Hybrid Approach to Modeling an Industrial Polyethylene Process

Mark Hinchliffe, Gary Montague, and Mark Willis
School of Chemical Engineering and Advanced Materials, University of Newcastle, Newcastle, UK.

Annette Burke
NOVA Chemicals Corporation, Calgary, Canada

A hybrid model of a polyethylene production process is developed. The mechanistic model utilizes fundamental material and energy balances to predict important process conditions, such as the reactor temperatures, conversions, and the molecular-weight distribution (MWD) of the polymer. Using plant data, it is shown that accurate MWD predictions are not obtained from the mechanistic model alone, despite efforts to accurately model the system and improve the accuracy of the input data. Because an accurate prediction of the MWD is required to predict end-use properties, a hybrid model was developed by adding an empirical layer to the mechanistic model. The empirical layer was developed by using an optimization algorithm to adjust the predicted MWD by manipulating multipliers of the key descriptors (states or functions of states) of the distribution. These multipliers were then predicted from plant data using feedforward artificial neural networks (FANNs). They are then combined with the mechanistic model to allow accurate MWD prediction.

Introduction

Improved product quality is one of the primary reasons for the use of advanced control in the polymer industry. Off-spec material must be sold at a reduced price, blended with other material, or wasted, resulting in reduced profits. The quality of a polymer resin is normally defined in terms of resin properties, such as the melt index (MI or I_2), molecular-weight distribution (MWD), and density. Unfortunately, these properties are measured off-line at infrequent intervals and there are substantial delays between the sampling and analysis of results. Delays in obtaining resin measurements make process control difficult and can result in significant quality variations. One solution is to improve on-line measurements, for example, by using viscosity or spectroscopic techniques to infer resin characteristics in the reactor outlet stream (Kiparissides and Morris, 1996). However, they are not direct measurements of resin properties and the interpretation of complex spectra can be problematic. An alternative is to infer quality variables using a dynamic model to predict the effect of changes in the reactor conditions upon the quality. The dynamic model can be based upon the mathematics–physics–chemistry of the process, be an empirical model developed using plant data, or a combination of both of these techniques.

There have been a number of notable advances in the area of control, monitoring, and modeling of polymerization reactors (Lines et al., 1992; Crowley and Choi, 1998; Ogawa et al., 1999; Seki et al., 2001), and there are a number of good reviews in the literature (MacGregor et al., 1984; Ray, 1985; MacGregor, 1986; Kiparissides and Morris, 1996; Kiparissides, 1996). Mechanistic models are now available to predict process conditions such as the reactor temperature and MWD. Although the methods used for developing these models are relatively well known, it is still difficult to reduce process model mismatch to obtain the needed degree of accuracy.

If a link is to be made to end-use properties, accurate predictions must be made not only of molecular-weight averages, but the entire MWD distribution. Predictions of resin processing and end-use properties can be based on manufacturing conditions and parameters that describe the MWD (Ohshima and Tanigaki, 2000; Latado et al., 2001), but Ari-

Correspondence concerning this article should be addressed to G. Montague.

awan et al. (2001) recently demonstrated that it may be more useful to consider the entire distribution. Ariawan et al. (2001) found that extensional and elastic properties of resins were dependent on the concentration of larger molecules, and in particular on the high molecular-weight tail of the distribution. Reducing process model mismatch for polymerization models is difficult in itself. Predicting the details of the entire MWD distribution with a high degree of accuracy is particularly challenging.

In this work, we develop a hybrid model that is a combination of mechanistic and empirical approaches for predicting the MWD from reactor inputs. A mechanistic description is used to maintain process understanding and model portability, while an empirical model is used to decrease process–model mismatch, which arises in the prediction of the MWD.

As with all modeling techniques, there are a number of options available for the development of a hybrid model. One approach is to couple artificial neural networks (ANNs), an empirical modeling technique, with a mechanistic model. Psichogios and Ungar (1992) proposed this approach to model batch and fed-batch fermentation processes. The ANN was used to model variations in the specific growth rate, which was then utilized as a parameter in the mechanistic model (adjusting mechanistic model parameters leads to a so-called serial hybrid model). Kramer et al. (1992) used a radial basis function (RBF) network to model the residual errors between actual plant data and the outputs from a fitted theoretical model (this is often referred to as a parallel-hybrid model). Schubert et al. (1994) considered a more sophisticated hybrid model, which incorporated a mathematical model in the form of balance equations and an ANN with fuzzy logic supervision for batch and fed-batch bioreactor. Tsen et al. (1996) developed a hybrid modeling strategy for a batch polymerization reactor. They utilized an augmented data set that integrated experimental data with a theoretical model to develop an accurate neural-network prediction. McKay et al. (1998) evolved a hybrid model of a fed-batch fermentation process. Simple mass-balance expressions were used for the mechanistic model, and genetic programming was used to find empirical expressions for the required reaction rates. Qi et al. (1999) combined a mechanistic model of a wall-cooled chemical reactor with a neural-network model. The neural-network was used to predict important process parameters, such as the overall heat-transfer coefficient. The hybrid model was shown to have greater accuracy than conventional neural networks. Furthermore, the model was found to be more reliable for extrapolation because of the mechanistic basis.

In this contribution a detailed mechanistic model of a polyethylene production process is developed. The model utilizes fundamental material and energy balances to predict important process conditions such as the reactor temperatures, conversions, and the MWD. Accurate prediction of MWD is required in order to fully characterize the polymer, and, therefore, predict the polymer properties. However, it is shown that accurate MWD predictions are not obtained using the mechanistic model. To overcome this problem, a feedforward artificial neural network (FANN) is used to correct the MWDs by adjusting model states (or functions of the states) that affect MWD (this is a serial approach to hybrid model development). This results in a more accurate MWD prediction, which may be used to estimate end-use properties and improve process control.

Process Description

The polyethylene process is a solution polyolefin technology consisting of two reactors in series (see Figure 1). The reactor system can be configured and operated differently so as to obtain specific resin characteristics. Ethylene and comonomer are fed to the first reactor along with the catalyst and hydrogen. The product from the first reactor is fed to the second reactor, with a second feed stream of ethylene, comonomer, catalyst, and hydrogen. The feeds to each reactor can be controlled independently, providing the flexibility to tailor resin properties. The reactors are operated at moderately high temperatures and high pressures. Process disturbances arise because of impurities in the feeds and variations in catalyst quality.

Traditionally, resin properties that are controlled are Melt Index (MI or I_2), density, and a measure of MWD breadth. However, given the flexibility in a dual-reactor process to tailor the distributional properties of the resin, these may no longer be sufficient. The development of on-line sensors could improve the situation by making additional information available in real-time. These would be complemented by the use of a mathematical model to predict difficult-to-measure process variables.

Polymer reactor model

The reactor system model can be configured as either single (reactor 2 alone) or dual reactors in series. In practice, both of these reactor configurations are used, depending on the required resin characteristics. Each reactor is modeled using a number of perfectly mixed "internal" continuous stirred-tank reactors (CSTRs) in series (see Figure 2.). A fresh feed stream is supplied to each reactor. The effluent from the first reactor and the fresh feed to the second are combined in a mixer before entering the second reactor. The mixer is used because the mechanistic reactor model is configured for only one feedstream.

Each CSTR in the reactor model consists of mass and heat balances. For the component mass balances, the instantaneous density is assumed to be constant. For the energy bal-

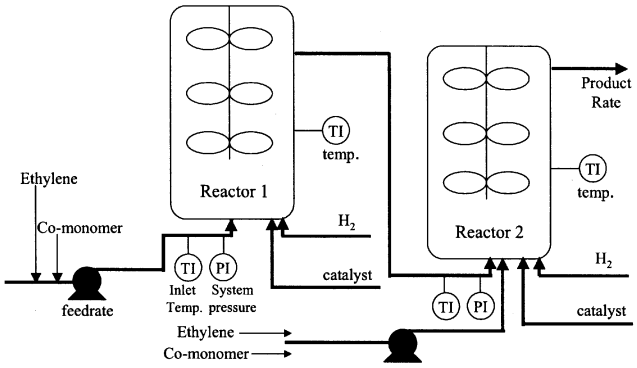


Figure 1. Polyethylene production process.

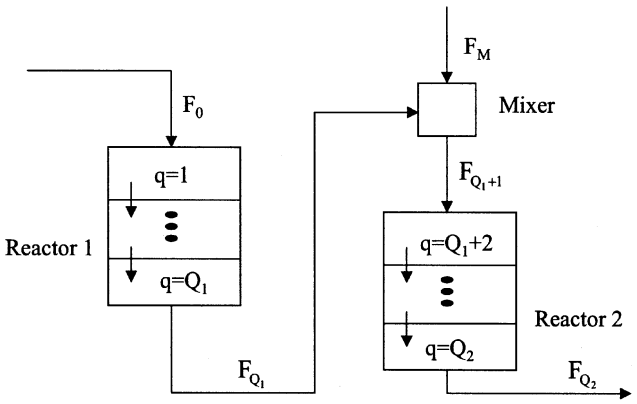


Figure 2. Dual reactor system.

ance, density is assumed to vary according to a correlation that is a function of temperature and pressure (Eq. 12). Given the inaccuracies associated with the identification of rate parameters, as well as an imperfect knowledge of the reaction mechanism, additional complexity was thought to be unjustified.

Component Balances. The component balance expressions are given by

$$V_q \frac{d[\psi]_q}{dt} = F_{q-1}[\psi]_{q-1} - F_q[\psi]_q + V_q r[\psi]_q \qquad (1)$$

where q is the number of internal CSTRs ($q = 1, \ldots, Q_1$ for reactor 1 and $q = Q_1 + 2, \ldots, Q_2$ for reactor 2); V is the volume of the CSTR (m^3), F is the volumetric flow rate (m^3/s); and $r[\psi]_q$ (kmol/m^3s) represents the change in the concentration of component $[\psi]_q$ in each CSTR due to the polymerization reactions, and Table 1 lists the components considered in the model (where ψ is replaced with the corresponding component name). It can be seen from Table 1 that rather than keeping a balance on the total catalyst concentration, separate balances are maintained for potential, vacant, dead, and occupied sites. Occupied sites are tracked using live zeroth-order moment balances. The zeroth live moments refer to the concentration of live chains with a terminal monomer

Table 1. Components Considered in the Mechanistic Model

$[\psi]_q$ (kmol/ m^3)	Component Name
$[M_1]_q$	Ethylene
$[M_2]_q$	Comonomer
$[H_2]_q$	Hydrogen
$[X]_q$	Poison
$[S]_q$	Solvent
$[PS]_q$	Potential catalyst sites
$[P_0]_q$	Vacant active catalyst sites
$[C_d]_q$	Dead catalyst site
$[\mu_{01}^k]_q$ $(k = 1,\ldots,4)$	Zeroth live moment (ethylene)
$[\mu_{02}^k]_q (k = 1,\ldots,4)$	Zeroth live moment (comonomer)
$[\lambda_{01}^k]_q (k = 1,\ldots,4)$	Zeroth bulk moment (ethylene)
$[\lambda_{02}^k]_q (k = 1,\ldots,4)$	Zeroth bulk moment (comonomer)
$[\lambda_{11}^k]_q (k = 1,\ldots,4)$	First bulk moment (ethylene)
$[\lambda_{12}^k]_q (k = 1,\ldots,4)$	First bulk moment (comonomer)

type, M_i $(i = 1,2)$. Bulk moments are used to maintain information about the properties of live and dead (that is, those not attached to catalyst sites) polymer chains. The zeroth bulk moment refers to the concentration of polymer chains, while the first bulk moment refers to the concentration of monomer, M_i $(i = 1,2)$, in the total population of chains.

Energy Balance. The energy-balance expression is given by

$$\rho_q V_q Cp_q \frac{dT_q}{dt} = \rho_{q-1} F_{q-1}(H_{q-1} - H_q) + H_R \qquad (2)$$

where T is the temperature of an internal CSTR (K); ρ (kg/m^3) is the density, C_p (kJ/kg K) is the specific heat capacity; H is the enthalpy (kJ/kg) relative to a datum temperature, and H_R (kJ/s) is the heat of reaction term where it is assumed that the monomer addition reaction is the only reaction that makes a significant contribution.

Kinetic Scheme. The kinetics used in the mechanistic model are for a multiple-site Ziegler-Natta catalyst. The kinetic mechanism is listed in the Appendix, where $P_{r,1}$ or $P_{r,2}$ refer to live polymer chains of length "r" ending in monomer 1 or 2, and all other terms are as defined in Table 1. Rate constants for the mechanisms were fitted from laboratory-scale reactor results where the reactor can be treated as an ideal CSTR. The mechanisms listed in the Appendix were deemed most appropriate for a solution polyolefins catalyst based on the laboratory data. During the data fitting, the optimum number of catalyst-site types was also determined to be four, based on deconvolution of measured MWD into Flory fractions.

Molecular-Weight Distributions. The MWD is extremely important because many polymer end-use properties are directly dependent upon the MWD. Consequently, an accurate prediction of the distribution is a desirable objective. The weight fraction of polymer of chain length r produced at site type k in CSTR q, is described by a Flory or most probable distribution, $w(r,k)_q$, given by

$$w(r,k)_q = \tau(k)_q^2 * r^* \exp(-\tau(k)_q^* r) \qquad (3)$$

where

$$\tau(k)_q = \frac{1}{DPN(k)_q} \qquad (4)$$

andz $DPN(k)_q$ is the overall number average degree of polymerization produced at site k in CSTR q which is given by

$$DPN(k)_q = \frac{(\lambda_1^k)_q}{(\lambda_0^k)_q} \qquad (5)$$

where

$$(\lambda_1^k)_q = V_q r[\lambda_{11}^k]_q + V_q r[\lambda_{12}^k]_q \quad \text{and}$$
$$(\lambda_0^k)_q = V_q r[\lambda_{01}^k]_q + V_q r[\lambda_{02}^k]_q \qquad (6)$$

The weight fraction or chain-length distribution produced in an ideally stirred reactor, $w(r)_q$, is the weighted sum of the

distributions for the individual catalyst sites

$$w(r)_q = \sum_{k=1}^{N_{ST}} m(k)_q w(r,k)_q \tag{7}$$

where $m(k)_q$ is the weight fraction of polymer produced at site type k in CSTR q, and N_{ST} are the number of catalyst sites. As there are several ideally stirred reactors, an additional summation is necessary to obtain the overall chain-length distribution of polymer produced. This is calculated as

$$w(r) = \sum_{q=1}^{Q} M(q) w(r)_q \tag{8}$$

where $M(q)$ is the weight fraction of polymer made in each reactor and $Q = Q_2 - 1$ is the total number of internal CSTRs used by the model.

As the model is tracking the bulk moments of the MWD, the overall chain-length distribution can be calculated by summing up the contributions of the polymer produced at each catalyst site in each CSTR. In order to compare the predicted distributions to measured gel permeation chromatographs (GPC), the chain-length distribution in terms of $w(r)$ vs. r must be converted to a molecular-weight distribution in terms of $dW(MW)/d\log_{10}(MW)$ vs.$\log_{10}(MW)$.

The calculation of molecular-weight distributions is also outlined in other references, such as Soares and Hamielec (1995).

Mixer Model. To model the mixer system, mass- and energy-balance dynamics are ignored and it is assumed that there is no significant heat of mixing. Referring to the notation, the expressions are

$$F_{Q_1}\psi_{Q_1} + F_M\psi_M = F_{Q_1+1}\psi_{Q_1+1} \qquad \forall\psi \tag{9}$$

$$\rho_{Q_1}F_{Q_1}H_{Q_1} + \rho_M F_M H_M = \rho_{Q_1+1}F_{Q_1+1}H_{Q_1+1} \tag{10}$$

where F_M is the flow rate of the fresh feed to the mixer (m^3/s), and H_M is the enthalpy of the fresh feedstream (kJ/kg) relative to the datum temperature. The energy balance expression, Eq. 10, must be solved iteratively to determine the mixer operating temperature.

Numerical solution of the mechanistic model

The first reactor is modeled with one internal CSTR, while the second uses two. The number of CSTRs chosen for each reactor was based upon the known characteristics of the mixing pattern in the vessel rather than varying the number of CSTRs to minimize model prediction errors. The semi-implicit Bulirsch-Stoer algorithm was used to solve the system of differential equations (see Press et al., 1992).

The following correlations for density and enthalpy are used

$$\rho = \alpha_1 + \alpha_2 P + \alpha_3 P^2 + \alpha_4 T + \alpha_5 T^2 \tag{12}$$

$$H = \beta_1 + \beta_2 T + \frac{\beta_3 T^2}{2} + \frac{\beta_4 T^3}{3} \tag{13}$$

where α_i $(i = 1,\ldots,5)$ and the $\beta_i (i = 1,\ldots, 4)$ are constants, P = pressure (kN/ m^2), and T is temperature (K), and there is a different correlation for each component.

Pressure is assumed to be at a constant value throughout the operation. The kinetic and physical property constants in the model have not been listed in order to protect commercially sensitive information.

Fundamental model predictions

A comparison between the results obtained from the mechanistic model and actual plant data revealed prediction errors in temperatures and conversions. It was thought that these errors were due to discrepancies in the plant heat and mass balance. Since these errors could contribute to errors in the MWD, every effort was made to improve the accuracy of the plant data. However, the MWD predicted by the mechanistic model was still not thought to be sufficiently accurate to predict resin or end-use properties (results of the uncorrected model predictions will be presented later). Therefore, a hybrid modeling strategy was considered.

Hybrid modeling strategy

As the MWD prediction error is multivariate, one approach is to use multivariate techniques such as partial least squares (PLS) to correct the distributions. However, this approach was not followed, as the mechanistic model concisely describes the states that affect the distribution. Adjusting the parameters of the predicted distribution at each catalyst site type can effectively modify the overall distribution. For instance, it is possible to modify the predicted mass fraction of polymer produced by each catalyst site, DPN for each site (which is a function of both the zeroth and first bulk moments), or either the zeroth or first bulk moments. In this work, adjustments were made using multipliers, as they are independent of the magnitude of the variables being adjusted. It was anticipated that these multipliers could then be predicted from plant data and used to improve the accuracy of the mechanistic model. In this work feedforward artificial neural networks were used for this purpose.

Figure 3 shows the hybrid model structure where FANN models are used to adjust key descriptors (states) of the mechanistic model. A multiple-input single-output FANN was used to predict each correction factor. Plant data are fed to

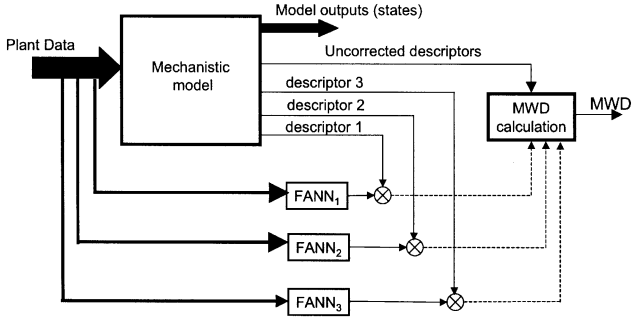


Figure 3. Hybrid model structure (showing correction of three descriptors using three separate FANN models with plant data as inputs).

the individual FANNs to determine the multiplication factors that modify certain descriptors. In the figure, three modified descriptors are shown, but in the general case the number may vary. The corrected descriptors are then fed to the MWD calculations to provide a more precise MWD.

Modeling Results

Since it was not obvious which correction technique would lead to the greatest improvement in the MWD prediction, the following three strategies were considered:

- Adjusting the mass fraction of polymer produced at certain catalyst sites.
- Adjusting DPN value for certain sites.
- Adjusting first bulk moment of ethylene (λ_{11}) concentrations for certain sites.

The correction factors were obtained by minimizing the squared error between the actual and the model MWDs for each point in the data set. In order to improve the predictions at the righthand tail of the MWD, an imaginary fifth site was included where necessary. As this site does not have moments, a DPN and mass fraction had to be assigned in order to generate the MWD. Both of these values were adjusted when fitting the distributions. Because an unconstrained optimization algorithm was used, constraints had to be hard-coded into the calculations to ensure, for example, that λ_{11} values were always positive and that mass fractions sum to unity. For the dual-reactor cases, each reactor was given a separate set of correction factors so that the MWD produced by each reactor could be determined independently.

Single reactor cases

Multiple optimization runs were carried out to determine the combination of catalyst sites and adjustment parameters that gave the best improvement in MWD prediction. The results are shown in Table 2. The sites have been numbered such that site one is producing the lowest molecular-weight fraction and site four the highest. It should be noted that the sum of the square errors (SSE) between the actual and predicted distributions is 62.0 when no correction factors are used, and that there were 46 distributions available for fitting.

Table 2. SSE Values Achieved Using Different Optimization Strategies

Sites Adjusted	Mass Fraction	DPN	λ_{11}
1	36.05	25.57	19.284
2	32.50	31.78	31.992
3	33.34	42.09	42.200
4	34.33	48.81	41.044
1,2	17.06	24.48	15.147
1,3	22.68	14.27	13.480
1,4	8.065	22.08	10.584
2,3	6.068	13.03	12.627
2,4	10.25	26.25	18.645
3,4	7.962	24.43	22.186
1,2,3	4.222	11.89	10.050
1,2,4	4.217	19.66	7.739
1,3,4	4.217	4.878	5.560
2,3,4	4.740	3.116	3.078
1,2,3,4	4.217	2.343	2.015
1,2,3,4,5	3.651	0.970	1.2104

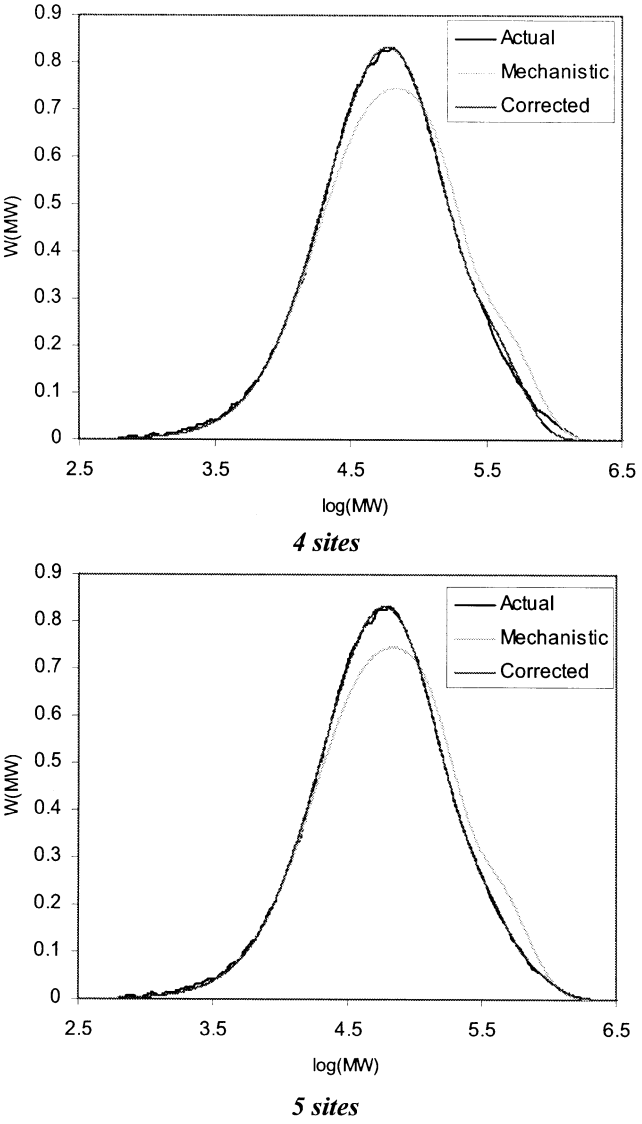


Figure 4. MWD comparison when using 4 and 5 sites.

In each case, the lowest errors are achieved when all five sites are included in the optimization. The lowest errors generally are obtained by adjusting the λ_{11} values, although modification of DPN is only slightly less effective. The mass-fraction correction factors give the worst results of the three strategies, therefore this strategy was not considered when fitting the neural networks.

Figure 4 illustrates how the addition of the fifth site improves the MWD fit at the righthand tail of the distribution (the displayed fits were obtained by adjusting DPN values). The decrease in SSE is small when the fifth site is added and may not justify the extra FANN parameters that will be needed to predict it. However, as discussed earlier, some regions of the MWD may have a greater influence on the physical properties of the polymer than others. Further analysis of the prediction of end-use properties from the MWD would help determine whether the inclusion of a fifth site is truly justified. Table 3 compares the prediction errors at the righthand tail of the distribution. It can be seen that there is a

Table 3. Prediction Errors for log(MW) in the Range [5.5

Correction Factor	Sites	SSE
DPN	1,2,3,4	1.520
DPN	1,2,3,4,5	0.153
λ_{11}	1,2,3,4	1.113
λ_{11}	1,2,3,4,5	0.261

Table 4. SSE Values Obtained Using Correction Factors Predicted by ANNs

Correction factor	Sites	Overall	Training	Validation
DPN	1,2,3,4	7.5144	3.4698	4.0446
DPN	1,2,3,4,5	6.6794	3.3167	3.3627
λ_{11}	1,2,3,4	6.8359	2.8388	3.9971
λ_{11}	1,2,3,4,5	6.0763	2.7790	3.2974

significant decrease in the errors when the fifth site is added. So if this region is required for an accurate prediction of a particular end-use property, the use of the fifth site may be justified.

Neural network prediction of correction factors

The next stage involves the use of artificial neural networks to predict the correction factors from plant data. FANNs with a single hidden layer of hyperbolic tangent activation functions were used to generate the predictions (see Figure 5). The main reason for choosing this network architecture is that it is one of the most commonly used approaches. However, there is no evidence to suggest that radial basis function networks, for example, would not provide the same degree of accuracy. The network inputs were all reactor inlet variables, including catalyst concentration, ethylene concentration, comonomer to ethylene ratio, hydrogen concentration, temperature, and total reactor feed flow.

The input–output data set was shuffled and split into training and validation sets. Twenty-five samples were used for network training, and the remaining twenty-one for validation. Multiple runs were performed using networks with different numbers of hidden-layer neurons. Due to the relatively small amount of data samples available, the networks were restricted to two, three and four hidden-layer neurons to keep the number of model parameters as low as possible. Network training was performed using an enhanced backpropagation algorithm. The "best" model was selected by choosing the network with the lowest error on the validation data set. The SSE values obtained using the correction factors predicted by the ANNs are summarized in Table 4. Note that the addition of the fifth catalyst site means that there are a total of six correction factors to be predicted, as there are two adjustable parameters for the fifth site (DPN and mass fraction).

Although the λ_{11} adjustment factors produce lower overall errors than the equivalent DPN approach, the difference in the errors on the validation data is very small.

Sample MWD Plots Using Network Predictions

The mean prediction error for the distributions in the validation set is 0.157 using λ_{11} corrections and 0.160 using DPN correction factors (adjusting all five sites). Figure 6 shows a distribution with this level of prediction accuracy.

Figure 7 compares the worst MWD fits obtained using the λ_{11} and DPN correction factors predicted using neural networks. The MWD predictions (from the validation data set) with the lowest errors are shown in Figure 8.

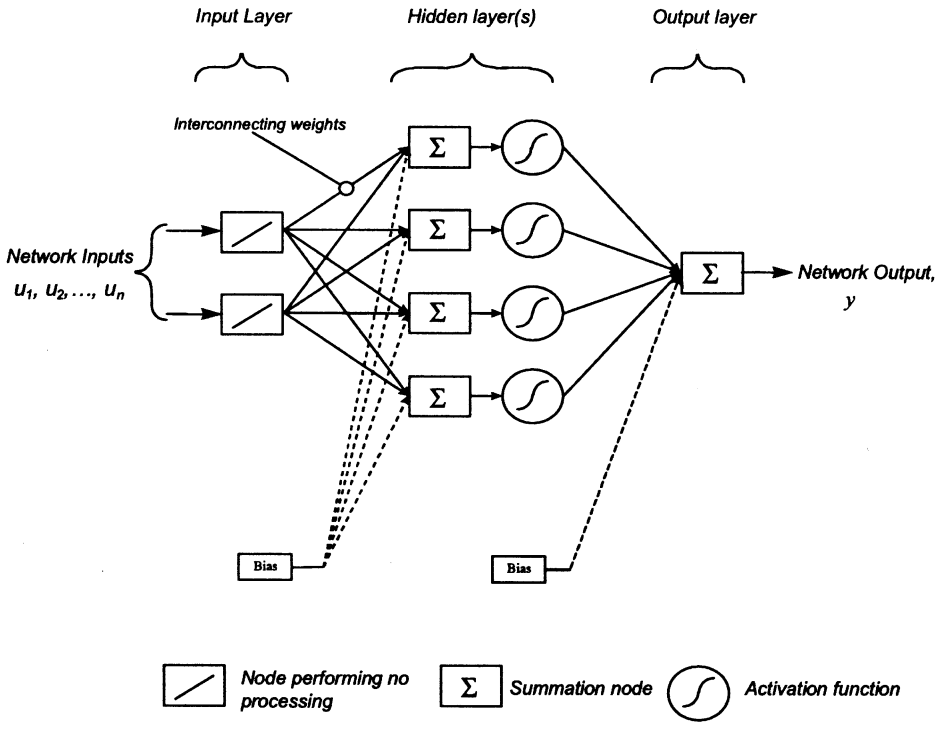


Figure 5. Feedforward artificial neural network architecture.

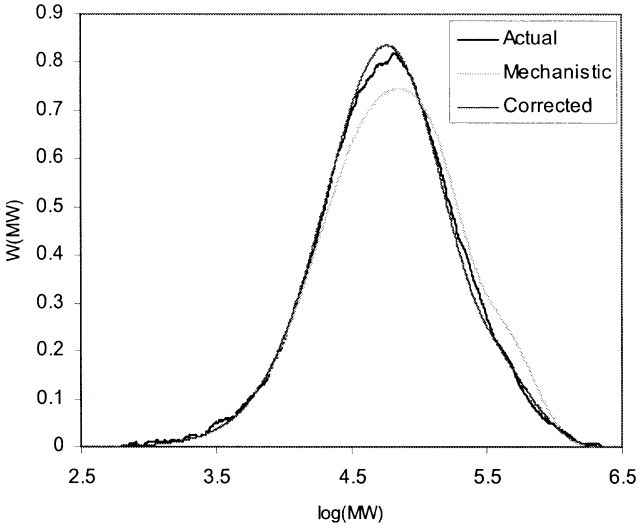
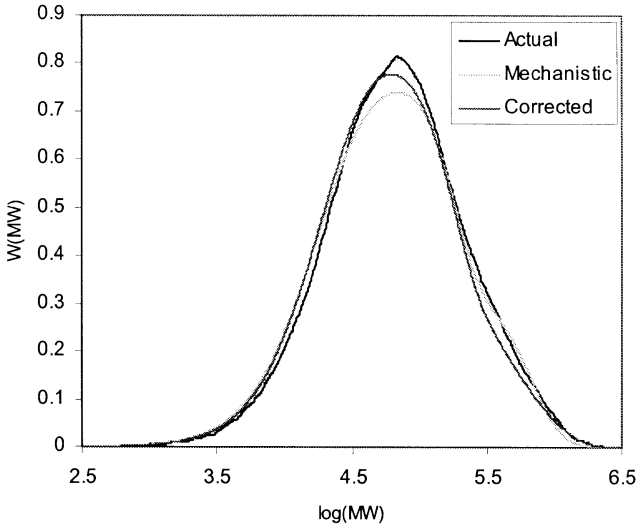
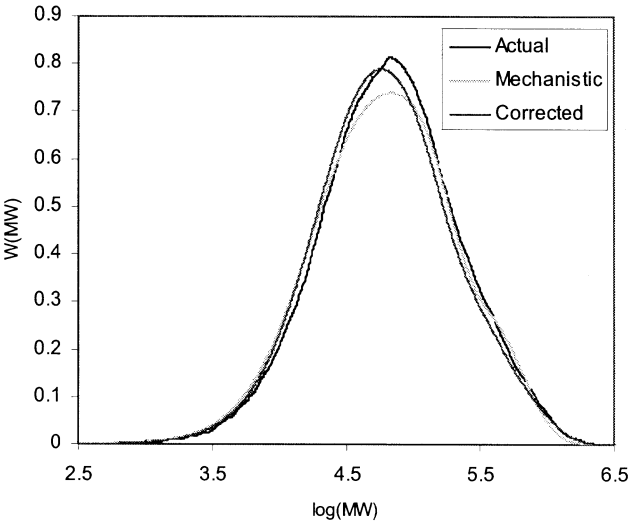


Figure 6. MWD fit with mean prediction error (λ_{11}1,2,3, 4,5; error = 0.157).



λ_{11} 1,2,3,4,5, error = 0.467



DPN 1,2,3,4,5, error = 0.588

Figure 7. Distributions with highest prediction errors.

Figure 9 shows the distributions of the errors for the λ_{11} 1,2,3,4,5 and DPN 1,2,3,4,5 corrected distributions compared to the distribution of errors for the uncorrected distributions.

Figure 10 shows that both strategies improve the accuracy of the predicted MWD. The λ_{11} correction strategy gives slightly lower SSE values when compared to the adjustment of DPN; however, the difference is not significant.

Dual reactor results

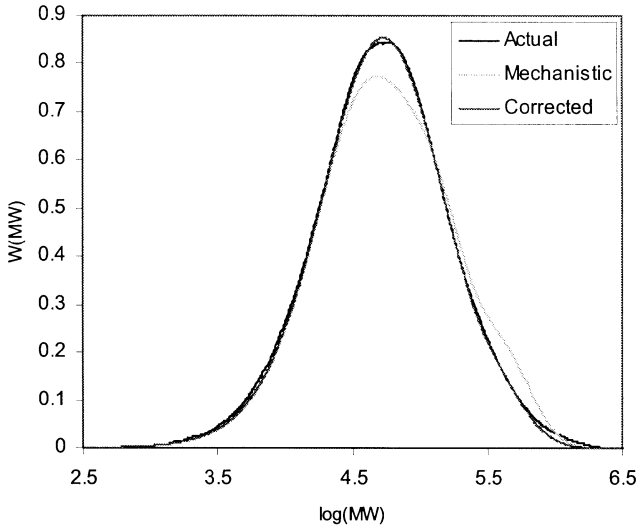
Because each reactor has its own adjustable parameters, the number of possible combinations of correction factors is much larger for the dual-reactor cases. Following the same approach as in the single-reactor work, each reactor has adjustable parameters for four catalyst sites and an additional imaginary fifth site. A selection of the optimization results are shown in Table 5. Sites 1–5 refer to the sites in the first reactor, and sites 6–10 refer to the five sites in reactor 2. As reactor 2 is modeled using two internal CSTRs, the correction factors are applied to the sites in both CSTRs. Table 6 shows the errors obtained by adjusting the λ_{11} and DPN values for all four sites in the second reactor, along with various combinations of the four sites in the first reactor.

Although it is possible to optimize a greater number of site parameters, the resulting increase in accuracy is small and is balanced by the fact that additional sets of neural-network predictions have to be developed. The errors associated with the neural-network predictions of these additional correction factors mean that the resulting distributions do not offer a significant improvement over those generated using fewer correction factors.

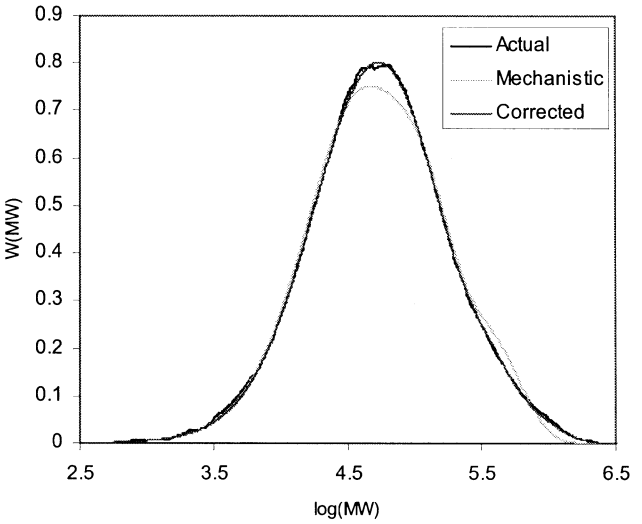
Table 7 shows how small improvements can be made by adding an imaginary fifth site to the second reactor. However, the problem with adding a fifth site is that two extra correction factors have to be estimated to yield a small improvement in accuracy. The additional neural-network weights required to predict these correction factors means that it is not necessarily better to use a fifth site.

Because the process of developing neural network predictions for the correction factors is extremely time-consuming, only a limited number of approaches could be investigated in detail. Preliminary studies indicated that the neural-network predictions for the λ_{11} correction factors resulted in more accurate distributions than those based on the DPN correction factors. This is in agreement with the single-reactor work, where the λ_{11} correction factors gave slightly better results. This may be because adjusting the λ_{11} values will also have an impact on the mass-fraction calculations, whereas adjusting the DPN has no impact on mass fractions.

Neural-network predictions were generated for the λ_{11} 3,4,6,7,8,9 and λ_{11} 4,6,7,8,9 correction factors. It was thought that these schemes provided the best trade-off between the



λ_{11} 1,2,3,4,5 (error=0.038)



(DPN 1,2,3,4,5 ,error=0.023)

Figure 8. Validation distributions with lowest prediction errors.

accuracy of the fitted distributions and the number of neural-network predictions that had to be generated. As with the single-reactor work, 20 runs were performed for a number of different neural-network architectures in order to discover the most suitable model structure. Sixty data samples were used for network training and a further 41 for validation purposes. Table 7 shows how the neural-network predictions for the λ_{11} 4,6,7,8,9 correction factors resulted in the most accurate predictions.

Examples of the λ_{11} 4,6,7,8,9 corrected distributions are shown in Figures 10, 11, and 12. The figures represent the minimum, mean, and maximum prediction errors, respectively.

Figure 13 compares the distribution of errors using the λ_{11} 4,6,7,8,9 and λ_{11} 3,4,6,7,8,9 correction schemes with that of

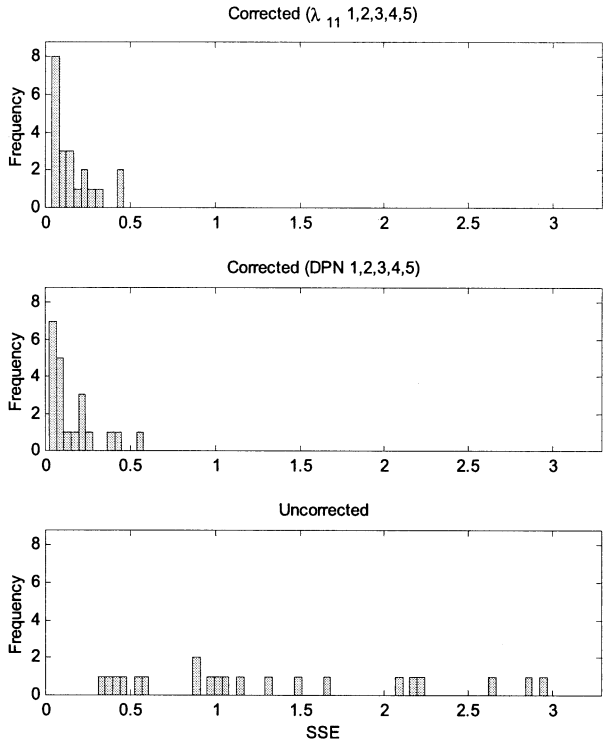


Figure 9. Comparison of error distributions (validation cases).

the uncorrected distributions. Although the λ_{11} 4,6,7,8,9 produces the more accurate predictions on the training data, there is little difference between the two techniques on the validation data. It can be concluded that the additional correction factor in the λ_{11} 3,4,6,7,8,9 scheme has not led to a significant improvement in performance.

Conclusions

In this contribution a hybrid model has been developed that allows the simulation of either a single- or dual-reactor

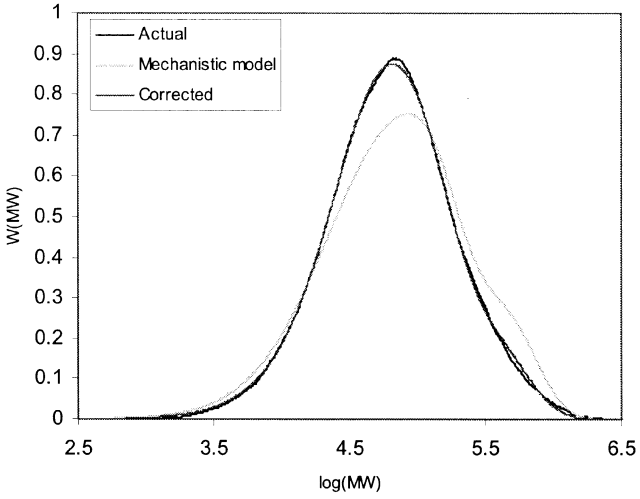


Figure 10. MWD (SSE = 0.03214).

Table 5. Optimizations Without the Addition of a Fifth Site

Sites adjusted	DPN	λ_{11}
6,7,8,9	7.236	6.389
1,6,7,8,9	6.626	6.116
2,6,7,8,9	7.098	6.131
3,6,7,8,9	6.640	5.656
4,6,7,8,9	4.014	4.327
1,2,6,7,8,9	6.483	5.808
1,3,6,7,8,9	6.076	5.359
1,4,6,7,8,9	3.465	4.014
2,3,6,7,8,9	6.548	5.363
2,4,6,7,8,9	3.830	4.060
3,4,6,7,8,9	2.153	2.637

Table 6. Optimizations Results with an Additional Fifth Site Added to Reactor 2

Sites Adjusted	DPN	λ_{11}
4,6,7,8,9,10	1.887	3.460
3,4,6,7,8,9,10	1.702	2.782

Table 7. Summary of Prediction Errors on Validation Data (SSE)

	Uncorrected	λ_{11}4,6,7,8,9	λ_{11}3,4,6,7,8,9
Minimum	0.2023	0.0511	0.0705
Mean	1.3286	0.4137	0.5212
Maximum	5.1276	2.5287	3.0548

configuration of a solution polymerization process. An empirical layer corrects the mechanistic model's MWD prediction by manipulating key descriptors (states or functions of states) used to construct the MWD. An optimization algorithm was used to adjust these parameters and minimize the error between the actual and predicted MWD. Feedforward artificial neural networks were then used to predict the correction factors from plant data. It was shown for both single- and dual-reactor simulations that the hybrid model produced more accurate MWD predictions than was possible with the mechanistic model alone.

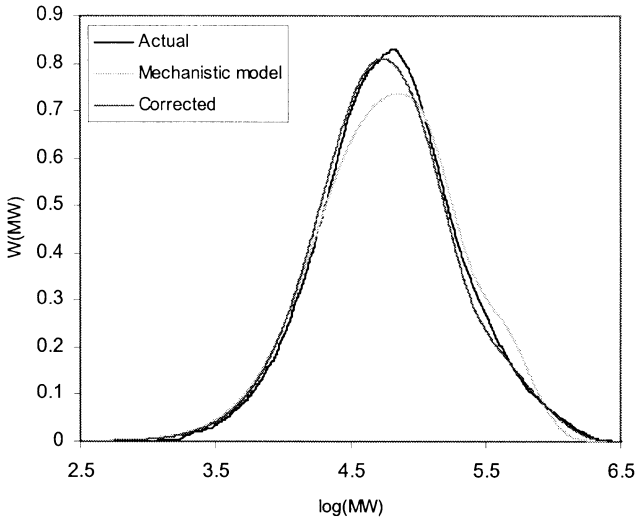


Figure 11. MWD (SSE = 0.3107).

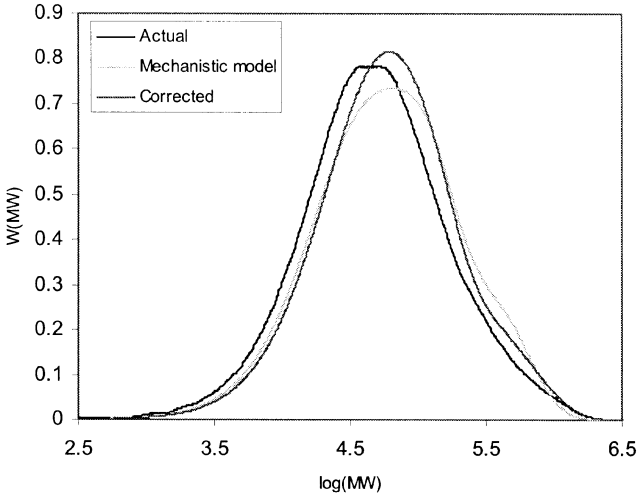


Figure 12. MWD (SSE = 2.8065).

The availability of more plant data would allow further investigation into alternative modeling strategies, which may improve the accuracy of the empirical layer. Alternative strategies could include using different combinations of site adjustment factors for improving the MWD directly or indirectly through improved predictions of other model states such as reactor temperatures.

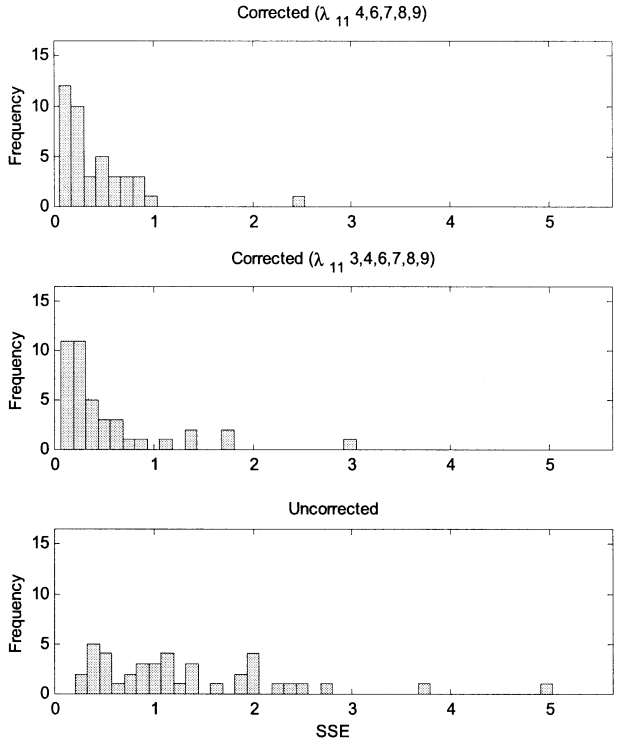


Figure 13. Comparison of prediction errors (validation cases).

While good MWD predictions have been obtained using off-line data and every effort has been made to minimize plant–model mismatch, on-line application may still require improved prediction accuracy. With the availability of an accurate hybrid model, state estimators/observers can be implemented to correct for unmeasured plant disturbances. Since the early 1970s there has been much work that has studied the problem of state estimation in polymerization reactors, and recent articles include Sirohi and Choi (1996), Crowley and Choi (1996), and Tatiraju et al. (1999).

Acknowledgments

The work described in this article evolved over 5 years of investigation. During this time, many people have contributed to the study. In particular, Brian Lines (formerly NOVA Chemicals and now with AspenTech) was influential in initiating the work and providing the initial vision regarding its potential. Ben McKay (formerly Newcastle University and now with Avantium BV) developed the basis of the fundamental model from which this article builds. The help of various people at NOVA Chemicals Corporation for provision of process data and knowledge is also gratefully acknowledged.

Notation

C_d = concentration of dead catalyst sites, (kmol/m^3)
C_p = specific heat capacity, (kJ/kgK)
DPN = overall number average degree of polymerization
F = volumetric flow rate, (m^3/s)
H = enthalpy, (kJ/kg)
H_2 = concentration of hydrogen, (kmol/m^3)
H_R = heat of reaction, (kJ/s)
k = number of catalyst sites
k_1 = rate constant for initiation reaction
K_p = rate constant for chain propagation
K_t = rate constant for chain transfer
K_D = rate constant for deactivation
m = weight fraction of polymer produced by a particular catalyst site
M = weight fraction of polymer produced in a reactor
M_1 = concentration of ethylene (kmol/m^3)
M_2 = concentration of comonomer, (kmol/m^3)
MW = molecular weight
N_{ST} = number of catalyst sites
P_0 = concentration of vacant active catalyst sites, (kmol/m^3)
$P_{r,i}$ = polymer chains of length r ending in monomer i
PS = concentration of potential catalyst sites, (kmol/m^3)
q = number of internal CSTRs
Q_1 = total number of internal CSTRs in reactor 1
Q_2 = total number of internal CSTRs in reactor 2
r = polymer chain length
S = concentration of solvent, (kmol/m^3)
t = time, (s)
T = temperature, (K)
V = volume of CSTR, (m^3)
$w(r,k)$ = weight fraction of polymer of chain length r produced at site k
X = concentration of poison, (kmol/m^3)

Greek Letters

α = parameters of density correlation
β = parameters of enthalpy correlation
λ_{01} = concentration of the zeroth bulk moment of ethylene, (kmol/m^{3}s)
λ_{02} = concentration of the zeroth bulk moment of comonomer (kmol/m^{3}s)
λ_{11} = concentration of the first bulk moment of ethylene, (kmol/m^{3}s)
λ_{12} = concentration of the first bulk moment of comonomer, (kmol/m^{3}s)
μ_{01} = concentration of the zeroth live moment of ethylene, (kmol/m^{3}s)
μ_{02} = concentration of the zeroth live moment of comonomer, (kmol/m^{3}s)
ρ = density, (kg/m^3)
τ = 1/DPN
$r[\psi]$ = the change in concentration of component ψ due to reaction, (kmol/m^{3}s)

Literature Cited

Ariawan, A. B., S. G. Hatzikiriakos, S. K. Goyal, and H. Hay, "Effects of Molecular Structure on the Rheology and Processability of Blow Molding High Density Polyethylene Resins," *Adv. in Polym. Technol.*, **20**(1), 1 (2001).
Crowley, T. J., and K. Choi, "On-Line Monitoring and Control of a Batch Polymerization Reactor," *J. Process Control.*, **6**(2/3), 119 (1996).
Crowley, T. J., and K. Y. Choi, "Experimental Studies on Optimal Molecular Weight Distribution Control in a Batch-Free Radical Polymerisation Process," *Chem. Eng. Sci.*, **53**(15), 2769 (1998).
Kiparissides, C., "Polymerization Reactor Modeling: a Review of Recent Developments and Future Directions," *Chem. Eng. Sci.*, **51**(10), 1637 (1996).
Kiparissides, C., and J. Morris, "Intelligent Manufacturing of Polymers," *Comput. Chem. Eng.*, **20**, S1113 (1996).
Kramer, M. A., M. L. Thompson, and P. M. Bhagat, "Embedding Theoretical Models in Neural Networks," ACC/WA14, 475 (1992).
Latado, A., M. Embirucu, A. G. M. Neto, and J. C. Pinto, "Modeling of End-Use Properties of Poly(propylene/ethylene) Resins," *Polym. Test* **20**(4), 419 (2001).
Lines, B., D. Hartlen, F. D. Paquin, S. Treiber, M. de Tremblay, and M. Bell, "Polyethylene Reactor Modelling and Control Design," *Hydrocarbon Process.,* 119 (1993)
MacGregor, J. F. "Control of Polymerisation Reactors," *Dycord '86*, Proc. IFAC Symp., Bournemouth, U.K., p. 31 (Dec. 8–10, 1986).
MacGregor, J. F., A. Penlidis, and A. E. Hamielec, "Control of Polymerisation Reactors: a Review," *Polymer Process Eng.*, **2**(2/3), 179 (1984).
McKay, B., C. Sanderson, M. J. Willis, J. Bartford, and G. Barton, "Evolving a Hybrid Model of a Fed-Batch Fermentation Process," *Trans. Inst. Meas. and Control*, **20** (1), 4 (1998).
Ogawa, M., M. Ohshima, K. Morinaga, F. Watanabe, "Quality Inferential Control of an Industrial High Density Polyethylene Process," *J. Process Control*, **9**, 51 (1999).
Ohshima, M., and M. Tanigaki, "Quality Control of Polymer Production Processes," *J. Process Control*, **10** (2/3), 135 (2000).
Press, W. H., B. P. Flannery, A. Saul, and W. T. Vetterling, *Numerical Recipes in C: The Art of Scientific Computing*, 2nd ed., Cambridge Univ. Press, New York (1992).
Psichogios, D. C., and L. H. Ungar, "A Hybrid Neural Network-1st Principles Approach to Process Modeling," *AIChE J.*, **38** (10), 1499, (1992).
Qi, H., X. Zhou, L. Liu, and W. Yuan, "A Hybrid Neural Network–First Principles Model for a Fixed Bed Reactor," *Chem. Eng. Sci.*, **54** 2521(1999).
Ray, W. H. "Polymerisation Reactor Control," *ACC*, 842 (1985).
Schubert, J., R. Simutis, M. Dors, I. Havlik, and A. Lubbert, "Hybrid Modelling of Yeast Production Processes—Combination of a Priori Knowledge on Different Levels of Sophistication," *Chem. Eng.Tech.* **17**(1), 10 (1994).
Seki, H., M. Ogawa, S. Ooyama, K. Akamatsu, M. Ohshima, and W. Yang, "Industrial Application of a Nonlinear Model Predictive Control to Polymerisation Reactors," *Control Eng. Pract.*, **9**, 819 (2001).
Sirohi, A., and K. Choi, "On-Line Parameter Estimation in a Continuous Polymerization Process," *Ind. Eng. Chem. Res.*, **35**, 1332 (1996).
Soares, J. B. P. and A. E. Hamielec," Deconvolution of Chain-Length Distributions of Linear Polymers Made by Multiple-Site-Type Catalysts,"*Polymer*, **36**(11), 2257 (1995).
Tatiraju, S., M. Soroush, and B. A. Ogunnaike, "Multirate Non-Linear State Estimation with Application to a Polymerization Reactor," *AIChE J.*, **45**(4), 769 (1999).

Tsen, A., Y.-D. S. S. Jang, D. S. H. Wong, and B. Joseph, "Predictive Control of Quality in Batch Polymerization Using Hybrid ANN Models," *AIChE J.*, **42**(2), 455 (1996).

Appendix 1: Kinetic Scheme for Polyethylene Model

Spontaneous site activation

$$C_P \rightarrow P_0 \quad \text{(A1)}$$

Initiation

$$P_0 + M_1 \xrightarrow{K_{I_1}} P_{1,1} \text{ (Addition of ethylene to an active vacant site)} \quad \text{(A2)}$$

$$P_0 + M_2 \xrightarrow{K_{I_2}} P_{1,2} \text{ (Addition of co-monomer to an active vacant site)} \quad \text{(A3)}$$

Chain propagation

$$P_{r,1} + M_1 \xrightarrow{K_{P_{11}}} P_{r+1,1} \text{ (Addition of ethylene to a chain ending in ethylene)} \quad \text{(A4)}$$

$$P_{r,2} + M_1 \xrightarrow{K_{P_{21}}} P_{r+1,1} \text{(Addition of ethylene to a chain ending in comonomer)} \quad \text{(A5)}$$

$$P_{r,1} + M_2 \xrightarrow{K_{P_{12}}} P_{r+1,2} \text{(Addition of comonomer to a chain ending in ethylene)} \quad \text{(A6)}$$

$$P_{r,2} + M_2 \xrightarrow{K_{P_{22}}} P_{r+1,2} \text{(Addition of comonomer to a chain ending in comonomer)} \quad \text{(A7)}$$

Chain transfer

$$P_{r,1} + M_1 \xrightarrow{K_{tM_{11}}} P_{1,1} + D_r \text{(Chain transfer to ethylene)} \quad \text{(A8)}$$

$$P_{r,2} + M_1 \xrightarrow{K_{tM_{22}}} P_{1,1} + D_r \text{(Chain transfer to ethylene)} \quad \text{(A9)}$$

$$P_{r,1} + M_2 \xrightarrow{K_{tM_{12}}} P_{1,2} + D_r \text{(Chain transfer to comonomer)} \quad \text{(A10)}$$

$$P_{r,2} + M_2 \xrightarrow{K_{tM_{22}}} P_{1,2} + D_r \text{(Chain transfer to co-monomer)} \quad \text{(A11)}$$

$$P_{r,1} + H_2 \xrightarrow{K_{tH_1}} P_0 + D_r \text{(Chain transfer to hydrogen)} \quad \text{(A12)}$$

$$P_{r,2} + H_2 \xrightarrow{K_{tH_2}} P_0 + D_r \text{(Chain transfer to hydrogen)} \quad \text{(A13)}$$

$$P_{r,1} \xrightarrow{K_{t\beta_1}} P_0 + D_r[\text{Spontaneous chain transfer } \beta\text{-hydride}] \quad \text{(A14)}$$

$$P_{r,2} \xrightarrow{K_{t\beta_2}} P_0 + D_r[\text{Spontaneous chain transfer}] \ (\beta\text{-hydride}) \quad \text{(A15)}$$

Deactivation

$$P_{r,1} + X \xrightarrow{K_{DX_1}} C_d + D_r \text{(Deactivation by poison)} \quad \text{(A16)}$$

$$P_{r,2} + X \xrightarrow{K_{DX_2}} C_d + D_r \text{(Deactivation by poison)} \quad \text{(A17)}$$

$$P_0 + X \xrightarrow{K_{DX_3}} C_d \text{(Deactivation by poison)} \quad \text{(A18)}$$

$$P_{r,1} \xrightarrow{K_{DSP_1}} C_d + D_r \text{(Spontaneous deactivation)} \quad \text{(A19)}$$

$$P_{r,2} \xrightarrow{K_{DSP_2}} C_d + D_r \text{(Spontaneous deactivation)} \quad \text{(A20)}$$

$$P_0 \xrightarrow{K_{DSP_3}} C_d \text{(Spontaneous deactivation)} \quad \text{(A21)}$$

Manuscript received Sept. 18, 2002, and revision received Mar. 26, 2003.