Chapter 14

Artificial Neural Networks in Production Scheduling and Yield Prediction of Semiconductor Wafer Fabrication System

Jie Zhang, Junliang Wang and Wei Qin

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/63444

Abstract

With the development of artificial intelligence, the artificial neural networks (ANN) are widely used in the control, decision-making and prediction of complex discrete event manufacturing systems. Wafer fabrication is one of the most complicated and high competence manufacturing phases. The production scheduling and yield prediction are two critical issues in the operation of semiconductor wafer fabrication system (SWFS). This chapter proposed two fuzzy neural networks for the production rescheduling strategy decision and the die yield prediction. Firstly, a fuzzy neural network (FNN)-based rescheduling decision model is implemented, which can rapidly choose an optimized rescheduling strategy to schedule the semiconductor wafer fabrication lines according to the current system disturbances. The experimental results demonstrate the effectiveness of proposed FNN-based rescheduling decision mechanism approach over the alternatives (back-propagation neural network and Multivariate regression). Secondly, a novel fuzzy neural network-based yield prediction model is proposed to improve prediction accuracy of die yield in which the impact factors of yield and critical electrical test parameters are considered simultaneously and are taken as independent variables. The comparison experiment verifies the proposed yield prediction method improves on three traditional yield prediction methods with respect to prediction accuracy.

Keywords: semiconductor wafer fabrication system, rescheduling, fuzzy neural networks, yield prediction, decision mechanism

1. The production scheduling and yield prediction of semiconductor wafer fabrication system (SWFS)

The semiconductor wafer fabrication system (SWFS) is one of the most sophisticated manufacturing systems. This kind of manufacture system is characterised by a different type of wafer process (batch and single process), hundreds of process steps, the large and expensive device, production unforeseen circumstances and re-entrant flow [1]. Semiconductor manufacturing orders are usually global, dynamic and customer driven since the 1990s. As a result, semiconductor manufacturers strive to achieve high-quality products using advance manufacturing technologies (such as process planning and scheduling and digitized indicators' prediction technologies) [2]. In recent years, production scheduling and yield prediction are always two issues above all in the complex SWFS.

An organization's competitive advantage is increasingly dependent on its response to market changes and opportunities, and in response to unforeseen circumstances (i.e. Machine breakdown, rush orders), so it is important to reduce inventory and cycle time, and improve resource utilization. Therefore, production scheduling is required to optimize the operation of SWFS and has been reviewed by Uzsoy and his colleagues [3]. SWFS operates in uncertain dynamic environments, facing with a lot of disturbances, such as machine failure, a lot of rework and rush orders [4]. Production rescheduling has been viewed as an efficient approach in responding to these uncertainties raised by the external environment and internal conditions of production [5]. In job shop and flow shop, heuristic algorithms and discrete event simulation methods are mainly applied in production scheduling problems [6–8]. However, the SWFS is large-scaled, complicated system with re-entrant flows, which is different from typical job and flow shop. Many rescheduling strategies improving traditional job shop rescheduling methods have been proposed and applied in SWFS in the recent decade [9, 10]. These methods using a single rescheduling strategy are not enough for the real-time dynamic manufacturing environment, which is more complex with disruptive events every day. For this reason, a layered rescheduling framework is needed to select rescheduling methodologies in SWFS according to the present system status.

Yield prediction plays an indispensable role in the semiconductor manufacturing factory for its powerful function of reducing cost, increasing production and maintaining a good relationship with customers. Before a malfunction is detected, the accurate prediction model of yield will serve as a warning role and help people take proactive measures to reduce the number of defect's wafers and increase the total yield of SWFS. An accurate prediction of yield plays a useful role in releasing the plan of production and optimizing the process of production, which will make the cycle time shorter and reduce fabrication cost of average units. To offer a reasonable and acceptable price and satisfy the customers, the prediction of manufacturing costs for products is necessary if they are still under development and the accurate prediction of yield can provide some advice for Ref. [11]. To maintain the good relationship with the customers, the order's due data should be guaranteed and the accurate prediction of yield is also useful in this aspect. Some organic problems located on the wafer such as microscopic particles, cluster defects, photo-resist, critical processing parameters would be the factors which affect the yield of the semiconductor wafer. With the statistic analysis models [12] and traditional artificial neural network (ANN) models [13], the prediction of semiconductor fabrication system's yield is difficult. A fuzzy neural network (FNN)-based yield model for yield prediction of semiconductor manufacturing systems is proposes in this chapter. In this system, the impacted factors, which are cluster defects, the defect's key attributed parameters, key electrical test parameters, should be considered in the same time. By this way, the precision of the wafer yield's prediction is improved.

2. The application of ANN in production scheduling and yield prediction of the SWFS

For selecting a scheduling strategy, the FNN approach is widely used. FNN is also an effective methodology for prediction of discrete event manufacturing systems, control and decision-making [14, 15]. For demonstrating the relationship between the monitoring features of a flexible manufacturing system and the conditions of tools, Li et al. [16] presented a fuzzy neural network approach. For controlling manufacturing process, Zhou et al. [17] used a fuzzy neural network approach. Chang et al. [18] created a FNN model of flow time estimation with data, which are generated from a foundry service company. The product design time was estimated with the FNN approach by Xu and Yan [19]. Chang et al. [20] used FNN approach to estimate the influence of the process on the results of the wafer fabrication in SWFS. However, the FNN approach has not been used to solve the problem of SWFS rescheduling problem. This chapter proposes the FNN-based rescheduling decision mechanism for SWFS. This methodology can solve the uncertainty problem and express the expert knowledge in weighted values. In the neural network, the evaluation of local weight values is the knowledge modelling of control rules. Rescheduling strategies, SWFS state parameters, disturbance parameters can be identified and analysed in this model. In this model, we can build the nonlinear relationship between these three components. With this approach, the layered rescheduling approach will be selected that make the yields rapid responsiveness and high productivity of the SWFS in an environment full of randomness.

To predict the wafer yield, Tong et al. [21] proposed a neural network-based approach through considering the clustering phenomenon of the defects in integrated circuit manufacturing. It was proved that the proposed approach was effective. For predicting wafer yield for integrated circuit with clustered defects, Tong and Chao [22] used a general regression neural network (GRNN) approach. Defect clustering patterns are simulated from three aspects: the size of chip, percentage of defects and the cluster pattern. A case study demonstrated the effectiveness of the approach of the model. For the lack of reliability and accuracy in the prediction of yield, an approach of a fuzzy set for yield learning was proposed by Chen and Wang [23]. A few of examples enhanced the reliability and precision of the forecasting of the yield. Chen and Lin [24] proposed a fuzzy-neural system with expert opinions, which can increase the precision of semiconductor yield prediction. The artificial intelligent-based yield forecasting models demonstrated above have some limitations that it only takes consideration of the physical parameters of wafer and the important attributed parameters of defects in wafer without considering the influence of variation of the key electrical test parameters. With the combining of neural network (NN) and memory-based reasoning (MBR), an integrated framework for a yield management system with techniques of hybrid machine learning was given by Chung and Sang [25]. In the forecasting model of the yield, some key electrical test parameters have been taken into consideration. With the use of wafer level electrical test data, a parametric neural forecasting model was constructed by Kim et al. [26] and Kim [27]. However, these yield forecasting models have not taken the attributed parameters of defects in wafer into consideration. This chapter proposes a yield forecasting model with the consideration of the wafer electrical test parameters and important attributed parameters of defects in wafer.

3. Artificial neural network for rescheduling decision mechanism in the SWFS

3.1. Layered rescheduling framework of SWFS

A layered rescheduling framework is proposed in order to reschedule the SWFS for the unstable environment which is shown in **Figure 1**. In the process of rescheduling framework, a three layers of rescheduling strategies are used. a three layers are machine group layer, machine layer and the system layer. The strategies of the rescheduling implement the dynamic scheduling, the global scheduling of SWFS and the machine scheduling. To choose the particular rescheduling strategy, the optimal rescheduling decision mechanism based on FNN approach. The layered rescheduling framework is described in detail in the following paragraph.

Global scheduling of SWFS. If there are some changes in the large-scale SWFS's condition or there are some disturbances, the rescheduling is needed and the global rescheduling of SWFS is managed for the adjustment of the global scheduling [28]. With the machine group layer's adjusted scheduling objectives, a local dynamic scheduling algorithm is applied for scheduling in the machine group layer [29]. In the end, with the machine group layer's adjusted scheduling objectives, machine scheduling is processed in real-time and the optimal machine real-time scheduling solutions are achieved.

Dynamic scheduling of SWFS. If there are some changes in the medium-scale SWFS's condition or there are some disturbances, the rescheduling in the machine group layer is needed and the local dynamic scheduling of SWFS is managed. In order to adjust the local scheduling of a machine group, a local dynamic scheduling algorithm is applied. With the adjusted scheduling objectives of the machine layer taken into consideration, machine scheduling of SWFS is processed.

Machine scheduling of SWFS. If there are some changes in the large-scale SWFS's condition or there are some disturbances, the rescheduling is just accomplished and in the same time, the machine scheduling is processed. Though they are same in the operation sequences of the lots, they are different in the operation start times of delayed lots.

FNN-based optimal rescheduling decision mechanism. With the consideration of the statuses and disturbances to SWFS, the rescheduling layer is chosen by optimal rescheduling decision mechanism. According to the fuzzy neural network, an algorithm for the system is stated in this paper.

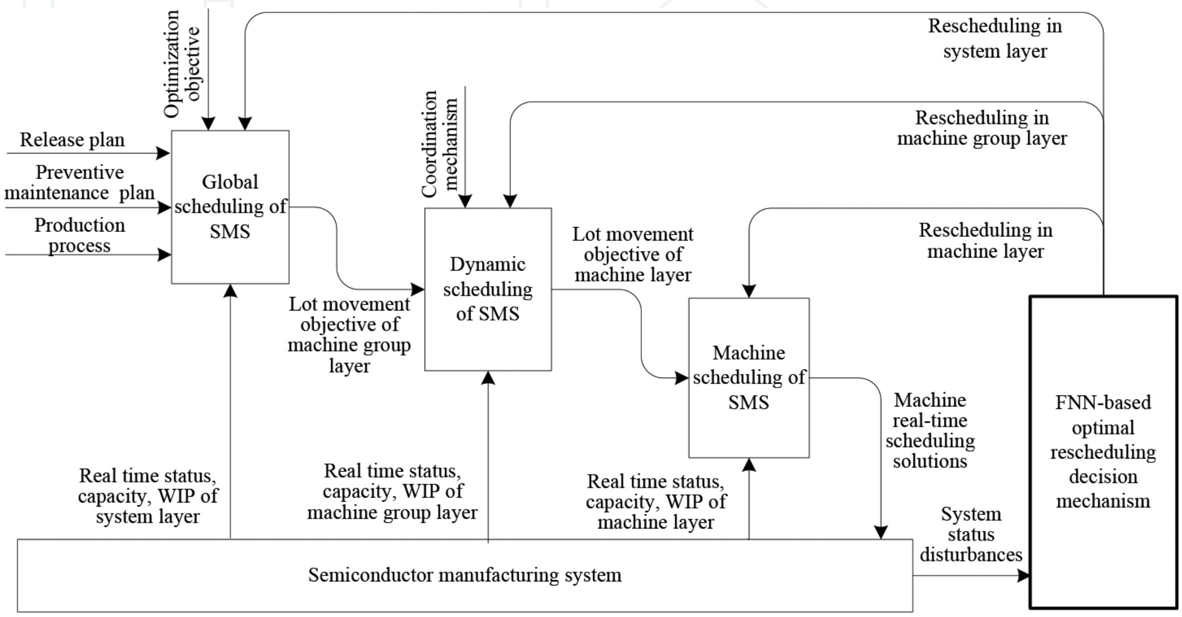


Figure 1. Layered rescheduling framework of SWFS.

3.2. FNN-based decision mechanism for rescheduling

Fuzzy neural network (FNN) is an ingenious combination of fuzzy logic and neural network, which inherits the advantages from both fuzzy system and neural network. The FNN has the characteristics of processing fuzzy information with fuzzy algorithms and learning with a high-speed parallel structure. The FNN approach is therefore adaptable and robust, and is well suited for the SMS rescheduling problem.

The FNN-based rescheduling decision model consists of an input layer, several hidden layers and an output layer. Input parameters connected with disturbances and state parameters are accepted in the input layer. The hidden layers calculate and transform the input parameters using fuzzy logic theory. The output layer produces the decision-making response of the rescheduling model. More details of this method are described.

3.2.1. Input factors in the proposed FNN model

The SMS's state and disturbance parameters are treated as input of the FNN, which can be detailed as: system disturbances parameter, average queue length, stability of SMS, average relative load and average slack time.

3.2.1.1. System disturbances parameter

Since the operating environments of SMS are uncertain and dynamic, disturbances mainly include: machine failures, lot reworks and rush orders. Once a disturbance has happened, an optimal rescheduling strategy must be selected and carried out to guarantee the stability and efficiency of SMS. Disturbances are converted into machine work times to quantify their effect. The mapping of disturbances to machine work times is defined as follows.

(1) Machine failures. The processing time in SMS increases if machine failures happen. Suppose that t^f refers to the increased process time caused by all machine failures, then,

$$t^f = \sum_{M_j^f \in F^f} \sum_{m_{ji} \in M_j^f} t^f{}_{ji} \tag{1}$$

where M_j^f is the failed machine group *j*, m_{ji} refers to the failed machine i of machine group j, t_{ji}^f represents the repair time of machine *i* of machine group *j*, and F^f is the set of failed machine group.

(2) Lot reworks. Lot reworks raise the output requirement of SMS. Suppose that t^r is the additional process time incurred by all lot reworks, then,

$$t^r = \sum_{M_j^r \in F^r} \sum_{p_{jk} \in R_j^r} t_{jk}^r \tag{2}$$

where M_j^r refers to the machine group *j* operating the rework lots, R_j^r refers to the set of the rework lots operated by the machine group *j*, p_{jk} refers to the rework lot k operated by the machine group j, t_{jk}^r refers to the process time of the rework lot k operated by the machine group j, F^r refers to the set of the machine group that operate rework lots.

(3) Rush orders. Rush orders also demand more of the production requirement of SMS. Suppose that t^o is the process time required by all rush orders, then,

$$t^o = \sum_{M_j^o \in F^o} \sum_{q_{jk} \in R_j^o} t_{jk}^o \tag{3}$$

where M_j^o represents the machine group j, that operates the rush orders in current plan time phase, R_j^o represents the set of the lots operated by machine group j in the rush orders; q_{jk} represents the lot *k* in the rush orders operated by machine group j, t_{jk}^o is the process time of the lot k operated by the machine group j in the rush orders, F^o represents the set of the machine group which are related with rush orders.

(4) System disturbances parameter. Suppose that td is the system disturbances parameter, denoting the total effect of disturbances on SMS scheduling. The formula to calculate td is shown as the (4).

$$td = t^f + t^r + t^o \tag{4}$$

3.2.1.2. Average queue length

Average queue length of machine groups reflecting the utility of the machine group is affected by disturbances. L is the average queue length of machine groups affected by disturbances; and the formula is shown in (5).

$$L = \frac{\sum_{M_j \in (F^r \cup F^f \cup F^o)} L_j}{N} \tag{5}$$

where M_j denotes the machine group j , L_j means queue length of machine group M_j, N refers to the number of machine group that affected by disturbances.

3.2.1.3. Stability of SMS

The stability of SMS is defined as the deviation in predicted average start time of a rescheduled strategy from the real start time. β denotes the stability of SMS, which is shown in (6).

$$\beta = \frac{\sum_{\substack{(i,s)\in S \\ tc_{is} \le tc}} \left| tc'_{is} - tc_{is} \right|}{\sum_{\substack{(i,s)\in S \\ tc_{is} \le tc}} q_{is}} \tag{6}$$

where tc'_{is} is practical start time of process stage s of product i, tc_{is} is computational start time of process stage s of product i which optimized with a global scheduling algorithm or re-scheduling strategy, q_{is} is the number of process stage s of product i, tc is the current time when disturbance happens, S is set of tasks of all machine group in SMS.

3.2.1.4. Average relative loads

Average relative loads denote the loads of machine groups measured from the current time to the end of the scheduling horizon which can be affected by disturbances. Let η represent the average relative loads, the formula for calculation is shown in (7).

$$\eta = \frac{\sum\limits_{\substack{(i,s)\in S_d \\ t_e \geq t_{is} \geq t_c}} tp_{is}}{\sum\limits_{M_j \in (F^r \cup F^f \cup F^o{}_3)} n_j (te - tc)} \tag{7}$$

where tp_{is} denotes process time of process stage s of product i, te denotes the time point when scheduling is ended, n_j denotes the number of machine of machine group M_j, S_d represents set of tasks of machine group which affected by disturbances.

3.2.1.5. Average slack time

Average slack time represents the space that the machine groups can be adjusted when disturbances happen. Suppose ts is the average slack time, shown in (8).

$$ts = \frac{\sum\limits_{\substack{(i,s)\in S_d \\ te \geq t_{is} \geq tc}} \left(t_{i(s+1)} - t_{is} - tp_{is} \right)}{\sum\limits_{\substack{(i,s)\in S_d \\ te \geq t_{is} \geq tc}} q_{is}} \tag{8}$$

3.2.2. Output variables

The output variables in the FNN output layer are related to the layered rescheduling strategies, which consists of the rescheduling in system layer, machine group layer, and machine layer. If a particular layered rescheduling strategy is selected, then the corresponding output variable is close to 1, otherwise it equals to 0. In FNN-based rescheduling decision model, suppose that y_1, y_2, y_3 are defined as output variables, then y_1, y_2, y_3 correspond to the rescheduling in system layer, rescheduling in machine group layer, and rescheduling in machine layer, respectively.

3.2.3. The structure of FNN

There are five layers in the rescheduling decision model based on FNN, as illustrated in **Figure 2**.

a. The input vector is X = [x1, x2, x3, x4, x5]T = [L, β, η, ts, td]T. The function of node input-output is:

$$f_{i^{(1)}} = x_{i(0)} = x_i ; x_{i(1)} = g_{i(1)} = f_{i^{(1)}} ; i = 1, 2, \cdots 5 \tag{9}$$

b. In the second layer which is the fuzzifer layer, the function of the Gauss membership is adopted.

http://dx.doi.org/10.5772/63444

$$u_{ij} = e^{-\frac{(x_i - c_{ij})^2}{\sigma_{ij}^2}}, i = 1, 2, \cdots, 5, \\ j = 1, 2, \cdots, k_i \tag{10}$$

In this formula, cij is the centre and σ_{ij} is width. The node input–output function is:

$$f_{ij(2)} = -\frac{(x_i - c_{ij})^2}{\sigma_{ij}^2};$$

$$x_{ij(2)} = u_{ij} = g_{ij(2)} = e^{f_{ij(2)}} = e^{-\frac{(x_i - c_{ij})^2}{\sigma_{ij}^2}}; \\ \boldsymbol{i = 1, 2, \cdots 5, j = 1, 2, \cdots, k_i} \tag{11}$$

c. In the third layer as the rule layer, each node in the layer is a fuzzy rule which not only matches the front part of the fuzzy rule but also calculates the adaptive of the rule,

$$a_j = \prod_{l=1}^{5} u_{li_l}(x_{li_l}), \\ j = 1, 2, \cdots, n \tag{12}$$

In this layer, the input-–output function is:

$$f_{j(3)} = \prod_{l=1}^{5} x_{li_l(2)} = \prod_{l=1}^{5} u_{li_l}(x_{li_l}); \; x_{j(3)} = g_{j(3)} = f_{j(3)}; \\ j = 1, 2, \cdots, n x_{j(3)} = g_{j(3)} = f_{j(3)}; j = \mathbf{1, 2,} \cdots, n \tag{13}$$

d. In the fourth layer which is the normalized layer. In this layer, the node numbers are the same in the third layer. It normalized the adaptive values of these rules. And the input-output function is:

$$f_{j(4)} = \frac{x_{j(3)}}{\sum_{i=1}^{n} x_{i(3)}} = \frac{a_j}{\sum_{i=1}^{n} a_i}; \\ x_{j(4)} = g_{j(4)} = f_{j(4)}; j = \mathbf{1, 2,} \cdots, n \tag{14}$$

e. The last layer is the output layer. It defuzzify the output variables. And each node describes a rescheduling strategy. While a rescheduling strategy is chose, the corresponding output is 1 or 0. The input-output function is:

$$f_{i(5)} = \sum_{j=1}^{n} w_{ij} x_{j(4)} \quad = \sum_{j=1}^{n} w_{ij} b_j ; x_{j(5)} = g_{j(5)} = f_{j(5)} \;\; ; i = 1,2,3$$
$$x_{j(5)} = g_{j(5)} = f_{j(5)} \;\; ; i = 1,2,3 \tag{15}$$
$$x_{j(5)} = g_{j(5)} = f_{j(5)} \;\; ; i = 1,2,3$$

where w_{ij} is the connection weight parameter.

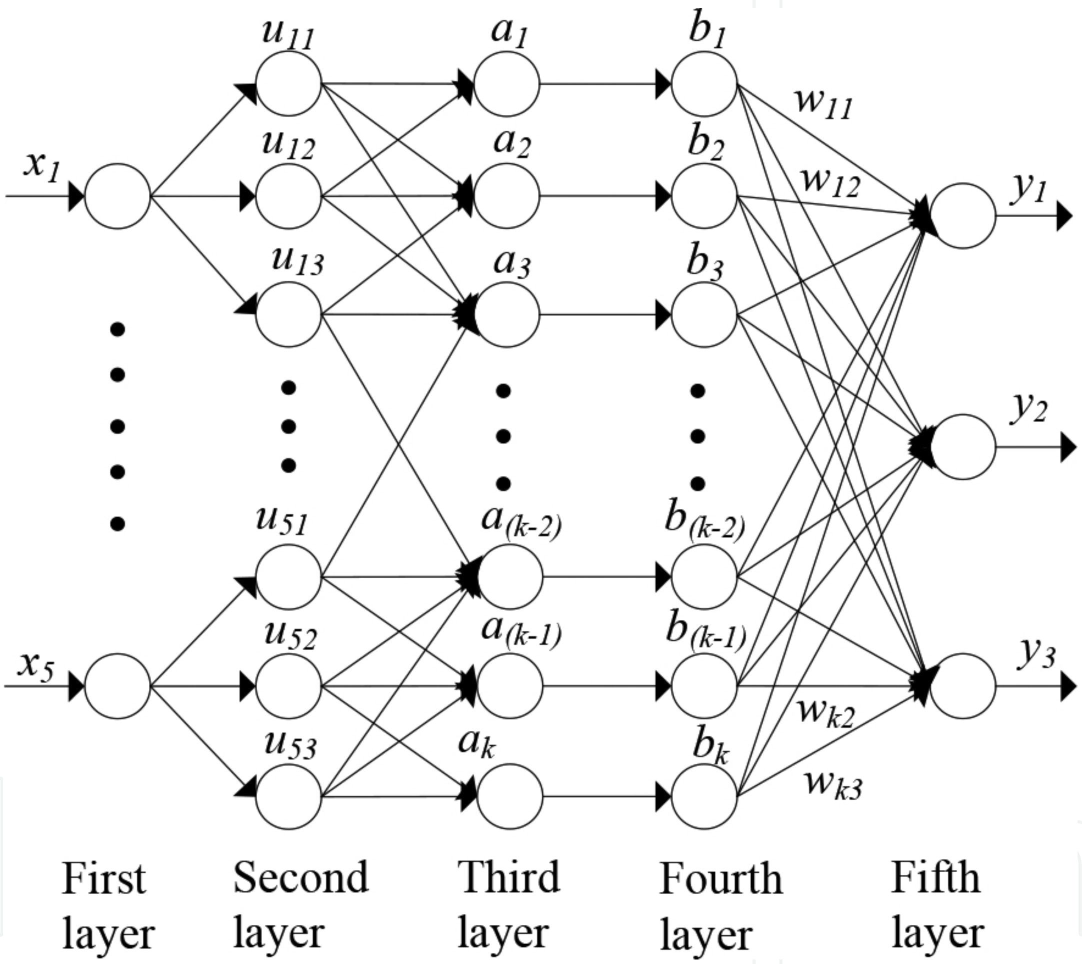


Figure 2. FNN structure.

3.2.4. The strategy of the fuzzy inference

The Mamdani-based fuzzy inference is applied in this FNN-based rescheduling decision model with a assumption that the fuzzy rule Ri describes the relationship between input and output. Then,

http://dx.doi.org/10.5772/63444

Ri:

IF x1 is A1i and x2 is A2i and … and xm is Ami,

THEN y1 is B1i and y2 is B2i and … and ym is Bki,

where

i = 1, 2, …, n.

n: number of rules.

m: number of input variables.

k: number of output variables.

A_{ji}: value of fuzzy linguistic variable xj.

B_{ji}: value of fuzzy linguistic variable yj.

3.3. Result and discussion

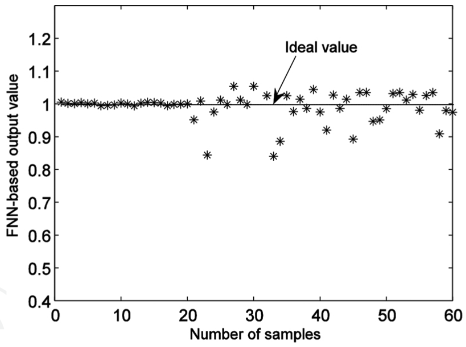
3.3.1. Experiment on the proposed FNN approach

In this section, the experiments are conducted to evaluate the effectiveness of the proposed FNN rescheduling decision mechanism. A discrete event simulation model is run to gather the experiment data, which is based on a 6-in. SWFS in Shanghai. This SWFS is composed by eleven machine groups, which add up to thirty-four machines in total. And three types of wafers are put into the SWFS. The processes of all three types of wafer lots are divided into dozes of stages, which is composed by a key step and several successive normal steps. One hundred and fifty records of rescheduling decision are collected from the simulation model, and shown in **Table 1**. Ninety records are used in model training, and 60 are taken to evaluate the model. The presented FNN approach is compared with the back propagation network (BPN) approach and the multivariate regression methodology, since the BPN and multivariate regression approaches are widely used in the rescheduling strategy decision and proven to be competitive [30, 31]. Furthermore, the detail numerical comparison of the FNN approach, BPNN approach and multivariate regression are demonstrated as follows.

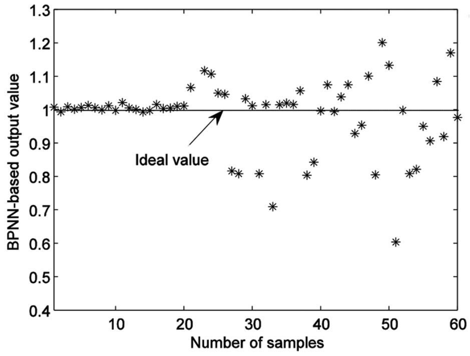
Now, it's going to compare the experimental results which are made by these three methods. **Figure 3** shows the optimal rescheduling decision value and the model outputs. It shows that the FNN rescheduling method has the best convergence. We also contrast the RMSE and the decision coefficients R^2 of these three methodologies in **Table 2**. The FNN has the best performance for the RMSE which is 0.042 and has the largest of the R^2 values which is 0.9941. Hence, the rescheduling decision based on FNN has the best performance in these three methods.

Samples no	Average queue length of disturbed machine stations x_1 (lot)	Stability of scheduling x_2 (h)	Average load of disturbed machine stations x_3 (100%)	Average slack time of disturbed machine stations x_4 (h)	Disturbance x_5 (h)	Optimal rescheduling decision objective		
						Rescheduling in machine layer y_1	Rescheduling in machine group layer y_2	Rescheduling in system layer y_3
1	1	1.10	0.59	6.42	2.14	1	0	0
2	2	0.78	0.52	6.51	2.01	1	0	0
3	0	0.81	0.46	5.16	1.76	1	0	0
4	0	0.48	0.1	5.81	1.21	1	0	0
5	2	0.49	0.14	4.62	1.42	1	0	0
6	1	0.52	0.12	5.54	1.79	1	0	0
7	2	0.74	0.17	5.16	1.13	1	0	0
8	0	0.76	0.07	4.49	1.64	1	0	0
9	6	0.38	0.29	4.86	1.4	1	0	0
10	5	0.37	0.26	5.17	2.21	1	0	0
..	..	..	..	..	..	..	..	..
141	6	4.23	0.75	3.81	9.81	0	0	1
142	4	4.15	0.76	5.64	9.18	0	0	1
143	5	4.01	0.76	4.97	9.77	0	0	1
144	2	0.81	0.38	1.87	9.87	0	0	1
145	4	0.87	0.37	2.41	10.11	0	0	1
146	2	0.68	0.42	2.18	9.42	0	0	1
147	2	0.72	0.35	2.18	9.76	0	0	1
148	2	0.91	0.23	2.7	9.13	0	0	1
149	1	0.87	0.21	2.97	9.73	0	0	1
150	2	0.87	0.31	2.77	10.18	0	0	1

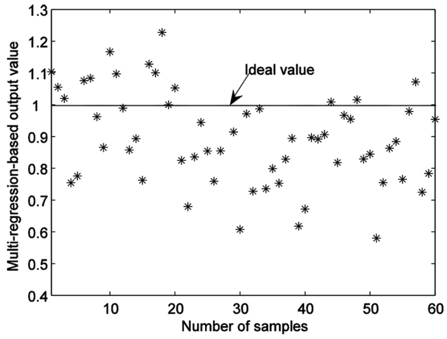
Table 1. One hundred and fifty records for numerical experiments.



(a) FNN based output value



(b) BPNN based output value



(c) Multivariate regression based output value

Figure 3. The relationship between the rescheduling strategy output and ideal target output for the FNN, BPNN and multivariate regression methods. (a) FNN-based output value, (b) BPNN-based output value and (c) multivariate regression-based output value.

Rescheduling strategy model	RMSE	R^2		
		$R^2_{Y_1}$	$R^2_{Y_2}$	$R^2_{Y_3}$
FNN	0.0042	0.9880	0.9762	0.9941
BPNN	0.0132	0.9745	0.9178	0.9274
Multivariate regression	0.0897	0.85887	0.75566	0.70813

Table 2. Comparison of RMSE and decision coefficients among the FNN, BPNN and multivariate regression methods.

3.3.2. Experiment on the proposed rescheduling decision mechanism

The FNN rescheduling decision mechanism is used in our layered rescheduling method (Method 1). There are two other different rescheduling methods. One is the monolayer-based rescheduling approach (Method 2). Another one is the first come first served (FCFS) approach (Method 23). In our method, the FNN rescheduling decision mechanism figures out the optimal rescheduling approaches which include the global scheduling of SWFS, the dynamic scheduling and the machine scheduling. By contrast, the Method 2 only considers the rescheduling of the machine group layer. But in practice, the Method 3 is widely used in the Fab. In order to prove the efficiency of our approach, we also compared these three rescheduling methods in terms of the machine utilization and the daily movement, which are the important system targets for SWFS.

In the case study, the data are collected from a 6-in. SWFS in Shanghai. It products three kinds of lots which are renamed as A, B and C. The whole process is shown in **Table 4**. This SWFS has eleven key machine groups (shown in **Table 3**). which has 34 machines with MTTF and MTTR parameters. They are explained in Section 5. The SWFS simulation model is built by eM-plant 7.0 software. In the simulation, it took 12 days, including a 5-day warm-up. Ten times repeated trials of the same stimulation, in which the initiated loads of machines were different, were performed (3 rescheduling methods 10 replications). The results are shown in **Figures 4** and **5**, which illustrate:

1. Method 1 performs well in the rescheduling decision in the SWFS.
2. Method 1 outperforms method 2 and 3, which indicates the layered rescheduling method is more suitable than the conventional FCFS rescheduling approach and monolayer-based rescheduling approach in the complex SWFS.

Machine group number	Processing type	Number of machine	Batch size	MTBF	MTTR
1	Ion implant	3	1	70	1
2	Ion implant	4	1	70	1
3	Diffusion	3	5	100	2
4	Diffusion	4	5	110	2
5	Etching	2	1	90	1
6	Etching	4	1	80	1
7	Etching	3	1	60	1
8	Etching	2	1	70	1
9	Lithography	4	1	90	1
10	Lithography	3	1	80	1
11	Lithography	2	1	100	1

Table 3. Configuration of SWFS.

http://dx.doi.org/10.5772/63444

Stage number	Number of time period by product *A*	Machine group number of product *A*	Process time of product *A* by key machine t/hour	Number of time period by product *B*	Machine group number of product *B*	Process time of product *B* by key machine t/hour	Number of time period by product *C*	Machine group number of product *C*	Process time of product *C* by key machine t/hour
1	1	8	1	1	10	1	1	10	1
2	1	7	1	1	8	1	1	8	1
3	1	10	1	4	5	1	2	5	1
4	1	9	1	1	10	1	1	8	1
5	5	0	1	1	1	1	1	0	1
6	3	0	1	1	6	1	4	3	6
7	4	2	4	1	10	1	1	8	1
8	1	8	1	1	9	1	2	5	1
9	1	5	1	6	0	1	2	2	3
10	1	3	3	1	0	1	1	8	1
11	1	0	1	1	4	1	2	5	3
12	2	10	1	1	8	1	1	0	1
13	1	0	1	2	1	1	2	3	6
14	4	4	2	3	2	3	1	8	1
15	1	6	1	1	10	1	2	9	1
16	2	1	1	1	8	1	2	4	1
17	3	4	2	2	5	2	3	0	1
18	1	9	1	1	3	1	1	8	1
19	1	5	1	2	5	1	3	6	1
20	1	8	1	1	9	1	2	1	1
21	3	3	3	1	1	1	1	8	1
22	1	5	1	2	5	1	3	7	1
23	4	2	4	1	9	1	2	2	3
24	1	9	1	3	0	1	2	5	1
25	3	4	2	3	2	3	1	0	1
26	2	0	1	2	10	1	1	9	1
27	1	8	1	2	7	1	2	4	1
28	1	7	1	–	–	–	1	1	1
29	2	3	3	–	–	–	–	–	–
30	1	2	1	–	–	–	–	–	–

Table 4. Lot products whole process.

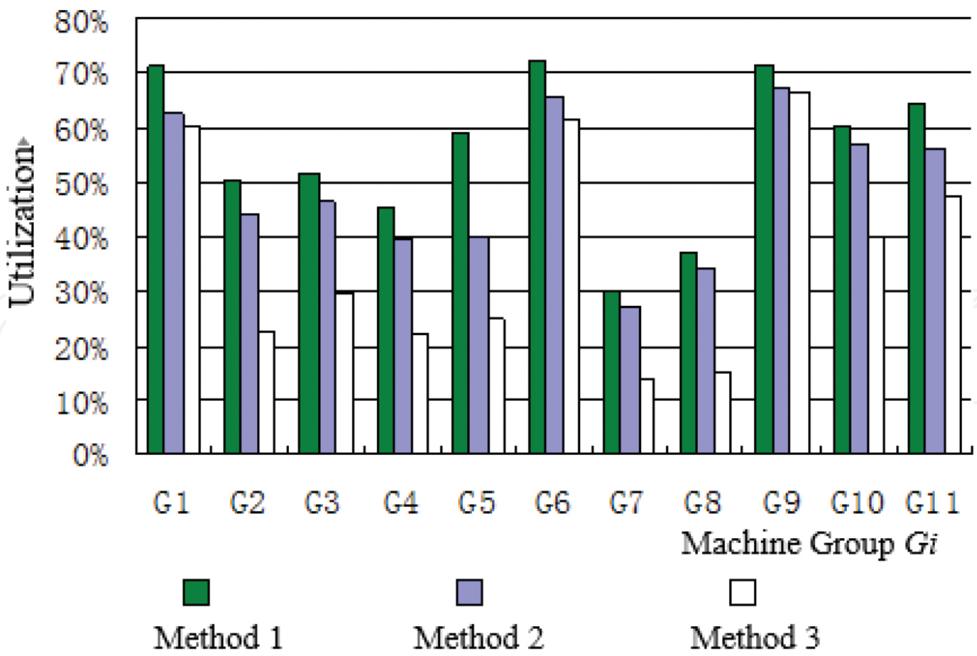


Figure 4. The utilization of machine group.

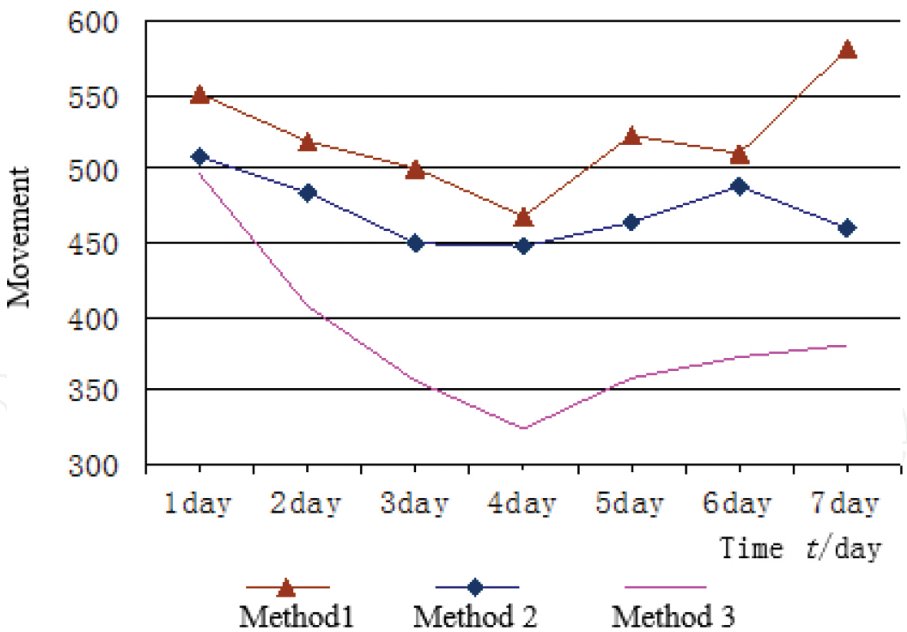


Figure 5. The utilization of machine group.

http://dx.doi.org/10.5772/63444

4. Artificial neural network approach for die yield prediction in the SWFS

4.1. FNN-based yield prediction model

The yield prediction model based on FNN is composed of three parts, which are an input layer, an output layer and several hidden layers. The three parts do the different jobs respectively. The input layer serves to accept input parameters connected with yield. The output layer does the job to get the yield response of the prediction model. The hidden layers are applied to compute and convert the input parameters which are on the basis of fuzzy logical theory. The following sections show a more detailed yield prediction model based on FNN.

4.1.1. Variables in FNN input layer

The input variables in the FNN prediction model include the following parameters: the critical electrical test parameters, wafer physical parameters and key parameters of defects in wafer. Critical process parameters refer to those electrical test parameters which are generally tested at the end of the wafer processing, and they have notable influences on the yield. Wafer physical parameters mainly refer to the size of the chip. Key parameters of defects in wafer Contain a number of defects, clustering parameter, mean number of defects in each chip and mean a number of defects in each unit area. Among these input variables, the critical electrical test parameters and clustering parameters are complex, and we will discuss them in the following sections.

4.1.1.1. Critical electrical test parameters

In the process of fabricating complex semiconductor wafer, there are more than one hundred electrical test parameters related to the probed wafer. This paper mainly does the research on establishing the exact relationship of a small number of critical electrical test parameters with yield. These critical electrical test parameters have significant influence on yield, and they have high correlating coefficients or exhibit a 'cliff' in the correlation graphs which means they can quickly improve the yield. Wong [32] proposed the hybrid statistical correlation analysis method, and the critical electrical test parameters are identified based on this method. Here, we remove some details of these electrical test parameters for confidentiality.

4.1.1.2. Clustering parameter

Clustering parameter displays cluster or clumps degrees of wafer defects in the defect map [33]. Suppose that the clustering parameter is expressed by c, shown in Eq. (1).

$$c = \min\left\{\frac{s_v^2}{\overline{v}^2}, \frac{s_w^2}{\overline{w}^2}\right\} \tag{22}$$

where the sample mean and variance of V_i is represented by $\bar{V}^2$ and S_v^2; and the sample mean and variance of W_i are represented by S_w^2. V_i and W_i are a series of defect intervals on the x and y axis defined as:

$$V_i = x_{(i)} - x_{(i-1)},\ i = 1, 2, \ldots, n \tag{23}$$

$$W_i = y_{(i)} - y_{(i-1)},\ i = 1, 2, \ldots, n \tag{24}$$

where $x_{(i)}$ refers to the ith smallest defect coordinates on x axis, and similarly, $y_{(i)}$ refers to the ith smallest defect coordinates on y axis, $x(0) = y(0) = 0$, and n refers to the quantity of defects on one wafer. If the defects are randomly scattered, the value of CI is close to 1, and when clustering of defects appears, the value of CI is likely to be greater than 1.

4.1.2. FNN structure

There are five layers in the rescheduling decision model based on FNN, as illustrated in **Figure 6**.

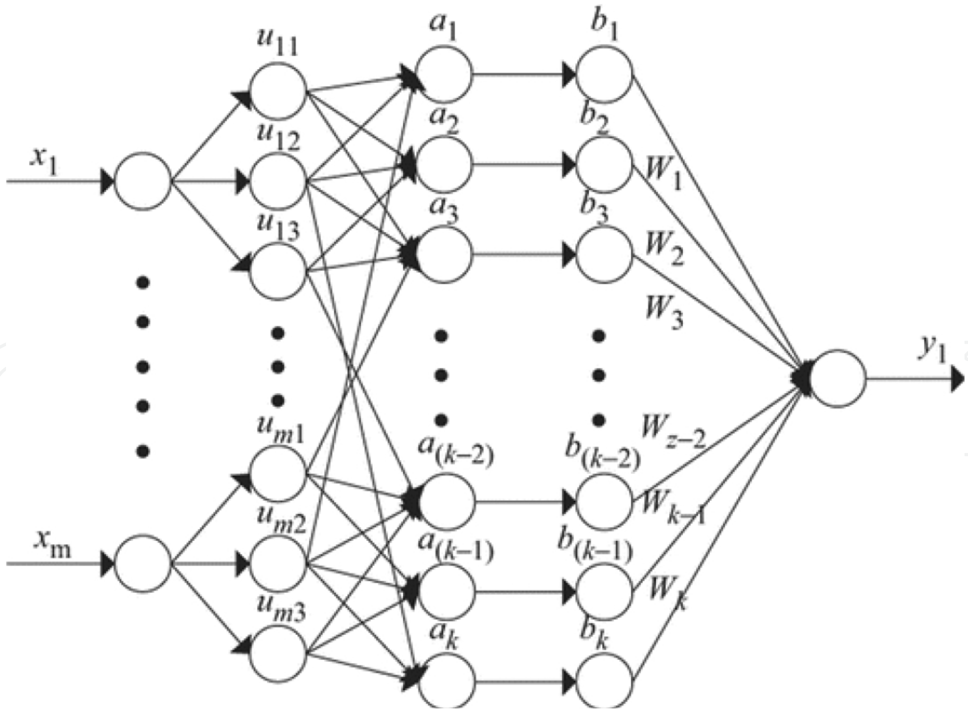


Figure 6. FNN model structure.

a. The input vector is $\boldsymbol{X} = [x_1, x_2, x_3, \ldots, x_m]$. The function of node input-output is:

$$f_i^{(1)} = x_i;\ x_i^{(1)} = g_i^{(1)} = f_i^{(1)}; i = 1, 2, \cdots m \tag{4-4}$$

b. In the second layer which is the fuzzifer layer, the function of the Gauss membership is adopted.

$$u_{ij}(x_i) = e^{-\frac{(x_i - c_{ij})^2}{\sigma_{ij}^2}} \tag{25}$$

In this formula, cij is the centre and σ_{ij} is width. The node input-output function is:

$$f_{ij}^{(2)} = -\frac{(x_i^{(1)} - c_{ij})^2}{\sigma_{ij}^2}; x_{ij}^{(2)} = u_{ij}(x_i^{(1)}) = g_{ij}^{(2)} = e^{f_{ij}^{(2)}} = e^{-\frac{(x_i - c_{ij})^2}{\sigma_{ij}^2}} \tag{26}$$

where $i = 1, 2, \cdots m$ and $j = \mathbf{1}, \mathbf{2}, \cdots, l_i$.

c. In the third layer as the rule layer, each node in the layer is a fuzzy rule which not only matches the front part of the fuzzy rule but also calculates the adaptive of the rule,

$$a_j = \prod_{i=1}^{m} u_{il_i}(x_i^{(1)}), j = 1, 2, \cdots, n \tag{27}$$

In this layer, the input-output function is:

$$f_j^{(3)} = \prod_{i=1}^{m} x_{il_i}^{(2)} = \prod_{i=1}^{m} u_{il_i}(x_i^{(1)});\ x_j^{(3)} = a_j = g_j^{(3)} = f_j^{(3)}; j = 1, 2, \cdots, n \tag{28}$$

d. In the fourth layer which is the normalized layer. In this layer, the node numbers are the same in the third layer. It normalized the adaptive values of these rules. And the input-output function is:

$$b_j = \frac{a_j}{\sum_{i=1}^{n} a_i}, j = 1, 2, \cdots, n \tag{29}$$

Node input-output function in this layer is as follows.

$$f_j^{(4)} = \frac{x_j^{(3)}}{\sum_{i=1}^{n} x_i^{(3)}} = \frac{a_j}{\sum_{i=1}^{n} a_i};\ x_j^{(4)} = b_j = g_j^{(4)} = f_j^{(4)}; j = 1, 2, \cdots, n \tag{30}$$

e. The last layer is the output layer. It defuzzify the output variables. And each node describes a rescheduling strategy. While a rescheduling strategy is chose, the corresponding output is 1 or 0. The input-output function is:

$$f^{(5)} = \sum_{j=1}^{n} w_j x_j^{(4)} = \sum_{j=1}^{n} w_j b_j;\ O_{out} = x^{(5)} = g^{(5)} = f^{(5)} \tag{31}$$

where W_j is connection weight parameter of output layer, and O_{out} is the output of FNN model.

4.2. Case study

In this section, the experiments are conducted to evaluate the effectiveness of the proposed FNN method. This section presents a numerical experiment study to demonstrate the effectiveness of the approach proposed. Seven hundred and twenty wafer samples are obtained from a 6 in. SWFS in Shanghai, and each sample includes 360 records of wafer yield. Five hundred and fifty-two records are used in model training, and 168 are taken to evaluate the model. The attributes contained in each record are, in order, number of defects, clustering parameter, die yield, mean number of defects per unit area, chip size parameter, mean number of defects per chip, and 28 electrical test parameters, which is shown in **Table 5**. Each feature is acquired by test during the critical manufacturing process. The presented FNN approach is compared with the Poisson model, negative binomial model and BPNN approaches, since the three approaches are widely used in research on yield predicting and have been proved to be competitive [34–36]. Furthermore, the detail numerical comparison of the FNN approach, Poisson model, negative binomial model and BPNN approaches are demonstrated as follows.

Record	Number of defects	Mean number of defects/chip	…	Chip size parameter (cm^2)	Clustering parameter	Process parameter 1	Process parameter 2	…	Process parameter 28	Yield (%)
1	21	0.14094	…	1	0.51836	322.7498	0.060865	…	1.169573	0.86577
2	45	0.30201	…	1	0.70814	324.4634	0.061573	…	1.172481	0.73826
3	16	0.10738	…	1	0.65597	322.1903	0.060648	…	1.176223	0.89262
4	21	0.14094	…	1	1.0277	313.9659	0.065036	…	1.162216	0.87248

http://dx.doi.org/10.5772/63444

Record	Number of defects	Mean number of defects/chip	…	Chip size parameter (cm^2)	Clustering parameter	Process parameter 1	Process parameter 2	…	Process parameter 28	Yield (%)
5	46	0.30872	…	1	0.59023	313.0953	0.068286	…	1.183461	0.73826
6	35	0.2349	…	1	0.73168	323.9832	0.061867	…	1.164384	0.81879
7	7	0.04698	…	1	0.73807	315.9001	0.059539	…	1.177436	0.95302
8	49	0.32886	…	1	0.75913	310.9356	0.060887	…	1.180799	0.72483
9	9	0.060403	…	1	0.57871	310.6571	0.06439	…	1.168414	0.9396
10	33	0.22148	…	1	0.83289	310.0921	0.068695	…	1.179983	0.80537
11	37	0.24832	…	1	0.69348	322.2567	0.068536	…	1.160716	0.77181
12	48	0.32215	…	1	0.9089	323.7277	0.069959	…	1.162259	0.73154
13	12	0.080537	…	1	0.6154	321.1246	0.067403	…	1.176334	0.91946
14	33	0.22148	…	1	0.85056	313.0574	0.068778	…	1.170839	0.78523
15	47	0.31544	…	1	1.0109	313.5133	0.063245	…	1.172714	0.72483
…	…	…	…	…	…	…	…	…	…	…
705	31	0.31959	…	1.44	1.0678	310.2484	0.065854	…	1.168734	0.82474
706	35	0.36082	…	1.44	0.67849	321.1613	0.06883	…	1.166041	0.83505
707	61	0.62887	…	1.44	1.00661	314.3752	0.066945	…	1.184936	0.75258
708	71	0.73196	…	1.44	0.95411	321.3472	0.061456	…	1.164627	0.7732
709	82	0.84536	…	1.44	1.5379	311.5562	0.060457	…	1.160369	0.73196
710	32	0.3299	…	1.44	1.2404	313.4114	0.067197	…	1.160229	0.80412
711	79	0.81443	…	1.44	1.5407	317.3192	0.067454	…	1.175912	0.7732
712	72	0.74227	…	1.44	2.4467	323.099	0.062285	…	1.171942	0.76289
713	58	0.59794	…	1.44	2.32601	318.6861	0.068344	…	1.174814	0.76289
714	57	0.58763	…	1.44	1.0014	316.3194	0.065981	…	1.173957	0.79381
715	73	0.75258	…	1.44	1.3139	322.1991	0.065962	…	1.180853	0.73196
716	46	0.47423	…	1.44	1.6882	320.4291	0.069517	…	1.166221	0.79381
717	80	0.82474	…	1.44	1.57021	316.6195	0.062461	…	1.16293	0.74227
718	38	0.39175	…	1.44	2.0034	317.0604	0.06606	…	1.171795	0.79381
719	18	0.18557	…	1.44	0.81646	323.2213	0.064985	…	1.167234	0.8866
720	72	0.74227	…	1.44	0.95479	319.6768	0.068621	…	1.175982	0.73196

Table 5. Partial wafer measurements parameters and yield.

4.2.1. Experiment on fuzzy neural network

The algorithm was programmed in Matlab 6.5, and ten factors were treated as input of the model, which are mean number of defects per chip, chip size, clustering parameter, mean number of defects per unit area, the number of defects per wafer and another five critical electrical test parameters.

Twenty-six rules for classification were identified the fuzzifier layer in the model. The 552 samples were utilized in the training of the FNN model with fivefold cross-validation. The learning process was explored in **Figure 7**. Afterward, the trained model was assessed by another 168 samples, which is demonstrated in **Table 6**. Furthermore, the linear regression analysis of the output of the FNN model is detailed in **Figure 8**.

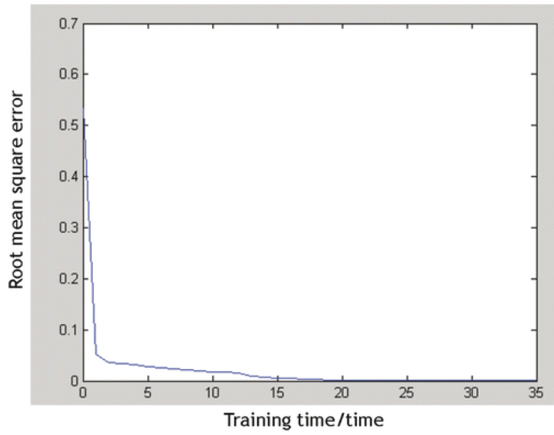


Figure 7. Fuzzy neural network learning curve.

Samples	The actual yield	The predicted yield	Relative error
1	0.72483	0.72514	0.000432
2	0.69128	0.69118	0.000146
3	0.8255	0.8337	0.009929
4	0.75168	0.74329	0.011168
5	0.75839	0.758	0.000513
6	0.73154	0.73031	0.001677
7	0.89933	0.89876	0.000632
8	0.75168	0.75211	0.000576
9	0.74497	0.74409	0.001179
10	0.75168	0.75002	0.002209
11	0.7651	0.7637	0.001824

Samples	The actual yield	The predicted yield	Relative error
12	0.85235	0.85658	0.004967
13	0.89933	0.90485	0.006143
14	0.9396	0.93918	0.000452
15	0.81208	0.81559	0.00432
…	…	…	…
160	0.75258	0.73323	0.025707
161	0.86598	0.84315	0.026365
162	0.74227	0.74376	0.002008
163	0.8866	0.88606	0.000609
164	0.73196	0.72817	0.005175
165	0.7732	0.76827	0.006381
166	0.75258	0.7456	0.009278
167	0.82474	0.83413	0.011391
168	0.83505	0.82249	0.01504

Table 6. The predicted yield based on FNN.

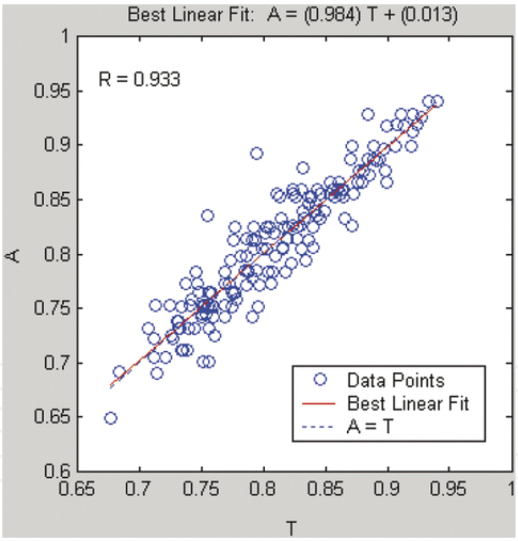


Figure 8. The linear regression analysis of the output of the FNN model.

4.2.2. Experiment of Poisson model

The Poisson model was built to predict wafer yield as follows.

$$Y = e^{-D_0 A} \tag{38}$$

In the model, Y means the wafer yield, D_0 denotes the defect density, and A is the chip size. The yield was forecasted by Poisson model, and the results of 168 samples can be found in **Table 7**. The lineal correlation analysis between the actual wafer yield and the prediction value is shown in **Figure 9**.

Samples	The actual yield	The predicted yield	Relative error
1	0.72483	0.57675	0.20429
2	0.69128	0.48117	0.30395
3	0.8255	0.79597	0.035769
4	0.75168	0.6552	0.12836
5	0.75839	0.66853	0.11849
6	0.73154	0.58064	0.20627
7	0.89933	0.89218	0.007953
8	0.75168	0.65081	0.13419
9	0.74497	0.61679	0.17206
10	0.75168	0.61267	0.18493
11	0.7651	0.59644	0.22044
12	0.85235	0.83988	0.014634
13	0.89933	0.87439	0.027733
14	0.9396	0.92883	0.011459
15	0.81208	0.74932	0.077284
…	…	…	…
160	0.75258	0.55222	0.26623
161	0.86598	0.75422	0.12905
162	0.74227	0.49038	0.33936
163	0.8866	0.84934	0.042026
164	0.73196	0.40431	0.44763
165	0.7732	0.48316	0.37512
166	0.75258	0.43547	0.42137
167	0.82474	0.75422	0.085502
168	0.83505	0.74311	0.1101

Table 7. The predicted yield based on the Poisson model.

http://dx.doi.org/10.5772/63444

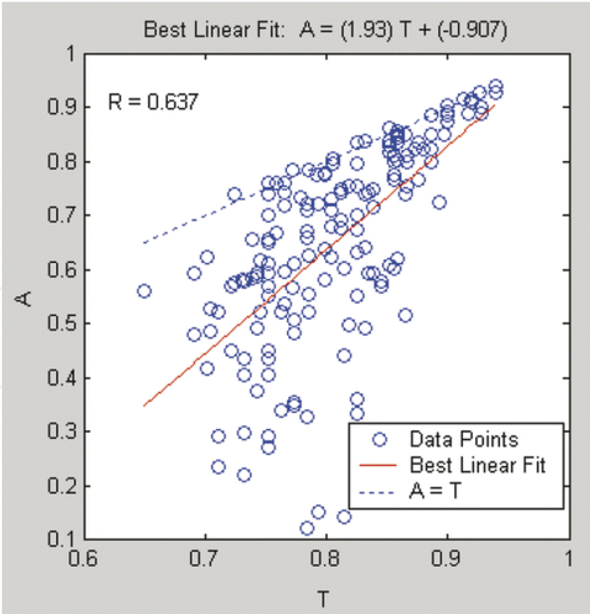


Figure 9. The linear regression analysis of the output of the Poisson model.

4.2.3. Experiment of negative binomial model

The negative binomial model is built to predict wafer yield as follows.

$$Y = \frac{1}{\left(1+D_0 A/a\right)^{a}} \tag{38}$$

In this model, Y means the defect-limited wafer yield, D_0 denotes the defect density, and A is the cluster coefficient. The yield was forecasted by negative binomial model and the results of 168 samples can be found in **Table 8**. The lineal correlation analysis between the actual wafer yield and the prediction value is shown in **Figure 10**.

Samples	The actual yield	The predicted yield	Relative error
1	0.72483	0.6395	0.11772
2	0.69128	0.57001	0.17543
3	0.8255	0.81257	0.015659
4	0.75168	0.69885	0.070289
5	0.75839	0.70919	0.064881
6	0.73154	0.64239	0.12187
7	0.89933	0.89708	0.002499
8	0.75168	0.69546	0.074787

Samples	The actual yield	The predicted yield	Relative error
9	0.74497	0.66949	0.10132
10	0.75168	0.66637	0.11349
11	0.7651	0.65417	0.14499
12	0.85235	0.85037	0.002326
13	0.89933	0.88098	0.020408
14	0.9396	0.93102	0.009135
15	0.81208	0.7737	0.047256
…	…	…	…
160	0.75258	0.61346	0.18486
161	0.86598	0.7748	0.10529
162	0.74227	0.5669	0.23626
163	0.8866	0.85746	0.032863
164	0.73196	0.50341	0.31225
165	0.7732	0.56152	0.27377
166	0.75258	0.52626	0.30072
167	0.82474	0.7748	0.06055
168	0.83505	0.76546	0.083336

Table 8. The predicted yield based on the negative binomial model.

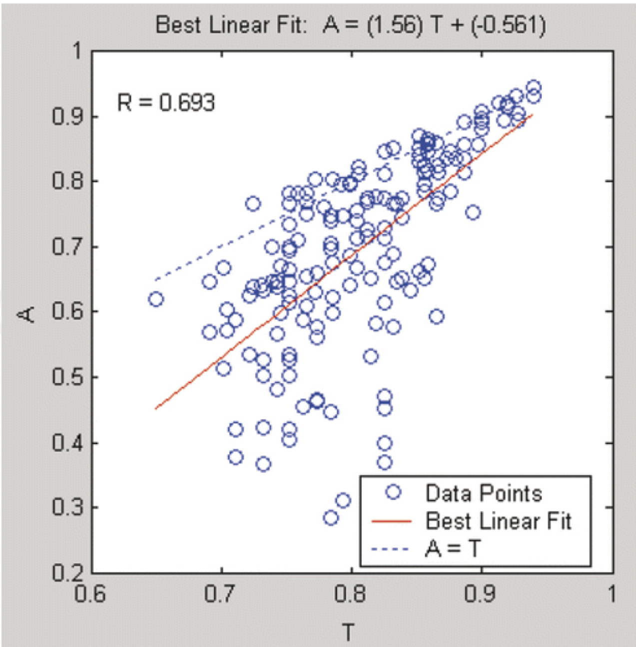


Figure 10. The linear regression analysis of the output of the negative binomial model.

http://dx.doi.org/10.5772/63444

4.2.4. Experiment of back-propagation neural network

A three layer BPNN is applied to predict wafer yield with ten input factors as same as the proposed FNN. The number of hidden neurons is determined by the empirical formula and selected to be 35. The yield was forecasted by BPNN and the results of 168 samples can be found in **Table 9**. The lineal correlation analysis between the actual wafer yield and the prediction value is shown in **Figure 11**.

Samples	The actual yield	The predicted yield	Relative error
1	0.72483	0.73495	0.013962
2	0.69128	0.7263	0.050661
3	0.8255	0.82385	0.001994
4	0.75168	0.76318	0.015293
5	0.75839	0.75985	0.001919
6	0.73154	0.75042	0.025807
7	0.89933	0.89876	0.000632
8	0.75168	0.76104	0.012456
9	0.74497	0.75932	0.019268
10	0.75168	0.73971	0.015921
11	0.7651	0.76666	0.002033
12	0.85235	0.83068	0.025423
13	0.89933	0.87383	0.028359
14	0.9396	0.93471	0.005204
15	0.81208	0.79469	0.021415
…	…	…	…
160	0.75258	0.72724	0.033671
161	0.86598	0.8581	0.009105
162	0.74227	0.7096	0.044007
163	0.8866	0.89151	0.005535
164	0.73196	0.68805	0.059996
165	0.7732	0.72722	0.059471
166	0.75258	0.71576	0.048929
167	0.82474	0.82683	0.00253
168	0.83505	0.82845	0.007907

Table 9. The predicted yield based on BPNN.

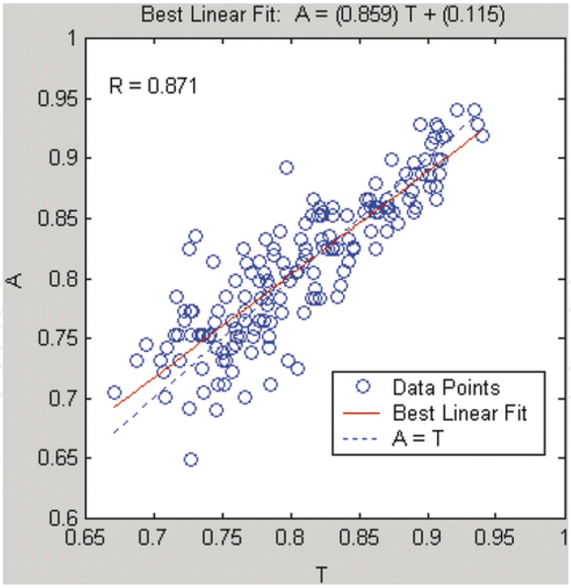


Figure 11. The linear regression analysis of the output of BPNN.

4.2.5. Results discussion

Aiming to assess the performance the proposed FNN methods, experiment with three contrast method was conducted for comparison. The lineal correlation analyses between the actual wafer yield and the prediction value of four methods are shown in **Figure 12**, which indicates that the FNN method outperforms other three methods from the view of convergence. The

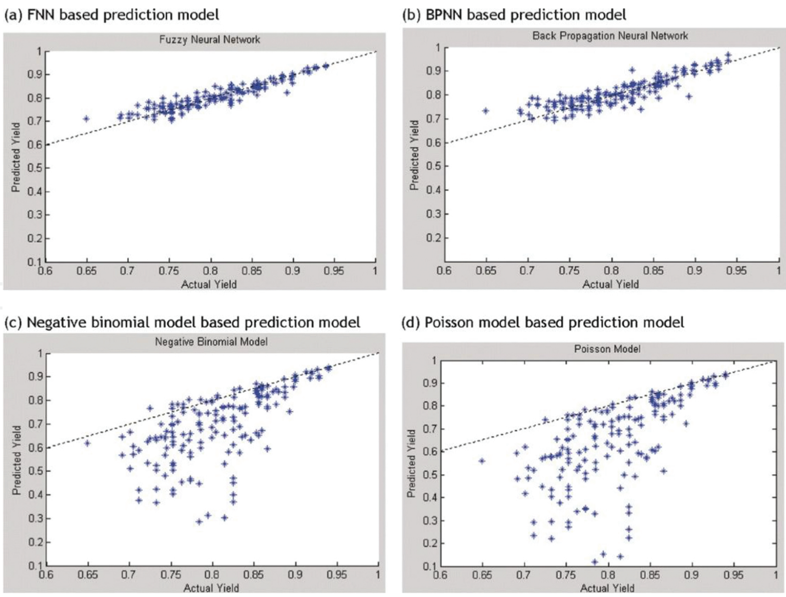


Figure 12. The relationship between the actual yields and predicted yields based on the FNN, BPNN and Poisson model and negative binomial model approach.

http://dx.doi.org/10.5772/63444

results of four methods in the RMSE and correlation coefficient *R* is presented in **Table 10**. The RMSE of the FNN method is 0.017, which is the smallest value above the four methods, and the *R* of the FNN-based model is 0.941, which is larger than other three methods. It indicates that the proposed FNN-based approach is more accurate and effective than other three methods, which are widely used in the yield predicting.

Yield prediction model	The actual yield		The predicted yield		RMSE	*R*
	Average	SD	Average	SD		
Poisson model	0.80864	0.06168	0.65394	0.18694	0.0169	0.637
Negative binomial model			0.70047	0.14789	0.0123	0.693
BPNN			0.80691	0.05736	0.0024	0.886
FNN			0.80838	0.05711	0.0017	0.941

Table 10. The comparisons of RMSE and correlation coefficients among the FNN, BPNN, Poisson model and negative binomial model.

5. Conclusion

The artificial neural networks (ANN) have a wide range of applications. For example, in complex discrete event manufacturing systems, they can be used to control, make decision and predict. SWFS is exactly such a complex manufacturing system. It has many characteristics, such as a mix of different process types, re-entrant flows, very expensive equipment and sequence dependent setup times and so on. In order to get more applications of ANN used in quality analysis and production scheduling in the semiconductor wafer fabrication system, this chapter implements two novel fuzzy neural networks that are used in the yield prediction of SWFS and rescheduling decision separately.

In the respect of rescheduling decision, this chapter puts forward a new method using a FNN model with which a system can make itself adapted to the current states and disturbances. In uncertain dynamic environments, current states and disturbances of the system are mathematically characterized. Rescheduling decision model, which assuming FNN builds the relationship between the inputs (i.e. disturbance, system state parameters) and the outputs (i.e. disturbance, system state parameters) of FNN. According to the current system disturbances, an optimal rescheduling method which can be used to schedule the semiconductor wafer fabrication lines is chosen by the make-decision model. We do experiment studies in Shanghai, which are based on 6-in. SWFS. The proposed rescheduling decision mechanism is proved to be effective by the linear regression between ideal targets and output of FNN. The rescheduling decision-making method which is proposed is demonstrated to be accurate by comparing with regression and traditional BPNN. We also do the comparison between the layered rescheduling method which is on the basis of FNN rescheduling decision mechanism and the two methods that are FCFS approach and the rescheduling approach based on the monolayer. The results indicate that, in respect of machine utilization and daily movement, layered rescheduling method, which is on the basis of FNN rescheduling decision mechanism, is superior to the other two approaches.

A yield prediction method for semiconductor manufacturing systems which is on the basis of new fuzzy neural networks is proposed for the yield prediction. This method builds the yield prediction model based on FNN by using the following parameters as input variables, which are the number of defects in each wafer, mean number of defects in each chip, mean number of defects in each unit area, clustering parameter, chip size and five critical electrical test parameters.

According to the data from the experiment studies in Shanghai which are based on 6-in. SWFS. The proposed rescheduling decision mechanism is proved to be effective by the linear regression between ideal targets and output of FNN. The rescheduling decision-making method which is proposed is demonstrated to be accurate by comparing with regression and traditional BPNN. The approach proposed in this paper has the advantage that it considers more variables' influences than other model such as negative binomial yield model, BPNN model and Poisson yield model. The variables here include physical parameters of wafer, key attributed parameters of defects and wafer electrical test parameters on wafer yield and so on. In a word, the model proposed in this paper is more accurate than the other traditional yield prediction approaches.

Acknowledgements

This work was supported by the State Key Program of the National Natural Science Foundation of China under Grant No. 51435009.

Author details

Jie Zhang*, Junliang Wang and Wei Qin

*Address all correspondence to: zhangjie_cims@hotmail.com

School of Mechanical Engineering, Shanghai Jiao Tong University, Shanghai, China

References

[1] Uzsoy, R., et al., A review of production planning and scheduling models in the semiconductor industry. Part II: shop-floor control IIE Transactions, 1994, 26(5): 44–55.

[2] Leachman, R.C., The competitive semiconductor manufacturing survey. IEEE international symposium on semiconductor manufacturing conference, 20–21 September 1993, Austin, Texas, USA. Piscataway, NJ: IEEE, 1993: 359–381.

[3] Uzsoy, R., et al., A review of production planning and scheduling models in the semiconductor industry. Part I: system characteristics, performance evaluation and production planning, IIE Transactions, 1992, 24(4): 47–60.

[4] Cheng, M., Sugi, M., Ota, J., Yamamoto, M., Ito, H., and Inoue, K., A fast rescheduling method in semiconductor manufacturing allowing for tardiness and scheduling stability. Proceeding of the 2006 IEEE, international conference on automation science and engineering, Shanghai, China, October, 2006: 7–10.

[5] Huang, H.P. and Chen, T.Y., A new approach to on-line rescheduling for a semiconductor foundry Fab. 2006 IEEE international conference on systems, man, and cybernetics, Taipei, China, October, 2006: 8–11.

[6] Kumar, R., Tiwari, M.K., and Allada, V., Modelling and rescheduling of a re-entrant wafer fabrication line involving machine unreliability, International Journal of Production Research, 2004, 42(21): 4431–4455.

[7] Maosn, S.J., Jin, S., and Wessels, M., Rescheduling strategies for minimizing total weighted tardiness in complex job shops, International Journal of Production Research, 2004, 42(3): 613–628.

[8] Toba, H., Segment-based approach for real-time reactive rescheduling for automatic manufacturing control, IEEE Transactions on Semiconductor Manufacturing, 2000, 13(3): 264–272.

[9] Tsai, C.J., and Huang, H.P., A real-time scheduling and rescheduling system based on RFID for semiconductor foundry fabs, Journal of the Chinese Institute of Industrial Engineers, 2007, 24(6): 437–444.

[10] Kumar, P.R., Scheduling semiconductor manufacturing plants, IEEE Control Systems Magazine, 1994, 14(6): 33–40.

[11] Kumar, N., et al., A review of yield modeling techniques for semiconductor manufacturing, International Journal of Production Research, 2006, 44(23): 5019–5036.

[12] Cunningham, S.P. and Spanos, C.J., Semiconductor yield improvement: results and best practices, IEEE Transactions on Semiconductor Manufacturing, 1995, 8(2): 103–109.

[13] Tong, L.-I. and Chao, L.-C., Novel yield model for integrated circuit with clustered defects, Expert Systems with Applications, 2008, 34: 2334–2341.

[14] Zhang, J., Qin, W., Wu, L.H., et al., Fuzzy neural network-based rescheduling decision mechanism for semiconductor manufacturing, Computers in Industry, 2014, 65(8): 1115–1125.

[15] Wu, L.H., Zhang, J., Fuzzy neural network based yield prediction model for semiconductor manufacturing system, International Journal of Production Research, 2010, 48(48): 3225–3243.

[16] Li, X.L., Yao, Y.X., and Yuan, Z.J., On-line tool condition monitoring system with wavelet fuzzy neural network, Journal of Intelligent Manufacturing, 1997, 8(4): 271–276.

[17] Zhou, Y.F., Li, S.J., Jin, R.C., A new fuzzy neural network with fast learning algorithm and guaranteed stability for manufacturing process control, Fuzzy Sets and Systems, 2002, 132(2): 201–216.

[18] Chang, P.C., Wang, Y.W., and Ting, C.J., A fuzzy neural network for the flow time estimation in a semiconductor manufacturing factory, International Journal of Production Research, 2008, 46(4): 1017–1029.

[19] Xu, D.M., and Yan, H.S., An intelligent estimation method for product design time, International Journal of Advanced Manufacturing Technology, 2006, 30(7–8): 601–613.

[20] Chang, Y.J., Kang, Y., Kang, Y., Hsu, C.L., Chang, C.T., and Chan, T.Y., Virtual metrology technique for semiconductor manufacturing. IEEE international conference on neural networks-conference proceedings, international joint conference on neural networks, Vancouver, BC, Canada, July, 2006: 5289–5293.

[21] Tong, L.-I., Lee, W.-I., and Su, C.-T., Using a neural network-based approach to predict the wafer yield in integrated circuit manufacturing, IEEE Transactions on Components, Packaging, and Manufacturing Technology – Part C, 1997, 20(4): 288–294.

[22] Tong, L.-I. and Chao, L.-C., Novel yield model for integrated circuit with clustered defects. Expert Systems with Applications, 2008, 34: 2334–2341.

[23] Chen, T. and Wang, M.J., A fuzzy set approach for yield learning modeling in wafer manufacturing, IEEE Transactions on Semiconductor Manufacturing, 1999, 12(2): 252–258.

[24] Chen, T. and Lin, Y.C., A fuzzy-neural system incorporating unequally important expert opinions for semiconductor yield forecasting, International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems, 2008, 16(1): 35–58.

[25] Chung, K.S. and Sang, C.P., A machine learning approach to yield management in semiconductor manufacturing, International Journal of Production Research, 2000, 38(17): 4261–4271.

[26] Kim, T.S., et al., Yield prediction models for optimisation of high-speed micro-processor manufacturing processes. 26th IEEE/CPMT international electronics manufacturing technology symposium, 2–3 October 2000, Santa Clara, CA, USA. Piscataway, NJ: IEEE, 2000: 368–373.

[27] Kim, T.S., Intelligent yield and speed prediction models for high-speed microprocessors. IEEE electronic components and technology conference, 28–31 May 2002, San Diego, CA, USA. Piscataway, NJ: IEEE, 2002: 1158–1162.

[28] Zhai, W.B., Chu, X.N., Zhang, J., Ma, D., Jin, Y., and Yan, J.Q., Research of combination auction based short-term scheduling technology of semiconductor fabrication line, Journal of Mechanical Engineering, 2004, 40(9): 95–99.

[29] Zhai, W.B., Zhang, J., Yan, J.Q., and Ma, D., Research of ETAEMS/GPGP-CN based on dynamic scheduling technology of semiconductor fabrication, Journal of Mechanical Engineering, 2005, 46(3): 53–58.

[30] Wang, J., and Malakooti, B., Feed-forward neural network for multiple criteria decision making, Computer & Operations Research, 1992, 19(2): 151–167.

[31] Malhotra, M.K., Sharma, S., and Nair, S.S., Decision making using multiple models, European Journal of Operational Research, 1999, 114(1): 1–14.

[32] Wong, A.Y., A statistical parametric and probe yield analysis methodology. Proceedings of IEEE international symposium on defect and fault tolerance in VLSI systems, 6–8 November 1996, Boston, MA, USA. Los Alamitos, CA: IEEE Comput. Soc. Press, 1996, 131–139.

[33] Jun, C.-H., et al., A simulation-based semiconductor chip yield model incorporating a new defect cluster index, Microelectronics Reliability, 1999, 39: 451–456.

[34] Cunningham, J.A., The use and evaluation of yield models in integrated circuit manufacturing, IEEE Transactions on Semiconductor Manufacturing, 1990, 3(2): 60–71.

[35] Aakash, T. and Bayoumi, M.A., Defect clustering viewed through generalised Poisson distribution, IEEE Transactions on Semiconductor Manufacturing, 1992, 5(3): 196–206.

[36] Koren, I., Koren, Z., and Stepper, C.H., A unified negative binomial distribution for yield analysis of defect-tolerant circuits. IEEE Transactions on Computers, 1993, 42: 724–734.