IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received May 7, 2019, accepted May 20, 2019, date of publication May 27, 2019, date of current version June 12, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2919110

Gold Recovery Modeling Based on Interval Prediction for a Gold Cyanidation Leaching Plant

ZHANG JUN[1], YAN HUA[1], YU HONGXIA[1], TIAN ZHONGDA[1], AND JIA RUNDA[2]
[1]College of Information Science and Engineering, Shenyang University of Technology, Shenyang 110870, China
[2]College of Information Science and Engineering, Northeastern University, Shenyang 110819, China

Corresponding author: Yan Hua (yanhua_01@163.com)

This work was supported in part by the National Natural Science Foundation of China under Grant 61803273, Grant 61873049, and Grant 61603172, and in part by the Natural Science Foundation Guidance Project in Liaoning of China under Grant 20180550970.

ABSTRACT The production index of gold cyanidation leaching process has an important influence on the economic benefits of the plant-wide hydrometallurgical process. In the actual leaching production process, due to the fluctuations of the previous procedure, the leaching process is often affected by the uncertainty and process disturbance. And hence, the prediction accuracy of the traditional point prediction models (such as ANN: Artificial Neural Network) decreases seriously and cannot provide any quantified information about the uncertainty or process disturbance. To solve the above problem, the interval prediction technique based on Radial Basis Function (RBF) ANN is proposed and used to model a gold cyanidation leaching plant encountering uncertainty and process disturbances in this paper. The objective function trained in the interval prediction model is not based on the prediction error, but the comprehensive measure index, namely, the coverage width criteria, which consists of the prediction interval coverage probability and the prediction interval normalized averaged width. Compared to the traditional point prediction, when the process uncertainty and disturbances are present, the interval prediction model will provide more helpful process information to the operators or process designers for determining further optimization and control strategies. The simulation and practical application results show that most of the real values of gold recovery can be covered between the upper and lower bounds with a predefined probability, which indicates the effectiveness and reliability of the interval prediction model, thus laying an important foundation for plant-wide optimization and control.

INDEX TERMS Data-driven model, gold cyanidation leaching process, hydrometallurgical process, interval prediction, point prediction, RBF ANN.

I. INTRODUCTION

Hydrometallurgy is generally effective at dealing with complex and low-grade ores and has been successfully applied to low-grade nonferrous metal metallurgy processes. For a typical gold ore processing plant, it often consists of the following fundamental unit operations: ore comminution, size classification, gravity concentration, pulp dewatering, gold leaching, gold recovery, gold elution, electrolytic extraction, and melting and casting [1]–[3]. At the present, although the hydrometallurgical technology is mature, advanced process optimization and control is hindered by the following complication. The hardwares for key parameter values are not reliable over time when running in the complex industrial conditions. And the maintenance cost is extremely high compared with the corresponding benefit resulted from using the advanced process optimization and control technique. To overcome the above shortcoming, the soft measurement technique can be employed to predict the key process parameter values instead of the hardwares [4]–[6]. The soft measurement technique is a method to estimate process parameter values hard to measure directly through the hardwares. And the common soft measurement methods are mathematical mechanism model, Neural Network, Support Vector Machine, Kernel Partial Least Square and so on.

As the most significant operation unit, leaching process has an important impact on the subsequent procedures. In the leaching process, the solid gold in ore is leached by the

The associate editor coordinating the review of this manuscript and approving it for publication was Mostafa Rahimi Azghadi.

leaching agent and dissolved into the liquid [7], [8]. The performance of the leaching process determines the gold recovery and hence has an significant impact on the overall yield and production efficiency. Perfect leaching performance can be achieved by implementing the advanced process optimization and control strategy, the premise of which is to establish an accurate soft measurement model for predicting the key process parameter values. Gold recovery is the most important performance indicator for the leaching process. Accurate prediction for gold recovery is conducive to determining the optimal set points of key control loops based on some production target. In the past two decades, researchers have investigated the gold cyanidation leaching process widely and some important results have been presented. In 1991, the kinetics of gold cyanidation leaching was explained in terms of charge transfer and diffusion model, and then a generalized equation for the leaching process was proposed [9]. As a useful concept to explain the observed results qualitatively, the mixed potential theory was applied to explain the kinetics of oxidative and reductive leaching processes [10]. Two fundamental kinetic equations based on the mixed potentials theory and the Habashi's cyanidation model were adopted and incorporated into the variable order kinetic equation. Both equations had good predictive performance for the gold extraction rate [11]. Afterwards, Crundwell and Godorr established an electrochemical mechanism model based on the shrinking-particle model [12]. A mechanism model was proposed by Khalid *et al.* [13], which was based on the idea that the surface reactions are thought to control the gold dissolution rate kinetics. In 2006, a mechanism model for gold cyanidation leaching process was established by De Andrade Lima and Hodouin [14], which is composed of the mass conservation equations (gold and cyanide) and the corresponding kinetic reaction expressions. The kinetic model parameters were identified based on the real industrial data from an Australian plant and the fitted model gave good predictive performance. The above established predictive models can be categorized as the pure mechanism models, which are obtained simply according to the process mechanism (physics, chemistry, and so on). Afterwards, the effects of cyanidation conditions on gold dissolution were studied by ANN modeling (pure data-driven modeling), in which the following six input parameters were considered, the leaching time, the solid percentage, the particle diameter, the NaCN content in cyanide media, the temperature of solution and the pH value [15]. Then a serial hybrid model combined with the mechanism and data-driven models was proposed, which made full use of the advantages of both the mechanism and data-driven models [16]. Compared with the mechanism model and the pure data-driven model, the hybrid model improved the predictive accuracy to some extent when the leaching process was in a stable operation.

In view of the complexity of the gold cyanidation leaching processes, some assumptions had to be made in order to simplify the mechanism model for a real application. And the model parameters especially in the kinetic reaction model are often unknown and difficult to be estimated accurately. Furthermore, inevitable process disturbances have a significant impact on the model predictive performance [17], [18]. The data-driven modeling techniques may be a suitable choice to overcome the above shortcomings. The data-driven model can be established completely only based on the process input and output data without process prior knowledge. Hence, the assumptions made for the mechanism models are not necessary using the data-driven modeling techniques [5], [19], [20]. Moreover, the data-driven modeling techniques ease the model solution procedures instead of solving the complex differential and algebraic equations, and hence decrease the computational load accordingly, which makes it more convenient for practical application. However, the traditional data-driven models may give rather worse prediction results when the disturbances or uncertainties are present in the gold cyanidation leaching process, which is due to the essential characteristics of the data-driven model. Moreover, the model cannot provide any useful information associated with the disturbances and uncertainties [21], [22].

Recently, the interval prediction technique has attracted more and more attention in practical process industries. The traditional model only provides a point predictive result, whereas the interval prediction model can give the interval including the plant output with a predefined probability (confidence level), which is defined with upper and lower bounds [21]–[24]. Compared to the point prediction, the interval prediction method will provide more helpful process information (such as the width of prediction interval, the upper and lower bounds) to the operators or process designers, which is significant for the consequent process optimization and control. For the gold cyanidation leaching plant investigated here, the real case is similar. In the real plant, to keep a stable leaching operation, the set points of the key control loops are not changed frequently, if the production index is in the allowable range, which is determined by engineering computations and operators' experience. Therefore, in view of practical applications, it is quite necessary to establish an interval prediction model of key process variable for gold cyanidation leaching process. In our study, an interval prediction model based on RBF ANN model is proposed in order to improve the model prediction accuracy when the process uncertainties and disturbances are present.

The rest of this paper is organized as follows: Section 2 introduces the gold cyanidation leaching process under investigation and gives the mechanism model in detail. And then a detail description of the interval prediction approach based on RBF ANN model is given in Section 3, which includes the traditional RBF ANN, the interval prediction and the corresponding developments especially for the practical application of gold cyanidation leaching. In Section 4, firstly, the mechanism model of Section 2 is used to simulate the real leaching plant and generate some necessary data sets used by the subsequent simulation tests. Secondly, the proposed interval prediction method is applied for predicting the upper and lower bounds of the gold recovery in a gold cyanidation

leaching plant. Finally, a detailed discussion of the above simulation test results is presented. Conclusions and some potential future research directions are presented in Section 5.

II. PROCESS DESCRIPTION

A. PROCESS DESCRIPTION

Gold cyanidation leaching has been one of the most important gold extraction methods for more than one century. A typical gold ore processing plant usually includes the following basic operation units: ore comminution, size classification, gravity concentration, pulp dewatering, gold leaching, gold recovery, gold elution, electrolytic extraction and melting and casting [7], [8], [14], [25]. In particular, gold cyanidation leaching process extracts the gold from the ore by cyanide solution, the performance of which has a significant impact on the production benefit. In the past few decades, the reaction mechanism and mathematical models of gold cyanidation leaching have been studied widely by researchers all over the world, most of whom suggest that pure gold cyanidation leaching is an electrochemical process [12], [13]. The solid gold in the ore is mostly oxidized and synthetized into a stable and soluble complex ion $[Au(CN)_2]^-$. Gold cyanidation leaching usually occurs in continuous stirred tank reactors (CSTRs) according to the following two chemical reactions [8], [14], [25], [26].

$$2Au+O_2+4CN^-+2H_2O\rightarrow 2[Au(CN)_2]^-+2OH^-+H_2O_2 \quad (1)$$

$$2Au+4CN^-+H_2O_2\rightarrow 2[Au(CN)_2]^-+2OH^- \quad (2)$$

The above two chemical equations can be readily summed for two sides and rewritten as the whole reaction equation:

$$4Au+8CN^-+O_2+2H_2O\rightarrow 4[Au(CN)_2]^-+4OH^- \quad (3)$$

In practice, to increase the final gold recovery and avoid the gold loss as much as possible, usually several leaching tanks are employed together in series. In each tank, the same reaction occurs as described above. A photograph of the gold cyanidation leaching plant in a gold refinery is shown in Fig.1. The block flow diagram of the cyanidation gold extraction technique in a gold refinery and the corresponding simplified plant flowsheet are given in Fig.2 and Fig.3, respectively. For the cyanidation gold extraction technique studied here, it consists of the following unit operations, namely, flotation, washing and conditioning, gold cyanidation leaching, two-stage washing and gold recovered by zinc. In advance of the leaching operation, the gold ores after the flotation operation are firstly pumped into the thickener and filter press, in which the remaining flotation reagents can be mostly removed. Then the gold bearing filter mass is mixed with the recycled waste water containing CN^- and conditioned to appropriate pulp concentration. Finally, the ore pulp is continuously pumped into the buffer tank with constant flowrate and prepared for the subsequent leaching operation. The ore pulp in the buffer tank overflows into the four pneumatic leaching tanks in series, respectively. In each tank,

FIGURE 1. Photograph of the gold cyanidation leaching plant in a gold refinery.

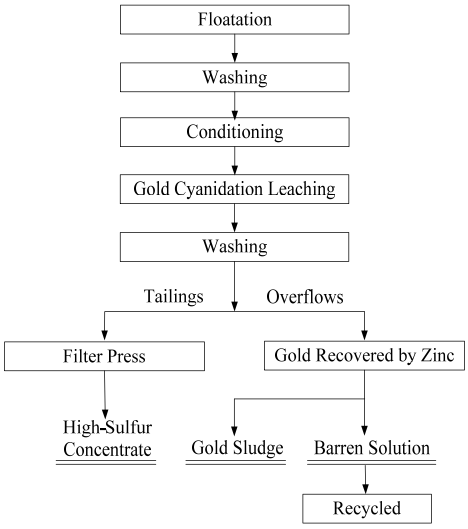


FIGURE 2. Block flow diagram of the cyanidation gold extraction technique in a gold refinery.

the same chemical reactions described as Eqs. (1) and (2) or Eq. (3) occur, namely, the gold particle in solid reacts with the cyanide ion and dissolved oxygen in liquid and synthetizes to a stable and soluble complex ion $[Au(CN)_2]^-$ in order to separate from other impurities. Herein, the CN^- is provided in two ways, that is, the waste water used for conditioning pulp and the newly added NaCN solution with 30% concentration in each leaching tank. Whereas the necessary dissolved oxygen is provided by compressed air that is pumped into the bottom of each tank, which also prevents pulp from heaping up at the bottom of tanks and makes the reactions more thorough by the moderate pneumatic stirring effect generated by the compressed air. After the leaching operation, the gold is extracted into the liquid and then the ore pulp overflows into the two-stage thickeners. After washing, the pregnant solution is stored in the tank of pregnant solution and gradually pumped into the subsequent process of gold recovery by zinc.

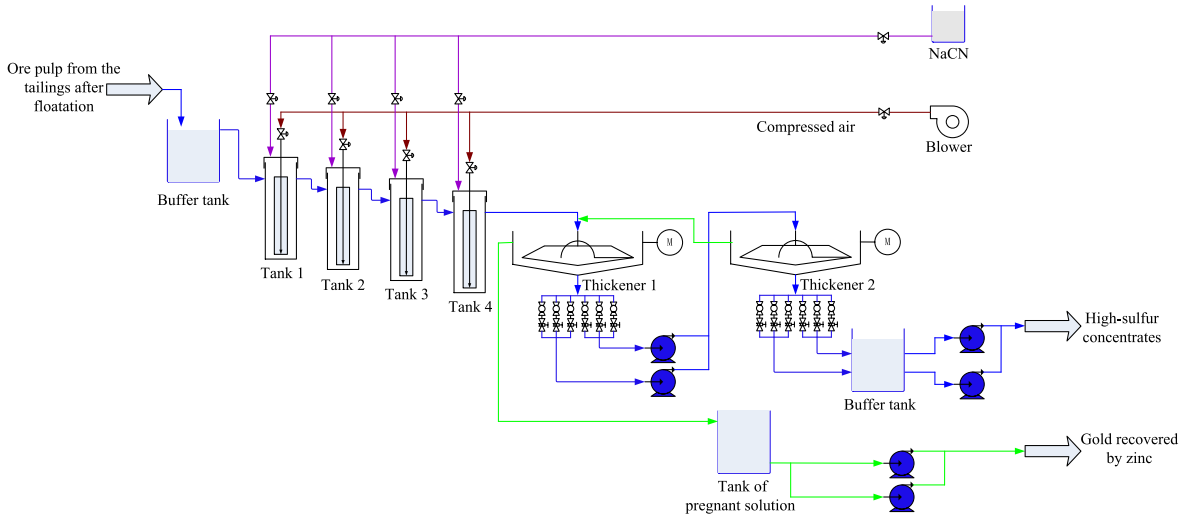


FIGURE 3. Simplified plant flowsheet of the cyanidation gold extraction plant in a gold refinery.

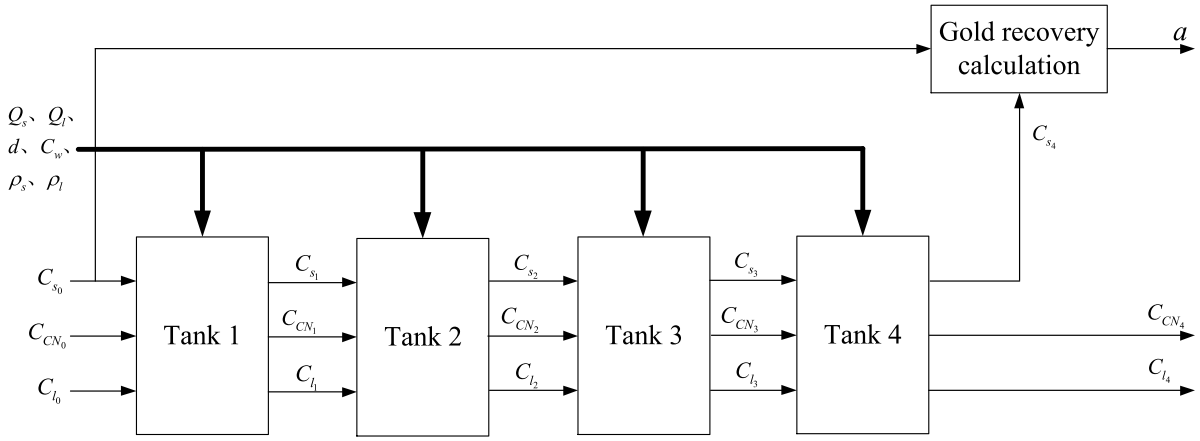


FIGURE 4. Schematic diagram of the mechanism model for gold cyanidation leaching process.

In gold cyanidation leaching, gold recovery is the most important performance indicator, which is influenced by the initial operation conditions, manipulated variables, process uncertainties and disturbances, and so on [14], [16]–[18], [25]. To keep the leaching process operating at the optimal state determined by the plant-wide optimization and control level, the optimal additive amount of NaCN into each tank need to be determined by operation optimization based on process model.

B. MECHANISM MODELING

In our study, the gold particles are assumed to be mono-sized and the cyanidation leaching tanks under investigation are regarded as ideal CSTRs. To simulate the existing process uncertainties and disturbances in a real leaching plant, in this paper, the established mathematical model developed by De Andrade Lima and Hodouin is used to simulate the reality of the gold cyanidation leaching process, in which the kinetic model parameters of the gold and cyanide have been identified and calibrated based on a set of real industrial data from an Australian gold processing plant [14], [25]–[27]. For each leaching tank, the steady-state mechanism model for simulating the leaching process is composed of the mass conservation equations of key components, namely, gold in the solid, gold in the liquid and cyanide in the liquid, as well as the corresponding kinetic reaction rate models, as shown in Eqs. (4)-(8), respectively [14], [25]–[28]. The schematic diagram of the mechanism model for gold cyanidation leaching is illustrated in Fig.4.

$$\frac{Q_{s_i}}{M_{s_i}}(C_{s_{i-1}} - C_{s_i}) - r_{Au_i} = 0 \tag{4}$$

$$\frac{Q_{l_i}}{M_{l_i}}(C_{l_{i-1}} - C_{l_i}) + \frac{M_{s_i}}{M_{l_i}} r_{Au_i} = 0 \tag{5}$$

$$\frac{Q_{l_i}}{M_{l_i}}(C_{CN_{i-1}} - C_{CN_i}) + \frac{Q_{CN_i}}{M_{l_i}} - r_{CN_i} = 0 \tag{6}$$

$$r_{Au_i} = (1.13 \times 10^{-3} - 4.37 \times 10^{-11}\bar{d}^{2.93}) \times (C_{s_i} - C_{s_\infty}(\bar{d}))^{2.13} \times C_{CN_i}^{0.961} \times C_{o_i}^{0.228} \tag{7}$$

$$r_{CN_i} = (\frac{1.69 \times 10^{-8}}{\bar{d}^{0.547} - 6.40}) \times C_{CN_i}^{3.71} \tag{8}$$

where the subscript " i " means the mark number of the leaching tanks, and in this paper, a serial four-tank gold cyanidation leaching process is investigated ($i = 1, 2, 3, 4$). It is noted that the subscript " 0 " means the initial variable values before entering the leaching process. $Q_{s_i}, M_{s_i}, C_{s_i}$ are the solid flow rate, holdup in the tanks and gold grade, respectively. $Q_{l_i}, M_{l_i}, C_{l_i}$ are the liquid flow rate, holdup in the tanks and gold concentration, respectively. Q_{CN_i}, C_{CN_i} are the added cyanide flow rate and the liquid cyanide ion concentration, respectively. Cs_∞ represents the ideal residual gold grade in the ore, which increases with the increase of the average particle diameter $\bar{d}$ of gold ore:

$$Cs_\infty(\bar{d}) = 0.357[1 - 1.49\exp(-1.76 \times 10^{-2}\bar{d})] \quad (9)$$

The kinetic reaction rates r_{Au_i} and r_{CN_i} represent the gold dissolution rate and the cyanide consumption rate, respectively, which are associated with the initial operation conditions (such as the dissolved oxygen concentration C_{o_i} in the liquid and $\bar{d}$) and the current state variables, C_{s_i} and C_{CN_i} [14], [28]. In a real leaching plant, to obtain as great a production profit as possible, the leaching process runs at steady state for most of the time, which is assured by the fact that the feed ore pulp is pumped into the leaching tank continuously and steadily and that the connection between the successive leaching tanks is in the overflow mode. Therefore, the material mass in the solid and liquid is conserved for each tank, namely, $Q_{s_{i-1}} = Q_{s_i}$ and $Q_{l_{i-1}} = Q_{l_i}$. In addition, it is assumed that all the reactants in the leaching tanks are perfectly mixed and hence the material segregation can be negligible. By introducing the parameter C_{w_i} as the weight concentration of solid in the pulp, the liquid flow rate can be calculated by $Q_{l_i} = Q_{s_i}(1/C_{w_i} - 1)$. Then the average residence time τ_i for the solid or the liquid staying in the tank can be determined according to the following formula [14], [16]–[18], [25]:

$$\tau_i = V_i / \left(Q_{s_i}/\rho_s + Q_{l_i}/\rho_l\right) \times 1000 \quad (10)$$

where V_i is the net active volume of tank i, ρ_s and ρ_l are the corresponding density for the solid and liquid, respectively. Afterwards, the constant holdup M_{s_i} and M_{l_i} for the solid and liquid in the tank can be determined as follows [14], [25]:

$$M_{s_i} = Q_{s_i} \times \tau_i \quad (11)$$

$$M_{l_i} = Q_{l_i} \times \tau_i \quad (12)$$

In conclusion, the mechanism model for simulating the gold cyanidation leaching process is composed of the mass conservation equations (4)-(6) of gold and cyanide, the corresponding kinetic reaction rate models given by Eqs. (7)-(8) as well as the calculation formulas of key variables given by Eqs. (9)-(12). In this paper, the above mechanism model is solved using an iterative numerical method based on the Matlab software [29], [30]. After the mechanism model is solved, we can obtain the corresponding variable values of the gold grade in the solid, and the gold and cyanide ion concentration in the liquid for each tank. Finally, the whole gold recovery after four-tank leaching can be calculated by:

$$a = \frac{C_{s_0} - C_{s_4}}{C_{s_0}} \times 100\% \quad (13)$$

Though the gold grade in the ore C_s is easy to measure by some methods, such as titration or atomic spectrometer, the above measure method is time-consuming due to the fact that the sampled ore pulp should be firstly kept in nature precipitation and filtered for a period of time in order to be apart from the liquid before measuring. Therefore, the current process operation information cannot be obtained in real time and hence cannot be used to decide the subsequent operation properly. From this point of view, the gold recovery (or gold grade in the ore) is hard to measure in real time and it is of significance for further optimization and control to establish the soft-sensor model of gold recovery.

III. INTERVAL PREDICTION MODELING BASED ON RBF ANN

As mentioned above, due to the complexity of the reaction kinetics, some necessary assumptions have to be made to simplify the mechanism model. The accurate kinetic model parameters are usually unknown. Moreover, the accurate kinetic model outputs (kinetic reaction rates) are unmeasurable or difficult to be obtained accurately. Therefore, the mechanism model does not match with the real plant completely and hence the corresponding model prediction result is more or less inaccurate. To provide more helpful process information to the operators or process designers, an interval prediction model based on RBF ANN model is developed in this paper, which aims to improve the model prediction accuracy and provides some useful guidance for process optimization and control when process uncertainties and disturbances are present.

A. RBF ANN MODELING

It is well known that ANN, such as RBF ANN, can discover the complex relationship between two data sets and approximate any complicated nonlinear function when the number of the neurons in the hidden layers are sufficient [31], [32], which is a local approximation network and hence has a fast learning convergence rate. In particular, the RBF ANN has good generalization ability and has been successfully applied to the areas of nonlinear function approximation, time series analysis, data classification, pattern recognition, information processing, image processing, system modeling, control, fault diagnosis and so on [31]–[35]. So far, there have been many successful real cases of modeling practical industrial processes using RBF ANNs, please refer to [36]–[41]. The schematic diagram of the RBF ANN model structure is shown in Fig.5. The RBF ANN is composed of the input layer, the hidden layer and the output layer. The basic idea behind the RBF ANN is that the RBF is used as the basis function of the hidden layer to transform the low-dimensional input vector into the high-dimensional hidden layer space without any weighted connections. According to the Cover

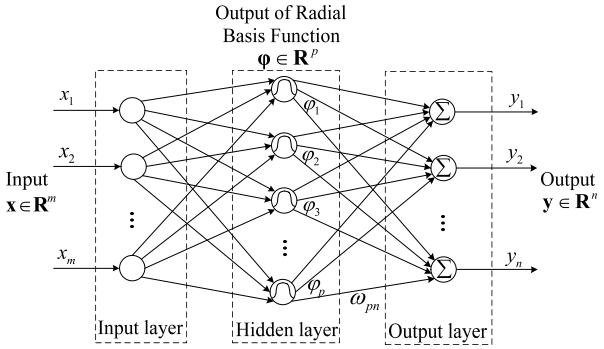


FIGURE 5. Schematic diagram of the RBF ANN model structure.

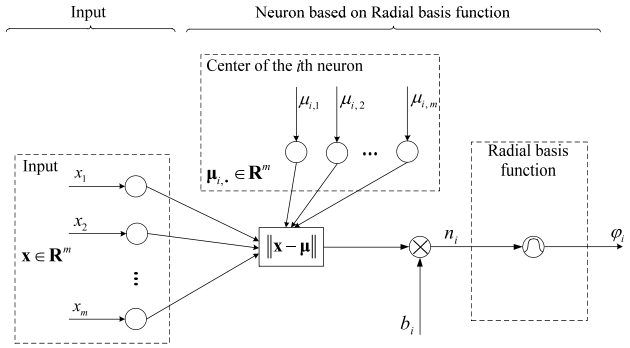


FIGURE 6. Schematic diagram of the model structure for the *i*th neuron (radial basis function).

theorem [42], the probability that classes are linearly separable increases when the features are nonlinearly mapped to a higher dimensional feature space. In other words, the essence of the RBF ANN is firstly to map the input data in the low-dimensional space into a high-dimensional space using a nonlinear RBF and then find a surface that can best fit the training data in the implicit high-dimensional space, which is different from the ordinary multi-layer perceptron (MLP) (such as Back Propagation (BP) ANN). For RBF ANN, the relationship between the input layer and the hidden layer is nonlinear, while the output layer is linear with regard to the hidden layer, which is helpful to improve the learning rate of RBF ANN and avoid the problem of local minimum [37], [41].

The schematic diagram of the model structure for the *i*th neuron (radial basis function) is shown in Fig.6. Each neuron in the hidden layer of RBF ANN acts as a radial basis function, which is a real-valued function whose value depends only on the distance from the origin or alternatively on the distance from some other point $\boldsymbol{\mu}_{i,\cdot}$, called a center.

The commonly used types of radial basis functions include Gaussian, Multiquadric, Polyharmonic Spline, Inverse Quadratic, Inverse Multiquadric functions and so on [32], [34], [43]. In view of the advantages of the Gaussian kernel function (such as: simple form, radially symmetric, good smoothness and analyticity), here the following Gaussian kernel function is selected:

$$\varphi_i = e^{-n_i^2} = e^{-(\|\mathbf{x}-\boldsymbol{\mu}_{i,\cdot}\|b_i)^2} = e^{-\frac{\|\mathbf{x}-\boldsymbol{\mu}_{i,\cdot}\|^2}{2\sigma_i^2}} \tag{14}$$

where n_i and φ_i is the input and output of the *i*th RBF, respectively. $\mathbf{x} = [x_1, x_2, \cdots, x_m]^T \in \mathbf{R}^m$ is the input vector, $\boldsymbol{\mu}_{i,\cdot} = [\mu_{i,1}, \mu_{i,2}, \cdots, \mu_{i,m}]^T \in \mathbf{R}^m (i = 1, 2, 3, \cdots, p)$ and $b_i = 1/\sqrt{2}\sigma_i = 0.8326/C_{spread}$ is the center vector and the threshold value (sensitiveness) of the *i*th RBF neuron in the hidden layer, respectively. $\|\bullet\|$ is the Euclidean 2-norm operator. C_{spread} is the expansion constant of RBF. σ_i is the width of the *i*th RBF and determines the width of the radial basis function surrounding the corresponding center. In practice, the methods to determine the width parameter σ_i are the sample method and the self-organizing selection methods [34], [37]. For example, the data center $\boldsymbol{\mu}_{i,\cdot}$ is selected from the sample. More samples should be collected for the densely areas. Each basis function uses a uniform width parameter $\sigma_i = d_{\max}/\sqrt{2M}$. It can be easily seen that the RBF is a radially symmetric and monotonically decreasing function and when $\mathbf{x}$ and $\boldsymbol{\mu}_{i,\cdot}$ are equal, the RBF output reaches its maximum 1. When the distance between $\mathbf{x}$ and $\boldsymbol{\mu}_{i,\cdot}$ is large, the RBF output is near 0. Whereas the RBF output is near 1, when $\mathbf{x}$ is close to $\boldsymbol{\mu}_{i,\cdot}$. That is to say, only a small number of neurons close to $\mathbf{x}$ can be activated [34], [44], [45].

Finally, the RBF ANN outputs are the linear combination of the above outputs of p RBF neurons, which is calculated as follows:

$$\mathbf{y} = \boldsymbol{\omega}^T \boldsymbol{\varphi} + \mathbf{c} \tag{15}$$

where $\mathbf{y} = [y_1, y_2, \cdots, y_n]^T \in \mathbf{R}^n$ is the output vector, the output vector of the hidden layer is $\boldsymbol{\varphi} = [\varphi_1, \varphi_2, \cdots, \varphi_p]^T \in \mathbf{R}^p$, $\boldsymbol{\omega}^T = [\omega_{ij}]_{i=1,2,\cdots,n;j=1,2,\cdots,p} \in \mathbf{R}^{n\times p}$ and $\mathbf{c} = [c_1, c_2, \cdots, c_n]^T \in \mathbf{R}^n$ are the output weight matrix and bias vector of the RBF ANN, respectively. To some extent, the introduction of Gaussian kernel function is helpful to improve the problem of curse of dimensionality for RBF ANN due to the fact that the output layer is linear with regard to the hidden layer in the high dimensionality. From the above, the training of the RBF ANN usually consists of determining the centers and widths of RBF, and the weight matrix and the bias vector from the hidden layer to the output layer. Many RBF ANN learning algorithms have been proposed and common approaches include back-propagation and simulated annealing [46]–[49]. The common objective function of training RBF ANN model is defined as follows:

$$\begin{aligned} E &= \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{n} e_{ij}^2 = \frac{1}{2}\sum_{i=1}^{N} \mathbf{e}_{i,\cdot}\mathbf{e}_{i,\cdot}^T \\ &= \frac{1}{2}\sum_{i=1}^{N} \left\|\mathbf{y}_i^* - \mathbf{y}_i\right\| \\ &= \frac{1}{2}\sum_{i=1}^{N} \left\|\mathbf{y}_i^* - (\boldsymbol{\omega}^T\boldsymbol{\varphi}_i + \mathbf{c})\right\| \end{aligned} \tag{16}$$

where N is the number of training data, the predictive error matrix is $\mathbf{e}^T = [e_{ij}]_{i=1,2,\cdots,N;j=1,2,\cdots,n} \in \mathbf{R}^{N\times n}$, $\mathbf{y}_i^*$ and $\mathbf{y}_i$ are the real and predictive output vector for the *i*th sample, respectively. $\boldsymbol{\varphi}_i$ is the RBF output vector corresponding to the

*i*th sample. For the RBF ANN model investigated here, when the number of the neurons in the hidden layer is excessively large or the net structure is extremely complex, maybe overfitting will occur. An overfitting model has poor predictive performance, as it overreacts to minor fluctuations in the training data [32], [34], [43].

B. INTERVAL PREDICTION MODELING BASED ON RBF ANN

Due to the fact that the traditional data-driven models (such as RBF ANNs) may give rather worse prediction results when disturbances or uncertainties are present and moreover the model cannot provide any useful information associated with the disturbances and uncertainties, there has been considerable motivation to extend RBF ANN. Here, the interval prediction technique proposed by Hosen and Khosravi [21], [22], [50]–[52] is used to model the gold cyanidation leaching process with the purpose of providing more helpful process state information to the operators or process designers compared to the traditional point prediction methods.

The traditional RBF ANN modeling method is trained based on some index (such as the sum of squared prediction errors shown in Eq.(16)) for the given data set and then is used to calculate the corresponding outputs for the new input data. What the interval prediction modeling method based RBF ANN provides is not the certain output values but its upper and lower bounds, between which the real outputs lie with a specified probability (Confidence Level, CL, $(1-\alpha)$) [21], [22]. In the past few decades, some methods to construct the prediction interval of ANN have been proposed, such as Bootstrap [53], Bayesian [54], Delta [55] and Mean-variance estimation techniques [56]. In this paper, the lower upper bound estimation (LUBE) method proposed by Khosravi *et al.* [21], [22] is employed to construct the prediction interval of RBF ANN, which is due to its low computational load, high computational efficiency and no extra data distribution assumption.

The main difference between the traditional RBF ANN modeling method and the interval prediction method based on RBF ANN is that the objective function used for training the latter is not the simple sum of squared prediction errors but a comprehensive measure index particular for the interval prediction, called the Coverage Width Criteria (*CWC*), which consists of two performance aspects, namely the Prediction Interval Coverage Probability (*PICP*) and the Prediction Interval Normalized Averaged Width (*PINAW*) [21], [22].

The *PICP* index is used to measure the total number of the output values lying between the upper and lower bounds of the prediction interval and can be easily calculated as follows:

$$PICP = \frac{1}{N}\sum_{i=1}^{N} c_i \tag{17}$$

where c_i is the indicator that describes if the ith output value lies between the upper and lower bounds of the prediction interval or not:

$$c_i = \begin{cases} 1, & y_i \in \left[\underline{y}_i, \bar{y}_i\right] \\ 0, & y_i \notin \left[\underline{y}_i, \bar{y}_i\right] \end{cases} \tag{18}$$

and y_i, $\underline{y}_i$ and $\bar{y}_i$ are the output value, its upper and lower bounds for the ith sample, respectively. It can be easily observed from Eq.(17) that a larger *PICP* index can be obtained by increasing the width of the prediction interval. However, a too wide prediction interval cannot provide any helpful process information to operators and is of little use. Therefore, the *PICP* index is not sufficient to measure the prediction accuracy of the interval prediction method.

The other performance index is *PINAW*, which is used to measure the normalized averaged width of the prediction interval and can be calculated as:

$$PINAW = \frac{1}{y_r}\left(\frac{1}{N}\sum_{i=1}^{N}\left(\bar{y}_i - \underline{y}_i\right)\right) \tag{19}$$

where y_r is the output range. A smaller *PINAW* means a narrower averaged width of the prediction interval and hence results in a lower *PICP*, namely a lower coverage probability of the prediction interval. Therefore, the above two measure indexes oppose each other and cannot reach optimal values simultaneously. The comprehensive measure index (*CWC*) considering *PICP* and *PINAW* together is given as:

$$CWC = PINAW \cdot (1 + \gamma(PICP) \cdot e^{-\eta\cdot(PICP-\phi)}) \tag{20}$$

where

$$\gamma(PICP) = \begin{cases} 0, & PICP \geq \phi \\ 1, & PICP < \phi \end{cases} \tag{21}$$

which means that when *PICP* is equal to or greater than the predefined coverage probability ϕ, only the *PINAW* index is considered. ϕ is usually equal to the nominal confidence level and can be set to $1-\alpha$ simply. η is a parameter used to change the penalty degree and smaller difference between *PICP* and ϕ will be enhanced due to the multiplication effect of η. The above comprehensive measure index (*CWC*) has two objectives, namely maximizing *PICP* ($\phi \leq PICP \leq 1$) and minimizing *PINAW* ($PINAW > 0$). Specially, the above interval prediction model can be also used to predict the actual real data value by some treatment. One possible way is to transform the upper and lower bounds into the actual predictive data value by computing the average of the upper and lower bounds.

IV. SIMULATION RESULTS AND PRACTICAL APPLICATION

In this section, the proposed interval prediction strategy based on RBF ANN is used to predict the interval of gold recovery for a gold cyanidation leaching plant. To simulate the general situations of process disturbance and uncertainty existing in practice, the mechanism model given in Section 2 is used to

simulate the real gold cyanidation leaching process and generate the necessary simulation data. In this case, the process disturbance or uncertainty can be easily realized by changing the corresponding operating conditions or model parameters in the mechanism model. For the leaching plant investigated here, the on-line measuring instruments can display and store the values of the pulp flow rate, the pulp concentration and the oxygen concentration in liquid in real time. However, in consideration of the reliability and the related maintenance cost, the gold grade and cyanide ion concentrations are sampled and then assayed by experimental titration every two hours. Unless specially indicated, all of the following simulation tests are performed by the MATLAB software package on an Intel i5 3.1GHz PC with 4G RAM.

A. MECHANISM MODEL SIMULATION AND ANALYSES

The Mechanism model given in Section 2 is obtained by analyzing the reaction mechanism of the leaching process and then derived from rigorous mathematical calculations. Moreover, the kinetic model parameters of the gold and cyanide have been identified based on a set of real industrial data from an Australian gold processing plant and the calibrated process model has better prediction performance under the condition without process disturbance or uncertainties. To demonstrate the agreement between the mechanism model and the plant investigated here, the above mechanism model is simulated systematically in order to determine the relationships between the model output (gold recovery) and the relevant variables, which determines the auxiliary variables of the RBF ANN data-driven model. The number of tanks is selected as 4, which is the same as the real leaching plant investigated here. To analyze the effect of a certain variable on the gold recovery, all the relevant simulations are carried out as follows: under the same operation conditions and inputs, namely that the other variables are kept the same as the nominal values, however, the investigated variable is changed continuously, the gold recovery values at the steady state are calculated and then the correlationship between them is obtained easily. The relevant variable and parameter values (nominal values) used to simulate the reality with the above mechanism model are listed in Table 1, which are obtained mainly by referring to the relevant literature [25]–[28] and the production reports in the real leaching plant. For simplicity, the added cyanide flow rates in each leaching tanks are set as the same values. The relationship curves between the gold recovery and its influencing factors, namely the solid flow rate, the weight concentration of solid in the pulp, the added cyanide flow rate, the dissolved oxygen concentration, the average particle diameter, the initial gold grade and cyanide ion concentration and the number of the leaching tanks n are shown in Fig.7 (a)—(h), respectively.

The following simulation results are all in accordance with the real phenomena occurring in the investigated leaching plant, which illustrates that the proposed mechanism model is an adequate description of the real gold cyanidation leaching process and hence is suitable for simulating the behavior of

TABLE 1. The relevant variable and parameter values used to simulate the reality with the mechanism model.

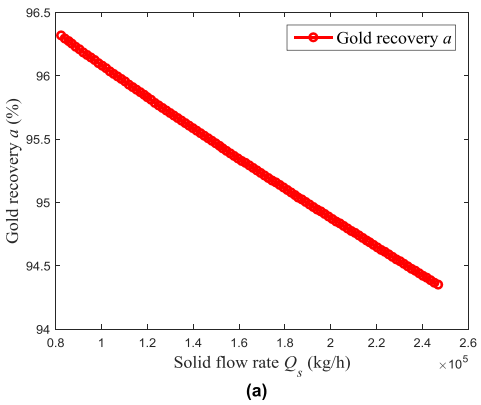
Symbol	Quantity	Value
Q_{s_i}	ore flow rate into the ith tank (kg/h)	1.645×10^5
Q_{CN_i} $(i=1,2,3,4)$	cyanide flow rate added into the ith tank (mg/h)	2×10^8
C_{s_0}	initial gold concentration in the ore (mg/kg)	10
C_{CN_0}	initial cyanide ion concentration in the liquid (mg/kg)	300
$\bar{d}$	average size of the ore particles (μm)	80
C_w	solid concentration in the pulp (kg/kg)	48.8%
V	net volume of the reactor (m^3)	412
ρ_s	ore density (g/cm^3)	2.8
ρ_l	liquid density (g/cm^3)	1.0
C_o	oxygen concentration in the liquid (mg/kg)	8

the real leaching process. For example, with the increase of the solid flow rate, the gold recovery decreases gradually, which is mainly due to the fact that the average residence time of the pulp holding in the leaching tanks is shorten and then the effective average leaching time decreases. Nevertheless, when the weight concentration of solid in the pulp increases, the content of solid ore in the pulp increase, which increases the contact area between gold and cyanide, causing an acceleration in the reaction and a sharp increase in gold recovery. However, with the continuous increase of the pulp concentration, the rate increase of gold recovery is hindered by slow diffusion rates due to higher pulp viscosity [2], [7].

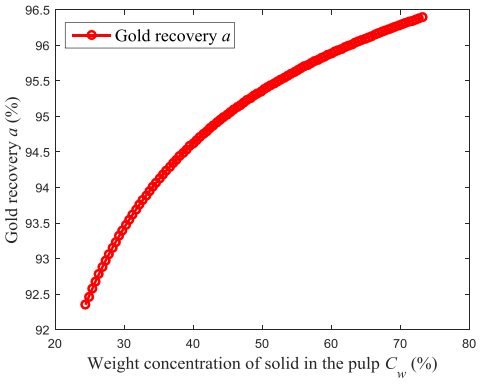
B. DATA-DRIVEN MODELING AND PREDICTION WITH TRADITIONAL RBF ANN

In this subsection, the traditional RBF ANN is used to capture the complicated nonlinear characteristic of the gold cyanidation leaching process. The proposed mechanism model in Section 2 is used to simulate the reality of the leaching process and generate the plant data necessary for the RBF ANN model training and testing. The data set are generated with and without process disturbances and uncertainties, respectively, which can be easily realized by changing the process operating conditions or parameters in the mechanism model. Moreover, for the sake of accordance with the reality, the Gaussian noise with zero mean and 5% standard deviation are also added to the output concentration measurements.

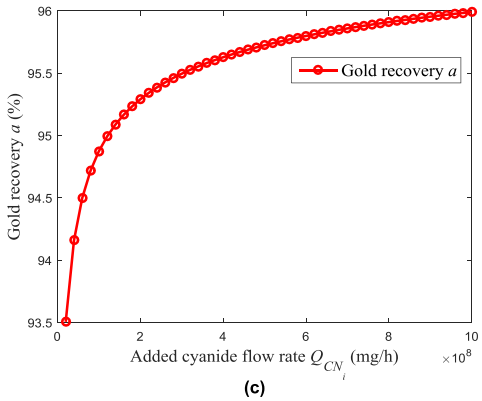
In the previous subsection, the individual effect of each operating variable on the gold recovery has been studied in detail by simulation, which has laid an important foundation for constructing the RBF ANN model structure and selecting the input variables. The structural diagram of the gold recovery prediction model based on the traditional RBF ANN is shown in Fig.8. Here, the variables Q_s, C_w, C_{s_0}, C_{CN_0}, Q_{CN_i}, and C_o are selected as the inputs of the data-driven RBF



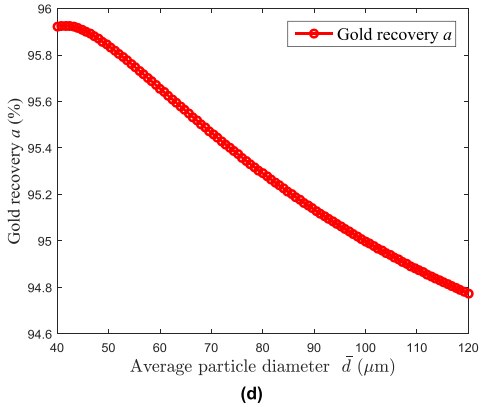
(a)



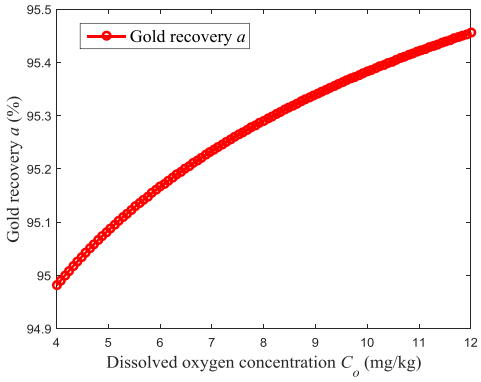
(b)



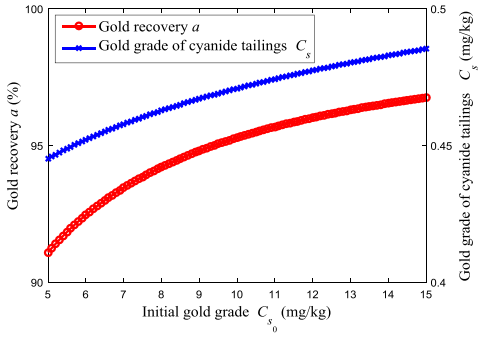
(c)



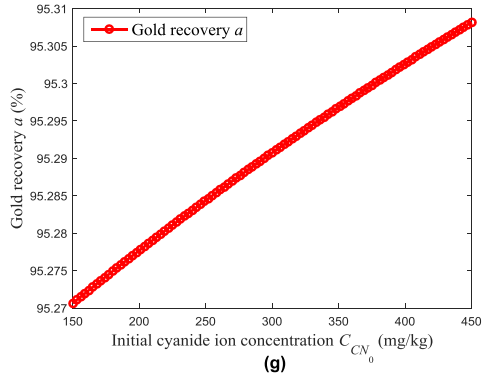
(d)



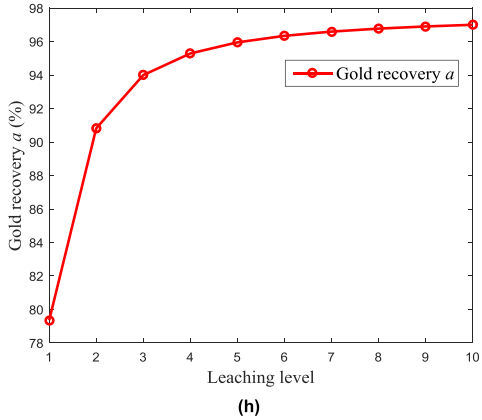
(e)



(f)



(g)



(h)

FIGURE 7. The relationship curves between the gold recovery and its influencing factors. (a) Q_S. (b) C_W. (c) Q_{CN_i}. (d) $\bar{d}$. (e) C_o. (f) C_{S_0}. (g) C_{CN_0}. (h) n.

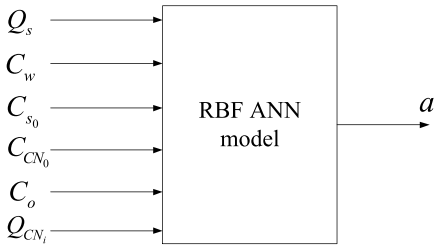


FIGURE 8. The structural diagram of the gold recovery prediction model based on the traditional RBF ANN.

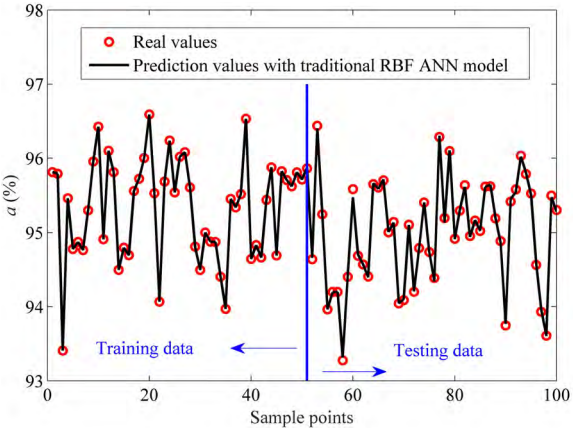


FIGURE 9. The prediction result with the traditional RBF ANN model under no process disturbance and uncertainty.

TABLE 2. The error analyses of the gold recovery predicted with the traditional RBF ANN model under no process disturbance and uncertainty.

Scenarios	Indices	Training	Testing	All
Outputs with no noise	RMSE	0.0003	0.0265	0.0133
	MAE	0.0001	0.0189	0.0049
	MAPE	1.31×10^{-6}	0.0002	5.14×10^{-5}

ANN model, whose variable ranges are ±20% of the nominal values, respectively. In total, 1000 samples (Data #1∼#1000) are generated with the above mechanism model, of which the first 75% are used to train the RBF ANN model and the rest are used to test the prediction performance of the trained model. Here, the expansion constant C_{spread} of RBF is selected as 1 and the neuron number N_n in the hidden layer is equal to the number of training data, namely 750. The prediction result with the traditional RBF ANN model under no process disturbance and uncertainty is shown in Fig.9, in which the measurement noise is considered to be a kind of uncertainty and is not taken into account in this scenario. Only parts of the prediction results are shown in Fig.9 for convenience, namely Data #1∼#50 in training data set and Data #751∼#800 in the testing data set.

And the corresponding prediction error analyses of the gold recovery predicted with the traditional RBF ANN model is listed in Table 2. Here, the following three performance indices are used to evaluate the model prediction performance, namely RMSE (Root mean square error), MAE (Mean absolute error) and MAPE (Mean absolute percentage error):

$$e_{\mathrm{RMSE}} = \sqrt{\frac{\sum_{i=1}^{N}(a_{real,i} - a_{pre,i})^2}{N}} \tag{22}$$

$$e_{\mathrm{MAE}} = \frac{\sum_{i=1}^{N}\left|a_{real,i} - a_{pre,i}\right|}{N} \tag{23}$$

$$e_{\mathrm{MAPE}} = \frac{1}{N}\sum_{i=1}^{N}\left|\frac{a_{real,i} - a_{pre,i}}{a_{real,i}}\right| \tag{24}$$

where $a_{real,i}$ is the ith measurement values (real values) of the gold recovery a and $a_{pre,i}$ is the corresponding prediction values with the prediction model. N is the number of data used to calculate the performance index.

It can be easily observed that when the process disturbance and uncertainty are not present, the predictive values with the traditional RBF ANN model for both the training and testing data are completely identical with the real values from the naked eyes, which is mainly due to the strong ability of the RBF ANN model to capture the complicated nonlinear relationship between the gold recovery and its influencing factors. Moreover, unknown process disturbance or uncertainty is not present and hence the training data includes sufficient process information. The similar results can be also found from the prediction error analyses in Table 2. The prediction errors (RMSE, MAE and MAPE) for training, testing and all data are all quite small and can be completely accepted from the point of engineering applications, which means that the RBF ANN model captures the nonlinear relationship between the gold recovery and its influencing factors when the process disturbance and uncertainty are all not present. However, in industrial applications, the above ideal situation never exists. Therefore, it is significant to investigate the corresponding model prediction performance for the process with disturbance and uncertainty.

Meanwhile, the effect of the number of neurons N_n in the hidden layer on the model prediction performance is also investigated and the corresponding prediction results as well as the prediction error analyses are shown in Fig.10 and Table 3, respectively. Only the 50 prediction results are shown in Fig.10 for convenience. When the process disturbance and uncertainty are both not present, the obvious difference of the model prediction performance resulted from changing the neuron number in the hidden layer of RBF ANN model is quite small from Fig.10 (a) and the detail information can be observed from Fig.10 (b) and the error analyses listed in Table 3. With the increase of N_n, the three performance indices for training data all decrease simultaneously, which is due to the fact that the increase of N_n complicates the RBF ANN model structure and hence enhances its ability to capture the complex nonlinear relationship between the input and output data. However, the corresponding performance indicators for testing data don't decrease all the time, but decrease

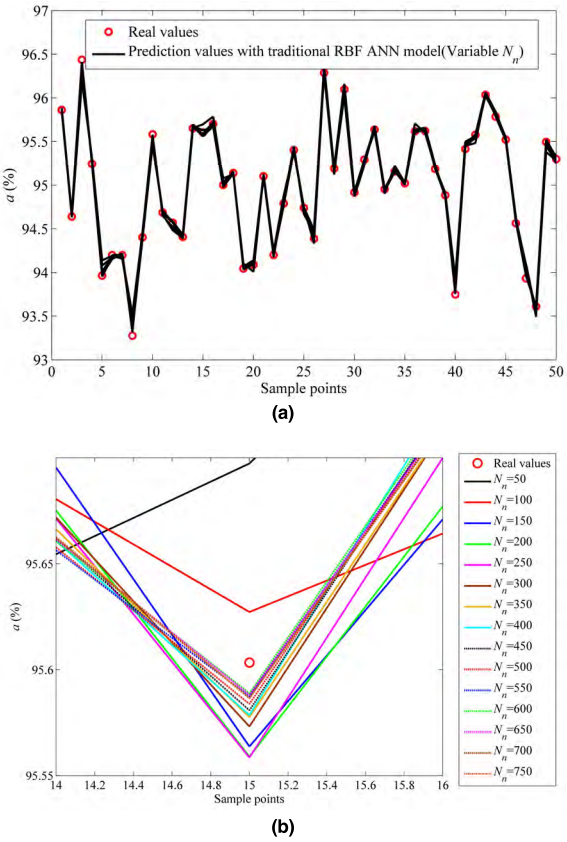


FIGURE 10. The effect of the neuron number in the hidden layer on the model prediction performance under no process disturbance and uncertainty. (a) Global graph. (b) Local graph.

to a certain value and then increase instead on the whole, which can be explained by the overfitting phenomenon of the training data. Here in the above simulations, an appropriate neuron number in the hidden layer of RBF ANN model is selected as 550. When the number of neurons in the hidden layer of RBF ANN model is larger, the degree of fitting is higher, but the time spent is correspondingly increased. In practical applications, the influence of the above factors should be considered comprehensively to choose a suitable neuron number.

It can be easily seen from the above simulations and analysis that the traditional RBF ANN model is a good tool to predict the gold recovery when the leaching process doesn't suffer from disturbance and uncertainty. However, in practical leaching plants, especially the continuous leaching plant investigated in this paper, strong disturbance and uncertainty indeed exist during the leaching operation process due to the unstable or faulty operation of the preceding grinding-flotation and thickening processes, and even the uncertain nature of the ore. Moreover, most of the disturbances are difficult or expensive to measure accurately. Therefore, simulations have been conducted which include process disturbance and uncertainty usually existing in the real industrial applications. Here in this paper, the process uncertainty is simply simulated by introducing the Gaussian measurement

TABLE 3. The effect of the neuron number in the hidden layer on the model prediction performance under no process disturbance and uncertainty.

N_n	Indices	Training	Testing	All
50	RMSE	0.0613	0.0930	0.0847
	MAE	0.0349	0.0715	0.0644
	MAPE	0.0004	0.0008	0.0007
100	RMSE	0.0330	0.0612	0.0488
	MAE	0.0190	0.0451	0.0366
	MAPE	0.0002	0.0005	0.0004
150	RMSE	0.0191	0.0402	0.0298
	MAE	0.0114	0.0307	0.0229
	MAPE	0.0001	0.0003	0.0002
200	RMSE	0.0129	0.0322	0.0219
	MAE	0.0078	0.0241	0.0164
	MAPE	8.15×10^{-5}	0.0003	0.0002
250	RMSE	0.0094	0.0282	0.0178
	MAE	0.0055	0.0202	0.0124
	MAPE	5.79×10^{-5}	0.0002	0.0001
300	RMSE	0.0073	0.0275	0.0161
	MAE	0.0043	0.0201	0.0107
	MAPE	4.48×10^{-5}	0.0002	0.0001
350	RMSE	0.0056	0.0255	0.0143
	MAE	0.0033	0.0182	0.0089
	MAPE	3.46×10^{-5}	0.0002	9.40×10^{-5}
400	RMSE	0.0044	0.0257	0.0138
	MAE	0.0025	0.0185	0.0080
	MAPE	2.67×10^{-5}	0.0002	8.42×10^{-5}
450	RMSE	0.0033	0.0252	0.0131
	MAE	0.0019	0.0181	0.0070
	MAPE	1.96×10^{-5}	0.0002	7.39×10^{-5}
500	RMSE	0.0025	0.0250	0.0128
	MAE	0.0014	0.0181	0.0064
	MAPE	1.49×10^{-5}	0.0002	6.75×10^{-5}
550	RMSE	0.0019	0.0244	0.0124
	MAE	0.0010	0.0175	0.0057
	MAPE	1.07×10^{-5}	0.0002	6.03×10^{-5}
600	RMSE	0.0013	0.0244	0.0123
	MAE	0.0007	0.0177	0.0053
	MAPE	6.97×10^{-6}	0.0002	5.59×10^{-5}
650	RMSE	0.0007	0.0250	0.0125
	MAE	0.0004	0.0177	0.0049
	MAPE	3.90×10^{-6}	0.0002	5.18×10^{-5}
700	RMSE	0.0004	0.0260	0.0130
	MAE	0.0002	0.0186	0.0049
	MAPE	1.83×10^{-6}	0.0002	5.14×10^{-5}
750	RMSE	0.0003	0.0265	0.0133
	MAE	0.0001	0.0189	0.0049
	MAPE	1.31×10^{-6}	0.0002	5.14×10^{-5}

noise into the corresponding concentration measurements. Moreover, the process disturbance is realized by changing the corresponding operating conditions (such as: $\bar{d}$, ρ_s and ρ_l) in the mechanism model used to simulate the realistic leaching process. The corresponding prediction results with the traditional RBF ANN model for the above situations are shown in Fig.11. The corresponding error analyses for the testing data are listed in Table 4. Due to the fact that the uncertainty and process disturbances are either unavailable or difficult and expensive to measure accurately in real plants, they are not included into the process input. And hence the traditional trained RBF ANN model not considering the uncertainty and disturbances gives poor prediction performance compared with the case of no uncertainty and disturbance. To be

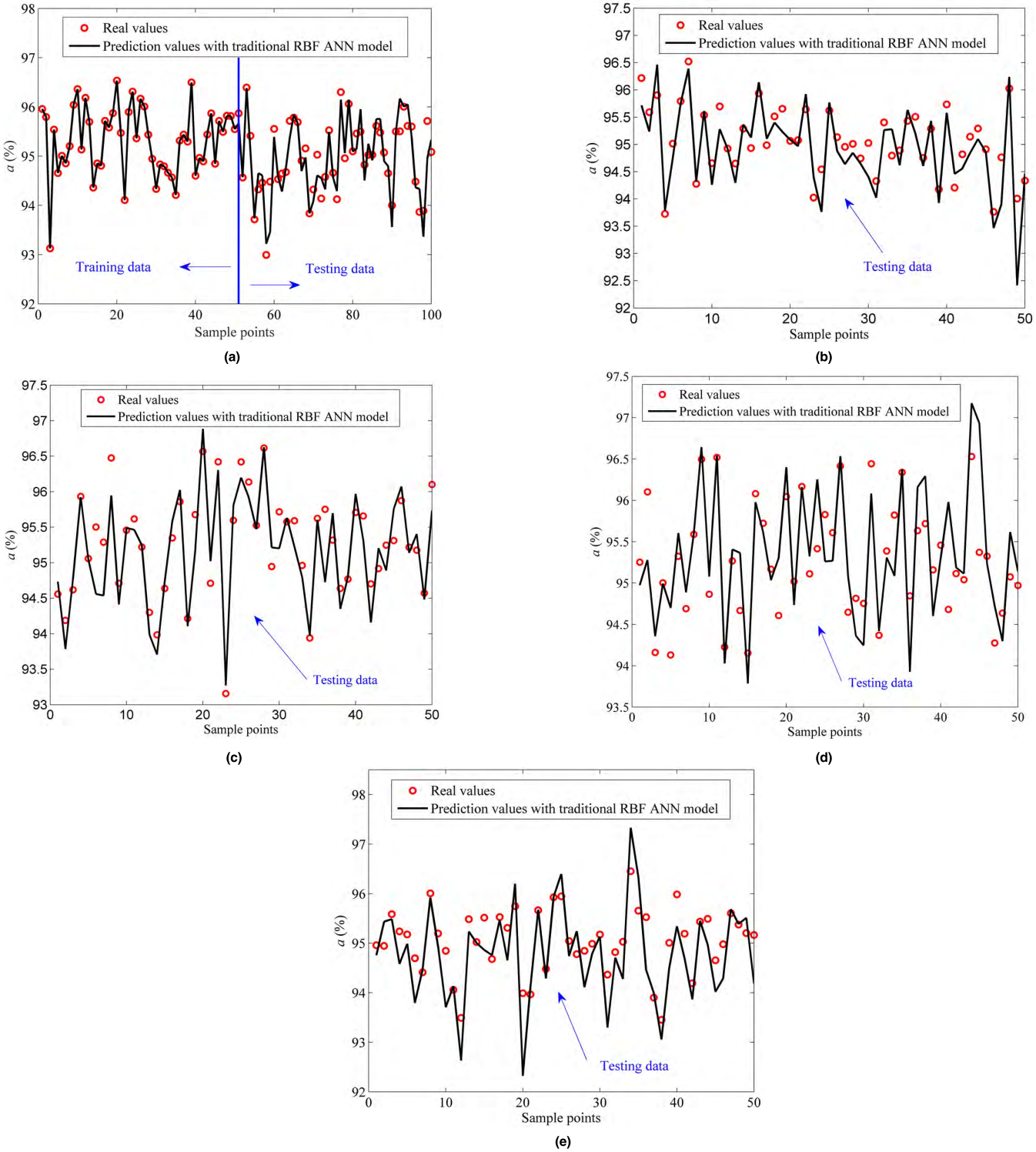


FIGURE 11. The prediction results with the traditional RBF ANN model under process disturbance and uncertainty. (a) Outputs with 5% noise and no process disturbance. (b) Outputs with 5% noise and disturbance $\bar{d}$ (c) Outputs with 5% noise and disturbance ρ_S (d) Outputs with 5% noise and disturbance ρ_I (e) Outputs with 5% noise and disturbances $\bar{d}$, ρ_S and ρ_I.

specific, for the case of 5% measurement noise, the RMSE and MAE values are approximately 11.4 and 12.3 times as much as those for the no noise case. Moreover, the model prediction performance is much worse when the unknown process disturbances are present. Due to the fact that the process uncertainty and disturbance are unmeasurable and cannot be included into the input variables, the traditional RBF ANN model trained based on the sum of squared

TABLE 4. The error analyses of the gold recovery predicted with the traditional RBF ANN model under process disturbance and uncertainty.

Scenarios	Indices	Testing
Outputs with 5% noise	RMSE	0.3026
	MAE	0.2334
	MAPE	0.0025
Outputs with 5% noise and disturbance $\overline{d}$	RMSE	0.5162
	MAE	0.3981
	MAPE	0.0042
Outputs with 5% noise and disturbance ρ_s	RMSE	0.4126
	MAE	0.3220
	MAPE	0.0034
Outputs with 5% noise and disturbance ρ_l	RMSE	0.3906
	MAE	0.3049
	MAPE	0.0032
Outputs with 5% noise and disturbance $\overline{d}$, ρ_s and ρ_l	RMSE	0.5112
	MAE	0.3972
	MAPE	0.0042

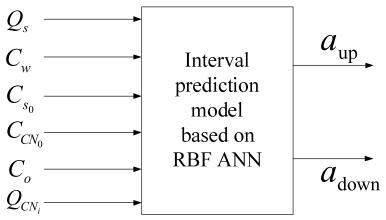


FIGURE 12. The structural diagram of the interval prediction model of the gold recovery based on RBF ANN.

errors cannot give accurate and acceptable prediction results. In conclusion, the traditional RBF ANN model is not suitable for predicting the gold recovery when the leaching process suffers from the uncertainty and process disturbances. The poor prediction results cannot provide more helpful process information to the operators or process designers and hence are not helpful for process optimization and control.

C. INTERVAL PREDICTION BASED ON RBF ANN

To deal with the effect of uncertainty and process disturbances, the interval prediction technique proposed in Section 3 is used to model the gold cyanidation leaching process, thereby improving the model prediction performance beyond that of traditional point prediction methods. The inputs of the interval prediction model based on RBF ANN are the same as those of the traditional RBF ANN model proposed in the last subsection, whereas the output is not simply the gold recovery but its upper and lower bounds. The structural diagram of the interval prediction model of the gold recovery based on RBF ANN is shown in Fig.12.

The comprehensive measure index *CWC* given in Eq. (20) is used as the performance index to optimize the parameters of the RBF ANN model, which is nonlinear, complex and non-differentiable. Due to the different objective function from the traditional RBF ANN model, the traditional RBF ANN training method is not appropriate for the interval prediction model. Therefore, in this paper, the SA (Simulated Annealing) numerical optimization algorithm is used to optimize the parameters of the interval prediction model based on RBF ANN [57], [58], namely used to train the RBF ANN, specifically, minimize the above performance index *CWC*. Here the corresponding parameters of the traditional RBF ANN model are used as the initial values of the SA optimization algorithm. And the confidence level (ϕ) is selected as 90%, which means the real gold recovery lies between the upper and lower bounds with the 90% probability. η is equal to 5 in order to enhance the penalty degree of the smaller difference between *PICP* and ϕ. η should be determined in comprehensive consideration of the two indexes: *PICP* and *PINAW*. If the constraint of $PICP \geq \phi$ cannot be violated, a larger η should be selected. However, too large η will lead to that the index of *PINAW* is too larger, namely the upper and lower bounds are too wide. Here in this paper, η is selected as a constant 5 by considering the order of magnitudes of the two parts in the objective function Eq. (20). Moreover, a heuristic approach to determine η is to regard η as a decision variable and then obtain it by training the interval prediction model based on RBF ANN simultaneously.

The prediction results with the interval prediction model based on RBF ANN are shown in Fig.13. And the corresponding performance indicator analyses of the gold recovery predicted with the interval prediction model based on RBF ANN are listed in Table 5. As above, for better visualization, here only the corresponding prediction results for the Data #751∼#800 in the testing data set are given in Fig.13. Here, the proposed interval prediction model is not trained based on the prediction error usually used in the traditional RBF ANN model, but the comprehensive measure index particular for the interval prediction, namely *CWC* consisting of *PICP* and *PINAW*. Moreover, the training data is enough and the initial point of the optimization objective function *CWC* is uniformly selected by the corresponding parameter values of the traditional RBF ANN model. The comprehensive measure index is effective to avoid the overfitting phenomenon in the traditional RBF ANN model, all of which can ensure the reproducibility of the simulation results to some extent. And the simulation results showed that the difference among successive 10 simulation runs is slight and the simulation results are possible to reproduce. Without loss of generality, only the simulation results for the first run are given here. It can be easily observed from Fig.13 (a) that the upper and lower bounds are very close to each other and most of the real values lie between the upper and lower bounds. And hence when the uncertainty and process disturbances are not present, the interval prediction method can provide similar prediction results as the traditional RBF ANN, if we use the average value of the upper and lower bounds as the real prediction value. Moreover, for the cases of measurement noise and process disturbances, the interval prediction method also shows excellent performance compared to the traditional RBF ANN model. It can be seen from Fig.13 (b)-(f) that most of the real values lie between the upper and lower bounds with the predefined confidence level, and the violation of ϕ less than 90% in Table 5 can be attributed to the presence of the measurement noise. Moreover, the widths of the prediction

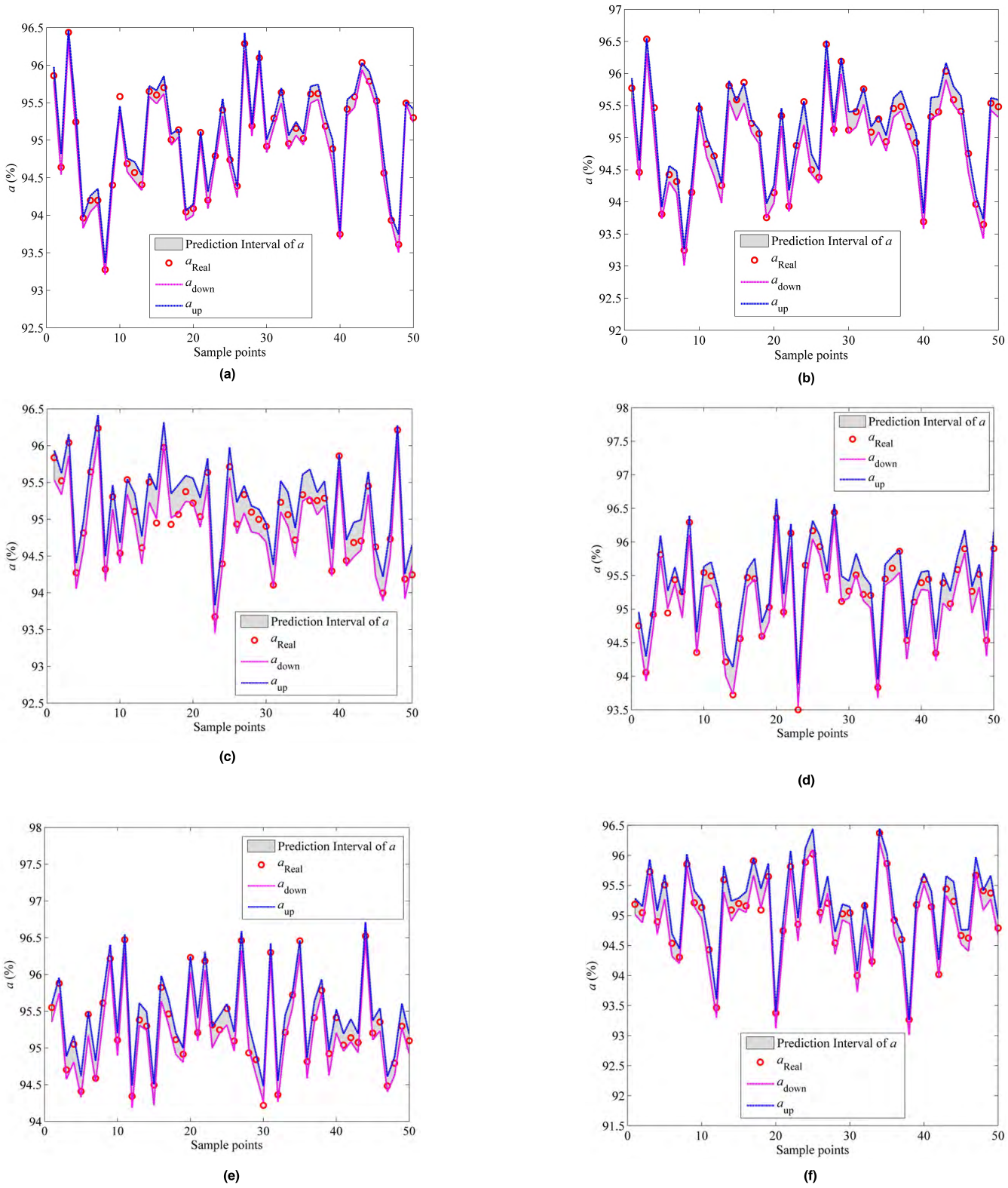


FIGURE 13. The prediction results with the interval prediction model based on RBF ANN. (a) Outputs with no noise and process disturbance. (b) Outputs with 5% noise and no process disturbance. (c) Outputs with 5% noise and disturbance $\bar{d}$. (d) Outputs with 5% noise and disturbance ρ_S. (e) Outputs with 5% noise and disturbance ρ_I (f) Outputs with 5% noise and disturbances $\bar{d}$, ρ_S and ρ_I.

interval, as helpful information, can be provided to the operators or process designers to estimate the degrees of uncertainty and disturbances and then take necessary measures to optimize the process performance. However, the traditional RBF ANN model only gives an inaccurate prediction result and does not provide any important process information about

TABLE 5. The performance index analyses of the gold recovery predicted with the interval prediction model based on RBF ANN.

Scenarios	Indices	Testing
Outputs with no noise	PICP	90.53%
	PINAW	2.65%
	CWC	0.026466
Outputs with 5% noise	PICP	89.73%
	PINAW	7.79%
	CWC	0.156764
Outputs with 5% noise and disturbance $\overline{d}$	PICP	90.4%
	PINAW	8.36%
	CWC	0.083616
Outputs with 5% noise and disturbance ρ_s	PICP	89.87%
	PINAW	8%
	CWC	0.160564
Outputs with 5% noise and disturbance ρ_l	PICP	89.73%
	PINAW	7.44%
	CWC	0.149779
Outputs with 5% noise and disturbance $\overline{d}$, ρ_s and ρ_l	PICP	90.13%
	PINAW	7.61%
	CWC	0.076098

TABLE 6. The effect of the confidence level (φ) on the prediction performance of the interval prediction model based on RBF ANN.

ϕ	Indices	Testing
85%	PICP	86%
	PINAW	1.92%
	CWC	0.019239
90%	PICP	90.53%
	PINAW	2.65%
	CWC	0.026466
95%	PICP	95.2%
	PINAW	4.41%
	CWC	0.044089
98%	PICP	98%
	PINAW	5.04%
	CWC	0.050446

the uncertainty and disturbances. In general, the results for all the scenarios are good (such as *PICP*) and only slight difference is present, which can be improved by changing the corresponding parameters in the objective function in fact.

Furthermore, the effect of the confidence level (ϕ) on the prediction performance of the interval prediction model based on RBF ANN is listed in Table 6. With the increase of ϕ, more real values need to be covered between the upper and lower bounds and hence the width of the prediction interval extends in order to cover more real values.

D. INTERVAL PREDICTION BASED ON RBF ANN FOR A GOLD CYANIDATION LEACHING PLANT

The effectiveness of the interval prediction method proposed in this paper has been validated by simulation tests in the previous subsection. In this subsection, the above interval prediction method is used to predict the upper and lower bounds of the gold recovery for a gold cyanidation leaching plant in China. To increase the final gold recovery, two-level leaching processes are designed in this plant. Without loss of generality, only the first-level leaching process is considered

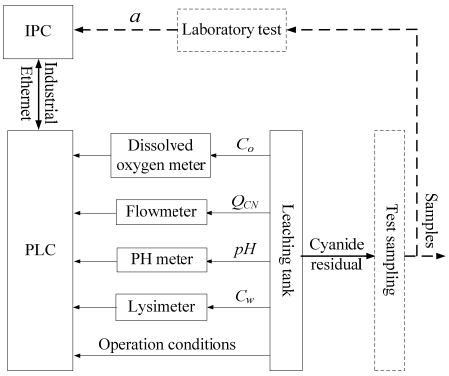


FIGURE 14. The schematic diagram of the gold recovery prediction system in a gold cyanidation leaching plant.

in this paper, which is more influenced by the preceding processes compared to the second-level leaching process. For example, the effect of the unsteady production or changed production index in the preceding process procedure on the gold cyanidation leaching process. The schematic diagram of the gold recovery prediction system in the gold cyanidation leaching plant is shown in Fig.14. The whole system consists of the leaching rate prediction system, PLC (Programmable Logic Controller), IPC (Industrial Personal Computer), the off-line laboratory test and the online measuring instruments. In all, 1198 samples between April 21, 2016 and July 31, 2016 for the case encountering process uncertainty and disturbance were continuously collected every two hours. Due to the precision and reliability of measuring instruments or the non-ideal industrial conditions, a data pretreatment procedure is needed to eliminate the inevitable abnormal data existing in the training and testing data set. Here, the notable 3σ criteria is used to eliminate the abnormal data among the data set, which is conducted as follows: for an individual sample datum y_i in a data set with n samples, y_i is considered to be an abnormal datum and should be replaced by a proper value, such as the average value of its previous and later data, if the following condition $|\varepsilon_i| = |y_i - \bar{Y}| > 3\sigma$ holds, where $\bar{Y} = \sum_{i=1}^{n} y_i/n$ and $\sigma = \sqrt{\sum_{i=1}^{n} (y_i - \bar{Y})^2/(n-1)}$ are the mean value and standard deviation of the data set, respectively.

Similarly as above, the pretreated samples are divided into two data sets (the training and testing data sets), the first 75% samples of which are used to train the interval prediction model based on the RBF ANN, whereas the rest are used to test the prediction performance of the proposed interval prediction model. Parts of the prediction results with the interval prediction model based on RBF ANN are shown in Fig.15, which is captured from the screen of the gold recovery prediction system in the real industrial plant. It can be easily observed from Fig.15 that the two bounds can cover most of the real values of gold recovery with a higher probability and moreover, the interval prediction method can track the variation trend of the gold recovery when the leaching process is subjected to process uncertainty or disturbances.

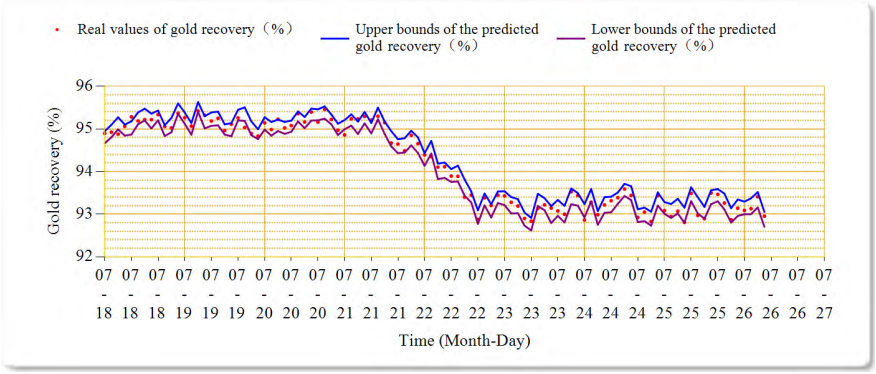


FIGURE 15. Parts of the prediction results with the interval prediction model based on RBF ANN in the real industrial plant.

In summary, the proposed interval prediction model is a good tool to predict the gold recovery of gold cyanidation leaching under process uncertainty and disturbances.

V. CONCLUSION

It has been accepted widely by researchers that ANNs (such as RBF ANN) are ideal modeling tools for most nonlinear industrial production processes, such as chemical process. However, their perfect nonlinear approximation capability is obtained ideally under the premise of good process data without uncertainty and process disturbances, which is impossible in real industrial applications, particularly for complicated hydrometallurgical processes. Under real situations with uncertainty and disturbances, the traditional point prediction method based on RBF ANN fails to provide accurate and reliable prediction results and hence not any helpful process information can be used by the process operators or designers. In this paper, the interval prediction technique based on RBF ANN has been investigated in detail and then considered as a good alternative modeling method for a gold cyanidation leaching plant encountering uncertainty and process disturbances. The interval prediction method can give the interval including the plant output with a predefined probability (confidence level), which is defined with the upper and lower bounds, respectively. Compared to the traditional point prediction, when the process uncertainty and disturbances are present, the interval prediction method will provide more helpful process information (such as the width of prediction interval, the upper and lower bounds) to the operators or process designers, which is significant for the process optimization and control. The simulation and the practical application results show that most of the real values of gold recovery can be covered between the upper and lower bounds with a predefined probability, which indicates the effectiveness and reliability of the interval prediction methods for leaching process with uncertainty and disturbances, thus laying an important foundation for the plant-wide optimization and control.

To decrease the training difficulty of the interval prediction model based on RBF ANN, in the future the focus may be on developing an alternative interval prediction model using other effective data-driven modeling methods, such as LSSVM (Least Square Support Vector Machine), KPLS (Kernel Partial Least Square) and so on. The aim is to make full use of the perfect advantages of the above modeling methods. For example, the LSSVM method is a good tool to avoid the curse of dimensionality and the KPLS method can deal with the multicollinearity well. However, if the traditional *CWC* performance index is used, the above and even other perfect characteristics for LSSVM or KPLS methods cannot be guaranteed. Therefore, our future work is to develop a novel comprehensive performance index for interval prediction model based on LSSVM or KPLS methods, which is meanwhile in agreement with the common index for the above two methods.

REFERENCES

[1] J. Y. Zhou and L. J. Cabri, "Gold process mineralogy: Objectives, techniques, and applications," *J. Minerals, Metals Mater. Soc.*, vol. 56, no. 7, pp. 49–52, 2004.
[2] J. Sun, *Gold and Silver Metallurgy*. Beijing, China: Metallurgical Industry Press, 2008, pp. 16–41.
[3] L. Michael, *Hydrometallurgy Fundamentals and Applications*. Hoboken, NJ, USA: Wiley, 2011, pp. 156–179.
[4] S. Gupta, P.-H. Liu, S. A. Svoronos, R. Sharma, N. A. Abdel-Khalek, Y. Cheng, and H. El-Shall, "Hybrid first-principles/neural networks model for column flotation," *AICHE J.*, vol. 45, no. 3, pp. 557–566, 1999.
[5] H. C. Aguiar and R. M. Filho, "Neural network and hybrid model: A discussion about different modeling techniques to predict pulping degree with industrial data," *Chem. Eng. Sci.*, vol. 56, no. 2, pp. 565–570, 2001.
[6] D. S. Lee, P. A. Vanrolleghem, and J. M. Park, "Parallel hybrid modeling methods for a full-scale cokes wastewater treatment plant," *J. Biotechnol.*, vol. 115, no. 3, pp. 317–328, 2005.
[7] Editorial Committee of the Gold Production Guide, *The Gold Production Guide*. Beijing, China: Geological Press, 2000, pp. 28–35.
[8] S. R. La Brooy, H. G. Linge, and G. S. Walker, "Review of gold extraction from ores," *Minerals Eng.*, vol. 7, no. 10, pp. 1213–1241, 1994.
[9] M. E. Wadsworth, "Rate processes in the leaching of gold and other metals forming stable complexes," in *Proc. HH Kellogg Int. Symp.*, Warrendale, PA, USA, 1991, pp. 197–216.
[10] J. Li, T.-K. Zhong, and M. E. Wadsworth, "Application of mixed potential theory in hydrometallurgy," *Hydrometallurgy*, vol. 29, nos. 1–3, pp. 47–60, 1992.
[11] D. H. Rubisov, V. G. Papangelakis, and P. D. Kondos, "Fundamental kinetic models for gold ore cyanide leaching," *Can. Metall. Quart.*, vol. 35, no. 4, pp. 353–361, 1996.

[12] F. K. Crundwell and S. A. Godorr, "A mathematical model of the leaching of gold in cyanide solutions," *Hydrometallurgy*, vol. 44, nos. 1–2, pp. 147–162, 1997.

[13] M. Khalid, F. Larachi, and A. Adnot, "Cyanidation of gold associated with silver minerals in sulfide mineral matrices," *Chem. Eng. Technol.*, vol. 41, no. 7, pp. 1282–1293, Jul. 2018.

[14] L. R. P. de Andrade Lima and D. Hodouin, "Simulation study of the optimal distribution of cyanide in a gold leaching circuit," *Minerals Eng.*, vol. 19, no. 13, pp. 1319–1327, 2006.

[15] S. Khoshjavan, M. Mazloumi, and B. Rezai, "Artificial neural network modeling of gold dissolution in cyanide media," *J. Central South Univ. Technol.*, vol. 18, no. 6, pp. 1976–1984, 2011.

[16] J. Zhang, Z.-Z. Mao, R.-D. Jia, and D.-K. He, "Serial hybrid modelling for a gold cyanidation leaching plant," *Can. J. Chem. Eng.*, vol. 93, no. 9, pp. 1624–1634, 2015.

[17] J. Zhang, Z.-Z. Mao, R.-da Jia, and D.-K. He, "Real time optimization based on a serial hybrid model for gold cyanidation leaching process," *Minerals Eng.*, vol. 70, pp. 250–263, Jan. 2015.

[18] J. Zhang, Z.-Z. Mao, and R.-D. Jia, "Real-time optimization based on SCFO for gold cyanidation leaching process," *Chem. Eng. Sci.*, vol. 134, no. 9, pp. 467–476, 2015.

[19] K. Mitra and M. Ghivari, "Modeling of an industrial wet grinding operation using data-driven techniques," *Comput. Chem. Eng.*, vol. 30, no. 3, pp. 508–520, 2006.

[20] K. Y. Rani and S. C. Patwardhan, "Data-driven model based control of a multi-product semi-batch polymerization reactor," *Chem. Eng. Res. Des.*, vol. 85, no. 10, pp. 1397–1406, 2007.

[21] A. Khosravi, S. Nahavandi, D. Creighton, and A. F. Atiya, "Lower upper bound estimation method for construction of neural network-based prediction intervals," *IEEE Trans. Neural Netw.*, vol. 22, no. 3, pp. 337–346, Mar. 2011.

[22] A. Khosravi, S. Nahavandi, D. Creighton, and A. F. Atiya, "Comprehensive review of neural network-based prediction intervals and new advances," *IEEE Trans. Neural Netw.*, vol. 22, no. 9, pp. 1341–1356, Sep. 2011.

[23] M. A. Hosen, A. Khosravi, D. Creighton, and S. Nahavandi, "Prediction interval-based modelling of polymerization reactor: A new modelling strategy for chemical reactors," *J. Taiwan Inst. Chem. Eng.*, vol. 45, no. 5, pp. 2246–2257, 2014.

[24] M. A. Hosen, A. Khosravi, S. Nahavandi, and D. Creighton, "Prediction interval-based neural network modelling of polystyrene polymerization reactor—A new perspective of data-based modelling," *Chem. Eng. Res. Des.*, vol. 92, no. 11, pp. 2041–2051, 2014.

[25] L. R. P. de Andrade Lima and D. Hodouin, "Optimization of reactor volumes for gold cyanidation," *Minerals Eng.*, vol.18, no. 7, pp. 671–679, 2005.

[26] L. R. P. de Andrade Lima, "Some remarks on the reactor network synthesis for gold cyanidation," *Minerals Eng.*, vol. 19, no. 2, pp. 154–161, 2006.

[27] L. R. P. De Andrade Lima, "Modeling, control and optimization applied to the gold hydrometallurgy," Ph.D. dissertation, Dept. Mines Metallurgy, Laval Univ., Québec, QC, Canada, 2001.

[28] L. R. P. de Andrade Lima and D. Hodouin, "A lumped kinetic model for gold ore cyanidation," *Hydrometallurgy*, vol. 79, no. 3, pp. 121–137, Oct. 2005.

[29] T. F. Coleman and Y. Li, "An interior trust region approach for nonlinear minimization subject to bounds," *SIAM J. Optim.*, vol. 6, no. 2, pp. 418–445, 1996.

[30] *MATLAB: The Language of Technical Computing. Desktop Tools and Development Environment*, MathWorks Inc, Natick, MA, USA, 2005, pp. 1–261.

[31] V. T. S. Elanayar and Y. C. Shin, "Radial basis function neural network for approximation and estimation of nonlinear stochastic dynamic systems," *IEEE Trans. Neural Netw.*, vol. 5, no. 4, pp. 594–603, Jul. 1994.

[32] H. Leung, T. Lo, and S. Wang, "Prediction of noisy chaotic time series using an optimal radial basis function neural network," *IEEE Trans. Neural Netw.*, vol. 12, no. 5, pp. 1163–1172, Sep. 2001.

[33] M. Dhimish, V. Holmes, B. Mehrdadi, and M. Dales, "Comparing Mamdani Sugeno fuzzy logic and RBF ANN network for PV fault detection," *Renew. Energy*, vol. 117, pp. 257–274, Mar. 2018.

[34] Y.-S Hwang and S.-Y. Bang, "An efficient method to construct a radial basis function neural network classifier and its application to unconstrained handwritten digit recognition," in *Proc. 13th Int. Conf. Pattern Recognit.*, Vienna, Austria, Aug. 1996, pp. 640–644.

[35] H. Sarimveis, A. Alexandridis, S. Mazarakis, and G. Bafas, "A new algorithm for developing dynamic radial basis function neural network models based on genetic algorithms," *Comput. Chem. Eng.*, vol. 28, nos. 1–2, pp. 209–217, 2004.

[36] S. Chen, S. A. Billings, C. F. N. Cowan, and P. M. Grant, "Non-linear systems identification using radial basis functions," *Int. J. Syst. Sci.*, vol. 21, no. 12, pp. 2513–2539, 1990.

[37] D. J. Swider, "A comparison of empirically based steady-state models for vapor-compression liquid chillers," *Appl. Thermal Eng.*, vol. 23, no. 5, pp. 539–556, 2003.

[38] H. Peng, T. Ozaki, Y. Toyoda, H. Shioya, K. Nakano, V. Haggan-Ozaki, and M. Mori, "RBF-ARX model-based nonlinear system modeling and predictive control with application to a NO_x decomposition process," *Control Eng. Pract.*, vol. 12, no. 2, pp. 191–203, 2004.

[39] W. You, Y.-X. Liu, B.-Z. Bai, and H.-S. Fang, "RBF-type artificial neural network model applied in alloy design of steels," *J. Iron Steel Res. Int.*, vol. 15, no. 2, pp. 87–90, 2008.

[40] H.-G. Han, J.-F. Qiao, and Q.-L. Chen, "Model predictive control of dissolved oxygen concentration based on a self-organizing RBF neural network," *Control Eng. Pract.*, vol. 20, no. 4, pp. 465–476, 2012.

[41] T. Sayahi, A. Tatar, and M. Bahrami, "A RBF model for predicting the pool boiling behavior of nanofluids over a horizontal rod heater," *Int. J. Thermal Sci.*, vol. 99, pp. 180–194, Jan. 2016.

[42] T. M. Cover, "Geometrical and statistical properties of systems of linear inequalities with applications in pattern recognition," *IEEE Trans. Electron. Comput.*, vol. EC-14, no. 3, pp. 326–334, Jun. 1965.

[43] M. D. Buhmann, *Radial Basis Functions: Theory and Implementations*. Cambridge, U.K.: Cambridge Univ. Press, 2003, pp. 156–179.

[44] R. J. Howlett, L. C. Jain, and J. Kacprzyk, *Radial Basis Function Networks 1: Recent Developments in Theory and Applications*. Berlin, Germany: Springer, 2001, pp. 269–284.

[45] F. Schwenker, H. A. Kestler and G. Palm, *Unsupervised and Supervised Learning in Radial-Basis-Function Networks in Self-Organizing Neural Networks*. Berlin, Germany: Springer, 2002, pp. 217–243.

[46] J. Moody and C. J. Darken, "Fast learning in networks of locally-tuned processing units," *Neural Comput.*, vol. 1, no. 2, pp. 281–294, 1989.

[47] M. Bianchini, P. Frasconi, and M. Gori, "Learning without local minima in radial basis function networks," *IEEE Trans. Neural Netw.*, vol. 6, no. 3, pp. 749–756, May 1995.

[48] J. Schmidhuber, *Deep Learning in Neural Networks*. Amsterdam, The Netherlands: Elsevier Science, 2015, pp. 16–24.

[49] J. M. Valls, I. M. Galván, and P. Isasi, "Lazy learning in radial basis neural networks: A way of achieving more accurate models," *Neural Process. Lett.*, vol. 20, no. 2, pp. 105–124, 2004.

[50] A. Khosravi, S. Nahavandi, and D. Creighton, "A prediction interval-based approach to determine optimal structures of neural network metamodels," *Expert Syst. Appl.*, vol. 37, no. 3, pp. 2377–2387, 2010.

[51] A. Khosravi, S. Nahavandi, and D. Creighton, *Load Forecasting and Neural Networks: A Prediction Interval-Based Perspective in Computational Intelligence in Power Engineering*. Berlin, Germany: Springer, 2010, pp. 131–150.

[52] A. Khosravi, S. Nahavandi, and D. Creighton, "Construction of optimal prediction intervals for load forecasting problems," *IEEE Trans. Power Syst.*, vol. 25, no. 3, pp. 1496–1503, Aug. 2010.

[53] J. G. Carney, P. Cunningham, and U. Bhagwan, "Confidence and prediction intervals for neural network ensembles," in *Proc. IJCNN*, Washington, DC, USA, Jul. 1999, pp. 1215–1218.

[54] D. J. C. MacKay, "The evidence framework applied to classification networks," *Neural Comput.*, vol. 4, no. 5, pp. 720–736, Sep. 1992.

[55] J. T. G. Hwang and A. A. Ding, "Prediction intervals for artificial neural networks," *J. Amer. Statist. Assoc.*, vol. 92, no. 438, pp. 748–757, Jun. 1997.

[56] D. A. Nix and A. S. Weigend, "Estimating the mean and variance of the target probability distribution," in *Proc. IEEE Int. Conf. Neural Netw.*, Orlando, FL, USA, Jul. 1994, pp. 55–60.

[57] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi, "Optimization by simulated annealing," *Science*, vol. 220, no. 4598, pp. 671–680, 1983.

[58] J. Vandekerckhove. *General Simulated Annealing Algorithm Version 1.0*. Accessed: Jun. 2, 2008. [Online]. Available: http://www.cogsci.uci.edu/~joachim/

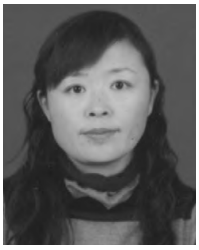
[59] L. Wei, Z. Zhang, D. Zhang, and S. C. H. Leung, "A simulated annealing algorithm for the capacitated vehicle routing problem with two-dimensional loading constraints," *Eur. J. Oper. Res.*, vol. 265, no. 3, pp. 843–859, 2018.

ZHANG JUN was born in Shenyang, Liaoning, China, in 1986. He received the B.S. degree in automation from the Shenyang University of Technology, Shenyang, China, in 2008, and the M.S. and Ph.D. degrees in control theory and control engineering from Northeastern University, Shenyang, China, in 2010 and 2015, respectively.

Since 2017, he has been a Postdoctoral Researcher with the College of Information Science and Engineering, Shenyang University of Technology, Shenyang, China, where he is currently a Lecturer. His current research interests include modeling, optimization, and control of complicated industrial production processes.

YAN HUA was born in 1964. She received the B.S. and M.S. degrees from the Shenyang University of Technology, Shenyang, China, in 1985 and 1988, respectively, and the Ph.D. degree from Northeastern University, Shenyang, China, in 1999.

She is currently a Professor and a Ph.D. Supervisor with the Shenyang University of Technology. Her main research interests include detection technology based on tomography, signal processing, and virtual instrument.



YU HONGXIA was born in Shenyang, Liaoning, China, in 1975. She received the B.S. degree in automation from the Liaoning University of Technology, Jinzhou, China, in 1998, the M.S. degree in control theory and control engineering from Northeastern University, Shenyang, China, in 2003, and the Ph.D. degree in mechatronic engineering from the Shenyang Institute of Automation (SIA), Shenyang, China, in 2012.

Since 2003, she has been a Lecturer with the College of Information Science and Engineering, Shenyang University of Technology, Shenyang, China. Her current research interests include motor control, smart grid dispatching, intelligent control, and adaptive control.

TIAN ZHONGDA was born in Shenyang, Liaoning, China, in 1978. He received the B.S. degree in communication engineering from Liaoning University, Shenyang, China, in 2001, the M.S. degree in communication and information system from Northeastern University, Shenyang, China, in 2004, and the Ph.D. degree in control theory and control engineering from Northeastern University, Shenyang, China, in 2013. He is currently an Associate Professor with the College of Information Science and Engineering, Shenyang University of Technology, China. His research interests include predictive control, delay compensation and scheduling for networked control systems, and time series prediction.

JIA RUNDA was born in Shenyang, Liaoning, China, in 1981. He received the B.S. degree in automation from the Dalian University of Technology, Dalian, China, in 2004, and the M.S. and Ph.D. degrees in control theory and control engineering from Northeastern University, Shenyang, China, in 2007 and 2011, respectively. He is currently an Associate Professor with the College of Information Science and Engineering, Northeast University. His research field includes modeling and soft sensors of complex systems.