How AI and ML can save $38bn for semiconductor manufacturers

Atonarp's CEO Prakash Murthy explains how artificial intelligence and machine learning can be used for semiconductor manufacturing equipment and process co-optimization.

Today, increasing throughput is the number-one priority for semiconductor fabs, as they work to overcome the challenges of the global chip shortage.

Looking beyond throughput, there are significant opportunities for long-term cost savings from optimizing, simplifying or removing processing steps. We call this approach EPCO – Equipment and Process Co-Optimization. It is a combination of good engineering and applying data-driven machine learning (ML) to the manufacturing process and equipment.

A 2021 paper by McKinsey argued that semiconductor manufacturing optimization, using artificial intelligence (AI) and machine learning (ML), could save $38bn, through improved yields and increased throughput.

Real-time, accurate and actionable data is vital to achieving this potential. McKinsey highlighted that the single most important point to address is the real-time, run-to-run adjustment of tool parameters, using live in-situ tool sensor data. This enables AI/ML algorithms to optimize the nonlinear relationship between process operations.

The problem: increasing process complexity

Today's high-volume, advanced logic processes — including Fin-FET and gate-all-around (GAA) transistors, as well as high-aspect-ratio etch techniques used in 3D-NAND memories — require a new approach to the established standards based on Intel's CopySmartly! methodology.

As process nodes have shrunk, new variables have emerged that affect process yield, and can cause deviations even on the exact same equipment. In Figure 1, shared in a study of machine learning for high-volume manufacturing metrology challenges, chamber-based effects on process critical dimensions (CD) can be clearly seen.

Some of these critical variables that can affect process performance include localized virtual vacuum leaks, subtle reaction gas partial pressure variations, wafer surface saturation due to changes in pumping performance, surface reactivity due to changing wafer temperature, chamber clean end-point, and chamber seasoning profile.

Additional challenges — inter-layer adhesion, 300mm wafer mechanical stresses, new atomic-level deposition and etch chemistries, exotic low-resistance contact and fill metals, stringent cross-contamination protocols, and maximizing throughput — all require greater insight into how the process and equipment are interacting. Optimizing advanced processes such as these now demand higher-accuracy metrology tools and add a new layer of in-situ molecular complexity.

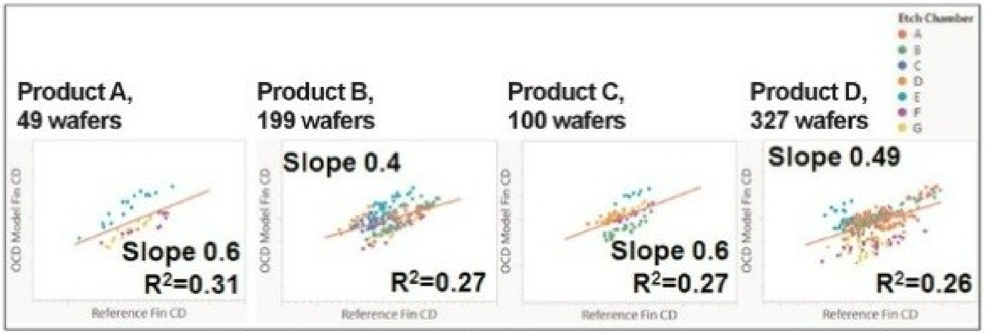

Figure 1. Wafers of four different products from eight process chambers were evaluated above. Each chamber is represented by a different colored dot. Based on the colored dot clustering, some chambers demonstrate significant variation in CD values between wafers, clearly showing a chamber-related effect on the unit process.

The solution

We can improve semiconductor metrology in two ways: either by capturing better data with more sensitive metrology tools or by extracting more value from existing data with new ML algorithms. Of course, if we can do both, we may well see the biggest improvements.

Figure 2. As process geometries shrink, the effects of equipment variation on process margins drive the need for equipment process co-optimization within the EPCO Zone.

Either way, for successful AI/ML deployment, it's vital to have truly actionable real-time data. This enables appropriate models to be created and tested with data correlation between real-world and ML model inputs and outputs.

For example, statistical process controls can look at the real effects of chamber-to-chamber, machine-to-machine and run-to-run performance variances, even on the exact same equipment with the same recipe. Chamber cleaning and seasoning have material effects on chamber performance and drift in process results (process margins) between cleans, and PM (performance management) cycles are common. The difference is that, at mature nodes like 40nm, the differences run-to-run are small compared with the process control limits. However, as process geometries shrink, so do process control margins and the chamber and equipment effects (sigma variation) become increasingly critical (see Figure 2).

Process control has become a lot more complicated as critical dimensions have shrunk, along with the margin for error. This means that individual chamber management is becoming fundamental to ensuring high line-yield, with tight statistical process control.

This is what EPCO is all about: leveraging ML to jointly optimize equipment, chambers and processes in unison.

In-situ, real-time data

There are three main types of data in the semiconductor process control environment:

1. in-situ data taken real-time on the process tool;
2. in-line data to measure results (usually immediately) after a processing step;
3. parametric or post-fab data (used for wafer line-yield and wafer ship acceptance criteria).

One of the fundamental changes needed to optimize fab management is the switch from in-line to in-situ metrology. Measurements taken after processing is completed are sequential in nature, costing throughput and cycle time, and lack the immediacy to affect meaningful real-time process change and optimization.

Measuring in-situ, real-time data at the molecular level gives true insight to how the process is set up and proceeding, offering rich, actionable and impactful data. Reactants, by-products and partial products can be identified and quantified, allowing for dynamic process control to ensure tight mean and standard deviation control for a given process module across run-to-run, chamber-to-chamber, tool-to-tool and even site-to-site.

Managing overall complex semiconductor process control and line-yield starts with having tight control on individual process steps and ensuring low variability and precise statistical process control (SPC). In-situ data from processing chambers can be used with machine learning to improve linearity and accuracy, and to achieve the control required.

Molecular sensor

Atonarp has spent a lot of time understanding the fab and equipment manufacturers' problems and challenges. The result of those efforts is the Aston, a robust molecular sensor.

Aston provides the accurate, actionable, real-time data that's critical for effective EPCO. This data enables