processes

MDPI

*Article*

# Design of Soft-Sensing Model for Alumina Concentration Based on Improved Grey Wolf Optimization Algorithm and Deep Belief Network

**Jianheng Li [1], Zhiwen Chen [1], Xiaoting Zhong [2], Xiangquan Li [2,*], Xiang Xia [3,*] and Bo Liu [2]**

[1] School of Mechanical and Electrical Engineering, Central South University, Changsha 410083, China; ljh20040813@126.com (J.L.); zhiwen.chen@csu.edu.cn (Z.C.)
[2] School of Information Engineering, Jingdezhen University, Jingdezhen 333000, China; zxt20250218@126.com (X.Z.); 21062@jdzu.edu.cn (B.L.)
[3] School of Information Engineering, Jingdezhen Ceramic University, Jingdezhen 333403, China
* Correspondence: 21021@jdzu.edu.cn (X.L.); xiaxiang200112@163.com (X.X.)

**Abstract:** To address the challenge of the real-time monitoring of alumina concentrations during the production process, this paper employs a Deep Belief Network (DBN) within the framework of deep learning to predict alumina concentration. Additionally, the improved Grey Wolf Optimizer (IGWO) is utilized to optimize key parameters of the DBN model, including the number of hidden layer nodes, reverse iteration count, and learning rate. An IGWO-DBN hybrid model is then constructed and compared against DBN models optimized by other techniques, such as the Sparrow Search Algorithm (SSA) and Particle Swarm Optimization (PSO), to evaluate the predictive performance. The comparative analysis reveals that, in terms of predictive accuracy, the IGWO-DBN model outperforms both the SSA-DBN and PSO-DBN models. Specifically, it achieves lower root mean square errors (RMSE) and mean absolute errors (MAE), alongside a higher coefficient of determination ($R^2$). Furthermore, the IGWO-DBN model exhibits a faster convergence rate and a lower final convergence value, indicating superior generalization ability and robustness. Furthermore, the IGWO-DBN model not only demonstrates significant advantages in prediction accuracy for alumina concentration but also substantially reduces model training time through its efficient parameter optimization mechanism. The successful implementation of this model provides robust support for the intelligent and refined management of the aluminum electrolysis industry, aiding enterprises in reducing costs, improving production efficiency, and advancing the green and sustainable development of the industry.

**Keywords:** aluminum electrolysis; alumina concentration; deep belief network; neural network; improved grey wolf optimization; the IGWO-DBN model

check for updates

Academic Editor: Blaž Likozar

Received: 16 January 2025
Revised: 17 February 2025
Accepted: 19 February 2025
Published: 20 February 2025

**Citation:** Li, J.; Chen, Z.; Zhong, X.; Li, X.; Xia, X.; Liu, B. Design of Soft-Sensing Model for Alumina Concentration Based on Improved Grey Wolf Optimization Algorithm and Deep Belief Network. *Processes* **2025**, *13*, 606. https://doi.org/10.3390/pr13030606

**Copyright:** © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

## 1. Introduction

In industrial production, certain variables are challenging to measure online due to technological limitations or the high cost of sensors. However, these variables play a critical role in determining both intermediate and final product quality, thereby requiring continuous monitoring and control. In such cases, mathematical models—commonly referred to as soft sensing models—are developed to estimate these difficult-to-measure variables based on easily measurable ones [1].

In the aluminum electrolysis industry, manual inspections are not only costly and inefficient but also present significant safety risks. As a result, the application of soft

measurement techniques, particularly the use of soft sensing models for real-time alumina concentration prediction, becomes crucial. Maintaining alumina concentration within the optimal range is essential for ensuring the stable and efficient operation of industrial aluminum production. Practical experience indicates that the optimal alumina concentration range is between 1.5% and 2.5%, at which point the current efficiency of aluminum electrolysis is maximized [2]. If the alumina concentration is too high, precipitation may occur at the bottom of the cell, leading to reduced current efficiency, increased resistance, and cathode voltage drop, potentially compromising the stability of the aluminum liquid layer. Conversely, if the alumina concentration is too low, anode effects are more likely to occur, resulting in a sharp increase in cell voltage and disrupting the energy balance of the cell. Therefore, to achieve higher current efficiency, it is imperative to maintain the alumina concentration within a narrow range that is low enough to prevent anode effects, yet high enough to ensure operational stability. Hence, the accurate prediction and control of alumina concentration are critical prerequisites for an efficient, stable, and energy-efficient aluminum electrolysis process.

Given the challenge of the real-time monitoring of alumina concentrations, a critical parameter in aluminum electrolysis processes, scholars worldwide have conducted extensive research into its online measurement and prediction, yielding significant advancements. Specifically, Wang et al. [3] propose a rapid prediction method for alumina concentration that combines Particle Swarm Optimization (PSO) with Least Squares Support Vector Machines (LS-SVM). Li et al. [4] introduces a Bayesian regression model, grounded in Bayesian theory, for predicting alumina concentration within aluminum electrolysis cells. Li et al. [5] presents an innovative fuzzy expert control approach that integrates online intelligent identification, multi-control modes with self-decision mechanisms, and variable universe techniques, enabling both prediction and real-time control of alumina concentration. Wang et al. [6] suggests a fast and effective method for alumina concentration prediction, utilizing a Genetic Algorithm (GA) to optimize the structure and parameters of a BP neural network, which is subsequently used for prediction. Zhu et al. [7] proposes a prediction method for alumina concentration in electrolysis cells based on Long Short-Term Memory (LSTM) neural networks optimized by Particle Swarm Optimization (PSO). Furthermore, Li et al. [8] introduces a prediction model based on wavelet neural networks, which tracks cell resistance parameters related to alumina concentration, dynamically adjusts the controller, and regulates alumina feeding through a feeding device to maintain concentration within the desired range. Finally, Zhang et al. [9] presents a detection method for alumina concentration based on kernel extreme learning machines.

To achieve precise prediction of alumina concentration in aluminum electrolysis processes, this paper employs a Deep Belief Network (DBN) for alumina concentration prediction. DBN, which has been widely applied in various machine learning tasks and is capable of addressing complex issues effectively, thereby enhancing learning performance. The DBN has also found broad applications across diverse domains, such as fault diagnosis in photovoltaic arrays [10], bearing fault diagnosis [11], and research on predicting muzzle vibration in artillery [12].

Given the significant impact of DBN model parameters on prediction accuracy, this paper adopts an improved Grey Wolf Optimizer (GWO) algorithm to optimize and adjust these parameters. Several experts and scholars have previously achieved promising results in optimizing DBN model parameters using intelligent optimization algorithms. For example, Liu et al. [13] presents a fault diagnosis model based on a Particle Swarm Optimization-optimized Deep Belief Network (PSO-DBN) for multi-sensor information fusion. Kong et al. [14], proposes a combined navigation fault diagnosis technique utilizing a Genetic Algorithm-optimized Deep Belief Network (GA-DBN). In a paper published by our

team, Li et al., we combined Empirical Mode Decomposition (EMD) with Particle Swarm Optimization (PSO) and DBN to propose a soft-sensing model for alumina concentration, the EMD-PSO-DBN model. This model effectively denoises and optimizes the structure of the prediction model.

With ongoing research and development in soft measurement techniques, significant progress has been made in the accurate prediction of alumina concentration in aluminum electrolysis processes. However, the theoretical framework of soft sensing remains incomplete, and effective methodologies to guide its practical application are still lacking [15]. To further refine and improve the accuracy of alumina concentration prediction in soft sensing models, this paper introduces a Deep Belief Network (DBN) optimized by an Improved Grey Wolf Optimization (IGWO) algorithm. This approach addresses the issue of local optima resulting from random initialization by optimizing key network parameters, including the number of hidden layer nodes, the backpropagation learning rate, and the backpropagation iterations of the DBN. The IGWO-DBN model has been successfully implemented, enhancing both prediction accuracy and convergence speed. This work further refines the theoretical framework for accurate alumina concentration prediction, providing a robust methodological guide for the practical application of soft measurement techniques. The effectiveness of this approach has been validated through simulations, significantly enhancing both the prediction accuracy and the stability of the model.

## 2. Optimization Algorithm and Prediction Model

### *2.1. Improved Grey Wolf Optimizer Algorithm*

#### 2.1.1. Grey Wolf Optimizer Algorithm

The Grey Wolf Optimizer (GWO), inspired by the collective hunting behaviour of wolves, is a novel swarm intelligence optimization algorithm [16]. It iteratively searches for optimal solutions by simulating the hunting behaviour of wolves, demonstrating strong applicability and search capabilities for parameter optimization problems [17]. He et al. [18] shows that the GWO achieves significant improvements in both prediction accuracy and computational efficiency compared to traditional prediction algorithms. However, the initial population generation method in GWO leads to limited population diversity, as random initialization cannot guarantee effective diversity. Moreover, the search mechanism of the GWO algorithm often results in slow convergence during later stages. Wolves primarily determine their distance from prey based on their proximity to the $\alpha$, $\beta$, and $\delta$ wolves, which contributes to a slower convergence rate in the later stages. Additionally, the algorithm is prone to becoming stuck in local optima, as the $\alpha$ wolf may not always represent the global optimum. As the $\omega$ wolves continuously approach the top three wolves during iterations, the GWO algorithm risks converging to local optimal solutions. To address the issues of premature convergence and low accuracy in the basic GWO, an improved version incorporating a Dimensional Learning Hunting (DLH) search strategy is proposed [19].

#### 2.1.2. Improved Grey Wolf Optimizer Algorithm

The Improved Grey Wolf Optimizer (I-GWO), proposed by Nadimi-Shahraki H M et al. [20], is designed to address global optimization and engineering design problems. The I-GWO incorporates a novel movement strategy, known as the Dimensionality Learning Hunting (DLH) search strategy, which is inspired by the individual hunting behaviour of wolves in nature. In the DLH approach, various techniques are employed to construct a neighbourhood for each wolf, enabling the sharing of information among neighbouring wolves. The dimensionality learning aspect of the DLH strategy enhances the balance

between local and global search capabilities while preserving population diversity. The IGWO based on the DLH search strategy primarily involves the following stages:

(1) Initialization Phase

In this phase, N grey wolf individuals are randomly distributed within the search space according to Equation (1), within the given range $[l_j, u_j]$.

$$X_{ij} = l_j + rand_j[0,1] \times (u_j - l_j), \quad i \in [1,N], j \in [1,D] \tag{1}$$

The position of the i-th grey wolf at the t-th iteration is represented by the vector $X_i(t) = \{X_{i1}, X_{i2}, \ldots, X_{iD}\}$, where D denotes the dimensionality of the problem. The entire pack of wolves is stored in a matrix Pop, which has N rows and D columns. The fitness value of $X_i(t)$ is calculated by the fitness function $f(X_i(t))$.

(2) The Motion Phase

Apart from pack hunting, individual hunting is another social behaviour of grey wolves, which serves as the motivation for improving the GWO. The IGWO incorporates an additional movement strategy known as the Dimension Learning-based Hunting (DLH) search. In DLH, each wolf's neighbours are recognized as potential candidates for its new position $X_i(t)$. The following steps describe how the canonical GWO and the DLH search strategy generate two distinct candidates.

Canonical GWO Search Strategy: In GWO, the top three best wolves from the population Pop are identified as $\alpha$, $\beta$, and $\delta$. Subsequently, the linearly decreasing coefficient a and coefficients A and C are calculated using respective formulas. Then, taking into account the positions of $X_\alpha$, $X_\beta$, and $X_\delta$, the encircling of prey is determined through the corresponding formula. Finally, the first candidate for the new position of the grey wolf $X_i(t)$, denoted as $X_{i-GWO}(t+1)$, is computed using the relevant formula. Dimension Learning-based Hunting (DLH) Search Strategy: In the original GWO, a new position is generated for each wolf with the assistance of the three leading wolves from Pop. This approach results in slow convergence of GWO, premature loss of diversity within the population, and wolves becoming trapped in local optima. To address these deficiencies, the proposed DLH search strategy considers the individual hunting of wolves to be learned from their neighbours. In the DLH search strategy, each dimension of the new position of $X_i(t)$ is calculated by Equation (4), where the individual learns from its various neighbours and randomly selected individuals from Pop. Furthermore, besides $X_{i-GWO}(t+1)$, the DLH search strategy computes a radius $R_i(t)$ based on the Euclidean distance between the current position of $X_i(t)$ and the candidate position $X_{i-GWO}(t+1)$.

$$R_i(t) = ||X_i(t) - X_{i-GWO}(t+1)|| \tag{2}$$

Then, the neighbourhood of $X_i(t)$, denoted by $N_i(t)$, is constructed using Equation (3) with respect to the radius $R_i(t)$, where $D_i$ represents the Euclidean distance between $X_i(t)$ and $X_j(t)$.

$$N_i(t) = \{X_j(t) | D_i(X_i(t), X_j(t) \leq R_i(t), X_j(t) \epsilon Pop)\} \tag{3}$$

Once the neighborhood of $X_i(t)$ is constructed, more than neighborhood learning is performed by Equation (3), where the d-th dimension of $X_{i-DLH,d}(t+1)$ is computed based on a randomly selected $X_{i,d}(t)$ from Pop.

$$X_{i-DLH,d}(t+1) = X_{i,d}(t) + rand \times (X_{n,d}(t) - X_{r,d}(t)) \tag{4}$$

(3) Selection and Update Phase

At this stage, firstly, the more outstanding candidate position is selected by comparing the fitness values of the two candidate positions $X_{i-GWO}(t+1)$ and $X_{i-DLH,d}(t+1)$ through Formula (5).

$$X_i(t+1) = \begin{cases} X_{i-GWO}(t+1) \\ X_{i-DLH,d}(t+1) \end{cases} \quad (5)$$

Then, in order to update the new position of $X_i(t+1)$, if the fitness value of the selected candidate is smaller than that of $X_i(t)$, $X_i(t)$ is updated with the selected candidate; otherwise, $X_i(t)$ remains unchanged in Pop.

Lastly, after this process is performed for all individuals, the iteration counter (iter) is incremented by 1, and the search can iterate until the predefined number of iterations (Maxiter) is reached. The process flow is shown in Figure 1.

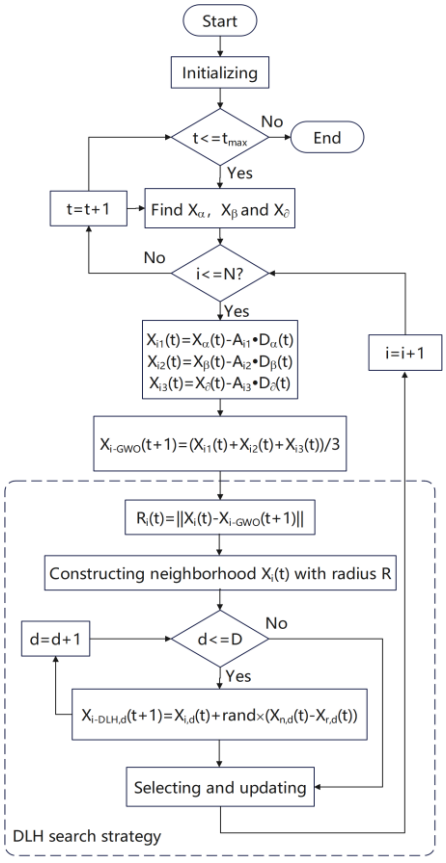



**Figure 1.** Flowchart of the IGWO algorithm.

*2.2. Deep Belief Network (DBN)*

DBN, proposed by Hinton et al. [21], is a probabilistic deep learning model that combines unsupervised and supervised learning. It integrates deep learning with feature learning, enabling efficient analysis of large-scale time-series data within a short time frame. With its superior data fitting capabilities, DBN addresses the limitations of backpropagation (BP) networks, which have stringent input data requirements and relatively long prediction cycles [22].

The DBN is structured as a multi-layer neural network, comprising an input layer, an output layer, and several hidden layers. Unlike a standard deep neural network (DNN),

DBN uses Restricted Boltzmann Machines (RBMs) as its fundamental building blocks. As shown in Figure 2, the network consists of a stack of multiple RBMs at the bottom, followed by a layer of Backpropagation (BP) network at the top. In an RBM, the weights and bias parameters of the visible and hidden layers are trained independently and in an unsupervised manner. Training concludes after the specified number of iterations, and the network is subsequently connected to the initial DBN model. However, this approach typically leads to a locally optimal solution, as each layer of the RBM reaches its optimal state individually [23]. The training process is divided into two stages: pre-training and fine-tuning. The pre-training stage involves training each RBM layer in the DBN progressively while preserving as much feature information as possible for subsequent layers [24]. The fine-tuning stage adjusts the parameters using the traditional BP algorithm to ensure that all model parameters converge to their optimal values.

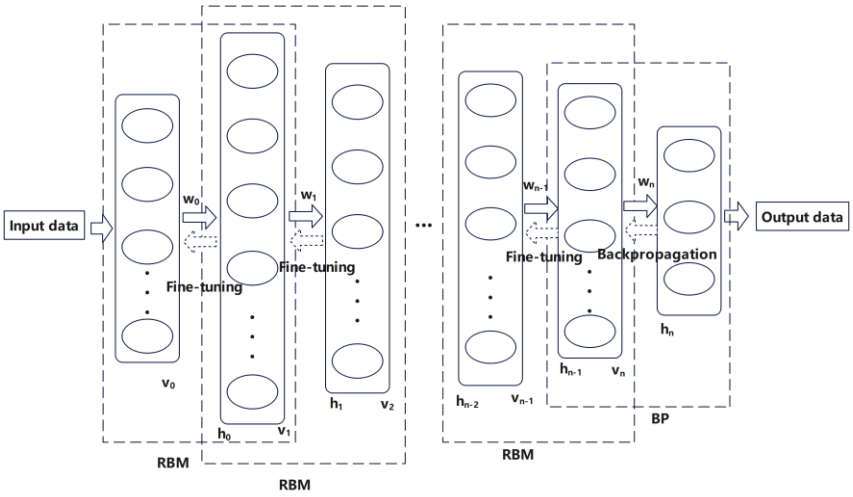



**Figure 2.** Diagram of the DBN structure.

### *2.3. Model Parameter Selection*

The successful application of Deep Belief Networks (DBNs) depends critically on the appropriate configuration of their architecture, particularly the optimal size, which is determined by the number of hidden neurons [25]. The IGWO-DBN model proposed in this paper is constructed with an input layer, two RBM layers, and an output layer. The pre-training process involves 150 epochs, with five sets of data trained per epoch. The pre-training learning rate is set to 0.005. The number of hidden layer nodes in the two RBM layers is optimized using the IGWO algorithm. Additionally, the number of iterations and the learning rate for backpropagation fine-tuning are optimized by IGWO, with five sets of data trained per iteration during the fine-tuning phase. The scale factor for the learning rate during fine-tuning is set to 1.

The prediction accuracy of the model is strongly influenced by the configuration of network parameter [26]. The performance of the DBN model is highly sensitive to the choice of hyperparameters. By optimizing these parameters through algorithms, the optimal settings can be obtained, thus enhancing the model's performance [27]. Determining the number of hidden layer nodes in the DBN does not follow a regular pattern, yet this choice is crucial for the proper functioning of the network. Similarly, the learning rate and number of iterations in the fine-tuning stage significantly impact network performance. The learning rate controls the speed of weight updates: a higher learning rate can accelerate convergence but may also cause overfitting. While increasing the number of training epochs can improve model accuracy, excessive training may also lead to overfitting. This paper

optimizes and automatically searches for the three critical parameters of the DBN model using the IGWO algorithm..

## 3. Experimental Study and Model Validation

### *3.1. Aluminum Electrolysis Data*

The experimental data in this study consist of voltage and current measurements collected from 24 anode rods on side A of electrolytic cell No. 2875 during the aluminum electrolysis production process at the Zunyi Aluminum Factory. The anode rod measurement device used in this experiment is shown in Figure 3 [28].



**Figure 3.** Schematic diagram of offline commissioning of anode rod measuring device.

The anode rod measurement device transmits data to the host computer via the RS485 communication protocol, while simultaneously displaying real-time data locally. The real-time data collected is shown in Figure 4 [28].




**Figure 4.** Data acquisition display diagram of anode current measuring device.

Due to the lack of suitable sensors for measuring alumina concentration in current technologies, alumina concentration data are primarily obtained through on-site sampling. This process involves multiple steps, including drilling sampling holes, collecting samples, cooling, packaging, and sending the samples to the laboratory for assay and chemical analysis. An overview of this process is illustrated in Figure 5 [28].

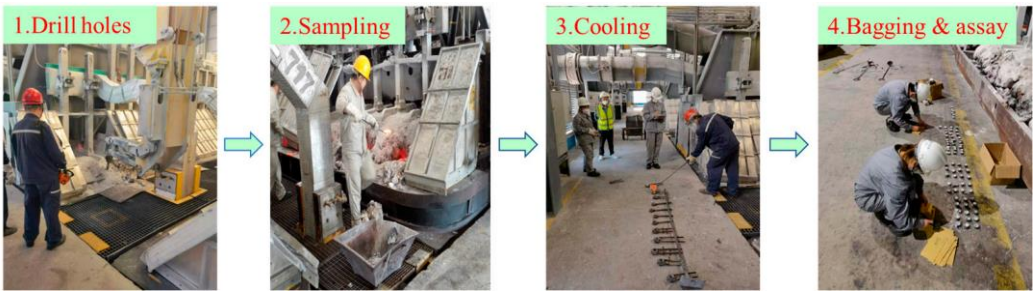



**Figure 5.** The collection process of electrolyte.

**Remark 1.** *The use of real-time voltage and current measurements from anode rods provides a continuous and accurate representation of the electrolysis process, which can be critical for optimizing operational parameters. However, the reliance on manual sampling for alumina concentration data remains a limitation in current practises, suggesting a need for the development of more advanced soft sensing models to improve the efficiency and accuracy of alumina concentration data collection in aluminum electrolysis operations.*

The data for the dominant variable of alumina concentration, along with the corresponding auxiliary variables of bipolar voltage and anode conductor current, were collected. Among these, 600 complete datasets were selected for verification, with 500 sets used as the training sample set and 100 sets reserved for testing. To eliminate the differences in the order of magnitude among the data and avoid large prediction errors due to significant disparities between input and output magnitudes, the data were normalized prior to DBN prediction. The data normalization was performed using the maximum-minimum standardization method, and the calculation formula is as follows:

$$X_{std} = \frac{x - x_{min}}{x_{max} - x_{min}} \tag{6}$$

In the formula: $X_{std}$ represents the processed number, X represents the initial number, $X_{max}$ represents the maximum number of selected samples, and $X_{min}$ represents the minimum number of selected samples.

### *3.2. Model Evaluation Metrics*

The performance of the DBN model is determined by its goodness-of-fit and accuracy, which can be described through metrics such as the coefficient of determination ($R^2$), root mean square error (RMSE), and mean absolute error (MAE). The formula for calculating $R^2$ is as follows:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \overline{y})^2} = 1 - \frac{RMSE}{Var} \in [0,1] \tag{7}$$

where *Var* represents the variance, and RMSE stands for the root mean square error.

The larger the $R^2$ value (approaching 1), the smaller the error between the predicted and actual values in the sample, indicating a better explanation of the dependent variable by the independent variable in the regression analysis, i.e., the more accurate the model's prediction results.

RMSE, standing for root mean square error, is a typical metric for regression models, representing the error generated by the model during prediction.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}, \in [0, +\infty) \tag{8}$$

where y represents the actual value and ŷ denotes the predicted value. Generally, the smaller the RMSE value, the more accurate the model's prediction results are.

MAE, standing for mean absolute error, is used to measure the average absolute difference between the predicted and actual values.

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i|, \in [0, \infty) \quad (9)$$

where y is the actual value and ŷ is the predicted value. Typically, the smaller the MAE value, the more accurate the model's prediction results are.

### 3.3. Design of the Soft Sensor Model

Based on the theoretical analysis presented above and leveraging the advantages of the Improved Grey Wolf Optimization (IGWO) algorithm, this study proposes a hybrid forecasting model, termed IGWO-DBN, which integrates the strengths of both methodologies. The model encompasses several critical stages, including data input and preprocessing, parameter optimization via IGWO, DBN training, prediction, and final output. The detailed framework is outlined as follows:

(1) Data Input and Preprocessing

The system begins by importing the relevant datasets. Subsequently, the input data undergoes preprocessing, with normalization applied in this study to ensure data quality and enhance the model's learning efficiency. Normalization standardizes the data range, thereby mitigating potential biases and improving the convergence rate during the training process.

(2) IGWO-Based Parameter Optimization

The optimization process begins with the initialization of the grey wolf population, where multiple candidate solutions are randomly generated as the starting points for the search. The fitness values of the current grey wolf positions are then computed, and the primary candidate position is identified. A local search is conducted within a defined radius around this candidate position to explore potential improvements. To further refine the search, a multi-neighbourhood learning strategy (DLH) is employed, yielding higher-quality candidate positions. The positions are updated by integrating the primary candidate position, the results from the DLH search, and the individual with the lowest fitness value, thereby enhancing search precision. The fitness values of all candidate individuals are compared, and the top three grey wolves are selected as the new leaders for the subsequent search. The algorithm iterates this process until the maximum iteration count is reached, at which point the optimal parameters are output.

(3) DBN Training and Prediction

The optimal parameters obtained from the IGWO process are utilized to train the Deep Belief Network (DBN). Once the DBN model is configured with these parameters, it undergoes training using the prepared dataset to learn the underlying patterns. Upon completion of the training phase, the model is applied to the test dataset to perform the prediction task, generating forecasts for the target variable, such as alumina concentration. The overall workflow of the proposed IGWO-DBN model is illustrated in Figure 6.

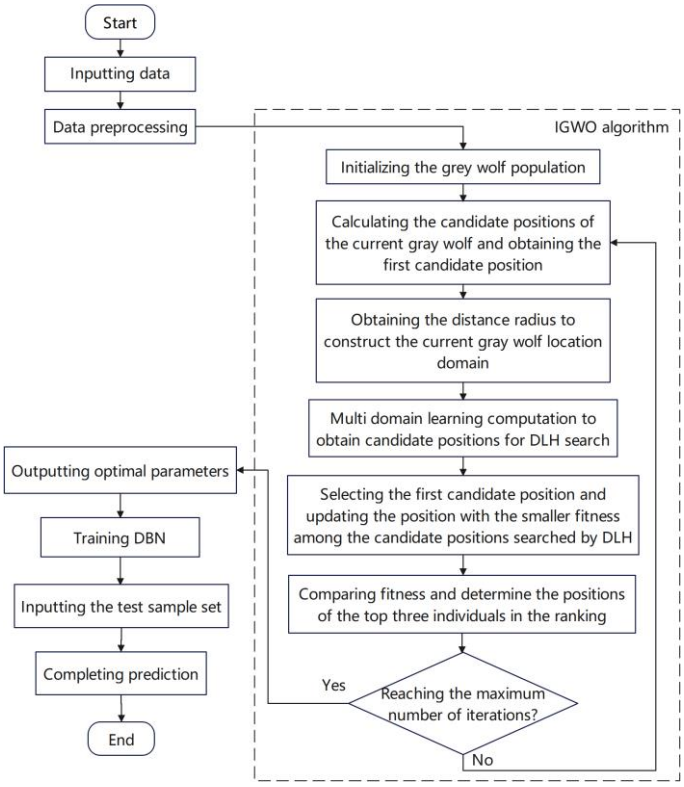



**Figure 6.** Flowchart of the soft sensor model design based on IGWO-DBN.

## 4. Results Analysis

### *4.1. Parameter Optimization Results*

The optimal number of hidden layer nodes, the number of backpropagation iterations, and the backpropagation learning rate obtained after optimizing the DBN through the IGWO, SSA, and PSO algorithms are shown in Table 1.

**Table 1.** Optimized network parameters for the DBN model.

| Composition Model | Hidden Layer Structure | Reverse Learning Rate | Reverse Iteration Times |
|---|---|---|---|
| PSO-DBN | 25–25 | 0.9675 | 200 |
| SSA-DBN | 28–64 | 0.9265 | 400 |
| IGWO-DBN | 64–60 | 0.6083 | 481 |

### *4.2. Comparison of Simulation Results*

To verify the improvement effect of IGWO on the DBN, the traditional DBN model is compared with the DBN model improved by IGWO, and the simulation results are shown in Figure 7. This paper also selects the models optimized by the Sparrow Search Algorithm (SSA) and Particle Swarm Optimization (PSO) for comparison with the IGWO-optimized model. After tuning the networks, the simulation results are shown in Figure 8.

As illustrated in Figure 7, the model optimized using IGWO exhibits noticeable improvements in prediction accuracy. Figure 8 further demonstrates that models optimized by different algorithms achieve varying degrees of accuracy enhancement. Among these, the DBN model optimized by the Improved Grey Wolf Optimization (IGWO) algorithm outperforms the other two models in overall performance, showing smaller deviations between predicted and actual alumina concentration values, higher prediction accuracy, and greater system stability. Additionally, in terms of convergence, the IGWO-DBN model demonstrates a significant advantage, achieving a faster convergence rate and a lower final convergence value compared to the other two models, as shown in Figure 9.

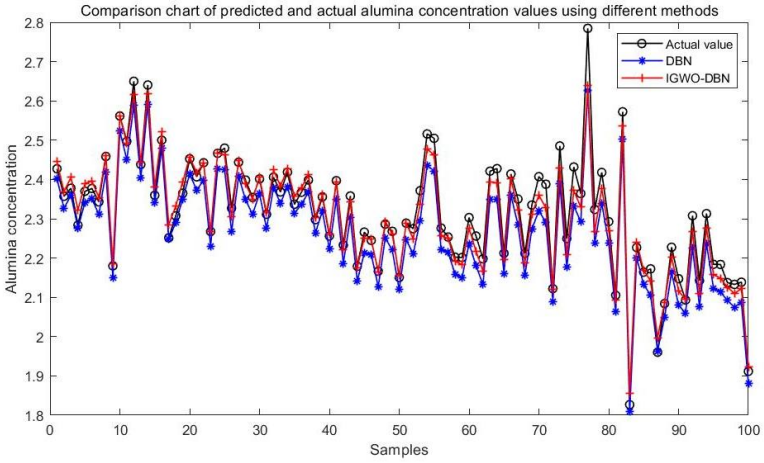



**Figure 7.** Comparison chart of prediction results between the traditional DBN model and the IGWO-DBN model.

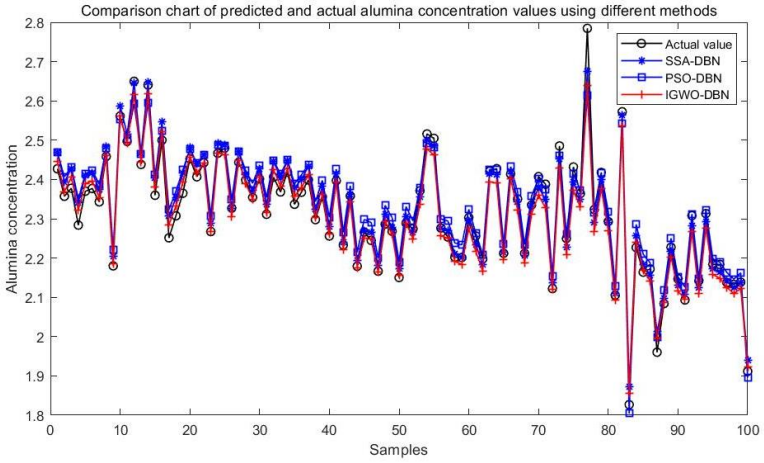



**Figure 8.** Comparison chart of predicted and actual alumina concentration values using different methods.

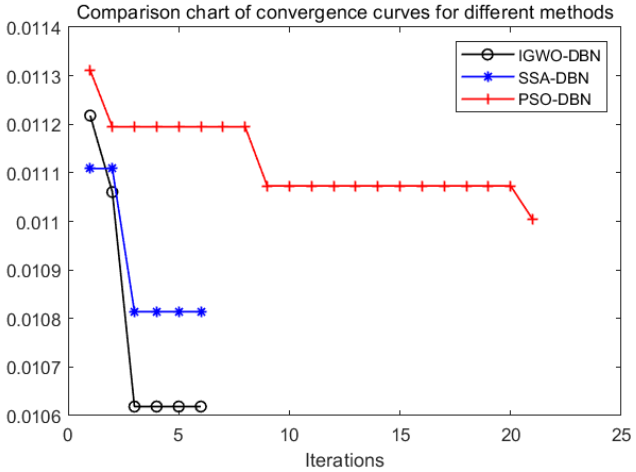



**Figure 9.** Convergence curve of different methods.

Table 2 presents the data of evaluation indicators corresponding to the four models. Among PSO-DBN, SSA-DBN, and IGWO-DBN, IGWO-DBN has the best fitting effect. Moreover, the MAE and RMSE values of the IGWO-DBN model are slightly lower than those of the other models, confirming the effectiveness and accuracy of this model.

**Table 2.** Comparison of prediction results optimized by different algorithms.

| Prediction Model | MAE | RMSE | $R^2$ |
|---|---|---|---|
| DBN | 0.048723 | 0.054227 | 0.97494 |
| PSO-DBN | 0.030706 | 0.037375 | 0.95899 |
| SSA-DBN | 0.021147 | 0.032541 | 0.96041 |
| IGWO-DBN | 0.020947 | 0.028466 | 0.96846 |

## 5. Conclusions

This study successfully introduces an alumina concentration prediction model based on an Improved Grey Wolf Optimization (IGWO)-optimized Deep Belief Network (DBN). By employing the IGWO algorithm to optimize the number of hidden layer nodes during the pre-training phase, as well as the learning rate and iteration count during fine-tuning, the model effectively minimizes the impact of manual parameter selection on prediction accuracy.

Experimental results demonstrate that the IGWO-DBN model surpasses traditional PSO-DBN and SSA-DBN models in both prediction accuracy and reliability. It achieves lower root mean square error (RMSE) and mean absolute error (MAE), a higher coefficient of determination ($R^2$), faster convergence speed, and a lower final convergence value. These improvements highlight the model's superior generalization capability and robustness, underscoring its practical significance in aluminum electrolysis production. The proposed approach enhances production efficiency and ensures more stable monitoring of alumina concentration, thereby providing stronger support for intelligent decision-making in process control.

Despite these promising results, certain limitations remain. The theoretical framework for accurately predicting alumina concentration in aluminum electrolysis is still evolving. Future research should focus on further minimizing prediction errors in soft-sensing models and improving their accuracy to enhance both robustness and practical applicability. Continued exploration in this direction will be crucial for advancing the field and addressing existing challenges.

**Author Contributions:** Conceptualization, Z.C. and X.L.; methodology, J.L. and X.Z.; validation, X.X.; formal analysis, B.L.; resources, X.L.; writing—original draft preparation, J.L. and X.Z.; writing—review and editing, X.X. and X.L.; visualization, B.L.; supervision, Z.C.; project administration, X.L.; funding acquisition, X.L. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work is jointly supported by Natural Science Foundation Project of JiangXi Province, Grant Number 20242BAB25091, Natural Science Foundation Project of Guizhou Province, Grant Number ZK [2023] Genera004, Science and Technology Project of Jiangxi Provincial Department of Education, Grant Number GJ2202404.

**Data Availability Statement:** The datasets presented in this article are not readily available because the data are part of an ongoing study or due to technical. Requests to access the datasets should be directed to 21021@jdzu.edu.cn.

**Conflicts of Interest:** The authors declare no conflicts of interest.

## References

1. Lawrence, N.P.; Damarla, S.K.; Kim, J.W.; Tulsyan, A.; Amjad, F.; Wang, K.; Chachuat, B.; Lee, J.M.; Huang, B.; Gopaluni, R.B. Machine Learning for Industrial Sensing and Control: A Survey and Practical Perspective. *Control Eng. Pract.* **2024**, *4*, 105841. [CrossRef]

2. Ding, X.; Xu, W.; Ma, J. Accurate Control Analysis of Alumina Concentration in Large Prebaked Aluminum Electrolysis Cell. *World Nonferrous Met.* **2005**, *9*, 26–28.
3. Wang, M.; Cao, B.; Huang, R.; Ao, Y. Prediction of Alumina Concentration Based on PSO Optimized LS-SVM. *Light Met.* **2020**, *10*, 33–40. [CrossRef] [PubMed]
4. Li, Y. Research on Alumina Concentration Prediction in Aluminum Electrolysis Cell. Master's Thesis, Inner Mongolia University of Science and Technology, Baotou, China, 2019. Available online: https://cdmd.cnki.com.cn/Article/CDMD-11973-1019663265.htm (accessed on 26 February 2020).
5. Li, J.; Zhang, W.; Ding, F.; Li, Y. Fuzzy Expert Control Method Based on Online Intelligent Identification and Its Application. *J. Cent. South Univ. (Sci. Technol.)* **2004**, *6*, 911–914.
6. Wang, S.; Li, B.; Lin, J. Prediction of Alumina Concentration Based on Genetic Neural Network. *Mod. Electron. Tech.* **2011**, *12*, 163–165+172.
7. Zhu, J.; Xu, Y.; Cao, B.; Cheng, J. Research on Prediction of Alumina Concentration in Electrolysis Cell Based on PSO+LSTM for Aluminum Industry. *Light Met.* **2023**, *7*, 21–27.
8. Li, J.; Wu, C.; Li, Y. Research on the Application of Neural Network Predictive Control Technology in Aluminum Electrolysis Process. *Instrum. Tech. Sens.* **2011**, *8*, 91–93. [CrossRef]
9. Zhang, S.; Zhang, T.; Yin, Y.; Xiao, W. Alumina Concentration Detection Based on the Kernel Extreme Learning Machine. *Sensors* **2017**, *9*, 2002. [CrossRef] [PubMed]
10. Peng, H.; Tian, C.; Zheng, Y.; Huang, J. Fault Diagnosis Method for Photovoltaic Power Generation Arrays Based on Deep Belief Network. *J. Nav. Univ. Eng.* **2023**, *12*, 1–8.
11. Sun, Y.; Jin, Z. Research on fault diagnosis method of bearing based on parameter optimization VMD and improved DBN. *J. Vibroengineering* **2023**, *6*, 1068–1082. [CrossRef]
12. Gao, X.; Gao, F.; Gao, Y.; He, Z.; Cheng, C.; Han, Z. Research on Prediction of Gun Muzzle Vibration Based on Deep Belief Network. *Fire Control Command Control* **2023**, *11*, 164–168.
13. Liu, F.; Zhang, D. Fault Diagnosis of Hydraulic System Cooler Based on PSO-DBN. *Comput. Mod.* **2023**, *6*, 7–14.
14. Kong, W. Research on Fault Diagnosis Technology of Integrated Navigation Based on GA-DBN. Master's Thesis, Harbin Institute of Technology, Harbin, China, 2022. Available online: https://kns.cnki.net/kcms2/article/abstract?v=F7lgbn8YeJruCXY69eVXW0hhGJq1sPTHa0_AlFhQxkTGYTtUjN0OuPutuCV1FLGJGaomGL4lMVoLRLQ9HqcIfVaTpdVyVxleV35A0EbaRlaRoGFwSUanXX7-ZAG3RDG4fHtMKGYJa9pwSDHir6ScR-Y6tQV4t7KmjV19R_eShM4=&uniplatform=NZKPT (accessed on 11 February 2022.).
15. Cui, J.; Zhang, N.; Yang, X.; Li, Q.; Cao, B.; Wang, M. Soft sensing of alumina concentration in aluminum electrolysis industry based on deep belief network. *Chin. Autom. Congr. (CAC)* **2020**, *3*, 1–5.
16. Mirjalili, S.; Mirjalili, S.M.; Lewis, A. Grey wolf optimizer. *Adv. Eng. Softw.* **2014**, *3*, 46–61. [CrossRef]
17. Liu, J.; Maimaitireyimu, A.; Xiang, Z.; Xie, L. Fault Diagnosis of Wind Turbine Bearings Based on Improved Grey Wolf Optimization Algorithm and SVM. *J. Mech. Transm.* **2023**, *9*, 160–169.
18. He, Y.; Shi, Z.; Guo, X.; He, W.; Han, T. Photovoltaic power prediction algorithm based on parameter-optimized multi-kernel support vector machine. *Acta Energiae Solaris Sin.* **2024**, *10*, 394–404.
19. Zhou, J.; Zheng, R.; Hou, H. Improved Grey Wolf Algorithm Based on Dimensional Learning Hunting Search Strategy. *Instrum. Tech. Sens.* **2023**, *12*, 88–93+98.
20. Nadimi-Shahraki, H.M.; Taghian, S.; Mirjalili, S. An Improved Grey Wolf Optimizer for Solving Engineering Problems. *Expert Syst. Appl.* **2021**, *3*, 113917. [CrossRef]
21. Hinton, G.E. Deep belief networks. *Scholarpedia* **2009**, *6*, 5947. [CrossRef]
22. Fang, N.; Chen, H.; Li, X.; Deng, X. Short-term Power Load Forecasting Based on CNN-PSO-DBN. *Comput. Simul.* **2023**, *10*, 118–122.
23. Guo, D.; Song, Z.; Liu, N.; Xu, T.; Wang, X.; Zhang, Y.; Su, W.; Cheng, Y. Performance Study of Hard Rock Cantilever Roadheader Based on PCA and DBN. *Rock Mech. Rock Eng.* **2024**, *1*, 2605–2623. [CrossRef]
24. Kuremoto, T.; Kimura, S.; Kobayashi, K.; Obayashi, M. Time series forecasting using a deep belief network with restricted Boltzmann machines. *Neurocomputing* **2014**, *8*, 47–56. [CrossRef]
25. Wang, G.; Yuan, G.; Hu, Z.; Cui, Y.; Jia, Q.; Qiao, J. Complexity-Based Structural Optimization of Deep Belief Network and Application in Wastewater Treatment Process. *IEEE Trans. Ind. Inform.* **2024**, *1*, 6974–6982. [CrossRef]

26. Liu, D.; Zhou, L.; Zheng, X. Ultra-Short-Term Power Load Forecasting Based on SA-DBN. *J. Guangxi Norm. Univ. (Nat. Sci. Ed.)* **2021**, *4*, 21–33.
27. Panigrahy, S.; Dash, S.; Padhy, S. Optimized Deep Belief Networks Based Categorization of Type 2 Diabetes using Tabu Search Optimization. *Int. J. Adv. Comput. Sci. Appl.* **2024**, *11*, 667. [CrossRef]
28. Li, X.; Liu, B.; Qian, W.; Rao, G.; Chen, L.; Cui, J. Design of Soft-Sensing Model for Alumina Concentration Based on Improved Deep Belief Network. *Processes* **2022**, *10*, 2537. [CrossRef]

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.