Predictive Modeling of Ball Mill Load Parameters Using Hybrid Physics-Informed Neural Networks

1st Yiwen Liu
College of Electrical and Power engineering
Taiyuan University of Technology
Taiyuan, China
liuyiwen0716@163.com

2nd Gaowei Yan*
College of Electrical and Power engineering
Taiyuan University of Technology
Taiyuan, China
yangaowei@tyut.edu.cn

3rd Yusong Pang
Section of Transport Engineering and Logistics
Delft University of Technology
Delft, Netherlands
y.pang@tudelft.nl

***Abstract*—Wet ball mill plays a key role in the grinding process, and its load state directly affects production efficiency, energy consumption and product quality. Aiming at the problems of poor interpretability of pure data-driven models and complex modeling of mechanism models under variable working conditions, a hybrid prediction model DAPINN combining deep learning and physical information is proposed. By introducing the deep hidden physics model principle and using the characteristics of neural networks to approximate arbitrary functions to simulate complex physical partial differential equations, the physical interpretability of the model is enhanced. At the same time, the model introduces domain adaptation technology to improve the prediction accuracy and generalization of the model under variable working conditions. Experiments were conducted on data collected from a small ball mill in the laboratory. The experimental results show that under variable working conditions, the prediction accuracy of the DAPINN model is better than that of the pure data-driven model.**

***Keywords*—Mill load parameters prediction, PINN, domain adaptation, deepHPM**

I. Introduction

Wet ball mills are widely used in the grinding process, mainly for ore grinding [1]. In the mineral processing process, wet ball mills grind ore into finer particles by means of the collision and friction of grinding media [2]. The working performance of the ball mill, especially its load state, directly affects production efficiency, energy consumption and product quality [3][4]. Therefore, accurately predicting the load parameters of wet ball mills is of great significance for improving industrial production efficiency, reducing energy consumption and ensuring product quality.

The load parameters of the wet ball mill usually include the material-to-ball volume ratio (MBVR), pulp density (PD) [5]. The changes in these parameters are closely related to the working efficiency and product quality of the ball mill. For example, too high or too low MBVR and PD will affect the grinding effect of the material, thereby reducing production efficiency or wasting energy [6]. In order to improve the product quality of the mill and reduce energy consumption, accurate mill load parameter prediction is particularly important. However, since the wet ball mill is a large rotating equipment in the industrial process, it is impossible to directly obtain the accurate mill load parameters [7]. Therefore, researchers use soft measurement to establish the mill load parameter prediction model [8][9].

Traditional mill load parameter prediction methods can be divided into two categories: mechanism models [10][11] and pure data-driven models [12][13]. Mechanism models mainly rely on physical laws and understanding of the internal mechanisms of industrial processes to establish models [14][15]. However, the grinding process is a highly nonlinear system involving multiple factors, and it is challenging to establish an accurate mill load parameter model. Compared with mechanism models, data-driven models, especially those based on deep learning, can make predictions by mining hidden information from data [16]. However, the purely data-driven approach may lead to unsatisfactory results [17].

In order to solve the above problems, researchers have begun to explore methods that combine deep learning with physical information in recent years, hoping to enhance the interpretability and accuracy of the model while maintaining the generalization ability of the data-driven model [18]. Among them, physical information neural network (PINN) is an emerging method that uses physical laws as constraints to enable the neural network to learn feature representations that are more in line with physical laws during training [19]. However, the PINN method requires partial differential equations as input, and the partial differential equations of mill load and vibration signals are often difficult to establish, which limits the application of PINN in mill load prediction. In addition, under variable operating conditions, there are differences in data distribution, which will cause the performance of the traditional PINN model to deteriorate under new operating conditions.

To address the above problems, this paper proposes a hybrid prediction model that combines domain adaptation and deep hidden physics model (DAPINN). The model draws on the principle of deep hidden physics model (deepHPM) [20] and uses the characteristics of neural networks to approximate arbitrary functions to simulate complex physical partial differential equations, thereby avoiding the difficulty of directly establishing partial differential equations. At the same time, the model also introduces domain adaptation technology to improve the accuracy of the model under variable working conditions.

II. Grinding process

Wet ball mill plays an important role in the grinding process. Its input mainly includes raw ore, steel ball and water, and its output is ore slurry. The ore is ground into finer particles mainly through the collision and friction of the grinding media, as shown in Fig. 1.

Mill load is an important parameter that reflects the state of a wet ball mill. However, since the ball mill is a large, closed, continuously rotating device, the mill load cannot be directly measured. In actual practice, the vibration signal of the mill is closely related to the mill state. Therefore, the mill load parameters can be predicted by soft measurement.

This work is supported by the National Natural Science Foundation of China (61973226), and the Key R&D Program Project of Shanxi Province (201903D121143).

However, the relationship between the vibration signal and the mill load parameters is complex, which makes it difficult to establish a mechanism equation that accurately describes the relationship between the two. Therefore, how to apply the traditional PINN model to the mill load parameter prediction task is a problem that needs to be solved.

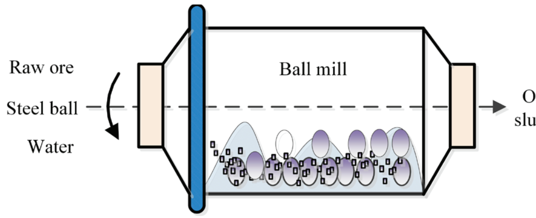


Fig. 1. Ball mill process diagram

III. METHODOLOGY

A. Problem description

Assume that the source domain data is $X^S = \{x_i^S\} \in \mathrm{R}^{N^S \times m}$, which contains the corresponding label $Y^S = \{y_i^S\} \in \mathrm{R}^{N^S \times 1}$. The distribution of the source domain data is P. The target domain data is $X^T = \{x_i^T\} \in \mathrm{R}^{N^T \times m}$, which follows the distribution Q, where $P \neq Q$. N^S and N^T represent the number of samples in the source domain and the number of samples in the target domain respectively, and m represents the feature dimension of the sample.

B. Network Structure

The hybrid model proposed in this paper combines DeepHPM and domain adaptation to improve the prediction accuracy and interpretability of wet ball mill load parameters. The model consists of three parts: feature extraction network, prediction network and physical network. The model framework is shown in Fig. 2.

The feature extraction network is responsible for extracting features useful for mill load parameters from the original input data. The network used is a three-layer convolutional layer. By inputting data into the convolutional layer, the data dimension is reduced while extracting effective information. The features of the source domain and target domain data obtained are shown in equation (1):

$$\begin{cases} Z^S = \mathrm{Conv}_3(\mathrm{Conv}_2(\mathrm{Conv}_1(X^S))) \\ Z^T = \mathrm{Conv}_3(\mathrm{Conv}_2(\mathrm{Conv}_1(X^T))) \end{cases} \tag{1}$$

where Z^S and Z^T represent the features of source domain data and target domain data respectively, and $\mathrm{Conv}_1(\cdot)$, $\mathrm{Conv}_2(\cdot)$, and $\mathrm{Conv}_3(\cdot)$ represent three convolutional layers with different convolution kernel sizes.

The prediction network outputs the predicted value of the mill load parameter based on the features obtained by the feature extraction network. Here, a fully connected neural network is used for prediction. The prediction result can be expressed as equation (2):

$$\begin{cases} \hat{Y}^S = W_p \cdot Z^S + B_p \\ \hat{Y}^T = W_p \cdot Z^T + B_p \end{cases} \tag{2}$$

Where $\hat{Y}^S$ and $\hat{Y}^T$ represent the predicted values of mill load parameters. W_p and B_p represent the weights and biases of the fully connected layer in the prediction network respectively.

The physical network uses a neural network to fit the physical partial differential equation, so that the model not only relies on the data learning model, but also fully follows the actual physical laws. Due to the complexity of the grinding process, it is difficult to establish the partial differential equation of the mill load and vibration signal. Therefore, with the help of the principle of deepHPM, the characteristics of neural networks approximating arbitrary functions can be used to simulate complex physical partial differential equations.

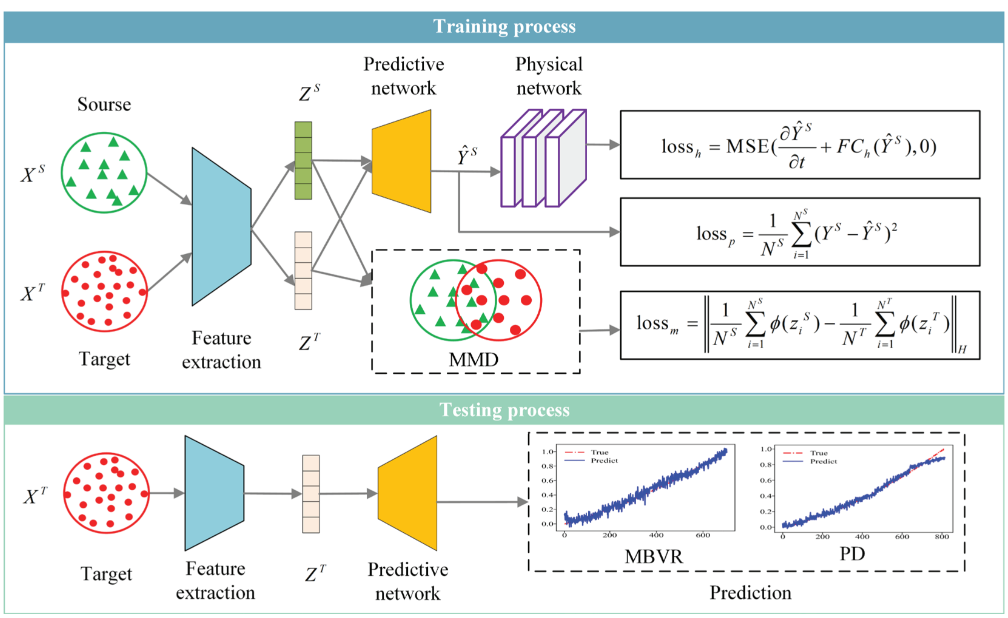


Fig. 2. Model framework diagram

In general, the partial differential equation is defined as equation (3)[21]:

$$u_t + N(u) = 0 \quad (3)$$

where u_t represents the partial derivative of u with respect to time, and $N(\cdot)$ represents a nonlinear operation.

For the convenience of analysis, equation (3) is rewritten as equation (4):

$$f = u_t + N(u) \quad (4)$$

where f represents the residual of the partial differential equation.

In the physical network, a neural network can be used as a nonlinear operator $N(\cdot)$. By optimizing the neural network, the residual f tends to zero, thereby accurately simulating the physical process of the system. This paper uses two fully connected layers as nonlinear operators in the physical network, and the residual of the partial differential equation can be expressed as equation (5):

$$f = \frac{\partial \hat{Y}^S}{\partial t} + FC_h(\hat{Y}^S) \quad (5)$$

where $FC_h(\cdot)$ represents the fully connected layer in the physical network.

C. Loss Function

The model optimizes the model training process through a joint loss function to ensure that both prediction accuracy and physical law constraints are met. The loss function consists of three parts: prediction loss, domain adaptation loss, and physical loss.

The prediction loss is used to measure the gap between the predicted value and the true value. By minimizing the prediction loss, the prediction accuracy of the model is improved. The mean square error is used as the prediction loss, which can be expressed as equation (6):

$$\text{loss}_p = \frac{1}{N^S}\sum_{i=1}^{N^S}(Y^S - \hat{Y}^S)^2 \quad (6)$$

where loss_p represents the prediction loss and Y^S represents the true value of the source domain data.

Domain adaptation loss is used to deal with data distribution differences under different working conditions. In a wet ball mill, differences in operating variables, set values, etc. may cause changes in data distribution. In order to make the model have stronger adaptability, domain adaptation loss is introduced. The maximum mean difference (MMD) is used as the domain adaptation loss, which can be expressed as equation (7):

$$\text{loss}_m = \left\| \frac{1}{N^S}\sum_{i=1}^{N^S}\phi(z_i^{\,S}) - \frac{1}{N^T}\sum_{i=1}^{N^T}\phi(z_i^{\,T}) \right\|_H \quad (7)$$

where loss_m represents the domain adaptation loss, $\phi(\cdot)$ represents the mapping function, $z_i^{\,S} \in Z^S$, $z_i^{\,T} \in Z^T$, H represents the reproducing kernel Hilbert space.

Physical loss embeds physical laws into the neural network training process to make the model output conform to physical laws. According to the principle of deepHPM, combined with equation (5), the physical loss formula is obtained as shown in equation (8):

$$\text{loss}_h = \text{MSE}(f, 0) = \text{MSE}(\frac{\partial \hat{Y}^S}{\partial t} + FC_h(\hat{Y}^S), 0) \quad (8)$$

where loss_h represents physical loss and $\text{MSE}(\cdot)$ represents the mean square error loss.

The total loss of the model can be expressed as:

$$\text{loss} = \text{loss}_p + \text{loss}_m + \text{loss}_h \quad (9)$$

The pseudo code of the model is shown in Algorithm 1:

Algorithm 1 DAPINN Algorithm

Input: source domain data X^S and the corresponding mill load parameter true value Y^S, target domain data X^T.

Output: target domain mill load parameter predicted value $\hat{Y}^T$.

1、Input the source domain data and the target domain data into the feature network, and calculate the features Z^S and Z^T according to equation (1).

2、Input the obtained features Z^S and Z^T into equation (7) and calculate the domain adaptation loss.

3、At the same time, the source domain feature Z^S is input into the prediction network, and the source domain mill load parameter prediction value $\hat{Y}^S$ is calculated according to equation (2).

4、Input the source domain prediction value $\hat{Y}^S$ and the mill load parameter true value Y^S into equation (6) to calculate the prediction loss loss_p .

5、At the same time, the source domain prediction value $\hat{Y}^S$ is input into the physical network to obtain $FC_h(\hat{Y}^S)$. The neural network is used to automatically calculate the partial derivative $\frac{\partial \hat{Y}^S}{\partial t}$ of $\hat{Y}^S$ with respect to time. The residual f of the partial differential equation is calculated according to equation (5).

6、Input the residual f into equation (8) to obtain the physical loss loss_h .

7、Input the obtained prediction loss, domain adaptation loss and physical loss into equation (9) to obtain the total loss, and obtain the trained model by minimizing the total loss.

8、Input the target domain data into the trained model to obtain the target domain mill load parameter prediction value $\hat{Y}^T$.

IV. Experiment

A. Dataset

The dataset used in this study was collected from a small laboratory ball mill to evaluate the prediction effect of the hybrid model based on domain adaptation and deepHPM. The experiment uses a fixed Ball charge volume ratio, that is, a certain steel ball mass, and changes the load parameters such as MBVR and PD in the mill by continuously adding materials. The experiment collects data under five working conditions. The ball charge volume ratio, the number of training set samples, and the number of test set samples under the five working conditions are shown in TABLE I.

TABLE I. EXPERIMENTAL PARAMETER

Dataset	Ball charge volume ratio	Number of training sets	Number of test sets
1	0.3	2780	1112
2	0.35	2060	824
3	0.4	1760	704
4	0.45	1900	760
5	0.5	2040	816

The data is normalized to improve the stability and efficiency of model training. The normalization method used is maximum and minimum normalization, and its calculation method is shown in equation (10):

$$X_{\text{nor}} = \frac{X - X_{\min}}{X_{\max} - X_{\min}} \tag{10}$$

where X is the input data, $X_{\min}$ is the minimum value of the feature in the data set, $X_{\max}$ is the maximum value of the feature in the data set, and X_{nor} is the normalized data.

B. Performance Metrics

In order to comprehensively evaluate the performance of the hybrid model, we used two evaluation indicators: root mean square error (RMSE) and determination coefficient (R^2). RMSE is used to measure the difference between the model prediction value and the true value. The smaller the RMSE, the higher the prediction accuracy of the model. The determination coefficient R^2 measures the fit of the model to the data. The closer the R^2 value is to 1, the better the model fits the data; the closer the R^2 value is to 0, the worse the model fits the data. The calculation formulas of RMSE and R^2 are shown in equation (11):

$$\begin{cases} \text{RMSE} = \sqrt{\dfrac{1}{N}\sum_{i=1}^{N}\left(y_i - \hat{y}_i\right)^2} \\ R^2 = 1 - \dfrac{\sum_{i=1}^{N}\left(\hat{y}_i - y_i\right)^2}{\sum_{i=1}^{N}\left(\overline{y}_i - y_i\right)^2} \end{cases} \tag{11}$$

where N represents the number of samples, y_i, $\hat{y}_i$ and $\overline{y}_i$ represent the true value, predicted value and average value at the i-th moment respectively.

C. Experimental Result

To evaluate the performance of the proposed model, we compared the proposed DAPINN model with a pure data-driven model. FCNN is a fully connected neural network. DAFCNN is a domain-adaptive fully connected neural network. A domain adaptation module is added to FCNN to enhance the model's adaptability to different working conditions. The prediction results of different models for MBVR are shown in TABLE II. and the prediction results for PD are shown in TABLE III.

TABLE II. EXPERIMENTAL RESULT OF MBVR

S	T	FCNN		DAFCNN		DAPINN	
		R^2	RMSE	R^2	RMSE	R^2	RMSE
1	2	0.832	0.121	0.869	0.106	0.892	0.097
	3	0.622	0.182	0.719	0.157	0.876	0.104
	4	0.759	0.144	0.787	0.136	0.869	0.106
	5	0.817	0.127	0.888	0.099	0.936	0.075
2	1	0.737	0.150	0.763	0.143	0.890	0.097
	3	0.941	0.071	0.964	0.056	0.978	0.043
	4	0.975	0.047	0.987	0.034	0.997	0.015
	5	0.972	0.050	0.985	0.036	0.997	0.016
3	1	0.591	0.188	0.624	0.176	0.722	0.155
	2	0.923	0.082	0.933	0.076	0.979	0.043
	4	0.956	0.062	0.959	0.060	0.987	0.033
	5	0.863	0.110	0.863	0.110	0.956	0.062
4	1	0.701	0.160	0.779	0.138	0.924	0.081
	2	0.968	0.052	0.986	0.035	0.992	0.026
	3	0.981	0.041	0.989	0.031	0.993	0.025
	5	0.950	0.066	0.975	0.047	0.983	0.039
5	1	0.884	0.100	0.904	0.091	0.954	0.063
	2	0.985	0.036	0.993	0.025	0.993	0.024
	3	0.876	0.104	0.914	0.087	0.964	0.056
	4	0.935	0.075	0.955	0.062	0.978	0.044

TABLE III. EXPERIMENTAL RESULT OF PD

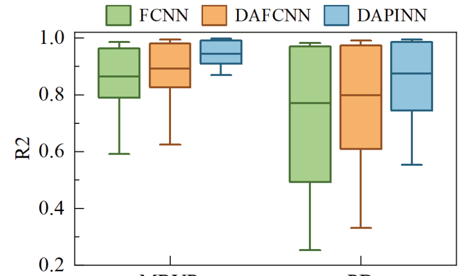
S	T	FCNN		DAFCNN		DAPINN	
		R^2	RMSE	R^2	RMSE	R^2	RMSE
1	2	0.711	0.155	0.768	0.138	0.827	0.120
	3	0.462	0.214	0.531	0.199	0.832	0.119
	4	0.456	0.211	0.483	0.205	0.729	0.149
	5	0.601	0.184	0.634	0.177	0.725	0.153
2	1	0.500	0.204	0.584	0.186	0.757	0.142
	3	0.906	0.089	0.913	0.086	0.958	0.059
	4	0.949	0.065	0.952	0.063	0.971	0.049
	5	0.958	0.060	0.971	0.050	0.992	0.026
3	1	0.252	0.250	0.330	0.236	0.577	0.188
	2	0.962	0.056	0.970	0.049	0.970	0.049
	4	0.979	0.041	0.986	0.034	0.993	0.023
	5	0.833	0.119	0.949	0.066	0.978	0.043
4	1	0.372	0.229	0.390	0.225	0.552	0.193
	2	0.979	0.042	0.966	0.053	0.986	0.034
	3	0.982	0.039	0.977	0.044	0.992	0.025
	5	0.982	0.039	0.978	0.043	0.989	0..031
5	1	0.484	0.207	0.650	0.171	0.716	0.154
	2	0.974	0.047	0.990	0.029	0.991	0.027
	3	0.951	0.065	0.955	0.061	0.966	0.053
	4	0.963	0.055	0.974	0.046	0.984	0.037

From the experimental results in the table, it can be seen that the DAFCNN and DAPINN models are superior to the FCNN model in the prediction accuracy of MBVR and PD. The experimental results show that under variable operating conditions, domain adaptation can help the model learn the common features under different operating conditions, thereby improving the generalization ability and prediction accuracy of the model. Among the three methods, the prediction accuracy of the model DAPINN proposed in this paper is the highest. In addition to using domain adaptation, the DAPIN model combines physical information with deep learning, so that the model can not only adapt to different working conditions, but also constrain the prediction results according to physical information, thereby further improving the prediction accuracy of the model. This result shows that the hybrid model based on domain adaptation and deepHPM has obvious advantages in accurately predicting mill load parameters, especially under variable operating conditions.

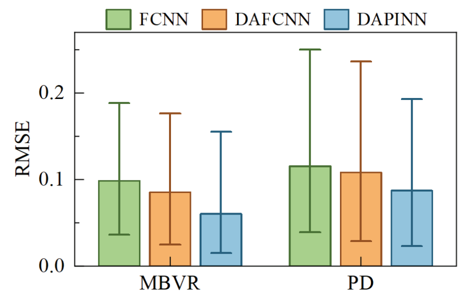
In order to improve the readability of the data and facilitate visual comparison of the performance of different models under various experimental conditions, we used a variety of visualization methods to present the R^2 and RMSE values, as shown in Fig. 3. As can be seen from the figure, there are significant differences in the performance of different models in R^2 values and RMSE values. Fig. 3 (a) shows the distribution of R^2 values of different models. The DPINN model shows a higher mean R^2 value in predicting MBVR and PD, and the box is more compact, indicating that its fitting performance is better and more stable. higher. The R^2 distributions of FCNN and DAFCNN are more scattered, indicating that the model fitting effect and stability are worse than DAPINN. Fig. 3 (b) shows the distribution of RMSE values of different models. It can be seen that the RMSE values of the DAPINN model on the two data sets are significantly lower than other models, further verifying its strong prediction accuracy. In contrast, the RMSE distribution of FCNN and DAFCNN is higher, indicating that their prediction results are poor. Taken together, the DPINN model is better than the other two models in both R^2 and RMSE, showing stronger prediction performance and stability.

D. Domain Adaptation Performance Analysis

In order to analyze the role of domain adaptation method in mill load parameter prediction, this paper draws a visualization of the data distribution of the source domain and the target domain. As shown in Fig. 4, the source domain data is dataset 3 and the target domain data is dataset 5. In order to observe the data distribution more intuitively, the principal component analysis (PCA) method is used to reduce the high-dimensional data to two dimensions, and the input data distribution and the feature distribution after maximum mean difference (MMD) processing are drawn respectively.



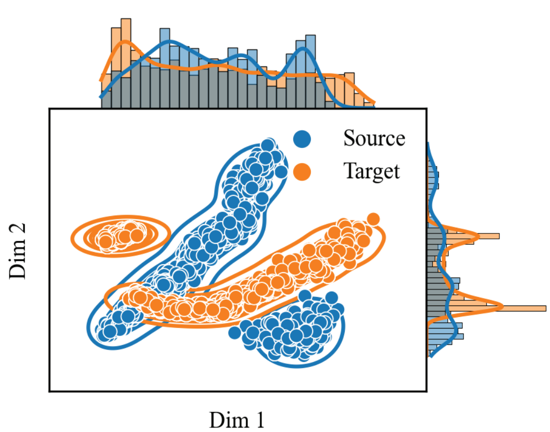
（a）Box plot of R2 values for different models



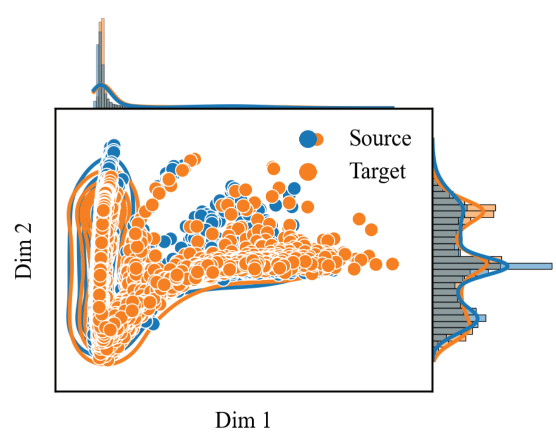
（b）Histogram of RMSE values of different models

Fig. 3. Visualization of different model evaluation indicators

Fig. 4 (a) shows the distribution of input data. The blue dots represent source domain data, and the orange dots represent target domain data. The horizontal axis represents the first dimension feature (Dim 1) after dimensionality reduction, and the vertical axis represents the second dimension feature (Dim 2) after dimensionality reduction. It can be seen that in the input data distribution, there is a relatively obvious separation between the source domain and the target domain data distribution, indicating that there is a certain difference in the distribution of the two. Fig. 4 (b) shows the feature distribution after applying MMD. It can be seen from the figure that the distribution of source domain and target domain data tends to be consistent, indicating that MMD has a good effect in reducing the difference between domains. It can be seen that the domain adaptation method plays an important role in reducing the difference in data distribution between the source domain and the target domain and improving the prediction accuracy of the target domain.



(a) Input data distribution



(b) Feature distribution

Fig. 4 Data distribution diagram

V. CONCLUSION

This paper combines physical information and pure data-driven models to establish a hybrid prediction model for mill load parameters that adapts to variable operating conditions. The model consists of three parts: feature extraction network, prediction network and physical network. Among them, the feature extraction network is used to extract the features of source domain data and target domain data. Considering the differences in data distribution, the prediction accuracy of the model is improved by minimizing the domain adaptation loss. At the same time, considering the complexity of the grinding process, the characteristics of neural networks that approximate arbitrary functions are used to simulate the physical partial differential equations of the grinding process. The interpretability of the model is improved by imposing physical constraints. The experimental part compares the prediction results of MBVR and PD by the hybrid model DAPINN proposed in this paper and the pure data-driven model. Experimental data show that the prediction effect of the model proposed in this paper is the best.

REFERENCES

[1] J. Tang, L. Zhao, J. Zhou, H. Yue, and T. Chai, "Experimental analysis of wet mill load based on vibration signals of laboratory-scale ball mill shell", Minerals Engineering, vol. 23, no. 9, pp.720-730, Aug. 2010.

[2] N. Matsanga, N. Willie, and C. Ngonidzashe, "A Review of the Grinding Media in Ball Mills for Mineral Processing", Minerals, vol. 13, no. 11, p.1373, Oct. 2023.

[3] P. Huang, J. Guo, G. Sang, Q. Miao, and M. Jia, "Soft measurement of ball mill load under variable working conditions based on deep transfer learning," Measurement Science and Technology, vol. 33, no. 7, p. 075009, Apr. 2022.

[4] L. Yang and J. Cai, "A method to identify wet ball mill's load based on ceemdan, rcmde and srnn classification," Minerals Engineering, vol. 165, p. 106852, May 2021.

[5] Y. Liu, G. Yan, R. Li, Y. Pang, and T. Qiao, "Multi-source domain adaptation method of mill load based on common and special characteristics," in 2023 42nd Chinese Control Conference (CCC). IEEE, 2023, pp. 6682–6688.

[6] P. Zhou, T. Chai, and H. Wang, "Intelligent optimal-setting control for grinding circuits of mineral processing process," IEEE Transactions on Automation Science and Engineering, vol. 6, no. 4, pp. 730–743, Oct. 2009.

[7] F. K. Mulenga, A. A. Mkonde, and M. M. Bwalya, "Effects of load filling, slurry concentration and feed flowrate on the attainable region path of an open milling circuit," Minerals Engineering, vol. 89, pp. 30–41, Apr. 2016.

[8] J. Tang, J. Qiao, Z. Liu, X. Zhou, G. Yu, and J. Zhao, "Mechanism characteristic analysis and soft measuring method review for ball mill load based on mechanical vibration and acoustic signals in the grinding process," Minerals Engineering, vol. 128, pp. 294–311, Nov. 2018.

[9] P. Huang, G. Sang, Q. Miao, Y. Ding, and M. Jia, "Soft measurement of ball mill load based on multi-classifier ensemble modelling and multi-sensor fusion with improved evidence combination," Measurement Science and Technology, vol. 32, no. 1, p. 015105, Oct. 2020.

[10] P. Jons´en, B. I. Palsson, J. F. Stener, and H.- A. aggblad, "A novel method for modelling of interactions between pulp, charge and mill structure in tumbling mills," Minerals Engineering, vol. 63, pp. 65–72, Aug. 2014.

[11] K. Mayank, M. Malahe, I. Govender, and N. Mangadoddy, "Coupled dem-cfd model to predict the tumbling mill dynamics," Procedia IUTAM, vol. 15, pp. 139–149, 2015.

[12] J. Tang, G. Yan, Z. Liu, Y. Liu, G. Yu, and N. Sheng, "Experimental analysis of wet mill load parameter based on multiple channel mechanical signals under multiple grinding conditions," Minerals Engineering, vol. 159, p. 106609, Dec. 2020.

[13] Y. Liu, G. Yan, R. Li, S. Xiao, M. Ren, and L. Cheng, "Multi-source unsupervised domain adaptive mill load forecasting method based on deep learning and fusion features," Minerals Engineering, vol. 209, p. 108650, Apr. 2024.

[14] H. Xia, J. Tang, W. Yu, and J. Qiao, "Hybrid simulator-based mechanism and data-driven for multidemand dioxin emissions intelligent prediction in the mswi process," IEEE Transactions on Industrial Electronics, vol. 71, no. 10, pp. 13 221–13 231, Oct. 2024.

[15] X. Liao, S. Chen, P. Wen, and S. Zhao, "Remaining useful life with self-attention assisted physics-informed neural network," Advanced Engineering Informatics, vol. 58, p. 102195, Oct. 2023.

[16] Z. Yang, B. Wu, J. Shao, X. Lu, L. Zhang, Y. Xu, and G. Chen, "Fault detection of high-speed train axle bearings based on a hybridized physical and data-driven temperature model," Mechanical Systems and Signal Processing, vol. 208, p. 111037, Feb. 2024.

[17] L. Rueden, S. Mayer, K. Beckh, B. Georgiev, S. Giesselbach, R. Heese, B. Kirsch, et al., "Informed machine learning – a taxonomy and survey of integrating prior knowledge into learning systems," IEEE Transactions on Knowledge and Data Engineering, vol. 35, no. 1, pp. 614–633, Jan. 2023..

[18] J. Sansana, M. N. Joswiak, I. Castillo, Z. Wang, R. Rendall, L. H. Chiang, and M. S. Reis, "Recent trends on hybrid modeling for industry 4.0," Computers & Chemical Engineering, vol. 151, p. 107365, Aug. 2021.

[19] A. Sutar, A. Kulkarni, A. Jain, P. Jadhav, and V. Gohokar, "Physics informed neural networks–a methodology review," in 2022 6th International Conference On Computing, Communication, Control And Automation (ICCUBEA. IEEE, 2022, pp. 1–4.

[20] M. Raissi, "Deep hidden physics models: Deep learning of nonlinear partial differential equations," Journal of Machine Learning Research, vol. 19, no. 25, pp. 1–24, Jan. 2018.

[21] X. Liao, S. Chen, P. Wen, and S. Zhao, "Remaining useful life with self-attention assisted physics-informed neural network," Advanced Engineering Informatics, vol. 58, p. 102195, Oct. 2023.