Bootstrap Application for Semiconductor Incoming Material SPC

Sheng Kang, Violet Shangguan, Lisa Yu, and Wei-Ting Chien

***Abstract*—In the industry of semiconductor manufacturing, various material sources, such as gas, chemical, wafer, target, PR (Photo Resist), are used in process. The incoming material control plays a very important role in the whole production process. For traditional SPC (Statistical Process Control), the factors or parameters controlled are requested to follow normal (Gaussian) distribution, whereas most of them are not so in practice. Another problem we faced is there are fewer amounts of data. In order to control incoming material risk and prevent excursion occurrence, we studied the Bootstrap method, a re-sampling method in statistical. This paper depicts how we combine the Bootstrap into traditional SPC for incoming material quality management. And a comparison of the common incoming material management with Bootstrap SPC is also described.**

***Index Terms*—Bootstrap, SPC, incoming quality control, non- parameter.**

I. Introduction

The quality control of final products must be the cornerstone of any efficient control system, because the integrity of incoming material has obvious influence on the integrity of the finished product as sold to the customer. Especially, the quality control of incoming material is very important in semiconductor manufacturing since hundreds of material types are used in whole process. And the supplier control includes appraisal of the supplier's ability, the systematic application of statistical sampling techniques on incoming products, maintenance of a comprehensive Supplier/Product Quality Record System and the continuous analysis of Supplier Quality Records and the formulation of a Supplier Quality Index System [1]. The quality control system tools include SPC (Statistical Process Control) and Process Capability Index likes Cpk which is general used in industry. One of the problems of incoming material control is that most parameters are non-parameter distribution which violates the basic assumption of SPC [2]. And another problem is data count, because the incoming material is sampling measured before being used or the measurement is done by suppliers.

In this paper, we present a re-sampling method, i.e., a bootstrap based method, to realize the statistic control and monitor with SPC chart and process capability index.

Manuscript received May 30, 2015; revised October 14, 2015. This work was supported in part by Semiconductor Manufacturing International (Shanghai) Corp.

The authors are with the Division of Semiconductor Manufacturing International (Shanghai) Corp, Shanghai, China (e-mail: Randy_Kang@smics.com).

II. Methodology

A. Bootstrap Re-sampling Method

Non-parameter estimation is a branch of Statistics. It can be used to deal with a general distribution which can't be described by a finite number of parameters. The methods generally used to estimate the important value of non-parameter distributions are bootstrapping or jackknifing [3]. Bootstrapping is a statistical method for estimating the sampling distribution of an estimator by sampling the replacement of the original samples. Mostly, the purpose is to derive the robust estimates of standard errors and confidence intervals of population parameters such as mean, variation, correlation coefficient or percentile.

The core principle of Bootstrap is to re-sample the original samples to obtain the new statistics, such as mean, standard deviation, variation, correlation coefficient, etc. Following the Central Limit Theorem (CLT), the obtained statistics should follow normal distribution, and the average of the statistics should be close to the expected values when the sufficient resampling times are obtained based on Law of Large Number (LLN) [4].

B. Step of Bootstrap

For example, we got the independent samples $X \sim \Omega$, $(X_1, X_2, \ldots, X_n)$, Ω is from one unknown distribution. And then, we want to estimate the median of this population using bootstrap method [5].

Step 1: Do resampling with the replacement from the independent sample X, and obtain a set of new sample $X^*, (X_1^*, X_2^*, \ldots, X_{\mathrm{m}}^*)$

$$X^* \in (X_1, X_2, \ldots, X_n)$$

thereby $X^* \propto \Omega$

Step 2: Calculate the statistic θ for the sample of X^*. The statistic is usual for mean, standard deviation, percentile or correlation coefficient.

Step 3: Repeat step 1 and 2 for k times. The k usually uses 1000. And then, the k piece of statistics collected should follow one type of distribution. Because the sample is sufficient large, the distribution should be normal distribution.

$$\theta \propto N(\mu, \sigma^2)$$

wherein, the μ is the expected value of θ, $E(\theta)$; the σ^2 is variation of θ, $\mathrm{Var}(\theta)$. Based on CLT, if the sampling size

doi: 10.18178/ijmmm.2016.4.4.270

is sufficient large, the E(θ), denoted as θ_{boot} , should infinitely approach the statistic of population.

C. SPC Chart

Regarding the general SPC, it is applied in order to monitor and control process stability, and ensure it operates at its full potential. An advantage is that it emphasizes early detection and prevention of problems. A key tool used in SPC is control chart. To use the control chart, we must calculate the statistic of the samples, such as mean and standard deviation. And the control limit of SPC can be calculated based on these statistics. As showed in equation (1)

$$\begin{cases} UCL = \mu + 3 \times \sigma \\ LCL = \mu - 3 \times \sigma \end{cases} \tag{1}$$

wherein,

$$\mu = \frac{\sum_{i=1}^{n} x_i}{n}$$

$$\sigma = \sqrt{\frac{\sum_{i=1}^{n} (x_i - \mu)^2}{n - 1}}$$

x_i are the observation samples, and n is the sample count. The concept of SPC chart is that, when a point falls outside of the limits established for a given control chart, those factors responsible for the underlying process are expected to be used to judge whether a special cause has occurred. We improved the SPC control limit definition for bootstrap estimation as below.

$$\begin{cases} UCL = \mu_{\text{boot}} + 3 \times \sigma_{\text{boot}} \\ LCL = \mu_{\text{boot}} - 3 \times \sigma_{\text{boot}} \end{cases} \tag{2}$$

wherein, μ_{boot} and σ_{boot} are estimated by bootstrap method described in above section. Equation 2 is adequate for the data of normal distribution.

Furthermore, to set the control for data following non-parameter, equation 3 can meet the requirement.

$$\begin{cases} UCL = \mathrm{P}_{99.865\bullet\text{boot}} \\ LCL = \mathrm{P}_{0.135\bullet\text{boot}} \end{cases} \tag{3}$$

wherein, the statistics, $\mathrm{P}_{99.865\bullet\text{boot}}$ and $\mathrm{P}_{0.135\bullet\text{boot}}$, are estimated 99.865% and 0.135% percentiles with bootstrap method. Equation 3, advantage over equation 2 is that, it has a fixed false alarm ratio (type I error) regardless of what kind of parameters distribution [6]. It needs be noted that, when the process is stable and controllable, SPC chart is adequate for use.

D. Process Capability Index

In general, we use Cpk to measure the process capability in statistical. The concept of process capability only holds meaning for processes that are in a state of statistical control. Process capability index measures how much "natural variation" in a process is relative to its specification limits. If we define the upper and lower specification limits of the process are USL and LSL. Cpk can be denoted as

$$Cpk = Min\left\{\frac{USL - \mu}{3\sigma}, \frac{\mu - LSL}{3\sigma}\right\} \tag{4}$$

Basing on bootstrap method, we can get the Cpk as equation 5 & 6.

$$Cpk = Min\left\{\frac{USL - \mu_{\text{boot}}}{3\sigma_{boot}}, \frac{\mu_{boot} - LSL}{3\sigma_{boot}}\right\} \tag{5}$$

Or

$$Cpk = Min\left\{\frac{USL - med_{\text{boot}}}{P_{99.86\bullet boot} - med_{boot}}, \frac{med_{boot} - LSL}{med_{boot} - P_{0.135\bullet boot}}\right\} \tag{6}$$

wherein, med_{boot} is the median estimated by bootstrap method. It's just like SPC chart, equation 5 is adequate to the data of normal distribution, and equation 6 is adequate to non-parameter distribution.

III. Practice

In semiconductor industry, the incoming material quality control is one of the most important subjects for foundry manufacturing process, especially with the advancing of semiconductor technology into nanometer nodes. Therefore, how to effectively detect abnormality at the early stage becomes one more challenge topic [7]. SPC is a most commonly used tool for early detection.

The data we want to monitor is incoming COA (Certificate of Assurance) which includes series of parameters measured by suppliers in their manufacturing process. Usually, we just check whether the parameters are out of the specification defined in COA. In order to make excursion be early detected, SPC concept needs to be used to monitor significant process shift, even the measurement results are in the specification limitation. But it's different from the assumption of general SPC. Most parameters of incoming materials can't meet the normal distribution. If they use the general SPC method in equation 1, the false alarm can increase. Usually, it will cause excess waste in manufacturing engineering.

Since bootstrap method can be used in non-parameter distribution, and keep a relatively good accuracy, we try to use this to setup a SPC management and process capability index review system for incoming quality control.

Another problem we can solve by bootstrap method is the sample size. Especially for the non-parameter distribution, we need a large number of samples to accurately estimate the empirical distribution. In bootstrap we can use relatively small amount of data to do the same thing with good accuracy.

For example, we have a set of discrete data X with 30 data counts, and the histogram plot is as shown in Fig. 1.

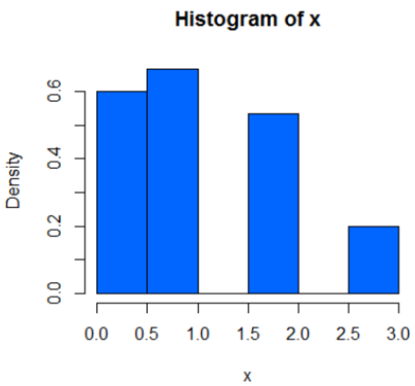


Fig. 1. Histograms of discrete data X.

And then, we want to compute its percentile like 90%, 95%, 98% and 99.9%. We use the nearest even order statistic method (SAS defined method) and bootstrap method to estimate these statistics.

TABLE I: THE COMPARISON OF PERCENTILE COMPUTE METHOD

Percentiles	90%	95%	98%	99.9%
SAS Method	2.0	3.0	3.0	3.0
Bootstrap Method	2.723574	2.724649	2.730213	2.731282

The comparison results are shown in Table I. For SAS method, when the percentage is larger than 95%, all percentiles are equal to 3. Its precision is low because of less data count. But bootstrap method has better precision to denote each percentile. Therefore, we can gain the precision control limit definition through bootstrap method.

But, bootstrap method has also its limitation in practice. Because it requires resampling thousands of samples from the original samples, it's hard to compute the process by manual. The computer program is required to do this job instead of traditional way. The R program language is used to do the calculation in this paper, because it has a convenient function to do resampling. The function can also be realized by SAS or Matlab.

For confidentiality reasons, the SPC & Cpk data presented below are fictitious.

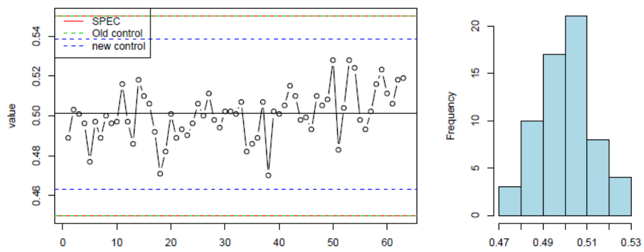
A. Practice in SPC Chart

We test the program with incoming material quality data. In this way, we selected COA data within 3 years to calculate the control limits. And then, we test the results within half year, which is compared with original alarm limit defined with experience of engineers.

The calculation database includes 6-categories of material type, 563 parameter counts and 1.66M data counts.

The samples of SPC chart for bootstrap method are shown in Fig. 2

The results are summarized in Table II. There are 3-types of results summarized after comparing the control limit values in two methods. The proportion of 3 types is showed in the table. "Equal" denotes the control limit set by bootstrap is equal to the one by original method, and "Loosen" denotes the control limit set by bootstrap is looser than that by original method, on the contrary "Tighten" denotes the control limit set by bootstrap is tighter than that by original method. The "Summary" column tells us about half of the SPC charts are tightened by bootstrap method and only about quarter of charts are loosened.



Red solid line denotes specification
Blue dotted line denotes new limit by bootstrap
Green dotted line denotes old limit by experience of engineers
Fig. 2. SPC chart and histogram plot.

TABLE II: THE COMPARISON FOR BOOTSTRAP SPC CONTROL VS. ENGINEERING EXPERIENCE CONTROL

Comment	*A*	*B*	*C*	*D*	*E*	*F*	*Summary*
Equal	14%	15%	8%	16%	9%	63%	25%
Loosen	33%	36%	24%	26%	22%	29%	28%
Tighten	53%	49%	68%	59%	69%	8%	47%

In order to further verify the results, another question is whether it's over engineering since so many charts are tightened. An important index to judge the effectiveness of SPC control chart is the alarm ratio or average run length when the process is stable. Therefore, we collect the alarm ratios for the control limits set by two methods within about half year. The alarm ratio comparison is showed in Table III.

The results show the alarm ratios are almost the same for both methods. 0.44% is very close to the alarm ratio of general SPC. The average run length is about 227, which means the process can continuous run without alarm in average 227 runs. The average run length of general SPC is 370. Considering there could be some minor shift, this alarm ratio is acceptable in practice.

TABLE III: THE ALARM RATIO COMPARISON FOR BOOTSTRAP SPC CONTROL VS. ENGINEERING EXPERIENCE CONTROL

Material Type	***Old***	***New (Bootstrap Control)***
A	0.43%	0.42%
B	0.68%	0.83%
C	0.27%	0.70%
D	0.44%	0.03%
E	1.05%	0.89%
F	0.08%	0.49%
Total Alarm Ratio	**0.43%**	**0.44%**

TABLE IV: THE DETECTION LEVEL IMPROVEMENT INDEX

Material Type	*ΔD*	
	Loosen	**Tighten**
A	-0.87	10.1
B	-1.36	4.3
C	-0.88	25.1
D	-2.03	50.2
E	-1.61	5
F	-1.57	1.7
Summary	**-1.4**	**14.6**

In the other way, we can check the detection level for both methods. For this index, we change the control limit value to a position far away from the data mean. And we use the ratio of standard deviation to express this distance, and it is denoted as D_{boot} for the distance by bootstrap method and D_{eng} for it by old method, and $\Delta D = D_{eng} - D_{boot}$ is defined as the difference of distance. A larger ΔD means a better detection level can be improved by new method. Table IV summarizes the results of detection level improvement.

The positive numbers show the improvement of detection ability for tightened charts and the negative numbers show loosen value of distance. One half of the charts are have detection distance improved by 14.6 times over the standard deviation in average. And a quarter of charts loosen the detection distance by 1.4 times over the standard deviation in average.

Combining the viewpoints in last section, we can get the conclusion that bootstrap method can provide the near average run length compared with the general SPC and old method. And it also provides higher detection level than old method does.

B. Practice in Process Capability Index

SPC is the real time control for single process shift or special pattern behaviors. In order to monitor the long term performance of process and evaluate the risk of incoming material by process variation, the process capability index Cpk based on bootstrap is utilized in manufacturing.

Fig. 3 shows the sample of monthly Cpk trend chart for one of the incoming parameters. In the period of March to August, 2013, the Cpk trend down occurred and the corresponding part is larger than the baseline variation. It seems that even the data could be non-parameter distribution, bootstrap method could still signal the abnormal fluctuation of reaction.

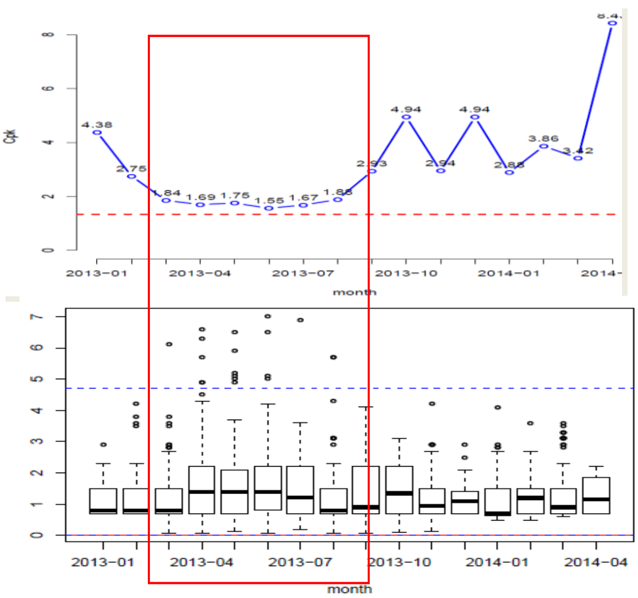


Fig. 3. Monthly Cpk trend by bootstrap.

IV. Conclusion

In the traditional field of semiconductor incoming material, less statistical method is used for monitoring since most of the data are non-parameter distribution and few data counts. The resampling method, bootstrap, provides relatively accurate estimation for non-parameter distribution. In practice, we combined bootstrap method into general SPC chart and process capability monitor, and present a new bootstrap-based non-parameter estimation method. The results show that, it can improve the detection level in about 14 times of standard variation but keep the same error level of general SPC. And the Cpk method with bootstrap-base can also detect the systemic process variation.

Although bootstrap can make the estimation on the small sample, it will still have a relatively high estimation error for very small sample size. The recommended sample size is larger than 30.

References

[1] R. Griffin, "Quality assurance and control of incoming material," *IET Journals & Magazines*, vol. 42, no. 9, pp. 463-467, 1964.

[2] D. L. Ma, Y. G. Gao, D. Shi, Y. M. Shen, Y. Z. Zhou, and Q. H. Yang, "Nonparametric bootstrap estimation of confidence interval in base station test," in *Proc. 5th Asia-Pacific Conference on Environmental Electromagnetics*, 2009, pp. 390-394.

[3] C. F. J Wu, "Jackknife, bootstrap and other resampling methods in regression analysis (with discussions)," *Annals of Statistics*, vol. 14, pp. 1261-1350, 1986.

[4] B. Efron, "Better bootstrap confidence intervals," *Journal of the American Statistical Association*, vol. 82, no. 397, pp. 171-185, 1987.

[5] R. C. H. Cheng, "Bootstrap methods in computer simulation experiments," in *Proc. Simulation Conference*, 1995, pp. 171-177.

[6] X. H. Niu and C. Spanos, "Statistical enhancement of a reflectometry metrology system," in *Proc. 1997 2nd International Workshop on Statistical Metrology*, 1997, pp. 40-43.

[7] V. S. Guan, S. July, and K. Chang, "The quality control application for abnormal raw material early detection," in *Proc. 2013 IEEE International Conference on IEEM*, 2013, pp. 1179-1183.

Sheng Kang is from SMIC PQE, who has total 10 years job experience for semiconductor process control. He graduated as master from the theory of probability and the mathematical statistic in East China Normal University in 2015. He focuses on statistical method application and SPC in semiconductor manufacturing. He is good at data analysis and statistical analysis. He has published several papers on IEEE.

Violet Shangguan is from SMIC IQE, who has total 9 years job experience for semiconductor manufacturing raw material incoming quality engineering which include incoming inspection and supplier management after graduated from ECUST in 2006 with master degree. She is familiar with raw material manufacturing process and application in fab, such as wafer, gas, chemical, PR, slurry, target, quartz, etc. She has abundant experience for supplier on-site audit and raw material related abnormal case handling. She knows more about supplier management method from new supplier introduce, approved vendor routine management, to supplier disqualify process as well as being familiar with supply chain management method from supplier to customer.

Lisa Yu graduated as a Ph.D. candidate from the School of Material Science in Harbin Institute Technology in 2003. She has worked in semiconductor manufacturing industry for more than 10 years, focusing on device level reliability, metrology tools MSA, raw materials quality control, supplier management etc. She is one of senior managers in quality and reliability engineering unit of Semiconductor Manufacturing International Corporation (SH).

Wei-Ting Chien was born in Taiwan in 1965. He received his Ph.D. degree in Industrial engineering from Texas A&M University in 1994. Before he came back to Taiwan for industrial jobs in semiconductor manufacturing in 1995, he did post-doc researches with HP, and IBM for a year on software reliability modeling, and highly reliable system optimization, respectively. After being with Nan-Ya Technology Corporation, Intel, and TSMC, he joined SMIC in Shanghai in 2001. In 2010, he was deputy operation director with STMicroelectronics in Singapore. He re-joined SMIC in 2011 as Corp Q&R vice president until now. His current assignments cover business process reengineering; quality system management; incoming, in-process, and outgoing quality control; subcontractor quality; reliability engineering; process and product qualification; failure analysis; and ESH (environment, safety, and hygiene). He has published over 170 papers, more than 60 patents, and is Senior Member of IEEE, Member of Phi-Tau-Phi Scholastic Honor Society, and a member of Shanghai 1000-Person Plan.