IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received February 28, 2020, accepted March 6, 2020, date of publication March 9, 2020, date of current version March 17, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2979611

Learning Adaptive Semi-Supervised Multi-Output Soft-Sensors With Co-Training of Heterogeneous Models

DONG LI[1], DAOPING HUANG[1], GUANGPING YU[2], AND YIQI LIU[1], (Member, IEEE)
[1]School of Automation Science and Engineering, South China University of Technology, Guangzhou 510640, China
[2]Shenyang Institute of Automation, Chinese Academy of Sciences, Guangzhou 511458, China

Corresponding author: Yiqi Liu (aulyq@scut.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 61873096 and Grant 61673181, and in part by the Science and Technology Program of Guangzhou, China, under Grant 201804010256.

ABSTRACT Soft-sensors are widely utilized for predictions of important but hard-to-measure variables in industrial processes. However, significant variations, process uncertainties, negative influence of external environment and insufficient use of unlabeled data always cause the attenuation of prediction performance. Thus, this paper proposed an adaptive semi-supervised multi-output soft-sensor by co-training recursive heterogeneous models. In the proposed strategy, a linear multi-output model, called recursive partial least square (MRPLS), and a nonlinear multi-output, called long short-term memory recurrent neural network (MLSTM), are co-trained to deal with inefficient use of label data adaptively. Ensemble of both models are not only able to address the linear and nonlinear hybrid behaviors in different time scale, but also able to deal with multiple tasks learning issues. In addition, the model proposed an odd-even grouping strategy to equalize two parts of the labeled data, which is able to capture the global variations of a process. To validate the prediction performance of the proposed soft-sensor, it was verified through a simulation benchmark platform (BSM1) and a real sewage treatment plant (UCI database). The results meant that co-training MRPLS-MLSTM achieved better performance compared with other existing co-training models in terms of the hard-to-measure variables.

INDEX TERMS Soft-sensors, semi-supervised, co-training, multi-output, recursive partial least square (RPLS), long short-term memory recurrent neural network (LSTM).

I. INTRODUCTION

In the process industries, soft-sensors are proposed to predict the hard-to-measure variables on the basis of easy-to-measure variables. Because of technical difficulty, large measurement delays, high investment cost and so on, soft-sensors gains more popularity recently. Typically, soft-sensors modeling methods are divided into mechanism-driven and data-driven. The mechanism-driven model is to establish a mathematical model to capture relationship between the input variables and the output variables by analyzing the physical and chemical reactions in the industrial process. However, a large number of industrial processes are too complex to be clearly grasped by the first-principal model, so the use of mechanism-driven is limited widely in industries [1], [2]. Recently, data-driven modeling has been gained popularity in the process industries, mainly due to the fact that data-driven modeling does not need to explore the complex process mechanism exactly, but resort to the collected data [3], [4]. Among them, Principal Component Regression (PCR), Partial Least Squares (PLS) [5], Gaussian process regression (GPR) [6], Support Vector Machine (SVM) [7], Deep learning networks [8] and other models have attracted extensive attentions in industrial and academic communities.

Typically, aforementioned methods are premised on the equal number of input and output variables. However, with increasing complexity and highly cost control requirement of industrial processes, it is well known that some output variables are harder to be obtained and the proportion of labeled and unlabeled data in acquired data is seriously unbalanced. In the field of machine learning, data samples containing input and output variables are often called labeled

The associate editor coordinating the review of this manuscript and approving it for publication was Xiao-Sheng Si.

data, while those only with input variables are referred as unlabeled data. Therefore, to develop soft sensors, traditional methods usually only use labeled data for model establishment and most unlabeled data information are fully taken for granted as useless [9]. To make full use of the information carried by unlabeled data, semi-supervised learning was proposed [3]. Semi-supervised learning method can be categorized as: graph-based method [10], generative models [11], transductive support vector machines (TSVM) [12], self-training [13] and co-training [14].

Different from other algorithms, co-training belongs to one of semi-supervised learning algorithms, it is able to take full use of the labeled data to establish two and more independent regression models. The data with high confidence from the unlabeled data are able to be selected to join the labeled data set, and the model, which is built based on the original labeled data, can be updated to improve the prediction performance. Finally, the calculation process will be repeated until the ending condition is met. Since Zhou and Li [15] proposed the co-training regression algorithm, the application of co-training algorithm to soft-sensors has received more and more attentions by a large number of scholars. Bao *et al* combined a co-training algorithm directly with the traditional PLS algorithm to obtain a co-training PLS model that can effectively solve the linear prediction problem [16]. However, the model is off-line trained and on-line used, the prediction model will degrade gradually over time. To address this issue, S. Goldman *et al* proposed an enhanced co-training algorithm, in which cross-validation was adopted when unlabeled data was selected [17]. Although unlabeled data with higher confidence can be selected, the computational cost of cross-validation process is intensive relatively. K. Nigan *et al* divided the labeled data into two groups through random sampling, which can improve the prediction ability of the model [18]. However, the grouping method tends to converge into local data selection, thereby reducing the model prediction performance. In this light, Zhou *et al* proposed a tri-training algorithm to improve the generalization ability of the model by establishing three mutually independent labeled data sets and regression models [19]. Unfortunately, when the size of labeled data is small, three pre-trained regression models may simultaneously select inappropriate unlabeled data into the original labeled data set and remove the unlabeled data with higher confidence. It is imperative to propose a new co-training algorithm to deal with aforementioned issues efficiently.

In addition, traditional soft-sensors models are mostly formulated as single-output models [20]. However, with the increase of hard-to-measure variables, the single-output models can no longer meet the prediction requirement. Multi-output models, also called multi-task learning, have gained significant attentions recently. The most typical multi-output models are mainly to transform specific single-output models into multi-output models with considering the co-relationship among the whole output variables without having to be independent necessarily. The independent assumption usually compromises the multivariate prediction due to ignorance of correlation-ship among the output variables. In this light, multi-output linear models, such as multivariate linear regression (MLR), are used to build a soft-sensor model [21]. Despite their advantages, nonlinearity in the industrial processes could degrade their prediction performance. Multi-output nonlinear models are utilized to soft-sensor models, such as multi-output gaussian process regression (MGPR) and multi-output support vector machine (MSVM) [22]. With the industrial processes becoming more and more complicated, the relationship between input variables and some output variables could be linear, whereas the relationship between input and remaining output variables would be nonlinear. Such hybrid behaviors will lead to purely linear or nonlinear multi-output being inadequate. Therefore, multi-output RPLS (MRPLS) algorithm and multi-output LSTM (MLSTM) algorithm are used to act as the sub-models for co-training semi-supervised learning by involving the output variables co-relationship. In this light, the original single-output co-training algorithm can be formulated for multi-output semi-supervised soft-sensors [23]. Similar with the supervised soft-sensors, the prediction performance will degrade as an industrial process evolution. Improving the semi-supervised soft-sensors to adapt to a process variation is an indispensable issue.

The paper proposed an adaptive semi-supervised multi-output soft-sensor, termed as co-training MRPLS-MLSTM. Firstly, different from the standard co-training models to split the labeled data set into two groups with the former half of labeled data being training sets and latter being testing sets, this paper proposed an odd-even grouping method, which divided the labeled data into two parts in an odd-even sequence. By doing so, global features can be taken into account, rather than local features as standard training data set selection. Secondly, a real industrial process is difficult to define to be linear or nonlinear during the entire operational stage. Co-training of the MRPLS and MLSTM will simultaneously adapt two diverse regression algorithms for soft-sensors. Also, two diverse regression methods can improve the modeling independence, thus potentially ensuring that one model can achieve an acceptable performance at least. Last but not least, since recursive multi-output models are used to establish prediction model, proper adaptation of a multivariate and dynamical process can be achieved. To the best of authors' study, it is the first attempt to propose an adaptive semi-supervised multi-output soft-sensor.

The rest of this paper is structured as follows. Section 2 is the detailed introduction to MRPLS and MLSTM algorithm. Section 3 describes co-training MRPLS-MLSTM soft-sensor in detail. Section 4 firstly applies the proposed soft-sensor to a well-established wastewater plant (WWTP) validation platform, Benchmark Simulation Model No.1 (BSM1). Then, the soft-sensor is utilized for a real WWTP with data collecting from the field. Section 5 provides a discussion. The conclusions suggest that the prediction performance can be indeed improved by the co-training MRPLS-MLSTM in Section 6.

II. PRELIMINARIES

A. MULTI-OUTPUT RECURSIVE PARTIAL LEAST SQUARE (MRPLS)

In this paper, MRPLS algorithm is considered as a sub-model for co-training algorithm. To ensure the multi-output model adapt to process variations, MPLS algorithm is improved by using moving window technique. In this strategy, when new data are coming, they will be added to the original labeled training data that is enveloped, and a few oldest data will be removed. Sequentially, the mean and variance of the samples will be updated accordingly. Finally, the remaining old data and the new data will work together to rebuild a new model [24]. The specific mathematical calculation of MRPLS algorithm is shown as follows:

According to the criterion of covariance maximization, input and output variables matrix, X and Y, are decomposed as follows:

$$X = TP + E = \sum_{h=1}^{a} t_h p_h + E \tag{1}$$

$$Y = UQ + F = \sum_{h=1}^{a} u_h q_h + F \tag{2}$$

where $X \in R^{n\times m}$ is input matrix, $Y \in R^{n\times l}$ represents output matrix. $T \in R^{n\times a}$ is the score matric of X and $U \in R^{n\times a}$ is the score matric of Y. n is the size of data, m is the number of input variables, a is the number of potential variables. t_h is the hth row of T, u_h is the hth row of U. $P(a \times m)$ and $Q(a \times l)$ are the loading matrices, p_h is the hth row of P,q_h is the hth row of Q. E and F are two noise matrices. Therefore, the relationship between u_h and t_h can be received:

$$u_h = b_h t_h \tag{3}$$

where b_h denotes the regression coefficient of the correlation between X space principal component, t, and Y space principal component, u. Therefore, the relationship between X and Y can be expressed.

$$Y = TBQ + F \tag{4}$$

where B is the regression matrix. After acquiring T, Q, B and so on though the labeled data, MRPLS algorithm will update the labeled data X and Y with the new labeled data x_t, y_t and the forgetting factor $\lambda(0 < \lambda < 1)$. Then, the updated $X = [\lambda X, x_t]$, $Y = [\lambda Y, y_t]$. Finally, T, Q and B in $Y = TBQ + F$ are updated by the changed labeled data X and Y.

MRPLS algorithm is a method applicable to the prediction of high-dimensional data. In addition. MRPLS, as a multi-output algorithm, can effectively improve the prediction efficiency.

B. MULTI-OUTPUT LONG SHORT-TERM MEMORY RECURRENT NEURAL NETWORK (MLSTM)

LSTM is a widely used enhanced recurrent neural network and has the better performance when processing data with strong time-series dependence. It has been applied for many fields recently, such as speech recognition and natural language processing [25]. In this paper, LSTM algorithm is used for soft-sensor modeling and extended to multi-output system, called MLSTM. The corresponding network structure diagram as shown in Fig. 1 (A).

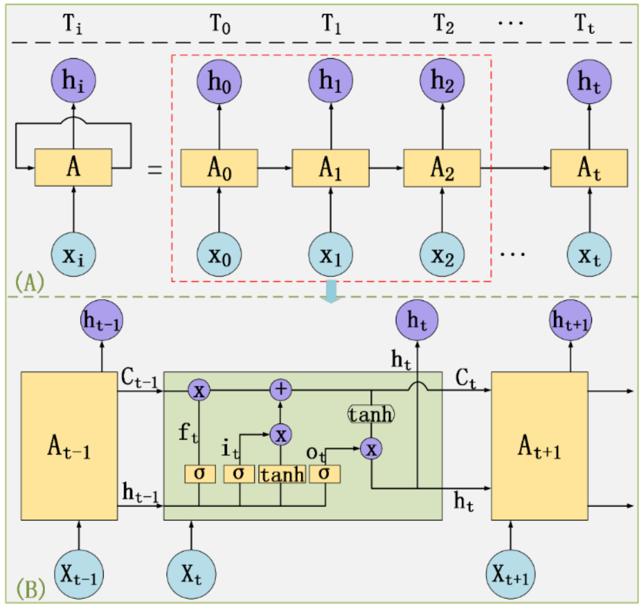

FIGURE 1. The structure diagram of MLSTM.

Fig. 1 (A) (left) reveals the network structure, with the processed information about the current moment being passed to the next moment. x_i is the input information in input layer at time i, h_i is the output information in output layer at time i and A presents the hidden layer. Fig. 1 (A) (right) displays the hidden layer structure in detail. As can be seen from the Fig. 1, the information at the previous moment can not only be passed to the output layer, but also to the sequential moment. This is called the "memory function". Fig. 1 (B) is the hidden layer of specific structure, detailed procedures of MLSTM algorithm is shown as follows:

Firstly, retained information at the previous moment need to be determined. Information is retained (output "1" represents completely retain) or discarded (output "0" represents completely discard) through the S-layer of "forgetting threshold". Suppose f_t is the reserved information, x_t is input information, h_{t-1} is the output information at the previous moment, σ represents the S-layer. w_f is the weight vector of "forgetting threshold", b_f is the threshold. Therefore, the retained information is:

$$i_t = \sigma(w_i \cdot [h_{t-1}, x_t] + b_t) \tag{5}$$

Then, to calculate candidate limit $\bar{C}_t$:

$$\bar{C}_t = tanh\,(w_c \cdot [h_{t-1}, x_t] + b_c) \tag{6}$$

Current new state C_t is:

$$C_t = f_t \times C_{t-1} + i_t \times \bar{C}_t \tag{7}$$

Secondly, the output information is determined necessarily. The output information is determined by multiplying the output of the *tanh*-layer and the output of the "output threshold" of the S-layer. Suppose the output unit state is O_t, w_o is the weight vector of the output state, b_o is the threshold, h_t represents output information. The state of output unit is:

$$O_t = \sigma(w_o \cdot [h_{t-1}, x_t] + b_o) \tag{8}$$

Output information h_t is:

$$h_t = O_t \times tanh\,(C_t) \tag{9}$$

In LSTM recurrent neural network, due to the "memory function", The historical information can be effectively utilized and the regression model can also be better established [26]. For MLSTM, it is improved based on LSTM, so the proposed algorithm can retain original advantages and can also predict multi-output variables simultaneously.

III. ADAPTIVE SEMI-SUPERVISED MULTI-OUTPUT SOFT-SENSORS

The purpose of co-training algorithm is to select appropriate unlabeled data with high confidence and then to optimize the performance of prediction model. In this section, an adaptive co-training MRPLS-MLSTM is proposed. The proposed soft-sensor can not only deal with dynamic multi-output learning in industrial process, but also be able to approach the hybrid behaviors of linearity and nonlinearity.

A. ADAPTIVE SEMI-SUPERVISED MULTI-OUTPUT SOFT-SENSOR BY CO-TRAINING MRPLS

The core idea behind co-training MRPLS is that, by combining the co-training paradigm with MRPLS, the semi-supervised soft-sensor can be built to deal with unlabeled data. The reasons for selecting MRPLS are mainly from following aspects. First of all, as an adaptive algorithm, MRPLS can update the model though the changed data and make the model adapt to the external environment. Also, MRPLS can predict multiple output variables simultaneously through one modeling, which improves the efficiency of prediction. Finally, because the co-training algorithm is an iterative method, the building of model should not be time-consuming. As everyone knows, PLS is a linear modeling method, modeling fast is its advantage, so is MRPLS.

The procedure of co-training MRPLS is depicted as follows. Let L denotes the labeled data set and U denotes the unlabeled data set. $L = \{X, Y\} = \left\{(x_1, y_1), (x_2, y_2) \ldots \hat{\mathrm{A}}(x_{|L|}, y_{|L|})\right\}$, X and Y are the input and output data respectively. $|L|$ and $|U|$ are the sizes of corresponding data sets.

Firstly, L is split into two parts L_1 and L_2, which represent two independent data sets. Then two regression models $h_1(L_1)$ and $h_2(L_2)$ can be built by the two labeled data sets L_1 and L_2. The root means squared error (RMSE) of the labeled data set L was calculated by $h_1(L_1)$ and $h_2(L_2)$, defined R. After that, every unlabeled data x_u was used to receive the predictive values $\hat{y}_u^1$ and $\hat{y}_u^2$. Then, let $\left(x_u, \hat{y}_u^1\right)$ and $\left(x_u, \hat{y}_u^2\right)$ serve as the new labeled data and be put in L_1 and L_2. two sets are updated and be used to build new models $h_1'(L_1, \left(x_u, \hat{y}_u^1\right))$ and $h_2'(L_2, \left(x_u, \hat{y}_u^2\right))$. Accordingly, the new RMSE R' was calculated by the new model $h_1'(L_1, \left(x_u, \hat{y}_u^1\right))$ and $h_2'(L_2, \left(x_u, \hat{y}_u^2\right))$ in the original labeled data set. By calculating the difference of R and R', we take the maximum value as the new labeled data with the highest confidence of. Finally, we put $(x_u, h_2(x_u))$ into L_1 and $(x_u, h_1(x_u))$ into L_2 by the cross-placement method. The specific confidence formula is as follows:

$$\nabla_u = \left| \sqrt{\frac{\sum_{x_i \in L} (y_i - h\,(x_i))^2}{|L|}} - \sqrt{\frac{\sum_{x_i \in L} (y_i - h'\,(x_i))^2}{|L|}} \right| \tag{10}$$

where $x_i \in L$ and $y_i \in L$ represent the labeled input and output data respectively. $|L|$ denotes the size of data. h is the initial regression model, h' is the new regression model when add new labeled data $(x_u, \hat{y}_u)$. In a word, by calculating the ∇_u, the unlabeled data with the highest consistency against the original labeled can be selected, and then the new labeled data set will be use to improve the accuracy of model.

Finally, until the ending condition is satisfied, the mean value of predictions from two new models which are built from the new labeled sample sets is taken as the final predicted value:

$$h\,(x) = \frac{1}{2}\,(h_1\,(x) + h_2\,(x)) \tag{11}$$

B. ADAPTIVE SEMI-SUPERVISED MULTI-OUTPUT SOFT-SENSOR BY CO-TRAINING MRPLS-MLSTM

For the prediction problem of nonlinear data, this paper proposed co-training MLSTM by replacing MRPLS with MLSTM. MLSTM, as the multi-output extension of LSTM, can not only retain the "memory function" and other merits, but also can predict multiple output variables through one modeling. These greatly improve the predictive performance and efficiency. In addition, the procedure of co-training MLSTM is basically consistent with the co-training MRPLS, except that MRPLS is replaced by MLSTM when establishing the prediction model.

co-training MLSTM is proposed to better solve the nonlinear problem that co-training MRPLS fails to involve. Through the two models, the co-training algorithm can be found the appropriate prediction model for both linear and nonlinear data.

C. ADAPTIVE SEMI-SUPERVISED MULTI-OUTPUT SOFT-SENSOR BY CO-TRAINING MLSTM

Based on the above co-training regression model, an adaptive semi-supervised multi-output soft-sensor is proposed in this paper, termed as co-training MRPLS-MLSTM. Different from the traditional co-training models, co-training MRPLS-MLSTM simultaneously uses two recursive multi-output prediction algorithms to build the predicted model with the labeled data. To widen and enhance the modeling performance, a linear multi-output model, MRPLS and a nonlinear multi-output model, MLSTM are involved in co-training. Also, to avoid converging into locally optimal model construction, labeled data are grouped into odd-even for training

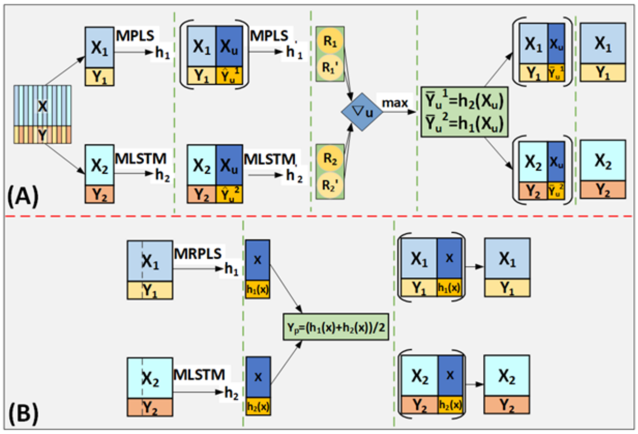

FIGURE 2. The schematic of co-training MRPLS-MLSTM soft-sensor.

and testing. Overall, the soft-sensor can attempt to solve the problems of data linearity-nonlinearity, time-varying and multiple-task learning and semi-supervised learning simultaneously in industrial processes.

The calculation procedures of co-training MRPLS-MLSTM are shown in Fig. 2 (where (A) and (B) represent the training and prediction procession respectively). In Fig. 2 (A), the odd-even grouping method is to mark the data number, and then divide the labeled data into two groups by the number. Also, the max represents to select the x_u corresponding to the largest ∇_u. It should be noticed that when update the labeled data set, we put $(x_u, h_2(x_u))$ into L_1 and $(x_u, h_1(x_u))$ into L_2 by the cross-placement method. In Fig. 2(B), when the training is ending, the new labeled data sets are used to establish the models h_1 and h_2 through MRPLS and MLSTM. Then, the final prediction value is determined by the mean of $h_1(x)$ and $h_2(x)$. At the same time $(x, h_1(x))$ and $(x, h_2(x))$ will be added into L_1 and L_2 to update labeled data sets. During the procession, labeled data L mean those data with input output variates, the residual are unlabeled data U. L and U will keep changing under different simulation conditions. To further clarify the calculation, the procedures are tabulated as follows:

In this soft-sensor, heterogeneous models, including MRPLS and MLSTM, are co-trained to make predictions for multiple hard-to-measure variables. The soft-sensor can select appropriate unlabeled data efficiently to update the prediction model. Then, the soft-sensor overcomes the disadvantages which divide labeled data locally. Moreover, two different types of regression algorithms can establish more genetic models to approach widen process variations. Additionally, by using the on-line regression algorithm, the prediction model can be updated with new collected data information efficiently.

IV. CASE STUDIES

In order to evaluate the prediction performance of co-training MRPLS-MLSTM, two simulation studies were provided. One case is a well-established wastewater plant (WWTP) validation platform, sampled every 15 minutes. Whereas the other is a real sewage treatment plant, with every sample being collected every day.

TABLE 1. Algorithm: Co-training MRPLS-MLSTM soft-sensor.

Input: labeled data set L. (include input variables x_i and output variables y_i)
unlabeled data set U. (only include input variables x_u)
test data set P, maximum number of learning iterations T
Process:
divided L into L_1 and L_2 by the odd-even grouping method
$L_1 = \{(x_1, y_1), (x_3, y_3) \dots (x_{|L|-1}, y_{|L|-1})\}$
$L_2 = \{(x_2, y_2), (x_4, y_4) \dots (x_{|L|}, y_{|L|})\}$
Repeat for T rounds:
$h_1 = MPLS(L_1), h_2 = MLSTM(L_2)$
For $\forall x_u \in U$
$h_1(x_u) = \hat{y}_u^1, h_2(x_u) = \hat{y}_u^2$
$h_1' = MPLS(L_1, (x_u, \hat{y}_u^1)), h_2' = MLSTM(L_2, (x_u, \hat{y}_u^2))$
$R = \sqrt{\frac{\sum_{x_i \in L_1}(y_i - h_1(x_i))^2 + \sum_{x_j \in L_2}(y_j - h_2(x_j))^2}{|L|}}$
$R' = \sqrt{\frac{\sum_{x_i \in L_1}(y_i - h_1'(x_i))^2 + \sum_{x_j \in L_2}(y_j - h_2'(x_j))^2}{|L|}}$
$\nabla_u = |R - R'|$
$x_n = argmax \nabla_u$
$\bar{y}_n^1 = h_1(x_n), \bar{y}_n^2 = h_2(x_n)$
$L_1 = L_1 \cup \bar{y}_n^2, L_2 = L_2 \cup \bar{y}_n^1$
$U = U - x_n$
End of for
End of repeat
Output new labeled data set L_1 and L_2
$x \in P$
$h_1 = MRPLS(L_1), h_2 = MLSTM(L_2)$
$y_p = \frac{1}{2}(h_1(x) + h_2(x))$

The prediction performance of co-training MRPLS-MLSTM is evaluated by RMSE, correlation coefficient (R) and D-values, the detailed formulas are as follows:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(\hat{y}_i - y_i)^2}{n}} \quad i = 1, 2, \dots n \tag{12}$$

$$R\left(Y, \hat{Y}\right) = \frac{cov(Y, \hat{Y})}{\sqrt{var[Y]\, var[\hat{Y}]}} \tag{13}$$

$$D - values = |predict\ values - real\ values| \tag{14}$$

where n is the size of sets, $\hat{y}_i$ and y_i denote predicted value and real value respectively, $\hat{Y} = (\hat{y}_1, \hat{y}_2, \dots, \hat{y}_n)$, $Y = (y_1, y_2 \dots, y_n)$.$cov(Y, \hat{Y})$ represents the covariance of Y and $\hat{Y}$, $var[Y]$ and $var\left[\hat{Y}\right]$ are variance of Y and $\hat{Y}$. Smaller RMSE represents better prediction performance of soft-sensors. R mainly falls inside [0,1], the value which is the

TABLE 2. RMSE, R and RMSSD of prediction variables.

Type	Model selection	Predicted variables	SS	SNH	SNO	COD	BOD_5	RMSSD
Non-adaptive	Co-training MPLS	RMSE	0.102	2.728	0.683	6.165	0.068	3.122
		R	0.820	0.921	0.894	0.949	0.986	
	Co-training MBP	RMSE	0.035	3.014	0.513	2.630	0.043	2.497
		R	0.735	0.904	0.923	0.951	0.983	
Adaptive	Co-training MRPLS	RMSE	0.017	3.033	1.615	1.427	0.038	2.476
		R	0.825	0.902	0.901	0.944	0.962	
	Co-training MLSTM	RMSE	**0.004**	**0.812**	0.823	11.864	0.046	3.681
		R	**0.963**	**0.994**	0.918	0.861	0.990	
	Co-training MRPLS-MLSTM	RMSE	0.014	2.617	**0.503**	**0.608**	**0.006**	**1.936**
		R	0.821	0.839	**0.926**	**0.971**	**0.992**	

closer to 1 means the better performance. $D - values$ means the loss between predicted value and real value.

To access the prediction performance for multiple responses, RMSSD detail formula is as follows:

$$RMSSD = \sqrt{\frac{1}{N} trace\left\{\left(Y - \hat{Y}\right)'\left(Y - \hat{Y}\right)\right\}} \tag{15}$$

where *trace* represents the trace of the matrix, N is the number of the output data set. Smaller RMSSD is better prediction performance of a soft-sensor.

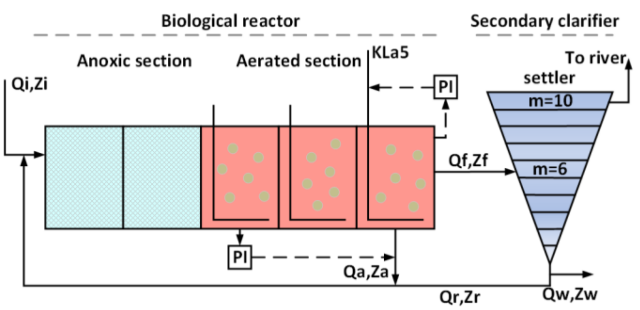

FIGURE 3. Schematic of the BSM1.

A. BENCHMARK SIMULATION MODEL 1 (BSM1)

1) BACKGROUND

The first International Association Water Quality Task Group developed Benchmark Simulation Model (BSM). The simulation benchmark plant is shown in Fig. 3, it consists of a bioreactor (5999 m^3) and a secondary sedimentation tank (4 m deep, 10 layers, 6000 m^3). There are five mixed small cell reaction tanks in the bioreactor, the first two compartments are non-aerated, whereas the others are aerated. There are two types of internal circulation: activated sludge circulation is from the bottom of the secondary sedimentation tank to the front end of the plant and nitrate internal circulation is from the last tank to the first tank. The average flow rate of sewage treatment is 20000 m^3/day, the average concentration of chemical oxygen demand (COD) is 300 mg/L. Nitrification and denitrification reactions are required to remove organic matter. This plant samples every 15 minutes for every variable and simulates the sewage treatment process under sunny conditions for 14 days, collecting 1344 sets of data [27].

The case study is to validate the adaptive semi-supervised multi-output soft-sensor co-training MRPLS-MLSTM to achieve suitable predictions. In this study, because readily biodegradable substrate effluent (SS-E), $NH_4^+ + NH_3$ nitrogen effluent (SNH-E), nitrate and nitrite nitrogen effluent (SNO-E), chemical oxygen demand for effluent (COD-E) and five-day biological oxygen demand effluent (BOD_5-E) are quality-related hard-to-measure variables, and they can represent the treatment efficiency of WWTP, we denote them as the output variables. To access the prediction performance, we compare co-training MRPLS-MLSTM with other models (co-training MPLS, co-training MBP, co-training MRPLS and co-training MLSTM). In addition, according to the mechanism analysis and expert experiences, 15 important variables are selected as input variables and the aforementioned 5 variables act as output variables shown as the appendix Table. 4 (the variables 16-20 are the output variables). Half of 672 data sets are taken as labeled data. Then, the remaining data sets are taken as unlabeled data, but their output variables need to be overwritten. Until the end of the iteration, the output variables of the unlabeled data are supplemented completely as test data to verify the prediction performance.

2) PREDICTION PERFORMANCE OF CO-TRAINING MRPLS-MLSTM SOFT-SENSO

To present the prediction performance, Table. 2 displays all criterion evaluation results of these models (co-training MPLS, MBP, MRPLS, MLSTM and MRPLS-MLSTM). According to the RMSSD, in non-adaptive soft-sensors, co-training MBP is more accurate than co-training MPLS. The main reason is that MBP is better than MPLS to approximate the nonlinear complex sewage treatment process. For adaptive soft-sensors, co-training MRPLS-MLSRM achieves the best prediction performance with RMSSD being 1.963. This value decreases by 21.81% and 47.41% compared with

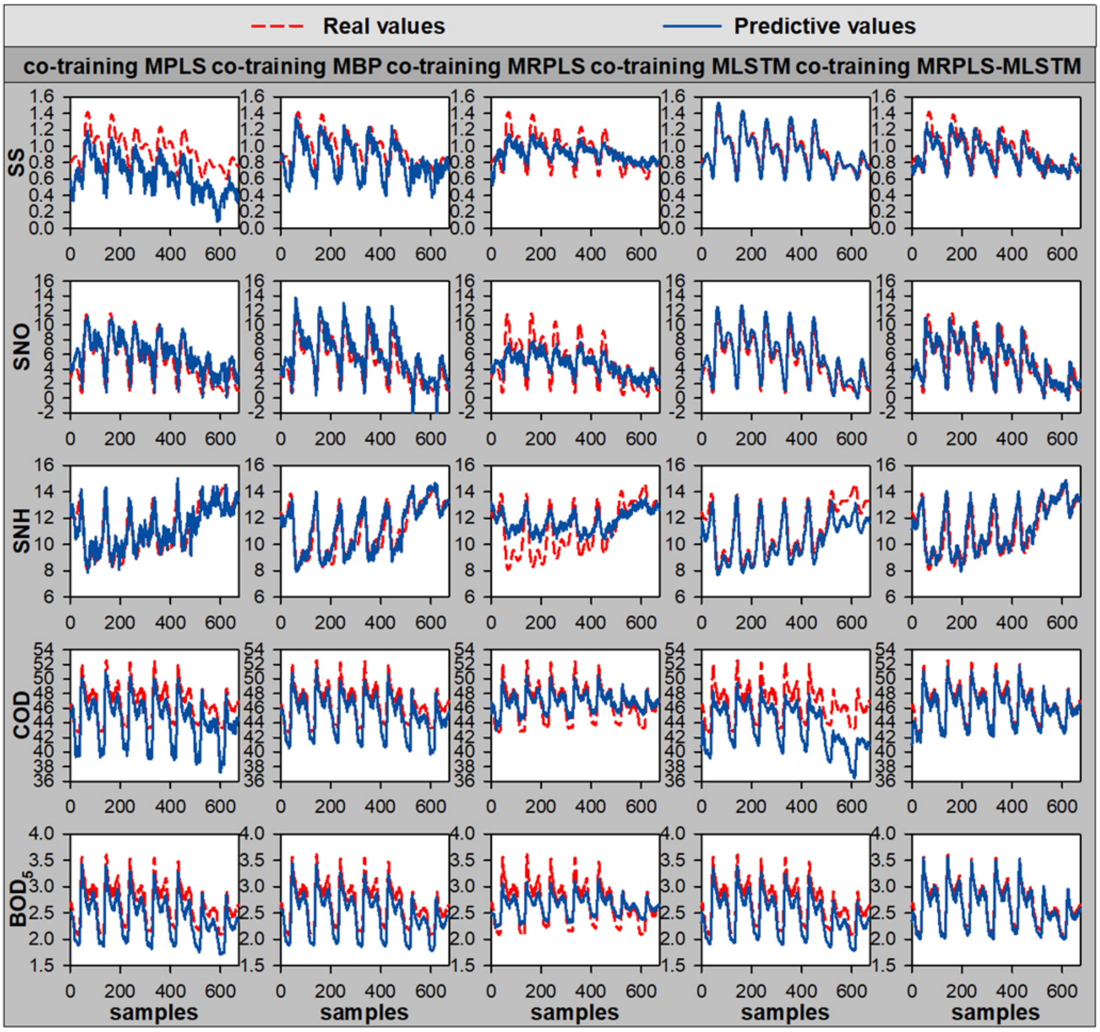

FIGURE 4. Fits of the prediction and real values with respect to output variables using non-adaptive and adaptive soft-sensors for the case study 1.

the adaptive co-training MRPLS and co-training MLSTM. Because co-training MRPLS-MLSRM is a heterogeneous soft-sensor, by linear and nonlinear models, it can make two models complementary. Moreover, by comparing RMSE and R between the same output variables, it is obvious that the prediction results by co-training MRPLS-MLSTM are most accurate. In particular, RMSE of BOD_5 is reduced by 90.74% and 85.39% than non-adaptive co-training MPLS and co-training MBP, respectively. However, for adaptive soft-sensors, it is noted that co-training MLSTM has the better prediction performance for SS and SNH, even better than the co-training MRPLS-MLSTM. This is mainly because the proposed soft-sensor and other models belong to multi-output models, they can satisfy the condition of the optimal overall output variables, rather than every output variable. In a word, we can know from the Table. 2 that co-training MRPLS-MLSTM has the best results for BSM1, which is usually a stable system under sunny conditions.

The predicted curves of non-adaptive and adaptive soft-sensors are compared with the real curve in Fig. 4. In non-adaptive soft-sensors co-training MPLS and co-training MBP, it can be seen that the predictive results of co-training MBP are better than co-training MPLS, especially the prediction of output variable SS. Also, by comparing with the adaptive soft-sensors, we found co-training MRRPLS-MLSTM can better track the change trend of the target. For example, the prediction curve of peaks and valleys of co-training MRPLS-MLSTM soft-sensor for COD can be completely consistent with the real curves, the prediction curve of co-training MLSTM is worst. This further proves the proposed soft-sensor having excellent prediction ability for important variables of sewage treatment plants under stationary conditions. However, it is worth noting the prediction curve of SS. Obviously, co-training MLSTM is the most accurate model to track the real curve, the performance of the proposed model is not as good as it. The main reason is that they belong to multi-output soft-sensors, by establish a prediction model to

TABLE 3. RMSE, R and RMSSD of prediction variables.

Type	Model selection	Predicted variables	RD-DBO-S	RD-DQO-S	DQO-S	DBO-S	RMSSD
Non-adaptive	Co-training MPLS	RMSE	**7.512**	**16.884**	116.558	10.090	12.290
		R	0.882	0.853	0.794	0.845	
	Co-training MBP	RMSE	11.420	34.676	**92.469**	14.852	12.386
		R	0.834	0.738	0.832	0.781	
Adaptive	Co-training MRPLS	RMSE	14.237	43.699	140.437	11.673	14.493
		R	0.876	0.808	0.807	0.775	
	Co-training MLSTM	RMSE	12.737	34.801	180.739	14.908	15.594
		R	0.800	0.665	0.624	0.676	
	Co-training MRPLS-MLSTM	RMSE	9.160	32.686	150.228	**8.753**	14.171
		R	0.848	0.738	0.730	0.829	

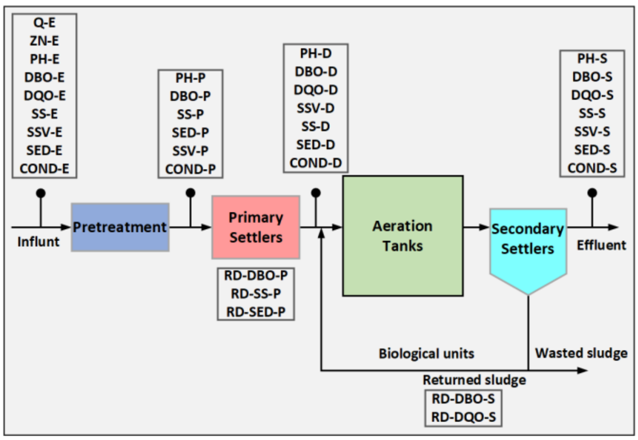

FIGURE 5. A wastewater plant for validation.

predict multiple output variables, it is impossible to achieve optimal results for every variable. Finally, compared with the prediction curves of the non-adaptive and adaptive, it is no hard to find that adaptive soft-sensors are better. This is because they use recursive regression algorithm to update the prediction model with new data information.

B. A REAL WASTEWATER TREATMENT PLAN

1) BACKGROUND

Different from the simulation platform in the first case study, the second case collected data from a real sewage treatment plant. As shown in Fig. 5, the proposed wastewater plant process (Blake and Merz, 1998) consists of five parts: pretreatment, primary precipitation, aeration tank, secondary precipitation and sludge reflux. The plant has a sewage treatment capacity of 35 000 m^3/day, more details can be seen in [28]. In this process, with the time going by, the population of microorganisms (both in quality and number of species) and the influent rate are varied. because of lack of instrumentation, the collecting period of the data is one day, a total number of data is 527.

In this case study, because the collected data belong to real wastewater treatment plants data, the predictive performance of co-training MRPLS-MLSTM can be evaluated in real wastewater treatment plants. The effluent chemical oxygen demand (DQO) and biological oxygen demand (DBO), DQO and DBO of secondary settlers (RD-DQO-S and RD-DBO-S) are selected as the output variables. because of the small amount of data set, the all 38 variables were selected from the measurable variables as the input and output variables. The detailed variables information is shown in appendix Table. 5 (the variables, 28-29 and 33-34, are the output). Due to the abnormal data points affected by rain-storm in the data and partial missing data, 126 groups of data were deleted before model training. To verify the prediction performance of the proposed soft-sensor for abrupt change data, some data points less affected by the environment were retained. The first

TABLE 4. Selected variables for modeling in BSM1.

No	variables	Comments
1	SS-1	Readily biodegradable substrate-1
2	SS-in	Readily biodegradable substrate influent
3	SNH-1	NH_4^++NH_3 nitrogen-1
4	SNH-2	NH_4^++NH_3 nitrogen-2
5	SNH-3	NH_4^++NH_3 nitrogen--3
6	SNH-in	NH_4^++NH_3 nitrogen- influent
7	SNO-1	Nitrate and nitrite nitrogen-1
8	SNO-2	Nitrate and nitrite nitrogen-2
9	SNO-3	Nitrate and nitrite nitrogen-3
10	SO-1	Oxygen-1
11	SO-2	Oxygen-2
12	SO-5	Oxygen-5
13	Q-tr	Flow rate internal recycling
14	Q-in	Flow rate influent
15	COD-in	Chemical oxygen demand for influent
16	**SS-E**	**Readily biodegradable substrate effluent**
17	**SNH-E**	**NH_4^++NH_3 nitrogen- effluent**
18	**SNO-E**	**Nitrate and nitrite nitrogen effluent**
19	**COD-E**	**Chemical oxygen demand for effluent**
20	**BOD_5-E**	**Five-day biological oxygen demand effluent**

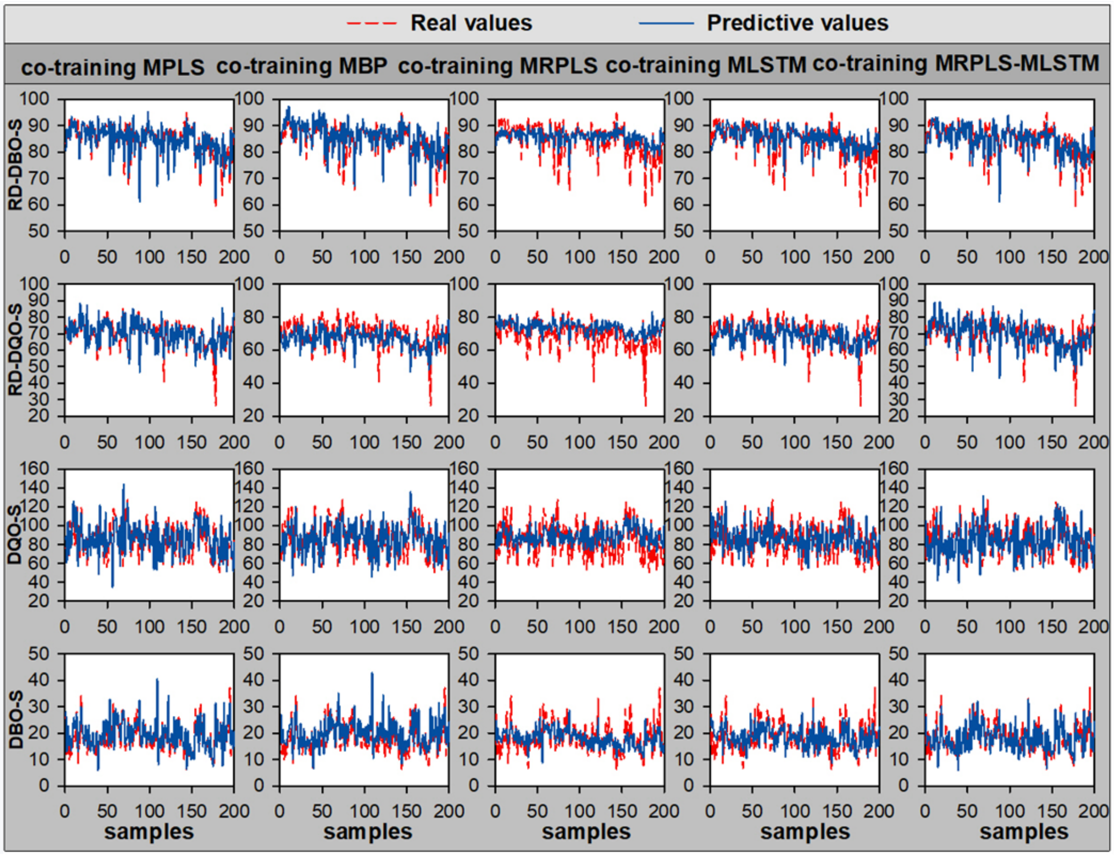

FIGURE 6. Fits of the prediction and real values with respect to output variables using non-adaptive and adaptive soft-sensors for the case study 2.

200 sets of data are taken as labeled data, and the remaining 200 sets of data are taken as unlabeled data after covering the output variables. When the termination condition is met, output variables of unlabeled data are supplemented as test data to evaluate the predicted performance.

2) PREDICTION PERFORMANCE OF CO-TRAINING MRPLS-MLSTM SOFT-SENSO

Table. 3 is the criterion evaluation results of prediction variables by the non-adaptive and adaptive models (co-training MPLS, MBP, MRPLS, MLSTM and MRPLS-MLSTM models). According to the RMSSD values, co-training MBP is still better than co-training MPLS in non-adaptive soft-sensors. This is because data of real wastewater treatment plants have more nonlinear characteristics, co-training MBP can be more accurate approximation. Also, we can see the value of co-training MRPLS-MLSTM alleviates 2.22% and 9.13% than adaptive co-training MRPLS and co-training MLSTM, but for co-training MPLS and co-training MBP, it increases 15.31% and 14.41%. the reason is that heterogeneous co-training MRPLS-MLSTM can make models complementary, it can solve more comprehensive data prediction problems. But, for the data with frequent fluctuations, heterogeneous soft-sensors will affect each other, making the prediction performance of the model worse. By comparing RMSE and R, it is shown that co-training MRPLS-MLSTM has better prediction results for DBO than other models. The RMSE values decreases 25.01% and 41.29% than co-training MRPLS and co-training MLSTM. In addition, for the other output variables, co-training MRPLS-MLSTM has the better predictive performance than other adaptive models. However, the non-adaptive models have the better prediction performance than adaptive models. Therefore, the proposed soft-sensor can only enhance predictive accuracy over the other adaptive models in the real wastewater treatment plant.

The prediction curves of co-training MRPLS-MLSTM and other soft-sensors for output variables are shown in Fig. 6 Obviously, the prediction curves of non-adaptive soft-sensors are more accurate than adaptive soft-sensors, especially the prediction of peaks and valleys. This proves that these adaptive soft-sensors cannot optimize the real data with frequent fluctuations. But the prediction curve of co-training MRPLS-MLSTM for DBQ-S is an exception, as can be seen in Fig. 6. In addition, we found that all soft-sensors could not track outliers well, but the prediction results of heterogeneous co-training MRPLS-LSTM was relatively accurate. In a word, the prediction of these model for the hard-to-measure variables of the real sewage plant need to be further study.

TABLE 5. The variables introduction in the real wastewater treatment plant.

Locations	No	Variable comments	Variables
Influent to WWTP	1	Water flow	Q-E
	2	Zinc	ZN-E
	3	PH	PH-E
	4	Suspended solids	SS-E
	5	Volatile suspended solids concentration	SSV-E
	6	Sedimentary solids	SED-E
	7	Electrical conductivity	COND-E
	8	Biological oxygen demand	DBO-E
	9	Chemical oxygen demand	DQO-E
Primary settlers	10	PH	PH-P
	11	Suspended solids	SS-P
	12	Volatile suspended solids concentration	SSV-P
	13	Sedimentary solids	SED-P
	14	Electrical conductivity	COND-P
	15	Biological oxygen demand	DBO-P
Secondary settlers	16	Suspended solids	SS-D
	17	Sedimentary solids	SED-D
	18	PH	PH-D
	19	Volatile suspended solids concentration	SSV-D
	20	Electrical conductivity	COND-D
	21	Biological oxygen demand	DOB-D
	22	Chemical oxygen demand	DQO-D
Effluent	23	Sedimentary solids	SED-S
	24	PH	PH-S
	25	Volatile suspended solids concentration	SSV-S
	26	Suspended solids	SS-S
	27	Electrical conductivity	COND-S
	28	**Biological oxygen demand**	**DBO-S**
	29	**Chemical oxygen demand**	**DQO-S**
Primary settlers	30	Suspended solids	RD-SS-P
	31	Biological oxygen demand	RD-DBO-P
	32	Sedimentary solids	RD-SED-P
Secondary settlers	**33**	**Biological oxygen demand**	**RD-DBO-S**
	34	**Chemical oxygen demand**	**RD-DQO-S**
Overall plant	35	Suspended solids	RD-SS-G
	36	Sedimentary solids	RD-SED-G
	37	Biological oxygen demand	RD-DBO-G
	38	Chemical oxygen demand	RD-DQO-G

V. DISCUSSION

An improved adaptive co-training MRPLS-MLSTM was proposed in this paper. The soft-sensor used recursive multivariate regression algorithms to predict the multi-output variables. It provided a simple but powerful tool for predicting the important variables of sewage treatment plants. Compared with the traditional co-training models, it can not only solve the linear prediction problem effectively, but also realize the nonlinear prediction well. Firstly, adopts the odd-even grouping method when dividing the labeled data. This effectively avoids making model training using the local data with unnecessary data fluctuation. Secondly, the linear MRPLS model and nonlinear MLSTM model are integrated to deal with the hybrid behaviors of a process, thus leading to widen process adaptation. Moreover, recursive models are used to predict multi-output variables in order to make better use of the information of new data and establish a more accurate prediction model efficiently.

The proposed soft-sensor is verified by a simulated sewage treatment plant (BSM1) and a real sewage treatment plant (UCI). The simulated sewage plant is sufficiently equipped with short sampling period and the data are considered to be sufficient, whereas for real sewage treatment plants, the process is complex and exposes to extreme conditions sometimes. Furthermore, there will be more hard-to-measure variables in real field. When co-training MRPLS-MLSTM soft-sensor is applied to two cases, satisfactory prediction results can be achieved better prediction results than other models. In non-adaptive soft-sensors, the performance of co-training MBP is better than co-training MPLS, the reason is that nonlinear MBP is better to approximate the nonlinear complex sewage treatment process. In adaptive soft-sensors, because sewage treatment is a complicated and dynamic process, the data usually exhibit strongly time-varying features and a real industrial process is hard to define to be linear or nonlinear. Thus, the heterogeneous co-training MRPLS-MLSTM can achieve more accurate prediction results. At last, we can draw a conclusion from two cases. Co-training MRPLS-MLSTM cannot only accurately predict stable large dataset, but also effectively solve the prediction problem of small dataset.

In our study, we assumed that labeled data is sufficient. In other words, labeled data can be used to establish an accurate initial regression model after being evenly divided into two parts. For the case of large sample data, it is obvious that the soft-sensor can achieve satisfactory prediction performance. Therefore, the proposed co-training MRPLS-MLSTM is a convenient and efficient multi-output model. However, for a small sample of time-varying data, the adaptive soft-sensors would often reduce the prediction performance during the process of updating the models. Therefore, in the proposed model, MRPLS and MLSTM can be replaced by other linear or nonlinear models [29], and the recursive method can also be used in others, such as just-in-time learning (JITL) and so on [30].

VI. CONCLUSION

An adaptive semi-supervised multi-output soft-sensor, termed as co-training MRPLS-MLSTM, is proposed to predict hard-to-measure variables in a simulated sewage treatment plant (BSM1) and a real sewage treatment plant (UCI). By integrating the linear recursive multi-output model (MRPLS) with nonlinear recursive multi-output model (MLSTM), the prediction performance of the proposed soft-sensor can be indeed improved, in terms of RMSSD decreasing 21.81% and 47.41% than other adaptive multiple-output soft-sensors. Moreover, the odd-even grouping methods can be able to select the global labeled data to build the accurate model. However, the proposed soft-sensor requires a large amount of labeled data when establishing the initial regression model. Future study can focus on the improvement of adaptive semi-supervised multiple-output models under the small sample data or mechanism-based modeling.

APPENDIX

See Tables 4 and 5.

REFERENCES

[1] W. Shao, S. Chen, and C. J. Harris, "Adaptive soft sensor development for multi-output industrial processes based on selective ensemble learning," *IEEE Access*, vol. 6, pp. 55628–55642, 2018.

[2] X. Yuan, C. Ou, Y. Wang, C. Yang, and W. Gui, "A layer-wise data augmentation strategy for deep learning networks and its soft sensor application in an industrial hydrocracking process," *IEEE Trans. Neural Netw. Learn. Syst.*, to be published.

[3] W. Yan, P. Guo, Y. Tian, and J. Gao, "A framework and modeling method of data-driven soft sensors based on semisupervised Gaussian regression," *Ind. Eng. Chem. Res.*, vol. 55, no. 27, pp. 7394–7401, Jul. 2016.

[4] Y. Ye, J. Ren, X. Wu, G. Ou, and H. Jin, "Data-driven soft-sensor modelling for air cooler system pH values based on a fast search pruned-extreme learning machine," *Asia–Pacific J. Chem. Eng.*, vol. 12, no. 2, pp. 186–195, Mar. 2017.

[5] S. Mahesh, D. S. Jayas, J. Paliwal, and N. D. G. White, "Comparison of partial least squares regression (PLSR) and principal components regression (PCR) methods for protein and hardness predictions using the near-infrared (NIR) hyperspectral images of bulk samples of canadian wheat," *Food Bioprocess Technol.*, vol. 8, no. 1, pp. 31–40, Jan. 2015.

[6] Y. Liu, H. Xiao, Y. Pan, D. Huang, and Q. Wang, "Development of multiple-step soft-sensors using a Gaussian process model with application for fault prognosis," *Chemometric Intell. Lab. Syst.*, vol. 157, pp. 85–95, Oct. 2016.

[7] Z. Ma, Y. Dong, H. Liu, X. Shao, and C. Wang, "Forecast of non-equal interval track irregularity based on improved grey model and PSO-SVM," *IEEE Access*, vol. 6, pp. 34812–34818, 2018.

[8] X. Yuan, J. Zhou, B. Huang, Y. Wang, C. Yang, and W. Gui, "Hierarchical quality-relevant feature representation for soft sensor modeling: A novel deep learning strategy," *IEEE Trans. Ind. Informat.*, vol. 16, no. 6, pp. 3721–3730, Jun. 2020.

[9] J. Zhu, Z. Ge, and Z. Song, "Quantum statistic based semi-supervised learning approach for industrial soft sensor development," *Control Eng. Pract.*, vol. 74, pp. 144–152, May 2018.

[10] E. Tu, Y. Zhang, L. Zhu, J. Yang, and N. Kasabov, "A graph-based semi-supervised k nearest-neighbor method for nonlinear manifold distributed data classification," *Inf. Sci.*, vols. 367–368, pp. 673–688, Nov. 2016.

[11] C. Li, J. Zhu, and B. Zhang, "Max-margin deep generative models for (semi-)supervised learning," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 40, no. 11, pp. 2762–2775, Nov. 2018.

[12] X. Li, X. Lyu, Y. Tong, S. Li, and D. Liu, "An object-based river extraction method via optimized transductive support vector machine for multi-spectral remote-sensing images," *IEEE Access*, vol. 7, pp. 46165–46175, 2019.

[13] J. Tanha, M. van Someren, and H. Afsarmanesh, "Semi-supervised self-training for decision tree classifiers," *Int. J. Mach. Learn. Cybern.*, vol. 8, no. 1, pp. 355–370, Feb. 2017.

[14] Y. Z. Zhan and Y. B. Chen, "Co-Training semi-supervised active learning algorithm with noise filter," *Pattern Recognit. Artif. Intell.*, vol. 22, no. 5, pp. 750–755, May 2009.

[15] Z. Zhou and M. Li, "Semi-supervised regression with co-training," in *Proc. Int. Joint Conf. Artif. Intell.* San Mateo, CA, USA: Morgan Kaufmann, vol. 5, 2005, pp. 908–913.

[16] L. Bao, X. Yuan, and Z. Ge, "Co-training partial least squares model for semi-supervised soft sensor development," *Chemometric Intell. Lab. Syst.*, vol. 147, pp. 75–85, Oct. 2015.

[17] S. Goldman and Y. Zhou, "Enhancing supervised learning with unlabeled data," in *Proc. ICML*, 2000, pp. 327–334.

[18] K. Nigam and R. Ghani, "Analyzing the effectiveness and applicability of co-training," in *Proc. 9th Int. Conf. Inf. Knowl. Manage. (CIKM)*, 2000, pp. 86–93.

[19] Z.-H. Zhou and M. Li, "Tri-training: Exploiting unlabeled data using three classifiers," *IEEE Trans. Knowl. Data Eng.*, vol. 17, no. 11, pp. 1529–1541, Nov. 2005.

[20] J. Wu, H. Cheng, Y. Liu, B. Liu, and D. Huang, "Modeling of adaptive multi-output soft-sensors with applications in wastewater treatments," *IEEE Access*, vol. 7, pp. 161887–161898, 2019.

[21] J.-M. Chiou, Y.-F. Yang, and Y.-T. Chen, "Multivariate functional linear regression and prediction," *J. Multivariate Anal.*, vol. 146, pp. 301–312, Apr. 2016.

[22] A. Boustati, T. Damoulas, and R. S. Savage, "Non-linear multitask learning with deep Gaussian processes," 2019, *arXiv:1905.12407*. [Online]. Available: http://arxiv.org/abs/1905.12407

[23] Y. Liu, B. Liu, X. Zhao, and M. Xie, "Development of RVM-based multiple-output soft sensors with serial and parallel stacking strategies," *IEEE Trans. Control Syst. Technol.*, vol. 27, no. 6, pp. 2727–2734, Nov. 2019.

[24] B. Lu, I. Castillo, L. Chiang, and T. F. Edgar, "Industrial PLS model variable selection using moving window variable importance in projection," *Chemometric Intell. Lab. Syst.*, vol. 135, pp. 90–109, Jul. 2014.

[25] J. Yan, Y. Qi, Q. Rao, H. He, and S. Qi, "LSTM-based with deterministic negative sampling for API suggestion," *Int. J. Softw. Eng. Knowl. Eng.*, vol. 29, no. 7, pp. 1029–1051, Jul. 2019.

[26] X. Yuan, L. Li, and Y. Wang, "Nonlinear dynamic soft sensor modeling with supervised long short-term memory network," *IEEE Trans Ind. Informat.*, vol. 16, no. 5, pp. 3168–3176, May 2020.

[27] Y. Liu, Y. Pan, and D. Huang, "Development of a novel adaptive soft-sensor using variational Bayesian PLS with accounting for online identification of key variables," *Ind. Eng. Chem. Res.*, vol. 54, no. 1, pp. 338–350, Jan. 2015.

[28] Y. Liu, D. Huang, and Y. Li, "Development of interval soft sensors using enhanced just-in-time learning and inductive confidence predictor," *Ind. Eng. Chem. Res.*, vol. 51, no. 8, pp. 3356–3367, Feb. 2012.

[29] Y. Liu, B. Liu, X. Zhao, and M. Xie, "A mixture of variational canonical correlation analysis for nonlinear and quality-relevant process monitoring," *IEEE Trans. Ind. Electron.*, vol. 65, no. 8, pp. 6478–6486, Aug. 2018.

[30] X. Yuan, Z. Ge, B. Huang, Z. Song, and Y. Wang, "Semisupervised JITL framework for nonlinear industrial soft sensing based on locally semisupervised weighted PCR," *IEEE Trans Ind. Informat.*, vol. 13, no. 2, pp. 532–541, Apr. 2017.

DONG LI was born in Kaifeng, Henan, China. He received the B.E. degree in computational mathematics from Henan Normal University, in 2016. He is currently pursuing the Ph.D. degree in control science and engineering with the Department of Automation, South China University of Technology, Guangzhou, China. His research interests include statistical theory research, machine learning, data-driven soft sensors, and semi-supervised learning in wastewater treatment.

DAOPING HUANG received the B.Eng. degree in chemical automation and instruments and the M.Eng. and Ph.D. degrees in automatic control theory and applications from the South China University of Technology (SCUT), Guangzhou, China, in 1982, 1986, and 1998, respectively.

In 1982, he started his teaching and research career with the Department of Automation, SCUT. From 1995 to 1996, he worked with the University of Ghent, Belgium, as a government sponsored Visiting Scholar. He is currently a Full Professor with the School of Automation Science and Engineering, SCUT. He has directed over ten research projects. He has authored three academic books and more than 200 conference and journal articles. His research interests include intelligent detection, control, soft-sensing technology, and fault diagnosis and accident prediction for industrial processes. He serves as the Vice Director of the Education Committee and a member of the Process Control and Application Committees with the Chinese Association of Automation. He received the Third Prize from the National Education Committee, in 1992, and the Second Prize from the Guangdong Provincial Government, in 2005, for his contributions to science and technology.

GUANGPING YU was born in February 1980. He received the Ph.D. degree from Shenyang Automation Institution, in 2009. He is currently the Director of the Intelligent-Factory-Technology Research Center, an Assistant Director, and a Full Professor with the Shenyang Institute of Automation. He has been engaged in the research on intelligent factory and application of intelligent environmental protection technology for a long time. He has joined in more than 50 projects, such as the National Natural Science Foundation of China, 863 Program, and so on. He received the National Wu-Yi Labor Medal, in 2019, a Star of Double Innovation, High-Level Talents in Guangzhou, Lingnan Outstanding Talents, and Nansha District High-End Leading Talent.

YIQI LIU (Member, IEEE) was born in Haikou, China, in 1983. He received the B.S. and M.S. degrees in control engineering from the Chemical University of Technology, Beijing, in 2009, and the Ph.D. degree in control engineering from the South China University of Technology, Guangzhou, China, in 2013. From 2013 to 2016, he was a Lecturer with the South China University of Technology, where he has been an Associate Professor with the Department of Automation, since 2016. He is the author of more than 60 articles. He holds three patents. His research interests include soft sensors, fault diagnosis, and wastewater treatment. He was a recipient of the Marie Skłodowska-Curie Individual Fellowships, in 2020, the Hong Kong Scholar Award, in 2016, the Chinese Scholarship Council Award, in 2011, and the Deutscher Akademischer Austausch Dienst Award, in 2015. He is an Associate Editor of IEEE Access and *ISA Transactions*.