

Overcoming APC Obstacles

How IBEX's DNC answers the tough challenges of APC in the real world



Overcoming APC Obstacles

- The world APC has to live in
- ROI
- What's next

In an ideal APC world:

- Metrology delays would have no effect on control capability, therefore no effect on yield
- Control could still be done with missing data
- Multiple recipes, steps, products, and layers could all be handled with the same model
- The results of maintenance actions would be understood and accounted for in the model



Oh, and one more thing . . .

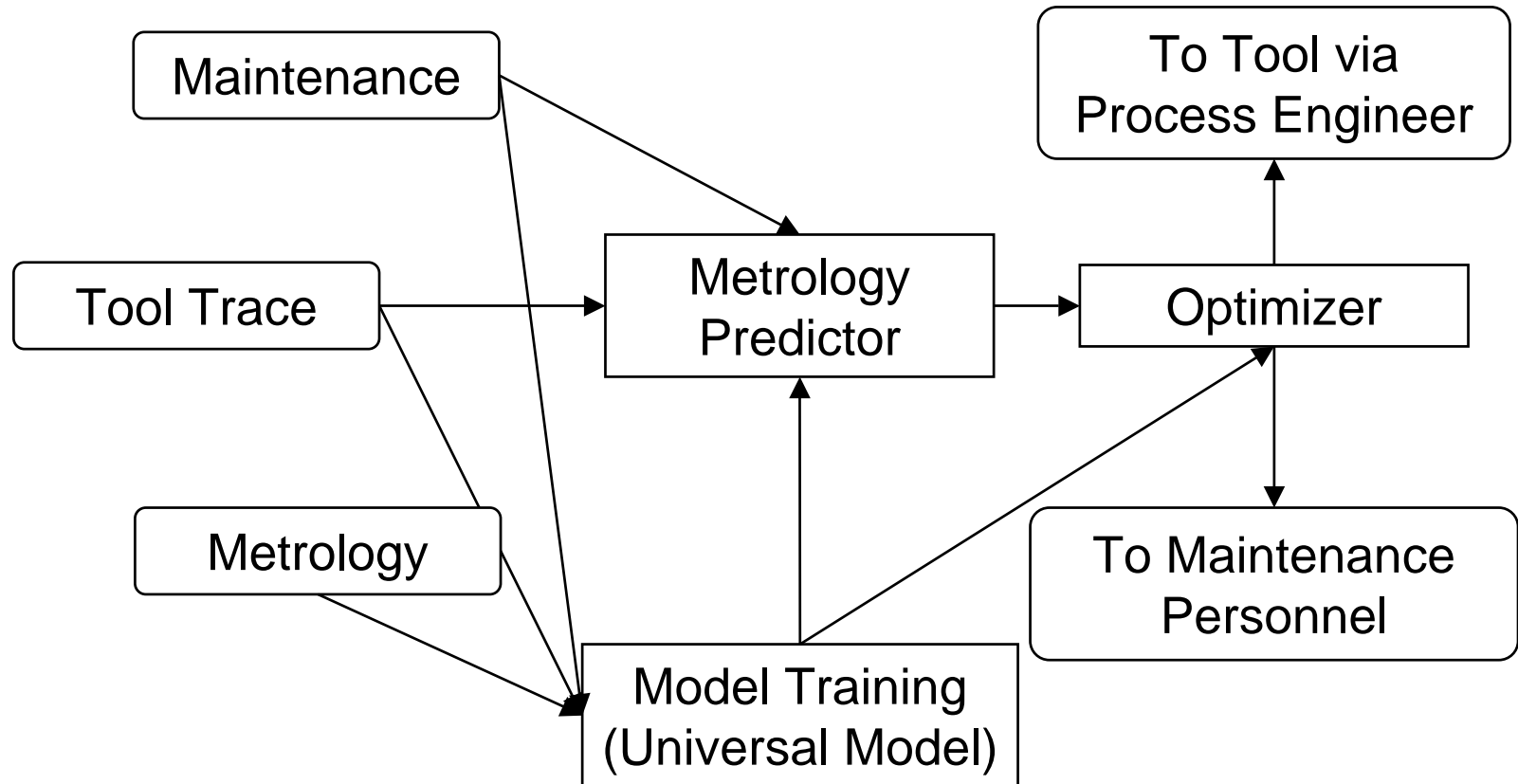
- If metrology wasn't measured, it would be predicted, so we wouldn't have to come up with complex sampling schemes to maximize control and throughput

So why not?

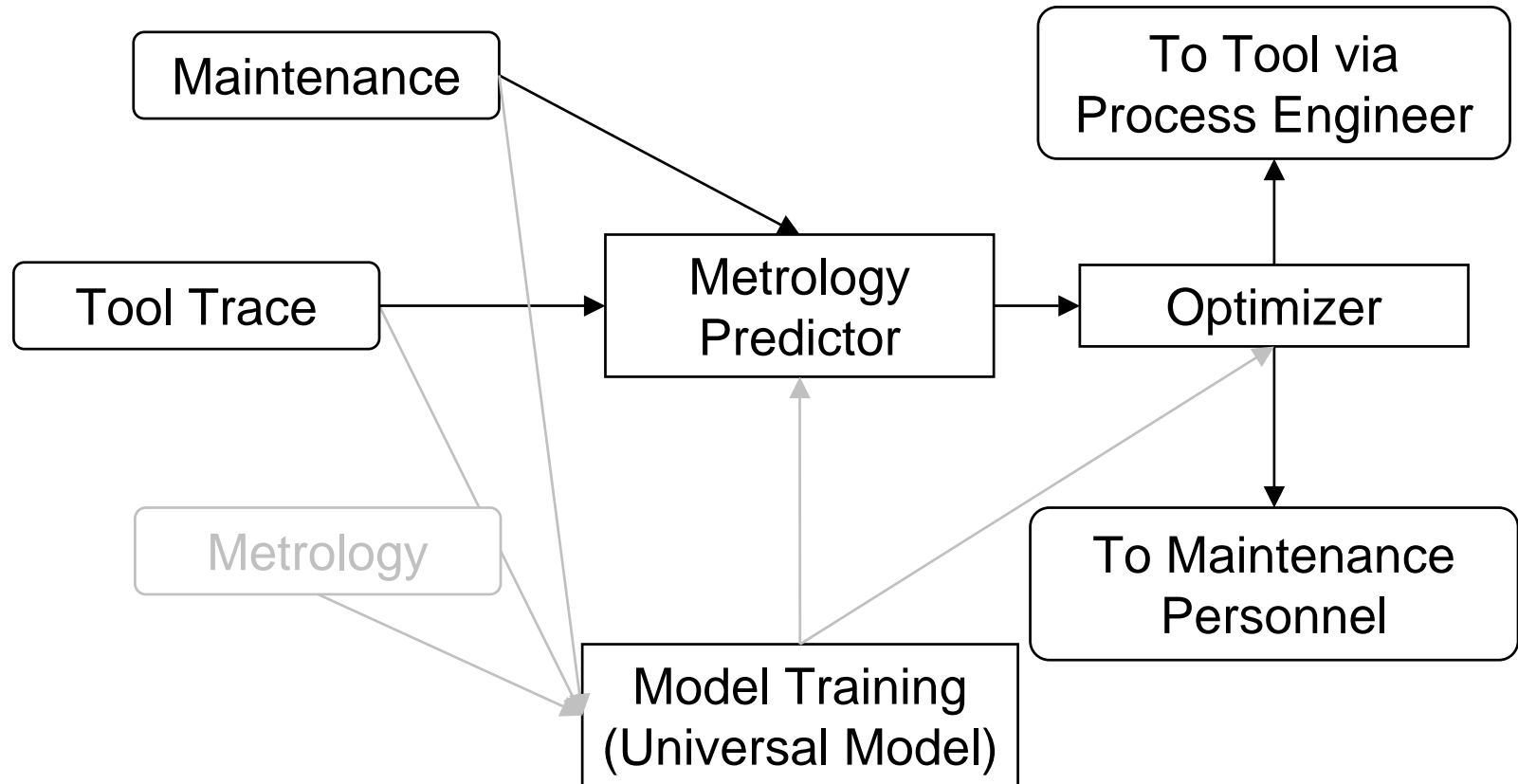
- Predictive Metrology™ handles most metrology data issues
- Complex non-linear modeling allows multiple recipes, multiple steps, in one model
- Adding maintenance data enables accurate predictions, creates a robust controller, and enables maintenance recommendations to be made



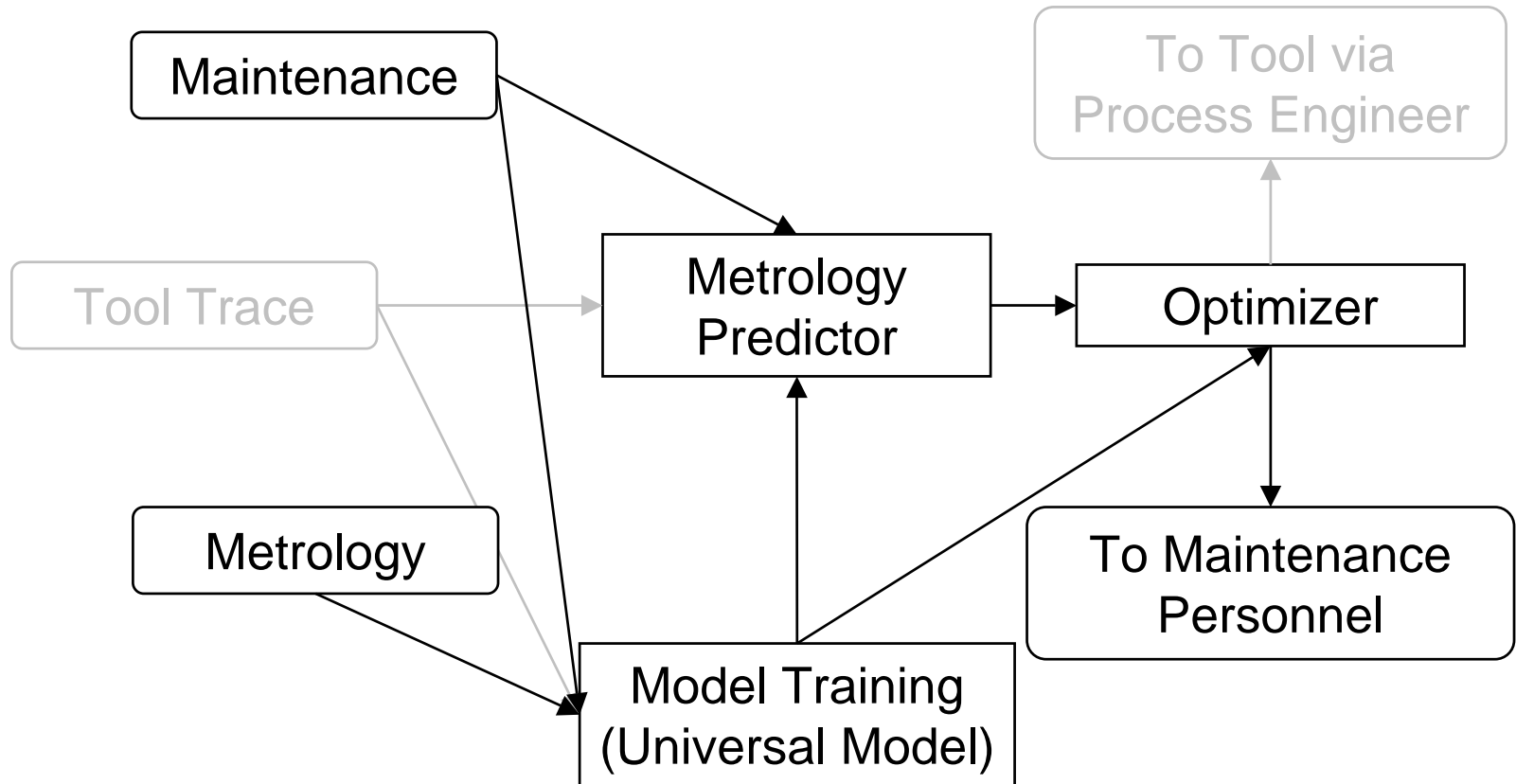
How does this happen?



How it works with missing metrology



How it works with missing trace data

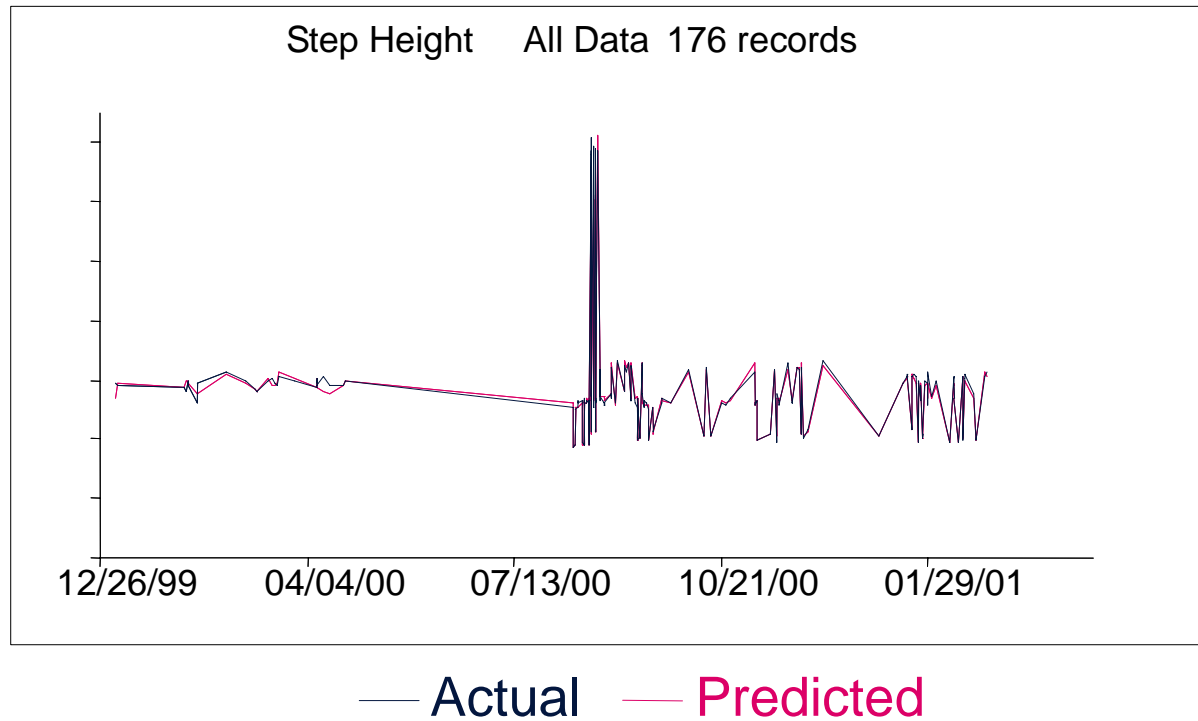


What is this “Universal Model”?

- Neural Networks detect patterns in the data to create an “intuited” model of the physics in the chamber
- With enough data, the neural network can incorporate any process that can be run in the chamber into the model—just like an engineer is able to see patterns across recipes



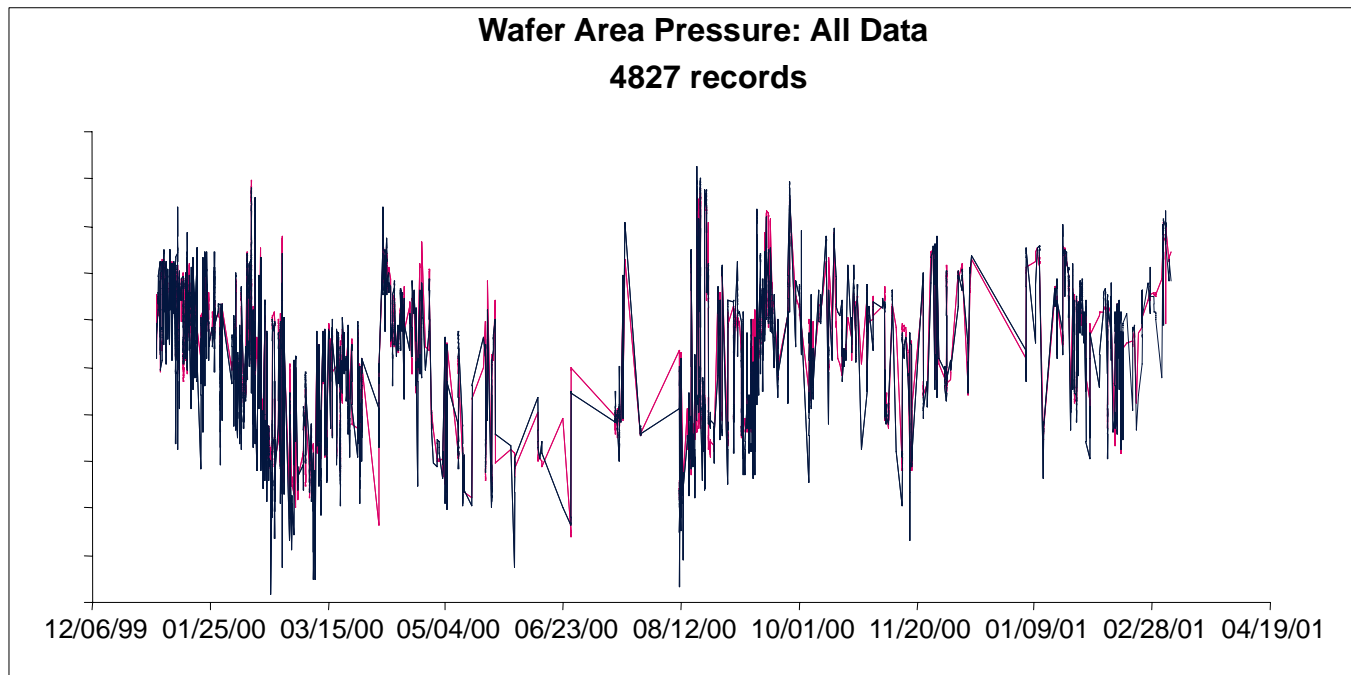
But you're from the "Show me" State . . .



Etch step height network results for 15 recipes shows $< 2\%$ error across all recipes.



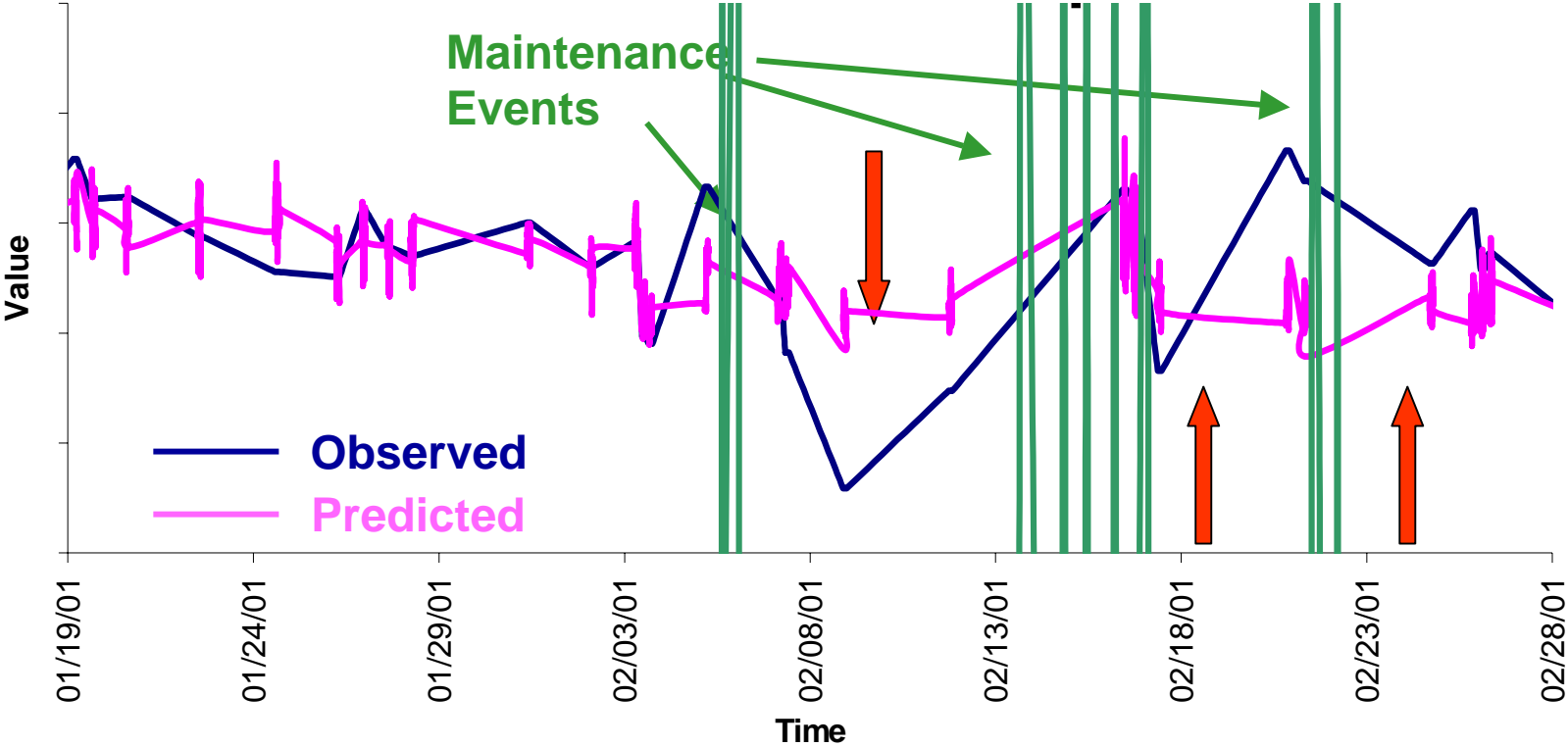
Wafer area pressure predictions for 15 recipes show 1% error across all recipes.



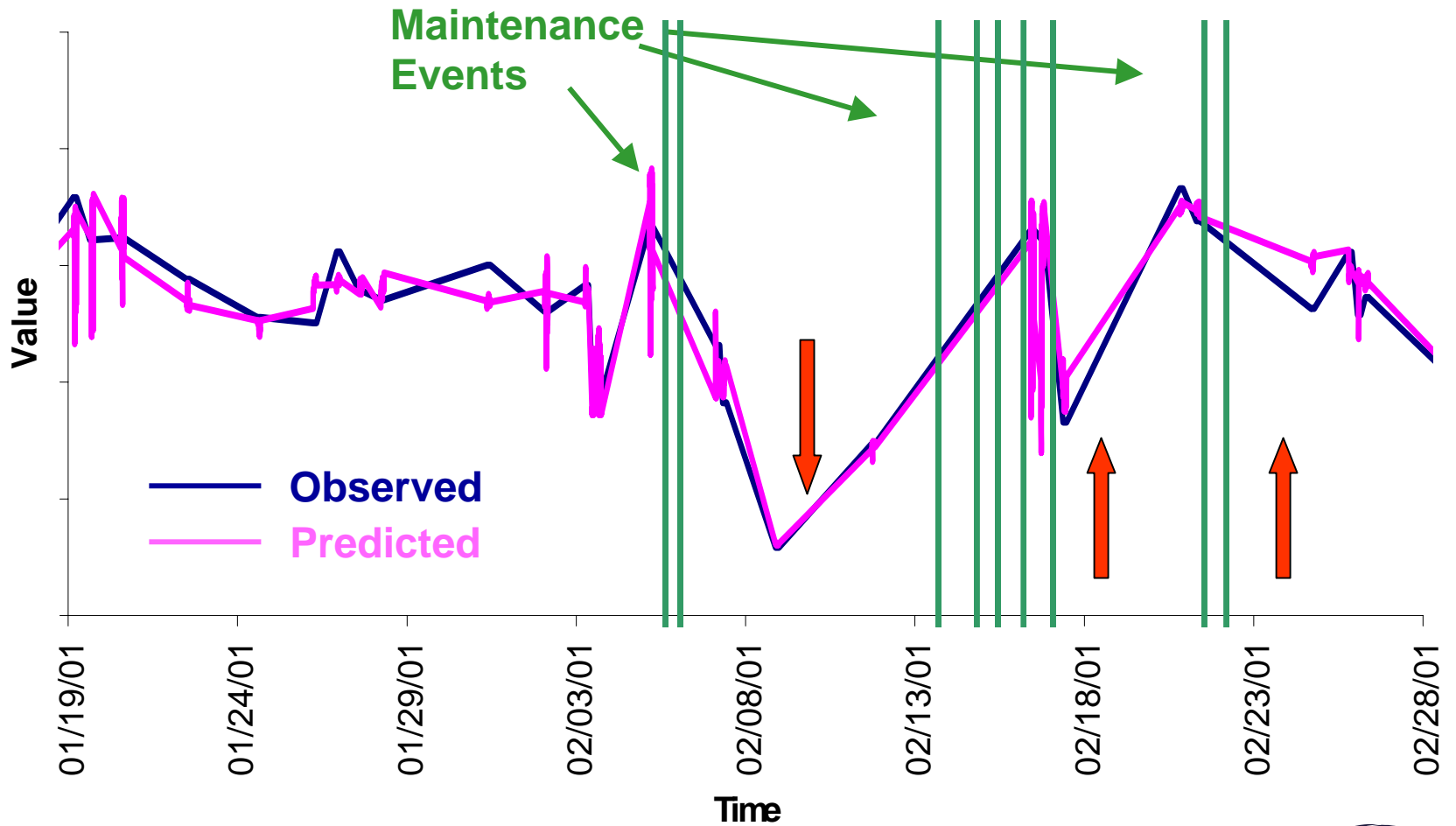
— Actual — Predicted



Thickness Difference without maintenance inputs

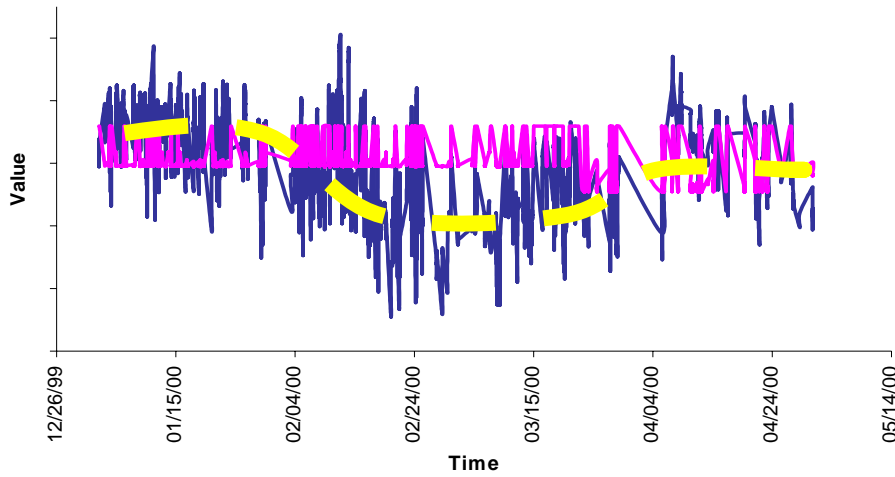


Thickness Difference with Maintenance Inputs

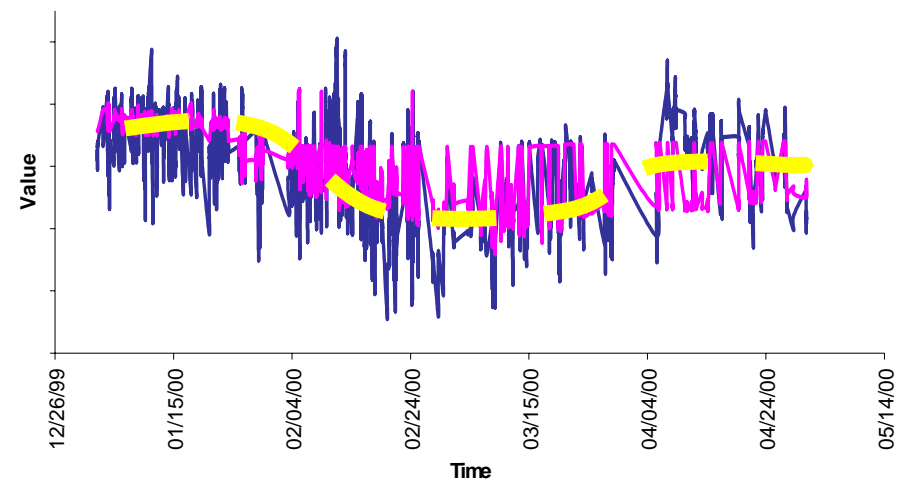


Long-term drift cannot be modeled by tool trace data alone

WAP 2 without maintenance inputs



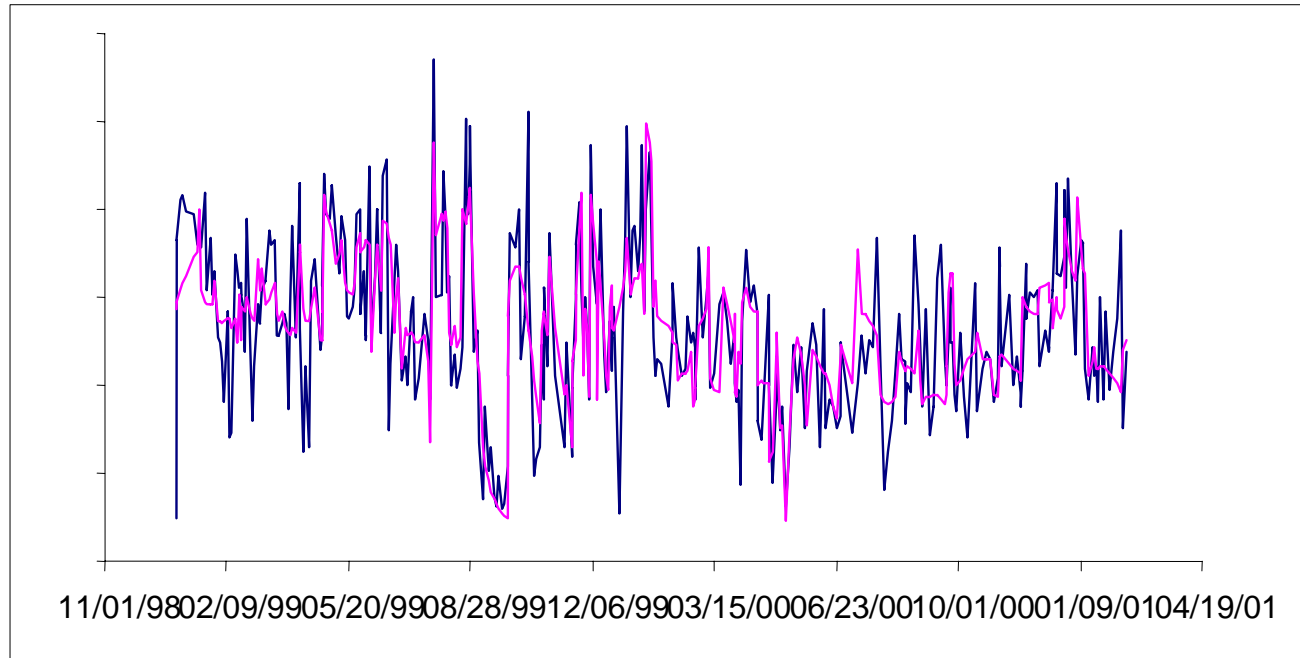
WAP 2 with maintenance inputs



— Observed

— Predicted

But etch rate can be predicted using only maintenance data



Average deviation of predicted value from observed value: 1.6%

— Observed

— Predicted

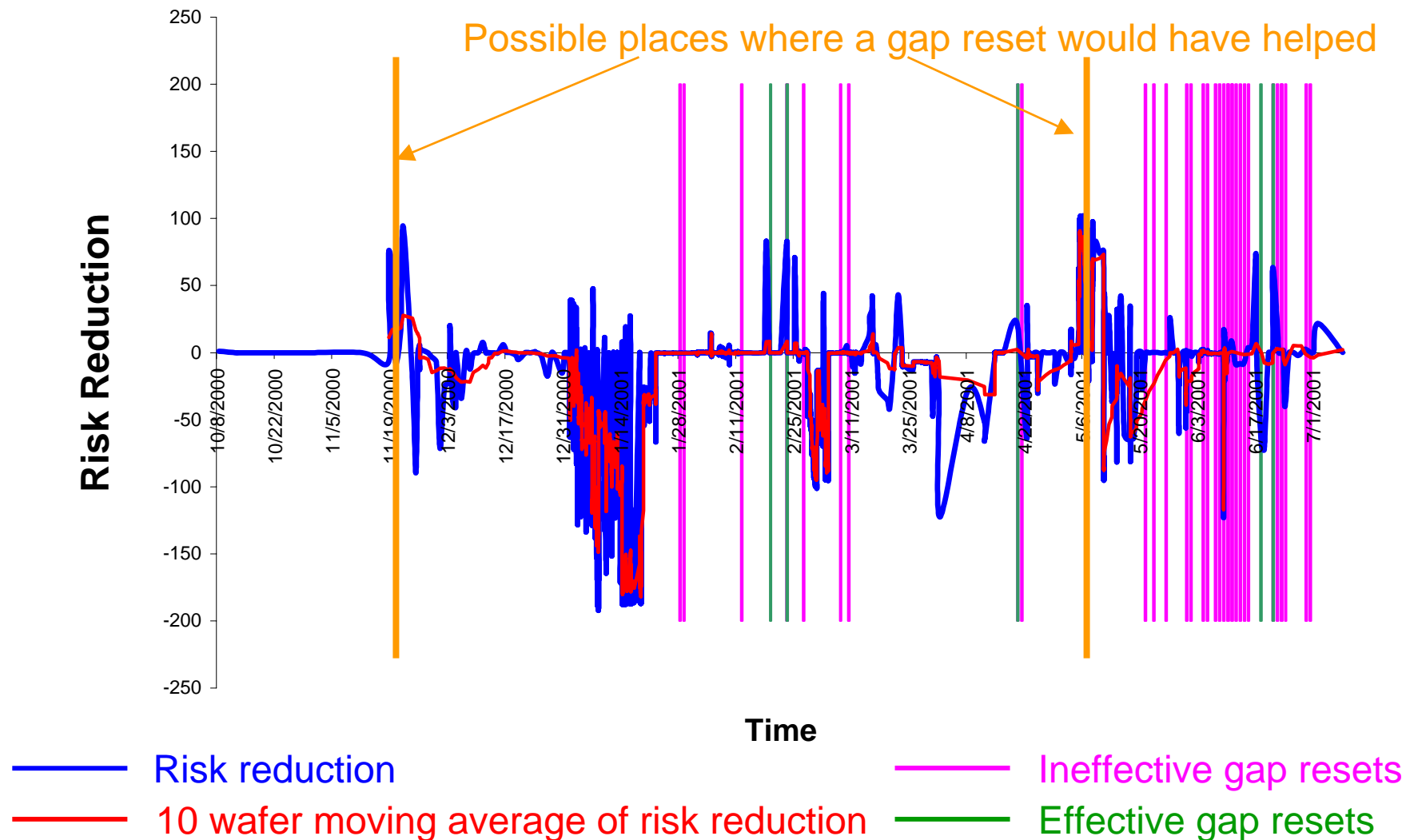


Justify it . . .

- Effort reduction on a Lam Etcher
 - Just eliminating wasted repeat maintenance actions would reduce the number of actions by 23%
 - For major cleans and gap resetting, another 40% - 50% of the actions appear unnecessary, using time based analysis of risk

Of a total of 41 gap resets, 23 (58%) were repeat actions. Repeat actions are actions that occur within two days of an identical action. None of the repeat resets were effective.

Reset-Gap



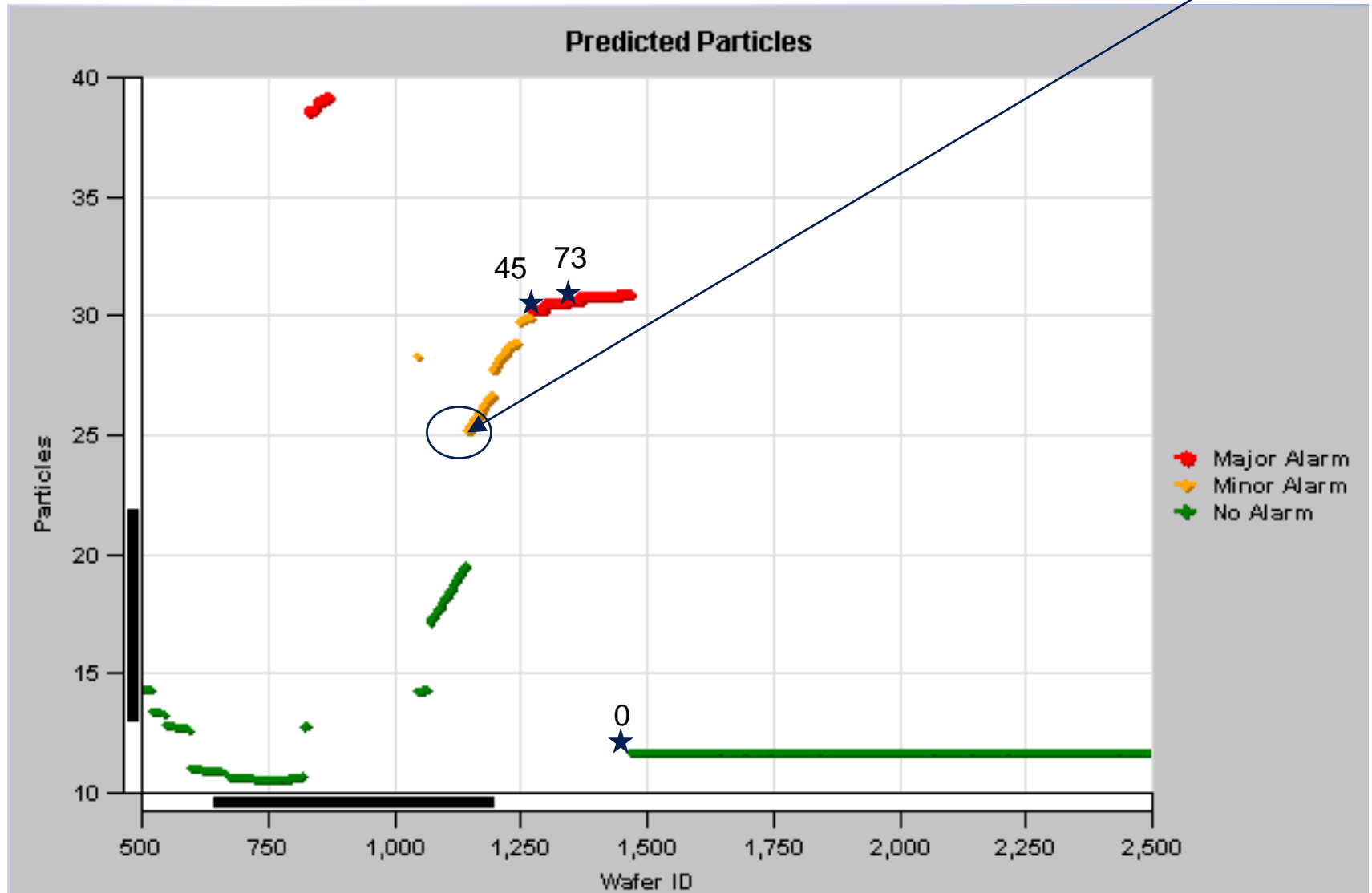
Cost Savings (Effective vs Ineffective Actions)

Repair Actions*	Major_clean	Reset_Gap
Effective Actions (performed and recommended)	4	5
Ineffective Actions (performed, not recommended)	3	13
Missed Actions (recommended, not performed)	4	2
% Missed Actions Missed/(Effective+Missed)	50%	29%
% Wasted Actions Ineffective/(Effective+Ineffective)	43%	58%

*These are in addition to unnecessary repeated actions



Know a day (100 wafers) early that your particles are going out of spec



Fix the problem before it reduces yield

- Early warning of a problem (100 wafers in the low level alarm area before particles go completely out of spec)
- Within 50 wafers (2 lots), the Neural Network has enough of a pattern to recommend a conditioning run
- 200 out of spec wafers could have been prevented. How much yield is lost when particles are too high?



So what's next?

- Closing the loop with integrated measurement
 - Getting information out of masses of numbers
 - Using that information to improve the model
- Integrating wafer to wafer control with fault detection and control

Closing the loop with integrated measurement

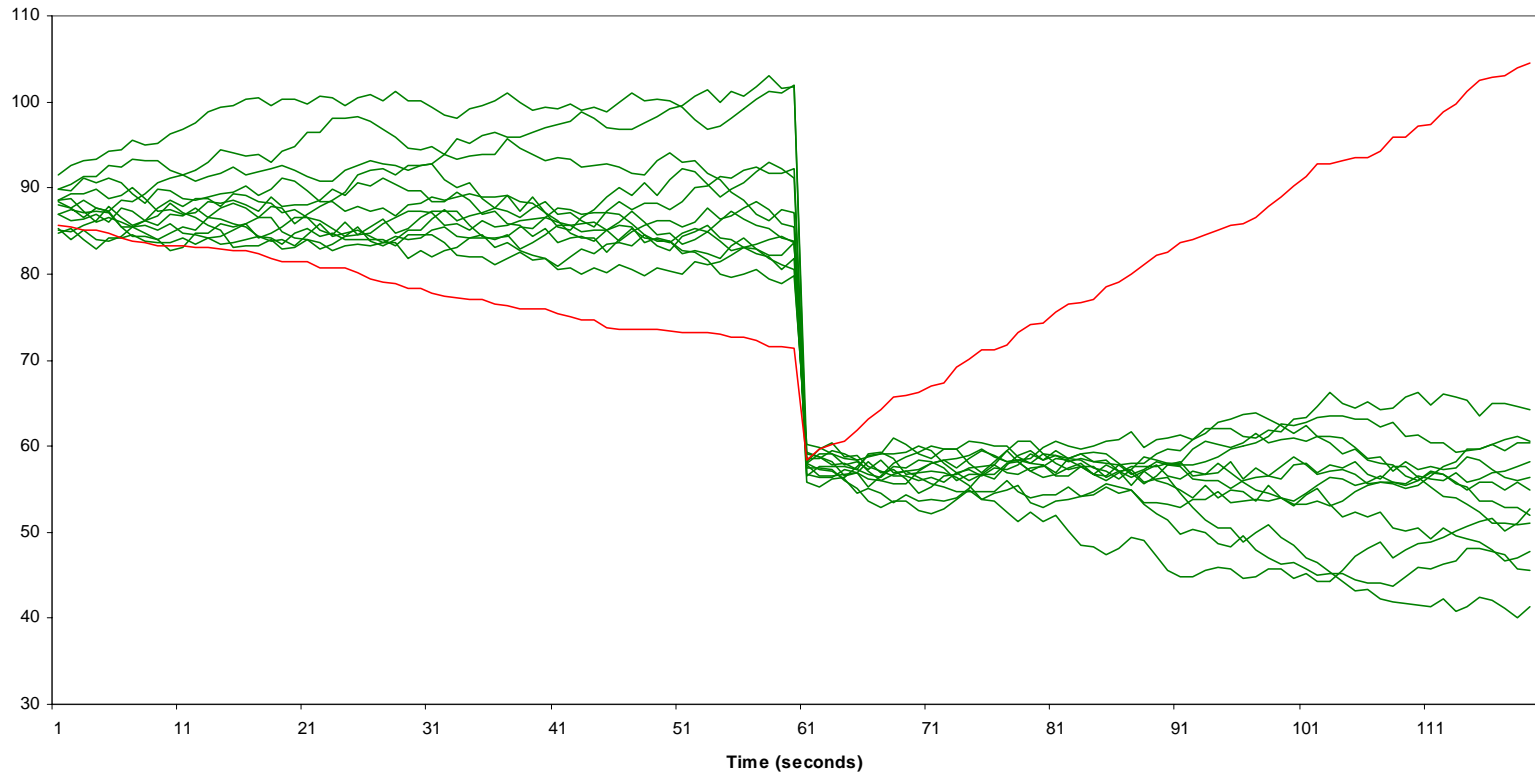
- Can OES data tell us more than end-point?
- Can an online ellipsometer tell us more about what's going on than thickness/refractive index?
- Can we identify a robotic misalignment by the particle pattern on the wafer?

Use neural networks to adaptively extract data

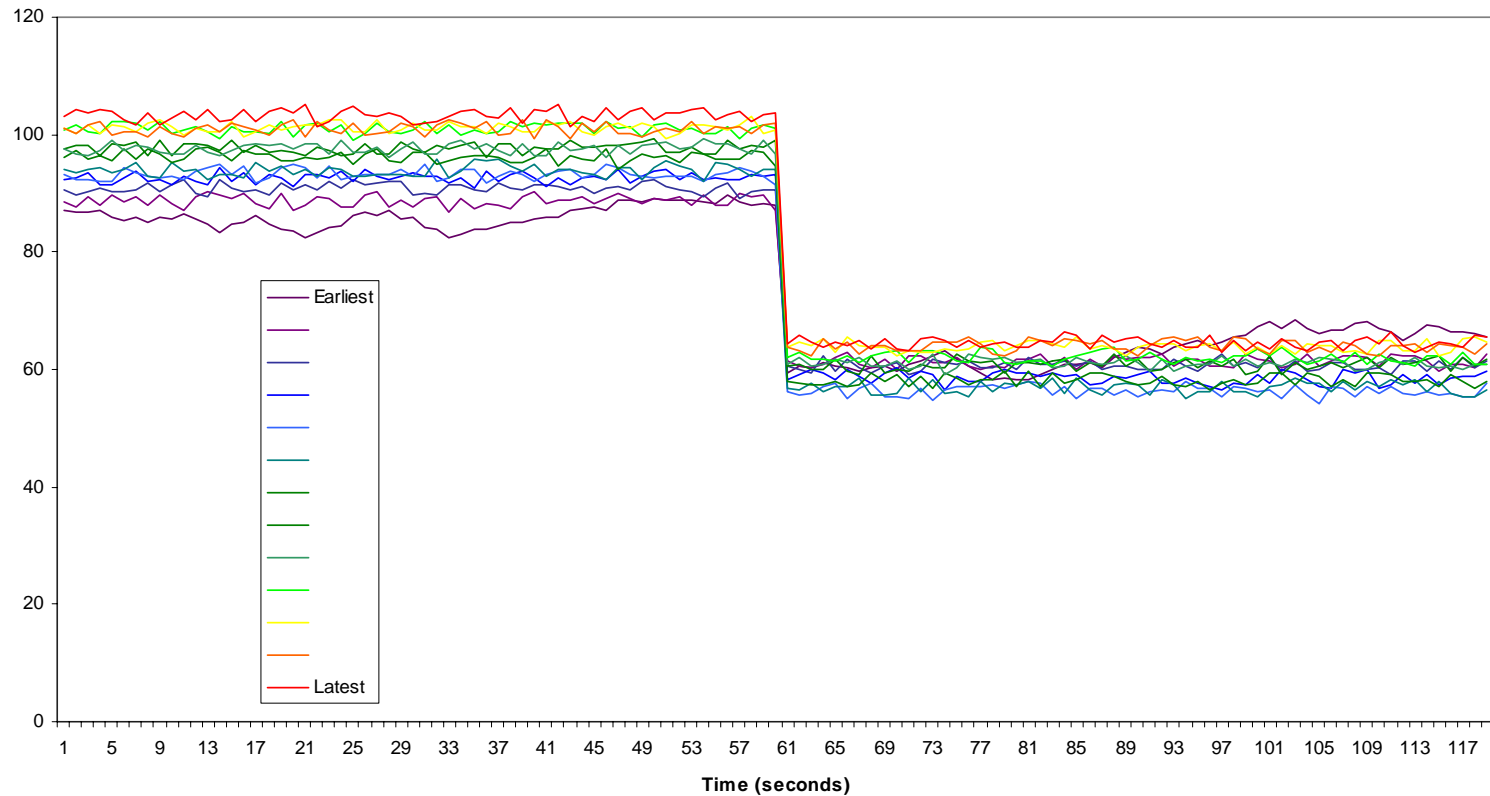
- Generic tool for any complex sensor or metrology tool
- Adaptive to changes in the data over time
- Reduces 1 or 2D images, wafer maps, or multivariate traces to critical **information**.
- Can feed into any analytical prediction, fault detection, or APC application.



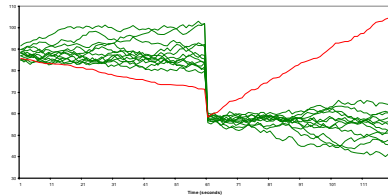
Illustrative sensor data showing fault



Illustrative sensor data showing drift



Extractor as fault classifier



Sensor Trace

Adaptive Critical Feature Information

1 54 7 32 5 7 3 4 2000

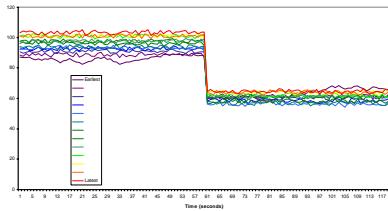
Maintenance

Tool Trace

Predictive Map

Fault Classification

Extractor as maintenance identifier



Sensor Trace

Adaptive Critical Feature Information

1 54 7 32 5 7 3 4 2000

Maintenance

Tool Trace

Predictive Map (DNC)

Drift identification and solution

Then combine them . . .

- Fault detection and wafer to wafer control must work together
- Fault detection identifies and helps correct catastrophic failures not identifiable by wafer to wafer control (wrong or unremoved photoresist, arcing, etc.)
- Wafer to wafer control identifies and corrects for drift



Taking on the tough challenges

- The DNC—APC for the real world
- ROI through Predictive Metrology and reduction of effort
- Looking to the future when we can get more out of integrated metrology and combine FDC with W2W control

