

DATA-DERIVED PROCESS MONITORING AND SUPERVISION  
SCUOLA DI DOTTORATO IN INGEGNERIA INDUSTRIALE  
POST-GRADUATE COURSE (JULY 16-20, 2012)

---

|                       |  |   |
|-----------------------|--|---|
| CONTACT INFO          | Teacher:   | Francesco Corona<br>Aalto University School of Science (Finland)<br>Dept. of Information and Computer Science |
|                       | Email:   | <a href="mailto:francesco.corona@aalto.fi">francesco.corona@aalto.fi</a>                                      |
|                       | Web:   | <a href="http://users.ics.tkk.fi/fcorona">users.ics.tkk.fi/fcorona</a>  |
| SCHEDULE AND LOCATION | Monday July 16 to Friday July 20, 2011. Lectures start at 9:15 AM and will finish around 1 PM. Dip. di Ingegneria Meccanica, Chimica e dei Materiali, Aula G. Mura.  |   |
| OVERVIEW              | Data-derived methods established themselves as a valuable alternative to the traditional means for the acquisition of critical process variables, process monitoring and other tasks related to process control. The focus is on data-derived soft-sensors because of their growing popularity, already demonstrated usefulness and huge, though yet not completely realised, potential. The course discusses the characteristics of industrial process data which are critical for the development of soft sensors. A comprehensive selection of methods for data preprocessing and of the most popular modeling techniques are the main topics of the lectures and tutorial sessions.  |   |
| SYLLABUS              | Process data and data-derived soft-sensors: <ul style="list-style-type: none"><li>• General characteristics of process data (Missing Data, Outliers, Drifting, Sampling and delays) and the anatomy of Soft-sensor development (Data inspection, Data pre-processing, Model design and Model maintenance).</li></ul> Methods for data preprocessing: <ul style="list-style-type: none"><li>• Sample selection: Multivariate selection based on Classical and Robust Principal Component Analysis, Measures of fit (Mahalanobis and Orthogonal distances and the <math>J</math>-statistic) and Outliers maps.</li><li>• Variable selection: Relevance criteria (Canonical correlation, Mutual information, Noise variance and Prediction error) and Search algorithms (Ranking, Exhaustive, Incremental and Randomised search).</li></ul> Methods for on-line prediction: <ul style="list-style-type: none"><li>• Linear methods: Least-squares estimates, Subset selection (Best-subset selection, Forward and Backward selection, Forward Stagewise regression), Shrinkage methods (Ridge Regression, Least Absolute Shrinkage Operator and Least Angle Regression) and methods on derived variables (Principal Components Regression and Partial Least Squares Regression).</li><li>• Nonlinear methods: Generalities and Locally linear regression using non-enclosing neighbourhoods (<math>k</math>-nearest neighbours and Correlation-based neighbours) and enclosing neighbourhoods (Enclosing nearest neighbours, Natural neighbours, Natural neighbours inclusive and Delaunay neighbours).</li></ul> Methods for process monitoring: <ul style="list-style-type: none"><li>• Continuous processes: Principal Component Analysis for Fault analysis (formulation, Solution and Quality control charts).</li><li>• Batch processes: Unfolding Principal Component Analysis and Multilevel Simultaneous Component Analysis.</li></ul> |   |
| PASSING THE COURSE    | To pass the course, students must actively participate to the lectures, and hand-in a coursework project in the format of conference article.<br>Further information and the course material will be given during the lectures.  |   |