Statistical key variable analysis and model-based control for improvement performance in a deep reactive ion etching process*

Chen Shan(陈山)[1], Pan Tianhong(潘天红)[1,†], Li Zhengming(李正明)[1], and Jang Shi-Shang(郑西显)[2]

[1]School of Electrical and Information Engineering, Jiangsu University, Zhenjiang 212013, China
[2]Department of Chemical Engineering, National Tsing-Hua University, Hsin-Chu, Taiwan 30013, China

Abstract: This paper proposes to develop a data-driven via's depth estimator of the deep reactive ion etching process based on statistical identification of key variables. Several feature extraction algorithms are presented to reduce the high-dimensional data and effectively undertake the subsequent virtual metrology (VM) model building process. With the available on-line VM model, the model-based controller is hence readily applicable to improve the quality of a via's depth. Real operational data taken from a industrial manufacturing process are used to verify the effectiveness of the proposed method. The results demonstrate that the proposed method can decrease the MSE from 2.2×10^{-2} to 9×10^{-4} and has great potential in improving the existing DRIE process.

Key words: deep reactive-ion etching; virtual metrology; through silicon via; key variable analysis; model-based control

1. Introduction

With the increasing demand for continuous and convenient access to information and communication in vehicles, more and more high speed, powerful and nomadic electronic devices such as laptops, smart phones and GPS, etc., have being incessantly renewed in the last few decades. However, suffering from the inherent technical limit of the standard two-dimensional (2D) integrated circuit (IC), powerful processor and huge memory capability integrated in 2D ICs have difficulty crossing those obstacles such as resistive–capacitive (RC) delay and thermal heating and power consumption[1]. Fortunately, 3D IC has been proposed as a promising means of mitigating those problems, which can allow high integration density, fast signal transmission, low manufacturing cost and lower power consumption. Beyond all these benefits, heterogeneous stacking can easily be performed with 3D technology, enabling more sophisticated than ever system on chip (SoC) integration. So, 3D ICs can be regarded as a new approach to improved IC performance and have attracted considerable attention in the past few years[2].

To realize 3D ICs and electrically connect the components in different layers, through-silicon-via (TSV) technology may be used to provide the electrical interconnect and to provide mechanical support, which can increase the bandwidth, reduce the footprint of the system, and achieve heterogeneous integration of the system[3]. In TSV technology, the critical step is to make the microvia in a silicon chip. The deep reactive ion etching (DRIE) process is most commonly used in microelectromechanical systems (MEMS) devices to "drill" the hole through the silicon substrates of individual chips. Vias are then filled with copper which interconnects chips forming a stack.

In TSV processing, via patterns must have smooth sidewalls and a deep pattern depth in order to deposit insulator, barrier, and seed layers, as well as full filling of metal. Unfortunately, there are insufficient etching results to form deep silicon via patterns. Thus, certain critical parameters of the DRIE process have been investigated in the literature[4−8]. Optimum etch processing can be obtained by careful adjustment of the process variables, such as passivation and etch times, wafer temperature, ion energy, and species fluxes. Chen *et al.* pointed out that the etch depth was a function of applied coil power and SF_6 flow rate[6]. Morgan *et al.* investigated that the sidewall roughness could be reduced by using a low bias power and reduced passivation cycle time[7]. What's more, the balance between passivation layer growth and removal also plays a critical role in controlling profiles[8]. Besides optimum parameter adjustment, many kinds of new or improved DRIE processes have also been proposed to improve the quality of via patterns[9−11]. However, the basic relationships between the plasma parameters and the etched depth have not been fully explored. Especially, there is a lack of in situ sensors to provide real time information about etch depth during the DRIE process. Therefore, etch depth control, which is not clearly understood, retards the development and optimization of the etching process for a deep silicon etch.

The objectives of this paper are to develop a systematic approach to model etch depth with cycle number and a few key factors. The modeling uses multi-variant statistical methods with real-time operational data. The acquired etch depth model can be used to show how to adjust cycle number to get
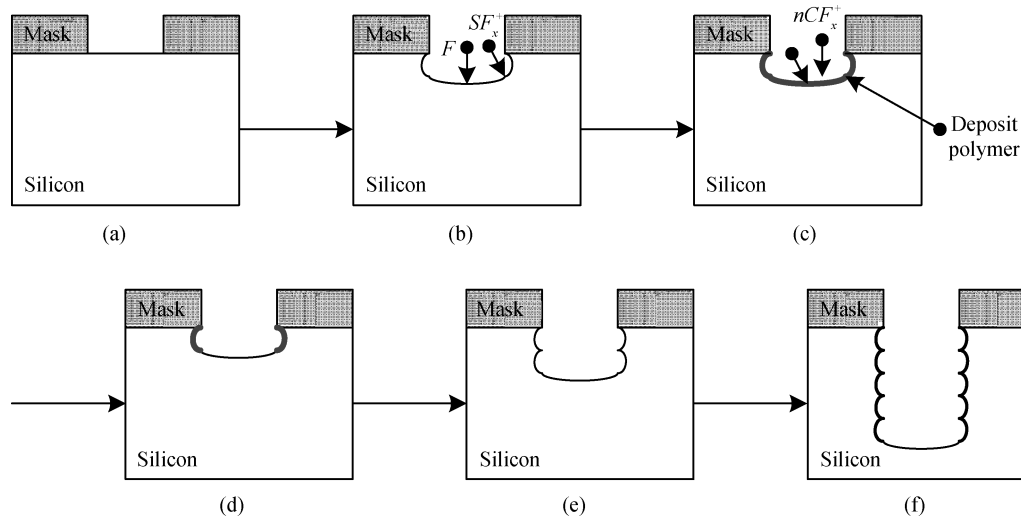
Fig. 1. The sequence step of a DRIE process. (a) Masking material (photoresist or silicon dioxide) is patterned on a silicon wafer. (b) Silicon etching step, a shallow, isotropic trench is formed. (c) Passivation step, a protective fluorocarbon film is deposited everywhere. (d) In the subsequent etch step, ion bombardment promotes the preferential removal of the film from all horizontal surfaces, allowing the profile to evolve in a highly anisotropic fashion. (e) Finally, the fluorocarbon film is removed from all horizontal surfces by directional ion bombardment, and another shallow trench is formed. (f) In the end, the alternating of etching and passivating cycles forms scallops.

a better deep silicon etch. The developed approach can be used not only for online virtual metrology operations, but also for controller design to improve etching quality to save production costs.

2. Process description and problem statement

2.1. Process description

To realize 3D integration, one important technology is TSV, which utilizes deep vias to provide the shortest interconnect in the semiconductor packing industry. For via drilling, DRIE is the most popular process which has been employed by the Bosch process. The technology relies on a sequential passivation and etching process to form vertical trenches and walls on silicon substrates with high aspect ratios, as shown in Fig. 1[6].

It can be seen that the DRIE process is conducted using the following steps:

(1) In the preparation step, pattern masking material (photoresist or silicon dioxide) on a silicon wafer;

(2) In the etching step, etch silicon wafer in SF_6 plasma and form a isotropic trench;

(3) In the passivation step, deposit polymer as a passivation layer in the TSV in a C_4F_8 plasma;

(4) In the subsequent etching step, ion bombardment promotes the preferential removal of the film from all horizontal surfaces, allowing the profile to evolve in a highly anisotropic fashion;

(5) Then, alternating of etching and passivating cycles forms scallops on the sidewalls of etched features;

(6) After completing the silicon etch, the passivation layer was removed in an O_2 plasma.

In the cyclic repetition of those etching and passivation steps, the quality of via (depth, sidewall slope and smoothness) is affected by several parameters, which include etchant gases, flow rate of the chosen gases, RF-power, bias voltage, process pressure, temperature, etc. Extensive studies have been performed on via formation, especially the effect of process parameters on via profile and sidewall roughness[5, 6]. Many instruments are installed to record those process parameters and corresponding control actions are taken to manipulate them to attain the target value. Unfortunately, there is no in situ sensor to provide real time information about the etching depth in each processed cycle. The quality of deep via is usually measured offline, which may lead to scrapped wafers. So, field engineers attend to the investigation of these process variables in order to enhance the yield rate.

2.2. Problem statement

In general, design of experiments (DOE) is used to optimize the DRIE process for depth of via, etch rate, sidewall slope and smoothness. In an industrial manufacturing process, those parameters are kept with their claimed value to achieve a consistent performance. In fact, wafers processed in the nominally identical DRIE process show different end-of-line quality (especially the depth), because the performances of DRIE will change with the passage of time or after preventative maintenance.

The study was focused on mining out key variables X' ($X' \subset X$), which represent the underlying causes of variation in depth of via, and deriving a variation model.

$$y = f(X'), \tag{1}$$

where y is the depth of via, X' is key variables vector, and X is all process variable vector.

Then, a model-based controller was designed to improve the performance of the DRIE process and achieve the same depth of via as much as possible.
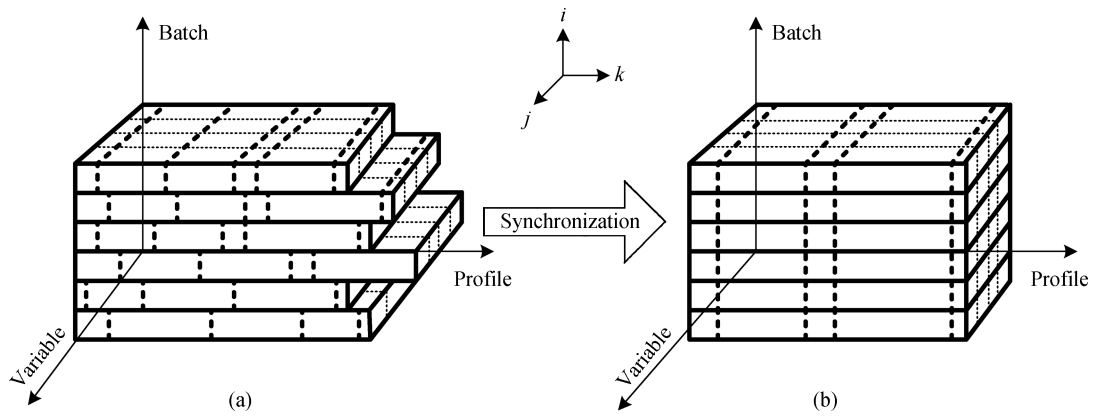
Fig. 2. Data of sensor variable and its synchronization. (a) Procession times vary from batch-to-batch. (b) Synchronized results.

3. Development of virtual metrology

The DRIE manufacturing stage can be regarded as a multi-step batch (wafer or lot) process. So, the data recorded by sensor variable within a batch should be regarded as a profile of variables. Consequently, the processing times vary from batch to batch (Fig. 2(a)) and a synchronization procedure is required. Figure 2 demonstrates the three-dimensional (variable-profile-batch) data structure of sensor variable output and its synchronized version[12]. In this work Akima spline is applied to each step for synchronization.

3.1. Feature extraction on FDC data

The idea of virtual metrology (VM) is based on the concept that the profile of sensors installed in process equipment can reflect process quality effectively and loyally to find out the excursion issues of process equipment and defect impact. But these raw data collected from a fault detection and classification (FDC) system are a large number of time serials. So, these massive and untreated data cannot be used to effectively undertake the subsequent VM model building process. How to project these large numbers of values of each profile onto the space spanned by the variable transformation is an important issue. In this work, the data treatment methods, which include the discrete Fourier transform (DFT) method shown in Fig. 3(b), the area method shown in Fig. 3(c) and the average method shown in Fig. 3(d), are proposed to effectively convert these massive values into the significant univariates that VM models need.

Taking Fig. 3 (a) for example, there are many cyclical profiles owing to the alternating etching and passivation steps. The DFT gives a solution to extracting features from those kinds of signals, which computes the frequency information of the equivalent time domain signal[13]. Since a cyclical signal in the DRIE process contains only real point values, we can make use of this fact and use a real-point Fast Fourier Transform (FFT) for increased efficiency. To access the spectral component of z, the DFT $Z: \Re \to C^p$ of the output trajectory $\{z_k\}_{k=1}^n$ could be computed as follows:

$$\text{DFT}: \quad Z(m\omega_s) = \sum_{k=1}^{n-1} z_k e^{-jm\omega_s k}, \quad \omega_s = \frac{2\pi}{n}, \quad (2)$$

where $k = 1, 2, \cdots, n$ is sampling time, $m = 0, 1, \cdots, n-1$ of the DFT computation.

Using a DFT (fft and fftshift command in Matlab) algorithm, the resulting output is shown in Fig. 3(b). The figure clearly reveals the nature of the cyclical time series, which contains both the magnitude and phase information of the original time domain signal (Fig. 3(a)). There are three clear spectral peaks, which indicates that there is more to this signal than just noise. Among three obvious components, one is located at zero frequency and the other two peaks are symmetric. So, two features (constant value and highest frequency value) are extracted.

Except for cyclical signals, there are also some other profiles. Some of them change with time (Fig. 3(c)), some of them remain constant (Fig. 3(d)). In this work, area and average methods are used to extract feature[14].

$$\text{Area:} \quad A = \int_{z_1}^{z_n} f(z)\mathrm{d}z = \sum_{k=1}^{n} z_k. \quad (3)$$

$$\text{Average:} \quad \bar{z} = \frac{\sum_{k=1}^{n} z_k}{n}. \quad (4)$$

Using feature extraction, three-dimensional data (variable-profile-batch) descends as two dimensional data (feature-batch). Then, a statistical analysis approach for building VM is used as follows.

3.2. Key variable selection based on a statistical method

Although the high dimensional data have been reduced to two dimensions by using feature extraction approaches during the preprocessing step, there are still too many input variables. Some of them are highly correlated, and some of them are even unnecessary for VM prediction models. Principal component analysis (PCA) and partial least square (PLS) give a solution to overcome this problem, which projects the original input variables onto a space defined by orthogonal principal components (PCs) or latent variables (LVs). However, the field engineer may not understand the physical meaning of PC/LV. Thus, a variable selection method is used in this paper. Unlike PCA and PLS, the variables selected by a stepwise regression method usually have physical meaning. Therefore, the inferential models built by the stepwise regression method are suitable for process control and process fault diagnosis. Besides,
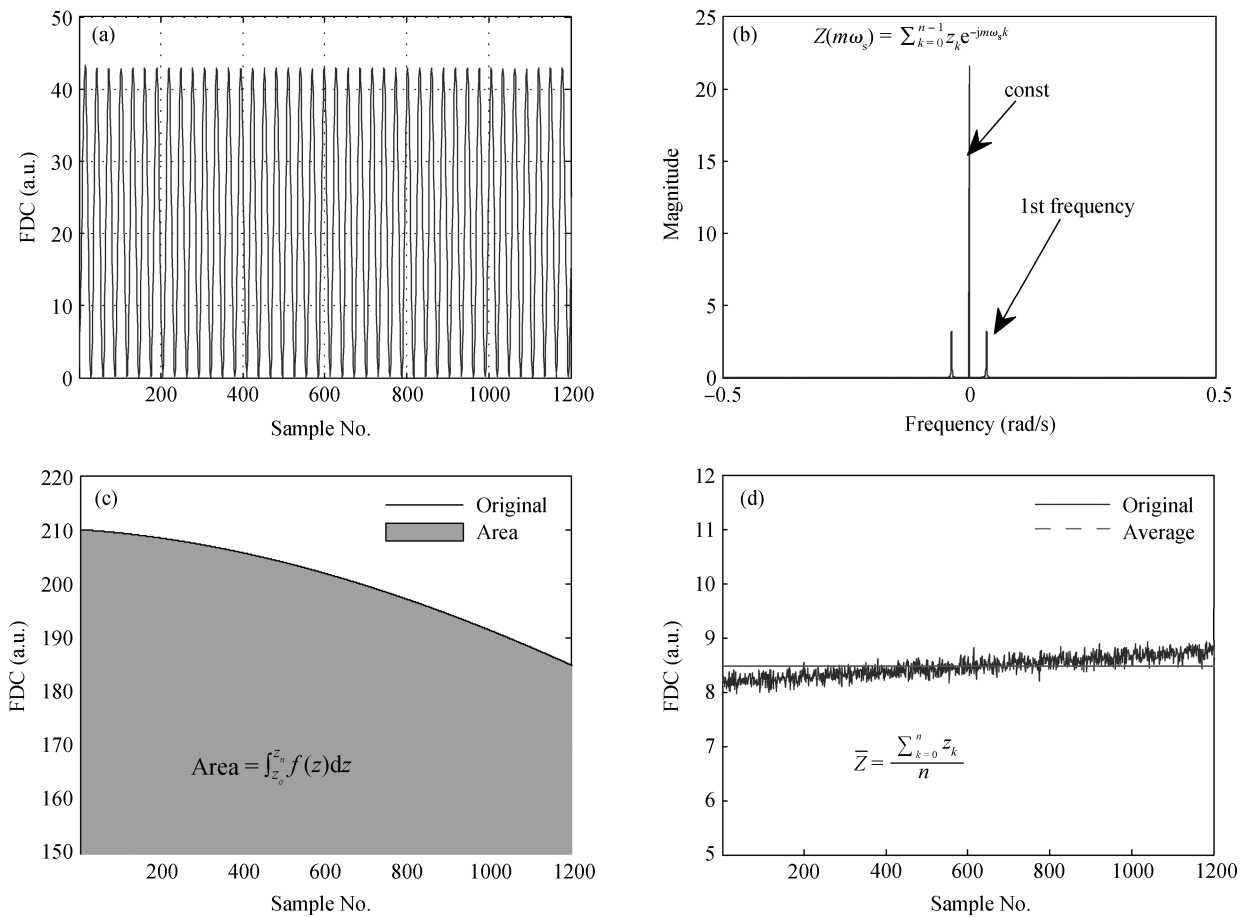
Fig. 3. Examples of the feature extraction on FDC data. (a) Original cyclical profile. (b) Feature extraction by DFT. (c) Feature extraction by area. (d) Feature extraction by avergae.

the implementation of the algorithm is easy and fast on digital computers.

The purpose of stepwise regression is to find an appropriate subset of input variables $\Omega = \{x_1, x_2, \cdots, x_p\}$ to fit the multiple regression model below to a set of data:

$$\hat{y} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_j x_j, \qquad (5)$$

where $x_1, x_2, \cdots, x_j$ are the selected key variables; $\beta_0, \beta_1, \cdots, \beta_j$ are the coefficients of regressor; and $\hat{y}$ is the predicted depth.

Stepwise regression begins with a single input variable selected from the candidate set Ω which can give a best linear relationship for input-output. Then, a new variable is added from the candidate set Ω (forward selection) if it can improve explanation of the developed model. An existing selected variable can be removed (backward elimination) if its absence increases or at least maintains the prediction accuracy. By repeating forward selection and backward elimination alternately until the candidate set is empty or prediction accuracy improvement is negligible, one can achieve a small subset of important input variables (see the authors' previous work[15]).

4. Results and discussion

To demonstrate our approach, two batches of data are collected (2010/2/23–2010/3/3 and 2010/5/14–2010/5/29)

from a normally identical DRIE process in a local fabrication unit. The recipe of the first batch data is 200 cycles, the second recipe is 210 cycles. The specification (target of depth) of both batches is 3.9 a.u. (arbitrary units). In the DRIE process, not all actual depth of wafers which are denoted as $y[i]$ ($i = 1, 2, \cdots, m$) are measured, only 1–2 wafers of a lot are selected as sample wafers, whose depths are measured to monitor the quality of the whole lot. Consequently, 36 wafers of process data and their corresponding actual metrology depths can be used as data in building and testing a VM model of the DRIE process. Among these 36 glasses, we use three-fold cross validation to evaluate the performance of developed VM model. That means the whole data set is randomly divided into three disjointed subsets of equal size 12, and the holdout method is repeated three times. Each time, one of subsets is used as the test set and the other two subsets are put together to form a training set.

4.1. Performance test for the VM model

It is obvious that the cycle is the most important parameter in the DRIE process: the cycle means the times of alternating etching and passivation steps. The more cycles, the deeper the depth of via. So, the cycle should be selected as a fixed parameter in the developed VM model. Then, the residual of this model is used for further analysis.

$$e[i] = y[i] - \hat{y}^c[i] = y[i] - \beta_0 - \beta_1 x_c[i], \qquad (6)$$
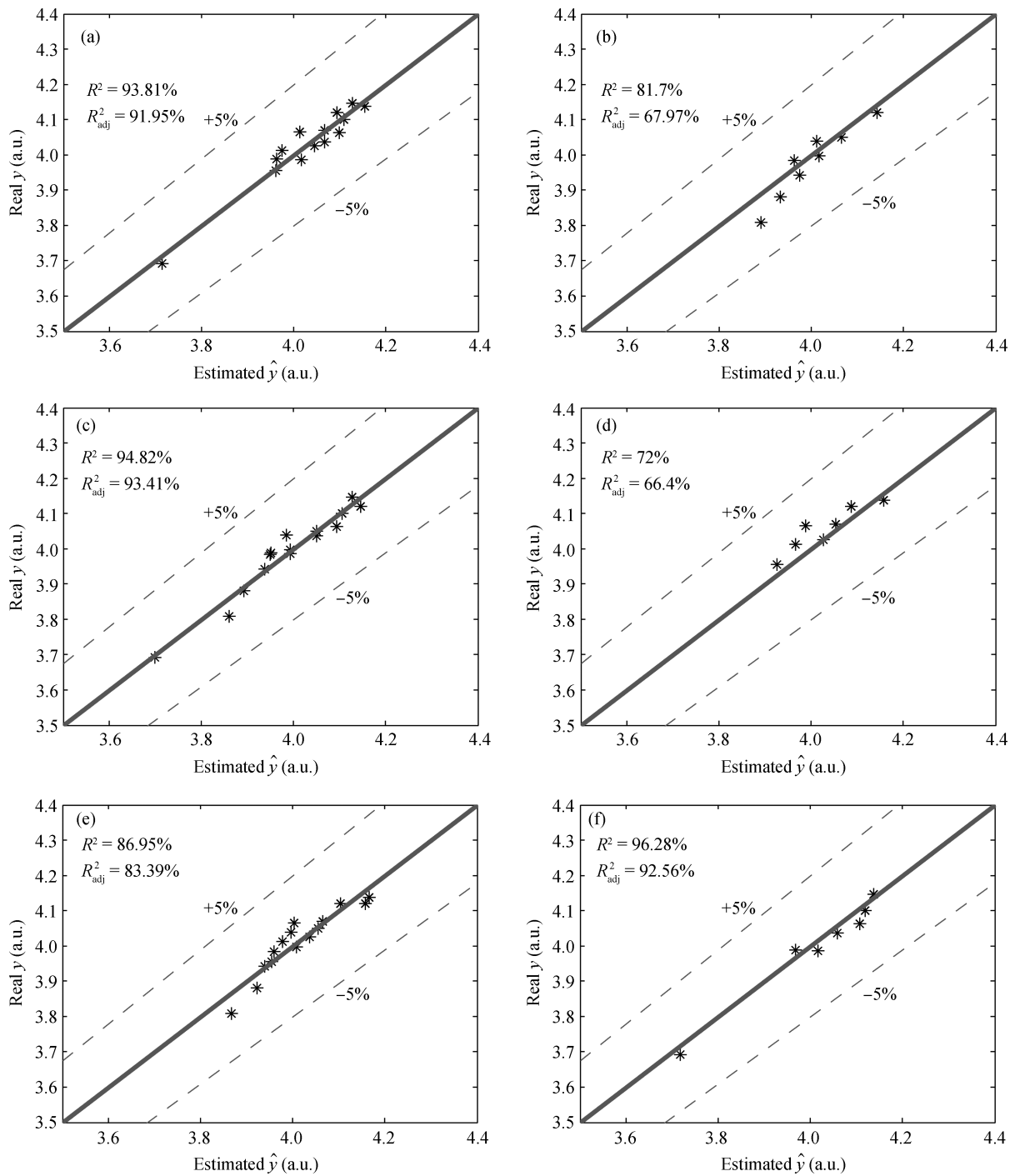
Fig. 4. Three-fold cross validation for VM. (a) First modeling part. (b) First predicted part. (c) Second modeling part. (d) Second predicted part. (e) Third modeling part. (f) Third predicted part.

where $y[i]$ is the real depth and $\hat{y}^c[i]$ is the estimated depth obtained by using only one cycle; $x_c[i] \in \{200, 210\}$ is the cycle; β_1 is the regression coefficient of linear regression and β_0 is a constant, and those coefficients can be obtained by standard least square estimation; and i is the serial number of the wafer.

As mentioned above, features extracted by the proposed methods may be inter-correlated. So, a stepwise regression approach is used to select the key variables.

$$e[i] = \beta_2 x_1[i] + \beta_3 x_2[i] + \cdots + \beta_{j+1} x_j[i], \qquad (7)$$

where $\beta_2, \cdots, \beta_{j+1}$ are the coefficients; j is the number of

key variables; $x_1[i], \cdots, x_j[i]$ are the selected key variables.

Table 1 gives the details about the developed VM model. In this work, the thresholds of the probability of type I error (i.e., $\alpha_{in} = 0.05, \alpha_{out} = 0.1$) for stepwise regression are selected. The VM model should give a good variance explanation of the depth which is usually evaluated by the R^2 statistics and adjusted R^2 statistics (R^2_{adj}). The p-value is also considered in this stepwise regression, which measures the significance of the best-fitting independent variable to be entered at an arbitrary step.

Shown in Table 1, the p-value is 0.052 and greater than

Table 1. Explanation of the VM model in different steps.

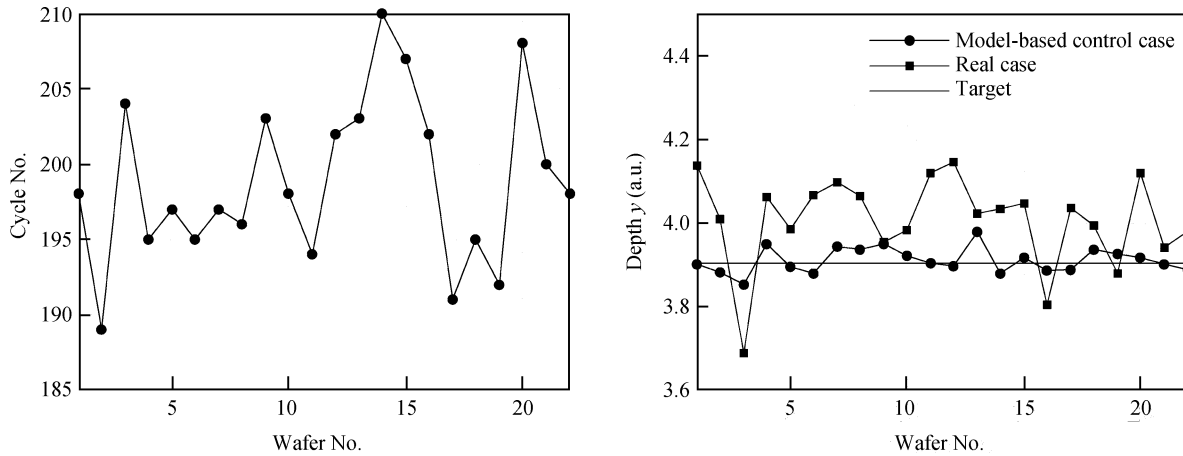| Criteria	Cycle No (%)	Residual for stepwise regression (%)					
		x_1	x_1, x_2	x_1, x_2, x_3	$x_1, \cdots, x_4$	$x_1, \cdots, x_5$	$x_1, \cdots, x_6$
R^2	52.51	46.88	67.69	73.15	80.15	83.71	86.46
R^2_{adj}	52.51	44.84	65.09	69.79	76.70	80.00	82.11
p-value	—	0.0001	0.0001	0.037	0.009	0.039	**0.052**

Fig. 5. Data of sensor variable and its synchronization.

0.05 when the VM model includes x_6, which means that x_6 should not be included in the VM model. As a result of stepwise regression, the bias RF voltage (x_1), the constant item of adjusted pressure (x_2), the ESC temperature (x_3), the chamber temperature (x_4) and the frequency item of IB2 (x_5) are chosen as key variables from all of the 44 features. The fitting capability in terms of R^2 and R^2_{adj} are maximized to 83.71% and 80.00% by setting the parameter set $[\beta_0, \beta_1, \cdots, \beta_6]$ in Eq. (5) to be $[9.76, 0.12, -1.55, -0.24, -0.30, 0.69]$.

Figure 4 shows the comparison between the real depth and the predicted one using three-fold cross validation. The minimum R^2 and R^2_{adj} of prediction results of this VM model are 72% and 66.4%. And the mean square error of prediction results are 0.0016, 0.001, and 0.0007 respectively. It can be seen that the VM predictor is accurate enough to be implemented in an actual DRIE process.

4.2. Model-based control of DRIE in a virtual plant

Although wafers were manufactured by the nominal identical DRIE process, most wafers were over-controlled (most depths are greater than target) due to the lack of an online measuring instrument. One shortcoming of the over-controlled process is that it will increase the manufacturing cost. The another is that it will increase the possibility of penetrating the whole wafer which can cause the defective wafer. So, the field engineers want to improve performance by adaptively adjusting the process parameters run by run based on the information obtained during the process[16].

Given the above real time estimator, i.e., the depth conjecture model, it is hence feasible to install a model-based controller to improve the quality of depth by estimating the optimal cycle number. However, the application of a model-based controller to real plant needs comprehensive considerations and

design. The virtual plant simulating the plant operational data can be deemed as real as a genuine plant. This example uses the following ways to build a virtual plant.

(1) First, a linear model based on the second batch of data (including 22 wafers with 210 cycles) can serve as a virtual plant based on the concept of reverse engineering.

(2) When the DIRE procedure has been implemented for 180 cycles, Equations. (2), (3) and (4) are used to extract the features from the process sensors.

(3) The following deadbeat control algorithm can be implemented.

$$x_c[i] = \frac{\tau - y[i-1]}{\beta_1}, \tag{8}$$

where $x_c[i]$ is the required cycles, β_1 is the slope linear model identified by the above algorithm, and τ is the target. The future depth $y[i-1]$ can be calculated by Eq. (5).

(4) Given the extracted features in the next operational data, along with the cycle number calculated from the previous step, obtain the depth from the virtual plant.

Figure 5(b) illustrates the improvement of the depth using the developed VM model. The figure reveals the results of how to increase performance of depth by controlling the cycle number through a model-based controller. Compared with a real DRIE process, the mean square error decreases from 2.2×10^{-2} to 9×10^{-4}. The control actions of the setting cycle number taken by the model-based control to the virtual plant are shown in Fig. 5(a). It can be seen that most of wafers don't need 210 cyclical alternative etching and passivation steps to reach the target. Production cost and implemented time will be obviously reduced.

5. Conclusion

A data-driven depth estimator of the DRIE process based on statistical identification of key variables is developed. Several feature extraction algorithms are proposed to reduce the dimensions of the original data collected from the FDC system. Standard stepwise linear regression is adopted to identify the key variables. The proposed algorithm is easy to maintain. This study verifies the effectiveness of the proposed method by using industrial examples. The application to the virtual plant shows that the proposed approach is valid and feasible for the industries. Substantial improvement of depth can be achieved using this approach.

References

[1] Shen L C, Chien C W, Cheng H C, et al. Development of 3-D chip stacking technology using a clamped through-silicon via interconnection. Microelectron Reliab, 2009, 50(4): 489

[2] Lau J H. Overview and outlook of through-silicon via (TSV) and 3D integrations. Microelectron International, 2011, 28(2): 8

[3] Choi K M. An industrial perspective of 3D IC integration technology—from the viewpoint of design technology. Proceeding of 15th Asia and South Pacific Design Automation Conference, Taipei, 2010: 536

[4] Ranganathan N, Lee D Y, Youhe L, et al. Influence of Bosch etch process on electrical isolation of TSV structures. IEEE Trans Comp Packa Manuf Technol, 2011, 1(10): 1497

[5] Abhulimen I U, Polamreddy S, Burkett S, et al. Effect of process parameters on via formation in Si using deep reactive ion etching. J Vac Sci Technol B: Microelectronics and Nanometer Structures, 2007, 25(6): 1762

[6] Chen K S, Ayon A A, Zhang X, et al. Effect of process parameters on the surface morphology and mechanical performance of silicon structures after deep reactive ion etching (DRIE). J Microelectromechan Syst, 2002, 11(3): 264

[7] Morgan B, Hua X, Iguchi T, et al. Substrate interconnect technologies for 3-D MEMS packaging. Microelectron Eng, 2005, 81(1): 106

[8] Blauw M A, Zijlstra T, Drift E. Balancing the etching and passivation in time-multiplexed deep dry etching of silicon. J Vac Sci Technol B: Microelectronics and Nanometer Structures, 2001, 19(6): 2930

[9] Ham Y H, Kim D P, Park K S, et al. Dual etch processes of via and metal paste filling for through silicon via process. Thin Solid Films, 2011, 519(20): 6727

[10] Abdolvand R, Ayazi F. An advanced reactive ion etching process for very-high aspect-ratio sub-micron wide trenches in silicon. Sensors and Actuators A: Physical, 2008, 144(1): 109

[11] Azimi S, Sandoughsaz A, Amirsolaimani B, et al. Three-dimensional etching of silicon substrates using a modified deep reactive ion etching technique. J Micromechan Microeng, 2011, 21(7): 1

[12] Chao A G, Tseng S T, Wong D S H, et al. Systematic applications of multivariate analysis to monitoring of equipment health in semiconductor manufacturing. Proceedings of the 2008 Winter Simulation Conference Miami, Florida, USA, 2008: 2330

[13] Mario D. Application of Fourier linear spectralanalyses to the characterization of smooth muscle contractile signals. Journal of Biochemical and Biophysical Methods, 2007, 70(5): 803

[14] Pan J C H, Tai D H E. A new strategy for defect inspection by the virtual inspection in semiconductor wafer fabrication. Computers & Industrial Engineering, 2011, 60(1): 16

[15] Chen S, Pan T H, Jang S S. Development of a virtual metrology for high-mix TFT-LCD manufacturing processes. Journal of Semiconductors, 2010, 31(11): 1160061

[16] Kang P, Kim D, Lee H J, et al. Virtual metrology for run-to-run control in semiconductor manufacturing. Expert Systems with Applications, 2011, 38(3): 2508