Chinese Journal of Chemical Engineering 89 (2026) 302–313

Contents lists available at ScienceDirect

Chinese Journal of Chemical Engineering

journal homepage: www.elsevier.com/locate/CJChE

Full Length Article

A hierarchical and role-driven digital twin system with applications to complex chemical process operation

Xinyu Tao [1], Runjie Yao [2], Lingyu Zhu [3,*], Changrui Xie [1], Muqian Zhang [1], Xi Chen [1,2,*]

[1] *State Key Laboratory of Industrial Control Technology, College of Control Science and Engineering, Zhejiang University, Hangzhou 310027, China*
[2] *Huzhou Institute of Industrial Control Technology, Huzhou 313098, China*
[3] *College of Chemical Engineering, Zhejiang University of Technology, Hangzhou 310014, China*

ARTICLE INFO

Article history:
Received 26 April 2025
Received in revised form
31 August 2025
Accepted 1 September 2025
Available online 30 October 2025

Keywords:
Digital twin
Hierarchical framework
Role-driven
Process operation
Process modeling & optimization
Operator training system

ABSTRACT

Digital twin technology brings more opportunities and challenges to chemical engineering in both academic and industry. A complex process could have multiple digitalization needs, including simulation, monitoring, operator training, *etc.*; thus, a hierarchical digital twin would be a comprehensive solution to that. In this study, a novel and general framework of the digital twin is proposed for operations in process industry. With the hierarchical structure, the framework can handle various tasks driven by different roles in process industry, including managers, engineers, and operators. To complete these tasks, the framework consists of three modules: OAS (Operation Analysis System), OMS (Operation Monitoring System), and OTS (Operator Training System). Each module focuses on one unique type of demand from the staff, as well as interactions among them enabling efficient data sharing. Based on the hierarchical framework, a digital twin system is applied for one complex industrial nitration process, which successfully enhances the operation efficiency and safety in several industrial scenarios with different demands.

1. Introduction

Since first used in NASA's Apollo space program in 2010 [1], interests in the concept “Digital Twin” have greatly increased across both academia and industry, accompanied by a growth in the number of related publications, processes, concepts, and envisaged benefits [2]. With the introduction of numerous advanced frameworks and algorithms related to digital twin, various industrial applications have ensured efficient and safe production. Undoubtedly, digital twin is becoming a conventional tool for complex large-scale manufacturing.

The definition of digital twins has evolved over time. NASA described it as “a simulation of a system to mirror the life of itself”. Grieves *et al.* [3] later introduced digital twin as “A complete virtual description of a physical product that is accurate to both micro and macro level.” In recent years, despite variations in architecture definitions, researches on digital twins share the same aim, that is linking physical object and digital object in an accurate and real-time manner. It is not a specific technology, but an idea that can be implemented with many advanced technologies [4,5]. Through digital twin objects, a more detailed understanding of the behavior of physical objects is achieved, enabling various applications using the same digital model.

Across different life phase and based on various purposes, digital twins have been extensively applied, especially in discrete manufacturing industries. For instance, in design phase, Liu *et al.* [6] introduced a methodology for rapid individualized designing of the automated flow-shop manufacturing system, and Chinesta *et al.* [7] proposed a hybrid method combining data and mechanism. In the manufacturing phase, Sun *et al.* [8] completed dynamic iterative optimization of the process-parameters and real-time iterative optimization of the assembly-commissioning processing through digital twin. In the service phase, digital twin performed well with dynamic Bayesian network for faults detection [9]. It is evident that the applications of digital twins have extended across all phases of production, and a mature digital twin system can benefit production in various ways.

Though the term “digital twin” is not widely used in process manufacturing industries as in discrete manufacturing industries, similar concept attracted lots of attention. Comparing to simulators in process industry, the idea of digital twin emphasizes on precise description of physical entity to complete a series of tasks within one

---

* Corresponding authors.
E-mail addresses: zhuly@zjut.edu.cn (L. Zhu), xi_chen@zju.edu.cn (X. Chen).

https://doi.org/10.1016/j.cjche.2025.09.017

system. Researchers found that digital twins are capable of improving the process in real-time monitoring, predictive maintenance, process design, optimization, control, resulting in better efficiency in production and enhanced safety by facilitating early risk detection [10]. There are several successful applications. For instance, He *et al.* [11] proposed a digital twin for monitoring, diagnostics, and control of the chemical process, Wajeh *et al.* [12] established a dynamic digital twin system for modeling, advanced control and estimation techniques for biodiesel production, and Spatenka *et al.* [13] designed digital twin system for catalytic reactors to complete monitoring, forecasting, and optimization. Besides a whole digital twin system, some researchers focus on specific questions that contribute to digital twins. For instance, Yu *et al.* [14] proposed a simulation framework to achieve high-fidelity simulation of thermal systems by making full use of online measurements; Gao *et al.* [15] offered strategies of hybrid models that integrate process mechanisms and data-driven approaches to facilitate implementation of the digital twin; Wei *et al.* [16] proposed an active disturbance rejection control strategy for overcoming disturbance, uncertainties, and strong nonlinear couplings, and Vaccari *et al.* [17] proposed a real-time optimization framework for chemical products in production, storage and sale phase.

Despite numerous impressive cases, a hierarchical digital twin system for one specific process is still a challenge. Researchers have offered ingenious strategies for several scenarios, such as modeling, monitoring and training operators. However, a complex process may have various demands from different roles in the industry work. If different digital twin systems are applied independently to each problem, it will lead to significant redundancy in numerical modeling and computation, while also hindering future extensible development. To avoid incompatible strategies and redundant data processing in different tasks, it is effective to serve all the roles in the industry work by one unified process system. Therefore, an efficient hierarchical digital framework for processing multiple types of information is essential. The hierarchical digital framework would ensure that the same models is fully reusable while also adapting to the needs of different roles. Multiple modules collectively form a complete digital twin system, providing hierarchical digital applications for the process.

In this work, a digital twin system framework for general process operation is introduced. It consists of three modules, OAS (Operation Analysis System), OMS (Operation Monitoring System), and OTS (Operator Training System), handling modeling, monitoring, and training tasks, respectively. Overall, the three modules together form a hierarchical system, providing assistance to various groups of staff within the process industry, and addressing its own issues while communicating with each other. Section 2 introduces the structure of the hierarchical digital twin system. Section 3 analysis a complex nitration process as an example. Case studies associated with this process will be employed to illustrate the practical implementation results of individual modules. OAS, OMS and OTS are introduced in Section 4, 5 and 6, respectively. Section 7 is the conclusion.

2. Hierarchical and Role-Driven Digital Twin System for Process Operation

A classic digital twin framework for process industry is shown in Fig. 1. The upper part of the image represents the virtual entity, while the lower part represents the physical entity. The bidirectional data flow connects the two, forming a loop. From the perspective of practical application, three key components of the virtual entity are: data, model, and software. The red lines (both solid line and dashed line) in the figure illustrate how the information or commands flow within the system. In the flow within virtual entity, data and mechanism are used to create models of the process, and based on the models, the digital twin software completes description, diagnosis, prediction, and decision-making tasks. In the flow across virtual entity and physical entity, the digital twin software generates instructions and conclusion, which will be applied on physical objects manually or automatically. After that, the physical entity may work on new conditions and provide new data for the virtual entity. An ideal digital twin system can harmonize data, models, and software to significantly enhance industrial production efficiency.

As for process operation, the digital twin must adopt a hierarchically structured framework to address demands in different phases from different roles in process industry. A hierarchical framework of digital twin focusing on general process operation is presented in Fig. 2. The whole work consists of three modules, the OAS for design phase, the OMS for runtime phase, and the OTS for service phase. To show how the digital twin system can benefit the process, the work in the industry is categorized into three roles: Managers, referring to those who made decisions for the process and the company;

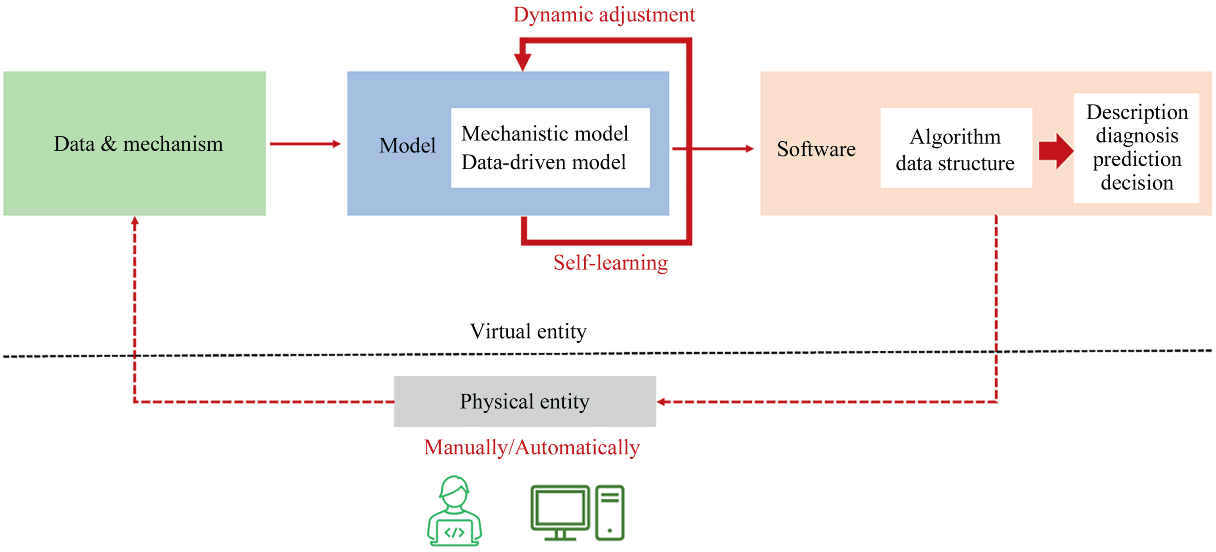


Fig. 1. Structure of a typical digital twin system for processes.

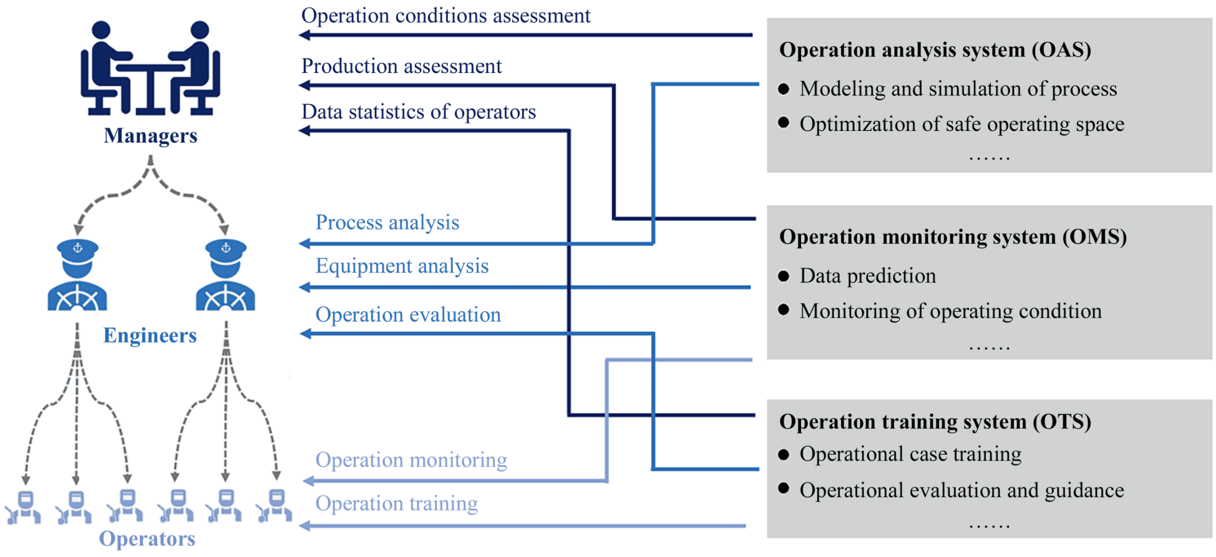


Fig. 2. The structure of hierarchical and role-driven digital twin system.

Engineers, referring to those focusing on the key technology of the process, such as working conditions, reaction yield and *etc.*; Operators, referring to those individuals who manually or semi-automatically perform chemical operations. Considering that the hierarchical framework of digital twin is designed to address the various demands from these roles in one system, it is defined as a "role-driven" digital twin system. Fig. 2 indicates that the three modules can benefit the work of the three roles from different perspectives.

1) Operation Analysis System (OAS)

OAS establishes the basic architecture of the digital twin system. It helps to validate the feasibility and reliability of the process from a mechanistic perspective. To achieve that, a generalized modeling platform is necessary. An idealized analysis system would be able to provide operational data using various unit models and under any operating conditions. It helps conduct simulation and optimization of the process in the design phase. Specifically, OAS can run operating conditions tests, which provide data for Managers to make decisions and for Engineers to analyze the process operation.

2) Operation Monitoring System (OMS)

OMS focuses on the monitoring work during the runtime phase. It establishes dynamic models capable of real-time monitoring of operating conditions and prediction of key variables in the process. To ensure the accuracy of dynamic data, a hybrid approach that combines mechanistic and data-driven modeling is a highly effective strategy. When the system is able to provide support with dynamic process information, Managers can use it for production assessment; Engineers can analyze equipment operation and Operators can obtain operation assistance.

3) Operator Training System (OTS)

OTS is a typical example for tasks in service phase of process industry. It provides virtual training scenarios for operators to help them become proficient in handling complex and hazardous process operations. The challenge in this module is to accurately extract key operating condition scenarios, and to provide feedback and guidance for different operations. For OTS, data interaction with OAS/OMS is necessary to establish an interactive dynamic training scenario. OTS can provide statistics data of operators for managers, operation evaluation for engineers, and operation training for operators.

Due to the various demands of different roles in the nitration process, it is a reasonable strategy to address different problems using different systems and technologies. Each subsystem has a unique structure to address its specific issues, and a different way for data interaction with the physical world. In addition, the digital twin system with three modules forms a clearly layered architecture. As the basement, OAS clarify digital rules and structure for variables, models and processes. Based on these rules, the process is built in a digital way. OMS provides dynamic models for monitoring tasks, mining actionable intelligence from historical data. And OTS makes use of static models and dynamic models to executes higher-level tasks atop simulation and monitoring results. In subsequent sections, OAS, OMS, and OTS will be introduced

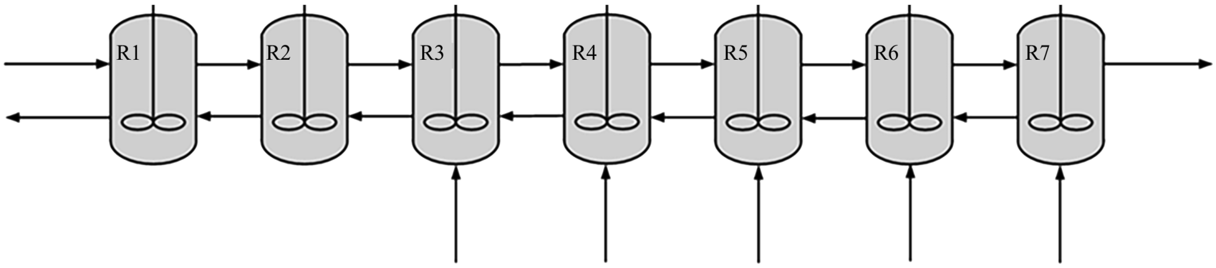


Fig. 3. The structure of a countercurrent nitration process.

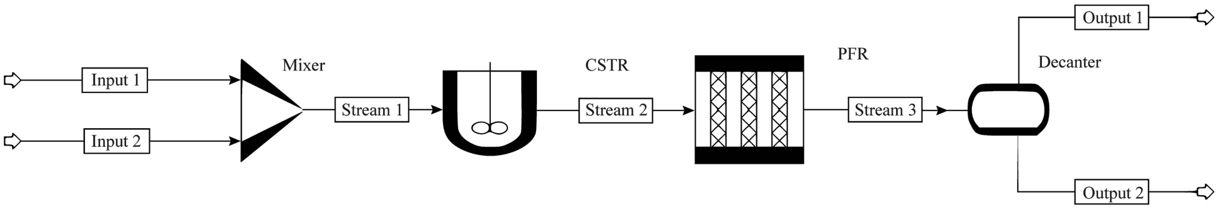


Fig. 4. Hierarchy of the nitration reactor.

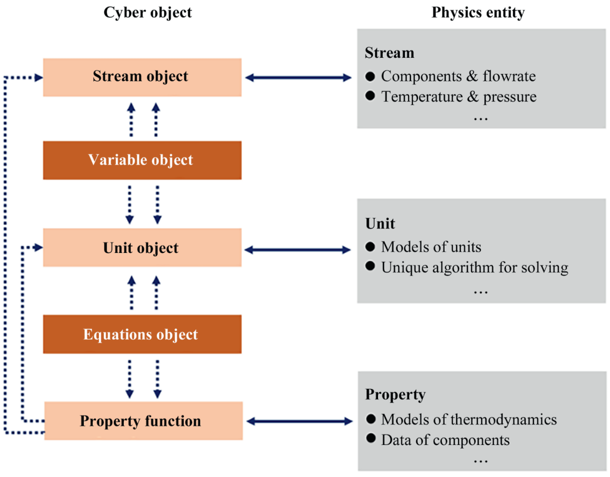


Fig. 5. Objects and their corresponding physical entities in OAS.

respectively, focusing on a) how the data structure of the subsystem is. b) how each system interacts with real-world data. c) how it interacts with other modules. Several application cases and practical performance related to the complex nitration process for different purposes will be presented in the corresponding sections.

3. Challenges and Demands in Complex Chemical Process

As an instance, nitration process is one of the most hazardous chemical processes, possessing diverse digital analysis needs across various aspects. Fig. 3 illustrates a complex nitration process which consists of seven cascaded nitration reactors in a countercurrent flow. One reactant, benzene, flows from left to right; the other reactant, nitric acid, mixed with the catalyst, sulfuric acid, flows in a reversed direction. Each nitration reactor is treated as a complex combination of a mixer, a continuous stirred-tank reactor (CSTR), a plug-flow reactor (PFR) and a decanter, as illustrated in Fig. 4. Two input streams are from the succeeding and preceding reactors, respectively; and the stream leaving the PFR is further separated by a decanter into two output streams flowing to the succeeding and preceding reactors. Nitration reactors adopt distinct parameter sets and connection configurations, rendering each nitration reactor unique within the process.

The continuous and countercurrent flow design significantly intensifies the reaction and increases the yield and productivity. However, it also poses stringent control requirements for temperature and composition in each reactor. Under improper operation, it may explode and cause catastrophic damage. Therefore, thermal safety is the most important issue in the process operation. In practice, the process was frequently shut down at disturbances to avoid further disaster, which brought great inconvenience for the process operation. To ensure thermal safety and efficiency throughout operation, industrial practitioners need to manage vast amounts of data and apply various strategies at different phases of a nitration process. That is where the digital twin system can play a crucial role.

The digital twin system could benefit the nitration process through its life cycle. In the design phase, digital twins provide platform for the digital process. The platform must accommodate customizable unit models to reflect the inherent variability in nitration process modeling. These customizable unit models would be used for simulation and other numerical analysis, such as optimization. In the runtime phase, any subtle variable variation in nitration processes may critically impact both process safety and operational efficiency. Thus, precise dynamic models are essential to monitoring the process, and even predict some key variables to provide early operation guidance. In the service phase, the digital process could help train the operators. After taking training courses in the system, human operators are able to gain a profound understanding of the characteristics of the process, master the operation methods proficiently, and avoid the occurrence of extreme accidents.

4. Module I: Operation Analysis System

As the first module, OAS provides modeling and simulation capabilities for the digital twin system. It plays a role as the basement of many applications. On one hand, by providing a digital process with adjustable units and structures, along with corresponding operational simulation data, OAS can offer strong support for the design and decision-making. On the other hand, by establishing the digital process corresponding to the physical entity, different operating conditions can be simulated to achieve various objectives, such as optimization and estimation.

The input of OAS includes information(data) related to streams, units, and process structure. That information can be either virtual, in the design phase, or real, in the runtime phase. What comes out of OAS is a fully transparent digital process. Data from any specific part of the digital process, no matter in streams, units, or process structure, is available for engineers and other modules of the whole system.

To complete the work of modeling and simulation, OAS has to be a generalized and open system: 1) The word “generalized” indicate that engineers should be able to establish models of vast majority of common chemical engineering units in OAS. To complete that, OAS needs to establish data structures for sufficiently fundamental concepts. That would make itself versatile and inclusive; 2) The word “open” indicate that engineers should be easy to operate their digital processes, to modify the streams, units or even functions in the system. To achieve that, OAS needs to have a clear modular design and various data APIs, to be friendly enough for modification and extension.

Fig. 5 illustrates the cyber-physical system designed for OAS. The double solid arrows are the correspondence while the dotted arrows are the connection within the cyber world. Our cyber system is based on two basic objects, variable object and equation object. Note that the equation object is designed to contain the relation between variables. For those who do not vary during a solving procedure, they are regarded as parameters and have no unique objects. Corresponding to physics entities, there are stream object based on variable, property function based on equation, and unit object based on equation (and variable if necessary).

A digital process is made of stream objects and unit objects. Every unit connects with several streams, and these streams connect with other units, forming a process structure. Since equations in a unit contain only the relation between variables linked to it, for example, a sum relation between instreams and outstreams in a mixer unit, there will be no extra connection equations in the system.

Fig. 6 illustrates the workflow of the OAS. User layer gives data and commands to the task platform. The platform then would build the digital process. In this phase, it instantiates stream objects and various unit objects from the model library. Then, the objects are linked to form the process. Meanwhile, the platform will ask the database for property information, like density and molecular weight. The task platform generates the algebraic system from user definition and interacts with the solver module for solution. Both equation-oriented and sequential-module solvers are available in the solving system. The results are stored in the corresponding stream and unit objects after the solving procedure is completed, available for users.

It is worth noting that such design in Figs. 5 and 6 gives the cyber-physical system two significant advantages. 1) Since the system is based on variables and equations rather than models of chemical devices, users are capable of selecting standard models or customizing non-standard models. Every detail in the process can be modeled in the cyber space, as long as it can be expressed in the form of variables/equations. This guarantees the platform to be general. 2) The object-oriented approach is widely applied in the platform design. Each unit or stream is an independent model, capable of data exchange or the addition of unique functional modules. Based on simulation of a process, the design makes the system open enough for missions like conditions tests, optimization, data integration and *etc*.

4.1. Application A: modeling in the digital twin system

Modeling is the first task in any cyber-physical system. In the countercurrent nitration process as shown in Fig. 3, one stream flows from R1 to R7 while the other flows in the opposite direction. The temperatures of each reactor are the key variables to infer whether the process is working safely and efficiently. Code 1 and Code 2 are displayed to show how to create process and customized units in OAS. They intuitively demonstrate that OAS has achieved openness and generalization: 1) The process is created totally based on object-oriented approach in Code 1, showing the modular design of OAS; 2) A customized unit is defined in Code 2, showing that engineers can easily create their own models and the models can be accepted by OAS.

Code 1: Modeling the process structure on OAS

```
# Define Process
Nitr = Process('Nitration', ComponentsList=components)

# Define Streams
SBZ_in = Nitr.DefStream('SBZ_0')
R1_in = Nitr.DefStream('R1_in')
SBZ1_out = Nitr.DefStream('SBZ_1')
SAC1_out = Nitr.DefStream('SAC_1')
# …

# Define Units
Mixer1 = Nitr.DefUnit('Mixer', 'Mixer1', [SBZ_in, SAC2_out, SW_1], R1_in)
R1 = Nitr.DefUnit('CSTR_with_decanter', 'R1', R1_in, [SBZ1_out, SAC1_out])
R1.temperature.SetValue(T1)

Mixer2 = Nitr.DefUnit('Mixer', 'Mixer2', [SBZ1_out, SAC3_out, SW_2], R2_in)
R2 = Nitr.DefUnit('CSTR_with_decanter', 'R2', R2_in, [SBZ2_out, SAC2_out])
R2.temperature.SetValue(T2)
# …

# Value for Feed Streams
# SBZ_in.SetValue(...)
# ...

# Value for InitialGuess
# SBZ1_out.InitialGuess(...)
# ...

# Simulation
Nitr.Solve()
```

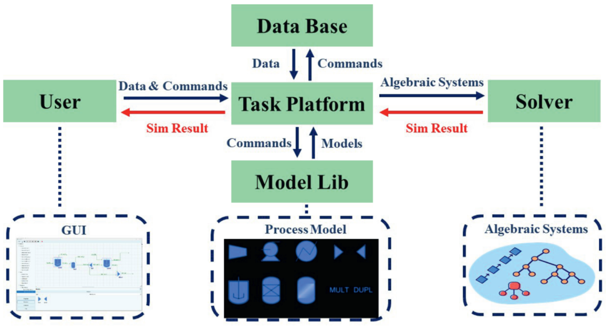


Fig. 6. The workflow of OAS modeling platform.

Code 1 shows how to establish such a process in OAS on our platform developed on Python. The procedure consists of five key steps: 1) Define the process/project, as line 2; 2) Define the streams in the process, as line 4–9; 3) Define the units with the stream connection, as line 11–19; 4) Input the data, as line 21–27; and 5) Run the simulation, as line 30.

Code 2: Create Customized Units on OAS

```
class CustomizedUnit(Unit):
    def DefInfo(self):
        self.EquNum = 2
        # Parameters
        self.var1 = 0.5
        self.var2 = 0.7

    def GetEquationFun(self):
        res = zeros(2)
        res[0] = self.instream.set[0].value - self.outstream.set[1].value / self.var1
        res[1] = self.instream.set[1].value - self.outstream.set[0].value / self.var2
        # x1 = y2 / par1
        # x2 = y1 / par2
        return res
```

On OAS, users have freedom to define customized units, as shown in Code 2. By inheriting from the basic class “Unit”, the specific unit has the same format and API as other standard units. Thus, only two functions are waiting to be rewritten. In “DefInfo”, users claim the number of equations and parameters in the unit; and in “GetEquationFun”, users provide the model of the unit. Note that OAS adopts the grammar of Python, which means not only equations but also logic relations, such as conditional judgment and third-party functions, are acceptable as a unit object. This feature allows various units definition in OAS and other modules in the digital twin.

4.2. *Application B: assessment and optimization on thermal safety*

OAS allows users to describe the complex structure precisely and define customized reactors in the network. Furthermore, OAS can help engineers to complete process optimization by defining objective function and constraints.

In the optimization of the nitration process, each nitration reactor is treated as a complex combination of a mixer, a continuous stirred-tank reactor (CSTR), a plug-flow reactor (PFR) and a decanter, as illustrated in Fig. 4. Note that the parameters of each reactor vary depending on its location in the process.

The models of CSTR and PFR in steady state can be described as:

$$F_{i,\mathrm{in}} - F_{i,\mathrm{out}} = V(-r_i) \tag{1}$$

$$V = F_{i,\mathrm{in}} \int_0^{x_i} \frac{\mathrm{d}x_i}{-r_i} \tag{2}$$

where F denotes the mole flowrate, V is the reactor volume, x_i is the conversion for the ith component; and r_i represents the reaction rate for the ith component as calculated by the following equations.

$$-r_i = \sum k_i^j C_i^n \tag{3}$$

where C represents mole concentration, j represents the jth reaction and n represents the order of the reaction. k_i^j is obtained by Arrhenius equation, pre-exponential factor equation and activation energy equation [18].

The model of the decanter can be formulated as:

$$x_i^a \gamma_i^a = x_i^o \gamma_i^o, i = 1, \cdots, C \tag{4}$$

$$\sum_i x_i^a = 1, \sum_i x_i^o = 1, i = 1, \cdots, C \tag{5}$$

$$\ln \gamma_i^a = \frac{\sum\limits_{j=1}^{C} x_j^a \tau_{ji} G_{ji}}{\sum\limits_{k=1}^{C} x_k^a G_{ki}} + \sum_{j=1}^{C} \frac{x_j^a G_{ij}}{\sum\limits_{k=1}^{C} x_k^a G_{kj}} \left(\tau_{ij} - \frac{\sum\limits_{m=1}^{C} x_m^a \tau_{mj} G_{mj}}{\sum\limits_{k=1}^{C} x_k^a G_{kj}} \right) \tag{6}$$

$$\ln \gamma_i^o = \frac{\sum\limits_{j=1}^{C} x_j^o \tau_{ji} G_{ji}}{\sum\limits_{k=1}^{C} x_k^o G_{ki}} + \sum_{j=1}^{C} \frac{x_j^o G_{ij}}{\sum\limits_{k=1}^{C} x_k^o G_{kj}} \left(\tau_{ij} - \frac{\sum\limits_{m=1}^{C} x_m^o \tau_{mj} G_{mj}}{\sum\limits_{k=1}^{C} x_k^o G_{kj}} \right) \tag{7}$$

$$G_{ij} = e^{-\alpha_{ij}\tau_{ij}}, i \neq j \tag{8}$$

$$\tau_{ij} = a_{ij} + \frac{b_{ij}}{T} + e_{ij} \ln T + f_{ij} T, i \neq j \tag{9}$$

$$\alpha_{ij} = c_{ij} + d_{ij} T, i \neq j \tag{10}$$

where x_i^a and x_i^o are the mole fraction of component i in the acid phase and the organic phase, respectively; γ_i^a and γ_i^o are the activity coefficient of component i in the acid phase and the organic phase, respectively; T is the reactor temperature; a_{ij}, b_{ij}, c_{ij}, d_{ij}, e_{ij}, and f_{ij} are the binary interaction parameters between components i and j.

Ensuring thermal safety is the top priority in the countercurrent nitration process. OAS can play a role in the assessment of thermal safety. A flexibility analysis problem is formulated as Eq. (11), aiming for quantitative safe boundaries on temperatures for three key reactors, R3, R4 and R5, in the nitration process. In Eq. (11), $\boldsymbol{x}$ represents the state variables; h represents the equality constraints, referring to the mechanism model; $g$represents the inequality constraints including: 1) constraints on nitration rates (including the first-stage and second-stage nitration rates); 2) constraints on cooling water flowrate, referring to the ultimate heat exchange capacity of the system; 3) constraints on cooling water outlet temperature, avoiding being too close to the temperatures of the streams in the process.

$$\begin{aligned} &\max_{\delta\in\mathbb{R}^+} \delta \\ &\text{s.t.}\quad g_j\left(\boldsymbol{T}^k, \boldsymbol{x}\right) \leq 0, \forall j \in J, \forall k \in K \\ &\quad h_i\left(\boldsymbol{T}^k, \boldsymbol{x}\right) = 0, \forall i \in I, \forall k \in K \\ &\quad \boldsymbol{T}^k = \boldsymbol{T}^N + \delta \boldsymbol{T}^i, \forall k \in K \end{aligned} \tag{11}$$

The aim of the flexibility analysis problem in Eq. (11) is to assess the thermal safety of the process and provide quantified safety boundaries in operation. The operating space of the temperatures, $\boldsymbol{T}^k$, $\forall k \in K$, is defined by $\boldsymbol{T}^N$ and δ in Eq. (11). $\boldsymbol{T}^N$ is the temperature setting value and the central point of the safe operating space; δ is the flexibility index of the problem, presenting the size of the safe operating space. The maximum δ that satisfies the constraints corresponds to the safety boundaries of the operating space. Thus, δ serves as the objective function in Eq. (11).

By setting $\boldsymbol{T}^N$ to be the same as one practical operating condition, the thermal safety assessment can be done on that condition. As a solution to Eq. (11), $\delta_{\max} = 8.60$. That indicates when the scaled temperatures of R3, R4 and R5 fluctuate within a range of ±8.60, the operating conditions of the process meet the safety standards, and the product also meets the quality requirements. Fig. 7 illustrates the safety space (denoted by the shaded regions) with the historical data. The temperature in the figure has been normalized and scaled to the range of 0 to 100, so does the data in the following text. The real temperatures may occasionally exceed the safety boundaries, resulting in potential safety hazards or non-conforming products.

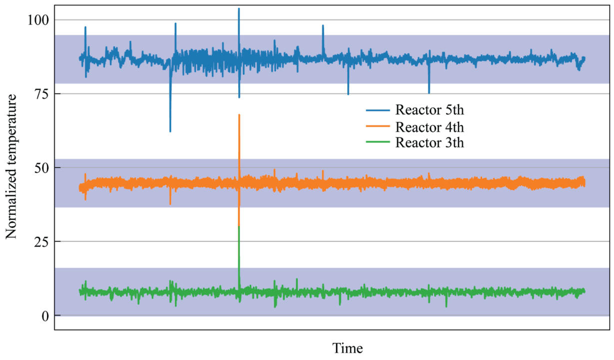


Fig. 7. Safe boundaries and industrial historical data.

To further enhance the operation safety, the temperature setting value is adjusted by adding a modification $\Delta\boldsymbol{T}^N$ to the previous $\boldsymbol{T}^N$, in order to obtain an increased safe operating space. Thus, an optimization problem as Eq. (12) is formulated.

$$\begin{aligned} &\max_{\Delta\boldsymbol{T}^N, \delta\in\mathbb{R}^+} \delta \\ &\text{s.t.}\quad g_j\left(\boldsymbol{T}^k, \boldsymbol{x}\right) \leq 0, \forall j \in J, \forall k \in K \\ &\quad h_i\left(\boldsymbol{T}^k, \boldsymbol{x}\right) = 0, \forall i \in I, \forall k \in K \\ &\quad \boldsymbol{T}^k = \left(\boldsymbol{T}^N + \Delta\boldsymbol{T}^N\right) + \delta \boldsymbol{T}^i, \forall k \in K \end{aligned} \tag{12}$$

In Eq. (12), since the aim is to optimize the temperature setting value $\boldsymbol{T}^N$, the operating space $\boldsymbol{T}^k$, $\forall k \in K$, is now defined by $\boldsymbol{T}^N$, δ and modification $\Delta\boldsymbol{T}^N$. $\Delta\boldsymbol{T}^N$ serves as the decision variables in Eq. (12), aiming for the maximum of the operating space δ. If the solution δ in Eq. (12) is bigger than it in Eq. (11), the new temperature setting value $\boldsymbol{T}^N + \Delta\boldsymbol{T}^N$ corresponds to larger safe operating space than $\boldsymbol{T}^N$, in which case the optimization is successful.

The optimal temperature change $\Delta\boldsymbol{T}^N$ and corresponding δ are obtained as:

$$\Delta\boldsymbol{T}^N_{\text{opt}} = [+0.39, -2.05, +1.40], \delta_{\text{opt}} = 14.91$$

As $\Delta\boldsymbol{T}^N_{\text{opt}}$ is accepted, δ is expanded from 8.60 in Eq. (11) to 14.91 in Eq. (12), which increases by 73.37%. Note that constraints g and h include not only safety constraints, but also efficient constraints, such as ensuring the conversion rate is no lower than 95%. Therefore, the result is optimal in both safety and efficiency. In the optimization case, OAS serves as a modeling and simulation platform, offering data of the process as the working conditions change in the iteration of the solver. In addition, as an open system, it is convenient for users to connect OAS to a nonlinear solver or an optimization module, as in the application.

In conclusion, OAS provide basic digital environment for model building, simulation and optimization tasks. The platform is open for most of the external software, such as solvers and computing modules. And it is general enough for users and other modules in the system to define customized models. By establishing a digital process on OAS, managers can test different working conditions to make decisions and engineers can analyze the process over a broader scope, such as quantified safe space.

5. Module II: Operation Monitoring System (OMS)

In the practice of chemical industry, offline static simulation and analysis alone do not meet the requirement of production. OMS is introduced as the second module to handle problems when engineers try to monitor a process. Typically, digital twin systems complete such missions by creating dynamic models, taking information from the sensors and providing computation results in real time. An ideal dynamic model running online in a process can benefit engineers in various fields. For instance, it can provide data that is not detected by actual sensors in the production, such as conversion rate in a particular reactor.

To describe the behavior of the process in a dynamic way, OMS takes real-time data and mechanistic knowledge as input. The main aim of OMS is to establish an accurate dynamic model, which is also the key challenge. Pure static mechanistic model fails to describe the dynamic characteristics. But pure data-driven and AI-assisted methods hardly perform well either for complex processes. It is necessary to make the full use of both real-time data and chemical engineering knowledge in the process industry. A hybrid approach that combines mechanistic and data-driven modeling is a highly effective strategy. Two applications are introduced in this section to show how OMS benefits the industry production.

5.1. Application C: temperature prediction

If abnormal fluctuations in temperatures can be predicted several minutes ahead of the operation, the operators can be warned and operate in time to save the process from an emergency and avoid shutdown. In light of this, the temperature prediction is developed in the OMS for the countercurrent nitration process.

In order to predict future temperature changes, a mechanistic model is first developed based on heat balance. The released heat from reactions is also included to correlate the temperature and concentration variables. Unscented Kalman filter(UKF) is then used to facilitate the co-estimation of both model states and parameters, and a Seq2Seq neural network is employed to compensate for the unmodeled dynamics. The whole framework is illustrated by Fig. 8.

The time-discretized mechanistic model can be represented as Eq. (13).

$$\begin{cases} x(k+1) = f(x(k), u(k), \theta) \\ y(k) = x(k) \end{cases} \tag{13}$$

where $x(k)$ is the state variables; $u(k)$ include raw material stream flowrates, inlet temperatures and flowrates of the cooling water; θ denotes the model parameters. In each iteration, UKF will do update and predict. By taking new observations, UKF updates the θ, and then computes $f(x(k), u(k), \theta)$. Such updates may cause new parameters to violate the physical constraints. To address this issue, a constrained optimization is incorporated to adjust the parameters in each iteration of the UKF.

In the UKF model, future states of the process are assumed to depend solely on the current state and control and are independent of the past states. However, it cannot be always true in the real-world industrial application. Various factors could violate this assumption, such as time lags in measurements, resulting in the deviation of the predicted values. Thus, the Seq2Seq network is employed to encapsulate historical data from a much longer past through its encoder. Therefore, by integrating a Seq2Seq network after the UKF, such deviations can be compensated. The specific algorithms and the effectiveness of the model can be found here [19].

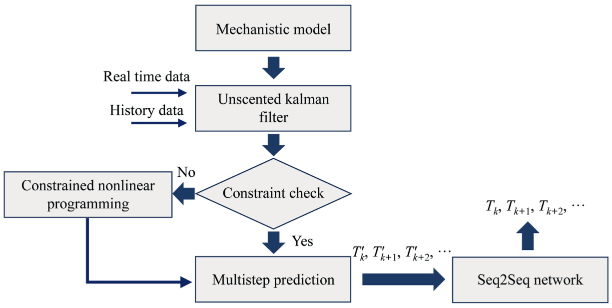


Fig. 8. The flowchart of the hybrid model.

Positive results are obtained when the framework is used in real-time monitoring. As shown in Fig. 9, it contains the temperature measurement denoted by black curve and the predicted temperatures of 1 to 5 step-ahead by color curves. Note that the temperature in the figure has been normalized and scaled to the range of 0 to 100. Each step corresponds to 1 min in the real world; thus, we can have the prediction by 1–5 min ahead, respectively. At 10:05, a disturbance occurred in the cooling water flowrate due to a wrong operation of the operator. The black curve records how the real temperature measurement fluctuated. The OMS also gives the temperature predictions. The predicted values accurately reflect the temperature change, providing early warning to operators to save the process back to the normal state. The total computational time for making the 5-step-ahead prediction once is usually no more than 3 s on a workstation with Intel Xeon Gold 5222 (3.80FHz). Since the 5-step-ahead prediction is conducted for every minute, the computational time consumption of a few seconds need not to consider.

5.2. Application D: nitric acid concentration monitoring

Nitric acid concentration is a key variable in the countercurrent nitration process that engineers and operators all rely on it to determine the current state of the reaction. In practice, however, the information can only be obtained through offline analysis for every 8 h. With the OMS, it is possible to build a soft-sensor to monitor the acid concentration in real-time. Mechanistic model with parameter estimation is used for this work.

The fixed conversion reactor mode [20] can be considered as consisting of two components, mass balance equations and thermal balance equations:

$$\begin{cases} MB(u, T, p_{\mathrm{m}}) = 0 \\ \\ Q(u, T, p_{\mathrm{R}}) = 0 \end{cases} \tag{14}$$

where u is the flowrate, T is the temperature, p_{m} and p_{R} is the parameters in the mass balance and thermal balance equations, respectively. When attempting to describe the reactor by this model in real time, it is challenging to establish accurate reaction kinetics equations that match real-world data. In addition, the combination of static mass balance equations and dynamic thermal balance equations requires a specific method of reconciliation.

Denote $x_{i,j}$ as the conversion rate of reaction$j$ in reactor i. It is selected as the key state variable in this work. It is defined as $u_{i,j}^{\mathrm{in}}(1 - x_{i,j}) = u_{i,j}^{\mathrm{out}}$, where $u_{i,j}$ is the flowrate of the limiting component of reaction j in reactor i. The thermal balance equations can be formulated as:

$$Q_i - Q_{R,i} + Q_{c,i} - Q_{sh,i} = 0 \tag{15}$$

where Q_i represents the heat change calculated from the temperature variation, as $Q_i = \rho_i c p_i V_i \frac{\mathrm{d}T_i}{\mathrm{d}t}$; $Q_{c,i}$ represents as heat exchanged by cooling water; $Q_{\mathrm{sh},i}$ represents the sensible heat due to the temperature difference between the feed and the reactor; $Q_{R,i}$ denotes the reaction released heat relating to the conversion rate, $x_{i,j}$. The thermal balance and mass balance are treated as two layers of the model, connected by the state variable $x_{i,j}$. In the real-time monitoring task, heat balance is first solved to estimate the conversion rate. Then the mass balance is used to estimate concentrations. Through this strategy, real-time monitoring of specific variables in the process is achieved. Fig. 10 illustrated the performance of OMS in acid concentration monitoring in practice.

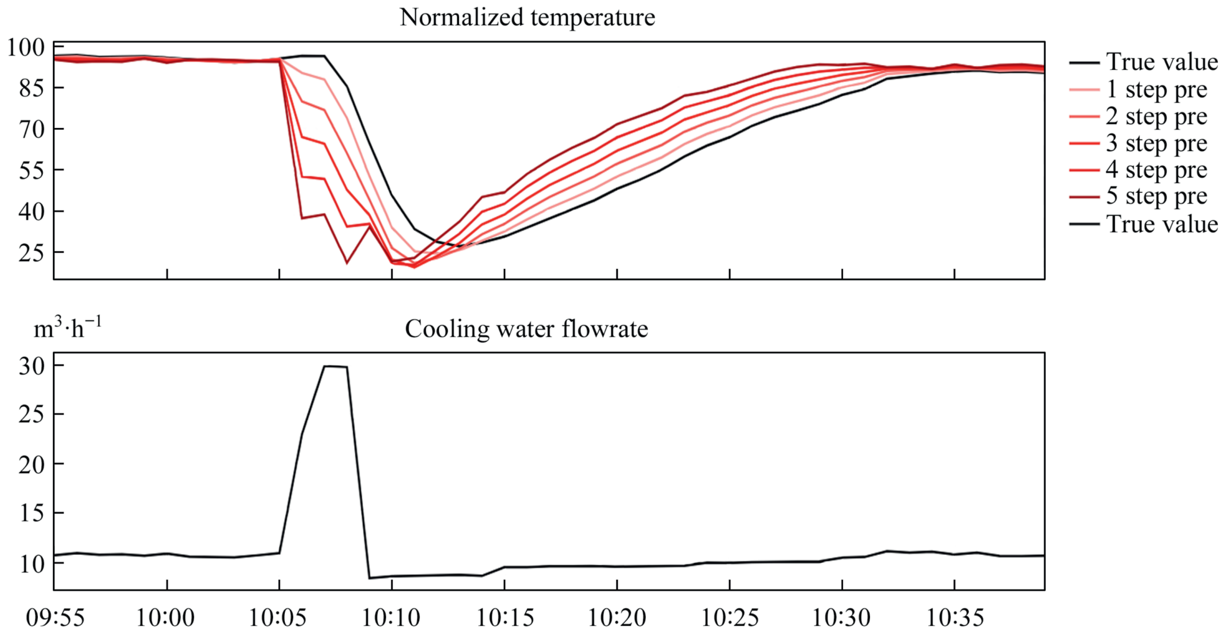


Fig. 9. The temperature prediction when cooling water Enters.

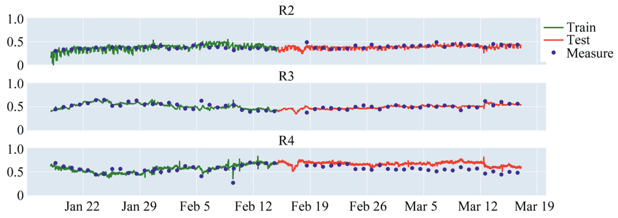


Fig. 10. Monitoring data of acid concentration in reactors 2, 3, and 4.

In Fig. 10, the horizontal axis denotes the timestamp, and the vertical axis denotes the nitric acid concentration in the reactors. Note that the data has been normalized and scaled to the range of 0 to 1.0. Measure data from the sensors in the industry is presented as blue dots. Data from the OMS is presented in green and red, with green for training data and red for test results. During the training stage, some parameters in the mass and heat balance equations are estimated and later fixed for the online prediction as shown in the test stage. The results show that the prediction well agrees with offline analysis, supporting that the OMS succeeds in monitoring key variables in real-time.

As a summary, the OMS serves as a module generate dynamic models in the nitration process. By applying hybrid models, it succeeds in real-time monitoring tasks in operation, such as variables monitoring and prediction. As a subsystem in the digital twin, it could provide equipment analysis for managers, trend analysis for engineers, and operation assistance for operators. Considering the open modeling platform of the digital twin system, OMS is also acceptable for further extension, like fault detection, troubleshooting methods and *etc*.

6. Module III: Operator Training System (OTS)

The process industry requires operators to execute corresponding operations based on different working scenarios. These scenarios are diverse, with various steady-state and dynamic characteristics, different working conditions, and distinct operational methods. To master the operations, operators are required to go through extensive manual training before taking up their positions.

Operator Training System (OTS) is designed to train operators using a simulation model and real data, as shown in Fig. 11. It involves the dynamic mathematical model development of the process under different scenarios, requiring the handling of process structure, control system, unit features,and reality data. When an operator executes operations on OTS, it sends input to the model and provides feedback caused by the operations. Meanwhile, OTS also records and verifies the operations, determines the completion status of each step, and subsequently provides an operational evaluation. To complete that goal, a data framework capable of managing various types of information, including simulation models and data, tags and step rules, and evaluation scores for operations, is essential. The framework is illustrated in Fig. 12. The modules are divided into functional and information modules, represented by light blue and grey blocks, respectively.

Tags Lib is at the center of the system. It stores the tag information of the entire process in the form of key-value pairs, consisting of the tag name and the current value. Almost all data interactions in OTS are carried out through Tags Lib. Operation UI is where operators execute operations and see the feedback in tags. Step information is stored in Steps Lib. Each step object

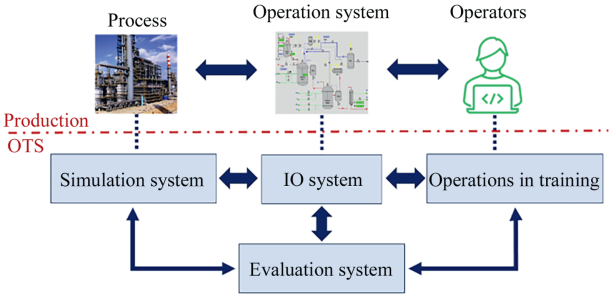


Fig. 11. The workflow of OTS.

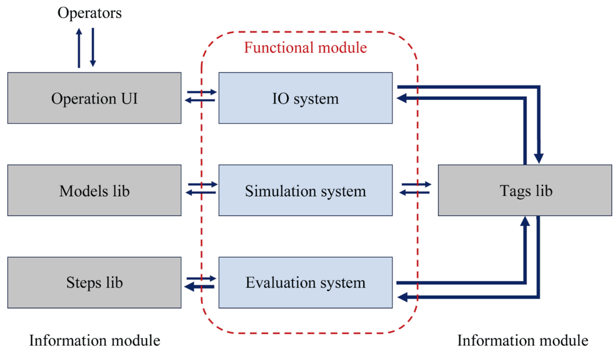


Fig. 12. The functional modules and information modules in OTS.

includes step status, associated operations, trigger tag numbers, completion conditions, and penalty conditions. Steps Lib also stores a hierarchical relationship diagram of all steps, enabling any combination of sequential and parallel operations, as well as operation prompts. Models Lib stores the models used for calculating the operational response. Each model contains the mathematical model, the input and output tags.

Corresponding to the information modules, there are three types of functional modules. IO System completes the data communication between Operation UI and Tags Lib. Simulation systems are designed to manage the models in the training. It calls the models in serial or in parallel to update values in tags. Evaluation System compares the operation with Steps Lib. It provides the completion status and penalty conditions of each step to assess the user's operational quality. Note that Evaluation System will monitor all the tags. When one tag value has been changed, Evaluation System locates all the corresponding steps in Steps Lib, and determines whether one step is complete. The steps can be structured in either a serial or parallel manner.

To meet specific training requirements, managers or engineers can define specific operational scenarios. Within a digital twin system, the OAS and OMS could help to build mechanistic models or hybrid models with historical data in specific scenarios. Steps with certain logical relation can be defined by scripts. Finally, an exam can be provided to the operators for each scenario under training.

OTS has been developed for the countercurrent nitration process, including startup procedure, shutdown procedure, emergency power outage, and emergency water shutdown. The startup procedure of the countercurrent nitration process is illustrated as an example to show how to build such a procedure in OTS. The workflow is shown in Fig. 13.

First, historical data is collected from the real plant. Next, by using a scenario extraction function, it is able to identify the data relating to the startup operations from the collected historical data and slice them into segments. Based on a predefined hybrid model structure developed in the OMS, parameter estimation can be conducted with the multi-segment data. Eventually, the model fitting with the startup operation data is built.

After the model is obtained, steps are created and stored in the Step Libs to form a step-tree for the startup scenario. As illustrated in Fig. 14, it shows a time slice during the test. There is a clear logical sequence among the steps, illustrated by black arrow connections. The completed steps are marked in blue while uncompleted ones are marked in grey. In this time slice, Step 5 is in progress and marked yellow. In a step object, step status, associated operations, trigger tag numbers, completion conditions, and penalty conditions are recorded. Each of the associated operations contains labels of itself, trigger tags, range of the target variable, completion mark, and operation score.

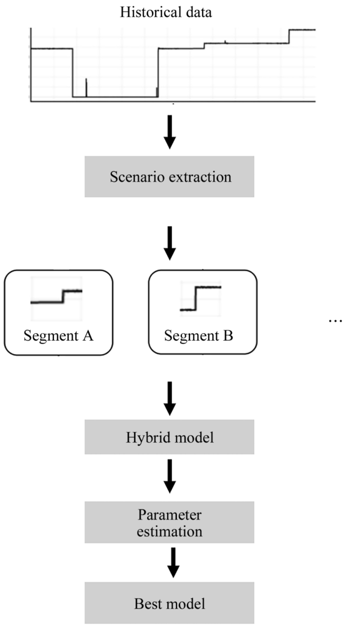


Fig. 13. The workflow to build models for startup procedure.

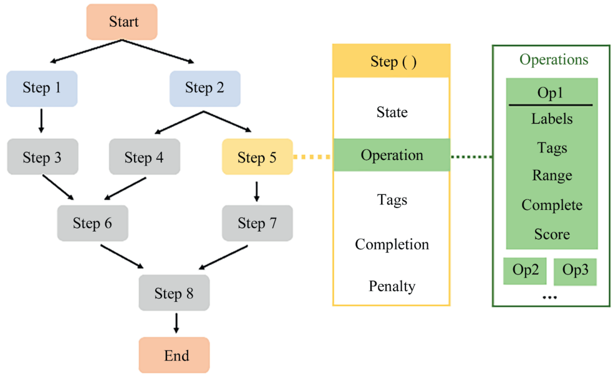
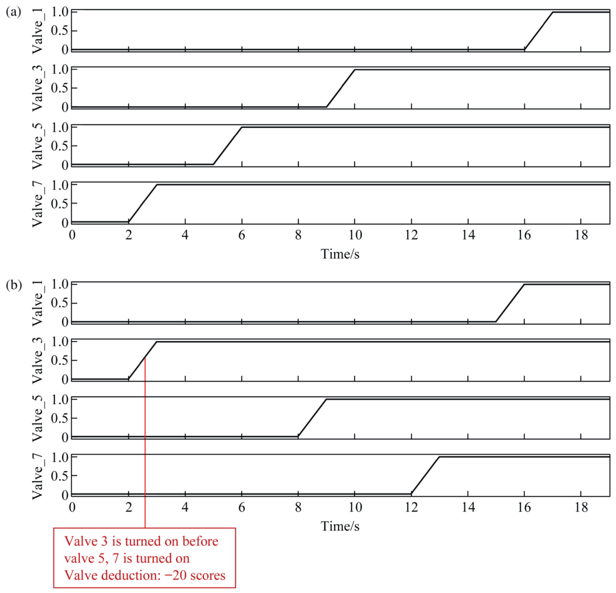


Fig. 14. The steps forming a test procedure in OTS.

With the predefined scripts for the model and steps, the OTS of the startup procedure is constructed and ready for operators. The operation tests done by operators will be recorded and scored. Fig. 15 presented a case of an incorrect operation sequence in the training test of the startup procedure. The step requires operators to turn on the circulation pump. The correct operating sequence is to turn on Valves 7, 5, 3, and 1, sequentially, as shown in subplot (a). The operator did a wrong operation with an incorrect sequence as shown in subplot (b). The evaluation system thus deducts 20 scores for the wrong operation according to the predefined rules in Step Lib.

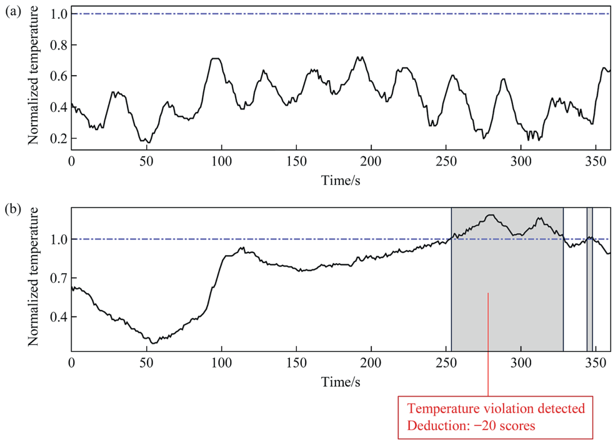
Fig. 16 presents another case of violating key constraints in the startup procedure. Note that the temperature in the figure has been normalized and scaled to the range of 0 to 1. The steps require the reactor temperature to be controlled below 1.0. Comparing to a standard operation recorded in historical data as shown in subplot (a), the operation done by the tested operator surpasses the limit as shown in subplot (b). As a penalty, a deduction of 20 scores is made by the evaluation system.

With the OTS, it is convenient for high-level managers and engineers to define specific scenarios according to actual needs, and train operators in a simulator environment. Besides, it can also



Valve 3 is turned on before valve 5, 7 is turned on
Valve deduction: −20 scores

Fig. 15. Standard operation and wrong operation for circulation pumps; (a) standard operation, (b) wrong operation.



Temperature violation detected
Deduction: −20 scores

Fig. 16. Standard operation and wrong operation for temperature control; (a) standard operation, (b) wrong operation.

reveal operators' level and take precautions against unsafe operations, which is essential for the high hazard equipment like the complex nitration process. OTS characterizes a working mode: by adding data structures based on steady-state and dynamic models, it enables the achievement of high-level task requirements during the service phase.

7. Conclusions

A hierarchical and role-driven framework of digital twin system for general process operation is presented in this work. With one digital twin, the framework is designed to address different demands from managers, engineers, and operators in the nitration industry. It consists of three modules, OAS, OMS, and OTS, handling tasks in design phase, runtime phase and service phase, respectively. The OAS creates the digital process using stream and unit objects, which can help high-level staff to make decisions on any process operation and optimization. The OMS monitors the dynamic changes in the operation with hybrid mechanistic and data-driven models, which can enhance the safety and profit of the process operation. The OTS creates a virtual training environment with different operating scenarios based on the models and historic data, which is useful for the operator to improve their operating skills. Communication between modules improves the efficiency of the digital twin system. Each module is illustrated with applications of different purposes in the real nitration process. These applications presented in this work demonstrate that the novel hierarchical framework could benefit the nitration process in different phase during the lifecycle, by the analysis, monitoring, and training tasks facilitated by one digital system.

Comparing to usual solutions to the tasks in industry, such as independent simulator or SCADA systems, an integrated digital twin system aiming for addressing issues of different roles at one time, thus playing a higher-level role than independent module. It ensures that the same modeling structure based on variables and equations is applied consistently across different application tasks. In processes with multiple tasks, the framework prevents redundant computation of data and incompatibility between different systems. Besides, the digital twin system can accommodate extension designs following the same structure oriented toward other tasks, such as other analysis or monitoring tasks.

Although the applications are within the nitration process, the framework is believed to be applicable universally for process operations. The framework proposed in this work is designed based on some basic concept: OAS is based on stream and unit objects; OMS is based on hybrid methods of static and dynamic models; and OTS is based on steps and single operation. Since these concepts are shared within process industry, the framework is general for process operation in most situations. Instead of how-to building models specifically, this work introduces a framework to build a specific digital twin system to arrange a large amount of data, aiming for addressing demands of different roles in general process operations.

CRediT Authorship Contribution Statement

Xinyu Tao: Writing – original draft, Software, Resources, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. Runjie Yao: Software, Methodology, Investigation. Lingyu Zhu: Visualization, Supervision, Methodology, Investigation, Conceptualization. Changrui Xie: Writing – original draft, Software, Methodology, Investigation. Muqian Zhang: Writing – original draft, Software, Methodology, Investigation, Conceptualization. Xi Chen: Writing – review & editing, Validation, Supervision, Methodology, Investigation, Conceptualization.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests that could have appeared to influence the work reported in this paper.

Acknowledgement

The financial support of the "Pioneer" and "Leading Goose" Research & Development Program of Zhejiang (2024C01028) and the State Key Laboratory of Industrial Control Technology, China (ICT2024C04) are gratefully acknowledged.

References

[1] M. Shafto, M. Conory, R. Dolye, E. Glaessgen, C. Kemp, J. LeMoigne, L. Wang, DRAFT Modeling, Simulation, Information Technology & Processing Roadamap, Technology Area 11, 2010.

[2] D. Jones, C. Snider, A. Nassehi, J. Yon, B. Hicks, Characterising the Digital Twin: a systematic literature review, *CIRP J. Manuf. Sci. Technol.* 29 (2020) 36–52.

[3] M. Grieves, J. Vickers, Digital twin: mitigating unpredictable, undesirable emergent behavior in complex systems, in: Transdisciplinary Perspectives on Complex Systems: New Findings and Approaches, Springer International Publishing, Cham, 2017, pp. 85–113.

[4] M.N. Liu, S.L. Fang, H.Y. Dong, C.Z. Xu, Review of digital twin about concepts, technologies, and industrial applications, *J. Manuf. Syst.* 58 (2021) 346–361.

[5] L. Peterson, I.V. Gosea, P. Benner, K. Sundmacher, Digital twins in process engineering: an overview on computational and numerical methods, *Comput. Chem. Eng.* 193 (2025) 108917.

[6] Q. Liu, H. Zhang, J.W. Leng, X. Chen, Digital twin-driven rapid individualised designing of automated flow-shop manufacturing system, *Int. J. Prod. Res.* 57 (12) (2019) 3903–3919.

[7] F. Chinesta, E. Cueto, E. Abisset-Chavanne, J.L. Duval, F. El Khaldi, Virtual, digital and hybrid twins: a new paradigm in data-based engineering and engineered data, *Arch. Comput. Methods Eng.* 27 (1) (2020) 105–134.

[8] X.M. Sun, J.S. Bao, J. Li, Y.M. Zhang, S.M. Liu, B. Zhou, A digital twin-driven approach for the assembly-commissioning of high precision products, *Robot. Comput. Integrated Manuf.* 61 (2020) 101839.

[9] C.Z. Li, S. Mahadevan, Y. Ling, S. Choze, L.P. Wang, Dynamic Bayesian network for aircraft wing health monitoring digital twin, *AIAA J.* 55 (3) (2017) 930–941.

[10] C.C. Pantelides, F.E. Pereira, P.J. Stanger, N.F. Thornhill, Process operations: from models and data to digital applications, *Comput. Chem. Eng.* 180 (2024) 108463.

[11] R. He, G.M. Chen, C. Dong, S.F. Sun, X.Y. Shen, Data-driven digital twin technology for optimized control in process systems, *ISA Trans.* 95 (2019) 221–234.

[12] M. El Wajeh, A. Mhamdi, A. Mitsos, Dynamic modeling and plantwide control of a production process for biodiesel and glycerol, *Ind. Eng. Chem. Res.* 62 (27) (2023) 10559–10576.

[13] S. Spatenka, M. Matzopoulos, Z. Urban, A. Cano, From laboratory to industrial operation: model-based digital design and optimization of fixed-bed catalytic reactors, *Ind. Eng. Chem. Res.* 58 (28) (2019) 12571–12585.

[14] J.X. Yu, P. Liu, Z. Li, Data reconciliation-based simulation of thermal power plants for performance estimation and digital twin development, *Comput. Chem. Eng.* 168 (2022) 108063.

[15] S.D. Gao, C.M. Bo, C. Jiang, Q.L. Zhang, G.K. Yang, J. Chu, Hybrid modeling for carbon monoxide gas-phase catalytic coupling to synthesize dimethyl oxalate process, *Chin. J. Chem. Eng.* 70 (2024) 234–250.

[16] W. Wei, P.F. Xia, Z.W. Liu, M. Zuo, A modified active disturbance rejection control for a wastewater treatment process, *Chin. J. Chem. Eng.* 28 (10) (2020) 2607–2619.

[17] M. Vaccari, R.B. di Capaci, E. Brunazzi, L. Tognotti, P. Pierno, R. Vagheggi, G. Pannocchia, Optimally managing chemical plant operations: an example oriented by industry 4.0 paradigms, *Ind. Eng. Chem. Res.* 60 (21) (2021) 7853–7867.

[18] P.A. Quadros, N.M.C. Oliveira, C.M.S.G. Baptista, Continuous adiabatic industrial benzene nitration with mixed acid at a pilot plant scale, *Chem. Eng. J.* 108 (1–2) (2005) 1–11.

[19] C.R. Xie, R.J. Yao, L.Y. Zhu, H. Gong, H.Y. Li, X. Chen, Hybrid dynamic modeling of an industrial reactor network with first-principles and data-driven approaches, *Chem. Eng. Sci.* 289 (2024) 119852.

[20] L.T. Biegler, I.E. Grossmann, A. Westerberg, Systematic methods of chemical process design. Mass and Energy Balances, 1997.