

Article

Not peer-reviewed version

Three-dimensional Fuzzy Modeling for Nonlinear Distributed Parameter Systems Using Simultaneous Perturbation Stochastic Approximation

[Xianxia Zhang](#)*, [Tangchen Wang](#), [Chong Cheng](#), [Shaopu Wang](#)

Posted Date: 5 August 2024

doi: 10.20944/preprints202408.0249.v1

Keywords: distributed parameter system; fuzzy model; simultaneous perturbation stochastic approximation; fuzzy modeling



Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Article

Three-dimensional Fuzzy Modeling for Nonlinear Distributed Parameter Systems Using Simultaneous Perturbation Stochastic Approximation

Xianxia Zhang ^{1,*} , Tangchen Wang ¹, Chong Cheng ² and Shaopu Wang ¹

¹ School of Mechatronics and Automation, Shanghai University, Shanghai 200444, China.

² China Mobile (Suzhou) Software Technology Co., Ltd.

* Correspondence: Xianxia Zhang, E-mail: xianxia_zh@t.shu.edu.cn

Abstract: Many systems in the manufacturing industry have spatial distribution characteristics, which correlate with both time and space. Such systems are known as distributed parameter systems (DPS). Due to the spatiotemporal coupling characteristics, the modeling of such systems is quite complex. The paper presents a new approach for three-dimensional fuzzy modeling using Simultaneous Perturbation Stochastic Approximation (SPSA) for nonlinear DPS. The Affinity Propagation clustering approach is utilized to dynamically determine the optimal number of fuzzy rules and construct a collection of preceding components for three-dimensional fuzzy models. Fourier space base functions are used in the resulting components of three-dimensional fuzzy models, and their parameters are learned by the SPSA algorithm. The proposed three-dimensional fuzzy modeling technique was utilized on a conventional DPS within the Semiconductor manufacturing industry, with the simulation experiments confirming its efficacy.

Keywords: distributed parameter system; fuzzy model; simultaneous perturbation stochastic approximation; fuzzy modeling

1. Introduction

The spatial distribution characteristics are widely present in many manufacturing industries such as petroleum, chemical, steelmaking, and rolling. The dynamic characteristics of the system change with time and space, and this type of system is called a distributed parameter system (DPS)[1]. The dynamic characteristics of DPS are usually described by partial differential equations and integral equations. Traditionally, the spatial distribution characteristics of such systems are often ignored, and the inherent DPS is approximated as lumped parameter systems (LPS), and then are modeled and controlled using the mature lumped parameter methods. With the development of the times, people's requirements for product quality are gradually increasing. With the advancement of sensors, actuators, and computing technology, the modeling, control, and optimization techniques of DPS have received increasing attention in the fields of science and engineering. For practical use in areas like optimization, control design, analyzing, and numerical simulations of DPS[2], it is necessary to use an appropriate mathematical model to describe its dynamic characteristics. Therefore, the modeling of DPS has become a very important research field.

The modeling methods for DPS can usually be divided into: mechanism modeling with known partial differential equations and grey box/black box modeling with unknown partial differential equations. Mechanism modeling is the process of obtaining partial differential equations for DPS from the point of view of physical or chemical reaction knowledge of the system, such as thermal energy balance, material balance, or momentum balance relationships. Considering the complexity of calculations and the restricted quantity of sensors and actuators, modeling such systems further requires the use of spectral methods[1], finite element[3], finite difference techniques, etc., to simplify the essentially infinite dimensional system into a finite dimensional system[4]. Nonetheless, in real-world scenarios, it will be challenging to obtain accurate mathematical descriptions of the system through mechanism modeling method due to incomplete knowledge of physical and chemical processes. The second modeling method is based on data, including gray box modeling and black box modeling[5]. The

model expressed by partial differential equations of the system is established through prior knowledge, and the unknown parameters of the model are obtained by identification from the input/output data of the system. This method is called grey box modeling. Directly identifying the structure and parameters of a system from input/output data is called black box modeling. In practical applications, black box modeling has a wider range of applications. How to extract the relationship between system inputs/outputs with spatial distribution characteristics from data is currently a challenge in modeling research.

In recent decades, researchers have conducted continuous research on black box modeling of DPS. Most methods decompose variables into a series of spatial functions and temporal coefficients[6], and then reconstruct the system. Traditional methods usually pre-set spatial functions, such as Jacobian polynomials, spline functions, etc., and prior knowledge also dictates the quantity of spatial functions. The time coefficients are estimated using traditional modeling methods for LPS. In the past 20 years, spatial functions have been estimated using the Karhunen-Loève (KL) decomposition method[7,8]. The time coefficient is still estimated using traditional modeling methods for LPS. In practical applications, essentially infinite dimensional systems require model simplification to obtain lower order ordinary differential equations. However, model simplification can lead to unmodeled dynamics and unknown nonlinearity[9].

Intelligent methods like neural networks and fuzzy systems have been incorporated into DPS modeling as well. References [8,10,11] combined neural networks with space-time separation methods and achieved good experimental results on parabolic DPS modeling problems. Wang et al. [10] used a nonlinear PCA method to obtain space base functions. After space-time separation and dimension reduction, a neural network was used to identify the time coefficients. Zhang et al.[11] used the PCA algorithm to obtain space base functions. After space-time separation and dimension reduction, the time model was decomposed into linear and nonlinear parts. The ARX model was used to model the linear part, and the RBF neural network was employed to model the nonlinear part. Fan et al.[8] used the KL decomposition method to obtain space base functions. After space-time separation and dimension reduction, a mixed time series model of long short-term memory network and multi-layer perceptron was used to model the time coefficients. However, the aforementioned modeling methods still require dimension reduction.

Three-dimensional fuzzy modeling[9,12] is a new type of DPS modeling method that has developed in the past five years. In three-dimensional fuzzy model, the functions of space-time separation and space-time reconstruction are naturally integrated into three-dimensional fuzzy rules. In each three-dimensional fuzzy rule, space-time separation is naturally achieved, where the preceding component is computed as the time factor and the resulting component is regarded as the space base function. In the combination operation of multiple excited three-dimensional fuzzy rules, space-time reconstruction is naturally achieved. Compared with traditional modeling methods of distributed parameter systems, three-dimensional fuzzy modeling method has two advantages: it does not rely on model dimension reduction and has language interpretability.

At present, the three-dimensional fuzzy model has been applied in the modeling of DPS such as rapid thermal processing and chemical reactors. Zhang et al.[9] introduced a three-dimensional fuzzy modeling approach utilizing clustering and support vector regression, where nearest neighbor clustering (NNC) method and similarity analysis technique were employed for mining and condensing the initial preceding components in the three-dimensional fuzzy rules, and support vector regression machines are employed to acquire knowledge about the space base functions of the resulting components in the three-dimensional fuzzy rules. However, the clustering results of the NNC algorithm are influenced by the initial representative point selection, and the structural simplification process is very cumbersome. Zhang et al.[12] proposed a method for three-dimensional fuzzy modeling that utilizes KL decomposition to derive space base functions for the resulting sets of the three-dimensional fuzzy rule. The resulting sets consist of a restricted amount of space base functions with higher weights. This

method pertains to the scope of dimension reduction and results in dynamic characteristics that have not been modeled.

Based on the above discussion, this paper introduces an innovative three-dimensional fuzzy modeling method for DPS. Utilizing the Affinity Propagation (AP) clustering algorithm, the preceding components of the fuzzy rule are derived within the three-dimensional fuzzy modeling framework. By considering every data point as a potential class representative point, the AP algorithm circumvents the issue of clustering outcomes being constrained by the choice of initial class representative points. Unlike existing methods such as KL decomposition or support vector regression to learn space base functions, this paper sets the Fourier space base function as the resulting component of three-dimensional fuzzy rules. The construction of Fourier space base functions is simple and can better approximate the actual spatial distribution characteristics. Furthermore, the parameters of the Fourier space base function are fine-tuned using the simultaneous perturbation stochastic approximation (SPSA) algorithm.

The key innovations of this study are outlined below:

- 1) The AP clustering approach is utilized to dynamically determine the most suitable quantity of fuzzy rules and construct preceding sets for the three-dimensional fuzzy model.
- 2) The Fourier space base functions are initially introduced as the resulting sets of the three-dimensional fuzzy model.
- 3) The SPSA technique is employed to dynamically update the coefficients of Fourier space base functions for the three-dimensional fuzzy model.

This paper is outlined as follows. The second section offers problem description. The third section offers an in-depth explanation of the three-dimensional fuzzy modeling method. The fourth section applies the proposed modeling method to a rapid thermal chemical vapor deposition system. In the final section, the conclusion is given.

2. Problem Description

2.1. Mathematical Description of Distributed Parameter Systems and Its Characteristics

Numerous processes, like the industrial chemical reaction[1], semiconductor production[14], transmission in cure or reflow processes[15] and Lithium-ion Battery Thermal Process[16,17], are spatially distributed in practice. They can be depicted using the subsequent mathematical formula:

$$\begin{aligned} \frac{\partial^m w(f,t)}{\partial t^m} &= Yw(f,t) + \hbar(w) + \lambda b(f)Q(t) \\ l_a \leq f \leq l_b, t &\geq 0 \\ m &= 1 \text{ or } 2 \end{aligned} \quad (1)$$

where, $w(f,t)$ represents the output; $\hbar(w)$ is a nonlinear function; Y is linear space differential operator, which includes first or second order spatial derivatives ($\partial/\partial f, \partial^2/\partial f^2$), and is dense in Hilbert space; λ is a constant vector; $Q = [q_1 \ q_2 \ \cdots \ q_L]^T \in R^{1 \times L}$ is manipulating input vector (i.e. multiple control sources); $b(f) = [b_1(f) \ b_2(f) \ \cdots \ b_L(f)]^T \in R^L$ is a known smooth function about f , which describes the distribution of Q in the spatial domain $\bar{F} \subset [l_a, l_b]$.

Usually, boundary and initial conditions also restrict the system depicted in the Equation (1). As an illustration, at $t=1$, the boundary and initial conditions can be depicted in the following equations:

$$\begin{aligned} \frac{\partial w(l_a,t)}{\partial f} &= g_a(w(l_a,t), Q(t)) \\ \frac{\partial w(l_b,t)}{\partial f} &= g_b(w(l_b,t), Q(t)) \end{aligned} \quad (2)$$

$$w(f,0) = w_0(f) \quad (3)$$

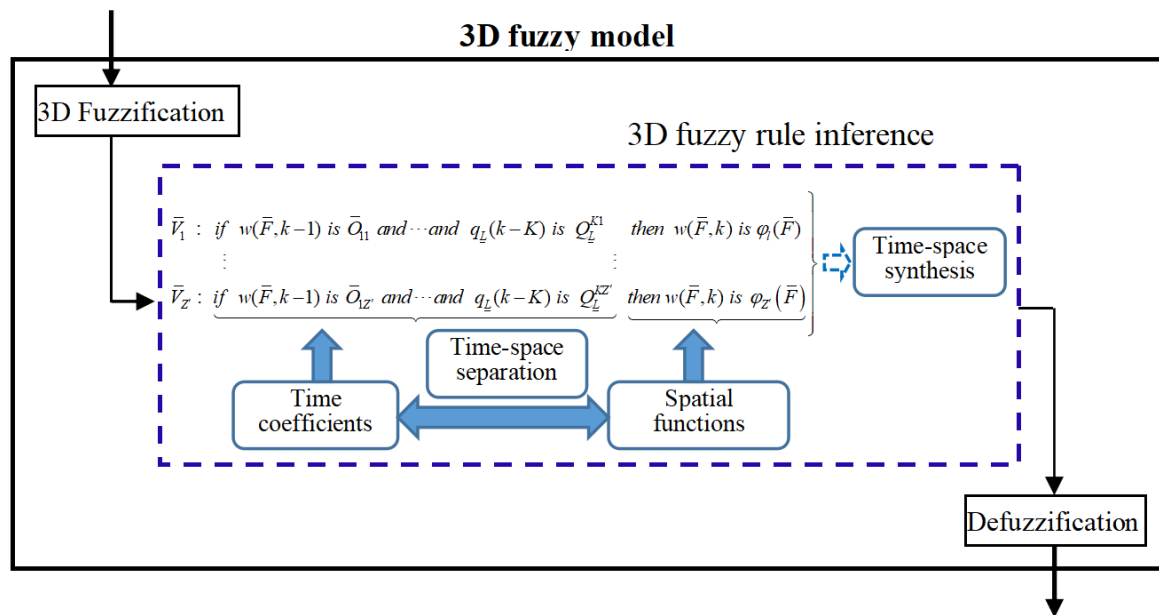


Figure 1. Framework of three-dimensional fuzzy model (three-dimensional abbreviated as 3D)

where, $g_a(w(l_a, t), Q(t))$ is dependent on $w(l_a, t)$ and $Q(t)$, and $g_b(w(l_b, t), Q(t))$ is dependent on $w(l_b, t)$ and $Q(t)$; $w_0(f)$ is a known function.

This type of system has spatiotemporal coupling characteristics and is an infinite dimensional system. $w(f, t)$ can be spatiotemporally separated into an infinite weighted sum of space base functions and time coefficients, as shown in Equation (4).

$$w(f, t) = \sum_{i=1}^{\infty} w_i \varphi_i(f) \quad (4)$$

Practically, it is necessary to utilize a finite quantity of sensors positioned at $f_1, f_2, \dots, f_p$; therefore, $\bar{F} = [f_1, f_2, \dots, f_p]$. $w(\bar{F}, t) = [w(f_1, t), w(f_2, t), \dots, w(f_p, t)]$ represents the output in space. The data-driven modeling challenge for the DPS, as outlined in Equations (1) - (3), entails the task of determining a spatiotemporal coupled model using the input $\{Q(k)\}_{k=1}^K$ and the output $\{w(f_i, k)\}_{i=1, k=1}^{P, K}$, where K is the time length. Essentially, systems with an infinite number of dimensions will be converted into the systems with a finite number of dimensions. When the mechanism model described in Equations (1) - (3) is unknown, utilizing the KL decomposition technique, the system's space base function can be derived from the gathered data set $\bar{S} = \{Q(k) w(f_i, k)\}_{i=1, k=1}^{P, K}$. Then, $w(f, t)$ can be approximately depicted as Equation (5).

$$w_n(f, t) = \sum_{i=1}^n w_i \varphi_i(f) \quad (5)$$

The time coefficient in Equation (5) can be identified using traditional methods.

However, the modeling approach of DPS using KL decomposition unavoidably results in neglected dynamic features and uncertainties in the identified model due to the decrease of dimensions.

2.2. Three-dimensional Fuzzy Modeling

The three-dimensional fuzzy system[18] was first established in 2007 and has since been used in the domain of DPS control[19]. Recently, there has been effective development of the three-dimensional fuzzy system for modeling DPS. The relevant theory of three-dimensional fuzzy system and black box modeling theory were combined to form a three-dimensional fuzzy modeling framework as shown in Figure 1[9]. Unlike traditional modeling methods, the three-dimensional fuzzy modeling method can model a DPS without reduction dimension.

Assume that K and J represent the order of the input variable $Q(k)$ and output variable $w(\bar{F}, k)$, respectively. Input variables of three-dimensional fuzzy system are $\{Q(k-1), Q(k-2), \dots, Q(k-K)\}$ and $\{w(\bar{F}, k), w(\bar{F}, k-1), \dots, w(\bar{F}, k-J)\}$. The three-dimensional fuzzy rule is given as shown in Equation (6).

$$\begin{aligned} \bar{V}_l : \\ \text{if } w(\bar{F}, k-1) \text{ is } \bar{O}_{1l} \text{ and } \dots \text{ and } w(\bar{F}, k-J) \text{ is } \bar{O}_{Jl} \\ \text{and } q_1(k-1) \text{ is } Q_1^{1l} \text{ and } \dots \text{ and } q_1(k-K) \text{ is } Q_1^{Kl} \\ \vdots \\ \text{and } q_{\bar{L}}(k-1) \text{ is } Q_{\bar{L}}^{1l} \text{ and } \dots \text{ and } q_{\bar{L}}(k-K) \text{ is } Q_{\bar{L}}^{Kl} \\ \text{then } w(\bar{F}, k) \text{ is } \varphi_l(\bar{F}) \end{aligned} \quad (6)$$

where $\varphi_l(\bar{F})$ is a space base function, $\bar{O}_{sl}$ is a three-dimensional fuzzy set, Q_g^{hl} is a conventional fuzzy set, $s = 1, 2, \dots, J$, $h = 1, 2, \dots, K$, $g = 1, 2, \dots, \bar{L}$, $l = 1, 2, \dots, \bar{Z}$, $\bar{Z}$ indicates the count of fuzzy rules.

For any three-dimensional fuzzy rule $\bar{V}_l (l = 1, 2, \dots, \bar{Z})$, its preceding set 'if $w(\bar{F}, k-1)$ is $\bar{O}_{1l}$ and $\dots$ and $q_{\bar{L}}(k-K)$ is $Q_{\bar{L}}^{Kl}$ ' corresponds to the time coefficient w_l . The resulting set 'then $w(\bar{F}, t)$ is $\varphi_l(\bar{F})$ ' corresponds to the space base function $\varphi_l(\bar{F})$, and space-time separation is ingeniously achieved through three-dimensional fuzzy rules. If Z' three-dimensional fuzzy rules are excited, the time coefficients and space base functions of Z' three-dimensional fuzzy rules will naturally realize space-time reconstruction. The final result is $\sum_{l=1}^{Z'} w_l \varphi_l(\bar{F})$.

The three-dimensional fuzzy modeling framework can be divided into three parts.

- (1) Acquire the preceding set of fuzzy rules in order to derive temporal coefficients.
- (2) Acquire knowledge about the space base functions of the resulting set.
- (3) Obtain predicted output through space-time reconstruction.

The study will explore an innovative approach to three-dimensional fuzzy modeling within this specific framework.

3. SPSA Learning Based Three-Dimensional Fuzzy Modeling

3.1. Modeling Methodology

As for an unknown DPS system, space base functions are frequently calculated by KL decomposition technique. However, implementing the KL decomposition technique is bound to result in the simplification of models. The three-dimensional fuzzy model has the innate ability to amalgamate space-time separation and space-time reconstruction into three-dimensional fuzzy rules and thereafter avoids the model reduction.

In this study, a novel SPSA learning based three-dimensional fuzzy model methodology is proposed, the scheme of which is shown in Figure 2. The modeling methodology is divided into two parts, namely, learning for preceding components of three-dimensional fuzzy rules and learning for resulting components of three-dimensional fuzzy rules. In the first part, by using AP clustering, preceding components of three-dimensional fuzzy rules are learned. In the second part, the Fourier function is initially selected as the space base function as the resulting components of the three-dimensional fuzzy rule. Then the SPSA method is used to determine the coefficients of the Fourier space base function adaptively. Three distinct characteristics are evident in the SPSA learning based three-dimensional fuzzy modeling method as blow.

1) AP clustering is employed to manage a collection of spatiotemporal data and extract the preceding components of three-dimensional fuzzy rules in an adaptable manner, without the need to define the number of clusters.

2) The Fourier space base functions are investigated in three-dimensional fuzzy modeling.

3) SPSA algorithm is used to acquire optimal parameters of Fourier space base functions.

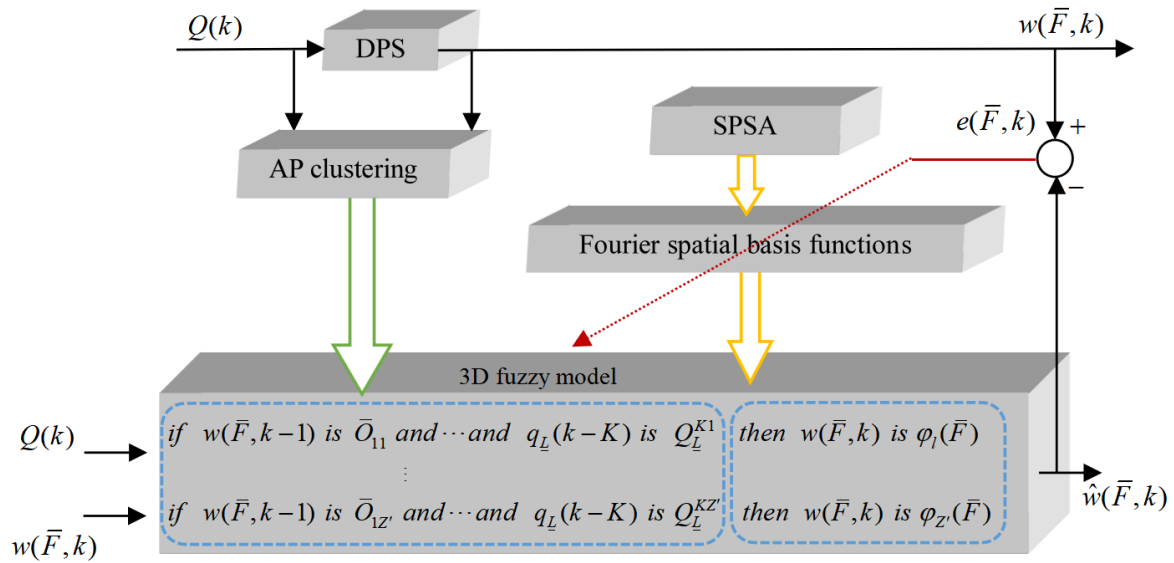


Figure 2. Scheme of SPSA learning based three-dimensional fuzzy modeling

3.2. AP Clustering Learning for Preceding Components of Three-dimensional Fuzzy Rule

Frey and Dueck introduced Affinity Propagation (AP) Clustering[13], a novel clustering technique, in 2007. It can handle with large-scale data sets in a short period of time without initializations of cluster number and cluster core. The process of AP clustering involves initially selecting specific samples known as exemplars, followed by linking each sample on the left to its closest examples. Two kinds of message, namely responsibility and availability, are passed in AP clustering. Let $V = [r(i, \nu)]_{n \times n}$ be responsibility matrix, $A = [a(i, \nu)]_{n \times n}$ be the availability matrix, and $S = [s(i, \nu)]_{n \times n}$ be the similarity matrix. The responsibility $r(i, \nu)$ reflects how strongly the sample i want to choose the candidate exemplar ν as its exemplar. The presence of $a(i, \nu)$ mirrors the collective data demonstrating the appropriateness of sample i in selecting candidate ν as the exemplar. The similarity $s(i, \nu)$ represents the degree of similarity between the sample i and its nearest exemplar ν . Then, the message between V and A is propagated so that the sample similarity S in the cluster is the largest and the sample similarity between the clusters is the smallest. The detailed algorithm description is presented in Appendix A and the algorithm procedure is shown in Figure A.1 of Appendix A.

For a spatiotemporal data set $\bar{S}$ from DPS, AP clustering algorithm is employed to cluster the dataset $\bar{S}$ to Z' clusters. Each cluster corresponds to the preceding components of a three-dimensional fuzzy rule. Hence, the preceding components of Z' three-dimensional fuzzy rules are achieved. Assumed that $o^1, \dots, o^{Z'}$ with $o^i = (o_{\chi}^i, o_c^i)$ are Z' cluster cores. These cores will guide to construct the preceding components of Z' three-dimensional fuzzy rules. Utilizing a Gaussian type membership function, the cluster's core o^i aligns with the Gaussian type MFs' center in the preceding components, while o_{χ}^i and o_c^i represent the centers of three-dimensional MFs and conventional MFs, respectively.

The breadth of Gaussian type membership function can be computed according to the interval range of the input variables from the input data set. Regarding the spatial output variable $w(\bar{F}, k)$, the breadth of the Gaussian type three-dimensional MF at an identical spatial point is calculated as in (7).

$$\sigma_w(f_j) \approx \max_{\eta \in [0.01, 2]} \left(\frac{w^{\max}(f_j) - w^{\min}(f_j)}{\eta} \right) \quad (7)$$

where $w^{\max}(f_j)$ and $w^{\min}(f_j)$ represent the highest and lowest $w(\bar{F}, k)$ values at the sensor's location $f = f_j$; η is a variable set by the users.

Regarding the variable $Q(k)$, the breadth of the conventional Gaussian MF is outlined below.

$$\sigma_Q = \begin{bmatrix} \sigma_{\mu_1} & \cdots & \sigma_{\mu_L} \end{bmatrix} \quad (8)$$

where $\sigma_{\mu_i} = \frac{q_i^{\max} - q_i^{\min}}{\eta_i}$, $\eta_i \in [0.01, 2]$; $q_i^{\max}$ and $q_i^{\min}$ represent the highest and lowest values of the i th input variable; η_i is a variable set by the users; $i = 1, 2, \dots, L$.

The coefficients η in Equation (7) and η_i in Equation (8) can be modified based on the present situation.

3.3. SPSA Learning for Resulting Components of Three-dimensional Fuzzy Rule

3.3.1. Fourier Space Base Function

Space base function is crucial to the performance of the modeling[6]. In three-dimensional fuzzy modeling, KL decomposition technique and SVR have been used to estimate space base functions. Traditional DPS modeling approaches allow for the selection of fixed form functions, including Spline functions, Legendre polynomials, and Jacobi polynomials[6], as space base functions, in addition to KL decomposition estimation. In this study, Fourier space base functions are taken as the resulting components of three-dimensional fuzzy rules. Spatial dynamics of a DPS can be represented by multiple Sine waves and Cosine waves with different amplitude and frequency. A Fourier space base function is expressed as in (9).

$$\varphi_l(\bar{F}) = a_l \sin(b_l \bar{F}) + c_l \cos(b_l \bar{F}) + d_l \quad (l = 1, 2, \dots, Z') \quad (9)$$

Therefore, the variable $w(\bar{F}, k)$ can be expressed as a linear sum of Fourier space base functions, with each space base function multiplied by its corresponding time coefficient, which is shown as in (10).

$$w(\bar{F}, k) = \sum_{l=1}^{Z'} \psi^l(k) \varphi_l(\bar{F}) \quad (10)$$

$$\begin{aligned} \psi^l(k) &= \sum_{h=1}^P a_h^l \prod_{s=1}^J E_A \times \prod_{i=1}^L \prod_{g=1}^K E_B \\ E_A &= \exp\left(-\left(\left(w(f_h, k - s) - c_{sh}^l\right) / \sigma_{sh}^l\right)^2\right) \\ E_B &= \exp\left(-\left(\left(q_i(k - g) - d_i^{gl}\right) / \delta_i^{gl}\right)^2\right) \end{aligned} \quad (11)$$

where $\psi^l(t)$ is called as fuzzy basis function[19] in the context of three-dimensional fuzzy system, also called time coefficient in a three-dimensional fuzzy model; c_{sh}^l and σ_{sh}^l are denoted as central point and breadth of the Gaussian type three-dimensional fuzzy set $\bar{O}_{sl}$ at the h th sensor position; d_i^{gl} and δ_i^{gl} represent the central point and breadth of the conventional Gaussian fuzzy set Q_i^{gl} ; a_h^l represents the weight assigned to the h th sensor location.

The SPSA method will be used to learn the coefficients of the Fourier space base function.

3.3.2. Parameter Learning Using SPSA Algorithm

The SPSA [20] is a highly effective technique for optimizing complex nonlinear problems. It possesses a robust theoretical basis and is frequently more effective and simpler to execute compared to widely recognized techniques like simulated annealing or genetic algorithms. In contrast with gradient descent method, SPSA does not necessitate the direct measurement of the gradient of the target function. Instead, it just requires two measurements of the target function at every iteration to approximate the gradient, regardless of the complexity of the optimization problem. The two measurements are

obtained by randomly and simultaneously altering all the variables in the problem in a correct manner, which is known as “simultaneous perturbation”.

In this study, in each iteration the approximation gradients of all parameters of Fourier space base functions are computed by two randomly disturbed measurements of target functions. Let $\hat{\theta}_k = [a_{1,k}, b_{1,k}, c_{1,k}, d_{1,k}, a_{2,k}, b_{2,k}, c_{2,k}, d_{2,k}, \dots, a_{Z',k}, b_{Z',k}, c_{Z',k}, d_{Z',k}] \in R^{4Z'}$ be all the parameters at the k th iteration, where $a_{i,k}, b_{i,k}, c_{i,k}$ and $d_{i,k}$ are parameters of i th space base function at the k th iteration. The target function is described as the L2 norm of the difference between the spatial output $w(\bar{F}, t)$ and the predicted spatial output $\hat{w}(\bar{F}, t)$ as follows.

$$J(\hat{\theta}) = \frac{1}{2} \sum_{i=1}^n (w_i(\bar{F}, t) - \hat{w}_i(\bar{F}, t))^2 \quad (12)$$

The algorithm of SPSA is described as follows.

Step 1: Initialization and selection of coefficients

Initialize the timer number as $k = 0$. Select an initial estimate $\hat{\theta}_0$ and choose non-negative constants b, d, A, α , and β for SPSA algorithm. Then, calculate sequences b_k and d_k as in (13) and (14).

$$b_k = b / (A + k + 1)^\alpha \quad (13)$$

$$d_k = d / (k + 1)^\beta \quad (14)$$

Step 2: Simultaneous perturbation vector generation

Create a random perturbation vector Δ_k in a $4Z'$ -dimensional space, where each component of Δ_k is individually produced from a zero-mean probability distribution that satisfies the specified constraints in [20]. Here, we use Bernoulli ± 1 distribution, which means that there is an equal probability of 0.5 for each event of ± 1 .

Step 3: Evaluation of the cost function

Calculate two values of the cost function $J(\hat{\theta})$ by perturbing the current $\hat{\theta}_k$ simultaneously: $J(\hat{\theta}_k + d_k \Delta_k)$ and $J(\hat{\theta}_k - d_k \Delta_k)$, where d_k and Δ_k are obtained from Step 1 and Step 2, respectively.

Step 4: Approximation of the gradient

Compute the simultaneous perturbation estimate for the gradient $\mathbf{g}(\hat{\theta}_k)$:

$$\hat{\mathbf{g}}_k(\hat{\theta}_k) = \frac{J(\hat{\theta}_k + d_k \Delta_k) - J(\hat{\theta}_k - d_k \Delta_k)}{2d_k} \begin{bmatrix} \Delta_{k1}^{-1} \\ \Delta_{k2}^{-1} \\ \vdots \\ \Delta_{kp}^{-1} \end{bmatrix} \quad (15)$$

where Δ_{ki} represents the i th element of the Δ_k vector.

Step 5: Revision of the θ estimation

$$\hat{\theta}_{k+1} = \hat{\theta}_k - b_k \hat{\mathbf{g}}_k(\hat{\theta}_k) \quad (16)$$

Using (16) to modify $\hat{\theta}_k$ to a fresh value $\hat{\theta}_{k+1}$.

Step 6: Iterating or terminating

Go back to Step 2 and substitute k with $k + 1$. Terminate the algorithm if there is minimal variation in several consecutive iterations or if the maximum number of iterations has been reached.

3.3.3. Three-dimensional Fuzzy Modeling Flowchart

The proposed three-dimensional fuzzy modeling is designed in seven stages, as illustrated in Figure 3.

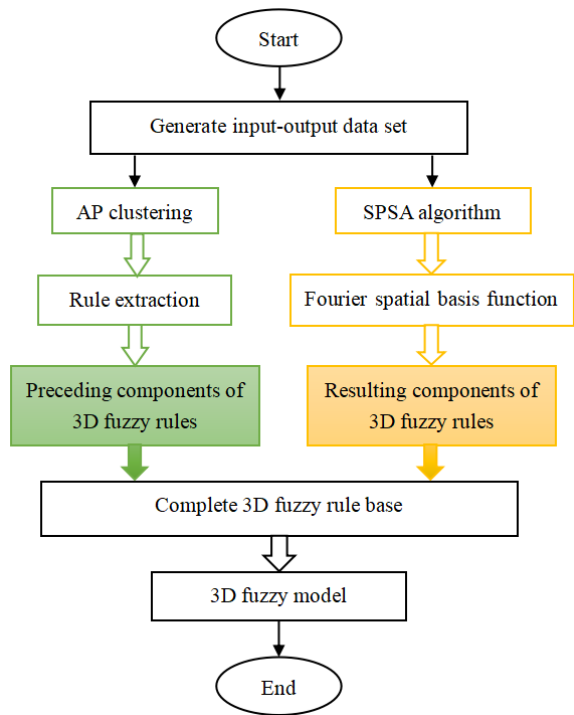


Figure 3. Flowchart of SPSA learning based three-dimensional fuzzy modeling (three-dimensional abbreviated as 3D)

Stage 1: Set up the elements of a three-dimensional fuzzy model, consisting of the structure of three-dimensional fuzzy rules, the method of three-dimensional fuzzification, the method of defuzzification, various types of fuzzy operators, and other related components.

Stage 2: Assign the Fourier space base functions, as depicted in Equation (9), to the resulting components of the three-dimensional fuzzy rules.

Stage 3: Create a collection of spatial and temporal input-output data from a DPS.

Stage 4: Cluster the spatiotemporal data using AP clustering algorithm.

Stage 5: Create the preceding components of three-dimensional fuzzy rules based on the clusters identified in Stage 4.

Stage 6: Utilizing the SPSA technique to determine the parameters of the Fourier space base functions in resulting components of three-dimensional fuzzy rules.

Stage 7: Combine the preceding components from Stage 5 with the resulting components of three-dimensional fuzzy rules to form the entire three-dimensional fuzzy rule foundation.

A three-dimensional fuzzy model is obtained after the completion of a seven-stage design process.

4. Application to RTCVD System

4.1. RTCVD System

Figure 4 displays the schematic of the rapid thermal chemical vapor deposition (RTCVD) system.

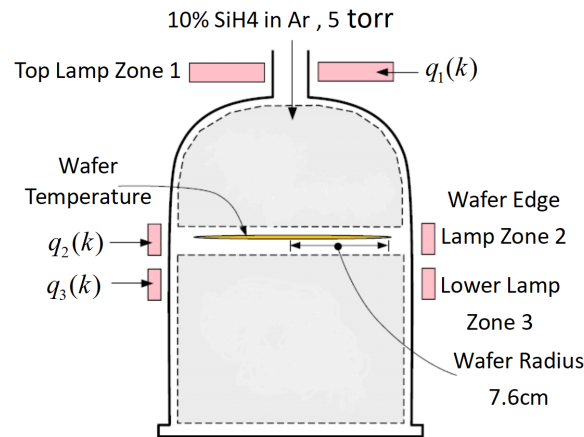


Figure 4. Sketch for RTCVD system

A 6-inch silicon wafer is positioned on a revolving support inside the system's chamber and subjected to heat from a heating system. The heating system consists of three Lamp banks. Lamp bank 1 uniformly warms the whole surface area of the wafer, Lamp bank 2 specifically heats the periphery of the wafer, and Lamp bank 3 equally heats the wafer as a whole. The Figure 5 displays the incident radiation flux emitted by the heating lights. A 10% concentration of silane gas (SiH_4) is fed into the reactor. SiH_4 undergoes decomposition to produce silicon (Si) and hydrogen gas (H_2). A layer of polysilicon, with a thickness of $0.5\mu\text{m}$, is applied onto the wafer by depositing it at temperatures of about 800K or more, which takes about 1 minute. During wafer processing, the support is turned to ensure that the temperature is evenly distributed in the azimuthal direction. Due to the thinness of the silicon wafer, all changes in temperature along the azimuth direction are disregarded. Hence, it is necessary to maintain consistent temperature distribution over the wafer radius in order to achieve a homogeneous deposition of polysilicon on the wafer. This may be achieved by regulating the power supplied to the three zones of lighting groups. A one-dimensional space model of the thermal dynamics may be expressed as the following PDE [21,22]:

$$\begin{aligned} \partial T'_f / \partial k' = & \kappa_0 \left[(1/f') \partial T'_f / \partial f' + \partial^2 T'_f / \partial f'^2 \right] \\ & + \sigma_0 (1 - T'^4_f) + \omega_f \varrho_1(f') q_1 \\ & + \omega_f \varrho_2(f') q_2 + \omega_f \varrho_3(f') q_3 \end{aligned} \quad (17)$$

Constrained by the specified border conditions.

$$\begin{cases} \partial T'_f / \partial f' = \sigma_{ed} (1 - T'^4_f) + \varrho_{ed} q_2 & \text{when } f' = 1 \\ \partial T'_f / \partial f' = 0 & \text{when } f' = 0 \end{cases} \quad (18)$$

where the non-dimensional wafer temperature, denoted by $T'_f = T_f / T_{amb}$, represents the ratio of the actual wafer temperature (T_f) to the ambient temperature (T_{amb}), which is 300K. The non-dimensional time, denoted by $k' = k / \tau$, represents the ratio of the actual time (k) to τ , which is 2.9s. The non-dimensional radius position, denoted by $f' = f / V_f$, represents the ratio of the actual radius position (f) to the wafer radius (V_f), which is 7.6cm. The variables q_1 , q_2 and q_3 represent the percentage of the light source power. The variables $\varrho_1(f')$, $\varrho_2(f')$, $\varrho_3(f')$ represent the wafer incident radiation flux from the three-zone warming lights, respectively. The parameters in (17) through (18) are listed as follows:

$$\begin{aligned} \kappa_0 &= 0.0021, \sigma_0 = 0.0012, \sigma_{ed} = 0.0037 \\ \varrho_{ed} &= 4.022, \omega_f = 0.0256 \end{aligned}$$

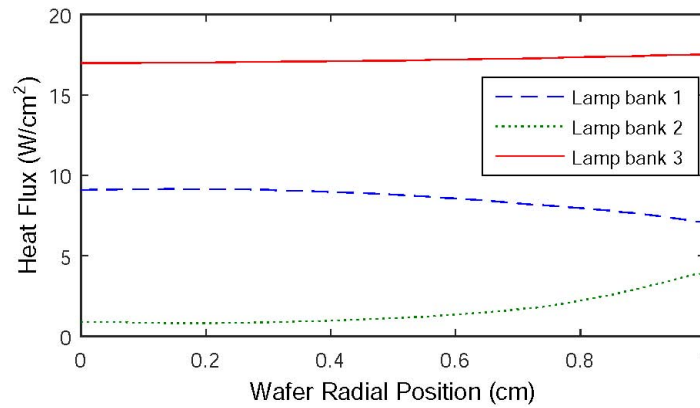


Figure 5. Radiation flux distribution of Lamp banks 1, 2, and 3

The RTCVD system is a complex system that exhibits both spatial and temporal characteristics and has an unlimited number of dimensions. For practical purposes, a certain number of sensors must be used. Assume that sensors positioned at $f_1, f_2, \dots, f_P$ are employed to gauge the system's output. Consider $\bar{F}$ as the spatial domain, where $\bar{F}$ is a vector containing the elements $f_1, f_2, \dots, f_P$. Let $w(\bar{F}, k)$ be a vector denoted by $w(f_1, k), w(f_2, k), \dots, w(f_P, k)$, where $w(f, k) = T_f(f, k)$. The temporal input $Q(k)$ is defined as $[q_1(k) \ q_2(k) \ q_3(k)]$, where $k \in \mathbb{R}$ is the time variable. The objective of the modeling problem is to determine a spatiotemporal model based on the input $\{Q(k)\}_{k=1}^K$ and the output $\{w(f_i, k)\}_{i=1, k=1}^{P, K}$, where K represents the duration of time.

In practical operation, 11 sensors are placed evenly along the radial direction. To simulate the impact of noise, 11 independent sets of Gaussian white noise signals were added to the data collected from 11 sensors. The average value is zero and the standard deviation is $\sigma(f) = A_d(f)n_d$, $A_d(f) = (\max(w(f, t)) - \min(w(f, t)))/3$, $n_d = 0.2\%$.

4.2. SPSA Learning Based Three-dimensional Fuzzy Modeling

In order to obtain sufficient dynamic information of the system, interference signals with amplitudes not exceeding 10% are added to the manipulated input variables q_1, q_2 , and q_3 , respectively. Thus, the three manipulated input variables affected by interference signals, namely the excitation signals, can be given by the following Equation (19).

$$\begin{aligned} q_1(t) &= q_{1a} + \Delta_{1a} * q_{1a} * \text{rand}() * S_r \\ q_2(t) &= q_{2a} + \Delta_{2a} * q_{2a} * \text{rand}() * S_r \\ q_3(t) &= q_{3a} + \Delta_{3a} * q_{3a} * \text{rand}() * S_r \\ S_r &= \text{sign}(\text{rand}() - 0.5) \end{aligned} \quad (19)$$

where q_{1a}, q_{2a} and q_{3a} are steady-state inputs at a furnace temperature of 1000K. Δ_{1a}, Δ_{2a} and Δ_{3a} are the disturbance amplitude of q_{1a}, q_{2a} and q_{3a} . In this study, q_{1a}, q_{2a} and q_{3a} are 0.2028, 0.1008 and 0.2245, respectively, Δ_{1a}, Δ_{2a} and Δ_{3a} are set to 10%.

A sample interval of 0.5 seconds was established in this investigation. The entire simulation time period was 7000 seconds. Therefore, 14000 samples were generated, in which 600 samples were randomly selected for training experiments and then 300 samples were randomly selected for testing experiments.

In this study, for simplicity, $J = 1$ and $K = 1$ were chosen. Thus, the dataset can be written as Equation (20).

$$S = \left\{ \chi^k, w_f^k \right\} = \left\{ \left(\chi_f^k, \chi_q^k \right), w_f^k \mid \chi_f^k \in \mathbb{R}^{11 \times 1}, \chi_q^k \in \mathbb{R}^{1 \times 3}, w_f^k \in \mathbb{R}^{11 \times 1}, k = 1, \dots, 600 \right\} \quad (20)$$

where

$$\begin{aligned}\chi^k &= (\chi_f^k, \chi_q^k) \\ \chi_f^k &= [w(\bar{F}, k-1)] \\ \chi_q^k &= [q_1(k-1), q_2(k-1), q_3(k-1)] \\ w_f^k &= w(\bar{F}, k) = [w(f_1, k), \dots, w(f_{11}, k)]^T\end{aligned}$$

For χ^k in the dataset S , AP clustering is used. The AP method sets the coefficient of dampness λ to 0.9 and the reference degree P to 0.5. Following the process of clustering learning, a total of Seven fuzzy rules and their accompanying clustering cores were successfully acquired. Equation (7) shows the breadth setting of the Gaussian spatial membership function, whereas Equation (8) shows the breadth setting of the conventional Gaussian membership function.

This paper uses Fourier space base functions to represent the fuzzy rule resulting set, and applies the SPSA method to improve the coefficients of Fourier space base functions.

In the SPSA algorithm, the relevant parameter settings are as follows: $\alpha = 0.602$, $\beta = 0.101$, $iter_{\max} = 50000$, $A = 5000$, $d = 0.003$, $b = 0.015$.

Based on the learning results of the AP algorithm, it is evident that there are a total of 7 fuzzy rules, so the parameter dimension that needs to be adjusted is 28, and the optimization range control of the parameters is $[-10, 10]$. The end point criterion is defined as either reaching the maximum number of iterations or when the change in iteration becomes less than 0.00001 for five consecutive occasions.

The initial value of $\hat{\theta}_0$ is obtained through gradient descent algorithm, as shown in Equation (21).

$$\begin{aligned}& [0.0376 \quad 29.4590 \quad -0.0307 \quad -1.1126 \\& -0.0343 \quad 26.6246 \quad -0.0037 \quad 0.7921 \\& -0.0272 \quad 26.6246 \quad -0.0464 \quad 0.1736 \\& 0.0048 \quad 25.0165 \quad 0.0339 \quad 0.4068 \\& 0.0007 \quad 18.4397 \quad -0.0079 \quad 0.9764 \\& 0.0016 \quad 18.4623 \quad -0.0041 \quad -0.0672 \\& -0.0323 \quad -8.0553 \quad -0.0381 \quad 0.2669]\end{aligned} \quad (21)$$

By using the SPSA algorithm, the optimal parameters for the Fourier space base function of the fuzzy rule resulting set can be obtained.

For the convenience of observation, Figure 6 shows seven space base functions.

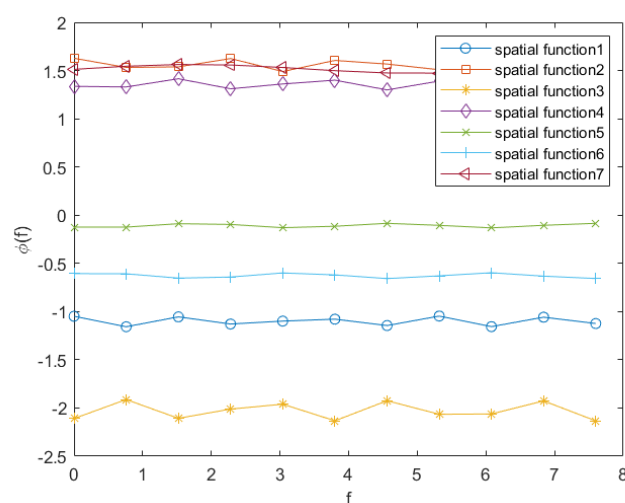


Figure 6. Seven space base functions learned by SPSA algorithm

The seven three-dimensional fuzzy rules derived from the data are listed as follows:

$\bar{V}_1$: if $w(\bar{F}, k-1)$ is very Positive Large and $q_1(k-1)$ is Positive Small and $q_2(k-1)$ is Positive Large and $q_3(k-1)$ is very Positive Large, then $w(\bar{F}, k)$ is [0.8147 0.9058 0.1270 0.9134 0.6324 0.0975 0.2785 0.5469 0.9575 0.9649 0.1576].

$\bar{V}_2$: if $w(\bar{F}, k-1)$ is more than Zero and $q_1(k-1)$ is less than Positive Medium and $q_2(k-1)$ is Positive Small and $q_3(k-1)$ is Positive Medium, then $w(\bar{F}, k)$ is [6.7941 6.7002 3.3976 5.6020 0.9932 2.9523 6.4101 5.5455 6.7164 4.5902 0.2500].

$\bar{V}_3$: if $w(\bar{F}, k-1)$ is less than Positive Small and $q_1(k-1)$ is more than Positive Medium and $q_2(k-1)$ is more than Zero and $q_3(k-1)$ is very Positive Large, then $w(\bar{F}, k)$ is [12.7369 -14.0099 -10.1810 -11.3661 -11.1470 -5.8834 -9.8322 -2.5678 -10.5907 -0.4775 -4.1538].

$\bar{V}_4$: if $w(\bar{F}, k-1)$ is very Positive Large and $q_1(k-1)$ is Positive Large and $q_2(k-1)$ is slightly Positive Small and $q_3(k-1)$ is Positive Large, then $w(\bar{F}, k)$ is [0.4617 0.9713 8.2346 6.9483 3.1710 9.5022 0.3445 4.3874 3.8156 7.6552 7.9520].

$\bar{V}_5$: if $w(\bar{F}, k-1)$ is more than Positive Medium and $q_1(k-1)$ is Positive Large and $q_2(k-1)$ is more than Zero and $q_3(k-1)$ is Positive Medium, then $w(\bar{F}, k)$ is [1.8687 4.8976 4.4559 6.4631 7.0936 7.5469 2.7603 6.7970 6.5510 1.6261 1.1900].

$\bar{V}_6$: if $w(\bar{F}, k-1)$ is more than Positive Medium and $q_1(k-1)$ is Positive Medium and $q_2(k-1)$ is slightly Positive Small and $q_3(k-1)$ is very Positive Large, then $w(\bar{F}, k)$ is [-2.4918 -4.7987 -1.7019 -2.9263 -1.1191 -3.7563 -1.2755 -2.5298 -2.5673 -2.3571 -3.125].

$\bar{V}_7$: if $w(\bar{F}, k-1)$ is Positive Large and $q_1(k-1)$ is Positive Small and $q_2(k-1)$ is Positive Small and $q_3(k-1)$ is less than Positive Medium, then $w(\bar{F}, k)$ is [-1.2263 -0.4110 -1.1889 -0.7443 -1.0026 -0.4407 -1.3535 -1.6484 -2.6840 -1.4376 -0.6583].

Finally, the complete three-dimensional fuzzy modeling of RTCVD has been achieved. The abbreviation for the suggested three-dimensional modeling technique, which combines AP clustering with Fourier space base functions, is AP Fourier SPSA-3D fuzzy modeling.

The simulation results of AP-Fourier-SPSA-3D are shown in the following figures. The predicted result of the model in the training data is displayed in Figure 7, while the discrepancies between the expected and actual values are illustrated in Figure 8. Figure 9 displays the anticipated outcome of the procedure in the test data, whereas Figure 10 illustrates the flaws in the predictions. The root mean square error (RMSE) performance indices are shown in Table 1.

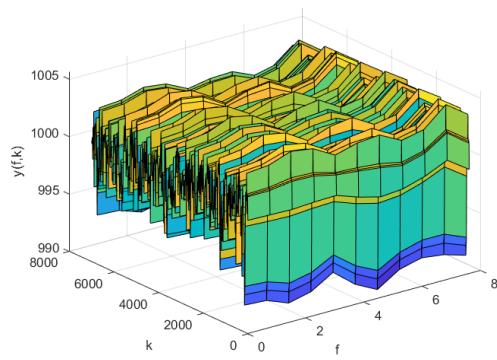


Figure 7. The predicted process output of the AP-Fourier-SPSA-3D model in training data

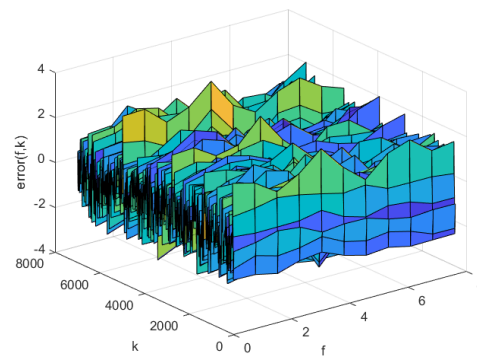


Figure 8. The predicted errors of the AP-Fourier-SPSA-3D model in training data

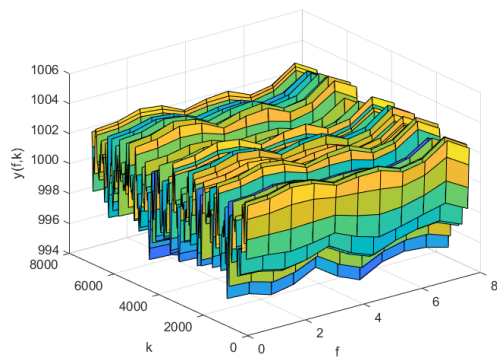


Figure 9. The predicted process output of the AP-Fourier-SPSA-3D model in test data

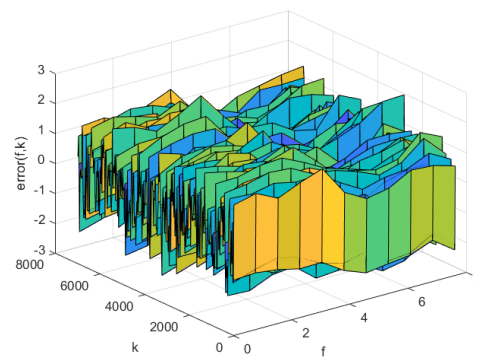


Figure 10. The predicted errors of the AP-Fourier-SPSA-3D model in test data

4.3. Simulation Comparison

In this study, three modeling methods were selected for comparison: KL-LS[23] and NNC-SVR-3D[9]. The KL-LS modeling approach utilizes the KL decomposition technique to compute space functions and decrease model dimensions. Subsequently, the least squares identification method is employed to determine low dimensional temporal coefficients. NNC-SVR-3D fuzzy modeling method is to use nearest neighbor clustering algorithm to learn the preceding components of three-dimensional fuzzy rules, which is reduced with the use of a similarity assessing approach, and use multiple single output support vector machines to learn the space base functions of the resulting components. Additionally, gradient descent approach is utilized to determine the coefficients of Fourier space base function for the suggested three-dimensional fuzzy modeling based on AP clustering and Fourier space base function, which is abbreviated as AP-Fourier GD-3D.

The prediction errors of KL-LS model, NNC-SVR model on training data and test data can be seen in Figure 11 and Figure 12, Figure 13 and Figure 14, respectively. The time normalized absolute errors $TNAE(f) = \sum |e(f, k)| / \sum \Delta k$ of these models on training and test data are shown in Figure 15 and Figure 16, respectively. The RMSE performance indexes of these models are shown in Table 1.

From Table 1 and Figure 7 to Figure 16, it can be evident that the three-dimensional fuzzy models outperform the KL-LS model. The KL-LS modeling method is traditional modeling methods based on KL decomposition, which require dimension reduction. Consequently, this will unavoidably result in the emergence of unmodeled dynamic characteristics and uncertainties. The suggested three-dimensional modeling approach seamlessly merges space-time separation and space-time reconstruction, eliminating the necessity for reducing dimensions. In three three-dimensional fuzzy models,

the proposed AP-Fourier SPSA-3D model has better performance. It discloses that the Fourier space base function has better spatial characteristics to disclose the spatial relationship of the DPS system, the AP clustering algorithm can adaptively select cluster cores and its numbers based on the distribution characteristics of data, and SPSA algorithm can successfully avoid getting stuck in local optima.

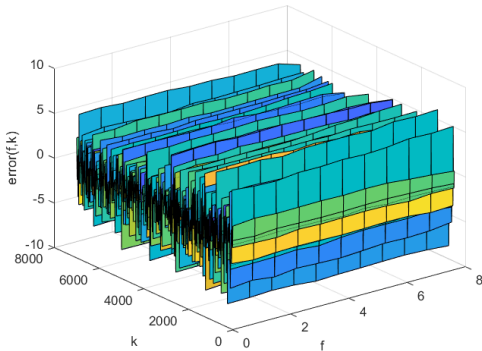


Figure 11. Prediction error of KL-LS model on training set

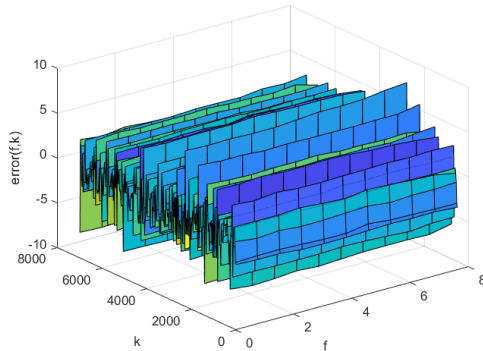


Figure 12. Prediction error of KL-LS model on test set

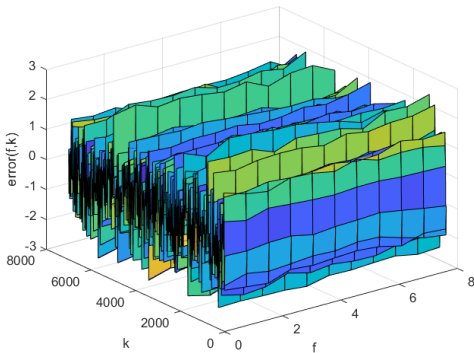


Figure 13. Prediction error of NNC-SVR-3D model on training set

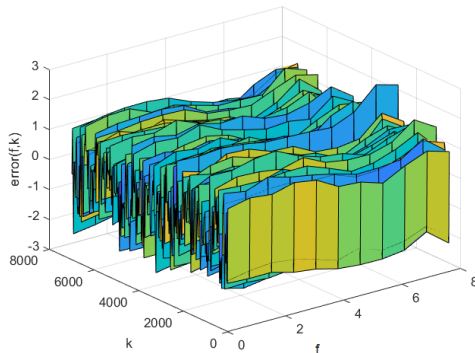


Figure 14. Prediction error of NNC-SVR-3D model on training set

Table 1. Comparison of RMSE among AP-Fourier-SPSA-3D, AP-Fourier-GD-3D, NNC-SVR-3D and KL-LS

model	AP-Fourier-SPSA-3D	AP-Fourier-GD-3D	NNC-SVR-3D	KL-LS
Training data	0.9229	0.9899	1.2862	2.4481
Test data	0.9027	0.9558	1.2214	2.2903

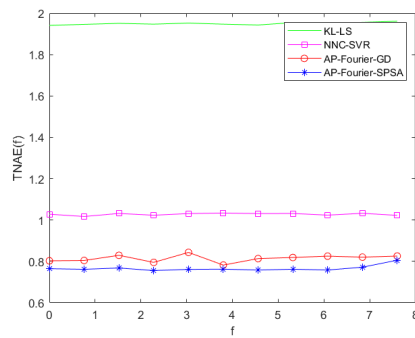


Figure 15. TNAE of four models on training data

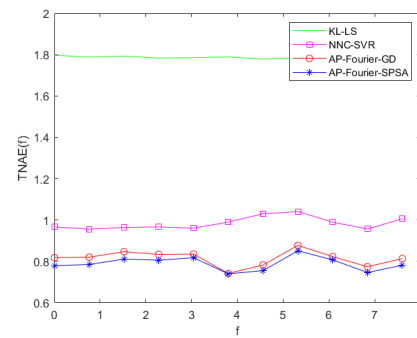


Figure 16. TNAE of four models on test data

5. Conclusions

The study introduced a three-dimensional fuzzy modeling technique using SPSA learning for DPS, which widely exists in the manufacturing industry. On the three-dimensional fuzzy modeling framework, AP clustering was used to learn the preceding components, and adaptively determine the quantity of fuzzy rules. The three-dimensional fuzzy model used the Fourier space base function as the resulting components, while the SPSA method was employed to dynamically acquire the coefficients of the Fourier space base function. The suggested approach was implemented on a conventional distributed parameter system in the semiconductor manufacturing industry. The simulation experiment demonstrates the superiority of the three-dimensional fuzzy modeling method over the conventional space-time separation modeling method. Additionally, it shows that the Fourier space base function is more effective in revealing the spatial distribution characteristics of distributed parameter systems compared to the spatial function obtained through KL decomposition.

Appendix A

Algorithm Description of AP Clustering and Its Flow Chart

Assumed that $\tilde{E} = \{\tilde{E}_1, \dots, \tilde{E}_Z\}$ is a data set, $\tilde{E}_i \in R^n (i = 1, \dots, Z)$. The procedure of AP clustering is described as below and its process diagram is displayed in Figure A.1.

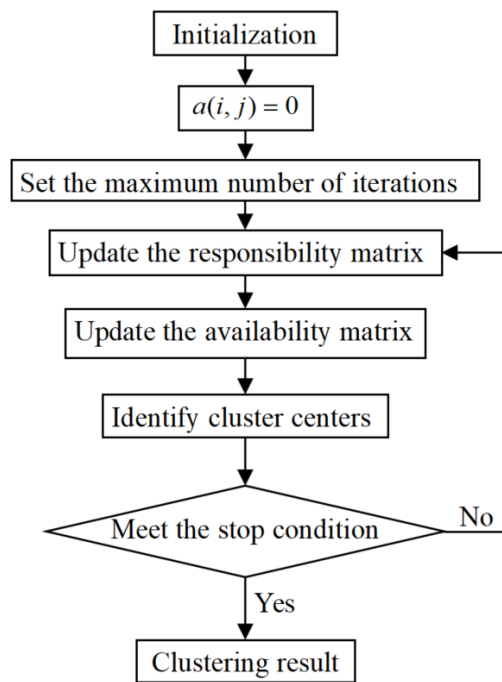


Figure 17. Flow chart of AP clustering algorithm

Step 1: Calculate the similarity matrix S of the data pair. The similarity between any two sample points $\tilde{E}_i$ and $\tilde{E}_v$ is measured by the negative Euclidean distance: $s(i, v) = -\|\tilde{E}_i - \tilde{E}_v\|^2$. Then, the similarity matrix S is normalized as in (A.1)

$$S = \frac{S - \min(\min(S))}{\max(\max(S) - \min(\min(S)))} \quad (\text{A.1})$$

Step 2: Set the maximum number of iteration T , damping factor λ , and reference degree $p(v) = s(v, v) = \text{median}(s(i, v))$.

Step 3: Update the responsibility matrix, the update rule is given as follows.

$$r(i, v) \leftarrow s(i, v) - \max_{v' \neq v} (a(i, v') + s(i, v')) \quad (\text{A.2})$$

$$r(v, v) = p(v) - \max_{j \neq v} \{a(v, j) + s(v, j)\} \quad (\text{A.3})$$

Using damping factor λ , the responsibility matrix $V^{(v)}$ at the k th iteration is updated as in (A.4)

$$V^{(k)} = (1 - \lambda)V^{(k)} + \lambda V^{(k-1)} \quad (\text{A.4})$$

Step 4: Update the availability matrix, the update rules are given as follows

$$a(i, v) \leftarrow \min\{0, r(v, v) + \sum_{i' \notin \{i, v\}} \max(0, r(i', v))\} \quad (\text{A.5})$$

$$a(v, v) \leftarrow \sum_{i' \neq v} \max(0, r(i', v)) \quad (\text{A.6})$$

Using damping factor λ , the availability matrix $A^{(j)}$ at the j th iteration is updated as in (A.7)

$$A^{(j)} = (1 - \lambda)A^{(j)} + \lambda A^{(j-1)} \quad (\text{A.7})$$

Step 5: The decision matrix is obtained by summing the availability matrix and the responsibility matrix. Find out the point in the decision matrix where the diagonal is greater than 0 as the cluster representative center. The sample points are classified according to the distance between the sample points and the cluster representative points.

Step 6: Determine the number of iterations and the iteration stop condition, and assign the data points to the corresponding cluster centers.

Author Contributions: Conceptualization, Tangchen Wang; Data curation, Shaopu Wang; Formal analysis, Chong Cheng; Funding acquisition, Xianxia Zhang; Investigation, Chong Cheng; Methodology, Xianxia Zhang; Project administration, Xianxia Zhang; Resources, Shaopu Wang; Software, Tangchen Wang; Supervision, Xianxia Zhang; Validation, Chong Cheng; Visualization, Shaopu Wang; Writing – original draft, Tangchen Wang; Writing – review & editing, Xianxia Zhang.

Funding: This work was supported by the project from the National Natural Science Foundation of China under Grant 62073210.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Christofides, P. D.; Chow, J. Nonlinear and Robust Control of PDE Systems: Methods and Applications to Transport-Reaction Processes. *Appl. Mech. Rev.* **2002**, *55*, B29-B30.
2. Wang, Z. P.; Wu, H. N.; Chadli, M. H_∞ Sampled-Data Fuzzy Observer Design for Nonlinear Parabolic PDE Systems. *IEEE Trans. Fuzzy Syst.* **2020**, *29*, 1262-1272.
3. Braess, D.; Schumaker, L. L. *Finite Elements: Theory, Fast Solvers, and Applications in Elasticity Theory*; Cambridge University Press: Cambridge, UK, 2007.
4. Lu, X. J.; Zou, W.; Huang, M. H. A Novel Spatiotemporal LS-SVM Method for Complex Distributed Parameter Systems with Applications to Curing Thermal Process. *IEEE Trans. Ind. Inf.* **2016**, *12*, 1156-1165.
5. Wang, M. L.; Li, N.; Li, S. Y. Model-Based Predictive Control for Spatially-Distributed Systems Using Dimensional Reduction Models. *Int. J. Autom. Comput.* **2011**, *8*, 1-7.
6. Li, H. X.; Qi, C. K. Modeling of Distributed Parameter Systems for Applications—A Synthesized Review from Time-Space Separation. *J. Process Control* **2010**, *20*, 891-901.
7. Li, H. X.; Qi, C. K. *Spatio-Temporal Modeling of Nonlinear Distributed Parameter Systems: A Time/Space Separation Based Approach*; Springer Science & Business Media: Berlin, Germany, 2011.
8. Fan, Y. J.; Xu, K. K.; Wu, H.; Zheng, Y.; Tao, B. Spatiotemporal Modeling for Nonlinear Distributed Thermal Processes Based on KL Decomposition, MLP and LSTM Network. *IEEE Access* **2020**, *8*, 25111-25121.
9. Zhang, X. X.; Zhao, L. R.; Li, H. X.; Ma, S. W. A Novel Three-Dimensional Fuzzy Modeling Method for Nonlinear Distributed Parameter Systems. *IEEE Trans. Fuzzy Syst.* **2018**, *27*, 489-501.
10. Wang, M. L.; Qi, C. K.; Yan, H. C.; Shi, H. B. Hybrid Neural Network Predictor for Distributed Parameter System Based on Nonlinear Dimension Reduction. *Neurocomputing* **2016**, *171*, 1591-1597.
11. Zhang, R. D.; Tao, J. L.; Lu, R. Q.; Jin, Q. B. Decoupled ARX and RBF Neural Network Modeling Using PCA and GA Optimization for Nonlinear Distributed Parameter Systems. *IEEE Trans. Neural Netw. Learn. Syst.* **2016**, *29*, 457-469.
12. Zhang, X. X.; Fu, Z. Q.; Li, S. Y.; Zou, T.; Wang, B. A Time/Space Separation Based 3D Fuzzy Modeling Approach for Nonlinear Spatially Distributed Systems. *Int. J. Autom. Comput.* **2018**, *15*, 52-65.
13. Frey, B. J.; Dueck, D. Clustering by Passing Messages Between Data Points. *Science* **2007**, *315*, 972-976.
14. Deng, H.; Li, H. X.; Chen, G. Spectral-Approximation-Based Intelligent Modeling for Distributed Thermal Processes. *IEEE Trans. Control Syst. Technol.* **2005**, *13*, 686-700.

15. Meng, X. B.; Chen, C. L. P.; Li, H. X. Confidence-Aware Multiscale Learning for Online Modeling of Distributed Parameter Systems with Application to Curing Process. *IEEE Trans. Ind. Electron.* **2023**, *70*, 9432-9440.
16. Wei, P.; Li, H. X. Two-Dimensional Spatial Construction for Online Modeling of Distributed Parameter Systems. *IEEE Trans. Ind. Electron.* **2022**, *69*, 10227-10235.
17. Feng, Y.; Zhu, X. Y.; Wang, Y. N.; Wang, B. C.; Zhang, H.; Wu, Z. G.; Yan, H. C. PDE Model-Based On-Line Cell-Level Thermal Fault Localization Framework for Batteries. *IEEE Trans. Syst. Man Cybern. Syst.* **2024**, *54*, 2507-2516.
18. Li, H. X.; Zhang, X. X.; Li, S. Y. A Three-Dimensional Fuzzy Control Methodology for a Class of Distributed Parameter Systems. *IEEE Trans. Fuzzy Syst.* **2007**, *15*, 470-481.
19. Zhang, X. X.; Jiang, Y.; Li, H. X.; Li, S. Y. SVR Learning-Based Spatiotemporal Fuzzy Logic Controller for Nonlinear Spatially Distributed Dynamic Systems. *IEEE Trans. Neural Netw. Learn. Syst.* **2013**, *24*, 1635-1647.
20. Spall, J. C. *Introduction to Stochastic Search and Optimization: Estimation, Simulation, and Control*; John Wiley & Sons: Hoboken, NJ, USA, 2005.
21. Theodoropoulou, A.; Adomaitis, R. A.; Zafiriou, E. Model Reduction for Optimization of Rapid Thermal Chemical Vapor Deposition Systems. *IEEE Trans. Semicond. Manuf.* **1998**, *11*, 85-98.
22. Adomaitis, R. A. A Reduced-Basis Discretization Method for Chemical Vapor Deposition Reactor Simulation. *Math. Comput. Model.* **2003**, *38*, 159-175.
23. Qi, C. K.; Li, H.-X. A Karhunen-Loève Decomposition-Based Wiener Modeling Approach for Nonlinear Distributed Parameter Processes. *Ind. Eng. Chem. Res.* **2008**, *47*, 4184-4192.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.