How to Swallow a Mouthful of Data and Avoid Indigestion

Barry M. Wise, Ph.D.
President
Eigenvector Research, Inc.

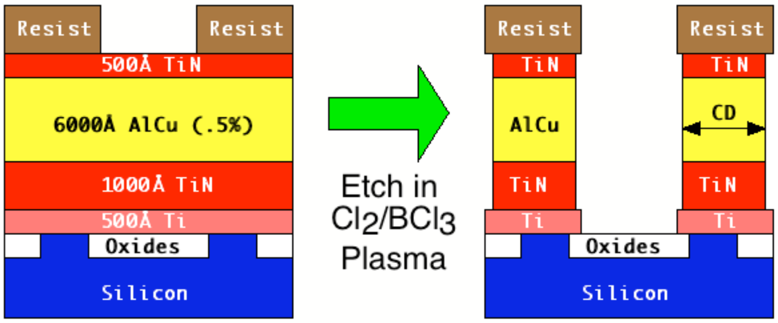
EIGENVECTOR RESEARCH INCORPORATED 1

Outline

- The process monitoring data problem
- Example data set
- Approaches to preprocessing
  - Variable means
  - Warping for multi-way
  - Data summary
- Comparison of results
- Conclusions

EIGENVECTOR RESEARCH INCORPORATED 2

Plasma Metal Etch



- Linewidth (Critical Dimension) Control
  - Constant linewidth reduction run to run and across wafer
  - Constant linewidth reduction for every material in stack
- Minimal damage to oxide

EIGENVECTOR RESEARCH INCORPORATED 3

Metal Etch Data Set

This data has been used many times and is available at www.eigenvector.com. A few examples:

- Barna, G.G., White, D., Wise, B.M., Gallagher, N.B., Sofge, D., “Development of Robust Fault Detection and Classification Techniques/SEMATECH J-88E Project at TI”, SEMATECH AEC/APC Workshop VIII, Santa Fe, New Mexico, Oct. 27-30, 1996.
- White, D., Barna, G.G., Butler, S.W., Wise B., and Gallagher, N., "Methodology for Robust and Sensitive Fault Detection," Electrochemical Society Meeting, Montreal, May, 1997.
- Gallagher, N.B.,Wise, B.M., Butler, S.W., White, D., and Barna, G.G., "Development and Benchmarking of Multivariate Statistical Process Control Tools for a Semiconductor Etch Process: Improving Robustness Through Model Updating", IFAC ADCHEM'97, Banff, Canada, 78–83, June, 1997.
- Wise, B.M., Gallagher, N.B., Butler, S.W., White, D., and Barna, G.G., "A Comparison of Principal Components Analysis, Multi-way Principal Components Analysis, Tri-linear Decomposition and Parallel Factor Analysis for Fault Detection in a Semiconductor Etch Process," *J. Chemometr.*, **13**, 379–396 (1999).
- Wise, B.M., Gallagher, N.B., Butler, S.W., White, D., and Barna, G.G., "Development and Benchmarking of Multivariate Statistical Process Control Tools for a Semiconductor Etch Process: Impact of Measurement Selection and Data Treatment on Sensitivity", IFAC SAFEPROCESS'97, 35–42, Kingston Upon Hull, U.K., Aug., 1997.
- Wise, B.M., Gallagher, N.B., and Martin, E.B., “Application of PARAFAC2 to Fault Detection and Diagnosis in Semiconductor Etch,” *J. Chemometr.*, **15**(4), 285–298 (2001).
- Wise, B.M., Gallagher, N.B., "Multi-way Analysis in Process Monitoring and Modeling," *AIChE Symposium Series*, **93**(316), 271–274 (1997).
- Warren, J. and Gallagher, N.B., “Heuristic and Statistical Methods for Fault Detection: Complementary or Competing Approaches?”, SEMATECH AEC/APC Symposium XVIII, Westminster, CO, Sept. 30-Oct. 5 (2006).

EIGENVECTOR RESEARCH INCORPORATED 4

Available Measurements

- Equipment has SECS-II Port
  - Provides traces with time stamp and step number
- Regulatory controller setpoints & controlled variable measured values
  - gas flows, pressure, plasma powers
- Regulatory controller manipulated variables
  - exhaust throttle valve, capacitors
  - mass flow controller do not provide valve position
- Additional process measurements
  - endpoint intensity (plasma emission at particular frequency)
  - impedance measurements
- ~~Optical emission spectra~~
- ~~RF plasma variables~~

Sensitivity of MSPC Models

- Three experiments performed with 21 “induced” faults on:
  - TCP top power
  - RF bottom power
  - Cl2 flow
  - BCl3 flow
  - Chamber pressure
  - Helium chuck pressure
- Goal: Compare ability of models considered for detecting faults

Global and Local Models

- Global models based on data over long period of time with considerable variance due to drift
- Local models build over narrower time windows, less drift variance

EIGENVECTOR RESEARCH INCORPORATED 7

Generating Faults

- Setpoints were changed for controlled process variables
- Data for the controlled variable was adjusted to have the original desired mean
- Result is data that looks like a sensor has developed a bias

EIGENVECTOR RESEARCH INCORPORATED 8

Batch Process Monitoring Data Problem

- Messy-typically includes start-up and shut-down phases that are not of interest
- Periods of “steady-state” where not much is changing
- Variable record lengths
- Lots of data!
- Reduce to a set of more compact descriptors?

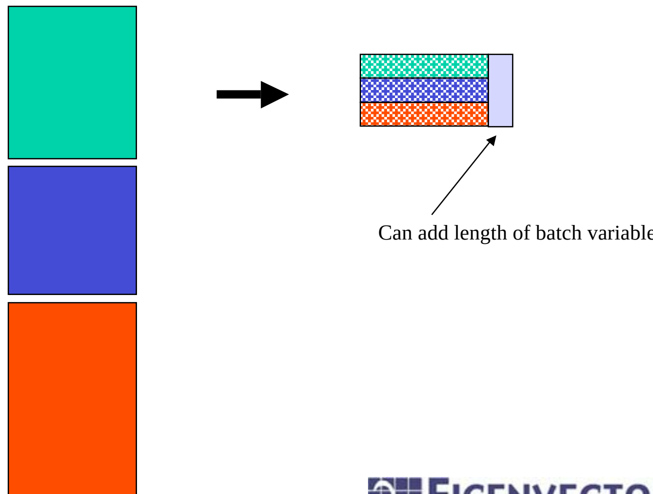
EIGENVECTOR RESEARCH INCORPORATED 9

Mean of Each Variable over Wafer Processing Time

- Take mean of each variable
- Pros:
  - Easy, conceptually simple
  - Large data reduction
  - No problem with variable record length
- Cons:
  - May miss some types of fault
  - Completely lose time information

EIGENVECTOR RESEARCH INCORPORATED 10

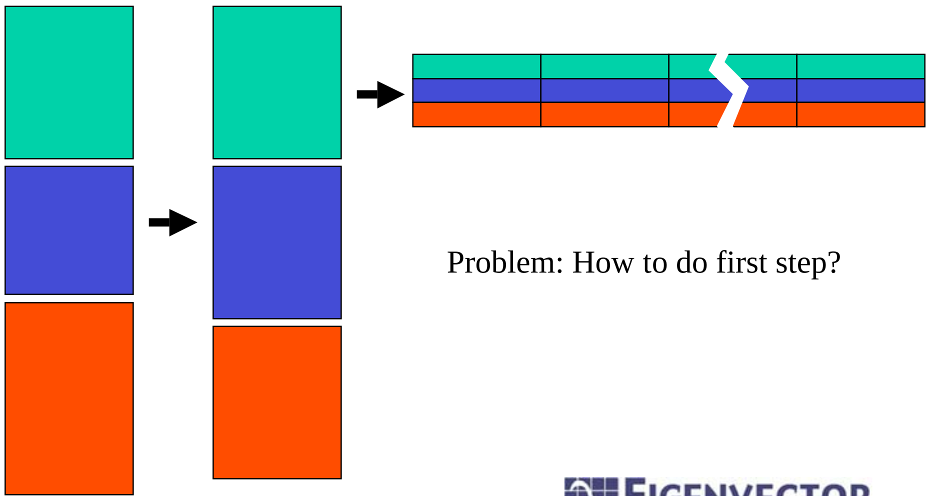
Mean Procedure



Arrangement of Data for Multiway Analysis

- Records must be the same length (except PARAFAC2)
- Requires some combination of alignment, truncation or warping
- Pros
  - Retains time information
- Cons
  - Results in many variables (parsimony problem)
  - Complex preprocessing step, requires interpolation
  - Model interpretation issues

Alignment for Multi-way



Problem: How to do first step?

EIGENVECTOR RESEARCH INCORPORATED 13

Alignment and Warping Methods

- A veritable smorgasbord of methods available
  - Dynamic Time Warping (DTW)
  - Correlation Optimized Warping (COW)
  - Indicator variable/step number
  - Linear interpolation
  - Align and truncate
  - Combinations and variations of the above
  - Etc.

EIGENVECTOR RESEARCH INCORPORATED 14

Correlation Optimized Warping

- Piecewise preprocessing method
- Allows limited changes in segment lengths, controlled by slack parameter
- Linear interpolation over segments
- Dynamic programming used to optimize correlation between warped sample and reference
- Less flexible than DTW (unless constrained)

EIGENVECTOR RESEARCH INCORPORATED 15

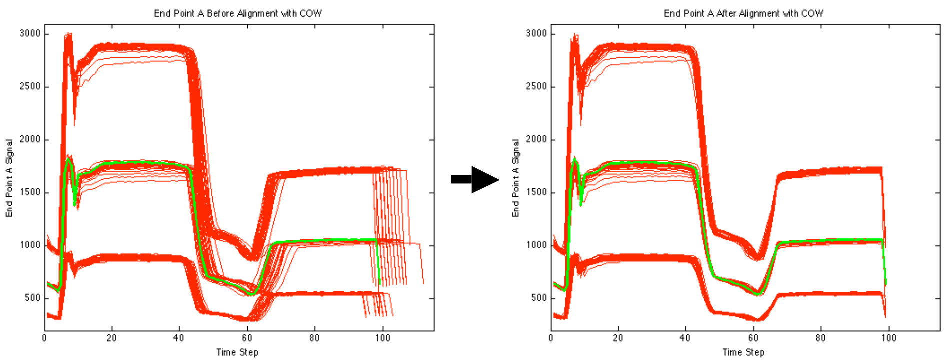
COW References

- N.P.V Nielsen, J.M. Carstensen and J. Smedsgaard, "Aligning of single and multiple wavelength chromatographic profiles for chemometric data analysis using correlation optimized warping," *J. Chromatogr. A*, **805**, 17-35, 1998.
- G. Tomasi, F. van den Berg and C. Andersson, "Correlation Optimized Warping and Dynamic Time Warping as Preprocessing Methods for Chromatographic Data," *J. Chemometrics*, **18**, 231-241, 2004.
- G. Tomasi, T. Skov and F. van den Berg, Warping Toolbox, see: http://www.models.life.ku.dk/source/DTW_COW/index.asp

EIGENVECTOR RESEARCH INCORPORATED 16

Example: COW

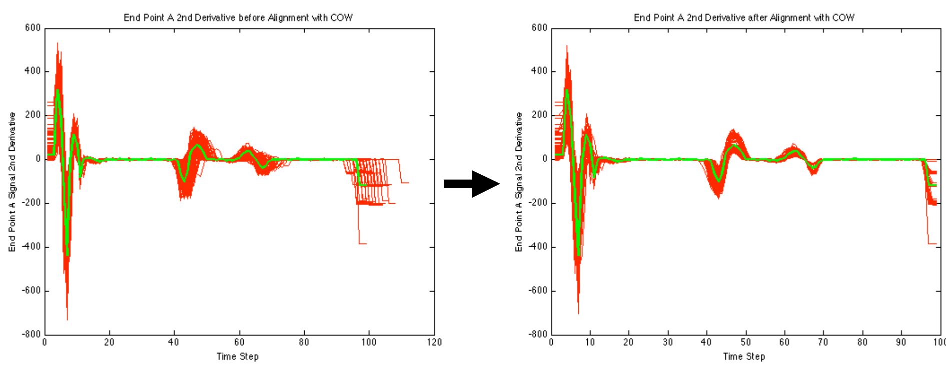
COW breaks signals into segments and linearly expands or contracts them to optimize correlation



EIGENVECTOR RESEARCH INCORPORATED 17

Hints on COW

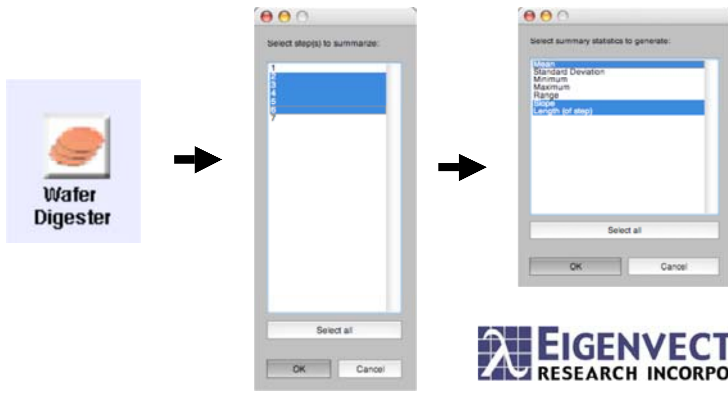
May be better to calculate warp with 2nd derivative
Apply calculated warp to other variables
Calculate warp on PCA scores or other latent variable



EIGENVECTOR RESEARCH INCORPORATED 18

Data Summary Approach

- Convert data into alternate set of descriptors
- If process has multiple steps, calculate parameters that describe each step



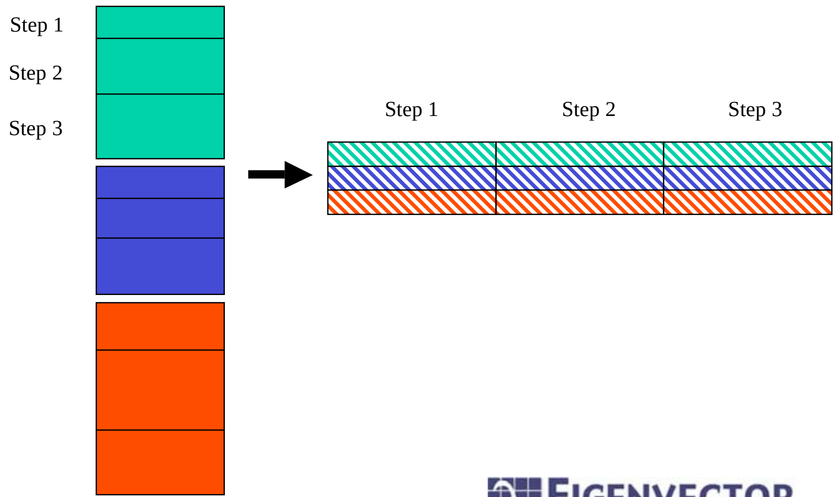
EIGENVECTOR RESEARCH INCORPORATED 19

Summary Variables

- Pros
  - Conceptually simple
  - Some time information retained
  - Noise reduction
  - Reduces number of variables (vs. MPCA)
- Cons
  - Further from original data
  - May not have step numbers to work with

EIGENVECTOR RESEARCH INCORPORATED 20

Summary Variables



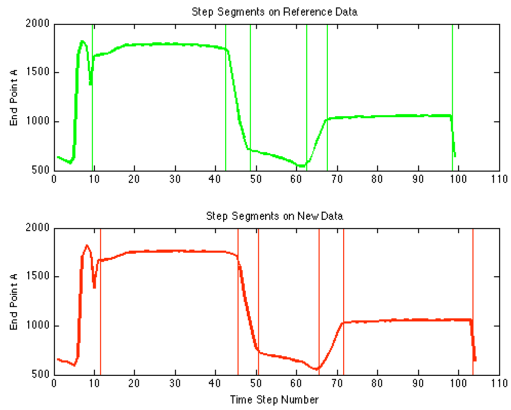
EIGENVECTOR RESEARCH INCORPORATED 21

Creation of Pseudo-steps

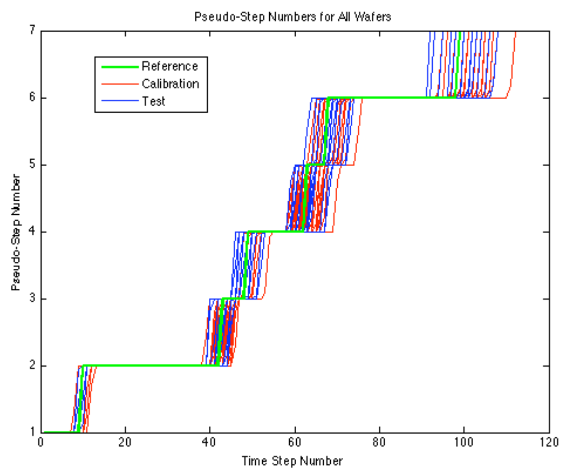
- Break reference process trace into "sensible" segments (manually)
- Assign step numbers
- Warp new data onto reference
- Reverse warp reference step numbers into new data
- Use summary variables

EIGENVECTOR RESEARCH INCORPORATED 22

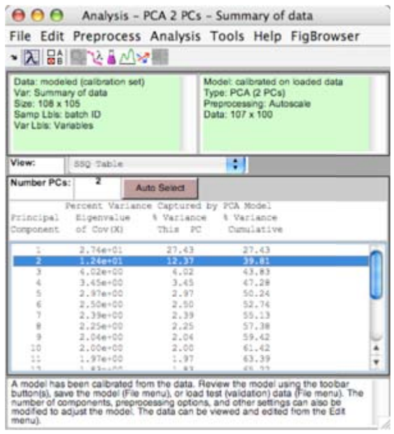
Example of Step Creation

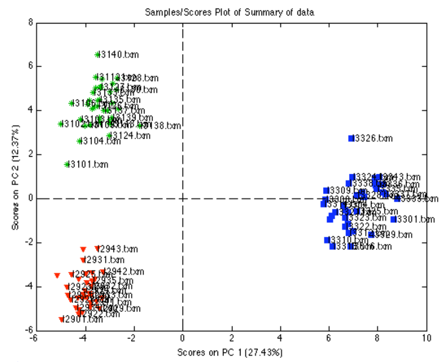


Variation in Step Length



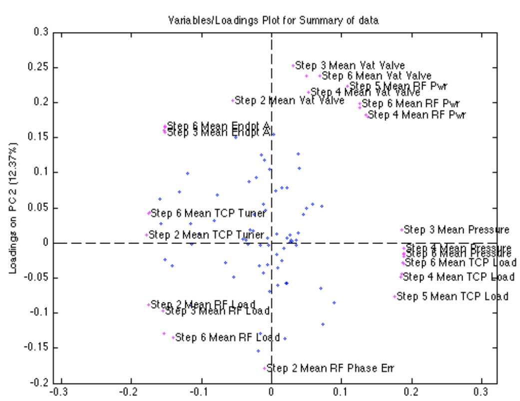
PCA on Summary Variables





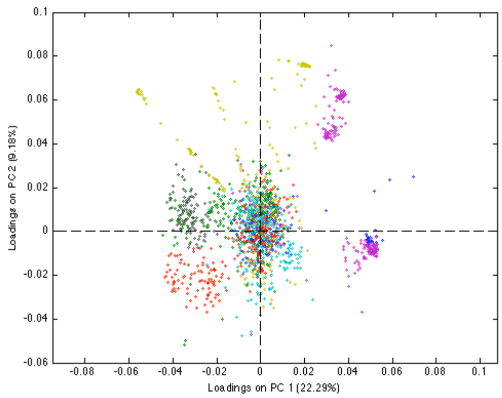
Decluttered

Loadings for PCA Model



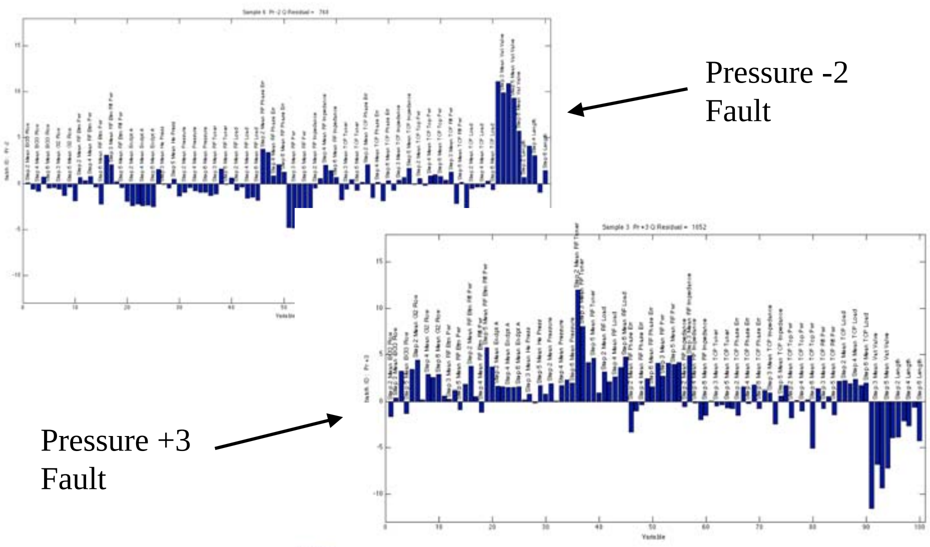
Decluttered

Compare to MPCA Loadings



EIGENVECTOR RESEARCH INCORPORATED 27

Contribution Plots



Pressure -2 Fault

Pressure +3 Fault

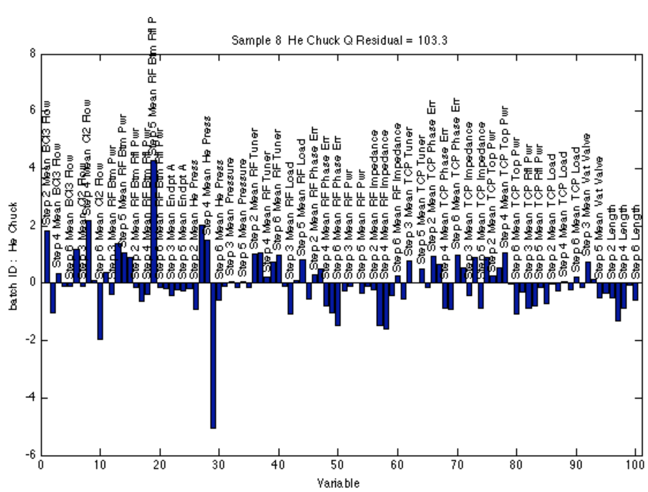
EIGENVECTOR RESEARCH INCORPORATED 28

Results

Faults Caught by PCA on Summary Variables Compared with TLD, PARAFAC, MPCA, PCA on the Means and PARAFAC2.

	TLD	PARAFAC	MPCA	PCA/Means	PARAFAC2	Summary
Global	11	12	10	10	7	**14**
Local	14	17	11	16	13	**18**

He Pressure Fault



Summary

- Digestion into summary variables creates smaller, more manageable and interpretable data sets
- Segments can be based on process steps
- Warping techniques such as COW can be used to create pseudo-steps
- Models based on summary variables at least as sensitive as MPCA but easier to work with

EIGENVECTOR RESEARCH INCORPORATED 31