Data-driven modeling for advanced process control: research directions and some examples

Riccardo Bacci di Capaci, Marco Vaccari, Gabriele Pannocchia

University of Pisa, Pisa, Italia

E-mail: *gabriele.pannocchia@unipi.it*

Advanced process control (APC) is a key element in the present digital era and is typically based on a reliable model of the system to be controlled. Among the possible types (first principles, data-driven, and hybrid), data-driven models leverage exclusively on historical data of the process. There are several milestones of designing a data-driven model: data collection, data (pre)processing, model structure selection, feature selection, model training and testing. The present work aims to give a brief overview of the recent activities of our research group at University of Pisa about data-driven modelling for APC. Research directions concerning system identification through data-driven models are for: i) PID retuning, ii) process monitoring, and iii) model predictive control (MPC), all by using our comprehensive open-source toolbox SIPPY.

As a first example of data-driven modeling for retuning of PID controllers, practical linear models are identified for electric power systems. The objective is to build reduced-order models that retain the most relevant modes of oscillations of the original large nonlinear systems. Input-output data are generated within a rigorous simulation software, and the corresponding models are identified and evaluated; different subspace methods are tested, and standard approaches for the selection of the best model order are investigated. Such reduced-order linear models are employed for the (re)tuning of different grid controllers as one of the main objectives of our recent EU Project DAEDALOS.

As an example for process monitoring, a multi-fidelity (MF) modeling approach is adopted to integrate dynamic information from online, low-fidelity (LF) data with infrequent, high-fidelity (HF) laboratory measurements. The proposed methodology is demonstrated on a composition monitoring problem derived from a real alkylation unit. The developed MF model, which combines diverse data sources and different model types, exhibits a significant improvement in the accuracy with respect to both LF data (online sensor) and the HF model (standard soft sensor).

Finally, as an example of data-driven MPC, a controller based on recurrent neural networks (RNNs) is used to address challenges posed by nonlinear process dynamics. A comparative analysis with the nominal scenario, based on the first-principles model with no plant/model mismatch, and with some traditional linear data-driven models, highlights the competitive performance of RNN-based MPC in the simulated control scenarios. A continuous flow stirred tank reactor characterized by complex equilibrium series kinetics is considered: informative data generation, model training and testing are discussed, and the proposed RNN-MPC proves to offer a competitive approach to enhance control strategies.

Keywords: *data-driven modeling, system identification, PID retuning, process monitoring, model predictive control, digital era.*