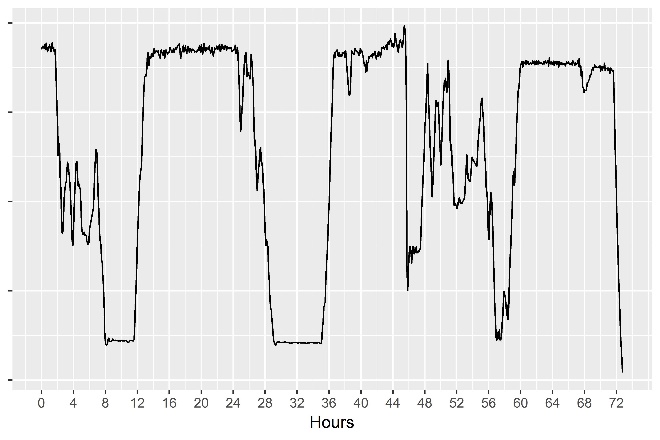
Introduction to PCA for Process Monitoring   
within the Griffin Toolkit

Jake Tuttle

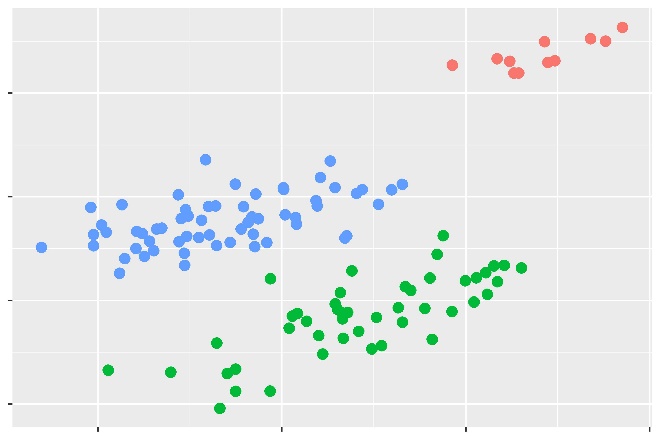
Principal Component Analysis (PCA) is a powerful statistical method used primarily for dimensionality reduction of large and highly complex correlated datasets. This method of dimensionality reduction creates new “components” from original factors within the dataset which maximize the variation present within the entirety of the original dataset.

Below we work through the analysis of a common problem at a coal-fired power plant – a tube leak. The load trend of the unit is displayed below 72 hours before the shutdown to repair the tube leak.



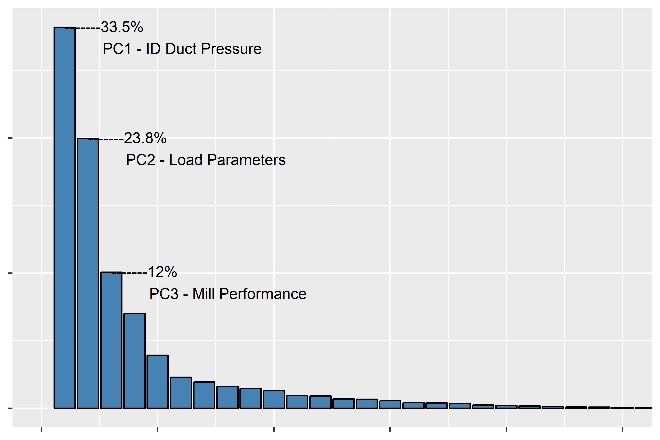
Up until coming offline, the load profile and most operating parameters appear relatively normal. The application of PCA Process Monitoring by the Griffin Toolkit tells a different story.

By representing hundreds of variables in an easily viewable and understood 2D plot, complex relationships and intercorrelations can be comprehended in seconds. When PCA is used, changes or variations in the data tend to for subgroups within the full dataset.

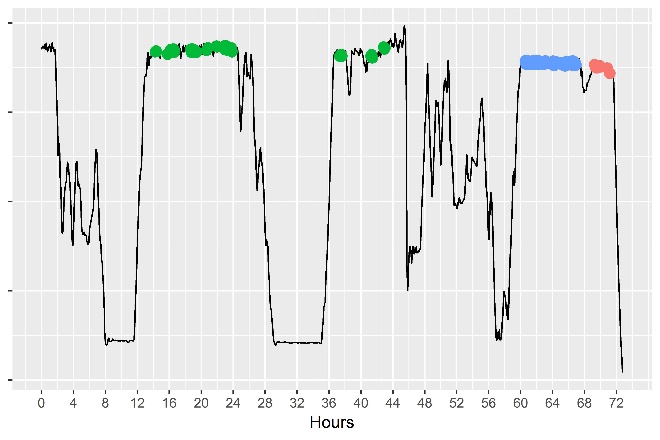


The image above shows the PCA results of monitoring full load operation of the station during the time period presented. There are clearly 3 different subgroups of operation present, as shown by the PCA results.

The makeup of each principal component (PC) tells us what factors are contributing to the x- and y-dimensions shown in the results. A scree plot helps show the percent of the original dataset represented by each PC, as well as its main contributing factors.



As seen above, nearly 70% of the original dataset is represented by the first three PCs, and over half by just the first two.



Finally, by observing when Griffin Toolkit PCA identifies the underlying subgroups of performance, advanced detection of a trouble event is identified, earlier than many traditional methods.

The green group shown on the MW trends represents business-as-usual. After a volatile load movement, the return to full load operation demonstrates atypical performance, which is identified by PCA at the 60-hour mark, almost immediately upon returning to full load. At nearly the 68-hour mark, the event likely occurs, causing the load swing. Upon recovering, PCA identifies the red subgroup, and clearly a problem event has occurred, followed quickly by unit shutdown.

