IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received December 10, 2019, accepted January 23, 2020, date of publication January 31, 2020, date of current version February 10, 2020.

*Digital Object Identifier 10.1109/ACCESS.2020.2970836*

# Spatiotemporal Modeling for Nonlinear Distributed Thermal Processes Based on KL Decomposition, MLP and LSTM Network

**YAJUN FAN[1], KANGKANG XU[2], HUI WU[3], YING ZHENG[3], (Senior Member, IEEE), AND BO TAO[1], (Member, IEEE)**

[1]State Key Laboratory of Digital Manufacturing Equipment and Technology, Huazhong University of Science and Technology, Wuhan 430074, China
[2]Key Laboratory of Mechanical Equipment Manufacturing and Control Technology of Ministry of Education, Guangdong University of Technology, Guangzhou 510006, China
[3]National Key Laboratory of Science and Technology on Multispectral Information Processing, Huazhong University of Science and Technology, Wuhan 430074, China

Corresponding author: Bo Tao (taobo@hust.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant U1501248, Grant 51575215, and Grant 51905109, and in part by the National Key Basic Research Program of China under Grant 2017YFB1301504.

**ABSTRACT** Estimation of absolute temperature distributions is crucial for many thermal processes in the nonlinear distributed parameter systems, such as predicting the curing temperature distribution of the chip, the temperature distribution of the catalytic rod, and so on. In this work, a spatiotemporal model based on the Karhunen-Loève (KL) decomposition, the multilayer perceptron (MLP), and the long short-term memory (LSTM) network, named KL-MLP-LSTM, is developed for estimating temperature distributions with a three-step procedure. Firstly, the infinite-dimensional model is transformed into a finite-dimensional model, where the KL decomposition method is used for dimension reduction and spatial basis functions extraction. Secondly, a novel MLP-LSTM hybrid time series model is constructed to deal with the two inherently coupled nonlinearities. Finally, the spatiotemporal temperature distribution model can be reconstructed through spatiotemporal synthesis. The effectiveness of the proposed model is validated by the data from a snap curing oven thermal process. Satisfactory agreement between the results of the current model and the other well-established model shows that the KL-MLP-LSTM model is reliable for estimating the temperature distributions during the thermal process.

**INDEX TERMS** Spatiotemporal modeling, nonlinear distributed thermal processes, Karhunen-Loève decomposition, multilayer perceptron, long short-term memory.

## I. INTRODUCTION

In the industrial field, many thermal processes can be characterized by the nonlinear distributed parameter systems (DPSs), in which the input and output states vary both in time and space domain [1]. There are plenty of these thermal processes, such as the curing process of a snap curing oven in the integrated circuit (IC) packaging industry, the exothermic catalytic reaction of a catalytic rod in the chemical industry, the endothermic catalytic reaction of a packed-bed reactor in the chemical industry, the thermal process of a lithium-ion battery in the automobile industry, and so on [2], [3]. Mathematically, such systems are often described by partial differential equations (PDEs) using the first-principle modeling [4]. Some efficient methods have been proposed to solve PDEs, for example, Anitescu *et al.* has presented a practical approach using artificial neural networks (ANNs) and an adaptive collocation strategy [5]. However, the complex process and nonlinear system usually make it difficult to obtain accurate PDEs for the industrial thermal processes. As well known, the data-driven spatiotemporal model based on the historical data can describe the spatiotemporal dynamics in DPS without the prior knowledge. Therefore, it has been widely applied in modeling nonlinear distributed parameter processes [6]–[11]. The modeling challenges in DPSs can be presented as follows: 1) the analytical model is hard to derive for the complex process; 2) the temperature distribution in DPSs is time/space coupled, which contain infinite-dimensional characteristics that is difficult to model; 3) the complex nonlinear dynamics often exist not

The associate editor coordinating the review of this manuscript and approving it for publication was Meng Huang.

only in the time direction, but also in the space direction; 4) the complicated nonlinear structure in the temporal model approximating the actual condition is hard to be determined.

A spatiotemporal variable of DPSs can be expressed by an infinite number of spatial basis functions and the corresponding temporal coefficients [12]. For the data-driven spatiotemporal model, the key is to choose proper spatial basis functions for dimension reduction and construct the nonlinear finite-order (low-order) temporal estimation model. Then spatiotemporal model can be constructed by spatiotemporal synthesis. Clearly, both spatial basis functions and temporal model influence the accuracy of the model.

Therefore, the selection of spatial basis functions is critical to the spatiotemporal modeling performance. There are a wide variety of methods, such as the weighted residual method (WRM) [13], the finite-difference method (FDM) [14], [15], the finite-element method (FEM) [16], the spectral method (SM) [17], and the Karhunen-Loève (KL) method [18], [19]. The KL decomposition method has been widely used by lots of researchers, and it turns out to be a suitable and effective approach for data modeling [20]–[22]. Accordingly, a data model based on the KL decomposition method is a powerful tool for complicated DPS [23]. In addition, by separating the time and space information of the nonlinear DPS, the KL decomposition can achieve a more accurate lower-order model, which can capture the dominant dynamics of the PDE system described by a few dominant (slow) modes.

After choosing the dominant spatial basis functions, it is easy to obtain the system temporal coefficients. Then the input and output information of the nonlinear temporal model can be derived from the system input and the system temporal coefficients. Finally, the temporal model can be constructed by using traditional system identification techniques, which mainly includes statistical inference and machine learning methods. For example, Montaseri *et al.* employed a Multilayer Neural Network (MNN) to construct the finite-dimensional model and applied the model to predict future states in the model predictive control (MPC), and the simulation results indicated this model satisfied all considered control objectives [24]. Qi *et al.* presented a fuzzy-based spatiotemporal multi-modeling approach for nonlinear DPS to improve model accuracy [25]. Liu *et al.* developed extreme learning machine (ELM) for the low-order model, and the temperature distributions of the whole battery can be estimated based on the identified low-order model in real time [26]. Lu *et al.* adopted a least-squares support vector machine (LS-SVM) to model nonlinear time dynamics successfully and applied this spatiotemporal method on a practical curing thermal process [27]. Xu *et al.* proposed a spatiotemporal model which integrated the finite Gaussian mixture model (FGMM) with principal component regression (PCR) for complex nonlinear DPSs, and this model showed strong ability to track and handle the complex nonlinear dynamics [28]. However, all the aforementioned temporal models do not consider the intrinsic structure of low-order models. As a result, they cannot deal with the situations with two coupled nonlinear dynamics such as the inputs and outputs of many DPS processes with complex nonlinearities. Xu *et al.* proposed a dual LS-SVM temporal model combined with the KL method to confirm the model structure [29]. This model turned out to be a high accuracy spatiotemporal model while it treated the two modules equally without considering their features and differences, respectively.

The complicated nonlinear structure can be decomposed into two parts: a nonlinear function relating to the input and a counterpart relating to the output [29]. The former function can be represented by a variety of models such as SVM [30], MLP [31], ELM [32], fuzzy [33], etc. As for the latter function, it is suitable to choose a time series model. As we know, the long short term memory (LSTM) is a specific recurrent neural network (RNN) architecture designed to model time series, and it is more accurate than the RNN [34]. Accordingly, the LSTM has been widely used in time series prediction [35]–[38], whereas it has not been adopted in DPS spatiotemporal model in the related literature. This paper will focus on the MLP-LSTM hybrid temporal model.

In this work, a novel data-driven spatiotemporal model, which combines the KL method with the MLP-LSTM temporal model is established to overcome these challenges aforementioned. The KL method is used for dimension reduction and extracting the spatial basis functions, and the MLP-LSTM temporal model is adopted to solve the two inherently coupled nonlinearities based on the temporal coefficients and system inputs. In this way, the KL-MLP-LSTM can approximate the real DPS situation of the two coupled nonlinear dynamics more accurately. The proposed spatiotemporal model is verified by a snap curing oven thermal process.

The main contributions of this paper are summarized as:

1) The KL method for spatiotemporal separation: With this modeling mechanism, the spatiotemporal variable can be decoupled into sets of unit orthonormal spatial basic functions and corresponding time coefficients.
2) The MLP-LSTM for temporal modeling: With the MLP-LSTM, the temporal model can satisfactorily describe the actual situations with two inherently coupled nonlinearities and therefore work more efficiently.
3) Comparisons with the experiment on a curing oven validate the effectiveness of the proposed method.

The remaining part of this article is organized as follows. The problem is described briefly in Sections II. In Section III, the spatiotemporal modeling method based on the KL-MLP-LSTM is constructed. Section IV gives the experiment result of a snap curing oven thermal process to verify the applicability of the proposed method. Finally, a brief summary is given in Section V.

## II. PROBLEM DESCRIPTION

According to the heat transfer laws, the general PDE descriptions of the industrial thermal processes can be

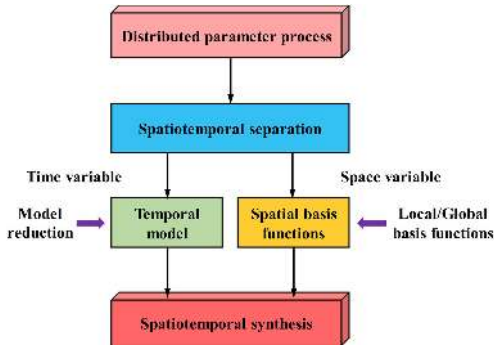


**FIGURE 1.** Schematic representation of the spatiotemporal separation.

expressed as follows:

$$\rho c \frac{\partial \left(y\left(S,t\right)\right)}{\partial t} = \nabla^2 \left(y\left(S,t\right)\right) + f_c\left(y\left(S,t\right)\right) + f_r\left(y\left(S,t\right)\right) + \rho Q\left(S,t\right), \quad (1)$$

where $y\left(S,t\right)$ is the spatiotemporal variable temperature at location $S = (x, y, z)$ and time $t$, $\rho$ is the density parameter, $c$ is the heat capacity, $\nabla^2 = \frac{\partial}{\partial x^2} + \frac{\partial}{\partial y^2} + \frac{\partial}{\partial z^2}$ is the Laplacian operator, $f_c\left(y\left(S,t\right)\right)$ and $f_r\left(y\left(S,t\right)\right)$ are unknown nonlinear effects of convection and radiation, respectively. $Q\left(S,t\right)$ is the heating source.

Based on the principle that a continuous function can be approximated by Fourier series [39], the DPS spatiotemporal variable can be expanded as:

$$y\left(S,t\right) = \sum_{i=1}^{\infty} \phi_i\left(S\right) a_i\left(t\right). \quad (2)$$

As only the first $n$ slow modes contribute to the whole system [23], equation (2) is reduced to the following finite system in practical implementation:

$$y\left(S,t\right) = \sum_{i=1}^{n} \phi_i\left(S\right) a_i\left(t\right). \quad (3)$$

The spatiotemporal variables can be separated into two parts: a set of spatial basis functions and corresponding temporal model. It is vital to select proper spatial basis functions and construct the corresponding finite-order temporal model. Then, the spatiotemporal system can be reconstructed through the spatiotemporal synthesis. The procedure is displayed in Fig. 1.

According to the reported theoretical derivation and analysis [29], the temporal variable $a_i\left(t\right)$ can be expressed as

$$a_i\left(t\right) = g^i\left(u(t)\right) + h^i\left(a_i\left(t-1\right)\right), \quad (4)$$

where the model $a_i\left(t\right)$ consists of two nonlinear blocks $g^i(\cdot)$ and $h^i(\cdot)$.

As described above, the main difficulties and challenges are summarized as follows:

1) The thermal process is time/space coupled, which needs to be transformed into finite dimensions, therefore the spatial basis functions should be well chosen.

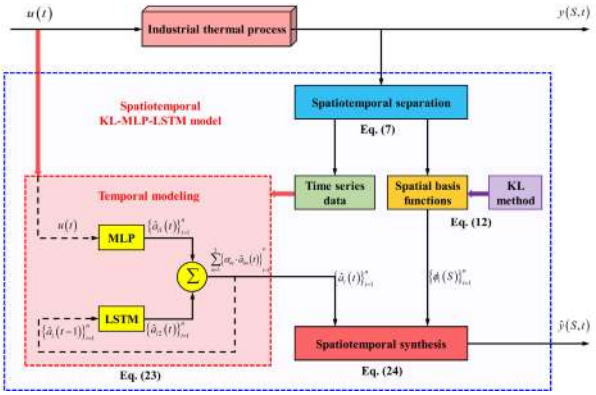


**FIGURE 2.** Schematic representation of the KL-MLP-LSTM based spatiotemporal model.

2) The thermal dynamics leads to two nonlinear blocks in the temporal model, which poses additional difficulty for traditional modeling methods to approximate the real situations.

## III. KL-MLP-LSTM SPATIOTEMPORAL MODEL

### A. THE WHOLE PROCESS OF THE PROPOSED MODEL

To address the above problems, a KL-MLP-LSTM hybrid model is developed for spatiotemporal estimation of the thermal process. The configuration of the proposed spatiotemporal modeling is depicted in Fig. 2. The detailed descriptions of the proposed model are shown as follows:

1) The KL decomposition method is used for extracting the spatial basis functions and reducing the infinite-dimensional model to a set of ordinary differential equations. Details of the KL decomposition method are described in Section III.B.
2) A data-based MLP-LSTM hybrid time series model is constructed to match the dual-model structure and deal with the two inherently coupled nonlinearities. Details of the MLP-LSTM hybrid temporal model can be found in Section III.C.
3) The spatiotemporal temperature distribution model can be reconstructed through spatiotemporal synthesis. Details of spatiotemporal synthesis are presented in Section III.D.

### B. KL DECOMPOSITION METHOD

The KL decomposition is a statistical analysis technique which is very effective to obtain a low-dimensional model from experimental or numerical data [18], [40]. It is a global linear method and can preserve the global Euclidean structure effectively [41]. The key idea of the KL method is to find the spatial basis functions with the minimal number to represent the dominant characteristics of the system.

For simplicity, suppose that $\{y(S_i,t)|i = 1, \cdots, N; t = 1 \cdot \cdot, L; S = (x, y, z)\}$ is the process output obtained from experiments distributed in time and space. The spatiotemporal variable can be expressed as

$$y\left(S,t\right) = \sum_{i=1}^{\infty} \phi_i\left(S\right) a_i\left(t\right), \quad (5)$$

where $\{\phi_i(S)\}_{i=1}^{\infty}$ are the spatial basis functions and $\{a_i(t)\}_{i=1}^{\infty}$ are the temporal coefficients. The spatial basis functions $\{\phi_i(S)\}_{i=1}^{\infty}$ possess the following orthogonal property:

$$\left(\phi_i(S), \phi_j(S)\right) = \int_{\Omega} \phi_i(S)\,\phi_j(S)dS = \begin{cases} 0 & i \neq j \\ 1 & i = j. \end{cases} \tag{6}$$

To obtain the most typical or characteristic spatial basis functions $\{\phi_i(S)\}_{i=1}^{\infty}$ of the system by the KL method, an optimization problem by minimizing the following objective function should be solved [42]:

$$\begin{cases} \min\limits_{\phi_i(S)} \left\langle \left[ y(S,t) - \sum\limits_{i=1}^{n} \phi_i(S)\left(\phi_i(S), y(S,t)\right) \right]^2 \right\rangle \\ s.t.\ (\phi_i(S), \phi_i(S)) = 1, \phi_i(S) \in L^2(\Omega), \quad i = 1, \cdots, n, \end{cases} \tag{7}$$

where $\langle f(S,t)\rangle = \frac{1}{L}\sum_{i=1}^{L} f(S,t)$ represents the ensemble average; $\phi_i(S)$ is unique in the given constraint $(\phi_i(S), \phi_i(S)) = 1$. Applying the Lagrange multiplier method, we have

$$J = \left\langle \left[ y(S,t) - \sum_{i=1}^{n} \phi_i(S)\left(\phi_i(S), y(S,t)\right) \right]^2 \right\rangle + \sum_{i=1}^{n} \lambda_1 \left( \left(\phi_i(S), \phi_i(S)\right) - 1 \right), \tag{8}$$

and the necessary condition of its extreme value can be expressed as

$$\begin{cases} \int_{\Omega} R(S,\zeta)\phi_i(\zeta)\,d\zeta = \lambda_i \phi_i(S) \\ s.t.\ (\phi_i(S), \phi_i(S)) = 1, \quad i = 1, \cdots, n, \end{cases} \tag{9}$$

where $R(S,\zeta) = \langle y(S,t), y(\zeta,t)\rangle$ is the spatial two-point correlation function.

Because of the discretization of the data in space, the numerical method should be used to solve (9), which can be transformed into an $N \times N$ matrix eigenvalue problem by discretizing the integral equation. Therefore, the maximum number of eigenvalues obtained at $N$ sampled spatial locations is $N$. Interpolation is applied to locations where the sample points are unavailable.

To solve (9) efficiently, the spatial basis function $\phi_i(S)$ is supposed to be expressed as follows:

$$\phi_i(S) = \sum_{t=1}^{L} \delta_{it} y(S,t). \tag{10}$$

Combining (9) and (10), we obtain

$$\int_{\Omega} \frac{1}{L} \sum_{t=1}^{L} y(S,t)\, y(\zeta,t) \sum_{k=1}^{L} \delta_{ik} y(\zeta,k)\, d\zeta = \lambda_i \sum_{t=1}^{L} \delta_{it} y(S,t). \tag{11}$$

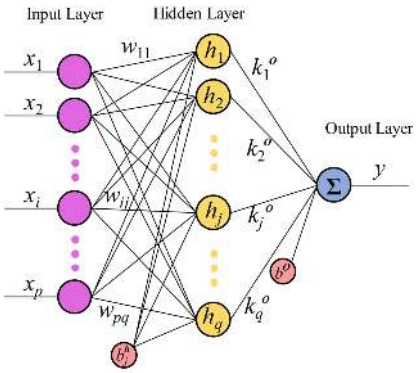


**FIGURE 3.** Schematic representation of the MLP.

Then, the former $N \times N$ matrix eigenvalue problem in (9) can be transformed into a $L \times L$ matrix eigenvalue problem as follows:

$$C\delta_i = \lambda_i \delta_i, \tag{12}$$

where $C_{tk} = \frac{1}{L}\int_{\Omega} y(\zeta,t)\, y(\zeta,k)d\zeta$ is the temporal two-point correlation function, and $\delta_i = [\delta_{i1}, \cdots, \delta_{iL}]^T$ is the $i$th eigenvector. Eigenvectors $\{\delta_i\}_{i=1}^{L}$ and the corresponding eigenvalues $\{\lambda_i\}_{i=1}^{L}$ are obtained by solving (12). Then, the orthogonal spatial basis functions $\{\phi_i(S)\}_{i=1}^{L}$ are constructed by (10).

The eigenvalues are assumed to satisfy $\lambda_1 > \lambda_1 > \cdots > \lambda_K$ and the corresponding spatial basis functions are $\{\phi_1(S), \phi_2(S), \cdots, \phi_K(S)\}$ where $K \leq \min(N, L)$ is the maximum number of nonzero eigenvalues. In theory, the more real situations of the spatiotemporal system are reflected, the more complicated the model will be built when taking the whole spatial basis functions into consideration. Therefore, it is necessary to choose $n$ $(n \leq K)$ dominant spatial basis functions to approximate the major dynamics of the spatiotemporal system. The parameter $\eta$ denotes the ratio of the sum of the $n$ largest eigenvalues to the total sum of all eigenvalues, which is written as:

$$\eta = \sum_{i=1}^{n} \lambda_i \Big/ \sum_{i=1}^{K} \lambda_i, \quad (n \leq K), \tag{13}$$

where $n$ is determined according to (13) as the ratio $\eta$ exceeds 99.9%.

### C. MLP-LSTM TEMPORAL MODEL

ANNs have been a topic of great interest in the machine learning community due to their ability to solve very difficult problems [43]. The MLP is a feedforward neural network which is trained with the backpropagation algorithm [44]. The MLP shown in Fig. 3 is a flexible and adjustable architecture whose integral structure general consisting of three layers, i.e., an input layer, a hidden layer, and an output layer [45]. The number of neuron in both input and output layers depends on the actual situations.

The vector $x = \left[x_1, x_2, \cdots, x_p\right]^T$ represents the $p$ input variables, and the numbers of neurons in the hidden layer and

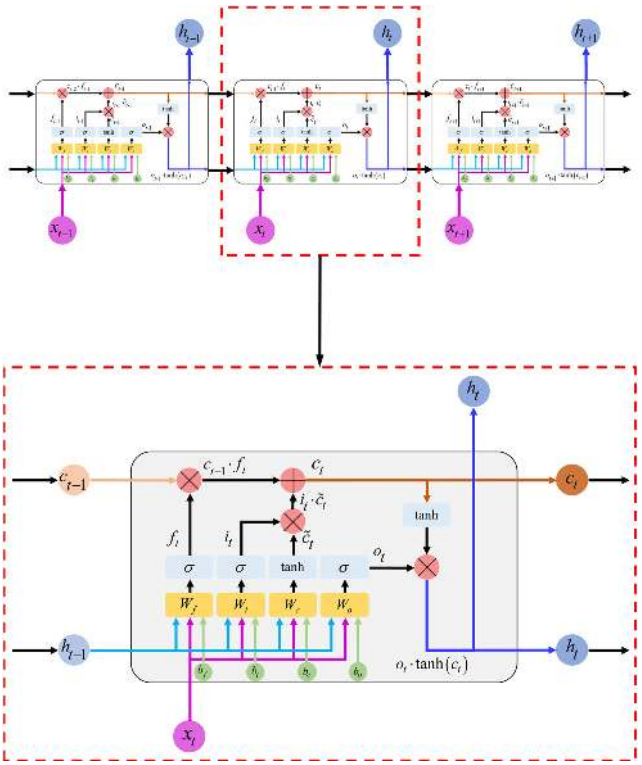


**FIGURE 4.** Schematic representation of the LSTM.

output layer are $q$ and 1, respectively. The weight $\mathbf{W} = [w_{ij}]$ is a $p \times q$ matrix that connects the $p$ inputs with the $q$ hidden layer nodes. The vector $\mathbf{b} = \left[b_j^h\right]$ $(j \in [1, q])$ is the biases for the hidden layer nodes. The weight $\mathbf{k} = \left[k_j^O\right]$ $(j \in [1, q])$ is the vector that connects the $q$ hidden layer nodes with the single output. Therefore, the input-output of the MLP can be expressed as:

$$y = g\left(\sum_{j=1}^{q} k_j^o f\left(\left(\sum_{i=1}^{p} w_{ij} x_i\right) + b_j^h\right) + b^o\right), \tag{14}$$

where $f(\cdot)$ and $g(\cdot)$ are the activation transfer functions of the neuron, which have many different forms such as step function, sigmoid function, etc. [46].

As a special kind of RNN, Long Short Term Memory (LSTM) is designed to overcome the limitations of long-term dependency [47]. The LSTM can be applied for modeling time series data with high nonlinearities and long interval [48]. A normal LSTM consists of an input layer, a recurrent hidden layer, and an output layer, and storage elements are employed to transfer information from past output to current output. Different from the other neuronal networks, the memory block is the basic unit of the hidden layer [49]. It is comprised of memory cells with self-connections memorizing the temporal state, and two adaptive gating units (named input gate and output gate) are used to control information flow in the block. Fig. 4 shows the architecture of the LSTM.

The LSTM has three control signals, each of which is a nonlinear function $\sigma$ activated by a weighted sum of the

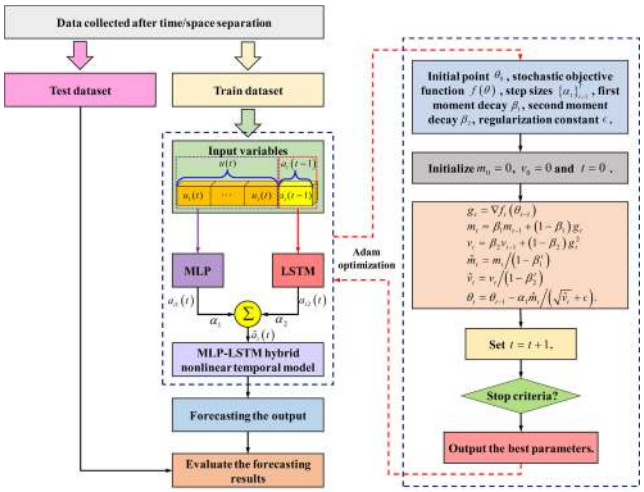


**FIGURE 5.** Procedure of the MLP-LSTM temporal model.

current input observation $x_t$ and previous output $h_{t-1}$. The forget gate $f_t$, input gate $i_t$, and output gate $o_t$ can be expressed as:

$$f_t = \sigma\left(W_f \cdot h_{t-1} + W_f \cdot x_t + b_f\right), \tag{15}$$
$$i_t = \sigma\left(W_i \cdot h_{t-1} + W_i \cdot x_t + b_i\right), \tag{16}$$
$$o_t = \sigma\left(W_o \cdot h_{t-1} + W_o \cdot x_t + b_o\right), \tag{17}$$

where $W_f$, $W_i$ and $W_o$ are the weight matrix of forget gate, input gate, and output gate, respectively; $b_f$, $b_i$ and $b_o$ are the bias of forget gate, input gate, and output gate, respectively. The forget gate $f_t$ determines whether to retain or forget the previous state $c_{t-1}$ of the LSTM. The input gate $i_t$ decides whether to update the state of the LSTM using the current input, and the output gate $o_t$ determines whether to pass the current state $h_t$ to the next iteration. The final output of the LSTM is determined by the new cell state $c_t$ as follows:

$$\tilde{c}_t = \tanh\left(W_c \cdot h_{t-1} + W_c \cdot x_t + b_c\right), \tag{18}$$
$$c_t = f_t \otimes c_{t-1} + i_t \otimes \tilde{c}_t, \tag{19}$$
$$h_t = o_t \otimes \tanh\left(c_t\right), \tag{20}$$

where $\tilde{c}_t$ is the new gated input at time $t$, $c_{t-1}$ is the gated previous state at time $t-1$, $W_c$ is the weight matrix of the input cell state, $b_c$ is the bias of the input cell state, tanh is the activation function, and $\otimes$ is the elementwise multiplication operator.

The full procedure of the MLP-LSTM temporal model is listed as follows and illustrated in Fig. 5.

*Step 1:* Maintain the two parts of the input data which comes from system inputs $u(t)$ and spatiotemporal separation $a_i(t-1)$.

*Step 2:* Apply the MLP to establish the dynamic relationship between the input $u(t)$ and the prediction $a_i(t)$. The output of this MLP part is $a_{i1}(t)$.

*Step 3:* Apply the LSTM to construct the effect of the output from the previous $a_i(t-1)$ on the prediction $a_i(t)$. The output of this LSTM part is $a_{i2}(t)$.

*Step 4:* Construct MLP-LSTM model.

Based on the inputs and outputs in the MLP and the LSTM, the output of the MLP-LSTM is $a_i(t) = \alpha_1 \cdot a_{i1}(t) + \alpha_2 \cdot a_{i2}(t)$, i.e., $a_i(t) = g^i(u(t)) + h^i(a_i(t-1))$. Then the Adam algorithm [50] is used to train the model and optimize the parameters to prepare for the following spatiotemporal synthesis procedure.

### D. SPATIOTEMPORAL SYNTHESIS

Combining the MLP-LSTM estimations of the output $\{\hat{a}_i(t)\}_{i=1}^{L}$ with spatial basis functions $\{\phi_i(S)\}_{i=1}^{L}$, we can obtain the KL-MLP-LSTM spatiotemporal model as follows:

$$\hat{a}_i(t) = \hat{g}^i(u(t)) + \hat{h}^i(a_i(t-1)), \tag{21}$$

$$\hat{y}(S,t) = \sum_{i=1}^{n} \phi_i(S)\hat{a}_i(t). \tag{22}$$

Based on all the procedures above, the KL-MLP-LSTM spatiotemporal model of the DPS can be set up according to the following steps:

*Step 1:* Impose systematic excitation $u(t)$ to the DPS, and the system output is $\{y(S_i,t)\}_{i=1,t=1}^{N,L}$, from which the spatial basis functions $\{\phi_i(S)\}_{i=1}^{n}$ and temporal coefficients $\{a_i(t)\}_{i=1,t=1}^{n,L}$ can be obtained by using the KL decomposition method.

*Step 2:* Apply the MLP-LSTM to identify the system inputs $u(t)$ and temporal coefficients $\{\hat{a}_i(t)\}_{i=1}^{L}$ to finish the model identification.

*Step 3:* Reconstruct the system model by spatiotemporal synthesis and get the model prediction.

### E. SUMMARY

To sum up, compared to the previous spatiotemporal models, the proposed method has a dual-model structure, in which the two inherently coupled nonlinearities are fully considered. Besides, compared to other dual-models [29], the proposed model utilizes a hybrid model containing two different sub-models to approximate the real situations.

## IV. EXPERIMENTAL RESULT

In this section, a snap curing oven thermal process is employed to demonstrate the effectiveness of the proposed model. Curing is one of the most critical processes in the semiconductor back-end packaging industry, and snap curing oven is a vital equipment to provide the required temperature distribution during the curing process. The curing quality of the chip profoundly affects the quality and the life of the final product. It is essential to obtain the temperature distribution model of the snap curing oven since the chip quality requires high accuracy of the temperature distribution. The snap curing oven system is shown in Fig. 6. As shown in Fig. 7, the oven has four heaters controlled by pulse-width modulation (PWM) signals which must be persistently excited for modeling. As can be seen from Fig. 8, sixteen thermocouples, i.e., 16 spatial nodes ($s$), distributed uniformly in the same horizontal plane are employed for temperature sensing in the process. Measurements from the

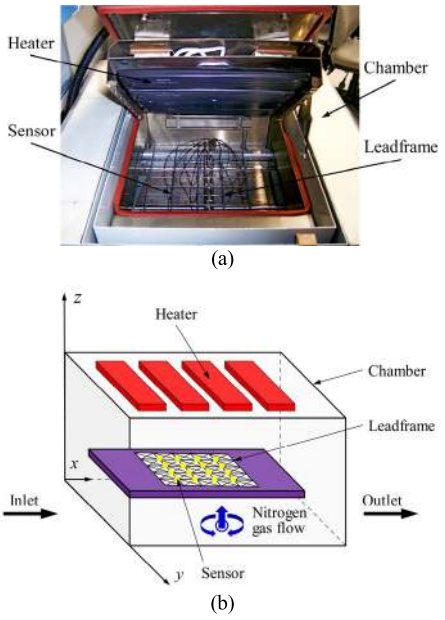


**FIGURE 6.** The snap curing oven system. (a) the actual snap curing oven system; (b) diagram of the internal structure.

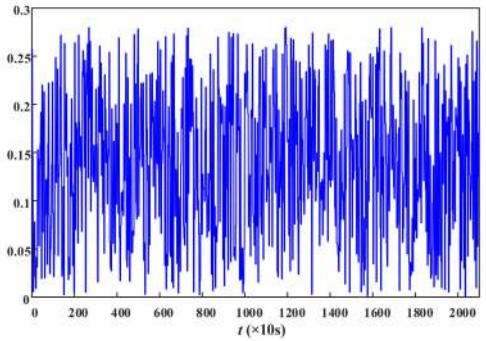


**FIGURE 7.** Input PWM signal of heater 1.

highlighted sensors ($s_1 - s_5, s_7 - s_{10}, s_{12} - s_{16}$) are used for modeling while measurements from the other two sensors ($s_6, s_{11}$) are used for model validation. The output temperature signal is collected under the dSPACE platform due to its strong real-time performance, high reliability, and good expansibility, etc. In total, about 2100 time series measurements are collected by all the sensors in a sampling interval $\Delta t = 10s$ in the experiment, which implies that the number of temporal nodes $L$ is 2100. About 1100 samples are adopted as training data and the other 1000 samples as testing data.

The performance of the proposed model is evaluated by the following five statistics which have different meanings and widely adopted to evaluate the performance of the prediction model. These five indicators can represent the comparative results of the methods from different perspectives. In particular, *RMSE* describes the total spatiotemporal error information of all sensors over all time periods. *TNAE* represents the total temporal error information at every sensor. *SNAE* means the total spatial error information at every moment.

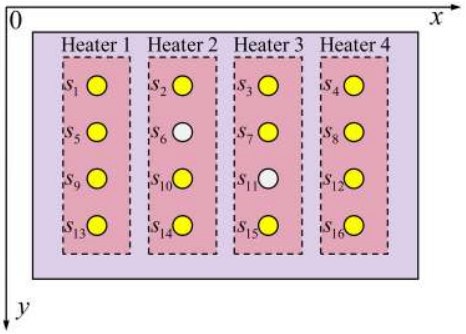


**FIGURE 8.** Sensor locations for modeling and validation.

1) Spatiotemporal prediction error ($e$)

$$e(S_i, t) = y(S_i, t) - \hat{y}(S_i, t), \quad i = 1, \cdots, N. \tag{23}$$

2) Absolute relative error ($ARE$)

$$ARE = \frac{\left|y(S_i, t) - \hat{y}(S_i, t)\right|}{y(S_i, t)}, \quad i = 1, \cdots, N. \tag{24}$$

3) Root mean square error ($RMSE$)

$$RMSE = \sqrt{\frac{1}{NL}\sum_{i=1}^{N}\sum_{t=1}^{L}\left(y(S_i, t) - \hat{y}(S_i, t)\right)^2}. \tag{25}$$

4) Temporal normalized absolute error ($TNAE$)

$$TNAE = \frac{1}{L}\sum_{t=1}^{L}\left|y(S_i, t) - \hat{y}(S_i, t)\right|, \quad i = 1, \cdots, N. \tag{26}$$

5) Spatial normalized absolute error ($SNAE$)

$$SNAE = \frac{1}{N}\sum_{i=1}^{N}\left|y(S_i, t) - \hat{y}(S_i, t)\right|, \quad t = 1, \cdots, L. \tag{27}$$

In the KL-MLP-LSTM spatiotemporal modeling process, three dominant spatial basis functions ($n = 3$) have been chosen according to the KL decomposition method. Fig. 9 shows the selected spatial basis functions.

After getting the three spatial basis functions, a series of low-dimensional time series data ($\{a_i(t)\}_{i=1}^{3}$) can be obtained by projecting high-dimensional spatiotemporal data onto the spatial basis functions. The most critical procedure is to use the MLP-LSTM to identify the model based on the input signal $u(t)$ and the temporal coefficients $a(t)$. The structure of the MLP has five layers, i.e., an input layer, three hidden layers, and an output layer. The MLP neural network structure is confirmed as 4-25-17-17-1. The structure of the LSTM network has three layers, and the number of the hidden layer nodes is 50. The time step length is 10, and the mini-batch size is 32. In Adam algorithm, the learning rate is 0.001, and the other parameters are set as default. The first two low-order actual temporal coefficients predicted by the MLP-LSTM model, i.e., $a_1(t)$, $a_2(t)$, are compared to those obtained by the MLP and the dual LS-SVM [29] under the same conditions in this example. Fig. 10 and Fig. 11 show the $a_1(t)$, $a_2(t)$ predicted by the two above-mentioned

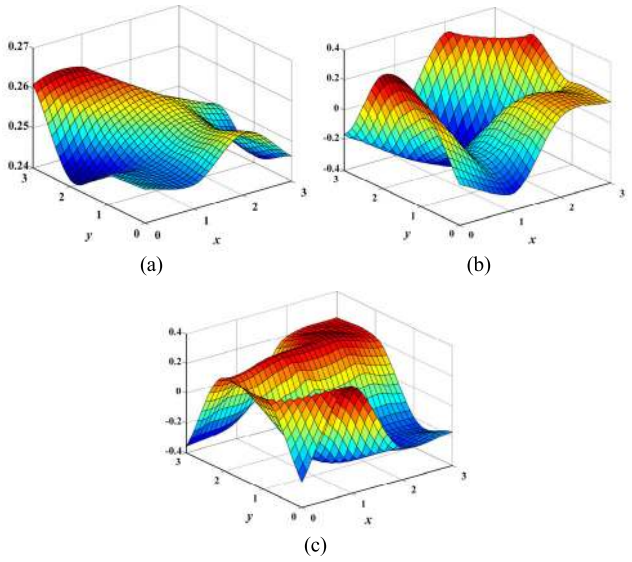


**FIGURE 9.** The selected spatial basis functions. (a) spatial basis function ($i = 1$); (b) spatial basis function ($i = 2$); (c) spatial basis function ($i = 3$).

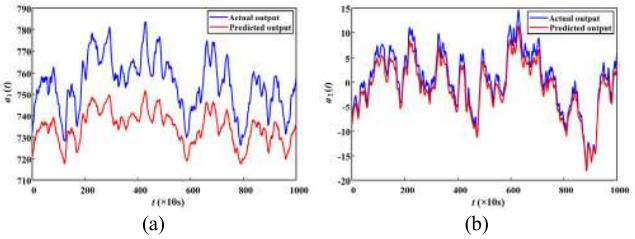


**FIGURE 10.** The comparison of the actual temporal coefficients and its prediction by MLP mode. (a) temporal coefficient $a_1$; (b) temporal coefficient $a_2$.

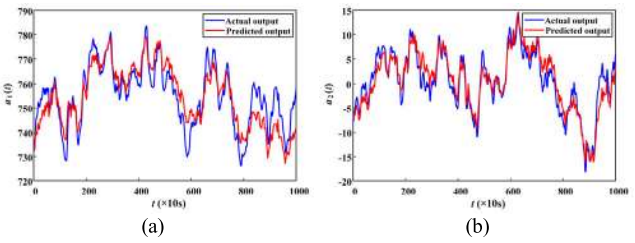


**FIGURE 11.** The comparison of the actual temporal coefficients and its prediction by dual LS-SVM mode. (a) temporal coefficient $a_1$; (b) temporal coefficient $a_2$.

methods, respectively, and the corresponding results of the proposed MLP-LSTM are demonstrated in Fig. 12. It is obvious that the predictions given by the MLP-LSTM can track the tendency of the actual results very well and show the best prediction performance among the three models.

Finally, the temperature distribution can be reconstructed by synthesizing the obtained spatial basis functions and the identified temporal model. Once the spatiotemporal model of thermal process is obtained, the temperature predicted by the temporal model under the designed testing input conditions are compared with the actual measured temperature at the indicated times to validate the proposed model. Fig. 13 shows

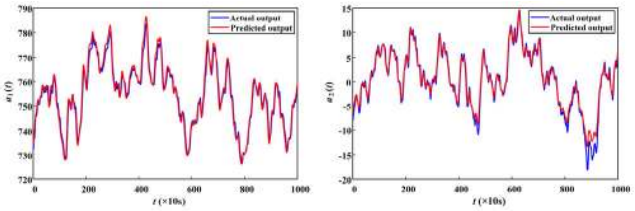


**FIGURE 12.** The comparison of the actual temporal coefficients and its prediction by MLP-LSTM mode. (a) temporal coefficient $a_1$; (b) temporal coefficient $a_2$.

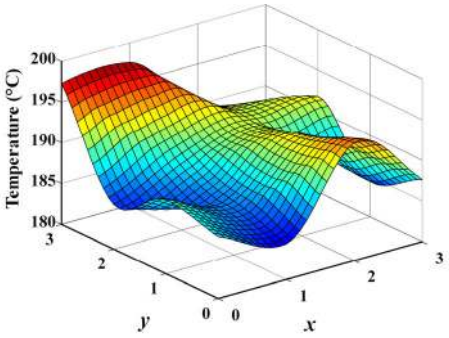


**FIGURE 13.** The real temperature distributions at the last sampling ($t = 21000s$).

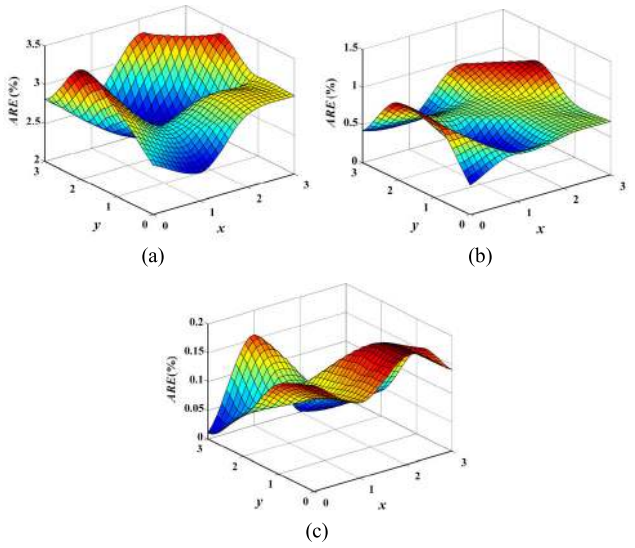


**FIGURE 14.** *ARE* distributions at the last sampling ($t = 21000s$). (a) *ARE* distribution of the KL-MLP; (b) *ARE* distribution of the KL-dual LS-SVM; (c) *ARE* distribution of the KL-MLP-LSTM.

the real temperature distributions over the spatial domain at the last sampling ($t = 21000s$) and Fig. 14 displays the *ARE* distributions of the predicted temperature by three different models at the same time. Compared with the KL-MLP and the KL-dual LS-SVM, the maximum *ARE* of the KL-MLP-LSTM is reduced from 3.42% and 1.04% to 0.19%, respectively, which means the proposed method has a higher prediction accuracy. The KL-MLP-LSTM can achieve the best prediction performance due to the high accuracy of the MLP-LSTM temporal model.

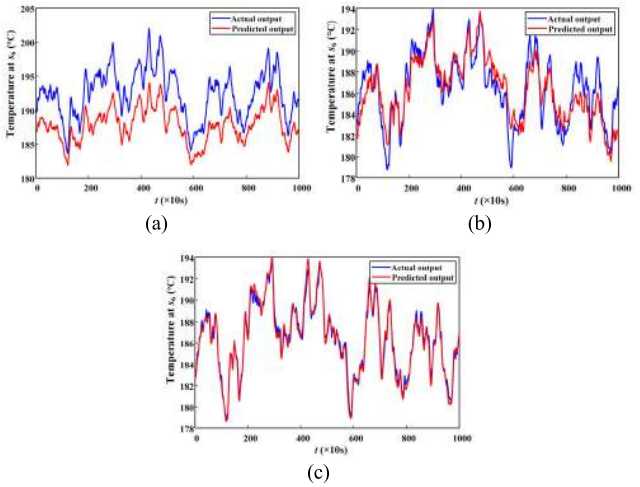


**FIGURE 15.** Performance comparisons of the three methods for sensor $s_6$. (a) performance of the KL-MLP for sensor $s_6$; (b) performance of the KL-dual LS-SVM for sensor $s_6$; (c) performance of the KL-MLP-LSTM for sensor $s_6$.

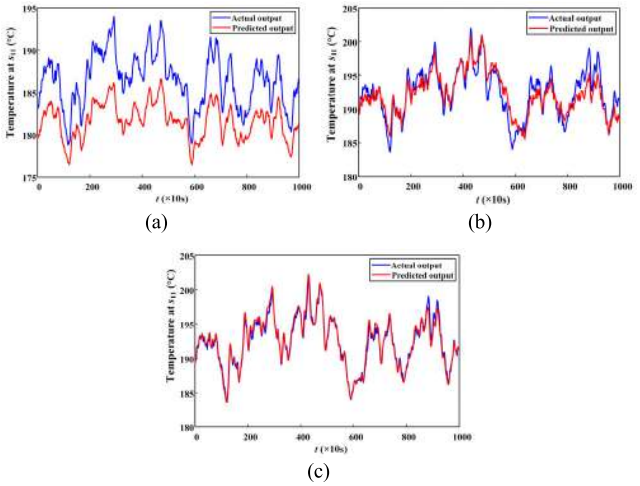


**FIGURE 16.** Performance comparisons of the three methods for sensor $s_{11}$. (a) performance of the KL-MLP for sensor $s_{11}$; (b) performance of the KL-dual LS-SVM for sensor $s_{11}$; (c) performance of the KL-MLP-LSTM for sensor $s_{11}$.

Two untrained sensors ($s_6, s_{11}$) outputs are compared to the experimental data for model performance validation. Fig. 15 and Fig. 16 show the predicted temperatures in comparison with actual measurements of the test samples for the sensor $s_6$ and $s_{11}$ by three different models. As can be seen from these figures, the prediction performance of the KL-MLP-LSTM is excellent for $s_6$ as well as at $s_{11}$. The *ARE* of the three modeling methods for the sensor $s_6$ and $s_{11}$ are shown in Fig. 17. Whether for the sensor $s_6$ or $s_{11}$, most of *ARE* values of the KL-MLP-LSTM are basically below 0.60%, while those of the KL-MLP and the KL-dual LS-SVM are below 4.50% and 1.50%.

For further comparisons of these models over the time and spatial domain, the criteria of *RMSE*, *SNAE* and *TNAE* are adopted. Table 1 shows the *RMSE* of the three modeling methods. It is obvious that the KL-MLP-LSTM has the best

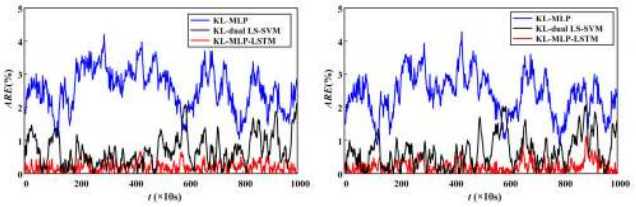


**FIGURE 17.** ***ARE*** **comparisons of the three methods for sensor $s_6$ and $s_{11}$. (a)** ***ARE*** **comparison for sensor $s_6$; (b)** ***ARE*** **comparison for sensor $s_{11}$.**

**TABLE 1.** ***RMSE*** **of three methods.**

| Method | *RMSE* |
|---|---|
| KL-MLP | 5.24 |
| KL-dual LS-SVM | 1.72 |
| KL-MLP-LSTM | 0.50 |

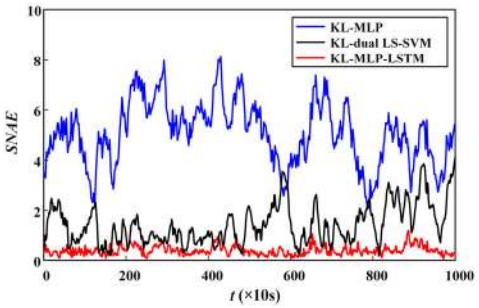


**FIGURE 18.** ***SNAE*** **comparisons of the three methods.**

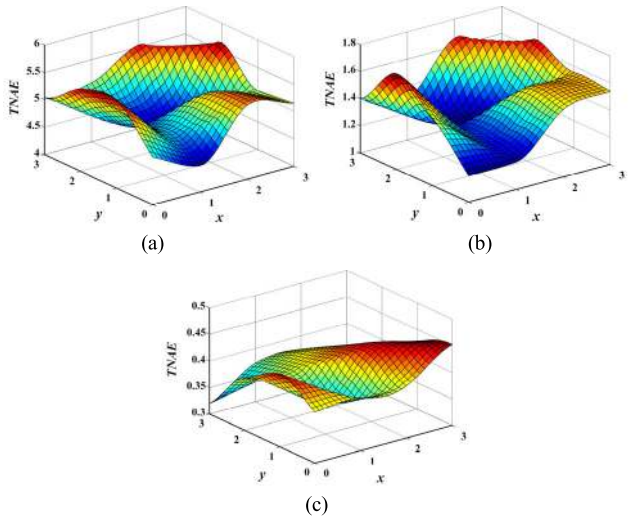


**FIGURE 19.** ***TNAE*** **distributions of the three methods. (a)** ***TNAE*** **distribution of the KL-MLP; (b)** ***TNAE*** **distribution of the KL-dual LS-SVM; (c)** ***TNAE*** **distribution of the KL-MLP-LSTM.**

prediction performance with minimum *RMSE* of 0.50 compared to those of the KL-MLP and the KL-dual LS-SVM. Fig. 18 displays the *SNAE* of three different models. The *SNAE* of the KL-MLP-LSTM is almost less than 0.50, which means that its performance is much better than the other two methods. Fig. 19 presents the *TNAE* of these models. The maximum *SNAE* of the KL-MLP-LSTM is 0.45, while those for the other two models are 5.65 and 1.71, respectively, which indicates the proposed method has a greater prediction capability.

All the results presented above indicate that the proposed KL-MLP-LSTM model can perform better than the other two models in the experiment, which demonstrates its excellent performance in both the time domain and the spatial domain. Since the KL-MLP model is a single-structure method, it is difficult to obtain an excellent result. Although the KL-dual LS-SVM model is a double-structure method with a higher prediction accuracy, it does not consider the time series factor. As the proposed KL-MLP-LSTM model in our paper is a double-structure method considering the time series problem, it is more consistent with the practical DPS, and then performs superior to the other methods.

## V. CONCLUSION

In this work, a data-driven spatiotemporal model named KL-MLP-LSTM is proposed. The model combines the KL decomposition for dimension reduction and the spatial basis functions extraction by a novel MLP-LSTM hybrid time series model to deal with the two inherently coupled nonlinearities. Data set from a snap curing oven thermal process is performed to validate the reliability of the spatiotemporal KL-MLP-LSTM model reconstructed through spatiotemporal synthesis. In addition, the proposed KL-MLP-LSTM is compared with the KL-MLP and the KL-dual LS-SVM. The results indicate that the KL-MLP-LSTM model presents the best prediction performance among the three methods. The proposed spatiotemporal method can be applied to a parabolic distributed parameter system with a feedback input signal in any field, especially for a nonlinear system.

The contributions of this paper can be summarized into the following aspects: 1) fast spatial basis functions computation; 2) perfect model structure matching; 3) high model performance; 4) easy implementation. Though this proposed model can work very well, we will continue this research in the future and put attention on the following aspects: 1) simplify the model structure; 2) optimize the number and the placement of sensors; 3) expand the model applications.

## ACKNOWLEDGMENT

The authors would like to thank Dr. Yue Zhang for many helpful discussions.

## REFERENCES

[1] H. Deng, H.-X. Li, and G. Chen, "Spectral-approximation-based intelligent modeling for distributed thermal processes," *IEEE Trans. Control Syst. Technol.*, vol. 13, no. 5, pp. 686–700, Sep. 2005.

[2] H.-X. Li and C. Qi, *Spatio-Temporal Modeling of Nonlinear Distributed Parameter Systems: A Time/Space Separation Based Approach*, vol. 50. New York, NY, USA: Springer-Verlag, 2011.

[3] Z. Ghasemi, M. Karrari, and M. Shafiee, "Recursive identification of nonlinear two-dimensional (2-D) systems," in *Proc. 2nd Int. Conf. Control, Instrum. Automat.*, Dec. 2011, pp. 1124–1129.

[4] H. Shang, J. F. Forbes, and M. Guay, "Model predictive control for quasilinear hyperbolic distributed parameter systems," *Ind. Eng. Chem. Res.*, vol. 43, no. 9, pp. 2140–2149, 2004.

[5] C. Qi and H.-X. Li, "A time/space separation-based Hammerstein modeling approach for nonlinear distributed parameter processes," *Comput. Chem. Eng.*, vol. 33, no. 7, pp. 1247–1260, Jul. 2009.
[6] C. Qi, H.-X. Li, X. Zhang, X. Zhao, S. Li, and F. Gao, "Time/space-separation-based SVM modeling for nonlinear distributed parameter processes," *Ind. Eng. Chem. Res.*, vol. 50, no. 1, pp. 332–341, Jan. 2011.
[7] C. Anitescu, E. Atroshchenko, N. Alajlan, and T. Rabczuk, "Artificial neural network methods for the solution of second order boundary value problems," *Comput., Mater. Continua*, vol. 59, no. 1, pp. 345–359, 2019.
[8] M. Wang, N. Li, and S. Li, "Local modeling approach for spatially distributed system based on interval type-2 T-S fuzzy sets," *Ind. Eng. Chem. Res.*, vol. 49, no. 9, pp. 4352–4359, May 2010.
[9] X.-X. Zhang, Y. Jiang, H.-X. Li, and S.-Y. Li, "SVR learning-based spatiotemporal fuzzy logic controller for nonlinear spatially distributed dynamic systems," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 24, no. 10, pp. 1635–1647, Oct. 2013.
[10] M. Wang, C. Qi, H. Yan, and H. Shi, "Hybrid neural network predictor for distributed parameter system based on nonlinear dimension reduction," *Neurocomputing*, vol. 171, pp. 1591–1597, Jan. 2016.
[11] Z. Liu and H.-X. Li, "A spatiotemporal estimation method for temperature distribution in lithium-ion batteries," *IEEE Trans. Ind. Informat.*, vol. 10, no. 4, pp. 2300–2307, Nov. 2014.
[12] M. Wang, X. Yan, and H. Shi, "Spatiotemporal prediction for nonlinear parabolic distributed parameter system using an artificial neural network trained by group search optimization," *Neurocomputing*, vol. 113, pp. 234–240, Aug. 2013.
[13] C. A. J. Fletcher, *Computational Galerkin Methods*. New York, NY, USA: Springer-Verlag, 1984, pp. 72–85.
[14] G. D. Smith, and G. D. Smith, *Numerical Solution of Partial Differential Equations: Finite Difference Methods*. Oxford, U.K.: Oxford Univ. Press, 1985.
[15] L. Guo and S. A. Billings, "State-space reconstruction and spatio-temporal prediction of lattice dynamical systems," *IEEE Trans. Autom. Control*, vol. 52, no. 4, pp. 622–632, Apr. 2007.
[16] K. Hollig, *Finite Element Methods With B-Splines*. Philadelphia, PA, USA: SIAM, 2003.
[17] M. Jiang, X. Li, J. Wu, and G. Wang, "A precision on-line model for the prediction of thermal crown in hot rolling processes," *Int. J. Heat Mass Transf.*, vol. 78, pp. 967–973, Nov. 2014.
[18] M. Meyer and H. G. Matthies, "Efficient model reduction in nonlinear dynamics using the Karhunen–Loève expansion and dual-weighted-residual methods," *Comput. Mech.*, vol. 31, nos. 1–2, pp. 179–191, 2003.
[19] H. Park and D. Cho, "The use of the Karhunen–Loève decomposition for the modeling of distributed parameter systems," *Chem. Eng. Sci.*, vol. 51, no. 1, pp. 81–98, Jan. 1996.
[20] H. Kim and M. D. Shields, "Modeling strongly non-Gaussian non-stationary stochastic processes using the iterative translation approximation method and Karhunen–Loève expansion," *Comput. Struct.*, vol. 161, pp. 31–42, Dec. 2015.
[21] C. Turchetti, G. Biagetti, and F. Gianfelici, "Nonlinear system identification: An effective framework based on the Karhunen–Loève transform," *IEEE Trans. Signal Process.*, vol. 57, no. 2, pp. 536–550, Nov. 2009.
[22] T. G. Ritto, F. S. Buezas, and R. Sampaio, "Proper orthogonal decomposition for model reduction of a vibroimpact system," *J. Brazilian Soc. Mech. Sci. Eng.*, vol. 34, no. 3, pp. 330–340, Nov. 2012.
[23] H. X. Li and C. Qi, "Modeling of distributed parameter systems for applications—A synthesized review from time-space separation," *J. Process Control*, vol. 20, no. 8, pp. 891–901, 2010.
[24] G. Montaseri and M. J. Yazdanpanah, "Predictive control of uncertain nonlinear parabolic PDE systems using a Galerkin/neural-network-based model," *Commun. Nonlinear Sci. Numer. Simul.*, vol. 17, no. 1, pp. 388–404, Jan. 2012.
[25] C. Qi, H.-X. Li, S. Li, X. Zhao, and F. Gao, "A fuzzy-based spatio-temporal multi-modeling for nonlinear distributed parameter processes," *Appl. Soft Comput.*, vol. 25, pp. 309–321, Dec. 2014.
[26] Z. Liu and H.-X. Li, "Extreme learning machine based spatiotemporal modeling of lithium-ion battery thermal dynamics," *J. Power Sources*, vol. 277, pp. 228–238, Mar. 2015.
[27] X. Lu, W. Zou, and M. Huang, "A novel spatiotemporal LS-SVM method for complex distributed parameter systems with applications to curing thermal process," *IEEE Trans. Ind. Informat.*, vol. 12, no. 3, pp. 1156–1165, Jun. 2016.
[28] K. Xu, H. Yang, C. Zhu, and L. Hu, "Finite Gaussian mixture model based multimodeling for nonlinear distributed parameter systems," *IEEE Trans. Ind. Informat.*, vol. 16, no. 3, pp. 1754–1763, Mar. 2020.
[29] K.-K. Xu, H.-X. Li, and H.-D. Yang, "Dual least squares support vector machines based spatiotemporal modeling for nonlinear distributed thermal processes," *J. Process Control*, vol. 54, pp. 81–89, Jun. 2017.
[30] R. Moazenzadeh, B. Mohammadi, S. Shamshirband, and K.-W. Chau, "Coupling a firefly algorithm with support vector regression to predict evaporation in northern Iran," *Eng. Appl. Comput. Fluid Mech.*, vol. 12, no. 1, pp. 584–597, Jan. 2018.
[31] W.-C. Wang, K.-W. Chau, L. Qiu, and Y.-B. Chen, "Improving forecasting accuracy of medium and long-term runoff using artificial neural network based on EEMD decomposition," *Environ. Res.*, vol. 139, pp. 46–54, May 2015.
[32] Z. M. Yaseen, S. O. Sulaiman, R. C. Deo, and K.-W. Chau, "An enhanced extreme learning machine model for river flow forecasting: State-of-the-art, practical applications in water resource engineering area and future research direction," *J. Hydrol.*, vol. 569, pp. 387–408, Feb. 2019.
[33] C.-T. Cheng, J.-Y. Lin, and Y.-G. Sun, "Long-term prediction of discharges in Manwan Hydropower using adaptive-network-based fuzzy inference systems models," in *Proc. Int. Conf. Natural Comput.*, 2005, pp. 1152–1161.
[34] H. Sak, A. Senior, and F. Beaufays, "Long short-term memory based recurrent neural network architectures for large vocabulary speech recognition," *Comput. Sci.*, pp. 338–342, 2014.
[35] S. Li, H. Fang, and B. Shi, "Multi-step-ahead prediction with long short term memory networks and support vector regression," in *Proc. 37th Chin. Control Conf. (CCC)*, Jul. 2018, pp. 8104–8109.
[36] S. Xingjian, Z. Chen, and H. Wang, "Convolutional LSTM network: A machine learning approach for precipitation nowcasting," in *Proc. Adv. Neural Inf. Process. Syst.*, 2015, pp. 802–810.
[37] Z. Zhao, W. Chen, X. Wu, P. C. Y. Chen, and J. Liu, "LSTM network: A deep learning approach for short-term traffic forecast," *IET Intell. Transp. Syst.*, vol. 11, no. 2, pp. 68–75, Mar. 2017.
[38] X. Ma, Z. Tao, Y. Wang, H. Yu, and Y. Wang, "Long short-term memory neural network for traffic speed prediction using remote microwave sensor data," *Transp. Res. C, Emerg. Technol.*, vol. 54, pp. 187–197, May 2015.
[39] D. G. Zill, and W. S. Wright, *Differential Equations With Boundary-Value Problems*. Boston, MA, USA: Cengage Learning, 2012.
[40] C. Qi and H.-X. Li, "A Karhunen–Loève decomposition-based Wiener modeling approach for nonlinear distributed parameter processes," *Ind. Eng. Chem. Res.*, vol. 47, no. 12, pp. 4184–4192, 2008.
[41] H. Bourlard and Y. Kamp, "Auto-association by multilayer perceptrons and singular value decomposition," *Biological*, vol. 59, nos. 4–5, pp. 291–294, 1988.
[42] P. Chang and Y. Yao, "Time/space separation based KL-MKELM modeling for nonlinear distributed parameter processes," in *Proc. Chin. Autom. Congr. (CAC)*, 2017, pp. 5641–5646.
[43] B. Najafi, S. Faizollahzadeh Ardabili, S. Shamshirband, K.-W. Chau, and T. Rabczuk, "Application of ANNs, ANFIS and RSM to estimating and optimizing the parameters that affect the yield and cost of biodiesel production," *Eng. Appl. Comput. Fluid Mech.*, vol. 12, no. 1, pp. 611–624, Jan. 2018.
[44] Y.-R. Chien, J.-W. Chen, and S. S.-D. Xu, "A multilayer perceptron-based impulsive noise detector with application to power-line-based sensor networks," *IEEE Access*, vol. 6, pp. 21778–21787, 2018.
[45] F. Fotovatikhah, M. Herrera, S. Shamshirband, K.-W. Chau, S. Faizollahzadeh Ardabili, and M. J. Piran, "Survey of computational intelligence as basis to big flood management: Challenges, research directions and future work," *Eng. Appl. Comput. Fluid Mech.*, vol. 12, no. 1, pp. 411–437, Jan. 2018.
[46] S. Bueno and J. L. Salmeron, "Benchmarking main activation functions in fuzzy cognitive maps," *Expert Syst. Appl.*, vol. 36, no. 3, pp. 5221–5229, Apr. 2009.
[47] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, 1997.
[48] J. Guo, Z. Xie, Y. Qin, L. Jia, and Y. Wang, "Short-term abnormal passenger flow prediction based on the fusion of SVR and LSTM," *IEEE Access*, vol. 7, pp. 42946–42955, 2019.
[49] O. A. Abidogun, "Data mining, fraud detection and mobile telecommunications: Call pattern analysis with unsupervised neural networks," M.S. thesis, Univ. Western Cape, Cape Town, South Africa, 2005.

[50] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," 2014, *arXiv:1412.6980*. [Online]. Available: https://arxiv.org/abs/1412.6980

**YAJUN FAN** received the M.S. degree from the South China University of Technology, Guangzhou, China, in 2014. He is currently pursuing the Ph.D. degree with the School of Mechanical Science and Engineering, Huazhong University of Science and Technology, Wuhan, China. His research interests include data-based modeling, machine learning, and soft sensor applications.

**KANGKANG XU** received the B.E. and Ph.D. degrees in mechatronics engineering from Central South University, Changsha, China, in 2012 and 2017, respectively. He is currently a Lecturer with the School of Electromechanical Engineering, Guangdong University of Technology, Guangzhou, China. His research interests include distributed parameter systems and intelligent modeling.

**HUI WU** received the B.E. degree in electrical engineering and automation from the China University of Mining and Technology, Xuzhou, China, in 2018. He is currently pursuing the master's degree with the Huazhong University of Science and Technology, Wuhan, China. His research interests include machine learning and soft sensor.

**YING ZHENG** (Senior Member, IEEE) received the B.S. degree in industrial electric automation engineering, in 1997, and the M.S. and Ph.D. degrees in control theory and engineering from the Huazhong University of Science and Technology, Wuhan, China, in 2000 and 2003, respectively. She has been a Postdoctoral Fellow at the Chemical Engineering Department, National Tsing-Hua University, Hsinchu, Taiwan, from 2004 to 2005, and a Faculty Member of the Department of Control Science and Engineering, Huazhong University of Science and Technology (Associate Professor in 2005 and Professor in 2010). Her research interests include process control, data-driven method, fault diagnosis, and networked control systems.

**BO TAO** (Member, IEEE) received the B.S. and Ph.D. degrees in mechanical engineering from the Huazhong University of Science and Technology (HUST), in 1999 and 2007, respectively. From 2007 to 2009, he held a postdoctoral position at the Department of Electronics Science and Technology, HUST. After that, he has been an Associate Professor (2009) and also a Professor (2013) with the School of Mechanical Science and Engineering, HUST. He has published more than 40 papers in international journals. His research interests include intelligent manufacturing and robotics technologies, as well as RFID technologies and applications.