Emerging trends in Statistical Process Control of Industrial Processes

Marco S. Reis

Guimarães, July 15th, 2016

Chemical Process Engineering and Forest Products Research Center – CIEPQPF

Department of Chemical Engineering
University of Coimbra

Outline

1. Introduction & Background on SPC
2. Current Trends & Examples
   - From univariate, to multivariate, to high-dimensional ("mega-variate")
   - From monitoring the mean, to dispersion, to correlation
   - From stationary, to dynamic, to non-stationary
   - From sensor data to higher-order profiles
   - From detection, to diagnosis, to prognosis
3. Conclusions
1. Introduction & Background

Statistical Process Control (SPC)

≈

(Statistical) Process Monitoring (SPM)

≈

Fault Detection and Diagnosis (FDD)

≈

(...)

• Statistical Process Monitoring (SPM): Goal
  – Verify if process behaviour is consistent with normal operating conditions.
    • Detection: rapidly detect abnormalities in process operation
    • Diagnosis: look for the root cause of abnormal behaviour
    • Fault criticality assessment
    • Decision: stop the process and fix the problem or accommodate the fault and proceed

Rapidly detect and act on abnormalities in process operation.
1. Introduction & Background

Walter Shewhart first control chart: Western Electric (Hawthorne Works, Chicago, 1924)

The beginning...
1. Introduction & Background

• Benefits from SPM
  – Increase safety of people
  – Protect critical industrial assets
  – Increase process efficiency: ↓ out-of-spec product, ↓ scrap, ...
  – Improve quality: ↑ product consistency, ↓ defects
  – Improve economic results
  – Reduce environmental impact
  – ...

Goal

Present an overview of some of the main trends on SPM over the last 90+ years

• From univariate, to multivariate, to high-dimensional
• From monitoring the mean, to dispersion, to correlation
• From stationary, to dynamic, to non-stationary
• From sensor data to higher-order profiles
• From detection, to diagnosis, to prognosis
Topics

• From univariate, to multivariate, to high-dimensional

• From monitoring the mean, to dispersion, to correlation

• From stationary, to dynamic, to non-stationary

• From sensor data to higher-order profiles

• From detection, to diagnosis, to prognosis

From univariate, to multivariate, to high-dimensional ("megavariate")
Hotelling’s $T^2$ (1931)

$$T_i^2 = n \left( \bar{x}_i - \bar{x} \right)^T S^{-1} \left( \bar{x}_i - \bar{x} \right)$$

H. Hotelling

PCA
The diagram shows a principal component analysis (PCA) plot with two principal components, PC1 and PC2. The data points are projected onto the plane defined by these components. The table below provides the principal components analysis:

<table>
<thead>
<tr>
<th>Number</th>
<th>Eigenvalue</th>
<th>Percent</th>
<th>20</th>
<th>40</th>
<th>60</th>
<th>80</th>
<th>Cum Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2.8235</td>
<td>94.115</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>94.115</td>
</tr>
<tr>
<td>2</td>
<td>0.1406</td>
<td>4.687</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>98.802</td>
</tr>
<tr>
<td>3</td>
<td>0.0359</td>
<td>1.198</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>100.000</td>
</tr>
</tbody>
</table>

The equation shown in the image is:

\[ X = PC1 + PC2 + E \]
PCA-MSPC (1959/1991)

Example:
Megavariate statistical process control in electronic devices assembling
(M. Reis; P. Delgado)

- Solder Paste Deposits (SPD’s) are of critical importance, because:
  - They provide the necessary fixation for all the electronic components
  - Functionalize the operation of electronic components

- Different shapes
- Different positions
- ...

The problem

- 100% inspection of Printed Circuit Boards (PCB’s).
- Each PCB has more than 3000 deposits (SPD’s) of different shapes.
- Operators have less than 1 min to decide about the status of each PCB.
- Each solder deposit is evaluated according to 5 parameters obtained through Moiré interferometry
  - Volume (V)
  - Area (A)
  - Height (H)
  - Offset in the X coordinate (X)
  - Offset in the Y coordinate (Y)

> 15 000 measurements for each PCB!

**FIRST LEVEL OF DETECTION**

**Combined approach**

**Multivariate Statistical Process Control using Principal Components Analysis (PCA-MSPC*)**

\[
\begin{aligned}
V & \xrightarrow{\text{PCA}_V} T^2_V & Q_V \\
H & \xrightarrow{\text{PCA}_H} T^2_H & Q_H \\
A & \xrightarrow{\text{PCA}_A} T^2_A & Q_A \\
X & \xrightarrow{\text{PCA}_X} T^2_X & Q_X \\
Y & \xrightarrow{\text{PCA}_Y} T^2_Y & Q_Y \\
\end{aligned}
\]

\[
\begin{aligned}
\text{PCA}_C & \xrightarrow{T^2_C} Q_C \\
\end{aligned}
\]

SECOND LEVEL OF DETECTION

- Analysis of the residuals from the projection of each multivariate observation to the PCA subspace.

![Residuals for observation 4](image)

RESULTS FOR THE FIRST LEVEL OF DETECTION

<table>
<thead>
<tr>
<th>Detection Statistics</th>
<th>Measurements used to compute the relative area under the ROC curve (values in %)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Height (H)</td>
</tr>
<tr>
<td>$T^2$</td>
<td>70.00</td>
</tr>
<tr>
<td>$Q$</td>
<td>93.13</td>
</tr>
</tbody>
</table>

10 PCB's classified as "good" were used to represent NOC data in SPC (estimate the PCA subspace, ...)

16 PCB's classified as "fail" (16) and "good" (5) were used to test the procedure
RESULTS FOR THE SECOND LEVEL OF DETECTION

<table>
<thead>
<tr>
<th>Detection Statistics</th>
<th>Measurements used to identify abnormal SPD’s (values in %)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Height (H)</td>
<td>Area (A)</td>
</tr>
<tr>
<td>Mean</td>
<td>80.33</td>
<td>65.82</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>20.07</td>
<td>29.97</td>
</tr>
</tbody>
</table>


From monitoring the mean, to dispersion, to correlation
Monitoring the mean and dispersion

• Most MSPC schemes are focused on location (mean level);
  – Shewhart
  – EWMA
  – CUSUM
  – Hotelling’s $T^2$
  – PCA-MSPC

Monitoring the correlation

• Most industrial processes are heavily controlled
  – Feedback control loops (PID)
  – Cascade control
  – Model predictive control
  – …
• When a fault arises, controllers fight to keep the mean levels on track:
  – Faults are “masked” by the controller actions!
  – But the correlation structure of the process variable changes!
SPM based on partial correlations

- Marginal correlations are unable to discern between direct and indirect associations between variables
- Partial correlations offer a better description of the process Normal Operating Conditions (NOC) network structure

$$\text{Corr}(A, B) \uparrow \uparrow \text{ but } \text{R}(A, B | Z) \downarrow$$

$$r_{AZ} = 0.8 \quad r_{AZ \cdot B} = 0.7295$$
$$r_{BZ} = 0.6 \quad r_{BZ \cdot A} = 0.4104$$
$$r_{AB} = 0.48 \quad r_{AB | Z} = 0$$

$$\text{RMAX} = \max \left\{ |w(r)| \right\}$$
$$\text{VnMAX} = \max \left\{ |w_s(v)| \right\}$$
SPM based on partial correlations

- PCA performs rather poorly in detecting localized changes in correlation
- A conventional approach, based on the marginal covariance (W), performs much better
From stationary, to dynamic, to non-stationary
PCA – assumes i.i.d. observations

<table>
<thead>
<tr>
<th>X1</th>
<th>X2</th>
<th>X3</th>
<th>X4</th>
<th>Xm</th>
</tr>
</thead>
<tbody>
<tr>
<td>X1(0)</td>
<td>X2(0)</td>
<td>X3(0)</td>
<td>X4(0)</td>
<td>Xm(0)</td>
</tr>
<tr>
<td>X1(1)</td>
<td>X2(1)</td>
<td>X3(1)</td>
<td>X4(1)</td>
<td>Xm(1)</td>
</tr>
<tr>
<td>X1(2)</td>
<td>X2(2)</td>
<td>X3(2)</td>
<td>X4(2)</td>
<td>Xm(2)</td>
</tr>
<tr>
<td>X1(3)</td>
<td>X2(3)</td>
<td>X3(3)</td>
<td>X4(3)</td>
<td>Xm(3)</td>
</tr>
<tr>
<td>X1(4)</td>
<td>X2(4)</td>
<td>X3(4)</td>
<td>X4(4)</td>
<td>Xm(4)</td>
</tr>
<tr>
<td>X1(5)</td>
<td>X2(5)</td>
<td>X3(5)</td>
<td>X4(5)</td>
<td>Xm(5)</td>
</tr>
<tr>
<td>X1(6)</td>
<td>X2(6)</td>
<td>X3(6)</td>
<td>X4(6)</td>
<td>Xm(6)</td>
</tr>
<tr>
<td>X1(7)</td>
<td>X2(7)</td>
<td>X3(7)</td>
<td>X4(7)</td>
<td>Xm(7)</td>
</tr>
<tr>
<td>X1(8)</td>
<td>X2(8)</td>
<td>X3(8)</td>
<td>X4(8)</td>
<td>Xm(8)</td>
</tr>
<tr>
<td>X1(9)</td>
<td>X2(9)</td>
<td>X3(9)</td>
<td>X4(9)</td>
<td>Xm(9)</td>
</tr>
</tbody>
</table>

Dynamic PCA – Ku et al. (1995)

<table>
<thead>
<tr>
<th>X1 (L=1)</th>
<th>X1 (L=2)</th>
<th>X2 (L=1)</th>
<th>X2 (L=2)</th>
<th>Xm (L=1)</th>
<th>Xm (L=2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>X1(0)</td>
<td>X1(0)</td>
<td>X2(0)</td>
<td>X2(0)</td>
<td>Xm(0)</td>
<td>Xm(0)</td>
</tr>
<tr>
<td>X1(1)</td>
<td>X1(1)</td>
<td>X2(1)</td>
<td>X2(1)</td>
<td>Xm(1)</td>
<td>Xm(1)</td>
</tr>
<tr>
<td>X1(2)</td>
<td>X1(2)</td>
<td>X2(2)</td>
<td>X2(2)</td>
<td>Xm(2)</td>
<td>Xm(2)</td>
</tr>
<tr>
<td>X1(3)</td>
<td>X1(3)</td>
<td>X2(3)</td>
<td>X2(3)</td>
<td>Xm(3)</td>
<td>Xm(3)</td>
</tr>
<tr>
<td>X1(4)</td>
<td>X1(4)</td>
<td>X2(4)</td>
<td>X2(4)</td>
<td>Xm(4)</td>
<td>Xm(4)</td>
</tr>
<tr>
<td>X1(5)</td>
<td>X1(5)</td>
<td>X2(5)</td>
<td>X2(5)</td>
<td>Xm(5)</td>
<td>Xm(5)</td>
</tr>
<tr>
<td>X1(6)</td>
<td>X1(6)</td>
<td>X2(6)</td>
<td>X2(6)</td>
<td>Xm(6)</td>
<td>Xm(6)</td>
</tr>
<tr>
<td>X1(7)</td>
<td>X1(7)</td>
<td>X2(7)</td>
<td>X2(7)</td>
<td>Xm(7)</td>
<td>Xm(7)</td>
</tr>
<tr>
<td>X1(8)</td>
<td>X1(8)</td>
<td>X2(8)</td>
<td>X2(8)</td>
<td>Xm(8)</td>
<td>Xm(8)</td>
</tr>
<tr>
<td>X1(9)</td>
<td>X1(9)</td>
<td>X2(9)</td>
<td>X2(9)</td>
<td>Xm(9)</td>
<td>Xm(9)</td>
</tr>
</tbody>
</table>
Dynamic Principal Components Analysis (DPCA)

- However, the DPCA scores still present autocorrelation ...

\[ X = [X(k) \ X(k-1) \ \ldots \ X(k-l)]; \]

\[ x = \begin{bmatrix} x^\# \\ x^* \end{bmatrix}, \quad S = PA^TP, \quad P = \begin{bmatrix} P^h \\ P^* \end{bmatrix} \]

\[ \hat{t}_{l:k} = \begin{bmatrix} I_k & 0_{k \times (n-k)} \end{bmatrix} A^*P^T \left( P^*A^*P^T \right)^{-1} x^* \]

Scores

\[ T^2_{PREV} = (t - \hat{t})^T S_{t-i}^{-1} (t - \hat{t}) \]

Residuals

\[ T^2_{RES} = (x - \hat{P}t)^T S_i^{-1} (x - \hat{P}t) \]
Dynamic DPCA-DR

- DPCA-MD scores for the same system present a significantly lower level of autocorrelation.

Figure. Sample autocorrelation function for $T^2_{\text{Prev}}$ and $T^2_{\text{Res}}$ statistics.
From sensor data to higher-order profiles

Monitoring Profiles (1D, 2D, 3D, ...)

“(…) We view the monitoring of process and product profiles as the most promising area of research in statistical process control. (…)”

• **Definition [Profile, \( P \)]:**

An array of data, indexed by time and/or space, that characterizes a given entity (product, process).

\[
P: \left\{ Y \left( i_x, i_y, i_z, i_t \right) \right\}_{i_x, i_y, i_z, i_t \in \Omega_x \times \Omega_y \times \Omega_z \times \Omega_t} \quad \{ Y \}_{i_x, i_y, i_z, i_t} \in \mathbb{R}^n
\]

---

**“Dual domain” classification**

- **Frequency Localized Random Profiles**
- **Fully Localized Profiles**
  - More:
    - Acoustic signals
    - Seismic signals
  - “Hard” constraint:
    \[ \sigma(g) \cdot \sigma(\hat{g}) \geq \frac{1}{2} \]

- **Delocalized Random Profiles**
- **Time-Space Localized Random Profiles**
  - More:
    - Analysis of Coherent Images
    - Perfusion experiments
    - fMRI, …
Monitoring paper formation*
(Reis, MS & Bauer, A.)

- Currently is evaluated off-line: few times per day (e.g. after each paper reel production)
  - Very high delay, regarding the production speed of current paper machines (~100 Km/h!)

* Level of uniformity in the way fibres are distributed across the paper surface.

Goal

Develop a technology for on-line monitoring of the paper formation.
Experimental: Measurement Sensor

- Digital camera
- Housing with a rotating head
- Moving web
- Light source

Image set: 24 images, representing different levels of formation quality.
Methods

Process Monitoring

Grade evaluation

Status of the process (normal / abnormal)

Quality Grade

Results (RQ1)

- PCA analysis of wavelet signatures

2PC’s – 97.96% of the overall variability
Results (RQ1)

- Prototypes of clusters 2A and 2B

2A – heterogeneity through smaller and more frequent irregularities.

2B – “cloudy” texture.

Results (RQ2)

Analysis 2 (sub-images)

Note: Half of the Grade 1 samples were used to estimate the NOC region.
From detection, to diagnosis, to prognosis

Detection

• Basic requirement: a good description of the Normal Operating Conditions
  – Mean levels and main correlations between variables
  – Non-causal associations
Diagnosis

• To find the root cause, causal information is needed!
• However, most SPM models are acausal, and therefore cannot provide all the information required for a thorough diagnosis
• They may also point to variables that are not directly involved in the fault
  – The smearing-out effect of PCA-MSPC is a well-known manifestation

One solution

• Plug-in causality into conventional NOC models through an adequate pre-processing of the variables

• This can be done using the concept of Sensitivity Enhancing Transformations (SET)
Sensitivity enhancing transformation (SET)

1. Network Identification
   \[ x_1 \rightarrow x_3 \]
   \[ x_2 \rightarrow x_3 \]
   \[ x_3 \rightarrow x_4 \]

2. Regress each variable onto its parents
   \[ y_1 = x_1 \]
   \[ y_2 = x_2 \]
   \[ y_3 = x_3 - x_1 b_{13} - x_2 b_{23} \]
   \[ y_4 = x_4 - x_3 b_{34} \]

3. Final model
   \[ Y = XB \]
   \[ B = \begin{bmatrix} 1 & 0 & -b_{13} & 0 \\ 0 & 1 & -b_{23} & 0 \\ 0 & 0 & 1 & -b_{34} \\ 0 & 0 & 0 & 1 \end{bmatrix} \]

4. Apply the Cholesky decomposition to the regression residuals thus obtained.

Plug-in approach:

1. Process Data
2. SET
3. Uncorrelated data
4. Monitoring statistics (PC, T²)
5. Statistics
6. MSPC
7. Process State (Normal / Abnormal)


**Figure.** Percentage of times that each variable was considered as the faults’ root cause by the marginal correlation procedure on a fault in the relationship between variables 1 and 8: (a) marginal correlation of the original variables; (b) marginal correlation of the transformed variables.
SPM in the big data era

Typology of SPM applications
Conclusions

- 90+ years after its introduction, SPC is still an exciting and evolving field!
- SPC should be complemented with effective Diagnosis tools
- New challenges include
  - Move focus from Detection to Diagnosis
  - Handling complex dynamics: multiscale methods
  - Integrating the structure of the system and existing domain knowledge: SET, Bayesian methods
  - Handling multiple data structures (profiles): multi-block methods
  - Monitoring time-varying systems: adaptive methods, ...

and ... making everything simple to use and robust!

Acknowledgements

Pedro Saraiva
Pedro Delgado
Tiago Rato
Ricardo Rendall
Maria Campos
Veronique Medeiros

http://www.eq.uc.pt/~marco/research/pclab/