320d On-Line and off-Line Fault Detection and Diagnosis in Fed-Batch Fermentation

Jon C. Gunther and Dale E. Seborg

The biotechnology industry stands to benefit significantly from improved batch monitoring techniques, because most high-valued products are manufactured using batch and fed-batch processes. Detection of process abnormalities and fault diagnosis increases the likelihood of avoiding these problems in the future and improving production consistency, which is a major facet of the FDA's new Process Analytical Technology (PAT) industry initiative.

One powerful monitoring technique, multivariate statistical process control (MSPC), capitalizes upon the structured information captured in historical process data. MSPC control charts, based upon principal component analysis (PCA) models, can be developed that identify abnormal batches and process faults. MSPC can be implemented in either an on-line or off-line mode. In particular, MSPC contribution plots can be used to determine which process variables are most affected by a process fault; subsequent investigation by plant personnel can result in fault remediation. However, MSPC charts generated for on-line and off-line monitoring are not always in agreement when detecting faults. For example, while on-line charts might detect a serious fault occurring at some point in the duration of the batch, off-line evaluation may deem the same batch to be within model confidence limits and thus normal.

To explore this and other issues, alternative monitoring strategies have been evaluated in simulation studies and an industrial application. A fed-batch fermentation model was developed in cooperation with Lund Institute of Technology in Sweden. This model provides close agreement with their experimental pilot-plant data. For monitoring purposes, data for 50 calibration batches, subject to normal variation, were generated from the simulation. A PCA model was constructed from the calibration data and validated using ten additional batches. Three MSPC metrics used: the Hotelling T2 statistic (on-line and off-line), the sum of squared residuals Q (off-line), and squared prediction error SPE (on-line). To test the fault detection capabilities of the PCA model and MSPC charts, a series of 11 process faults was simulated. An extensive sensitivity analysis was performed to determine the detection thresholds for each fault and for each MSPC metric. Furthermore, SPE contribution plots were developed to diagnose each specific fault.

Practical monitoring strategies must be both flexible and robust in order to accommodate the many different operating conditions that occur in industrial applications. For example, in pilot plants where new processes are being developed, the data captured from these experiments can be quite variable from batch to batch. Two inherent limitations of conventional process monitoring strategies are the assumptions that all batches being evaluated have equal duration and are synchronized in time. However, it is common in industrial practice to have a set of batches where the total duration of the batches, and the duration of individual process stages, are not constant. For example, the time at which the initiation of the fed-batch phase, protein production induction, and the end of the batch are achieved for E. coli fermentation can be quite variable from batch to batch.

In partnership with Amgen, Inc., fermentation data from a set of pilot plant batches were analyzed using MSPC. The data provided were subject to the phenomena described in the previous paragraph. To improve the monitoring results for this data, synchronization of batch trajectories was performed. To align the batches, a priori operating knowledge was used, specifically when the onset of each new operating phase occurred. Initial monitoring results are quite promising, indicating the successful identification of bad batches can be achieved while avoiding false positives.