



---

*Research article*

## **Soft-sensing modeling of mother liquor concentration in the evaporation process based on reduced robust least-squares support-vector machine**

**Xiaoshan Qian<sup>1,\*</sup>, Lisha Xu<sup>2,\*</sup> and Xinmei Yuan<sup>1</sup>**

<sup>1</sup> College of Physical Science and Engineering Technology, Yichun University, Yichun 336000, China

<sup>2</sup> College of Information Science and Engineering, Hunan Women's University, Changsha 410004, China

\* **Correspondence:** Email: [qxs@jxycu.edu.cn](mailto:qxs@jxycu.edu.cn), [xlscsu2021@126.com](mailto:xlscsu2021@126.com).

**Abstract:** The evaporation process is vital in alumina production, with mother liquor concentration serving as a critical control parameter. To address the challenge of online detection, we propose the introduction of a soft measurement strategy. First, due to the significant fluctuations in the production process variables and inter-variable coupling, comprehensive grey correlation analysis and kernel principal component analysis are employed to reduce the input dimension and computational complexity of the data, enhancing the efficiency of the soft sensing model. The reduced robust least-squares support-vector machine (LSSVM), with its commendable predictive performance, is used for modeling and predicting the principal components. Concurrently, an improved Pattern Search-Differential Evolution (PS-DE) algorithm is proposed for optimizing the pivotal parameters of the LSSVM network. Lastly, on-site industrial data validation indicates that the new model offers superior tracking capabilities and heightened accuracy. It is deemed aptly suitable for the online detection of mother liquor concentration.

**Keywords:** grey relational analysis (GRA); evaporation process; least squares support vector machine (LSSVM); PS-DE algorithm; soft sensor

---

### **1. Introduction**

Alumina, serving as a vital intermediary in aluminum production, underpins numerous industries, from aerospace to consumer electronics, with global demand surpassing 120 million tonnes annually [1, 2]. Within the complex alumina multi-effect evaporation process, a major procedure in alumina production, the precise assessment of mother liquor concentration is paramount, dictating both product quality and operational efficiency [3]. Directly gauging this concentration is fraught with

obstacles, notably the extreme processing conditions such as elevated temperatures and corrosiveness. In response, soft measurement methodologies have gained prominence, harnessing readily available process data to estimate this crucial parameter [4, 5]. These techniques, offering real-time concentration insights, not only facilitate rapid process adjustments but also serve as a cornerstone for implementing advanced control strategies, thereby optimizing the entire production process.

Traditional methods for measuring mother liquor concentration often grapple with inherent challenges. On-site mother liquor concentration detection uses manual periodic sampling and laboratory analysis, with feedback to the production site several hours later [6]. This is offline analysis and cannot meet the real-time requirements of control. Simultaneously, the evaporation process exhibits strong correlation coupling [7], encounters uncertain external environmental factors and displays non-linearity [8]. These issues often compromise accuracy and reliability, prompting a search for alternative approaches. Enter the least-squares support-vector machine (LSSVM) [9–11]: a machine learning technique that stands out for its robustness against non-linearities and noise. In the realm of soft measurement applications, LSSVM not only demonstrates improved predictive accuracy but also stands out for its adaptability, rendering it an appealing choice for the dynamic context of alumina production.

In order to achieve online real-time detection of mother liquor concentration at the outlet, we established a soft measurement model for outlet mother liquor concentration. It consists of the following contributions:

- We adopt a combination of comprehensive grey relational analysis and kernel principal component analysis to reduce the input dimensionality and computational complexity of the data.
- A predictive performance-oriented, reduced-robust LSSVM is employed for modeling and forecasting the principal components.
- An improved PS-DE algorithm is proposed to optimize the key parameters of the LSSVM network.

The remainder of this article is organized as follows: Section 2 delves into the fundamentals of soft-sensing modeling and the application of LSSVM in soft measurement. The alumina evaporation process is clarified in Section 3. Section 4 elucidates the combination of comprehensive grey relational analysis and kernel principal component analysis. In Section 5, we introduce the reduced robust LSSVM and its significance in modeling and forecasting. The innovative Pattern Search-Differential Evolution (PS-DE) algorithm for optimizing the LSSVM network is elaborated in Section 6. Finally, Section 7 concludes the article, highlighting the key findings.

## 2. Literature review

### 2.1. Soft measurement for mother liquor concentration

Soft sensing, also known as soft sensors, primarily involves estimating hard-to-measure or costly-to-measure variables using variables that are easily measurable (like temperature, pressure, etc.) [12]. Soft sensing techniques utilize mathematical models, statistical methods and data-driven approaches to establish these connections [13]. Taking the mother liquor concentration as an example: In certain chemical or bioprocesses, directly measuring the concentration of the mother liquor might

require complex, expensive or real-time analytical methods. Soft sensing can utilize data from other measurable process variables, such as temperature, flow rate, conductivity and more, to estimate or predict the concentration of the mother liquor using a pre-established model [14–16].

To tackle the challenge of monitoring alkaline solution concentration in the alumina production's evaporation process, a soft sensor model was developed employing a recursive partial least squares approach incorporating a forgetting factor, as detailed in Wang's study [14]. By utilizing historical data from the initial stages, the model effectively handles the periodicity and slow dynamic nature of the process. Distributed by Damour et al. [4], a model-based soft sensor was tailored for the final stage of industrial sugar crystallization, obviating the necessity for population balance computations. The soft-sensor's effectiveness is confirmed by employing real plant data derived from an industrial crystallization process. Meng et al. [5] devised a data-driven soft sensor model based on twin support vector regression. The model chooses seven easily measurable variables as inputs to estimate the challenging-to-measure variables of mother liquor purity and supersaturation. More recently, an innovative first-principles model-based soft sensor approach was introduced, which integrates two distinct models: a supersaturation model and a simplified last-stage crystallization model [17]. This study addresses the challenge of monitoring and controlling the degree of supersaturation in the ultimate step of industrial sugar crystallization. However, a persistent challenge lies in effectively managing nonlinearity, high-dimensionality and noise, which frequently results in overfitting or limited generalization in real-world scenarios.

## 2.2. *LSSVM in soft measurement*

The LSSVM is an offshoot of the standard Support Vector Machine (SVM), distinguished by its use of a squared loss function instead of the hinge loss, rendering the optimization problem linear [18, 19]. This unique adaptation ensures computational efficiency while preserving the model's robustness against non-linearities. LSSVM functions by identifying the optimal hyperplane within a higher-dimensional space to separate data points into distinct classes, a process aided by the utilization of kernel functions [20, 21]. Its resilience against overfitting, capability to handle large feature spaces, and adaptability make it particularly suitable for soft measurement.

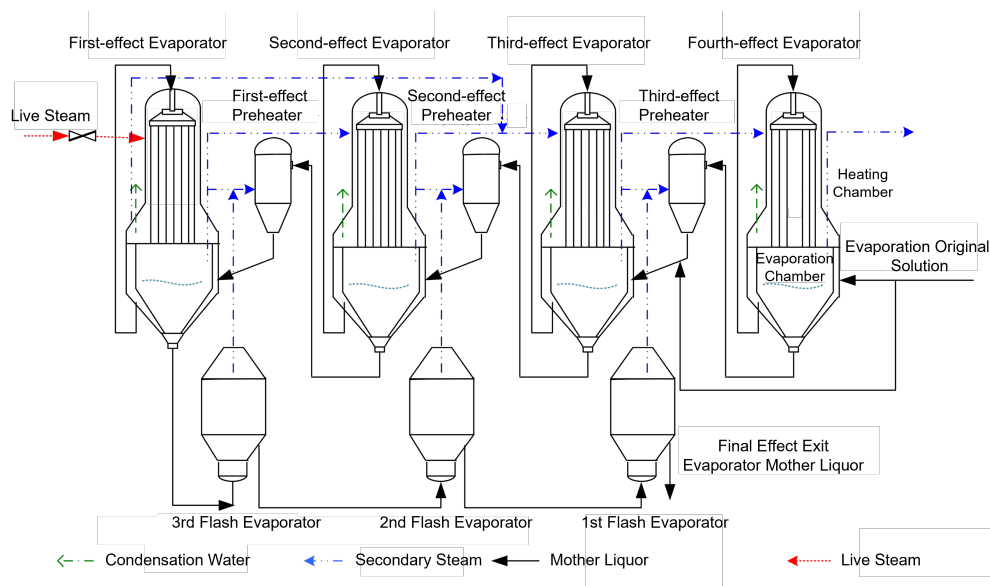
In 2017, based on the rough set's theory reduction capability and the nonlinear adaptive prowess of SVM, Wang and Chen [22] proposed an RS-LSSVM-based soft sensor model for accurately determining the burning zone temperature in rotary kilns. Through global discretization and attribute reduction, coupled with immune evolutionary algorithm optimization, the resulting model surpasses traditional LSSVM in accuracy and resilience against interference. Moreover, Zheng et al. [23] tackled the constraints of static soft measurement models in capturing dynamic information while monitoring temperature in the firing zone of cement rotary kilns. By integrating the LSSVM with the autoregressive moving average model (ARMA), and using cross-validation and grid search for optimization, a more dynamic and responsive model was developed. In 2022, the ILGSSA-LSSVM model was developed, leveraging the Improved Logistic Chaos Mapping and Golden Sine Algorithm to predict the surface temperature of continuous casting billets [24]. When compared to traditional methods like Gray Wolf optimized LSSVM and Backpropagation (BP) neural network, the proposed model exhibited superior accuracy with an average error of 0.05805 °C. In [25], Liu et al. applied LSSVM to model and predict effluent COD (chemical oxygen demand) levels in an anaerobic wastewater treatment system. While the steady-state LSSVM model demonstrated satisfactory results

in predicting effluent COD, the dynamic-state models excelled under various shock load scenarios, especially under the absence of a bicarbonate buffer.

### 3. Alumina evaporation process

In the Bayer process of alumina production, the evaporation step is used to evaporate the excess water from the seed mother liquor and washing filtrate, ensuring that the alkali concentration in the outlet mother liquor meets the requirements for the dissolution process. This promotes the recycling of the alkali solution and reduces its discharge.

In the actual alumina evaporation process, if single-effect evaporation is used, there are problems such as low evaporation capacity, high fresh steam consumption and low secondary steam utilization rate. Simultaneously, the original evaporation solution contains plenty of impurity salts and has high viscosity. As the concentration of the original evaporation solution increases, impurities tend to crystallize. If the crystallized salt cannot be discharged from the equipment in time, it will cause blockages. Hence, in practical production, enhancing salt solubility, minimizing precipitation in equipment or pipelines, and achieving improved evaporation outcomes are achieved through the adoption of a two-stage evaporation process, combining multi-effect flash evaporation with counter-current operation. In the multi-effect counter-current evaporation process, the order in which materials and steam enter each device is opposite. This article studies the four-effect three-flash falling film evaporation process of an alumina plant's evaporation workshop, with the specific process flow shown in Figure 1.



**Figure 1.** Process flow diagram of the four-effect triple flash evaporation in alumina production.

The four-effect three-flash evaporation process of alumina mainly includes four evaporators, three preheaters, three flash evaporators, and some condensate water tanks. According to the process flow shown in Figure 1, the original evaporation solution first enters the third-effect evaporator and the

fourth-effect evaporator. The solution flows from the fourth-effect evaporator to the third-effect preheater, raising the solution temperature through the preheater. When the solution temperature approaches the boiling point temperature of the third-effect evaporator, the solution is sent from the preheater to the third-effect evaporator. Following this pattern, the solution successively flows through the third-effect, second-effect and first-effect evaporators, with its temperature and concentration continuously increasing. After exiting the first-effect evaporator, the evaporated mother liquor flows through the three flash evaporators in sequence, and the concentrated mother liquor is pumped to the blending tank through the evaporation process outlet pump.

The primary heat source for the evaporation process comes from the fresh steam of the thermal power plant. The fresh steam first enters the first-effect evaporator, indirectly heating the solution outside the heating tube, simultaneously producing first-stage steam condensate. Subsequently, the secondary steam produced by the first-effect evaporator enters the second-effect evaporator, the secondary steam from the second-effect evaporator enters the third-effect evaporator and the secondary steam from the third-effect evaporator enters the fourth-effect evaporator. Finally, the secondary steam produced by the fourth-effect evaporator enters the water cooler to ensure smooth steam discharge from the evaporation system.

#### 4. Selection of auxiliary variables and data dimensionality reduction

The evaporation process is a nonlinear, uncertain, time-varying, highly coupled and lengthy production process. There are many input parameters affecting the mother liquor concentration, and directly constructing a soft measurement model with these input-output parameters will inevitably impact modeling accuracy. It's imperative to understand that while having an abundance of data might seem advantageous, not all of it adds value. Redundant and collinear data can distort the model's predictive power, leading to inefficiencies and inaccuracies. By simplifying the model and focusing on significant variables, we can achieve more accurate predictions with faster computation. Therefore, this paper first performs data dimensionality reduction on production data samples based on transfer entropy-based grey relational (GRA) analysis and kernel principal component analysis (KPCA).

##### 4.1. Variable selection based on transfer entropy grey relational analysis

GRA [26,27] is a relative ranking analysis aiming to quantitatively measure the level of connection between different factors in a system. Its primary objective is to evaluate the significance of factors impacting the concentration in the evaporation process, serving as a foundational step towards simplifying our model. The fundamental idea behind GRA, as introduced by Ozgur [28], centers on evaluating the degree of association by comparing sequences of geometric shapes within a spatial context. This process entails the identification of both the output sequence reflecting system behavior characteristics and the input sequence influencing system behavior, followed by a subsequent standardization step. Let  $\xi_{io}(k)$  and  $r_{io}$  be the correlation coefficient and the number of correlation degrees of the input sequence  $\{X_i(k)\}$  and output sequence  $Y_o(k)$  at time  $t = k$ , respectively. We have:

$$\xi_{io}(k) = \frac{\Delta \max + \Delta \min}{\Delta_{io}(k) + \lambda \Delta \max}, \quad (4.1)$$

$$r_{io} = \frac{1}{L} \sum_{k=1}^L \xi_{io}(k). \quad (4.2)$$

Where  $\lambda$  is the resolution coefficient,  $0 < \lambda < 1$ ;  $\Delta \min$  and  $\Delta \max$ , respectively, represent the minimum and maximum absolute differences of the comparison sequence at each moment.  $\Delta \min$  is the absolute difference between each point on the input sequence  $\{X_i(k)\}$  curve and each point on the output sequence  $Y_o(k)$  curve, and  $L$  represents the sampling moment.

The classic GRA overlooks the direct influence degree of different factors in different output sequences, i.e., the weight size. Sometimes all correlation degrees are very close and the distribution interval is small, making it hard to discern the similarity between the standard sequence and the sequence to be tested, and it's challenging to evaluate objectively and accurately. Therefore, through the integration of subjective weighting using the Analytic Hierarchy Process (AHP) [29] and objective weighting employing the transfer entropy method, the GRA comprehensively addresses the differences in importance among various evaluation indicators. This ensures that the determined weight coefficients have both subjective and objective information. The specific calculation steps are as follows:

- 1) First, use the Analytic Hierarchy Process to determine the weight of each evaluation indicator  $S_k$ ;
- 2) Then, use the transfer entropy method according to formula (4.3) to calculate the objective weight  $O_k$ ;

$$O_k = \sum_{y_n, \tilde{y}_{n-1}^{(l)}, \tilde{x}_{n-1}^{(k)}} P(y_n, \tilde{y}_{n-1}^{(l)}, \tilde{x}_{n-1}^{(k)}) \log_2 \frac{P(y_n, \tilde{y}_{n-1}^{(l)}, \tilde{x}_{n-1}^{(k)})}{P(y_n, \tilde{y}_{n-1}^{(l)})}. \quad (4.3)$$

Here,  $Q_{ik}$  represents the standardization of the evaluation matrix.

- 3) Finally, apply AHP to provide subjective weighting for each indicator and combine it with the objective weight of the entropy method to ultimately determine the weight of each indicator. To amplify the importance of the differences between the indicators, a multiplication synthesis method is used to combine weights for the evaluation indicators. This means multiplying the weight coefficients determined by both subjective and objective weighting methods and then normalizing the product results. The comprehensive weight coefficient is shown in formula (4.4).

$$w_k = \frac{S_k O_k}{\sum_{k=1}^L S_k O_k}. \quad (4.4)$$

By combining AHP and entropy weighting with the grey relational analysis method, the improved grey relational degree is obtained, as shown in formula (4.5).

$$r'_{io} = \frac{1}{L} \sum_{k=1}^L w_k \bullet \xi_{io}(k). \quad (4.5)$$

- 4) The calculated correlation degrees are then sorted in descending order.

Taking the four-effect countercurrent falling film evaporation process of an alumina plant as an example, numerous factors influence the outlet mother liquor concentration. Only relying on the

qualitative analysis of the on-site operator's experience and mechanism analysis results to determine the main influencing factors is not highly credible. Therefore, based on the mechanism analysis and the on-site operator's qualitative analysis experience, we use the comprehensive grey relational analysis method to determine the impact degree of each factor on the soft measurement model from a quantitative analysis perspective. The analysis yielded the new steam temperature  $TX$ , new steam flow rate  $LX$ , original liquid temperature  $TY$ , original liquid flow rate  $LY$ , original liquid concentration  $CY$ , and the liquid temperatures from I effect to IV effect as  $TI$ ,  $TII$ ,  $TIII$ ,  $TIV$ , respectively, and the vapor pressures from I effect to IV effect as  $PI$ ,  $PII$ ,  $PIII$ ,  $PIV$ , and the heat transfer coefficient  $C2$  as input sequences  $\{X_i(k)\}$ . The output concentration  $C0$  is the output sequence  $\{X_o(k)\}$ . Selecting 500 sets from the production data as samples and using the mean method to process the samples to eliminate the impact of dimensions, we calculated the results using the proposed transfer entropy-based grey correlation method.

$$\begin{aligned} QI(0, 1) &= 0.8281, & QI(0, 2) &= 0.8364 \\ QI(0, 3) &= 0.9647, & QI(0, 4) &= 0.8213 \\ QI(0, 5) &= 0.8296, & QI(0, 6) &= 0.4614 \\ QI(0, 7) &= 0.8196, & QI(0, 8) &= 0.8327 \\ QI(0, 9) &= 0.8218, & QI(0, 10) &= 0.4014 \\ QI(0, 11) &= 0.8453, & QI(0, 12) &= 0.9342 \\ QI(0, 13) &= 0.5993, & QI(0, 14) &= 0.4905 \end{aligned}$$

The correlations are arranged from highest to lowest as follows:

$$TY > PIII > C2 > PII > LX > TIII > TX > TIV > CY > LY > TII > PIV > TI > PI$$

We can see from the results that if we choose a correlation degree greater than 0.85 or 0.9, we would exclude the new steam flow rate  $LX$  and the original liquid flow rate  $LY$ , which are two crucial control variables. Therefore, combining the actual operating experience of the alumina evaporation process and the results of mechanism modeling, we selected 10 variables with a correlation degree greater than 0.81 as influencing factors. We then used the kernel principal component analysis to extract features from these 10 selected variables.

#### 4.2. Kernel principal component analysis

KPCA [30, 31] employs a nonlinear kernel function to map sample data from the original space to a linear high-dimensional feature space. In essence, KPCA aids in distilling the essence of our data, focusing on principal components that significantly contribute to the model's predictive power. Through nonlinear mapping, samples formed from the 10 variables selected via gray relational analysis constitute an input space  $X_i (i = 1, 2, \dots, m) \in R^M$  and are subsequently mapped to a high-dimensional feature space  $F$ . Samples in  $F$  are denoted as  $\phi(x_i)$ . The covariance matrix in the high-dimensional feature space is calculated as per Eq (4.6).

$$C = \frac{1}{m} \sum_{i=1}^m \phi(x_i) \phi(x_i)^T. \quad (4.6)$$

The covariance matrix  $C$  undergoes eigenvalue decomposition as shown in Eq (7),

$$\lambda v = CV, \quad (4.7)$$

where  $\lambda$  represents the eigenvalues of the covariance matrix  $C$  and  $v$  corresponds to the eigenvectors. Multiplying both sides of Eq (4.7) by the kernel sample  $\phi(x_k)$  results in:

$$\lambda \phi(x_k) \cdot v = \phi(x_k) \cdot CV, k = 1, 2, \dots, m. \quad (4.8)$$

For all eigenvectors  $v$  of  $\lambda \neq 0$ , there exists a relation described by  $\alpha_i (i = 1, \dots, m)$ :

$$v = \sum_{i=1}^m \alpha_i \phi(x_i). \quad (4.9)$$

Introducing the kernel function  $K_{ij}$ , we get:

$$K_{ij} = K(x_i, x_j) = \phi(x_i) \phi(x_j). \quad (4.10)$$

The eigenvectors and eigenvalues of the kernel function matrix  $K$  are given by:

$$m\lambda\alpha = K\alpha, \quad (4.11)$$

where  $\alpha$  represents the eigenvector of matrix  $K$ , and  $m$  denotes the total number of samples. For any vector  $x$ , its projection on the principal component direction  $\phi(x)$  in the feature space is:

$$v \cdot \phi(x) = \sum_{i=1}^m \alpha_i \phi(x_i) \phi(x) = \sum_{i=1}^m \alpha_i K(x_i, x). \quad (4.12)$$

The radial basis function (RBF) kernel is chosen as:  $K(x_i, x) = \exp\left(-\frac{\|x_i - x\|^2}{2\sigma^2}\right)$ .

The number of principal components  $s$  is typically selected based on the following rule:

$$\left( \sum_{i=1}^s \lambda_i / \sum_{i=1}^m \lambda_i \right) > E. \quad (4.13)$$

The value of  $E$  is usually greater than 85%.

To compare the data dimensionality reduction effects, KPCA is applied to both the 14 variables before gray relational analysis and the 10 variables after. The results are shown in Tables 1 and 2.

Based on the KPCA results from Tables 1 and 2, principal components with cumulative contribution rates greater than 95% are chosen. After gray relational analysis, the 6 principal components, compared to the 8 before the analysis, streamline the data dimensionality. This reduction not only accelerates model training time but also hones the model's focus on the most impactful variables, amplifying its predictive capabilities.

## 5. Reduced robust LSSVM modeling based on PS-DE

As mentioned earlier, due to characteristics such as nonlinearity, uncertainty, time variability and strong coupling, it's challenging to accurately model the evaporation process using conventional mathematical methods. Mechanism-based soft measurement modeling methods often fall short of the



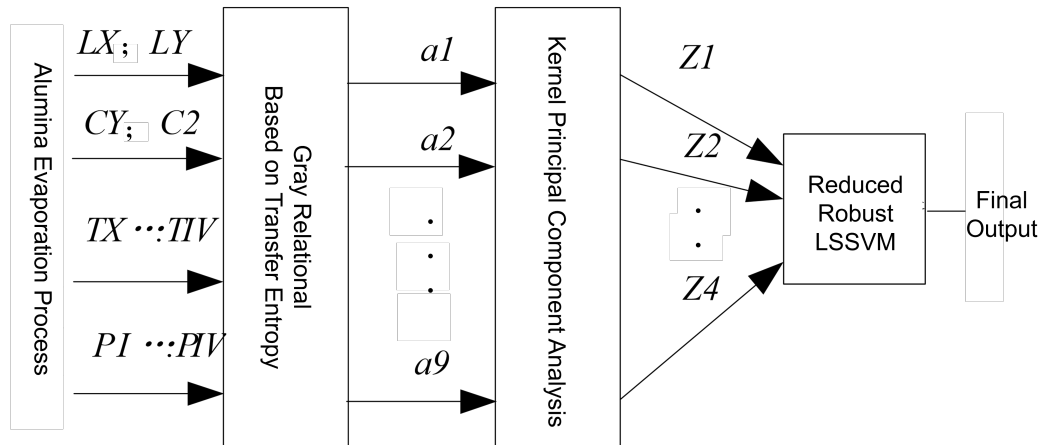
**Table 1.** Fourteen variable kernel principal component analysis results before gray relational analysis.

No.	Eigenvalue	Contribution rate %	Accumulating contribution rate %
1	8.7765	51.6263	51.6263
2	2.679	15.759	67.3854
3	1.4299	8.411	75.7964
4	1.0179	5.9876	88.9979
5	0.6486	2.8154	91.8134
6	0.1748	1.0028	95.5196
7	0.1369	0.9054	96.425
8	0.0677	0.8981	97.3231
9	0.0582	0.7424	98.0655
10	0.0383	0.7251	98.7906
11	0.0115	0.6674	99.458
12	0.0005	0.0029	99.9985
13	0.0003	0.0010	99.9995
14	0.0001	0.0005	100

**Table 2.** Ten variable kernel principal component analysis results after gray relational analysis.

No.	Eigenvalue	Contribution rate %	Accumulating contribution Rate %
1	6.1527	62.5273	62.5273
2	2.6894	20.6875	73.2487
3	1.3836	13.8676	76.3949
4	1.0585	10.5853	87.9802
5	0.7484	7.4844	95.4646
6	0.3470	2.4700	97.9346
7	0.2180	1.1802	99.1149
8	0.0729	0.7289	99.8437
9	0.0091	0.0915	99.9352
10	0.0044	0.0648	100

predictive accuracy required for actual industrial production. Therefore, this paper introduces a soft measurement model based on data transfer entropy, gray relational analysis and KPCA, named the reduced robust least squares support vector machine (LSSVM). The detailed soft measurement modeling process is presented in Figure 2.



**Figure 2.** Modeling process diagram of the reduced robust LSSVM based on PS-DE.

### 5.1. Reduced robust LSSVM model

Suykens et al. [9, 32] proposed the least squares support vector machine (LSSVM), which substitutes the squared loss function in place of the “ $\mathcal{E}$ -insensitive loss function”. This change replaces the inequality constraints with equality constraints. Solving these problems is equivalent to solving a set of linear equations in the dual space, thereby reducing computational complexity. However, this also diminishes the noise resistance of the LSSVM. If the noise distribution doesn't follow a Gaussian distribution, the obtained solution might deviate significantly from the actual value. To enhance the prediction robustness of the LSSVM, a weighting factor  $\mu_i$  is introduced into the error term  $\alpha_i = \gamma_i \xi_i$  of the standard LSSVM. The Lagrange equation becomes:

$$L(\omega, b, \xi, \alpha) = \frac{1}{2} \omega^T \omega + \frac{C}{2} \sum_{i=1}^N \mu_i \xi_i^2 - \sum_{i=1}^N \alpha_i (\omega \phi(x) + b + \xi_i - y_i). \quad (5.1)$$

The optimization problem derived from the KKT conditions is:

$$\min_{\omega, b, \xi} J(\omega, \xi) = \frac{1}{2} \omega^T \omega + \frac{C}{2} \sum_{i=1}^N \mu_i \xi_i^2. \quad (5.2)$$

Defining a kernel function  $K$ , and eliminating unknown parameters, we get:

$$\begin{bmatrix} 0 & 1_\mu^T \\ 1_\mu & K + CU \end{bmatrix} \begin{bmatrix} b \\ \alpha \end{bmatrix} = \begin{bmatrix} 0 \\ y \end{bmatrix}. \quad (5.3)$$

Where  $K$  represents the kernel matrix,  $CU = \text{diag} \left( \frac{1}{C\mu_1}, \dots, \frac{1}{C\mu_N} \right)$ .

The weighting factor  $\mu_i$  is adaptively determined by the error  $\xi_i = \alpha_i/C$ .

$$\mu_i = \begin{cases} 1 & \text{if } \xi_i/s' \leq c_1 \\ \frac{c_2 - \xi_i/s'}{c_2 - c_1} & \text{if } c_1 \leq \xi_i/s' \leq c_2. \\ 10^{-4} & \text{Other} \end{cases} \quad (5.4)$$

Where  $S'$  is an estimated value of the standard deviation of the error  $\xi$ , and  $c_1$  and  $c_2$  are constants, typically set to values such as  $c_1 = 2.5$  and  $c_2 = 3$ . If the error is below a preset threshold, the loss function of the robust LSSVM can be simplified to the standard LSSVM form. If the error lies within a predefined range, the penalty function's weighting factor will adjust adaptively with the error. If it exceeds a preset upper limit, a small constant is used to suppress it. This enhances the prediction accuracy and improves the model's resistance to interference.

Given the multiple factors affecting the outlet mother liquor concentration in the alumina evaporation process and the numerous input variables in the reduced robust LSSVM model, one of the flaws of the LSSVM is that its solutions lack sparsity, i.e., each input sample is represented by a support vector value. To overcome this flaw, [33–35] introduced various methods such as introducing weighting factors, matrix reduction by extracting feature vectors, introducing insensitive zone bandwidths, and redundant sample information. These efforts aim to increase the sparsity of the solutions, accelerate modeling speed, and enhance the model's robustness. In this section, we employ the Schmid orthogonalization method [36] to reduce the kernel matrix within the robust LSSVM model, introducing a degree of sparsity to the solution.

For each training sample  $\mathbf{x}_i$ , its mapping can be represented by  $\varphi(\mathbf{x}_i)$  ( $1 \leq i \leq N$ ), and all samples can be mapped to matrix  $\{\varphi(\mathbf{x}_1), \dots, \varphi(\mathbf{x}_N)\}$ . If there exists a base matrix  $\{\varphi(\tilde{\mathbf{x}}_1), \dots, \varphi(\tilde{\mathbf{x}}_M)\}$  in this mapping matrix, then any mapped vector [37] can be represented as:

$$\begin{bmatrix} \varphi(\mathbf{x}_1) \\ \vdots \\ \varphi(\mathbf{x}_N) \end{bmatrix} = \begin{bmatrix} \alpha_{11} & \cdots & \alpha_{1M} \\ & \ddots & \\ \alpha_{N1} & \cdots & \alpha_{NM} \end{bmatrix} \begin{bmatrix} \varphi(\tilde{\mathbf{x}}_1) \\ \vdots \\ \varphi(\tilde{\mathbf{x}}_M) \end{bmatrix}. \quad (5.5)$$

It is evident that the base matrix in the mapping matrix is of paramount importance. References [38–40] aimed to find a linearly independent vector group in the hyper-space with the smallest Euclidean distance as the target. Rosipal et al. [41] introduced a technique that utilizes the Schmid orthogonalization method to reduce the kernel matrix, resulting in a set of linearly independent vectors that form the basis matrix.

The theoretical foundation of the Schmid orthogonalization method [42] is relatively mature. This section uses this algorithm to construct the base of the model kernel matrix. Following the principles of Schmid orthogonalization theory, the orthogonalization of column vector  $\varphi(\mathbf{x}_a)$  in the mapping matrix in hyper-space can be expressed as:

$$\varphi_{t+1}(\mathbf{x}_a) = \varphi_t(\mathbf{x}_a) - (\varphi_t(\mathbf{x}_a)^T \mathbf{v}_t) \mathbf{v}_t. \quad (5.6)$$

Where  $\mathbf{v}_t = \frac{\varphi_t(\mathbf{x}_i)}{\sqrt{\varphi_t(\mathbf{x}_i)^T \varphi_t(\mathbf{x}_i)}}$  and

$$\varphi(\mathbf{x}_i) \quad (1 \leq i \leq N) \quad (5.7)$$

are the selected vectors.

For the kernel matrix  $\mathbf{G}(a, b)$ :

$$\mathbf{G}(a, b) = \varphi(\mathbf{x}_a)^\top \varphi(\mathbf{x}_b) = K(\mathbf{x}_a, \mathbf{x}_b). \quad (5.8)$$

From which, the Gram form is:

$$\mathbf{G}_{t+1}(a, b) = \varphi_{t+1}(\mathbf{x}_a)^\top \varphi_{t+1}(\mathbf{x}_b) = \mathbf{G}_t(a, b) - \frac{\mathbf{G}_t(a, \mathbf{x}_i) \mathbf{G}_t(b, \mathbf{x}_i)}{\mathbf{G}_t(\mathbf{x}_i, \mathbf{x}_i)}. \quad (5.9)$$

During the construction of the reduced kernel matrix, a greedy algorithm is used to select vectors one by one. The magnitude of  $\mathbf{G}(i, i)$  is used as the basis for selecting vectors, i.e., each time the column  $\mathbf{x}_p$  where  $\mathbf{G}(i, i)$  is the largest is selected. The subsequent column vectors in the original data matrix are then orthogonalized. The algorithm is as follows:

Step 1: Let  $\hat{\mathbf{G}}_0(p, p) = \mathbf{K}(\mathbf{x}_p, \mathbf{x}_p)$ .

Step 2: for  $t = 0 : (d - 1)$ ,  $\hat{\mathbf{G}}_0(t, p) = \mathbf{K}(\mathbf{x}_t, \mathbf{x}_p)$ , end.

(Where:  $d$  is the rank of the matrix, representing the maximum number of linearly independent vector groups.)

Step 3: for  $t = 0 : (d - 1)$ , based on the selection criteria, a vector  $\mathbf{x}_i$  is chosen, denoted as  $\text{index}(t) = i$ ,

for  $s = (t + 1) : (d - 1)$ ,

$$\hat{\mathbf{G}}_{t+1}(s, p) = \hat{\mathbf{G}}_t(s, p) - \frac{\hat{\mathbf{G}}_t(t, p) \mathbf{G}_t(\text{index}(t), \text{index}(s))}{\mathbf{G}_t(\text{index}(t), \text{index}(t))}, \quad (5.10)$$

end

$$\hat{\mathbf{G}}_{t+1}(p, p) = \hat{\mathbf{G}}_t(p, p) - \frac{\hat{\mathbf{G}}_t(t, p) \hat{\mathbf{G}}_t(t, p)}{\mathbf{G}_t(\text{index}(t), \text{index}(t))}, \quad (5.11)$$

end

(Note: In the algorithm,  $d$  can either be the rank of the matrix or can be predetermined.)

## 5.2. Parameter optimization of reduced LSSVM based on the PS-DE algorithm

The integration of two optimization algorithms, pattern search (PS) and differential evolution (DE), forms the crux of the PS-DE algorithm. These algorithms, when combined, offer a powerful tool for optimizing the parameters of our reduced LSSVM model.

The pattern search (PS) algorithm, introduced by Hooke and Jeeves in the early 1960s, is classified as a direct search algorithm, as outlined in their work [43]. The main idea is to generate an iterative sequence without relying on any derivative information. During each iteration, if a better optimal solution can be derived from the iteration point, it is accepted; otherwise, the search continues. The PS algorithm begins from an initial point and alternates between two search operations: axial search and pattern search. The axial search method systematically explores along the  $n$  coordinate axes, aiming to identify a new base point by favoring directions that reduce the function's value. The pattern search operates along the direction connecting adjacent base points, aiming to accelerate the function value reduction. The steps of the PS algorithm are as follows:

Step 1: Set the initial point as  $x^{(0)}$ , initial step length as  $\delta$ , acceleration factor as  $\alpha$ , reduction factor as  $\beta \in (0, 1)$ , computation precision as  $\varepsilon > 0$  and  $e_1, e_2, \dots, e_n$  as the unit coordinate vector. Let  $y^{(1)} = x^{(0)}$  (initial point), set  $k = 0$ ,  $j = 1$ .

Step 2: If  $f(y^{(j)} + \delta e_j) < f(y^{(j)})$ , then the forward probe is successful, let  $y^{(j+1)} = y^{(j)} + \delta e_j$  and move to Step 3; Otherwise, the forward probe is unsuccessful.

If  $f(y^{(j)} - \delta e_j) < f(y^{(j)})$ , let  $y^{(j+1)} = y^{(j)} - \delta e_j$ , and move to Step 3;

If  $f(y^{(j)} - \delta e_j) \geq f(y^{(j)})$ , let  $y^{(j+1)} = y^{(j)}$ , and move to Step 3.

Step 3: If  $j < n$ , then let  $j = j + 1$  and return to Step 2. Otherwise, if  $j = n$  and  $f(y^{(n)}) < f(x^{(k)})$ , the probing move is successful, and proceed to Step 4. If not, move to Step 5.

Step 4: Let  $x^{(k+1)} = y^{(n)}$ ,  $y^{(1)} = x^{(k+1)} + \alpha(x^{(k+1)} - x^{(k)})$ ,  $k = k + 1$ ,  $j = 1$ , and return to Step 2.

Step 5: if  $\delta < \varepsilon$ , take  $x^* = x^{(k)}$ , the algorithm terminates. Otherwise, let  $\delta = \beta\delta$ ,  $y^{(1)} = x^{(k)}$ ,  $x^{(k+1)} = x^{(k)}$ ,  $k = k + 1$ ,  $j = 1$ , and return to Step 2.

The differential evolution (DE) algorithm [44–46] is an intelligent search algorithm for global optimization proposed by Storn and Price in 1995. It was designed specifically for real-number coded genetic individuals and uses differential operations to implement crossover and mutation in the context of the genetic algorithm framework. The core idea of the DE algorithm is to use the differential quantity from two randomly chosen individual vectors from the population as a perturbation to a third random base vector, resulting in a mutated vector. This mutated vector then undergoes crossover operations with the base or target vector to produce a trial vector. Finally, the base vector and trial vector compete, and the better one is retained in the next generation.

Select NP initial population solutions  $x_i^G (i = 1, 2, \dots, NP)$ , where  $i$  is the population size and  $G$  is the current evolutionary generation. For the  $G$ th individual  $x_i^G$ , perform mutation according to Eq (5.12) to obtain a new individual  $R_i^{G+1}$ :

$$R_i^{G+1} = x_{h3}^G + F \cdot (x_{h1}^G - x_{h2}^G). \quad (5.12)$$

Where,  $x_{h1}^G$ ,  $x_{h2}^G$ ,  $x_{h3}^G$  are three distinct individuals randomly chosen from generation  $G$ , and  $F$  is the mutation rate.

Next, carry out crossover using Eq (5.13) to produce the trial individual  $S_i^{G+1}$ :

$$S_{ij}^{G+1} = \begin{cases} R_{ij}^{G+1}, & \text{rand}(j) \leq CR \text{ or } j = \text{randn}(i) \\ x_{ij}^G, & \text{rand}(j) > CR \text{ and } j \neq \text{randn}(i) \end{cases}. \quad (5.13)$$

In this equation,  $CR$  is the crossover probability constant,  $\text{rand}(j)$  is a random number uniformly distributed between 0 and 1, and  $\text{rand}(i)$  is a randomly chosen integer.

Finally, decide whether to retain  $x_i^{G+1}$  using Eq (5.14):

$$x_i^{G+1} = \begin{cases} S_i^{G+1}, & f(S_i^{G+1}) < f(x_i^G) \\ x_i^G, & f(S_i^{G+1}) \geq f(x_i^G) \end{cases}. \quad (5.14)$$

Here,  $f$  represents the fitness function.

Building upon the PS-DE algorithm and the analysis of the reduced robust LSSVM, the PS-DE optimization algorithm is applied for parameter tuning of the reduced robust LSSVM. This is a critical juncture as it amplifies the prediction performance of our model. The PS-DE algorithm optimizes the

penalty coefficient and kernel width coefficient of the reduced robust LSSVM. Initially, it utilizes the DE algorithm for global optimization, and subsequently, it employs the PS algorithm for more localized exploration. The optimization process alternates between these methods, eventually converging to a global optimum. The detailed steps are as follows:

Step 1: Initialize the population size  $NP$ , mutation operator  $F$ , crossover factor  $CR$ , maximum iteration for the DE algorithm  $G_{max}$ , initial step length  $\delta$ , acceleration factor  $\partial$ , reduction rate  $\beta \in (0, 1)$  and precision  $\varepsilon > 0$ .

Step 2: For the initial population, perform mutation, crossover and selection operations as per Eq (14)–(16), resulting in new individuals  $X_i^{G+1}$ . Update the objective function values  $f(X_i^{G+1})$  accordingly.

Step 3: Check if the initial step length meets the computational precision. If it does, move to Step 4; otherwise, return to Step 2.

Step 4: Execute the pattern search algorithm, recording the optimal objective function value  $f(\bar{X}_i^{G+1})$  and the corresponding new individual  $\bar{X}_i^{G+1}$ .

Step 5: Check the termination criteria. If met, the algorithm terminates and outputs the optimal parameters  $(C, \sigma)$ . Otherwise, return to Step 2.

## 6. PS-DE used for verification and analysis of the reduced robust LSSVM model

### 6.1. Verification and analysis of the PS-DE algorithm

To assess the efficacy and practicality of the PS-DE algorithm, we conducted experiments on three standard function optimization problems (aimed at minimizing the function values). We evaluated the PS-DE algorithm's performance and compared it with that of the DE algorithm.

1) Sphere function:

$$f_1(x) = \sum_i^n x_i^2. \quad (6.1)$$

Optimal value:  $x_i = 0, f_1(x) = 0, x_i \in [-100, 100]$ .

2) Rosenbrock function:

$$f_2(x) = \sum_{i=1}^n \left( 100 \times (x_{i+1} - x_i^2)^2 + (1 - x_i)^2 \right), x_i \in [-30, 30]. \quad (6.2)$$

Its optimal value and its corresponding optimal state are:

$$\min f(x^*) = f(1, 1, \dots, 1) = 0.$$

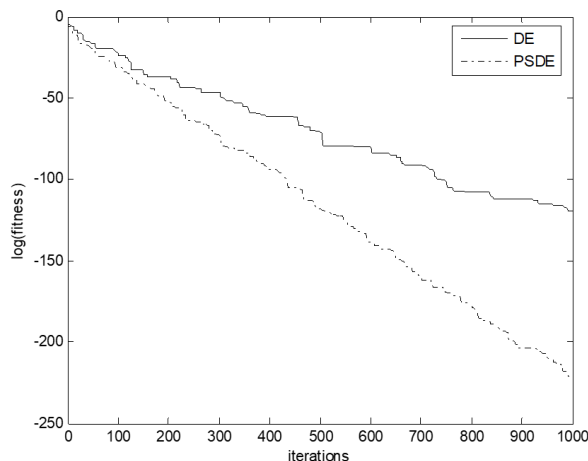
3) Rastrigin function:

$$f_3(x) = \sum_{i=1}^n \left( x_i^2 - 10 \cos(2\pi x_i) + 10 \right), x_i \in [-5.12, 5.12]. \quad (6.3)$$

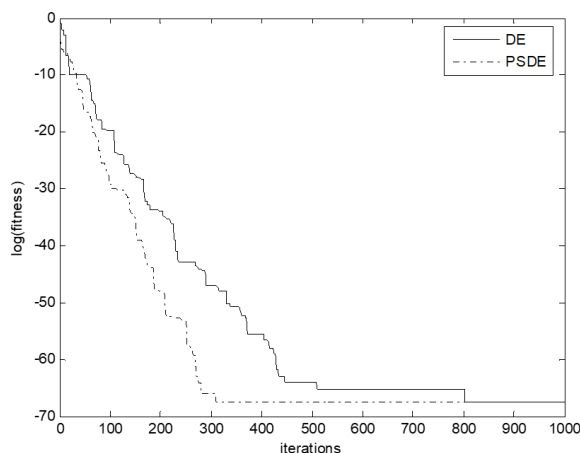
Its optimal value and its corresponding optimal state are:  $\min f(x^*) = f(0, 0, \dots, 0) = 0$ .

The simulations were run on a platform based on Windows XP and MATLAB 7.0, with a 2.1 GHz CPU and 1 GB RAM. The simulation was conducted in a 30-dimensional space, with the termination condition being the maximum number of iterations. Initial parameters were set as: initial step length

$\delta = 1.618$ , acceleration factor  $\alpha = 0.618$ , reduction factor  $\beta = 0.382$ , precision  $\varepsilon = 10^{-5}$ , optimization count of 50,  $NP = 100$ ,  $F = 0.6$ ,  $CR = 0.9$ ,  $G_{\max} = 1000$ . The simulation results are shown in Figures 3–5, and Table 3.

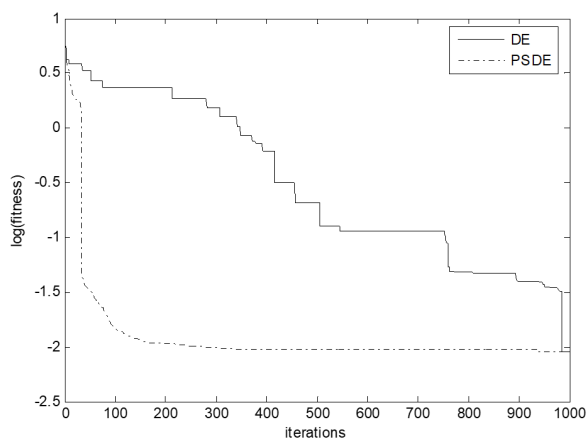


**Figure 3.** Optimization comparison for the Sphere function.



**Figure 4.** Optimization comparison for the Rosenbrock function.

From Table 3 and Figures 3–5, it can be observed that for the three typical test functions, the PS-DE algorithm generally outperforms the DE algorithm in terms of optimization results, convergence speed and success rate. During the experimental testing, the PS-DE algorithm, which incorporates the PS algorithm with robust local search capabilities into the DE algorithm, effectively preserved the global search ability of the novel algorithm. By observing the dynamic search graphs of the particles, it was found that for the three test functions, the PS-DE algorithm could quickly reach the optimal point and hardly fluctuated around it. In contrast, the DE algorithm reached the optimal point more slowly and oscillated around it before slowly converging to the global optimum, sometimes even failing to reach the optimum. Especially in the high-dimensional test experiment for the multi-peak Rastrigin function, the PS-DE algorithm showed a significant improvement in optimization for the Rastrigin function.



**Figure 5.** Optimization comparison for the Rastrigrin function.

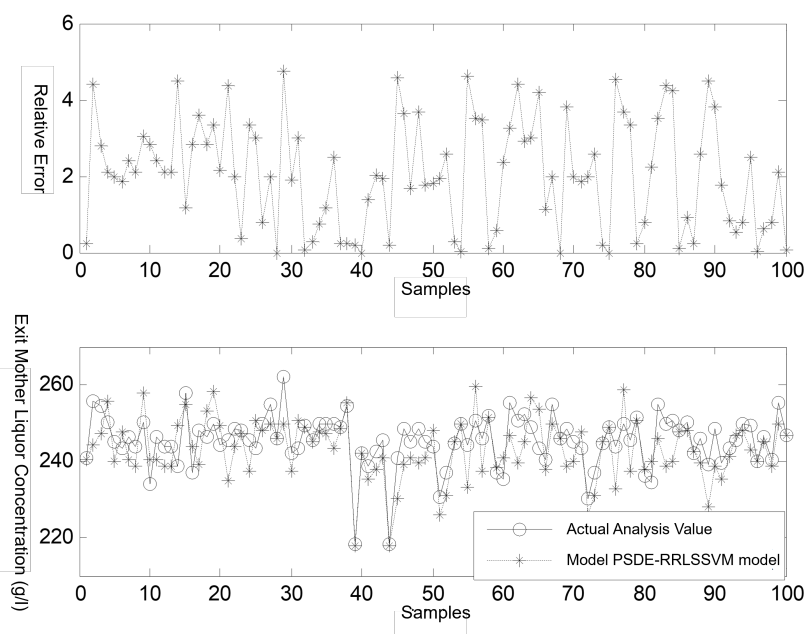
**Table 3.** Optimization comparison between PS-DE algorithm and DE algorithm.

Test function	Optimization algorithm	Average value	Minimum number of iterations	Success rate
Sphere	DE	$1.56 \times 10^{-5}$	640	87.30%
	PSDE	$1.34 \times 10^{-7}$	550	98.70%
Rosenbrock	DE	$6.73 \times 10^{-1}$	760	80.00%
	PSDE	$4.22 \times 10^{-11}$	220	99.60%
Rastrigin	DE	$4.11 \times 10^{-1}$	790	77.60%
	PSDE	$3.26 \times 10^{-3}$	186	93.50%



## 6.2. Model verification and analysis

We selected 500 sets of industrial field data after conducting steady-state detection and data coordination. These were employed as training samples, while an additional 200 sets were reserved for testing. Before training, the data was normalized. During parameter optimization for the reduced robust LSSVM model, the penalty coefficient  $C$  and the kernel function width  $\sigma$  had initial selection ranges of  $[0, 1000]$  and  $[0, 10]$ , respectively. Algorithm parameters were set as:  $NP = 100$ ,  $F = 0.6$ ,  $CR = 0.9$ ,  $G_{\max} = 1000$  and precision  $\varepsilon = 10^{-5}$ . The PS-DE algorithm was employed to choose the optimal parameters for the reduced robust LSSVM (denoted as PSDE-RRLSSVM), resulting in optimal parameter pairs of  $(158.7, 1.264)$ . The DE algorithm (denoted as DE-RRLSSVM) yielded optimal parameter pairs of  $(C, \sigma) = (158.7, 1.264)$ . Simulation results are shown in Figures 6 and 7, and error analysis is presented in Table 4.

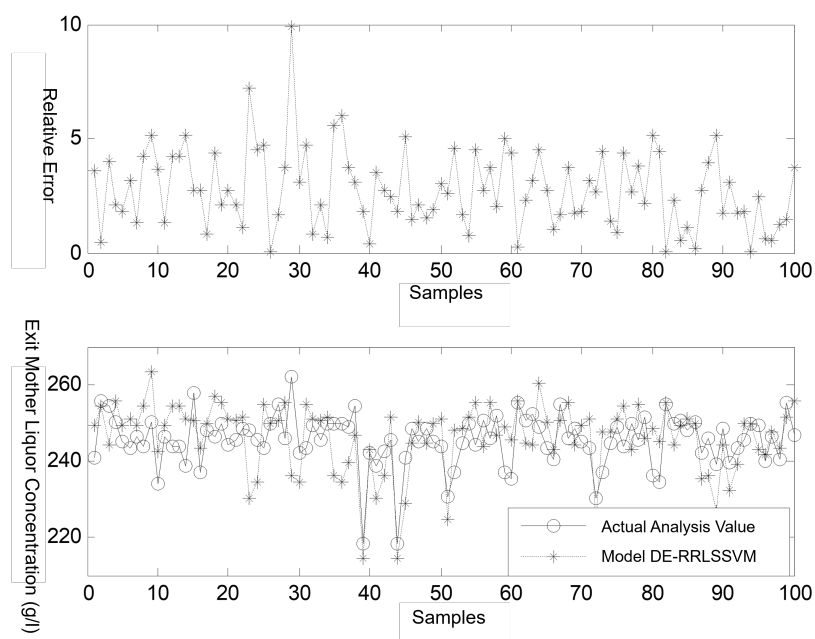


**Figure 6.** Soft measurement simulation results of the exit mother liquor concentration based on PSDE-RRLSSVM.

**Table 4.** Error result analysis.

	Emax%	RMSE	RRMSE%
PSDE-RRLSSVM	4.7243	0.0061404	7.84
DE-RRLSSVM	9.8798	0.0082234	9.07
Mechanism model	12.0345	0.023653	13.56

Figures 5 and 6, along with the calculations presented in Table 4, reveal that the PSDE-reduced robust LSSVM integrated model outperforms the DE-reduced robust LSSVM in soft measurement. The maximum relative error stands at 4.7243%, the root mean square error is 0.0061404, and the root mean square relative error is 7.84%. These simulation results demonstrate the model's exceptional



**Figure 7.** Soft measurement simulation results of the exit mother liquor concentration based on DE-RRLSSVM.

accuracy in soft measurement, aligning with the stringent requirements of the production process and providing a solid foundation for real-time operational optimization in the evaporation process.

## 7. Conclusions

The concentration of the exit mother liquor is a vital control index in the alumina evaporation process. Given the multitude of influencing factors and the challenge of online detection, this study proposes a mother liquor concentration soft measurement model based on the PS-DE reduced robust least squares support vector machine, integrated with comprehensive grey relation and kernel principal component analysis. As a result, potential pitfalls such as data redundancy and collinearity, common in evaporation process data, are effectively circumvented, streamlining the model's input samples. On one hand, through grey relation and kernel principal component analysis, auxiliary variables are screened and sample characteristic information is extracted. This eliminates the redundancy and collinearity of data samples from the evaporation process, simplifying the model's input samples. On the other hand, the introduction of the PS algorithm in the new method addresses both the slow convergence speed and the tendency of the differential evolution algorithm to become stuck in local optima. It also fine-tunes the reduced LSSVM parameters, ensuring the algorithm's convergence speed and precision.

The verification and analysis using actual production data indicate that the proposed method is effective. The established model for soft measurement of exit mother liquor concentration exhibits superior learning and generalization capabilities when compared to the DE-LSSVM and LSSVM models. Given the robustness and adaptability of the proposed method, it has potential applications beyond the alumina evaporation process. Specifically, industries that involve complex chemical processes, such as petrochemicals, pharmaceuticals and food processing, could benefit from the soft

measurement capabilities of this method. Furthermore, processes that require precise concentration control, like wastewater treatment or fermentation processes, might also find the model advantageous in ensuring product quality and process efficiency.

### Use of AI tools declaration

The authors declare they have not used Artificial Intelligence (AI) tools in the creation of this article.

### Acknowledgments

This study was supported in part by the National Natural Science Foundation Project under Grant No. 61963036, in part by Jiangxi Provincial Department of Education Science and Technology Project under Grant No. GJJ2201719.

### Conflict of interest

The authors declare there is no conflict of interest.

### References

1. Z. Huda, N. I. Taib, T. Zaharinie, Characterization of 2024-t3: An aerospace aluminum alloy, *Mater. Chem. Phys.*, **113** (2009), 515–517. <https://doi.org/10.1016/j.matchemphys.2008.09.050>
2. J. Zhao, Y. Lv, Output-feedback robust tracking control of uncertain systems via adaptive learning, *Int. J. Control Autom. Syst.*, **21** (2023), 1108–1118. <https://doi.org/10.1007/s12555-021-0882-6>
3. A. Smirnov, D. Kibartas, A. Senyuta, A. Panov, *Miniplant tests of hcl technology of alumina production*, Light Metals, Springer, (2018), 57–62. <https://doi.org/10.1007/978-3-319-72284-9>
4. C. Damour, M. Benne, B. Grondin-Perez, J. P. Chabriat, Soft-sensor for industrial sugar crystallization: On-line mass of crystals, concentration and purity measurement, *Control Eng. Pract.*, **18** (2010), 839–844. <https://doi.org/10.1016/j.conengprac.2010.03.005>
5. Y. Meng, Q. Lan, J. Qin, S. Yu, H. Pang, K. Zheng, Data-driven soft sensor modeling based on twin support vector regression for cane sugar crystallization, *J. Food Eng.*, **241** (2019), 159–165. <https://doi.org/10.1016/j.jfoodeng.2018.07.035>
6. S. Jouenne, G. Heurteux, B. Levache, Online monitoring for measuring the viscosity of the injected fluids containing polymer in chemical eor, in *SPE EOR Conference at Oil and Gas West Asia*, 2022. <https://doi.org/10.2118/200209-MS>
7. J. Tran, M. Linnemann, M. Piper, E. Kenig, On the coupled condensation-evaporation in pillow-plate condensers: Investigation of cooling medium evaporation, *Appl. Thermal Eng.*, **124** (2017), 1471–1480. <https://doi.org/10.1016/j.applthermaleng.2017.06.050>
8. A. Peters, W. Durner, Simplified evaporation method for determining soil hydraulic properties, *J. Hydrology*, **356** (2008), 147–162. <https://doi.org/10.1016/j.jhydrol.2008.04.016>
9. J. A. Suykens, J. Vandewalle, Least squares support vector machine classifiers, *Neural Process. Lett.*, **9** (1999), 293–300. <https://doi.org/10.1023/A:1018628609742>

10. Z. Liu, D. Yang, Y. Wang, M. Lu, R. Li, Eggn: Graph structure learning based on evolutionary computation helps more in graph neural networks, *Appl. Soft Comput.*, **135** (2023), 110040. <https://doi.org/10.1016/j.asoc.2023.110040>
11. Y. Wang, Z. Liu, J. Xu, W. Yan, Heterogeneous network representation learning approach for ethereum identity identification, *IEEE Trans. Comput. Social Syst.*, 2022. <https://doi.org/10.1109/TCSS.2022.3164719>
12. P. Kadlec, B. Gabrys, S. Strandt, Data-driven soft sensors in the process industry, *Comput. chem. Eng.*, **33** (2009), 795–814. <https://doi.org/10.1016/j.compchemeng.2008.12.012>
13. M. L. Fravolini, G. Del Core, U. Papa, P. Valigi, M. R. Napolitano, Data-driven schemes for robust fault detection of air data system sensors, *IEEE Trans. Control Syst. Technol.*, **27** (2017), 234–248. <https://doi.org/10.1109/TCST.2017.2758345>
14. Y. Wang, J. Ding, T. Chai, Soft-sensor for alkaline solution concentration of evaporation process, in *2008 7th World Congress on Intelligent Control and Automation*, (2008), 3476–3480. <https://doi.org/10.1109/WCICA.2008.4594499>
15. H. Su, W. Qi, Y. Schmirander, S. E. Ovrur, S. Cai, X. Xiong, A human activity-aware shared control solution for medical human–robot interaction, *Assembly Autom.*, **42** (2022), 388–394. <https://doi.org/10.1108/AA-12-2021-0174>
16. W. Qi, H. Su, A cybertwin based multimodal network for ecg patterns monitoring using deep learning, *IEEE Trans. Industr. Inform.*, **18** (2022), 6663–6670. <https://doi.org/10.1109/TII.2022.3159583>
17. H. Morales, F. di Sciascio, E. Aguirre-Zapata, A. N. Amicarelli, A model-based supersaturation estimator (inferential or soft-sensor) for industrial sugar crystallization process, *J. Process Control*, **129** (2023), 103065. <https://doi.org/10.1016/j.jprocont.2023.103065>
18. H. Wang, D. Hu, Comparison of svm and ls-svm for regression, in *2005 International conference on neural networks and brain*, **1** (2005), 279–283. <https://doi.org/10.1109/icnnb.2005.1614615>
19. W. Qi, H. Fan, H. R. Karimi, H. Su, An adaptive reinforcement learning-based multimodal data fusion framework for human–robot confrontation gaming, *Neural Networks*, **164** (2023), 489–496. <https://doi.org/10.1016/j.neunet.2023.04.043>
20. H. Xu, G. Chen, An intelligent fault identification method of rolling bearings based on lssvm optimized by improved pso, *Mech. Syst. Signal Process.*, **35** (2013), 167–175. <https://doi.org/10.1016/j.ymsp.2012.09.005>
21. W. Qi, S. E. Ovrur, Z. Li, A. Marzullo, R. Song, Multi-sensor guided hand gesture recognition for a teleoperated robot using a recurrent neural network, *IEEE Robot. Autom. Lett.*, **6** (2021), 6039–6045. <https://doi.org/10.1109/LRA.2021.3089999>
22. Y. Wang, X. Chen, On temperature soft sensor model of rotary kiln burning zone based on rs-lssvm, in *2017 36th Chinese Control Conference (CCC)*, (2017), 9643–9646. <https://doi.org/10.23919/chicc.2017.8028894>

23. T. Zheng, Q. Li, Soft measurement modeling based on temperature prediction of lssvm and arma rotary kiln burning zone, in *2019 IEEE 3rd Advanced Information Management, Communicates, Electronic and Automation Control Conference (IMCEC)*, (2019), 642–647. <https://doi.org/10.1109/imcec46724.2019.8983824>
24. J. Liu, L. Yang, X. Nan, Y. Liu, Q. Hou, K. Lan, et al., A soft sensing method of billet surface temperature based on ilgssa-lssvm, *Sci. Reports*, **12** (2022), 21876. <https://doi.org/10.1038/s41598-022-26478-3>
25. Z. J. Liu, J. Q. Wan, Y. W. Ma, Y. Wang, Online prediction of effluent cod in the anaerobic wastewater treatment system based on pca-lssvm algorithm, *Environ. Sci. Pollut. Res.*, **26** (2019), 12828–12841. <https://doi.org/10.1007/s11356-019-04671-8>
26. Y. Kuo, T. Yang, G. W. Huang, The use of grey relational analysis in solving multiple attribute decision-making problems, *Comput. Industr. Eng.*, **55** (2008), 80–93. <https://doi.org/10.1016/j.cie.2007.12.002>
27. N. Tosun, Determination of optimum parameters for multi-performance characteristics in drilling by using grey relational analysis, *Int. J. Adv. Manuf. Technol.*, **28** (2006), 450–455. <https://doi.org/10.1007/s00170-004-2386-y>
28. E. Özgür, E. C. Sabir, Ç. Sarpkaya, Multi-objective optimization of thermal and sound insulation properties of basalt and carbon fabric reinforced composites using the taguchi grey relations analysis, *J. Natural Fibers*, **20** (2023), 2178580. <https://doi.org/10.1080/15440478.2023.2178580>
29. R. W. Saaty, The analytic hierarchy process—what it is and how it is used, *Math. Model.*, **9** (1987), 161–176. [https://doi.org/10.1016/0270-0255\(87\)90473-8](https://doi.org/10.1016/0270-0255(87)90473-8)
30. Q. Jiang, X. Yan, Parallel pca-kpca for nonlinear process monitoring, *Control Eng. Pract.*, **80** (2018), 17–25. <https://doi.org/10.1016/j.conengprac.2018.07.012>
31. J. Liu, J. Wang, X. Liu, T. Ma, Z. Tang, MwrsPCA: online fault monitoring based on moving window recursive sparse principal component analysis, *J. Intell. Manuf.*, (2022), 1–17. <https://doi.org/10.1007/s10845-020-01721-8>
32. J. Suykens, Least squares support vector machines for classification and nonlinear modelling, *Neural Network World*, **10** (2000), 29–48.
33. J. A. Suykens, J. De Brabanter, L. Lukas, J. Vandewalle, Weighted least squares support vector machines: robustness and sparse approximation, *Neurocomputing*, **48** (2002), 85–105. [https://doi.org/10.1016/S0925-2312\(01\)00644-0](https://doi.org/10.1016/S0925-2312(01)00644-0)
34. C. F. Lin, S. D. Wang, Training algorithms for fuzzy support vector machines with noisy data, *Patt. Recogn. Lett.*, **25** (2004), 1647–1656. <https://doi.org/10.1016/j.patrec.2004.06.009>
35. D. Tsujinishi, S. Abe, Fuzzy least squares support vector machines for multiclass problems, *Neural Networks*, **16** (2003), 785–792. [https://doi.org/10.1016/S0893-6080\(03\)00110-2](https://doi.org/10.1016/S0893-6080(03)00110-2)
36. X. Q. Zeng, G. Z. Li, Incremental partial least squares analysis of big streaming data, *Patt. Recognit.*, **47** (2014), 3726–3735. <https://doi.org/10.1016/j.patcog.2014.05.022>
37. K. Bennett, M. Embrechts, An optimization perspective on kernel partial least squares regression, *Nato Sci. Series sub series III computer and systems sciences*, **190** (2003), 227–250.

38. J. Valyon, G. Horváth, A sparse least squares support vector machine classifier, in *2004 IEEE International Joint Conference on Neural Networks*, **1** (2004), 543–548. <https://doi.org/10.1109/IJCNN.2004.1379967>
39. D. R. Heisterkamp, J. Peng, H. K. Dai, Adaptive quasiconformal kernel metric for image retrieval, in *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, **2** (2001), 543–548. <https://doi.org/10.1109/CVPR.2001.990987>
40. G. Baudat, F. Anouar, Kernel-based methods and function approximation, in *IJCNN'01. International Joint Conference on Neural Networks*, **2** (2001), 1244–1249. <https://doi.org/10.1109/IJCNN.2001.939539>
41. R. Rosipal, L. J. Trejo, Kernel partial least squares regression in reproducing kernel hilbert space, *J. Mach. Learn. Res.*, **2** (2001), 97–123.
42. R. Sun, X. Qian, Soft sensor of concentration of sodium aluminate solution based on reduction robust lssvm, *J. Syst. Simul.*, **27** (2015), 2203.
43. M. C. Chen, D. M. Tsai, A simulated annealing approach for optimization of multi-pass turning operations, *Int. J. Product. Res.*, **34** (1996), 2803–2825. <https://doi.org/10.1080/00207549608905060>
44. N. Mughees, M. H. Jaffery, A. Mughees, E. A. Ansari, A. Mughees, Reinforcement learning-based composite differential evolution for integrated demand response scheme in industrial microgrids, *Appl. Energy*, **342** (2015), 121150. <https://doi.org/10.1016/j.apenergy.2023.121150>
45. H. Su, W. Qi, J. Chen, D. Zhang, Fuzzy approximation-based task-space control of robot manipulators with remote center of motion constraint, *IEEE Trans. Fuzzy Syst.*, **30** (2022), 1564–1573. <https://doi.org/10.1109/tfuzz.2022.3157075>
46. H. Su, W. Qi, Y. Hu, H. R. Karimi, G. Ferrigno, E. De Momi, An incremental learning framework for human-like redundancy optimization of anthropomorphic manipulators, *IEEE Trans. Industr. Inform.*, **18** (2020), 1864–1872. <https://doi.org/10.1109/TII.2020.3036693>



AIMS Press

©2023 the Author(s), licensee AIMS Press. This is an open access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>)