Shewhart and Profile Monitoring for Industry 4.0
Geoff Vining
Virginia Tech

Abstract:

An important current area within statistical process control/monitoring is profile monitoring, which assumes that the underlying profile of the data over time is some linear, nonlinear, or nonparametric model. Let y be the characteristic of interest, and let $f(y; \boldsymbol{\theta}, \boldsymbol{x})$ be the underlying model, where $\boldsymbol{x}$ is the $p \times 1$ vector of variables that explain the behavior of $\boldsymbol{y}$ over time and $\boldsymbol{\theta}$ is an unknown vector of model parameters relating $\boldsymbol{x}$ to $\mathbf{y}$. The standard approach taken by the profile monitoring community uses the following algorithm:

(1) Estimate $\boldsymbol{\theta}$ for each individual value of y. Let $\widehat{\boldsymbol{\theta}}_i$ be the resulting vector of estimates associated with each individual y_i. Let $\widehat{\boldsymbol{\theta}}_{avg}$ be the average value of the $\widehat{\boldsymbol{\theta}}_i$s.

(2) Estimate the variance of $\widehat{\boldsymbol{\theta}}_i$ by computing estimates of every variance and covariance involving the components of $\boldsymbol{x}$.

(3) Construct control limits using some variation of Hotelling's T^2 statistic.

This approach historically assumes that p is very small. There are many serious issues from a linear-models perspective to this approach, not the least of which is an unnecessary need to estimate $p + \binom{p}{2}$ variance components, which typically require much larger sample sizes to estimate than averages.

Industry 4.0 increases the number of sensors that can provide huge amounts of information in real time. As a result, there are opportunities to align data on those variables known to impact a critical quality characteristic to improve the monitoring of that characteristic. The case study that underlies this talk has very good information on 40 variables known to impact the performance of the critical quality characteristic. The current profile monitoring approach requires the estimation of 780 variances/covariances, which is completely unrealistic.

This talk outlines how to incorporate the extra information efficiently and effectively. Ironically, this approach has its origins in Shewhart's original ideas underlying control charts. Seeing the connection is important for advances in statistical process monitoring within Industry 4.0.