

## Machine-Learning Assisted Prediction of Spectral Power Distribution for Full-Spectrum White Light-Emitting Diode

Volume 12, Number 1, February 2020

Jiajie Fan, *Senior Member, IEEE*

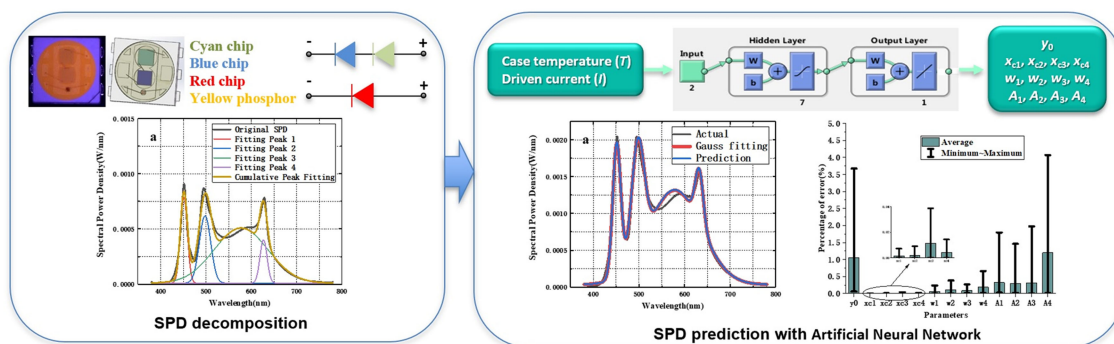
Yutong Li

Irena Fryc

Cheng Qian, *Member, IEEE*

Xuejun Fan, *Fellow, IEEE*

Guoqi Zhang, *Fellow, IEEE*



DOI: 10.1109/JPHOT.2019.2962818

# Machine-Learning Assisted Prediction of Spectral Power Distribution for Full-Spectrum White Light-Emitting Diode

Jiajie Fan <sup>1,2,3</sup> *Senior Member, IEEE*, Yutong Li,<sup>1,2</sup> Irena Fryc <sup>4</sup>,  
Cheng Qian <sup>5</sup> *Member, IEEE*, Xuejun Fan <sup>6</sup> *Fellow, IEEE*,  
and Guoqi Zhang,<sup>3</sup> *Fellow, IEEE*

<sup>1</sup>College of Mechanical and Electrical Engineering, Hohai University, Changzhou, 213022, China

<sup>2</sup>Changzhou Institute of Technology Research for Solid State Lighting, Changzhou, 213161, China

<sup>3</sup>Department of Microelectronics, Delft University of Technology, Delft, 2628 CD, Netherlands

<sup>4</sup>Faculty of Electrical Engineering, Bialystok University of Technology, Bialystok, 15-351, Poland

<sup>5</sup>School of Reliability and Systems Engineering, Beihang University, Beijing, 100191, China

<sup>6</sup>Departments of Mechanical Engineering, Lamar University, Beaumont, TX 77710 USA

DOI:10.1109/JPHOT.2019.2962818

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see <https://creativecommons.org/licenses/by/4.0/>

Manuscript received November 24, 2019; revised December 22, 2019; accepted December 25, 2019. Date of publication December 30, 2019; date of current version January 28, 2020. This work was supported in part by the National Natural Science Foundation of China under Grant 51805147, in part by the Six Talent Peaks Project in Jiangsu Province under Grant GDZB-017, in part by Changzhou Sci&Tech Program under Grant CC20191008 and in part by the Postgraduate Student Practice Innovation Project of Hohai University (Changzhou Campus) under Grant SJCX19\_0159. Corresponding author: Jiajie Fan (e-mail: jay.fan@connect.polyu.hk).

**Abstract:** The full-spectrum white light-emitting diode (LED) emits light with a broad wavelength range by mixing all lights from multiple LED chips and phosphors. Thus, it has great potentials to be used in healthy lighting, high resolution displays, plant lighting with higher color rendering index close to sunlight and higher color fidelity index. The spectral power distribution (SPD) of light source, representing its light quality, is always dynamically controlled by complex electrical and thermal loadings when the light source operates under usage conditions. Therefore, a dynamic prediction of SPD for the full-spectrum white LED has become a hot but challenging research topic in the high quality lighting design and application. This paper proposes a dynamic SPD prediction method for the full-spectrum white LED by integrating the SPD decomposition approach with the artificial neural network (ANN) based machine learning method. Firstly, the continuous SPDs of a full-spectrum white LED driven by an electrical-thermal loading matrix are discretized by the multi-peak fitting with Gaussian model as the relevant spectral characteristic parameters. Then, the Back Propagation (BP) and Genetic Algorithm-Back Propagation (GA-BP) NNs are proposed to predict the spectral characteristic parameters of LEDs operated under any usage conditions. Finally, the dynamically predicted spectral characteristic parameters are used to reconstruct the SPDs. The results show that: (1) The spectral characteristic parameters obtained by fitting with the Gaussian model can be used to represent the emission lights from multiple chips and phosphors in a full-spectrum white LED; (2) The prediction errors of both BP NN and GA-BP NN can be controlled at low level, that is to say, our proposed method can achieve a highly accurate SPD dynamic prediction for the full-spectrum white LED when it operates under different operation mission profiles.

**Index Terms:** Full-spectrum white LED, spectral power distribution, BP neural network, genetic algorithm, machine learning.

## 1. Introduction

With the improvement of living standards, people's requirements on lighting have gradually shifted from environmental protection and energy saving to the pursuit of health and comfort. However, several traditional light-emitting diode (LED) products have high proportion of blue light with short wavelength and low proportion of red light at long wavelength, which will lead to visual fatigue and vision loss [1]. Thus, the design of next generation LED light source will be challenged not only by low cost and high light efficiency, but also need to meet the demands of health, comfort, high color quality, low frequency flash, high reliability and so on [2]. As natural sunlight is the most suitable and comfortable light for human beings, the full-spectrum white LED has great potential applications in indoor lighting, medical and human centric lighting, special display, plant lighting and other fields [3]. Currently, a common design of the full-spectrum white LED is to mix all emission lights from multiple LED chips and phosphors [4]. As a fundamental performance indicator, the spectral power distribution (SPD) of a full-spectrum white LED is complex and always interfered by electrical and thermal loadings [5]. Therefore, achieving dynamic prediction of SPD to simulate natural light spectrum becomes one of essential but challenging research topics in future human centric lighting design and application.

The SPD represents a functional relationship between spectral density with wavelength and the spectral density represents the radiation energy per unit wavelength range. Because there are many internal and external factors to determine the SPD of a LED light source, it is difficult to achieve its dynamic prediction with high accuracy. At present, there are some researches relating the SPD prediction. For instance, J. C. C. Lo *et al.* [6] proposed a mathematical model, in which the excitation and emission spectra of various phosphors were used as input parameters to predict the emission spectra of multiple LED phosphors and compared with the experimental results. It is found that the proposed model can accurately estimate the emission spectrum of the LED with mixed multiple phosphors. P. Dupuis *et al.* [7] developed an SPD prediction method by using the low-order polynomial function with current as the only input parameter. Moreover, the SPD decomposition approach with statistical models is used to discretize the continuous spectrum of a LED and the statistical characteristic parameters is extracted to represent the whole information in an SPD. The commonly used statistical models include Gaussian model, asymmetric double sigmoidal model and Lorentz model and so on [8]–[11]. J. J. Fan [12] *et al.* used the Gaussian and Lorentz models to extract SPD features, and achieved a dynamic and accurate predicting of the color coordinates, correlation color temperature (CCTs), CRIs and estimating the residual life of phosphor converted white LED (PC-wLED). C. Qian [13] *et al.* modeled the SPD of PC-wLED by superimposing two asymmetric double sigmoidal (Asym2sig) functions. H.T. Chen [14] *et al.* provided a method to predict the instantaneous changes of CCT and CRI when the power of LED system was changed, in which the Gaussian function was also used to model the SPD of LED system. M. H. Chang [15] *et al.* used multiple peak fitting method to extract features of SPD for PC-wLED and applied the principal component analysis (PCA) to reduce the dimensionality of features, finally achieved a shorten of the LED qualification test period from 6000 hours to 1200 hours. Generally, the statistical characteristic parameters extracted from the SPD decomposition model have been proved as an effective way to discretize the continuous spectrum, however, the multidimensional data mining and processing of the extracted statistical characteristic parameters become another challenge in SPD prediction.

In general, machine learning (ML) is a set of methods that can be used to learn and detect patterns from input data samples and use the uncovered patterns for further decision making in prognostics or predicting future data [16]. ML always includes supervised learning, unsupervised learning and semi-supervised learning. In the supervised learning, a labeled or classified set of input data is used to estimate and predict the resulting output pattern. As a result, it depends on the learning method to discover the group of input data or desired pattern. As one of most popular supervised learning methods, Artificial Neural Network (ANN) has been proven as an effective way on data mining and processing. It abstracts the human brain neuron network from the perspective of information processing, establishes a simple model, and forms different networks

according to different connection modes [17]. As it has the advantages of function approximation, self-learning, complex classification, associative memory, fast optimization, and strong robustness and fault tolerance brought by highly parallel distributed information storage [18], [19], ANN has been widely used in many fields, such as medicine [20], [21], biology [22], [23], physics [24], and even in the LED field [25]. For example, K. Y. Lu [26] *et al.* proposed a lifetime prediction method based on the multi-dimensional back propagation neural network (BP-NN), in which the Adaboost algorithm improved BP-NN could lower the life prediction error but the operation time was increased. Aiming at the high cost of reliability prediction and evaluation of high-power LED, an intelligent prediction method based on dynamic neural network was proposed by Yan [27]. It is proved that this method has good extrapolation ability and robustness and it can successfully predict the life of high-power LED in a short time with prediction error as less than 5%. Liu *et al.* [28] combined the finite element modeling (FEM) analysis of a single heat transfer physics field with the ANN to present a more efficient method for heat dissipation analysis of multi-chip LED light sources. However, the limitation of BP-NN is that its convergence speed is slow and it is easy to fall into local minimum [29]. Genetic Algorithms (GA) is a computational model to simulate the natural selection and genetic mechanism of Darwin's biological evolution theory. It is a method to search the optimal solution by simulating the natural evolution process [30], [31]. The GA can be used to optimize the initial weights and thresholds of the neural network to prevent the BP-NN from falling into local minimum in the training process. L. Liu *et al.* [32] used the GA-BP neural network to recognize the alphabet. The results show that the network optimized by genetic algorithm has a more accurate and faster convergence compared with BP-NN. J. W. Gao *et al.* [33] predicted the short-term traffic flow, and applied the GA to optimize the weights and thresholds of BP-NN. The results also show that the new model has higher superiority in convergence speed and prediction accuracy.

Taking the dynamic change and complexity of SPDs of a full-spectrum white LED operated under different electrical and thermal loadings into consideration, this paper combines the SPD decomposition method with the ANN based machine learning method (i.e., BP-NN and GA-BP NN) to realize the dynamic prediction of SPDs for a full-spectrum white LED. The remaining of this paper is organized as follows: Section 2 proposes the SPD decomposition model and ANN theory. Section 3 introduces the test sample, experimental setups and collected data used in this study. Section 4 compares and discusses the prediction results obtained by using the BP-NN and GA-BP NN. Finally, the concluding remarks are presented in Section 5.

## 2. Theory and models

### 2.1 The SPD Decomposition Model

Generally, the SPD of full spectrum white LED contains many peaks, which can be modeled by the multi-peak function. In the SPD decomposition model proposed in this paper, a summation of symmetric Gaussian functions was chosen to fit and disassemble the SPD of a full-spectrum white LED, that is expressed as Eq. (1).

$$y = y_0 + \frac{A_1}{w_1\sqrt{\pi/2}} e^{-2\frac{(x-x_{c1})^2}{w_1^2}} + \frac{A_2}{w_2\sqrt{\pi/2}} e^{-2\frac{(x-x_{c2})^2}{w_2^2}} + \dots + \frac{A_n}{w_n\sqrt{\pi/2}} e^{-2\frac{(x-x_{cn})^2}{w_n^2}} \quad (1)$$

Where  $y_0$  represents the initial value,  $A$ ,  $x_c$  and  $w$  are the peak area, peak wavelength and half-wave width of each emitted spectrum respectively.

### 2.2 BP-NN and GA-BP NN

As mentioned before, ANN refers to a complex network structure formed by interconnection of a large number of processing units (neurons), which is an abstraction, simplification and simulation of the structure and operation mechanism of human brain. Fig. 1 is a structure diagram of a single hidden layer neural network, which contains  $n$  inputs and  $m$  outputs. ANN does not show the exact

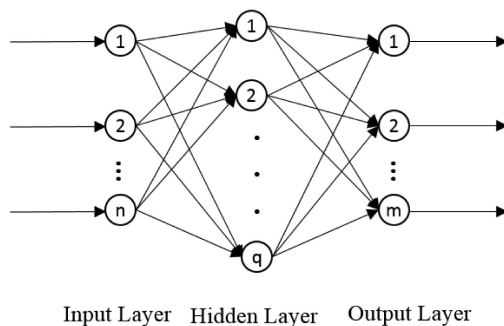


Fig. 1. Single hidden layer neural network structure.

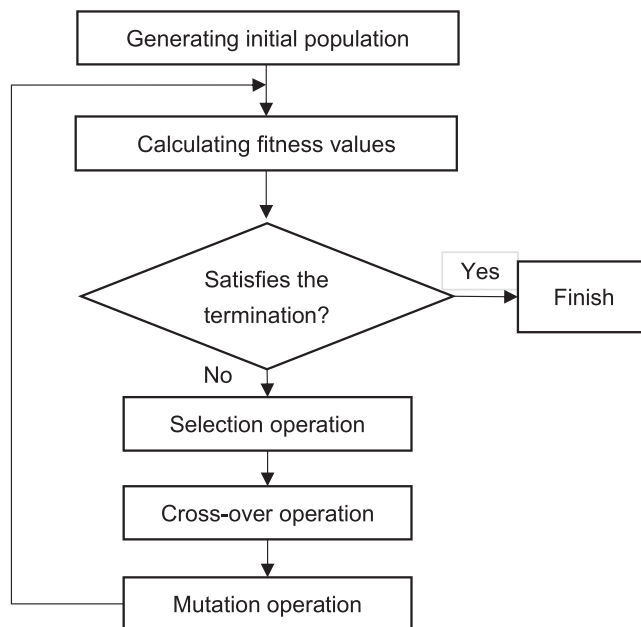


Fig. 2. The flowchart of GA optimization.

relationship between input and output. It only presents the unsteady factors that cause output changes, i.e., non-constant parameters. In this paper, BP-NN was firstly used to predict the SPD of a full-spectrum white LED. BP-NN is a multi-layer feedforward neural network trained by error back propagation algorithm, and is the most widely used one at present. Its basic idea is the gradient descent method, which uses the gradient search technique to minimize the mean square error between the actual output value and the expected output value of the network.

However, BP-NN also has some drawbacks, such as easy to fall into local minimum values, the number of network layers and the number of neurons without corresponding theoretical guidance. Currently, there are some improved BP-NN approaches to accelerate the convergence speed of the network and avoid falling into local minimum. To optimize the BP-NN, the GA model was integrated in this paper with BP-NN to reduce the blindly searching process in the early stage of BP-NN algorithm and find the optimal weights and thresholds. Fig. 2 shows the flowchart of GA optimization.



Fig. 3. A full-spectrum white LED package.

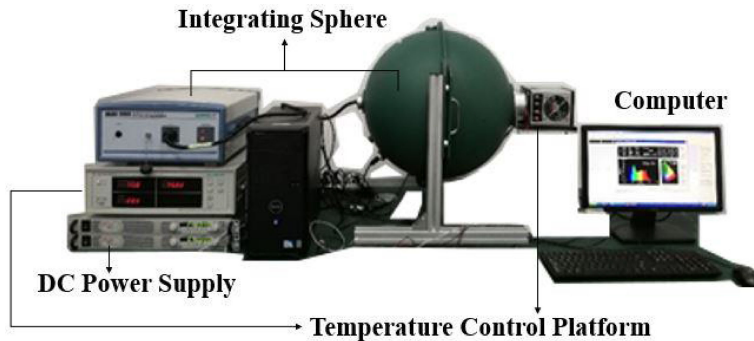


Fig. 4. Experimental setup for SPD data acquisition.

TABLE 1

The Ranges of Case Temperature and Drive Current Selected for Test Sample

Parameters	Initial values	Termination values	Change gradients
Case temperature (°C)	25	80	5
Red LED's driven current (mA)	10	40	1
Cyan-blue LEDs' driven current (mA)	50	200	5

### 3. Test and Data Acquisition

This paper chooses a full-spectrum white LED package with high CRI ( $R_a = 90\sim 92$ ,  $CCT = 4800\sim 5200$  K). It is packaged with a cyan LED chip ( $I_{F_C} = 150$  mA), a blue LED chip ( $I_{F_B} = 150$  mA), a red LED chip ( $I_{F_R} = 30$  mA,  $V_{F_R} = 2.0\sim 3.0$  V) and coated with a yellow phosphor layer, and the cyan chip is connected in series with the blue chip ( $I_{F_{CB}} = 150$  mA,  $V_{F_{CB}} = 5.0\sim 6.0$  V). The package size is  $5.2$  mm  $\times$   $5.4$  mm as shown in Fig. 3.

In order to realize the dynamic prediction of SPD of the full-spectrum white LED operated under different electrical and thermal conditions, we first used the integrating sphere to measure the SPD of test sample driven by an electrical-thermal loading matrix. As shown in Fig. 4, the experimental setup consists of an integrating sphere (Model: EVERFINE HASS20) to collect the SPD data, a DC power supply (Model: KEYSIGHT N5751A) to provide driven current, and a temperature control platform (Model: EVERFINE CL-200) to control the case temperature for test sample. The ranges of case temperature and driven current selected for the test sample are shown in Table 1. The selected full-spectrum white LED package needs two driven currents to power-on the red LED chip and series-connected cyan-blue LED chips separately, and a synchronous change of driven currents is recommended for the test samples. Usually, the limit operation temperature of most LEDs is between  $80\sim 100$  °C. In addition, the excessively high temperature and large driven current may cause serious degradation and even bring catastrophic failure to LEDs. Considering the rated

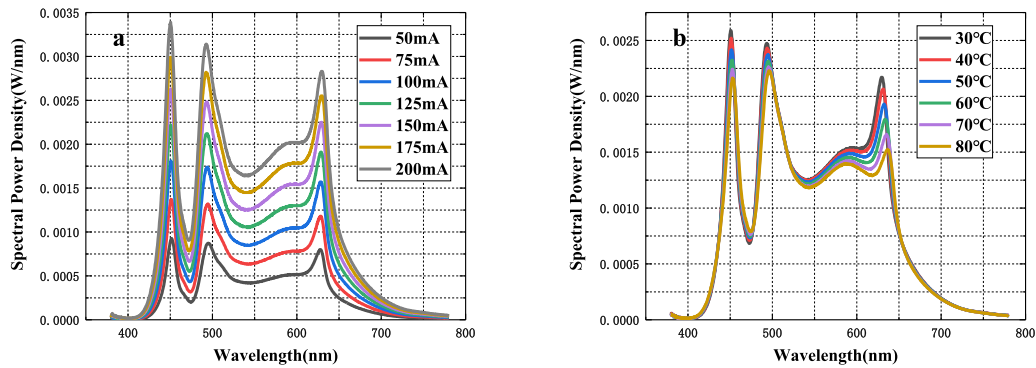


Fig. 5. Effects of case temperature and driven current on the SPD of full-spectrum white LED. (a) 25 °C and (b) 150 mA.

TABLE 2  
The Selected Four Sets of Test Conditions

Test conditions	Case temperature(°C)	Driven current(mA)
a	30	50
b	45	100
c	60	150
d	75	200

currents and actual operation temperature limit, the maximum case temperature was selected as 80 °C and the maximum driven currents were set as 40 mA and 200 mA for the red LED chip and series-connected cyan-blue LED chips respectively. Fig. 5(a) shows the SPD change trend of test samples operated under different driven currents when the case temperature is controlled as 25 °C, where its SPD increased when the driven current rising. Meanwhile, Fig. 5(b) shows the SPD changed under different case temperatures when a drive current was fixed as 150 mA, which indicates that the case temperature has negative impact on SPD.

## 4. Results and Discussion

### 4.1 SPD Decomposition Results

It can be seen from Section 2, the number of fitted peaks is selected according to the spectral power distribution of different LEDs. The full-spectrum LED used in this paper should select four fitted peaks. As calculated from Table 1, 372 sets of test condition are considered in this paper. Firstly, the SPD decomposition model described in Eq. (1) is validated by using the data collected under four sets of test conditions as shown in Table 2. The SPD decomposition modeling with the Gaussian function fitting is present in Fig. 6, which includes the original SPD, four extracted individual spectrum, and cumulative peak-fitting model, respectively. The Goodness-of-Fit is evaluated by using the coefficient of determination ( $R^2$ ). The maximum value of  $R^2$  is 1, and the closer the value of  $R^2$  to 1, the better the fitting degree of regression line to the observation value is. Here, we fit sixty sets of SPD data and obtain all  $R^2$  values. As shown in Table 3, all  $R^2$  values are larger than

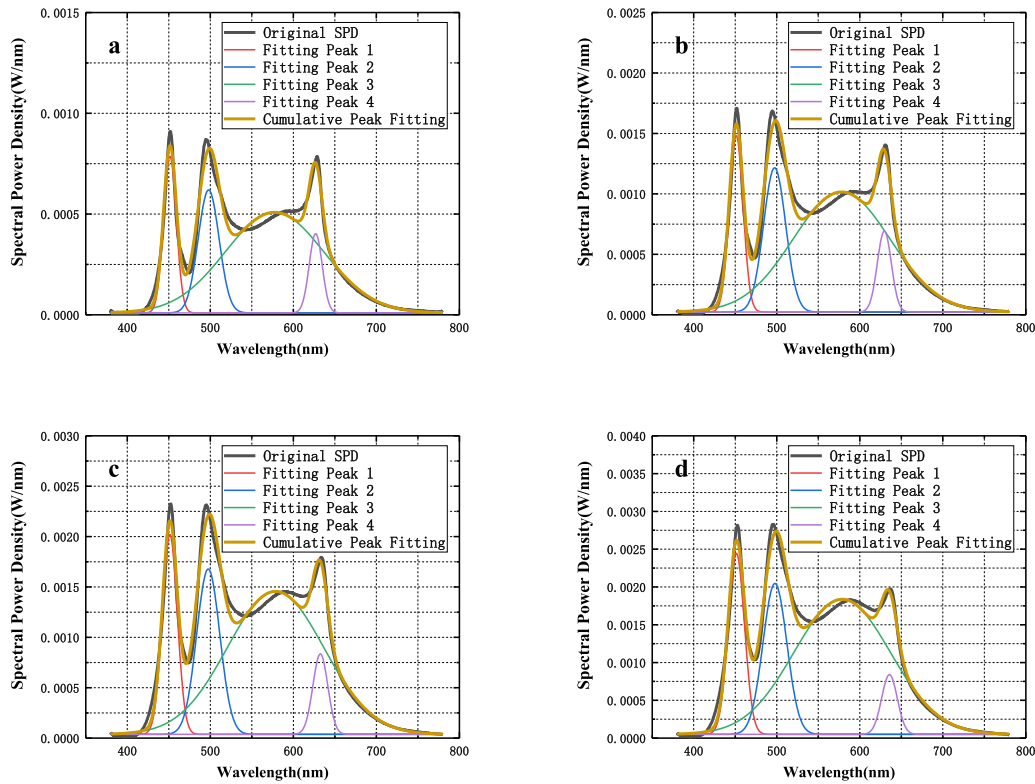


Fig. 6. SPD decomposition modeling with the Gaussian function fitting.

0.98, which indicates that the Gaussian based SPD decomposition model used in this study is appropriate to represent the original SPD of the full-spectrum white LED.

#### 4.2 SPD Prediction With BP-NN

According to above analysis, it can be seen that the SPD of the full-spectrum LED package is highly controlled by the case temperature and driven current. The neural network modeling in this paper is implemented by MATLAB. As the junction temperature ( $T_j$ ) is not effectively monitored during operation, the case temperature ( $T$ ) and driven current ( $I$ ) are selected as the inputs of the BP-NN model used in this study. And the outputs are the characteristic parameters extracted from the Gaussian function fitting, such as  $y_0$ ,  $x_{c1}$ ,  $x_{c2}$ ,  $x_{c3}$ ,  $x_{c4}$ ,  $w_1$ ,  $w_2$ ,  $w_3$ ,  $w_4$ ,  $A_1$ ,  $A_2$ ,  $A_3$ ,  $A_4$ . BP-NN is used to predict each output parameters and each network has two input neurons and one output neurons. Eq. (2) is used to calculate the number of hidden layers of the network.

$$q = \sqrt{n + m} + a \quad (2)$$

Where  $q$  is the number of hidden layers,  $n$  is the number of input neurons,  $m$  is the number of output neurons and  $a$  is the adjustment constant between 1 and 10.  $q = 7$  is chosen in this paper. The designed neural network structure is shown in Fig. 7.

Next, we arrange the 372 sets of collected SPD data to validate the BP-NN in SPD prediction, in which 352 sets are randomly selected as training set through the above network shown in Fig. 7. The remaining 20 sets of collected SPD data are used in testing. The gradient descent method is used to learn the neural network. The learning rate is usually 0.01 or 0.1, here, 0.01 is chosen. And the maximum number of network training is set as 1000. Table 4 shows the input data (case temperature and driven current) of the randomly selected test samples used in prediction. It can



TABLE 3  
Goodness-of-Fit Results ( $R^2$ )

Case temperature(°C) \ Driven current(mA)	30	40	50	60	70	80
125	0.9871	0.9879	0.98867	0.98946	0.99008	0.99078
130	0.98726	0.98801	0.98885	0.98947	0.9902	0.99081
135	0.98736	0.9881	0.98887	0.98958	0.9902	0.99088
140	0.98742	0.9882	0.98891	0.98967	0.99026	0.99091
145	0.98756	0.98833	0.98906	0.98976	0.99039	0.99101
150	0.9876	0.98832	0.98913	0.98979	0.99035	0.99105
155	0.9877	0.9885	0.98921	0.98985	0.99046	0.99107
160	0.98781	0.98852	0.98925	0.98994	0.99053	0.99116
165	0.98787	0.9886	0.98933	0.98998	0.99057	0.99125
170	0.98802	0.98865	0.98934	0.99008	0.99064	0.99124

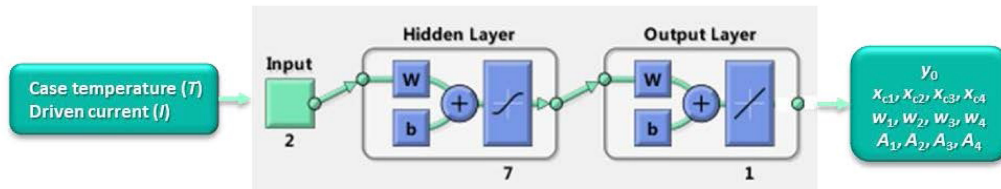


Fig. 7. Neural network structure designed for this study.

be seen from the table that the test samples used in prediction has a good discreteness, covering a variety of case temperatures and driven currents.

Furthermore, run the network and get the prediction result. In order to evaluate the prediction accuracy, the percentage of prediction error described in Eq. 3 is used, in which  $E_p$  is the percentage of error,  $R$  represents the actual characteristic parameters of Gaussian function and  $P$  are the predicted values by using the NN.

$$E_p = \left( \frac{R - P}{P} \right) \times 100\% \quad (3)$$

Fig. 8 shows the absolute percentage of prediction error, which reveals that all maximum prediction errors can be controlled under 5%, and the averaged values are lower than 1.2%. Except for  $A_4$  and  $y_0$ , the average error percentage of other parameters is less than 0.5%. This

TABLE 4

The Input Data of the Randomly Selected Test Scenarios Used in Prediction

Scenarios	Case Temperature(°C)	Driven Current (mA)
1	60	135
2	75	165
3	35	140
4	55	180
5	40	170
6	35	170
7	50	120
8	75	100
9	25	50
10	75	145
11	40	60
12	60	60
13	75	175
14	55	70
15	80	160
16	50	190
17	35	145
18	55	80
19	30	180
20	80	50

result demonstrates that the BP-NN can achieve a good prediction accuracy for the characteristic parameters of SPD. In order to illustrate the SPD prediction more intuitively, we plot the measured SPD, the Gaussian function fitting model and the predicted SPD by BP-NN in Fig. 9 with four sets of conditions as examples. It can be seen that the predicted SPDs by the BP-NN coincide with those modeled by the Gaussian function. To see the two curves more clearly, we bold the SPD predicted by the Gaussian function. Meanwhile, according to the fitting coefficient  $R^2 > 0.98$  obtained in Section 4.1, it also proves that the BP-NN has high accuracy of predicting the SPD for the full-spectrum LED package.

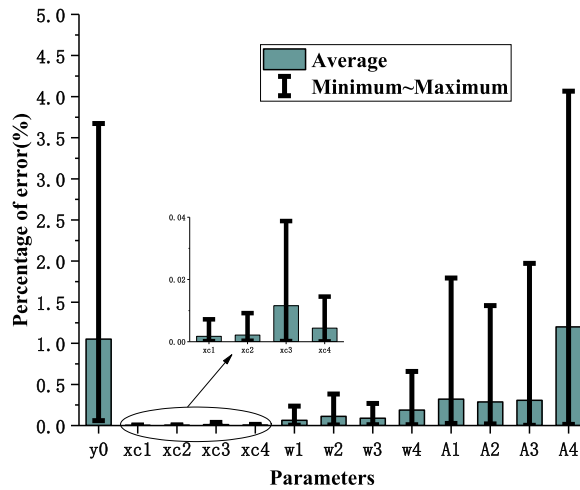


Fig. 8. Prediction errors of characteristic parameters in Gaussian model with BP-NN.

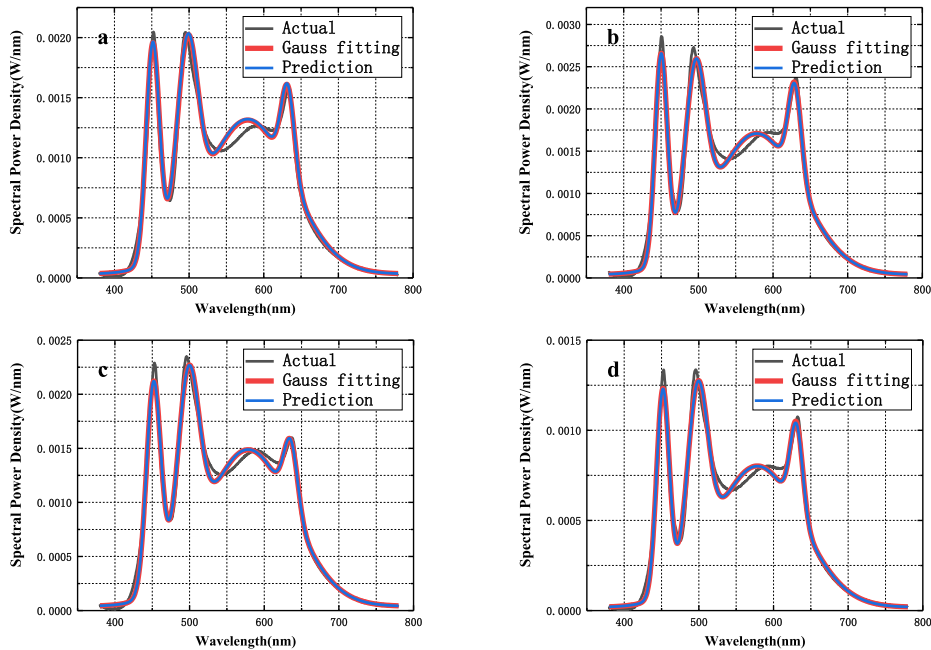


Fig. 9. SPD prediction results with BP-NN at conditions of (a) 60 °C, 135 mA; (b) 35 °C, 170 mA; (c) 80 °C, 160 mA; (d) 55 °C, 80 mA.

In order to furtherly illustrate the accuracy of predicting SPD, we calculate the Root-Mean-Square-Error (RMSE) and chromaticity difference ( $\Delta xy$ ) between predicted SPD and actual SPD, those can be expressed as Eq. (4) and Eq. (5).

$$RMSE = \sqrt{\frac{\sum (I_m(\lambda_j) - I_e(\lambda_j))^2}{n}} \tag{4}$$

TABLE 5  
The Prediction RMSE and Chromaticity Difference

Scenarios	Case Temperature(°C)	Driven Current (mA)	$\Delta xy$	RMSE( $10^{-5}$ )
1	60	135	0.001726268	8.4751
2	75	165	0.001910497	6.8895
3	35	140	0.001897367	7.6785
4	55	180	0.001649242	8.3135
5	40	170	0.001941649	8.7253
6	35	170	0.001931321	8.9524
7	50	120	0.00235372	6.2582
8	75	100	0.001964688	4.5966
9	25	50	0.003736308	3.1781
10	75	145	0.001627882	6.2430
11	40	60	0.002325941	3.5189
12	60	60	0.002370654	3.1913
13	75	175	0.001726268	7.2192
14	55	70	0.0018	3.7493
15	80	160	0.00210	6.52
16	50	190	0.002061553	8.9151
17	35	145	0.001708801	7.9041
18	55	80	0.002024846	4.2328
19	30	180	0.001923538	9.5946
20	80	50	0.002983287	2.3835
<b>Average values</b>	-	-	0.0020883	6.3270

where  $I_m(\lambda_j)$  is real measured SPD,  $I_e(\lambda_j)$  is predicted value of SPD, and  $n$  is number of measurements.

$$\Delta xy = \sqrt{(x_m - x_e)^2 + (y_m - y_e)^2} \quad (5)$$

where  $x_m$ ,  $x_e$  and  $y_m$ ,  $y_e$  are respectively chromaticity coordinates calculated from measured and expected value at CIE1931 color space.

As shown in Table 5, the averaged RMSE and chromaticity difference are  $6.33 \times 10^{-5}$  and 0.0021, respectively, that furtherly confirms the high prediction accuracy of the proposed BP-NN method.

#### 4.3 SPD Prediction With GA-BP NN

In order to improve the SPD prediction accuracy, we integrate the GA with BP-NN to optimize the weights and thresholds of the BP-NN. In the GA-BP NN, the number of iterations of the GA, the population size and the crossover probability are set as 20, 10 and 0.4 separately. The network

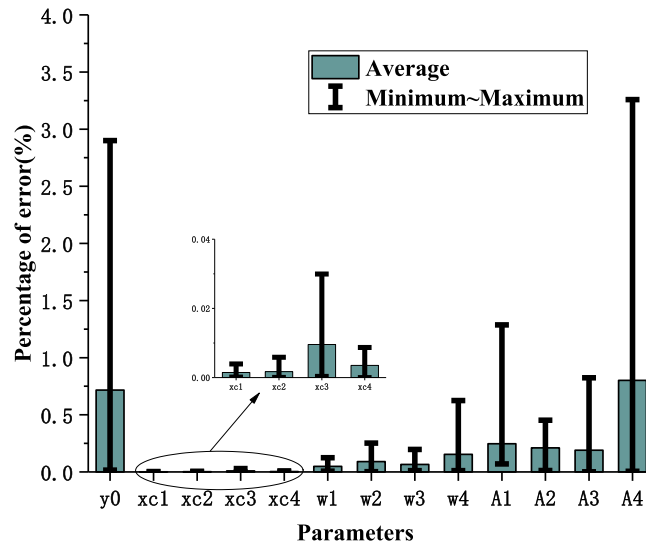


Fig. 10. Prediction errors of characteristic parameters in Gaussian model with GA-BP NN.

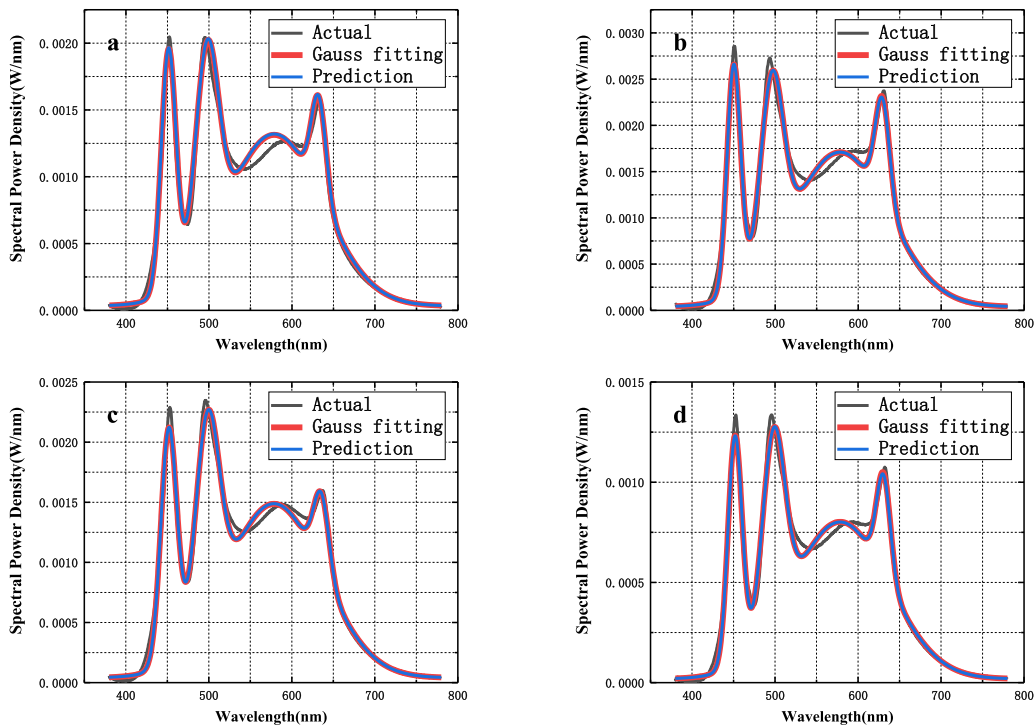


Fig. 11. SPD prediction results with GA-BP NN at conditions of (a) 60 °C, 135 mA; (b) 35 °C, 170 mA; (c) 80 °C, 160 mA; (d) 55 °C, 80 mA.

structure is the same as Fig. 7. The prediction results of characteristic parameters in Gaussian model with GA-BP NN are shown in Fig. 10, which presents that the maximum error percentage is 3.3%, the maximum average error percentage is 0.8%. Comparing the prediction results in the Fig. 8 and Fig. 10, it can be seen that both the average and maximum values of prediction error percentage with the GA-BP NN are lower than those of BP-NN. This is because the GA-BP NN

TABLE 6  
The Prediction RMSE and Chromaticity Difference

Scenarios	Case Temperature(°C)	Driven Current (mA)	$\Delta xy$	RMSE( $10^{-5}$ )
1	60	135	0.001649242	7.2719
2	75	165	0.001910497	6.8893
3	35	140	0.001649242	7.6646
4	55	180	0.001552417	8.3125
5	40	170	0.001941649	8.7277
6	35	170	0.001931321	8.9545
7	50	120	0.002024846	6.2621
8	75	100	0.002039608	4.5859
9	25	50	0.002308679	3.1484
10	75	145	0.001627882	6.2399
11	40	60	0.002376973	3.5204
12	60	60	0.002325941	3.1810
13	75	175	0.001627882	7.2197
14	55	70	0.001900000	3.7499
15	80	160	0.001910497	6.5228
16	50	190	0.001868154	8.9132
17	35	145	0.001676305	7.9000
18	55	80	0.002024846	4.2330
19	30	180	0.001772005	9.6012
20	80	50	0.001708801	2.3253
Average values	-	-	0.001891339	6.2612

reduces the blindly searching process at the early stage of BP-NN implement and find the optimal weights and thresholds. However, when the prediction error is reduced, the program running time is increased because of the more 20 iterations implemented in GA optimization. The running time of BP-NN is 4s, while the running time of GA-BP NN is 23s.

In order to illustrate the SPD prediction more intuitively, we also compare the measured SPD, the Gaussian fitting model and the predicted one by the GA-BP NN in Fig. 11 with four sets of conditions as examples.

Similarly, we can also get the RMSE and chromaticity difference. As shown in Table 6, the averaged RMSE and  $\Delta xy$  are  $6.26 \times 10^{-5}$  and 0.0019, respectively. According to the above analysis, it is known that the percentage of error, root mean square error, and chrominance difference predicted by GA-BP NN are smaller than those predicted by BP-NN. It can be explained that GA-BP NN has played an optimization role to a certain extent.

#### 4.4 Robustness Study of the Methods

To evaluate the universal applicability of the proposed methods, the robustness study with one two additional cases is presented and discussed in this section. Both cases are based on the BP-NN predictions.

TABLE 7  
Neural Network Training Dataset in Case 1

Parameters	Initial values	Termination values	Change gradients
Case temperature (°C)	25	65	5
Red LED's driven current (mA)	10	40	1
Cyan-blue LEDs' driven current (mA)	50	200	5

TABLE 8  
Neural Network Test Dataset in Case 1

Parameters	Interval one	Interval two	Change gradients
Case temperature (°C)	70	75	-
Cyan-blue LEDs' driven current (mA)	75~120	125~170	5

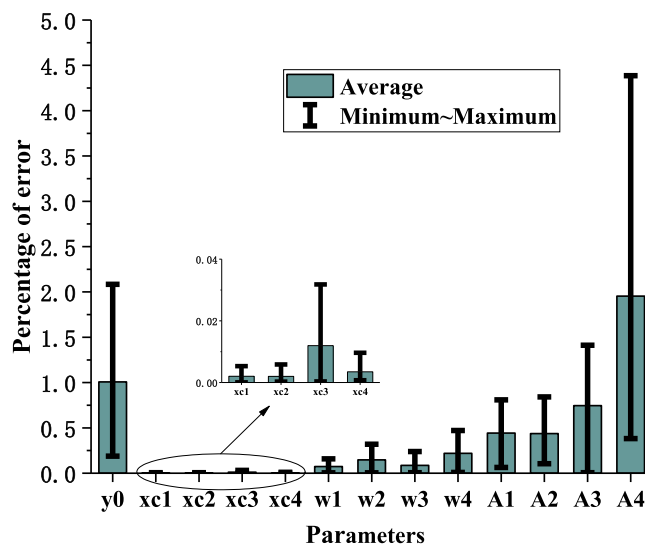


Fig. 12. Prediction errors of characteristic parameters in Gaussian model with BP-NN in Case 1.

**4.4.1 Case 1: Predict the Data Outside the Experimental Measurement:** In order to verify that the method can still predict the data outside the experimental dataset, we select some data to prove it. The prediction method is chosen in the same way as in Section 4.2. Select the input training set case temperature range from 25 °C to 65 °C, a total of 279 sets of training data. Used to predict the SPD with case temperatures of 70 °C and 75 °C, a total of 20 sets prediction data. The specific training set and test set input data are shown in Tables 7 and 8.

Fig. 12 shows the absolute percentage of prediction error, which reveals that all maximum prediction errors can be controlled under 5%. Table 9 shows RMSE and  $\Delta xy$ , from which we can see that the average RMSE and  $\Delta xy$  are  $5.50 \times 10^{-5}$  and 0.0018, respectively. This result proves that this method can still achieve good prediction for data outside the test data interval.

**4.4.2 Case 2 Prediction With Small Amount of Training Data:** In order to prove that the method can still achieve prediction when the training data is small, we choose 48 sets of data to prove. Table 10 shows their drive current and case temperature, of which sets of data as the training set and the remaining 6 sets as the test set. The input data of the test set is shown in Table 11. The prediction method is chosen in the same way as in Section 4.2.

TABLE 9  
RMSE and Chromaticity Difference in Case 1

Scenarios	Case Temperature(°C)	Driven Current (mA)	$\Delta xy$	RMSE( $10^{-5}$ )
1	70	75	0.00210238	3.6440
2	70	80	0.00212132	3.8500
3	70	85	0.002022375	4.0942
4	70	90	0.00210238	4.2847
5	70	95	0.001910497	4.5013
6	70	100	0.00190263	4.7039
7	70	105	0.001711724	4.9495
8	70	110	0.001868154	5.1692
9	70	115	0.001992486	5.3530
10	70	120	0.001772005	5.5148
11	75	125	0.0017	5.5829
12	75	130	0.001655295	5.7806
13	75	135	0.001886796	6.0299
14	75	140	0.001664332	6.2004
15	75	145	0.001565248	6.3228
16	75	150	0.001486607	6.4628
17	75	155	0.001615549	6.6442
18	75	160	0.001552417	6.8301
19	75	165	0.001581139	6.9568
20	75	170	0.001523155	7.1587
Average values	-	-	0.001786824	5.5017

TABLE 10  
Neural Network Training Dataset in Case 2

Parameters	Initial values	Termination values	Change gradients
Case temperature (°C)	30	80	10
Red LED's driven current (mA)	10	38	4
Cyan-blue LEDs' driven current (mA)	50	190	20

Fig. 13 shows the absolute percentage of prediction error, which shows that all maximum prediction errors can be controlled under 6%, and the averaged are lower than 2.2%. Table 12 shows RMSE and  $\Delta xy$ , from which we can see that the average RMSE and  $\Delta xy$  are  $5.50 \times 10^{-5}$  and 0.0018, respectively. This result proves that this method can still achieve prediction when the training data is small. Accuracy of prediction depends on nature and number of samples. This shows that the prediction accuracy has a relationship with nature of LED and amount of test data.



TABLE 11  
The Input Data of the Randomly Selected Test Scenarios Used in Case 2

Scenarios	Case Temperature(°C)	Driven Current (mA)
1	70	170
2	40	130
3	40	50
4	50	110
5	40	70
6	40	110

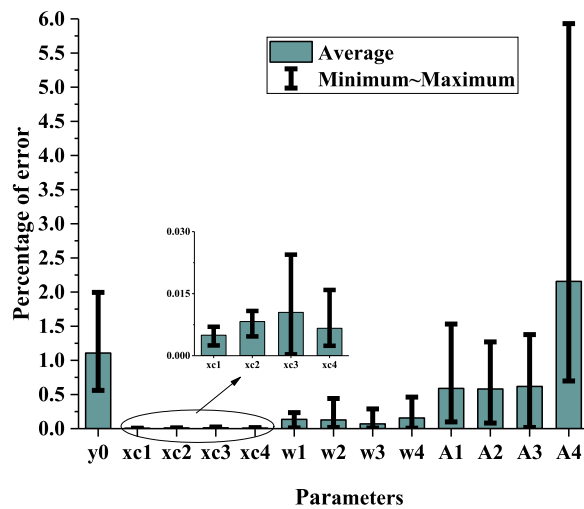


Fig. 13. Prediction errors of characteristic parameters in Gaussian model with BP-NN in Case 2.

TABLE 12  
RMSE and Chromaticity Difference in Case 2

Scenarios	Case Temperature(°C)	Driven Current (mA)	$\Delta xy$	RMSE( $10^{-5}$ )
1	70	170	0.001802776	7.294645302
2	40	130	0.002088061	7.058605549
3	40	50	0.00261725	2.992878148
4	50	110	0.002024846	5.778648403
5	40	70	0.002109502	4.122237567
6	40	110	0.002002498	6.125870549
Average values	-	-	0.002107489	5.562147587

## 5. Conclusion

This paper proposes an SPD prediction method for a full-spectrum white LED package by combining the SPD decomposition model with ANN base machine learning method. The Gaussian based SPD decomposition model is firstly proposed to discretize the continuous SPD of a full-spectrum white LED package as many spectral characteristic parameters. Then case temperature and driven current as inputs in both BP-NN and GA-BP NN models are used to estimate the spectral characteristic parameters, and finally a highly accurate SPD prediction is achieved and validated with different case studies. The prediction results show that: (1) The averaged SPD prediction RMSE and chromaticity difference ( $\Delta xy$ ) in both BP-NN and GA BP-NN can be controlled at around  $6 \times 10^{-5}$  and 0.002 respectively. (2) The robustness study also proofs the effective potential of our proposed method applied in dynamically accurate SPD prediction for the full-spectrum white LED operated under different mission profiles.

## Disclosures

The authors declare no conflicts of interest.

---

## References

- [1] L. Wei, Y. S. Chen, Q. Q. Guo, Y. H. Li, and K. J. He, "Discussion on development strategy of healthy lighting industry," in *Proc. 14th China Int. Forum Solid State Lighting: Int. Forum Wide Bandgap Semicond. China (SSLChina: IFWS)*, 2017, pp. 117–120.
- [2] J. G. Liu, W. Tang, C. Shen, Y. Ke, and G. Sun, "Advances in higher color quality and healthier white LEDs," in *Proc. 14th China Int. Forum Solid State Lighting: Int. Forum Wide Bandgap Semicond. China (SSLChina: IFWS)*, 2017, pp. 13–16.
- [3] W. Hertog, A. Llenas, J. M. Quintero, C. E. Hunt, and J. Carreras, "Energy efficiency and color quality limits in artificial light sources emulating natural illumination," *Opt. Express*, vol. 22, no. S7, pp. A1659–A1668, 2014.
- [4] C. Liang, J. F. Xu, Y. B. Fu, X. M. Teng, K. Liu, and B. Wu, "Study on phosphor schemes in full spectrum WLED," in *Proc. 15th China Int. Forum Solid State Lighting: Int. Forum Wide Bandgap Semicond. China (SSLChina: IFWS)*, 2018, pp. 1–3.
- [5] R. Dubnička, A. Rusnák, and P. Janiga, "Change of the spectral power distribution of luminaires under various electrical conditions," in *Proc. 16th Int. Scientific Conf. Electric Power Eng. (EPE)*, 2015, pp. 555–559.
- [6] J. C. C. Lo, S. W. R. Lee, X. G. Guo, and H. S. Zhao, "Numerical prediction and experimental validation of multiple phosphor white LED spectrum," in *Proc. Int. Conf. Electron. Pack. (ICEP)*, 2016, pp. 57–61.
- [7] P. Dupuis, E. Purwanto, N. I. Sinisuka, and G. Zissis, "LED spectrum optimal modelization," in *Proc. IEEE Industry Appl. Soc. Annu. Meeting (IAS)*, 2018, pp. 1–12.
- [8] B. M. Song and B. Han, "Spectral power distribution deconvolution scheme for phosphor-converted white light-emitting diode using multiple Gaussian functions," *Appl. Opt.*, vol. 52, no. 5, pp. 1016–24, 2013.
- [9] J. J. Fan, C. H. Yu, and C. Qian, "Thermal/luminescence characterization and degradation mechanism analysis on phosphor-converted white LED chip scale packages," *Microelectronics Rel.*, vol. 74, pp. 179–185, 2017.
- [10] M. E. Raypah, M. Devarajan, and F. Sulaiman, "Modeling spectra of low-power SMD LEDs as a function of ambient temperature," *IEEE Trans. Electron. Devices*, vol. 64, no. 3, pp. 1180–1186, Mar. 2017.
- [11] H. Jin, S. Jin, K. Yuan, and S. Cen, "Two-part gauss simulation of phosphor-coated LED," *IEEE Photon. J.*, vol. 5, no. 4, Aug. 2013, Art no. 1600110.
- [12] J. J. Fan, M. G. Mohamed, C. Qian, X. J. Fan, G. Q. Zhang, and M. Pecht, "Color shift failure prediction for phosphor-converted white LEDs by modeling features of spectral power distribution with a nonlinear filter approach," *Materials*, vol. 10, no. 7, p. 819, 2017.
- [13] C. Qian, J. J. Fan, X. J. Fan, and G. Q. Zhang, "Prediction of lumen depreciation and color shift for phosphor-converted white light-emitting diodes based on a spectral power distribution analysis method," *IEEE Access*, vol. 5, pp. 24054–24061, 2017.
- [14] H. T. Chen and S. Y. Hui, "Dynamic prediction of correlated color temperature and color rendering index of phosphor-coated white light-emitting diodes," *IEEE Trans. Ind. Electron.*, vol. 61, no. 2, pp. 784–797, Feb. 2014.
- [15] M. H. Chang, C. Chen, D. Das, and M. Pecht, "Anomaly detection of light-emitting diodes using the similarity-Based metric test," *IEEE Trans. Ind. Informat.*, vol. 10, no. 3, pp. 1852–1863, Aug. 2014.
- [16] T. M. Mitchell, *Machine Learning (Engineering/Math)*. New York, NY, USA: McGraw-Hill, pp. 81–126, 1997.
- [17] M. T. Hagan, H. B. Demuth, and M. Beale, *Neural Network Design*. Boston: PWS Publishing Co., 1996.
- [18] A. Géron, *Hands-On Machine Learning with Scikit-Learn and TensorFlow*. California, CA, USA: O'Reilly Media, CA, USA: 2017, pp. 253–273.
- [19] D. L. Hudson and Maurice E. Cohen, *Neural Networks and Artificial Intelligence for Biomedical Engineering*. Hoboken, NJ, USA: Wiley-IEEE Press, 2000, pp. 13–56.

- [20] Y. C. Chen, Y. C. Chang, W. C. Ke, and H. W. Chiu, "Cancer adjuvant chemotherapy strategic classification by artificial neural network with gene expression data: An example for non-small cell lung cancer," *J. Biomed. Informat.*, vol. 56, pp. 1–7, 2015.
- [21] A. Esteva, B. Kuprel, R. A. Novoa, J. Ko, S. M. Swetter, H. M. Blau, and S. Thru, "Dermatologist-level classification of skin cancer with deep neural networks," *Nature*, vol. 542, no. 7639, pp. 115–118, 2017.
- [22] M. Wenning, N. R. Büchl, and S. Scherer, "Species and strain identification of lactic acid bacteria using FTIR spectroscopy and artificial neural networks," *J. Biophotonics*, vol. 3, no. 8-9, pp. 493–505, 2010.
- [23] P. E. Larsen, D. Field, and J. A. Gilbert, "Predicting bacterial community assemblages using an artificial neural network approach," *Nature Methods*, vol. 1260, no. 6, pp. 33–43, 2012.
- [24] P. Bärtschi, C. Galloni, C. Lange, and B. Kilminster, "Reconstruction of  $\tau$  lepton pair invariant mass using an artificial neural network," *Nuclear Instruments Methods Phys. Res. Section A: Accelerators, Spectrometers, Detectors Associated Equip.*, vol. 929, pp. 29–33, 2019.
- [25] I. Ashdown, "Neural networks for LED color control," *Proc. SPIE - Int. Soc. Opt. Eng.*, vol. 5187, pp. 215–226, 2008.
- [26] K. Y. Lu, W. J. Zhang, and B. Sun, "Multidimensional data-driven life prediction method for white LEDs based on BP-NN and improved-adaboost algorithm," *IEEE Access*, vol. 5, pp. 21660–21668, 2017.
- [27] Q. Y. Yan, "Reliability prediction of the high-powered LED based on dynamic neural network," in *Proc. First Int. Conf. Electron. Instrum. Inf. Syst. (EIIIS)*, 2017, pp. 1–4.
- [28] H. W. Liu *et al.*, "Multi-Physical field analysis of multi-chip LED light source based on finite element method and artificial neural network," *Chin. J. Lumin.*, vol. 40, no. 06, pp. 795–802, 2019.
- [29] S. F. Ding and C. Y. Su, "Application of optimizing BP neural networks algorithm based on genetic algorithm," in *Proc. 29th Chin. Control Conf.*, 2010, pp. 2425–2428.
- [30] M. Eremia, C. C. Liu, and A. A. Edris, *Advanced Solutions in Power Systems: HVDC, FACTS, and Artificial Intelligence*. Hoboken, NJ, USA: Wiley, 2016, pp. 755–781.
- [31] M. D. Vose, *The Simple Genetic Algorithm*. Cambridge, Ma: Mit Press, 1999, vol. 1, pp. 31–57.
- [32] L. Liu, H. J. Qi, and J. Liu, "The research of alphabet identification based on genetic BP neural network," in *Proc. 4th Int. Conf. Intell. Human-Mach. Syst. Cybern.*, 2012, pp. 25–28.
- [33] J. W. Gao, Z. W. Leng, B. Zhang, G. Q. Cai, and X. Liu, "Short-Term forecasting of traffic flow based on genetic algorithm and BP neural network," *Lecture Notes Elect. Eng.*, vol. 254, pp. 745–752, 2013.
- [34] E. F. Schubert, *Light-Emitting Diodes*, Second edition. Cambridge, U.K.: Cambridge University Press, 2012, pp. 334–335.