AEC/APC Symposium Asia 2009

IDMS-O-018

VM Capability Evaluation on Neural Network and PLS– Tomonori Kishimoto

Noriyuki Inomata

Tnkisimo@micron.com, Inomata@micron.com

Micron Technology

302-2, Hirano-Cho, Nishiwaki, Hyogo, Japan

Phone: +81 -795-23-6608 Fax: +81-795-23-6629

BACKGROUND

The process control spec is getting tighter on the latest technology node of the semiconductor production. However, it is no longer realistic to measure and analyze each production wafer. The Virtual Metrology (=VM) technique is becoming a more important method for the future because it is expected to predict the process outputs accurately without actual measurement, thus reducing the number of the test wafers. In this paper, we describe the practical VM model by the back-propagation Neural Network (=NN) [1] that is used as prediction model for the other industry field. In addition, the advantage of the NN is represented by the comparison of prediction capability with the Partial Least Square (=PLS) method, which is commonly used for the VM technique in the semiconductor manufacturing environment.

METHODOLOGY

To build a robust VM prediction model, tool parameters that have a strong correlation with the process outputs must be determined in advance. However, the NN generally does not have the capability to extract such parameters. We prepare two networks for the “analysis” and the “prediction”. First, the analysis network evaluates the sensitivity of each tool parameters to the process output. If the sensitivity is big, then that tool parameter has a strong correlation with the process output. Second, the prediction network sequentially excludes the tool parameter of the lowest sensitivity and evaluates the prediction accuracy from the training dataset, composed of the tool parameters and the process outputs. When this accuracy is the best, the prediction network determines that the remaining tool parameters are the most optimal to predict. And the NN prediction model is created on the prediction network from these parameters that have strong correlations with the process output. At this time, we use the trial-and-error method [2] to determine the most optimal number of the middle layer units (Fig1). In addition, we create another prediction model by the PLS technique. As well as the NN model, the PLS model is composed of the optimal tool parameters. These models are applied to the prediction dataset to compare prediction capability.

EVALUATION AND RESULTS

We applied these prediction models for two Chemical Vapor Deposition (=CVD) processes. One is process-A, which generally shows much stable performance. Another is process-B, which is comparatively unstable and supposed to be difficult for the process output prediction. The common dataset for the training and the prediction were used by both prediction models. 80 runs of the training dataset were applied for process-A and 160 runs for process-B. The number of prediction dataset was 250 runs at both processes. As the result, both models showed comparable determination coefficient R2 (77%) on process-A, as shown in Fig 2. However, the PLS model did not show good R2 (0.3%) on process-B, in contrast with the excellent result on the NN model (81%), as shown in Fig 3. This was caused by 1) a big drop of one tool parameter (RF Reflect Power), as shown in Fig 4 and 2) the parameter had quite a high contribution for the process output in the PLS model. That is, the PLS model could not predict the process output very well when the deviation level of the tool parameter went beyond that of the training dataset. On the other hand, the NN model kept good predictability, as shown in Fig 5, in spite of the RF power deviation because it was not considered as a high contribution parameter for the process output, as shown in Fig 6.

CONCLUSION

We could verify that the practical VM model could be designed by the analysis NN and the prediction NN. Both PLS and NN models showed comparable prediction capability as long as the deviation of the tool parameters was well studied in the training dataset. However, if the deviation goes beyond that which was studied in training dataset, NN showed much better prediction capability compared to that of PLS model.

REFERENCES

[1]Study and its Algorism – Neural Network
IEE, morikita publication, 8/28/2002

[2]Design and Application of Neural Networks
Bahman Kermanshahi, shoko-do, 6/30/1999

Fig 1. The number of middle units by Trial & Error

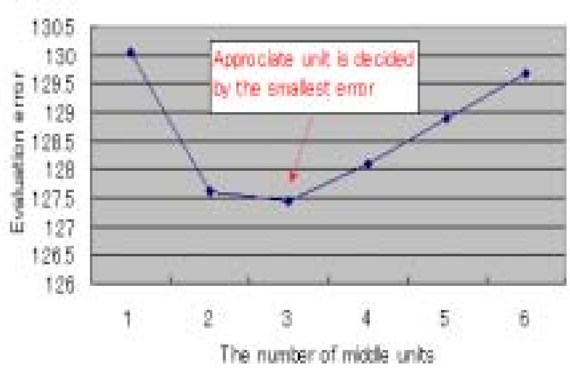

Fig 2. Neural Network vs. PLS at Process A

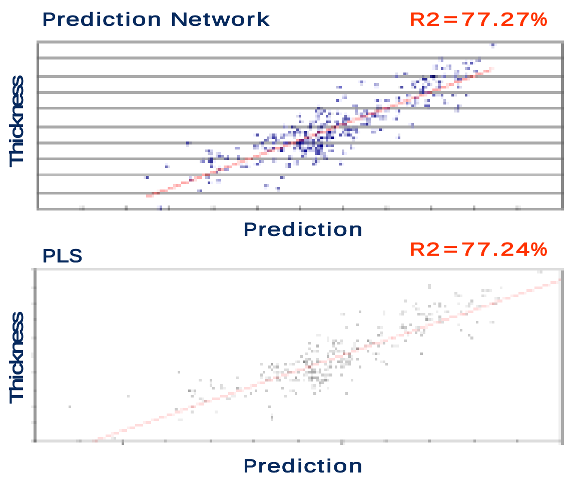

Fig 3. Neural Network vs. PLS at Process B

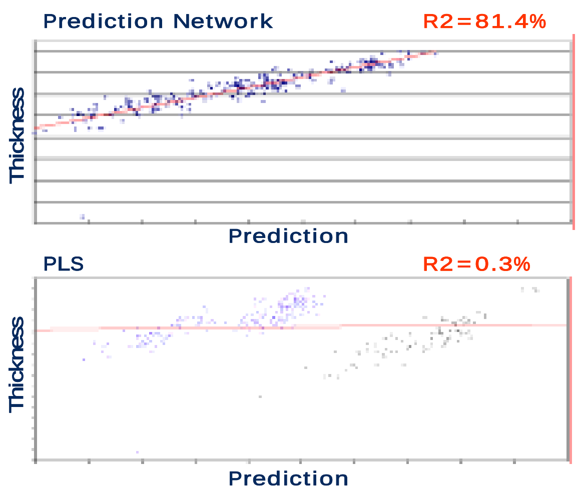

Fig 4. ReflectPower trend at Process B

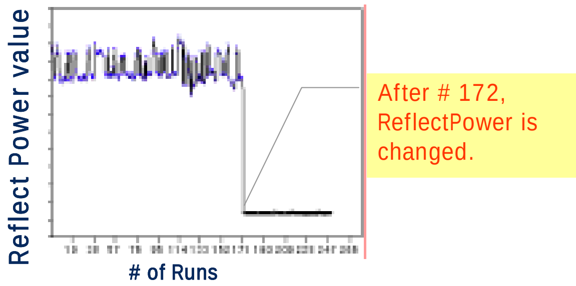

Fig 5. Prediction trend at Process B

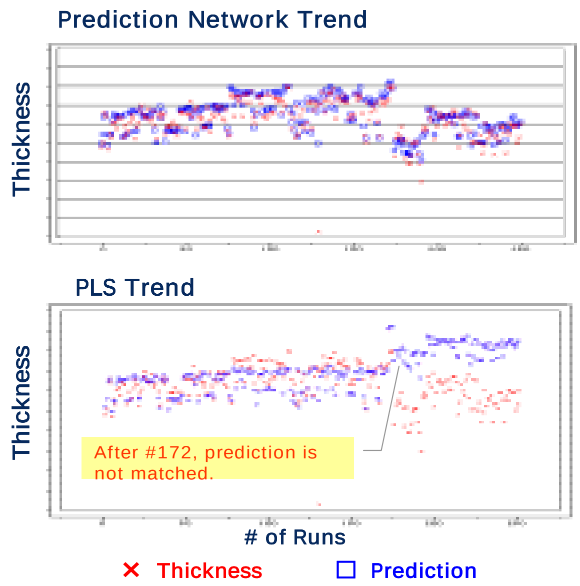

Fig 6. Tool parameter sensitivity at Process B