACT: Modelling polymerization processes involves considerable uncertainties due to the intricate polymerization reaction mechanism involved. The complex reaction kinetics results in highly nonlinear process dynamics. Available conventional models are limited in applicability and cannot describe accurately the actual physico-chemical characteristics of the reactor dynamics. The usual practice for operating polymerization reactors is to optimize the reactor temperature profile because the end use properties of the product polymer depend highly on temperature. However, to obtain accurate models in order to optimize the temperature profile, the kinetic parameters (i.e. frequency factors and activation energies) for a specific reactor must be determined accurately. Kinetic parameters vary considerably in batch reactors because of its high sensitivity to other reactor design and operational variables such as agitator geometry and speed, gel effects, heating systems, etc. In this work, the kinetic parameters were estimated for a styrene-free radical polymerization conducted in an experimental batch reactor system using a nonlinear least squares optimization algorithm. The estimated kinetic parameters were correlated with respect to reactor operating variables including initial reactor temperature ($T_o$), initial initiator concentration ($I_o$) and heat duty ($Q$) using artificial neural network (ANN) techniques. The ANN kinetic model was then utilized in combination with the conventional mechanistic model. The experimental validation of the model revealed that the new model has high prediction capabilities compared with other reported models. © 2010 Curtin University of Technology and John Wiley & Sons, Ltd.

KEYWORDS: polystyrene batch reactor; modelling polymerization reactor; kinetic parameters; parameter estimation; optimization; artificial neural network

INTRODUCTION

Synthesizing industrial polymers is achieved through a multitude of reaction mechanisms and processes, including addition (e.g. free-radical and group-transfer) and step-growth polymerization. Every reaction mechanism has to be studied thoroughly and treated by considering the variables that affect polymer conversion and product grade. Polymer engineers are continuously challenged by the difficulties involved in optimizing the performance of these reactors. In this respect, much attention is directed towards understanding the major factors affecting the physical transport processes (e.g. mass, heat transfer and mixing), reactor configuration and reactor operating conditions which consequently affect the macromolecular architecture (e.g. molar mass, molecular weight distribution, etc.). As the polymer industry becomes more competitive, polymer manufacturers face increasing pressures for cost reductions in production and more stringent ‘polymer quality’ requirements. To achieve these goals one needs to develop comprehensive mathematical models capable of predicting the product quality indicators in terms of reactor configuration and operating conditions.

Usually polymerization reactors are modelled dynamically using a lumped parameter approach. In this case, the model consists of differential equations that describe the dynamic behaviour of the system, such as mass and energy balances, and algebraic equations that describe physical and thermodynamic relations. However, polymerization reactors are difficult to study numerically due to the fact that steady state is never achieved practically. In order to improve their performance and safety conditions, polymerization reactors generally require knowledge about the dynamic behaviour, for instance, through a mathematical description of the kinetics. A thorough investigation in the literature, for the effort of developing such models for batch reactors reveals that

*Correspondence to: Mohammad Anwar Hosen, Chemical Engineering Department, University of Malaya, Kuala Lumpur, Malaysia. E-mail: anwar.buet97@gmail.com

© 2010 Curtin University of Technology and John Wiley & Sons, Ltd.
researchers tend to enforce a lot of assumptions regarding the kinetics and the reactor environment. Available models are limited in applicability and cannot describe the actual reactor dynamics accurately. Consequently, it is necessary to develop new accurate models that best represent the characteristics of the reactor and can be used confidently for process control aspects.

Elicabe and Meira\textsuperscript{[1]} reported three basic problems related to polymerization mathematical models: (1) These models are of high nonlinear nature and therefore, all the control theory developed for linear systems is not strictly applicable. (2) Model parameters are in many instances unknown, and their estimation may be extremely difficult. (3) Disturbances such as impurities may have significant influence on the polymerization process. Furthermore, most of the disturbances are immeasurable. Westerhout et al.\textsuperscript{[2]}, in their review, have concluded that all published polymerization models are limited to narrow conversion ranges and reaction conditions. In their opinion, the use of simple first-order kinetic models is not appropriate to describe the kinetics over large conversion ranges. In addition, kinetic parameters are frequently taken from the literature in which experimental conditions may be considerably different and the results may vary significantly. Although direct experimental determination of kinetic parameters is not a trivial task, it must be stressed the necessity to calibrate each model for each specific experimental conditions as well as reactor condition instead of using values of parameters reported in previous works.

Ray et al.\textsuperscript{[3]} developed a complete mechanistic model for styrene polymerization batch reactor consisting of mass and energy balances and kinetic rate relations. This basic modelling study was the first successful attempt in this area and was used by many other researchers as a starting point for related modelling studies.\textsuperscript{[4–6]} However, all these researchers used a fixed set of kinetic rate constants in all their runs at different operating conditions.

To overcome the difficulty of model calibration, many algorithms were introduced to estimate the unknown model parameters using numerical optimization techniques. This includes using genetic algorithms,\textsuperscript{[7,8]} simulated annealing,\textsuperscript{[9]} particle swarm optimization (PSO)\textsuperscript{[10]} and using inverse methods\textsuperscript{[11]} for different systems such as ecological modelling, estuarine eutrophication modelling, etc.

Pagano\textsuperscript{[12]} proposed a new methodology for parameter estimation from kinetic model of polymeric resin using a differential algebraic approach. The PSO was applied to minimize the least squares function and to find the parameters from an autocatalytic model for describing cure kinetics of thermosetting resins. Experimental results revealed that this method gives satisfactory results for parameter estimation.

Dirion et al.\textsuperscript{[13]} presented a general methodology to determine kinetic models of solid thermal decomposition with thermo gravimetric analysis (TGA) instruments. The goal is to determine a simple and robust kinetic model for a given solid with the minimum of TGA experiments. From this last point of view, this work can be seen as an attempt to find the optimal design of TGA experiments for kinetic modelling. Two computation tools were developed. The first is a nonlinear parameter estimation procedure for identifying parameters in nonlinear dynamical models. The second tool computes the thermogravimetric experiment (here, the programmed temperature profile applied to the thermobalance) required in order to identify the best kinetic parameters, i.e. parameters with a higher statistical reliability. The combination of the two tools can be integrated in an iterative approach generally called sequential strategy. The application concerns the thermal degradation of cardboard in a Setaram TGA instrument and the results that were presented demonstrate the improvements in the kinetic parameter estimation process.

However, recently artificial intelligence approach using artificial neural networks (ANNs) has been extensively used in process systems\textsuperscript{[14–16]} for modelling its unknown parameters\textsuperscript{[17]} as well as reaction kinetics.\textsuperscript{[18]} For a given set of inputs, ANNs are able to produce a corresponding set of outputs according to some mapping relationship. This relationship is encoded into the network structure during a period of training (also called learning), and is dependent upon the parameters of the network, i.e. weights and biases. Once the network has been trained (on the basis of known sets of input/output data), the input/output mapping is produced in a time, i.e. orders of magnitude lower than the time needed for rigorous deterministic modelling.\textsuperscript{[19]} This recent upsurge in research on neural networks has led to their popular application for identifying nonlinear processes and widely considered to be a powerful tool in the identification of highly nonlinear systems as they are able to learn from process data even when noise and uncertainty are present.

In our work, the complex kinetics characteristics of the batch chemical reactor model are captured using ANN models that can be embedded within the mechanistic model of the reactor. In this case, the high function approximation property of the ANN is exploited in improving the prediction capability of the mechanistic batch reactor polymerization model.

Hybrid modelling approach is different from the conventional clear or black-box models and should offer advantages over those methodologies. It has the merit of retaining the physical significance of the mechanistic model as well as incorporating the prediction efficiency of the black-box-based parameter models.\textsuperscript{[20]}

In the present work, a systematic experimental investigation was conducted and the collected temperature profile data were used in a parameter estimation algorithm to identify the major kinetic model parameters.
The estimated kinetic parameters were then utilized to develop a black-box ANN-based kinetic model with respect to reactor operating. Finally, the developed ANN kinetics model was combined with the first principles mechanistic model for predicting the dynamic transient of the polystyrene batch reactor system.

**MATHEMATICAL MODELLING**

In the first step, the free-radical polymerization of monomer styrene, solvent (toluene) and a monomer soluble initiator benzoyl peroxide (BPO) is modelled using a combined ANN-mechanistic modelling strategy. Traditionally, this process was modelled using a kinetic model derived from mass balances of the polymerization system with the reaction temperature. Further modification of this model involves considering the energy balance effects around the reactor and adding the relevant energy equations onto the kinetic model to get a more detailed model. In this work, the later modelling strategy is adopted, in addition to implementing ANN representation for the kinetic parameters.

**Kinetic mechanism**

The polystyrene polymerization is a free-radical reaction mechanism, i.e. three main reactions occur simultaneously, which involve the initiator, monomer, free-radical and polymer compounds. The initiation step is considered to involve two reactions. The first is the production of free radicals by any one of a number of reactions. The usual case is the homolytic dissociation of an initiator species $I$ to yield a pair of radicals $R$. The second part of the initiation involves the addition of this radical to the first monomer molecule to produce the chain initiating radical $P_1^-$. Propagation consists of the growth of $P_1^-$ by the successive additions of large numbers (hundreds and perhaps thousands) of monomer molecules. Each addition creates a new radical that has the same identity as the one previously, except that it is larger by one monomer unit. At some point, the propagating polymer chain stops growing and terminates. Termination occurs by bimolecular reaction between radicals. Two radicals react with each other by combination (coupling) or, more rarely, by disproportionation, in which a hydrogen radical that is beta to one radical centre is transferred to another radical centre. This results in the formation of two polymer molecules; one saturated and the other unsaturated. Termination can also occur by a combination of coupling and disproportionation. The term dead polymer signifies the cessation of growth for the propagating radical.$^{[21]}$

### Table 1. Kinetic mechanism of styrene polymerization.

<table>
<thead>
<tr>
<th>Reaction Type</th>
<th>Mechanism</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Initiation</strong></td>
<td>$I \xrightarrow{k_i} 2R^*$</td>
</tr>
<tr>
<td><strong>Propagation</strong></td>
<td>$P_n + M \xrightarrow{k_p} P_{n+1}$</td>
</tr>
<tr>
<td><strong>Termination by combination</strong></td>
<td>$P_n + P_m \xrightarrow{k_{td}} M_n + M_m$</td>
</tr>
<tr>
<td><strong>Termination by disproportionation</strong></td>
<td>$P_n + P_m \xrightarrow{k_{td}} M_n + M_m$</td>
</tr>
<tr>
<td><strong>Chain transfer to monomer</strong></td>
<td>$P_n + M \xrightarrow{k_{ts}} M_n + P_1$</td>
</tr>
<tr>
<td><strong>Chain transfer to solvent</strong></td>
<td>$P_n + S \xrightarrow{k_n} M_n + S^*$</td>
</tr>
<tr>
<td><strong>Chain transfer to transfer agent</strong></td>
<td>$P_n + T \xrightarrow{k_n} M_n + T^*$</td>
</tr>
<tr>
<td><strong>Chain transfer to polymer</strong></td>
<td>$P_n + M_m \xrightarrow{k_{tp}} M_n + P_m$</td>
</tr>
</tbody>
</table>

Although the chemical reality is much more complex, a simplified reaction scheme is often referred to in process studies. The complete mechanism can be written as given in Table 1.$^{[3]}$

**First principles model**

The theoretical description of the rate of polymerization is dependent on the assumed mechanism of polymerization and on the mathematical simplifications used to obtain analytical expressions. As the number of distinct reactions is increased, such as the various transfer reactions, the mathematical expressions can become quite complex. In general, the equations for the rate of polymerization are the most difficult to describe. Hence, it is necessary to make several assumptions for expressing the kinetic rate equations. The assumptions considered in developing the mechanistic model are as follows:

a. Quasi-steady-state approximation (QSSA) for live radicals and long-chain hypothesis (LCH) are valid.$^{[22]}$

b. All the reaction steps are irreversible.

c. Perfect mixing and constant-reacting heat capacity.

d. The jacket temperature is uniform and the heat losses with the ambient surrounding are negligible.

On the basis of the free-radical-initiated chain polymerization mechanism shown in Table 1 and the assumptions above, the reactor mass balance equation can be described as follows.$^{[4, 23]}$

The initiation reaction in polymerization is composed of two steps (Table 1) as discussed previously. The
second step (the addition of the primary radical to monomer) is much faster than the first step. The homolysis of the initiator is the rate-determining step in the initiation sequence, and the rate of initiation is then given by

\[
\frac{d(IV)}{dt} = -k_{d}IV \tag{1}
\]

\[
\frac{1}{V}\frac{d(RV)}{dt} = 2k_{d}I - k_{i}RM \tag{2}
\]

The rate of monomer disappearance as well as rate of polymerization is given by

\[
\frac{1}{V}\frac{d(MV)}{dt} = -k_{p}MP - k_{fim}MP \tag{3}
\]

For growing polymer

\[
\frac{1}{V}\frac{d(P_{1}V)}{dt} = 2k_{fim}M^{3} + k_{i}RM - k_{p}MP \\
+ (k_{fim} + k_{fis} + k_{fT})P \tag{4}
\]

\[
\frac{1}{V}\frac{d(P_{n}V)}{dt} = k_{p}M(P - P_{t}) - k_{fim}M + k_{fis} + k_{fT}P_{n} (n \geq 2) \tag{5}
\]

\[
\frac{1}{V}\frac{d(PV)}{dt} = -k_{p}MP - k_{fim}MP \tag{6}
\]

where \( P = \sum_{n=1}^{\infty} P_{n} \) is the total concentration of growing polymer.

The rate of dead polymer

\[
\frac{1}{V}\frac{d(M_{n}V)}{dt} = (k_{fim} + k_{fis} + k_{fT})P_{n} + k_{c} \sum_{m=1}^{n-1} P_{n-m}P_{m} (n \geq 2) \tag{7}
\]

From these rate expressions, one can derive the following equations:

\[
P = \left\{ \frac{2k_{d}I}{k_{c} + k_{d}} \right\}^{1/2} \tag{8}
\]

\[
P_{1} = (1 - \alpha)P \tag{9}
\]

\[
P_{n} = (1 - \alpha)P \alpha^{n-1} \tag{10}
\]

where

\[
\alpha = \frac{k_{p}M}{k_{p}M + k_{fim} + \alpha k_{T} + (k_{c} + k_{d})P} \tag{11}
\]

So, the final equation for dead polymer

\[
\frac{1}{V}\frac{d(M_{n}V)}{dt} = [k_{fim}S + k_{fim}M + k_{fT} + k_{d}P] \tag{12}
\]

\[
(1 - \alpha)P \alpha^{n-1} + \frac{1}{2}k_{c}P^{2} \\
(1 - \alpha)^{2} \alpha^{n-2}(n - 1)
\]

where \( M_{n}(0) = M_{n0}; \ n \geq 2 \).

Zeroth moment of dead polymer

\[
\frac{1}{V}\frac{d(\xi_{0}V)}{dt} = (k_{fim} + k_{fim} + k_{fT} + k_{d}P) \tag{13}
\]

\[
\alpha P + \frac{1}{2}k_{c}P^{2}
\]

First moment of dead polymer

\[
\frac{1}{V}\frac{d(\xi_{1}V)}{dt} = \frac{P}{1 - \alpha} \tag{14}
\]

\[
\{ [k_{fim} + k_{fim} + k_{fT} + k_{d}P] \}
\]

\[
\alpha(2 - \alpha) + k_{c}P \tag{15}
\]

Second moment of dead polymer

\[
\frac{1}{V}\frac{d(\xi_{2}V)}{dt} = \frac{P}{(1 - \alpha)^{2}} \tag{16}
\]

\[
2\alpha[k_{fim} + k_{fim} + k_{fT} + k_{d}P]
\]

\[
+ k_{c}P(2\alpha + 1) + \frac{1}{V}\frac{d(\xi_{1}V)}{dt} \tag{17}
\]

The kinetic coefficients are calculated with Arrhenius temperature dependency and they are expressed as \( k_{i} = A_{i} \times \exp[-E_{i} / (RT)] \) and the gel effect is considered\(^{34}\) as follows:

\[
g_{t} = \frac{k_{i}}{k_{i0}} = \exp[-2(BX + CX^{2} + DX^{3})] \tag{18}
\]

where \( X \) and \( 0k_{i0} \) denote the monomer conversion and the termination rate constant at zero monomer conversion, respectively, and

\[
B = 2.57 - 5.05 \times 10^{-3} \text{T (K)} \tag{19}
\]

\[
C = 9.56 - 1.76 \times 10^{-2} \text{T (K)} \tag{20}
\]

\[
D = -3.03 + 7.85 \times 10^{-3} \text{T (K)} \tag{21}
\]

Due to the density difference between polymer and monomer, the reaction volume of the polymerizing medium changes during the polymerization. Thus, the rate of volume change is described by the following equation:

\[
\frac{1}{V}\frac{dV}{dr} = -\left( \frac{\epsilon}{M_{b} + \epsilon M} \right) \frac{dM}{dr} \tag{22}
\]
where $M_0$ is the initial concentration of monomer and $\varepsilon$ the volume contraction factor defined by

$$\varepsilon = \frac{V_{X=1} - V_{X=0}}{V_{X=0}}$$  \hspace{1cm} (21)

and the densities of styrene monomer and polymer described are as follows\(^\text{24}\):

$$\rho_m = 924 - 0.918(T - 273.1)$$ \hspace{1cm} (22)

$$\rho_p = 1084.8 - 0.605(T - 273.1)$$ \hspace{1cm} (23)

The reactor is presumed to be perfectly well mixed. According to the energy balance of reactor contents, the reactant temperature depends on the following equations:

Reactor dynamics:

$$\frac{dT}{dt} = \frac{Q + (-\Delta H)R_m V - UA(T - T_j)}{V \rho C_p} - \frac{T}{V} \frac{dV}{dt}$$ \hspace{1cm} (24)

Jacket dynamics:

$$\frac{dT_{jo}}{dt} = \frac{M_c C_p c(T_{ji} - T_{jo}) + UA(T - T_j)}{V_c \rho_c C_p}$$ \hspace{1cm} (25)

where $R_m = \frac{dM}{dt}$; $T_j = \frac{T_{ji} + T_{jo}}{2}$. The overall rate of heat production by the reaction:

$$R_H = (-\Delta H) \frac{dM}{dt}$$ \hspace{1cm} (26)

The number of average chain length and conversion can be completely determined in terms of operation variables, i.e.

Conversion:

$$X = \frac{M_o - M}{M_o}$$ \hspace{1cm} (27)

Number average chain length:

$$X_n = \frac{M_o - M}{\xi_0}$$ \hspace{1cm} (28)

As the polymerization reaction proceeds, the concentration of the polymer chain increases. This increase in the polymer concentration primarily causes the viscosity of reacting mixture to increase significantly. Consequently, the notable increase results in a sharp decrease of the overall heat transfer coefficient and these changes give a sort of disturbance to batch polymerization reactors. Unless stated, an initial constant overall heat transfer coefficient is used in the polymerization model. In literature, the decrease of the overall heat transfer coefficient has been modelled through the use of empirical correlations. Although several correlations have been reported for the heat transfer coefficient in the literature, the one recently proposed by Erdogan et al.\(^\text{25}\) will be adopted in this work as follows:

$$U(X) = U(0) - \alpha X$$ \hspace{1cm} (29)

where $U(0)$ is the coefficient at $X_n = 0$ and $X$ is the percent conversion.

**BATCH POLYSTYRENE REACTOR SYSTEM**

A schematic diagram of the experimental batch polymerization reactor\(^\text{4,26}\) can be seen in Fig. 1. A 2000-ml jacketed glass reactor was used with 12 cm inside diameter and 20 cm depth. The working volume capacity of the reactor is 1.5 l. Thermocouples were used for measuring the reactor, jacket inlet and outlet temperature. The mixture inside the reactor is stirred using a 25-mm diameter turbine agitator located 7 cm above the base of the reactor. The motor speed of the agitator can be manually adjusted in the range 50–2000 rpm. Intensive bubbling of nitrogen was applied from the bottom of reactor to remove dissolved oxygen. The vaporized toluene was collected by a reflux condenser in the system. A heater of 500 W is used to provide the required heating energy to the reaction medium. The apparatus is equipped with data acquisition and logging system to monitor temperatures and heating load.

In this study, styrene (99.9 mol%), toluene (99.9 mol%) and benzoyl peroxide (BPO) from Sigma–Aldrich were used as monomer, solvent and initiator, respectively. The inhibited styrene was activated by removing the inhibitor using an inhibitor remover (aluminium oxide) from Sigma–Aldrich. The

![Figure 1. Experimental setup for the polystyrene batch reactor. This figure is available in colour online at www.apjChemEng.com.](image-url)
PROPOSED METHOD FOR MODELLING POLYSTYRENE REACTOR

Psichogios and Ungar\(^{[27]}\) have introduced the idea of using first principles-neural network methodology for modelling chemical processes. Such a methodology attempts to utilize all accessible process knowledge possible by implementing black-box correlations for predicting process parameters. A similar methodology was considered in this work for modelling the polystyrene batch reactor. This methodology includes the following steps:

i. Develop a first principles model involving temperature varying kinetic parameters.

ii. Generate reactor temperature profiles by conducting several reaction runs using predefined operating conditions.

iii. Use the generated temperature profiles and the mechanistic model to estimate the kinetic parameters.

iv. Use ANN modelling methodology to model the estimated kinetic parameters with reactor operating conditions.

v. Combine the trained ANN-based kinetic model into the mechanistic model and validate it with experimental profiles generated in step (ii).

Parameter estimation

A nonlinear least squares optimization algorithm was adopted to estimate the reaction kinetic parameters.

Experimental temperature profiles were used for estimating the kinetic parameters. The parameters are evaluated using an optimization search algorithm. At each iteration, the model output is consolidated with the experimental data for a given set of inputs and the sum square of the deviations is optimized.

Nonlinear least squares is the form of least squares analysis which is used to fit a set of \( m \) observations with a model that is nonlinear in \( n \) unknown parameters \((m > n)\). Consider a set of \( m \) data points, \((x_1, y_1), (x_2, y_2), \ldots, (x_m, y_m)\), and a curve (model function) \( y = f(x, \beta) \), that in addition to the variable \( x \) also depends on \( n \) parameters, \( \beta = (\beta_1, \beta_2, \ldots, \beta_n) \) with \( m \geq n \). It is desired to find the vector \( \beta \) of parameters such that the curve fits best the given data in the least squares sense, i.e. the sum of squares

\[
S = \sum_{i=1}^{m} r_i^2
\]

is minimized, where the residuals (errors) \( r_i \) are given by

\[
r_i = y_i - f(x_i, \beta) \quad \text{for} \quad i = 1, 2, \ldots, m.
\]

The minimum value of \( S \) occurs when the gradient is zero. Since the model contains \( n \) parameters, there are \( n \) gradient equations:

\[
\frac{\partial S}{\partial \beta_j} = 2 \sum_i r_i \frac{\partial r_i}{\partial \beta_j} = 0, \quad (j = 1, \ldots, n)
\]

In a nonlinear system, the derivatives \( \frac{\partial r_i}{\partial \beta_j} \) are functions of both the independent variable and the parameters, so these gradient equations do not have a closed solution. Instead, initial values must be chosen for the parameters. Then, the parameters are refined iteratively, i.e. the values are obtained by successive approximation,

\[
\beta_j \approx \beta_j^{k+1} = \beta_j^k + \Delta \beta_j
\]

Here, \( k \) is an iteration number and the vector of increments, \( \Delta \beta_j \) is known as the shift vector. At each iteration, the model is linearized by approximation to a first-order Taylor series expansion about \( \beta^k \)

\[
f(x_i, \beta) \approx f(x_i, \beta^k) + \sum_j \frac{\partial f(x_i, \beta^k)}{\partial \beta_j} (\beta_j - \beta_j^k)
\]

\[
\approx f(x_i, \beta^k) + \sum_j J_{ij} \Delta \beta_j
\]

The Jacobian, \( J \), is a function of constants, the independent variable and the parameters, so it changes from
one iteration to the next. Thus, in terms of the linearized model, \( \frac{\partial r_i}{\partial \beta_j} = -J_{ij} \) and the residuals are given by

\[
r_i = \Delta y_i - \sum J_{ij} \Delta \beta_j; \quad \Delta y_i = y_i - f(x_i, \beta^k)
\]

Substituting these expressions into the gradient equations, it becomes

\[
-2 \sum_{i=1}^{m} J_{ij} \left( \Delta y_i - \sum_{s=1}^{n} J_{is} \Delta \beta_s \right) = 0
\]

which, on rearrangement, become \( n \) simultaneous linear equations, the normal equations

\[
\sum_{i=1}^{m} \sum_{j=1}^{n} J_{ij} J_{is} \Delta \beta_s = \sum_{i=1}^{m} J_{ij} \Delta y_i, \quad (J = 1, n)
\]

The normal equations are written in matrix notation as

\[
(J^T J) \Delta \beta = J^T \Delta y
\]

When the observations are not equally reliable, a weighted sum of squares may be minimized,

\[
S = \sum_{i=1}^{m} W_{ii} r_i^2
\]

Each element of the diagonal weight matrix, \( W \) should, ideally, be equal to the reciprocal of the variance of the measurement. The normal equations are then

\[
(J^T W J + \lambda I) \Delta \beta = (J^T W \Delta y)
\]

These equations form the basis for the Levenberg–Marquardt algorithm for nonlinear least squares problems. In Eqn (40), \( \lambda \) is the Marquardt parameter and \( I \) is an identity matrix. The Levenberg–Marquardt (LM) algorithm is an iterative technique that locates the minimum of a function that is expressed as the sum of squares of nonlinear functions. It has become a standard technique for nonlinear least squares problems and can be thought of as a combination of steepest descent and the Gauss–Newton method.[28]

**Artificial neural networks design**

Commonly neural networks are adjusted or trained so that a particular input leads to a specific target output. The neural networks can be trained for complex mappings because the hidden layer elements learn to respond to characteristics found in the input. This refers to correlations of activities among different input spots, allowing an input information abstract representation in the hidden layers. They have the abstraction and generalization abilities, and capable of classifying a complex pattern correctly, even if it does not belong to the network training set. They are also immune to little noises present in the inputs.

A schematic diagram of a feedforward neural network is shown in Fig. 2. In the architecture, the neurons are connected to all posterior layer neurons, with the information from an anterior layer suffering a ponderation by a weight \( (w_{ij}) \) that is sent to all neurons of the next layer.

Backpropagation training algorithm refers to the way the weights are adjusted. This way is also known as the general delta rule, based on the descending gradient optimization and has been used in the majority of the works applied to chemical processes.[29]

Processing elements of the same layer act in parallel and layer-to-layer processing is sequential. The equations that administrate the feedforward processing are

\[
s_j^{(k)} = w_{ij}^{(k)} + \sum_{i=1}^{N_k} w_{ij}^{(k)} x_i^{(k-1)}
\]

\[
x_j^{(k)} = f(s_j^{(k)})
\]

In this relation, \( x_i^{(k)} \) refers to the activating function input of \( k \) layer \( i \) element, \( s_j^{(k)} \) refers to the pondered sum of the weights through the inputs and \( w_{ij}^{(k)} \) refers to the synaptic connection weights at the \( k \)th layer \( j \)th element input, where \( i \) is the connection index and \( N_k \) is the \( k \)th layer processing element number.

In the input, \( x_i = x_i^{(0)} \) are the components of the input vector \( X \) and, in the output, \( y_i = y_i^{(m)} \) are the components of the output vector \( Y \). The feedforward neural network input and output neurons can be related by the sigmoid transfer functions, given by

\[
y_i = f(s_j) = \frac{1}{1 + e^{-s_j}}
\]

**Figure 2.** Structure of a feedforward neural network.
However, other transfer functions can be used depending on the characteristics of the problem being studied. The purpose of training a network is the adjustment of its weights in such a way that the application of a pattern produces an output value and in this sense the general delta rule aims to reduce the network quadratic error indicated by

$$
\varepsilon = \sum_{j=1}^{m} (d_j - y_j)^2 \quad (44)
$$

In Eqn (44), \(d_j\) stands for the experimental or real value and \(y_j\) represents the value predicted from the neural model.

In this work, ANN was used for correlating kinetic parameters with reactor operating variables. The three operating conditions were used as inputs, whereas the estimated kinetic parameters were used as outputs for training the neural network.

**First principles-neural networks model**

The conventional approach for chemical process modelling is based on the mass and energy balance equations. This form of modelling requires further knowledge about reaction kinetics. For polymer processes, the predictive ability of conventional model is quite limited. This is mainly and simply due to the nonlinear time varying characteristics of the polymerization reactions, with kinetic structures which often are partially known, or even completely unknown. In first principles-neural network approach, the poorly known or unknown parameters such as kinetic parameters is modelled using artificial neural network which is combined with first principles model.

With this approach, the conventional mechanistic model is used to describe mass and heat transfer phenomena, while ANNs are utilized to predict to kinetic parameters. The use of neural networks in this case is justified by the complex and highly nonlinear nature of the reaction mechanism. In the previously described process model, the six kinetic parameters (\(A_d\), \(A_p\), \(A_t\), \(E_d\), \(E_p\) and \(E_t\)) are calculated from pretrained ANNs with high prediction capabilities. A schematic diagram representing the structure of such a model is shown in Fig. 3.

**RESULTS AND DISCUSSION**

From a practical point of view, the most effective operating variables for this process are the reactor initial temperature (\(T_o\)), initial initiator concentration (\(I_o\)) and heater load (\(Q\)).

<table>
<thead>
<tr>
<th>Variables</th>
<th>(T_o) (K)</th>
<th>(I_o) (mol/l)</th>
<th>(Q) (W)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Run 1</td>
<td>360.00</td>
<td>0.016</td>
<td>100.00</td>
</tr>
<tr>
<td>Run 2</td>
<td>356.00</td>
<td>0.016</td>
<td>100.00</td>
</tr>
<tr>
<td>Run 3</td>
<td>360.00</td>
<td>0.024</td>
<td>100.00</td>
</tr>
<tr>
<td>Run 4</td>
<td>356.00</td>
<td>0.024</td>
<td>100.00</td>
</tr>
<tr>
<td>Run 5</td>
<td>360.00</td>
<td>0.016</td>
<td>90.00</td>
</tr>
<tr>
<td>Run 6</td>
<td>356.00</td>
<td>0.016</td>
<td>90.00</td>
</tr>
<tr>
<td>Run 7</td>
<td>360.00</td>
<td>0.024</td>
<td>90.00</td>
</tr>
<tr>
<td>Run 8</td>
<td>356.00</td>
<td>0.024</td>
<td>90.00</td>
</tr>
<tr>
<td>Run 9</td>
<td>358.00</td>
<td>0.016</td>
<td>93.00</td>
</tr>
<tr>
<td>Run 10</td>
<td>358.00</td>
<td>0.018</td>
<td>93.00</td>
</tr>
<tr>
<td>Run 11</td>
<td>358.00</td>
<td>0.021</td>
<td>93.00</td>
</tr>
<tr>
<td>Run 12</td>
<td>358.00</td>
<td>0.024</td>
<td>93.00</td>
</tr>
<tr>
<td>Run 13</td>
<td>360.00</td>
<td>0.016</td>
<td>96.00</td>
</tr>
<tr>
<td>Run 14</td>
<td>360.00</td>
<td>0.018</td>
<td>96.00</td>
</tr>
<tr>
<td>Run 15</td>
<td>360.00</td>
<td>0.021</td>
<td>96.00</td>
</tr>
<tr>
<td>Run 16</td>
<td>360.00</td>
<td>0.024</td>
<td>96.00</td>
</tr>
</tbody>
</table>

Investigation. Three levels for the variables initial reactor temperature \(T_o\) (\(83, 85\) and \(87\)). However, four measurement levels were considered for the heater load \(Q\) (\(90, 93, 96\) and \(100\)) and the initiator concentration \(I_o\) (\(0.016, 0.018, 0.021\) and \(0.024\)). At each
The polymerization process mechanistic model described above was simulated and solved using Matlab simulation environment. The reactor operating conditions and design parameters used in the simulations are given in Table 3.

Parameter estimation analysis was then conducted. This was done by reconciling the experimental temperature profile data against the model predicted profiles. A nonlinear least square algorithm was used to optimize the model performance with the experimental temperature profiles. The estimated kinetic parameters involve frequency factors and activation energies for the decomposition, propagation and termination reaction steps \(A_d, A_p, A_t, E_d, E_p\) and \(E_t\). Table 4 gives the estimated values for these parameters as well as the mean square error (MSE) of estimation for the first eight runs. The estimation errors are relatively small and acceptable for all runs.

The ANN design was started by selecting a proper architecture which is not large in size and that gives good prediction capabilities. The preliminary selected design was a three-layer network with three neurons in the input layer, six tangent sigmoidal neurons in the hidden layer and six linear neurons output layer. The three operating conditions were used as inputs, whereas the estimated kinetic parameters \((A_d, E_d)\) were used as outputs for training the neural network. Figure 5 depicts the estimated kinetic parameter data which were used for ANN training testing and validation. The total data set size used was 48. These data were generated for the three tested variables (three initial reactor temperature

### Table 3. Operating conditions and reactor specifications.

<table>
<thead>
<tr>
<th>Name of the parameters</th>
<th>Symbol</th>
<th>Value</th>
<th>Units</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reactant specific heat</td>
<td>(C_p)</td>
<td>1.96886</td>
<td>J/g K</td>
</tr>
<tr>
<td>Coolant specific heat</td>
<td>(C_{pc})</td>
<td>4.29</td>
<td>J/g K</td>
</tr>
<tr>
<td>Heat of reaction, exothermic</td>
<td>(H)</td>
<td>-57766.8</td>
<td>J/g K</td>
</tr>
<tr>
<td>Coolant flowrate</td>
<td>(m_c)</td>
<td>0.51</td>
<td>g/s</td>
</tr>
<tr>
<td>Gas constant</td>
<td>(R)</td>
<td>8.314</td>
<td>J/mol K</td>
</tr>
<tr>
<td>Coolant inlet temperature</td>
<td>(T_{bi})</td>
<td>303.14</td>
<td>K</td>
</tr>
<tr>
<td>Overall heat transfer coefficient</td>
<td>(U)</td>
<td>55.1</td>
<td>W/(m² K)</td>
</tr>
<tr>
<td>Reactor volume</td>
<td>(V)</td>
<td>1.2</td>
<td>l</td>
</tr>
<tr>
<td>Reactor jacket volume</td>
<td>(V_c)</td>
<td>1</td>
<td>l</td>
</tr>
<tr>
<td>Coolant density</td>
<td>(\rho_c)</td>
<td>998.00</td>
<td>g/l</td>
</tr>
<tr>
<td>Reactant density</td>
<td>(\rho_i)</td>
<td>983.73</td>
<td>g/l</td>
</tr>
</tbody>
</table>

### Table 4. Estimated kinetic parameters for different experimental runs.

<table>
<thead>
<tr>
<th>Run</th>
<th>(A_d)</th>
<th>(A_p)</th>
<th>(A_t)</th>
<th>(E_d)</th>
<th>(E_p)</th>
<th>(E_t)</th>
<th>MSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>(1.82 \times 10^{16})</td>
<td>(3.66 \times 10^6)</td>
<td>(1.77 \times 10^9)</td>
<td>(1.397 \times 10^5)</td>
<td>(2.76 \times 10^4)</td>
<td>(2.10 \times 10^4)</td>
<td>0.44</td>
</tr>
<tr>
<td>2</td>
<td>(1.81 \times 10^{16})</td>
<td>(1.11 \times 10^6)</td>
<td>(1.59 \times 10^9)</td>
<td>(1.393 \times 10^5)</td>
<td>(2.78 \times 10^4)</td>
<td>(1.93 \times 10^4)</td>
<td>0.45</td>
</tr>
<tr>
<td>3</td>
<td>(4.65 \times 10^{16})</td>
<td>(6.43 \times 10^7)</td>
<td>(2.01 \times 10^9)</td>
<td>(1.412 \times 10^5)</td>
<td>(3.04 \times 10^4)</td>
<td>(6.00 \times 10^3)</td>
<td>0.62</td>
</tr>
<tr>
<td>4</td>
<td>(3.17 \times 10^{16})</td>
<td>(8.12 \times 10^6)</td>
<td>(1.40 \times 10^9)</td>
<td>(1.414 \times 10^5)</td>
<td>(2.72 \times 10^4)</td>
<td>(1.89 \times 10^4)</td>
<td>1.08</td>
</tr>
<tr>
<td>5</td>
<td>(2.34 \times 10^{16})</td>
<td>(1.06 \times 10^7)</td>
<td>(1.57 \times 10^9)</td>
<td>(1.399 \times 10^5)</td>
<td>(2.89 \times 10^4)</td>
<td>(6.60 \times 10^3)</td>
<td>1.21</td>
</tr>
<tr>
<td>6</td>
<td>(2.12 \times 10^{16})</td>
<td>(1.25 \times 10^7)</td>
<td>(2.62 \times 10^9)</td>
<td>(1.400 \times 10^5)</td>
<td>(2.77 \times 10^4)</td>
<td>(5.57 \times 10^3)</td>
<td>0.76</td>
</tr>
<tr>
<td>7</td>
<td>(3.31 \times 10^{16})</td>
<td>(1.20 \times 10^7)</td>
<td>(1.40 \times 10^9)</td>
<td>(1.397 \times 10^5)</td>
<td>(2.92 \times 10^4)</td>
<td>(5.24 \times 10^3)</td>
<td>1.13</td>
</tr>
<tr>
<td>8</td>
<td>(4.14 \times 10^{16})</td>
<td>(1.97 \times 10^7)</td>
<td>(1.43 \times 10^9)</td>
<td>(1.399 \times 10^5)</td>
<td>(2.94 \times 10^4)</td>
<td>(5.03 \times 10^3)</td>
<td>0.93</td>
</tr>
</tbody>
</table>

MSE, mean square error.

levels × four heater load levels × four initiator concentration levels). This set was subdivided for training, testing and validation in a ratio of 2 : 1 : 1. However, our exploratory experiments showed that having a three-input six-output network structure does not guarantee the network prediction quality, hence three ANNs were designed with three inputs and two kinetic parameters (one frequency factor and one activation energy for each of initiation, propagation and termination mechanisms) as outputs. This reduction in the complexity of the ANN structure resulted in a much better prediction quality. The design specifications for neural network model are given in Table 5. The MSE estimates for the three ANNs indicate very good degree of

Figure 5. Kinetic parameters ANN training data sets (training data: 1–24, testing data: 25–36 and validation data: 37–48).
networks predictions. The three trained ANNs represent the kinetic model of the polymerization process. The prediction capability of this model is demonstrated by plotting the experimental values of the kinetic parameters vs the ANNs predictions as shown in Fig. 6. The Pearson correlation coefficient\[30\] was used to verify the quality of the results obtained. The correlation coefficients of the six parameters 0.996, 0.984, 0.998, 0.997, 0.977 and 0.983, which are above 0.97, indicate a high prediction quality. This is further indicated by
plotting the residual of the model temperature prediction from the experimental temperature profiles for the eight runs. Fig. 7 shows that the reactor temperature residual ($T_{\text{expt}} - T_{\text{predicted}}$) profiles are of reasonable variability and low values. The trained ANNs were implemented within the mechanistic model structure as explained earlier. The experimental runs were simulated using the developed hybrid model. The experimental temperature profiles and the model predicted ones are shown in Fig. 8. It is clear that the new model predictions follow the experimental temperature trend for all runs. Experimentally determined reactor conversions are compared with model prediction as shown in Fig. 9. The model predictions are in good agreement with the experimentally determined profiles.

To check the efficiency of the new model as compared with four published model, the kinetic parameters reported by previous studies (as given in Table 6) were simulated and compared with the new hybrid model. Figure 10 shows that the current model attained the closest predictions to the experimental temperature profiles. The model of Meyer was the best among the other four models, whereas that of Odian was the worst. Generally speaking, although the predicted trend was correct, all previous models failed to explain the temperature experimental profiles efficiently. However,
### Table 6. Literature reported kinetic parameters values for styrene polymerization reaction.

<table>
<thead>
<tr>
<th>Kinetic parameters</th>
<th>( A_d )</th>
<th>( A_p )</th>
<th>( A_t )</th>
<th>( E_d )</th>
<th>( E_p )</th>
<th>( E_t )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alphaz et al. [4]</td>
<td>( 2.60 \times 10^{16} )</td>
<td>( 1.05 \times 10^7 )</td>
<td>( 1.26 \times 10^9 )</td>
<td>143 161</td>
<td>29 553</td>
<td>7023</td>
</tr>
<tr>
<td>Novakovic et al. [31]</td>
<td>( 4.59 \times 10^{13} )</td>
<td>( 7.47 \times 10^7 )</td>
<td>( 2.67 \times 10^8 )</td>
<td>122 400</td>
<td>34 460</td>
<td>2084</td>
</tr>
<tr>
<td>Meyer et al. [32]</td>
<td>( 2.58 \times 10^{16} )</td>
<td>( 4.27 \times 10^7 )</td>
<td>( 2.20 \times 10^9 )</td>
<td>141 152</td>
<td>32 500</td>
<td>6500</td>
</tr>
<tr>
<td>Odian [21]</td>
<td>( 6.70 \times 10^{13} )</td>
<td>( 0.45 \times 10^7 )</td>
<td>( 5.80 \times 10^8 )</td>
<td>124 000</td>
<td>260 00</td>
<td>8000</td>
</tr>
</tbody>
</table>

Figure 10. Performance of this work model with published models. This figure is available in colour online at www.apjChemEng.com.

The hybrid-based model developed in this work achieved a very high accuracy of prediction.

### CONCLUSION

Polymerization reactors operating in batch mode are difficult to model and simulate. This is due to the complex and partially understood kinetic reaction mechanism involved in their operation. Consequently, there is a great need for developing simple and reliable kinetic models. In this work, an ANN modelling strategy was used in combination with conventional mechanistic modelling methodologies to explain the complex dynamics of a batch polymerization reactor.

An optimization-based parameter estimation technique was used to calculate the kinetic parameters of styrene polymerization batch reaction from several experimental reactor temperature profiles. The calculated kinetic parameters were then used to train the ANNs. The three trained ANNs representing the correlation between operating variables and the six kinetic parameters attained a high degree of prediction accuracy. The current hybrid model is superior in prediction compared with published models. The resulting hybrid model can be utilized for the purpose of design, analysis and control of this and similar polymerization reactors.

### NOMENCLATURE

- \( A \) heat transfer area (m\(^2\))
- \( A_d \) initiator decomposition frequency factor (1/s)
- \( A_p \) propagation frequency factor (1/mol s)
- \( A_t \) termination frequency factor (1/mol s)
- \( C_p \) specific heat of reactor mixer (J/g K)
- \( C_{pc} \) specific heat of jacketed water (J/g K)
- \( E_d \) initiator decomposition activation energy (J/mol)
- \( E_p \) propagation activation energy (J/mol)
- \( E_t \) termination activation energy (J/mol)
- \( f \) initiator efficiency
- \( I \) initiator concentration, initial initiator concentration (mol/l)
- \( (-\Delta H) \) heat of reaction (J/mol)
- \( M \) monomer concentration (mol/l)
- \( M_o \) initial monomer concentration (mol/l)
- \( M_n \) dead polymer (mol/l)
- \( X \) conversion (%)
- \( M_c \) coolant flow rate (g/s)
- \( Q \) heat given from the heater (W)
- \( R_m \) rate of polymerization (mol/s)
- \( T \) reactor temperature (K)
- \( T_j \) Jacket temperature (K)
- \( T_{ji} \) Jacket inlet temperature (K)
- \( T_{jo} \) Jacket outlet temperature (K)
- \( U \) overall heat transfer coefficient [W/(m\(^2\) K)]
- \( V \) volume of the reactor (l)
- \( V_c \) Jacket volume (l)
- \( v \) constant
- \( \rho \) density of reactor mixer (g/l)
- \( \rho_c \) density of water (g/l)
- \( \rho_m \) density of monomer (g/l)
- \( \rho_p \) density of polymer (g/l)
- \( \xi_i \) moment of dead polymer (mol/l)

### Subscripts

- \( o \) initial condition
- \( c \) coolant

© 2010 Curtin University of Technology and John Wiley & Sons, Ltd.

DOI: 10.1002/apj
REFERENCES
