12th IFAC International Symposium on
Advanced Control of Chemical Processes
July 14-17, 2024. Toronto, Canada

Novel distributed broad seasonal trend learning system for industrial soft sensing application

Peng-Fei Wang, Qun-Xiong Zhu, Yan-Lin He*

** College of Information Science and Technology, Beijing University of Chemical Technology, Beijing, China; Engineering Research Center of Intelligent PSE, Ministry of Education of China, Beijing, China (e-mail: heyl@mail.buct.edu.cn)*

Abstract: Industrial soft sensing has gained widespread use in industrial processes due to its advantages in terms of low cost and easy maintenance. However, as industrial processes become increasingly complex, characterized by high dimensionality, coupling, and nonlinearity in the data, traditional data-driven soft sensing models often fall short of achieving the required level of accuracy. In this paper, a novel and enhanced variant of the Broad Learning System (BLS) called the Distributed Broad Seasonal Trend Learning System (DBSTLS) is proposed for the development of industrial soft sensing with improved accuracy. In the proposed DBSTLS, a distributed structure based on Seasonal-Trend Decomposition Procedure Based on LOESS (STL) is established. Through STL, dynamic process data can be mainly separated into two distinct components: the trend feature and the season feature. The distributed structure is built separately for the trend feature and the season feature. Subsequently, a fast learning strategy, based on BLS, is applied to both the distributed trend feature and the season feature. This integrated approach culminates in the development of the DBSTLS for industrial soft sensing. To validate the effectiveness of the proposed DBSTLS-based industrial soft sensing, the process data collected from the Pure Terephthalic Acid production process are used. Simulation results confirm that the DBSTLS-based industrial soft sensing outperforms other models in terms of accuracy.

Keywords: Industrial soft sensing; soft sensor; data-driven modeling; broad learning system; process industry.

1. INTRODUCTION

Soft sensing has been applied to predict product quality, monitor machinery health, and optimize energy usage, all of which help ensure product consistency and quality, minimize waste, and improve manufacturing efficiency (Kadlec et al., 2009; Jiang et al.,2020). Industrial soft sensing often refers to virtual sensing, signifying a shift from conventional, sensor-based data acquisition to a more flexible and sophisticated approach. Instead of relying solely on physical sensors, soft sensing harnesses the power of computational methods, algorithms, and data analytics to estimate, predict, or infer critical process variables, product characteristics, or system states (Zhu et al., 2020; Chen et al.,2023). Traditional sensors may have limitations in their ability to capture all relevant information accurately, often resulting in incomplete data. Soft sensing addresses this limitation by amalgamating diverse data sources, incorporating historical data, machine learning, and statistical models, enabling industries to monitor and control processes with high accuracy (Chin et al., 2020; Shih et al., 2020). Moreover, the installation and maintenance of physical sensors can be expensive and logistically challenging, especially in hazardous or remote environments. Therefore, soft sensing has gained widespread use in industrial processes.

Generally, for industrial soft sensing, there are two types: mechanism based soft sensing and data-drive soft sensing (Guo et al., 2022). As industrial systems become more intricate and interconnected, it is difficult to build mechanism based soft sensing models due to the complexity in obtaining process knowledge and solving differential equations. In modern process industry, distributed control systems have been widely utilized (He et al., 2023). Therefore, more and more historical process data can be easily collected. The data-driven soft sensing models can be effectively established using the collected process data. Hence, data-driven soft sensing has become a hot research filed (Yan et al., 2016; Curreri et al., 2020). Data-driven soft sensing uses computational methods, algorithms, and data analytics to estimate, predict, or infer critical process variables. Among the data-driven models, neural networks has been widely studied and utilized as industrial soft sensing (Rani et al., 2013). Back propagation neural network (BPNN) has been used as soft sensing to predict key process variables. Radial basis function neural network shortened as RBFNN was developed as industrial soft sensing in industrial processes (Wang et al., 2022). Although the former mentioned two neural networks have been successfully used as industrial soft sensing, there is still limitation. In BPNN and RBFNN, much training time is required for training, which reduces the application ability of BPNN and RBFNN as industrial soft sensing. Extreme learning machine shortened as ELM as a feed forward neural network solves this limitation. ELM has extreme learning speed, which make the training process fast. Zhang et. al has used ELM as soft sensing in industrial processes (Zhang et al., 2018). The ELM based soft sensing achieved high accuracy with fast training speed. However, with the increasing complexity of process data in the term of high dimension, coupling and nonlinearity, the traditional ELM cannot obtain acceptable accuracy in soft sensing modeling due to the shallow structure of ELM. Luckily, broad learning system (BLS) can handle complex data with a broad manner. On one hand, BLS has the good feature of fast learning. On the other hand, BLS can achieve high accuracy when dealing with complex data, which has the handling ability of deep learning. BLS has been successfully and widely used in the fields of classification and modeling (Chen and Liu, 2017; Gong et al., 2021).

For industrial soft sensing, there is still limitation in BLS. BLS can well handle process data in the static stage. For time series process data, BLS cannot well handle this kind of data. In industrial processes, most of the process data are dynamic. Thus, it is necessary to avoid this limitation when developing soft sensing using BLS. In this paper, we propose a novel improved BLS named distributed broad seasonal trend learning system (DBSTLS) to handle this limitation. In the proposed DBSTLS, there is a distributed structure based on Seasonal-Trend Decomposition Procedure Based on LOESS (STL). The STL is an effective method to extract the features of dynamic process data from the following aspect: the trend feature and the season feature (Cleveland et al., 1990; Li et al., 2020). The trend feature is a linear relationship hided in the process data. The season feature contains the dynamic information of the process data. Hence, a distributed structure can be built using the trend feature and the season feature, respectively. Then, the fast learning strategy in BLS is utilized in the distributed trend feature and season feature. Finally, the proposed DBSTLS can be developed as industrial soft sensing. For the sake of verifying the effectiveness of the proposed DBSTLS based industrial soft sensing, the process data collected from the Pure Terephthalic acid process is used. The simulation results verified that the proposed DBSTLS based industrial soft sensing could achieve higher accuracy compared with other models.

The rest part of this paper is organized as follows: section 2 briefly introduces the BLS and STL; section 3 provided the details of the methodology of the proposed DBSTLS based industrial soft sensing; section 4 gives the simulation results and analyses; section 5 contains the conclusions.

2. PRELIMINARIES

In this section, two basic models are briefly introduced: STL and BLS.

2.1 Brief introduction to STL

The Seasonal-Trend Decomposition Procedure Based on LOESS (STL) is a time series decomposition algorithm relying on local weighted regression. It can decompose time series data into trend, seasonal components, and residual elements. Specifically, the trend component captures long-term trends within the data, the seasonal component identifies cyclic patterns, and the residual component accounts for random noise that is difficult to attribute to either trend or seasonality. The inner loop process of STL is illustrated in Figure 1.

The inner loop process of STL is depicted in Figure 1. To alleviate the impact of outliers in the residuals, robust weights η, are introduced in the outer loop. In the successive iteration steps (2) and (6) within the inner loop, when executing Loess smoothing, adjacent weights are multiplied by the robust weight η to constrain outliers. This adjustment serves to minimize the influence of previously identified outliers, consequently fortifying the algorithm's robustness. The utilization of STL for decomposing industrial process data enables a more comprehensive elucidation of their temporal, non-linear, and non-stationary characteristics.

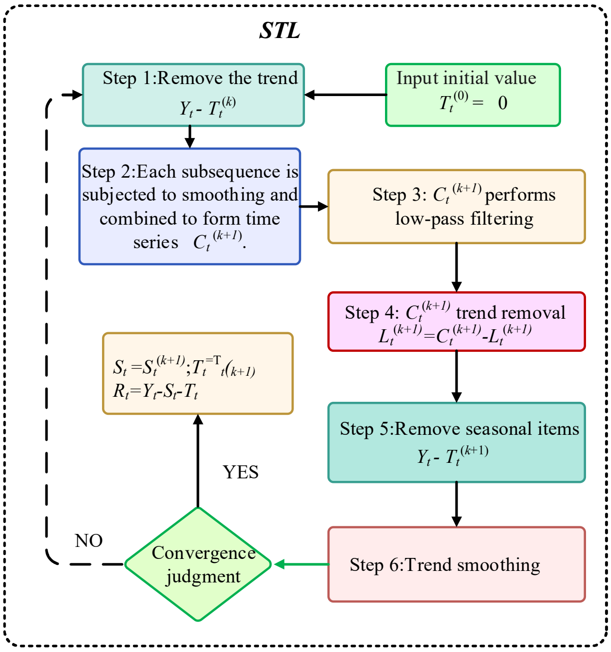

Figure 1.The procedure of STL sequence decomposition.

2.2 Brief introduction to BLS

BLS can be considered an alternative approach to deep learning models, which is optimized on the basis of a chain neural network with random vector functions as the carrier. Its capacity for rapid system updates and modifications during the training process is facilitated by an incremental learning algorithm. The structure of BLS is visually depicted in Figure 2, revealing its components, which consist of a feature layer, an enhancement layer, and an output coefficient matrix.

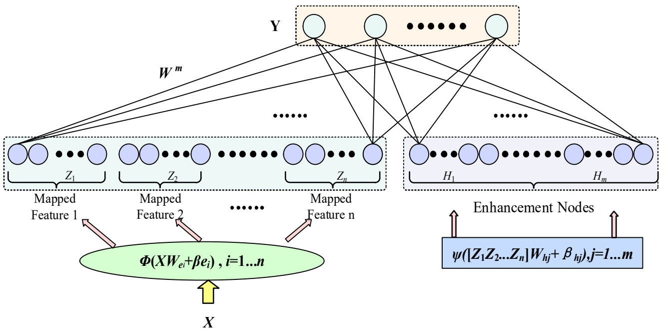

Figure 2. Description of the BLS .

Assume that the input data set is $X \in R^{m\times n}$ and the output data is $Y \in R^{m\times 1}$,the n feature mappings can be represented as follows:

$$\mathrm{Z}_i = \phi(XW_{e_i} + \beta_{e_i}), \quad i = 1,...,n \tag{1}$$

where W_{e_i} and β_{e_i} are randomly determined weights and bias terms, $\phi(\cdot)$ indicates linear transformation. While each feature mapping $\boldsymbol{Z}_i$ consists of k nodes, collectively represented as $Z^n \equiv [Z_1, Z_2, ..., Z_n]$.

At the enhancement layer, Z^n serves as the input to the nonlinear function of the i-th enhancement node H_j, resulting in its expression as follows:

$$H_j = \psi(Z^n W_{h_j} + \beta_{h_j}), \quad j = 1,...,m \tag{2}$$

where W_{h_j} and β_{h_j} are randomly determined weights and bias terms, $\psi(\cdot)$ indicates nonlinear transformation. The total number of enhancement nodes is denoted as m, with all of them defined as $H^m \equiv [H_1, H_2, ..., H_m]$.

The enhancement nodes and feature nodes are combined to obtain A, the predicted output $\widetilde{Y}$ is formulated as follows:

$$\widetilde{Y} = [Z^n | H^m] W^m = AW^m \tag{3}$$

$$W^m = [A]^+ Y \tag{4}$$

Where W^m is the output coefficient matrix, which can be obtained by pseudo-inverse calculation.

3. THE PROPOSED METHOD

Contemporary industrial process data display intricate temporal and coupled characteristics, frequently encumbered by residual errors and noise stemming from data acquisition lag and equipment limitations. To tackle these challenges, this paper introduces a soft sensing model grounded in DBSTLS. The STL algorithm is employed to scrutinize the dynamic properties of the data, thereby proficiently mitigating noise errors. Following this, the amalgamation of trend components, seasonal components, and the original input is parallelly fed into the width learning system, facilitating swift and accurate regression prediction of the output variable. The structure of the DBSTLS model is visually portrayed in Figure 3.

Assuming that there are m-dimensional input variables and 1-dimensional output variables comprising the set of samples $\{(x_i, \ y), i = 1, 2, ..., m\}$, where x_i represent n samples of the i-th dimension input variable, and y represents the output variable. x_i is subjected to STL sequence decomposition as follows:

$$x_i = t_i + s_i + r_i \tag{5}$$

Where s_i, t_i and r_i respectively represent the seasonal component, trend component, and residual of t the i-th dimension input variable.

$$S^m = [s_1, s_2, ..., s_m] \tag{6}$$

$$T^m = [t_1, t_2, ..., t_m] \tag{7}$$

The application of n-group linear eigen-transform to the trend component T^m, is effectuated in the following:

$$\mathrm{Z}_i = \phi(T^m W_{e_i} + \beta_{e_i}), \quad i = 1,...,n \tag{8}$$

$$Z^n \equiv [Z_1, Z_2, ..., Z_n] \tag{9}$$

where W_{e_i} and β_{e_i} are randomly determined weights and bias terms, $\phi(\cdot)$ indicates linear transformation function. Z_i represents the i-th set of feature transform outputs and Z^n represents all feature transforms.

Parallelization of the seasonal component S^m with the original input X results in the generation of S^k, which is subsequently subjected to nonlinear enhancement:

$$S^k = [X | S^m] \tag{10}$$

$$H_j = \psi(Z^n W_{h_j} + \beta_{h_j}), \quad j = 1,...,m \tag{11}$$

$$H^m \equiv [H_1, H_2, ..., H_m] \tag{12}$$

where W_{h_j} and β_{h_j} are the weight matrices and bias terms pertaining to the corresponding dimensions, while $\psi(\cdot)$ signifies the nonlinear transformation function. The output of the j-th augmented node is denoted as H_j, and the aggregate output across all augmented nodes is formally defined as H^m.

Hypothesis that the output matrix of the linear feature mapping and enhancement layer is denoted as A_m^n, the predicted output of DBSTLS can be formulated as follows:

$$\widetilde{Y} = A_m^n W_m^n = [Z^n | H^m] W_m^n \tag{13}$$

The output coefficient matrix, represented as W_m^n, can be rapidly calculated through an approximation of the pseudo-inverse of A_m^n in the context of ridge regression applied to the actual output Y. The process is the following:

$$\begin{aligned} W_m^n &= [A_m^n]^+ Y \\ &= \{\lambda I + A_m^n (A_m^n)^T\}^{-1} (A_m^n)^T Y \end{aligned} \tag{14}$$

$$[A_m^n]^+ = \lim_{\lambda \to 0} \{\lambda I + A_m^n (A_m^n)^T\}^{-1} (A_m^n)^T \tag{15}$$

In summary, the computational development process of the DBSTLS model proposed in this paper is as follows: Firstly, the input data is decomposed into three components, namely, trend, seasonality, and residuals, using the STL algorithm. Secondly, considering the highly dynamic and coupled nature of the process data, a distributed structure is established. The trend component undergoes linear feature mapping, and the seasonality component, in parallel with the original input, is subjected to nonlinear transformation in the enhancement layer. Finally, the combination of the feature layer output and the enhancement layer output undergoes a pseudo-inverse operation to yield the predictive output.

$$MAE = \frac{1}{N_{test}} \sum_{i=1}^{N_{test}} \left| \hat{y}_{i,test} - y_{i,test} \right| \tag{16}$$

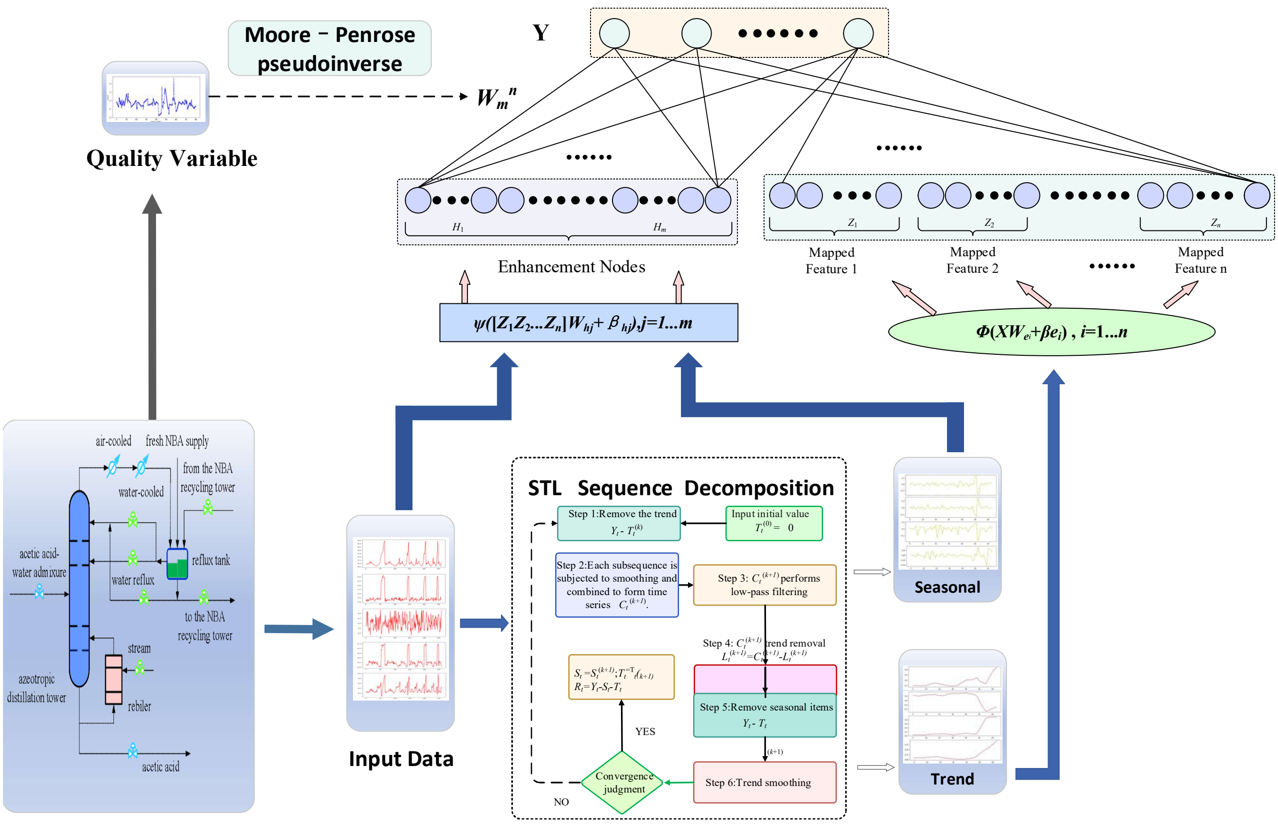

Figure 3. Structure of the DBSTLS

$$RMSE = \sqrt{\sum_{i=1}^{N_{test}} \left(\hat{y}_{i,test} - y_{i,tset} \right)^2 \Big/ N_{test}} \tag{17}$$

$$R^2 = 1 - \sum_{i=1}^{N_{test}} \left(y_{i,test} - \hat{y}_{i,test} \right)^2 \Big/ \sum_{i=1}^{N_{test}} \left(y_{i,test} - \bar{y}_{i,test} \right)^2 \tag{18}$$

To quantitatively analyze the prediction accuracy of the soft-sensing model, three evaluation indices are chosen in this paper: mean absolute error (MAE) and root mean square error (RMSE). The calculation formulas for these three indices are presented above.

4. CASE STUDY

In this section, to further assess the effectiveness of the proposed soft sensing model based on DBSTLS, an actual industrial process known as the solvent system used for purifying terephthalic acid (PTA) was employed. The proposed method and other approaches were utilized for simulation validation, and the results were subsequently analyzed.

4.1 Brief introduction to PTA

PTA is a crucial raw material in the production of polyester fibers. The process production unit for PTA primarily consists of three components: the solvent dehydrator, reboiler, and reflux tank. The detailed composition of PTA can be found in the literature (Tian et al., 2020). During the production process, the consumption of acetic acid reflects the progress of the oxidation process. Therefore, online measurement of acetic acid content at the top of the dehydrator tower is of paramount importance. However, due to the presence of residual acetic acid in the tower top wastewater, direct measurement is challenging. Therefore, it is highly necessary to establish a soft sensor for predictive purposes.

In the PTA production process, there are 17 variables, including flow rate, pressure, and temperature, which are closely associated with the acetic acid content at the top of the tower. This study utilizes a total of 260 sets of PTA sample data to assess the effectiveness of the soft sensing model based on DBSTLS. The data is divided into a training set, comprising 80% of the total (208 data points), and a test set, consisting of the remaining 20% (52 data points).

4.2 Simulation result analyses

To ascertain the decomposition periods of process variables within the STL decomposition module, the Step was defined as {5, 9, 13, 17, 21, 25}. Post-decomposition, input variables underwent normalization via min-max scaling to facilitate model training. The training process for BLSTL involved the application of a grid search method to identify optimal hyperparameters. Consequently, the learning rate γ was established at 0.05, the number of feature nodes k at 51, the number of mapping groups n at 130, and the number of enhancement nodes m at 68.

For a comprehensive evaluation of the soft sensing model based on BSTLS, this study conducted a comparative analysis with other soft sensing models, namely the Extreme Learning Machine (ELM), traditional BLS, GRU, PCA-based PCA-BLS, and PCA-ELM. The efficacy of the six soft sensing models was gauged using three metrics: Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and R-squared (R^2).

Table 1. Prediction performance of six models

	R^2	RMSE	MAE
ELM	0.645	0.358	0.286
BLS	0.686	0.337	0.275
GRU	0.749	0.307	0.247
PCA-ELM	0.707	0.326	0.264
PCA-BLS	0.735	0.311	0.251
DBSTLS	0.801	0.268	0.216

Table 2 outlines the evaluation results of the six models on the PTA dataset. As delineated in Table 1, the soft sensing model predicated on the traditional ELM manifests the least favorable fit. This outcome is attributed to the inherent simplicity of ELM's structure, which precludes the extraction of temporal features integral to the nature of PTA.

In contrast to the soft sensing model rooted in PCA-BLS, the innovative soft sensing model proposed in this study, based on DBSTLS, attains superior accuracy. A pivotal disparity between these models lies in the fact that while the PCA-BLS model undertakes linear dimension reduction on the data, DBLSTL dynamically decomposes the data, effectively eliminating residual noise.

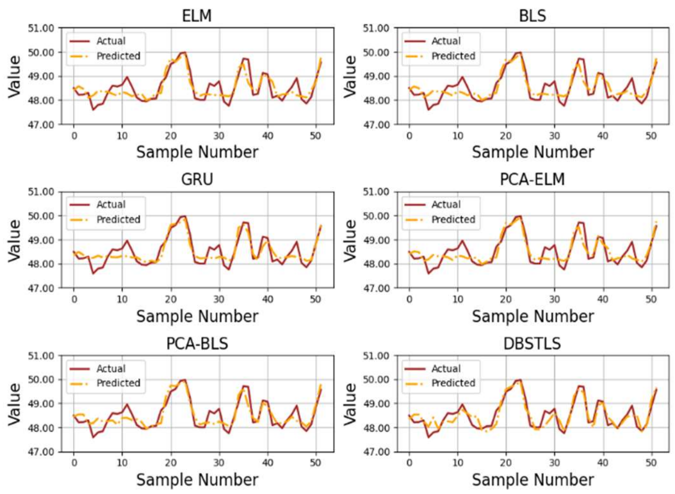

Figure 4. Predicted results of the six soft sensing models

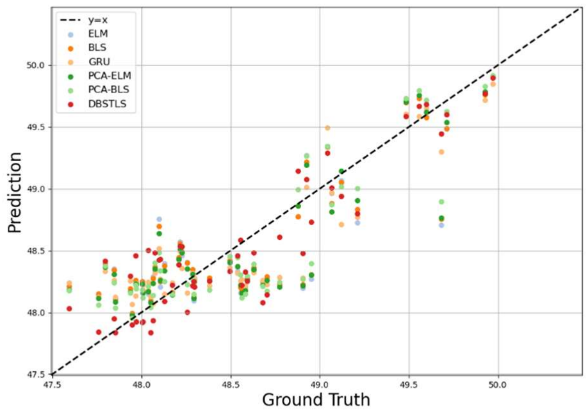

Figure 5. scatter plot of the residual errors by six models

To lucidly demonstrate the predictive performance, the outcomes of the six soft sensing models are visually depicted in Figure 4. Unlike the remaining five soft sensing models, the model grounded in DBSTLS consistently exhibits the closest approximation to the actual values, thus providing the most robust fit. Moreover, Figure 5 illustrates the distribution of absolute residuals for each model. Notably, the residuals of DBSTLS are prominently concentrated near the diagonal, indicative of the proposed model possessing the smallest prediction error range and the most tightly focused error distribution.

5. CONCLUSIONS

In this paper, a novel improved BLS named distributed broad seasonal trend learning system (DBSTLS) is proposed to develop industrial soft sensing with acceptable accuracy. In the proposed DBSTLS, there is a distributed structure based on Seasonal-Trend Decomposition Procedure Based on LOESS(STL). Through STL, the dynamic process data can be mainly decomposed into the following two sections: the trend feature and the season feature. The distributed structure can be built using the trend feature and the season feature, respectively. Then, the fast learning strategy in BLS is utilized in the distributed trend feature and season feature. Finally, the proposed DBSTLS can be developed as industrial soft sensing. The process data collected from the Pure Terephthalic acid process is used to verify the effectiveness of the proposed DBSTLS based industrial soft sensing. The simulation results verified that the proposed DBSTLS based industrial soft sensing could achieve higher accuracy compared with other models. The proposed DBSTLS provides an effective way to develop accurate industrial soft sensing for complex industrial processes.

REFERENCES

Chen, L., Xu, Y., Zhu, Q. X., & He, Y. L. (2023). Adaptive Multi-Head Self-Attention Based supervised VAE for Industrial Soft Sensing With Missing Data. IEEE Transactions on Automation Science and Engineering.

Chen, C. P., & Liu, Z. (2017). Broad learning system: An effective and efficient incremental learning system without the need for deep architecture. IEEE transactions on neural networks and learning systems, 29(1), 10-24.

Chin, K., Hellebrekers, T., & Majidi, C. (2020). Machine learning for soft robotic sensing and control. Advanced Intelligent Systems, 2(6), 1900171.

Cleveland, R. B., Cleveland, W. S., McRae, J. E., & Terpenning, I. (1990). STL: A seasonal-trend decomposition. J. Off. Stat, 6(1), 3-73.

Curreri, F., Graziani, S., & Xibilia, M. G. (2020). Input selection methods for data-driven Soft sensors design: Application to an industrial process. Information Sciences, 537, 1-17.

Gong, X., Zhang, T., Chen, C. P., & Liu, Z. (2021). Research review for broad learning system: Algorithms, theory, and applications. IEEE Transactions on Cybernetics, 52(9), 8922-8950.

Guo, R., Liu, H., Xie, G., Zhang, Y., & Liu, D. (2022). A self-interpretable soft sensor based on deep learning and multiple attention mechanism: From data selection to sensor modeling. IEEE Transactions on Industrial Informatics.

He, Y. L., Wang, P. F., & Zhu, Q. X. (2023). Improved Bi-LSTM With Distributed Nonlinear Extensions and Parallel Inputs for Soft Sensing. IEEE Transactions on Industrial Informatics.

Jiang, Y., Yin, S., Dong, J., & Kaynak, O. (2020). A review on soft sensors for monitoring, control, and optimization of industrial processes. IEEE Sensors Journal, 21(11), 12868-12881.

Kadlec, P., Gabrys, B., & Strandt, S. (2009). Data-driven soft sensors in the process industry. Computers & chemical engineering, 33(4), 795-814.

Li, Y., Bao, T., Gong, J., Shu, X., & Zhang, K. (2020). The prediction of dam displacement time series using STL, extra-trees, and stacked LSTM neural network. IEEE Access, 8, 94440-94452.

Rani, A., Singh, V., & Gupta, J. R. P. (2013). Development of soft sensor for neural network based control of distillation column. ISA transactions, 52(3), 438-449.

Shih, B., Shah, D., Li, J., Thuruthel, T. G., Park, Y. L., Iida, F., ... & Tolley, M. T. (2020). Electronic skins and machine learning for intelligent soft robots. Science Robotics, 5(41), eaaz9239.

Tian, Y., He, Y. L., & Zhu, Q. X. (2020). Soft sensor development using improved whale optimization and regularization-based functional link neural network. Industrial & Engineering Chemistry Research, 59(43), 19361-19369.

Wang, G., Jia, Q. S., Zhou, M., Bi, J., Qiao, J., & Abusorrah, A. (2022). Artificial neural networks for water quality soft-sensing in wastewater treatment: a review. Artificial Intelligence Review, 55(1), 565-587.

Yan, W., Tang, D., & Lin, Y. (2016). A data-driven soft sensor modeling method based on deep learning and its application. IEEE Transactions on Industrial Electronics, 64(5), 4237-4245.

Zhu, X., Rehman, K. U., Wang, B., & Shahzad, M. (2020). Modern soft-sensing modeling methods for fermentation processes. Sensors, 20(6), 1771.

Zhang, X., Zhu, Q., Jiang, Z. Y., He, Y., & Xu, Y. (2018). A novel ensemble model using PLSR integrated with multiple activation functions based ELM: Applications to soft sensor development. Chemometrics and Intelligent Laboratory Systems, 183, 147-157.