Article

A Confidence Interval-Based Process Optimization Method Using Second-Order Polynomial Regression Analysis

Jungwon Yu ^{1,*}, Soyoung Yang ¹, Jinhong Kim ¹, Youngjae Lee ¹, Kil-Taek Lim ¹, Seiki Kim ², Sung-Soo Ryu ² and Hyeondeok Jeong ^{2,*}

¹ Electronics and Telecommunications Research Institute, Daegu-Gyeongbuk Research Center, Daegu 42994, Korea; syyang@etri.re.kr (S.Y.); jinhong@etri.re.kr (J.K.); lyj4295@etri.re.kr (Y.L.); ktl@etri.re.kr (K.-T.L.)

² Korea Institute of Ceramic Engineering & Technology, Icheon Branch, Engineering Ceramics Center, Icheon-si 17303, Korea; kimseiki@kicet.re.kr (S.K.); ssryu@kicet.re.kr (S.-S.R.)

* Correspondence: gardenyoo@etri.re.kr (J.Y.); jhd526@kicet.re.kr (H.J.)

Received: 28 August 2020; Accepted: 19 September 2020; Published: 24 September 2020

Abstract: In the manufacturing processes, process optimization tasks, to optimize their product quality, can be performed through the following procedures. First, process models mimicking functional relationships between quality characteristics and controllable factors are constructed. Next, based on these models, objective functions formulating process optimization problems are defined. Finally, optimization algorithms are applied for finding solutions for these functions. It is important to note that different solutions can be found whenever these algorithms are independently executed if a unique solution does not exist; this may cause confusion for process operators and engineers. This paper proposes a confidence interval (CI)-based process optimization method using second-order polynomial regression analysis. This method evaluates the quality of the different solutions in terms of the lengths of their CIs; these CIs enclose the outputs of the regression models for these solutions. As the CIs become narrower, the uncertainty about the solutions decreases (i.e., they become statistically significant). In the proposed method, after sorting the different solutions in ascending order, according to the lengths, the first few solutions are selected and recommended for the users. To verify the performance, the method is applied to a process dataset, gathered from a ball mill, used to grind ceramic powders and mix these powders with solvents and some additives. Simulation results show that this method can provide good solutions from a statistical perspective; among the provided solutions, the users are able to flexibly choose and use proper solutions fulfilling key requirements for target processes.

Keywords: second-order polynomial regression analysis; confidence interval; process optimization; ball mill

1. Introduction

In manufacturing, one of the major concerns faced by the industry is in saving production costs and time, and at the same time, improving the quality of the products. These can be achieved by optimizing crucial controllable factors (also known as input, process, or design parameters) in target processes [1,2]. To obtain the best values of the controllable factors, we must first formulate functional relationships (i.e., process models) between these factors and quality characteristics; regression analysis and machine learning methods can be used to construct these relationships. After finishing the model-building, objective functions are defined based on the models, and then optimization techniques are applied to these functions to find optimal solutions (also known as optimal recipes).

To build the process models, we prepare a process dataset, composed of input-output observation pairs, beforehand. Experimental designs [3], such as full factorial, central composite, and Box-Behnken designs have been widely employed to collect these datasets in systematic ways. Second-order polynomial regression analysis [2,4,5] is the most popular statistical technique for constructing the process models from the collected datasets; this technique has been successfully used for process optimization in various fields [6–15]. The advantages of this regression technique can be summarized as follows. First, since the number of levels (i.e., the number of points at which experiments are performed) in each factor, are in general, set as 3 or 4, the second-order polynomial regression models can capture the nonlinearities contained in the process data adequately; as the order of the models become larger than 2, the model complexity (i.e., the number of regression coefficients) exponentially increase, and as a result the resulting models overfit the target datasets. Second, when the input-output relations are described by these models, an output variable is nonlinear in input variables but is linear in the coefficients; one can easily obtain least squares estimates for the unknown coefficients, which are closed-form solutions minimizing residual sum-of-squares. Last, the constructed models, based on the method of least squares, are highly interpretable; by examining the estimated coefficients closely, one can understand which terms affect the output more than the others.

After building the polynomial regression models mimicking the input-output relationships of target processes, one can try to find input values (i.e., the values of controllable factors) that make the model outputs (i.e., quality characteristics) equal to desired values (e.g., maximum, minimum, or target values) through optimization procedures. These models are also capable of predicting the output values corresponding to arbitrary input points at which experiments were not conducted. If the models have a single output and the aim of optimization procedures is to maximize or minimize it, we make use of regression functions themselves as the objective functions. If there are user-defined target values for the output values, one should define the objective functions by modifying the regression functions suitably.

One of the major difficulties encountered in the optimization procedure is that a unique solution may not exist. In particular, when the goal is to find a solution that makes the output values equal to user-defined target values, one cannot find a unique solution. This can be easily confirmed from surfaces of regression functions (also known as response surfaces). In this case, different solutions can be found whenever optimization algorithms are independently executed; this is common to both derivative-based and derivative-free optimization methods. In the derivative-based methods (e.g., numerical optimization methods [16]), if searching the solutions starts at different initial points, discovered solutions may be changed; in the derivative-free methods (e.g., metaheuristics [17]), stochastic mechanisms are employed to find the solutions, and thus different solutions are obtained each time. In summary, as explained above, a unique solution may not exist in process optimization problems, and therefore, different solutions are discovered when running optimization algorithms multiple times; what seems to be lacking is the study to determine which of these different solutions to be selected and then recommended for users, such as process operators and engineers.

In this paper, we propose a confidence interval (CI)-based process optimization method using second-order polynomial regression analysis; it aimed at assessing the quality of the different solutions from a statistical point of view and selecting statistically significant solutions to be provided to the users. The second-order polynomial regression models give us closed-form expressions able to calculate CIs in which the model outputs for the solutions are included. The proposed method uses the length of these CIs as a measure to evaluate the different solutions discovered by optimization techniques; among the different solutions, only a few solutions whose lengths of CIs are shorter than others, are selected and recommended for the users.

To verify the performance, we apply the proposed method to a process dataset collected from a ball mill. In ceramic industries, this equipment has been popularly used to pulverize ceramic powders and to mix these powders with solvents and some additives [18]; finishing the operation of a ball mill, one can obtain a slurry where they are mixed with each other. Among various measurable variables

from the slurry, particle size distribution and viscosity have been regarded as key quality characteristics. The focus of this study is limited to the slurry viscosity. Central composite design (CCD) [2,3] is used to prepare the experiment dataset required to construct process models. The purpose of applying the proposed method to the ball mill is to find the values of its controllable factors, achieving the values of slurry viscosity, equal to target values. The importance of achieving its desired values are attributed to the fact that it plays a vital role in a next stage process, i.e., a spray dryer. This is a unit process to produce granules by spray drying the slurry. Therefore, the slurry should satisfy the viscosity values required to obtain high-quality (i.e., uniform and free-flowing) granules [18]. The proposed method is able to recommend several reliable solutions (i.e., the values of controllable factors achieving desired values of slurry viscosity) from a statistical viewpoint. In accordance with practical situations, the users can flexibly choose and use proper solutions among them.

The remainder of this paper is organized as follows. Section 2 explains the second-order polynomial regression method for process modeling, CIs enclosing model outputs, and Monte Carlo (MC)-based method to estimate the importance of controllable factors; one can prioritize these factors according to the estimated importance values, and their priorities can be quite useful for the operations of target processes. Section 3 outlines particle swarm optimization (PSO), one of the widely used global optimization methods, and the proposed CI-based process optimization method; in fact, any optimization methods can be employed in the proposed method. Section 4 describes a target process (i.e., a ball mill) and a process dataset gathered from it via CCD. Section 5 presents the simulation results and discussion, and finally, we give our conclusions in Section 6.

2. Process Modeling Using Second-Order Polynomial Regression

In this section, the second-order polynomial regression analysis and the MC-based method inspired from Ref. [10,19] are explained in detail. From now on, scalars and vectors are written in italics, and bold lowercase, respectively, and matrices are written in bold capitals.

2.1. Second-Order Polynomial Regression Analysis

The 2nd-order polynomial regression models describe the functional relationships between p input variables (also called independent, explanatory, or predictor variable) $x_1, \dots, x_p$ and a single output variable (also called dependent, target, or response variable) y as follows,

$$y = \beta_0 + \sum_{j=1}^p \beta_j x_j + \sum_{j < k} \beta_{jk} x_j x_k + \sum_{j=1}^p \beta_{jj} x_j^2 + \varepsilon \quad (1)$$

where β_0 is an intercept, and β_j , β_{jk} , and β_{jj} are regression coefficients relevant with main effect, interaction effect, and quadratic effect terms, respectively; in Equation (1), the number of model parameters (i.e., regression coefficients) to be estimated is $p' = 1 + 2p + p(p-1)/2$; ε is an error term and it includes all possible errors, such as the measurement error for y , the error incurred by setting regression functions as second-order polynomial functions and the error incurred by omitting the other input variables except for the p inputs, and so on. Based on the central limit theorem, it is usually assumed that the error term follows Gaussian distribution with mean 0 and variance σ_ε^2 .

Since the output variable y is linear in the p' regression coefficients, one can use the method of least squares to estimate them. Let $\mathbf{y} = [y_1, \dots, y_n]^T \in \mathbb{R}^n$ and $\mathbf{Z} = [\mathbf{z}_1, \dots, \mathbf{z}_n]^T \in \mathbb{R}^{n \times p'}$ be the output vector consisting of n measurements for y and the design matrix, respectively; the i th row of the matrix $\mathbf{Z}$ is $\mathbf{z}_i^T = \mathbf{z}(\mathbf{x}_i)^T = [1, x_{i1}, \dots, x_{ip}, x_{i1}x_{i2}, \dots, x_{i(p-1)}x_{ip}, \dots, x_{i1}^2, \dots, x_{ip}^2]$. Least squares estimates for the parameter vector $\beta \in \mathbb{R}^{p'}$ composed of the p' coefficients can be obtained as follows: $\hat{\beta} = (\mathbf{Z}^T \mathbf{Z})^{-1} \mathbf{Z}^T \mathbf{y}$. After estimating the regression coefficients, the output value (also known as predicted value) of the

model for an arbitrary input $\mathbf{x}_0 = [x_{01}, \dots, x_{0p}]^T$ can be calculated as $\hat{y}_0 = \mathbf{z}_0^T \hat{\boldsymbol{\beta}}$. A $100(1 - \alpha)\%$ CI for the predicted value $\hat{y}_0$ is defined as,

$$\hat{y}_0 \pm t_{1-\alpha/2}(n-p')s \sqrt{\mathbf{z}_0^T (\mathbf{Z}^T \mathbf{Z})^{-1} \mathbf{z}_0} \quad (2)$$

where $s = \sqrt{\frac{1}{n-p'} \sum_{i=1}^n (y_i - \hat{y}_i)^2}$, and $\hat{y}_i = \mathbf{z}_i^T \hat{\boldsymbol{\beta}}$, $i = 1, \dots, n$; s is an unbiased estimator for standard deviation σ_ε of the error term ε ; $t_{1-\alpha/2}(n-p')$ is the $1 - \alpha/2$ percentile of the t distribution with degree of freedom $n - p'$. More details regarding the (2nd-order polynomial) regression methods are available elsewhere [2,4,5].

2.2. Importance Estimation of Input Variables

This subsection describes a MC-based method to estimate the variable importance of the p controllable factors; this is a modified version of the method presented in [10,19]. In general, when using second-order polynomial regression models, one can confirm how much each term contributes to the output by closely examining the magnitude of the corresponding estimated coefficient and its t -statistic. The larger the magnitude of a certain estimated coefficient, the greater the amount of output variation according to an increase or decrease of the corresponding term. In addition, as the t -statistic of a certain estimated coefficient becomes larger, statistical significance of the hypothesis that the corresponding term clearly influences the output variable increases. By doing so, we can identify the importance of individual terms (i.e., x_j , $x_j x_k$, and x_j^2), but we cannot confirm which original input variables (i.e., x_j , $j = 1, \dots, p$) are highly significant for predicting the output values. In this paper, the MC-based method is performed to estimate the importance of the original input variables (not individual terms) in Equation (1) as follows:

- Step 1:** Generate N uniform random vectors $\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(N)}$ with p dimensions; the j th element of these vectors follows uniform distribution over the interval $[-1, 1]$.
- Step 2:** Calculate N outputs $\hat{y}^{(1)}, \dots, \hat{y}^{(N)}$ by substituting the N vectors $\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(N)}$ in the constructed model, and then obtain the median y_{med} of $\hat{y}^{(i)}$, $i = 1, \dots, N$.
- Step 3:** Divide the N vectors $\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(N)}$ into two sets X_{larger} and X_{smaller} as follows: $X_{\text{larger}} = \{\mathbf{x}^{(i)} | \hat{y}^{(i)} \geq y_{\text{med}}, i = 1, \dots, N\}$ and $X_{\text{smaller}} = \{\mathbf{x}^{(i)} | \hat{y}^{(i)} < y_{\text{med}}, i = 1, \dots, N\}$.
- Step 4:** Estimate two cumulative distribution functions (CDFs) for each input variable empirically on the basis of the vectors belonging to X_{larger} and X_{smaller} , respectively; that is, x_j , $j = 1, \dots, p$, has two estimated CDFs related with X_{larger} and X_{smaller} , respectively.
- Step 5:** Calculate the total area of the region(s) surrounded by the two CDFs of each input variable by numerical integration techniques.

As the output variations become larger according to the changes of x_j , the calculated area for x_j becomes wider. In this paper, the calculated areas in Step 5 are regarded as the importance values of controllable factors, and variable ranking can also be carried out based on these values.

3. Process Optimization

After constructing data-driven process models as described in Section 2, the values of controllable factors can be optimized based on these models. The input vectors that maximize or minimize the model outputs are located at the boundary of the p -dimensional input space or at the points where the gradient vector $\nabla_{\mathbf{x}} \hat{y}(\mathbf{x})$ of the regression functions $\hat{y}(\mathbf{x}) = \mathbf{z}(\mathbf{x})^T \hat{\boldsymbol{\beta}}$ is equal to zero vector. The input vectors satisfying the condition $\nabla_{\mathbf{x}} \hat{y}(\mathbf{x}) = 0$ are called stationary points, and analyzing the output behaviors at these points is known as canonical analysis [2].

In multi-stage manufacturing processes [20,21], finding the values of controllable factors that make them the same as pre-determined target values is much more needed than maximizing or minimizing quality characteristics. In these processes, some (or all) of quality characteristics in the current

stage process are considered as design parameters for the next stage process. Therefore, the target values in the current stage process should be decided by taking the requirements for the next stage process into account; the objective function formulating these optimization problems can be defined as $F(\mathbf{x}) = \frac{1}{2}(y_{\text{target}} - \hat{y}(\mathbf{x}))^2$ or $F(\mathbf{x}) = |y_{\text{target}} - \hat{y}(\mathbf{x})|$, where y_{target} denotes the target value. The goal of process optimization is to find the solution $\mathbf{x}^* = \operatorname{argmin}_{\mathbf{x}} F(\mathbf{x})$, and to do this, various optimization algorithms [16,17,22–25] can be applied. In this paper, we make use of PSO algorithm whose implementation is relatively simple and mechanisms for searching the solutions are intuitive. Since the estimated regression functions by second-order polynomial regression analysis have smooth response surfaces, basic optimization algorithms such as PSO can well be applicable to find the solutions for $F(\mathbf{x})$ defined by the regression functions.

3.1. Particle Swarm Optimization

PSO mimicking the social behavior of bird flocks is a population-based search algorithm to obtain a global optimum; in PSO, particles composing a swarm try to find a global optimum for objective functions through flying in their high-dimensional domain space. The behavior of each particle depends on both its own experiences (i.e., cognitive component) and those of the others (i.e., social component). Let $\mathbf{x}_i(t)$ be a position of the i th particle at time t ; at each time, this position is changed by,

$$\mathbf{x}_i(t+1) = \mathbf{x}_i(t) + \mathbf{v}_i(t+1), \quad i = 1, \dots, n_s, \quad (3)$$

where $\mathbf{x}_i(t+1)$ is the changed position, $\mathbf{v}_i(t+1)$ is the velocity vector that changes its position, and n_s is the number of particles belonging to the swarm. In general, the particle positions are initialized to follow uniform distribution as follows: $x_{ij}(0) \sim U(x_{j,\min}, x_{j,\max})$, $i = 1, \dots, n_s$, $j = 1, \dots, p$, where $x_{ij}(0)$ is the j th component of $\mathbf{x}_i(0)$; $x_{j,\min}$ and $x_{j,\max}$ are the upper and lower limits of the j th components, respectively. The velocity vector driving their stochastic movements are updated as,

$$\mathbf{v}_i(t+1) = w(t)\mathbf{v}_i(t) + c_1(t)\mathbf{r}_1(t)^T(\mathbf{x}_i^{pbest}(t) - \mathbf{x}_i(t)) + c_2(t)\mathbf{r}_2(t)^T(\mathbf{x}^{gbest}(t) - \mathbf{x}_i(t)) \quad (4)$$

where $\mathbf{x}_i^{pbest}$ is the best position discovered by itself so far (i.e., personal best position), and $\mathbf{x}^{gbest}$ is the best position discovered by entire particles so far (i.e., global best position); $c_1(t)$ and $c_2(t)$ are positive acceleration constants to assign weights to cognitive, and social components, respectively; $\mathbf{r}_1(t)$ and $\mathbf{r}_2(t)$ are random vectors and each of their elements is sampled from uniform distribution with the range $[0, 1]$; $w(t)$ is the inertia weight to control the contribution of $\mathbf{v}_i(t)$ to $\mathbf{v}_i(t+1)$. The velocity vectors are usually initialized to zero vectors, i.e., $\mathbf{v}_i(0) = \mathbf{0}$, $i = 1, \dots, n_s$. At each time step, $\mathbf{x}_i^{pbest}$ and $\mathbf{x}^{gbest}$ are updated, respectively, as follows:

$$\mathbf{x}_i^{pbest}(t+1) = \begin{cases} \mathbf{x}_i^{pbest}(t) & \text{if } F(\mathbf{x}_i(t+1)) \geq F(\mathbf{x}_i^{pbest}(t)) \\ \mathbf{x}_i(t+1) & \text{if } F(\mathbf{x}_i(t+1)) < F(\mathbf{x}_i^{pbest}(t)) \end{cases} \quad (5)$$

$$\mathbf{x}^{gbest}(t+1) = \operatorname{argmin}_{\mathbf{x} \in X^{pbest}} F(\mathbf{x}), \quad \text{where } X^{pbest} = \{\mathbf{x}_i^{pbest}(t+1) \mid i = 1, \dots, n_s\}. \quad (6)$$

In population-based research algorithms, it is well-known that the diversity of particles should be guaranteed in the early stages and, at the same time, they should have good convergence properties in the later stages. In Equation (4), the inertia weight and acceleration coefficients influence the abilities of global and local searches (also known as exploration, and exploitation, respectively). Increasing the inertia weight improves the ability of global search (i.e., the diversity of the swarm); on the other hand, the smaller the inertia weight, the better the ability of local search. In this paper, the inertia weight $w(t)$ is decreased over time, as follows,

$$w(t) = (w(0) - w(t_{\max})) \times \frac{t_{\max} - t}{t_{\max}} + w(t_{\max}) \quad (7)$$

where $t_{\max}$ is the maximum number of iterations, and $w(0)$ and $w(t_{\max})$ are the initial and final values of $w(t)$, usually set as 0.9, and 0.5, respectively. PSO algorithms with Equation (7) focus on the global search in the early stages and on the local search in the later stages. In the case of acceleration coefficients, as c_2 becomes larger, a more weight is assigned to the social component and this is effective for objective functions with smooth surfaces; when the functions have rough multi-modal surfaces, it is beneficial to increase the value of c_1 . In this paper, these coefficients are changed at every time step, as follows,

$$c_1(t) = (c_{1\min} - c_{1\max}) \frac{t}{t_{\max}} + c_{1\max}, c_2(t) = (c_{2\max} - c_{2\min}) \frac{t}{t_{\max}} + c_{2\min} \quad (8)$$

where $c_{1\max}$ and $c_{2\max}$ are set as 2.5, and $c_{1\min}$ and $c_{2\min}$ are set as 0.5. For more details of PSO, readers are invited to see Refs. [26–29].

3.2. Proposed Method: Confidence Interval-Based Process Optimization

As described in Section 1, process optimization problems may not have a unique solution because of the nature of their objective functions. In these problems, whenever a PSO algorithm is applied, different solutions can be found. The initial starting point $\mathbf{x}_i(0)$ of the i th particle is initialized randomly, and random vectors $\mathbf{r}_1$ and $\mathbf{r}_2$ are involved in the update of its velocity vector $\mathbf{v}_i(t+1)$. If the different solutions are obtained whenever optimization algorithms are executed, this leads to confusion about what solutions are preferable relative to the others. Therefore, among these different solutions, it is critical to decide which solutions should be recommended for users (e.g., process operators and engineers). In this paper, the CI in Equation (2) is employed to evaluate the quality of these different solutions; as the length of CI for a solution becomes shorter, its uncertainty is reduced from a statistical point of view.

Algorithm 1 summarizes the proposed CI-based process optimization procedure. In the proposed method, first, a 2nd-order polynomial regression model is fitted to a prepared experiment dataset; based on this model, the objective function can be defined. Second, by applying the PSO algorithm to the objective function repeatedly (e.g., 100 times), many different solutions are obtained. Although, a PSO algorithm is employed here, it does not matter what optimization algorithms are used. Third, according to the lengths of their CIs, they are sorted in ascending order; these CIs can be calculated using Equation (2). Finally, the first few solutions whose lengths of CIs are shorter than the others are recommended for the users; among the recommended solutions, the users can select and use proper solutions flexibly through considering, for example, the amount of remaining materials and/or required time for process operation.

Algorithm 1. The proposed CI-based process optimization procedure.

- 1: **Input:** Experiment dataset gathered from a target process, $D = \{(\mathbf{x}_i; y_i) | i = 1, \dots, n\}$
 - 2: $\alpha \leftarrow$ significant level for CI
 - 3: $y_{\text{target}} \leftarrow$ target value for a quality characteristic
 - 4: $L \leftarrow$ the number of repetitions of optimization algorithm
 - 5: $L' \leftarrow$ the number of solutions to be recommended, where $L' < L$
 - 6: Construct 2nd-order polynomial regression model $\hat{y}(\mathbf{x}) = \hat{f}(\mathbf{x} | \hat{\beta}) = \mathbf{z}(\mathbf{x})^T \hat{\beta}$ based on the dataset D
 - 7: Define objective function $F(\mathbf{x})$ formulating process optimization problem; for example, $F(\mathbf{x}) = |y_{\text{target}} - \hat{y}(\mathbf{x})|$
 - 8: **For** l from 1 to L
 - 9: Find a solution $\mathbf{x}_l^* = \arg\min_{\mathbf{x}} F(\mathbf{x})$ by applying optimization algorithm to $F(\mathbf{x})$; here, different solutions are found each time because of its stochastic mechanism
 - 10: Obtain CI for $\mathbf{x}_l^*$ using Equation (2), and then calculate its length, i.e.,

$$W_l = 2 \times t_{1-\alpha/2}(n-p')s \sqrt{\mathbf{z}(\mathbf{x}_l^*)^T (\mathbf{Z}^T \mathbf{Z})^{-1} \mathbf{z}(\mathbf{x}_l^*)}$$
 - 11: **end**
 - 12: Sort $W_l, l = 1, \dots, L$, in ascending order, i.e., $W_{(1)} < W_{(2)} < \dots$
 - 13: **return** $\{\mathbf{x}_{(1)}^*, \mathbf{x}_{(2)}^*, \dots, \mathbf{x}_{(L')}^*\}$
-

4. Description of the Target Process: Ball Mill

This section provides the description of the target process (i.e., ball mill) and the process dataset collected from it through CCD. Figure 1 shows a simplified schematic diagram for the target process. Ball mills are the most widely used equipment to grind ceramic powders to make their size appropriate, and to mix these powders with solvents and additives; after finishing ball milling, we obtain a slurry in which they are mixed together. To do this, after putting grinding media (e.g., balls), powders, solvents, and some additives into a closed cylindrical container (also called drum), in order, we rotate this container horizontally at a steady speed. While, the container rotates, the balls repeat climbing up its inner wall and then falling to the ground (i.e., they cascade); these behaviors enable us to evenly blend powders, solvents, and additives together. In addition, as powder particles move between the balls and between the balls and the inner wall, they are broken into smaller particles effectively. The quality of the slurry is usually evaluated, based on its viscosity and particle size distribution, before the slurry is fed into the next-stage unit process (i.e., spray dryer) to produce granules. In this study, we only focus on slurry viscosity that has significant impact on the efficiency of the spray dryer. Figure 2 shows the ball mill with an internal volume of 50 L and an internal diameter of 40 cm equipped for collecting the experiment dataset, and a rotating disk viscometer (DV2TLV) used to measure the values of slurry viscosity.

Figure 1. Simplified schematic diagram for the target ball mill. Adapted from FIGURE 12.11 in [18], with the permission of ASM International.

Figure 2. Target ball mill equipped for collecting experiment dataset and viscometer used to measure slurry viscosity: (a) Ball mill with the internal volume of 50 L and the internal diameter of 40 cm; (b) rotating disk viscometer (DV2TLV).

In this paper, volume percentage of slurry (vol%), solid content (wt%), milling speed (rpm), and milling time (h) are considered as key controllable factors for slurry viscosity. The volume percentage of slurry is the ratio of the volume of slurry to the volume of balls; the solid content equals the weight of powders divided by the weight of slurry; the milling speed and time are the rotation speed and the operation time of the ball mill, respectively. To gather input-output observations from the ball mill, ranges and levels for the three factors except for the milling time are set as tabulated in Table 1, and as shown in Figure 3a, fifteen experiments are designed; the range of each factor was determined by taking the opinions from field experts in ceramic industry. In every experiment, balls with the diameter of 10 mm were used, and the total weight and volume of these balls are 58.8 kg and 25 L, respectively. In each of the 15 experiments, together, the precise amounts of powders

(i.e., Al_2O_3), solvent (i.e., ion-exchanged water), and additives (i.e., dispersant and binder) that fulfill the corresponding experimental conditions were supplied to the container, and it was operated for 24 h. During its operation, the values of slurry viscosity were measured every 4 h, and because of this, the milling time was excluded in the design of experiment. That is, the milling time is not included in the experimental design, but considered as a key controllable factor affecting the viscosity. Small amounts of samples were taken from the solenoid valves (see Figure 2a) installed in the bottom of the drum in 4-h intervals. Their values of viscosity were measured by the viscometer (see Figure 2b). Figure 3b shows the behavior of the measured viscosity values when the values of the three controllable factors are set as the center point in Figure 3a indicated by '2'; in this case, as can be seen from this figure, the slurry viscosity goes up steadily over time. Table 2 presents the experimental dataset collected from the target ball mill by conducting 15 experiments designed by CCD; these experiments were performed in random order. From the collected dataset, we can prepare 90 data points, each of which consists of 4 inputs and single output, to train the 2nd-order polynomial model.

Table 1. Ranges and levels for the three controllable factors.

Controllable Factor	Level		
	−1	0	1
x_1 : slurry volume (vol%)	40	50	60
x_2 : solid content (wt%)	50	60	70
x_3 : milling speed (rpm)	34	41	48

Figure 3. (a) Arrangement of 15 tests designed by CCD; (b) measured values of slurry viscosity at the center point.

Table 2. Experiment dataset collected from the target ball mill by conducting 15 tests designed by CCD.

Test ID	x_1	x_2	x_3	Slurry Viscosity (cP)					
				After 4 h	After 8 h	After 12 h	After 16 h	After 20 h	After 24 h
Test 1	40	60	41	258.0	416.4	565.8	579.3	592.8	574.8
Test 2	50	60	41	319.8	368.4	417.0	457.2	471.0	514.2
Test 3	60	60	41	216.0	233.4	295.8	315.0	272.4	377.4
Test 4	50	60	34	474.6	423.0	371.4	414.6	438.6	566.4
Test 5	50	60	48	204.6	341.4	428.4	442.2	489.0	506.4
Test 6	50	50	41	99.6	451.2	399.0	331.2	254.4	359.4
Test 7	50	70	41	304.8	451.2	552.6	590.4	528.0	465.6
Test 8	40	50	34	148.8	297.6	408.0	364.2	377.4	439.2
Test 9	40	70	34	443.3	599.4	593.6	676.2	733.0	656.4
Test 10	40	50	48	119.4	405.6	479.4	552.6	516.6	516.6
Test 11	40	70	48	418.2	488.2	638.2	766.4	774.6	753.4
Test 12	60	50	34	206.4	277.2	348.0	295.2	242.4	278.4
Test 13	60	70	34	364.8	405.0	445.2	485.4	500.4	567.0
Test 14	60	50	48	148.2	186.6	246.0	267.0	277.2	329.4
Test 15	60	70	48	198.6	354.6	510.6	537.0	575.4	549.6

5. Experimental Results and Discussion

In this section, to verify the performance, the proposed method summarized in Algorithm 1 is applied to the process dataset gathered from the target ball mill. Before describing the results of process modeling and optimization, let us look at scatter plots and main effect plots for the target dataset. Scatter plot is a visualization tool to intuitively check the relationships (e.g., positive or negative; linear or nonlinear) between input and output variables, here, the input-output coordinate points are displayed at two-dimensional planes. The main effect plot displaying level means and overall mean of a quality characteristic is effective for visualizing the effects of controllable factors on it based on a dataset collected by experimental design. Level means are the sample means of output variables calculated only from the observations whose j th input values are equal to a certain level; overall mean is the sample mean of output variable calculated from all observations.

Figure 4 shows the scatter plots between four input variables $x_1, \dots, x_4$, and output variable y ; here, reference lines indicated by black lines are the 2nd-order polynomial curves, least squares fitted to two-dimensional coordinate points. As shown in Figure 4a,b, y is negatively and positively correlated with x_1 and x_2 , respectively. In Figure 4c, we can see the nonlinear relation between x_3 and y , which can be captured by 2nd-order polynomials, but this is relatively weaker than those presented in Figure 4a,b. As Figure 4d indicates, there exists nonlinear and positive correlations between x_4 and y ; it is observed that as the milling time becomes longer, the increase of slurry viscosity is saturated.

Figure 4. Scatter plots between input and output variables: (a) x_1 vs. y ; (b) x_2 vs. y ; (c) x_3 vs. y ; (d) x_4 vs. y .

Figure 5 presents the main effect plots for four controllable factors, and the ranges of level means (i.e., differences between maximum and minimum of level means) for them. In these main effect plots, level means are denoted by black circles and connected by black lines, and overall means are indicated by horizontal dashed black lines. In Figure 5e, the values of the ranges are described by a bar graph. It can be expected that if the level means are close to the overall mean, a controllable factor has no apparent effect on a quality characteristic. We can also expect that as the range values becomes larger, the effects of controllable factors on a quality characteristic becomes greater [30]. As can be seen visually from Figure 5c, the level means and overall mean relevant with x_3 are extremely close in comparison with the other variables; in Figure 5a,b,d, the maximum and minimum of level means are far away from overall means compared with Figure 5c. As shown in Figure 5e, the range value relevant with x_4 is the largest, and followed by those for x_2 and x_1 ; the range value of x_3 is very small compared to those of others.

Figure 5. (a) main effect plot for x_1 ; (b) main effect plot for x_2 ; (c) main effect plot for x_3 ; (d) main effect plot for x_4 ; (e) ranges of level means for $x_1, \dots, x_4$.

From now on, we present the results of applying the proposed CI-based process optimization method to the ball mill dataset and the discussion about the advantages of the proposed method. Before performing 2nd-order polynomial regression analysis, each input variable should be standardized to be bounded in the interval $[-1, 1]$.

5.1. Results of Second-Order Polynomial Regression Analysis

In this paper, to formulate the functional relationship between the four controllable factors and the slurry viscosity of the target ball mill, we make use of the 2nd-order polynomial regression model presented in Equation (1). On the basis of the experiment dataset listed in Table 2, the model parameters can be estimated using the method of least squares; the total number of regression coefficients to be estimated is $p' = 1 + 2p + p(p - 1)/2 = 15$ because the number of input variables is $p = 4$. Table 3 summarizes the results of estimating these coefficients; in this table, we list the estimates of coefficients, their standard errors (SEs), their t -statistics that equal to the values of estimates divided by their SEs, and the p -values of these t -statistics. The estimated coefficient $\hat{\beta}_1$ for x_1 has the negative value, -80.80 , since the volume percentage of slurry is negatively correlated with the slurry viscosity (see Figure 4a); on the contrary, the estimated coefficients for x_2 and x_4 are positive values, i.e., $\hat{\beta}_2 = 105.08$ and $\hat{\beta}_4 = 105.01$, respectively, because as the solid content or the milling time increase, the slurry viscosity also increases (see Figure 4b,d).

Table 3. Results of estimating 15 regression coefficients using the method of least squares.

Coefficient	Estimate	Standard Error	t -Statistic	p -Value
β_0	439.00	15.10	29.08	0.0000
β_1	-80.80	7.75	-10.43	0.0000
β_2	105.08	7.75	13.57	0.0000
β_3	3.01	7.75	0.39	0.6989
β_4	105.01	9.26	11.34	0.0000
β_{12}	-10.91	8.66	-1.26	0.2115
β_{13}	-19.32	8.66	-2.23	0.0286
β_{14}	-30.38	11.34	-2.68	0.0091
β_{23}	-5.55	8.66	-0.64	0.5233
β_{24}	21.31	11.34	1.88	0.0640
β_{34}	38.42	11.34	3.39	0.0011
β_{11}	-3.70	15.27	-0.24	0.8091
β_{22}	3.82	15.27	0.25	0.8031
β_{33}	29.92	15.27	1.96	0.0538
β_{44}	-81.38	15.84	-5.14	0.0000

Determining whether it is possible to exclude a certain term from the full model with p' coefficients can be determined based on the relevant p -value; the regression coefficients with large p -values have large variance, and thus, their precision is degraded. In general, the regression terms with the p -values larger than 0.05 or 0.01 are regarded to be statistically insignificant, so that we can drop these insignificant terms from the model. As tabulated in Table 3, the p -values of the regression coefficients β_3 , β_{12} , β_{23} , β_{11} , and β_{22} are larger than 0.05. The coefficient β_3 is not significant since y is not linear but weakly nonlinear in x_3 (see Figure 4c). The p -values of the coefficients for x_1 and x_2 are very small, but the coefficients for their quadratic terms x_1^2 and x_2^2 have very large p -values; this tells us that y is affected by the input variables x_1 and x_2 not nonlinearly but linearly (see Figure 4a,b). Since the p -values of β_{12} and β_{23} are very large, it is obvious that the interaction terms x_1x_2 and x_2x_3 should be excluded from the full model, the remaining interaction terms with the p -values smaller than 0.05 should remain in the regression model.

After dropping the terms, x_3 , x_1x_2 , x_2x_3 , x_1^2 , and x_2^2 with the p -values of their coefficients larger than 0.05 from the full model, and then fitting the reduced model to the same dataset, we can obtain the following functional relationship:

$$\hat{y}(x) = \underset{(12.99)}{439.05} - \underset{(7.61)}{80.80}x_1 + \underset{(7.61)}{105.08}x_2 + \underset{(9.10)}{105.01}x_4 - \underset{(8.51)}{19.32}x_1x_3 - \underset{(11.14)}{30.38}x_1x_4 + \underset{(11.14)}{21.31}x_2x_4 + \underset{(11.14)}{38.42}x_3x_4 + \underset{(13.18)}{29.97}x_3^2 - \underset{(15.57)}{81.38}x_4^2 \quad (9)$$

where in the bottom of each estimated coefficient, its SE is also presented in parentheses; note that the values of coefficients and SEs in Equation (9) are slightly different from those in Table 3, since they are re-estimated without the terms, x_3 , x_1x_2 , x_2x_3 , x_1^2 , and x_2^2 . By modifying the full model in this way, the number of coefficients has been reduced from 15 to 10. In addition, the value of R^2 has been reduced by 0.004 (from 0.865 to 0.861), but the value of adjusted R^2 has increased by 0.006 (from 0.84 to 0.846). The R^2 and adjusted R^2 are the statistics used to determine how well the regression models are fitted to the target dataset; these values lie in the interval $[0, 1]$, and as they get closer to 1, the goodness of fit of the regression model becomes better. It is well-known that the adjusted R^2 is more suitable for preventing the model from overfitting than the R^2 [5]; thus, it makes sense to exclude the insignificant terms from the full model because the adjusted R^2 has increased. Moreover, the F statistic used to test the statistical significance of the regression model has increased by 20.8 (from 34.5 to 55.3); the larger the F statistic, the greater the statistical significance of these models. In summary, it is obvious that by removing the insignificant terms from the full model, we can obtain the more statistically significant and transparent model described in Equation (9).

Next, to validate the resulting model in Equation (9), we present the scatter plot between actual outputs y_i and predicted outputs $\hat{y}_i$, and the results of analyzing the residuals $r_i = y_i - \hat{y}_i$, $i = 1, \dots, 90$, one by one. Figure 6a shows the scatter plot between y_i and $\hat{y}_i$, and Figure 6b–f describe visualization tools to check the independence, homoscedasticity (i.e., constant variance), white noise properties, symmetry about zero, and normality of the error term ε in Equation (1). In Figure 6a, black circles correspond to coordinate points $(y_i, \hat{y}_i)$, $i = 1, \dots, 90$, and dotted black and thick blue lines are the straight line $X = Y$ and least squares fitted line to these points, respectively. Since there exists strong linear correlation between y_i and $\hat{y}_i$, it is apparent that the regression model describes the target process dataset adequately. Figure 6b is the scatter plot of the points (r_{i-1}, r_i) , $i = 2, \dots, 90$, to investigate the independence of ε ; as can be seen from this figure, since there is no clear positive or negative correlation, the error term ε meets the independence assumption. Figure 6c shows the scatter plot between the fitted values $\hat{y}_i$ and the residuals r_i to check the homoscedasticity of ε ; it is observed that the residual values do not significantly increase or decrease with the fitted values. Figure 6d indicates the line plot of the residuals to confirm their white noise properties. In this figure, we can see that the sample mean of them is approximately equal to zero and they also show irregular behaviors; the characteristics of the residuals are roughly similar with those of white noise. Figure 6e,f describe the histogram and normal probability plots of the residuals, respectively; these figures indicate that the residuals do not depart from the symmetry and normality severely. To sum up, from Figure 6, we can conclude that the calculated residuals, r_i , $i = 1, \dots, 90$, based on Equation (9) do not violate the general assumptions for the error term ε considerably.

Now, let us look at the results of estimating the importance values of input variables using the MC-based method (see Section 2.2), presented in Figure 7. To do this, we generate $N = 5000$ random vectors $\mathbf{x}^{(i)} \in \mathbb{R}^4$. Figure 7a–d show the histograms for each component of $\mathbf{x}^{(i)} \in \mathbb{R}^4$, $i = 1, \dots, 5000$; here, histograms related with those belonging to X_{large} and X_{small} are displayed as pink and shaded blue colors, respectively. Figure 7e–h depict empirical (red and blue) CDFs corresponding to each of the two histograms. In Figure 7a–h, for convenience of interpretation, the horizontal axis ranges have been recovered to their original ranges. Figure 7i is the bar graph for the estimated importance values of four controllable factors. As can be viewed from Figure 7a,b, as the values of x_1 or x_2 change, the frequencies increase or decrease nearly linearly; in Figure 7c,d, the frequencies change with the

values of x_3 and x_4 nonlinearly. From Figure 7g, one can easily notice that the two CDFs of x_3 are extremely similar to each other, thus, it is valid to say that the input x_3 will have little impact on the output y . The input x_4 has the greatest difference between the red and blue CDFs, and followed by x_2 and x_1 . In other words, x_4 has the highest estimated importance value, and followed by x_2 and x_1 , and that of x_3 is the lowest. Comparing Figures 5e and 7i, we can see that these two bar graphs are slightly different; this may be due to the fact that the former based on main effect plots does not consider interaction effects between input variables, and the latter was obtained after dropping the insignificant terms, x_3 , x_2x_3 , x_1^2 , and x_2^2 , from the full model.

Figure 6. (a) scatter plot between y_i and $\hat{y}_i$; (b) scatter plot between r_{i-1} and r_i ; (c) scatter plot between $\hat{y}_i$ and r_i ; (d) line plot for r_i ; (e) histogram for r_i ; (f) normal probability plot for r_i .

Figure 7. Results of estimating the importance values of controllable factors using the MC-based method: (a) histograms for x_1 ; (b) histograms for x_2 ; (c) histograms for x_3 ; (d) histograms for x_4 ; (e) ECDFs for x_1 ; (f) ECDFs for x_2 ; (g) ECDFs for x_3 ; (h) ECDFs for x_4 ; (i) estimated importance values of four controllable factors.

Next, let us take a look at the response surface and contour plots for Equation (9), described in Figures 8 and 9, respectively. To obtain each plot, we first prepare grid coordinates based on X-axis and Y-axis variables. Then, by plugging these coordinates into Equation (9), we calculate its output values. In doing this, their median values are substituted into Equation (9), in the case of those inputs not included in the coordinates. In common with Figure 7, the ranges of X- and Y-axes have returned to their original ranges. By examining the response surface and contour plots closely, we can see how the pairs of axes variables interact and then affect the output variable y visually. As shown in Figures 8a and 9a, as the value of x_1 increases or that of x_2 decreases, the value of y decreases; this was also confirmed in the scatter plots in Figure 4a,b. That is, these two variables linearly influence on y , and x_2 has a greater effect on y than x_1 . In Figures 8b and 9b, as the value of x_1 becomes larger, the value of y becomes smaller. Also, as the value of x_3 increases, the value of y decreases and then increases. Since the value of y changes more when the value of x_1 is changed than when that of x_3 is changed, it is valid to say that x_1 is the more important controllable factor than x_3 . From Figures 8c and 9c, one can easily notice that the value of y changes much more when the value of x_4 increases than when the value of x_1 increases. There exists clear nonlinear relationship between y and x_4 , and the increase of the value of x_4 tends to dampen the increase of the value of y . When the value of x_4 is small (i.e., in the early stage of milling operation), x_4 influences much more on y than x_1 ; as the milling time passes, the effect of x_1 on y becomes larger than that of x_4 . Figures 8d and 9d indicate that x_2 has a much greater impact on y than x_3 ; in these figures, the shapes of contours indicated by different colors are very similar to each other because the interaction term x_2x_3 was removed from the full model. Figures 8e and 9e tell us that as the value of x_4 becomes larger, this slows down the increase of the value of y ; here, we can also confirm that as the value of x_4 increases, the influence of x_2 on y gets larger. In Figures 8f and 9f, we can see that x_4 influences much more on y than x_3 , and as the value of x_4 becomes larger, the value of y is saturated.

Figure 8. Response surface plots for Equation (9): (a) (x_1, x_2, y); (b) (x_1, x_3, y); (c) (x_1, x_4, y); (d) (x_2, x_3, y); (e) (x_2, x_4, y); (f) (x_3, x_4, y).

Figure 9. Contour plots for Equation (9): (a) (x_1, x_2) ; (b) (x_1, x_3) ; (c) (x_1, x_4) ; (d) (x_2, x_3) ; (e) (x_2, x_4) ; (f) (x_3, x_4) .

5.2. Results of Confidence Interval-Based Process Optimization

In this subsection, we describe the results of applying the PSO algorithm to the objective function $F(\mathbf{x}) = |y_{\text{target}} - \hat{y}(\mathbf{x})|$ defined based on Equation (9). The values of α , L , and L' in Algorithm 1 are set as 0.05, 100, and 10, respectively, and Table 4 lists the user-defined parameters used in the PSO algorithm.

Table 4. User-defined parameters in the PSO algorithm.

n_s	$t_{\max}$	$w(0)$	$w(t_{\max})$	$c_{1\min}$	$c_{1\max}$	$c_{2\min}$	$c_{2\max}$
20	200	0.9	0.4	0.5	2.5	0.5	2.5

First, let us take a look at the results of finding the values of controllable factors achieving the slurry viscosity of 550 cP (i.e., $y_{\text{target}} = 550$); as can be seen from Figure 9, it is obvious that there exists an infinite number of solutions for this process optimization problem. Table 5 lists $L' = 10$ solutions found by the proposed method, CIs enclosing their outputs of Equation (9), and their lengths. Among these solutions, the most statistically significant solution is $\mathbf{x}_{(1)}^* = [47.07, 63.93, 45.59, 16.29]^T$. Figure 10 shows the trajectories of each component in $\mathbf{x}^{g^{\text{best}}}$, and the trajectory of $\hat{y}(\mathbf{x}^{g^{\text{best}}})$ during the search for the $\mathbf{x}_{(1)}^*$ using the PSO algorithm; similar to Figure 8, the Y-axis ranges have been recovered to their original ranges. In the initiation stage for the search, there are some fluctuations in these trajectories; as the updates are repeated and then the number of iterations become larger than about 120, $\mathbf{x}^{g^{\text{best}}}$ and $\hat{y}(\mathbf{x}^{g^{\text{best}}})$ have converged to $\mathbf{x}_{(1)}^*$ and 550, respectively. Figure 11 shows the contour plots where the solution $\mathbf{x}_{(1)}^*$ is indicated by red asterisks; as described in this figure, $\mathbf{x}_{(1)}^*$ is located at the coordinate points whose height values are exactly equal to the target value $y_{\text{target}} = 550$. Based on Table 5, one can easily notice that the lengths of CIs for the 10 solutions are slightly different from each other; however, the values of controllable factors vary from solution to solution considerably. It is worthwhile to emphasize that although $\mathbf{x}_{(1)}^*$ is the best solution from a statistical viewpoint, this may not always fulfill the requirements for the users. The proposed method recommends several statistically significant solutions so that among them the users can flexibly choose and use the suitable solutions for process conditions. For example, if the users want the milling time to be shorter than 16.29 h in $\mathbf{x}_{(1)}^*$ and, at the same time, want the slurry viscosity to be equal to the target value, i.e., $y_{\text{target}} = 550$, they can set

the values of controllable factors on the basis of $\mathbf{x}_{(8)}^*$ instead of $\mathbf{x}_{(1)}^*$; in this case, the milling time can be reduced by about 2 h, the value of the volume percentage of slurry and solid content become higher by 0.9 vol% and 0.36 wt%, respectively, and the milling speed becomes faster than before. The difference in the lengths of CIs between $\mathbf{x}_{(1)}^*$ and $\mathbf{x}_{(8)}^*$ is 3.9, and it is negligible. If it is desirable to decrease the value of solid content to be smaller than 63.93 wt% in $\mathbf{x}_{(1)}^*$, the 6th solution $\mathbf{x}_{(6)}^*$ can be used instead; in this case, the milling time increases by 0.77 h, the value of the volume percentage of slurry decreases by 2.77 vol%, and the milling speed is similar to each other; the difference in the lengths of CIs between $\mathbf{x}_{(1)}^*$ and $\mathbf{x}_{(6)}^*$ is 3.9, and it is also negligible.

Table 5. ($y_{\text{target}} = 550$) Results of finding $L' = 10$ solutions using the proposed method.

Order	x_1	x_2	x_3	x_4	CI	Length
1	47.07	63.93	45.59	16.29	[528.63, 571.37]	42.73
2	46.41	61.80	44.66	19.29	[527.83, 572.17]	44.34
3	50.53	66.91	34.83	17.13	[527.74, 572.26]	44.52
4	48.38	65.23	35.07	19.36	[527.70, 572.30]	44.60
5	50.00	62.60	48.00	18.94	[527.63, 572.37]	44.75
6	44.30	60.71	45.52	17.06	[527.44, 572.56]	45.12
7	49.92	66.38	34.16	16.40	[527.16, 572.84]	45.67
8	46.17	63.57	48.00	14.32	[526.69, 573.31]	46.63
9	48.10	65.21	35.59	20.29	[526.68, 573.32]	46.64
10	47.75	65.28	37.37	19.10	[526.67, 573.33]	46.65

Figure 10. Trajectories of $\mathbf{x}^{g\text{best}}$ and $\hat{y}(\mathbf{x}^{g\text{best}})$ during the search for the $\mathbf{x}_{(1)}^*$ using the PSO algorithm:

(a) trajectory of $x_1^{g\text{best}}$; (b) trajectory of $x_2^{g\text{best}}$; (c) trajectory of $x_3^{g\text{best}}$; (d) trajectory of $x_4^{g\text{best}}$; (e) trajectory of $\hat{y}(\mathbf{x}^{g\text{best}})$.

Figure 11. Contour plots with the solution $\mathbf{x}_{(1)}^*$ achieving $\hat{y}(\mathbf{x}_{(1)}^*) = 550$: (a) (x_1, x_2) ; (b) (x_1, x_4) ; (c) (x_2, x_4) .

Tables 6 and 7 summarize the results of applying the proposed method to obtain the values of controllable factors achieving the slurry viscosities of 500 cP and 450 cP, respectively; in these two tables, the differences in the lengths between the first and last solutions are 3.21, and 3.16 (typically negligible), respectively. This means that the statistical significances of the solutions in each table are similar with each other, and thus, the users can select and use the most appropriate solutions for field situations among them. If the users want to operate the target ball mill for a relatively short period of time

and, at the same time, want to obtain the slurry with the viscosity value of 500 cP, they are able to utilize the 6th solution $\mathbf{x}_{(6)}^*$ in Table 6; if it is desirable for the solid content to be relatively low, the 5th solution $\mathbf{x}_{(5)}^*$ in Table 6 can be used to set the values of process parameters. When the target value for slurry viscosity is equal to $y_{\text{target}} = 450$, the users can choose the proper solutions satisfied with their requirements from the recommended solutions listed in Table 7.

Table 6. ($y_{\text{target}} = 500$) Results of finding $L' = 10$ solutions using the proposed method.

Order	x_1	x_2	x_3	x_4	CI	Length
1	49.13	61.62	34.54	16.27	[480.23, 519.77]	39.54
2	47.52	61.56	35.00	14.87	[480.08, 519.92]	39.83
3	48.65	61.69	34.53	15.38	[480.01, 519.99]	39.98
4	51.15	61.61	44.95	19.51	[479.39, 520.61]	41.22
5	47.57	59.92	45.10	16.56	[479.18, 520.82]	41.64
6	49.02	65.43	34.96	11.66	[479.07, 520.93]	41.86
7	51.34	65.18	46.18	14.20	[479.02, 520.98]	41.97
8	51.95	61.87	47.80	16.75	[478.97, 521.03]	42.06
9	49.32	61.09	34.00	17.95	[478.64, 521.36]	42.72
10	48.56	61.56	37.42	19.87	[478.63, 521.37]	42.75

Table 7. ($y_{\text{target}} = 450$) Results of finding $L' = 10$ solutions using the proposed method.

Order	x_1	x_2	x_3	x_4	CI	Length
1	48.46	57.58	35.03	14.71	[430.31, 469.69]	39.38
2	50.32	62.94	36.25	11.07	[430.17, 469.83]	39.65
3	52.68	60.31	46.15	15.49	[430.12, 469.88]	39.77
4	53.64	60.34	35.50	18.25	[429.68, 470.32]	40.64
5	48.25	61.67	36.81	11.37	[429.42, 470.58]	41.16
6	47.11	56.06	46.66	14.36	[429.20, 470.80]	41.60
7	49.02	57.46	45.16	15.60	[429.05, 470.95]	41.89
8	46.63	59.84	45.36	12.04	[428.94, 471.06]	42.12
9	50.21	60.35	37.04	13.90	[428.90, 471.10]	42.20
10	48.37	62.18	47.88	10.79	[428.73, 471.27]	42.54

5.3. Discussion

In this subsection, we shall summarize the strengths of the proposed CI-based process optimization method. The main strengths include that when a unique solution for process optimization problems does not exist, and thus, different solutions can be found whenever optimization algorithms are executed, the proposed method can assess the qualities of these different solutions from a statistical point of view. Recall that, as described in Figure 11, there is an infinite number of different solutions achieving the same quality characteristic as a desired target value y_{target} . In the proposed method, the qualities of them are evaluated on the basis of the CI in Equation (2); it is important to remark that as the lengths of their CIs become shorter, their uncertainties caused by such things as measurement errors and/or uncontrollable factors decrease. The second strength is that since the proposed method recommends the L' statistically significant solutions for the users (see Tables 5–7), it can provide them with a variety of choices. As explained in Section 5.2, it is undesirable to recommend only the first solution $\mathbf{x}_{(1)}^*$ for the users because it may not fulfill the requirements for them. The proposed method enables the users to flexibly select and use the solutions suitable for practical situations among the L' solutions.

Prior to completing this subsection, let us describe how the importance values of controllable factors presented in Figures 5e and 7i can help the process operators and engineers. First, these important values provide guidance on which controllable factors should be set more correctly so as to optimize quality characteristics; it is entirely fair to say that the controllable factors with larger importance

values should be more precisely set to the values of the preferred solution. For example, in the case of the target ball mill, the milling time should be set most accurately because it has the biggest importance value; the milling speed with the smallest importance value may be loosely set. Second, the importance values can serve as a good reference for performing additional experimental designs. For example, unimportant factors may be excluded from these additional experimental designs, and thus one can only focus on important factors when planning them. In addition, one can increase the number of levels for important factors to closely examine their effects on quality characteristics while reducing the number of levels for unimportant factors.

6. Conclusions

In this paper, we proposed a CI-based process optimization method using second-order polynomial regression analysis. Recall that if the goal of process optimization is to make the quality characteristics equal to the user-defined target values, we cannot obtain a unique solution; in this case, whenever optimization algorithms are executed, different solutions can be obtained. In this paper, after obtaining several different solutions by applying the PSO algorithm repeatedly, the qualities of these solutions are evaluated on the basis of the CIs (see Equation (2)); the solutions are sorted in ascending order according to the lengths of their CIs, and then the first few solutions are recommended for the users. To verify the performance, the proposed method was applied to the process dataset collected from the target ball mill; here, the aim of the proposed method is to obtain the values of the controllable factors achieving the target values of slurry viscosity. The simulation results showed that the proposed method can provide several statistically significant solutions; among these solutions, the users can flexibly select and use the proper solutions for practical situations.

In future research, we will consider improving the proposed method to be applicable even when machine learning techniques [31] are used to formulate the input-output relations. In this case, the number of collected input-output observations should be large enough to prevent the overfitting problem. The regression surfaces based on machine learning techniques are, in general, more uneven than those based on the second-order polynomial regression analysis; therefore, instead of PSO algorithm, more advanced optimization algorithms presented in [22–25] should be applied.

Author Contributions: Conceptualization, J.Y., H.J. and S.K.; methodology, J.Y. and H.J.; software, J.Y.; validation, S.Y., J.K. and Y.L.; formal analysis, J.Y.; investigation, J.Y. and H.J.; resources, K.-T.L. and S.-S.R.; data curation, S.K., S.-S.R. and H.J.; writing—original draft preparation, J.Y.; writing—review and editing, H.J., S.Y. and J.K.; visualization, J.Y.; supervision, K.-T.L. and S.-S.R.; project administration, Y.L., K.-T.L. and S.-S.R.; funding acquisition, Y.L. and K.-T.L. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Ministry of Trade, Industry and Energy (MOTIE), South Korea, through the i-Ceramic platform construction project-i-Ceramic manufacturing innovation platform (Development of Cloud Big Data Platform for the Innovative Manufacturing in Ceramic Industry) under Grant 20004367.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Rao, R.V. *Advanced Modeling and Optimization of Manufacturing Processes: International Research and Development*; Springer Science & Business Media: Berlin/Heidelberg, Germany, 2010. [\[CrossRef\]](#)
2. Castillo, E.D. *Process Optimization: A Statistical Approach*; Springer Science & Business Media: Berlin/Heidelberg, Germany, 2007. [\[CrossRef\]](#)
3. Dejaegher, B.; Heyden, Y.V. Experimental designs and their recent advances in set-up, data interpretation, and analytical applications. *J. Pharm. Biomed. Anal.* **2011**, *56*, 141–158. [\[CrossRef\]](#) [\[PubMed\]](#)
4. Draper, N.R.; Smith, H. *Applied Regression Analysis*; John Wiley & Sons: Hoboken, NJ, USA, 1998. [\[CrossRef\]](#)
5. Hastie, T.; Tibshirani, R.; Friedman, J. *The elements of Statistical Learning: Data Mining, Inference, and Prediction*; Springer Science & Business Media: Berlin/Heidelberg, Germany, 2009. [\[CrossRef\]](#)
6. Ebadnejad, A.; Karimi, G.R.; Dehghani, H. Application of response surface methodology for modeling of ball mills in copper sulphide ore grinding. *Powder Technol.* **2013**, *245*, 292–296. [\[CrossRef\]](#)

7. Parida, M.K.; Joardar, H.; Fout, A.K.; Routaray, I.; Mishra, B.P. Multiple response optimizations to improve performance and reduce emissions of Argemone Mexicana biodiesel-diesel blends in a VCR engine. *Appl. Thermal Eng.* **2019**, *148*, 1454–1466. [[CrossRef](#)]
8. Adesina, O.A.; Abdulkareem, F.; Yusuff, A.S.; Lala, M.; Okewale, A. Response surface methodology approach to optimization of process parameter for coagulation process of surface water using Moringa oleifera seed. *S. Afr. J. Chem. Eng.* **2019**, *28*, 46–51. [[CrossRef](#)]
9. Costa, N.; Garcia, J. Using a multiple response optimization approach to optimize the coefficient of performance. *Appl. Thermal Eng.* **2016**, *96*, 137–143. [[CrossRef](#)]
10. Chen, Z.; Shi, Y.; Lin, X.; Yu, T.; Zhao, P.; Kang, C.; He, X.; Li, H. Analysis and optimization of process parameter intervals for surface quality in polishing Ti-6Al-4V blisk blade. *Results Phys.* **2019**, *12*, 870–877. [[CrossRef](#)]
11. Li, Y.X.; Xu, Q.Y.; Guo, R.T.; Wang, Z.Y.; Liu, X.Y.; Shi, X.; Qiu, Z.Z.; Qin, H.; Jia, P.Y.; Qin, Y.; et al. Removal of NO by using sodium persulfate/limestone slurry: Modeling by response surface methodology. *Fuel* **2019**, *254*. [[CrossRef](#)]
12. Aggarwal, A.; Singh, H.; Kumar, P.; Singh, M. Optimizing power consumption for CNC turned parts using response surface methodology and Taguchi's technique-a comparative analysis. *J. Mater. Process. Technol.* **2008**, *200*, 373–384. [[CrossRef](#)]
13. Mohanty, S.; Mishra, A.; Nanda, B.K.; Routara, B.C. Multi-objective parametric optimization of nano powder mixed electrical discharge machining of AlSiCp using response surface methodology and particle swarm optimization. *Alexandria Eng. J.* **2018**, *57*, 609–619. [[CrossRef](#)]
14. Zhang, C.; Chen, Z.; Mei, Q.; Duan, J. Application of particle swarm optimization combined with response surface methodology to transverse flux permanent magnet motor optimization. *IEEE Trans. Magn.* **2017**, *53*, 1–7. [[CrossRef](#)]
15. Hasanien, H.M. Particle swarm design optimization of transverse flux linear motor for weight reduction and improvement of thrust force. *IEEE Trans. Ind. Electron.* **2010**, *58*, 4048–4056. [[CrossRef](#)]
16. Nocedal, J.; Wright, S.J. *Numerical Optimization*; Springer Science & Business Media: Berlin/Heidelberg, Germany, 2006. [[CrossRef](#)]
17. Talbi, E.G. *Metaheuristics: From Design to Implementation*; John Wiley & Sons: Hoboken, NJ, USA, 2009. [[CrossRef](#)]
18. Richerson, D.W.; Lee, W.E. *Modern Ceramic Engineering Properties, Processing, and Use in Design*; CRC Press: Boca Raton, FL, USA, 2018. [[CrossRef](#)]
19. Deng, B.; Shi, Y.; Yu, T.; Kang, C.; Zhao, P. Multi-response parameter interval sensitivity and optimization for the composite tape winding process. *Materials* **2018**, *11*, 220. [[CrossRef](#)] [[PubMed](#)]
20. Yin, X.; He, Z.; Niu, Z.; Li, Z.S. A hybrid intelligent optimization approach to improving quality for serial multistage and multi-response coal preparation production systems. *J. Manuf. Syst.* **2018**, *47*, 199–216. [[CrossRef](#)]
21. Bera, S.; Mukherjee, I. A multistage and multiple response optimization approach for serial manufacturing system. *Eur. J. Oper. Res.* **2016**, *248*, 444–452. [[CrossRef](#)]
22. Precup, R.E.; David, R.C. *Nature-inspired Optimization Algorithms for Fuzzy Controlled Servo Systems*; Butterworth-Heinemann: Oxford, UK, 2019. [[CrossRef](#)]
23. Wang, G.; Guo, L. A novel hybrid bat algorithm with harmony search for global numerical optimization. *J. Appl. Math.* **2013**. [[CrossRef](#)]
24. Haber, R.E.; Beruvides, G.; Quiza, R.; Hernandez, A. A simple multi-objective optimization based on the cross-entropy method. *IEEE Access* **2017**, *5*, 22272–22281. [[CrossRef](#)]
25. Tharwat, A.; Hassanien, A.E. Quantum-behaved particle swarm optimization for parameter optimization of support vector machine. *J. Classif.* **2019**, *36*, 576–598. [[CrossRef](#)]
26. Engelbrecht, A.P. *Computational Intelligence: An Introduction*; John Wiley & Sons: Hoboken, NJ, USA, 2007. [[CrossRef](#)]
27. Kennedy, J.; Eberhart, R. Particle swarm optimization. In *Proceedings of ICNN'95-International Conference on Neural Networks*; IEEE: Piscataway, NJ, USA, 1995; Volume 4, pp. 1942–1948. [[CrossRef](#)]
28. Ratnaweera, A.; Halgamuge, S.K.; Watson, H.C. Self-organizing hierarchical particle swarm optimizer with time-varying acceleration coefficients. *IEEE Trans. Evol. Comput.* **2004**, *8*, 240–255. [[CrossRef](#)]

29. Del Valle, Y.; Venayagamoorthy, G.K.; Mohagheghi, S.; Hernandez, J.C.; Harley, R.G. Particle swarm optimization: Basic concepts, variants and applications in power systems. *IEEE Trans. Evol. Comput.* **2008**, *12*, 171–195. [[CrossRef](#)]
30. Kwak, J.S. Application of Taguchi and response surface methodologies for geometric error in surface grinding process. *Int. J. Mach. Tools Manuf.* **2005**, *45*, 327–334. [[CrossRef](#)]
31. Castaño, F.; Beruvides, G.; Villalonga, A.; Haber, R.E. Self-tuning method for increased obstacle detection reliability based on internet of things LiDAR sensor models. *Sensors* **2018**, *18*, 1508. [[CrossRef](#)] [[PubMed](#)]

© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).