

Article

A Soft Measurement Method for the Tail Diameter in the Growing Process of Czochralski Silicon Single Crystals

Lei Jiang ^{1,2,*}, Da Teng ¹ and Yue Zhao ^{1,2}

¹ School of Automation and Information Engineering, Xi'an University of Technology, Xi'an 710048, China; tengdaxaut@163.com (D.T.); zhaoyue@xaut.edu.cn (Y.Z.)

² Crystal Growth Equipment and System Integration National & Local Joint Engineering Research Center, Xi'an University of Technology, Xi'an 710048, China

* Correspondence: jianglei@xaut.edu.cn

Abstract: In the Czochralski silicon single crystal growth process, the tail diameter is a key parameter that cannot be directly measured. In this paper, we propose a real-time soft measurement method that combines a deep belief network (DBN) and a support vector regression (SVR) network based on system identification to accurately predict the crystal diameter. The main steps of the proposed method are as follows: First, we address the delay problem of the effects of the temperature and crystal pulling speed on the tail diameter growth by using a back propagation (BP) neural network based on the mean impact value (MIV) method to determine the optimal delay time. Second, we construct a prediction model of the tail diameter by using the DBN network with the temperature and crystal pulling speed as input variables in the crystal growth process. Third, we improve the DBN network by using the SVR network to enhance its linear regression capability. We also employ the ant colony optimization (ACO) algorithm to obtain the optimal parameters of the SVR network. Finally, we compare the performance of the DBN-ACO-SVR network based on system identification with the DBN and SVR networks, and the results show that our method can effectively deal with the delay problem and achieve the accurate prediction of the tail diameter in the Czochralski silicon single crystal growth process.

Keywords: soft measurement; Czochralski crystal growth; neural networks; system identification; the diameter of tail growth



Citation: Jiang, L.; Teng, D.; Zhao, Y. A Soft Measurement Method for the Tail Diameter in the Growing Process of Czochralski Silicon Single Crystals. *Appl. Sci.* **2024**, *14*, 1569. <https://doi.org/10.3390/app14041569>

Academic Editor: Alexandre Carvalho

Received: 28 November 2023

Revised: 11 February 2024

Accepted: 14 February 2024

Published: 16 February 2024



Copyright: © 2024 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 the ever-evolving landscape of semiconductor and solar photovoltaic technology, monocrystalline silicon stands as a material of paramount importance. Its production predominantly relies on the Czochralski process, a sophisticated and multi-staged method encompassing seeding, shoulder formation, body growth, and tailing. The basic principle is to use the molten silicon material to contact a small monocrystalline silicon seed crystal, and slowly pull the seed crystal at a specific temperature, so that the molten silicon solidifies on the surface of the seed crystal to form monocrystalline silicon. During the whole process, the temperature of the molten silicon is kept slightly higher than the melting point of silicon, and the uniformity and crystallization quality in the crystal growth process are promoted by rotating the seed crystal and/or furnace. The tail process reduces the crystal diameter to a certain extent by slowly increasing the temperature and crystal pulling speed, and finally leaves the silicon melt level. In the tail process, if the crystal diameter is not properly controlled, it will lead to crystal dislocation and crystal growth failure [1–3]. A schematic diagram of the Czochralski growth of monocrystalline silicon is shown in Figure 1.

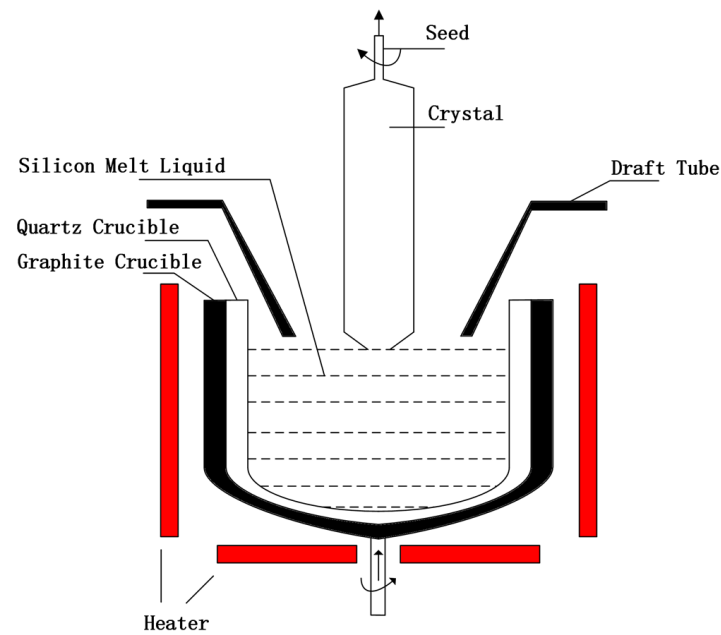


Figure 1. A schematic diagram of the Czochralski growth of monocrystalline silicon.

For crystal growth, it is necessary to ensure the integrity of crystal growth. Integrity means that in the tail of crystal growth, the tail diameter is continuously reduced until it is separated from the silicon melt. The integrity of the crystal tail can effectively prevent the reverse extension of crystal dislocations, thus damaging the quality of the wafer at the body stage. The different tail shapes are shown in Figure 2. Therefore, the tail diameter must be effectively controlled, and the premise of control is to measure the tail diameter in real time. Generally, we want the tail length to be as short as possible to save silicon. However, in the tail process, the crystal diameter obtained by image measurement [4] needs to ensure a longer tail length so that the crystal meniscus is visible. However, a longer tail length not only wastes silicon but also takes a large amount of time. Similarly, the method of using high-temperature infrared pyrometers to obtain the crystal diameter has the same disadvantages. Also, taking the crystal mass variation to estimate the crystal diameter of the tail growth still has many shortcomings. On the one hand, due to the detection accuracy and observation noise of the sensor, and on the other hand, due to the time-varying shape of the solid–liquid interface in the tail process [5], the estimation results are inaccurate. Therefore, we need to study the indirect measurement technology of the crystal diameter.



Figure 2. Tail shape of a crystal.

The realm of soft measurement has witnessed significant advancements in recent years. In various studies, innovative approaches have been adopted, such as introducing dropout regularization into deep convolution networks for enhanced accuracy in health status identification [6], employing mutual information-based methods for effective biochemical oxygen demand prediction [7], and utilizing fuzzy least squares support vector machines optimized by adaptive genetic algorithms for torque measurement in AC dynamome-

ters [8]. In addition, in references [9,10], the forecast models and wavelet neural network models have been applied to the field of soft measurement. However, these existing techniques, while groundbreaking, fall short in addressing the specific challenges posed by the Czochralski process, particularly in accurately measuring the tail diameter amid fluctuating process conditions. This highlights a significant gap in the existing methodologies, which need a more precise, non-intrusive, and efficient approach to measure the tail diameter.

To address these challenges, this paper introduces a novel soft measurement method, DBN-ACO-SVR, leveraging the synergistic potential of a deep belief network (DBN) and a support vector regression (SVR) network, enriched by system identification techniques so as to precisely predict the crystal diameter. We begin by addressing the delay effect of the temperature and crystal pulling speed on the tail diameter growth, utilizing a backpropagation (BP) neural network based on the mean impact value (MIV) method to determine the optimal delay time. This critical step ensures the synchronization of the process variables with the actual growth dynamics of the crystal tail. At the heart of our method is the DBN network, which leverages the temperature and crystal pulling speed as inputs to capture intricate relationships between process variables and the tail diameter, and offers a sophisticated predictive model. To refine this model, we augment the DBN network with an SVR network, enhancing its linear regression capabilities and infusing robustness against non-linear variations in the data. The ant colony optimization (ACO) algorithm is employed for fine-tuning the SVR network's parameters, ensuring accuracy and adaptability under varying data conditions. Extensive computational experiments benchmark our DBN-ACO-SVR network against traditional models, demonstrating a superior capability in addressing the delay problem and achieving precise tail diameter predictions.

The contributions of this study are manifold and significant. Primarily, it pioneers a soft measurement technique for the tail diameter in the Czochralski process, a task previously marred by technical and practical challenges. This technique not only enhances the accuracy, but also negates the need for direct measurement methods that are often cumbersome and inefficient. The study further elucidates the relationship between process variables (temperature and pulling speed) and the tail diameter by quantifying their impact through the MIV method. This contributes to a deeper understanding of the crystal growth process. Moreover, the integration of DBN and SVR networks, optimized by ACO, marks an innovative approach in the realm of predictive modeling, offering a novel solution to a longstanding problem in semiconductor manufacturing. This research, therefore, holds substantial promise for advancing the precision and efficiency of monocrystalline silicon production, with potential ramifications in the semiconductor and solar photovoltaic industries.

2. Data Collection

Generally, the tail-growing process in the Czochralski silicon single-crystal growing process includes basic processing parameters and control processing parameters. Basic processing parameters include the crystal rotation, crucible rotation, argon flow, magnetic field strength, furnace pressure, liquid level position, and other process parameters. The control processing parameters include the temperature and pulling speed, which are used to adjust the crystal diameter in the tail-growing process.

Due to the lack of technology to measure the tail diameter at present, we have selected four measurable and directly related variables—temperature, pulling speed, tail length, and mass increment—as data sources for predicting the crystal tail diameter.

The data in this article were collected from crystal-growing equipment on-site. The model of the single-crystal furnace was TDR-120CZ, as shown in Figure 3. This equipment can produce CZ silicon single crystals of 100–310 mm diameter. The diameter measurement range is 4–350 mm. The equipment has a maximum power of 180 kW. The crucible rotation speed can be adjusted within the range of 0–15 rpm, while the crystal rotation speed can be adjusted within the range of 0–20 rpm. The crucible lifting speed can be adjusted within the

range of 0–0.5 mm/min, and the crystal lifting speed is within the range of 0–6 mm/min. The ultimate vacuum degree is 0.4 Pa, and the inlet air flow can be adjusted within the range of 0–150 L/min, with a liquid level adjustment range of 0–100 mm. The crystal lifting stroke is 2.8 m, and the crucible lifting stroke is 600 mm. In terms of on-site process parameters, the target set value for the diameter is 308 mm, and the casting speed can be adjusted within the range of 0.2–1.1 mm/min. The temperature adjustment is determined based on the input of the silicon material quantity. The general rule is to first cool and then raise the temperature, with the adjustment range typically being 2100–2500 (dimensionless).

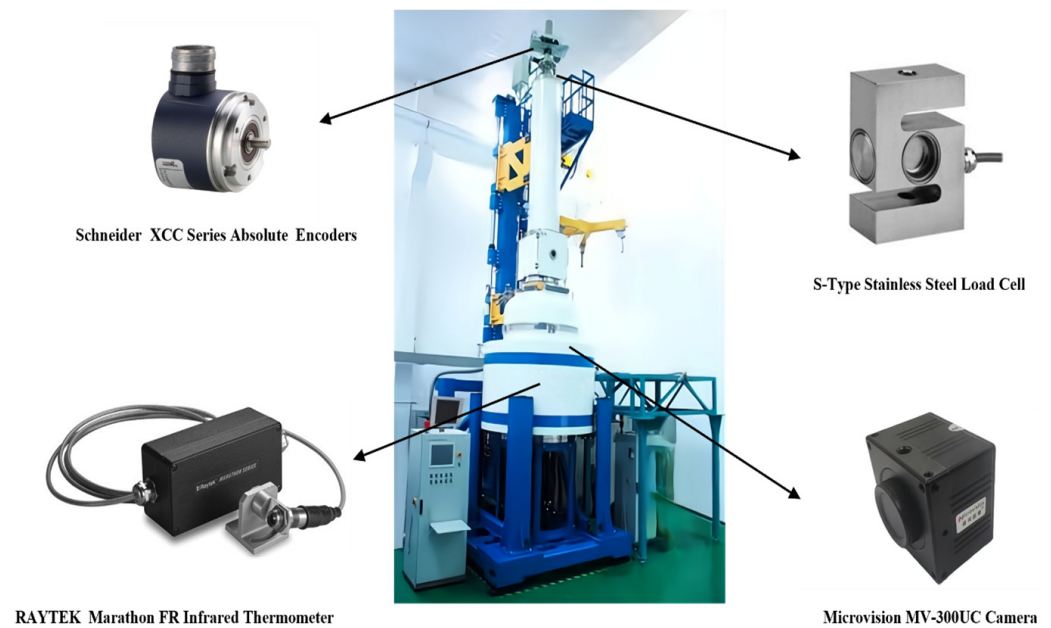


Figure 3. Capture device images.

The crystal diameter signal used for the data fusion was acquired using a Microvision MV-300UC Camera. The temperature signal in the growing process of the crystal was obtained through a RAYTEK Marathon FR Infrared Thermometer sensor. The pulling speed and tail length of the crystal were acquired using Schneider XCC Series Absolute Encoders. The mass increment of the crystal was obtained using an S-Type Stainless Steel Load Cell. The data of the temperature, pulling speed, tail length, and mass increment were sampled every 2 s. All signal-acquisition devices are shown in Figure 3.

3. Overall Architecture of the Proposed Method

The proposed prediction method development process is shown in Figure 4. It includes some stages. The first stage is the processing of input data. For the field data measured by crystal-growing equipment on site, the grey correlation analysis (GRA) method [11] is used to perform correlation analysis on data with a significant impact on the tail diameter. For the temperature and pulling speed, the MIV algorithm [12] is used to select delaying orders that have a significant impact on the tail diameter, and then the data are randomly divided into training and testing datasets. The second stage is the establishment of a prediction model. The deep belief network is employed to extract the feature information in the data. Then, the extracted feature information is used as an input of the SVR. The ant colony optimization (ACO) algorithm [13] is employed to automatically select the optimal parameter of the SVR, and apply it to the soft measurement model. The final stage is to predict the tail diameter based on the DBN-ACO-SVR model, compare the predicted results with those of other methods, and validate the data of different lots based on this model.

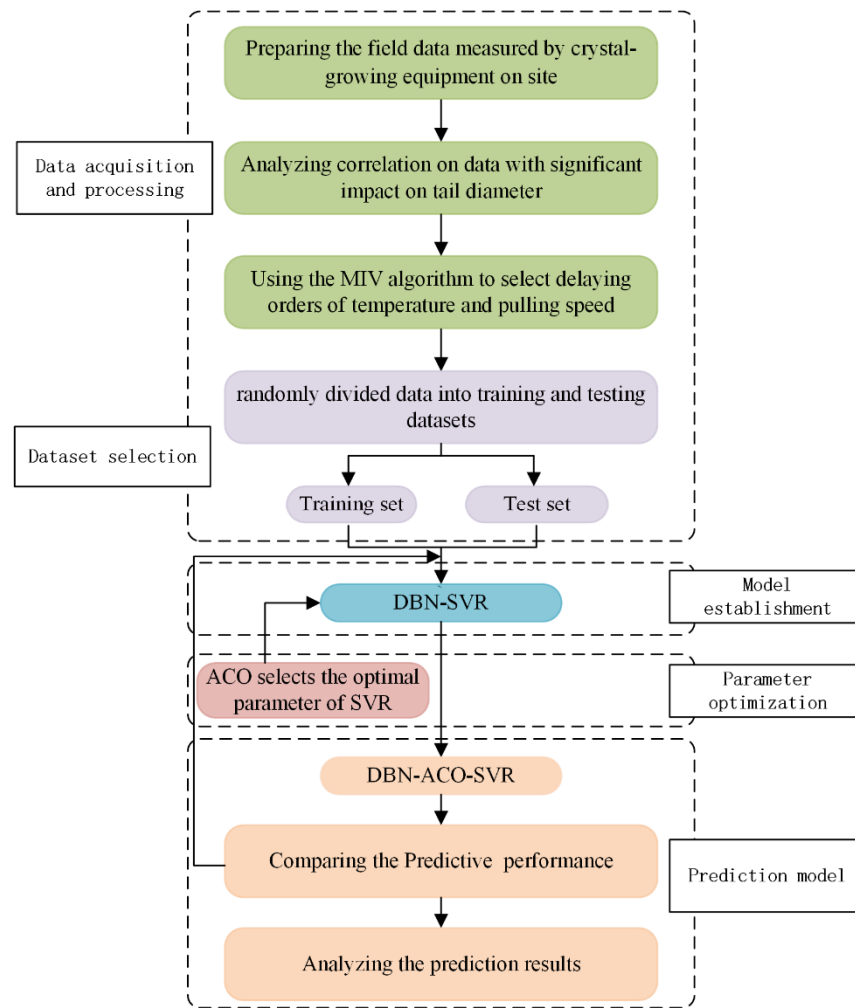


Figure 4. The overall architecture of the proposed soft measurement method.

In the tail process of Czochralski crystal growth, the variables affecting the crystal diameter include temperature, pulling speed, etc. Other parameters such as the crystal rotation speed and crucible rotation speed are mostly kept constant at this stage. Therefore, in order to ensure that the data used in modeling reflect the actual growth status of the tail diameter as much as possible, it is necessary to adopt a method to determine the signals associated with the tail diameter. This article adopts the grey correlation analysis method, which refers to the geometric comparison of data sequences that reflect the changing characteristics of various factors. In the gray correlation analysis method, the reference series and the comparison sequence are respectively $X_0 = [X_0(k)|k = 0, 1, 2, \dots, n]$ and $X_i = [X_i(k)|k = 0, 1, 2, \dots, n]$. The calculation formula for the correlation degree of each series is Formula (1):

$$r_i = \frac{1}{n} \sum_{k=1}^n \frac{\min_{i=1, n} \min_{k=1, n} |X_0(k) - X_i(k)| + \rho \max_{i=1, n} \max_{k=1, n} |X_0(k) - X_i(k)|}{|X_0(k) - X_i(k)| + \rho \max_{i=1, n} \max_{k=1, n} |X_0(k) - X_i(k)|} \quad (1)$$

where ρ is the resolution coefficient, the value range is 0–1, and the general value is 0.5.

When predicting the tail crystal diameter, due to the delaying time of the temperature and pulling speed, the current crystal diameter cannot be directly trained and predicted by using the temperature and pulling speed at the current moment. It is necessary to determine the delaying order of the temperature and pulling speed on the diameter.

In this paper, the MIV algorithm [14] is used to choose the significant input items from the temperature and pulling speed. Input variables that significantly affect the output

are retained, and those that do not play a significant role are excluded. Finally, the above temperature and pulling speed input items are used to become the DBN-ACO-SVR network inputs. The delaying of temperature and pulling speed needs to be analyzed independently. Taking the temperature delaying time as an example, we establish a three-layer multi-input and single-output temperature-diameter dynamic neural network model [14]. The model is shown in Figure 5, and can be expressed as Formula (2):

$$D(k) = f(T(k), T(k-1), \dots, T(k-n_{max})) \quad (2)$$

where $D(k)$ is the crystal diameter at moment k , $T(k-i)$ is the temperature at moment $k-i$, n_{max} is the maximum delaying order of the temperature input, and $f(\cdot)$ is a nonlinear function.

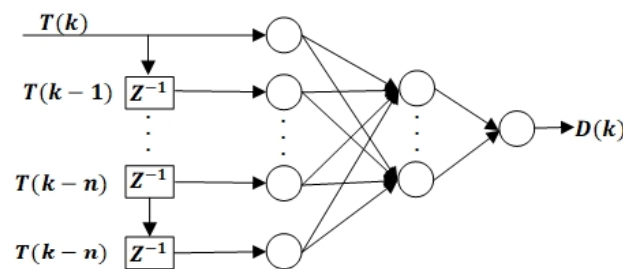


Figure 5. Three-layer multi-input and single-output temperature-diameter dynamic neural network model.

First, the neural network shown in Figure 5 is trained by using the dataset. When the accuracy of the training results is satisfied, the network training parameters are saved. Second, the original input values of $T(k)$ are increased to $T(k-i)$ by 10% and decreased by 10% one by one, and the original values of the other input items are kept unchanged at the same time. This results in new i datasets. The new dataset is substituted into the trained network to obtain network outputs D_{i+} and D_{i-} . Finally, we use D_{i+} and D_{i-} to make a difference, and the result is the impact value (IV). The MIV of this input is obtained by calculating the average of IV. The $|MIV|$ corresponding to the i -th input item is as follows:

$$MIV = \frac{1}{M} \sum_{j=1}^M (D_{i+} - D_{i-}) \quad (3)$$

All $|MIV|$ from $T(k)$ to $T(k-i)$ are calculated. The larger $|MIV|$ is, the greater the impact of this input on the output $D(k)$ of the network. δ_t is the temperature threshold. Finally, we select all inputs $T(k-n)$ that satisfy $|MIV| > \delta_t$, where n is the delaying order. Similarly, the delaying order m of the pulling speed is also obtained by this method.

After determining the delaying order of the influence of the temperature and pulling speed on the crystal tail diameter, a neural network model is established to predict the diameter. The model also needs to introduce the crystal length and the crystal mass as input variables. The tail length is introduced considering that the thermal radiation will change as the crystal length increases [3], and the introduction of a mass increment indirectly reflects the change in diameter. The model can be expressed using Formula (4):

$$D(k) = F(T(k-l), \dots, T(k-n), V(k-p), \dots, V(k-m), L(k), M(k)) \quad (4)$$

where $T(k-l), \dots, T(k-n)$ are satisfied temperature inputs that meet $|MIV| > \delta_t$, $V(k-p), \dots, V(k-m)$ are satisfied pulling speed inputs that meet $|MIV| > \delta_v$, δ_v is the pulling speed threshold, $L(k)$ is the crystal tail length at the time of k , $M(k)$ is the mass increment from the time $k-1$ to k , and $F(\cdot)$ is a nonlinear function.

To improve the accuracy of the tail diameter prediction, it is necessary to fully extract the characteristic information in the data. Although several soft measurement models based on deep learning can achieve this goal, those models still have limitations. For

example, in the case of a large amount of data, complex physical structures, and increased dimensions, the SVR modeling method is generally inefficient [15,16]. Also, this method only uses a shallow network for training rather than repeatedly training through a multi-layer network, so it cannot extract the characteristic information in the data, which will affect the model's prediction accuracy. In addition, the BP neural network has a higher efficiency in a single hidden layer. When there are multiple hidden layers, the efficiency of back propagation will be greatly reduced. Therefore, this paper uses a DBN with multiple hidden layers for soft measurement modeling [17,18]. A DBN can learn complex data features by combining multi-layer networks to realize multi-level representation learning and represent the implicit information of data, thus making the prediction more accurate.

The DBN is composed of multiple restricted Boltzmann machines (RBM) [19]. The structure of the DBN model is shown in Figure 6. An RBM visible layer is used to describe the features of the input data, and the hidden layer can be regarded as a feature-extraction layer.

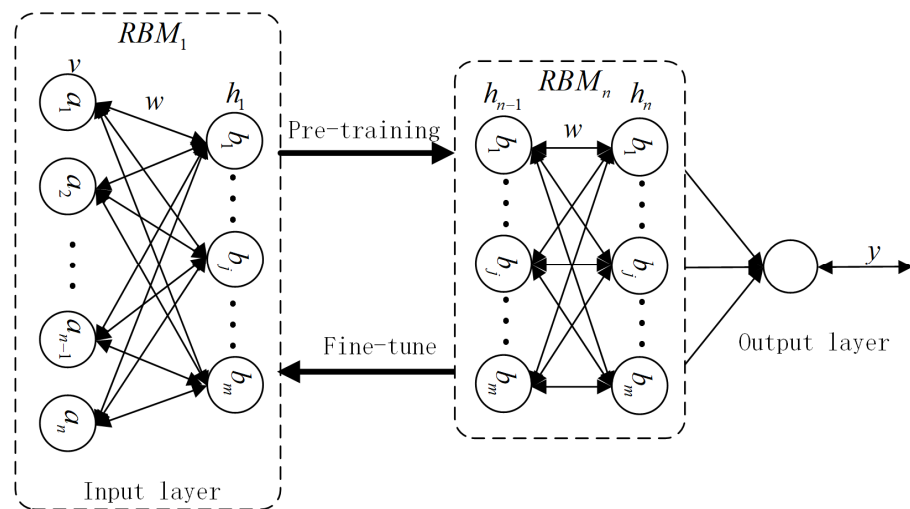


Figure 6. Structure of the DBN model.

The visible layer unit is $v = \{v_1, v_2, \dots, v_i\}$, the hidden layer unit is $h = \{h_1, h_2, \dots, h_j\}$, and the internal parameter set is $\theta = \{w, a, b\}$, and then the probability distribution based on the energy probability model is as follows:

$$P(v, h; \theta) = \frac{1}{Z(\theta)} e^{-E(v, h; \theta)} \quad (5)$$

where Z is the regularization factor and $Z(\theta) = \sum_v \sum_h e^{-E(v, h; \theta)}$. Therefore, the energy function of RBM is as follows:

$$E(v, h) = -\sum_i a_i v_i - \sum_j b_j h_j - \sum_{i,j} w_{ij} v_i h_j \quad (6)$$

where a_i and b_j are visible layer offsets and hidden layer offsets, respectively, and w_{ij} are the connection weights between the visible layer v_i and hidden layer h_j .

The joint probability distribution of visible layer v is as follows:

$$P(v; \theta) = \sum_h \frac{e^{-E(v, h; \theta)}}{\sum_{v, h} e^{-E(v, h; \theta)}} = \frac{1}{Z(\theta)} e^{(b^T v)} \prod_{j=1}^F \left(a + \exp \left(a_j + \sum_{i=1}^D w_{ij} v_i \right) \right) \quad (7)$$

When the input signal of the visible layer is given, the activation conditions of each node of the hidden layer are independent; conversely, when the hidden layer condition is determined, the activation conditions of each node of the visible layer are also independent.

Therefore, according to the above formula, when the input condition and visible layer v are known, the access probability of hidden layer h is as follows:

$$P(v|h;\theta) = \prod_i P(v_i|h) \quad (8)$$

Similarly, when the output data of the hidden layer and the output unit vector h are known, the output probability of the visible layer v is as follows:

$$P(h|v;\theta) = \prod_j P(h_j|v) \quad (9)$$

Taking the activation function Sigmoid function as an example, the activation probability of h and v is obtained as follows:

$$P(h_j = 1|v) = \text{sigmoid}\left(a_j + \sum_i v_i w_{ij}\right) \quad (10)$$

$$P(v_i = 1|h) = \text{sigmoid}\left(b_i + \sum_j h_j w_{ji}\right) \quad (11)$$

The DBN network is obtained by stacking multiple RBMs and a BP layer. Then, the achieved network is pre-trained and adjusted accurately layer by layer. The unsupervised greedy learning algorithm and the contrast divergence algorithm are used to complete the pre-training layer by layer. Afterwards, the gradient descent algorithm is used to complete the accurate adjustment of the network. Considering the weak nonlinear regression ability of the DBN, this paper replaces the top structure of the DBN with SVR to strengthen the nonlinear regression ability and uses the data features extracted by the DBN as the prediction input of the SVR model. Additionally, the ACO is used to select the appropriate parameters for SVR, and the SVR algorithm is used to fine-tune the entire network. The structure of the proposed soft measurement model named DBN-ACO-SVR is shown in Figure 7.

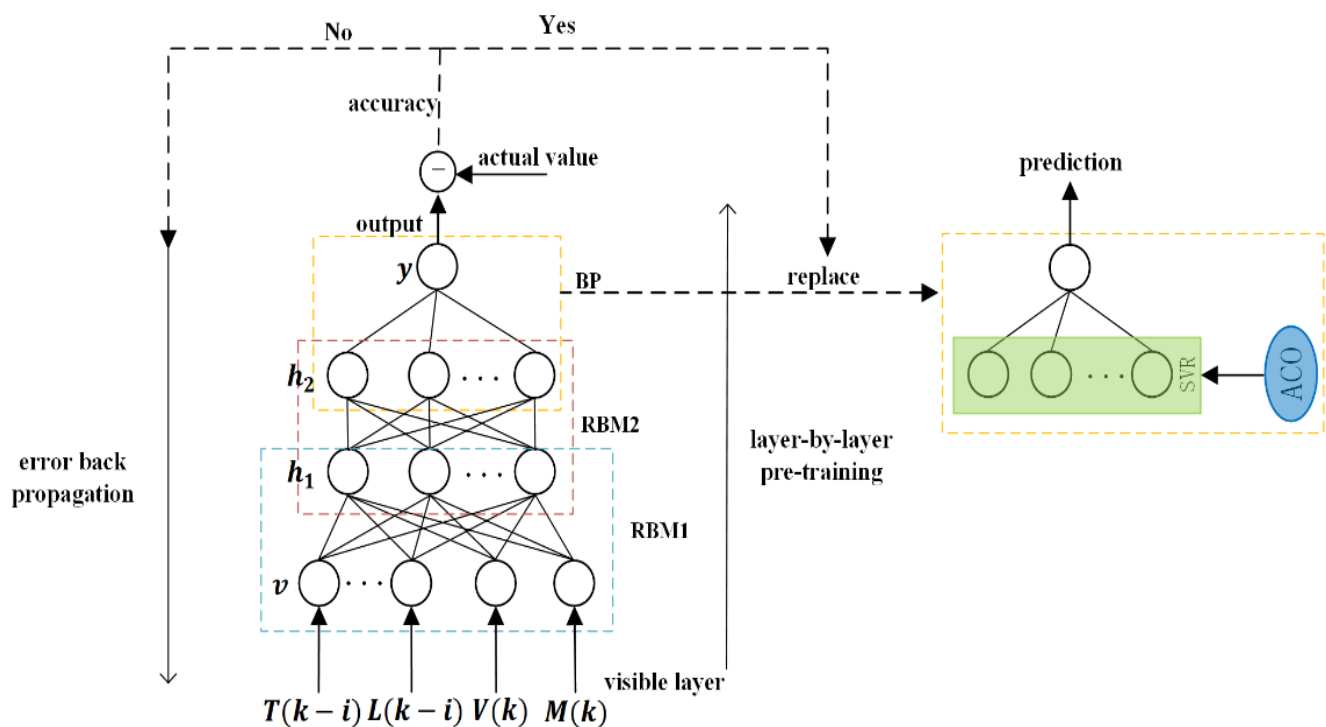


Figure 7. The structure of the proposed soft measurement model.

For the training set sample $\{(x_1, y_1), (x_2, y_2), \dots, (x_m, y_m)\}$ of the SVR model, we hope to obtain a regression model to make $f(x)$ and y as close as possible. w and b are the regression parameters to be determined. Assuming that the deviation between $f(x)$ and y can be tolerated up to ε , the SVR problem is finally converted as shown in Equation (12):

$$\min_{w,b} \frac{1}{2} \|w\|^2 + C \sum_i^m l_e(f(x_i), y_i) \quad (12)$$

where C is the regularization parameter and l_e is the insensitive loss function, as shown in Equation (13):

$$l_e = \begin{cases} 0 & \text{if } |f(x) - y| < \varepsilon \\ |f(x) - y| - \varepsilon & \text{else} \end{cases} \quad (13)$$

The final SVR solution is obtained as follows:

$$f(x) = \sum_i^m (\alpha_i^* - \alpha_i) x_i^T x + b \quad (14)$$

Then, the kernel function is introduced:

$$k(x_i, x_j) = \varphi(x_i)^T \varphi(x_j) \quad (15)$$

In this manner, the SVR solution can be rewritten as follows:

$$f(x) = \sum_i^m (\alpha_i^* - \alpha_i) k(x_i^T x) + b \quad (16)$$

4. Experimental Results and Analysis

Grey correlation analysis was conducted using data obtained from industrial sites. The correlation analysis method was used to analyze the temperature and casting speed variables that are closely related to the tail diameter. The correlation between the two comparison sequences and the tail diameter is shown in Figure 8.

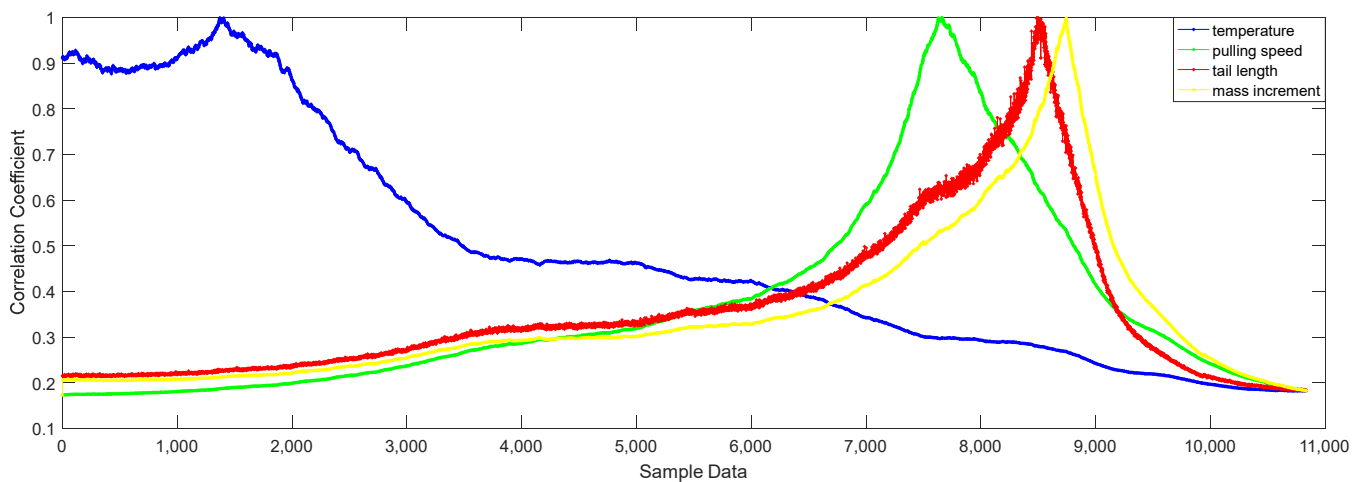


Figure 8. The correlation between the two comparison sequences and the tail diameter.

As shown in Figure 8, the temperature, pulling speed, tail length, and mass increment have a high correlation with the tail diameter in the tail process. The mean correlation coefficients of the temperature, pulling speed, tail length, and mass increment are 0.4896, 0.3606, 0.3588, and 0.3452, respectively.

The growth process data of 300 mm monocrystalline silicon were obtained for the Czochralski furnace of the Crystal Growth Equipment and System Integration National & Local Joint Engineering Research Center. These data were used to validate the soft measurement method. According to experience, the delay time of the temperature is

about 10~13 min, and the delay time of the pulling speed is less than 1 min. The preset values of the temperature delaying order n , the pulling speed delaying order m , and the maximum delaying time all need to be satisfied. The data-sampling frequency is 0.2 Hz, so let $n = 100, m = 20$. Taking the delaying order identification of temperature as an example, the inputs of the neural network in the MIV algorithm are $[T(k - 301), T(k - 302), \dots, T(k - 400)]$, and the number of neurons in the hidden layer is $\sqrt{N_i + N_o} + a$, where N_i is the number of inputs and N_o is the number of outputs. a is between 1 and 10, so let $a = 10$. The Levenberg–Marquardt algorithm is used to obtain the relationship between the delaying order of the temperature and pulling speed and the absolute value of the MIV after 1000 iterations. The relationship diagram is shown in Figure 9.

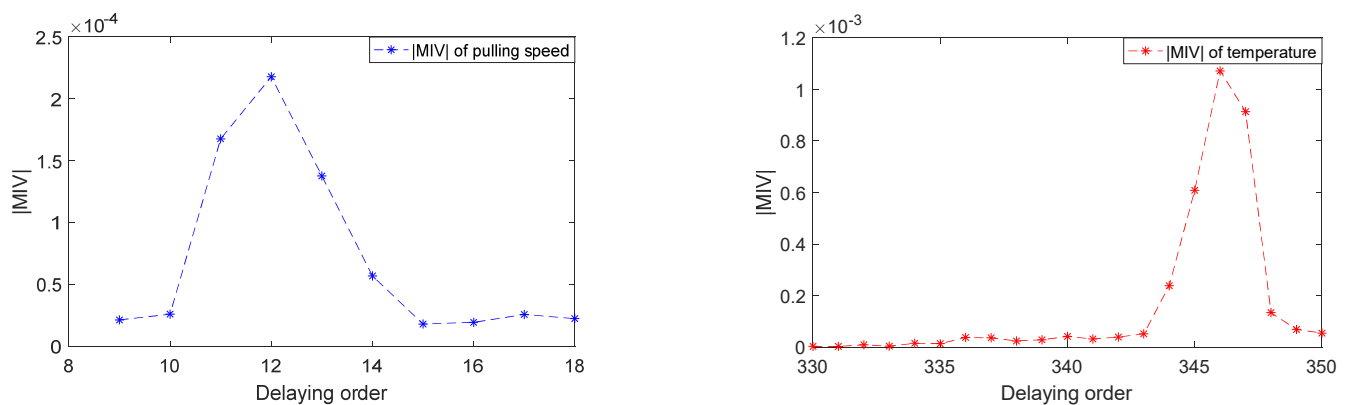


Figure 9. Delaying order and the $|MIV|$ curve of the pulling speed and temperature.

The delaying order maximum MIV counterparts of the effect of the temperature and pulling speed on the crystal diameter were $T(k - 346)$ and $V(k - 12)$. In order to prevent the omission of other input items, the final selection of the temperature input item was $T(k - 345), T(k - 346), T(k - 347)$ and the pulling speed input item was $V(k - 11), V(k - 12), V(k - 13)$.

After the delaying order of the temperature was obtained, the DBN-ACO-SVR network was constructed. The network model can be expressed by Formula (17):

$$D(k) = F(T(k - 345), T(k - 346), T(k - 347), V(k - 11), V(k - 12), V(k - 13), L(k), M(k)) \quad (17)$$

From the above formula, the number of input nodes in the network was 8, and the number of output nodes was 1.

We determined the number of layers and nodes of the DBN through continuous experimental verification. The number of network layers of the DBN was set as four layers, and the number of RBMs was two. The number of nodes in each layer was selected as 8-20-6-1. The learning rate was set as 0.02, and the activation function was selected as Sigmoid. The error results of each RBM iteration and prediction error results of the DBN iteration are shown in Figure 10.

ACO is used to select the optimal parameter of SVR and apply it to soft sensing models. The results of the iterative optimization after processing 200 iterations are shown in Figure 11. We substituted the optimal parameter of the SVR algorithm obtained from the ACO algorithm into the DBN model to obtain a new soft measurement model, DBN-ACO-SVR, and predict the crystal tail diameter using this model.

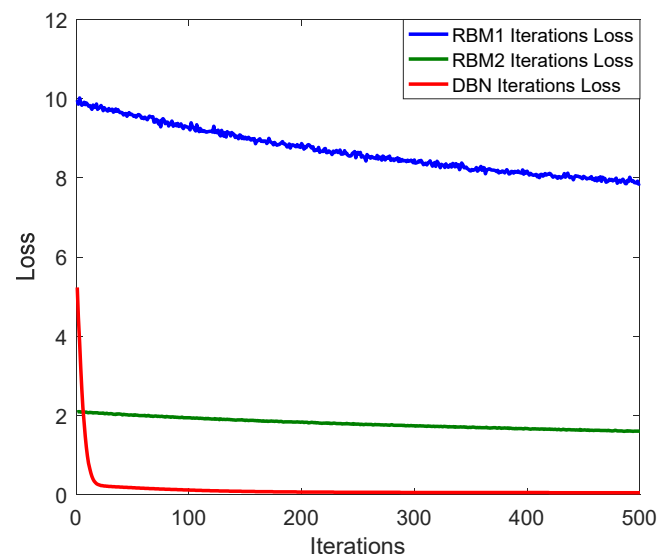


Figure 10. The error results of the DBN iteration.

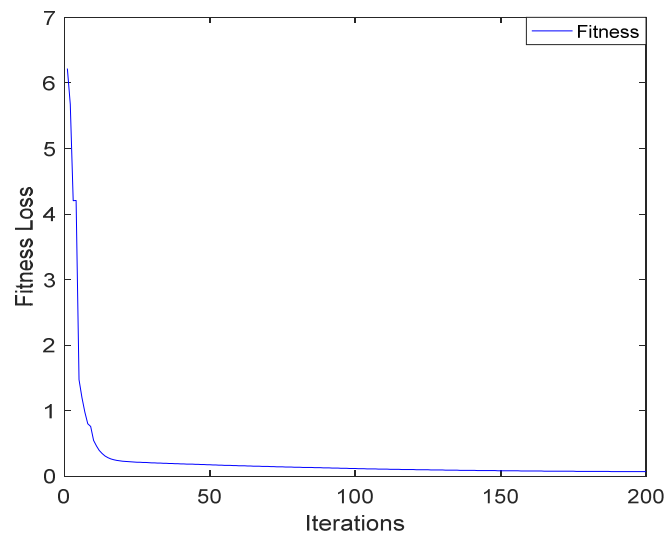


Figure 11. Iterative search results graph.

This paper proposes a soft measurement method for the tail diameter of the DBN-ACO-SVR model. The training dataset of the verification conclusion in this paper includes two kinds of datasets. In the first dataset, the basic processing parameters are the same, and the control processing parameters are different. Different temperatures and pulling speeds produce different tail diameters and mass increments. In the second dataset, some basic processing parameters and control processing parameters are different.

For the results in Figure 12, this paper uses the dataset with the same basic processing parameters for prediction. For the results of mass estimation, the predicted results deviate significantly from the actual results. The maximum error is tens of millimeters. One reason is that mass data are affected by changes in motor torque while reflecting changes in crystal mass. Another reason is that the accuracy of the weighing signal is also restricted by the accuracy and sensitivity of the weighing sensor itself.

In Figure 12, we establish a prediction model with mass input and a prediction model without mass input. The same dataset is used for network training at the same time. Then, we use the trained network model to predict the crystal tail diameter. The effect diagram and error diagram between the two prediction results and the real value are shown in Figure 12. By comparing the results of two prediction models, the method DBN-ACO-SVR

considers the weighing signal, which can significantly reduce the prediction error, which is less than 10 mm. It can be used for the prediction and control of the tail process.

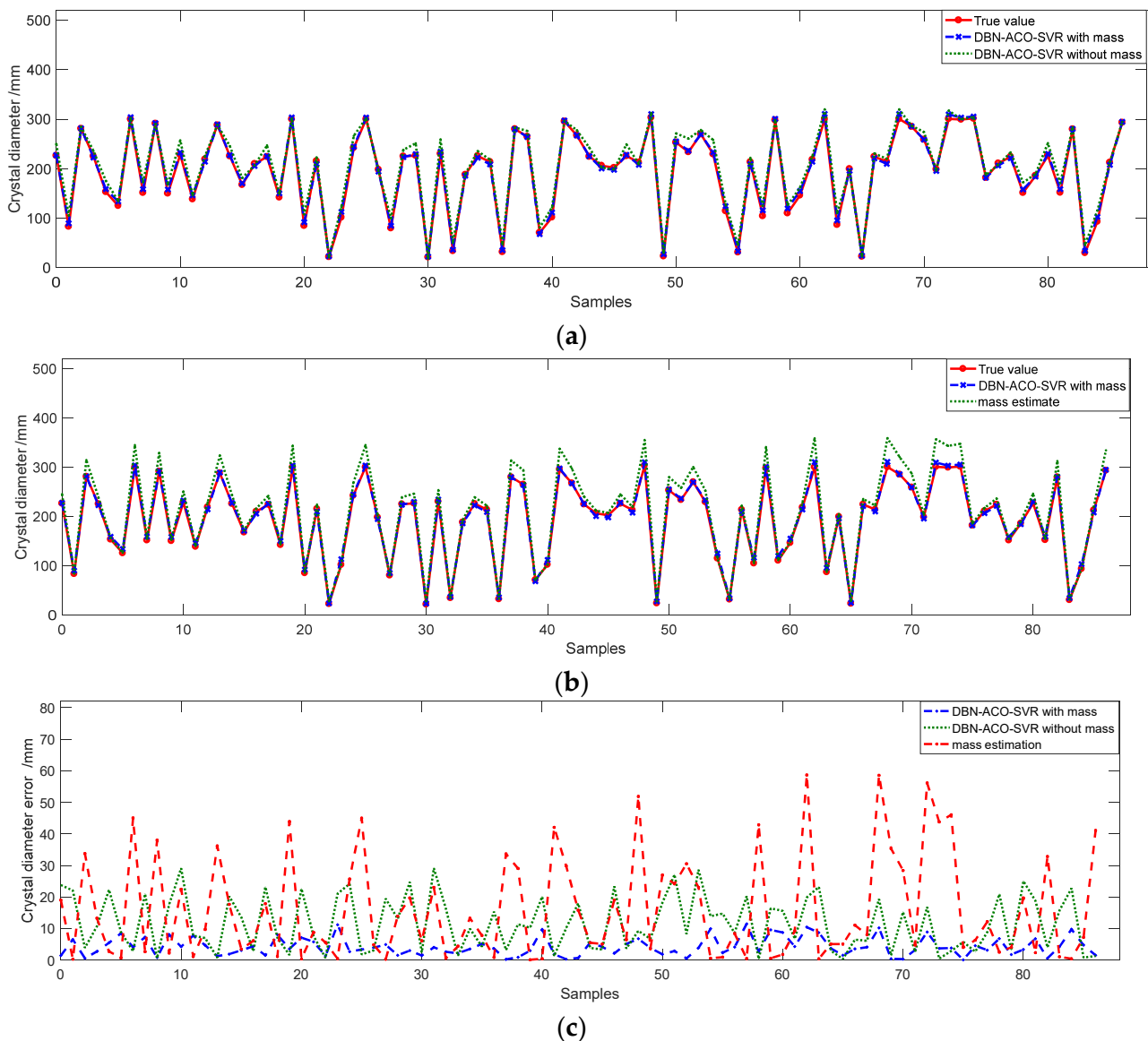


Figure 12. The results of different methods based on the same dataset. (a) Comparison of prediction results with or without mass. (b) Comparison of prediction results with mass. (c) Comparison of prediction errors.

For the results in Figure 13, the training set and test set used in this paper are not only different in control processing parameters, but also different in some basic processing parameters. Because the difference of basic processing parameters is not reflected in the prediction model, the prediction results become worse, which is still smaller than the error of directly using the mass signal to predict the tail diameter. The results show that the DBN-ACO-SVR prediction model with parameter identification has a low error and better prediction capability.

To better visualize the superior performance of the soft measurement model based on DBN-ACO-SVR in predicting the crystal tail diameter, we compared a soft measurement model combined with DBN and SVR with a DBN-ACO-SVR synthetic model. The test set, according to the changing trend of crystal length, was selected for prediction. The final experimental results based on different models are shown in Figure 14.

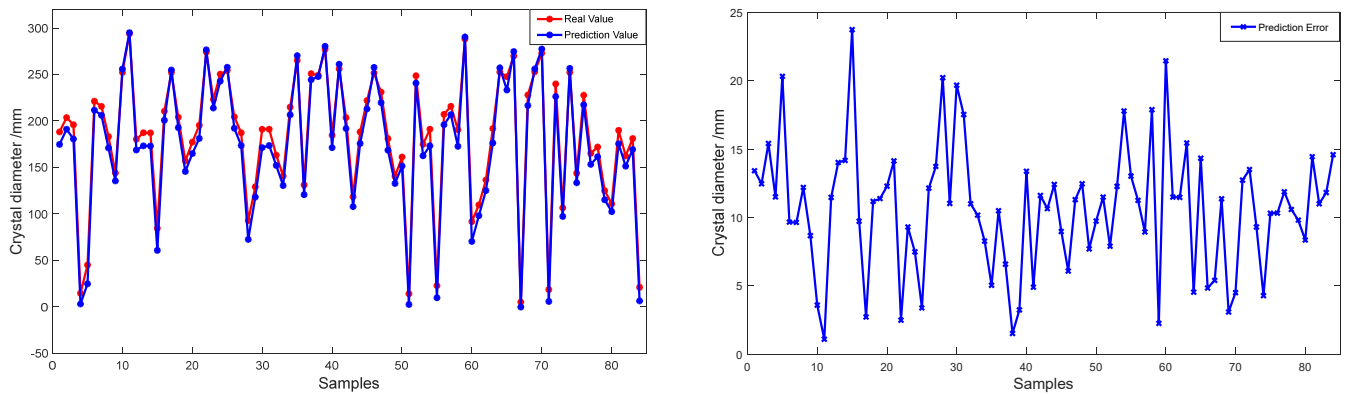


Figure 13. The error of the DBN-ACO-SVR prediction model with parameter identification based on the different dataset.

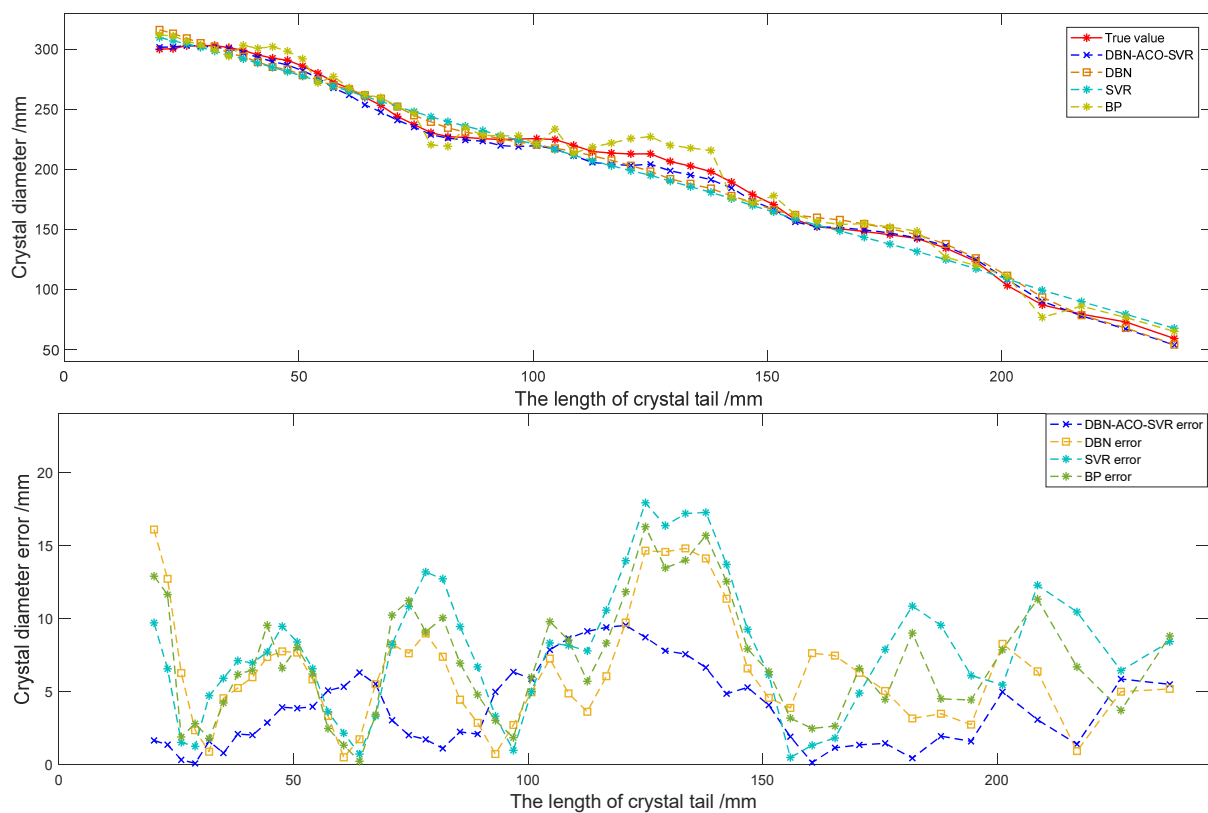


Figure 14. The results of different network models based on the same dataset.

The root mean square error (RMSE) is a common index used to measure the difference between the predicted value and the actual observation value of the model. It calculates the mean value of the square of the difference between the predicted value and the actual value, and then takes its square root to obtain a measure of the overall error. A smaller RMSE value usually indicates that the model has better prediction accuracy, while a larger RMSE value indicates that the prediction performance of the model is poor [20]. DBN-ACO-SVR, DBN, and SVR were used to predict the tail diameters of crystals in different lots. The RMSE of the predicted results is shown in Table 1.

The simulation results show that the accuracy of the DBN-ACO-SVR prediction model is better than that of other prediction models, and the model's prediction accuracy is higher.

Table 1. RMSE of prediction results based on different networks for different lots.

	DBN-ACO-SVR	DBN	SVR	BP	Mass Estimate
Lot #1	4.7692	7.4286	8.9509	8.3634	23.1228
Lot #2	5.5249	7.8367	10.5931	9.7127	25.3933
Lot #3	4.5077	4.9082	7.2256	6.8695	20.2498
Lot #4	5.1634	6.4839	8.5635	7.9351	22.4974
Lot #5	5.0157	6.7349	9.0331	8.9334	21.4200
Lot #6	4.4461	5.4053	7.5639	8.7108	20.0879
Lot #7	4.7147	6.5398	8.4546	9.5147	22.9657
Lot #8	5.5660	6.0257	8.4222	7.2050	25.8389

5. Conclusions

In addressing the challenge of diameter measurement during the tail process of Czochralski silicon single crystal growth, this study introduced a soft measurement model, DBN-ACO-SVR, for accurately determining the diameter of the crystal. The proposed model is built based on DBN, which intricately integrates key parameters such as temperature, crystal pulling speed, crystal mass, and crystal length. A BP neural network, informed by the MIV method, is adeptly utilized for the identification of system parameters, which is pivotal in determining the latency in the impact of the temperature and pulling speed on the crystal diameter. Furthermore, to enhance the regression capabilities of the model, the ACO is employed for fine-tuning the SVR network's parameters. The efficacy of the proposed model was rigorously validated using actual data from crystal growth processes, demonstrating its effectiveness in the soft measurement of the crystal diameter during the tail process. Comparative analysis with existing methods, including standalone DBN and SVR, underscored the superiority of the proposed methodology. The results distinctly highlight the model's advanced capabilities in providing precise diameter measurements, affirming its potential as a robust tool in the field of silicon single crystal manufacturing.

Author Contributions: L.J., D.T. and Y.Z. designed the work, analyzed the data, and completed the simulation. L.J. and D.T. wrote the first manuscript and everyone participated in the revision. All authors contributed to the article. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Key Research and Development Project of Shaanxi Province of China (No.2022QCY-LL-27).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The author will provide raw data supporting the conclusion of this paper without reservation.

Conflicts of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. Feigelson, R.S. Crystal growth History: Theory and melt growth processes. *J. Cryst. Growth* **2022**, *594*, 126800. [[CrossRef](#)]
2. Ren, J.C.; Liu, D.; Wan, Y. Modeling and application of Czochralski silicon single crystal growth process using hybrid model of data-driven and mechanism-based methodologies. *J. Process. Control* **2021**, *104*, 74–85. [[CrossRef](#)]
3. Zhao, Y.; Liu, D.; Chen, D.W.; Wang, L. Study on the tail growth control parameters of Czochralski Silicon based on the numerical simulation. *J. Synth. Cryst.* **2010**, *39*, 208–213.
4. Wang, X.Y.; Xiang, S.W.; Xiang, K.; Pan, F. A novel method for diameter measurement of silicon single crystal. *Measurement* **2018**, *121*, 286–293. [[CrossRef](#)]
5. Qi, X.F.; Ma, W.C.; Dang, Y.F.; Su, W.J.; Liu, L.J. Optimization of the melt/crystal interface shape and oxygen concentration during the Czochralski silicon crystal growth process using an artificial neural network and a genetic algorithm. *J. Cryst. Growth* **2020**, *548*, 125828. [[CrossRef](#)]

6. Zhang, G.; Liang, W.G.; She, B.; Tian, F.Q. Soft Measurement Method for Health State of Typical Complex Feeding and Ramming Mechanism Based on Time-frequency Data Drive. *Acta Armamentarii* **2022**, *43*, 737–747.
7. Li, W.J.; Li, M.; Qiao, J.F. Effluent BOD soft measurement based on mutual information and self-organizing RBF neural network. *CIESC J.* **2019**, *70*, 687–695.
8. Zhong, D.Q.; Wang, A.L.; He, Q. Soft measurement model on torque of alternating current electrical dynamometer including copper loss and iron loss. *J. Cent. South Univ.* **2019**, *26*, 2272–2280. [[CrossRef](#)]
9. Wang, J.L.; Liu, X.R.; Su, Y.Y.; Ye, W. A real-time soft measurement method based on the furnace system parameters of the forecast model. In Proceedings of the 2013 IEEE 12th International Conference on Cognitive Informatics and Cognitive Computing, New York, NY, USA, 16–18 July 2013; pp. 461–466.
10. Cao, J.; Gao, L.M. MAGA-WNN soft measurement model in application of gelatin concentration. *Comput. Eng. Appl.* **2012**, *48*, 231–234.
11. Qiu, M.Y.; Wu, Z.R.; Zhang, Y.; Ma, Y.H.; Ma, F.; Li, X.X. Analysis of Road Performance of Modified Asphalt Mixture of Waste Cigarette Butts Based on Grey Correlation Method. In Proceedings of the 2021 7th International Conference on Hydraulic and Civil Engineering & Smart Water Conservancy and Intelligent Disaster Reduction Forum (ICHCE & SWIDR), Nanjing, China, 6–8 November 2021; pp. 1274–1296.
12. Zhang, Y.; Wang, Z.Y. A hybrid model for blood pressure prediction from a PPG signal based on MIV and GA-BP neural network. In Proceedings of the 2017 13th International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery (ICNC-FSKD), Guilin, China, 29–31 July 2017; pp. 1989–1993.
13. Armond, A.M.; Prasetyo, Y.D.; Ediningrum, W. Application of Ant Colony Optimization (ACO) Algorithm to Optimize Trans Banyumas Bus Routes. In Proceedings of the 2022 IEEE International Conference on Cybernetics and Computational Intelligence (CyberneticsCom), Malang, Indonesia, 16–18 June 2022; pp. 132–137.
14. Zeng, J.; Zhang, C.; Xie, N.; Yang, P.; Xu, C.Y.; Zhang, Z. A hybrid model for short-term wind power forecasting based on MIV, Tversky model and GA-BP neural network. In Proceedings of the 2016 the 3rd International Conference on Manufacturing and Industrial Technologies, Online, 11 August 2016; Volume 70, p. 10002.
15. Zou, J.F.; Li, C.L.; Yang, Q.; Li, Q. Fault prediction method based on SVR of improved PSO. In Proceedings of the 27th Chinese Control and Decision Conference (2015 CCDC), Qingdao, China, 23–25 May 2015; pp. 1671–1675.
16. Ji, Y.J.; Xu, K.P.; Zeng, P.; Zhang, W.F. GA-SVR algorithm for improving forest above ground biomass estimation using SAR data. *IEEE J. Stars* **2021**, *14*, 6585–6595. [[CrossRef](#)]
17. An, Y.; Wang, Y.H.; Liu, Y.X. Research on gas pressure regulator fault diagnosis based on deep confidence network (DBN) theory. In Proceedings of the 2017 Chinese Automation Congress (CAC), Jinan, China, 20–22 October 2017; pp. 3821–3825.
18. Gao, Y.; Zhang, B.; Wang, S.; Ma, A.X. DBN based cloud service response time prediction method. In Proceedings of the 2019 21st International Conference on Advanced Communication Technology (ICACT), PyeongChang, Republic of Korea, 17–20 February 2019; pp. 42–46.
19. Hinton, G.E.; Osindero, S.; The, Y.W. A fast learning algorithm for deep belief nets. *Neural Comput.* **2006**, *18*, 1527. [[CrossRef](#)] [[PubMed](#)]
20. Isham, M.I.M.; Haron, H.N.H.; Mohamed, F.B.; Siang, C.V. VR Welding Kit: Accuracy Comparison Between Smartphone VR and Standalone VR Using RMSE. In Proceedings of the 2021 IEEE International Conference on Computing (ICOCO), Kuala Lumpur, Malaysia, 17–19 November 2021; pp. 341–346.

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.