HIGH POWER LASER
SCIENCE AND ENGINEERING

RESEARCH ARTICLE

Neural network modeling and prediction of HfO$_2$ thin film properties tuned by thermal annealing

Min Gao [1,2], Chaoyi Yin[2], Jianda Shao[2,3,4], and Meiping Zhu[1,2,3,4]

[1] *School of Microelectronics, Shanghai University, Shanghai, China*

[2] *Laboratory of Thin Film Optics, Key Laboratory of Materials for High Power Laser, Shanghai Institute of Optics and Fine Mechanics, Chinese Academy of Sciences, Shanghai, China*

[3] *Center of Materials Science and Optoelectronics Engineering, University of Chinese Academy of Sciences, Beijing, China*

[4] *Hangzhou Institute for Advanced Study, University of Chinese Academy of Sciences, Hangzhou, China*

Abstract

Plasma-enhanced atomic layer deposition (PEALD) is gaining interest in thin films for laser applications, and post-annealing treatments are often used to improve thin film properties. However, research to improve thin film properties is often based on an expensive and time-consuming trial-and-error process. In this study, PEALD-HfO$_2$ thin film samples were deposited and treated under different annealing atmospheres and temperatures. The samples were characterized in terms of their refractive indices, layer thicknesses and O/Hf ratios. The collected data were split into training and validation sets and fed to multiple back-propagation neural networks with different hidden layers to determine the best way to construct the process–performance relationship. The results showed that the three-hidden-layer back-propagation neural network (THL-BPNN) achieved stable and accurate fitting. For the refractive index, layer thickness and O/Hf ratio, the THL-BPNN model achieved accuracy values of 0.99, 0.94 and 0.94, respectively, on the training set and 0.99, 0.91 and 0.90, respectively, on the validation set. The THL-BPNN model was further used to predict the laser-induced damage threshold of PEALD-HfO$_2$ thin films and the properties of the PEALD-SiO$_2$ thin films, both showing high accuracy. This study not only provides quantitative guidance for the improvement of thin film properties but also proposes a general model that can be applied to predict the properties of different types of laser thin films, saving experimental costs for process optimization.

Keywords: laser-induced damage threshold; laser thin film; neural network; plasma-enhanced atomic layer deposition

1. Introduction

Optical thin films are key components of laser systems, and their optical properties and laser-induced damage threshold (LIDT) directly affect their output energy[1–3]. Traditional preparation methods for laser thin films include electron-beam evaporation[4–6] and ion-beam sputtering[7]. Recently, plasma-enhanced atomic layer deposition (PEALD) has attracted attention because of its precise thickness controllability[8], excellent conformality[9], low-temperature growth properties[10] and high LIDT[11]. Furthermore, post-treatment annealing improves the properties of thin films grown via PEALD[12]. However, owing to the diversity and wide range of process parameters, process optimization and thin film performance improvement often require extensive, expensive and time-consuming experiments.

Back-propagation neural networks (BPNNs), a subset of machine learning, have shown potential for mapping the relationship between experimental parameters and material properties[13,14]. This approach can identify underlying regularities in the training data by updating the internal weight parameters[15,16]. In recent years, researchers have begun to study the application of neural networks in the field of thin films to predict the growth rate[17–20], hydrophobicity[21], permeate flux and foulant rejection[22]. Although these reports demonstrate the application of BPNNs in various thin films, studies on the properties of laser thin films are lacking. Furthermore, the adopted models were mainly shallow structures with single or double hidden layers. Shallow-structure neural networks can meet most modeling and prediction needs but may require a large number of neurons to accurately represent the relationship between the input

Correspondence to: Meiping Zhu, Shanghai Institute of Optics and Fine Mechanics, Chinese Academy of Sciences, Shanghai 201800, China. Email: bree@siom.ac.cn

1

and output[23], which increases the likelihood of errors in models[22]. In 2022, Mengu *et al.*[24], while studying the emerging symbiotic relationship between deep learning and optics, reported the advantages of deep neural networks with three or more hidden layers in terms of approximation and generalization capability. However, as the number of hidden layers increases, deep neural networks may suffer from poor performance or training failure owing to issues such as vanishing/exploding gradients[25]. Therefore, it is necessary to determine the optimal number of hidden layers for solving a special task.

In this study, we employ several BPNN models to establish the relationship between the annealing process and the properties of PEALD-HfO$_2$ thin films for laser applications. Firstly, comparing the performance of BPNN models with different numbers of hidden layers, it is deduced that the three-hidden-layer back-propagation neural network (THL-BPNN) performs best. The THL-BPNN model was then used to model and predict the relationship between the annealing process and the PEALD-HfO$_2$ thin film properties and was compared with the other two models. Finally, the LIDT of the PEALD-grown thin films and the properties of the PEALD-SiO$_2$ thin films were predicted using the THL-BPNN model, and the applicability of the THL-BPNN model was verified. We believe that the THL-BPNN model can help predict the properties of other laser thin films.

2. Materials and methods

2.1. Data preparation

The HfO$_2$ thin films used to construct the annealing process–thin film property relationship were grown on Si substrates using a commercial PEALD device (Picosun Advanced R200, Finland) with an integrated remote plasma source. HfO$_2$ thin films were grown by alternating exposure to the precursor tetrakis-ethylmethylamino hafnium (Hf(N(CH$_3$)(CH$_2$CH$_3$))$_4$, TEMAH) and O$_2$/Ar gas mixture plasma reactant at a deposition temperature of 150°C. The number of deposition cycles was 500, and the pulse sequence for each HfO$_2$ growth cycle was as follows: TEMAH feeding (1.6 s), N$_2$ purging (19 s), Ar/O$_2$ mixture feeding (11 s) and Ar purging (10 s). The samples were then

annealed in quartz tube annealing equipment (RS 80/300/11, Nabertherm) for 3 h. The annealing process included a combination of three atmospheres (vacuum, O$_2$ and N$_2$) and six annealing temperatures (300°C to 800°C in 100°C increments). For vacuum annealing, the pressure in the tubular annealing chamber was approximately 1×10^{-4} Pa. For O$_2$ and N$_2$ atmosphere annealing, the gas flow rate was 150 SCCM for both O$_2$ and N$_2$. The HfO$_2$ thin films were measured using an ellipsometer (Horiba Uvisel 2), and the thicknesses and refractive indices were extracted using the Tauc-Lorentz model in DeltaPsi2 software, neglecting the extinction coefficient (k). The O/Hf ratio of the HfO$_2$ thin films was analyzed using X-ray photoelectron spectroscopy (XPS) (Thermo Scientific) with a monochromatic Al Kα (1486.6 eV) X-ray source. The data used to construct the annealing process–thin film property relationship consisted of 19 samples, including 1 as-deposited sample and 18 annealed samples.

The HfO$_2$ thin film data used for LIDT modeling and prediction come from Ref. [26], including 12 samples treated by different annealing process parameters. Among them, six samples were annealed in an O$_2$ atmosphere, and the other six samples were annealed in a N$_2$ atmosphere. The annealing temperature ranged from 300°C to 800°C.

The SiO$_2$ thin film data used for property modeling and prediction come from Ref. [27], including 10 samples grown by different deposition process parameters. Among them, four samples were grown at different temperatures ranging from 50°C to 200°C, and six samples were grown with different precursor source exposure times ranging from 0.2 to 0.7 s.

Table 1 lists the detailed parameters of the datasets used to model and predict the properties of HfO$_2$ and SiO$_2$ thin films, including the refractive index, thickness and stoichiometric ratio. As the annealing temperature increases, the thickness of the HfO$_2$ thin film decreases and the refractive index increases. In a vacuum environment, O$_2$ environment and N$_2$ environment, the thickness of HfO$_2$ thin films annealed at different temperatures changes in the range of 34.7–42.7, 38.5–49.1 and 36.3–46.7 nm, respectively, while the refractive index (at 355 nm) of HfO$_2$ thin films annealed at different temperatures changes in the range of 1.99–2.24, 1.83–1.97 and 1.88–2.00, respectively. This means that the

Table 1. Datasets for property prediction of HfO$_2$ and SiO$_2$ thin films.

	HfO$_2$ thin films		SiO$_2$ thin films	
	Variables	Range	Variables	Range
Input	Annealing atmosphere*	0–3	Deposition temperature (°C)	50–200
	Annealing temperature (°C)	0–800	Precursor exposure time (s)	0.1–0.7
Output	Refractive index (at 355 nm)	1.83–2.24	Refractive index (at 355 nm)	1.48–1.49
	Thickness (nm)	34.7–50.3	Thickness (nm)	69.0–88.1
	O/Hf ratio	1.80–2.04	O/Si ratio	1.94–2.01

*Note: 0, 1, 2 and 3 represent the as-deposited sample, O$_2$, N$_2$ and vacuum, respectively.

Table 2. Datasets for LIDT prediction of HfO$_2$ and SiO$_2$ thin films.

|		Range	
	Variables	HfO$_2$ thin films	SiO$_2$ thin films
Input	Type[*]	1	2
	Total impurity content (%, atomic fraction)	5.4–13.5	0.6–1.1
	Absorption (ppm)	211–10892	3.8–5.8
	Stoichiometric ratio	1.81–2.06	1.94–2.01
Output	LIDT (J/cm^2)	1.2–6.3	22.0–39.4

[*]Note: 1 and 2 represent HfO$_2$ samples and SiO$_2$ samples, respectively.

packing density of the HfO$_2$ thin film increases with increasing annealing temperature[26]. In addition, the O/Hf ratio of HfO$_2$ thin films annealed in an O$_2$ environment fluctuates slightly around the ideal value of 2.0. However, the O/Hf ratio of HfO$_2$ thin films annealed in vacuum and N$_2$ environments decreases with increasing annealing temperature.

Table 2 lists the detailed parameters of the datasets used for LIDT modeling and prediction. Compared with PEALD-HfO$_2$ thin films, PEALD-SiO$_2$ thin films have lower absorption and impurity content. Furthermore, properties such as absorption, impurity content and stoichiometric ratio influence each other. Detailed relationships are described in Refs. [26,27]. The LIDT was tested in one-on-one mode according to ISO 21254 using a Gaussian-shape 3ω neodymium-doped yttrium aluminum garnet (Nd:YAG) laser (355 nm, 7.8 ns). The LIDT test was performed under normal incidence, and the maximum laser fluence with zero damage probability was determined as the LIDT. It is worth mentioning that the LIDT of HfO$_2$ thin films is lower than that of SiO$_2$ thin films, which is attributed to the fact that the bandgap of HfO$_2$ is lower than that of SiO$_2$.

2.2. Models

Six models, namely four BPNN models with different numbers of hidden layers (single-hidden-layer BPNN, double-hidden-layer BPNN, three-hidden-layer BPNN and four-hidden-layer BPNN), a support vector machine regression (SVR) model[28] using a Gaussian kernel function and a linear regression (LR) model[29], were used to establish the correlation between the annealing process and the refractive index, layer thickness and O/Hf ratio of PEALD-HfO$_2$ thin films. Except for the LR model, which belongs to the category of linear regression fitting, the other models belong to the category of nonlinear regression fitting. All models performed regression fitting by training on a training set, tuning the modeling parameters to achieve the highest accuracy (i.e., lowest error) and then validating on a validation set. When constructing the relationship between the annealing process and the properties of the PEALD-HfO$_2$ thin films, 6 samples were randomly selected as the
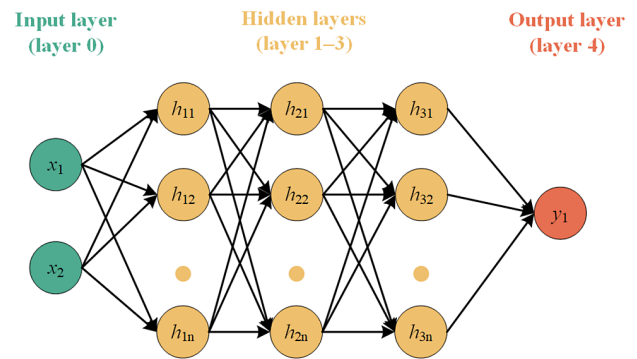
Figure 1. THL-BPNN model with all neurons in adjacent layers connected, where $x = [x_1; x_2]$, y_1 and h_{ij} represent the input, output and intermediate processing signals, respectively.

validation set, and the remaining 13 samples (12 annealed samples and 1 as-deposited sample) were used as the training set. When predicting the LIDT of PEALD-grown thin films, 6 samples (3 HfO$_2$ samples and 3 SiO$_2$ samples) were randomly selected as the validation set, and the remaining 16 samples (9 HfO$_2$ samples and 7 SiO$_2$ samples) were used as the training set. When predicting the properties of PEALD-SiO$_2$ thin films, the leave-one-out cross-validation method was adopted owing to limited data. For each test, one sample was used as a validation set, and the remaining samples were used as a training set until every sample was used as a validation set. Subsequently, the average performance deviation was calculated for each model.

Figure 1 shows a schematic of the THL-BPNN model, including an input layer (layer 0), three hidden layers (layers 1–3) and an output layer (layer 4), with each layer containing one or more neurons. The number of neurons in the input and output layers was determined by the number of input and output variables in the dataset, whereas the number of neurons in the hidden layers was initially determined using Equation (1) (an empirical formula) and finally determined by a global traversal search:

$$l = \sqrt{u + v} + a, \tag{1}$$

where u, v and l are the numbers of neurons in the input, output and hidden layers, respectively, and a is a random number between 1 and 10.

The neurons receive input signals from the previous layer and generate output signals for the next layer[30,31]. For example, the first neuron in layer 1 (from top to bottom), the circle where h_{11} is located, receives input signals, $x = [x_1; x_2]$, from layer 0. Then x undergoes linear transformation to get the weighted sum, z, which is expressed as follows:

$$z = w^T x + b, \tag{2}$$

where $w = [w_1; w_2] \in \mathbf{R}$ is a weight vector between the neurons, and $b \in \mathbf{R}$ is a bias.

Subsequently, z passes through a nonlinear activation function $f(\cdot)$[32], and the output signal h_{11} is generated as follows:

$$h_{11} = f(z). \tag{3}$$

These processes were performed for each neuron in each layer to form the final output signal, y_1[33]. Obviously, mapping from the input space to the output space is initially established through layer-by-layer information transfer.

To further improve the mapping accuracy, a training loss was constructed in the output layer, and an appropriate training algorithm is selected to update the relevant parameters (weights w and bias b) in combination with the chain rule[34] until the loss or the number of iterations reaches the preset threshold[35]. The Levenberg–Marquardt algorithm[36] was used to solve the nonlinear least squares problem. The hyperbolic tangent function was selected as the activation function for all hidden layers. The initialization state of each run was fixed to avoid interference from other factors.

2.3. Model specification and evaluation

2.3.1. Variable scaling
Considering that different distribution ranges of the input and output values may lead to biased assessments, Equation (4) is used to scale the input and output of the data to [−1, 1]:

$$X_{norm} = \frac{(Y_{max} - Y_{min})(X - X_{min})}{X_{max} - X_{min}} + Y_{min}, \tag{4}$$

where X is the input or output vector; X_{max} and X_{min} are the maximum and minimum values of the input or output vector, respectively; and Y_{max} and Y_{min} are the maximum and minimum values after normalization, respectively.

2.3.2. Model evaluation metrics
The coefficient of determination (R^2)[37] was used to evaluate the overall performance of each model. The average accuracy (AA) was used to evaluate the performance of each model on a validation set with only a single sample. The root mean square error (RMSE)[38] was used to measure the deviation between the predicted and measured values:

$$R^2 = 1 - \frac{\sum (Y_i - T_i)^2}{\sum (Y_i - \overline{Y})^2}, \tag{5}$$

$$AA = \frac{1}{n} \sum_{i=1}^{n} \left(1 - \frac{|Y_i - T_i|}{|Y_i|}\right), \tag{6}$$

$$RMSE = \left[\frac{1}{n} \sum_{i=1}^{n} (Y_i - T_i)^2\right]^{1/2}, \tag{7}$$

where n is the size of the dataset; Y_i and T_i are the measured and predicted values of the ith sample in the dataset, respectively; and $\overline{Y}$ is the average of the measured values.

A lower RMSE (close to 0) and higher R^2 and AA (close to 1) indicate smaller differences between the measured and predicted values.

3. Results and discussion

3.1. Analysis of the number of hidden layers of the BPNN model

The influence of the number of hidden layers in the BPNN model on the modeling accuracy was studied using the measured data of the refractive index, layer thickness and O/Hf ratio of the PEALD-HfO$_2$ thin films treated with different annealing process parameters. The optimal number of neurons in each hidden layer was determined by a global traversal search on the training set corresponding to the lowest mean absolute error, and then the optimal model was applied to the validation set. For the refractive index and layer thickness datasets, the total number of neurons in the BPNN model with multiple hidden layers was consistent
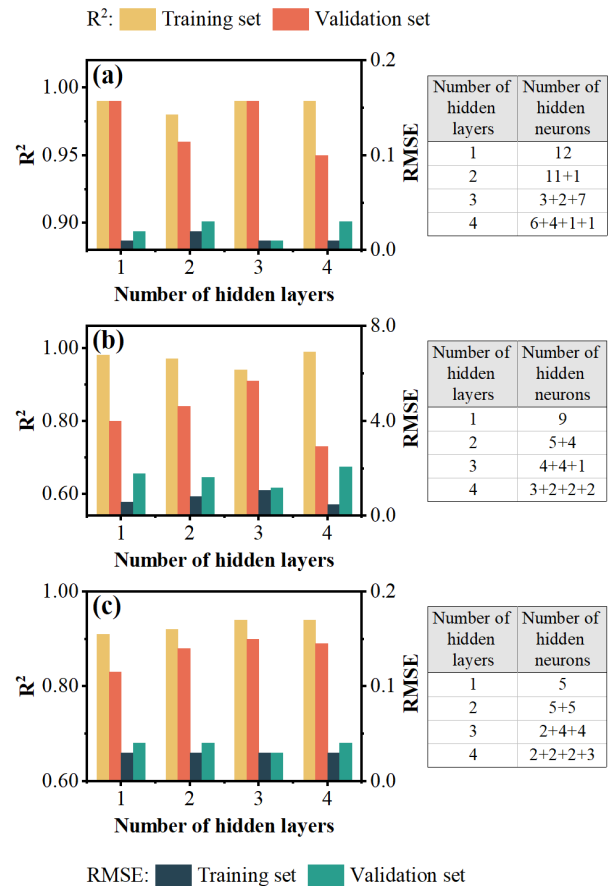
Figure 2. Accuracy of BPNNs with one to four hidden layers based on (a) the refractive index (at 355 nm), (b) layer thickness and (c) O/Hf ratio of PEALD-HfO$_2$. The four columns in each subgraph represent the R^2 values of the model in the training and validation sets and the RMSE values in the training and validation sets, respectively. The table indicates the number of neurons in each hidden layer of each model.
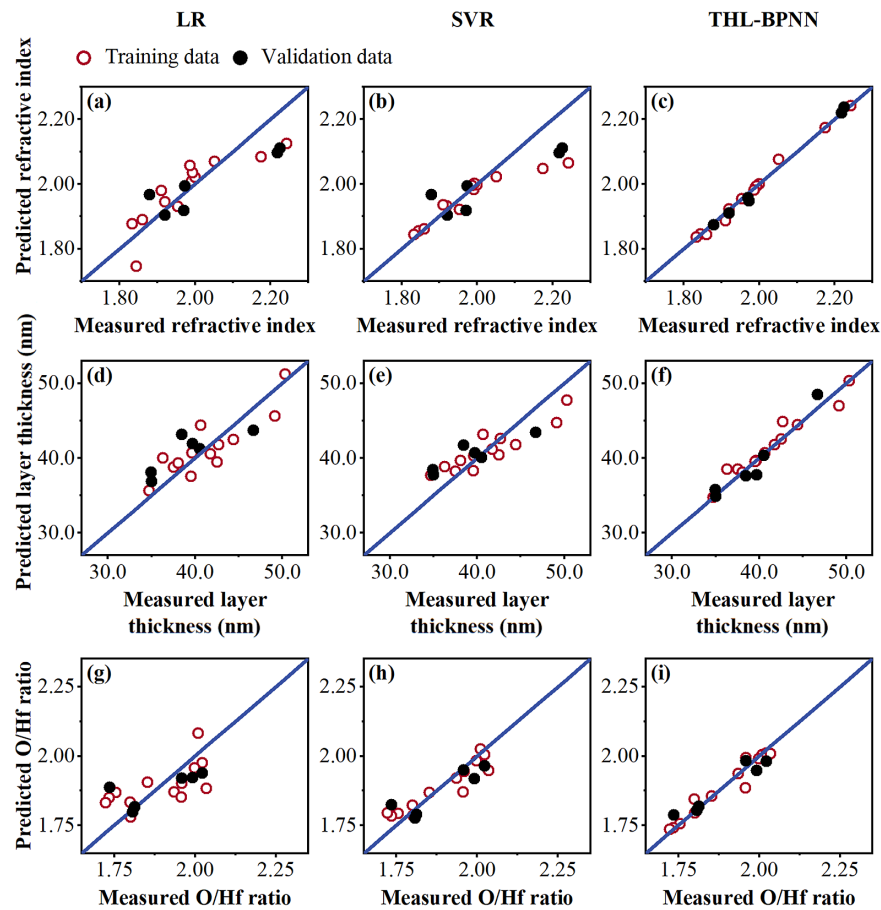
Figure 3. Measured and predicted (a)–(c) refractive index, (d)–(f) layer thickness and (g)–(i) O/Hf ratio of HfO$_2$ thin films. The data in the left-hand, middle and right-hand columns are predicted by the LR model, SVR model and THL-BPNN model, respectively. The blue line (with a slope of 1) serves as a guideline for perfect prediction.

with that of the single-hidden-layer BPNN model. For the O/Hf ratio dataset, because the optimal number of neurons in the single-hidden-layer BPNN model is only five, this value is set as the maximum number of neurons in each hidden layer in the BPNN model with multiple hidden layers. The modeling and prediction accuracies are shown in Figure 2. Overall, as the number of hidden layers increased from one to three, the difference between the R^2 and RMSE in the training and validation sets decreased, indicating that the model moved from inexact to exact fitting. However, as the number of hidden layers was further increased to four, the difference between the R^2 and RMSE in the training and validation sets increased. This may be due to the fact that the combination of neurons in each layer grows exponentially with the number of hidden layers, which introduces the risk of overfitting while potentially obtaining better solutions. The only exception is the modeling of the refractive index, where a single-hidden-layer BPNN also exhibits good performance, which could be attributed to the small variation in the properties and the uncomplicated relationship between the input and output. With the three-hidden-layer BPNN model, the R^2 values of the refractive index, layer thickness and

O/Hf ratio were higher than 0.90 in both the training and validation sets. The THL-BPNN model was selected for the follow-up study.

3.2. Comparison of the THL-BPNN model with other models

The performance of the THL-BPNN model was further evaluated and compared with the LR and SVR models. The refractive index, layer thickness and O/Hf ratio of the HfO$_2$ thin films predicted by the three models were compared with the measured values, as shown in Figure 3 and Table 3. As shown in Figures 3(a), 3(d) and 3(g), the poor performance of the LR model on all three datasets indicates a nonlinear relationship between the annealing process and the thin film properties. As shown in Figures 3(b), 3(e) and 3(h), the SVR model obtains a better fit than the LR model on the layer thickness and O/Hf ratio datasets, but it still does not perform well enough on the refractive index dataset. As shown in Figures 3(c), 3(f) and 3(i), the predicted and measured values of most samples are in good agreement, particularly for the refractive index dataset, indicating that the THL-BPNN

Table 3. Evaluation of the LR, SVR and THL-BPNN models.

	Refractive index				Layer thickness				O/Hf ratio			
	Training data		Validation data		Training data		Validation data		Training data		Validation data	
	R^2	RMSE	R^2	RMSE	R^2	RMSE	R^2	RMSE	R^2	RMSE	R^2	RMSE
LR	0.72	0.06	0.66	0.08	0.74	2.24	0.48	2.88	0.43	0.08	0.48	0.08
SVR	0.71	0.06	0.52	0.10	0.75	2.22	0.56	2.64	0.84	0.04	0.74	0.05
THL-BPNN	0.99	0.01	0.99	0.01	0.94	1.08	0.91	1.18	0.94	0.03	0.90	0.03

model has a high accuracy in modeling and predicting the relationship between the annealing process parameters and HfO$_2$ thin film properties.

Table 3 lists the specific performance of all models on the training and validation sets. The THL-BPNN model performs best among the three regression models, with R^2 values not lower than 0.90 for the refractive index, layer thickness and O/Hf ratio datasets. High R^2 values and low RMSE values indicate that the THL-BPNN model can capture the patterns and extend them to unknown data. In short, the THL-BPNN model shows good stability in constructing the relationship between the annealing process and HfO$_2$ thin film properties under several conditions.

3.3. Evaluation of the THL-BPNN model for other thin film applications

3.3.1. Prediction of the LIDT of PEALD-HfO$_2$ and PEALD-SiO$_2$ thin films

The LIDT value is a key specification for thin films used in laser systems[39,40]. Firstly, we analyzed the main factors affecting the LIDT. According to Ref. [26], the main factors affecting the LIDT of HfO$_2$ thin films are the C impurity content, N impurity content, absorption and O/Hf ratio. Pearson's correlation coefficient was used to further analyze the correlation between the main influencing factors and the LIDT. The results shown in Figure 4 indicate that, except for the O/Hf ratio, which is positively correlated with the LIDT, all other parameters are negatively correlated with the LIDT. The change in the C and N impurity contents can be represented by the total impurity content. Likewise, for SiO$_2$ thin films, factors affecting the LIDT include the total impurity contents, absorption and O/Si ratio. Then, we applied the THL-BPNN to the quantitative prediction of the LIDT based on these factors. The total impurity contents, absorption, stoichiometric ratio and type of thin film were fed into the THL-BPNN as input variables, and the LIDT was derived as the output variable.

Furthermore, the predicted LIDT and measured LIDT of each sample are shown in Figure 5. It is observed that the THL-BPNN model performs well in both training and validation sets with high accuracy and low error, which is smaller than the relative error of the LIDT. The relative error of damage probability is about ±15%, mainly due to the uncertainty of the nonuniformity among the samples
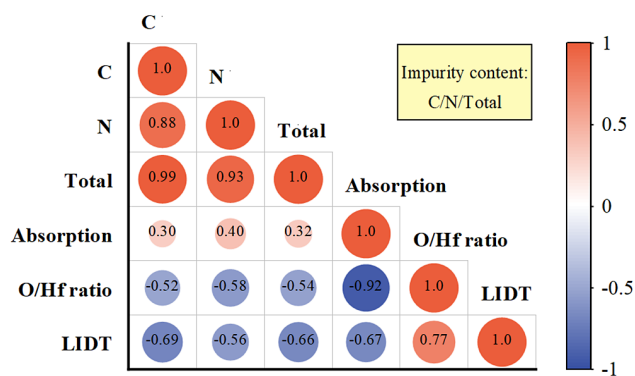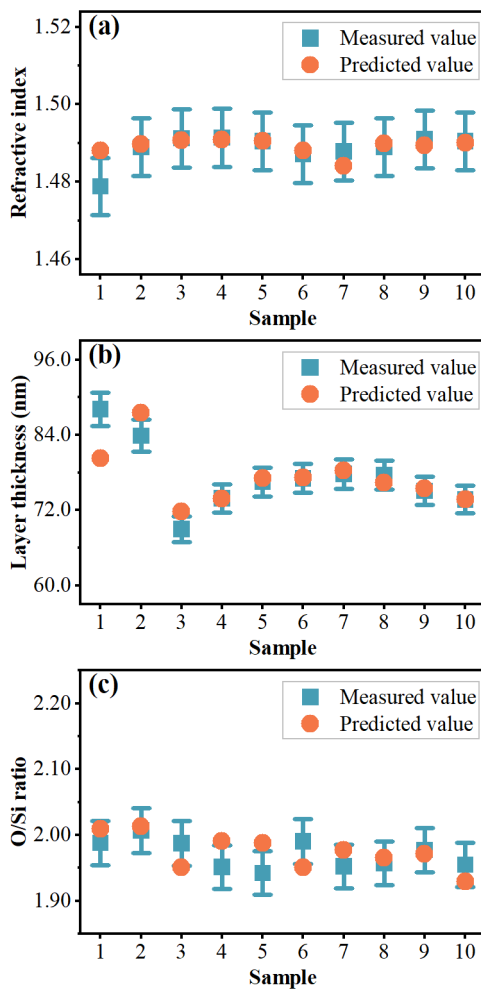

Figure 4. Correlations between properties of HfO$_2$ thin films used in this section. Blue indicates a negative correlation, whereas red indicates a positive correlation. Darker colors and larger circles indicate higher correlations. The numbers inside the circles indicate the corresponding correlation coefficients of the two features.
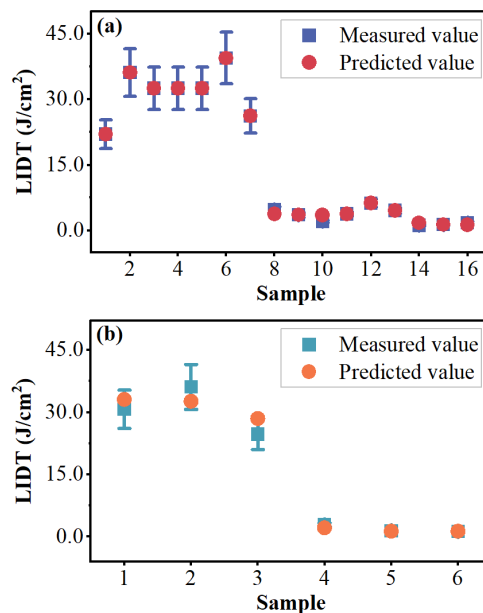
Figure 5. Comparison of measured and predicted LIDT values on the (a) training set and (b) validation set.

(3%), the measurement of the laser spot area (5%) and the fluctuation of laser energy (5%)[41]. For the training set and validation set, the R^2 values are 1.00 and 0.97, respectively, and the RMSE values are 0.48 and 2.32, respectively. The results show that the THL-BPNN model is effective for predicting LIDT values of HfO$_2$ and SiO$_2$ thin films.

Figure 6. Comparison of measured and predicted values of (a) the refractive index (at 355 nm), (b) the layer thickness and (c) the O/Si ratio for SiO$_2$ thin films in the validation set.

Table 4. Evaluation of the THL-BPNN model for SiO$_2$ thin film properties.

	Training data		Validation data	
	R^2	RMSE	AA	RMSE
Refractive index	0.99	0.00	1.00	0.00
Layer thickness	0.99	0.43	0.98	1.72
O/Si ratio	0.81	0.01	0.99	0.03

refractive index and layer thickness in the training set and the AA values of the three properties in the validation set, are higher than 0.98. Although the THL-BPNN model did not perform sufficiently well on the O/Si ratio training set, it still provided accurate predictions on the corresponding validation set. This could be attributed to the successful learning of correlations by the THL-BPNN model through training. Therefore, the THL-BPNN model can be used to construct the relationship between the deposition parameters and PEALD-SiO$_2$ thin film properties, thus proving the universality of the THL-BPNN model in studying the nonlinear relationship between the process parameters and thin film properties.

4. Conclusions

In this study, BPNN models with different numbers of hidden layers were used to establish the correlation between the properties of PEALD-HfO$_2$ thin films and annealing parameters. For modeling, the annealing parameters, including the annealing atmosphere and temperature, were used as inputs, and measured thin film properties, including the refractive index, layer thickness and O/Hf ratio, were used as outputs. The data were split into two categories: a training set and a validation set. Firstly, BPNN models with different numbers of hidden layers were compared. The results demonstrated that as the number of hidden layers was increased to achieve higher accuracy on the training sets, the risk of overfitting also increased. Considering the fitting accuracy and model stability, the THL-BPNN model was adopted in a follow-up study. The performance of the THL-BPNN model was then compared with that of the LR and SVR models. The poor performance of the LR model on most datasets indicated that the effect of the two input features on the dependent output variable was nonlinear. The THL-BPNN model achieved a high accuracy of not less than 0.90 on all training and validation datasets, confirming that the THL-BPNN model outperforms the SVR model, which also belongs to the category of nonlinear regression fitting. Finally, the THL-BPNN model was used to predict the LIDT of PEALD-HfO$_2$ and PEALD-SiO$_2$ thin films, and the mapping relationship between deposition parameters and PEALD-SiO$_2$ thin film properties was constructed. The modeling results showed that the predicted values are consistent with the measured values, proving that the THL-BPNN model is a reliable

3.3.2. Prediction of other properties of PEALD-SiO$_2$ thin films

SiO$_2$ is the most common low-refractive-index material used for laser thin films in the ultraviolet to near-infrared wavelength region. It is of great significance to study the correlation between the properties of SiO$_2$ thin films and the deposition parameters. Therefore, we applied the THL-BPNN model to evaluate the relationship between the deposition parameters and the properties of PEALD-SiO$_2$ thin films. Figure 6 shows the excellent performance of the THL-BPNN model in predicting the properties of PEALD-SiO$_2$ thin films on the validation set, including the refractive index, layer thickness and O/Si ratio. For most samples, the prediction deviation was smaller than the measurement error.

Table 4 lists the R^2, AA and RMSE values of the THL-BPNN model for SiO$_2$ thin film properties. Except for the average R^2 value of the O/Si ratio on the training set of 0.81, the other values, including the average R^2 value of the

predictive learning-based model. We believe that the THL-BPNN model can be used to predict the properties of different types of thin films, thereby reducing the experimental cost of process optimization.

Acknowledgements

References

1. N. Xu, M. Zhu, Y. Chai, B. Roshanzadeh, S. T. P. Boyd, W. Rudolph, Y. Zhao, R. Chen, and J. Shao, Opt. Lett. **43**, 4538 (2018).
2. B. Ma, J. Q. Han, J. Li, K. Wang, S. Guan, X. S. Niu, H. R. Li, J. L. Zhang, H. F. Jiao, X. B. Cheng, and Z. S. Wang, Chin. Opt. Lett. **19**, 081403 (2021).
3. Z. Xing, W. Fan, D. Huang, H. Cheng, and T. Du, High Power Laser Sci. Eng. **10**, e35 (2022).
4. E. S. Field, B. Galloway, D. Kletecka, P. Rambo, I. Smith, V. E. Gruzdev, C. W. Carr, D. Ristau, and C. S. Menoni, Proc. SPIE **11173**, 1117314 (2019).
5. K. Ye, T. Xu, Q. Zhong, Y. Dong, S. Zheng, Z. Xu, and T. Hu, Opt. Express **30**, 24852 (2022).
6. K. Shuai, X. Liu, Y. Zhao, K. Qiu, D. Li, H. Gong, J. Sun, L. Zhou, Y. Jiang, Y. Dai, J. Shao, and Z. Xia, High Power Laser Sci. Eng. **10**, e42 (2022).
7. S. Malobabic, M. Jupe, and D. Ristau, Light Sci. Appl. **5**, e16044 (2016).
8. C. Mahata, Y. C. Byun, C. H. An, S. Choi, Y. An, and H. Kim, ACS Appl. Mater. Interfaces **5**, 4195 (2013).
9. T. Faraz, H. C. M. Knoops, M. A. Verheijen, C. A. A. van Helvoirt, S. Karwal, A. Sharma, V. Beladiya, A. Szeghalmi, D. M. Hausmann, J. Henri, M. Creatore, and W. M. M. Kessels, ACS Appl. Mater. Interfaces **10**, 13158 (2018).
10. L. H. Kim, K. Kim, S. Park, Y. J. Jeong, H. Kim, D. S. Chung, S. H. Kim, and C. E. Park, ACS Appl. Mater. Interfaces **6**, 6731 (2014).
11. H. Liu, L. Jensen, P. Ma, and D. Ristau, Appl . Surf. Sci. **476**, 521 (2019).
12. G. Abromavičius, S. Kičas, and R. Buzelis, Opt. Mater. **95**, 109245 (2019).
13. W. Li, P. Chen, B. Xiong, G. Liu, S. Dou, Y. Zhan, Z. Zhu, T. Chu, Y. Li, and W. Ma, J. Phys. Mater. **5**, 014003 (2022).
14. E. J. Liu, Z. M. Yu, Z. Q. Wan, L. Shu, K. X. Sun, L. L. Gui, and K. Xu, Chin. Opt. Lett. **19**, 113901 (2021).
15. A. Lininger, M. Hinczewski, and G. Strangi, ACS Photonics **8**, 3641 (2021).
16. L. Xia, Y. Z. Hu, W. Y. Chen, and X. G. Li, High Power Laser Sci. Eng. **8**, e28 (2020).
17. G. Kimaev and L. A. Ricardez-Sandoval, J. Phys. Chem. C **124**, 18615 (2020).
18. Y. D. Ko, P. Moon, C. E. Kim, M. H. Ham, M. K. Jeong, A. Garcia-Diaz, J. M. Myoung, and I. Yun, Surf. Interface Anal. **45**, 1334 (2013).
19. Y. D. Ko, P. Moon, C. E. Kim, M. H. Ham, J. M. Myoung, and I. Yun, Expert Syst. Appl. **36**, 4061 (2009).
20. A. Bahramian, Surf. Interface Anal. **45**, 1727 (2013).
21. M. Jafari Gukeh, S. Moitra, A. N. Ibrahim, S. Derrible, and C. M. Megaridis, ACS Appl. Mater. Interfaces **13**, 46171 (2021).
22. M. Fetanat, M. Keshtiara, R. Keyikoglu, A. Khataee, R. Daiyan, and A. Razmjou, Sep . Purif. Technol. **270**, 118383 (2021).
23. G. F. Montufar, Neural Comput. **26**, 1386 (2014).
24. D. Mengu, M. S. Sakib Rahman, Y. Luo, J. Li, O. Kulce, and A. Ozcan, Adv. Opt. Photonics **14**, 209 (2022).
25. M. Liu, L. Chen, X. Du, L. Jin, and M. Shang, IEEE Trans. Neural Network Learn. Syst. **34**, 2156 (2023).
26. Z. Lin, M. Zhu, C. Song, T. Liu, C. Yin, T. Zeng, and J. Shao, J. Alloys Compd. **946**, 169443 (2023).
27. C. Yin, M. Zhu, T. Zeng, C. Song, Y. Chai, Y. Shao, R. Zhang, J. Zhao, D. Li, and J. Shao, J. Alloys Compd. **859**, 157875 (2021).
28. C. Cortes and V. Vapnik, Mach. Learn. **20**, 273 (1995).
29. Y. Wang, W. Wu, X. Zheng, Y. Zeng, M. Ding, and C. Zhang, J. Therm. Spray Technol. **20**, 1177 (2011).
30. Y. Xu, X. Zhang, Y. Fu, and Y. Liu, Photonics Res. **9**, B135 (2021).
31. L. Ma, J. Li, Z. Liu, Y. Zhang, N. Zhang, S. Zheng, and C. Lu, Chin. Opt. Lett. **19**, 011301 (2021).
32. X. Qiu, *Neural Networks and Deep Learning* (China Machine Press, 2020), p. 89.
33. P. R. Wiecha, A. Arbouet, C. Girard, and O. L. Muskens, Photonics Res. **9**, B182 (2021).
34. Y. LeCun, Y. Bengio, and G. Hinton, Nature **521**, 436 (2015).
35. X. Guo, T. D. Barrett, Z. M. Wang, and A. I. Lvovsky, Photonics Res. **9**, B71 (2021).
36. M. T. Hagan and M. B. Menhaj, IEEE Trans. Neural Netw. **5**, 989 (1994).
37. D. Chicco, M. J. Warrens, and G. Jurman, PeerJ Comput. Sci. **7**, e623 (2021).
38. T. Chai and R. R. Draxler, Geosci. Model Dev. **7**, 1247 (2014).
39. T. Y. Pu, W. W. Liu, Y. L. Wang, X. M. Pan, L. Q. Chen, and X. F. Liu, High Power Laser Sci. Eng. **9**, e19 (2021).
40. W. Du, M. Zhu, J. Shi, T. Liu, J. Sun, K. Yi, and J. Shao, High Power Laser Sci. Eng. **11**, e61 (2023).
41. W. Liu, C. Wei, J. Wu, Z. Yu, H. Cui, K. Yi, and J. Shao, Opt. Express **21**, 22476 (2013).