Models and data for future automation of biomanufacturing processes

Krist V. Gernaey

kvg@kt.dtu.dk

Technical University of Denmark, Department of Chemical and Biochemical Engineering, Denmark

Industry 4.0 comes with the promise of potentially transforming biomanufacturing, yielding better control of processes and opening up towards increased automation. Digital twins, virtual copies of a physical process interacting with that process, are assumed to play an important role in this transformation. This presentation has focus on two important aspects of such a digital twin: models, and data.

A digital twin of a process (or part of a process) can be considered a virtual copy of the process that exchanges data/information with the process in real-time: data are sent from the physical process to the virtual copy, and are processed to yield predictions about the ongoing biomanufacturing process. Subsequently, the model predictions are sent back from the digital twin to the real process, with the purpose of achieving improved operation. As such, there is a clear similarity with model predictive control. In this presentation, the focus will be on different modelling approaches that can be used to obtain a Digital Twin of a process, and issues related to model validation are highlighted.

Furthermore, the need for improved data collection is addressed, and examples of the use of more advanced measurement techniques on a biomanufacturing process are presented, including spectroscopy and electrochemical sensors.

Finally, interactions between model and data are discussed. Biomanufacturing operations at industrial scale are traditionally known to be relatively poor in on-line data, compared to the lab environment that is used to develop such processes. In that respect, it is interesting to reflect on the data collection aspect, where availability of more and better on-line data at industrial scale could have a strong impact on the application of digital twins in biomanufacturing. A data-poor environment might require advanced algorithms to extract information from available data; increased availability of relevant and information-rich on-line data might results in a decreased need for such advanced algorithms.