Practical Approaches to Photolithography Run-to-Run Control in Leading Edge High-Mix Semiconductor Manufacturing

Yulei Sun, Ph.D.

Rudolph Technologies, Inc.

RUDOLPH TECHNOLOGIES

Rudolph Software Solutions

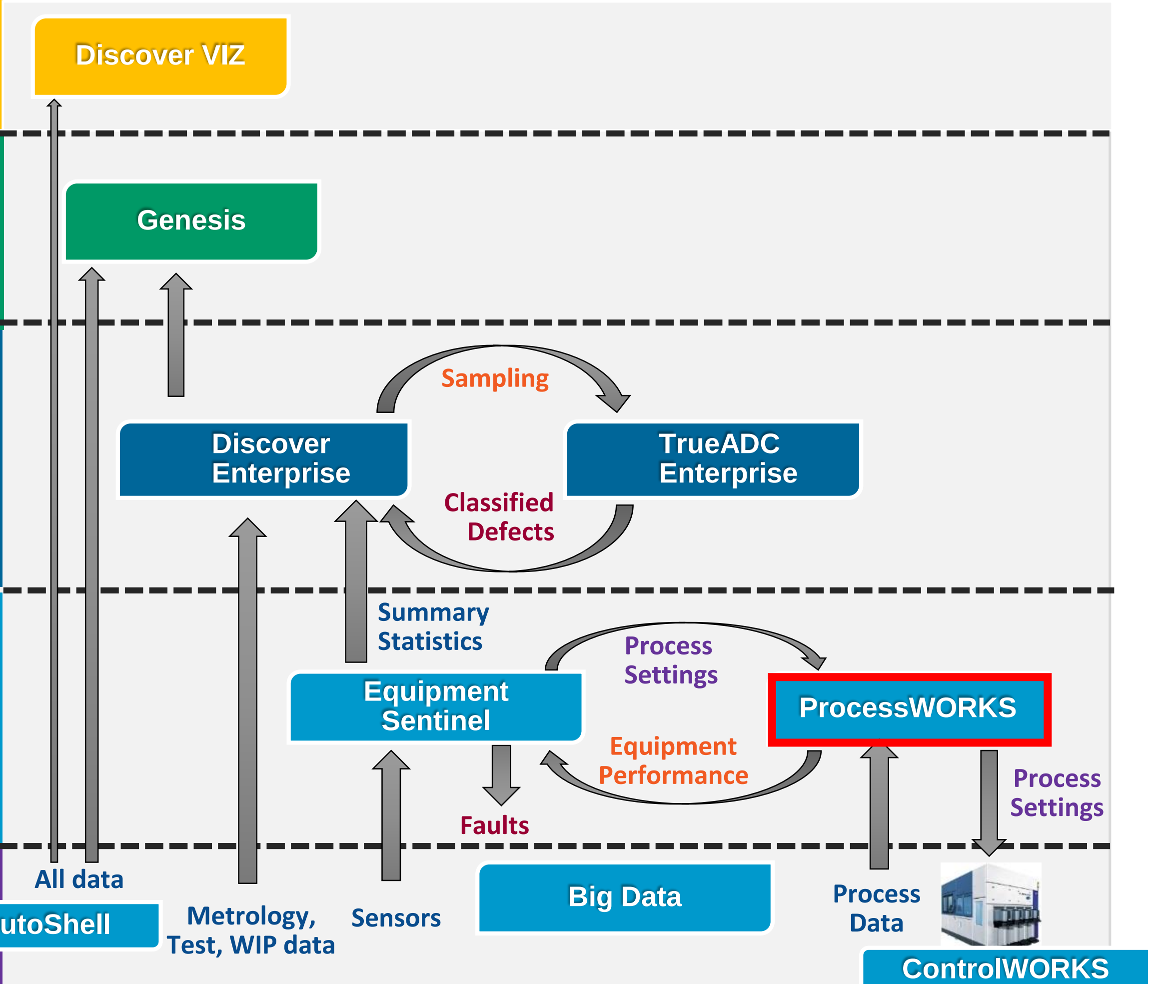
VISUALIZATION
- Interactive drill down
- Configurable dashboards

MODELING
- Use traditional statistics and MVA techniques to enhance process models
- Automate yield learning

MONITORING
- Focus on yield limiting defects to decrease yield production times
- Use process data to improve analysis and tighten limits

CONTROL
- Use R2R and FDC to control and monitor process
- Prevent model adjustments on runs with faults

COLLECT
- Gather wafer & tool data at every step



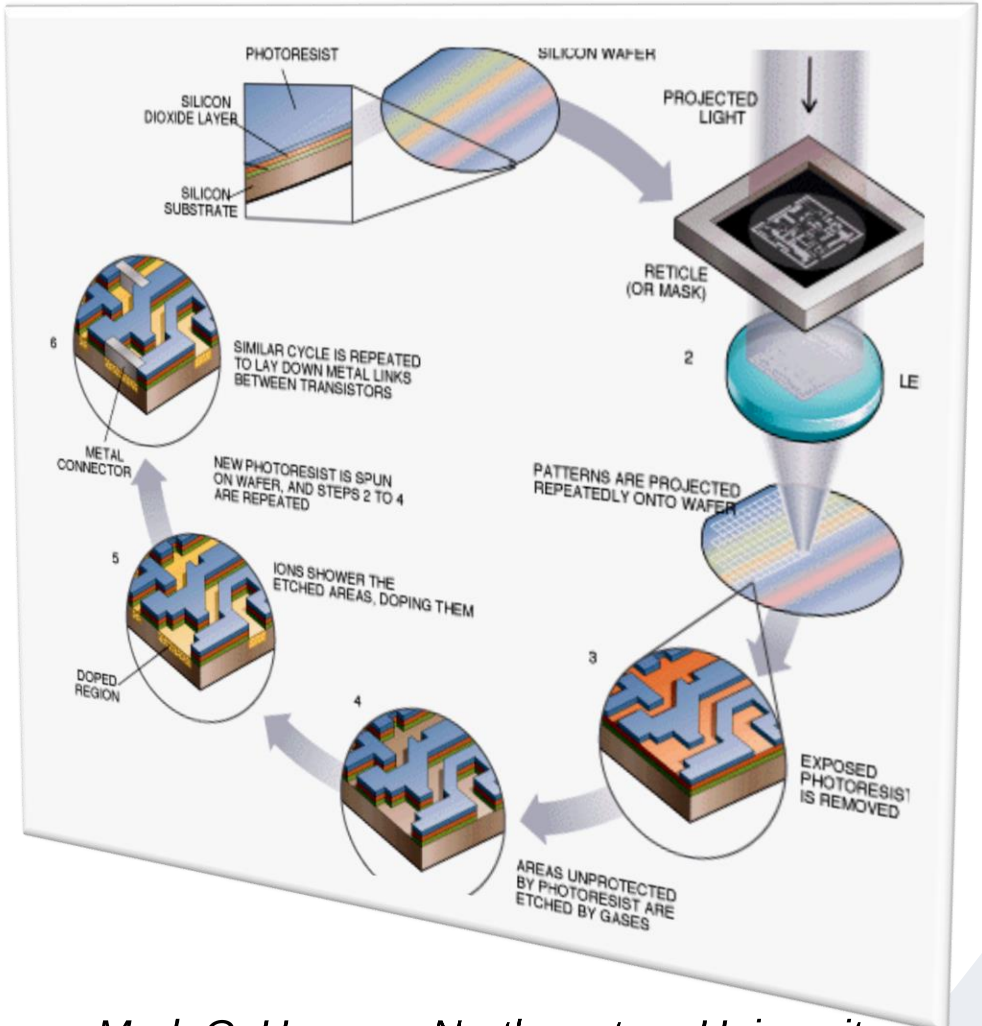
	Defect Analysis	Yield Analysis	Auto Defect Classification	Adv. Process Control	Fault Detection	Factory Automation	Equipment Control
Rudolph	✓	✓	✓	✓	✓	✓	✓

Outline

- Lithography R2R overview
  - High-mix challenges and solutions
- CD control
  - CD control with adaptive slope
- Overlay control
  - High order overlay control
  - Dynamic CPE Control
- Conclusion

Photolithography in Semiconductor Manufacturing

- Semiconductor manufacturing is the most sophisticated and unforgiving volume production technology
  - Consists of complex series of unit process steps
- Photolithography is the most critical step
  - Transfers designed patterns onto silicon wafers
  - Determines minimum feature size and ultimately performance of ICs

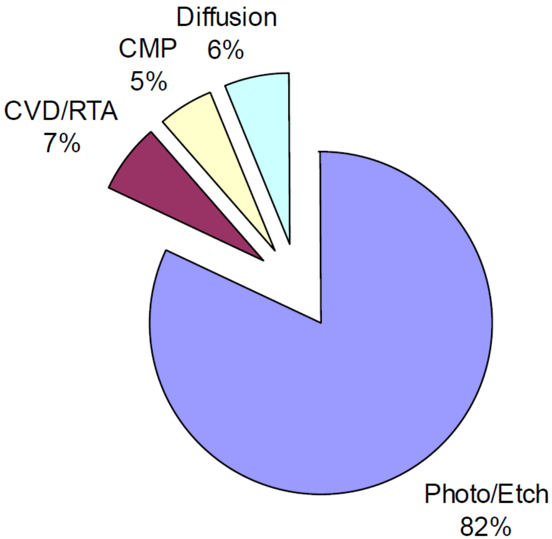


Mark C. Hersam, Northwestern University

Run-to-Run Control of Photolithography

- Lithography benefits most from R2R
  - Achieve proper CD and overlay - increased process capability
  - Reduce scrapped wafers – increased yield
  - Reduce rework and send-ahead wafers – reduced cycle time
- Litho R2R becomes increasingly important and a must-have in leading edge technology nodes and high-mix manufacturing environment
  - Decreasing feature size and introduction of multiple patterning technologies

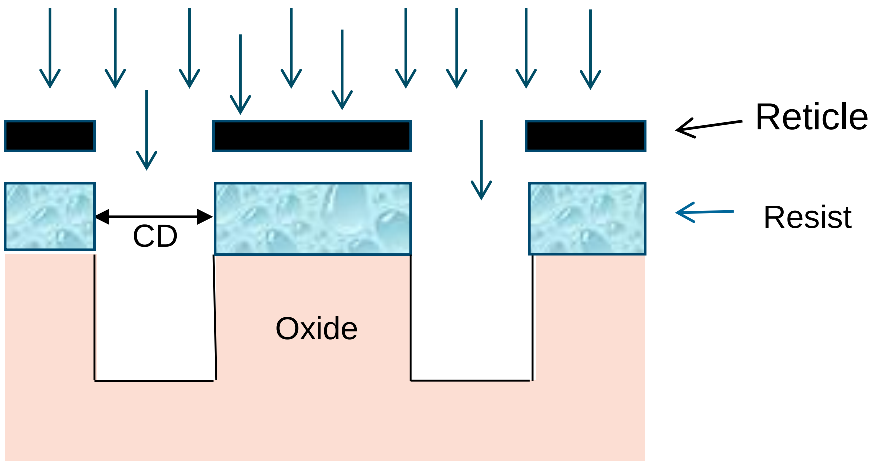
Run to Run (APC) Result	Benefit $M/year for 20k wafer start / month logic fab			
	Photo/ Etch	CVD/ RTA	CMP	Diffu-sion
Increased Process Capability	$360	$24		$24
Increased Process Accuracy				
Reduced Rework	$5			
Reduced Process Time	$11			
Reduced Downtime		$0	$1	
Reduced Cost of Consumables		$1.5	$5	
Engineering head count reduction	$0			
Reduced Operator Intervention			$1	$1
Reduced # of Wafers Scrapped	$2	$2	$7	$2
Reduced Send-ahead Wafers	$5		$11	

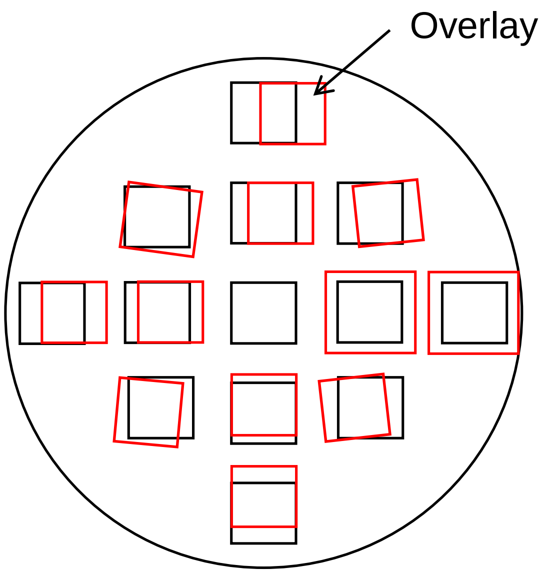


Stanley et al, SPIE, 2003

Litho R2R Strategies

- CD and overlay control
  - CD measures size of a particular feature, e.g., line, space, etc.
  - Overlay measures alignment of current pattern with respect to pattern from previous layer
  - Both are essential for high yield and finished IC performance

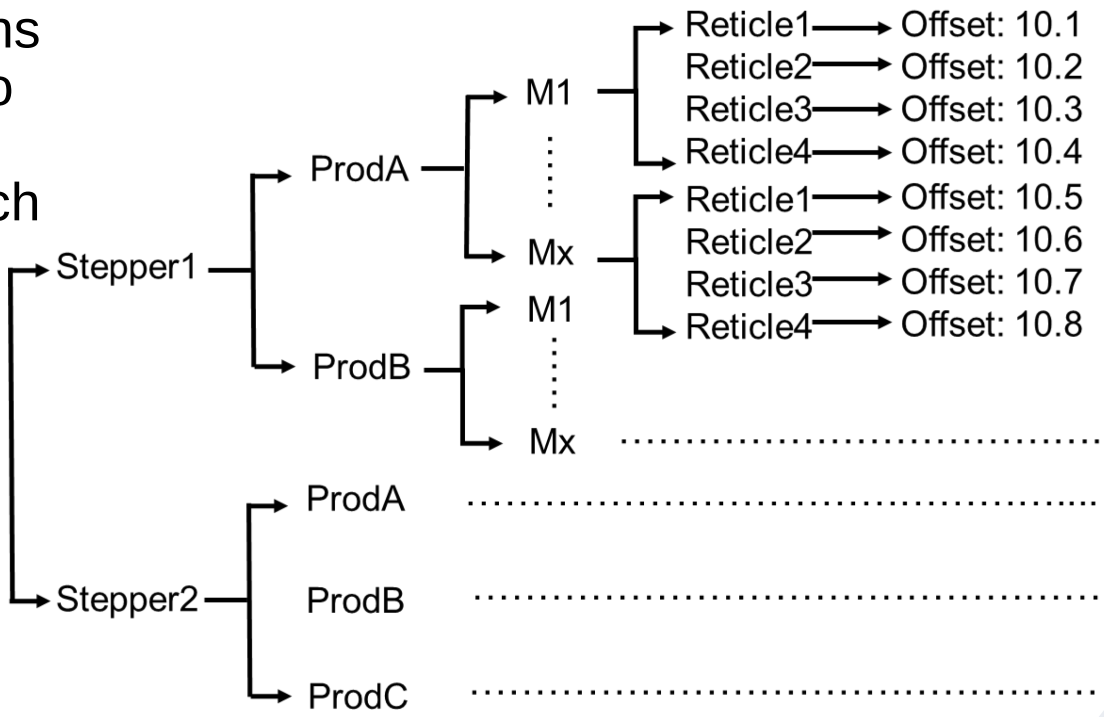




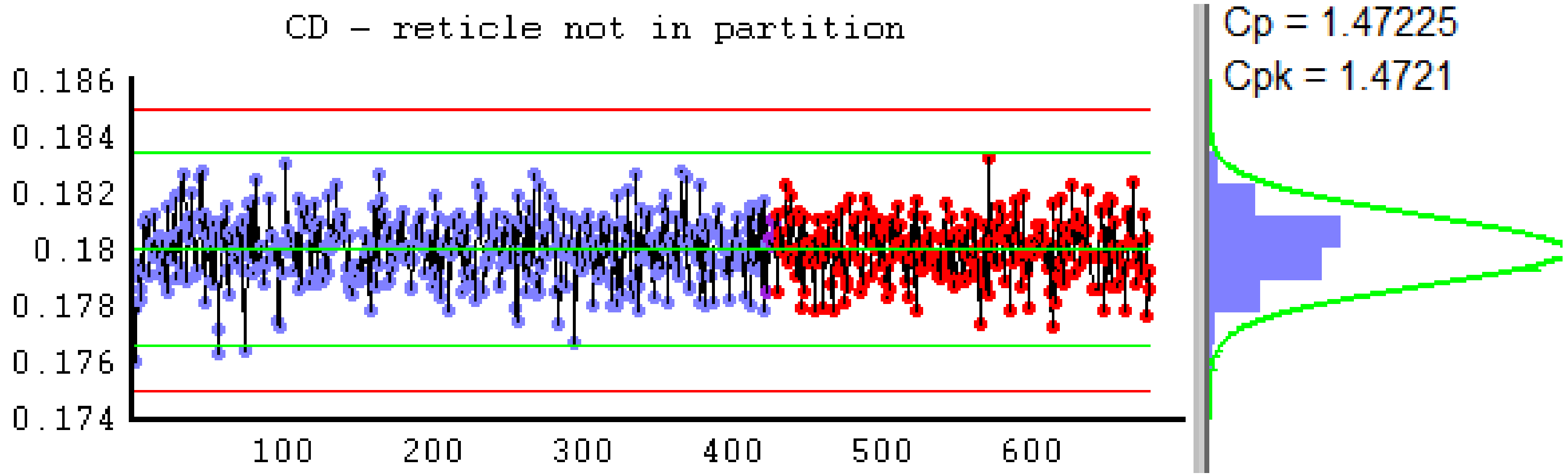
RUDOLPH TECHNOLOGIES

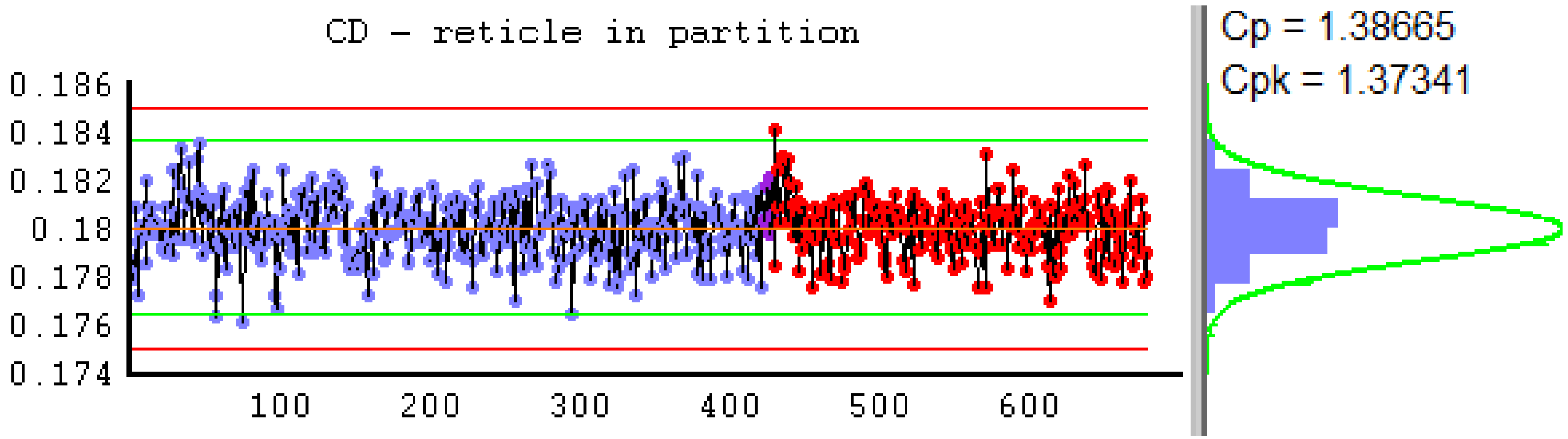
Partitions (Threads)

- Unique combinations of manufacturing context attributes, e.g., machine, product, layer, etc.
- Each partition has individual control loop using data only from itself.
- Proper definition of partitions separates disturbances into different groups (partitions) so that variability within each partition should be much smaller than the overall variability.
- Over-definition of partitions may undermine controller performance and lead to large number of partitions and data poverty.



Over-Partitioning Example

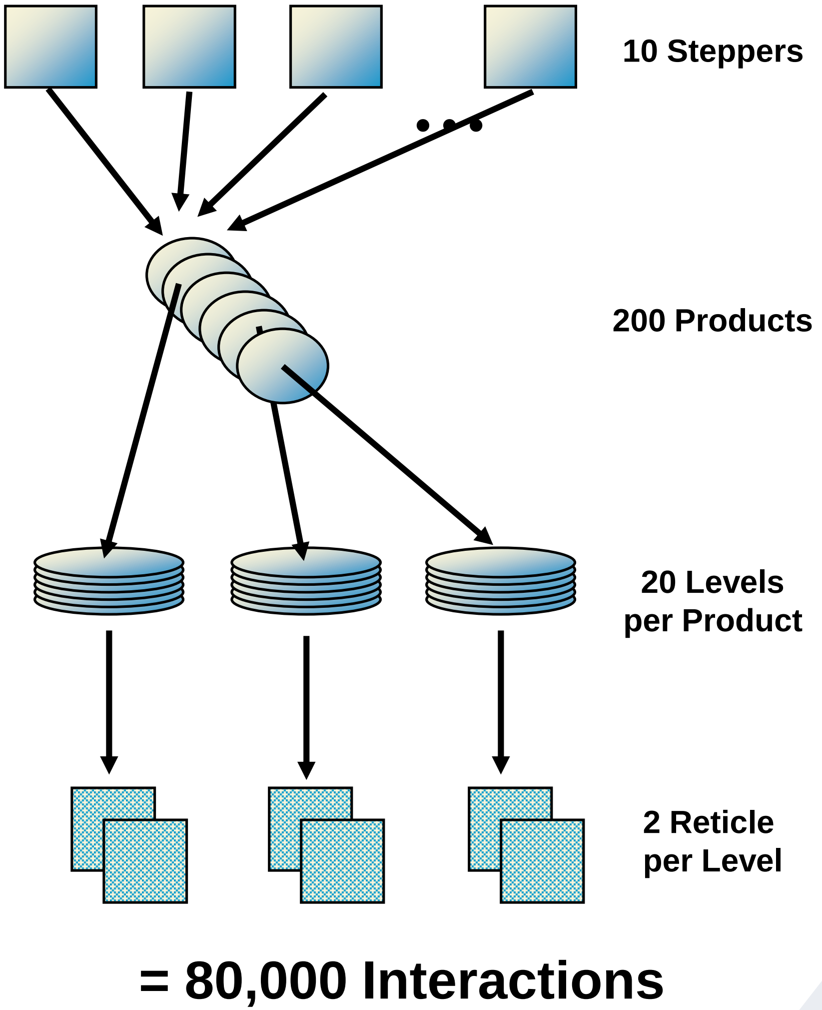




*Results are worse when reticle is in partition for modeled CD offset

High-Mix Challenges and Solutions

- Hard to keep low-running products updated if they require their own partition
  - Track time and number of wafers/runs since partition was last tuned
    - Control-oriented dispatching to help ensure partition state will be updated *(e.g., Anderson and Hanish, IEEE Trans. Semicond. Manuf., 2008)*
    - Require lookahead if last tuning was long ago *(e.g., Krumanocker and Yelverton, APC Conference, 2015)*
  - Similar partitions (e.g., reticles) can be combined into a partition group that share data with each other
  - Controller flexible enough to allow partition criteria change
  - Hierarchical partition definitions *(e.g., Yelverton and Agrawal, SPIE, 2014)*



Hierarchical Partitions and Dynamic State Estimation

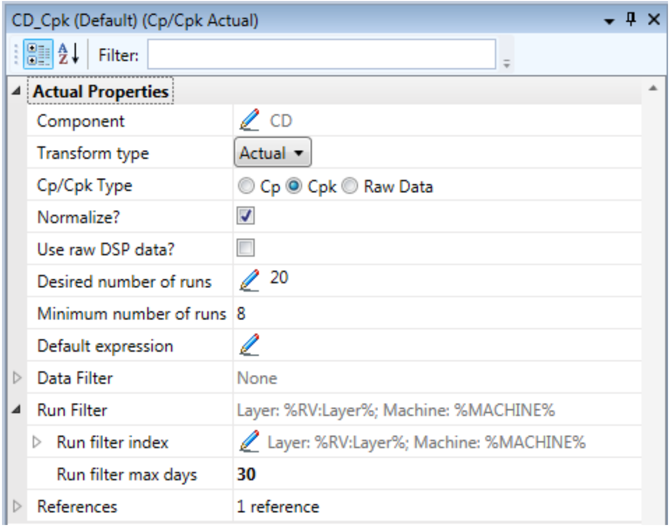
MP_AlignmentOffset_Prim1	Layer, MetToTool, Product, RefTool, Reticle
MP_AlignmentOffset_Prim2	Layer, Product, RefTool, Reticle
MP_AlignmentOffset_CR1	Layer, Product, Reticle
MP_AlignmentOffset_CR2	Layer, Product
MP_AlignmentOffset_CR3	Customer, Layer, Process

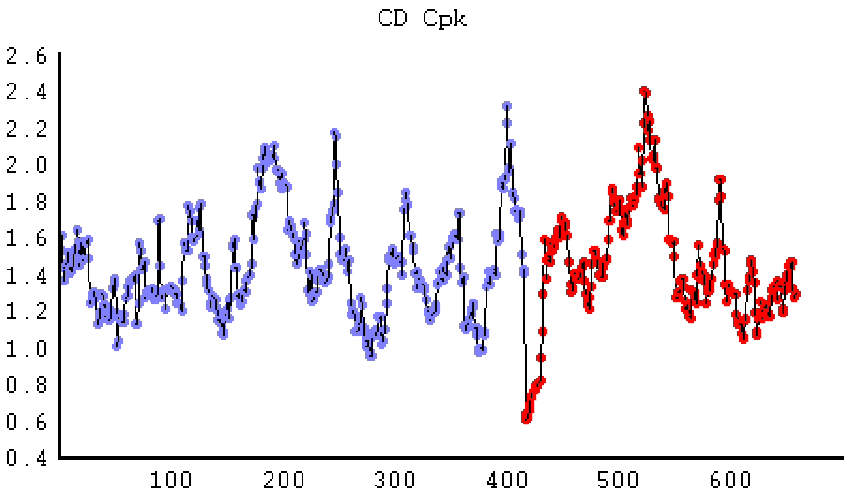
- Controller can start with the most specific partition and switch dynamically to partitions with relaxed criteria if not enough data is found
- Multiple hierarchical levels of partitions can be defined and take pre-determined precedence during run-time

```
if(timeSinceLastTune(MP_AlignmentOffset_Prim1)<(3600*24*180)) then
(MP_AlignmentOffset_Prim1) else(

if(timeSinceLastTune(MP_AlignmentOffset_Prim2)<(3600*24*180)) then
(MP_AlignmentOffset_Prim2) else(

if(timeSinceLastTune(MP_AlignmentOffset_CR1)<(3600*24*180)) then
(MP_AlignmentOffset_CR1) else(

if(timeSinceLastTune(MP_AlignmentOffset_CR2)<(3600*24*180)) then
(MP_AlignmentOffset_CR2) else(

if(timeSinceLastTune(MP_AlignmentOffset_CR3)<(3600*24*180)) then
(MP_AlignmentOffset_CR3) else(

C_AlignmentSettingsTable)))))
```

Smart Sampling

- R2R can provide additional information as guidelines for smart sampling
  - Tells sampling system to measure first few lots of a newly initialized partition to build up state estimate
  - Requests sampling on a particular partition if it hasn't been measured for certain period of time or certain number of lots
  - Calculates real-time metrics (e.g., Cpk) to evaluate process health by partition and determine sampling rate based on it *(Jones & Sun, APC Conference, 2016)*

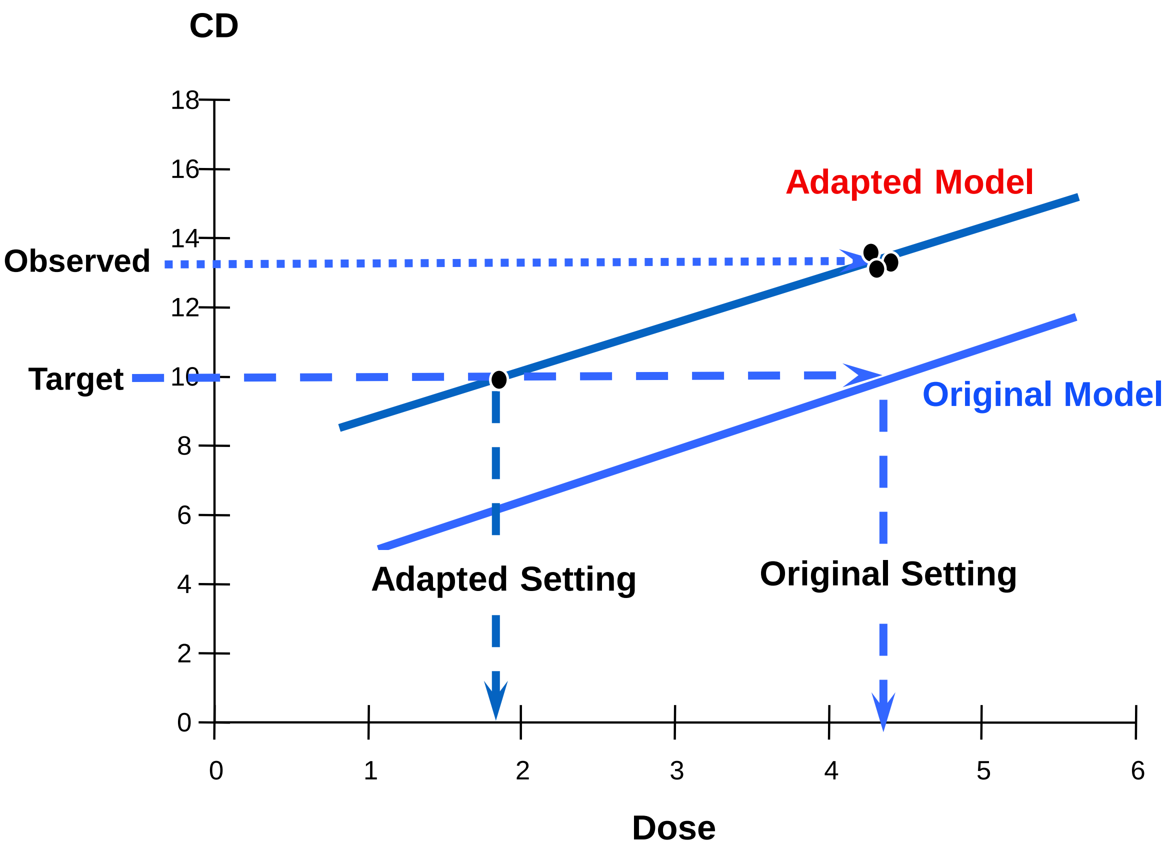


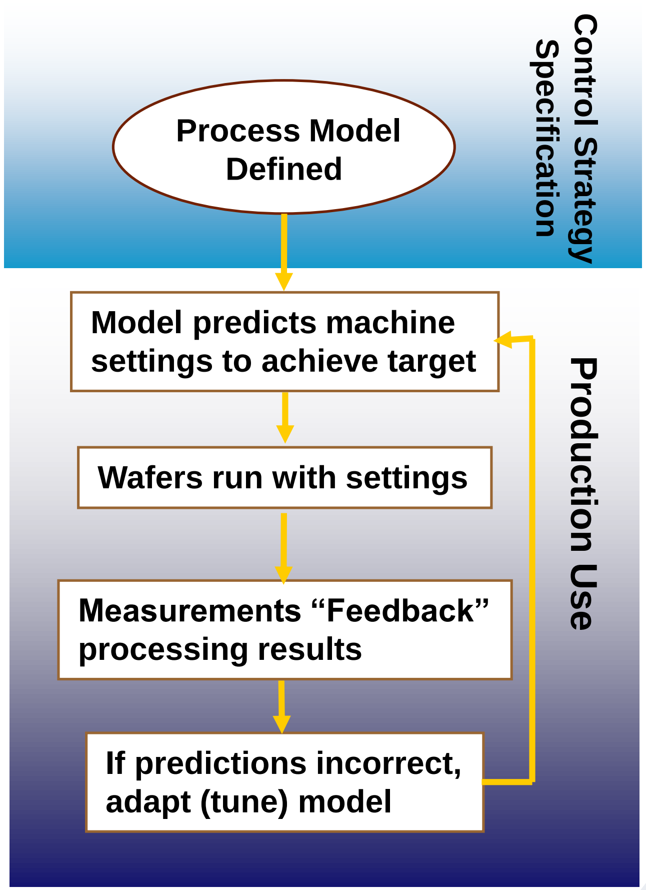


CD Control

Model: CD = Slope*Dose + Offset

where Offset is the State Variable



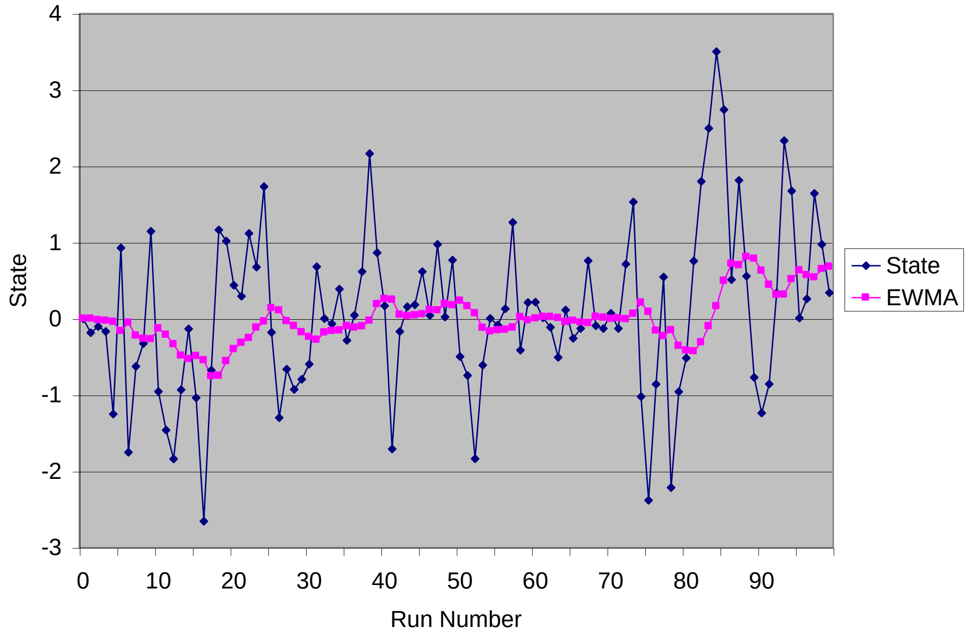


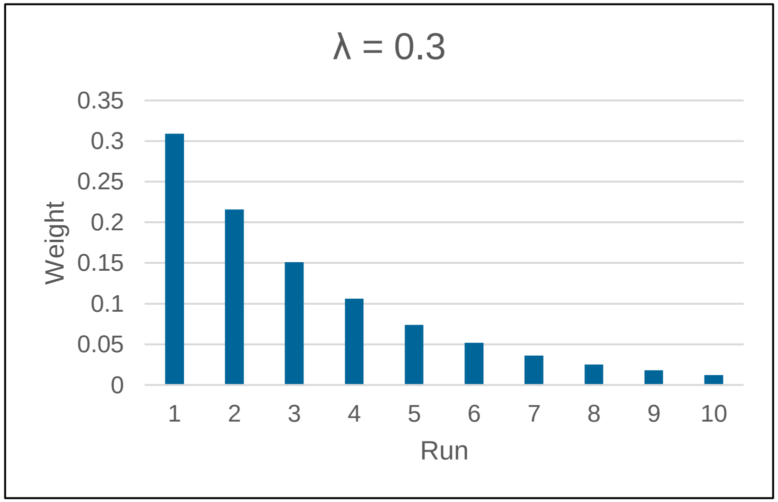
EWMA Filter

- Tuning algorithm adjusts state variable to reflect current processing conditions
- EWMA filter captures real trend from noisy process data
  - Exponentially Weighted Moving Average

    $$\hat{y}_{k+1} = (1-\lambda)\hat{y}_k + \lambda y_{k+1}$$

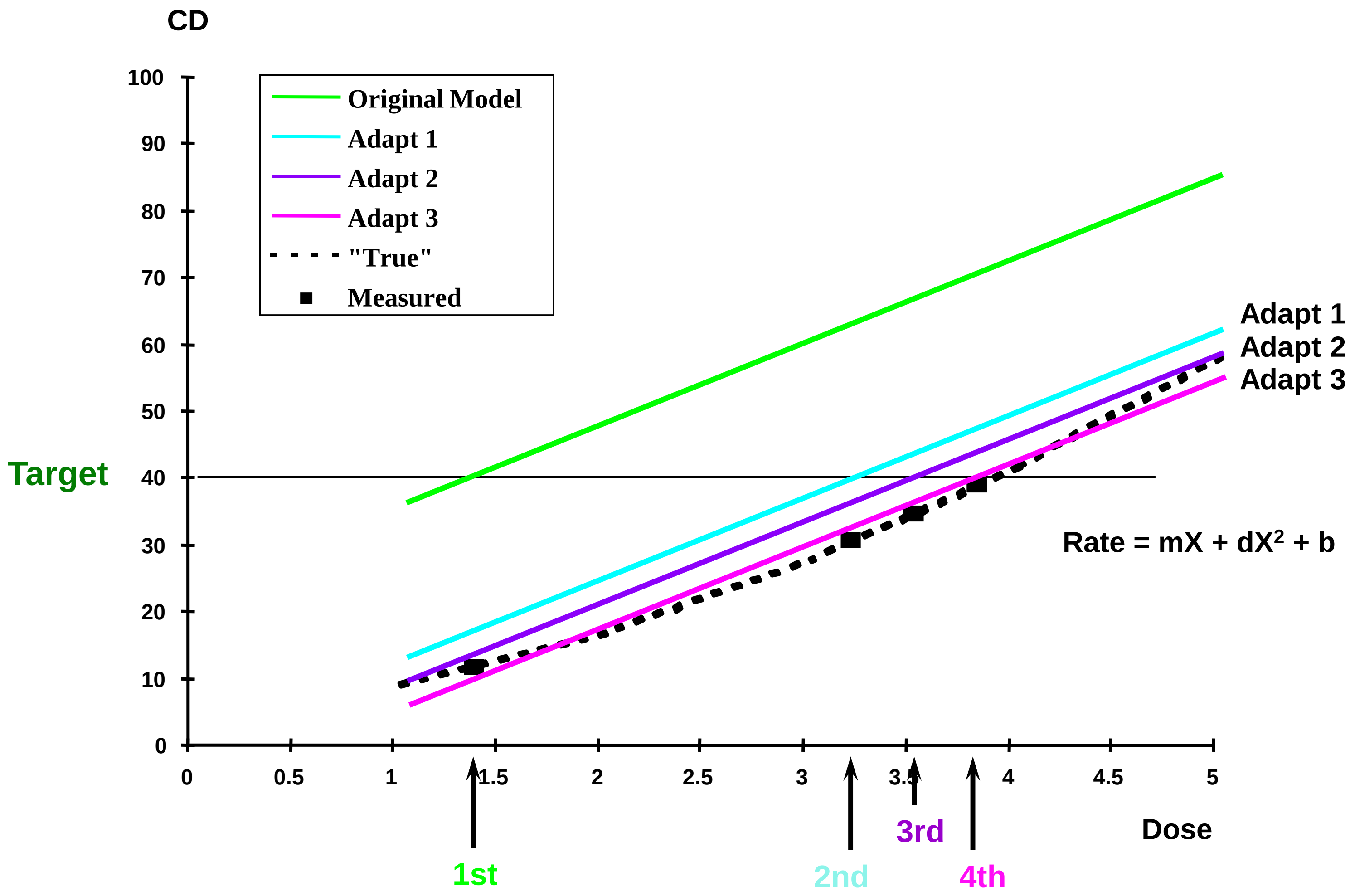
  - Higher weight given to most recent data
  - Weight decreases exponentially
  - Optimal λ value can be determined by simulation





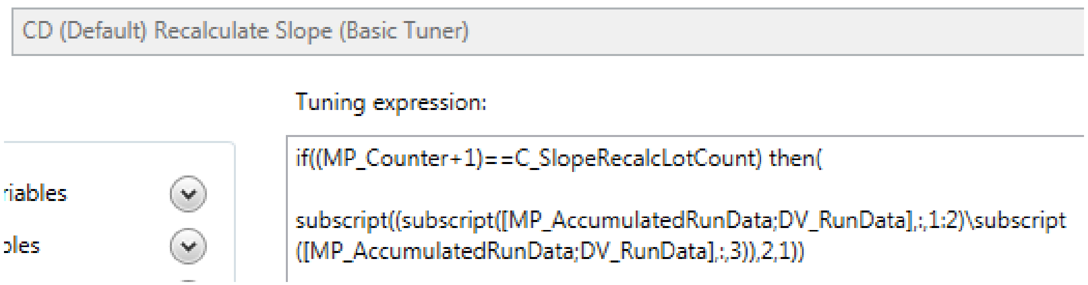
RUDOLPH TECHNOLOGIES

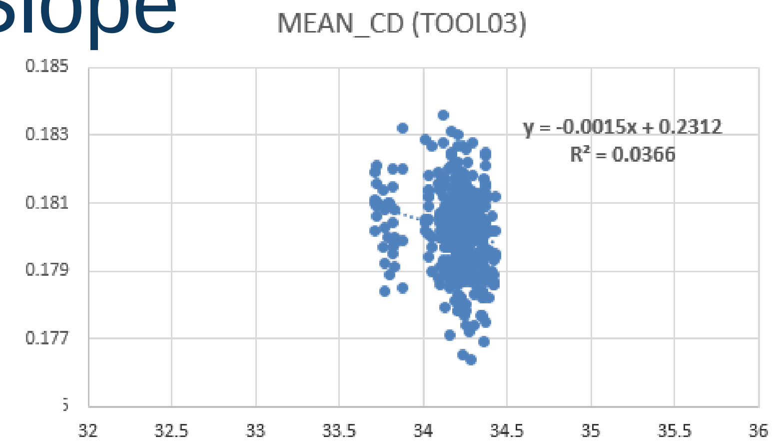
More Realistic Example

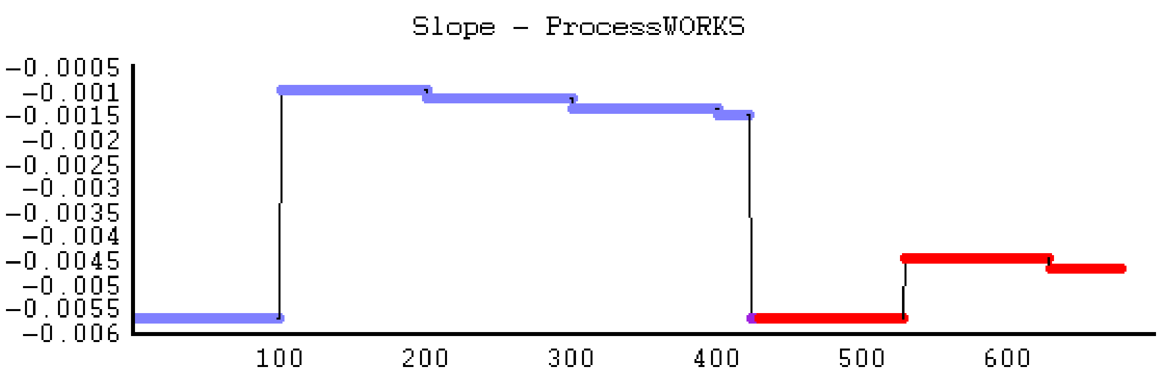


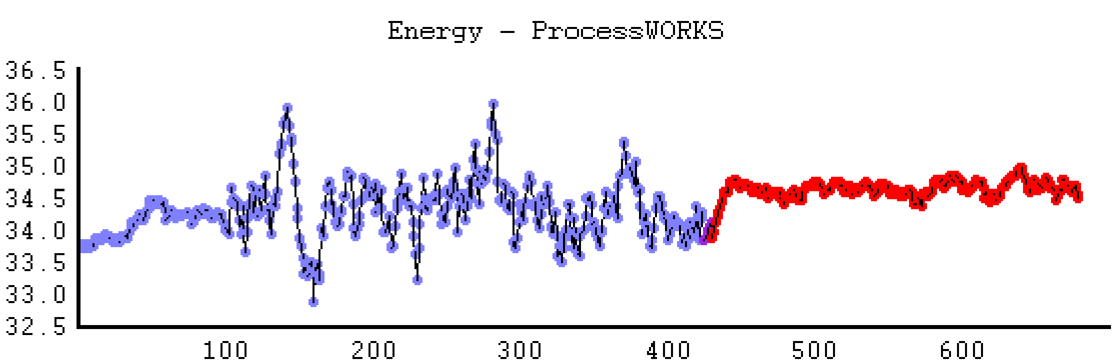
RUDOLPH TECHNOLOGIES

CD Control with Adaptive Slope







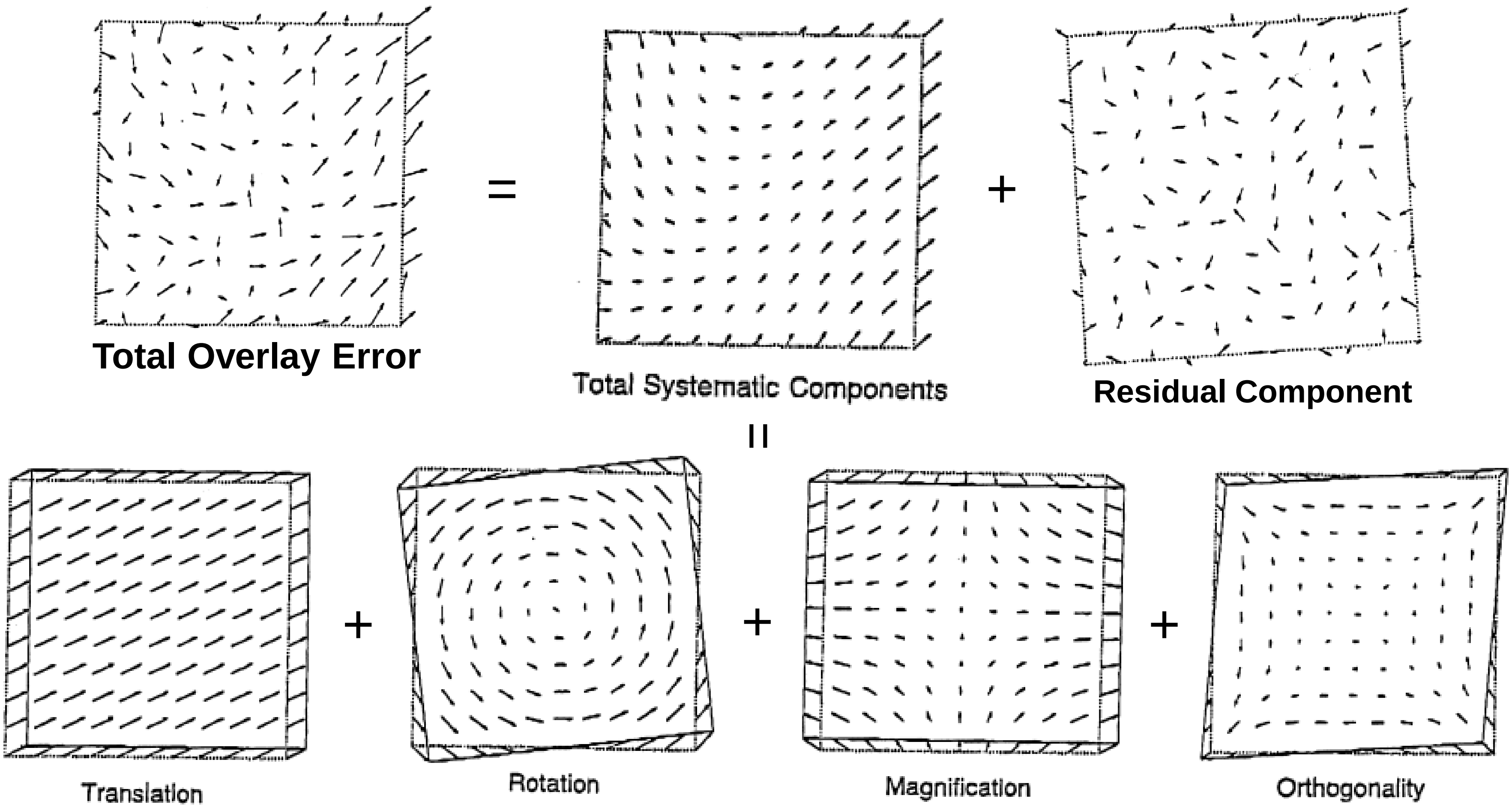


- Slope optimized using machine learning algorithms once enough new datasets are collected
  - e.g., Automatically recalculate slope every 100 runs and apply in dose calculation

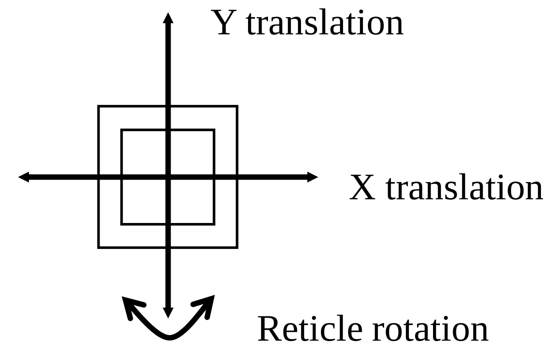
RUDOLPH TECHNOLOGIES

Overlay Measurements

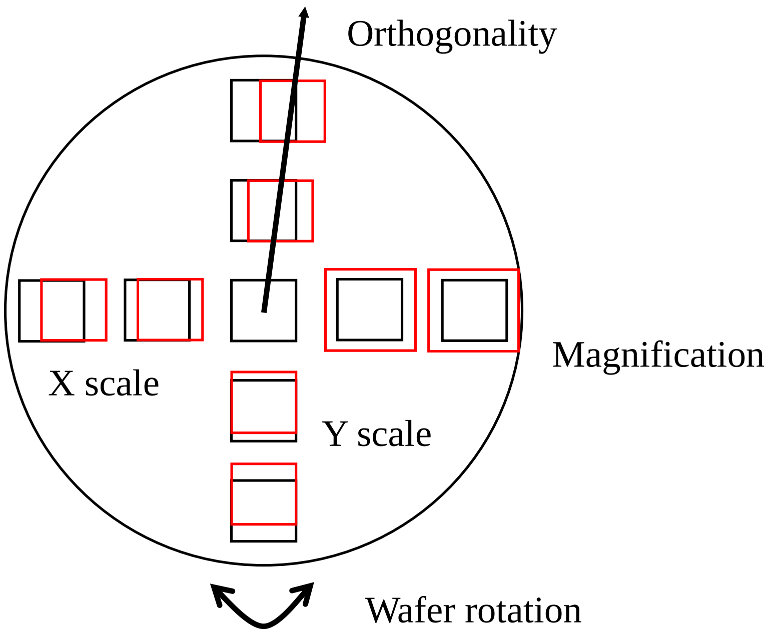
- Total overlay error is decomposed into a few main systematic error modes. *(Armitage and Kirk, SPIE, 1988)*



Sample Overlay Controller



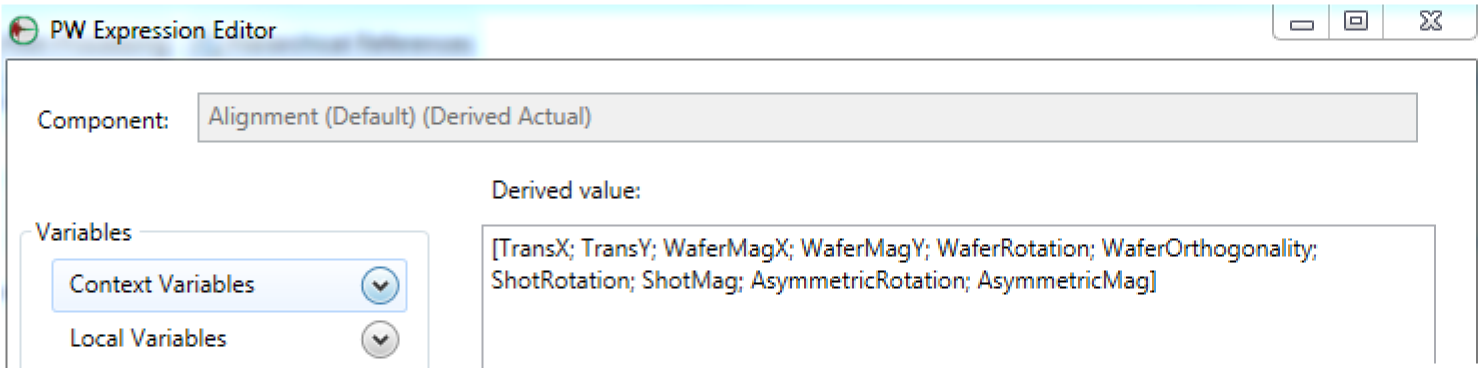
- Process model
  - **AlignmentMeasurement** = **AlignmentSetting** – ***Offset***
  - Alignment = $\begin{bmatrix} \text{TransX} \\ \text{TransY} \\ \text{WaferMagX} \\ \text{WaferMagY} \\ \text{WaferRotation} \\ \text{WaferOrthogonality} \\ \text{ShotRotation} \\ \text{ShotMag} \\ \text{AsymmetricRotation} \\ \text{AsymmetricMag} \end{bmatrix}$

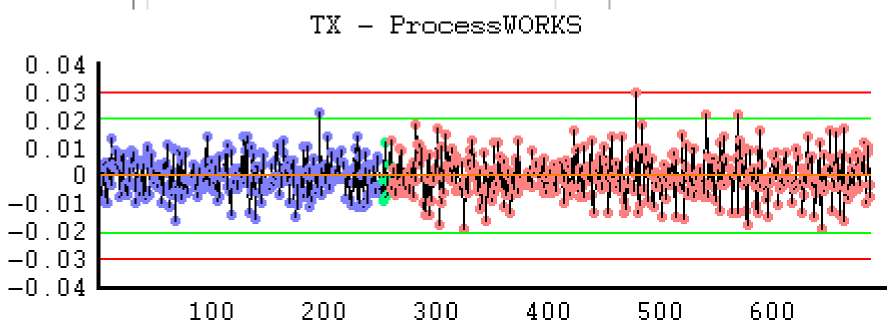


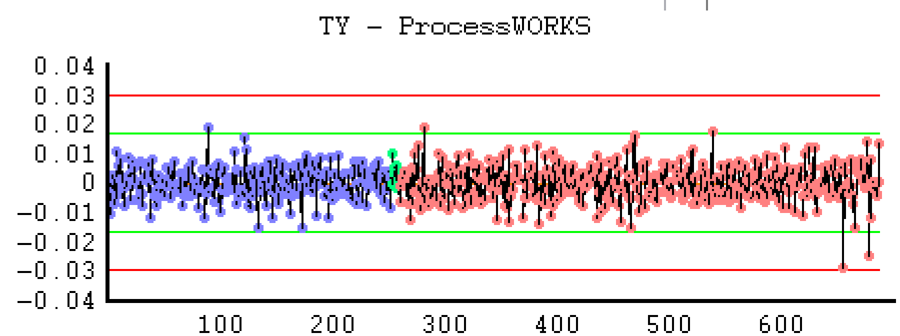
- EWMA tuner
  - ***PredOffset***(n+1) = (1−λ)****PredOffset***(n) + λ***ObservedOffset**(n)

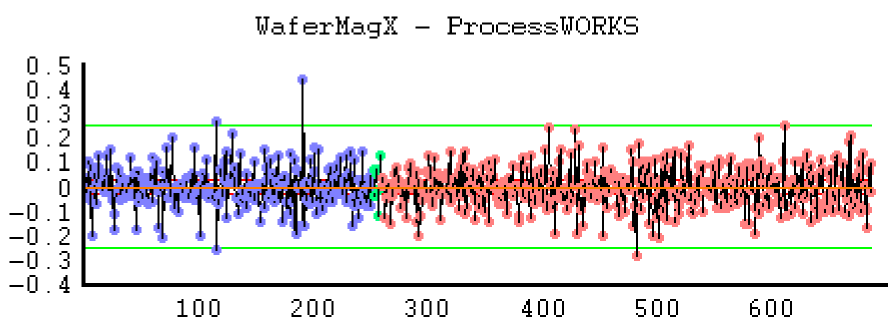
Basic Overlay Control

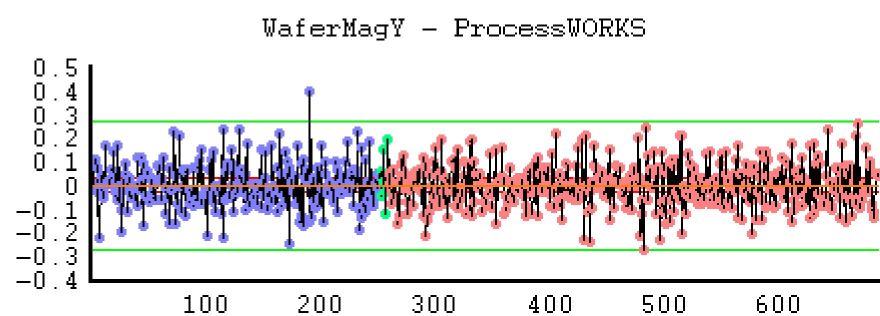
10 first-order terms





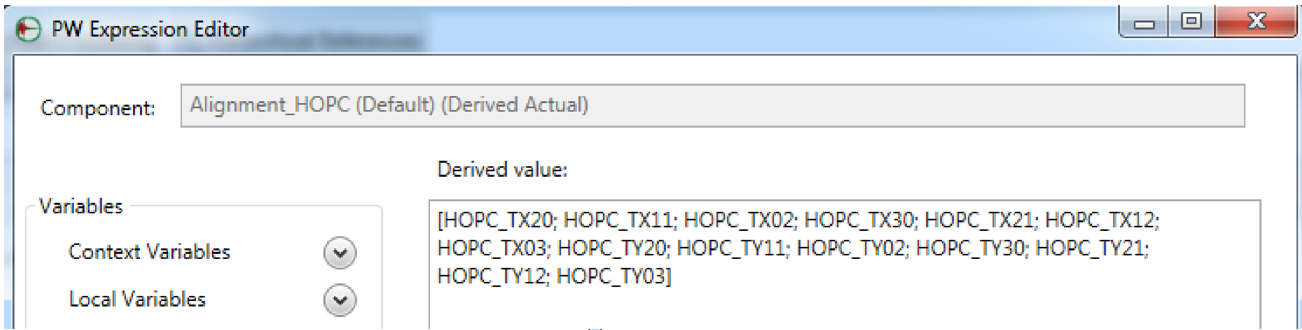


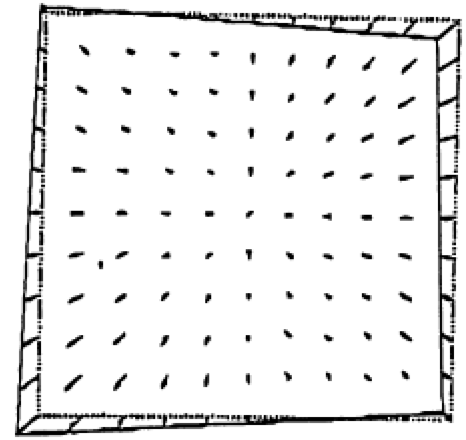




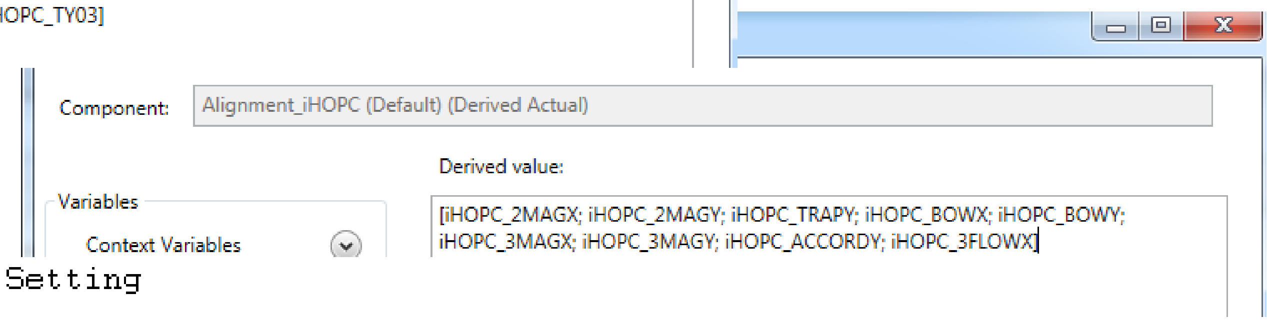
HOPC/iHOPC R2R Control

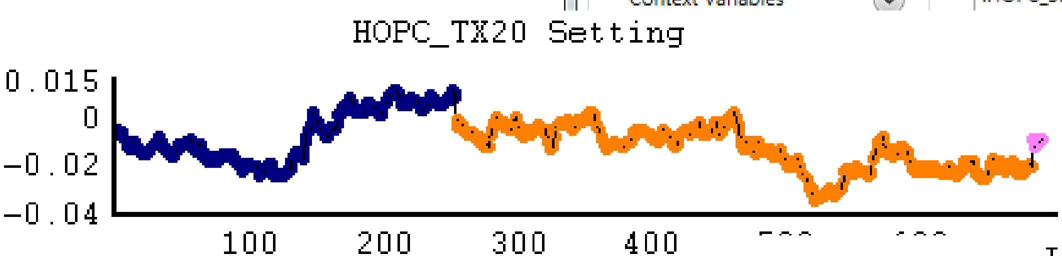
14 HOPC terms & 9 iHOPC terms

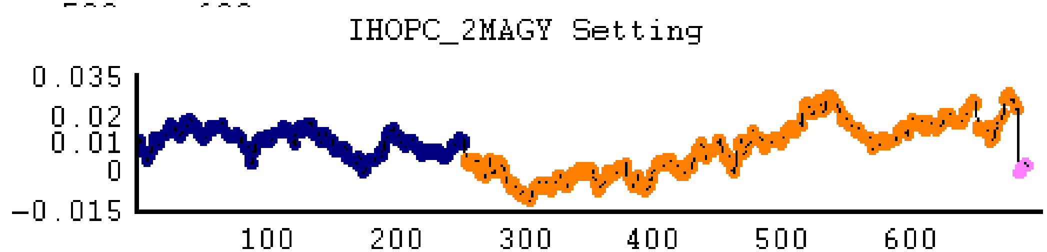




Trapezoid



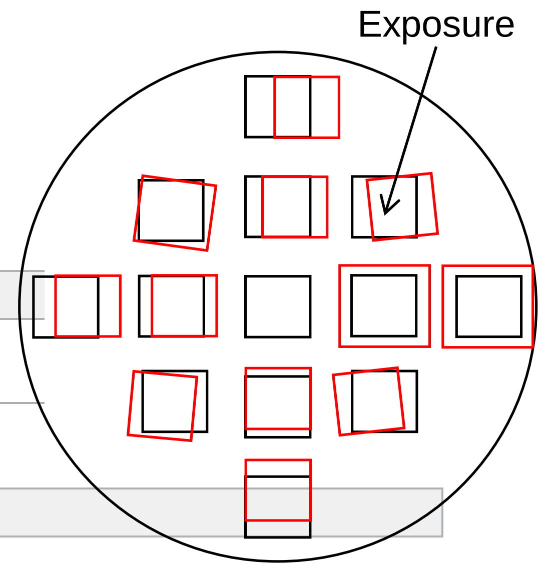




- More terms can be added

Dynamic CPE Control

6 CPE terms X 66 exposures = 396 settings adjusted separately



Component: Alignment_CPE (Default) (Derived Actual)

Variables

Context Variables

Derived value:

[CPE_WTRX; CPE_WTRY; CPE_FMAG; CPE_FROT; CPE_AMAG; CPE_AROT]

Component: CPE_FINDX (Layer: 130; Product: 5974)

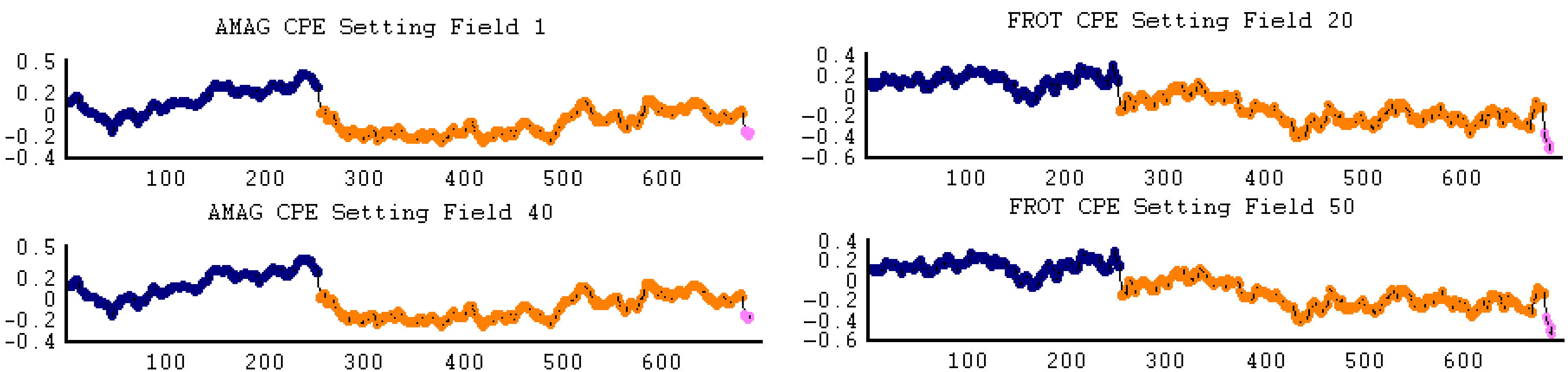
Variables

Context Variables

Local Variables

Derived value:

[(-3) (-2) (-1) 0 1 2 3 (-4) (-3) (-2) (-1) 0 1 2 3 4 (-4) (-3) (-2) (-1) 0 1 2 3 4 (-5) (-4) (-3) (-2) (-1) 0 1 2 3 4 5 (-5) (-4) (-3) (-2) (-1) 0 1 2 3 4 5 (-4) (-3) (-2) (-1) 0 1 2 3 4 (-3) (-2) (-1) 0 1 2 3 (-1) 0 1]



RUDOLPH TECHNOLOGIES

Photolithography R2R Summary

- Scalable enough to handle various number of settings ranging from tens to hundreds
- Flexible enough to allow each control technique (e.g., CD, basic overlay, HOPC, iHOPC, CPE)
  - To be turned on/off independently
  - To have unique partition definition
- Hierarchical partitioning scheme can be set up and switched dynamically

MP_AlignmentOffset_Prim1	Layer, MetToTool, Product, RefTool, Reticle
MP_AlignmentOffset_Prim2	Layer, Product, RefTool, Reticle
MP_AlignmentOffset_CR1	Layer, Product, Reticle
MP_AlignmentOffset_CR2	Layer, Product
MP_AlignmentOffset_CR3	Customer, Layer, Process

- Settings
  - AlignmentSettings (Default)
  - AlignmentSettings_CPE
    - AlignmentSettings_CPE (Layer: 130; Product: 5974)
  - AlignmentSettings_HOPC
    - AlignmentSettings_HOPC (Layer: 130; Product: 5974)
  - AlignmentSettings_iHOPC
    - AlignmentSettings_iHOPC (Layer: 130; Product: 5974)

MP_AlignmentOffset_HOPC	Layer, Product, Reticle
MP_AlignmentOffset_iHOPC	Layer, Product, Reticle
MP_AlignmentOffset_CPE	Layer, Product, Reticle

```
if(timeSinceLastTune(MP_AlignmentOffset_Prim1)<(3600*24*180)) then
(MP_AlignmentOffset_Prim1) else(

if(timeSinceLastTune(MP_AlignmentOffset_Prim2)<(3600*24*180)) then
(MP_AlignmentOffset_Prim2) else(

if(timeSinceLastTune(MP_AlignmentOffset_CR1)<(3600*24*180)) then
(MP_AlignmentOffset_CR1) else(

if(timeSinceLastTune(MP_AlignmentOffset_CR2)<(3600*24*180)) then
(MP_AlignmentOffset_CR2) else(

if(timeSinceLastTune(MP_AlignmentOffset_CR3)<(3600*24*180)) then
(MP_AlignmentOffset_CR3) else(

C_AlignmentSettingsTable)))))
```

谢谢 | 謝謝

danke

ありがとう

Thank You!

감사합니다

merci

obrigado

info@rudolphtech.com
www.rudolphtech.com

RUDOLPH TECHNOLOGIES