2016 6th IEEE International Conference on Control System, Computing and Engineering, 25–27 November 2016, Penang, Malaysia

In-situ Particle Monitor using Virtual Metrology System for Measuring Particle Contamination during Plasma Etching Process

Mohd Firdaus Abdullah, Muhammad Khusairi Osman, Noratika Mohammad Somari , Adi Izhar Che Ani ,
Faculty of Electrical Engineering
UiTM Pulau Pinang Pulau Pinang, Malaysia

Sooria Pragash Rao S.Appanan, Loh Kwang Hooi,
Infineon Technologies (Kulim) Sdn.Bhd.
Lot 10 & 11, Jalan Hi-Tech 7, Industrial Zone Phase II,
Kulim Hi-Tech Park, Kulim, Kedah, Malaysia

Abstract— This paper present an analysis on in-situ particle monitor using virtual metrology system for particle contamination measurement. In the process of manufacturing semiconductor devices, detecting particle contamination in process tools is a vital factor for determining the product yield. In-situ monitoring of particle contamination can be as accurate and cost effective method of contamination control depend on type of particle monitoring sensor selection and its data used for virtual metrology. In this study, data samples are obtained from three system; Statistical Process Control (SPC) data base, Advanced Process Control (APC) data base and Hamamatsu Multiband Plasma Process Monitor system. Then, an artificial neural network based classifier called multilayer perceptron (MLP) network is applied to measure the particle contamination level from the given dataset. The performance of MLP network is compared using two different algorithms namely Levernberg-Marquad (LM) and resilient back-propagation (RP) algorithm. Based on the simulation results, it can be concluded that the MLP network using LM algorithm gives the best regression result of 0.999 and 0.54 during the training and testing respectively. The outcome of this project is in-situ particle monitor would be able to detect particle in the oxide etch chamber as alternative for Surf-scan methodology for each processed wafers.

***Keywords*- In-situ Particle Monitor,Virtual Metrology System, Artificial Neural Network, Particle Contamination Measurement, Plasma Etching.**

I. INTRODUCTION

Plasma etching is the main process used to carry out precise control of wafer etching in modern semiconductor manufacturing facilities. In developed integrated circuits, plasma processes play a crucial role in depositing or patterning thin films [1]. Plasma etching offers process improved and simplification dimensional tolerances compared to wet-chemical etching technology [2]. Radio frequencies (RF) accelerate the generated plasma toward the electrode where it interacts with the masked wafer surface both mechanically and chemically to etch away the exposed surface. Optical emission spectroscopy (OES) can monitor the chemical composition of the exhaust gases from the chamber. The challenge with operating plasma etchers to successive wafers processed in the same etch tool and consistent etch rate for a given wafer to maintain. Etch rate variations cause many reasons include variability in the RF current discharge, the chemical interactions with chamber wall that effect chamber seasoning, non-uniformity in the composition of the plasma gases and temperature changes in the chamber during the etching step

The sensitivity to disturbances and their complex nonlinear performance of plasma etch processes sensitivity to disturbances makes them very difficult to control and model. Additional, in real time etch rate not available to measure. Before they are available for process adjustment, costly non-value added post-etch metrology step leading to significant delays needed to obtain etch rate measurement. Hence, in practice, there a small number of metrology measurement with pre-determined fine –tuned recipes need to performed during facilities process and statistical process control that used open loop fashion in plasma etch processes.

Valuable indicator of quality and properties of particulate materials and performance influences by particle size [3-5]. Examples are aerosols, emulsions, suspensions and powders. The shape and size of the powder influence compaction properties and flow. Larger, more spherical particles will typically flow more easily than smaller or high aspect ratio particles. Smaller particles lead to higher suspension viscosities and dissolve more quickly than larger one. The classification of particles according to the source of particle generation may be divided into three categories which are particles induced from the clean room environment, particles arising from the handling of wafers by workers and robots, and particles contributed by manufacturing tools and the actual processes themselves. A reduction in particles generated from the fabrication atmosphere has been obtained by improvements in the design and operation of clean rooms. Similarly, particle contribution from wafer handling has been reduced by the practice of workers and taking greater care of clean room garments. As a result, the particles induced from manufacturing tool predominate. This is due to the nature of the process by-product accumulate on process chamber. These by-product turn into particle on wafers and methods of reduction remain key challenge [6].

Virtual metrology (VM) is main factor in the wafer manufacturing community as advanced solution in critical process quality parameters by estimating from other accessible in-line process dimensions [2, 7]. Addition, referred as inferential estimation or soft-sensing, offers the possible to consider improve yield and enhance process capability in wafer production. Ping Hsu *et al.* [8] there improvement more than 65% shown in process capability by implement VM model integrated together with advanced process control (APC). Researches build virtual metrology (VM) models by using data mining based approaches estimating important parameters in production processes [9]. In addition, this is a historical process measurement in order to build the bonding between the important parameter and process variables. Then, trained model are used to approximate the important parameters on-line of control environment in a real-time [9].

In the process of manufacturing semiconductor devices, detecting particle contamination in process tools is a vital factor for determining the product yield. In-situ monitoring of particle contamination can be as accurate and cost effective method of contamination control depend on type of particle monitoring sensor selection and it is data used for virtual metrology. A Virtual Metrology (VM) system that fills to require physical measurement by prediction that enable consider an improvement in process control especially cut off the operational cost[9]. The purpose of this paper is to provide an effective method for early detection of the particle in the process chamber during oxide etch process using VM system.

II. METHODOLOGY

This section describes the purpose of ANN based virtual metrology system. The system is highly reproducible and repeatable at any time due to the deterministic nature of the using software. Thus, it can provide enhancement to the conventional physical metrology method. Figure 1 shows overview of the proposed Virtual Metrology system.

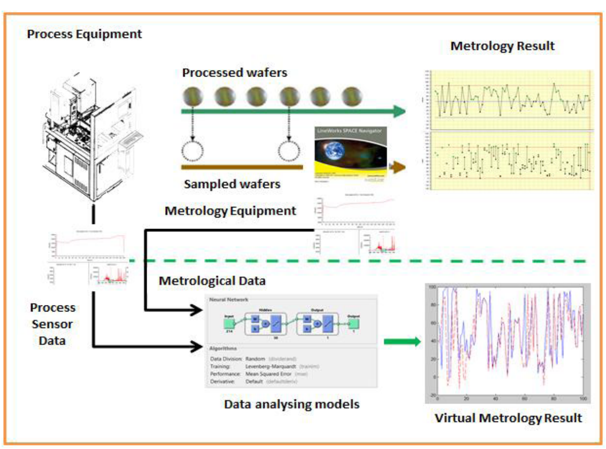

Figure 1: Overview of Virtual Metrology as enhancement of physical metrology

A. Data set Description

The data samples used from particle contamination count are collected from the SPC database, APC database and Hamamatsu Multiband Plasma Process Monitor. Hamamatsu Multiband Plasma Process Monitor guides the plasma radiation emitted from the main chamber of semiconductor manufacturing equipment into a polychromator through the optical fiber as in Figure 2. The CCD linear image sensor converts optical signals into electrical signals, passes them through the analog-to-digital (A/D) converter and stores them in the memory as digital signals (spectrum data). The spectrum data for intensity of light versus time which refer to dynamic light scattering will be collected.

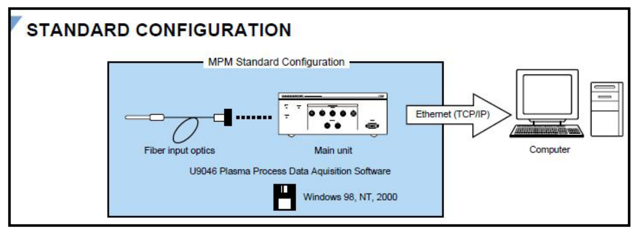

Figure 2: Standard configuration for Hamamatsu Multiband Plasma Process Monitor

Dynamic light scattering is also recognized as quasi-elastic light scattering or photon correlation spectroscopy and it a technique to verify the size distribution profile of small particles in suspension or polymers in solution [3]. In addition, it also used to probe the behavior of complex fluids as example, concentrated polymer solutions. Monochromatic and coherent light source if the light source is a laser and it scattering intensity fluctuates over time. Small molecules cause fluctuation as undergoing Brownian motion, the scatters between distances is constantly changing with time.

This scattered light then undergoes either destructive or constructive interference by the surrounding particles, and within this intensity fluctuation, information is contained on the time scale of movement of the scatters. Sample preparation either by centrifugation or filtration is important to remove artifacts and dust from the solution.

B. Data analysis

Target data samples can be selected to discover the desired knowledge and information. Data mining comprises the selection of suitable methods, the alignment with the defined goals according to the data selection and the selection of appropriate models including their parameters. The critical analysis if the previous steps match to the data mining technique and, if necessary, the initiation of adequate adjustments is also parts of the data mining step.

Data samples for a period time of 6 months were prepared and purified according to the data preparation steps. A total of 130 data samples were collected from SPC database, APC database and also Hamamatsu Multiband Plasma Process Monitor. Space navigator used for defect count for each wafer and total mean of defects for production wafer 5, 10 and 20. Then, equipment number and date were collected from the APC based on lot number from SPC database. Finally, by referring to the equipment number and date from APC data samples. Graph for light intensity against time was obtained from Hamamatsu Multiband Plasma Process Monitor. All the measurement ware set to the wavelength 532nm based on the scattered light measurement system by the CCD Camera inside Hamamatsu Multiband Plasma Process Monitor [4, 10].

C. Measuring Particle Contamination using Multilayer perceptron network with Levenberg Marquardt learning algorithm.

The final stage of this project is to measure the particle contamination using Artificial Neural Network. For plasma etching process, takes 212 second and therefore produce 212 data are used as inputs to the ANN to determine the number of particles. ANNs become an alternative modeling method as problem-solving tools to non-physical and physical systems with mathematical basis or scientific. It mimics the process of human learning using a relatively crude electronic models. It can be trained as human brains and master to gain task by some trained. It also master task through experiential knowledge but problem solving tools with patterns can be trained by ANNs and need to complete by optimization performance through training process. Neuron is consisting of simple processing units. MLP network is one type of ANN and its can known as a system of massively distributed parallel processor. It have natural tendency to utilizing and storing experiential knowledge by a system of massive distributed parallel [11]. Normally, the MLP learns based on relationship set of inputs and outputs that used the back-propagation learning algorithm by weight. Weight is updating internal connections [12]. To solve a nonlinear least square, attempts the algorithm by minimization problem in the form of:

$$f(x) = \frac{1}{2}[r(x)]^2 \qquad (1)$$

where r is the residual vector. Learning rule will take larger steps in flat to pass over plateaus in short time quickly when through for the minimum on the error surface, and when encounter a large gradient it take smaller steps because pretend overstepping in local minima. The LM algorithm does this by combining curvature as well as gradient information based on two update rules, the Gauss-Newton (GN), and the gradient descent (GD) rule. The rule for LM algorithm is given by:

$$x_{i+1} = x_i - (H - \mu diag(H))^{-1} \Delta f(x_i) \qquad (2)$$

when H is the approximated Hessian matrix, $\Delta f(x_i)$ is the gradient of the error function, and μ is the learning rate. LM algorithm designed to become second-order training speed and not having complex problem with Hessian matrix [13]. The LM algorithm used approximate Hessian Matrix by GN method when μ value is zero. Then, small step size by VGD algorithm was used when μ is large value. The GN method shows the accurateness within an error minimum and faster, then to achieve shift towards in Newton's method can obtain. Therefore μ increased tentative step could increase the performance and μ is decreased after each successful step. This could increase the performance of the function. In this step, each iteration of the algorithm always reduced the performance function [13].

The implementation of LM algorithm can be summarized as step below:

1. Set an update by following step (2).
2. Set up new parameter vector by evaluating r(x).
3. Set new weight to previous values and increasing μ is by some significant factor, α when r(x) is increased by referring updated result. Do repeat step 1.
4. Continue with same weights at new values when r(x) result is decreased. Then decrease value μ by a factor of β, and redo step 1.
5. Do continue above step until met the condition needed.

III. Result And Discussion

All the measurement was set to the wavelength 532nm [4, 14]. A total 100 data sample was collected and divided into pats; training and testing sets. 70% from the data sets will be used as training set, 25% used as testing set, while the remaining as validation set. The intensity versus time data for each lot number was tested using MLP network train by LM and further compared with the Resilient back propagation algorithm. Both algorithms were trained and tested to find the best performing algorithm. The training process was implemented using MATLAB Neural Network Toolbox. Training of the MLP network was performed with the function of 'trainlm' for LM algorithm and 'trainrp' for Resilient algorithm. Table 1 shows results of MLP network for both algorithms with different number of hidden nodes.

TABLE 1: PERFORMANCE COMPARISON OF MLP NETWORK TRAINED USING LM AND RESILIENT BACKPROPAGATION FOR DIFFERENT NUMBER HIDDEN NODES

Hidden No	Algorithm	R		RMSE	
		Training	Testing	Training	Testing
10	Trainlm	0.754	0.426	27.458	35.5
10	Trainrp	0.376	0.015	31.023	36.9
12	Trainlm	0.866	-0.109	26.28	46.9
12	Trainrp	0.534	-0.311	30.4	41.4
14	Trainlm	0.603	0.255	29.67	39.22
14	Trainrp	0.507	0.043	28.04	34.4
16	Trainlm	0.953	-0.601	33.13	79.3
16	Trainrp	0.695	0.045	26.29	37.06
18	Trainlm	0.454	-0.231	34.89	40.8
18	Trainrp	-0.099	-0.167	53.14	55.5
20	Trainlm	0.877	0.064	24.45	38.86
20	Trainrp	0.237	-0.1	40.25	45.7
40	Trainlm	0.967	0.206	26.3	47.7
50	Trainlm	0.94	0.526	20.25	26.9
50	Trainrp	-0.05	-0.55	87.85	95.96

To ease the performance comparison of both algorithms, two performance metrics namely correlation coefficient, R and root mean square error were used. The regression coefficient is used to determine how nearly the points fall on a straight line, or how nearly linear they are. A perfect correlation will have a regression coefficient of R = 1.000. Normally in the physical sciences, we would like to have a "confidence level' of 0.01 or better. That means that a coefficient of R =0.990 or higher gives us the confidence to say that a relationship is linear within a margin of tolerable error.

Based on Table 1, the MLP network with 50 hidden nodes and trained using LM algorithm gives the best R value of 0.94 and 0.526 for training and testing respectively. It can also be seen that LM algorithm outperformed the Resilient back propagation algorithm in both training and testing result, the LM algorithm always gives the best results. It is often the fastest back propagation algorithm in the toolbox, and is highly recommended as a first-choice supervised algorithm, although it does require more memory than other algorithms.

The root-mean-square error (RMSE) is a frequently used measure of the differences between values predicted by a model or an estimator and the values actually observed. Basically, the RMSE represents the sample standard deviation of the differences between predicted values and observed values. RMSE is a good measure of accuracy, but only to compare forecasting errors of different models for a particular variable and not between variables, as it is scale-dependent. Based on Table 1, both algorithms give higher value of RMSE. Further studies need to be done to reduce the value of RMSE.

Since LM algorithm gives the best result compare to Resilient back propagation algorithm, training has been done using Levernberg Marquad training algorithm by adding another 30 new data. Hidden number was varied from 10 to 50 to find the best results. The parameters shown in Table 2 shows the value of parameter to training LM algorithms. All parameters set up as Table 2 and settings as default values in MATLAB during MLP Training.

TABLE 2: THE LM SETTING PARAMETERS VALUE IN TRAINING THE MLP NETWORK

Parameter	Value
Maximum epochs	1000
Training goal	0
Minimum Δf (xi)	1×10^{-16}
μ	1×10^{-3}
A	0.1
B	10
Maximum μ	1×10^{16}

For the LM algorithm, when the number of iterations beat the maximum epochs the training will stops and not perform cause effect the training goal. More than that, magnitude takes part to stop when the gradient is less than minimum Δf (xi) or exceed the maximum. Table 3 shows results for LM algorithm when varying the hidden nodes. Based on table 3, the best regression result for training and testing is 30 hidden nodes.

TABLE 3: VARYING PARAMETER (HIDDEN NODES)

Hidden No	Regression		No. of Iteration
	Training	Testing	
10	0.798	0.485	9
20	0.727	0.626	8
30	0.999	0.537	8
40	0.878	0.571	8
50	0.587	0.274	9

Figure 3 shows comparison between desired numbers of particle (blue line) and predicted by the MLP network (red) against number of samples. It shows a huge different between this two graph. Figure 4 shows the regression plot of training and testing. Training result gives the best regression is 0.999, while during testing the regression is 0.537.

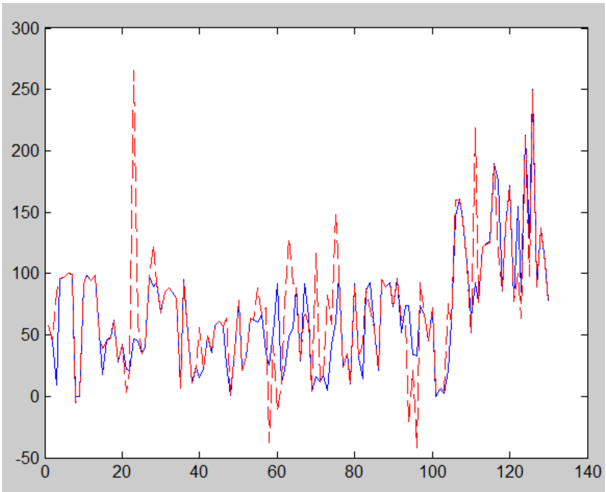

Figure 3: Comparison between predicted values (red) with the actual value (blue) for Trainlm with 30 hidden nodes.

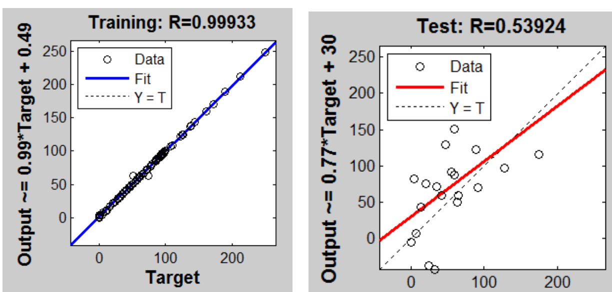

Figure 4: Neural Network Training Regression for Trainlm with 30 hidden nodes.

IV. CONCLUSIONS

In-situ particle monitor using virtual metrology system for particle measurement was presented in this paper. Based on the results, it can be concluded that Trainlm algorithm gives the best regression result, 0.999 when the hidden node is 30 during the training, while during testing the regression is 0.537. This result will be improved so that the predicted output will be same as actual output. The outcome of this project is in-situ particle monitor would able to detect particle in the oxide etch chamber as alternative for Surf-scan methodology for each processed wafers.

V. ACKNOWLEDGEMENT

The authors would like to acknowledge Collaborative Research in Engineering, Science and Technology Center (CREST), RAGS 600-RMI/RAGS 5/3, Infineon Technologies Sdn. Bhd and Universiti Teknologi MARA, Penang, Malaysia for their financial and equipment assistant for the project research. Special thanks to Tam Siew Chin and New Chi Eng from Infineon Technologies in sharing the process knowledge on plasma etching.

VI. REFERENCES

[1] D. W. L. B. Kim, K. Y. Park, Serk Rim Choia, Seongjin Choi, "Prediction of plasma etching using a randomized generalized regression neural network," Surface Engineering, Surface Instrumentation & Vacuum Technology, vol. 76, pp. 37-43, 2004.

[2] I. Shin-ichi, "Virtual metrology for plasma particle in plasma etching equipment," in Semiconductor Manufacturing, 2007. ISSM 2007. International Symposium on, 2007, pp. 1-4.

[3] N. I. F. Uesugi, T. Moriya, H. Doi, S. Sakamoto, Y. Hayashi, "Real-time monitoring of scattered laser light by a single particle of several tens of nanometers in the etching chamber in relation to its status with the equipment," Journal of Vacumm Science and Technology, vol. 16, 1998.

[4] M. T. e. al, "Particle Reduction and Control in Plasma Etching Equipment," in IEEE Transaction on Semiconductor Manufacturing,, 2005, pp. 477- 486.

[5] T. M. T.Moriya, H.Nakayama, H.Nagaike, E.Sugawara, Y.Kobayashi, M.Shimada and K.Okuyama, "Reduction of particle contamination in an actual plasma etching process," Japan, 2005.

[6] T. K. H. Miyashita, Y. Kawasaki, Y. Katakura, N. Ohsako, "Particle measurements in vacuum tools by in situ particle monitor," J. Vac. Sci. Technol. A 17, vol. 3, 1999.

[7] Y. K. Y. J. Chang, C.-L.Hsu, C.-T. Chang and T. Y. Chan,, "Virtual Metrology Technique for Semiconductor Manufacturing Neural Networks," in IJCNN' 06, International Joint Conferance 2006, pp. 5289-5293.

[8] W. S. PingHsu C, Junshien L, Ko F, Wang J, Yu C. H, and Liang M.S., "Virtual metrology: a solution for wafer to wafer advanced process control," in IEEE International Symposium on Semiconductor Manufacturing, 2005, pp. 155 -157.

[9] B. Lenz and B. Barak, "Data Mining and Support Vector Regression Machine Learning in Semiconductor Manufacturing to Improve Virtual Metrology," in 46th Hawaii International Conference, 2013, pp. 3447-3456.

[10] J. Simmons, M. Scottney-Castle, and C. Maskiell, "Successful Integration of in-situ Particle Monitoring into a Volume 300mm High Current Implant Manufacturing System," in Proceedings of the 14th International Conference onIon Implantation Technology USA, 2002.

[11] Z. F. M. Nadiah, J. Aiman, M. Y. Ihsan, Z. Azlee, A. H. Hasliza, "Comparison between Levernberg- Marquad and Scaled Conjugate Gradient Training Algorithms for Breast Cancer Diagnosis using MLP," in 2010 6th International Colloquium on Signal Processing & Its Applications (CSPA), pp. 162-168.

[12] A.Ranganathan, "The Levenberg-Marquardt Algorithm," Georgia Institute of Technology, 2004.

[13] H. Demuth and M. Beale, MATLAB Neural Network Toolbox v4 User's Guide.Mathworks Inc., 2005.

[14] W.C.A.Youthapolnavee,A. Chaisawadi,, "A construction of Particle Counter by Using Laser Light Scattering," in 6th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology, 2009, pp. 452-455.