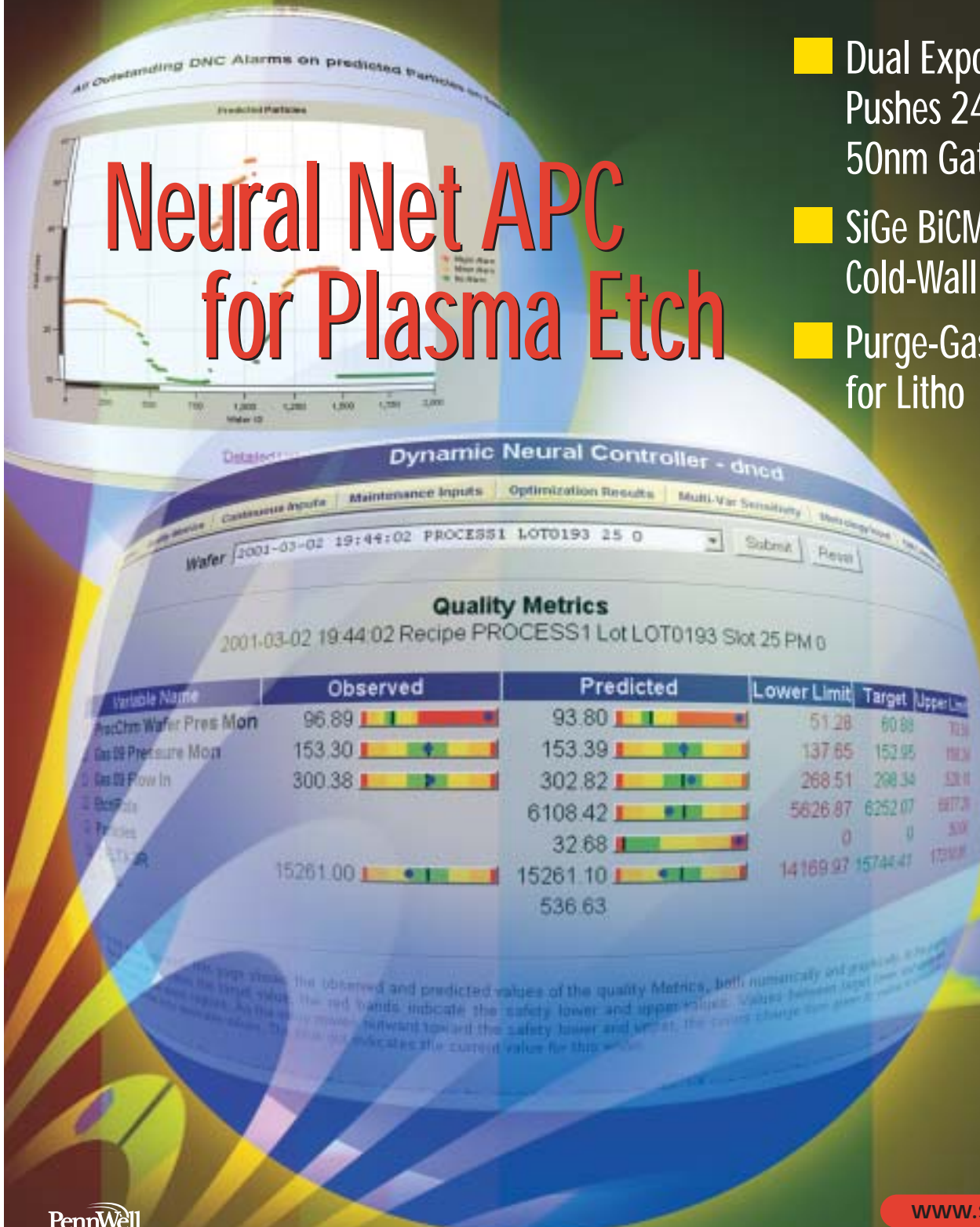


## Neural Net APC for Plasma Etch

- Dual Exposure PSM Pushes 248nm to 50nm Gates
- SiGe BiCMOS with Cold-Wall UHV-CVD
- Purge-Gas Purification for Litho



# Using neural networks for intelligent plasma etch process control

## OVERVIEW

When applied to real plasma etch data, neural network-based process control that ties in maintenance data predicted maintenance requirements nearly as well as fab engineers. The uniqueness of this particular method of advanced process control is its ability to rank process parameter settings by risk, thus allowing engineers to make more intelligent decisions. The net result is a reduction in tool downtime, increased effective capacity, and reduced operating expenses, perhaps amounting to a saving of \$500,000/etch chamber annually.

**A**dvanced process control (APC) involves

- detecting specific problems before they jeopardize wafers (i.e., fault detection),
- identifying excursions from known good tool-states (i.e., tool health monitoring), and
- using tool states and process results to feed back process parameters for subsequent runs (i.e., run-to-run control).

Run-to-run control — the ability to guide a process based on process results rather than just tool settings — has the greatest potential impact on wafer processing productivity.

Plasma etch tools, in particular, have been difficult to control because small variations in equipment set-up may cause unpredicted process changes. Plasma etch reactions vary with device and film types, and the reaction for a given process often changes during the process as underlying areas are exposed to the plasma (the basis for endpoint detection).

Run-to-run controllers, which allow engineers to monitor and modify plasma etch processes to compensate for variations in chamber condition, work by building a model of the process. As each process occurs, the controller compares the results to its model and adjusts parameters for the next run. Fabs using run-to-run control have reported as much as a 40% increase in process capability [1].

A complexity left out of most run-to-run controllers is access to maintenance data. Significant process change caused by chamber cleaning, for example, requires that models from conventional controllers be reset each time major



**Table 1** Partial list of monitored etch-chamber state and maintenance variables

Etch-chamber state variables	Maintenance tasks
RF forward/reverse powers	Major/minor wet cleans
DC bias	Replace MFCs
Chamber pressure	Adjust gap
Wafer area pressure	Electrostatic chuck
Lower/upper electrode temperatures	Confinement ring/screws
He clamp flow (3 steps)	Replace O-ring
He gas flow	Manometer
CF <sub>4</sub> /CHF <sub>3</sub> gas flows	RF matching network
Ar/O <sub>2</sub> /N <sub>2</sub> gas flows	RF generator calibration/replacement
	Electrodes
	Clean load locks/elevator
	Perimeter lifter
	Clamp ring
	Gas baffles

maintenance occurs. Another drawback is that conventional run-to-run controllers create a separate model for each process recipe and cannot be used on a tool running more than one recipe. In addition, recipes with multiple etch steps are difficult to model.

To handle complexities associated with modeling vast amounts of tool, process, and maintenance data generated by advanced process tools, some control experts are turning to neural networking and advanced mathematical methods that “learn as they go.” Such capability is found in the DNC family of advanced neural network controllers and can include the ability to calculate risk involved in a subsequent run, offering the user a

choice of least-risk, least-effort solutions.

## Neural network with etching

Applied to plasma etching, for example, the DNC neural network controller models the process to predict output variables, such as

etch step height and etch rate. Its dynamic database merges information from tool settings, such as gas flows and pressure, and output variables collected from in situ equipment monitors and downstream metrology tools.

The controller's learning algorithm, using cascade-correlation architecture, trains the neural network. The learning begins with random weighting between inputs and outputs. As processes run, the controller begins to weigh different variables until predicted result matches actual result. In the cascade correlation neural model, hidden nodes are added sequentially. The weights to a hidden node are optimized to minimize output error and fixed prior to the addition of another node. New nodes have unexplained error from the preceding node passed to them and are trained to reduce that error. When a successive hidden node fails to reduce overall error, the hidden layer is complete. As new data are obtained, the network continues to retrain itself.

In addition to driving output variables closer to target, an integrated optimizer recommends changes that minimize overall risk and require the least amount of operator or maintenance effort. Each variable value is risk-weighted by its deviation from its target value. After each run, the optimizer lists several options for tool set-up, each of which drives the output closer to target and reduces the overall risk of tool operation. It then ranks options by risk, listing the lowest risk/wafer option first.

This controller can also analyze information from a maintenance database. The database record for any wafer processing tool includes the age of all parts and the time since the last clean or calibration. Since all variables can have safe limits applied by an engineer, the controller displays a flag for any part that is overdue for service or for any setpoint that is above or below the limit.

Because data are an integral part of the model, the model can predict results after any maintenance event that has occurred several times. In particular, the model predicts significant process shift after a chamber clean. As long as maintenance events are recorded, the controller evaluates whether or not to improve the model after each event. If it develops a better model, it implements it immediately, learning throughout its life.

**Table 2 Output variables predicted for the etch tool**

	Product Wafers	Monitor Wafers
Ex situ	Etch step height Particle count	Monitor particles Monitor etch rate Etch rate stan-
dard		deviation
In situ	Wafer area pressure He clamp flow (1) He clamp flow (2)	

With sufficient data, the controller's initial level of control can be upgraded so that it automatically adjusts the model for each maintenance event and includes maintenance activities in its options list for optimum, lowest-risk processing. Since the need for maintenance on one particular tool may not always be the most pressing need in a fab, the controller offers several options, enabling intelligent maintenance scheduling. If the lowest risk solution requires a part change, for example, the tool can be scheduled for the part change later in the shift, and the next-lowest-risk solution implemented in the meantime.

### Real fab data

In a recent study, we used a DNC neural network controller to analyze data on a Lam plasma etch tool in large semiconductor facility in New England, collecting data from 10 product-wafer and two moni-

tor-wafer recipes for 15 months. Although the controller was designed for run-to-run monitoring, this evaluation was a retrospective study in which we entered data and compared system recommendations with actual maintenance records and tool performance metrics.

For product wafers, the controller monitored 60 input set-up parameters and maintenance variables (Table 1) and predicted six output variables. For monitor wafers, the controller used 35 input variables and predicted three output variables. The analyzed data were used to predict ex situ and in situ output variables for the following run (Table 2). To train the neural network, we entered data from

**Table 3 DNC recommendations for maintenance actions\***

	Corresponded with actual maintenance (events)	No recomm., but mainten. done (events)	Recomm. before actual maintenance done (avg. days)
Adjust gap	5	4	18
Clean load locks/elevator	1	1	11
Confine ring	1	2	4
Confine screws	3	0	5.7
Electrostatic chuck	0	1**	
Gap repair	1	0	0
Hot edge ring	1	0	4
Major clean	16	2	5.1
Manometer	1	0	5
Minor clean	14	3	3.1
Network match	1	0	4
O-rings	2	0	10.5
Perimeter lifter	1	0	14
PM	1	0	1
Repair CF <sub>3</sub> leak	1	0	35
Replace CF <sub>4</sub> MFC	1	0	88
RF cable	1	0	10
Temperature probe	2	0	13
Total events	53	14	

\*corresponded to actual replacement of parts 80% of the time

\*\*catastrophic failure

the first 50% of the runs into the controller. We used the remaining 50% to test and validate the accuracy of the trained neural network.

### Fab results

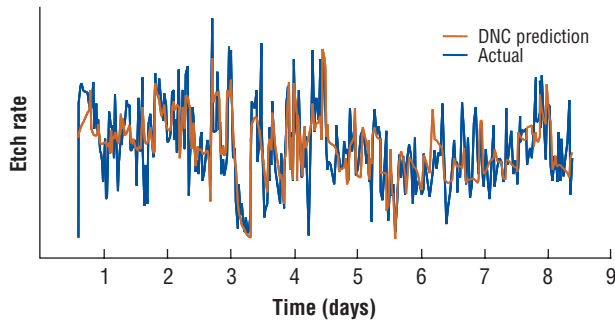
In our tests, the controller's optimizer recommended maintenance 53 times. When we compared the optimizer's suggestions for parts replacement, cleaning, and adjustments with actual maintenance records, we found that maintenance had always occurred, but frequently much later than when the optimizer would have flagged the need (Table 3). Our data also revealed that maintenance occurred 14 times when it was not recommended by the optimizer. We assume that the difference between the opti-

mizer's suggestions and actual maintenance actions were attributable to the following causes:

- The action recommended by the optimizer may not have been necessary.
- The maintenance records used to train the network were not initially as accurate as typically needed for the neural model.
- The optimizer suggested only actions that both increased quality outcome and decreased overall risks; actions taken by the maintenance crew may not have been the most risk-effective at the time taken, although they improved quality.
- The optimizer may have combined multiple maintenance events into a single event.

Our data also show that the electrostatic chuck failed catastrophically, but this was not predicted by the optimizer. This was not unexpected, because the optimizer is not designed to predict a catastrophic failure. We expect, however, that with longer use a neural network will recognize subtle patterns indicative of failure modes, and catastrophic failures should become less frequent.

Even within the limited scope of this evaluation, though, the optimizer's success rate was nearly 80% in predicting recorded maintenance events. More significant, optimizer recommendations occurred days or even weeks before actual maintenance occurred. Optimizer recommendations substituted major cleans and minor



DNC predictions for etch rate compared to actual results.

cleans with specific parts replacements, avoiding repeat clean operations and providing savings on parts required. In addition to the optimizer's ability to identify parts needing replacement, the neural network demonstrated its ability to predict the outcome of an etch process with specific settings and with a known age on tracked parts.

We were able to test the accuracy of the neural network by looking at its predictions across multiple recipes and different families, comparing them with the actual results. Results were excellent for all output parameters. The figure shows a typical example, with etch rate data from 358 wafers; prediction was within 98%, on average, of the actual results. The etch rate model was computed using only maintenance events to predict etch rate. The eventual addition of trace variables in the input vector will only enhance the fit of the neural model etch rate predictions. Other output variables gave similarly good results (Table 4).

In this particular test, optimizing maintenance actions and recipe parameter setpoints primarily yielded maintenance recommendations. The controller suggested few recipe changes, chiefly gas

Table 4 Output variable predictions\*

	Product wafers		Monitor Wafers	
Ex situ	Etch step height	98.08%	Monitor particles	72%
	Particle count	72%	Monitor etch rate	98.43%
In situ	Wafer area pressure	98.97%	Etch rate std. dev.	76.50%
	He clamp flow (1)	98.73%		
	He clamp flow (2)	98.75%		
	He clamp flow (3)	98.80%		

\*most were within 95% of actual results as defined by  $(\text{observed} - \text{predicted})/\text{observed} \times 100$

flows, etch duration, and chamber pressure adjustment.

The optimized maintenance recommendations would yield cost savings primarily attributable to reduction of maintenance activities. Timely replacement of parts reduces parts consumption and both scheduled and unscheduled downtime. For each run, the controller assigns a risk coefficient, weighing the suggested activities based on run results. Using these risk coefficients, we estimate that average risk savings/optimizer intervention would have been 22.84% (0.51% to 51.93%).

### Conclusion

Preliminary results with a neural network controller show that it predicted maintenance requirements for a plasma etch tool nearly as well as fab engineers. The controller also predicted the outcome of 10 product recipes using a single model. With its ability to analyze drift, rather than excursions outside preset limits, the controller can identify equipment failures before they cause scrap. Ranking of parameter settings by risk allows engineers to make intelligent decisions about process and maintenance changes.

By reducing tool downtime, the controller increases the effective capacity of the tool. We expect that implementation of this capability could increase capacity significantly, save operating expenses, and increase throughput, perhaps amounting to saving \$500,000/etch chamber annually. ■

### Acknowledgments

We thank Frank Hoppensteadt at the Center for Systems Science and Engineering, Arizona State University (ASU) for providing information on the fundamentals of neural networking, and Jennie Si at ASU's Department of Electrical Engineering and Deana Delp for providing information on neural network applications in the semiconductor manufacturing industry.

DNC is a trademark of IBEX Process Technology.

### Reference

1. A. Toprac, et al., *Proceedings of AEC/APC Symposium XII*, Sept. 2000, p. 128.

**JILL CARD** received her BS in biology natural resources from Cornell University and her MS in theoretical and applied statistics from Florida State University. She is founder, chair, and chief scientist at *IBEX Process Technology, 40 Church St., Suite 301, Lowell, MA; ph 978/453-3988, fax 978/452-7515, jcard@ibexprocess.com.*

**LISE LAURIN** received her BS in physics from Yale University. She is director of product marketing at IBEX Process Technology.