ACS OMEGA

Open Access

http://pubs.acs.org/journal/acsodf

Article

Adaptive Temporal−Spatial Pyramid Variational Autoencoder Model for Multirate Dynamic Chemical Process Soft Sensing Application

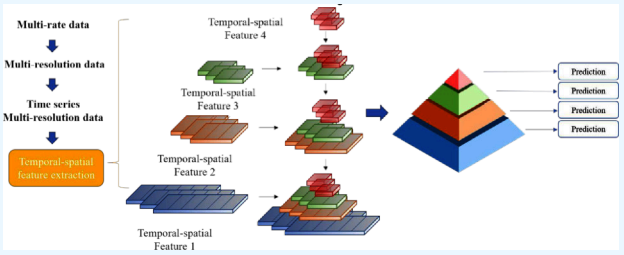
Bingbing Shen, Zeyu Yang,* and Le Yao*

Cite This: *ACS Omega* 2024, 9, 23021−23032

Read Online

ACCESS | Metrics & More | Article Recommendations

ABSTRACT: Data-driven soft sensors play an important role in practical processes and have been widely applied. They provide real-time prediction of quality variables and then guide production and improve product quality. In practical chemical production processes, nonlinear dynamic multirate data is widespread and challenging to model. This paper innovatively proposes a temporal−spatial pyramid variational autoencoder (TS-PVAE) model for the nonlinear temporal−spatial feature pyramid extraction from multirate data. This structure not only selectively utilizes multirate data but also handles complex nonlinear time-series data. Based on this, integrated with just-in-time (JIT) learning, an adaptive TS-PVAE (ATS-PVAE) model is developed. In this model, historical data are used for real-time fine-tuning of the model, leading to the development of an adaptive model. Finally, the proposed models are validated by an industrial case of a methanation furnace, demonstrating a superior estimation performance.



INTRODUCTION

As chemical processes develop, production processes become increasingly intricate, with a heightened level of integration in production equipment. Real-time monitoring is of great significance for increasing production, reducing energy consumption, and ensuring safety. At the same time, with the development of Internet of Things (IoT) technology, data acquisition has become easier. Therefore, there is a growing focus and challenge in analyzing and modeling process data to develop data-driven process monitoring models.[1−4] Quality variables are important indicators that can reflect the status of the production process.[5] However, their measurement is usually expensive and has time delays. Establishing data-driven soft sensor models for quality variables can predict them quickly and at low cost and improve production efficiency.

In the collected process data, there are various types of data, such as flow rate, pressure, temperature, and valve data. This data is obtained from different devices with varying sampling rates.[6] However, most existing models downsample process data according to the sampling rate of the quality variables. Information from the process data is lost during down-sampling. To effectively utilize a large amount of process variable information, some scholars have proposed semi-supervised learning methods.[7−9] For example, Yao and Ge proposed a hierarchical extreme learning machine based soft sensor for semisupervised process data.[10] These methods involve establishing semisupervised learning models by combining a large amount of unlabeled process data with a small amount of labeled process data, further extracting more data information and establishing more effective soft sensing models. However, semisupervised learning can only handle dual-rate data, remaining powerless when it comes to multirate data. To address multirate data modeling problems, some scholars have proposed multirate fault diagnosis models and soft sensing models.[11−13] Zhou et al. developed a multirate principal component regression model for soft sensor modeling.[14] Shen et al. proposed a pyramid variational autoencoder (PVAE) model for nonlinear multirate industrial data.[15]

The application of deep learning models in the field of soft sensing has become a highly regarded research direction.[16−18] These models can extract complex feature information from large-scale, high-dimensional process data and possess strong nonlinear modeling capabilities. Depending on the specific soft sensing problem and data characteristics, suitable deep learning model architectures can be developed, including deep neural networks (DNN), convolutional neural networks (CNN),[19−21] stacked autoencoder (SAE), variational autoencoder

Received: March 19, 2024
Revised: April 26, 2024
Accepted: May 7, 2024
Published: May 16, 2024

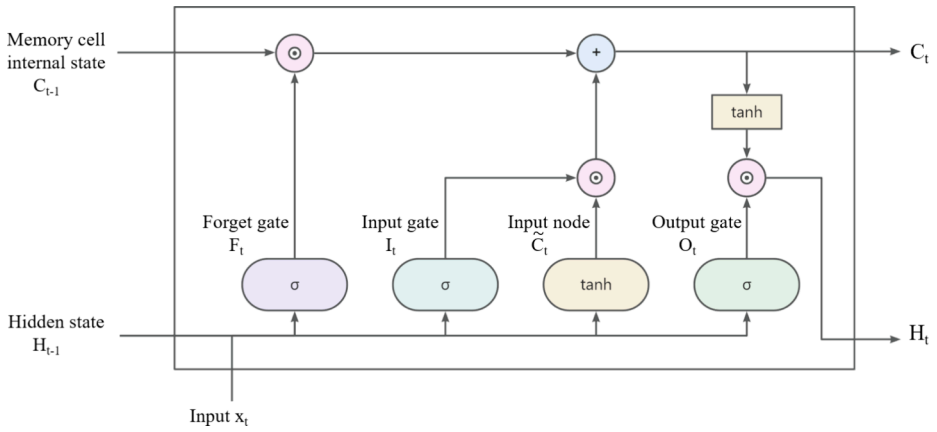


Figure 1. Structure of a LSTM.

(VAE),[22−24] etc. Yuan et al. designed a variable-wise weighted SAE for soft sensor modeling.[25]

In chemical processes, the process data often exhibits dynamic temporal characteristics, requiring the establishment of specific soft sensing models.[26−28] As a result, deep time-series soft sensing models based on methods such as recurrent neural networks (RNN), long short-term memory networks (LSTM),[29,30] Hidden markov models (HMM),[31,32] etc., have been developed. Jiang et al. built a dynamic temporal dependency model based on the similarities between the short-term load forecasting and the transformer architecture.[27] Dai et al. developed a time series denoising diffusion probabilistic model to enhance time-series samples for industrial soft sensing.[28] Liu et al. proposed a soft sensor based on correntropy long short-term memory for industrial polyethylene process.[33] Zheng et al. proposed a supervised hybrid CNN-LSTM network for nonlinear dynamic data soft sensor modeling.[34] Geng et al. designed a gated convolutional neural network based transformer regression model for industrial dynamic processes.[35]

Although dynamic network structures can retain long-term data information, memorable information is limited. In a prolonged dynamic process, it is necessary to combine existing dynamic information with similar historical information to adjust the model in real time to obtain the best adaptive soft sensor model. Just-in-time (JIT) learning[26,36] methods can identify similar historical data for test data, establish local models, and then achieve better prediction accuracy. Since JIT models can build models based on the target data, they are suitable for complex nonlinear data and multimode processes. In recent years, soft sensor models based on JIT learning have been widely developed and applied.[37] Recently, Xie et al. proposed a novel JIT model by combining with non-Gaussian information on the operation data.[38] Guo et al. developed a JIT learning soft sensor based on evolutional multiobjective optimization and VAE.[39] In order to solve the modeling problem of nonlinear dynamic multirate data, this paper needs to establish complex deep temporal networks. However, due to the complexity of the model and the abundance of data, JIT learning methods can help select a small amount of data for targeted adjustments and optimizations of existing networks.

In this paper, a temporal−spatial pyramid variational autoencoder (TS-PVAE) model is proposed to deal with the nonlinear dynamic multirate data modeling problem. The existing PVAE model is capable of effectively learning and establishing soft sensor models for multirate nonlinear data. TS-PVAE further dynamically extends multirate data, capturing the dynamic features of the data through LSTM and utilizing them for soft sensor modeling. Since dynamic modeling requires a large amount of data information, leading to the forgetting of historical information, to address this issue the concept of JIT learning is introduced, and an adaptive TS-PVAE (ATS-PVAE) model is proposed. This involves selecting historically similar data to adaptively adjust the existing model, thereby obtaining a more accurate prediction model.

The remainder of this work is organized as follows. In the Preliminaries, a brief introduction to LSTM and a review of the PVAE model are provided. Subsequently, this article presents a detailed introduction to the TS-PVAE model. Furthermore, the ATS-PVAE model is introduced. Finally, an industrial example is used to validate the effectiveness of the model.

PRELIMINARIES

LSTM. Long Short-Term Memory (LSTM) networks, a variant of RNNs, are tailored for processing sequential data by effectively capturing temporal dependencies with strong nonlinear capabilities. Comprising input gates, forget gates, output gates, and memory cells, LSTM architectures address the vanishing or exploding gradient issues inherent in traditional RNNs.

Mathematically, the behavior of an LSTM unit is described as follows:

$$i_t = \sigma(W_{ix}x_t + W_{ih}h_{t-1} + b_i) \quad (1)$$

$$f_t = \sigma(W_{fx}x_t + W_{fh}h_{t-1} + b_f) \quad (2)$$

$$o_t = \sigma(W_{ox}x_t + W_{oh}h_{t-1} + b_o) \quad (3)$$

$$g_t = \tanh(W_{gx}x_t + W_{gh}h_{t-1} + b_g) \quad (4)$$

$$c_t = f_t \odot c_{t-1} + i_t \odot g_t \quad (5)$$

$$h_t = o_t \odot \tanh(c_t) \quad (6)$$

Here, x_t is the input at time t; h_t is the hidden state; c_t is the cell state; and i_t, f_t, o_t, and g_t are the input gate, forget gate, output gate, and cell gate activations, respectively. W is weight matrices; b is bias vectors; σ denotes the sigmoid function; and $\odot$ denotes element-wise multiplication. The structure of an LSTM is given in Figure 1.

In the context of time series data, LSTM networks excel in capturing intricate patterns and dependencies, even in the

https://doi.org/10.1021/acsomega.4c02681
ACS Omega 2024, 9, 23021−23032

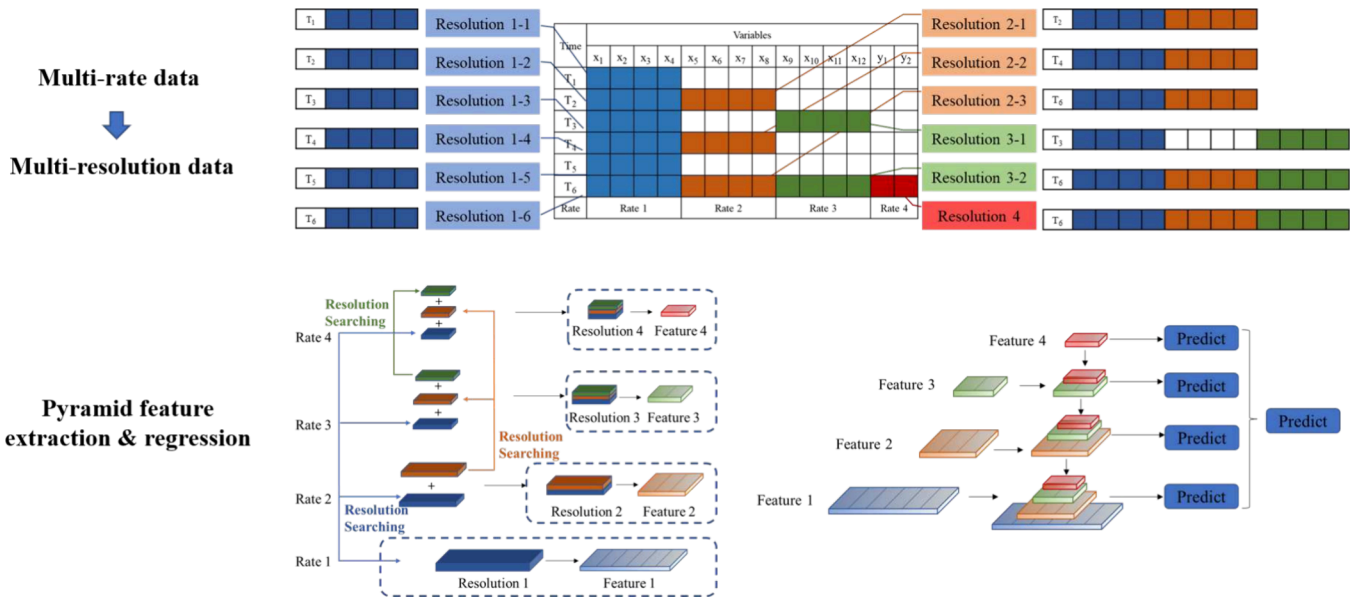


Figure 2. Diagram of the PVAE model.

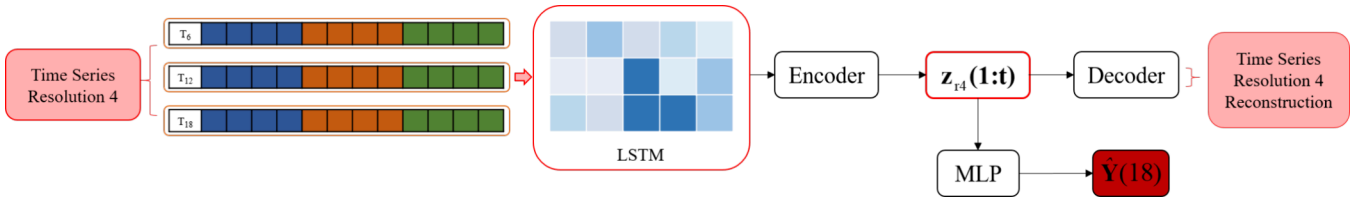


Figure 3. Structure of the TS-PVAE-r4 model.

presence of noise and irregularities. By leveraging their inherent nonlinear processing capabilities and robust memory management, LSTM networks offer a reliable framework for extracting salient features from time-varying signals, while preserving temporal dynamics.

Revisit of PVAE. The PVAE model first utilizes a multirate filter to restructure irregular multirate data into regular multiresolution data. Then, by utilizing VAEs to extract data features from different resolutions, feature pyramids with different resolutions are constructed for regression modeling. Figure 2 illustrates a specific example with four resolutions. In this figure, it is assumed that process variables x_1−x_4 are sampled every minute; x_5−x_8 are sampled every 2 min; x_9−x_{12} are sampled every 3 min; and quality variables y_1 and y_2 are sampled every 6 min. In this example, starting from each sampling rate, variables with higher sampling frequency at that time are found and concatenated to form four multiresolution data sets.

Afterward, multiresolution features are extracted from these reorganized four multiresolution data sets. Feature pyramids are constructed through these extracted features, along with lower-resolution features, and are used separately for prediction. The detailed model derivation and implementation can be found in ref 15.

PROPOSED TS-PVAE FOR SOFT SENSOR MODELING

To address the modeling problem of dynamic nonlinear multirate data, this paper proposes a TS-PVAE model. This model extends the PVAE model to a dynamic version. LSTM is embedded into encoders to learn and extract nonlinear features from time series data.

First, the multirate data filter of the PVAE model is used to reorganize the data into multiresolution data, and then the data are dynamically expanded to time series multiresolution data. Similar to the PVAE model, establishing a temporal−spatial pyramid model requires extracting each resolution data point separately to integrate low-resolution features into higher resolutions. Taking an example from the PVAE model, after restructuring, there are four resolution data sets. First, it is necessary to construct a temporal−spatial feature extraction model for temporal resolution 4. In this paper, LSTM is employed to process the temporal resolution 4 data; then temporal−spatial features for resolution 4 are extracted through the VAE model. These temporal−spatial features are used for data reconstruction, and a regression model is established using a multilayer perceptron (MLP). The specific training model diagram of TS-PVAE for the resolution 4 (TS-PVAE-r4) data set is shown in Figure 3.

Detailed derivation of the marginal likelihood for TS-PVAE-r4 in Figure 3 is shown in Appendix 1. The loss function of TS-PVAE-r4 can be given as

$$
\begin{aligned}
J_{\text{TS-PVAER}}(r_4(1:t)) &= -\mathbb{E}_{\mathbf{z}_{r_4}(1:t)\sim q(\mathbf{z}_{r_4}(1:t)|\mathbf{x}_{r_4}(1:t))} \\
&[\ln p(\mathbf{x}_{r_4}(1:t)|\mathbf{z}_{r_4}(1:t)) + \ln p(\mathbf{y}_{r_4}|\mathbf{z}_{r_4}(1:t))] \\
&+ D_{KL}(q(\mathbf{z}_{r_4}(1:t)|\mathbf{x}_{r_4}(1:t))\|p(\mathbf{z}_{r_4}(1:t))) \\
&= \|\mathbf{x}_{r_4}(1:t) - \boldsymbol{\mu}_{\mathbf{x}_{r_4}(1:t)}\|^2 + \|\mathbf{y}_{r_4} - \boldsymbol{\mu}_{\mathbf{y}_{r_4}}\|^2 \\
&+ D_{KL}(q(\mathbf{z}_{r_4}(1:t)|\mathbf{x}_{r_4}(1:t))\|N(0, \mathbf{I}))
\end{aligned} \tag{7}
$$

https://doi.org/10.1021/acsomega.4c02681
ACS Omega 2024, 9, 23021−23032

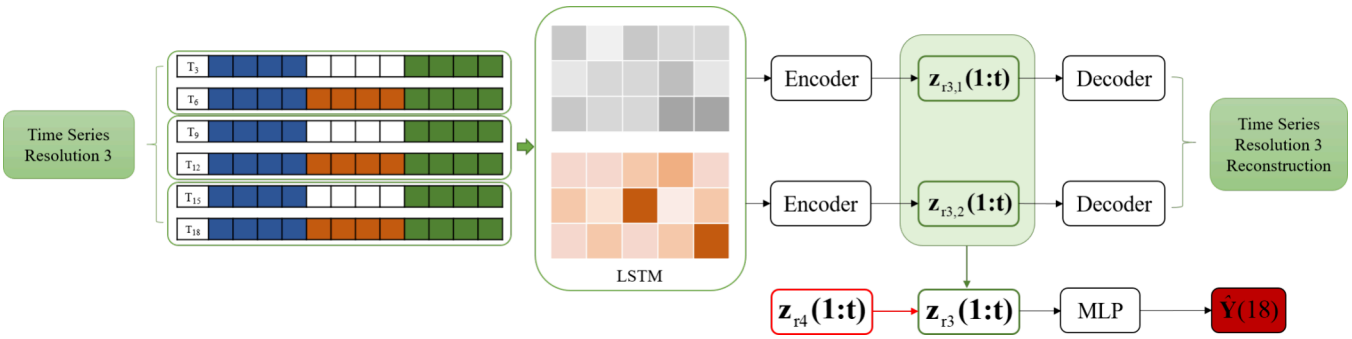


Figure 4. Structure of the TS-PVAE-r3 model.

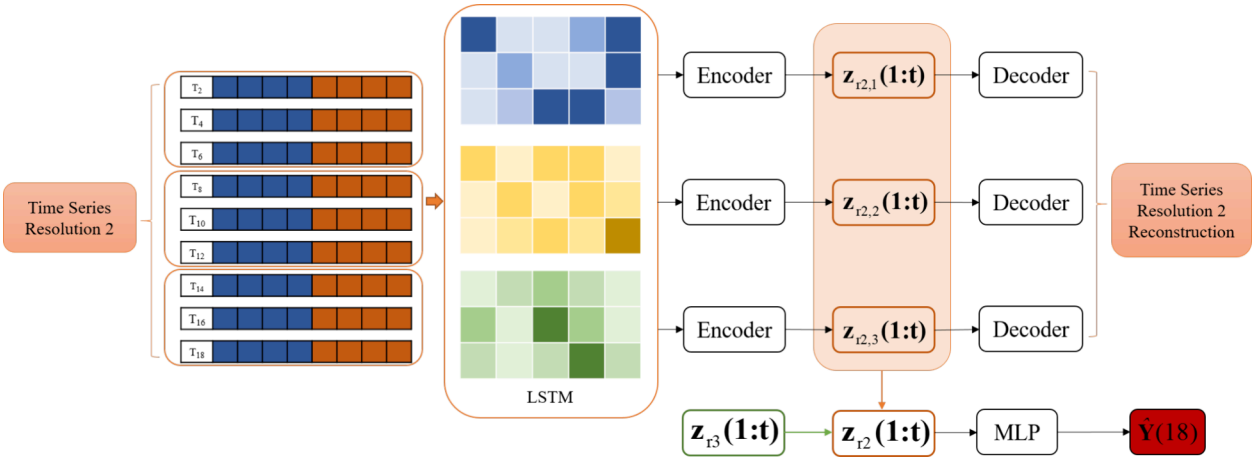


Figure 5. Structure of the TS-PVAE-r2 model.

where $\mathbf{x}(1{:}t)$ represents the process time series variables; $\mathbf{y}(1{:}t)$ represents the quality time series variable; $\mathbf{z}(1{:}t)$ represents the latent variables from time series input; subscript $r_4(1{:}t)$ indicates time series resolution 4; and it is assumed that the relationships of $\mathbf{x}(1{:}t)$ and $\mathbf{y}(1{:}t)$ are independent of each other conditionally.

The trained TS-PVAE-r4 model provides temporal−spatial features from the time series resolution 4 data set, serving as the first layer of the feature pyramid. These top-level temporal−spatial features are further introduced into the TS-PVAE for resolution 3 (TS-PVAE-r3) for modeling. The specific model structure is illustrated in Figure 4.

According to the model structure in Figure 4, the loss function for TS-PVAE-r_3 can be derived as

$$
\begin{aligned}
J_{\text{TS-PVAER}}(r_3(1{:}\,t)) = &-\alpha\sum_{i=1}^{2}\mathbb{E}_{q(\mathbf{z}_{r_{3,i}}(1:t)|\mathbf{x}_{r_{3,i}}(1:t))}\\
&[\ln p(\mathbf{x}_{r_{3,i}}(1{:}\,t)(1{:}\,t)|\mathbf{z}_{r_{3,i}}(1{:}\,t))] - \mathbb{E}_{q(\mathbf{z}_{r_3}(1:t)|\mathbf{z}_{r_{3,1}}(1:t),\mathbf{z}_{r_{3,2}}(1:t),\mathbf{z}_{r_4}(1:t))}\\
&[\ln p(\mathbf{y}_{r_4}|\mathbf{z}_{r_3}(1{:}\,t))] + \sum_{i=1}^{2} D_{KL}(q(\mathbf{z}_{r_{3,i}}(1{:}\,t)|\mathbf{x}_{r_{3,i}}(1{:}\,t)(1{:}\,t))\\
&\|p(\mathbf{z}_{r_{3,i}}(1{:}\,t))) = \alpha\sum_{i=1}^{2}\|\mathbf{x}_{r_{3,i}}(1{:}t)(1{:}\,t) - \boldsymbol{\mu}_{\mathbf{x}_{r_{3,i}}(1:t)(1:t)}\\
&\|^2 + \|\mathbf{y}_{r_4} - \boldsymbol{\mu}_{\mathbf{y}_{r_4}}\|^2 + \sum_{i=1}^{2} D_{KL}(q(\mathbf{z}_{r_{3,i}}(1{:}\,t)|\mathbf{x}_{r_{3,i}}(1{:}\,t)(1{:}\,t))\\
&\|N(0,\mathbf{I}))
\end{aligned}
\tag{8}
$$

where subscript $r_3(1{:}t)$ indicates the time series resolution 3 data set and subscript $r_{3,i}(1{:}t)$ represents time series resolution $3-i$ $(i=1,2)$. Detailed derivation of the marginal likelihood for TS-PVAE-r3 is given in Appendix 2. To address the data imbalance problem caused by different sampling rates in different networks, a balance coefficient α is introduced into this loss function.

In the TS-PVAE-r3 model, a two-layer temporal−spatial feature pyramid is constructed and used for predicting quality variables. Furthermore, the two-layer temporal−spatial features are introduced into the TS-PVAE-r2 model to extract the third layer temporal−spatial features and build the regression model. The model structure is illustrated in Figure 5.

Then, the loss function of TS-PVAE for resolution 2 (TS-PVAE-r2) in Figure 5 can be given as

$$
\begin{aligned}
J_{\text{TS-PVAER}}(r_2(1{:}\,t)) = &-\beta\sum_{i=1}^{3}\mathbb{E}_{q(\mathbf{z}_{r_{2,i}}(1:t)|\mathbf{x}_{r_{2,i}}(1:t))}\\
&[\ln p(\mathbf{x}_{r_{2,i}}(1{:}\,t)|\mathbf{z}_{r_{2,i}}(1{:}\,t))] - \mathbb{E}_{q(\mathbf{z}_{r_2}|\mathbf{z}_{r_{2,1}}(1:t),\mathbf{z}_{r_{2,2}}(1:t),\mathbf{z}_{r_{2,3}}(1:t),\mathbf{z}_{r_3}(1:t))}\\
&[\ln p(\mathbf{y}_{r_4}|\mathbf{z}_{r_2})] + \sum_{i=1}^{3} D_{KL}(q(\mathbf{z}_{r_{2,i}}(1{:}\,t)|\mathbf{x}_{r_{2,i}}(1{:}\,t))\|p(\mathbf{z}_{r_{2,i}}(1{:}\,t)))\\
&= \beta\sum_{i=1}^{3}\|\mathbf{x}_{r_{2,i}}(1{:}t) - \boldsymbol{\mu}_{\mathbf{x}_{r_{2,i}}(1:t)}\|^2 + \|\mathbf{y}_{r_4} - \boldsymbol{\mu}_{\mathbf{y}_{r_4}}\|^2\\
&+ \sum_{i=1}^{3} D_{KL}(q(\mathbf{z}_{r_{2,i}}(1{:}\,t)|\mathbf{x}_{r_{2,i}}(1{:}\,t))\|N(0,\mathbf{I}))
\end{aligned}
\tag{9}
$$

Here, $r_2(1{:}t)$ indicates the time series resolution 2 data set, and subscript $r_{2,i}(1{:}t)$ represents time series resolution $2-i$ $(i=1,2,3)$. Similar to the above derivation, β is introduced as a balance coefficient. Detailed derivation of the marginal likelihood for TS-PVAE-r2 is given in Appendix 3.

https://doi.org/10.1021/acsomega.4c02681
ACS Omega 2024, 9, 23021−23032

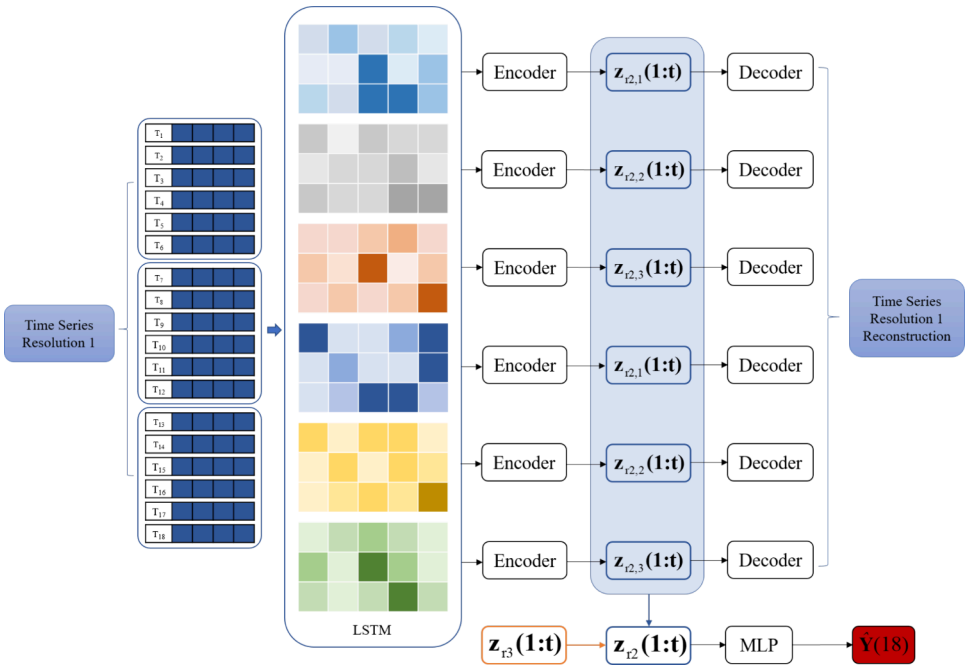


Figure 6. Structure of the TS-PVAE-r1 model.

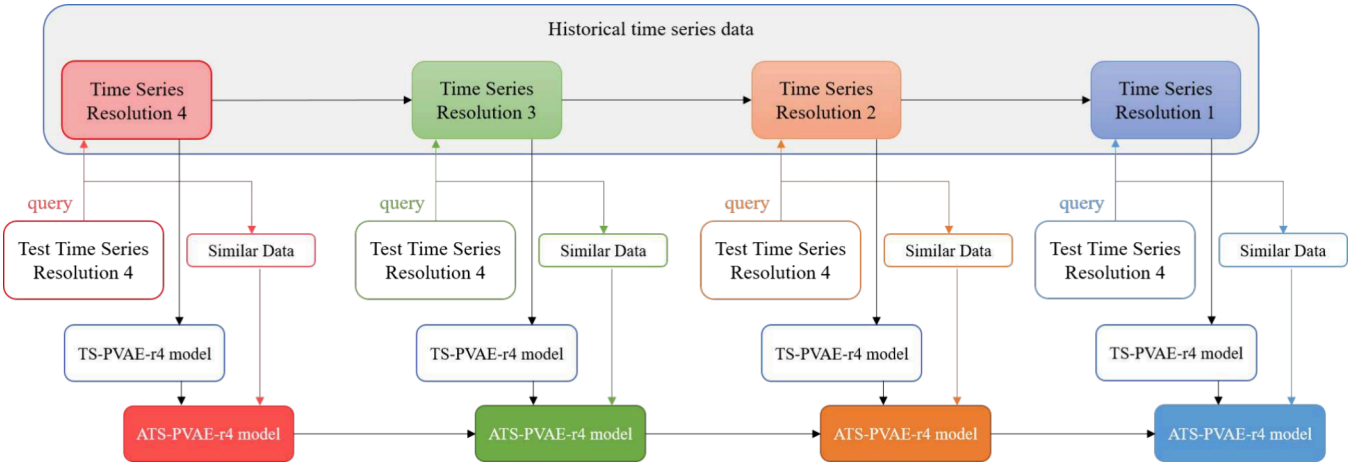


Figure 7. Flowchart of the ATS-PVAE model.

In the TS-PVAE for resolution 1 (TS-PVAE-r1), a three-layer temporal−spatial feature pyramid is introduced. The model structure is shown in Figure 6.

Further, the loss function of TS-PVAE-r1 can be written as

$$
\begin{aligned}
J_{\text{TS-PVAER}}(r_1(1{:}\,t)) = &-\gamma\sum_{i=1}^{6}\mathbb{E}_{q(\mathbf{z}_{r_{1,i}}(1:t)|\mathbf{x}_{r_{1,i}}(1:t))}\\
&[\ln p(\mathbf{x}_{r_{1,i}}(1{:}\,t)|\mathbf{z}_{r_{1,i}}(1{:}\,t))] - \mathbb{E}_{q(\mathbf{z}_{r_1}(1:t)|\mathbf{z}_{r_{1,1}}(1:t),\ldots,\mathbf{z}_{r_{1,6}}(1:t),\mathbf{z}_{r_2}(1:t))}\\
&[\ln p(\mathbf{y}_{r_4}|\mathbf{z}_{r_1}(1{:}\,t))] + \sum_{i=1}^{6} D_{KL}(q(\mathbf{z}_{r_{1,i}}(1{:}\,t)|\mathbf{x}_{r_{1,i}}(1{:}\,t))\|p(\mathbf{z}_{r_{1,i}}(1{:}\,t)))\\
&= \gamma\sum_{i=1}^{6}\|\mathbf{x}_{r_{1,i}}(1{:}t) - \boldsymbol{\mu}_{\mathbf{x}_{r_{1,i}}(1:t)}\|^2 + \|\mathbf{y}_{r_4} - \boldsymbol{\mu}_{\mathbf{y}_{r_4}}\|^2\\
&+ \sum_{i=1}^{6} D_{KL}(q(\mathbf{z}_{r_{1,i}}(1{:}\,t)|\mathbf{x}_{r_{1,i}}(1{:}\,t))\|N(0,\mathbf{I}))
\end{aligned}
\tag{10}
$$

where subscript $r_1(1{:}t)$ indicates the time series resolution 1 data set; subscript $r_{1,i}(1{:}t)$ represents time series resolution 1 − i (i = 1, ..., 6); and γ is also a balance coefficient. Detailed derivation of the marginal likelihood for TS-PVAE-r1 is given in Appendix 4.

PROPOSED ATS-PVAE FOR SOFT SENSOR MODELING

To further utilize historical data for adjusting and optimizing the model, an ATS-PVAE model (Figure 7) based on just-in-time (JIT) learning is proposed in this section. This model adjusts and optimizes the trained model by searching for similar data sets in historical data. For the TS-PVAE-ri model (i represents the resolution, and n_{r_i} data are selected by comparing the Euclidean distance between the current process data with prediction and the historical process data).

$$
D_{r_i} = \sqrt{(\mathbf{x}_{r_i,j} - \mathbf{x}_{r_i,q})^T(\mathbf{x}_{r_i,j} - \mathbf{x}_{r_i,q})} \tag{11}
$$

https://doi.org/10.1021/acsomega.4c02681
ACS Omega 2024, 9, 23021−23032

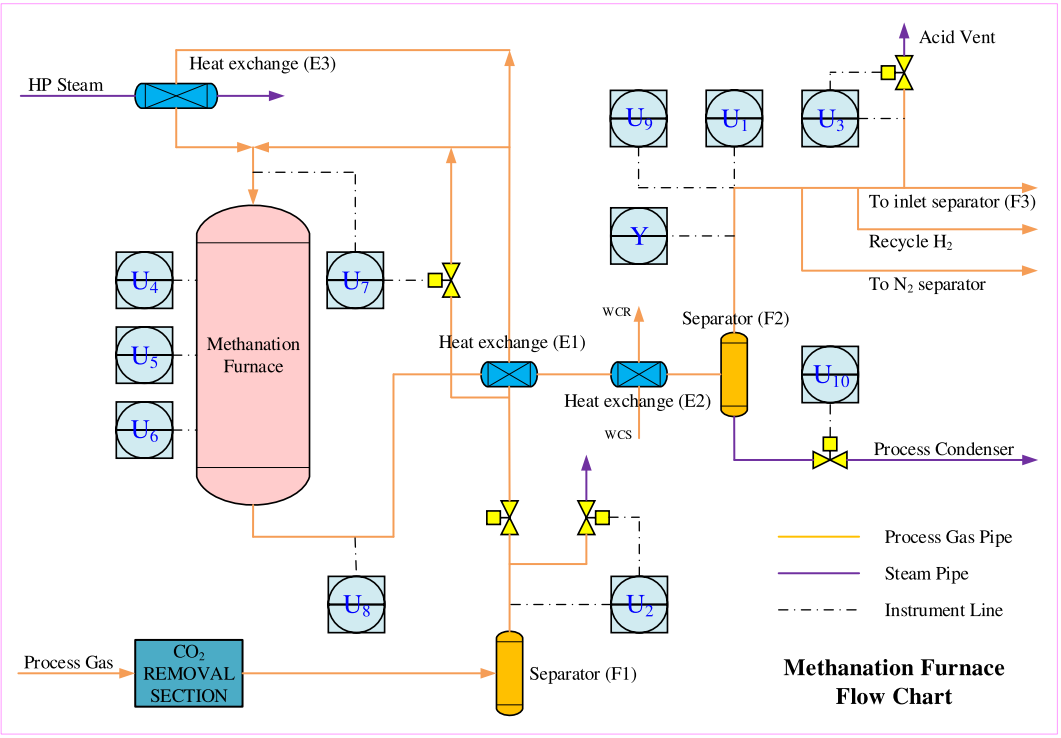


Figure 8. Flowchart of the methanation furnace.

where q represents the current time-series test data to be queried, while j represents the time-series data from the historical training set.

CASE STUDY

Methanation Furnace. In the production process of synthetic ammonia, nitrogen (N_2) and hydrogen (H_2) are the main feed gases, typically mixed in a certain ratio and fed into the reactor for the Haber−Bosch process. Hydrogen is usually prepared from natural gas or other hydrocarbons through methods such as steam reforming or partial oxidation, which may produce carbon monoxide and carbon dioxide as byproducts.

Before entering the ammonia synthesis reactor, it is necessary to remove these byproducts through the methanation furnace unit. In the methanation furnace unit, catalyzed by nickel and under high-temperature conditions, the following chemical reactions occur:

$$\begin{aligned} CO + 3H_2 &\leftrightarrow CH_4 + H_2O + Q \\ CO_2 + 4H_2 &\leftrightarrow CH_4 + 2H_2O + Q \end{aligned} \quad (12)$$

The purpose of this unit is to remove carbon monoxide and carbon dioxide from the feed gases as much as possible, making the residual levels of these gases at the outlet crucial quality variables. In practical production processes, precise measurement of carbon monoxide and carbon dioxide residuals relies on expensive mass spectrometers in laboratory settings, which is not feasible for large-scale synthetic ammonia processes.

Hence, it is possible to establish a soft sensor model linking easily measurable process variables with quality variables representing carbon monoxide and carbon dioxide residuals, enabling timely assessment and adjustment of the methanation furnace unit's production process and quality conditions. Operators of this facilitation can adjust feed gas proportions effectively based on the soft sensor to prevent production accidents. Figure 8 illustrates the production process of the methanation furnace unit, and Table 1 presents the process variables collected from the methanation furnace unit.

Table 1. Description of the Process Instruments in Methanation Furnace

Tags	Descriptions
U1	Flow rate of process gas at MF's exit
U2	Pressure of process gas at MF's entrance
U3	Pressure of process gas into F3's entrance
U4	Temperature of upper bed at MF
U5	Temperature of middle bed at MF
U6	Temperature of lower bed at MF
U7	Temperature of process gas at MF's exit
U8	Temperature of process gas at MF's exit
U9	Temperature of synthesis gas at F2's exit
U10	Liquid level of F2
Y	Content of CO and CO_2 at MF's exit

Results and Analysis. In this process, due to the different technical conditions of the measuring instruments and the discrepant nature of the sensors, the process variables are in multiple sampling rates. The time span of the collected data set for modeling is 12000 min. Among the 11 variables, the U_1, ⋯, U_5 are sampled every minute; U_6, ⋯, Ux_8 are sampled every 2 min; U_9, ⋯, U_{10} are sampled every 3 min; and the quality variable y is sampled every 30 min. In this case, the data of the first 6000 min are used for training, while the rest of the 6000 data points are used for testing. Then, in total 6000 U_1 to U_5, 3000 samples of U_6 to U_8, 2000 samples of U_9 to U_{10}, and 200 samples of y are prepared for model training. The remaining 6000 samples of U_1 to U_5, 3000 samples of U_6 to U_8, 2000 samples of U_9 to U_{10}, and 200 samples of y are taken as the

https://doi.org/10.1021/acsomega.4c02681
ACS Omega 2024, 9, 23021−23032

testing data set. The balance coefficients α, β, and γ are also set to 1/2, 1/3, and 1/6, respectively. According to the principle of the multirate data filter in the PVAE model, we selected the least common multiple of the four sampling rates to partition the data. In this vase, every 6 time steps form one group of data. Since y is sampled every 30 min, 5 groups of data can be used to predict y using a dynamic network. The specific composition of the time-series data is shown in Table 2.

Table 2. Reorganized Time Series Multi-Resolution Data

	Subresolution	1:t = 1:5	Reorganized data
Time Series Resolution 4		T_6, T_{12}, ⋯, T_{30}	U_1−U_{10}
Time Series Resolution 3	Resolution 3−1	T_3, T_9, ⋯, T_{27}	U_1−U_5, U_9−U_{10}
	Resolution 3−2	T_6, T_{12}, ⋯, T_{30}	U_1−U_{10}
Time Series Resolution 2	Resolution 2−1	T_2, T_8, ⋯, T_{26}	U_1−U_8
	Resolution 2−2	T_4, T_{10}, ⋯, T_{28}	U_1−U_8
	Resolution 2−3	T_6, T_{12}, ⋯, T_{30}	U_1−U_8
Time Series Resolution 1	Resolution 1−1	T_1−T_{25}	U_1−U_5
	Resolution 1−2	T_2−T_{26}	U_1−U_5
	Resolution 1−3	T_3−T_{27}	U_1−U_5
	Resolution 1−4	T_4−T_{28}	U_1−U_5
	Resolution 1−5	T_5−T_{29}	U_1−U_5
	Resolution 1−6	T_6−T_{30}	U_1−U_5

The prediction performances of the proposed TS-PVAE model and ATS-PVAE model are compared with dimensionality-reduced LSTM and LSTM trained on all multirate data (Mr-LSTM). The detailed model structure and parameters of the proposed models and comparison models are shown in Table 3. The mean absolute error (MAE) and root means square error (RMSE) are taken as indicators to measure the experimental results.

$$\mathrm{MAE} = \sum_{i=1}^{N} |y_i - \hat{y}_i|/N \tag{13}$$

$$\mathrm{RMSE} = \sqrt{\sum_{i=1}^{N} (y_i - \hat{y}_i)^2/N} \tag{14}$$

where y_i is the real value; $\hat{y}_i$ is the prediction value; N is the length of query samples; L is the length of labeled samples; and $\overline{y_l}$ is the mean value of the labeled outputs. MAE and RMSE can measure the predicted error. Generally, the best predictions correspond to the lowest MAE and RMSE.

The ATS-PVAE model and the TS-PVAE model share the same network structure, but the network parameters are fine-tuned for the TS model based on 50 selected similar data sets. The evaluation indices are listed in Tables 4−Table 6. The prediction results of TS-PVAE-r4, TS-PVAE-r2, ATS-PVAE-r4, and ATS-PVAE-r2 are shown in Figures 9−12.

Table 4. MAE of the Proposed Models in the Methanation Furnace

	r4	r3	r2	r1
TS-PVAE	0.1218	0.1223	0.1209	0.1202
ATS-PVAE	0.1143	0.1131	**0.1035**	0.1092

Table 5. RMSE of the Proposed Models in the Methanation Furnace

	r4	r3	r2	r1
TS-PVAE	0.1489	0.1480	**0.1465**	0.1483
ATS-PVAE	0.1389	0.1426	**0.1307**	0.1329

Table 6. MAE and RMSE of LSTM and MR-LSTM in the Methanation Furnace

	MAE	RMSE
LSTM	0.1703	0.2178
MR-LSTM	0.1617	0.1990

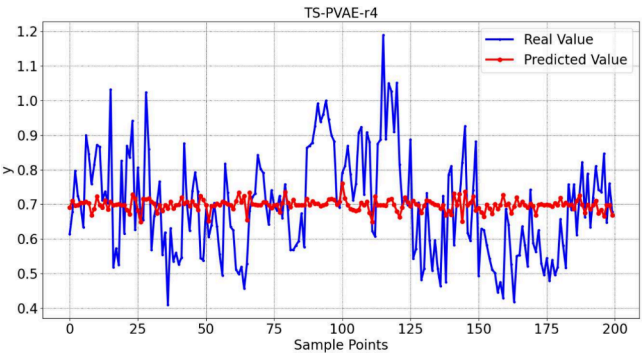


Figure 9. Prediction performance (TS-PVAE-r4).

From Figures 9−12 and Table 5, it can be observed that as the temporal−spatial feature pyramid is constructed the TS-PVAE-r2 model achieves the best estimation results when using a three-layer pyramid. The LSTM model only uses down-sampled data, and the model with less data is worse. The MR-LSTM model uses all the multirate data, although the model effect has been improved, but the data has not been reorganized. The temporal−spatial relationship of the data is

Table 3. Model Structure and Parameters

	LSTM	Encoder	Latent variable	Decoder	Regression	Epoch
TS-PVAE-r4	[20,10]	[15,10]	[8]	[10,15]	[10,5]	300
TS-PVAE-r3	[20,10]	[10,8] × 2	[8] × 2	[8,10] × 2	[10,5]	300
TS-PVAE-r2	[20,10]	[8,8] × 3	[8] × 3	[8,8] × 3	[10,5]	400
TS-PVAE-r1	[20,10]	[8,8] × 6	[8] × 6	[8,8] × 6	[10,5]	500
LSTM	[20,10]	/	/	/	/	1000
MR-LSTM	[20,10]	/	/	/	/	1000

https://doi.org/10.1021/acsomega.4c02681
ACS Omega 2024, 9, 23021−23032

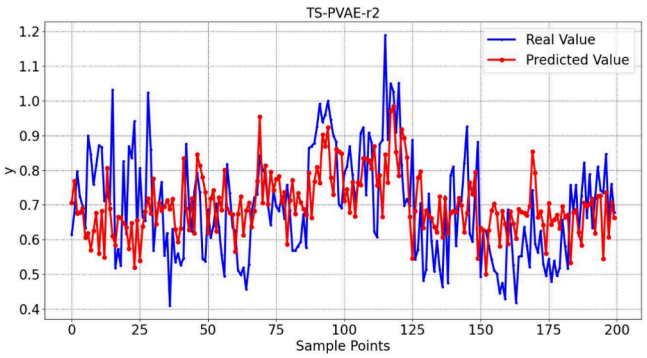


Figure 10. Prediction performance (TS-PVAE-r2).

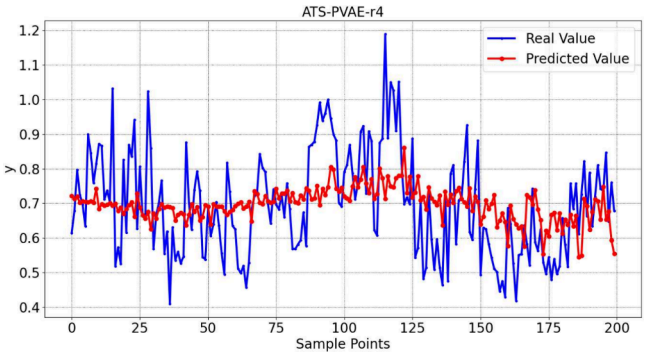


Figure 11. Prediction performance (ATS-PVAE-r4).

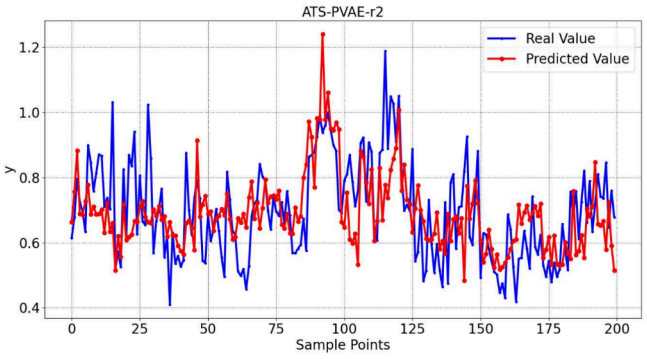


Figure 12. Prediction performance (ATS-PVAE-r2).

complicated, so the model effect is not good. Moreover, it can be found that compared to TS-PVAE mdoels ATS-PVAE models have achieved an overall improvement in performance. Additionally, ATS-PVAE-r2 achieves the best prediction performance. In other words, fine-tuning the model with a small amount of similar time series data can enhance the prediction performance of the model.

CONCLUSIONS

In this paper, a TS-PVAE model is proposed to build a soft sensor for nonlinear dynamic multirate data. The multirate data can be reorganized through the basic PVAE model. Then, the TS-PVAE model extracts nonlinear temporal−spatial features from multiple-time series resolution data. These features construct multiple temporal−spatial feature pyramids for regression modeling. As the feature pyramid expands, the prediction accuracy of the regression model improves. When there is sufficient temporal−spatial feature information, the model achieves an optimal prediction performance. However, when there is information redundancy, the model prediction performance deteriorates. To obtain the best prediction performance of the model, this paper integrated the JIT strategy, utilizing a small amount of similar data for model fine-tuning, and developed the ATS-PVAE model. Finally, a real industrial case has validated the effectiveness and superiority of the ATS-PVAE model. Furthermore, this model can be used for a quality related model. Related feature selection can be carried out through the correlation with quality variables in model building.

APPENDIX

1. The detailed derivation of the marginal likelihood for TS-PVAE-r4 can be given as

$$\ln p(\mathbf{x}_{r_4}(1:t), \mathbf{y}_{r_4}) = \int_{\mathbf{z}_{r_4}(1:t)} q(\mathbf{z}_{r_4}(1:t)) \left[\ln \frac{p(\mathbf{x}_{r_4}(1:t), \mathbf{y}_{r_4}, \mathbf{z}_{r_4}(1:t))}{q(\mathbf{z}_{r_4}(1:t))} - \ln \frac{p(\mathbf{z}_{r_4}(1:t)|\mathbf{x}_{r_4}(1:t), \mathbf{y}_{r_4})}{q(\mathbf{z}_{r_4}(1:t))} \right] \mathrm{d}\mathbf{z}_{r_4}(1:t) = \int_{\mathbf{z}_{r_4}(1:t)} q(\mathbf{z}_{r_4}(1:t))$$
$$\ln \frac{p(\mathbf{x}_{r_4}(1:t), \mathbf{y}_{r_4}, \mathbf{z}_{r_4}(1:t))}{q(\mathbf{z}_{r_4}(1:t))} \mathrm{d}\mathbf{z}_{r_4}(1:t) - \int_{\mathbf{z}_{r_4}(1:t)} q(\mathbf{z}_{r_4}(1:t)) \ln \frac{p(\mathbf{z}_{r_4}(1:t)|\mathbf{x}_{r_4}(1:t), \mathbf{y}_{r_4})}{q(\mathbf{z}_{r_4}(1:t))} \mathrm{d}\mathbf{z}_{r_4}(1:t) = \int_{\mathbf{z}_{r_4}(1:t)} q(\mathbf{z}_{r_4}(1:t))$$
$$\ln \frac{p(\mathbf{x}_{r_4}(1:t), \mathbf{y}_{r_4}, \mathbf{z}_{r_4}(1:t))}{q(\mathbf{z}_{r_4}(1:t))} \mathrm{d}\mathbf{z}_{r_4}(1:t) + D_{\mathrm{KL}}(q(\mathbf{z}_{r_4}(1:t)) \| p(\mathbf{z}_{r_4}(1:t)|\mathbf{x}_{r_4}(1:t), \mathbf{y}_{r_4})) = \mathrm{ELBO}_{\mathrm{TS-PVAE}}(r_4) + D_{\mathrm{KL}}(q(\mathbf{z}_{r_4}(1:t))$$
$$\| p(\mathbf{z}_{r_4}(1:t)|\mathbf{x}_{r_4}(1:t), \mathbf{y}_{r_4})) \geq \mathrm{ELBO}_{\mathrm{TS-PVAE}}(r_4) = \int_{\mathbf{z}_{r_4}(1:t)} q(\mathbf{z}_{r_4}(1:t)) \ln \frac{p(\mathbf{x}_{r_4}(1:t)|\mathbf{z}_{r_4}(1:t)) p(\mathbf{y}_{r_4}|\mathbf{z}_{r_4}(1:t)) p(\mathbf{z}_{r_4}(1:t))}{q(\mathbf{z}_{r_4}(1:t))} \mathrm{d}\mathbf{z}_{r_4}(1:t)$$
$$= \mathbb{E}_{\mathbf{z}_{r_4}(1:t) \sim q(\mathbf{z}_{r_4}(1:t)|\mathbf{x}_{r_4}(1:t))} [\ln p(\mathbf{x}_{r_4}(1:t)|\mathbf{z}_{r_4}(1:t)) + \ln p(\mathbf{y}_{r_4}|\mathbf{z}_{r_4}(1:t))] - D_{KL}(q(\mathbf{z}_{r_4}(1:t)|\mathbf{x}_{r_4}(1:t)) \| p(\mathbf{z}_{r_4}(1:t))) \quad (15)$$

2. The detailed derivation of the marginal likelihood for TS-PVAE-r3 can be given as

https://doi.org/10.1021/acsomega.4c02681
ACS Omega 2024, 9, 23021−23032

$$\ln p(\mathbf{x}_{r_3}(1:t), \mathbf{y}_{r_4}, \mathbf{z}_{r_4}(1:t)) = \ln p(\mathbf{x}_{r_{3,1}}(1:t), \mathbf{x}_{r_{3,2}}(1:t), \mathbf{y}_{r_4}, \mathbf{z}_{r_4}(1:t)) = \int_{\mathbf{z}_{r_{3,1}}(1:t)} \int_{\mathbf{z}_{r_{3,2}}(1:t)} q(\mathbf{z}_{r_{3,1}}(1:t))$$

$$q(\mathbf{z}_{r_{3,2}}(1:t)) \times \left[\ln \frac{p(\mathbf{x}_{r_{3,1}}(1:t), \mathbf{x}_{r_{3,2}}(1:t), \mathbf{y}_{r_4}, \mathbf{z}_{r_{3,1}}(1:t), \mathbf{z}_{r_{3,2}}(1:t), \mathbf{z}_{r_4}(1:t))}{q(\mathbf{z}_{r_{3,1}}(1:t)) q(\mathbf{z}_{r_{3,2}}(1:t))} - \ln \frac{p(\mathbf{z}_{r_{3,1}}(1:t) | \mathbf{x}_{r_{3,1}}(1:t), \mathbf{x}_{r_{3,2}}(1:t), \mathbf{y}_{r_4}, \mathbf{z}_{r_4}(1:t))}{q(\mathbf{z}_{r_{3,1}}(1:t))} \right.$$

$$\left. - \ln \frac{p(\mathbf{z}_{r_{3,2}}(1:t) | \mathbf{x}_{r_{3,1}}(1:t), \mathbf{x}_{r_{3,2}}(1:t), \mathbf{y}_{r_4}, \mathbf{z}_{r_4}(1:t))}{q(\mathbf{z}_{r_{3,2}}(1:t))} \right] d\mathbf{z}_{r_{3,1}}(1:t) d\mathbf{z}_{r_{3,2}}(1:t) = \int_{\mathbf{z}_{r_{3,1}}(1:t)} \int_{\mathbf{z}_{r_{3,2}}(1:t)} q(\mathbf{z}_{r_{3,1}}(1:t))$$

$$q(\mathbf{z}_{r_{3,2}}(1:t)) \times \ln \frac{p(\mathbf{x}_{r_{3,1}}(1:t), \mathbf{x}_{r_{3,2}}(1:t), \mathbf{y}_{r_4}, \mathbf{z}_{r_{3,1}}(1:t), \mathbf{z}_{r_{3,2}}(1:t), \mathbf{z}_{r_4}(1:t))}{q(\mathbf{z}_{r_{3,1}}(1:t)) q(\mathbf{z}_{r_{3,2}}(1:t))} d\mathbf{z}_{r_{3,1}}(1:t) d\mathbf{z}_{r_{3,2}}(1:t)$$

$$- \sum_{i=1}^{2} \int_{\mathbf{z}_{r_{3,i}}(1:t)} q(\mathbf{z}_{r_{3,i}}(1:t)) \ln \frac{p(\mathbf{z}_{r_{3,i}}(1:t) | \mathbf{x}_{r_{3,1}}(1:t), \mathbf{x}_{r_{3,2}}(1:t), \mathbf{y}_{r_4}, \mathbf{z}_{r_4}(1:t))}{q(\mathbf{z}_{r_{3,i}}(1:t))} d\mathbf{z}_{r_{3,i}}(1:t) = \int_{\mathbf{z}_{r_{3,1}}(1:t)} \int_{\mathbf{z}_{r_{3,2}}(1:t)} q(\mathbf{z}_{r_{3,1}}(1:t))$$

$$q(\mathbf{z}_{r_{3,2}}(1:t)) \times \ln \frac{p(\mathbf{x}_{r_{3,1}}(1:t), \mathbf{x}_{r_{3,2}}(1:t), \mathbf{y}_{r_4}, \mathbf{z}_{r_{3,1}}(1:t), \mathbf{z}_{r_{3,2}}(1:t), \mathbf{z}_{r_4}(1:t))}{q(\mathbf{z}_{r_{3,1}}(1:t)) q(\mathbf{z}_{r_{3,2}}(1:t))} d\mathbf{z}_{r_{3,1}}(1:t) d\mathbf{z}_{r_{3,2}}(1:t)$$

$$+ \sum_{i=1}^{2} D_{KL}(q(\mathbf{z}_{r_{3,i}}(1:t)) \| p(\mathbf{z}_{r_{3,i}}(1:t) | \mathbf{x}_{r_{3,1}}(1:t), \mathbf{x}_{r_{3,2}}(1:t), \mathbf{y}_{r_4}, \mathbf{z}_{r_4}(1:t))) = \mathrm{ELBO}_{\mathrm{TS-PVAE}}(r_3(1:t))$$

$$+ \sum_{i=1}^{3} D_{KL}(q(\mathbf{z}_{r_{3,i}}(1:t)) \| p(\mathbf{z}_{r_{3,i}}(1:t) | \mathbf{x}_{r_{3,1}}(1:t), \mathbf{x}_{r_{3,2}}(1:t), \mathbf{y}_{r_4}, \mathbf{z}_{r_4}(1:t))) \geq \mathrm{ELBO}_{\mathrm{TS-PVAE}}(r_3(1:t))$$

$$= \int_{\mathbf{z}_{r_{3,1}}(1:t)} \int_{\mathbf{z}_{r_{3,2}}(1:t)} q(\mathbf{z}_{r_{3,1}}(1:t)) q(\mathbf{z}_{r_{3,2}}(1:t)) \times \ln \frac{p(\mathbf{x}_{r_{3,1}}(1:t), \mathbf{x}_{r_{3,2}}(1:t), \mathbf{y}_{r_4}, \mathbf{z}_{r_{3,1}}(1:t), \mathbf{z}_{r_{3,2}}(1:t), \mathbf{z}_{r_4}(1:t))}{q(\mathbf{z}_{r_{3,1}}(1:t)) q(\mathbf{z}_{r_{3,2}}(1:t))} d\mathbf{z}_{r_{3,1}}(1:t) d\mathbf{z}_{r_{3,2}}(1:t)$$

$$= \int_{\mathbf{z}_{r_{3,1}}(1:t)} \int_{\mathbf{z}_{r_{3,2}}(1:t)} q(\mathbf{z}_{r_{3,1}}(1:t)) q(\mathbf{z}_{r_{3,2}}(1:t)) \times \left[\ln \frac{p(\mathbf{x}_{r_{3,1}}(1:t) | \mathbf{z}_{r_{3,1}}(1:t)) p(\mathbf{z}_{r_{3,1}}(1:t)) p(\mathbf{x}_{r_{3,2}}(1:t) | \mathbf{z}_{r_{3,2}}(1:t)) p(\mathbf{z}_{r_{3,2}}(1:t))}{q(\mathbf{z}_{r_{3,1}}(1:t)) q(\mathbf{z}_{r_{3,2}}(1:t))} \right.$$

$$\left. + \ln p(\mathbf{y}_{r_4} | \mathbf{z}_{r_{3,1}}(1:t), \mathbf{z}_{r_{3,2}}(1:t), \mathbf{z}_{r_4}(1:t)) p(\mathbf{z}_{r_4}(1:t)) \right] d\mathbf{z}_{r_{3,1}}(1:t) d\mathbf{z}_{r_{3,2}}(1:t)$$

$$= \sum_{i=1}^{2} \int_{\mathbf{z}_{r_{3,i}}(1:t)} q(\mathbf{z}_{r_{3,i}}(1:t)) \ln \frac{p(\mathbf{x}_{r_{3,i}}(1:t) | \mathbf{z}_{r_{3,i}}(1:t)) p(\mathbf{z}_{r_{3,i}}(1:t))}{q(\mathbf{z}_{r_{3,i}}(1:t))} d\mathbf{z}_{r_{3,i}}(1:t) + \int_{\mathbf{z}_{r_{3,1}}(1:t)} \int_{\mathbf{z}_{r_{3,2}}(1:t)} q(\mathbf{z}_{r_{3,1}}(1:t))$$

$$q(\mathbf{z}_{r_{3,2}}(1:t)) \times \ln p(\mathbf{y}_{r_4} | \mathbf{z}_{r_{3,1}}(1:t), \mathbf{z}_{r_{3,2}}(1:t), \mathbf{z}_{r_4}(1:t)) d\mathbf{z}_{r_{3,1}}(1:t) d\mathbf{z}_{r_{3,2}}(1:t) + C \geq \sum_{i=1}^{2} \mathbb{E}_{\mathbf{z}_{r_{3,i}}(1:t) \sim q(\mathbf{z}_{r_{3,i}}(1:t) | \mathbf{x}_{r_{3,i}}(1:t))} [-\ln p(\mathbf{x}_{r_{3,i}}(1:t) | \mathbf{z}_{r_{3,i}}(1:t))]$$

$$- \sum_{i=1}^{2} D_{KL}(q(\mathbf{z}_{r_{3,i}}(1:t) | \mathbf{x}_{r_{3,i}}(1:t)) \| p(\mathbf{z}_{r_{3,i}}(1:t))) + \mathbb{E}_{\mathbf{z}_{r3}(1:t) \sim q(\mathbf{z}_{r3}(1:t) | \mathbf{z}_{r_{3,1}}(1:t), \mathbf{z}_{r_{3,2}}(1:t), \mathbf{z}_{r_4}(1:t))} [\ln p(\mathbf{y}_{r_4} | \mathbf{z}_{r_{3,1}}(1:t), \mathbf{z}_{r_{3,2}}(1:t), \mathbf{z}_{r_4}(1:t))] \quad (16)$$

3. The detailed derivation of the marginal likelihood for TS-PVAE-r2 can be given as

https://doi.org/10.1021/acsomega.4c02681
ACS Omega 2024, 9, 23021−23032

$$\ln p(\mathbf{x}_{r_2}(1{:}\,t), \mathbf{y}_{r_4}, \mathbf{z}_{r_3}(1{:}\,t)) = \ln p(\mathbf{x}_{r_{2,1}}(1{:}\,t), \mathbf{x}_{r_{2,2}}(1{:}\,t), \mathbf{x}_{r_{2,3}}(1{:}\,t), \mathbf{y}_{r_4}, \mathbf{z}_{r_3}(1{:}\,t)) = \int_{\mathbf{z}_{r_{2,1}}(1:t)} \int_{\mathbf{z}_{r_{2,2}}(1:t)} \int_{\mathbf{z}_{r_{2,3}}(1:t)} q(\mathbf{z}_{r_{2,1}}(1{:}\,t)) q(\mathbf{z}_{r_{2,2}}(1{:}\,t)) q(\mathbf{z}_{r_{2,3}}(1{:}\,t))$$
$$\times \left[\ln \frac{p(\mathbf{x}_{r_{2,1}}(1{:}\,t), \mathbf{x}_{r_{2,2}}(1{:}\,t), \mathbf{x}_{r_{2,3}}(1{:}\,t), \mathbf{y}_{r_4}, \mathbf{z}_{r_{2,1}}(1{:}\,t), \mathbf{z}_{r_{2,2}}(1{:}\,t), \mathbf{z}_{r_{2,3}}(1{:}\,t), \mathbf{z}_{r_3}(1{:}\,t))}{q(\mathbf{z}_{r_{2,1}}(1{:}\,t)) q(\mathbf{z}_{r_{2,2}}(1{:}\,t)) q(\mathbf{z}_{r_{2,3}}(1{:}\,t))} \right.$$
$$\left. - \sum_{i=1}^{3} \ln \frac{p(\mathbf{z}_{r_{2,i}}(1{:}\,t) | \mathbf{x}_{r_{2,1}}(1{:}\,t), \mathbf{x}_{r_{2,2}}(1{:}\,t), \mathbf{x}_{r_{2,3}}(1{:}\,t), \mathbf{y}_{r_4}, \mathbf{z}_{r_3}(1{:}\,t))}{q(\mathbf{z}_{r_{2,i}}(1{:}\,t))} \right] \mathrm{d}\mathbf{z}_{r_{2,1}}(1{:}\,t) \mathrm{d}\mathbf{z}_{r_{2,2}}(1{:}\,t) \mathrm{d}\mathbf{z}_{r_{2,3}}(1{:}\,t) = \int_{\mathbf{z}_{r_{2,1}}(1:t)} \int_{\mathbf{z}_{r_{2,2}}(1:t)} \int_{\mathbf{z}_{r_{2,3}}(1:t)} q(\mathbf{z}_{r_{2,1}}(1{:}\,t))$$
$$q(\mathbf{z}_{r_{2,2}}(1{:}\,t)) q(\mathbf{z}_{r_{2,3}}(1{:}\,t)) \times \left[\ln \frac{p(\mathbf{x}_{r_{2,1}}(1{:}\,t), \mathbf{x}_{r_{2,2}}(1{:}\,t), \mathbf{x}_{r_{2,3}}(1{:}\,t), \mathbf{y}_{r_4}, \mathbf{z}_{r_{2,1}}(1{:}\,t), \mathbf{z}_{r_{2,2}}(1{:}\,t), \mathbf{z}_{r_{2,3}}(1{:}\,t), \mathbf{z}_{r_3}(1{:}\,t))}{q(\mathbf{z}_{r_{2,1}}(1{:}\,t)) q(\mathbf{z}_{r_{2,2}}(1{:}\,t)) q(\mathbf{z}_{r_{2,3}}(1{:}\,t))} \right] \mathrm{d}\mathbf{z}_{r_{2,1}}(1{:}\,t) \mathrm{d}\mathbf{z}_{r_{2,2}}(1{:}\,t) \mathrm{d}\mathbf{z}_{r_{2,3}}(1{:}\,t)$$
$$- \sum_{i=1}^{3} \int_{\mathbf{z}_{r_{2,i}}(1:t)} q(\mathbf{z}_{r_{2,i}}(1{:}\,t)) \ln \frac{p(\mathbf{z}_{r_{2,i}}(1{:}\,t) | \mathbf{x}_{r_{2,1}}(1{:}\,t), \mathbf{x}_{r_{2,2}}(1{:}\,t), \mathbf{x}_{r_{2,3}}(1{:}\,t), \mathbf{y}_{r_4}, \mathbf{z}_{r_3}(1{:}\,t))}{q(\mathbf{z}_{r_{2,i}}(1{:}\,t))} \mathrm{d}\mathbf{z}_{r_{2,i}}(1{:}\,t) = \int_{\mathbf{z}_{r_{2,1}}(1:t)} \int_{\mathbf{z}_{r_{2,2}}(1:t)} \int_{\mathbf{z}_{r_{2,3}}(1:t)} q(\mathbf{z}_{r_{2,1}}(1{:}\,t))$$
$$q(\mathbf{z}_{r_{2,2}}(1{:}\,t)) q(\mathbf{z}_{r_{2,3}}(1{:}\,t)) \times \left[\ln \frac{p(\mathbf{x}_{r_{2,1}}(1{:}\,t), \mathbf{x}_{r_{2,2}}(1{:}\,t), \mathbf{x}_{r_{2,3}}(1{:}\,t), \mathbf{y}_{r_4}, \mathbf{z}_{r_{2,1}}(1{:}\,t), \mathbf{z}_{r_{2,2}}(1{:}\,t), \mathbf{z}_{r_{2,3}}(1{:}\,t), \mathbf{z}_{r_3}(1{:}\,t))}{q(\mathbf{z}_{r_{2,1}}(1{:}\,t)) q(\mathbf{z}_{r_{2,2}}(1{:}\,t)) q(\mathbf{z}_{r_{2,3}}(1{:}\,t))} \right] \mathrm{d}\mathbf{z}_{r_{2,1}}(1{:}\,t) \mathrm{d}\mathbf{z}_{r_{2,2}}(1{:}\,t) \mathrm{d}\mathbf{z}_{r_{2,3}}(1{:}\,t)$$
$$+ \sum_{i=1}^{3} \mathrm{D}_{\mathrm{KL}}(q(\mathbf{z}_{r_{2,i}}(1{:}\,t)) \| p(\mathbf{z}_{r_{2,i}}(1{:}\,t) | \mathbf{x}_{r_{2,1}}(1{:}\,t), \mathbf{x}_{r_{2,2}}(1{:}\,t), \mathbf{x}_{r_{2,3}}(1{:}\,t), \mathbf{y}_{r_4}, \mathbf{z}_{r_3}(1{:}\,t))) = \mathrm{ELBO}_{\mathrm{TS-PVAE}}(r_2(1{:}\,t)) +$$
$$\sum_{i=1}^{3} \mathrm{D}_{\mathrm{KL}}(q(\mathbf{z}_{r_{2,i}}(1{:}\,t)) \| p(\mathbf{z}_{r_{2,i}}(1{:}\,t) | \mathbf{x}_{r_{2,1}}(1{:}\,t), \mathbf{x}_{r_{2,2}}(1{:}\,t), \mathbf{x}_{r_{2,3}}(1{:}\,t), \mathbf{y}_{r_4}, \mathbf{z}_{r_3}(1{:}\,t))) \geq \mathrm{ELBO}_{\mathrm{TS-PVAE}}(r_2(1{:}\,t)) =$$
$$\int_{\mathbf{z}_{r_{2,1}}(1:t)} \int_{\mathbf{z}_{r_{2,2}}(1:t)} \int_{\mathbf{z}_{r_{2,3}}(1:t)} q(\mathbf{z}_{r_{2,1}}(1{:}\,t)) q(\mathbf{z}_{r_{2,2}}(1{:}\,t)) q(\mathbf{z}_{r_{2,3}}(1{:}\,t)) \times \ln \frac{p(\mathbf{x}_{r_{2,1}}(1{:}\,t), \mathbf{x}_{r_{2,2}}(1{:}\,t), \mathbf{y}_{r_4}, \mathbf{z}_{r_{2,1}}(1{:}\,t), \mathbf{z}_{r_{2,2}}(1{:}\,t), \mathbf{z}_{r_{2,3}}(1{:}\,t), \mathbf{z}_{r_3}(1{:}\,t))}{q(\mathbf{z}_{r_{2,1}}(1{:}\,t)) q(\mathbf{z}_{r_{2,2}}(1{:}\,t)) q(\mathbf{z}_{r_{2,3}}(1{:}\,t))}$$
$$\mathrm{d}\mathbf{z}_{r_{2,1}}(1{:}\,t) \mathrm{d}\mathbf{z}_{r_{2,2}}(1{:}\,t) \mathrm{d}\mathbf{z}_{r_{2,3}}(1{:}\,t) = \int_{\mathbf{z}_{r_{2,1}}(1:t)} \int_{\mathbf{z}_{r_{2,2}}(1:t)} \int_{\mathbf{z}_{r_{2,3}}(1:t)} q(\mathbf{z}_{r_{2,1}}(1{:}\,t)) q(\mathbf{z}_{r_{2,2}}(1{:}\,t)) q(\mathbf{z}_{r_{2,3}}(1{:}\,t)) \times [\sum_{i=1}^{3} \ln \frac{p(\mathbf{x}_{r_{2,i}}(1{:}\,t) | \mathbf{z}_{r_{2,i}}(1{:}\,t)) p(\mathbf{z}_{r_{2,i}}(1{:}\,t))}{q(\mathbf{z}_{r_{2,i}}(1{:}\,t))}$$
$$+ \ln p(\mathbf{y}_{r_4} | \mathbf{z}_{r_{2,1}}(1{:}\,t), \mathbf{z}_{r_{2,2}}(1{:}\,t), \mathbf{z}_{r_{2,3}}(1{:}\,t), \mathbf{z}_{r_3}(1{:}\,t)) p(\mathbf{z}_{r_4}(1{:}\,t))] \mathrm{d}\mathbf{z}_{r_{2,1}}(1{:}\,t) \mathrm{d}\mathbf{z}_{r_{2,2}}(1{:}\,t) \mathrm{d}\mathbf{z}_{r_{2,3}}(1{:}\,t) = \sum_{i=1}^{3} \int_{\mathbf{z}_{r_{2,i}}(1:t)} q(\mathbf{z}_{r_{2,i}}(1{:}\,t))$$
$$\ln \frac{p(\mathbf{x}_{r_{2,i}}(1{:}\,t) | \mathbf{z}_{r_{2,i}}(1{:}\,t)) p(\mathbf{z}_{r_{2,i}}(1{:}\,t))}{q(\mathbf{z}_{r_{2,i}}(1{:}\,t))} \mathrm{d}\mathbf{z}_{r_{2,i}}(1{:}\,t) + \int_{\mathbf{z}_{r_{2,1}}(1:t)} \int_{\mathbf{z}_{r_{2,2}}(1:t)} \int_{\mathbf{z}_{r_{2,3}}(1:t)} q(\mathbf{z}_{r_{2,1}}(1{:}\,t)) q(\mathbf{z}_{r_{2,2}}(1{:}\,t)) q(\mathbf{z}_{r_{2,3}}(1{:}\,t))$$
$$\times \ln p(\mathbf{y}_{r_4} | \mathbf{z}_{r_{2,1}}(1{:}\,t), \mathbf{z}_{r_{2,2}}(1{:}\,t), \mathbf{z}_{r_{2,3}}(1{:}\,t), \mathbf{z}_{r_3}(1{:}\,t)) \mathrm{d}\mathbf{z}_{r_{2,1}}(1{:}\,t) \mathrm{d}\mathbf{z}_{r_{2,2}}(1{:}\,t) \mathrm{d}\mathbf{z}_{r_{2,3}}(1{:}\,t) + C \geq \sum_{i=1}^{3} \mathbb{E}_{\mathbf{z}_{r_{2,i}}(1:t) \sim q(\mathbf{z}_{r_{2,i}}(1:t) | \mathbf{x}_{r_{2,i}}(1:t))}$$
$$[-\ln p(\mathbf{x}_{r_{2,i}}(1{:}\,t) | \mathbf{z}_{r_{2,i}}(1{:}\,t))] - \sum_{i=1}^{3} D_{KL}(q(\mathbf{z}_{r_{2,i}}(1{:}\,t) | \mathbf{x}_{r_{2,i}}(1{:}\,t)) \| p(\mathbf{z}_{r_{2,i}}(1{:}\,t))) + \mathbb{E}_{\mathbf{z}_{r_2} \sim q(\mathbf{z}_{r_2} | \mathbf{z}_{r_{2,1}}(1:t), \mathbf{z}_{r_{2,2}}(1:t), \mathbf{z}_{r_{2,3}}(1:t), \mathbf{z}_{r_3}(1:t))}$$
$$[\ln p(\mathbf{y}_{r_4} | \mathbf{z}_{r_{2,1}}(1{:}\,t), \mathbf{z}_{r_{2,2}}(1{:}\,t), \mathbf{z}_{r_{2,3}}(1{:}\,t), \mathbf{z}_{r_3}(1{:}\,t))] \tag{17}$$

4. The detailed derivation of the marginal likelihood for TS-PVAE-r1 can be given as

https://doi.org/10.1021/acsomega.4c02681
ACS Omega 2024, 9, 23021−23032

$$\ln p(\mathbf{x}_{r_1}(1:t), \mathbf{y}_{r_4}, \mathbf{z}_{r_2}(1:t)) = \ln p(\mathbf{x}_{r_{1,1}}(1:t), ..., \mathbf{x}_{r_{1,6}}(1:t), \mathbf{y}_{r_4}, \mathbf{z}_{r_2}(1:t)) = \int_{\mathbf{z}_{r_{1,1}}(1:t)} ... \int_{\mathbf{z}_{r_{1,6}}(1:t)} q(\mathbf{z}_{r_{1,1}}(1:t)) ... q(\mathbf{z}_{r_{1,6}}(1:t))$$

$$\times \left[\ln \frac{p(\mathbf{x}_{r_{1,1}}(1:t), ..., \mathbf{x}_{r_{1,6}}(1:t), \mathbf{y}_{r_4}, \mathbf{z}_{r_{1,1}}(1:t), ..., \mathbf{z}_{r_{1,6}}(1:t), \mathbf{z}_{r_2}(1:t))}{q(\mathbf{z}_{r_{1,1}}(1:t)) ... q(\mathbf{z}_{r_{1,6}}(1:t))} - \sum_{i=1}^{6} \ln \frac{p(\mathbf{z}_{r_{1,i}}(1:t)|\mathbf{x}_{r_{1,1}}(1:t), ..., \mathbf{x}_{r_{1,6}}(1:t), \mathbf{y}_{r_4}, \mathbf{z}_{r_2}(1:t))}{q(\mathbf{z}_{r_{1,i}}(1:t))} \right] d\mathbf{z}_{r_{1,1}}(1:t)$$

$$... d\mathbf{z}_{r_{1,6}}(1:t) = \int_{\mathbf{z}_{r_{1,1}}(1:t)} ... \int_{\mathbf{z}_{r_{1,6}}(1:t)} q(\mathbf{z}_{r_{1,1}}(1:t)) ... q(\mathbf{z}_{r_{1,6}}(1:t)) \times \left[\ln \frac{p(\mathbf{x}_{r_{1,1}}(1:t), ..., \mathbf{x}_{r_{1,6}}(1:t), \mathbf{y}_{r_4}, \mathbf{z}_{r_{1,1}}(1:t), ..., \mathbf{z}_{r_{1,6}}(1:t), \mathbf{z}_{r_2})}{q(\mathbf{z}_{r_{1,1}}(1:t)) ... q(\mathbf{z}_{r_{1,6}}(1:t))} \right] d\mathbf{z}_{r_{1,1}}(1:t)$$

$$... d\mathbf{z}_{r_{1,6}}(1:t) - \sum_{i=1}^{6} \int_{\mathbf{z}_{r_{1,i}}(1:t)} q(\mathbf{z}_{r_{1,i}}(1:t)) \times \ln \frac{p(\mathbf{z}_{r_{1,i}}(1:t)|\mathbf{x}_{r_{1,1}}(1:t), ..., \mathbf{x}_{r_{1,6}}(1:t), \mathbf{y}_{r_4}, \mathbf{z}_{r_2}(1:t))}{q(\mathbf{z}_{r_{1,i}}(1:t))} d\mathbf{z}_{r_{1,i}}(1:t) = \int_{\mathbf{z}_{r_{1,1}}(1:t)} ... \int_{\mathbf{z}_{r_{1,6}}(1:t)} q(\mathbf{z}_{r_{1,1}}(1:t))$$

$$... q(\mathbf{z}_{r_{1,6}}(1:t)) \times \left[\ln \frac{p(\mathbf{x}_{r_{1,1}}(1:t), ..., \mathbf{x}_{r_{1,6}}(1:t), \mathbf{y}_{r_4}, \mathbf{z}_{r_{1,1}}(1:t), ..., \mathbf{z}_{r_{1,6}}, \mathbf{z}_{r_2})}{q(\mathbf{z}_{r_{1,1}}(1:t)) \cdots q(\mathbf{z}_{r_{1,6}}(1:t))} \right] d\mathbf{z}_{r_{1,1}}(1:t)$$

$$... d\mathbf{z}_{r_{1,6}}(1:t) + \sum_{i=1}^{6} D_{KL}(q(\mathbf{z}_{r_{1,i}}(1:t)) || p(\mathbf{z}_{r_{1,i}}(1:t)|\mathbf{x}_{r_{1,1}}(1:t), ..., \mathbf{x}_{r_{1,6}}(1:t), \mathbf{y}_{r_4}, \mathbf{z}_{r_2}(1:t))) = \mathrm{ELBO}_{\mathrm{TS-PVAE}}(r_1(1:t))$$

$$+ \sum_{i=1}^{6} D_{KL}(q(\mathbf{z}_{r_{1,i}}(1:t)) || p(\mathbf{z}_{r_{1,i}}(1:t)|\mathbf{x}_{r_{1,1}}(1:t), ..., \mathbf{x}_{r_{1,6}}(1:t), \mathbf{y}_{r_4}, \mathbf{z}_{r_2}(1:t))) \geq \mathrm{ELBO}_{\mathrm{TS-PVAE}}(r_1(1:t)) = \int_{\mathbf{z}_{r_{1,1}}(1:t)} ... \int_{\mathbf{z}_{r_{1,6}}(1:t)} q(\mathbf{z}_{r_{1,1}}(1:t))$$

$$... q(\mathbf{z}_{r_{1,6}}(1:t)) \times \ln \frac{p(\mathbf{x}_{r_{1,1}}(1:t), ..., \mathbf{y}_{r_4}, \mathbf{z}_{r_{1,1}}(1:t), ..., \mathbf{z}_{r_{1,6}}(1:t), \mathbf{z}_{r_2}(1:t))}{q(\mathbf{z}_{r_{1,1}}(1:t)) \cdots q(\mathbf{z}_{r_{1,6}}(1:t))} d\mathbf{z}_{r_{1,1}}(1:t) ... d\mathbf{z}_{r_{1,6}}(1:t) = \int_{\mathbf{z}_{r_{1,1}}(1:t)} ... \int_{\mathbf{z}_{r_{1,6}}(1:t)} q(\mathbf{z}_{r_{1,1}}(1:t)) ... q(\mathbf{z}_{r_{1,6}}(1:t))$$

$$[\sum_{i=1}^{6} \ln \frac{p(\mathbf{x}_{r_{1,i}}(1:t)|\mathbf{z}_{r_{1,i}}(1:t)) p(\mathbf{z}_{r_{1,i}}(1:t))}{q(\mathbf{z}_{r_{1,i}}(1:t))} + \ln p(\mathbf{y}_{r_4}|\mathbf{z}_{r_{1,1}}(1:t), ..., \mathbf{z}_{r_{1,6}}(1:t), \mathbf{z}_{r_2}(1:t)) p(\mathbf{z}_{r_4}(1:t)(1:t))] d\mathbf{z}_{r_{1,1}}(1:t)$$

$$... d\mathbf{z}_{r_{1,6}}(1:t) = \sum_{i=1}^{6} \int_{\mathbf{z}_{r_{1,i}}(1:t)} q(\mathbf{z}_{r_{1,i}}(1:t)) \ln \frac{p(\mathbf{x}_{r_{1,i}}(1:t)|\mathbf{z}_{r_{1,i}}(1:t)) p(\mathbf{z}_{r_{1,i}}(1:t))}{q(\mathbf{z}_{r_{1,i}}(1:t))} d\mathbf{z}_{r_{1,i}}(1:t) + \int_{\mathbf{z}_{r_{1,1}}(1:t)} ... \int_{\mathbf{z}_{r_{1,6}}(1:t)} q(\mathbf{z}_{r_{1,1}}(1:t)) ... q(\mathbf{z}_{r_{1,6}}(1:t))$$

$$\times \ln p(\mathbf{y}_{r_4}|\mathbf{z}_{r_{1,1}}(1:t), ..., \mathbf{z}_{r_{1,6}}(1:t), \mathbf{z}_{r_2}(1:t)) d\mathbf{z}_{r_{1,1}}(1:t) ... d\mathbf{z}_{r_{1,6}}(1:t) + C \geq \sum_{i=1}^{6} \mathbb{E}_{\mathbf{z}_{r_{1,i}}(1:t) \sim q(\mathbf{z}_{r_{1,i}}(1:t)|\mathbf{x}_{r_{1,i}}(1:t))}[-\ln p(\mathbf{x}_{r_{1,i}}(1:t)|\mathbf{z}_{r_{1,i}}(1:t))]$$

$$- \sum_{i=1}^{6} D_{KL}(q(\mathbf{z}_{r_{1,i}}(1:t)|\mathbf{x}_{r_{1,i}}(1:t)) || p(\mathbf{z}_{r_{1,i}}(1:t))) + \mathbb{E}_{\mathbf{z}_{r_1}(1:t) \sim q(\mathbf{z}_{r_1}(1:t)|\mathbf{z}_{r_{1,1}}(1:t), ..., \mathbf{z}_{r_{1,6}}(1:t), \mathbf{z}_{r_2}(1:t))}[\ln p(\mathbf{y}_{r_4}|\mathbf{z}_{r_{1,1}}(1:t), ..., \mathbf{z}_{r_{1,6}}(1:t), \mathbf{z}_{r_2}(1:t))] \quad (18)$$

AUTHOR INFORMATION

Corresponding Authors

Zeyu Yang − *Huzhou Key Laboratory of Intelligent Sensing and Optimal Control for Industrial Systems School of Engineering, Huzhou University, Huzhou 313000, China*; orcid.org/0000-0003-0253-9053; Email: yangzeyu@zjhu.edu.cn

Le Yao − *School of Mathematics, Hangzhou Normal University, Hangzhou 311121, China*; orcid.org/0000-0002-0881-213X; Email: yaole@hznu.edu.cn

Author

Bingbing Shen − *School of Mathematics, Hangzhou Normal University, Hangzhou 311121, China*

Complete contact information is available at: https://pubs.acs.org/10.1021/acsomega.4c02681

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This work was supported in part by the National Natural Science Foundation of China (NSFC) (62303146, 62203169, 62003300), the Natural Science Foundation of Zhejiang Province (LQ23F030004, LQ22F030009), and Huzhou Key Laboratory of Intelligent Sensing and Optimal Control for Industrial Systems (grant 2022-17).

REFERENCES

(1) Ding, S. X. Data-driven design of monitoring and diagnosis systems for dynamic processes: A review of subspace technique based schemes and some recent results. *Journal of Process Control* **2014**, *24*, 431−449.

(2) Jiang, Y.; Yin, S.; Dong, J.; Kaynak, O. A review on soft sensors for monitoring, control, and optimization of industrial processes. *IEEE Sensors Journal* **2021**, *21*, 12868−12881.

(3) Yang, Z.; Ge, Z. On paradigm of industrial big data analytics: From evolution to revolution. *IEEE Transactions on Industrial Informatics* **2022**, *18*, 8373−8388.

(4) Yao, L.; Ge, Z. Scalable learning and probabilistic analytics of industrial big data based on parameter server: Framework, methods and applications. *Journal of Process Control* **2019**, *78*, 13−33.

(5) Yao, L.; Ge, Z. Variable selection for nonlinear soft sensor development with enhanced binary differential evolution algorithm. *Control Engineering Practice* **2018**, *72*, 68−82.

(6) Yuan, X.; Xu, N.; Ye, L.; Wang, K.; Shen, F.; Wang, Y.; Yang, C.; Gui, W. Attention-Based Interval Aided Networks for Data Modeling of Heterogeneous Sampling Sequences With Missing Values in Process Industry. *IEEE Transactions on Industrial Informatics* **2024**, *20*, 5253.

(7) Lu, B.; Chiang, L. Semi-supervised online soft sensor maintenance experiences in the chemical industry. *Journal of Process Control* **2018**, *67*, 23−34.

(8) Shao, W.; Li, X.; Yao, Y.; Chen, J.; Zhao, D. Semi-supervised local manifold regularization model based on dual representation for industrial soft sensor development. *Chemometrics and Intelligent Laboratory Systems* **2023**, *242*, No. 104937.

(9) Yao, L.; Shen, B.; Cui, L.; Zheng, J.; Ge, Z. Semi-supervised deep dynamic probabilistic latent variable model for multimode process soft sensor application. *IEEE Transactions on Industrial Informatics* **2023**, *19*, 6056−6068.
(10) Yao, L.; Ge, Z. Deep learning of semisupervised process data with hierarchical extreme learning machine and soft sensor application. *IEEE Transactions on Industrial Electronics* **2018**, *65*, 1490−1498.
(11) Lyu, Y.; Zhou, L.; Cong, Y.; Zheng, H.; Song, Z. Multirate Mixture Probability Principal Component Analysis for Process Monitoring in Multimode Processes. *IEEE Transactions on Automation Science and Engineering* **2024**, *21*, 2027.
(12) Song, B.; Zhou, Y.; Shi, H.; Tao, Y.; Tan, S. A Soft Sensor for Multirate Quality Variables Based on MC-CNN. *IEEE Transactions on Neural Networks and Learning Systems* **2024**, 1.
(13) Huang, K.; Wu, S.; Li, Y.; Yang, C.; Gui, W. A multi-rate sampling data fusion method for fault diagnosis and its industrial applications. *Journal of Process Control* **2021**, *104*, 54−61.
(14) Zhou, L.; Wang, Y.; Ge, Z. Multi-rate principal component regression model for soft sensor application in industrial processes. *Science China. Information Sciences* **2020**, *63*, No. 149205.
(15) Shen, B.; Yao, L.; Ge, Z. Predictive modeling with multiresolution pyramid VAE and industrial soft sensor applications. *IEEE Transactions on Cybernetics* **2023**, *53*, 4867−4879.
(16) LeCun, Y.; Bengio, Y.; Hinton, G. Deep learning. *nature* **2015**, *521*, 436−444.
(17) Shang, C.; Yang, F.; Huang, D.; Lyu, W. Data-driven soft sensor development based on deep learning technique. *Journal of Process Control* **2014**, *24*, 223−233.
(18) Wang, Y.; Yan, P. RegGAN: A Virtual Sample Generative Network for Developing Soft Sensors with Small Data. *ACS omega* **2024**, *9*, 5954.
(19) Jiang, X.; Ge, Z. Augmented multidimensional convolutional neural network for industrial soft sensing. *IEEE Transactions on Instrumentation and Measurement* **2021**, *70*, 1−10.
(20) Yuan, X.; Wang, Y.; Wang, C.; Ye, L.; Wang, K.; Wang, Y.; Yang, C.; Gui, W.; Shen, F. Variable Correlation Analysis-based Convolutional Neural Network for Far Topological Feature Extraction and Industrial Predictive Modeling. *IEEE Transactions on Instrumentation and Measurement* **2024**, *73*, 1.
(21) Yuan, X.; Huang, L.; Ye, L.; Wang, Y.; Wang, K.; Yang, C.; Gui, W.; Shen, F. Quality Prediction Modeling for Industrial Processes Using Multiscale Attention-Based Convolutional Neural Network. *IEEE Transactions on Cybernetics* **2024**, *54*, 2696.
(22) Kingma, D. P.; Welling, M. Auto-encoding variational bayes. *arXiv* **2013**. DOI: 10.48550/arXiv.1312.6114
(23) Doersch, C. Tutorial on variational autoencoders. *arXiv* **2016**. DOI: 10.48550/arXiv.1606.05908
(24) Shen, B.; Yao, L.; Ge, Z. Nonlinear probabilistic latent variable regression models for soft sensor application: From shallow to deep structure. *Control Engineering Practice* **2020**, *94*, No. 104198.
(25) Yuan, X.; Huang, B.; Wang, Y.; Yang, C.; Gui, W. Deep learning-based feature representation and its application for soft sensor modeling with variable-wise weighted SAE. *IEEE Transactions on Industrial Informatics* **2018**, *14*, 3235−3243.
(26) Dai, Y.; Yang, C.; Zhu, J.; Liu, Y. Adversarial transferred data-assisted soft sensor for enhanced multigrade quality prediction. *ACS omega* **2023**, *8*, 19900−19911.
(27) Jiang, B.; Yang, H.; Wang, Y.; Liu, Y.; Geng, H.; Zeng, H.; Ding, J. Dynamic Temporal Dependency Model for Multiple Steps Ahead Short-term Load Forecasting of Power System. *IEEE Trans. Ind. Appl.* **2024**, 1.
(28) Dai, Y.; Yang, C.; Liu, K.; Liu, A.; Liu, Y. TimeDDPM: Time series augmentation strategy for industrial soft sensing. *IEEE Sensors Journal* **2024**, *24*, 2145.
(29) Yuan, X.; Li, L.; Shardt, Y. A.; Wang, Y.; Yang, C. Deep learning with spatiotemporal attention-based LSTM for industrial soft sensor model development. *IEEE Transactions on Industrial Electronics* **2021**, *68*, 4404−4414.
(30) Yang, Z.; Jia, R.; Wang, P.; Yao, L.; Shen, B. Supervised attention-based bidirectional long short-term memory network for nonlinear dynamic soft sensor application. *ACS omega* **2023**, *8*, 4196−4208.
(31) Shao, W.; Xiao, C.; Wang, J.; Zhao, D.; Song, Z. Real-time estimation of quality-related variable for dynamic and non-Gaussian process based on semisupervised Bayesian HMM. *Journal of Process Control* **2022**, *111*, 59−74.
(32) Zhu, J.; Ge, Z.; Song, Z. HMM-driven robust probabilistic principal component analyzer for dynamic process fault classification. *IEEE Transactions on Industrial Electronics* **2015**, *62*, 3814−3821.
(33) Liu, Q.; Jia, M.; Gao, Z.; Xu, L.; Liu, Y. Correntropy long short term memory soft sensor for quality prediction in industrial polyethylene process. *Chemometrics and Intelligent Laboratory Systems* **2022**, *231*, No. 104678.
(34) Zheng, J.; Ma, L.; Wu, Y.; Ye, L.; Shen, F. Nonlinear dynamic soft sensor development with a supervised hybrid CNN-LSTM network for industrial processes. *ACS omega* **2022**, *7*, 16653−16664.
(35) Geng, Z.; Chen, Z.; Meng, Q.; Han, Y. Novel transformer based on gated convolutional neural network for dynamic soft sensor modeling of industrial processes. *IEEE Transactions on Industrial Informatics* **2022**, *18*, 1521−1529.
(36) Fujiwara, K.; Kano, M.; Hasebe, S.; Takinami, A. Soft-sensor development using correlation-based just-in-time modeling. *AIChE J.* **2009**, *55*, 1754−1765.
(37) Zheng, W.; Liu, Y.; Gao, Z.; Yang, J. Just-in-time semi-supervised soft sensor for quality prediction in industrial rubber mixers. *Chemometrics and Intelligent Laboratory Systems* **2018**, *180*, 36−41.
(38) Xie, L.; Zeng, J.; Gao, C. Novel just-in-time learning-based soft sensor utilizing non-Gaussian information. *IEEE Transactions on Control Systems Technology* **2014**, *22*, 360−368.
(39) Guo, F.; Xie, R.; Huang, B. A deep learning just-in-time modeling approach for soft sensor based on variational autoencoder. *Chemometrics and Intelligent Laboratory Systems* **2020**, *197*, No. 103922.

https://doi.org/10.1021/acsomega.4c02681
ACS Omega 2024, 9, 23021−23032