A Contribution to Chemical Process Operation Support: New Machine Learning and Surrogate Models-Based Approaches for Process Optimization, Supervision and Control

By	*Ahmed Shokry Abdelaleem Taha Zied*
Host	*Department of Chemical Engineering, Escola d'Enginyeria de Barcelona Est – EEBE, Universitat Politècnica de Catalunya – UPC*
Activity period	*December 2015- May 2021*
Email	*ahmed.shokry@polytechnique.edu*
Advisor	*Prof. Antonio Espuna*
Full thesis is available at	*https://upcommons.upc.edu/handle/2117/349568*

Abstract

In the chemical process industry, the decision-making hierarchy is inherently model-based. The scale and complexity of the considered models (e.g., enterprise, plant or unit model) depend on the decision-making level (e.g., planning, scheduling, operation) and the allowable time slot (hours, seconds) within which model simulation/optimization runs must be performed and their output analyzed to support the decision making. The use of high-fidelity models, which include detailed physics-based description of the process, is attracting wide interest of the process engineers. Since, these First Principle Model (FPMs) are able to accurately predict the behavior of the process, leading to realistic optimal decisions. However, their use is hindered by practical challenges as the high computational time required for their simulation and the unguaranteed reliability of their consistent convergence. The challenges become prohibitive at lower levels of the decision-making hierarchy (i.e., operation), where decisions are required online within time slots of minutes or seconds, entailing lots of simulation runs using complex and highly nonlinear FPMs. Machine Learning (ML) and Surrogate modelling techniques are potential solution for these challenges, which rely on developing simplified, but accurate, data-driven or machine learning models using data generated by FPM simulations, or collected from a real process. This Thesis presents a framework for the proper and effective use of surrogate models and ML techniques in different phases of the process operation. The objective is to provide efficient methodologies, each supporting the decision making in a specific phase of the process operation, namely; steady-state operation optimization, Model Predictive Control (MPC), multivariate system identification and multistep-ahead predictions, dynamic optimization, Fault Detection and Diagnosis (FDD) and soft-sensing.

1 Context & challenges

Process operation is an important layer in the decision-making hierarchy of chemical plants management. It receives, as inputs, the outcomes and decisions coming from higher level layers (i.e., scheduling), such as forecasts of prices and demands, and production rate targets over long time periods[1]. Then, the process operation optimization layer provides as output: i) the real-time optimal values of the process variables (pressures, temperatures, etc.) at which the plant must operate to achieve the required performance, considering quality, safety and environmental restrictions and, more importantly, reacting to sudden and unexpected variations of the process or external uncertainties (e.g., demand), ii) detailed and timely orders to the basic equipment control systems to implement actions to maintain the plant units functioning at these set-points (or reference trajectories) against expected disturbances and iii) timely information about the process functionality state, i.e., if it is functioning under normal or abnormal conditions, and about the possible type of fault that impacts the process leading to these abnormal conditions [2].

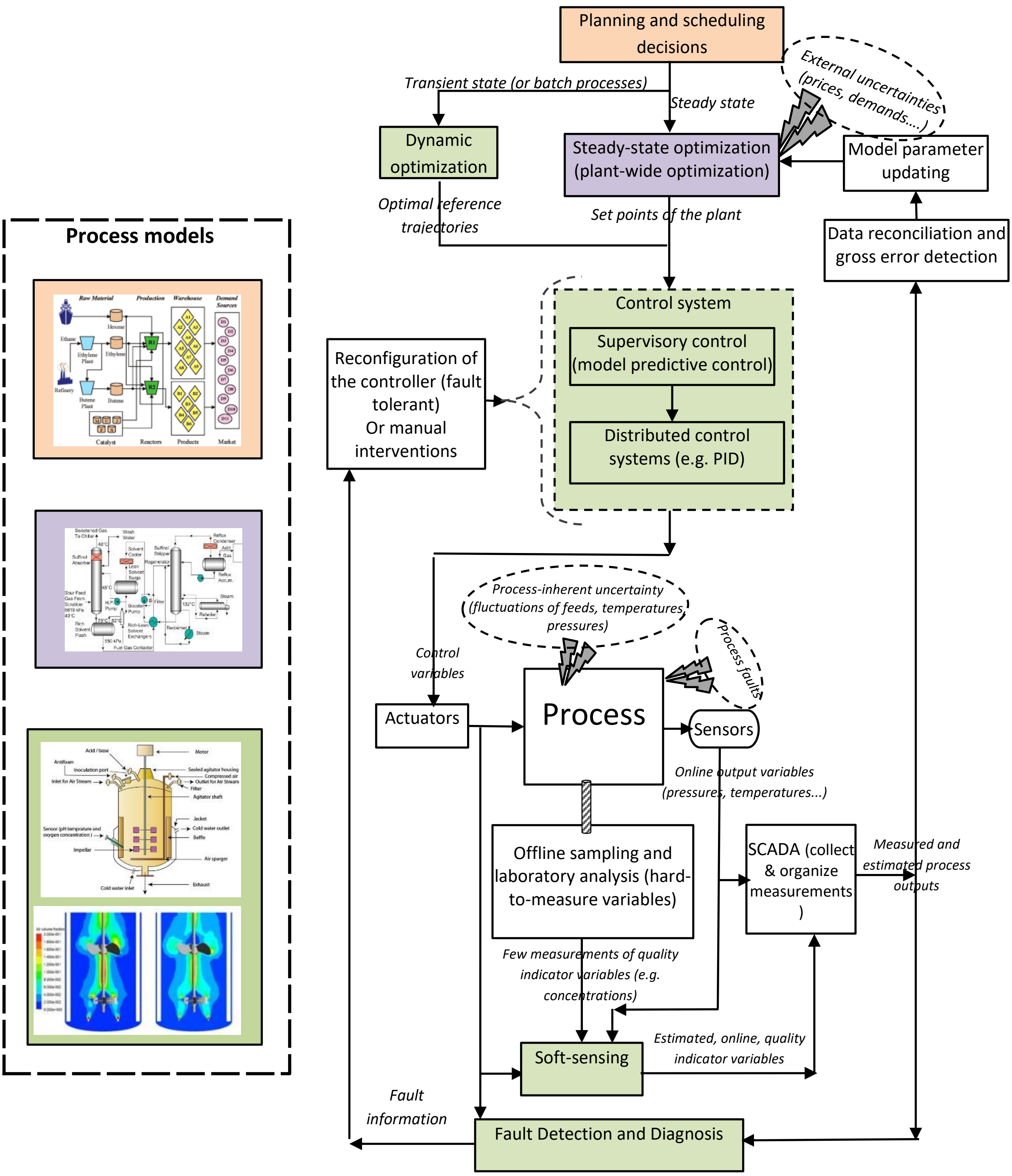

Figure 1. Process operation modules (right) and associated process model scales (left)).

Figure 1 sketches the main modules/tasks required for such functionalities, their usual activation sequence and the scales of the process models considered in each module.

1.1 Steady-state optimization

Context: Steady-State Optimization (SSO) aims at obtaining the optimal values of the process variables at which the plant must operate to maximize certain performance criteria and satisfy all requirements, by solving, in real time, an optimization problem based on a detailed and rigorous steady-state model of the process. On another hand, the presence of uncertainty sources in the system is unavoidable, including process-inherent (e.g., feed stream properties) and external uncertainties (e.g., demands) [3, 4]. To handle these types of uncertainty in SSO, reactive approaches are preferred, because they are able to promptly provide online update of the optimal values of the decision variables in response to real-time changes of the uncertain parameters value. Among the reactive methods, Multi-Parametric Programming (MPP) offers outstanding capabilities [5], since its solution provides simple mathematical functions mapping the optimal decisions over the entire space of the uncertain parameters. Hence, the optimal decisions can be immediately calculated by these functions avoiding huge computational cost required by repetitive optimization. There is a growing interest in using high-fidelity models of the process based on "first principles" (FPMs) in the SSO [6]. However, their development is a challenging task due to the required deep knowledge, effort and time. As a result, specialized simulation software tools have been developed to model such complex processes, e.g., Aspen [7].

Challenge 1: FPMs may pose many practical obstacles when they are used for SSO [8, 9], such as high nonlinearity, complex architectures, expensive computational cost, and noisy calculations [7].

Challenge 2: application of the most efficient reactive approaches (e.g.:, MPP) to handle uncertainty is not always possible, since it requires a well-constructed white-box model of the process [10].

1.2 Dynamic optimization

Context: for transient situations of continuous processes and also for batch processes, dynamic optimization is, instead, carried out relying on a dynamic FPM of the process [11]. It identifies, in a fast and accurate way, the optimal time-profiles of the control variables that must be applied over a specific period of time in order to drive the process to the required state at the end of this period [12]. Direct methods have been widely used for solving dynamic optimization problems, which are classified according to the variables to be discretized [11]. Control Vector Parameterization (CVP)) approach discretizes only the control variables and then a NonLinear Programming (NLP) problem is carried out in the space of the discretized control variables. Simultaneous approach discretizes both control and state variables resulting in a large-scale NLP problem [12]. CVP approach is straightforward and easy to implement, resulting in a NLP problem of a much reduced size [11].

Challenge 3: the CVP approach requires a large number of evaluations of the FPM, where each evaluation implies the integration of this model using expensive integration techniques [12]. Which, again, can result in a huge computational effort.

1.3 Model predictive control

Context: after obtaining the optimal set-points of the plant, they are sent to the supervisory control module that holds the plant operating at these set-points against expected fluctuations. Model Predictive Control (MPC) is, nowadays, the backbone of the supervisory control modules in chemical industries [5], as they offer efficient capabilities such as handling multivariable processes and constraints on the state and control variables [5]. In the MPC scheme, an online dynamic optimization problem is solved at each sampling period, based on a dynamic model of the process. Firstly, the dynamic model is updated by the current real measurements of the state/output variables representing the initial conditions [13]. Secondly, the dynamic optimization problem is solved to find the optimal profile of the control variables over the entire prediction horizon [10]. Then, only the first values of the calculated optimal control profile are implemented in the plant, and so on.

Challenge 4: solution of a dynamic optimization problem at each sampling period implies repetitive evaluation of the process model. This could require time larger than the sampling period, which makes the MPC application useless [5].

1.4 Dynamic modelling

In most control, monitoring and supervision modules, an accurate dynamic model of the process is a must [22].

Challenge 5: in frequent situations, accurate dynamic FPMs are complex and computationally expensive, while in other cases they are not attainable and only the process measurement data are available.

1.5 Fault detection and diagnosis

Context: a fault is an unexpected change in the process behavior that hampers the process normal operation, causing unacceptable deterioration of its performance. A Fault Detection and Diagnosis (FDD) system performs two functions: detecting the fault occurrence and, then, diagnosing its type [14]. This guarantees the safety, reliability and availability of the plant by avoiding sudden shutdowns and breakdowns [15]. Model-based FDD methods rely on monitoring the matching between the actual process measured features and the corresponding normal features estimated by a dynamic state-space FPM (an observer). This results in error signals whose values can be used to detect and diagnose faults [16].

Challenges 6: model-based FDD methods are applicable to processes of low nonlinearity and dimensionality. Therefore, their accuracy is reduced when applied to highly nonlinear processes, while their applications to large-scale processes result in a high number of observers that require an unaffordable computational effort [17].

1.6 Soft-sensing

Context: most of the process operations modules require the real-time measurements of the process variables. But, for what called Quality Indicator Variables (QIV), online measurements are not attainable due to technological and/or economic limitations [18]. Since QIV values are obtained through expensive and time-consuming offline sampling and laboratory analysis, which leads to delays preventing reliable control, monitoring and supervision of the process. Soft-sensing methods can be an effective alternative in these situations [18]. They are computational techniques that provide online and continuous "estimations" of the QIV values by exploiting the measurements of other variables of the process that are continuously recorded online by the physical sensors network. Model-based soft sensors use dynamic FPM to predict the process QIV, either solely or using data provided by physical sensors (e.g., for adjusting their parameters).

Challenge 7: further to the deep knowledge, cost and effort required to build reliable FPMs of chemical processes [18], available ones are often developed under the assumption of ideal working conditions. As a results, when these FPM-based soft sensors are used at industrial scale, uncontrolled disturbances and, possibly, different units/reactors geometries, can easily lead to the deterioration of their prediction accuracy [19].

2 Existing solutions and their shortcomings

2.1 Steady-state optimization

Existing solution: the use of Surrogate Based Optimization (SBO) methods have been proposed and received a big deal of attention to tackle **Challenge 1** [7, 6].

Drawbacks: current SBO methods proposed in the chemical engineering area show some limitations: i) they iteratively discard the previous training datasets and generate new ones (by the complex FPM simulations) for fitting new surrogate models, which can be computationally prohibitive in an online environment, ii) they do not consider the surrogate models uncertainty during the optimization search [20, 7] and, iii) they do not difficulty of handling uncurtaining of the constraints satisfaction (i.e., feasibility) [21].

Up to the author's knowledge, the literature, yet, doesn't include proposals for tackling the challenges that face the applications of MPP approaches for handling uncertainty in the operation optimization of steady-state processes for which the available model is complex, highly nonlinear and/or black box (i.e., **Challenge 2**).

2.2 Model predictive control

Existing solution: to overcome **Challenge 3**, explicit MPC methods have been developed [5], which solves the MPC problem offline providing a solution in the form of simple mathematical expressions able to calculate the optimal values of the control inputs, that should be applied in the next sampling period, as a function of the current values of the process state variables [13]. The obtained explicit functions are employed online to calculate the optimal values of the control inputs, in a simple and computationally cheap way.

Drawbacks: application of explicit MPC is conditioned by the availability of a linear discrete-time state-space process dynamic model with moderate dimensionality, which hinders its smooth usage in cases where the available process dynamic FPM is highly nonlinear, high dimensional, and/or black box [10].

2.3 Dynamic optimization

To the best of the authors' knowledge, use of data-driven approaches has been rarely proposed in the literature to tackle the challenges facing CVP methods in solving dynamic optimization problems (i.e. **Challenge 4**).

2.4 Data-driven dynamic modelling

Existing solution: ML techniques have been widely used to develop empirical dynamic models that capture the behaviors of nonlinear chemical processes (**Challenge 5**) [23]. These empirical dynamic models are built using data which are either generated from complex FPM simulations or measured from the real process [24].

Drawbacks: i) most of these methods are limited to univariate modeling, ii) they are validated using processes characterized by steady dynamics that involve simple changes in the control inputs [25, 26], iii) they provide simple Markovian models, while their ability to provide dynamic models with lagged inputs is not addressed, and iv) their robustness to handle different case studies and to integrate different ML types is not explored.

2.5 Fault detection and diagnosis

Existing solution: to tackle **Challenge 6**, data-based FDD approaches have been proposed, which rely on using data-driven Classification Techniques (CTs) that are trained using process historical data including information about normal and different faulty situations [19]. Then these CTs can be used for process supervision to detect and diagnose possible faults from the process output measurements. These approaches have shown a great flexibility and robustness for the FDD of nonlinear chemical processes without requiring any mathematical model of the process [19].

Drawbacks: CTs suffer from serious limitations, which is that the classification of faults is based only on the measurements of the process outputs, lacking any knowledge about the system dynamics. In dynamic situations, CTs could easily produce false alarms by diagnosing the changes in the processes outputs as faults.

2.6 Soft-sensing

Existing solution: to tackle **Challenge 7**, data-driven soft-sensing methods have been proposed. They are based on the construction of a data-driven model able to accurately approximate the relation between the QIV and the other online variables. In the literature, data-driven soft-sensors have been vastly applied to continuous processes, in order to predict the process steady-state behavior [26]. Comparatively, their application to batch processes, which are always in transient state, have been found to be relatively more complicated [27].

Drawbacks: most data-driven soft-sensing approaches for batch processes proposed in the literature have not considered the initial conditions of the batches, since they have been tailored for batch processes operated under fixed initial conditions.

3 Objectives

Directed by the challenges and drawbacks of current solution methods, the objectives of the Thesis include:

- **Objective 1:** the implementation of different state-of-art techniques for Design Of Computer Experiments (DOCE), ML models (also called data-driven or surrogate models), and ML models validation procedures.
- **Objective 2:** the development of a framework for data-based modeling of steady-state processes, which integrates the previously implemented techniques (in **Objective 1**). This framework is aimed at the flexible and robust construction of accurate ML or surrogate models of different types, and also the comparison between them to select the best ML model type for the case study to be addressed.
- **Objective 3:** the development of new methods for steady-state operation optimization of processes based on surrogate models, which enable the optimization of complex chemical processes that are difficult to be optimized through classical optimization methods. These difficulties can be due to the complexity and high nonlinearly of the process model and/or the existence of uncertainty sources.
- **Objective 4:** the development of an efficient and generic framework for data-driven multivariate dynamic modelling and emulation of complex and nonlinear chemical processes. The framework aims at providing dynamic models able to accurately and speedily predict the process outputs over large time horizons.
- **Objective 5:** the integration of these data-driven dynamic models in efficient methodologies for the enhancement of the process monitoring (e.g., a soft-sensing methodology), control (e.g., a dynamic optimization methodology) and supervision (e.g., a FDD methodology).

4 Contributions

The main contributions of the thesis are novel methodologies and algorithms, where each contribution is shown in an independent chapter.

- **Chapter 3** (References no. [10, 11], in the thesis publication list)
  - **Contribution I:** a new SBO methodology for the steady-state operation optimization of complex nonlinear chemical processes has been developed, in which the objective function and/or the constraints

are represented by black-box functions. The proposed approach consists in replacing the complex, nonlinear, black-box model of the processes built based on first principles with global kriging surrogate models. Then, an active optimization strategy involving a sequential sampling procedure, based on the Constrained Expected Improvement techniques, is used to explore the search space of the decision variables and to adapt, accordingly, the surrogate models, so as to reach a global solution for the problem. This contribution tackles **Challenge 1** and is related to **Objectives 1**, **2** and **3**.

- **Applications**: the methodology is applied to three benchmark examples and to two case studies of operation optimization of chemical processes modeled by modular black-box simulators.

➢ **Chapter 4** (References no. [1, 6, 7 12, 13, 14, 15, 17], in the thesis publication list)

- **Contribution II**: a novel ML-based methodologies for the multiparametric solution of continuous and mixed integer optimization of chemical processes operation, influenced by traceable uncertainty sources, has been developed. The methodologies are aimed at providing very accurate and fast-running data-based models that approximate the multiparametric behavior of the optimal solution over the uncertain parameters space. This contribution tackles **Challenge 2** and is related to **Objectives 1**, **2** and **3**.

- **Applications**: the methodologies are demonstrated through their applications to different benchmark examples from the MPP literature and to cases studies of process and unit operations optimization.

➢ **Chapter 5** (References no. [18, 27], in the thesis publication list)

- **Contribution III:** it consists in the development of a novel Data-Based explicit MPC methodology, which enables simple implementations of explicit MPC in situations when the available FPM model of the process is complex. This contribution tackles **Challenge 4** and is related to **Objectives 1, 2** and **5**.

- **Applications :** the effectiveness of the methodology is proven by its application to two benchmark examples adopted from the explicit MPC literature.

➢ **Chapter 6** (References no. [2, 19], in the thesis publication list)

- **Contribution IV:** a novel methodology for data-driven multivariate dynamic modelling and multistep-ahead prediction of nonlinear processes using ML models is developed. The proposed methodology utilizes machine learning techniques for building a group of NARX models, each of them able to predict the evolution of one process output as a function of the other inputs and outputs of the process, over a suitable time lag. The set of multivariate dynamic models, then, forecast the process outputs along larger time intervals (multistep-ahead prediction), through a recursive and inter-coordinated prediction scheme. This contribution tackles **Challenge 5** and is related to O**bjectives 1**, **2** and **4**.

- **Application**: the methodology is illustrated through its application to three case-studies of nonlinear dynamic processes selected from the process industry literature, including a bioreactor, three-interconnected-tanks, and an oil-shale pyrolysis batch reactor.

➢ **Chapter 7** (References no. [5], in the thesis publication list)

- **Contribution V:** a novel data-driven methodology for the sequential dynamic optimization applicable to solve open loop optimal control problems of complex highly nonlinear processes is presented. The method is based on the construction of a set of multivariate dynamic surrogate models (**Chapter 6**), which are able to accurately and rapidly predict the process output behavior corresponding to any time-profile of the process control inputs. Then, a sequential dynamic optimization procedure is tuned to integrate this set of dynamic surrogate models. This contribution tackles **Challenge 3** and is related to **Objectives 1**, **2, 4** and **5**.

- **Application**: the methodology is applied to three well-known problems from the process systems engineering area, including a plug-flow reactor, batch reactor, and a parallel reaction problem.

➢ **Chapter 8** (References no. [3,20, 27], in the thesis publication list)

- **Contribution VI:** proposes a novel data-based methodology for FDD that combines i) a dynamic observer based on multivariate dynamic surrogate models (**Chapter 6**) capable of estimating the expected normal outputs of the process and ii) ML classification techniques which are trained with residuals created from the comparison between the estimated outputs by the observer and the real outputs of the process. This contribution tackles **Challenge 6** and is related to **Objectives 1, 2, 4** and **5**.

- **Application**: the performance of the method is illustrated through its application to the well-known three-tank benchmark case study

➢ **Chapter 9** (References no. [4, 25,26], in the thesis publication list)

- **Contribution VII:** presents a new ML-based soft-sensing methodology for the online prediction of QIV of batch processes operated under changeable initial conditions (e.g., processes manage raw materials whose specifications differ from one batch to another, or when different product qualities/quantities are to be generated). This contribution tackles **Challenge 7** and is related to **Objectives 1, 2, 4** and **5**.

- **Application:** the methodology is demonstrated by its application to two simulation case-studies (a batch reactor and a fed-batch fermenter for Penicillin production) and to a real photochemical pilot plant.

5 Conclusions

This Thesis addresses two of the main challenges that can be faced at the different levels of chemical plants operation, which are the complexity of the available physics-based models of the processes and, in other situations, the lack of the process model. To tackle these challenges, the Thesis developed a platform of novel and efficient computer-aided methodologies relying on AI and ML. Each methodology is tailored to tackle these challenges at specific module/task of operation, such as SSO, MPC , multivariate dynamic modelling, dynamic optimization, FDD, and soft-sensing. The methodologies are developed through integrating a large number of tools and techniques in a highly innovative way, such as Design of Computer Experiments, ML regression models, ML classification models, clustering methods, mathematical programming, heuristic optimization, and modelling and simulation tools. The developed methodologies are assessed by comparing their performances to those of the state-of-the-art solution procedures, via their application to a wide range of benchmark examples and case studies of different characteristics, such as batch, fed-batch, continuous, single-unit, multiunit, simulated and real pilot processes. Moreover, the examples and case-studies belong to diverse domains including energy, reaction, biochemical, and environmental engineering. The obtained results confirms the remarkable effectiveness, high robustness and wide applicability of the developed methodologies. The methodologies could find real industrial applications in a wide range of problems, from process control and monitoring to process design, as well as sectors, from bulk chemicals to specialty and fine molecules.

6 References

[1] D. Muller, B. Dercks, E. Nabati, M. Blazek, T. Eifert, J. Schallenberg, U. Piechottka and K. Dadhe, "Real-Time Optimization in the Chemical Processing Industry," *Chem. Ing. Tech.,* vol. 89, pp. 1464-1470, 2017.

[2] D. E. Seborg, T. F. Edgar, D. A. Mellichamp and F. J. Doyle, Process Dynamics and Control, 4th ed., USA: John Wiley & Sons, 2016.

[3] Z. Li, PROCESS OPERATIONS WITH UNCERTAINTY AND INTEGRATION CONSIDERATIONS, New Brunswick, New Jersey: Rutgers, The State University of New Jersey, 2010.

[4] Y. Jiao, H. Su, W. Hou and Z. Liao, "Optimization of refinery hydrogen network based on chance constrained programming," *Chemical Engineering Research and Design,* vol. 90, pp. 1553-1567, 2012.

[5] J. Katz, I. Pappas, S. Avraamidou and E. N. Pistikopoulos, "Integrating deep learning models and multiparametric programming," *Computers and Chemical Engineering,* vol. 136, p. 106801, 2020.

[6] O. Kajero, T. Chen, Y. Yao, Y.-C. Chuang and D. S. H. Wong, "Meta-modelling in chemical process system engineering," *Journal of the Taiwan Institute of Chemical Engineers,* vol. 73, pp. 135-145, 2017.

[7] N. Quirante, J. Javaloyes-Antón and J. A. Caballero, "Hybrid simulation-equation based synthesis ofchemical processes," *Chemical Engineering Research and Design,* vol. 132, pp. 766-784, 2018.

[8] A. Norbert, S. Blagov, R. Böttcher, J. Schwientek, J. Burger, E. v. Harbou and M. Bortz, "Simulation and Multi-criteria Optimization under Uncertain Model Parameters of a Cumene Process," *Chemie Ingenieur Technik,* vol. 89, pp. 665-674, 2017.

[9] J. D. Kelly and D. Zyngier, "Unit-operation nonlinear modeling for planning and scheduling applications," *Optimization and Engineering,* vol. 18, pp. 133-154, 2017.

[10] P. Rivotti, R. S. C. Lambert and E. N. Pistikopoulos, "Combined model approximation techniques and multiparametric programming for explicit nonlinear model predictive control," *Computers & Chemical Engineering* , vol. 42, pp. 277-287, 2012.

[11] L. Wang, X. Liu and Z. Zhang, "A new sensitivity-based adaptive control vector parameterization approach for dynamic optimization of bioprocesses," *Bioprocess Biosyst Eng,* vol. 40, p. 181–189, 2017.

[12] L. Biegler, "An overview of simultaneous strategies for dynamic optimization," *Chemical Engineering and Processing: Process Intensification,* vol. 46, pp. 1043-1053, 2007.

[13] E. N. Pistikopoulos, "Perspectives in Multiparametric Programming and Explicit Model Predictive Control," *AIChE Journal,* vol. 55, pp. 1918-1925, 2008.

[14] O. Nelles, Nonlinear System Identification: From Classical Approaches to Neural Networks and Fuzzy Models, Berlin: Springer, 2001.

[15] Y.-J. Park, S.-K. S. Fan and C.-Y. Hsu, "A Review on Fault Detection and Process Diagnostics in Industrial Processes," *processes,* vol. 8, pp. 1123,1-26, 2020.

[16] M. Amozeghar and K. Khorasani, "An ensemble of dynamic neural network identifiers for fault detection and isolation of gas turbine engines," *Neural Networks*, vol. 76, p. 106–121, 2016.

[17] A. Elhsoumi, R. E. Harabi, S. B. Hadj, A. Naoui and M. N. Abdelkrim, "Monitoring of Chemical Processes Using Model-Based Approach," in *Numerical Simulations of Physical and Engineering Processes*, InTech, 2011, pp. 413-435.

[18] V. Venkatasubramanian, R. Rengaswamy, K. Yin and S. N. Kavuri, "A review of process fault detection and diagnosis Part I: Quantitative model-based methods," *Computers and Chemical Engineering,* vol. 27, pp. 293-311, 2003a.

[19] H. Jin, X. Chen, L. Wang and K. Yang, "Adaptive Soft Sensor Development Based on Online Ensemble Gaussian Process Regression for Nonlinear Time-Varying Batch Processes," *Ind. Eng. Chem. Res.,* vol. 54, p. 7320−7345, 2015.

[20] M. Askarian, G. Escudero, M. Graells, R. Zarghami, F. J. Farahani and N. Mostoufi, "Fault diagnosis of chemical processes with incomplete observations: A comparative study," *Computers & Chemical Engineering,* vol. 84, p. 104–116, 2016.

[21] N. Quirante and J. Caballero, "Large scale optimization of a sour water stripping plant usingsurrogate models," *Computers and Chemical Engineering,* vol. 92 , pp. 143-162, 2016.

[22] Y. Zhang, Z.-H. Han and K.-S. Zhang, "Variable-fidelity expected improvement method for efficient global optimization of expensive functions," *Structural and Multidisciplinary Optimization*, vol. 58, p. 2018, 2018.

[23] L. Zhou, J. Chen and Z. Song, "Recursive Gaussian Process Regression Model for Adaptive Quality Monitoring in Batch Processes," *Mathematical Problems in Engineering,* vol. 2015, pp. 1-9, 2015.

[24] W. J. Lee, J. Na, K. Kim, C.-J. Lee, Y. Lee and J. Lee, "NARX modeling for real-time optimization of air and gas compression systems in chemical processes," *Computers & Chemical Engineering,* vol. 115, pp. 262-274, 2018.

[25] A. F. Hernandez and M. A. Grover, "Stochastic dynamic predictions using Gaussian process models for nanoparticle synthesis," *Computers & Chemical Engineering,* vol. 34, pp. 1953-1961, 2010.

[26] F. Boukouvala, F. J. Muzzio and M. G. Ierapetritou, "Dynamic Data-Driven Modeling of Pharmaceutical Processes," *Ind. Eng. Chem. Res,* vol. 50, p. 6743–6754, 2011.

[27] L. Wang, H. Jin, X. Chen, J. Dai, K. Yang and D. Zhang, "Soft Sensor Development Based on the Hierarchical Ensemble of Gaussian Process Regression Models for Nonlinear and Non-Gaussian Chemical Processes.," *Ind. Eng. Chem. Res.,* vol. 55, pp. 7704-7719, 2016.

[28] Y. Liu, Z. Gao, P. Li and H. Wang, "Just-in-Time Kernel Learning with Adaptive Parameter Selection for Soft Sensor Modeling of Batch Processes," *Ind. Eng. Chem. Res.,* vol. 51, p. 4313–4327, 2012.

7 List of publications produced from the Thesis

Published Journal papers:

[1] ***Ahmed Shokry****, Sergio Medina, Piero Baraldi, Enrico Zio, Eric Moulines, Antonio Espuña. A Machine Learning-based Methodology for Multi-Parametric Solution of Chemical Processes Operation Optimization under Uncertainty.* ***Chemical Engineering Journal****, 2021, 131632.*

[2] ***Ahmed Shokry****, Piero Baraldi, Enrico Zio, Antonio Espuña. Dynamic Surrogate Modelling for Multistep-Ahead Prediction of Multivariate Nonlinear Chemical Processes.* ***Industrial & Engineering Chemistry Research****, 59, 15634–15655, 2020.*

[3] ***Ahmed Shokry****, Mohammad Hamed Ardakani, Gerard Escudero, Moisès Graells, Antonio Espuña. Dynamic Kriging based Fault Detection and Diagnosis Approach for Nonlinear Noisy Dynamic Processes.* ***Computers & Chemical Engineering****, 106, 758-776, 2017.*

[4] ***Ahmed Shokry****, Patricia Vicente, Gerard Escudero, Montserrat Perez Moya, Moisès Graells, Antonio Espuña. Data-driven Soft-Sensors for Online Monitoring of Batch Processes with Different Initial Conditions.* ***Computers & Chemical Engineering****,118, 159-179, 2018.*

[5] ***Ahmed Shokry****, Antonio Espuña. Sequential Dynamic Optimization of Complex Nonlinear Processes based on Kriging Surrogate Models. Procedia Technology, 15, 376–387, 2014.*

[6] *Sergio Medina-González,* ***Ahmed Shokry****, Javier Silvente, Gicela Lupera Calahorrano, Antonio Espuña. Optimal management of bio-based energy supply chains under parametric uncertainty through a data-driven decision-support framework.* ***Computers & Industrial Engineering****, 139, January 2020, 105561.*

[7] *Gicela J. Lupera Calahorrano,* ***Ahmed Shokry****, Gerard Campanya, Antonio Espuña. Application of the Meta-MultiParametric Methodology to the Control of Emissions in the Industry under Continuous and Discrete Uncertain Parameters. Chemical Engineering Research and Design, 155 B, 365-373, 2016.*

Journal papers in preparations

[8] ***Ahmed Shokry****, Eric Moulines.* Health-Constrained Explicit Model Predictive Control Based on Deep-Neural Networks Applied to Real-Time Charging of Li-Ion Batteries. ***(to be submitted to*** *IEEE transaction on industrial informatics)*

[9] ***Ahmed Shokry****, Mehdi Abou El Qassime, Eric Moulines.* Deep Neural Networks-based Control Laws for Explicit MPC of High-Dimensional Complex Chemical Processes. ***(to be submitted to*** *Computers & Chemical engineering)*

Published Book Chapters & Proceedings

[10] ***Ahmed Shokry****, Antonio Espuña.* Applying Metamodels and Sequential Sampling for Constrained Optimization of Process Operations. ***Lecture Notes in Computer Science*****:** ***Artificial Intelligence and Soft Computing****, 8468, 396-407, 2014.*

[11] ***Ahmed Shokry****, Aaron D. Bokhariski, Antonio Espuña.* Using Surrogate Models for Process Design and Optimization. ***Uncertainty Modeling in Knowledge Engineering and Decision Making****, 483-488, 2012*

[12] ***Ahmed Shokry****, Sergio Medina, Antonio Espuña.* Mixed-Integer MultiParametric Approach based on Machine Learning Techniques. ***Computer Aided Chemical Engineering****, 40, 451-456, 2017.*

[13] ***Ahmed Shokry****, Antonio Espuña.* Dynamic Optimization of Batch Processes under Uncertainty via Meta-MultiParametric Approach. ***Computer Aided Chemical Engineering*** *40, 2215-2220, 2017.*

[14] ***Ahmed Shokry****, Antonio Espuña.* Optimization Under Uncertainty Based on Multiparametric Kriging Metamodels. ***Operations Research Proceedings****, 2016, Springer.*

[15] *Gicela Lupera Calahorrano,* ***Ahmed Shokry****, Georgios M. Kopanos, Antonio Espuña.* Mixed-integer Multiparametric Meta-Modeling: a Machine Learning Tool Applied to Reactive Scheduling. ***Computer Aided Chemical Engineering****, 43, 163-168, 2018.*

[16] *Gicela Lupera Calahorrano,* ***Ahmed Shokry****, Sergio Medina-González, Eduardo Vyhmeister,Antonio Espuña.* Ordinary Kriging: a Machine Learning Tool Applied to Mixed-integer Multiparametric Approach. ***Computer Aided Chemical Engineering****, 43, 531-536, 2018.*

[17] *Sergio Medina,* ***Ahmed Shokry****, Javier Silvente, Antonio Espuña*. A meta-multiparametric framework: Application to the operation of bio-based energy supply chains. ***Computer Aided Chemical Engineering****, 37, 1955–1960, 2015.*

[18] ***Ahmed Shokry****, Canan Dombayci, Antonio Espuña.* Multiparametric Metamodels for Model Predictive Control of Chemical Processes. ***Computer Aided Chemical Engineering****,38, 937–942, 2016.*

[19] ***Ahmed Shokry****, Antonio Espuña*. The Ordinary Kriging Metamodel in Multivariate Dynamic Modeling and Multistep Ahead Prediction. ***Computer Aided Chemical Engineering*** *43, 265-270, 2018.*

[20] ***Ahmed Shokry****, Mohammad Hamed Ardakani, Gerard Escudero, Moisès Graells, Antonio Espuña.* Kriging based Fault Detection and Diagnosis Approach for Nonlinear Noisy Dynamic Processes. ***Computer Aided Chemical Engineering****, 38, 55–60, 2016.*

[21] *Mohammad Hamed Ardakani,* ***Ahmed Shokry****, Gerard Escudero, Moisès Graells, Antonio Espuña*. Online Quantification of the Concept Drift Using Incremental Learned Classifier and Non-automatic Clustering. ***Computer Aided Chemical Engineering****, 43, 1069-1074, 2018*

[22] *Mohammad Hamed Ardakani,* ***Ahmed Shokry****, Gerard Escudero, Moisès Graells, Antonio Espuña.* Unsupervised Automatic Updating of Classification Models of Fault Diagnosis for Novelty Detection. ***Computer Aided Chemical Engineering****, 43, 1123-1128, 2018.*

[23] *Mohammad Hamed Ardakani, Mahdieh Askarian,* ***Ahmed Shokry****, Gerard Escudero, Moisès Graells, Antonio Espuña.* Optimal Feature Selection for Designing a Fault Diagnosis System. ***Computer Aided Chemical Engineering****, 38, 1111-1116, 2016.*

[24] *Mohammad Hamed Ardakani,* ***Ahmed Shokry****, Ghazal Saki, Gerard Escudero, Moisès Graells, Antonio Espuña.* Imputation of Missing Data with Ordinary Kriging for Enhancing Fault Detection and Diagnosis. ***Computer Aided Chemical Engineering****, 38, 1377-1382, 2016.*

[25] ***Ahmed Shokry****, Montserrat Pérez-Moya, Moisès Graells, Antonio Espuña.* Data-Driven Dynamic Modeling of Batch Processes Having Different Initial Conditions and Missing Measurements. ***Computer Aided Chemical Engineering*** *40, 433-438, 2017.*

[26] ***Ahmed Shokry****, Francesca Audino, Patricia Vicente, Gerard Escudero, Montserrat Perez Moya, Moisès Graells, Antonio Espuña*. Modeling and Simulation of Complex Nonlinear Dynamic Processes Using Data Based Models: Application to Photo-Fenton Process. ***Computer Aided Chemical Engineering****, 37, 191–196, 2015.*

[27] *Mohammad Hamed Ardakani, Ahmed Shokry, Gerard Escudero, Moisès Graells, Antonio Espuña*. A Framework for Unsupervised Fault Detection and Diagnosis Based on Clustering Assisted Kriging Observer. ***The 3rd Conference on Control and Fault-Tolerant Systems (SysTol)****, 2016, pp. 183-188.*

Submitted Book Chapters & Proceedings

[28] ***Ahmed Shokry****, Mehdi Abou El Qassime, Eric Moulines.* Leveraging Deep Learning for Efficient Explicit MPC of High-Dimensional and Highly Non-linear Chemical Processes. *Submitted to* ***ESCAPE32,*** *2022, Toulouse, France.*