

Multivariate Fault Detection Method for Non-Gaussian Distributed Implant Data

Authors: John Mao, Patrick Holland

When we applied Principal Component Analysis (PCA) based fault detection and classification (FDC) to implant data, we had to deal with non-Gaussian distributed data. This paper discusses a new method to perform FDC on non-Gaussian data with correct control limits. We applied the method to implanter monitoring. The implanter equipment data are highly non-Gaussian distributed data. The method is well fit for the case.

PCA-based multivariate fault detection and classification methods require that data are close to a multivariate Gaussian distribution. When using T2 and SPE (Squared Predictive Error) charts for multivariate fault detection, it is a problem that the multivariate data is often far different from a Gaussian distribution. Therefore, in PCA-based multivariate T2 and SPE charts, the control limits are not correct because control limits are computed assuming that they follow a multivariate Gaussian distribution.

To deal with non-Gaussian data, a commonly used method is to transform each variable into a normally distributed variable. However, the transformation function is usually unknown. Also, when the data change to a different pattern, the transformation function should be changed. Therefore, the transformation approach is hard to use practically.

In this paper, we present a new density function estimation method to solve this problem. The method estimates the probability density function using kernel density function. By estimating the density function, we can compute the control limits of multivariate T2 and SPE charts. We used non-Gaussian equipment data to compare the method and the original method. The resulting control limits are much more useful to detect faults.

We applied the method to simulated Gaussian and actual implant process data. The study verified the correctness and usability of the method. We found that this method creates the control limits close to the original method (which assumes a Gaussian distribution) for Gaussian data. Also this new method created much more reasonable control limits for non-Gaussian distributed implant process data.

We demonstrated how the non-Gaussian method for calculating control limits detects problems for a large pool of monitors generated by the Implanters during run-to-run batch mode processing. This method of calculating the limits is practically applied to an offline monitoring system in which T2 and SPE multivariate control charts utilize constantly updating models as new data are added. The use of filter limits on the univariate parameters is an efficient method of tuning the multivariate model for a practical application. We also demonstrated how the multivariate method with non-Gaussian control limits easily lends itself to FDC in a high volume production environment.