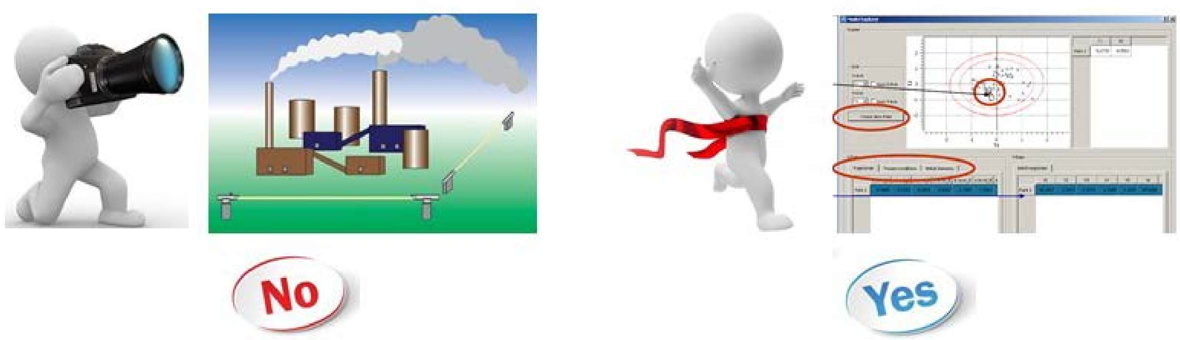
Statistical Process Monitoring with Advanced Data Analytics Tools

Dr. Fengqi You, Roxanne E. and Michael J. Zak Professor
Process-Energy-Environmental Systems Engineering
fengq.you@cornell.edu | you.cbe.cornell.edu

In industrial processes, it is crucial to detect and diagnose faults, process upsets, and other abnormal events to achieve safe and efficient operations. In the last two decades, machine learning and multivariate statistical process monitoring approaches have been extensively studied and implemented in the process control community. In general, they resort to massive process data measured in industrial processes rather than first-principle knowledge. Typical data-driven approaches such as principal component analysis (PCA) can be applied to derive latent variable subspaces and calculate monitoring statistics. They provide reduced dimensional subspaces to describe the essential sources of variations, and allow direct visualization and monitoring statistics design.



In this project, we intend to adopt recent advances in machine learning to analyze industrial big data and develop novel process monitoring schemes. Specifically, we will adopt the concept of Wasserstein metric to formulate the PCA model, which has received considerable attentions recently in the machine learning literature as an advanced data analytics tool. By virtue of the Wasserstein metric, the probability distribution of process data can be well described, with desirable robustness to potential outliers. Monitoring statistics can be defined on the derived low-dimensional subspace. Both numerical examples and the Tennessee Eastman benchmark process will be used to testify the effectiveness of the new data-driven approach.

This project will be helpful for enhancing one's knowledge and programming skills in machine learning and industrial data processing.