ADAPTIVE MULTI-OBJECTIVE OPTIMIZATION OF PROCESS CONDITIONS FOR INJECTION MOLDING USING THE GAUSSIAN PROCESS APPROACH

Jian Zhou\(^1\) and Lih-Sheng Turng\(^2\)*

\(^1\)Phoenix Integration Inc., Blacksburg, VA 24060, jzhou@phoenix-int.com;
\(^2\)Polymer Engineering Center, Department of Mechanical Engineering, University of Wisconsin–Madison, Madison, WI 53706, turng@engr.wisc.edu

Abstract

This paper presents an integrated simulation-based optimization system that incorporates the design of computer experiments, the Gaussian process (GP) for regression, a multi-objective genetic algorithm (GA), and levels of adjacency to adaptively and automatically search Pareto-optimal solutions for different objectives. Selecting the optimal process conditions for the injection molding process is a multi-objective optimization problem. Different objectives, such as minimizing the injection pressure, volumetric shrinkage, warpage, cycle time, etc., exhibit trade-off behaviours, and thus various optima may exist in the objective space. Since the GP approach is capable of providing both the predictions and the estimation of the predictions simultaneously, a non-dominated sorting of the predicted variances at each iteration step is performed to intelligently select extra samples that can be used as additional training samples to improve the GP surrogate models. At the same time, user-defined adjacency criterion has been implemented for evaluation of the convergence of model iterations. The applications in this paper demonstrate that the proposed multi-objective optimization system can help injection molding engineers efficiently and effectively identify and set up the optimal process conditions.

1 Introduction

Injection molding is one of the most important manufacturing processes for mass production of complex plastic parts [1]. During the whole manufacturing process, various process conditions have to be properly set up to reduce the production cost and time-to-market and to ensure the quality of the molded components. Traditionally, the process conditions are often determined by experienced engineers or based on reference handbooks and later improved and fine-tuned by trial-and-error on the shop floor. This method depends highly on the experience of molding operators and could potentially be costly and time consuming, especially with new resins or new applications.

With advances in numerical modeling and computer simulation, many engineers employ sophisticated computer-aided engineering (CAE) simulation tools to facilitate injection molding design and process setting [2-5]. However, the CPU-expensive evaluations of simulations, the huge amount of computer generated data, and the complex nonlinear interactions among the design and process variables frequently make it difficult to effectively determine the optimal process conditions. Moreover, even though an acceptable design or a set of process conditions have been decided after numerous numerical iterations and user interventions, it still remains unknown whether they can be further optimized [6]. Therefore, there still remains a gap between the numerical simulation predictions and the objectives of employing simulation, which is to achieve the optimal design and process conditions.

Many researchers have developed and employed different optimization schemes for determining the optimal design and process parameters for polymer processing. In particular, Gaspar-Cunha et al. developed a multi-objective evolutionary algorithm, which was denoted as Reduced Pareto Set Genetic Algorithm with elitism (RPSGAe), for the optimization of the processing conditions of a co-rotating twin-screw extruder, and further for the optimization of the screw geometry in order to maximize a prescribed performance [7, 8]. Deng et al. applied a genetic algorithm (GA) to optimize injection molding process conditions with user-definable objective functions. They implemented a modified simple weighting method to deal with multi-objective optimization, in which the objective functions can be defined with different criteria and/or weight-vectors according to the designer’s preference [9]. In their optimization procedure, every fitness value of the objective functions was evaluated with a complete CAE simulation, which may require extensive simulation runtime. Also, the trade-off solutions obtained by using the pre-defined strategy would be sensitive to the weight factors chosen in converting the multi-objective to a single-objective function. Castro et al. used an approach comprised of computer simulation, artificial neural network (ANN), and data envelopment analysis (DEA) to determine the proper operating conditions for finding the best compromise among several conflicting performance measures. The approach they presented also allowed for the identification of robust variable settings that might help to define a starting point for negotiation between multiple decision makers [10]. In their work, full factorial DOE were employed to generate the initial dataset for creating the surrogate models (metamodels). Recently, Zhou and Turng proposed a CAE optimization...
platform that can quickly determine the optimal processing conditions for injection molding. This approach uses support vector regression (SVR) to establish a surrogate model (or surrogate models in the case of multi-objective optimization) to approximate the CPU-intensive 3D simulations. GAs are then implemented to evaluate the surrogate model(s) for searching for the single or multiple optimal solutions, respectively. In addition, the performance and capabilities of using different modeling approaches to establish the surrogate models, such as artificial neural networks (ANNs), polynomial regression (PR), and support vector regression (SVR), were also investigated [11].

Although the above-mentioned studies have achieved various levels of success, it remains desirable to have the ability of offering an intelligent optimization strategy. Zhou and Turng employed the Gaussian process (GP) regression method to develop an adaptive optimization system for injection molding [12, 13]. The Gaussian processing method belongs to the kernel method and uses covariance functions as the kernel functions. Traditionally, GP has been used for specific applications such as spatial models in meteorology and geology, the analysis of computer experiments, and time series analysis. Recently, researchers in the engineering field began to use this method to deal with specific engineering classification and regression applications. To the best knowledge of the authors, this is the first attempt that the GP regression method was employed for optimization of injection molding. However, in this paper the GP regression method was only applied to solve single objective optimization problems. The purpose of this study is to develop an integrated simulation-based optimization system that can adaptively and automatically find out the Pareto-optimal solutions for different objectives. The idea is to use the Gaussian process (GP), a nonlinear statistical regression technique, coupled with design of computer experiments to establish surrogate models that can approximate and substitute computational expensive simulations while using minimum computing resources and intelligently selecting additional sampling points in the design space. These additional training samples will be used to adaptively update and improve the GP surrogate models for multi-objective optimization, so that a large amount of process (or design) alternatives can be evaluated in a reasonable timeframe during a system-level optimization process due to the surrogate models’ accurate approximation and instantaneous executions.

The rest of this paper is organized as follows. Section 2 gives a brief mathematical description of Gaussian process and multi-objective genetic algorithms (MOGA) used in this study for optimization. The levels of adjacency are used to evaluate the convergence of the adaptive optimization system. Section 3 presents the framework implementation of the adaptive optimization system, in which the methodology of design of experiments (DOE) for computer simulation using the Latin hypercube sampling (LHS) is employed to generate the samples for training and enhancing the surrogate models. In Section 4, a multi-objective optimization application will be presented to illustrate the performance and capabilities of the proposed optimization strategy. Finally, Section 5 gives the conclusions of this paper.

2 Adaptive Multi-Objective Optimization System for Injection Molding

2.1 Gaussian Process for Regression

Design and process condition optimization with simulation can be expensive and time consuming as they may require excessive computing time for full-fledged numerical simulation and numerous design/process iterations. Using statistical methods that can extract useful information from the simulation results and formulate surrogate models to correlate process conditions and objectives will dramatically reduce the computational time/resource requirements and provide the maximum benefit for optimization. Furthermore, explicit mathematical expressions for response surface topologies that are obtained from such models can not only predict unsimulated scenarios but can also convert the task of searching for optima into an efficient global optimization procedure.

The sample data in this study feature challenges of high-dimensionality, complex correlations within a large space, complicated interactions between factors, and small data samples. Therefore, a kernel-based regression technique called the Gaussian process (GP) method is proposed and implemented. GP is a highly nonlinear regression technique based on the Bayesian probability and inference approach. Taking into account prior information that is available from the data, and assuming the associated error to be normally distributed, the GP regression method finds the most likely value as well as the variance for the desired response directly through the statistical approach. The GP method demands much weaker regression model assumptions compared with parametric regression methods. This gives more flexibility to find the best surrogate models from a much larger objective space.

Given a training data set $D$ consisting of $N$ pairs of $L$-dimensional inputs $x_n$ and scalar outputs $t_n$ for $n = 1...N$, a GP model is concerned with evaluating the probability $P(t_{x_{n+1}}, D, x_n)$, where the new input vector is $x_{n+1}$ and the corresponding output is $t_{x_{n+1}}$. In this process, the $P(t_{x_{n+1}} | C, [x_n])$ is assumed to follow a Gaussian distribution as given by:

$$P(t_{x_{n+1}} | C, [x_n])$$
\[
P(t_x | \mathbf{C}_x, \{x_i\}) = \frac{1}{\sqrt{(2\pi)^{N} | \mathbf{C}_x |}} \exp \left[ -\frac{1}{2} ( \mathbf{t}_x - \mathbf{\mu})^T \mathbf{C}_x^{-1} ( \mathbf{t}_x - \mathbf{\mu}) \right]
\]

where \( \mathbf{t}_x = (t_1(x_1), t_2(x_2), \ldots, t_N(x_N)) \), \( \mathbf{C}_x \) is the covariance matrix for \( P(t_x | \{x_i\}) \), and \( \mathbf{\mu} \) is the mean, which will be zero for properly normalized data. In order to actually calculate either the mean or variance from these equations, some type of prior information that allows inference on the values of the covariance is needed. At the same time, the covariance matrix of the training input data should be properly chosen because this function determines the correlations among the training data outputs. Once the GP is trained, the optimal values of the hyperparameters are then used to construct the covariance matrix, which, in turn, is inverted and used to give predictions for both the predicted response and the predicted variance simultaneously through the following formulas:

\[
\begin{align*}
\hat{t}_{x,t} &= k^T \mathbf{C}_x^{-1} \mathbf{t}_x \\
\sigma^2_{x,t} &= \mathbf{\kappa} - k^T \mathbf{C}_x^{-1} k
\end{align*}
\]

where \( k = (C(x_t, x_{t,1}), C(x_t, x_{t,2}), \ldots, C(x_t, x_{t,N})) \), and \( \mathbf{\kappa} = C(x_t, x_{t,t}) \). The application of using the covariance function (Equation 3) produces a covariance matrix of \( N \times N \) dimension. This results in a training time of GP scales to \( N^3 \). The predicted variance can be used as a measure of importance for each prediction to reflect its accuracy in the objective space as to its significance in the surrogate models. For the details of the GP method, the reader is referred to [14].

2.2 Nondominated Sorting of Predicted Variances and Convergence of an Adaptive Multi-Objective Optimization System

Unlike single-objective optimization, the GP surrogate models for multi-objective optimization problems provide multiple predictions corresponding to the desired objective values. At the same time, they also output multiple corresponding variances for the predictions. In order to intelligently select additional training samples in the design spaces to improve the surrogate models and adaptively attain the Pareto-optimal solutions, nondominated sorting of the predicted variances and levels of adjacency have been implemented for the adaptive optimization system. The goals are to select the samples with the largest variances via nondominated sorting as the new training samples and to evaluate the degree of convergence of the adaptive multi-objective optimization process.

Since the variances of probability offers the estimations of the corresponding predictions, large predicted variances indicate poor predictions at these positions. Therefore, additional simulations need to be run at those positions with large predicted variances in order to improve the surrogate models. Essentially, selecting the additional training samples based on multiple predicted variances is similar to performing the maximum multi-objective optimization evaluation at the predicted variance space, in the sense that the samples at the forefront of nondominated sorting fronts have the largest variances. Therefore, a fast nondominated sorting approach is performed for the predicted variances at each iteration step [15], and the samples will be randomly selected from the forefront nondominated sorting front. After that, these selected samples will be evaluated with the simulations and act as the additional training samples to update the surrogate models.

Considering the need of additional CPU-intensive simulations in each iteration step for updating the surrogate models, the number of iteration steps used in the adaptive optimization procedure becomes a major issue. This issue deals with ensuring the quality of the Pareto-optimal front while using the minimum allowed number of iteration steps. By investigating the convergence of the adaptive optimization system, it may be possible to find an interaction between the number of iteration steps evaluated and the quality of the Pareto-optimal fronts. Understanding this interaction will enable effective evaluations of optimization problems. To deal with the convergence of the optimal solution from different surrogate models, a method to represent the Pareto-optimal front with user defined indifference thresholds is proposed and described in the following steps [16]:

1. Setup indifference thresholds for different objectives. This threshold refers to the change in each objective function within which the optimal solutions are indifferent to each other.
2. Discretize the objective space using the defined indifference thresholds. Using the indifference thresholds to establish the discretization sizes for each objective, the objective space is then divided into a collection of "hyperboxes." For example, for a problem with three objectives, the performance space (now in three-dimensions) would be discretized into a set of rectangular cuboids, where the size of each cuboid is the indifference threshold value for each objective.
3. Represent the Pareto-optimal front with a collection of hyperboxes in the objective space. Within the discretized objective space, it is easy to capture the Pareto-optimal front for the current GP surrogate models.

*The Polymer Processing Society 23rd Annual Meeting*
After capturing and representing the Pareto-optimal fronts with a collection of "hyperboxes" for different iteration steps, the concept of levels of hyperbox adjacency is proposed to evaluate the convergence of iteration stages. Using different adjacency constraints, it becomes possible to provide information about how well the different iteration steps capture the behavior of the optimal fronts. For example, the iteration process is considered converged if the change in the configuration of the hyperboxes that represent Pareto-optimal fronts meets the termination criterion based on levels of adjacency.

Two hyperboxes can be considered adjacent to each other as long as the absolute difference in any given dimension is no greater than one. Using this rule, there are four levels of adjacency possible in a three-dimensional objective space [16]: (1) the first level of adjacency is when all indices of the two hyperboxes being compared are the same in all dimensions; (2) when the two hyperboxes compared are the same in all but one dimension, the two hyperboxes that are adjacent share a common face; (3) when the indices comparison holds for all but two dimensions, the two compared hyperboxes share a common edge; and (4) when the indices of the hyperboxes are the same in all but three dimensions, the two compared hyperboxes share only a common point. This concept can be expanded into n-dimensional objective space for considering how much flexibility in adjacency the designer is willing to allow. Therefore, applying the user defined adjacency constraints to the Pareto-optimal fronts optimized using different surrogate models allows for evaluation of the convergence in the adaptive iteration stage.

3. Implementation of the Adaptive Optimization Procedure

Based on the GP regression method, a novel adaptive multi-objective optimization system is implemented. The system requires the user to specify the optimization criteria in terms of objective functions and constraints. This will allow a variety of optimization objectives, such as shorter cycle time, minimal injection pressure, better temperature uniformity, reduced residual stresses, minimal part shrinkage and warpage, etc, depending on the designers' requirements [6]. It should be pointed out that, due to the generic nature of this optimization scheme, any of the commercial or research-oriented simulation programs can be incorporated into this procedure.

In injection molding, many different parameters affect the outputs of the process. However, as training samples grow, the efficiency of data collection deteriorates due to the extensive run-time of computer simulations. It is observed that more representative training data would give the surrogate models better approximation accuracy, and therefore save the sample collecting and computing time. In this proposed procedure, the latin hypercube sampling (LHS) method, which is a stratified random sampling technique, is used for generating the samples needed for the initial training of the GP surrogate models [17]. LHS allows exceedingly sparse M point samples to be selected for an input design space with N dimensions, so that the number of sampling data is reduced while the behavior of the injection molding process simulation can be adequately captured. The criterion for the selection of the LHS that best represents the objective space is to maximize the minimum (Euclidean) distance of the sample points, thus making the sample points more scattered in the objective space. Therefore, LHS is useful when a simulation scheme uses samples that are random but relatively uniformly distributed over each dimension.

In the case of multi-objective optimization for injection molding, there are trade-off behaviors and thus various optimal solutions may exist in the objective space. These solutions are known as Pareto-optimal solutions, in which none of the solutions are better than the other without any further information. Since the genetic algorithms (GA) can simultaneously deal with a diverse population of sample points in each generation, it can adequately find the multiple Pareto-optimal solutions in one single simulation run while moving toward the true Pareto-optimal region. Once the entire set of Pareto-optimal solutions is achieved, the designers can choose one solution over the others according to the problem knowledge and/or based on a number of problem-related factors.

In this work, an elitist nondominated sorting genetic algorithm (NSGA-II) is coupled with the proposed procedure to perform the multi-objective optimization [15, 18]. The multi-objective optimization procedure includes the following two stages: the training stage and the iteration stage. The user-defined different adjacency constraints percentage will be used to evaluate the convergence of termination criteria.

Training Stage (cf. Figure 1):

The steps in the training stage include the following:

1. Determine the ranges of injection molding process conditions and the multiple objectives.
2. Perform the design of computer experiments over the expected feasible design space to achieve the simulation schemes. Each set of process conditions generated by the Latin hypercube sampling (LHS) method will be analyzed by simulation.
3. Execute simulation programs for all of the samples generated by LHS; collect the simulation results and extract desired data for establishing the preliminary GP surrogate models.
4. Fit a GP model to the results.

After the training stage, the GP models can produce predictions for both the multiple responses and the variances of responses simultaneously. These predicted variances can serve as the estimations for the accuracy of the predictions, and therefore, they can be used for concurrently updating and improving the GP surrogate models.

![Figure 1: Training stage using LHS selection and GP.](image)

**Figure 1: Training stage using LHS selection and GP.**

**Iteration Stage (cf. Figure 2):**

1. Assess the current surrogate models with the randomly generated validation samples within the design space.
2. Select the additional training data set for updating the GP surrogate models:
   a. Perform a feasibility check for the validation samples according to the constraints. If the samples are known to be infeasible, discard them from the current dataset.
   b. Perform a nondominated sort of the predicted variances, and randomly select the samples at the forefront of the nondominated sorting front.
3. Run additional simulations with the selected sets of process conditions from the above step, and update the GP surrogate models with the newly available training samples.
4. Use NSGA-II to find the Pareto-optimal solutions for current GP models.
5. Discretize the objective space into hyperboxes with user-defined indifference thresholds, and then represent the current Pareto-optimal front with a collection of hyperboxes.
6. Setup the percentages of the different levels of adjacency, and use the adjacency levels to evaluate the convergence of termination criteria. If the termination criteria are not met, repeat the process or stop after a predetermined number of iterations.

A major aspect of the new modeling approach in this iteration stage is the inclusion of runtime surrogate models with improved capabilities derived from the Bayesian nature of the GP method. Therefore, after a number of iteration steps, the GP surrogate models will become more accurate and consequently proven, and the Pareto-optimal solutions will be identified with genetic generations using the successively updated surrogate model.

In summary, the proposed adaptive multi-objective optimization system has the following characteristics: (1) It uses a specific design of computer experiments, the Latin hypercube sampling technique, to generate initial samples needed for training the surrogate models, which can perfectly capture the general behaviours of injection molding simulation with a relative small number of training samples; (2) It is capable of giving both predictions and estimations of the prediction simultaneously, thus providing direction as to where additional training samples could be added to improve the surrogate models; (3) It enables an adaptive process, in which the procedure will perform a nondominated sort of the predicted variances at each iteration step, thus adaptively updating the surrogate models whenever new training data become available; and (4) It applies different user-defined adjacency constraints to automatically evaluate the convergence of Pareto-optimal fronts for the optimization procedure.
4. Multi-objective Process Conditions Optimization for Optical Precision Lens

In this example, the proposed adaptive optimization system will be used to perform the multi-objective process conditions optimization for injection molding of a high-precision optical lens part (cf. Figure 3).

The outside diameter of the lens part is 96.19 mm and the height of the part at the center is 19.87 mm. The thickest part of the lens is at the outer rim of the lens and is 10.50 mm, while the thickness area at the center of the part is 6.00 mm. The weight of the full shot lens part is 69.8 g. In this illustrative case, there are three concerns regarding the production economics and part quality: (1) cycle time, which should be kept as short as possible in order to increase output and decrease manufacturing cost; (2) maximum injection pressure, which should be minimized to prolong machine life and reduce power consumption; and (3) volumetric shrinkage, which should be minimized to improve molded part quality.
Three outputs from simulation results—cooling time, maximum injection pressure, and volumetric shrinkage—are selected as the objective values to represent the above criteria, respectively. Thus, the optimization problem with the ranges of process conditions to be optimized are defined as follows:

\[
\begin{align*}
\text{Minimize :} & \quad \text{cooling time} \\
\text{Minimize :} & \quad \text{maximum injection pressure} \\
\text{Minimize :} & \quad \text{volumetric shrinkage}
\end{align*}
\]

Subject to:
\[
\begin{align*}
220^\circ \text{C} & \leq T_m \leq 250^\circ \text{C} \\
60^\circ \text{C} & \leq T_w \leq 80^\circ \text{C} \\
6 \text{s} & \leq t_{inj} \leq 10 \text{s}
\end{align*}
\]

In this application, three independent process conditions, namely, melt temperature \(T_m\), mold temperature \(T_w\), and injection time \(t_{inj}\), are optimized to achieve the desired objectives. Since this molded part has a complex curved geometry and relatively thick regions, traditional 2.5-dimensional simulation methods based on Hele-Shaw mid-plane models may not provide correct solutions. Therefore, the relationships between the process conditions and the objective functions are mapped with Moldex3D, a commercial 3D injection molding simulation program [19].

At first, 15 process condition combinations generated with the LHS technique were simulated for establishing the preliminary GP surrogate models. After the training stage, the initial surrogate models were adaptively updated to achieve the improved Pareto-optimal solutions during the iteration stage. Using the GP surrogate models, it took less than 3 seconds to perform the multi-objective function evaluations on a Dual Xeon workstation with 1 GB RAM. According to the proposed optimization procedure, at each iteration step, 500 validation samples that would potentially be evaluated for the performance of the GP models were randomly generated within the objective space. Subsequently, the corresponding predicted variances for these validation samples were nondominated sorted. After randomly selecting the samples at the forefront of the nondominated sorting front, additional simulations were performed with these selected process conditions and the corresponding simulation results were collected for improving the overall predictive capability of the GP surrogate models. At the same time, the NSGA-II was used to evaluate the GP models to find the current Pareto-optimal front, and then with the user-defined indifference thresholds, the current front was captured with the collection of hyperboxes. Following the iteration stage introduced in Section 3, and setting up the adjacency constraints percentages for different levels of adjacency, after five steps, the iteration stage stopped as the termination criterion was reached. The finalized GP models that satisfied the termination criterion were built with overall 38 training samples. Table 1 shows the corresponding optimization parameters in each of the iteration steps. The resultant GP surrogate models were validated over all available data to estimate the regression accuracy. The predicted objective values from the surrogate models are compared with the training data from simulation in Figure 4, in which the solid line represents the best fit whereas the dotted line corresponds to the unity slope. It shows that the predictions agree well with the data from the simulations, thus it confirms that the predictive ability of model constructions is adequate enough for further optimization with NSGA-II to achieve the Pareto-optimal solutions for this application.

| Table 1. Number of training samples and optimization parameters for each iteration step. |
|-------------------------------------------------|-------|-------|-------|-------|-------|-------|
| Training samples                                | 15    | 20    | 24    | 29    | 35    | 38    |
| Samples at the forefront of the nondominated sorting front | 22    | 16    | 15    | 19    | 13    | –     |
| Additional training samples                    | 5     | 4     | 5     | 6     | 3     | –     |
| Iteration parameters (percentages of the different levels of adjacency) | Exact overlap: 65%; All but one dimension: 15%; All but two dimensions: 15%; All but three dimensions: 5% |
| Optimization parameters for NSGA-II            | Initial population: 200 | Number of generations: 100 |
|                                                | Crossover: 0.9      | Mutation: 0.2               |

The Polymer Processing Society 23rd Annual Meeting
The three-dimensional Pareto-optimal front is displayed in Figure 5. The predicted responses are nonlinearly correlated to the process conditions, which feature complex interactions. Based on the achieved GP surrogate models, it shows that in this specific application there exists direct trade-off behaviors between the maximum injection pressure and the volumetric shrinkage as expected, in which the volumetric shrinkage will decrease with increasing maximum injection pressure, and vice versa. Through the analysis of variance, the importance of the process conditions can ranked according to their significance in the surrogate models, and interactions between process conditions and their influence on the corresponding objectives can assessed. For example, the maximum injection pressure is largely affected by the injection time and melt temperature. Meanwhile, all three process condition factors have significant main and interaction effects for cycle time and volumetric shrinkage. After establishing the finalized GP surrogate models and employing the NSGA-II to achieve the Pareto-optimal solutions, different combinations of optimal trade-off solutions can be selected from these solutions according to the designer’s preference for setting up the process conditions. Table 2 shows a comparison of four Pareto-optimal solutions based on the GP model predictions and computer simulation under the corresponding optimal process conditions. The average difference percentage is relatively small (6.59%), which confirms the superior predictive abilities of GP and the effectiveness of the proposed optimization scheme. Figure 6 shows the simulation results of the sprue pressure and the filling time with the process conditions ($T_m = 225.23^\circ C$, $T_o = 62.8^\circ C$, and $t_{inj} = 8.60s$) that can achieve the trade-off objectives: the lower maximum injection pressure but longer cooling time.

Compared to the optimization process for the same application with the SVR surrogate modeling approach that was reported previously [11], the current approach executed less simulations for objective function evaluations and achieved better solutions, thus the whole optimization time is greatly reduced due to the system’s adaptive and automatic properties. Therefore, with the help of this adaptive simulation-based optimization system, the optimization task specified in this application can be accomplished with a reasonably small amount of computing resources while still yielding reasonable results. Although the procedure needs a relatively long time to execute 3D simulations for obtaining the initial training data, the subsequent optimization process could realize a lot of benefits from the trained GP models.
<table>
<thead>
<tr>
<th>Table 2. Comparison between the GP model with simulation under the corresponding process conditions.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Optimal process conditions</td>
</tr>
<tr>
<td>$T_{melt}$ ($^\circ$C)</td>
</tr>
<tr>
<td>$T_{mold}$ ($^\circ$C)</td>
</tr>
<tr>
<td>$t_{inj}$ (s)</td>
</tr>
<tr>
<td>Cooling time (s)</td>
</tr>
<tr>
<td>Surrogate model’s prediction</td>
</tr>
<tr>
<td>Simulation results</td>
</tr>
<tr>
<td>Diff (%)</td>
</tr>
<tr>
<td>Maximum injection pressure (MPa)</td>
</tr>
<tr>
<td>Surrogate model’s prediction</td>
</tr>
<tr>
<td>Simulation results</td>
</tr>
<tr>
<td>Diff (%)</td>
</tr>
<tr>
<td>Volumetric shrinkage (%)</td>
</tr>
<tr>
<td>Surrogate model’s prediction</td>
</tr>
<tr>
<td>Simulation results</td>
</tr>
<tr>
<td>Diff (%)</td>
</tr>
</tbody>
</table>

5. Conclusions

In this study, a new multi-objective injection molding optimization procedure based on the Gaussian process (GP) surrogate modeling approach and design of computer experiments is presented and shown to be theoretically sound and practically applicable to the multi-objective process optimization of injection molding. Capable of enabling an intelligent optimization strategy and revealing the complicated correlations between process conditions and objectives, the present GP method is explored with data samples gained from the Latin hypercube sampling method. The foundation in the Bayesian probability and inference approach allow the GP surrogate models to accurately provide predictions and estimations of the confidence of predictions simultaneously. Using the nondominated sorting procedure for the predicted variances at each iteration step, and the user-defined adjacency constraints percentage for evaluation of the convergence of adaptive iteration stage, the new optimization system can intelligently determine the optimization direction and adaptively select the additional samples that can greatly improve the surrogate models. Based on the results of the illustrative application, it can be demonstrated that the GP method is capable of providing useful information through a data-driven model-building process in optimization problems that feature limited data samples and high dimensionality as well as complex and nonlinear interactions between process conditions and objectives. The new system proposed here can effectively establish the surrogate models, and adaptively and automatically search for the optimal process conditions for injection molding within a reasonable timeframe. Although the detailed GP surrogate models must be constructed with necessary computing resources, the extra benefits resulting from the quick global search for multi-objective optimal solutions make it worthwhile. Furthermore, the abundant information derived from the surrogate models may help in discovering key underlying mechanisms, and therefore greatly improve the worth of the optimization study.

6. Acknowledgments

This work has been supported by the 3M Precision Optics, Inc., and the National Science Foundation (DMI-0323509). The authors wish to thank Roy Auerbach, James P. Ruth, and Kathleen Allen of 3MPO for the experimental data.

References