Literature Review Basics:

Should I be including details on the actual NC plant in this document? Diagrams…inputs/outputs etc.

How many literature sources should I review? I have these (4+1book) that cover most of the literature. I can of course include more.

1. **On-line batch process monitoring using batch dynamic kernel principal component analysis**
2. **Batch process monitoring based on support vector data description method**

|  |  |
| --- | --- |
| *Objective* | Differentiate the normal data samples from the faulty ones to create an efficient one-class classification method for batch process monitoring. |
| *Approach* | Novel method called support vector data description (SVDD).  Does not require the Gaussian assumption of the process data and effective for nonlinear process modelling unlike PCA and PLS. SVDD only incorporates a quadratic optimization step for practical implementation.  The method is further extended to multiphase and multimode batch processes.  Two case studies are provided to evaluate the monitoring performance. |
| *Techniques Used* | SVDD (evolution of SVM/PCA/PLS) |
| *Computational Intensity* | Large datasets  Data pre-processing: Between each phase (offline phase and online phase monitoring), the data needs to be re-organized and/or processed. |
| *Data Used* | <dataset size>  <simulation data collected>  <frequency of data> |
| *Other Notes* | SVDD requires no Gaussian assumption.  Addresses nonlinear relationships between process variables well.  By searching a hypersphere in the feature space, a tight boundary of the data distribution is formulated improving the monitoring sensitivity for process abnormalities. |
| *Results success* |  |

1. **An Introduction to Statistical Learning Gareth James Daniela Witten Trevor Hastie Robert Tibshirani with Applications in R**

Statistical process control (SPC) (e.g. PCA and PLS), PCA, partial least squares, SVM, non-linear SVM-based feature selection algorithm

1. **Reprint of: Big data approach to batch process monitoring: Simultaneous fault detection and diagnosis using nonlinear support vector machine-based feature selection**

|  |  |
| --- | --- |
| *Objective* | Develop a novel data-driven framework for process monitoring in batch processes offering simultaneous fault detection and diagnosis. |
| *Approach* | Use high dimensional process data with nonlinear SVM-based feature selection to retrieve the most informative process measurements for accurate and simultaneous fault detection and diagnosis.  Train fault and time-specific models on the pre-aligned batch data trajectories with three distinct time horizon approaches: one-step rolling, two-step rolling and evolving which varies the amount of data incorporation during modelling.  Application made to batch process data adopted from an extensive simulation data set for penicillin production. The data is expanded with sensor noise. |
| *Techniques Used* | Non-linear SVM, RFE-SVM (recursive feature elimination) classification algorithm.  Novel optimisation-backed feature selection algorithm  3D process data to 2D via batch-wise unfolding. |
| *Computational Intensity* | Large datasets  Data pre-processing: Between each phase (offline phase and online phase monitoring), the data needs to be re-organized and/or processed. |
| *Data Used* | Extensive benchmark data set which includes process data describing 22,200 batches with 15 faults.  <simulation data collected>  <frequency of data> |
| *Other Notes* | Big Data computation.  Data pre-processing into the 3 time horizons or sample bins.  Unfold 3-dimensional batch process data into 2-dimension bia batch-wise unfolding. |
| *Results success* | Two-step rolling and evolving time horizon approaches are superior.  Regardless, the proposed framework provides a promising decision support tool. |

1. **Multi-model based real-time final product quality control strategy for batch processes**

|  |  |
| --- | --- |
| *Objective* | Develop a novel real-time final product quality control strategy for batch operations. |
| *Approach* | Periodically predict the final product quality and adjust process variables at pre-specified decision points.  Employ multiple models, one for each decision point, to capture the time-varying relationships.  Acts on combination of real-time process data and historical data from prior batches.  Design of experiments is performed to generate informative data that reveal the relationship between process conditions and the final product quality at various times.  Control action is taken at pre-specified decision points – the MV are calculated by solving an optimal control problem similar to MPC.  Illustrated by example on a simulated batch reaction case study. |
| *Techniques Used* | Novel technique.  Combination of many techniques (e.g. MLR, PLS, multi-block and multi-way PLS [MBMWPLS], linear quadratic Gaussian regulator [LQG]) |
| *Computational Intensity* |  |
| *Data Used* | Combine real-time batch information (process variables, initial conditions) with information from prior batches.  <simulation data collected>  <frequency of data>  <volume of previous historical data> |
| *Other Notes* | Key benefit – missing data imputation is obviated.  There is no need to estimate missing data using this approach, reducing uncertainty from data imputation. |
| *Results success* |  |

1. **On-line batch process monitoring using dynamic PCA and dynamic PLS models**

|  |  |
| --- | --- |
| *Objective* | Develop an on-line batch process monitoring scheme based on multivariate statistical process control to overcome lack of on-line key quality measurements when using batch processes. |
| *Approach* | Multivariate statistical process control.  Compare BDPCA and BDPLS with traditional on-line PCA and MPLS algorithms for 3 different examples. |
| *Techniques Used* | BD-PCA/BD-PLS integrates time-lagged windows with the PCA and PLS for on-line batch process monitoring. |
| *Data Used* | Only data required to setup the control chart is the historical data collected from the past successful batches.  <simulation data collected>  <frequency of data>  <volume of previous historical data> |
| *Computational Intensity* |  |
| *Other Notes* | Using only historical data collected from past successful batches enables simple monitoring charts, easy tracking of progress and monitoring of observable upsets.  BDPCA and BDPLS only collect the previous data during the batch run without expensive computations to anticipate the future measurements. |
| *Results success* |  |