APPLICATION OF HYBRID MODELING IN POLYMER PROCESSING

Christian Marschik, Wolfgang Roland, Bernhard Löw-Baselli, and Georg Steinbichler
Institute of Polymer Extrusion and Compounding, Johannes Kepler University, Linz, AUSTRIA

Abstract

For many decades, the setup and solution of polymer processing models involved use of analytical or numerical methods. These characteristics have changed with the recent digitization of polymer processes and the collection of enormous amounts of data. It is increasingly common to use data-driven modeling techniques to analyze processes, for which analytical and numerical models may not fully describe the process behavior in operational situations. These techniques have significantly extended the set of tools available to the engineer, providing new possibilities of how to develop more accurate process models. As a result, the setup of an appropriate modeling strategy more than ever requires a thorough understanding of the individual modeling techniques. This article was designed to address the potentials and limits of analytical, numerical, and data-based modeling techniques when modeling polymer processes. Moreover, we show how these methods can be combined into one hybrid approach to solve polymer process models not solvable so far. The findings are further illustrated by means of a particular use case, which models the flow of polymer melts in single-screw extruders.

Introduction

Modeling and simulation has been an important discipline in the field of polymer processing. In order to better understand and optimize polymer processes, a large variety of mathematical models have been developed, attempting to represent the actual process with equations. While experimental design and optimization procedures generally require use of cost-intensive prototypes, it is often less expensive and time consuming to develop a mathematical model of the physical process. Due to the complex material behavior of polymers, one of the main concerns of the engineer has been to strike a reasonable balance between level of accuracy and degree of sophistication. In many cases, it is possible to simplify the physical process model sufficiently in order to be able to find an analytical solution. The less complex the model, however, the less accurately it describes the actual process.

With the steadily growing requirements to decrease costs and improve product quality, numerical methods have been increasingly applied to solve more realistic polymer processing models. By relaxing certain assumptions, the complexity of the model increases and analytical solutions become elusive. Numerical techniques provide a useful tool for obtaining approximate solutions to those problems.

To date, various complex polymer processes have been realistically simulated ranging from mold filling with fiber orientation to extrusion with viscoelastic effects [1]. Due to the increased complexity of the governing equations to be solved, however, numerical analyses tend to be time consuming and generally require expert knowledge.

When using analytical or numerical techniques, the aim is to predict the physical phenomena of a process. With the recent computerization of our society and the explosive growth of data over the last few years, data-based modeling has grown enormously in importance. This approach has already been successfully implemented in highly digitized applications such as online-marketing or social media; however, it is still at an early stage of development in the field of engineering. Rather than solving the governing equations of a process model, the general objective is to extract knowledge implicitly captured in large datasets [2]. As a strongly interdisciplinary field, data-based modeling provides a large variety of techniques comprising the use of statistics, machine learning, database technology, or high-performance computing.

The applicability of the previous modeling techniques strongly depends on the complexity of the problem at hand. While analytical or numerical models try to explain the physical behavior of a process, data-based modeling can be used if there is no information about the functioning of a system at all. In order to combine the advantages of these methods, we have recently applied a hybrid modeling approach to (i) calculating the pressure loss of meltfiltration systems in polymer recycling processes [3-4], (ii)
optimizing the manufacturing of powder coatings [5], and (iii) predicting the flow in the metering zone of single-screw extruders [6-12]. This method incorporates all available knowledge about the process into one approach, thereby offering a novel strategy to solve polymer processing models not solvable so far (Figure 1). As a characteristic feature, the novel modeling procedure provides fast and stable analytical relationships even for highly complex physical problems.

This research compares the advantages and disadvantages of analytical, numerical, and data-based techniques when modeling polymer processes in general and the flow of polymer melts in single-screw extruders in particular. For the latter use case, we moreover illustrate how these methods were recently combined into one hybrid modeling approach to predicting the pumping capability and viscous dissipation of metering zones.

### Analytical Modeling

An equation is said to have an analytical solution if at least one solution can be written as a closed-form expression, which may contain constants, variables, arithmetic operations, and a specific set of functions. When analyzing engineering systems, analytical solutions offer several advantages. Because solutions are represented as analytical expressions, they can be mathematically interpreted. With parameter dependencies being expressed explicitly, solutions provide a clear view of how process variables affect the behavior of the underlying system. Moreover, if derived, analytical solutions are typically fast to compute.

The method of applying analytical techniques for the analysis of the melt conveying in screw pumps was employed in many theoretical studies [13-21]. The general procedure was as follows: Using the fundamentals of fluid mechanics, the flow of polymer melt in the screw channel was described mathematically by means of a physical process model which included the conservation equations of mass, momentum, and energy and constitutive equations. These governing equations were then simplified by means of modeling assumptions. Traditionally, the process was represented by a one-dimensional isothermal flow of a Newtonian fluid between two parallel plates in linear movement. Assuming the viscosity of the polymer melt to be constant, analytical solutions were obtained for down-channel and cross-channel velocity profiles, flow rate, power consumption, and other process variables.

Very few practical processing problems lead to exact analytical solutions. These are usually restricted to simple geometries with simple material properties and physical conditions. In the flow analysis of metering zones, analytical solutions are mainly reserved for Newtonian fluids, whereas shear-thinning polymer melts generally require use of numerical methods. Even for a one-dimensional temperature-independent flow of a power-law fluid between two parallel plates in relative movement, no exact analytical closed-form solution has been found to date [1]. Table 1 summarizes the advantages and disadvantages of analytical modeling techniques.

**Numerical Modeling**

Numerical methods can be used to derive approximate solutions to process models, for which analytical solutions are not available. In contrast to their counterparts, numerical solutions cannot be expressed in the form of complete mathematical expressions. Rather, they are given by discrete numerical values, which must be recalculated every time the parameter set changes. A major advantage of numerical procedures is their capability of handling large equations systems with different degrees of nonlinearities. In engineering systems, nonlinear physics are often found in combination with complex three-dimensional geometries. Without oversimplifying the real physical process, numerical methods provide detailed insights into the behavior of these systems by quantifying process variables that cannot be easily measured.

In the flow analysis of single-screw extruders, numerical methods were employed to include shear-thinning flow behavior of the polymer melt. In this case, the governing transport equations are coupled due to the dependency of viscosity on shear rate. Physically this means that the drag and pressure flows affect each other at each position of the flow field. The complexity is further increased by the combined effect of shear in the down- and cross-channel directions. To refine the understanding of melt-conveying in single-screw extruders, efforts have been directed towards numerical analyses of power-law model-based flows. Two approaches were mainly applied: For a long time, the use of numerical methods was limited by the computational power available to the engineer. As a result, early numerical analyses simulating shear-thinning flows in single-screw extruders developed stand-alone solution algorithms optimized for the task at hand [22-29]. These computational barriers have shifted tremendously with the advent of more advanced computers. The trend of more recent flow analyses is towards the use of commercial software packages [30-31]. In contrast to stand-alone
solving algorithms, these typically provide a large set of numerical procedures. However, since they usually solve the full set of conservation equations, they are significantly more time consuming and computationally expensive. In addition, use of commercial simulation software often involves acquisition of expensive licenses. Table 2 provides an overview of the potentials and limits of numerical modeling techniques.

Table 2. Advantages and disadvantages of numerical modeling techniques.

<table>
<thead>
<tr>
<th>Advantages</th>
<th>Disadvantages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Applicable to complex geometries</td>
<td>Offer approximate solutions</td>
</tr>
<tr>
<td>Applicable to arbitrary nonlinear physical problems</td>
<td>Solving process can be time consuming</td>
</tr>
<tr>
<td>Allow a more accurate representation of reality</td>
<td>Solving process can be computationally expensive</td>
</tr>
</tbody>
</table>

**Data-Based Modeling**

Data-based modeling techniques can be used to model polymer processes, where even numerical methods fail to deliver solutions. This may be the case if the complexity of the physical process model goes beyond a critical level or if there is a lack of process information. Many processes can be physically modeled, but solving the models may pose practical challenges if, for example, input parameters are unknown or initial and boundary conditions are not available. In these situations, data-driven models can fill the gap left by analytical or numerical process models.

From a general viewpoint, data-driven modeling is an iterative process that involves several steps [2]: (i) data cleaning, (ii) data integration, (iii) data selection, (iv) data transformation, (v) data mining, (vi) pattern evaluation, and (vii) knowledge representation. Data mining is one of the most important steps in the knowledge discovery process, because it uncovers patterns and knowledge from large amounts of data. For polymer processes, these may not only include experimental lab or production data, but also data resulting from analytical and numerical modeling. A wide variety of machine learning techniques can be applied to identify patterns in datasets. These can be classified into two main categories, as shown in Figure 2: (i) supervised and (ii) unsupervised learning algorithms. Supervised learning algorithms can be applied to discover relationships between input and target attributes [32]. The main intention is to learn from labeled training data, where the targeted features are known, how to predict labels of unseen instances. Supervised learning algorithms can be further divided into classification and regression algorithms. Unsupervised learning, in contrast, refers to modeling the distribution of instances in a typical high-dimensional input space [32]. These learning techniques, which are used when there is no class to be predicted, include clustering and association.

![Figure 2. Overview of data-mining techniques.](image)

Although providing great potential, data-based modeling is still at its infancy in the field of polymer processing. So far, the method is mainly applied for quality control and predictive maintenance systems. Only a few analyses have employed data-based modeling techniques for developing new process models [3-11, 33]. When using experimental data for data-based analyses, the models are usually restricted to the specific type of process under consideration. As a result, new models have to be constructed if for example the size of the processing machine or the behavior of the material changes. Table 3 compares advantages and disadvantages of data-based modeling techniques.

Table 3. Advantages and disadvantages of data-based modeling techniques.

<table>
<thead>
<tr>
<th>Advantages</th>
<th>Disadvantages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arbitrary source for datasets</td>
<td>Accuracy of predictions depends on data quality</td>
</tr>
<tr>
<td>Knowledge about the underlying process is not required</td>
<td>Correlations are only valid within the range of the underlying dataset</td>
</tr>
<tr>
<td>Applicable to problems that cannot be analyzed physically</td>
<td>Ignore the physics of the underlying system</td>
</tr>
</tbody>
</table>

**Hybrid Modeling**

Our research group has recently proposed a hybrid modeling approach to solve process models not solvable so far [3-12]. The following section illustrates the main characteristics of the modeling approach. To demonstrate its novelty and usefulness for a given example, the development of analytical models for predicting the pumping capability and viscous dissipation in metering channels is shown. Figure 3 represents a flow chart of the modeling procedure, which combines analytical, numerical, and data-based modeling techniques. It is pointed out that the modeling procedure is not restricted to particular types of problems. In our recent studies, the approach was mainly applied to analyze flows of polymer melts in processing machines. However, it can also be used
to build new models for predicting the mechanical behavior of plastics products during design and manufacturing.

Figure 3. Flow chart for our hybrid modeling approach.

**Analytical modeling**

In a first step, a mathematical model of the physical process is developed, which requires a fundamental analysis of the governing process equations. For the example shown here, these included the conservation equations of mass, momentum, and energy and a constitutive equation for the rheological behavior.

![Figure 4. Representation of the screw channel.](image)

To describe the stress responses of the polymer melt being processed, a constitutive equation was introduced:

\[
\tau = 2\eta (\dot{\gamma}) D, \quad (6)
\]

\[
D = \frac{1}{2} (\nabla \nu + \nabla \nu^T). \quad (7)
\]

The shear-thinning flow behavior of the polymer melt was modeled as a power-law fluid:

\[
\eta = K |\dot{\gamma}|^{n-1}. \quad (8)
\]
This well-known relationship relates the viscosity and the shear rate via the consistency and the power-law index. Further, we assume an incompressible polymer melt.

### Table 4. Velocity boundary conditions.

<table>
<thead>
<tr>
<th>x</th>
<th>y</th>
<th>v_x</th>
<th>v_y</th>
<th>v_z</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>y</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>w</td>
<td>y</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>x</td>
<td>h</td>
<td>v_{b,x}</td>
<td>0</td>
<td>v_{b,z}</td>
</tr>
</tbody>
</table>

In the second step, the governing process equations are rewritten in dimensionless form. The general objective is to detect the physically independent dimensionless parameters of the system. Two problems that are defined by identical dimensionless parameters are similar in terms of their underlying physics. It is thus possible to recognize operational situations that may run under different sets of processing conditions, but are governed by the same physics. Moreover, the number of influencing parameters can be significantly reduced. Using dimensional analysis and the theory of similarity, we transformed our physical process model into dimensionless form and showed that the flow equations were governed by four independent dimensionless input parameters: (i) the aspect ratio of the screw channel h/w, (ii) the screw-pitch ratio v/D_b, (iii) the power-law index n, and (iv) a dimensionless down-channel pressure gradient \( \Pi_{p,z} \) defined as:

\[
\Pi_{p,z} = \frac{p}{\nu h^{1+n}} \frac{t}{w},
\]

These characteristic dimensionless parameters were then varied to create a large set of roughly 90,000 physically independent modeling setups.

### Numerical modeling

In the third step, the physically independent design points derived in the previous section are solved numerically. For the example presented here, we evaluated the volume flow and the total viscous dissipation in the screw channel. These quantities were obtained from the velocity field and the cross-channel pressure distribution in the screw channel:

\[
\dot{V} = i \int v_x (x, y) \, dA,
\]

\[
\dot{Q}_{\text{diss}} = i \int q_{\text{diss}} (x, y) \, dA.
\]

For modeling, a dimensionless form of the output variables was introduced:

\[
\Pi_v = \frac{2\dot{V}}{i \, w \, h \, v_{b,z}},
\]

\[
\Pi_{\text{diss}} = \frac{\dot{Q}_{\text{diss}}}{i \, w \, K \, v_{b,z}^{n+1}}.
\]

For each modeling setup, the target variables were evaluated numerically by means of a parametric design study. Figure 5 shows numerical solutions for the volume flow rate for a given set of input parameters. Similar results were obtained for the entire set of modeling setups. At this stage, the numerical solutions are represented by discrete values, whereas a mathematical relationship between input and output parameters is still not available.

### Data-based modeling

In the fourth step, the numerical results of the parametric design study are approximated using data-based modeling techniques. This can be done by using supervised learning algorithms. While the aim of classification is to learn a function that assigns one of several predefined classes to each data item, the objective of regression is to find a function that predicts a real numerical value for each input object [37]. Several techniques are available to derive correlations between input and output data including e.g., decision trees, neural networks, support vector machines, or symbolic regression.

For the problem presented here, we used symbolic regression based on genetic programming to derive regression models for the dimensionless volume flow rate and the total viscous dissipation as a function of the characteristic dimensionless input parameters:

\[
\Pi_v = f \left( \frac{h}{w}, \frac{t}{D_b}, n, \Pi_{p,z} \right),
\]

\[
\Pi_{\text{diss}} = f \left( \frac{h}{w}, \frac{t}{D_b}, n, \Pi_{p,z} \right),
\]

Figure 5. Dimensionless flow rate as a function of screw-pitch ratio and dimensionless pressure gradient for \( n = 0.5 \) and \( t/D_b = 2.0 \).
\[ \Pi_m = f \left( \frac{h}{w}, \frac{t}{D_n}, n, \Pi_e \right). \]  \hspace{1cm} (16)

Unlike other regression methods, this type of regression analysis requires neither model structure nor model parameters of the regression model to be predefined. Rather, by employing evolutionary computation, symbolic regression involves finding the best model structure and its coefficients simultaneously. The following settings need to be predefined:

- terminal set (set of input variables or constants)
- function set (functions used for regression)
- fitness function (quality measure)
- algorithm control parameters
- termination criterion (e.g., error tolerance)

The solving process is driven as follows: By randomly combining the mathematical building blocks that are predefined by a function set, initial expressions are derived, which are then recombined to yield new models. For producing new solutions, the algorithm applies two genetic operators \[ 36 \]: Crossover takes two individuals (parents) and produces new individuals (offspring) by combining parts of the parents. Mutation is an arbitrary modification, which creates new points in the search space.

We divided the dataset into three subsets: (i) a training set, (ii) a validation set, and (iii) a test set. Whereas the first two subsets were employed to develop the regression models, the third was used to validate the relationships for unseen instances by means of error predictions. In this manner, three analytical regression models were developed for predicting the pumping capability and viscous dissipation of metering zones [7,12]:

\[ \Pi_m \left( \frac{h}{w}, \frac{t}{D_n}, n, \Pi_e \right) = \frac{A_1 A_4}{A_4} \left( A_1 + A_4 + \frac{A_4 A_5}{A_5 + A_4 / A_4} \right) + A_4, \]  \hspace{1cm} (17)

\[ \Pi_m \left( \frac{h}{w}, \frac{t}{D_n}, n, \Pi_e \right) = B_1 + B_2 + B_4 + B_4 + \frac{1}{B_5 + B_6 + B_7}, \]  \hspace{1cm} (18)

\[ \Pi_m \left( \frac{h}{w}, \frac{t}{D_n}, n, \Pi_e \right) = C_1 + C_2 + C_3 + C_4 \left( C_5 + C_6 \right), \]  \hspace{1cm} (19)

As a useful feature, our approximation models show a simple structure. The subfunctions \( A_1 \) to \( A_{14} \), \( B_1 \) to \( B_7 \), and \( C_1 \) to \( C_7 \) depend on the dimensionless influencing parameters, model coefficients, basic arithmetic operations, and simple analytical functions (e.g., square function, exponential function). The models consider the influence of both the shear-thinning flow behavior of the polymer melt and the effect of the three-dimensional channel geometry. Estimation is possible for pressure-generating and pressure-consuming metering zones. Compared to time-consuming and computationally expensive numerical analyses, the models are significantly faster and do not require large computational power. The quality of the developed regression models was evaluated by means of an error analysis. Figure 6 shows scatter plots of the symbolic regression models, comparing the numerically evaluated results with the approximated solutions of the test set.

![Scatter plots](image-url)
This article introduced a hybrid modeling approach to solving polymer processing models not solvable so far. The method incorporates analytical, numerical, and data-based modeling techniques into one approach. Combining the advantages of each modeling technique, the procedure considers the physics of the underlying system and provides simple, fast, and stable analytical relationships for complex processes. The usefulness of the hybrid modeling approach has been demonstrated in various applications [3-11]. Due to their simple algebraic structure, these models can not only be implemented in traditional fields – such as process optimization and troubleshooting – but also in emerging applications – such as digital twins and soft sensors. The above outlined features of models have the potential to save resources, reduce timelines, and improve manufacturing.

### Acknowledgement

This work was financially supported by the COMET Center CHASE. The COMET Centre CHASE is funded within the framework of COMET - Competence Centers for Excellent Technologies by BMVIT, BMDW, the Federal Provinces of Upper Austria and Vienna. The COMET program is run by FFG.

### References