“R” Open Source Statistics for Semiconductor Processing

D. Youlton, D. Fitzpatrick, P.F. Byrne

Brookside Software

Overview

- “R” Statistics Package
- Application in Semiconductor Manufacturing
- Sample Usage
- Open Source Software
- Conclusion

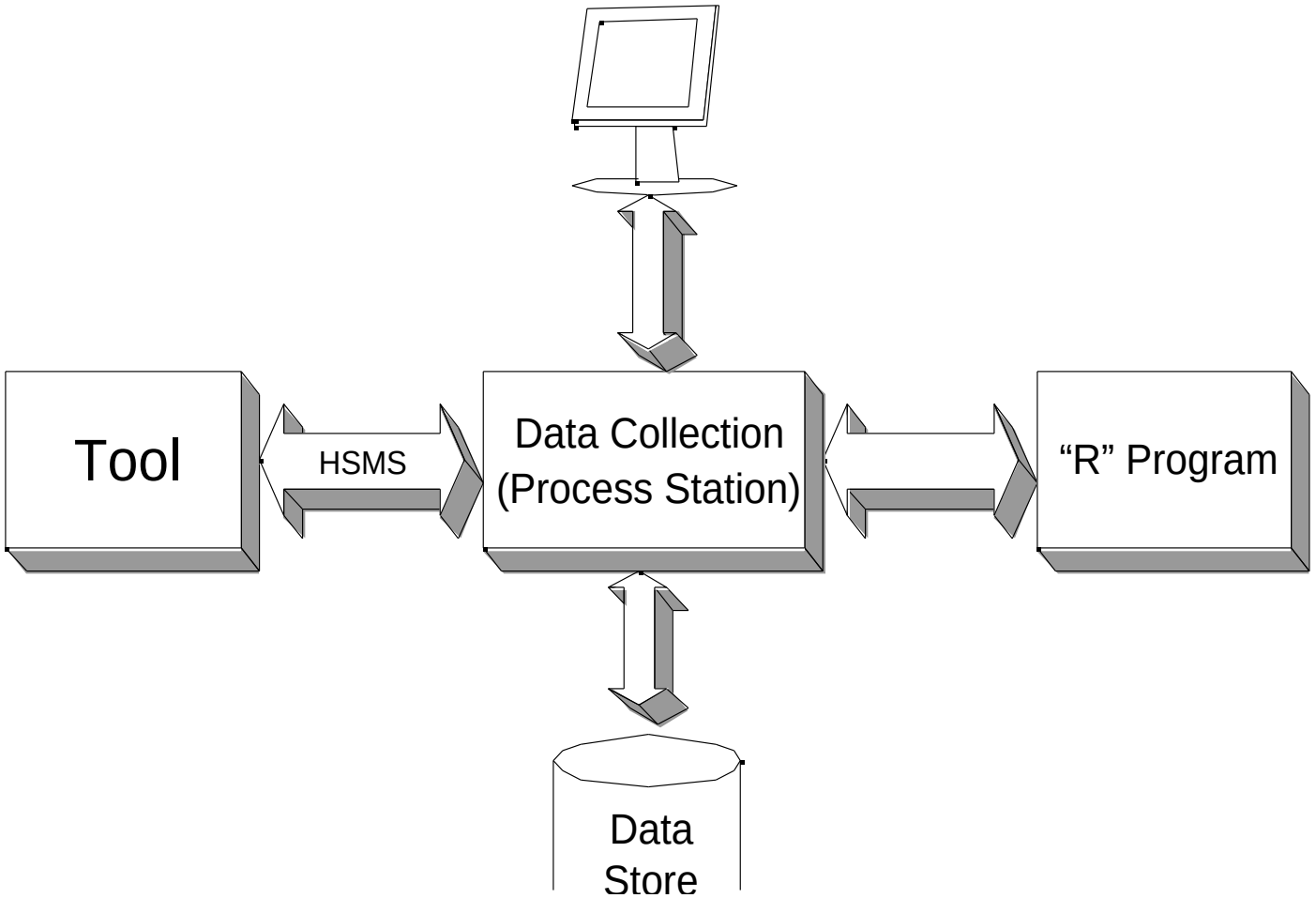
“R” Statistics

- Interactive Statistics Platform
- Widely used in Industry and Academia
  - Engineers trained in “R”
  - Many Active Users
- Standard Statistical Models
- Active Development
- Open source version of “S”

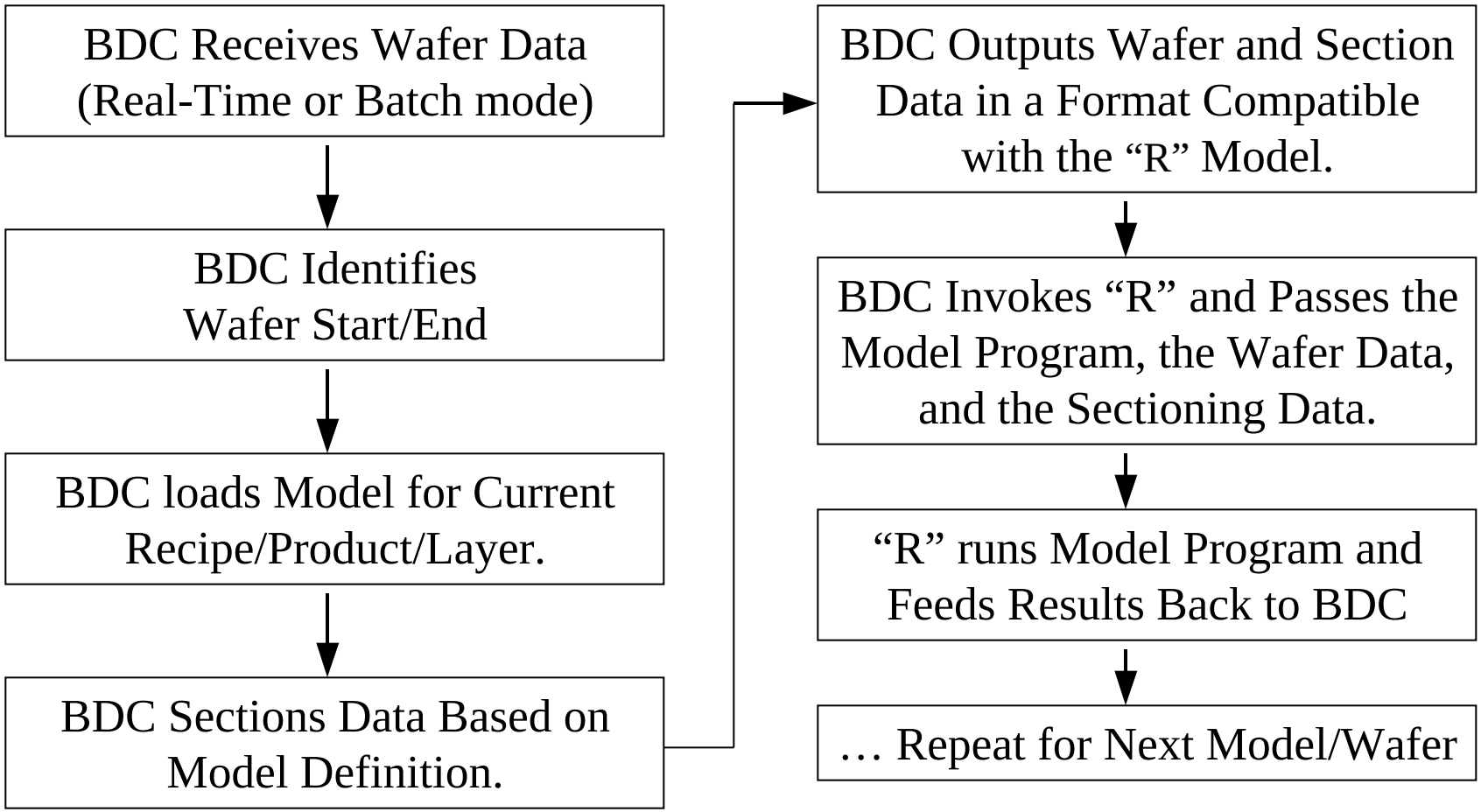
Applications in Semiconductor

- Real Time Fault Detection
  - Data Imported from Brookside Software Data Collection System
  - Wafer-by-Wafer data Analysis
  - Statistical Tests Provided as Templates
- Works “out-of-the-box”
- Modifiable by customer

System Architecture



Using “R” To Analyze Process Data



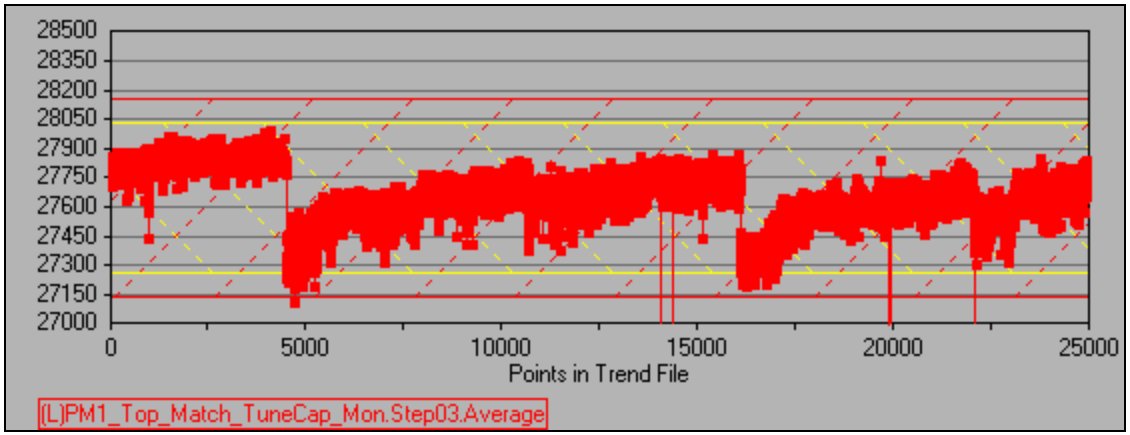
(*) BDC = Brookside Data Collector

Example “R” Program Commands

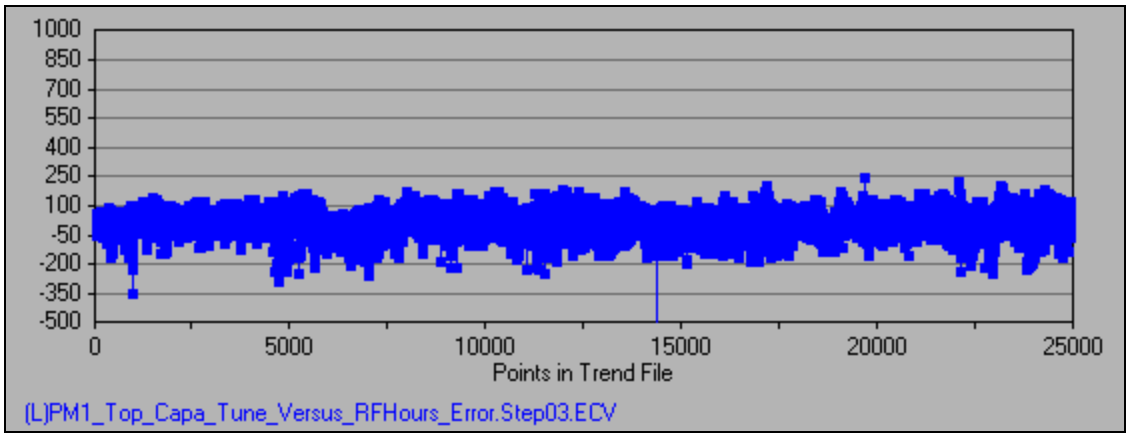
- Load a Data or Section Table into Memory
  ```
  DateTable=read.table(filename,TRUE,sep="\t")
  ```
- Calculate the Mean For 1 Variable
  ```
  var=mean(DataTable[StartRow:EndRow,ColumnIdx])
  ```
- Calculate Correlation Table Based on a Data Table
  ```
  CorrTable=cor(DataTable)
  ```
- Calculate Covariance Matrix
  ```
  CovMatrix=cov(DataTable)
  ```
- Invert the Covariance Matrix
  ```
  InvCovMatrix=solve(CovMatrix)
  ```

Monitoring Signal Drift Between Cleans

Signal drift between cleans forces wide control limits



Use “R” to create a normalized signal.



RESULT:

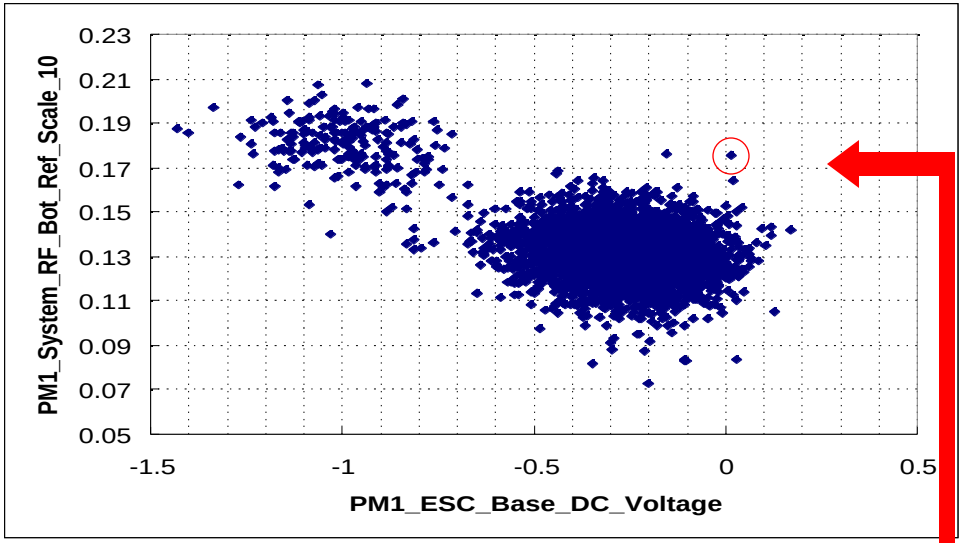
Much tighter control limits can be used.

Greater sensitivity to wafer to wafer variations.

Multivariate Analysis

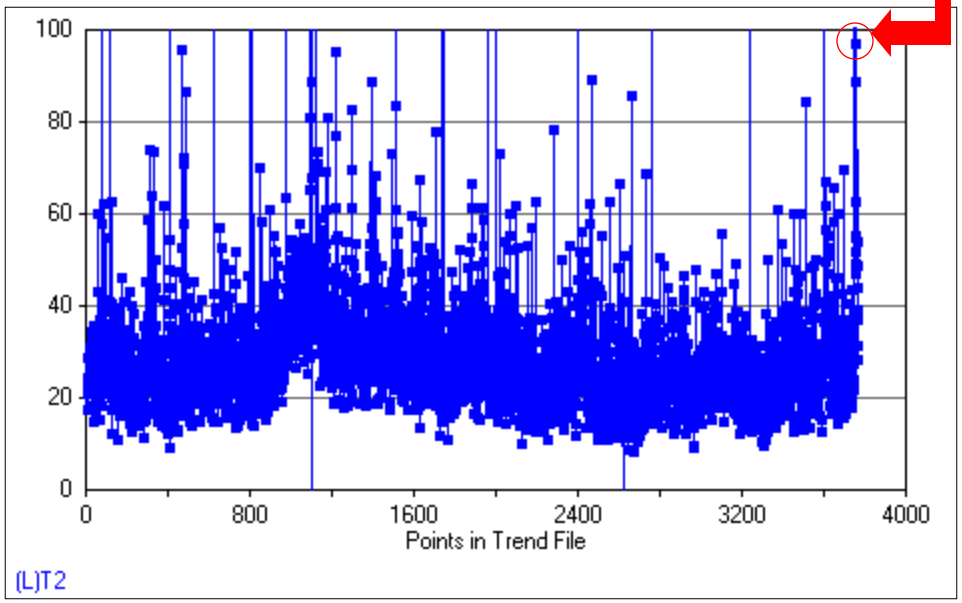
This data point is within it's normal range for each parameter separately.

But, clearly it's outside of it's normal range when considering both parameters combined.



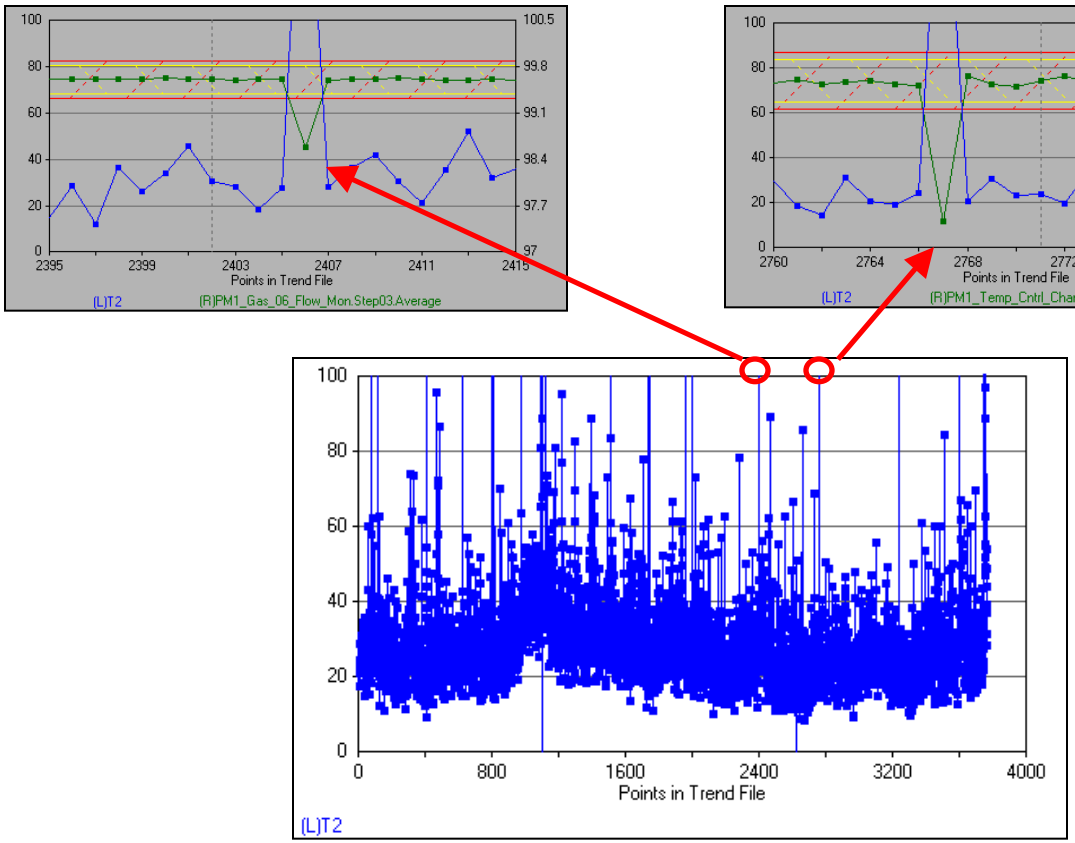
The Hotelling T^2 multivariate statistic detects this point as an outlier.

In this example, the T^2 statistic is calculated using 30 process parameters.



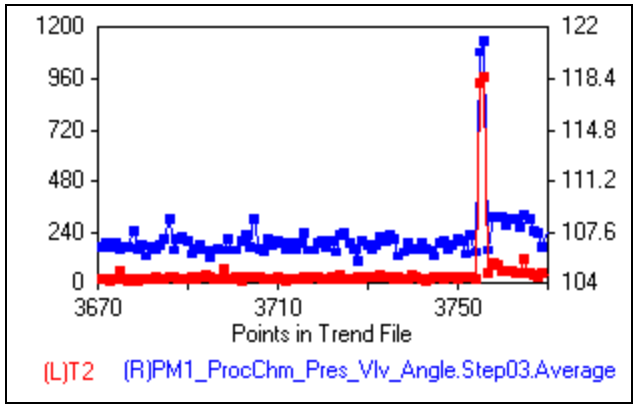
Process Monitoring Using 1 Control Chart

- Use a single Multivariate SPC control chart, drill down as needed.
- Hotelling T^2 can highlight both multivariate (previous example) and single variable faults (shown below)



Identify the Faults

< Column 37 > T2
27.5948667786530
34.2737149963767
28.4036877371761
23.9449410713582
936.929122764276
970.781992255485
47.1411893264346
97.025516651528
89.100148754251
62.9239621195477
55.1059181640305



< Column 7 > PM1_Bot_Match_PeakVolt_Mon Step03 Average
-2.038429375
-2.02193066666667
-2.004006
-2.011938
-1.62114266666667
-1.65116066666667
-2.032728125
-2.05193533333333
-2.07044875
-2.05216733333333
-1.99192333333333

< Column 17 > PM1_ProcChm_Pres_Vlv_Angle Step03 Average
107.125
106.133333333333
107.533333333333
106.2
120.333333333333
121.133333333333
106.375
108.8
108.6875
108.733333333333
108.2

< Column 21 > PM1_Top_Match_TuneCap_Mon Step03 Average
27687.375
27629.7333333333
27662.5333333333
27682.9333333333
26804.2666666667
26695.5333333333
27720.3125
27717.6
27722.625
27773.4
27658.7333333333

Advantages of Open Source Software

- Cost-effective
- Robust with ongoing validation
- Source code is available
- Large Installed Base
- Ongoing development
- Upgrades
- Technology that has come of age
  - ~ 70% of web servers in use are based on open source

Conclusion

- “R” Provides a Cost Effective Statistics Platform for the Manufacturing Environment
- Open Statistics Platforms provide Flexibility and Ease of Use
- Open source software is mainstream

References

- **http://www.r-project.org/**
  - **Home page for the R-Project**
- **http://cran.r-project.org/**
  - **Download the latest releases**
- **http://www.brooksidesoftware.com**
  - **For latest data on this project**

Acknowledgements

We wish to thank STMicroelectronics for providing data for this presentation.