COMPLEX SYSTEM MODELING AND SIMULATION
ISSN 2096-9929 04/06 pp 218−231
Volume 1, Number 3, September 2021
DOI: 10.23919/CSMS.2021.0015

# Evaluation System and Correlation Analysis for Determining the Performance of a Semiconductor Manufacturing System

Qingyun Yu, Li Li, Hui Zhao, Ying Liu, and Kuo-Yi Lin*

**Abstract:** Numerous performance indicators exist for semiconductor manufacturing systems. Several studies have been conducted regarding the performance optimization of semiconductor manufacturing systems. However, because of the complex manufacturing processes, potential complementary or inhibitory correlations may exist among performance indicators, which are difficult to demonstrate specifically. To analyze the correlation between the performance indicators, this study proposes a performance evaluation system based on the mathematical significance of each performance indicator to design statistical schemes. Several samples can be obtained by conducting simulation experiments through the performance evaluation system. The Pearson correlation coefficient method and canonical correlation analysis are used on the received samples to analyze linear correlations between the performance indicators. Through the investigation, we found that linear and other complex correlations exist between the performance indicators. This finding can contribute to our future studies regarding performance optimization for the scheduling problems of semiconductor manufacturing.

**Key words:** semiconductor manufacturing system; system modeling; evaluation system; correlation analysis

## 1 Introduction

Semiconductor wafer fabrication is one of the most complex production processes in the industry[1, 2]. The primary fabrication process generally involves 250–500 processing steps and hundreds of different machines[3, 4] . The workflow of semiconductor manufacturing is displayed in Fig. 1.

Numerous performance factors act as crucial indicators for evaluating the semiconductor manufacturing system[5]. Several studies have been conducted on performance indicators and performance evaluation systems in Industry 4.0. For a performance-driven approach, the performance indicators of the manufacturing system include productivity, production capacity, production equilibrium, work-in-process (WIP) amount, queue length (QL), waiting time, equipment utilization (EU), equipment availability, equipment maintainability, reliability, flexibility, and integration. Qiao et al.[6] improved the performance system for fabrication plants (fabs), but they only introduced the concept of performance indicators. Hinrichs et al.[7] presented a solution to develop the performance of the measurement system. Hinrichs and Barke[8] focused on the performance management of semiconductor design processes.

Owing to the importance and current limitation of industrial performance, several studies have been conducted regarding the system development of semiconductor manufacturing. To prevent bottleneck starvation and high WIP amount on nonbottleneck machines, Zhou and Rose[9] investigated a bottleneck detection and dynamic dispatching strategy to regulate

• Qingyun Yu, Li Li, Hui Zhao, and Kuo-Yi Lin are with the Department of Control Science and Engineering, College of Electronics and Information Engineering, Shanghai Institute of Intelligent Science and Technology, and also with Shanghai Research Institute for Intelligent Autonomous Systems, Tongji University, Shanghai 201804, China. E-mail: 953717916@qq.com; lili@tongji.edu.cn; huizhao@tongji.edu.cn; 19603@tongji.edu.cn.
• Ying Liu is with the Department of Mechanical Engineering, School of Engineering, Cardiff University, The Parade, CF24 3AA, UK. E-mail: LiuY81@cardiff.ac.uk.
* To whom correspondence should be addressed.
Manuscript received: 2021-04-29; revised: 2021-06-18; accepted: 2021-06-25

© The author(s) 2021. The articles published in this open access journal are distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/).

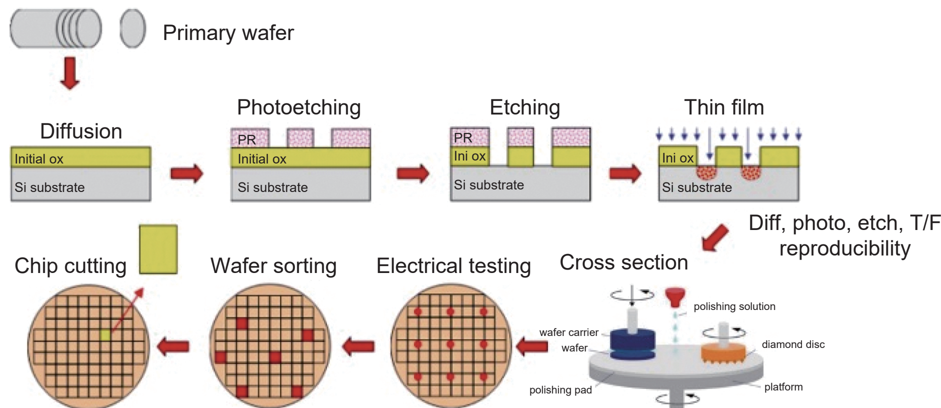


**Fig. 1 Workflow of semiconductor manufacturing.**

the workload of bottleneck and nonbottleneck machines. To achieve a satisfactory performance, Zhang et al.[10] designed a dynamic bottleneck dispatching policy to fabricate adaptive schemes according to the real-time conditions of the production line. Tsai and Chen[11] proposed a self-adaptive agent-based approach to enhance the performance of a semiconductor manufacturing factory. Every agent developed and modified its scheduling algorithm to adapt to the production environment in this approach. Yoon and Kim[12] proposed heuristic scheduling policies for semiconductor wafer fabrication facilities to reduce the variations in cycle times (CTs). These policies include a sequence control policy that involves a set of advanced operation due dates and an input release control policy for realizing an adaptive constant WIP level. Nemoto et al.[13] examined the financial benefits of CT reduction for developing a novel process using a stochastic simulation. The results demonstrated that even small decreases in CTs significantly influence the life cycle of a process. These existing studies have selected their optimization objectives through randomization or practical requirements and overlooked performance evaluation and correlations between performance indicators. These correlations are anfractuous and include complementary, inhibitory, and other complex correlations. Nevertheless, few systematic studies have conducted performance evaluations and correlation analyses in the semiconductor industry[14]. However, such studies exist in other manufacturing industries. Lin et al.[15] constructed a relationship framework of several key performance indicators by applying data mining techniques.

Currently, only a few scholars have studied the potential correlations between various performance indicators in the semiconductor manufacturing system. In the performance optimization process, the possible correlations among performance indicators should be first analyzed. All performance attributes can be balanced using this method, and thus, an optimal algorithm for performance optimization can be designed[16]. Many common methods have been developed for correlation analyses, such as the Pearson correlation coefficient method and canonical correlation analysis (CCA)[17]. The Pearson correlation coefficient method is a statistical method that is usually used in trend and classification analyses[18]. The CCA can measure the linear dependency between two sets of signal vectors obtained from two different data sources[19]. The linear dependence in the CCA is calculated using the cosines of principal angles between linear subspaces spanned by the respective signal vectors. Therefore, if two signal vector sets are linearly correlated, then a high correlation is obtained because the separate subspaces are very close together[20].

Data driving is widely used for performance optimization in manufacturing industries. Thus, the selection of data that should be chosen as the "driving source" is a significant scientific problem. To solve this problem, a correlation analysis on performance indicators should be conducted to provide a "data-driven/performance-driven" theoretical basis. Based on the aforementioned theoretical analysis, related literature, and actual investigations on semiconductor production in the industry, this study proposes a performance evaluation system for the semiconductor manufacturing system and conducted potential correlation analyses between performance indicators. In this study, the Pearson correlation coefficient method and CCA were used to analyze potential

correlations among performance indicators. The rest of this paper is organized as follows: Section 2 introduces the framework of the performance evaluation system and correlation analysis. Section 3 defines the performance of two categories and discusses the development of the performance evaluation system. Section 4 introduces the Pearson correlation coefficient method and CCA. Section 5 explains the acquisition of samples through simulation experiments and enumerates the simulation results generated in different manufacturing environments. Section 6 presents the conclusion.

## 2 Performance Evaluation and Correlation Analysis Framework

Performance evaluation is generally based on simulation and mathematical models, and performance is a summarization index used for determining the optimal releasing or scheduling policies. Performance evaluation results can be used to guide the manufacturing process. We implemented a performance evaluation method based on the simulation model of an industrial semiconductor manufacturing system. The evaluation method can provide the statistical values of performance indicators, assist managers in grasping the overall scheduling effect, and determine appropriate scheduling or releasing policies. Short-term performance indicators can be divided into two categories: fab-wide performance indicators and equipment-related performance indicators. These indicators are introduced in Sections 3.1 and 3.2, respectively.

The WIP amount refers to the number of wafers put into the production line but they have not undergone all processing operations. Researchers tend only to consider the WIP amount of the entire production line for common industrial manufacturing systems. Studies indicate that wafers must undergo multiple procedures for semiconductor manufacturing, which results in unbalanced workload levels (WLs) of particular equipment of each section or processing area. Thus, the homogeneous consideration of the WIP in the entire production line in semiconductor manufacturing cannot accurately measure the performance of the manufacturing system. We determined the WIPs of each processing section and the entire production line according to practical necessity. A processing section’s WIP includes wafers being processed and waiting to be processed in the buffers of the processing section. The WIP of the entire production line includes all wafers in all processing sections and transport. The evaluation and correlation analysis framework is illustrated in Fig. 2.

Our previous study developed a parallel simulation model based on numerous historical production data obtained from an industrial semiconductor manufacturing fab. This model is briefly introduced in Section 5. We can obtain an abundant amount of simulation data (performance samples) through this simulation model and common scheduling policies and use them for performance evaluation and potential correlation analysis.

## 3 Performance Evaluation

This study proposes a performance evaluation system to overcome the scheduling problem of semiconductor wafer manufacturing, as shown in Fig. 3.

Performance indicators can be classified into two categories based on their statistical period and physical meaning: short-term and long-term performance indicators. Short-term performance indicators are used to evaluate the influence of the daily scheduling scheme. They can be further classified into two categories: fab-wide performance indicators and equipment-related performance indicators. Long-term

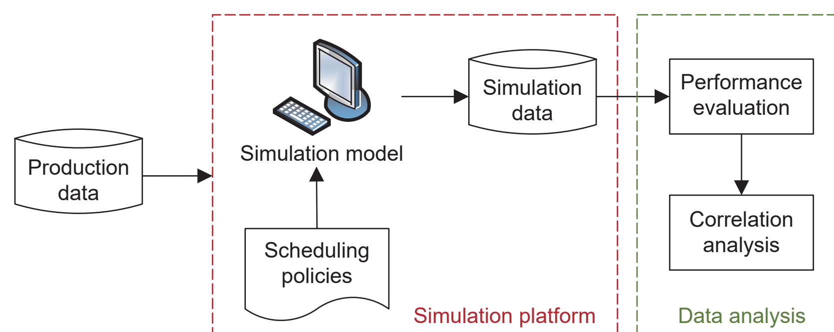


**Fig. 2 Framework of the performance evaluation and correlation analysis.**

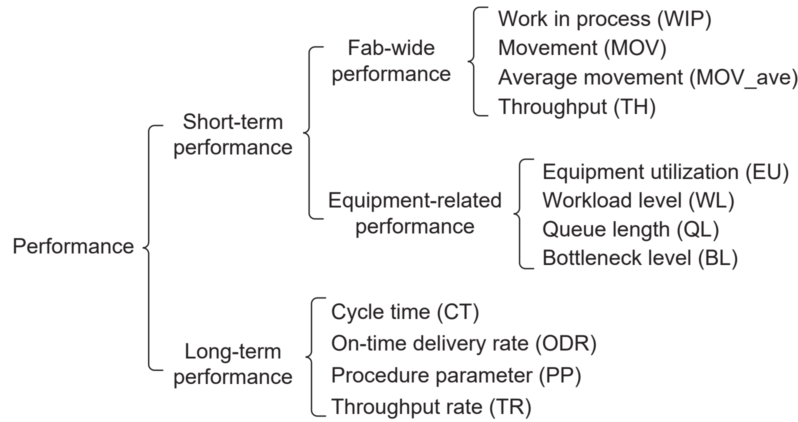


**Fig. 3 Performance evaluation system.**

performance indicators are measured or a year. The indicators are developed to supporta week, a month, several months, or a year and are used to guide the development of daily releasing policies.

## 3.1 Short-term performance indicators

Owing to the particularity of the semiconductor production line, the concept of processing sections is first introduced. In industrial production processes, the entire production line can be divided into different processing sections based on the corresponding production procedures, such as photolithography, oxidation, and etching sections. Particular equipment belongs to only one specific processing section, and equipment in a processing section may be functionally interchangeable or partially interchangeable. Therefore, if a wafer is processed in a processing section, then it is sent to another processing section. In this section, the wafer being processed waits in a queue by any ideal or relative ideal equipment, not a stationary one. This method can fully improve EU and production efficiency. Thus, several performance indicators related to processing areas exist, such as WIP amount, WL, and bottleneck level (BL).

### 3.1.1 WIP

WIP refers work in process. WIP is expressed as follows.

$$W_i = \sum_{j=1}^{n_i} W_{i,j} + W_{i,\text{wait}} \tag{1}$$

$$W_\text{f} = \sum_{i=1}^{n} W_i + W_\text{b} \tag{2}$$

where $W_i$ represents the number of wafers in the processing section $i$, $W_{i,j}$ represents the number of wafers being processed on equipment $j$ in the processing section $i$, $n_i$ represents the number of equipment in the processing section $i$, $W_{i,\text{wait}}$ represents the number of wafers waiting to be processed in the buffers of the processing section $i$, $W_\text{f}$ represents the total number of wafers in the entire production line, and $W_\text{b}$ represents the number of wafers in transport.

For the entire production line, the number of WIP should be medium built. If the WIP amount is very small, then the idle time of the equipment is high, and the EU is reduced. If the WIP amount is very large, then productivity is reduced, and the CT is prolonged.

### 3.1.2 Movement

Movement (MOV) represents the number of MOV steps. An “MOV step” refers to one processing step of a wafer on equipment. In a physical sense, each MOV of a wafer implies that one processing event is completed on the related equipment. The relationship between the MOV and equipment is shown in Eq. (3).

$$m_\text{o} = \sum_i \sum_j \delta_{ij} P_{ij} \tag{3}$$

where $m_\text{o}$ represents the MOV of all wafers, $\delta_{ij}$ represents the status of the $j$-th procedure on equipment $i$ (if the $j$-th procedure is completed, then $\delta_{ij} = 1$; else, $\delta_{ij} = 0$), and $P_{ij}$ denotes the number of wafers that have completed the $j$-th procedure on equipment $i$.

For a certain period of time and at a constant WIP value, the length of tasks’ waiting time and corresponding CT are reduced as the value of MOV increases.

### 3.1.3 Average MOV

Average MOV ($\text{MOV}_\text{ave}$) refers to the average MOVs of all wafers and can be obtained using MOV and the total WIP of the entire production line. $\text{MOV}_\text{ave}$ is expressed in Eq. (4). Owing to the changes in the WIP number over time, the “ total WIP” refers to the

"average WIP" during the scheduled period.

$$m_{\mathrm{a}} = \frac{m_{\mathrm{o}}}{W_{\mathrm{f}}} \tag{4}$$

where $m_{\mathrm{a}}$ represents the $\mathrm{MOV_{ave}}$ of all wafers. In the same schedule period, the scheduling policy improves with an increase in $m_{\mathrm{a}}$.

#### 3.1.4 Throughput

Throughput (TH) represents the number of wafers wholly processed in a certain period. A higher TP value indicates a higher machine utilization and lower waiting time for wafers.

#### 3.1.5 EU

EU refers to the ratio of the actual processing time to the available time of the equipment and is expressed in Eq. (5):

$$\mathrm{EU} = \frac{\sum_{h=1}^{m} T_h}{T_{\mathrm{oP}}} \times 100\% \tag{5}$$

where EU represents equipment utilization, $T_h$ represents the elapsed time of the $h$-th procedure, $m$ represents the number of procedures on the equipment, and $T_{\mathrm{oP}}$ represents the available time of the equipment.

The semiconductor manufacturing industry is a capital-intensive industry that requires a considerable amount of investment for equipment and has high operating costs. Thus, investors assign considerable significance to investment efficiency, are unwilling to waste any equipment production capacity, and constantly pursue full utilization of all equipment.

#### 3.1.6 WL

The WL refers to the processing time required for the queuing wafers in the buffer of the equipment. The WL for the non-batch-processing equipment can be computed by the sum of the processing time needed by each queuing wafer in the buffer. For the batch-processing equipment, queuing wafers have to be put in batches using a specific algorithm. Then, the total required processing time should be computed.

In actual semiconductor manufacturing, equipment in the same processing section can be functionally interchanged or partially interchanged. Thus, the ideology of the WL can be extended to processing sections. In this study, we introduce another concept, i.e., available capacity[21]. In addition to downtime and preventive maintenance, processing abnormal wafers in tasks, such as engineering jobs and profile control jobs, can also influence the capacity from the overall perspective of equipment efficiency. Thus, in the computation of the available capacity, we must deduct the capacity of abnormal wafers. Moreover, we should reserve some protective capacity to maintain the stability of the manufacturing system. The available daily capacity of equipment can be denoted as $A_{\mathrm{c}}$ and can be computed using Eq. (6)[22]:

$$A_{\mathrm{c}} = \left(1 - D_{\mathrm{t}} - P_{\mathrm{m}} - E_{\mathrm{g}} - M_{\mathrm{d}} - P_{\mathrm{c}}\right) \times 1440 \tag{6}$$

where $D_{\mathrm{t}}$ represents the proportion of downtime, $P_{\mathrm{m}}$ represents the proportion of preventive maintenance, $E_{\mathrm{g}}$ denotes the proportion of the processing time of engineering wafers, $M_{\mathrm{d}}$ represents the proportion of the processing time of profile control wafers, and $P_{\mathrm{c}}$ indicates the proportion of the protective capacity.

If the WL is larger than the corresponding available capacity, then the equipment or the processing section is regarded as overloaded. Moreover, if the WL value is zero, then the equipment or the processing section is considered to be underloaded. Otherwise, the equipment or processing section is deemed to have a balanced load.

#### 3.1.7 QL

QL refers to the number of waiting wafers in the buffer of the equipment. The concept of QL can also be extended to processing sections. Here we expound on the concept of" available capacity QL". The relationship between available capacity QL and available capacity is expressed in Eq. (7):

$$A_{\mathrm{c}} = A_{\mathrm{q}} \cdot \bar{t}_{\mathrm{pro}} \tag{7}$$

where $A_{\mathrm{q}}$ represents the available capacity QL and $\bar{t}_{\mathrm{pro}}$ represents the average processing time of queuing wafers.

If the capacity of the equipment of a processing section is less than the equipment of other processing sections or the equipment of a processing section has many wafers that require high-speed processing or time-consuming processing procedures, then the equipment of the processing section is identified as a bottleneck. In this case, the processing time of the wafers largely depends on the BL of the equipment or processing section and dispatching rules. In an actual manufacturing system, the bottleneck phenomenon implies that the equipment of a processing section is overloaded because most equipment in various processing sections is underloaded or has a balanced load. Bottlenecks are often associated with an unbalanced EU and prolonged CT. Thus, the bottleneck phenomenon should be avoided as much as possible. We can determine whether the equipment of a

processing area is a bottleneck or not using Eq. (8):

$$B(t) = \begin{cases} 1, & L \geqslant A_c; \\ 0, & L < A_c \end{cases} \tag{8}$$

where $B(t) = 1$ denotes that the equipment of a processing area is a bottleneck and $B(t) = 0$ denotes that the equipment of a processing area is not a bottleneck.

**3.1.8 BL**

BL is the ratio of the bottleneck time to the available time. It can be computed using Eq. (9):

$$P_b = \left[ \frac{\int_0^1 B(t)\mathrm{d}t}{t_1 - t_0} \right] \times 100\% \tag{9}$$

where $P_b \leqslant 1$ is always tenable. If $P_b = 1$, then the equipment of a processing area is an abiding bottleneck. If $P_b < 1$, then the equipment of a processing area is an instantaneous bottleneck. In a complex semiconductor manufacturing system, some equipment is initially regarded as abiding bottlenecks due to the high prices and lengthy processing time. The BL of abiding bottlenecks should be alleviated, and instantaneous bottlenecks should be avoided by adjusting scheduling algorithms.

### 3.2 Long-term performance indicators

**3.2.1 CT**

CT refers to the time span from when a wafer is put into the production line to when all procedures of the wafer are completed. CTs are different for different wafers. Due to the inevitable waiting time in the industrial manufacturing system, the actual CT is often several times longer than its net processing time. In semiconductor manufacturing, CT is a key factor for enterprises to maintain competitiveness in the market. Thus, one of the most crucial targets for wafer fabrication enterprises is to reduce CT by using excellent release policies, dispatching rules, scheduling policies, and other effective methods.

**3.2.2 On-time delivery rate**

On-time delivery rate (ODR) can reflect the completion degree of production tasks, which can be computed using Eq. (10). ODR denotes the ratio of the number of the on-time delivered wafers to all completed wafers:

$$\mathrm{ODR}_{z,T_d} = \frac{n_1}{n_1 + n_2} \tag{10}$$

where $\mathrm{ODR}_{z,T_d}$ refers to the ODR of Product $z$ in the time period of $T_d$ (the time span of the simulation period), $n_1$ refers to the number of products delivered on time, and $n_2$ refers to the number of products lingeringly delivered.

**3.2.3 Procedure parameter**

Here, we adapted the ratio of the actual CT to the net processing time as the procedure parameter (PP), which can be computed using Eq. (11):

$$\mathrm{PP} = \frac{T_{pro}}{\sum_{l=1}^{p} T_l} \tag{11}$$

where PP represents the procedure parameter, $T_{pro}$ represents the actual CT, $T_l$ represents the required processing time of the $l$-th procedure, and $p$ represents the number of all procedures. Clearly, $\mathrm{PP} \geqslant 1$. The smaller the PP value, the shorter the waiting time.

**3.2.4 Through rate**

Through rate (TR) denotes the number of completed wafers in one unit time and can directly reflect the effect of release policies. TR can be computed using Eq. (12):

$$\mathrm{TR} = \frac{P_{out}}{T_d} \tag{12}$$

where TR represents the throughout rate and $P_{out}$ represents the number of completed wafers during the simulation period. The improvement of TR results in a high CT. That is, a tradeoff exists between the improvement in the TR and the reduction in the CT.

Then, the CCA maximizes the following function.

## 4 Correlation Analysis Methods

Numerous performance indicators are involved in a fab. Thus, the potential correlations between the indicators should be analyzed. In this study, we used two correlation analysis methods to analyze the correlations between the indicators: the Pearson correlation coefficient method and CCA.

### 4.1 Pearson correlation coefficient method

The Pearson correlation coefficient method is a statistical method that is used to accurately measure the linear correlation and correlation degree between two variables. For two variables $X$ and $Y$, two sets of data are collected through simulation experiments: $X = [x_1, x_2, \ldots, x_n]$ and $Y = [y_1, y_2, \ldots, y_n]$. The correlation coefficients of the two sets can be obtained by solving the formula of the Pearson correlation coefficient. Many forms of formulas are used to express correlation coefficients. The most common one is covariance, which is expressed in Eq. (13):

$$\rho_{xy} = \frac{E(XY) - E(X)E(Y)}{\sigma_X \sigma_Y} \tag{13}$$

where $E(XY)$ represents the average value of the sum obtained after multiplying the corresponding elements in the two sets of data, $E(X)$ represents the average value of sample $X$, $E(Y)$ denotes the average value of sample $Y$, $\sigma_X$ represents the standard deviation of sample $X$, $\sigma_Y$ represents the standard deviation of sample $Y$, and $\rho_{xy}$ is the correlation coefficient between $X$ and $Y$ (i.e., Pearson correlation coefficient).

The value of the Pearson correlation coefficient is between −1 and 1. When $\rho_{xy}$ is larger than 0 ($\rho_{xy} > 0$), a positive correlation is observed between $X$ and $Y$. That is, when the value of one variable increases, the value of the other also increases. When $\rho_{xy}$ is less than 0 ($\rho_{xy} < 0$), a negative correlation is observed between $X$ and $Y$. Hence, the value of the other decreases when the value of one variable increases[23].

The correlation is generally defined as follows:

$0.8 < \rho_{xy} \leqslant 1.0$ means an extremely strong correlation;

$0.6 < \rho_{xy} \leqslant 0.8$ means a strong correlation;

$0.4 < \rho_{xy} \leqslant 0.6$ means a moderate correlation;

$0.2 < \rho_{xy} \leqslant 0.4$ means a weak correlation;

$0 \leqslant \rho_{xy} \leqslant 0.2$ means an extremely weak or no correlation.

### 4.2 CCA

The CCA was first introduced to analyze linear relations between two sets of variables[24]. Canonical correlation problems occur in multivariate data analyses that can be naturally divided into two blocks of observed variables.

Consider two variables $x \in \mathbb{R}^p$ and $y \in \mathbb{R}^q$. Then, the CCA finds pairs of directions $w_x$ and $w_y$ that maximize the correlation between the projections $u = w_x^{\mathrm{T}} x$ and $v = w_y^{\mathrm{T}} y$. These projections $u$ and $v$ are named as canonical variates. The pairs of directions $w_x$ and $w_y$ are also known as canonical loadings. Formally, we seek two linear combinations, $u$ and $v$, which correlate the most.

$$u = w_{x_1} x_1 + \cdots + w_{x_p} x_p = w_x^{\mathrm{T}} x \tag{14}$$

$$v = w_{y_1} y_1 + \cdots + w_{y_p} y_p = w_y^{\mathrm{T}} y \tag{15}$$

Then, the CCA maximizes the following function:

$$\rho = \frac{E[xy]}{\sqrt{E[x^2]E[y^2]}} = \frac{E\left[w_x^{\mathrm{T}} x y^{\mathrm{T}} w_y\right]}{\sqrt{E\left[w_x^{\mathrm{T}} x x^{\mathrm{T}} w_x\right] E\left[w_y^{\mathrm{T}} y y^{\mathrm{T}} w_y\right]}} \tag{16}$$

$$\rho = \frac{w_x^{\mathrm{T}} C_{xy} w_y}{\sqrt{w_x^{\mathrm{T}} C_{xx} w_x w_y^{\mathrm{T}} C_{yy} w_y}} \tag{17}$$

where $C_{xx} \in \mathbb{R}^{p\times p}$ and $C_{yy} \in \mathbb{R}^{q\times q}$ are the within-set covariance matrices of $x$ and $y$, respectively. Moreover, $C_{xy} \in \mathbb{R}^{p\times q}$ denotes their between-set covariance matrix. The CCA finds the directions $w_x$ and $w_y$ that maximize the correlation between corresponding projections by solving the following formula:

$$\begin{aligned} &\max_{w_x, w_y} \quad w_x^{\mathrm{T}} C_{xy} w_y \\ &\text{subject to} \quad w_x^{\mathrm{T}} C_{xx} w_x = 1 \ \&\ w_y^{\mathrm{T}} C_{yy} w_y = 1 \end{aligned} \tag{18}$$

By applying the Lagrange multiplier technique to the optimization formula Eq. (5), we obtain the following:

$$\max \begin{aligned} &w_x^{\mathrm{T}} C_{xy} w_y - \frac{\lambda_x}{2}\left(w_x^{\mathrm{T}} C_{xx} w_x - 1\right) \\ &-\frac{\lambda_y}{2}\left(w_y^{\mathrm{T}} C_{yy} w_y - 1\right) \end{aligned} \tag{19}$$

By taking derivatives with respect to $w_x$ and $w_y$, Eq. (20) is obtained:

$$C_{xy} w_y - \lambda_x C_{xx} w_x = 0 \ \&\ C_{yx} w_x - \lambda_y C_{yy} w_y = 0 \tag{20}$$

By subtracting $w_x^{\mathrm{T}}$ times the first from $w_y^{\mathrm{T}}$ times the second in Eq. (20), the following equation is obtained:

$$\lambda_x w_x^{\mathrm{T}} C_{xx} w_x - \lambda_y w_y^{\mathrm{T}} C_{yy} w_y = 0 \tag{21}$$

When the two constraints in Formula (18) are considered, $\lambda_x = \lambda_y$. By using $\lambda$ to denote this value, we obtain the following equation for computing the directions of the maximal correlation:

$$\begin{pmatrix} 0 & C_{xy} \\ C_{yx} & 0 \end{pmatrix} \begin{pmatrix} w_x \\ w_y \end{pmatrix} = \lambda \begin{pmatrix} C_{xx} & 0 \\ 0 & C_{yy} \end{pmatrix} \begin{pmatrix} w_x \\ w_y \end{pmatrix} \tag{22}$$

The eigenvectors $w_x^i$ and $w_y^i$ represent the directions, where $i = 1, 2, \ldots, I$ and $I = \min\{p, q\}$. The eigenvector $\lambda_i$ is the correlation coefficient $\rho_i$ because of the following:

$$\rho = w_x^{\mathrm{T}} C_{xy} w_y = \lambda_x w_x^{\mathrm{T}} C_{xx} w_x = \lambda \tag{23}$$

The CCA is capable of finding a pair of linear combinations for two sets that have a maximum correlation coefficient, $\rho_1$. Then, the second pair is uncorrelated with the first pair of canonical variables and has the second biggest correlation coefficient $\rho_2$, which can be determined using the CCA.

## 5 Numerical Experiments and Analysis

### 5.1 Experiment description

A simulation model of an industrial production line (a mixed 5-inches and 6-inches production line of a semiconductor manufacturer in Shanghai) was used to acquire simulation samples. The production line has an average WIP of more than 80 000 workpieces, and it

has 11 processing sections: oxidation, photolithography, injection, epitaxial growth, dry etching, deposition, sputtering, wet cleaning, and three non-processing areas (i.e., virtual machine, testing, and outsourcing). The oxidation and photolithography sections are bottleneck processing sections. Moreover, the line has a large number (over 800) of machines with five different types: single-workpiece processing machines, batch-processing machines, multi-workpiece processing machines, cluster tools, and tanks.

To make a comprehensive correlation analysis for the performance indicators of a semiconductor scheduling system, in this study, we designed four kinds of numerical experiment scenarios: different processing areas, different working conditions, different scheduling rules, and different products. In addition, we simultaneously adopted the Pearson correlation coefficient method and CCA method to avoid the possible contingency of a single approach.

### 5.2 Correlation analysis based on different processing areas

In semiconductor manufacturing systems, two main processing areas cover most processing steps: oxidation area and photolithography area. This section will analyze correlations among EU and correlations among queuing length in two main processing areas.

We obtained the average daily utilization of all equipment. We selected 15 pieces of equipment whose average utilization surpasses 85%. The 15 selected equipment belonged to two processing sections: photolithography and oxidation sections. The 15 selected pieces of equipment are listed in Table 1.

We computed the correlation coefficients among the 15 EUs presented in the following symmetrical matrix based on the Pearson correlation coefficient method, as shown in Table 2.

Here, $a_{ij}$ in the matrix denotes the correlation coefficient between the $i$- th and $j$- th elements. For example, $a_{53} = 0.850$ denotes that the correlation coefficient between the utilization of equipment_5ME07 and equipment_8ME05 is 0.850 (solid correlation).

From the above matrix, we can obtain the following conclusions:

(1) The equipment that belongs to the same processing section has a higher correlation coefficient of EU. Considering the photolithography section as an example, the correlation coefficients of the EU are

**Table 1 15 selected equipment.**

| Photolithography area | | Oxidation area | |
|---|---|---|---|
| No. | Equipment ID | No. | Equipment ID |
| 1 | 7001 | 9 | 3201 |
| 2 | 4V01 | 10 | 3206 |
| 3 | 8ME05 | 11 | 3207 |
| 4 | 4ME20 | 12 | 2FU03 |
| 5 | 5ME07 | 13 | 2FU62 |
| 6 | 8ME11 | 14 | 2FU26 |
| 7 | 8SU09 | 15 | 2FU38 |
| 8 | 7135 | | |

**Table 2 Correlation coefficient among the 15 equipment.**

| No. | Correlation coefficient | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
| 1 | 1 | 0.725 | 0.924 | 0.826 | 0.823 | 0.772 | 0.828 | 0.015 | 0.215 | 0.004 | 0.140 | 0.184 | 0.326 | 0.191 | 0.183 |
| 2 | 0.725 | 1 | 0.697 | 0.792 | 0.770 | 0.696 | 0.762 | 0.096 | 0.029 | 0.186 | 0.018 | 0.089 | 0.140 | 0.147 | 0.245 |
| 3 | 0.924 | 0.697 | 1 | 0.808 | 0.850 | 0.783 | 0.844 | 0.289 | 0.342 | 0.060 | 0.395 | 0.116 | 0.214 | 0.318 | 0.140 |
| 4 | 0.826 | 0.792 | 0.808 | 1 | 0.850 | 0.883 | 0.917 | 0.327 | 0.010 | 0.065 | 0.198 | 0.043 | 0.093 | 0.354 | 0.321 |
| 5 | 0.823 | 0.770 | 0.850 | 0.850 | 1 | 0.797 | 0.873 | 0.371 | 0.330 | 0.356 | 0.272 | 0.436 | 0.037 | 0.108 | 0.045 |
| 6 | 0.772 | 0.696 | 0.783 | 0.883 | 0.797 | 1 | 0.910 | 0.119 | 0.043 | 0.112 | 0.167 | 0.067 | 0.049 | 0.400 | 0.344 |
| 7 | 0.828 | 0.762 | 0.844 | 0.917 | 0.873 | 0.910 | 1 | 0.190 | 0.275 | 0.138 | 0.174 | 0.234 | 0.264 | 0.298 | 0.155 |
| 8 | 0.015 | 0.096 | 0.289 | 0.327 | 0.371 | 0.119 | 0.190 | 1 | 0.813 | 0.770 | 0.653 | 0.816 | 0.743 | 0.821 | 0.883 |
| 9 | 0.215 | 0.029 | 0.342 | 0.010 | 0.330 | 0.143 | 0.275 | 0.813 | 1 | 0.694 | 0.659 | 0.782 | 0.657 | 0.646 | 0.836 |
| 10 | 0.004 | 0.186 | 0.060 | 0.065 | 0.356 | 0.112 | 0.138 | 0.770 | 0.694 | 1 | 0.750 | 0.844 | 0.715 | 0.785 | 0.853 |
| 11 | 0.140 | 0.081 | 0.395 | 0.198 | 0.272 | 0.167 | 0.174 | 0.653 | 0.659 | 0.750 | 1 | 0.780 | 0.643 | 0.763 | 0.779 |
| 12 | 0.184 | 0.089 | 0.116 | 0.043 | 0.436 | 0.067 | 0.234 | 0.816 | 0.782 | 0.844 | 0.780 | 1 | 0.831 | 0.848 | 0.676 |
| 13 | 0.326 | 0.140 | 0.214 | 0.093 | 0.037 | 0.049 | 0.264 | 0.743 | 0.657 | 0.715 | 0.643 | 0.831 | 1 | 0.792 | 0.644 |
| 14 | 0.191 | 0.147 | 0.318 | 0.354 | 0.108 | 0.400 | 0.298 | 0.821 | 0.646 | 0.785 | 0.763 | 0.848 | 0.792 | 1 | 0.905 |
| 15 | 0.183 | 0.245 | 0.140 | 0.321 | 0.045 | 0.344 | 0.155 | 0.883 | 0.836 | 0.853 | 0.779 | 0.676 | 0.644 | 0.905 | 1 |

higher than 0.7 (strong correlation or extremely strong correlation). The same result is observed for the oxidation section. Because equipment in the same processing section is functionally interchangeable or partially interchangeable, their EUs are closely related.

(2) The equipment that belongs to the disparate processing sections has lower correlation coefficients of the EU. Consider the photolithography area and oxidation area as an example. The correlation coefficients of the EU are lower than 0.5 (weak correlation or extremely weak correlation or uncorrelated).

(3) Based on the strong correlation of EU in the same processing section, the utilization of the processing section can be reasonably defined as the average utilization of all equipment in the processing section.

### 5.3 Correlation analysis based on different WLs

To analyze the correlations among the performance indicators in detail, we obtained the values of WIP, MOV, TH, average EU in the photolithography section ($EU_P$), average EU in the oxidation section ($EU_O$), CT, and ODR under different WLs (underloaded, fully loaded, and overloaded). The results are presented in Table 3.

Similarly, to avoid occasionality, we recalculated correlations based on the simple coefficient method and recorded the corresponding performances and correlations.

We can draw the following conclusions from the experimental results:

(1) The correlation coefficients between performance indicators exhibit a slight change irrespective of the WL.

(2) The correlation coefficient between the MOV and CT decreases with an increase in the WL.

(3) The correlation coefficients between the short-term and long-term performances are influenced by the WL.

(4) MOV has the highest correlation with other performance indicators, especially with ODR. By contrast, CT has the lowest correlation with other performance indicators. More than 100 products exist in the manufacturing system, and their CTs are considerably different from one another. The CT in this study refers to the average CT of all products. Even with different WL, the value of CT does not change considerably.

(5) The overall correlation coefficients in the overloaded WL are smaller than those in the underloaded and fully loaded levels. The presence of

**Table 3 Correlations among performances in different workload levels.**

| Workload level | | WIP | MOV | TH | $EU_P$ | $EU_O$ | CT | ODR |
|---|---|---|---|---|---|---|---|---|
| Under-loaded | WIP | 1 | | | | | | |
| | MOV | 0.372 | 1 | | | | | |
| | TH | 0.346 | **0.595** | 1 | | | | |
| | $EU_P$ | 0.255 | **0.685** | 0.128 | 1 | | | |
| | $EU_O$ | 0.226 | **0.643** | −0.149 | 0.178 | 1 | | |
| | CT | −0.188 | 0.418 | 0.282 | 0.106 | 0.165 | 1 | |
| | ODR | −0.249 | **0.784** | 0.313 | 0.272 | 0.232 | **0.530** | 1 |
| Full-loaded | WIP | 1 | | | | | | |
| | MOV | 0.345 | 1 | | | | | |
| | TH | 0.276 | **0.562** | 1 | | | | |
| | $EU_P$ | 0.177 | **0.651** | 0.075 | 1 | | | |
| | $EU_O$ | 0.172 | **0.615** | −0.133 | 0.145 | 1 | | |
| | CT | −0.135 | 0.386 | 0.254 | 0.074 | 0.138 | 1 | |
| | ODR | −0.234 | **0.755** | 0.288 | 0.245 | 0.265 | **0.558** | 1 |
| Over-loaded | WIP | 1 | | | | | | |
| | MOV | 0.226 | 1 | | | | | |
| | TH | 0.152 | 0.445 | 1 | | | | |
| | $EU_P$ | 0.058 | **0.537** | −0.042 | 1 | | | |
| | $EU_O$ | 0.052 | 0.498 | −0.254 | 0.021 | 1 | | |
| | CT | −0.254 | 0.262 | 0.137 | −0.041 | 0.017 | 1 | |
| | ODR | −0.359 | **0.634** | 0.161 | 0.129 | 0.083 | 0.417 | 1 |

many wafers in the production line queue causes congestion that leads to nonlinear characteristics, which reduce the value of correlation coefficients.

### 5.4 Correlation analysis based on different scheduling rules

Similar to Section 5.3, we obtained values of MOV, TH, EU, CT, and ODR under different scheduling policies (first in, first out; earliest due date; shortest processing time; critical ratio; and least slack). The results of the correlation coefficients are listed in Table 4.

Similarly, to avoid occasionality, we recalculated the correlations based on the simple coefficient method and recorded the corresponding performances and correlations, as follows.

$$\text{FIFO}:\begin{bmatrix} 1 & 0.46 & 0.63 & -0.48 & 0.49 \\ 0.46 & 1 & 0.23 & -0.27 & -0.32 \\ \mathbf{0.63} & 0.23 & 1 & -0.23 & 0.4 \\ -0.48 & -0.27 & -0.23 & 1 & -0.65 \\ 0.49 & -0.32 & 0.40 & \mathbf{-0.65} & 1 \end{bmatrix},$$

$$\text{EDD}:\begin{bmatrix} 1 & 0.69 & 0.59 & -0.57 & 0.28 \\ \mathbf{0.69} & 1 & 0.17 & -0.20 & -0.26 \\ \mathbf{0.59} & 0.17 & 1 & -0.28 & 0.27 \\ \mathbf{-0.57} & -0.20 & -0.28 & 1 & -0.61 \\ 0.28 & -0.26 & 0.27 & \mathbf{-0.61} & 1 \end{bmatrix},$$

$$\text{SPT}:\begin{bmatrix} 1 & 0.52 & 0.53 & -0.66 & 0.47 \\ \mathbf{0.52} & 1 & 0.09 & -0.23 & -0.27 \\ \mathbf{0.53} & 0.09 & 1 & -0.20 & 0.19 \\ \mathbf{-0.66} & -0.23 & -0.20 & 1 & -0.65 \\ 0.47 & -0.27 & 0.19 & \mathbf{-0.65} & 1 \end{bmatrix},$$

$$\text{CR}:\begin{bmatrix} 1 & 0.52 & 0.54 & -0.45 & 0.42 \\ \mathbf{0.52} & 1 & 0.04 & -0.09 & -0.29 \\ \mathbf{0.54} & 0.04 & 1 & 0.04 & 0.16 \\ -0.45 & -0.09 & 0.04 & 1 & -0.75 \\ 0.42 & -0.29 & 0.16 & \mathbf{-0.75} & 1 \end{bmatrix},$$

$$\text{LS}:\begin{bmatrix} 1 & 0.59 & 0.49 & -0.58 & 0.40 \\ \mathbf{0.59} & 1 & 0.25 & -0.23 & -0.26 \\ 0.49 & 0.25 & 1 & -0.26 & 0.28 \\ \mathbf{-0.58} & -0.23 & -0.26 & 1 & -0.76 \\ 0.40 & -0.26 & 0.28 & \mathbf{-0.76} & 1 \end{bmatrix}.$$

**Table 4 Correlation coefficients among performances in different scheduling policies.**

| Scheduling policie | | MOV | TH | EU | CT | ODR |
|---|---|---|---|---|---|---|
| FIFO (first in first out) (prioritize the wafer which enters the buffer first) | MOV | 1 | | | | |
| | TH | **0.51** | 1 | | | |
| | EU | **0.58** | 0.15 | 1 | | |
| | CT | **−0.55** | −0.17 | −0.32 | 1 | |
| | ODR | 0.40 | −0.34 | 0.31 | **−0.67** | 1 |
| EDD (earliest due date) (prioritize the wafer with the earliest due date) | MOV | 1 | | | | |
| | TH | **0.60** | 1 | | | |
| | EU | **0.53** | 0.23 | 1 | | |
| | CT | **−0.51** | −0.17 | −0.18 | 1 | |
| | ODR | 0.36 | −0.28 | 0.28 | **−0.70** | 1 |
| SPT (shortest processing time) (prioritize the wafer with the shortest processing time) | MOV | 1 | | | | |
| | TH | **0.54** | 1 | | | |
| | EU | 0.46 | 0.06 | 1 | | |
| | CT | **−0.56** | −0.17 | −0.11 | 1 | |
| | ODR | 0.41 | −0.29 | 0.21 | **−0.71** | 1 |
| CR (critical ratio) (prioritize the wafer with the smallest critical ratio) | MOV | 1 | | | | |
| | TH | **0.59** | 1 | | | |
| | EU | **0.63** | −0.03 | 1 | | |
| | CT | **−0.53** | −0.14 | −0.06 | 1 | |
| | ODR | 0.47 | −0.30 | 0.24 | **−0.75** | 1 |
| LS (least slack) (prioritize the wafer with the shortest waiting time) | MOV | 1 | | | | |
| | TH | **0.59** | 1 | | | |
| | EU | **0.54** | 0.18 | 1 | | |
| | CT | **−0.52** | −0.16 | −0.27 | 1 | |
| | ODR | 0.40 | −0.27 | 0.31 | **−0.76** | 1 |

We can draw four conclusions from the above numerical experiments:

(1) The results of the correlation coefficients are inconspicuously different under different policies. Different scheduling policies could cause different characteristic performances owing to the distinctive scheduling characteristic.

(2) For wafers, the possibility of a timely delivery decreases with an increase in the CT. Thus, negative correlations exist between CT and ODR.

(3) MOV has a positive moderate or strong correlation with TH and EU. Thus, MOV can represent the entire manufacturing system's operation condition and can improve other performance indicators.

(4) CT refers to the cycle time of wafers, and it has negative correlations with MOV, TH, EU, and ODR. The performance of CT improves with an increase in the value of CT.

## 5.5 Correlation analysis based on different products

CT and ODR are the two most important performances related to wafers in the semiconductor manufacturing system. To further study the two performances, we conducted a correlation analysis on the two performances based on different kinds of products.

Based on the obtained numerous samples, the monthly TH was counted and arranged in descending order. Eight kinds of products with the highest monthly TH were selected as the objects and denoted as Product A, Product B, Product C, Product D, Product E, Product F, Product G, and Product H, respectively.

### 5.5.1 CT

Based on the Pearson correlation coefficient method, we computed the correlation coefficients of the CTs of the eight kinds of products. The results are shown in the following symmetrical matrix:

$$\begin{bmatrix} 1 & 0.39 & 0.18 & 0.20 & 0.22 & 0.29 & 0.29 & 0.23 \\ 0.39 & 1 & 0.21 & 0.38 & 0.34 & 0.26 & 0.44 & 0.01 \\ 0.18 & 0.21 & 1 & 0.47 & 0.13 & 0.38 & 0.44 & 0.02 \\ 0.20 & 0.38 & 0.47 & 1 & 0.34 & 0.07 & 0.40 & 0.33 \\ 0.22 & 0.34 & 0.13 & 0.34 & 1 & 0.13 & 0.33 & 0.41 \\ 0.29 & 0.26 & 0.38 & 0.07 & 0.13 & 1 & 0.17 & 0.09 \\ 0.29 & 0.44 & 0.44 & 0.40 & 0.33 & 0.17 & 1 & 0.18 \\ 0.23 & 0.01 & 0.02 & 0.33 & 0.41 & 0.09 & 0.18 & 1 \end{bmatrix},$$

From the above matrix, we can obtain the following conclusions:

(1) The correlation among CTs of different kinds of products is extremely weak and uncorrelated. This result is attributed to the CT's performance of each Product, which was mainly determined by its technological processes rather than the CTs of other products.

(2) Most of the correlations among different products' CTs are positive.

(3) With the increase in CTs, the wafers to be processed in the production line will gradually become crowded, thus extending the CTs of other products. Therefore, the CTs of different products present positive correlations rather than negative correlations.

### 5.5.2 ODR

Based on the Pearson correlation coefficient method, we computed the correlation coefficients of the ODRs of the eight kinds of products. The results are shown in the following symmetrical matrix:

$$\begin{bmatrix} 1 & 0.13 & 0.21 & 0.18 & -0.06 & 0.17 & -0.07 & 0.16 \\ 0.13 & 1 & 0.21 & 0.12 & 0.09 & -0.09 & 0.07 & -0.02 \\ 0.21 & 0.21 & 1 & 0.06 & 0.17 & 0.17 & 0.12 & 0.00 \\ 0.18 & 0.12 & 0.06 & 1 & 0.05 & 0.01 & 0.21 & 0.12 \\ -0.06 & 0.09 & 0.17 & 0.05 & 1 & -0.10 & 0.06 & 0.13 \\ 0.17 & -0.09 & 0.17 & 0.01 & -0.10 & 1 & 0.19 & 0.12 \\ -0.07 & 0.07 & 0.12 & 0.21 & 0.06 & 0.19 & 1 & 0.22 \\ 0.16 & -0.02 & 0.00 & 0.12 & 0.13 & 0.12 & 0.22 & 1 \end{bmatrix}$$

From the above matrix, we can obtain the following conclusions:

(1) The correlation among ODRs of different kinds of products is extremely weak and uncorrelated. This result is attributed to the performance ODR of each Product, which was mainly determined by its technological processes and the congestion degree of the manufacturing system, rather than ODRs of other products.

(2) Most of the correlations among the ODRs of different products are positive, except for a few positive correlations with absolute values less than 0.1.

(3) With the increase in ODRs, the congestion degree of the wafers to be processed is gradually reduced, which will contribute to the smooth processing of other products and increase the ODRs of other kinds of products.

## 5.6 Further discussion

The correlations of MOV with WIP and QL in a certain period are displayed in Figs. 4 and 5, respectively. The results reveal the following:

(1) The correlations of MOV with WIP and QL are nonlinear.

(2) The correlations of MOV with WIP and QL are complex, and they cannot be described by any

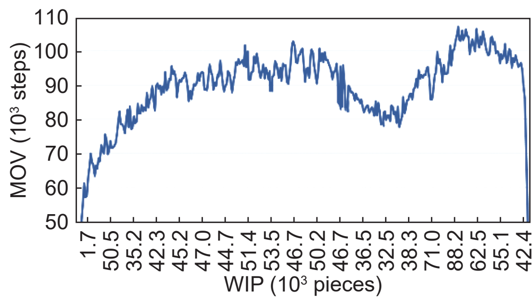


**Fig. 4 Trend line of MOV and WIP.**

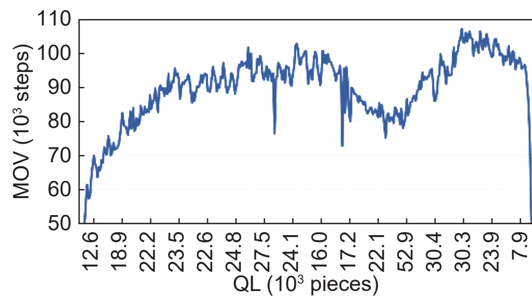


**Fig. 5 Trend line of MOV and QL.**

established linear formula.

(3) The trend lines of MOV with WIP and QL are similar.

(4) We should study some fuzzy models or other similar models to illustrate the exact correlations of MOV with WIP and QL.

Based on the similarity of the trend lines of MOV with WIP and QL, we analyzed the values of WIP and QL in one year, as shown in Fig. 6.

Figure 6 reveals that the value of QL varied with the value of WIP. The greater the value of WIP, the severer the congestion, and the greater the value of QL. Thus, QL is closely linked to WIP.

## 6 Conclusion

A performance evaluation system that can describe the performance of a semiconductor manufacturing system in detail is proposed in this study. Using the proposed

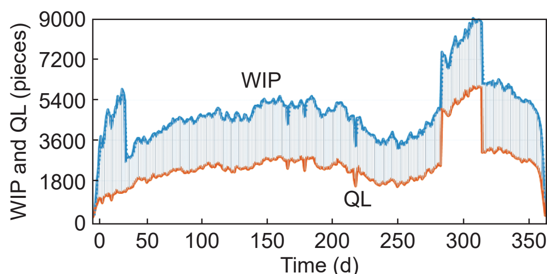


**Fig. 6 Trend line of WIP and QL.**

performance evaluation system, we can obtain the values of each performance indicator[25] through a simulation model to enhance the performance[26]. The ultimate goal of a semiconductor manufacturing system is to optimize the system performance as much as possible.

However, complex correlations exist among the performance indicators. Thus, we analyzed correlations among some common performance indicators through the Pearson correlation coefficient methods and CCA to obtain the correlation coefficients between the performance indicators. For example, CT has a negative correlation with various performance indicators, such as ODR, owing to its statistical significance. MOV has a positive correlation with indicators, such as TH, EU, and ODR. It also has complex correlations with indicators WIP and QL, which cannot be analyzed using common linear methods.

In the future, we will study fuzzy models or other suitable methods to further demonstrate the complex nonlinear correlations between the performance indicators of a semiconductor manufacturing system.

## References


[1] Q. Yu, H. Yang, K. Yi. Lin, and L. Li, A self-organized approach for scheduling semiconductor manufacturing systems, *J. Int. Manuf.*, vol. 32, no. 3, pp. 689–706, 2021.

[2] R. Singh and M. Mathirajan, Experimental investigation for performance assessment of scheduling policies in semiconductor wafer fabrication-a simulation approach, *Int. J. Adv. Manuf. Technol.*, vol. 99, nos. 5–8, pp. 1503–1520, 2018.

[3] H. Y. Sang, P. Y. Duan, and J. Q. Li, An effective invasive weed optimization algorithm for scheduling semiconductor final testing problem, *Swarm Evol. Comput.*, vol. 38, pp. 42–53, 2018.

[4] F. J. Yang, N. Q. Wu, Y. Qiao, M. C. Zhou, and Z. W. Li, Scheduling of single-arm cluster tools for an atomic layer deposition process with residency time constraints, *IEEE Trans. Syst. Man Cybern.: Syst.*, vol. 47, no. 3, pp. 502–516, 2017.

[5] J. H. Pang, H. M. Zhou, Y. C. Tsai, and F. D. Chou, A scatter simulated annealing algorithm for the bi-objective scheduling problem for the wet station of semiconductor manufacturing, *Comput. Ind. Eng.* , vol. 123, pp. 54–66, 2018.

[6] F. Qiao, X. Xu, M. Fang, and Q. Wu, Performance evaluation system for scheduling semiconductor wafer product line, (in Chinese), *J. Tongji Univ.*, vol. 35, no. 4, pp. 537–542, 2007.

[7] N. Hinrichs, P. Leppelt, and E. Barke, Building up a

performance measurement system to determine productivity metrics of semiconductor design projects, in *Proc. 2007 IEEE Int. Engineering Management Conf.*, Lost Pines, TX, USA, 2007, pp. 327–329.
[8] N. Hinrichs and E. Barke, Applying performance management on semiconductor design processes, in *Proc. 2008 IEEE Int. Conf. Industrial Engineering and Engineering Management*, Singapore, 2008, pp. 278–281.
[9] Z. G. Zhou and O. Rose, A bottleneck detection and dynamic dispatching strategy for semiconductor wafer fabrication facilities, in *Proc. 2009 Winter Simulation Conf.*, Austin, TX, USA, 2009, pp. 1646–1656.
[10] H. Zhang, Z. B. Jiang, Y. F. Lee, C. P. Ko, C. O. T. Luke, and L. P. Lim, An approach of dynamic bottleneck machine dispatching for semiconductor wafer fab, in *Proc. 2007 Int. Symp. Semiconductor Manufacturing*, Santa Clara, CA, USA, 2007, pp. 1–4.
[11] H. R. Tsai and T. Chen, Self-adaptive agent-based dynamic scheduling for a semiconductor manufacturing factory, in *Proc. 10th Int. Conf. Simulation of Adaptive Behavior*, Osaka, Japan, 2008, pp. 519–528.
[12] H. J. Yoon and J. G. Kim, Heuristic scheduling policies for a semiconductor wafer fabrication facility: minimizing variation of cycle times, *Int. J. Adv. Manuf. Technol.*, vol. 67, nos. 1–4, pp. 171–180, 2013.
[13] K. Nemoto, E. Akcali, and R. M. Uzsoy, Quantifying the benefits of cycle time reduction in semiconductor wafer fabrication, *IEEE Trans. Electron. Packag. Manuf.*, vol. 23, no. 1, pp. 39–47, 2000.
[14] H. X. Zhong, M. Liu, J. H. Hao, and S. L. Jiang, An operation-group-based soft scheduling approach for uncertain semiconductor wafer fabrication system, *IEEE Trans. Syst. Man Cybern.: Syst.*, vol. 48, no. 8, pp. 1332–1347, 2018.
[15] K. Y. Lin, A. Yu, P. C. Chu, and C. F. Chien, User-experience-based design of experiments for new product development of consumer electronics and an empirical study, *J. Proc. Ind. Eng.*, vol. 34, no. 7, pp. 504–519, 2017.
[16] H. Kwon and N. M. Nasrabadi, Kernel matched subspace detectors for hyperspectral target detection, *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 2, pp. 178–194, 2006.
[17] Y. T. Kao, S. Dauzère-Pérès, J. Blue, and S. C. Chang, Impact of integrating equipment health in production scheduling for semiconductor fabrication, *Comput. Ind. Eng.*, vol. 120, pp. 450–459, 2018.
[18] Y. J. Gu, J. Xu, Q. Q. Li, H. F. Li, and D. C. Chen, Fuzzy comprehensive evaluation method for peak shaving capability of coal-fired power units, (in Chinese), *Thermal Power Generation*, vol. 46, no. 2, pp. 15–21, 2017.
[19] S. J. Qin and Y. Zheng, Quality-relevant and process-relevant fault monitoring with concurrent projection to latent structures, *AIChE J.*, vol. 59, no. 2, pp. 496–504, 2013.
[20] M. B. Blaschko, C. H. Lampert, and A. Gretton, Semi-supervised laplacian regularization of kernel canonical correlation analysis, in *Proc. European Conf. Machine Learning and Knowledge Discovery in Databases*, Antwerp, Belgium, 2008, pp. 133–145.
[21] F. Pan and X. S. Qian, Methods of optimization for throughput and cycle time in semiconductor wafer fabrication, (in Chinese), *Semiconductor Technology*, vol. 29, no. 2, pp. 41–45, 2004.
[22] S. H. Chung and H. W. Huang, Loading allocation algorithm with machine capability restrictions for wafer fabrication factories, *J. Chin. Inst. Ind. Eng.*, vol. 18, no. 4, pp. 82–96, 2001.
[23] H. B. Li, G. Z. He, and Q. T. Guo, Similarity retrieval method of organic mass spectrometry based on the Pearson correlation coefficient, (in Chinese), *Chem. Anal. Met.*, vol. 24, no. 3, pp. 33–37, 2015.
[24] H. Hotelling, Relations between two sets of variates, *Biometrika*, vol. 28, nos. 3&4, pp. 321–377, 1936.
[25] L. Li, Y. Wang, and K. Y. Lin, Preventive maintenance scheduling optimization based on opportunistic production-maintenance synchronization, *J. Intell. Manuf.*, vol. 32, no. 2, pp. 545–558, 2021.
[26] C. F. Chien, K. Y. Lin, J. B. Sheu, and C. H. Wu, Retrospect and prospect on operations and management journals in Taiwan: From Industry 3.0 to Industry 3.5, (in Chinese), *J. of Mgt.*, vol. 33, no. 1, pp. 87–103, 2016.

**Hui Zhao** received the MS degree in transportation management and planning from Beijing Jiaotong University, Beijing, China, in 2016, and the PhD degree in computer science from Université de Technologie de Belfort-Montbéliard, Franche, France, in 2020. Currently, she is the postdoctoral researcher in the Department of Control Science and Engineering, College of Electronics and Information Engineering, Shanghai Institute of Intelligent Science and Technology, and Shanghai Research Institute for Intelligent Autonomous Systems, Tongji University, Shanghai, China. Her research interests include multi-agent system, decision-making model, smart city, and intelligent transportation.

**Li Li** received the BS and MS degrees from Shenyang Agriculture University, China, in 1996 and 1999, respectively; and the PhD degree from Shenyang Institute of Automation, Chinese Academy of Sciences, Shenyang, China, in 2003. She joined the Department of Control Science and Engineering, College of Electronics and Information Engineering, Shanghai Institute of Intelligent Science and Technology, and Shanghai Research Institute for Intelligent Autonomous Systems, Tongji University, Shanghai, China, in 2003, and is presently a professor. Her research interests include data-driven modeling and optimization, computational intelligence.

**Qingyun Yu** received the BS degree from Jiangnan University, Wuxi, China, in 2013, and the PhD degree from Tongji University, Shanghai, China, in 2021. Currently, she is the postdoctoral researcher in the Department of Control Science and Engineering, College of Electronics and Information Engineering, Shanghai Institute of Intelligent Science and Technology, and Shanghai Research Institute for Intelligent Autonomous Systems, Tongji University, Shanghai, China. Her research interests include data-driven modeling and optimization, intelligent manufacturing, and computational intelligence. She has published 2 monographs, more than 10 academic papers, one patent, one software-copyright, and has participated in 2 national scientific research projects and 2 municipal major projects.

**Ying Liu** is currently a reader in intelligent manufacturing and the group leader for high-value manufacturing with the Department of Mechanical Engineering, School of Engineering, Cardiff University, UK. His research interests focus primarily on engineering informatics, digital and intelligent (smart) manufacturing, AI and machine learning for engineering, design methodology and process and advanced ICT in engineering design and manufacturing, in which areas he has published over 160 scholarly articles, one edited book, and more than 15 journal special issues. He serves as an associate editor with *ASME JCISE, IEEE T-ASE, the Journal of Industrial and Production Engineering* (Taylor & Francis), *CCF Transactions on Pervasive Computing and Interaction* (Springer), and *Autonomous Intelligent Systems* (Springer), and is on the editorial board of *Advanced Engineering Informatics* (Elsevier).

**Kuo-Yi Lin** received the BS degree in statistics from Cheng Kung University, Taiwan, China, in 2007; and the MS and PhD degrees in industrial engineering and engineering management from Tsing Hua University, Taiwan, China, in 2009 and 2014, respectively. He is currently a professor of the Department of Control Science and Engineering, College of Electronics and Information Engineering, Shanghai Institute of Intelligent Science and Technology, and Shanghai Research Institute for Intelligent Autonomous Systems, Tongji University, Shanghai, China. He is the director of China Excellent Business Decision Making Society, member of Intelligent Simulation Optimization and Scheduling Committee of China Simulation Society, member of Natural Computing and Digital Intelligent City Committee of China Artificial Intelligence Society, and member of Industrial Big Data and Intelligent System Branch of China Mechanical Engineering Society. He is mainly engaged in intelligent manufacturing, federated learning, quantum algorithm, and transfer learning.