Santillán Martínez, Gerardo; Sierla, Seppo; Karhela, Tommi; Vyatkin, Valeriy

**Automatic Generation of a Simulation-based Digital Twin of an Industrial Process Plant**

*Published in:*
Proceedings of the 44th Annual Conference of the IEEE Industrial Electronics Society, IECON 2018

*DOI:*
10.1109/IECON.2018.8591464

Published: 01/01/2018

*Document Version*
Peer reviewed version

*Please cite the original version:*

This material is protected by copyright and other intellectual property rights, and duplication or sale of all or part of any of the repository collections is not permitted, except that material may be duplicated by you for your research use or educational purposes in electronic or print form. You must obtain permission for any other use. Electronic or print copies may not be offered, whether for sale or otherwise to anyone who is not an authorised user.
Automatic Generation of a Simulation-based Digital Twin of an Industrial Process Plant

Gerardo Santillán Martínez\textsuperscript{1}, Seppo Sierla\textsuperscript{1}, Tommi Karhela\textsuperscript{1}, Valeriy Vyatkin\textsuperscript{1,2}
\textsuperscript{1} Aalto University, Helsinki, Finland
\textsuperscript{2} Luleå University of Technology, Luleå, Sweden

Gerardo.santillan@ieee.org, seppo.sierla@aalto.fi, tommi.karhela@aalto.fi, vyatkin@ieee.org

Abstract— A Digital Twin (DT) of a production plant is a digital replica of the plant’s physical assets which contains the structure and the dynamics of how the devices and process operate. Simulation-based DTs (SBDTs) are those based on online first-principles simulation models. In these systems, model parameter estimation techniques keep an online plant simulator in the same state as the targeted device or process. As a result, non-measured information of the current state of the plant can be obtained from the model. SBDTs can be used for a number of important applications and they have various advantages compared to DTs based on data-driven models. However, wider industrial adoption of SBDTs is hindered by laborious development of their underlying first-principles simulation model as well as by a lack of integrated lifecycle-wide implementation methods and simulation architectures. This paper focuses on applying previously presented methods for reducing implementation effort of SBDTs. Firstly, laborious simulation model development is tackled by applying an automatic model generation method. Secondly, an integrated implementation methodology of a lifecycle-wide online simulation architecture is followed for developing the SBDT. A SBDT of a laboratory-scale process is implemented to demonstrate the proposed method. The results show a higher level of fidelity compared to previous publications.

Keywords— digital twin; dynamic process simulation; first-principles model lifecycle, simulation-based digital twin;

I. INTRODUCTION

Recent advances in modelling and simulation technology and in industrial interoperability standards have resulted in the development of Digital Twins of production plants \cite{1, 2}. A Digital Twin (DT) is a digital replica of the physical assets of an industrial plant which contains the structure and the dynamics of how the devices and the process operate \cite{3}. They are a powerful application for decision support of operational process plants in sectors such as chemical, power generation, mineral processing, pulp & paper and oil & gas. Commercial DT solutions for process plants are commonly based on data-driven models developed purely from the measured data of the targeted industrial plant \cite{4, 5}. DTs based on data-driven approaches rely on black-box models built to capture relations between the inputs and outputs of the plant \cite{5}. These systems can be applied to obtain production forecasts or to detect certain production anomalies. However, since they are based only on measurement data obtained from the plant, they cannot be used to forecast abnormal plant operation states not covered by the available collected data. Additionally, they require expert interpretation and are thus difficult to scale up \cite{5}. Moreover, applications based on data-driven DTs depend entirely on the automation and the monitoring systems data to provide information of the current plant state.

In contrast, simulation-based DTs (SBDTs) are based on online first-principles simulation models \cite{4–6}. First-principles models (FPMs) rely on engineering, physics or chemical descriptions to represent the behavior of the plant \cite{7}. As shown in Fig. 1, in SBDTs, a simulation model runs together with the plant, while estimation techniques keep the simulation state in the same state as the targeted device or process \cite{8}. These simulation configurations are also known as online model-based applications \cite{5}. A SBDT can be used to obtain high-fidelity predictions, including production forecasts of operating regions from which no measurement data is available \cite{9}. Furthermore, SBDTs can be used for developing operator training simulation systems, for production optimization, or for troubleshooting and failure diagnoses. SBDTs are a holistic and powerful application for plant operation support of modern industrial plants. However, developing the FPM of a SBDT can be time-consuming and expensive \cite{5}. Although research on the reuse of existing simulation models may lead to a significant reduction of effort and cost \cite{10}, development of FPMs remains laborious \cite{5}. Moreover, the lack of systematic approaches for SBDTs implementation, which address complex integration of the process with simulation systems and methods, limit wider industrial adoption of SBDTs \cite{8}.

In this paper, these SBDT shortcomings are addressed by applying a combination of previously studied implementation methods. In particular, laborious FPM development is tackled by exploring the utilization of automatic model generation (AMG) methods. Existing AMG methods utilize data from engineering sources that are accessible already during the process design phase. These sources include piping and instrumentation diagrams (P&ID), equipment technical data sheets and control application programs \cite{11–13}. This paper builds on our recent work, which proposes an AMG method based on the utilization of 3D plant model information \cite{14}. This method is applied for automatically generating the FPM. Furthermore, a lifecycle-wide online simulation architecture \cite{8} is followed to integrate the SBDT with the physical plant. The goal of this paper is to propose and test a method for generating a SBDT from an automatically generated FPM. This paper is structured as follows. Section II provides an overview of related work. Section III presents the proposed method for automatic generation of SBDTs. The presented method is tested by implementing a SBDT of a laboratory-scale process. The description of the example implementation and its results are presented in Section IV. The conclusions are presented in Section V.
II. Related work

A. The digital twin concept and its applications

A DT can be described as a digital copy of a physical asset that contains models of its structure and of its behavior [3], [6]. According to the definitions presented in [15], [16], DTs fulfill the following characteristics: 1) They are a real-time reflection of the physical asset. This is achieved through continuous synchronization between the real asset and the DT states. 2) They are fully integrated with the real asset and they can interact with current and historical data of the physical asset, enabling continuous improvement of the DT. 3) A DT can directly compare and analyze predicted and measured values of the physical asset. As a result, a DT can be used for simulating, monitoring, optimizing and verifying various activities in the entire asset lifecycle.

Various examples have been implemented in different industrial domains since the concept of DT was firstly introduced over a decade ago [17]. In the automotive and aerospace industries, DTs have been used as an ultra-high fidelity simulated replica of a vehicle, which can be used for anomaly diagnoses and for predicting future states and remaining useful life [4], [18]. In construction industry, DTs have been implemented after combining virtual models with physical data in order to obtain more accurate structure fatigue information of buildings [19]. In the industrial manufacturing domain, DTs of production assets have been applied for product lifecycle management. Consequently, DTs implemented for the manufacturing industry aim to mirror the entire lifecycle of end-products. As a result, DT applications in fracturing have been focused on product lifecycle design and services [2], [20]. On the other hand, DTs of the manufacturing plant [16], [17], [21] have been mainly based on plant virtualization to achieve improved flexibility, scalability and efficiency.

B. Digital twins for process industry

The industrial process domain is comprised mainly by the chemical, power generation, mineral processing, food processing, pulp & paper and oil & gas industries [22]. In this domain, simulation-based applications, which fulfill the DT characteristics previously listed, have been extensively used during decades. These applications, known as online model-based applications (OMBAs) [5] are online simulation systems based on an up-to-date condition of the physical plant. Although they can be implemented early during the plant design, they are mainly used during the operation and maintenance phases when it is possible to interface them with the real plant through connections between the simulation system and the control application of the plant. OMBAs are able to continuously update their state in order to closely represent the current state of the process. This is achieved through online and offline estimation of their underlying model parameters, as shown in Fig. 1. OMBAs have become the DTs of process plants as they have been used for various applications, including plant monitoring [23] and production forecasting [8].

Commercial DTs for industrial processes [24], [25] focus on building the model of the physical process following data-driven techniques [3], [4]. Although these models are fast to develop, drawbacks of data-driven-based DTs presented in Section I hinder their utilization for important industrial applications such as virtual sensors and operator training simulation. In contrast, SBDTs can be used for these critical applications but there are few commercial SBDT examples. The work in [26] is one of the earliest SBDT implementations. However, since it mainly focuses on proposing a method for online FPMs estimation, important SBDT implementation details such as its simulation system architecture are not tackled by this study. The work presented in [9] resulted in a commercial SBDT solution [27] which utilizes a combination of FPMs and statistical models for representing the physical system. However, industrial adoption of SBDT such as [26], [27], has been hampered by high development costs of FPMs and by a lack of implementation methodologies which ease development of SBDTs [5], [8].

Automatic generation of the simulation model of SBDT could be applied to increase cost-efficiency and to reduce development and maintenance time of the underlying FPMs [22]. Furthermore, lifecycle-wide implementation methodologies of OMBAs could be applied to tackle laborious integration between the simulation system and the process plant [8]. This paper leverages on the integration of these approaches in order to propose a method for enabling automatic generation of SBDTs of an industrial process plant for operation support.

Fig. 1. Simulation-based digital twin and its applications during the operation and maintenance phases of the plant lifecycle. Virtual Sensor refers to the application in which non-measured information of the current state of the process can be derived from the underlying FPM of the SBDT.
III. METHOD

In the proposed method, automatic generation of SBDTs is achieved by applying methods for automatic generation of the underlying FPM. AMG approaches use information mapping algorithms to generate a simulation model based on the targeted system information. These algorithms automatically map the accessed data into the model logic specified by the simulation language utilized [28]. There are a number of AMG methods available. These methods utilize different data sources for automatic model generation, including P&ID [12], [29], [30], 3D plant models [14] and control application programs [31]. Fig. 2 shows how, in the proposed method, different AMG methods can be applied and combined in order to enable automatic SBDT development. The phases depicted with numbers in Fig. 2 are described as follows:

0) During the preliminary planning of the plant, the steady state simulation model is created based on the initial process design and on nominal equipment information available from equipment data sheets. Steady state simulations are used for preliminary process design. They can only determine time-independent system response to a specific set of inputs [32]. For this reason, dynamic simulation models are required for implementing SBDTs.

1) At the process plant design phase, the dynamic simulation model, including a preliminary version of the control application model, can be automatically generated based on the steady state model [30]. The steady state model can also be used for providing initial conditions to the dynamic simulation model.

2) Information from P&ID and I/O tables of the automation system design are utilized for generating a more detailed model of the process and of the control application can be automatically generated from P&ID, equipment data sheets and I/O tables as it is proposed in [11], [12], [29].

3) 3D plant model information is utilized for detailed configuration and parametrization of the process simulation model. In particular, parameters such as the structural and geometrical data of the process equipment and especially of its pipeline layout is used, for example, to calculate head losses due to elbows or branches in the pipeline, thereby increasing the accuracy of the simulation results. Alternatively and similarly as in step 2, the simulation model of the process can be automatically generated using information only from the 3D plant model and process nominal information extracted from the equipment data sheet, as it is presented in [14].

4) During the plant commissioning, the control application model is utilized for generating a preliminary version of the real control application program of the process plant. The preliminary version of the control application can be utilized for virtual commissioning and eventually improved for its deployment on the distributed control system of the plant [33].

5) During operation and maintenance phases of the process plant lifecycle, the automatically generated simulation model of the plant is connected to the automation system, to the monitoring application as well as to the manufacturing execution system (MES) and enterprise resource planning (ERP) systems of the plant. This communication enables continuous synchronization of the simulation model and the current process state.

After the automatic generation of the FPM, during step 5, the SBDT requires integration with other plant information systems. This integration is a non-trivial task as the simulation system must be connected in a non-disruptive manner while the process plant is under operation. Furthermore, different simulation methods, required for model initialization and for model estimation must also be integrated to the simulation
Consequently, in the proposed method, the online simulation architecture presented in [8] is utilized to achieve efficient system integration. This architecture, originally designed for adaptation of previously existing models, is applied for the automatic generation of SBDTs.

AMG tackles laborious FPM development. However, the results of the automatically generated FPM of the SBDT does not always fully correspond to the current process measurements. One possible reason for this is that the physical system was not built exactly according to the design specification that was used to generate the FPM, or if the design specification lacked some detail. Another possible reason is that the physical system has been in use and some parameters such as friction have changed over time. Thus, in the proposed method, FPM parameters are optimized using the physical system historical data for the simulation model to represent the current behavior of the SBDT.

Finally, the SBDT needs to remain an accurate twin of the physical plant throughout the operation and maintenance phases of the plant lifecycle. Therefore, continuous alignment of the SBDT plant states is needed. This is accomplished by applying online model parameter estimation. This technique is required to achieve a permanent state synchronization of the SBDT with the targeted process by dynamically calibrating the FPM state based on a comparison of process measurements with simulation model results. Different parameter estimation methods can be utilized. However, one of the most frequently implemented techniques in industrial systems is implicit dynamic feedback, due to the relative ease of implementation and general applicability [34]. Fig. 3 shows the described steps of the method proposed in this work.

IV. EXPERIMENTS AND RESULTS

The proposed method was tested on the laboratory-scale heat production plant (HPP) described in detail in [8]. In order to assess the proposed methodology, design information of the HPP is used to automatically generate its FPM. Available design data sources of the HPP include the P&ID (see Fig. 5), equipment data sheets and the 3D plant model (see Fig. 6). The 3D model of the physical system, developed in AutoCAD Plant 3D [35], was created after measuring physical dimensions of the real process. The AMG method followed to automatically generate the FPM of the HPP is presented in [14]. In this method, process structure, dimensioning and component connection information is extracted from the machine-readable export of the 3D design tool. This information is combined with equipment nominal information obtained from their data sheets and then used to automatically generate and configure a dynamic thermal-hydraulic simulation model. The SBDT was implemented on Apros [36], a flowsheet-based tool for modelling and dynamic simulation of thermal-hydraulic processes.

After the simulation model is automatically generated, the implementation approach described in [8] is followed to integrate the simulation tool with the HPP process. Then, the simulation model is optimized utilizing the offline optimization method of the architecture described in [8]. As previously explained, this optimization is needed to readjust model parameters so that its behavior corresponds to the one described by historical data of the process. The method for the selection of parameters to be adjusted by the optimization and online estimation methods is described in [37]. Fig. 7 shows the comparison of the process measurements with the automatically generated model (AGM) results before and after
its optimization. These results correspond to transients caused by a change on the level setpoint L200 of the tank B200 and its corresponding inflow F100. Results show that the model generated following the AMG method presented in [14] closely follows the process results. However, the optimization method significantly improves the AGM fidelity. As a result, the AGM can be used as the underlying model of a SBDT. This is a significant improvement compared to the current state of the art on AMG [11], [12], [29], [30] in which the achieved fidelity of the resulting model limits its application only for factory acceptance tests and for virtual commissioning.

In order to complete the SBDT generation, the optimized AGM is connected to the real plant and initialized following the approach in [8]. Finally, the online parameter estimation method is started. The estimation method utilized is based on dynamic feedback estimation and it is described in detail in [34]. This method continuously adjusts model parameters in order to keep the simulated state in the same state as the operational process. Fig. 8 shows the online model parameter estimation results of the automatically generated SBDT. These results correspond to transients caused by changes on the L200 level setpoint of the tank B200 and its corresponding inflow F100. Results show that the SBDT based on the optimized AGM is able to closely follow the real process. At this point, the optimized AGM becomes the SBDT of the plant and it can be used for operation and maintenance support, as a virtual sensor or to obtain high fidelity predictions based on the current state of the process.

V. CONCLUSIONS

This paper presented a method for automatic generation of SBDTs of industrial process plants. In this work, laborious FPM development is addressed by utilizing an AMG method [14] which uses data from the 3D plant model to automatically generate the FPM of the SBDT. Furthermore, time-consuming system integration is addressed by applying a lifecycle-wide online simulation architecture [8]. This architecture enables also the possibility to apply model optimization and online estimation methods. The experimental results show that the fidelity of the AGM can be significantly improved after applying the optimization method of the architecture presented in [8]. The optimized AGM can be utilized for implementing a SBDT due to the high-fidelity achieved after applying the proposed approach. As discussed at the results section, this is a significant improvement on the current state of the art on AMG. Furthermore, the method proposed in this work shows that it is possible to reduce implementation effort required for the development of industrial SBDTs in order to increase industrial adoption of these systems.
ACKNOWLEDGMENTS
This work was supported by the research project Engineering Rulez: [38], partially funded by Business Finland [39].

REFERENCES


