Sensors and Materials, Vol. 34, No. 7 (2022) 2503–2515
MYU Tokyo

S & M 2986

IC Package Warpage Reduction Based on Fuzzy Adaptive Particle Swarm Optimization Algorithm and Neural Network

Te-Jen Su, Wen-Rong Yang, Yu-Chenge Lee, and Yi-Feng Chen*

Department of Electronic Engineering, National Kaohsiung University of Sciences and Technology,
415, Chien-Kung Rd., Sanming Distr. Kaohsiung City 80778, Taiwan

(Received Decembar 30, 2021; accepted April 27, 2022; online published May 23, 2022)

Keywords: IC warpage, fuzzy-PSO, neural network

The warpage of ICs in IC packaging manufacturing causes the production of defective ICs that can short-circuit or malfunction, including those in sensor devices. Applicable research results that predict IC warpage using a neural network have not been many, although many technologies have been proposed to prevent the warpage. It is necessary to understand the properties of IC materials as each material has a different coefficient of thermal expansion (CTE) for predicting the occurrence of the warpage. To provide a means to predict the warpage, a neural network with fuzzy adaptive particle swarm optimization (FAPSO) is proposed in this study based on the proposed architecture of the neural network and the defined weights of each layer in the IC. As the three layers of epoxy molding compound (EMC), die, and substrate (SBT) in IC packaging have different CTEs, nine conditional variables, namely, die thickness, glass transition temperature (T_g), CTEs ($\alpha 1$, $\alpha 2$), filler size, filler content, total height, post mold cure (PMC) temperature, and PMC time, are defined for predicting the warpage, and their parameters are found for training the neural network. In the comparison of the actual and predicted data of the neural network with FAPSO, the correlation coefficient is 0.9878, and the similarity between the two data sets is 99.7% in training. After the training, the validation is carried out for six data sets, the result of which shows that the correlation coefficient (R^2) is 0.8658 and the mean absolute percentage error (MAPE) is 29.74%, which is acceptable for applying the proposed neural network. The result of this study helps to improve the IC packaging process by preventing the warpage.

1. Introduction

3C products (computer, communication, and consumer electronics) are found everywhere in our daily life, offering many conveniences. The increased use of mobile devices demands a rapid electronic product manufacturing process. Important components of electronic products are the CPU, memory, power management device, sensor device, communication module, IC, and so on. Among them, the IC is critical as it amplifies signals from the components and plays an

*Corresponding author: e-mail: 1104405107@nkust.edu.tw
https://doi.org/10.18494/SAM3820

ISSN 0914-4935

important role as switches to control their functions. It also affects the overall reliability of the components' operation.

In manufacturing, the IC is soldered to the circuit board of the product, which is part of the surface mount technology (SMT). SMT refers to a technique by which electronic and electrical components are mounted on the surface of a printed circuit board (PCB). In SMT, high-temperature treatment is required to solidify the solder paste, that is, reflow soldering. The operating temperature reaches 260 °C, which causes the ICs to warp owing to incomplete soldering. It is important to keep an appropriate distance between solder balls and the PCB as the distance affects the quality of soldering and the degree of warpage. Thus, how to prevent the warpage in the packaging process is important, as the warpage causes wastage of time, manpower, and money in SMT. Recently, many researchers have conducted studies on the simulation and improvement of IC warpage. Loh *et al.* defined failures for the IC warpage on the PCB as an open circuit and solder ridging.(1) Shen *et al.* indicated that the difference in the coefficient of thermal expansion (CTE) between the die and the PCB is the main cause of the reliability problem in the flip-chip.(2) Tang *et al.* proposed that package warpage largely depends on the CTE of the materials, cure shrinkage, and glass transition temperature (T_g), and suggested changing materials such as a flexible hardener to optimize the CTE to prevent the IC warpage.(3) Sun *et al.* stated that the degree of warpage can be calculated according to the Timoshenko–Ehrenfest beam theory. However, they indicated that the actual occurrence is much more complicated than with the theory, so it is difficult to quantify the effects of design and materials on warpage.(4) Tzeng *et al.* proposed a smaller IC package size, a smaller solder ball, a thicker die, a thicker basic substrate, and a lower CTE in manufacturing plastic ball grid arrays (PBGAs) than in the conventional production design.(5) Tan *et al.* proposed that the filler of the sealant resin has a significant effect on warpage.(6) They claimed that plastic molds with a thicker thin fine-pitch ball grid array (TFBGA) improve warpage and solder ball stress. Thus, the higher the filler content, the more the warpages occur, whereas the warpages at different temperatures are not much different. Lin *et al.* considered that the effects of special additives on warpage are different and found that PCB design improves the degree of IC warpage by adjusting the thickness of the copper layer for example.(7) Wang *et al.* proposed that the die size is the main key factor, and a larger die size causes a greater warpage during thermal cycling in the eight-die package.(8) Thus, they suggested the asymmetrical material design of the lower layer on the PCB. Liu *et al.* studied the relationship between the different CTEs and mold thicknesses to simulate different warpages.(9) Ramirez *et al.* used different thicknesses and amounts of silicon dies to observe the behavior of the curvature for different packaging design parameters.(10) Kim *et al.* found that the built-in stress of copper, fiber-reinforced polymer (FRP), and solder resist (SR) affects the warpage of the IC package substrate at room temperature.(11) Wei *et al.* proposed reducing the CTE of the epoxy molding compound (EMC) and increasing the CTE of the core to effectively reduce product warpage.(12) Moran *et al.* used finite element analysis (FEM) coupled with direct optimization to reduce warpage in an AMB IC package.(13) Tan *et al.* proposed that during molding and post mold cure (PMC) processes, EMC will undergo polymerization conversion and chemical aging that will change its constitutive behavior and thus affect package warpage.(14)

On the basis of the previous research results, the IC warpage is considered to be caused by the different properties of materials such as glass transition temperature and CTE in IC packaging. Therefore, to prevent the IC warpage, it is necessary to have an accurate prediction model by considering the different properties of materials. The current method of analyzing and simulating the IC warpage is mainly to use FEM to simulate warpage, as the structure of an IC is complex and the influence factor has a nonlinear relationship with warpage, which is not a simple mathematical operation. Thus, we tried to use a neural network and a fuzzy particle swarm optimization algorithm to predict the degree of IC warpage after training the neural network with experimental data.

Nowadays, as the product development cycle becomes rapidly shortened, the warpage needs to be prevented to keep up with the speed of the product development. Therefore, the proposed algorithm contributes to the improvement of the product development of mobile devices. The algorithm is also expected to be applied to other manufacturing processes.

2. Methods

2.1 Neural network

A neural network is an artificial intelligence (AI) algorithm and helps to solve complex and nonlinear mathematical problems through training based on the neural network architecture. Figure 1 shows the architecture of a neural network with three layers, namely, an input layer, a hidden layer, and an output layer. The layers are connected by weights that are adjusted in the training of a neural network through the error back-propagation method. This method effectively defines the accurate weights, although it has a disadvantage of the local optimal solution.[15] In this study, the fuzzy particle swarm optimization algorithm is used to train the neural network for the optimal weights.

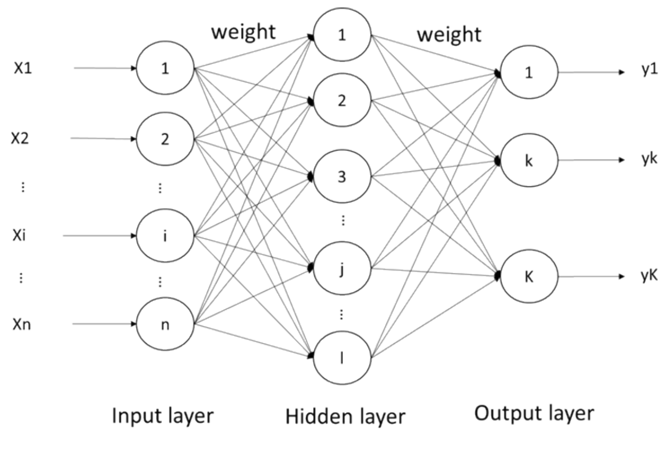


Fig. 1. Neural network architecture.

2.2 Fuzzy adaptive particle swarm optimization (FAPSO)

Optimization refers to finding the best solution to solve a complex problem. There are algorithms based on mimicking the biological behaviors of animals, such as the genetic algorithm,[16] immune algorithm,[17] ant colony optimization algorithm,[18] and particle swarm optimization (PSO) developed by Kennedy and Eberhart.[19] PSO is the most widely used among them, as the overall architecture is for simulating the collective foraging of birds or fish. It considers that the best solution is to find the exact location of food by a flock of birds or a school of fish in a certain space. All birds and fish do not know where the food is but in which area they currently are. The best strategy is to optimize the simplest and most effective way to reach food, as the whereabouts of food are searched and shared by birds closer to the food. Then, birds or fish move in the given direction to the food. While searching, an individual bird or fish adjusts its path according to the past and group experience, which is continuously updated until the best path is found. The mathematical expression of PSO is as follows.

$$v_i(t+1) = wv_i + c_1 r_1 \left(P_{ibest}(t) - x_i(t)\right) + c_2 r_2 \left(G_{ibest}(t) - x_i(t)\right) \tag{1}$$

$$x_i(t+1) = x_i(t) + v_i(t+1) \tag{2}$$

Here, v is the velocity of a single particle (bird or fish), x is the position of a single particle, i = 1, 2, …, n, n is the total number of particles, t is the number of iterations, c_1 and c_2 are the acceleration constants, usually between 0 and 4, r_1 and r_2 are independent random numbers between 0 and 1, P_{ibest} is the best position of a single particle, and G_{ibest} is the current best position of all particles.

Equation (1) reveals that the inertia weight w is a fixed value, and each weight of iteration is fixed during the iteration of calculation, which affects the speed of convergence. It is possible to search the food and expand the search area with a better solution, which is not affected by the experience of an individual particle (P_{ibest}) and other particles (G_{ibest}). However, if the weight is too large for the iteration to end, the particles do not converge. Therefore, many researchers propose different methods to adjust the weights to solve this problem.

Shi and Eberhart proposed a fuzzy system to dynamically adapt to the weight of a PSO algorithm and to adjust the weight in a nonlinear manner.[20] The fuzzy system has two input variables, namely, the current best performance evaluation (CBPE) and the current weight w. Then, the output variable is the adjustment weight w'. To use CBPE as an input for a fuzzy system for various optimization problems, CBPE is converted into a standardized format. Then, the FAPSO algorithm becomes effective and the best for dynamic environments. The best result is obtained by FAPSO than by a linear decreasing method. Equation (3) describes how to obtain CBPE.

$$NCBPE = \frac{CBPE - CBPE_{min}}{CBPE_{max} - CBPE_{min}} \tag{3}$$

Table 1
Weight-adjusted fuzzy membership function variables.

Input	*NCBPE*	Left triangle	0	0.06
		Triangle	0.05	0.4
		Right triangle	0.3	1
	w	Left triangle	0.2	0.6
		Triangle	0.4	0.9
		Right triangle	0.6	1.1
Output	*w′*	Left triangle	−0.12	−0.02
		Triangle	−0.04	0.04
		Right triangle	0	0.05

Table 2
Fuzzy rule of FAPSO in this study.

NCBPE	Low	Medium	High
Low	Medium	Low	Low
Medium	High	Medium	Low
High	High	Medium	Low

Here, *NCBPE* is standardized *CBPE*, $CBPE_{min}$ is the current minimum *CBPE*, and $CBPE_{max}$ is the current maximum *CBPE*. The integration of *NCBPE*, $CBPE_{min}$, and $CBPE_{max}$ establishes fuzzy membership function variables as listed in Table 1.

On the basis of variables, a fuzzy rule is deduced as shown in Table 2. When *NCBPE* is low, *w* becomes low, and *w′* becomes medium.

2.3 Neural network with FAPSO

A neural network (Fig. 2) with the FAPSO algorithm has the structure shown in Fig. 3. The main structure is a neural network, and the PSO algorithm is used to adjust and optimize the weights of the neural network. The fuzzy theory is used to adjust the weights of the particles in the PSO algorithm in the connection layer.

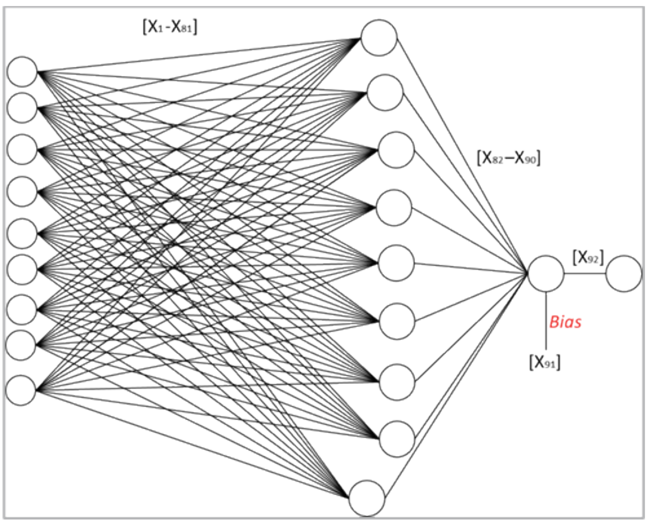


Fig. 2. (Color online) Neural network architecture design of this study.

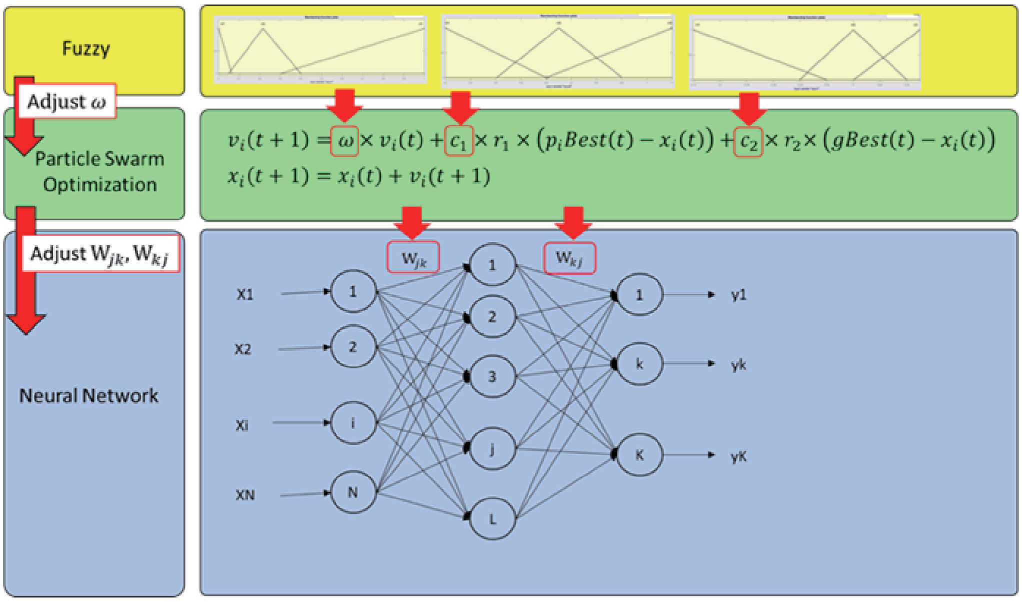


Fig. 3. (Color online) Architecture of neural network combined with fuzzy particle swarm optimization.

2.4 Neural network training

IC packaging is completed by bonding die and wires and sealing the layer of epoxy and the gold wire element (Fig. 4). The glue is a thermoset material that needs to be heated. Heating causes the warpage and shrinkage of the EMC and substrate (SBT) as shown in Fig. 5. Figure 6 shows the different shrinkage rates of each layer of the IC according to the different CTEs (α1 and α2) of the materials in the layers.

To find the appropriate amount of sealer (molding compound), the control IC warpage is necessary to test for various sealer volumes, the thickness of the sealant, and the grain size of the filler. The appropriate baking time and temperature also need to be found. Under the same sealant thickness and substrate material, a thicker die layer and a larger sealant volume cause a higher shrinkage rate after sealing in the PMC process, which causes warpage (Fig. 7).

Considering the factors that affect the occurance of warpage, the input data of the neural network, namely, CTE (α1), CTE (α2), filler content, and filler size (Figs. 8 and 9), are mainly for materials such as the sealant (T_g). Figure 10 shows the flow chart of how to train the neural network with FAPSO in this study. Sealing the IC package protects the die, gold wires, and other components after completing the die and wire bonding processes. As the sealant is a thermoset material, it needs to be heated.

Through the experiment, the conditional variables in Table 3 are defined for training the neural network. In training data processing, the above input data are converted so that the values are in the range from −1 to 1.

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

Fig. 4. (Color online) Layers of different materials in IC package.

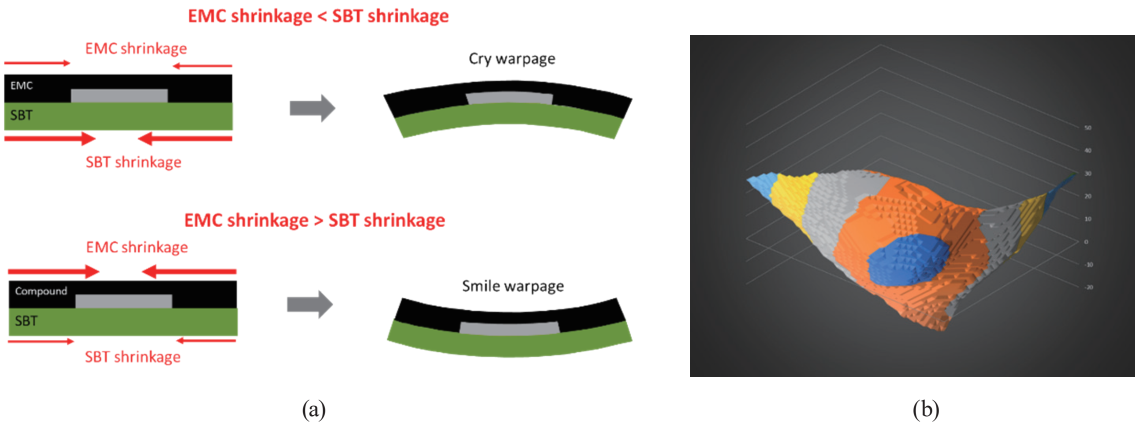


Fig. 5. (Color online) (a) IC warpage and shrinkage of layers with different materials. (b) 3D image of IC warpage.

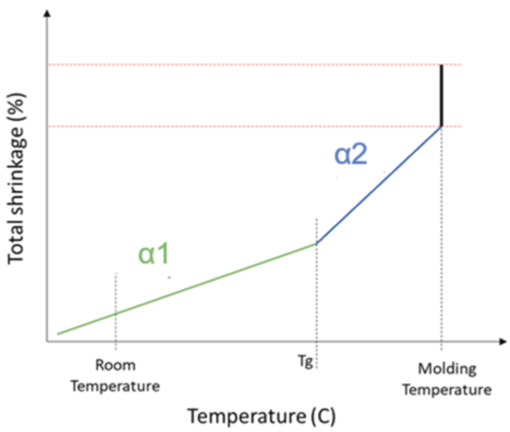


Fig. 6. (Color online) Shrinkage rates of materials with different CTEs at various temperatures.

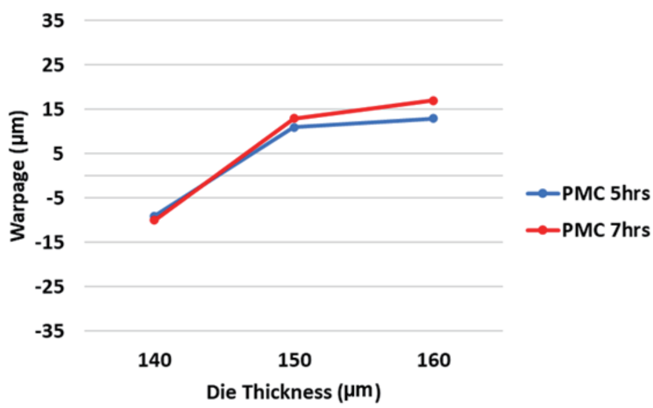


Fig. 7. (Color online) Warpage caused by different die thicknesses under different PMC conditions.

Here, Y_{max} is the maximum value of the target value, Y_{min} is the minimum value of the target value, X_{max} is the maximum value of the original value, X_{min} is the minimum value of the original value, X is the original value, and Y is the converted value.

As the value of the warpage is positive or negative, the activation function uses the *tansig* function with the value between −1 and 1 as shown in Eq. (5).

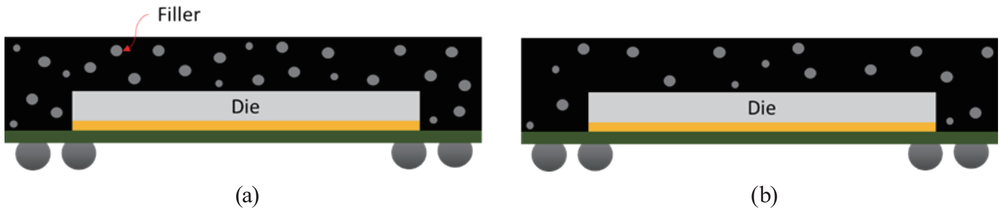


Fig. 8. (Color online) Filler content with different sealants. (a) More filler content. (b) Less filler content.

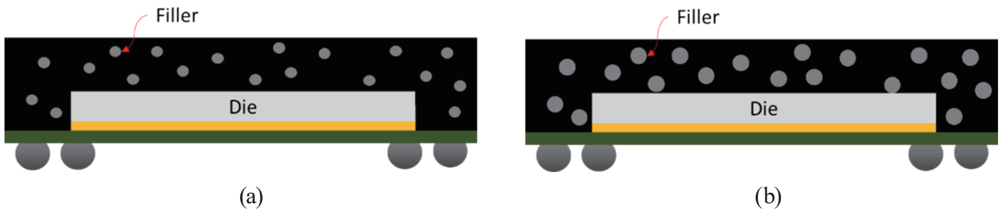


Fig. 9. (Color online) Filler size with different sealants. (a) Smaller filler size. (b) Larger filler size.

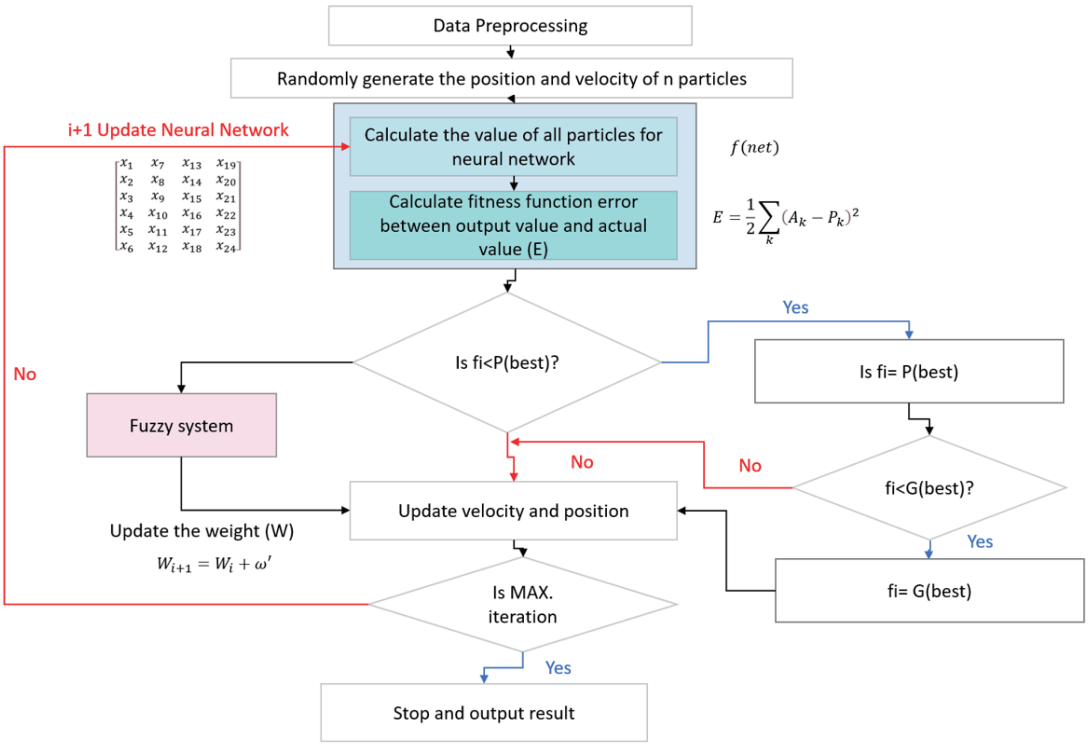


Fig. 10. (Color online) Flow chart of training the neural network with FAPSO.

$$tansig(x) = \frac{2}{1+e^{-2x}} - 1. \tag{5}$$

Then, the input and target value are converted into a value between −1 and 1 before being used for training the neural network. Finally, the parameters for training the neural network are set as shown in Table 4.

Table 3
Conditional variables for training the neural network.

Item	Value
Die thickness (μm)	60−80
T_g (°C)	125−167
CTE α1	8−9
CTE α2	30−40
Filler size (μm)	30−55
Filler content (%)	75−88
Total height (μm)	840−880
PMC temp. (°C)	175−185
PMC time (h)	1.75−7.50

Table 4
Parameters set by fuzzy particle swarm algorithm.

Item	Parameter setting
Number of variables	92
Maximum variable	4.2
Minimum variable	−4.2
Total number of particles	300
Number of iterations	120
c1	0.9
c2	1.1
Weight	1

A PSO algorithm has an adaptive function to calculate the value of each particle at the current position. When training the neural network, we set the adaptive function with the predicted values and their errors as shown in Eq. (6) to minimize the error value and make the target solution, $E = 0$.

$$E = \frac{1}{2}\sum_{i=1}^{k}\left(d_k - y_k\right)^2 \tag{6}$$

Here, d_k is the kth actual value and y_k is the kth value calculated using the neural network.

3. Results and Discussion

The FAPSO algorithm shows that the best cost decreases from 0.38 to 0.03 after 10 iterations, which clearly shows a convergence effect (Fig. 11). The weight is reduced from 1.0 to 0.4 after 10 iterations, so the initial weight is converged after the iteration (Fig. 12).

When comparing the actual and predicted data after training, the correlation coefficient (R^2) between them is 0.9878, which indicates that the neural network has significant accuracy (Fig. 13). In the regression equation, the slope $p1$ is 0.9968. Thus, the actual and predicted data show 99.7% similarity. Thus, the trained neural network can be applied to the experiment for validating the effectiveness of the neural network with FAPSO in this study.

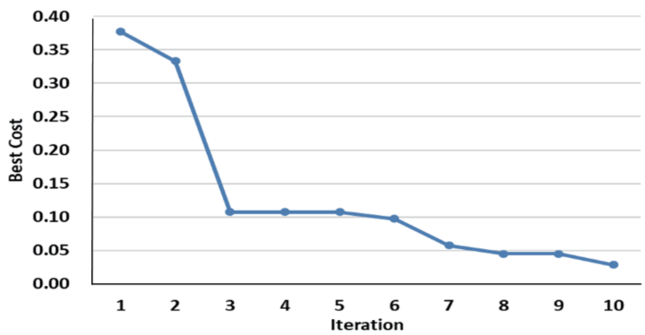


Fig. 11. (Color online) Iterative training result of FAPSO.

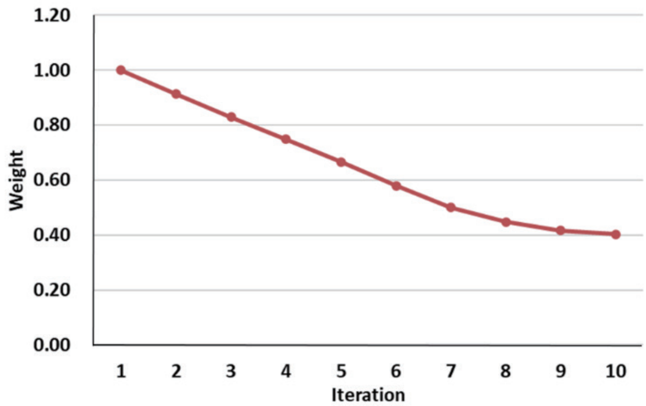
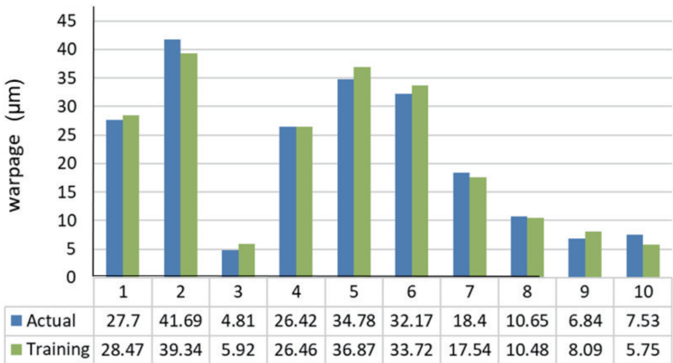
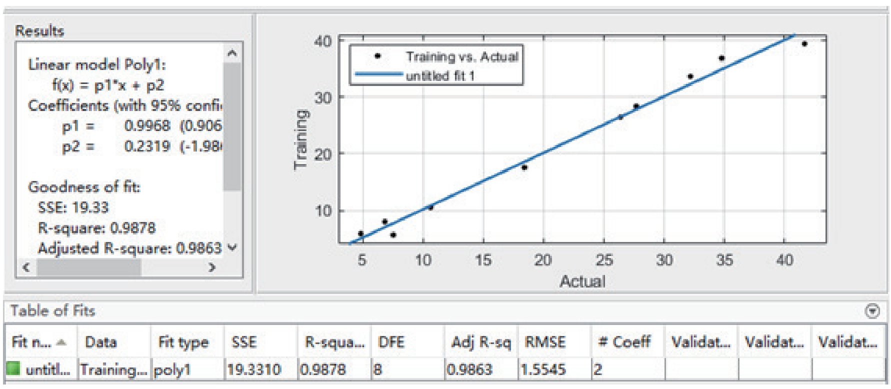


Fig. 12. (Color online) Weight changes by each iteration.



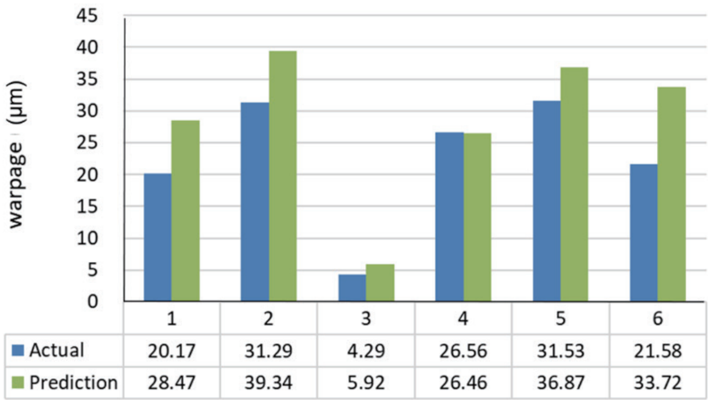
(a)



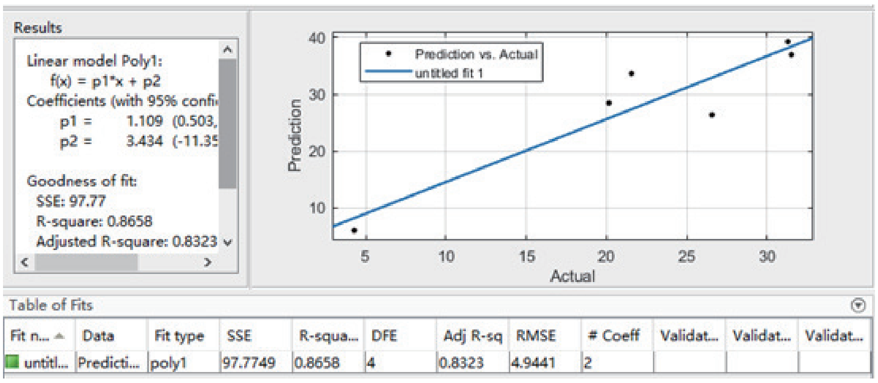
(b)

Fig. 13. (Color online) Comparison of actual and predicted data with regression analysis. (a) Actual and predicted data after training. (b) Regression analysis result of the comparison in (a).

After the neural network is trained, six data sets are used to validate the neural network. The verification results are shown in Fig. 14. When comparing the actual and predicted data for the validation data sets, the correlation coefficient (R^2) is 0.8658 and the slope $p1$ is 1.109. The result has more deviation between the actual and predicted data, but the mean absolute percentage



(a)



(b)

Fig. 14. (Color online) Comparison of actual and predicted data of six data sets for validation of neural network with regression analysis. (a) Actual and predicted data of six data sets for validation. (b) Regression analysis result of the comparison in (a).

error [MAPE, Eq. (7)] is 29.74%. Considering the accuracy in Table 5, the value predicted using the proposed neural network with FPOS is within an acceptable range.

$$MAPE = \frac{100\%}{n}\sum_{i=1}^{n}\left|\frac{y_i - \hat{y}_i}{y_i}\right| \quad (7)$$

Here, y_i is the actual value and $\hat{y}_i$ is the predicted value.

Table 5
Prediction accuracy by MAPE.

MAPE (%)	Accuracy
<10	Highly accurate forecasting
10–20	Good forecasting
20–50	Reasonable forecasting
>50	Inaccurate forecasting

4. Conclusions

To predict and prevent the IC warpage efficiently, a neural network with an FAPSO algorithm is proposed. FAPSO is used to predict the possible warpage in IC packaging for reducing the waste of time and resources due to the warpage. In this study, we establish the architecture of the neural network and define the weights of each layer. To optimize the weights, CBPE is used with FAPSO. Then, the architecture of the neural network combined with FAPSO is defined. Considering the three layers of EMC, die, and SBT in IC packaging and warpages due to different CTEs in each layer, we define the nine conditional variables to train the neural network, namely, die thickness, glass transition temperature (T_g), CTE ($\alpha 1$, $\alpha 2$), filler size, filler content, total height, PMC temperature, and PMC time. Then, the parameters are set for training the neural network. In training, 10 iterations (epochs) of training lead to the best cost and converged weights. In the comparison of the actual and predicted data in training, the correlation coefficient is 0.9878, and the similarity between the two data sets is 99.7%. After the training, FAPSO is applied to six data sets for its validation, and the result shows that the correlation coefficient (R^2) is 0.8658 and the MAPE is 29.74%, which is acceptable for applying the proposed FAPSO. The proposed FAPSO in this study provides a way of improving the IC packaging process and is applied to other manufacturing processes.

Acknowledgments

This research work was supported by the Ministry of Science and Technology, R.O.C., Grant No. MOST 110-2221-E-992-093.

References

1 W. K. Loh, R. Kulterman, H. Fu, and M. Tsuriya: Proc. 2016 Int. Conf. Electronics Packaging (ICEP) (IEEE, 2016) 89–93.
2 Y. Shen, L. Zhang, and X. Fan: Proc. 2015 IEEE 65th Electronic Components and Technology Conf. (ECTC) (IEEE, 2015) 1546–1552.
3 H. Tang, J. Nguyen, J. Zhang, and I. Chien: Proc. 2007 Int. Symp. High Density Packaging and Microsystem Integration (IEEE, 2007) 1–5.
4 W. Sun, W. Zhu, C. K. Wang, A. Sun, and H. Tan: Proc. EuroSimE 2008-Int. Conf. Thermal, Mechanical and Multi-Physics Simulation and Experiments in Microelectronics and Micro-Systems (IEEE, 2008) 1–8.
5 Y. L. Tzeng, N. Kao, E. Chen, J. Y. Lai, Y. P. Wang, and C. Hsiao: Proc. 2007 9th Electronics Packaging Technology Conf. (IEEE, 2007) 482–487.
6 W. Tan, F. Zhou, X. Cheng, D. Ding, and J. Wu: Proc. 2007 9th Electronics Packaging Technology Conf. (IEEE, 2007) 722–724.
7 W. Lin, B. Baloglu, and K. Stratton: Proc. 2014 IEEE 64th Electronic Components and Technology Conf. (ECTC) (IEEE, 2014) 1401–1406.
8 D. W. Wang, H. M. Huang, S. C. Ho, A. H. Liu, and D. S. Liu: Proc. 2010 60th Electronic Components and Technology Conf. (ECTC) (IEEE, 2010) 1968–1973.
9 S. W. Liu, C. H. Tsai, and K. N. Chiang: Proc. 2019 Int. Conf. Electronics Packaging (ICEP) (IEEE, 2019) 242–245.
10 J. L. Ramirez, R. T. Yoshioka, C. C. Nunes, I. F. Namba, and C. Coral: Proc. 2020 33rd Symp. Integrated Circuits and Systems Design (SBCCI) (IEEE, 2020) 1–5.
11 M. S. Kim, H. J. Jung, S. H. Park, T. H. Kim, and E. C. Ahn: Proc. 2020 IEEE 22nd Electronics Packaging Technology Conf. (EPTC) (IEEE, 2020) 73–76.

12 Q. Wei, C. Ting, W. Liu, N. Sun, Q. Fang, J. Liu, H. Yang, and X. J. Ma: Proc. 2021 22nd Int. Conf. Electronic Packaging Technology (ICEPT) (IEEE, 2021) 1–5.
13 R. L. Moran, E. Arriola, N. R. E. Lim, J. P. Mercado, R. Dimagiba, J. Gonzaga, and A. Ubando: Proc. 2019 IEEE 11th Int. Conf. Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment, and Management (HNICEM) (IEEE, 2019) 1–6.
14 L. Tan, J. Li, X. Cheng, Q. Wang, J. Cai, H. Shen, G. Yu, T. Zhang, and S. Wang: Proc. 2013 14th Int. Conf. Electronic Packaging Technology (IEEE, 2013) 343–346.
15 K. W. Chau: J. Hydrol. **329** (2006) 636. https://doi.org/10.1016/j.jhydrol.2006.02.025
16 H. Qiu, W. Zhou, and H. Wang: Proc. 2009 5th Int. Conf. Natural Computation (IEEE, 2009) 81–85.
17 W. Jiang, D. Luo, Y. Xu, and X. Sun: Proc. 5th World Congr. Intelligent Control and Automation (IEEE, 2004) 2122–2126.
18 B. Liu and P. Meng: Proc. 2008 9th Int. Conf. Young Computer Scientists (IEEE, 2008) 1819–1824.
19 J. Kennedy and R. Eberhart: Proc. ICNN'95–Int. Conf. Neural Networks (IEEE, 1995) 1942–1948.
20 Y. Shi and R. C. Eberhart: Proc. 2001 Congr. Evol. Comput. (IEEE, 2001) 101–106.