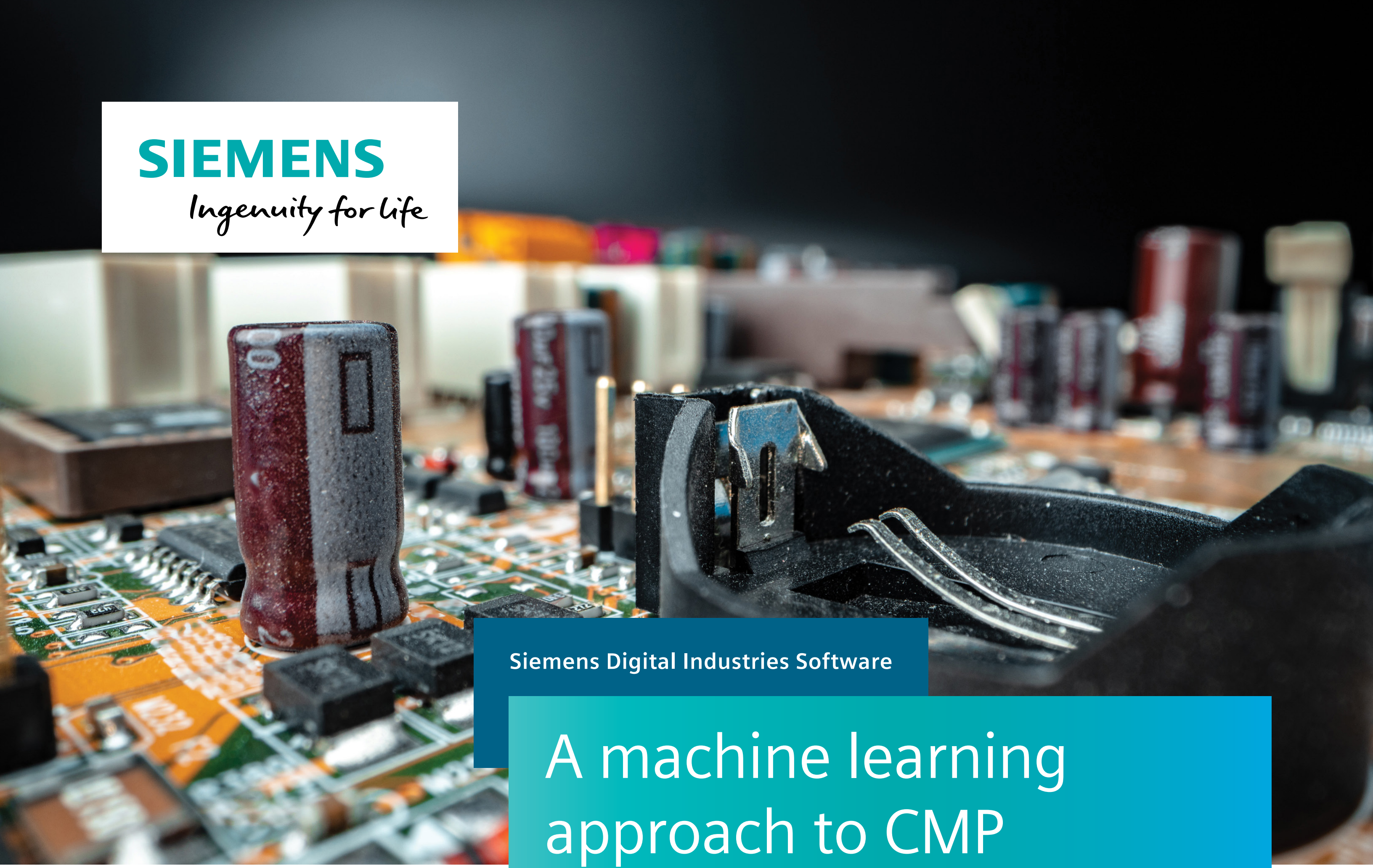

Siemens Digital Industries Software

A machine learning approach to CMP modeling

Executive summary

Most IC manufacturers use CMP modeling to detect potential hotspots as part of their DFM flow. However, building physics-based or compact models for FCVD and eHARP CMP processes has proven challenging, since these processes include several deposition and annealing steps to fill up trenches. Experiments show that using machine learning and neural networks for oxide deposition profile modeling for these and other CMP processes is a promising and exciting use of this technology.

Ruben Ghulghazaryan and Jeff Wilson
Siemens Digital Industries Software

Introduction

Chemical-mechanical polishing (CMP) is a key activity in today's integrated circuit (IC) manufacturing process. With designs so tightly packed, and scaled down to the most advanced process technology nodes, post-CMP planarity variations can have a significant impact on manufacturing success.

In an attempt to mitigate any negative impacts from the CMP process, most IC manufacturers use CMP modeling to detect potential hotspots in front-end-of-line (FEOL) and back-end-of-line (BEOL) layers as part of their design for manufacturing (DFM) flow. CMP hotspot analysis looks for areas of the design that have a higher than average probability of experiencing defects post-CMP. Since different materials exhibit different erosion rates under the CMP process, it is important to maintain a constant density balance across the die to prevent bumps and dishing that can cause shorts and opens in the metal interconnects. CMP analysis measures various aspects of the layout to ensure even planarity as the chip is built up over multiple layers.

The introduction of high-k metal gate (HKMG) technology with additional CMP steps[1, 2], the high cost of lithography due to double and triple patterning, strong depth of focus (DOF) requirements, and improved accuracy of CMP models have all increased interest in CMP modeling[3-6].

Building a CMP model

CMP modeling has a long history that includes modeling of single and two-material polishing, as well as numerous deposition and etch processes [6]. The main concept behind CMP modeling is to extract geometrical properties of the pattern on the layout, generate a pre-CMP surface profile after etch and numerous deposition steps, and predict the post-CMP surface profile for different patterns on the layout.

A chip is divided into tiles of fixed size, and for each tile, the average geometric characteristics of a pattern (such as width, space, pattern density and perimeter) are extracted and passed for etch, deposition, and CMP simulations. An effective trench approximation is used to model the structure defined for each tile (figure 1). A tile is considered to represent a trench with given geometric characteristics and data for two heights (Z_T and Z_{NT}) that define the height of material inside the trench and the height of material outside the trench.

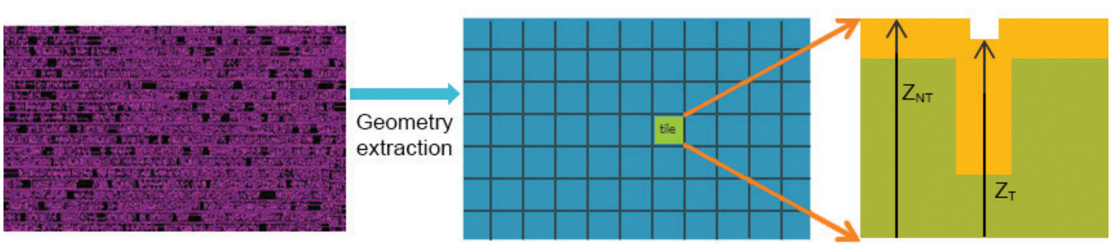

Figure 1. Geometry data extraction and effective trench approximation.

During CMP simulation, etch, deposition and polishing models simulate the change in Z_T and Z_{NT} data, as well as the geometry data change for each tile. The polishing model uses pre-CMP surface profile data as input, which is generated by either a deposition model or a previous polishing step. The first polishing step always uses the post-deposition profile as input. The surface profile after deposition is not planar and contains variations. Thus, for high quality CMP modeling, it is important to have a set of deposition models corresponding to the deposition processes used by manufacturers to generate the correct input profile for CMP simulation.

A critical step in CMP modeling is model building using measurement data from test chips. A CMP test chip usually consists of periodically-placed array blocks of parallel trenches of different widths with differing spaces between them (figure 2). The size of the test chip and the number of structures must be selected in a way that provide good coverage of width, space, perimeter, and pattern density values supported by the technology node, without violating design rule checks (DRC). An atomic force microscope (AFM) scanner or other profiler tool is often used to collect erosion and dishing data from line scans over test patterns, as shown in figure 2. By knowing the layer stack information and material thicknesses, erosion and dishing data may be converted into Z_T and Z_{NT} surface profile heights data.

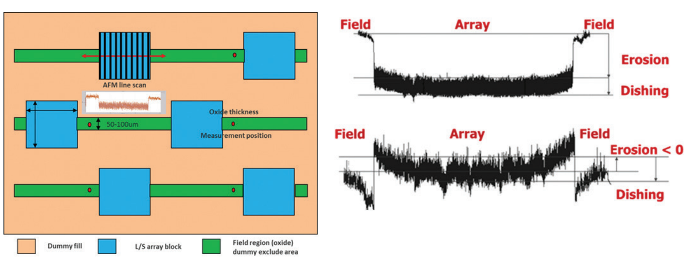

Figure 2. CMP test chip and AFM line scans with erosion and dishing definition.

Generation of a high-quality pre-CMP surface profile is crucial for accurate CMP model building, due to the complicated nature and long-range effects of CMP. Even with advanced deposition processes, the post-deposition (pre-CMP) profile on a patterned wafer is non-uniform, and may contain large variations that can effect on-surface planarity after CMP. Analysis of three-dimensional (3D) atomic force microscopy (AFM) and transmission electron microscopy (TEM) data shows complicated pre-CMP profile height dependence on the underlying pattern geometry for:

- high-density plasma CVD (HDP-CVD)
- spin-on dielectric (SOD)
- flowable CVD (FCVD)
- enhanced high-aspect-ratio processes (eHARP)

Shallow trench isolation (STI) and CMP modeling of FEOL layers showed successful application of HPD-CVD and SOD deposition models to CMP modeling.

Meanwhile, building physics-based or compact models for FCVD and eHARP processes is challenging, since these processes include several deposition and annealing steps to fill up trenches. We investigated the use of sensitivity analysis of measurement data with machine learning (ML) algorithms, which showed that the post-deposition surface profile for these processes depends mainly on the underlying pattern geometry, while long-range effects are secondary. With this information, we can apply neural networks (NN) regression calculation to the modeling of the pre-CMP surface profile, using as input the geometric characteristics of the underlying pattern. Later, the pre-CMP profile is used as input for CMP modeling.

Neural network configuration

ML, NNs, and deep learning have many applications in different areas of modern industry and life. Their ability to "learn" how to analyze and predict imprecise data has dramatically improved the state-of-the-art in areas of speech recognition and language translation, genomics and drug discovery, computer vision, autonomous vehicles, and many others[7].

One potential new use is the application of NNs to post-deposition surface profile modeling. The method works as follows: Local geometric characteristics of a pattern (width, space, pattern density, and perimeter) are extracted from the layout using the Calibre® CMP ModelBuilder and Calibre CMPAnalyzer products, and used as input to a multi-layer NN to generate surface profile height data predictions (figure 3).

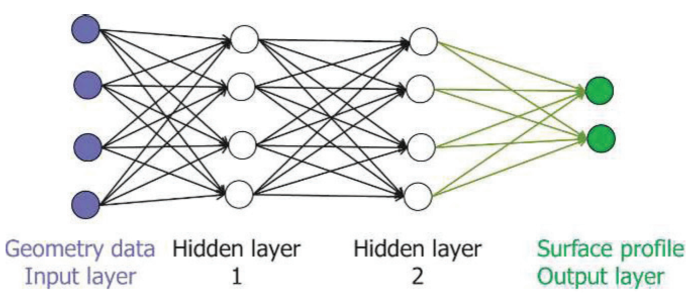

Figure 3. Multi-layer feed-forward neural network with two hidden layers.

The input layer of the NN takes this geometry data, and the network generates predicted erosion and dishing data on an output layer. For simplicity in this paper, any geometry change after deposition is not considered. The NN is trained using an advanced learning algorithm with a training set constructed of measurements collected from CMP test chips after deposition. The trained NN is used to run on a test chip or production designs for testing and validation. We considered an NN with one, two, or more hidden layers for fitting erosion and dishing data. To obtain good generalization of the model to unknown data, one should start from a minimum number of hidden layers and neurons, and increase them continuously to get better fitting on validation data, but over-fitting should be avoided. We determined that for our model, two hidden layers are optimal for modeling the surface profile of the mentioned processes, and there is no need to use deeper architectures (i.e., an NN with many hidden layers).

First, we used NN to model pre-CMP profiles of HDP-CVD and SOD processes for which compact models were available in the Calibre CMP ModelBuilder tool. Using measured data, we generated a training set used for NN training. Next, we performed validation on measured data and simulated data generated by models using the Calibre CMP ModelBuilder tool. Finally, we applied the NN to modeling of pre-CMP profiles for FCVD and eHARP processes.

Application of neural networks to profile modeling for CMP

To test the practicality and accuracy of using ML and NNs to generate CMP models, we experimented using the following four deposition processes: HDP-CVD, SOD, FCVD, eHARP.

HDP-CVD process modeling: The HDP-CVD process was originally used for STI, but nowadays it is widely used for the deposition of different oxides in combination with high aspect ratio deposition processes in chip manufacturing. Because both deposition and ion sputtering processes happen simultaneously during HDPCVD, triangular and trapezoidal shapes occur over active areas. The geometry of those shapes varies with the underlying active area pattern geometry, resulting in variation in deposited oxide thickness (figure 4).

Figure 4. Cross-section view of post-HDP-CVD surface profile.

Erosion and dishing data are collected from AFM line scan data from the test chip, and a training data set (table 1) is constructed for NN input in the format:

Input				Output	
Width	Space	Pattern density	Perimeter	Erosion	Dishing

Table 1 – Training data set

Input and output data normalization is performed for NN training. We used a tanh(x) activation function between input and hidden layers, and a linear activation function from the last hidden layer to the output layer. We then applied a back-propagation gradient-based optimization algorithm for NN training. Figure 5 shows fitting of 150 sites of the training set with a NN of two neurons per hidden layer. An absolute error per site is less than 3%. Due to the small fitting error, the simulated and measured data are indistinguishable on the plot.

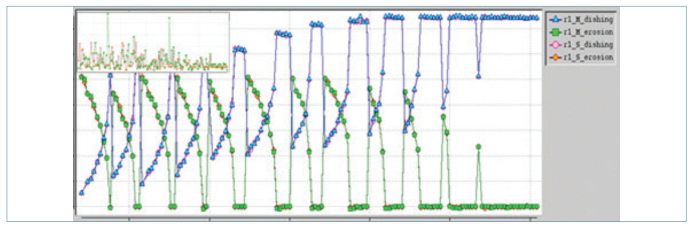

Figure 5. HDP-CVD erosion and dishing data fit during NN training.

Next, we ran a model validation using a different CMP test chip. Figure 6 displays the fitting of erosion and dishing data of the NN model versus compact model data for 33 structures. The error of fitting erosion and dishing data is larger than for the training set. As expected, the error is larger for the input values that are farther from the values used for training. Because the NN model is not physics-based, it may report non-physical results like small negative dishing, or erosion for some narrow or wide trench sites. While it is safe to set these data to zero, in general, this may be avoided if a large-enough data set is used for training. We observed that the NN model prediction has 98-99% correlation with compact model data, and the root-mean-square (RMS) error is roughly equivalent to the measurement error. The NN model predictions are in good agreement with both AFM line scan data and the HDP-CVD compact model.

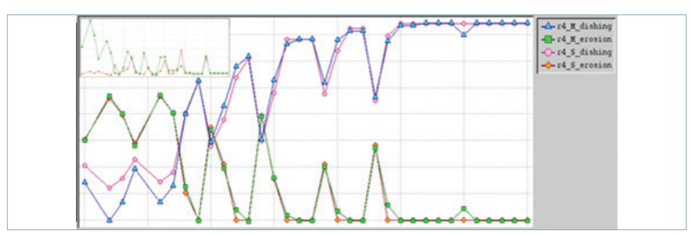

Figure 6. HDP-CVD validation error per site is about 10%, and for some sites dishing error is about 40%. Error subplots are not in scale with the fitting plots.

SOD process modeling: The SOD deposition process is used to coat the wafer with oxide material, which is originally in liquid form. Liquid is dispensed onto the wafer surface in a pre-determined amount, and the wafer is rapidly rotated. During spinning, the liquid is uniformly distributed over the wafer surface by centrifugal forces, and fills trenches. The material is then solidified by a low-temperature bake. In semiconductor processing, the SOD process is usually used to apply photoresist or to deposit pre-metal dielectric.

For NN modeling of the SOD surface, a NN with two hidden layers and four neurons per hidden layer (figure 3) was trained on a training set of 145 sites (figure 7).

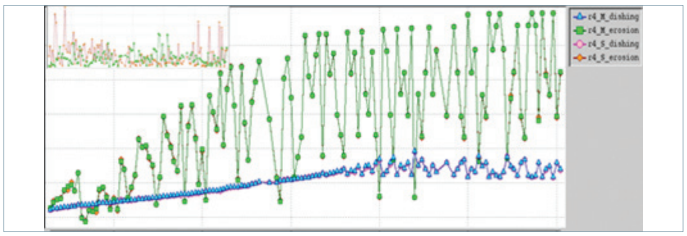

Figure 7. SOD process erosion and dishing fit after training.

The validation of the model was done on a different CMP test chip with SOD compact model simulated data for 35 sites (figure 8). Here also, the overall data fitting looks good except for a few sites, which can be corrected by extending the training set with more data.

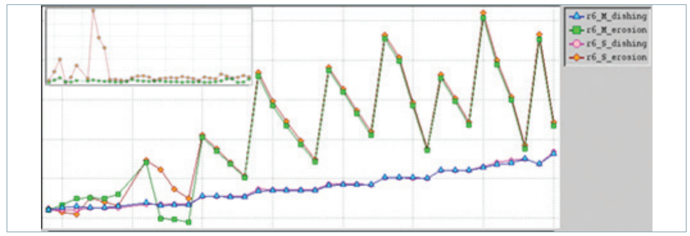

Figure 8. SOD model validation with error subplots. Validation error per site is less than 5% and for some sites, erosion error is about 50%. Error subplots are not in scale with the fitting plots.

Modeling FCVD and eHARP surface profiles with neural networks

The FCVD process, developed by Applied Materials[8], deposits a high-quality dielectric film in a liquid-like state on the wafer surface, allowing the film to readily flow into the gap, filling it completely and without voids or seams. eHARP is a non-plasma-based CVD oxide film deposition process that addresses the gap filling requirements for STI at the 4xnm node and beyond[8]. The eHARP process may be used before HDP-CVD to fill narrow trenches. These processes are too complicated for building a physics-based or compact model with a reasonable runtime and accuracy for CMP modeling. Thus, an NN is used to model surface profiles after FCVD, and the HDP-CVD profile after eHARP.

An NN consisting of two hidden layers with six neurons per layer is used for modeling (figure 9).

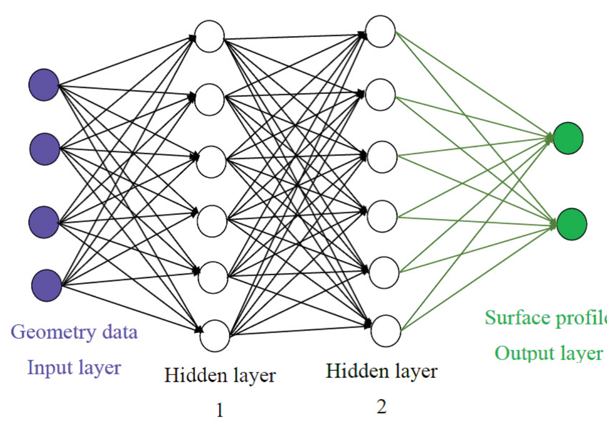

Figure 9. Multi-layer feed-forward neural network with two hidden layers and six neurons per hidden layer.

In figure 10, fitting of erosion and dishing data of the FCVD process on a training set with #163 measured data is shown.

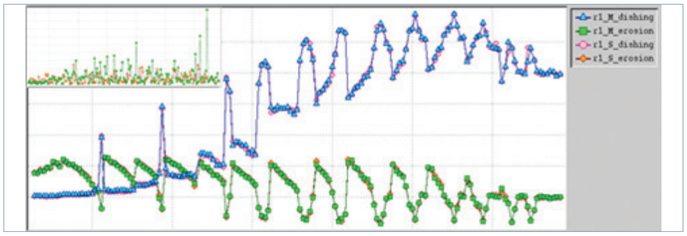

Figure 10. FCVD process erosion and dishing data fitting after training.

Correlation of dishing and erosion data on a validation set with 38 sites is 98% and 95% respectively (figure 11). The average error is about 15-20%.

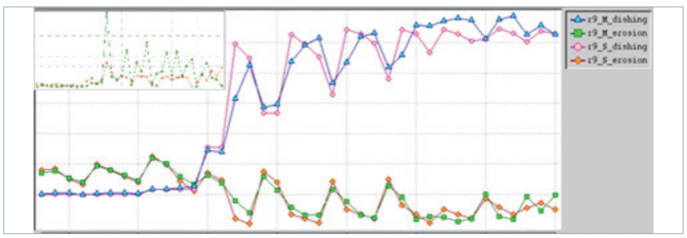

Figure 11. FCVD validation. Error subplots are not in scale with fitting plots.

In figure 12, the results of an NN model of HDPCVD surface profile after eHARP are shown, since it is the profile that is used for CMP modeling. The same NN configuration is used as for the FCVD case (figure 9). Training was done by fitting of erosion and dishing data on a training set with 61 measured data.

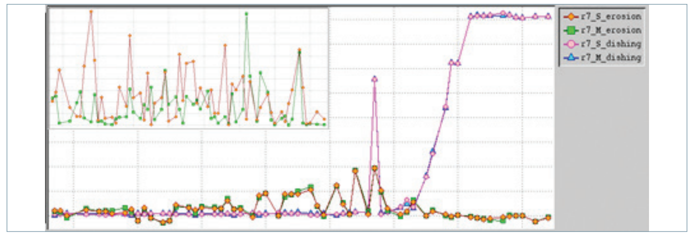

Figure 12. HDP-CVD after eHARP process erosion and dishing data fitting after training.

Figure 13 shows the validation set with 18 measurements with error subplots. The average error is about 10-20%.

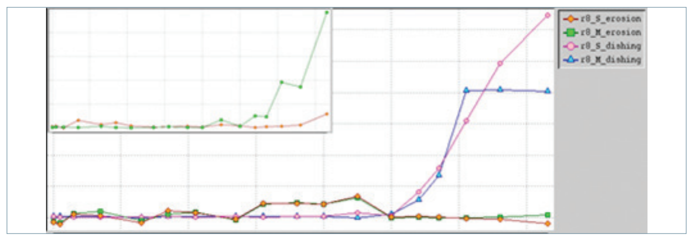

Figure 13. eHARP model validation. Error subplots are not in scale with fitting plots.

Conclusion

The application of ML and feed-forward NNs to post-deposition surface profile modeling of advanced deposition processes presents an exciting opportunity to develop accurate CMP models for complex CMP processes. In our experiments, a simple feed-forward NN with two hidden layers and a small number of neurons on the hidden layers was successfully used for surface profile modeling of HDP-CVD, SOD, FCVD and eHARP deposition processes, with more than 95% correlation and a small error per site. The challenge of using an NN is that the NN model is not physics-based, and it may report non-physical results (like small negative dishing) for some patterns that are not expected for the given deposited processes.

In the modeling presented here, simple correction was done, but for the future, a special activation function must be used. Using NNs for the modeling of geometry data change after deposition process simultaneously with heights data is straightforward. The challenge here is to get geometry data from AFM line scans or other measurements for the training dataset. In principle, it may be done, but it will require a more detailed analysis of measurement data. Based on the results of our research, the application of NNs to deposition profile modeling looks promising, and it is under active investigation.

References

1. K. Mistry, C. Allen, et al. "A 45nm logic technology with high-k+metal gate transistors, strained silicon, 9 Cu interconnect layers, 193nm dry patterning, and 100% Pb-free packaging", IEEE International Electron Devices Meeting, Washington D.C., 247, (2007).
2. M. Krishnan, J. W. Nalaskowski and L. M. Cook. "Chemical mechanical planarization: slurry chemistry, materials, and mechanisms", Chem. Rev. 110, 178 (2010).
3. U. Katakamsetty, D. Koli, Y. Sky, H. Colin, R. Ghulghazaryan, B. Aytuna, J. Wilson. "20nm CMP model calibration with optimized metrology data and CMP model applications", Proc. of SPIE Vol. 9427, doi: 10.1117/12.2085728, (2015).
4. U. Katakamsetty, J. J. Chee, et al. "Cutting-edge CMP modeling for front-end-of-line (FEOL) and full stack hotspot detection for advanced technologies", Proc. of SPIE Vol. 10148, 101480Q, (2017).
5. R. Ghulghazaryan, J. Wilson, A. Abouzeid. "FEOL CMP modeling: progress and challenges", Proc. of ICPT 2015, 34, (2015).
6. R. Ghulghazaryan, J. Wilson, N. Takeshita. "CMP model building and hotspot detection by simulation", Proceedings 158th Meeting of Planarization CMP Committee, 55, (2017).
7. Y. LeCun, Y. Bengio, G. Hinton. Deep Learning, Nature 521, 436, (2015).
8. Producer® Eterna™ FCVD™, http://www.appliedmaterials.com/products/producer-eterna-fcvd

Siemens Digital Industries Software

Headquarters
Granite Park One
5800 Granite Parkway
Suite 600
Plano, TX 75024
USA
+1 972 987 3000

Americas
Granite Park One
5800 Granite Parkway
Suite 600
Plano, TX 75024
USA
+1 314 264 8499

Europe
Stephenson House
Sir William Siemens Square
Frimley, Camberley
Surrey, GU16 8QD
+44 (0) 1276 413200

Asia-Pacific
Unit 901-902, 9/F
Tower B, Manulife Financial Centre
223-231 Wai Yip Street, Kwun Tong
Kowloon, Hong Kong
+852 2230 3333

About Siemens Digital Industries Software

Siemens Digital Industries Software is driving transformation to enable a digital enterprise where engineering, manufacturing and electronics design meet tomorrow. The Xcelerator™ portfolio, the comprehensive and integrated portfolio of software and services from Siemens Digital Industries Software, helps companies of all sizes create and leverage a comprehensive digital twin that provides organizations with new insights, opportunities and levels of automation to drive innovation. For more information on Siemens Digital Industries Software products and services, visit siemens.com/software or follow us on LinkedIn, Twitter, Facebook and Instagram. Siemens Digital Industries Software – Where today meets tomorrow.

83303-C2 11/17 H