

# Optimal Design for a Ball Grid Array Wire Bonding Process Using a Neuro-Genetic Approach

Chao-Ton Su and Tai-Lin Chiang

**Abstract**—This study presents an integrated method in which neural networks, genetic algorithms, and exponential desirability functions are used to optimize the ball grid array (BGA) wire bonding process. As widely anticipated, the BGA package will become the fastest-growing semiconductor package and push integrated circuit (IC) packaging to higher level of compactness and density. However, wire bonding in BGA is difficult owing to its high input/output (I/O) count, fine pitch wire bonds, and long wire lengths. This study addresses two fundamental issues in the semiconductor assembly facility on its quest toward a defect-free manufacturing environment. First, the problem of exploring the nonlinear multivariate relationship between parameters and responses and second, obtaining the optimum operation parameters with respect to each response in which the process should operate. The implementation for the proposed method was carried out in an IC assembly factory in Taiwan; results in this study demonstrate the practicability of the proposed approach.

**Index Terms**—Ball grid array (BGA), exponential desirability function, genetic algorithms, neural networks, wire bonding.

## I. INTRODUCTION

**B**GA packages provide a high interconnect density and lead count using standard pitch dimensions. Wire bonding designs include ultra fine pitch and cavity-up, which conduct heat from the die through the substrate and interconnect. Owing to their intrinsic design, BGAs are technically complex to bond, and are designed for high I/O counts, i.e., up to 500 leads is common. They also demand fine-pitch ( $\leq 85 \mu\text{m}$ ) wire bonding and require long wire lengths, straight loops and small first and second bond areas [1]. With high I/O count, fine pitch wire bonds, and long wire lengths, wire bonding in the BGA assembly is difficult. Exploring a manufacturing solution for the BGA requires an integrated study for wire bonding parameters. Fig. 1 depicts the thermal design architecture of the BGA package.

The wire bonding process begins from targeting the capillary on the bond pad and positioned above the die with ball formed on the end of the wire and pressed against the face of the capillary. The capillary descends bringing the ball in contact with the die. The inside cone, or radius, grips the ball in forming the bond. In a thermosonic system, ultrasound vibration is then applied. After the ball is bonded to the die, the capillary raises

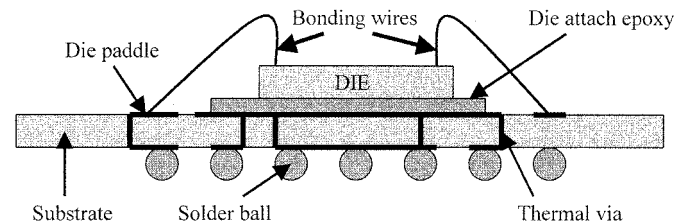


Fig. 1. Typical BGA thermal design.

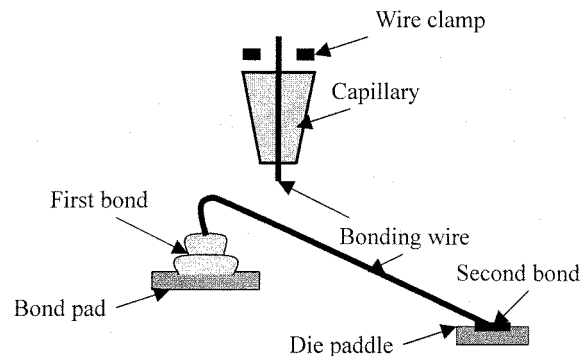


Fig. 2. Wire bonding process mechanism.

to the loop height position. The clamps are open and wire is free to feed out the end of the capillary. The lead of the device is positioned under the capillary, which is then lowered to the lead. Wire is fed out the end of the capillary, forming a loop. The capillary deforms the wire against the lead, producing a wedge-shaped bond, which has a gradual transition into the wire. In a thermosonic machine, ultrasonic vibration is then applied. The capillary raises off the lead. Leaving the stitch bond. At a pre-set height, the clamps are closed, while the capillary is still rising with the bonding lead. This prevents the wire from feeding out the capillary and pulls at the bond. The wire will break at the thinnest cross section of the bond. A new ball is formed on the tail of the wire, which extends from the end of the capillary. A hydrogen flame or an electronic spark may be used to form the ball. The cycle is completed and ready for the next ball bond. Fig. 2 depicts the mechanism of wire bonding.

Wire bonding is used throughout the semiconductor industry as a means of interconnecting the dies, substrates and I/O pins. Ultrasonic metal welding technology can be used for many different applications by appropriately utilizing its sound wave and high frequency mechanical energy characteristics. Ultrasonic energy is used to improve the structure of materials in metallurgy. The acoustic irradiation of molten mass improves degasification and the finer grain structure during the hardening process.

Manuscript received December 4, 2000; revised December 20, 2001. This paper was supported by the National Science Council of Taiwan, R.O.C., under Contract NSC-90-2218-E-009-008.

C.-T. Su is with the Department of Industrial Engineering and Management, National Chiao Tung University, Hsinchu, Taiwan, R.O.C.

T.-L. Chiang is with the Department of Business Administration, Minghsin Institute of Technology, Hsinchu, Taiwan, R.O.C.

Publisher Item Identifier S 1521-334X(02)02656-3.

The wire bonding operation attempts to develop a high yield interconnect and low wire sweep process with a sufficient long-term reliability to satisfy customers requirements [2]. To achieve a high level of wire bonding performance and quality, the appropriate bonding process parameters must be accurately identified and controlled. The task of the process engineers is to identify and control these parameters to obtain desired wire bonding quality for optimizing multiple responses (e.g., maximum ball shear strength, wire pull strength, and appropriate ball size), based on their experience or equipment provider's recommendations iteratively. However, this task is complicated and difficult due to coupled multivariable system, which makes it impossible to adjust a single parameter without affecting the others. Therefore, this multivariate operation requires an intelligent system capable of evaluating the process and determining the optimum adjustment [3].

The use of statistical experimental design techniques in semiconductor manufacturing has been proven very beneficial in process modeling, optimization and control. This approach in process has yielded fairly good empirical models for processes such as plasma etching and LPCVD [4]. However, statistical modeling in semiconductor manufacturing relies on response surface methods (RSM) to construct process models following experimentation. Himmel and May [5] demonstrated that RSM models are lack accurate and robust than models constructed using neural networks.

This study presents an integrated approach not only for exploring empirical models between process parameters and responses via neural networks, but also for optimizing the process through certain parameter settings using genetic algorithms and exponential desirability function for the BGA wire bonding process. A comparison through confirmatory trials between RSM and proposed approach with respect to each response is conducted as well.

The rest of this paper is organized as follows. Neural networks, genetic algorithms, and exponential desirability functions are briefly described is made. The next section presents an integrated procedure for optimizing the BGA wire bonding process. An experimental design for the implementation of proposed procedure is then illustrated, followed by a comparison of proposed procedure and RSM in terms of process performance. Concluding remarks are finally made.

## II. MODELING AND OPTIMIZATION APPROACH FOR BGA WIRE BONDING

### A. Neural Networks

Major progress in studying neural networks has been made since 1980. Neural networks are increasingly used to model complex manufacturing processes, generally for process and quality control [6]. Frequently these models are used to identify optimal process setting. An approximated model can be constructed using a back-propagation neural network. In addition, the output can be predicted by using statistical methods. However, according to Seet and Boullart's, such methods tend to be generally less accurate [7]. Neural networks possess the unique capability of learning arbitrary nonlinear mappings between noisy sets of input and output patterns [4]. Basically, a

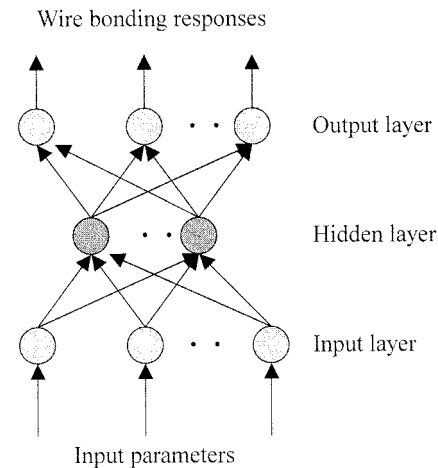


Fig. 3. Topology of the back-propagation neural network.

neural network approach can typically be constructed without assuming anything about the functional form of the relationship between predictors and response [8]. In addition to learning and extracting the process behavior from previous operating information. This approach can also be used as a model for process optimization. Neural networks have demonstrated strong capability of learning nonlinear and complex relationship between process parameters and responses without any prior knowledge regarding the process. The neural network approach holds a major advantage over the statistical method in that the neural network is explicitly nonlinear through hidden layers. It is a more general mapping procedure in which a specific function format is not required in model building [9]. This particularly fits the highly complex process of BGA wire bonding [10].

Neural networks have recently emerged as a highly promising alternative to physically based models and statistical methods of semiconductor process modeling. Fig. 3 displays the general structure of a feedforward, multilayer neural network used for semiconductor process modeling that is typically trained via back-propagation [4]. The back-propagation networks have already been applied to a wide range of problems (e.g., speech synthesis, and pattern recognition) and, in most cases, exhibit a good behavior and results [11]. Once trained, a back-propagation network can be evaluated quickly, which is an advantage during the optimization phase. Recent overviews of neural network applications in manufacturing industry were compiled by Zhang and Huang [12]. More related applications can also be found in ([3]–[7], [10], [13]–[17]).

The BP neural networks consist of layers of neurons interconnected such that information is stored in the weight assigned to the connections. Network learning aims to determine an appropriate set of connection strengths which facilitate the activation of these processing units to achieve a desired state that imitates a given set of sampled patterns. In addition, a sigmoid activation function determines the activation level of a neuron.

### B. Genetic Algorithms (GAs)

More conventional optimization methods start from one point in the search area and then move sequentially to achieve the optimum solution, thereby operating rather locally and highly prone

to falling inside a coincidental local optimum. GAs counteract entrapment in a local optimal solution to imitate the principles of natural genetics and natural selection to constitute search and optimization procedure. They perform a global, random, parallel search for an optimal solution using simple computations.

GAs are efficient local search methods based on natural selection and population genetics. These algorithms use randomized operators operating on a population of candidate solution to generate a new population of candidates in the search space [18]. Owing to that large dimensions are involved in the parameters-to-responses function and a mathematical formulation is unavailable, this study applies genetic algorithms, one of the promising approaches for optimizing the complicated production system. GAs are known for their robustness and effective overall search capabilities [19]. Huang and Adeli [15], Sette *et al.* [7], as well as Hsu and Su [20] have demonstrated the ability of Gas to perform an optimum search through GAs. A GA in its simplest form uses three operators: reproduction, crossover, and mutation.

### C. Exponential Desirability Function

It's not unusual to deal with multiresponses in a manufacturing process. Optimizing the process with respect to any single response often results in nonoptimum values for the remaining characteristics. A simple and intuitive approach to multiresponse problem is to superimpose the response contour plots and to determine an optimal solution by visual inspection. Such a method is severely limited by the number of input variables and/or responses [21]. The desirability function approach attempts to transform a multiresponse problem into a single response one by mathematical transformation [22]. Kim and Lin [21] develops an approach based on maximizing exponential desirability functions that do not require any assumptions regarding the form or degree of the estimated response models, such an approach is robust to the potential dependencies between response variables. Their approach aims to identify the settings of the input variables to maximize the degree of overall satisfaction with respect to all the responses. The exponential desirability function has been extensively used to simultaneously optimize several responses. The benefits of the exponential desirability are that they are easily understood, intuitively, and allow the user to weigh the response according to their importance.

In order to achieve an overall optimization with respect to all the responses, a "minimum" operator for aggregating the responses can be stated as

$$\underset{x}{\text{maximize}} \lambda \quad (1)$$

subject to

$$d_j\{\hat{y}(x)\} \geq \lambda, \quad j = 1, 2, \dots, r. \quad (2)$$

This formulation aims to identify  $x^*$  which maximizes the minimum degree of satisfaction ( $\lambda$ ) with respect to all the responses within the experimental region ( $\Omega$ ), i.e.,

$$\underset{x \in \Omega}{\text{maximize}} (\min\{d_1\{\hat{y}_1(x)\}, d_2\{\hat{y}_2(x)\}, \dots, d_r\{\hat{y}_r(x)\}\}). \quad (3)$$

The exponential desirability function can be formed as

$$d(z) = \begin{cases} \left( \frac{\exp(t) - \exp(t|z|)}{\exp(t) - 1} \right), & \text{if } t \neq 0 \\ 1 - |z|, & \text{if } t = 0 \end{cases} \quad (4)$$

where  $t$  is the exponential constant ( $-\infty < t < \infty$ ), and  $z$  is a standardized parameter representing the distance of the estimated responses from its target in units of the maximum allowable deviation. For example,  $z$  for a response with symmetric desirability function is defined as

$$z = \frac{\hat{y}_j(x) - T_j}{y_j^{\max} - T_j} = \frac{\hat{y}_j(x) - T_j}{T_j - y_j^{\min}}, \quad \text{for } y_j^{\min} \leq \hat{y}_j \leq y_j^{\max}. \quad (5)$$

Similarly, for the smaller-the-better (STB) type or larger-the-better (LTB) type response, the following transformations for the  $z$  value

$$z = \frac{\hat{y}_j(x) - y_j^{\min}}{y_j^{\max} - y_j^{\min}}, \quad \text{for } y_j^{\min} \leq \hat{y}_j \leq y_j^{\max} \quad (6)$$

$$z = \frac{y_j^{\max} - \hat{y}_j(x)}{y_j^{\max} - y_j^{\min}}, \quad \text{for } y_j^{\min} \leq \hat{y}_j \leq y_j^{\max}. \quad (7)$$

$z$  ranges between  $-1$  and  $+1$  for an NTB-type response and between  $0$  and  $1$  otherwise. In both cases the value of  $d(z)$  achieves its maximum value when  $z = 0$ . The function  $d(z)$  given in (4) has been proven to provide a reasonable and flexible representation of human perception and is convenient to handle analytically.

The exponential desirability function has several methodological advantages over the available methods in terms of optimizing multiresponse. First, the "maximin" approach is robust to the potential dependence between responses. Such dependence is extremely difficult to detect or model in practice. Second, this approach achieves a better balance between all the responses compared with the existing methods. Third, the objective function value allows a satisfactory physical interpretation in terms of degree of satisfaction. Fourth, the approach can also be viewed as a fuzzy logic approach. The "maximin" approach is equivalent to using the logical and operator in fuzzy logic, denoting the intersection of the corresponding membership functions. Related studies have demonstrated that this optimization scheme is quite effective in compromising multiple conflicting objectives [22]–[24].

### D. Proposed Optimization Procedure

This study proposes an integrated neuro-genetic-exponential desirability function algorithm capable of optimizing the parameter settings in a BGA wire bonding process. The proposed approach consists of two stages. The first stage procedure involves using of a BP network to derive the relationship model between input parameters and output responses. Notably, the trained network can accurately predict the behavior of possible parameter combinations. Thus, tuning the input parameters in the trained network allow us to obtain the corresponding responses. The exponential desirability is then used to transform the multiple responses into a single response. During the second stage, GA is applied to obtain the optimum degree of satisfaction ( $\lambda$ ). Herein, the chromosome is used to represent the possible solution. Each

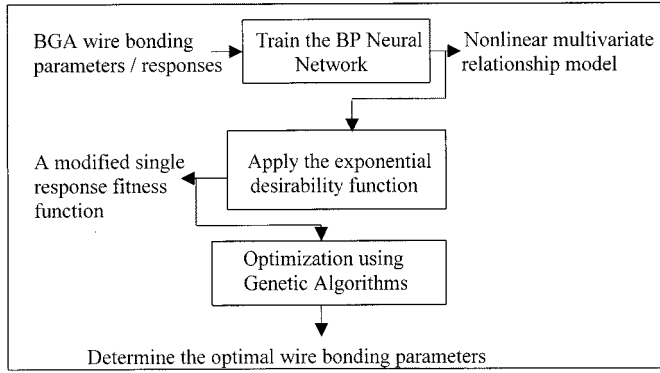


Fig. 4. Schematic diagram for determining the optimal wire bonding parameters.

gene in the chromosome represents the value of the input parameter. For example, a manufacturing process has three input parameters  $S$ ,  $T$ , and  $U$ . A chromosome can represent the value of the three parameters  $(S, T, U)$ , respectively. The essential genetic operators during the iterative procedure can be found in the previous section. These operations are conducted to obtain the optimal response, which is evaluated by the fitness function. Therefore, the optimal parameter of the problem can be obtained. Fig. 4 schematically depicts the proposed optimization procedure. The detailed procedure is summarized as follows.

- Step 1) Collect the input parameters and corresponding responses.
- Step 2) Develop a BP network model to obtain the relationship between the input parameters and output responses.
- Step 3) Apply the exponential desirability function to transform the multiple responses into a single one. The trained network with a modified single response is referred to as a fitness function.
- Step 4) Set the GA operating conditions (e.g., population size, generation size, parameter number, crossover rate, and mutation rate).
- Step 5) Create an initial population by randomly selecting the value of the input parameters.
- Step 6) Repeat steps 7-11 until the stopping condition is reached.
- Step 7) Calculate the fitness value by inputting the parameter values to the fitness function.
- Step 8) Select the parameter values according to the computed responses.
- Step 9) Crossover the fitness parameter values.
- Step 10) Mutate the parameter values to yield the next generation.
- Step 11) Obtain the current optimal parameter values.
- Step 12) Obtain the optimal parameter settings.

### III. EXPERIMENTAL RESULTS

#### A. Training of Neural Networks

An engineering experiment on the 52- $\mu\text{m}$  fine pitch BGA wire bonding process is conducted to optimize the wire bonding process with respect to each response, which is shown

TABLE I  
RESPONSES OF THE 52- $\mu\text{m}$  FINE PITCH BGA

Output Responses	Range
Wire Pull ( $y_1$ )	$\geq 4.0$ g
Ball Shear ( $y_2$ )	$\geq 10.0$ g
Ball Size ( $y_3$ )	$40 \pm 2$ $\mu\text{m}$

TABLE II  
PROCESS PARAMETERS AND THEIR LEVELS

Parameters	Levels	
	1	2
Temperature ( $x_1$ )	150	170
Force1 ( $x_2$ )	6	10
Pre-Bleed Ratio ( $x_3$ )	30	50
Contact Threshold ( $x_4$ )	40	70
CV1 ( $x_5$ )	0.2	0.4
Tip1 ( $x_6$ )	10	25
Time ( $x_7$ )	5	25
USG Mode1 ( $x_8$ )	60	80

TABLE III  
OPTIONS FOR NEURAL NETWORKS

Architecture	RMSE	
	Training	Testing
8-3-3	0.11221	0.12911
8-4-3	0.09431	0.10253
8-5-3	0.11082	0.12772
8-6-3	0.12676	0.14365
8-7-3	0.11883	0.13329

in Table I. Table II lists the process parameters and value for each level. Thirty-two trials are conducted by a well-structured orthogonal array  $L_{32}$ . The experimental data are then used for constructing the relationship model between parameters and responses through the BP neural network in which 80% (approximately 25 samples) are used for training the neural networks while the remaining 20% (approximately seven samples) are used for testing.

Table III lists several options of the neural network architecture; in which the structure 8-4-3 under the best convergence criterion of the root of mean square (RMSE) is selected to obtain a better performance.

#### B. Determination of the Fitness Function

In this study, responses  $y_1$ , and  $y_2$  have lower specifications and  $y_3$  has the corresponding target value. Herein, the exponential desirability function is used to solve the multi-response problem. We have

$$\lambda = \min(d_1, d_2, d_3) \quad (8)$$

where  $d_i$  is calculated from (3) and (4). The engineering management agrees on employing a convex, convex, and concave exponential desirability function for  $y_1, y_2$ , and  $y_3$  with  $t = -2, -1, 2$ , respectively, according to their importance. Therefore,  $\lambda$  is set as the fitness function of the GA as further explored in the next section.

TABLE IV  
IMPLEMENTATION RESULTS OF GA

Item	Data
The largest $\lambda$ value in 20 runs	0.8812
The smallest $\lambda$ value in 20 runs	0.7324
Average $\lambda$ value	0.8220
Standard deviation	0.0374

### C. Optimization Using Genetic Algorithms

Each input parameter in the BGA wire bonding process is normalized to the value between 0 and 1 and they are combined into one string. For example, the input parameters listed in Table II are transformed into the chromosome representation  $(x_1, x_2, x_3, \dots, x_8)$  in a string. Strings are randomly generated to form the initial population. When GA is applied to optimize the BGA wire bonding parameter selection, the essential operators, including reproduction, crossover and mutation should be determined in advance. Herein, a roulette wheel approach is adopted as the selection procedure. The crossover rate and mutation rates are set as 0.5 and 0.01, respectively. Fifty strings are randomly generated to establish the initial population. Notably, 5 000 generations were processed.

### D. Results

The above information is used and the GA is executed twenty runs. Table IV summarizes the implementation results. The higher value of  $\lambda$  implies a better degree of satisfaction in terms of compromised solution. The largest  $\lambda$  value is 0.8812 and its optimum chromosome is (167.1, 6.7, 36.7, 46.5, 0.21, 24.8, 5.9, 62.3). These settings are the optimal condition for the eight process parameters.

### E. Comparison

Conventionally, process engineers handle a multiresponse problem to apply the RSM, polynomial models fitted to each of the responses. They then superimpose the response contour plots to determine optimal parameter settings by overlaying of contour plots along with a separate response surface analysis. The best subset models and their  $R^2$ -values using RSM are

$$\begin{aligned} \hat{y}_1(x) &= 4.57 - 0.21x_1 + 0.27x_3 + 0.26x_7 \\ &\quad + 0.36x_8 - 0.22x_1x_8 \\ &\quad + 0.19x_2x_6 - 0.38x_3x_7 \quad (R^2 = 0.805) \\ \hat{y}_2(x) &= 11.58 + 0.56x_2 + 0.64x_3 + 1.24x_7 + 1.73x_8 \\ &\quad + 0.57x_1x_2 + 0.71x_2x_8 + 0.59x_3x_5 + 0.65x_4x_7 \\ &\quad + 0.6x_5x_7 + 0.62x_6x_7 \quad (R^2 = 0.874) \\ \hat{y}_3(x) &= 39.85 + 0.96x_1 + 1.85x_5 + 0.64x_1x_2 \\ &\quad + 0.84x_1x_3 - 0.89x_1x_8 \\ &\quad + 0.7x_2x_3 + 0.72x_3x_5 \quad (R^2 = 0.866). \end{aligned}$$

The optimal process settings can be obtained as values of (159.2, 8.2, 40.5, 51.8, 0.31, 21.4, 5.3, 67.1).

This study conducted a comparison between the RSM and the proposed approach for benchmarking purposes. According to the comparison on Table V, the proposed approach reveals better

TABLE V  
COMPARISON BETWEEN THE RSM AND THE PROPOSED APPROACH

Parameter	Results of the RSM approach	Results of the proposed approach
1. Optimal parameter settings:	(159.2, 8.2, 40.5, 51.8, 0.31, 21.4, 5.3, 67.1)	(167.1, 6.7, 36.7, 46.5, 0.21, 24.8, 5.9, 62.3)
2. Sample size of confirmation trials:	21	21
3. Mean and variance of response $(\bar{x}, s)$ :		
$y_1$	(5.9, 0.42)	(6.5, 0.45)
$y_2$	(13.4, 0.78)	(14.1, 0.80)
$y_3$	(39.8, 0.39)	(40.1, 0.40)
4. Degree of satisfaction $\lambda$ :	0.823	0.846
5. P-value of t-test for:		
$H_0: \mu_{y1/\text{proposed approach}} \leq \mu_{y1/\text{RSM}}$		*0.0001 (T= 4.26)
$H_0: \mu_{y2/\text{proposed approach}} \leq \mu_{y2/\text{RSM}}$		*0.0009 (T= 3.35)
$H_0: \mu_{y3/\text{RSM}} = \mu_{\text{Target}}$	0.051(T=-3.15)	----
$H_0: \mu_{y3/\text{proposed approach}} = \mu_{\text{Target}}$	----	0.09 (T=1.78)
6. Short term process capability:		
$y_1$	1.51	1.67
$y_2$	1.45	1.61
$y_3$	1.54	1.58

Note: the numbers shown in bold denote statistically significant to reject the null hypothesis at 5% levels of significance.

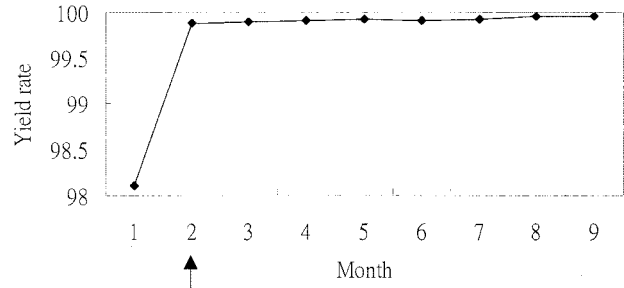


Fig. 5. Yield rate trend.

performance more than 10% on the wire pull and ball shear in terms of short term process capability. Because the values of wire pull and ball shear are the most important characteristics which will highly affect the electrical function of the device in later applications, these results are highly satisfied the line engineers.

This paper was also employed the  $t$ -tests of the mean values for the wire pull and ball shear between the two approaches, respectively, the  $t$  statistics are 4.26 with a  $P$  value of 0.0001 and 3.35 with a  $P$  value of 0.0009. Thus, there are strong evidences to indicate that the means for wire pull and ball shear by the proposed approach are greater than the means by the RSM approach.

The effectiveness of the proposed approach is conducted at a semiconductor assembly line in Taiwan that was undertaken to optimize the BGA wire bonding parameters. The implementation results under mass production over eight months confirm that the proposed approach outperforms the conventional RSM method in optimizing a BGA wire bonding process. According to the quality trend chart (Fig. 5) from the shop floor, the wire bonding yield has been risen to an average around 99.92% over

eight months from 98.1% which equivalents to a reduction of 18 200 DPPM (defect parts per million). The annual cost saving is expected to exceed 1.1 million US dollars from implementing the proposed approach in Month 2, whereas the expenditure for the experiment was below USD 2,000.

#### IV. CONCLUSION

This study has demonstrated that integrating BP neural networks, genetic algorithms, and exponential desirability function can optimize the BGA wire bonding process. Although statistical experimental design techniques in semiconductor manufacturing have greatly benefited in process modeling, optimization and control, statistical modeling that heavily relies on response surface methods (RSM) to construct process models following experimentation are less accurate and robust than neural networks models. The neural network approach is better than statistical method largely owing to that the neural network is explicitly nonlinear through hidden layers. It is a more general mapping procedure in which a specific function format is not required in model building. This study also demonstrated the superiority of the proposed approach over RSM base on criteria such as degree of satisfaction ( $\lambda$ ), testing the difference about two population means, and short term process capability. The proposed approach can easily achieve optimization of the complex process with multiple responses. These settings facilitate process engineers in achieving acceptable process control during the production. In addition, the improvement in process performance allows the factory to more easily fabricate products with superior quality in the IC assembly industry.

#### ACKNOWLEDGMENT

The authors would like to thank Process Engineering Manager D. Hsiao, AdvanTech, Ltd., for his full support, and B. Thomas, for permitting the use of the BGA technical data in this study.

#### REFERENCES

- [1] C. Leroy, L. Lee, and M. Eshelman, "Bonding BGA packages requires integrated solutions," in *Proc. IMAPS Conf.*, Philadelphia, PA, Oct. 1997.
- [2] M. Kuzawinski, "High density package applications for wire bond and flip chip: Small, fine pitch BGA packages," in *Proc. Semiconduct. Packag. Symp.-Session I: SEMICON W.*, 1999.
- [3] K. M. Tay and C. Butler, "Modeling and optimizing of a MIG welding process-A case study using experimental designs and neural networks," *Qual. Rel. Eng. Int.*, vol. 13, pp. 61-70, 1997.
- [4] K. K. Lee, T. Brown, G. Dagnall, R. Bicknell-Tassius, A. Brown, and G. S. May, "Using neural networks to construct model of the molecular beam epitaxy process," *IEEE Trans. Semiconduct. Manufact.*, vol. 13, pp. 34-45, Feb. 2000.
- [5] C. Himmel and G. May, "Advantages of plasma etch modeling using neural networks over statistical techniques," *IEEE Trans. Semiconduct. Manufact.*, vol. 6, pp. 103-111, May 1993.
- [6] D. W. Coit, B. T. Jackson, and A. E. Smith, "Static neural network process model: Considerations and cases studies," *Int. J. Prod. Res.*, vol. 36, no. 11, pp. 2953-2967, 1998.
- [7] S. Sette, L. Boullart, L. V. Langenhove, and P. Kiekens, "Optimizing the fiber-to-yarn production process with a combined neural network/genetic algorithm approach," *Textile Res. J.*, vol. 67, no. 2, pp. 84-92, 1997.
- [8] H. S. Stern, "Neural networks in applied statistics," *Technometrics*, vol. 38, no. 3, pp. 205-220, Aug. 1996.

- [9] C. A. Chang and C.-T. Su, "A comparison of statistical regression and neural network methods in modeling measurement errors for computer vision inspection systems," *Comput. Ind. Eng.*, vol. 28, no. 3, pp. 593-603, 1995.
- [10] J. Chen, P. T. P. Chu, D. Shan, H. Wong, and S. S. Jang, "Optimal design using neural network and information analysis in plasma etching," *J. Vac. Sci. Technol. B*, vol. 17, no. 1, pp. 145-153, Jan./Feb. 1999.
- [11] R. P. Lipmann, "An introduction to computing with neural nets," *ASSP*, vol. 40, 1987.
- [12] H. C. Zhang and S. H. Huang, "Applications of neural networks in manufacturing: A state-of-the-art survey," *Int. J. Prod. Res.*, vol. 33, pp. 705-728, 1995.
- [13] C. C. Chiu, C. T. Su, G. S. Yang, J. S. Huang, S. C. Chen, and T. N. Cheng, "Selection of optimal parameters in gas-assisted injection moulding using neural network model and the Taguchi method," *Int. J. Qual. Sci.*, vol. 2, no. 2, pp. 106-120, 1997.
- [14] F. L. Chen and S. F. Liu, "A neural-network approach to recognize defect spatial pattern in semiconductor fabrication," *IEEE Trans. Semiconduct. Manufact.*, vol. 13, pp. 366-373, Aug. 2000.
- [15] S. L. Hung and H. Adeli, "A parallel genetic/neural network learning algorithm for MIMD shared memory machines," *IEEE Trans. Neural Networks*, vol. 5, pp. 900-909, Nov. 1994.
- [16] T. K. Meng and C. Butler, "Solving multiple response optimization problems using adaptive neural networks," *Int. J. Adv. Manufact. Technol.*, vol. 13, pp. 666-675, 1997.
- [17] S. Y. S. Lam, K. L. Petri, and A. E. Smith, "Prediction and optimization of a ceramic casting process using a hierarchical hybrid system of neural networks and fuzzy logic," *IIE Trans.*, vol. 32, pp. 83-91, 2000.
- [18] D. E. Goldberg, *Genetic Algorithms in Search, Optimization, and Machine Learning*. Reading, MA: Addison-Wesley, 1989.
- [19] Z. Khan, B. Prasad, and T. Singh, "Machining condition optimization by genetic algorithms and simulated annealing," *Comput. Oper. Res.*, vol. 24, no. 7, pp. 647-657, 1997.
- [20] C.-M. Hsu and C.-T. Su, "Multiobjective machine-component grouping in cellular manufacturing: A genetic algorithm," *Prod. Planning Contr.*, vol. 9, no. 2, pp. 155-166, 1998.
- [21] K. J. Kim and K. J. D. Lin, "Simultaneous optimization of mechanical properties of steel by maximizing exponential desirability functions," *Appl. Stat.*, pt. 3, vol. 49, pp. 311-325, 2000.
- [22] E. Castillo, D. C. Montgomery, and D. R. Maccarville, "Modified desirability functions for multiple response optimization," *J. Qual. Technol.*, vol. 28, no. 3, pp. 337-345, July 1996.
- [23] R. Bellman and L. Zadeh, "Decision-making in a fuzzy environment," *Manag. Sci.*, vol. 17, pp. 141-164, 1970.
- [24] M. Laviolette, J. Seaman, J. Barrett, and W. Woodall, "A probabilistic and statistical view of fuzzy methods," *Technometrics*, vol. 37, pp. 249-261, 1995.



**Chao-Ton Su** received the Ph.D. degree in industrial engineering from the University of Missouri, Columbia.

He is a Professor in the Department of Industrial Engineering and Management, National Chiao-Tung University, Hsinchu, Taiwan, R.O.C. His research is in the areas of quality management and neural networks in industrial applications.



**Tai-Lin Chiang** received the Ph.D. degree in industrial engineering and management from National Chiao-Tung University, Hsinchu, Taiwan, R.O.C.

He is currently a Lecturer in the Department of Business Administration, Mingshin Institute of Technology, Hsinchu. His research interests include quality engineering, neural networks applications, and semiconductor manufacturing engineering.