IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 31 July 2023, accepted 11 September 2023, date of publication 15 September 2023, date of current version 3 October 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3316028

RESEARCH ARTICLE

Enhancing Operational Resilience of Critical Infrastructure Processes Through Chaos Engineering

PANAGIOTIS DEDOUSIS[1], GEORGE STERGIOPOULOS[2], GEORGE ARAMPATZIS[3], AND DIMITRIS GRITZALIS[1]

[1]Department of Informatics, Athens University of Economics and Business, GR-10434 Athens, Greece
[2]Department of Information and Communication Systems Engineering, University of the Aegean, GR-83200 Samos, Greece
[3]School of Production Engineering and Management, Technical University of Crete, GR-73100 Chania, Greece

Corresponding author: Dimitris Gritzalis (dgrit@aueb.gr)

This work was supported in part by the Hellenic Ministry of Digital Governance through the Research Grant Offered to the Research Center of the Athens University of Economics and Business.

ABSTRACT Modern cyber-physical systems (CPS) are interdependent, mechanical and IT components that support operations in most of society's critical infrastructures. Time and again history has proven that CPS are vulnerable to numerous types of threats, ranging from safety accidents to cybersecurity malicious attacks. Current research focuses on analyzing the various input vectors in CPS, whether mechanical or IT, to detect and patch flaws and vulnerabilities to mitigate potential impact in operation. Still, there is little work that can inherently analyze the software-based implementation of modern CPS with complex behavior and failure modes. The only vaguely relevant approach involves an operations-based experimentation methodology from NetFlix named chaos engineering (CE) that tests use cases on complex Content Delivery Networks (CDN) to build confidence in their capability to withstand turbulent conditions in production. Conditions can range from hardware failures to DoS attacks, to a malformed injection appearing in a runtime configuration parameter. Yet this approach was only tested on software based CDN, and not on CPS with industrial actuators and mechanical parts that control physical processes. In this paper, we introduce a novel framework that combines CE with digital twin (DT) technology to enhance the detection of operational vulnerabilities and increase the resilience of CPS. To achieve this objective, we integrated CE experimentation into the simulation phase of DT models using material flow networks. This allows us to assess system resilience during the operational stage without disrupting critical operations and identify vulnerable flows and processes within the modeled system. To evaluate the effectiveness of our approach, we conduct experiments on a DT that models a real-world Liquefied Petroleum Gas purification process from an existing oil refinery in the Mediterranean area. The results of these experiments demonstrate the method's effectiveness in capturing the heightened susceptibility of the Gas purification process to adverse events.

INDEX TERMS Cyber-physical systems, critical infrastructure protection, chaos engineering, digital twins, resilience, simulation, LPG purification process.

I. INTRODUCTION

Cyber-physical systems (CPS) are complex systems that integrate critical physical processes, computation, networking, and information and communication technologies (ICT) [1]. A key requirement for CPS is increased resilience during operation since most CPS support industrial plants and critical infrastructure operations. Resilience is defined as "the ability of a system to withstand various disruptions, such as natural disasters, cyber-attacks, equipment failures, or human

The associate editor coordinating the review of this manuscript and approving it for publication was Agostino Forestiero.

errors, by maintaining an acceptable level of performance, mitigating the severity of unstable states, and responding to disruptive events'' [2].

In spite of their importance, safety and (cyber)security concerns continue to hinder their deployment [3], [4]. Recent cyber-physical incidents, such as the Colonial Pipeline shutdown and Oldsmar's water treatment cyberattack, have highlighted the vulnerabilities of CPS, resulting in economic, social, legal, and political consequences [5], [6].

One of the critical challenges in enhancing CPS resilience lies in effectively integrating various resilience technologies to enhance the system's ability to withstand interference while minimizing the impact on system availability [6]. Current research focuses on analyzing the various input vectors in CPS, whether mechanical or IT, to detect and patch flaws and vulnerabilities to mitigate potential impact in operation. Still, there is little work that can inherently analyze the software-based implementation of modern CPS with complex behavior and failure modes. Some techniques exist to assess the resilience of physical systems from an engineering perspective, including simulation approaches that often rely on experts to input new parameters, and experimentation methods, such as hardware-in-the-loop experiments [7], [8], [9]. These are typically specific to the application and safety scenarios under test and do not follow systematic experimentation steps nor can they detect vulnerabilities that stem from the complex nature of such systems [10], [11], [12], [13].

A. THE CHAOS ENGINEERING METHODS

To our knowledge, the only work that is slightly relevant to the presented approach involves an operations-based experimentation methodology from Netflix named chaos engineering (CE) that tests use-cases on complex Content Delivery Networks (CDNs) to build confidence in their capability to withstand turbulent conditions in production [14]. Conditions can range from hardware failures to DoS attacks, to a malformed injection appearing in a runtime configuration parameter. Yet, this approach was only tested on software-based CDN, and not on CPS with industrial actuators and mechanical parts that control physical processes.

Our work demonstrates the potential benefits of using CE to enhance the operational resilience of CPS, providing decision makers with valuable insights into the weaknesses of the system.

The CE process involves three basic steps to control testing (see Fig. 1): (i) Steady state definition, (ii) hypothesis formulation, and (iii) experimentation.

The first step in any CE method [15], [16] is to describe and model the steady state of the system or under test, which serves as a baseline for its expected behavior under a wide range of normal conditions. The second step involves the formulation of a hypothesis based on this steady-state using measurable system outputs. The third step involves designing the experiment and varying the inputs to the system to reflect realistic adverse conditions. Finally, the experiments

FIGURE 1. Key CE experimentation steps.

are run within the system's production environment to test the hypothesis and identify any differences between the steady state and experimental conditions. The main goal of these experiments is to disprove the hypothesis, highlighting any weaknesses or vulnerabilities in the system that need to be addressed to enhance its overall reliability and resilience.

B. DIGITAL TWINS AND MATERIAL FLOW NETWORKS

During the experimentation stage, the CE methodology requires tests to be conducted in the production environment. However, testing chaotic variations in real-world critical processes is infeasible due to complexity and safety concerns. Therefore, we performed our CE experiments on a DT model of the physical system. A DT is a virtual replica of a physical product or system used to understand and anticipate the behavior of the original copy [17]. The DT model is built using Material Flow Networks (MFN), which are directed graphs with nodes representing the locations of material and energy transformations and storage areas within a production line. The edges connecting these vertices represent the material and energy flows between them [18], [19]. The DT is frequently updated with real-time information from the physical system to enable nondisruptive real-time testing, allowing for proactive observation of the physical system's behavior and performance.

C. CONTRIBUTION

Our study builds upon previous research on applying CE principles to CPS [14] and aims to demonstrate the benefits of using CE to enhance a CPS's operational resilience. To achieve this objective, we integrate CE experimentation [15], [16], into the simulation phase of DT models using material flow networks (MFNs). This allows us to assess system resilience during the operational stage without disrupting critical operations and identify vulnerable flows and processes within the modeled system.

To evaluate the effectiveness of the framework, we conduct experiments on a DT that models a real-world Liquefied Petroleum Gas (LPG) purification process from an existing oil refinery in the Mediterranean area. The results of these experiments demonstrate the methods effectiveness to capture the heightened susceptibility of the Gas purification process to adverse events.

In summary, our study makes the following contributions:

1. We propose a novel framework that combines DTs and CE using an MFNs to evaluate and improve CPS resilience in CPS without affecting the physical

system's normal operations and availability. Specifically, our approach:

   a. Integrates CE experimentation into the simulation phase of DT models to assess the system's resilience during the operational stage of CPS using real-time data.
   b. Facilitates the identification and prioritization of vulnerable flows and network nodes, allowing operators to prioritize their security efforts and develop targeted mitigation measures.
   c. Supports the validation of the developed mitigation strategies and contingency plans through continuous evaluation and analysis of system behavior, ensuring the seamless operation of the system.

2. We validate the effectiveness and reliability of the proposed approach through CE experiments conducted on a real-world LPG purification unit from a plant operating in the wider Mediterranean area.

To the best of our knowledge, this study is the first to combine real-world plant DT and CE using MFNs to evaluate and improve CPS resilience.

D. STRUCTURE

The rest of the paper is organized as follows. Section II discusses related work and compares approaches, methods, and frameworks addressing CPS resilience. Section III describes the proposed CE application framework. Section IV presents an implementation of the methodology in a real-world example and discusses our findings to validate the methodology. Finally, the results and potential future research are discussed in Section V.

II. RELATED WORK

In this section, we examine various approaches, techniques, and algorithms found in the literature to assess the resilience of CPS. The primary objective of these methodologies is to evaluate and understand the effects of adverse events, including cyber threats, and the multidimensional impact of disruptive incidents on the resilience of CPS across various sectors [7], [20].

Resilient CPSs require appropriate metrics for evaluating their critical components. Reference [21] suggests a resilience metric that quantifies the ability of the system to recover from an attack within a fixed time interval, as well as the cost of recovery. Although this metric provides real-time awareness of the system's state, it focuses mainly on recovery and overlooks other essential capabilities when evaluating CPS resilience. In contrast, [22] proposed a resilience metric that considers the resist, sustain, and recovery phases of an underlying CPS. This metric enables system operators to take corrective actions to minimize damage and enhance resilience. However, it primarily emphasizes the resilience of the ICS network against cyber threats, neglecting the assessment of the resilience, performance, and operational aspects of the physical system. Furthermore, the focus on cyber threats limits its assessment to incomplete cyber-physical threat scenarios, thus restricting the comprehensive evaluation of the system's overall resilience and potential vulnerabilities.

Graph theory is commonly employed to model CPS and identify vulnerabilities in ICS. In [23], a framework was proposed for automating security risk analysis and restructuring complex interconnected sensors and devices, which identifies critical components using dependency risk graphs, graph minimum spanning trees, and network centrality metrics. Similarly, in [24], a graph theory model was developed to evaluate the security and resilience of a naval water distribution CPS by assessing the cascading impact of different anomalies. However, both these methods lack quantitative analysis to measure the specific impact of these anomalies on the system output(s) and performance. Without quantitative analysis, it is challenging to set appropriate resilience targets and identify areas for resilience improvement. Furthermore, graph-based modeling methods face the challenge of determining the appropriate level of detail for representing the component relationships.

Simulation-based approaches have been developed to analyze the reliability and resilience of CPS under attack scenarios. For example, [25] proposed a Monte Carlo simulation method to model an ICS threatened by information attacks, which can significantly disrupt train operations. However, one challenge with statistical model checking is estimating the probability of rare events, which often is unfeasible. Other approaches utilize data analysis, machine learning, and simulations to ensure CPS can withstand adverse events [26]. For instance, in [27], the authors suggested using DT and machine learning (ML) algorithms to simulate an operational CPS and perform real-time analysis for the early detection and mitigation of cyber anomalies and threats. However, ML algorithms trained on the steady system state may face challenges in modeling system processes with values that deviate from normal operating conditions (i.e., adverse events), as they exhibit limitations, including a bias towards normalcy and limited extrapolation capability. Recent approaches have advanced the concept of DTs by incorporating fundamental aspects of cognition, leading to cognitive DTs that facilitate production resilience [28]. Although this modeling approach shares similarities with our methodology, it focuses on ensuring resilience by optimizing the system rather than proactively studying and evaluating resilience by identifying different system states before, during, and after disruptions.

Numerous techniques exist to detect faults and attacks in ICS. Dynamic fault testing approaches involve targeted technical attacks using test inputs to detect program execution errors and understand the causal relationships among events in Programmable Logic Controller (PLC) code, as demonstrated in [29]. Other approaches deploy network-protocol attacks and monitor target systems for unexpected behaviors [30]. Similarly, [31] suggested an efficient and accurate method for estimating performance errors caused by

denial-of-service (DoS) attacks in networked controlled systems. In [32], a new method was introduced to account for delays in ICS components, proposing an algorithm to ensure system stability and dissipation using fuzzy logic. Mining invariant values from ICS devices has also been used for fault detection. Recent approaches have aimed to extract operational conditions from system logs using data mining [33]. However, most proposed approaches use sensor data as datasets; therefore, manipulating sensors and data is a concern, as implicit confidence in the obtained data may lead to cyber incidents.

CE experiments address the limitations of functional charts and specifications for capturing the scale and complexity of user behavior in modern distributed systems. In addition, by treating the entire system under test as a single entity, CE experiments uncovered previously unknown and unwanted execution states [15]. While CE experiments may employ fuzzing techniques for input, they do not target specific software or components but instead consider all ICS components and field devices as a unified system, observing the effects of real-world, high-level input on the ICS boundary. In addition, unlike fault analysis and testing techniques that rely on binary assertion logic or predefined restrictions, CE experiments generate new input data and identify previously unrecognized system states, diverging from the traditional true/false evaluation approach [15], [34].

Regarding existing CE approaches, the authors of [15] first introduced CE techniques to verify the reliability of distributed systems experimentally. They achieved this by manipulating the boundary state of the components and analyzing system behavior using an internally developed service called Chaos Monkey. The concept of Chaos Monkey was also utilized in [35], where researchers proposed a balanced use of the service to introduce varying levels of failure into the network while maintaining its connection and evaluating its performance based on network-invariant metrics. Other research has focused on using CE to analyze the execution states of infrastructure-as-a-service (IaaS) cloud platforms. For instance, CloudStrike implemented CE principles by introducing system failures and defects to cloud resources to study cybersecurity breaches caused by human errors and misconfigured resources [36], [37]. In [38], the authors introduced an information technology (IT) service management framework that integrates CE techniques with DTs to test the resiliency of complex IT services deployed in hybrid cloud scenarios. This integration allowed them to assess the robustness of deployed IT services and identify potential weaknesses. Overall, CE techniques have proven effective in identifying vulnerabilities, enhancing system resilience, and providing valuable insights into IT and software systems.

The use of CE experiments in industrial CPS is new, and to the best of our knowledge, it was first introduced in [14]. In their approach, the authors effectively applied CE principles to an industrial CPS testbed, demonstrating its capability to predict environmental changes and implement mitigation measures that control the severity of adverse events. However, significant challenges still need to be addressed when conducting CE experiments on real-world operational CPS without disrupting normal operations and system availability. Similar to the work in [14], our approach employs CE to assess industrial CPS resilience by conducting experiments that introduce random and unpredictable behaviors into the system. We built upon this previous work by incorporating MFN to create a DT model of the production system. DT enable us to execute CE experiments in a safe and controlled environment without any negative impact on the operating CPS [28]. Additionally, drawing on insights from [39], the use of MFN enables us to model the underlying CPS, simulate real-world operational conditions, and quantitatively evaluate system responses and performance [18], [19]. Finally, acknowledging the importance of optimal control strategies for CPS resilience [40], we integrate this process into a framework that guides decision-making for resilience mitigation, enabling the enhancement of CPS resilience.

Consequently, existing approaches primarily focus on analyzing faults and their impacts on ICS components and networks. However, they lack a comprehensive evaluation of industrial CPS resilience, including proactive testing, cyber-physical threat scenarios, and quantitative insights into physical processes and system outputs. A more holistic approach is needed to bridge these gaps and analyze the resilience of industrial CPS against adverse events under real-world conditions. To address these challenges, we present a resilience management framework that integrates into operating industrial CPS. Our approach simulates interconnected physical components by considering potential failures and flow disruptions caused by cyber-physical threats. With the ability to quantitatively evaluate the system responses and performance, our approach facilitates the identification and prioritization of vulnerable flows and processes, enabling operators to prioritize their security efforts and develop targeted mitigation measures.

III. A CHAOS ENGINEERING FRAMEWORK FOR LPG SYSTEMS

The work presented in this paper demonstrates the use of CE by testing a system that purifies LPG, a valuable energy product consisting of a mixture of liquefied hydrocarbon gases C3-C4 (propane and butane). LPG is widely used in various industries and in transportation. It is produced as a by-product of refinery processes such as Crude Distillation (CDU), Hydrocracking (HYC), and Fluid Catalytic Cracking (FCC). However, LPG contains impurities that must be removed through purification. The main purification processes involve:

(i) removing naphtha (C5) using debutanizer columns,
(ii) removing ethane using deathanizer columns, and
(iii) removing hydrogen sulfur compounds (H2S).

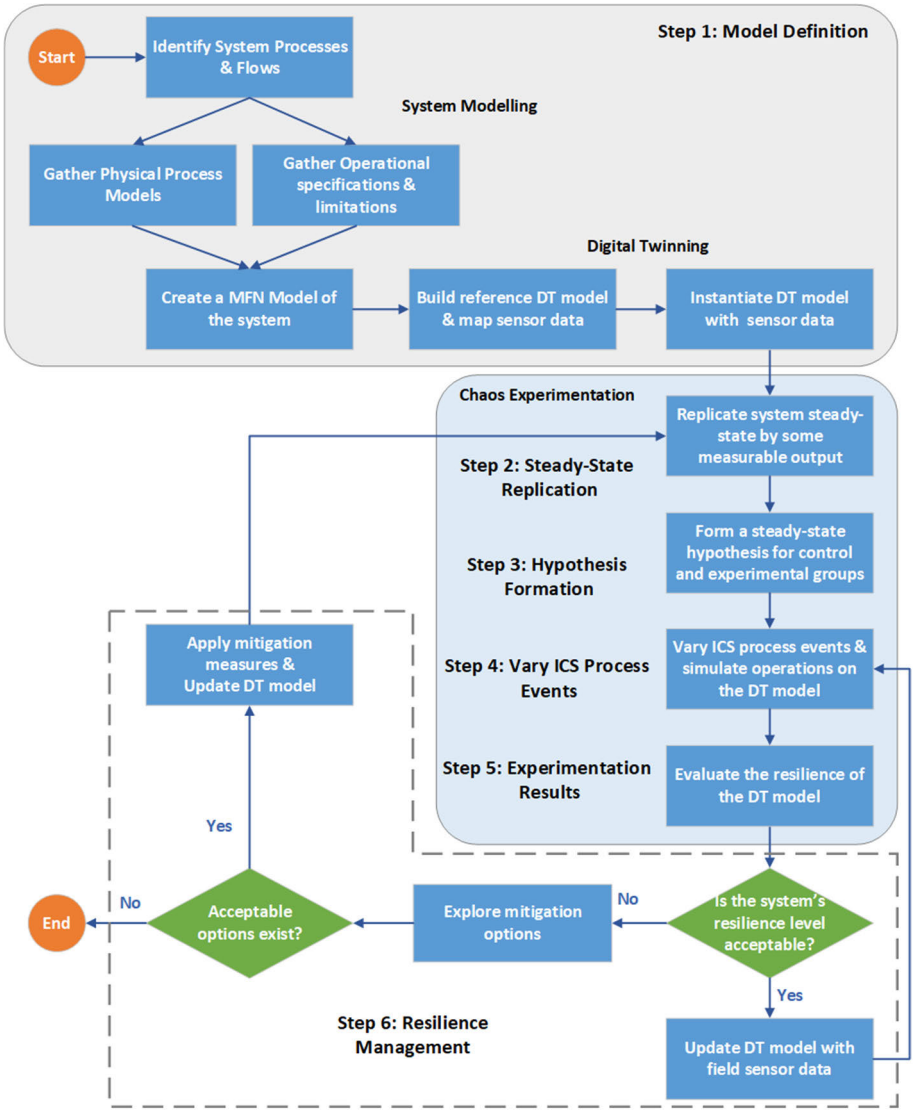

FIGURE 2. High-level overview of the proposed CE application framework for chemical processes (i.e., LPG purification processes).

Our framework models this process into a DT and then implements the CE methodology onto this LPG purification process using the following steps (Fig. 2):

1. Initially, we identify the system processes, specifications, and operational limits from the plant and create an MFN that enables the simulation of the CPS operation.
2. Next, we code a DT model of the physical system by mapping the sensor data and control inputs to the processes within the MFN. The DT model replicates the steady-state and nominal operation of the system.
    a. The sensor data, control inputs and mathematical model for the LPG mechanical process serve as a reference point for subsequent analysis and comparison with adverse conditions.
    b. The DT model is instantiated with ICS process signals to simulate the steady state of the physical system and establish a baseline understanding of its expected behavior under normal conditions.
3. Next, we formulate potential hypotheses on the created DT.
    a. Various hypotheses are developed as scenarios to test the system's behavior under steady-state conditions, as modeled during step 2.
4. We experiment on the DT using the formulated hypotheses. During the experimentation, the following automated procedures are performed:
    a. We execute these hypotheses in the form of use cases by automatically modifying ICS process signals to observe and analyze the DT's response to the introduced adverse events, including evaluating stability and indications of vulnerability or failure.

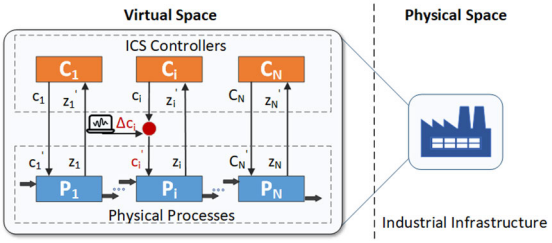

FIGURE 3. Conceptual representation of CPS. We assume that the system is in a steady-state, with no anomalies in the send-control and received-measurement signals, c'_i and z'_i, respectively, when $c'_i = c$ and $z'_i = z_i, i \in \Theta$.

The performance of our approach relies on the computational complexity of the mass balance calculations performed during the simulation of the physical system's operation using the MFN during Step 2 (see Section III-A for more details). Fig. 2 illustrates the overall steps and workflow of the proposed roadmap for implementing CE using DTs in industrial processes. This roadmap guides engineers and experts to evaluate resilience during the operational phase of CPS and identify weak points and vulnerabilities for implementing mitigation measures. It is important to note that the development of mitigating measures and the improvement of the system is not covered in this study. However, it is intended that the results of this study are used as input for those steps. In the following sections, we discuss how we incorporate CE experimentation into the simulation phase of DT models in alignment with the fundamental principles of CE.

A. STEP 1: STEADY STATE DEFINITION

The first step in implementing CE is to determine the steady state of the system. This involves establishing setpoints and relevant flow states to maintain specific system properties within a defined range. To do this, we need to identify the system boundaries that define the processes requiring analysis in terms of material and energy flows. Our objective is to model the processes and flows of the system, simulate its operation, and accurately define its steady state. In Figure 3, we illustrate a generic LPG purification infrastructure with a number (N) of interconnected processes ($P_i, i \in \Theta$).

These processes are assumed to be controlled by local controllers ($C_i, i \in \Theta$). In this system, the controllers send control-vector signals $c = [c_1 \dots c_N]^T$ to the processes and receive sensor measurements $z = [z_1 \dots z_N]^T$ from them.

The CPS model can be seen as a nonlinear equation system in the following state-space form:

$$\dot{x}(t) = g(x(t), c(t), z(t)) \tag{1}$$

$$\Delta y(t) = h(x(t), c(t), z(t)) \tag{2}$$

where the state $x(t) \in \mathbb{R}^n$ with $x(0) = x_0$, the measurement output deviations $\Delta y(t) \in \mathbb{R}^m$, the control input values $c(t) \in \mathbb{R}^m$ that represent the set points for state regulation during LPG purification, and $z(t) \in \mathbb{R}^n$ denotes the vector of measurement signals including disturbances, measurement noise, and unknown control variables. The system is assumed to be defined over domains $\mathcal{V} \subseteq \mathbb{R}^m \times \mathbb{R}^n$ and $\mathcal{W} \subseteq \mathbb{R}^n$ for which [14]:

- g and h are Lipschitz continuous functions on $\mathcal{W} \times \mathcal{V}$,
- a Lipschitz is a continuously differentiable function φ_w on $\mathcal{W}$ where $\varphi_w : \mathcal{V} \to \mathcal{W}$ and $\forall (c, d) \in \mathcal{V}$ satisfies $g(\varphi_w(c, z), c, z) = 0$, with $\varphi_w(c, z) \in \mathcal{W}$
- there exist constants $c_j > 0, j \in \mathbb{N} : j \in [1, 4]$, and
- a Lyapunov continuously differentiable function f in $(x, c), f : \mathcal{W} \times \mathcal{V} \to \mathbb{R}_{\geq 0}, (x, (c, z)) \longmapsto f(x, c, z)$ for which $\forall x \in \mathcal{W}$ and $\forall (c, z) \in \mathcal{W}$.

The above formulation forms a singular perturbation problem in which relaxation and other variants exist. Additionally, the input-output equilibrium mapping $\Delta\bar{y}(c, z) = h(\varphi_w(c, z), c, w) : \mathcal{V} \to \mathbb{R}^m$ has the form of:

$$h(\varphi_w(c, z), c, w) = \left(\frac{1}{\gamma}\right) * \mathbb{I}_m * \left(\mathbb{I}_m^\top \bar{c} - d\right) \tag{3}$$

where $\gamma > 0$, $\mathbb{I}_m$ represents the indicator function of size m, and $d \in \mathbb{R}$ is an unmeasured constant disturbance. Eq. (3) demonstrates that synchronization of output measurements is possible in steady-state conditions, with the deviation being the same at all ICS components [14].

Steady State in LPG Purification:

In our approach, we establish the steady state of the system by considering a network of processes that represent the LPG purification unit. This network is a directed graph $G = (V, E, A)$ with nodes $V = \{1, \dots, 12\}$, which consists of two hydrocrackers, one fluid catalytic cracker, three debutanizers, two deethanizers, three DEA absorption units, and one storage tank. The edges $E \subseteq V \times V$ within the graph indicate the flows of various resources (i.e., materials and energy) and capture the consumption rates (i.e., input flows) and production rates (i.e., output flows) of these resources. To quantify the material flows within the LPG purification unit, such as LPG, C2, C5, and S, we expressed them as volumetric or mass flow rates. These flow rates are measured in cubic meters per second (m^3/s) or kilograms per second (kg/s). By augmenting this graph with resource flows, we create an MFN that accurately models the LPG purification unit and its processes.

To determine the material flow rates of the LPG purification processes, we combine the debutanizer and deethanizer into a single column with one feed and two products, along with a reboiler and condenser. To achieve this, we employ a mechanistic model based on first principles, which incorporates a set of algebraic equations known as MESH equations. The MESH equations, introduced by Wang and Henke [41], consist of material balances (M), equilibrium relationships (E), summation equations (S), and heat or enthalpy balances (H). These equations are provided in detail in Appendix A. In addition, we refer to the debutanizer/deethanizer specification data (Table 1).

Using the input flow of LPG, along with temperature, pressure, and reflux rate, the model calculates the resulting

TABLE 1. Debutanizer-3 (P5) specification data.

Specification	Value
Number of trays	35
Feed tray-stage number	25
Type of tray used	Valve
Column Diameter	1.4 m
Column Length	22.9 5m
Type of Condenser	Partial

flow rate of purified LPG and the remaining impurity flows (such as C2 or C5).

The flow rate of a material (e.g., C2) in the distillation column depends on the flow rates of the input materials and process variables (i.e., temperature, pressure, and reflow variables). Therefore, we can consider the flow rate of the material G as the output of a distillation column with a single input flow (i.e., LPG mixed with impurities) $c_k > 0, c_k \in \mathbb{N}$ and three distillation variables: $c_p > 0, c_p \in \mathbb{N} : \mathrm{p} \in [1, 3]$. This is achieved by controlling the process using a vector signal $c = [c_1 \dots c_4]^T$ and by receiving sensor measurements $z = [z_1 \dots z_4]^T$ from the process.

Output deviations $\Delta y\,(t)$ in the output material G and control inputs $c_j\,(t)$ for state regulation are defined as members of the same set $\mathbb{R}^m$. $\mathcal{V} \subseteq \mathbb{R}^m \times \mathbb{R}^n$ includes the nominal control input ranges that induce acceptable states/members of the set $\mathcal{W} \subseteq \mathbb{R}^n$. These control input ranges are limited by the constraints of the LPG purification processes, which operate within specific domains as shown in Table 2. Table 2 presents the shutdown ranges for Debutanizer-3 (P5), which represent the limitations on the physical aspects of the LPG purification unit. Therefore, at each time t, the state $x_j\,(t)$ of the *jth* process (e.g., Debutanizer-3) that participates in the LPG purification exists within the state space $\mathbb{R}^n$, which is represented by the normal and shutdown values of the four control vector signals that control the process variables and input flow rate (Table 2).

The process variables and operating constraints listed in Table 2 were determined through a combination of manufacturer data, past sensor readings, and information provided by the oil refinery plant operators and engineers. The selection of these variables was based on the capabilities of the physical system. Typically, variables that are related to monitoring, control, and optimization of the process are considered suitable candidates, as they directly influence its operation and performance.

The shutdown operating limits for each process are established based on the equipment specifications and operating guidelines provided by the manufacturer. These specifications include the recommended operating ranges and limits for various process variables such as temperature, pressure, flow rates, and other relevant parameters.

The normal operating limits for each process were derived from past sensor readings. These readings provide historical data on the range of values observed during normal operation. By analyzing these readings, we can establish the typical

TABLE 2. Debutanizer-3 (P5) operating constraints.

Process variables	Normal operating limits		Shutdown operating limits	
	Low limit	High limit	Low limit	High limit
Temperature	60 °C	70 °C	20 °C	140 °C
Pressure	7 kg/cm^2	9 kg/cm^2	0	20 kg/cm^2
Reflow	5400 m^3/h	6000 m^3/h	0	7000 m^3/h
LPG input flow rate	3200 m^3/h	4000 m^3/h	0	5000 m^3/h

boundaries within which the process variables operate under normal conditions.

B. STEP 2: DT DEVELOPMENT FOR STEADY STATE SIMULATION

To simulate the operation model of the physical system described above, we utilize a DT in the form of a solver that enables us to compute the flows for all the modeled processes within the system. This includes tracking the production flows of the system output over time, which allows us to capture the state of the system. Our solver employs an iterative approach to implement and analyze the flows within the Mass Flow Network (MFN). The algorithmic steps used in our solver are as follows:

1. Mass balance calculation: For each process in the MFN the solver performs a mass balance calculation using specified equations. This calculation ensures that the inputs and outputs of each process are properly accounted for in terms of mass.
2. Energy balance calculation: In addition to the mass balance, the solver computes the energy balance for each process in the MFN. This calculation takes into account the energy inputs and outputs of each process, utilizing the defined equations.
3. Flow aggregation: Following the calculation of the mass and energy balances for each process, our solver consolidates the results to provide an overview of the material and energy flows throughout the MFN. This step combines the individual process calculations to give a comprehensive view of the overall flow patterns within the system.

The overall time complexity required to solve all processes in the network depends on the complexity of each process. For instance, if each process takes a constant time to calculate its outputs using simple equations with constant time complexity, the total time complexity to solve all processes would be O(P + F), where P is the number of processes in the network, and F is the number of flow dependencies between the processes.

However, if each process employs nonlinear equations and iterative methods, such as Newton's method, to calculate output flows, the time complexity can be approximated as $O(I * P * F * n^2)$, where I represents the number of iterations

required for convergence, P represents the number of processes, F is the number of flow dependencies between the processes, and n represents the size of the problem (i.e., number of variables or equations).

In a CE experiment, operational metrics such as the input or output flow rate of a process material (e.g., C5 and C2) can be utilized to monitor possible deviations y_t during the process operation. Similar metrics can also be used to capture the state of a system, which encompasses a network of processes. In our case study, the main activity was the removal of impurities from the LPG. Therefore, the concentrations of C2, C5, and S in the tank are suitable candidates for the CE operational metrics. Because we focused on the debutanizer and deethanizer columns, we utilized the C2 and C5 concentrations in the tank to capture the steady state of the system and detect possible deviations during our experimentation. Specifically, the C2 and C5 concentrations in the tank were calculated using steady operating values for the system process variables and input flows. To that end, variations within acceptable ranges may still occur in the C2 and C5 concentrations of the system, even if the process operating variables are within their normal limits (Table 2).

C. STEP 3: FORMULATING HYPOTHESES ON THE DT

Once the metrics and steady-state behavior of the ICS are determined and simulated, the next step is to define explicit hypotheses that clearly articulate the expected outcomes of the experiment, focusing on the selected metrics of C2 and C5 concentrations. The objective is to assess the system's response when different events are introduced into the ICS of the LPG purification unit.

In our experimentation, we formulate a CE hypothesis by assuming that the controllers of the LPG purification processes follow a Distributed Averaging Proportional Integral (DAPI) control, which is often deployed in multi-agent networked systems to control a system's output towards a consensus objective based on the error between the desired value and the actual value [42]. This requires that state vectors $x_j(t)$ are considered as a linear combination of control inputs $c_j(t)$, where $j \in N = \{1, \ldots, n\}$, and N represents the index set of the processes. The outputs $\Delta y(t)$ are observed with a sampling period of $\Delta t = 1$ sec, computed from data from the initial startup of the LPG purification unit.

Based on these assumptions, we hypothesize that, *given that the system's control inputs remain within their shutdown operational limits, the concentrations of C2 and C5 in the final LPG output flow will remain within acceptable bounds (i.e., the system will remain stable)* (see Table 3). Table 3 lists the product specifications of the LPG mixture. C5 hydrocarbons must contain no more than 2% of the LPG volume and the sum of C2 + C5 hydrocarbons must not exceed 5% of the LPG volume.

Since we are interested in the concentrations of C2 and C5 in the storage tank, we introduce a key performance indicator, denoted as Impurity Concentration Divergence (ICD),

TABLE 3. Product specification thresholds of LPG.

Specification	Threshold
Pentane (C5)	2% vol/vol
Sulfur (S)	50 mg/kg
Hydrogen Sulfide (H2S)	Negative
Total Impurities (C2 + C5)	5% vol/vol

to quantify the difference between the steady-state of the system and the experimental outcomes:

$$ICD(t) = 0.5\left(\frac{C_{C2}(t) - C'_{C2}(t)}{C'_{C2}(t)}\right) + 0.5\left(\frac{C_{C5}(t) - C'_{C5}(t)}{C'_{C5}(t)}\right) \quad (4)$$

where C_{C2}, C_{C5} represent the steady-state concentrations of C2 and C5 in the storage tank and C'_{C2}, C'_{C5} represent the concentrations of C2 and C5 during the experiment. When the system is in a normal operating condition, the ICD will be close to zero because there will be no or negligible deviation between the steady-state concentrations of C2 and C5 and those observed during the experiment. However, any variation from the steady state (i.e., major deviation in the ICD value) indicates that the system has moved towards an unwanted state with a deviation in production output (i.e., the concentrations of C2 and C5).

The concentration of each impurity in the final LPG output stored in the tank can be represented as a single output vector z_t in $\mathbb{R}^m$. These vectors are determined using the MESH equations based on the constraints and specification data provided in Tables 1-2. This formulation leads to a closed-form solution, ensuring that the system behavior remains close to steady-state conditions unless an adverse event affects the system. Therefore, *we utilize the ICD indicator to detect any deviation, positive or negative, from any steady-state operation of the system toward a state that does not satisfy the nominal system output.* In the case of an adverse event, the ICD indicator will allow us to detect any deviation in the concentrations of C2 and C5 during our experiment, and validate our hypothesis, thus, the assumption of stability.

D. STEP 4: TESTING HYPOTHESES AS EVENTS

Lastly, we want to test executions of processes with the aim to intentionally induce controlled failures or disruptions in the system and subsequently analyze the resulting outcomes. An example list of such experiments includes but is not limited to the following:

- Physical process failures
- Overload, or failure (either software or mechanical) of individual components (e.g., RTU) to test the overall system reaction
- Malicious data injections on sensory inputs or inconsistencies in measurement signals

- Environmental disturbances impacting the system's operation (e.g., temperature variations, electromagnetic interference)
- Human operator errors or incorrect configuration settings

Such disruptions can be simulated by manipulating the system's control signals, c_i'. We introduce an offset Δc_i to the control signals c_i corresponding to a process variable. Consequently, a vector $\Delta c = [\Delta c_1 \dots \Delta c_N]^T$ is defined, where the entries are nonzero for the manipulated signal and zero for all other signals. By incorporating this manipulation, a general model for the control signals received by the control elements of the processes can be expressed as $c_i' = c_i + \Delta c_i$.

Typical disruption experiments last 24-48 hours to capture information and calculate the relevant CE operational metrics [43]. In our case, the CE experiment lasted 24 h, whereas the control signal manipulation started at t_0 and lasted 12 h. During the simulation, each system process evolved according to the state vector $x(t)$ and control signals $c(t)$. Using (5), we can effectively simulate fluctuations as step changes in a chemical process [14], [44]. By manipulating the ICS control signals in this manner, we can simulate the effects of an adverse event that causes the system to enter an unstable or undesired state. The $\Delta c_i(t)$ is calculated as the difference between the nominal control signal $c_i(t)$ and a random value $R(t)$ selected from the range between the upper and lower shutdown operating limits of the control variables.

$$c_i'(t) = \begin{cases} c_i(t), & t < t_0 \\ c_i(t) + \Delta c_i(t), & t_0 \le t < (t_0 + 6h) \\ c_i(t) - \Delta c_i(t), & t_0 + 6h \le t < (t_0 + 12h) \\ c_i(t), & t \ge (t_0 + 12h) \end{cases} \quad (5)$$

The random value $R(t)$ is determined using the upper shutdown operating limits during the first six hours of signal manipulation and the lower limits for the remaining hours. For example, considering the temperature operating constraints (Table 2) for process P5 and a nominal temperature control signal of 52°C, the manipulated control signal could be 105°C during the initial 6 hours of signal manipulation.

The ICS incorporates hardware and software mechanisms to ensure the reliable operation of LPG purification processes, even in the presence of failures or disruptions. However, it is essential to note that if certain variable thresholds are surpassed in processes within the LPG purification unit, the system may be unable to maintain its intended nominal output.

To establish seamless execution of events between virtual and physical spaces, engineers and experts must consider the following key aspects upon DT execution of events:

- *Sensors and data acquisition systems*: The physical system must be equipped with sensors and data acquisition systems capable of communicating over protocols, such as Modbus, OPC UA, and MQTT.
- Establishing connectivity: A reliable and secure connection must be established to enable real-time communication between sensors, data acquisition systems, and DT. Depending on the application requirements, this can be achieved by using wired or wireless connections.
- Data processing and storage: The DT must be able to process and store data received from the physical system in real-time. This can be achieved using cloud-based or edge-computing solutions that can handle large volumes of data and provide real-time analytics.
- Monitor and troubleshoot: Regular monitoring and troubleshooting must be performed to ensure that the communication between the physical system and DT is functioning correctly. This can include the use of monitoring tools to detect issues (e.g., communication errors or data inconsistencies) and taking corrective actions when necessary.

Engineers and experts can ensure that the DT receives real-time updates from a physical system. Furthermore, validating the consistency between the DT and physical system is essential to ensure that the DT accurately represents the behavior and performance of the physical system.

Therefore, when the DT model retrieves data from the physical system, it should request these data through a secure connection to the relevant endpoint on the server. This request includes any required authentication credentials and parameters that define the data (i.e., sensor measurements) to be retrieved. In this manner, the DT can securely communicate and continuously update to reflect changes in its physical counterpart [45], [46], [47].

IV. EXPERIMENTS

The oil refinery under study comprises more than 20 physical processes. Process specifications and shutdown operational limits are assigned based on thorough literature research on similar systems and the information provided by system operators [48], [49], [50]. To ensure security, the company's name, and all associated data and component names were anonymized and sanitized. The flow network model utilized in this study represents a typical LPG purification unit found in oil refineries [51]. In the following sections, we discuss our CE experiment on the DT of the LPG purification unit in detail.

A. STEP 1: MODEL DEFINITION

Based on the data provided by the system operators, we identified eight internal processes, three internal inputs, and one output node for the LPG purification unit (Table 4). In addition, 44 flows of resources were identified (Table 5). We then created an MFN that modeled the LPG purification unit (Fig. 4). Table 4 lists the MFN nodes and process variables. The depicted MFN nodes use generic terms and ID. In Table 5, we list the modeled flows, their respective resources, and the measuring units.

Next, we gathered the shutdown operational limits and constraints associated with the physical components and

TABLE 4. Mapped MFN nodes and node IDs' associations with their respective variables.

Name	ID	Process variables
Hydrocracker-A	I1	-
Hydrocracker-B	I2	-
Fluid Catalytic Cracking (FCC)	I3	-
Debutanizer-1	P1	Temperature, Pressure, Reflow
Deethanizer-1	P2	Temperature, Pressure, Reflow
Deethanizer-2	P3	Temperature, Pressure, Reflow
Debutanizer-2	P4	Temperature, Pressure, Reflow
Debutanizer-3	P5	Temperature, Pressure, Reflow
DEA Absorption-1	P6	Temperature, Pressure
DEA Absorption-2	P7	Temperature, Pressure
DEA Absorption-3	P8	Temperature, Pressure
Tank	O1	-

equipment used for each mapped process. Finally, for each process, we mapped the sensors of the LPG purification unit to the flows and process variables, thereby creating the DT of the physical system (see Section III-D).

Table 6 illustrates the mapping between the MFN process attributes (i.e., variables and flows) and the physical system sensors. Virtual space refers to a digital representation of the physical system that is used to monitor and optimize production. On the other hand, physical space refers to the actual physical environment in which the ICS (Industrial Control System) is deployed [52], [53].

The ICS relies on the interaction between virtual and physical spaces to detect and react to real-time changes in the physical environment. In our CE approach, real-time data from mapped sensors were used to capture the steady state of the system. Access to real-time data is essential for creating an accurate DT, as it can significantly improve its effectiveness in analyzing physical systems.

B. STEP 2 AND 3: STEADY STATE AND HYPOTHESIS FORMULATION

The DT model simulates the operation of the LPG purification unit. The sensor data from the ICS had a frequency of 1 data point per second. In our simulation, we utilized a one-hour time step. The selected time step is short enough to capture the system behaviors of interest and sufficiently long to allow for reasonable execution times [54]. To process the data, we calculated an hourly average from the stored one-second sensor readings. The overall simulation time was 24 hours. First, the DT model flows and process variables (i.e., control signals) were updated before each simulation cycle based on the mapping in Table 6. Specifically, we updated the process variables for all processes (i.e., set the ICS control signals) and the input LPG flows for processes P1, P2, and P3 using the API in each cycle. Fig. 5a-5b illustrate the nominal control signals for the temperature, pressure, and reflow variables for Debutanizer-3 (P5) used in our CE experiment.

TABLE 5. Modeled MFN flows, associated resources, and units.

Source	Target	Resource	Units
I1	P1		
I2	P3		
I3	P5		
P1	P2		
P3	P4		
P2	P6	C2	
P4	P7		
P5	P7		
P6	O1		
P7	O1		
P7	O1		
I1	P1		
I2	P3		m^3/h
I3	P5		
P1	P2		
P3	P4		
P2	P6	C5	
P4	P7		
P5	P7		
P6	O1		
P7	O1		
P7	O1		
P1	P2		
P3	P4	LPG	
P2	P6		
P4	P7		
P5	P7		
P6	O1		
P7	O1		
P7	O1		
I1	P1		
I2	P3	Raw LPG	
I3	P5		
I1	P1		
I2	P3		
I3	P5		
P1	P2		
P3	P4		
P2	P6	S	
P4	P7		
P5	P7		
P6	O1		
P7	O1		
P7	O1		

In each simulation cycle, our solver computed the output flows for each process, as explained in Section III-A. In our CE experiment, we used the concentrations of C2 and C5 that accumulated in the initially empty storage tank (O1) as our operational metrics to monitor potential deviations. Therefore, after each simulation cycle, our solver determined the concentrations of C2 and C5 in the storage tank. Figures 6a-6b illustrate how these concentrations changed over time, indicating the system's steady state.

The results consistently showed that the levels of C2 and C5 concentrations (see Fig. 6) remained below the limits specified for LPG products in Table 3. These results were confirmed through an internal audit using LPG specification thresholds and prior laboratory measurements from the tank, providing confidence in the accuracy of our process specifications and sensor mappings.

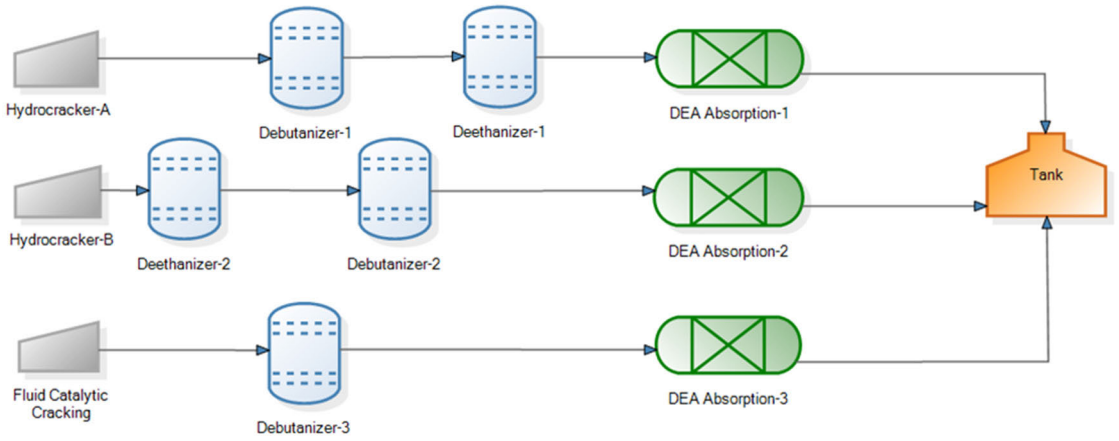

FIGURE 4. Graphical representation of the material flow network (MFN).

TABLE 6. Mapping between the material network flows and process Variables with the physical system sensors.

Process	Virtual space		Physical space
	Process attributes		Sensor ID
P1	Variables	Temperature	47TI1532
		Pressure	47PIC1504
		Reflow	47FIC1524
	Flows	LPG In	47FIC1507
P2	Variables	Temperature	47TI1550
		Pressure	47PIC1507
		Reflow	47FIC1524
	Flows	LPG In	47FIC1545
P4	Variables	Temperature	147TI1071
		Pressure	147PIC1028
		Reflow	147FIC1054
	Flows	LPG In	147FI1060
P3	Variables	Temperature	147TI1066
		Pressure	147PIC1022
		Reflow	147FIC1051
	Flows	LPG In	147FIC1003
P5	Variables	Temperature	7TI129
		Pressure	7PIC103
		Reflow	7FIC14
	Flows	LPG In	7FI126
P8	Variables	Temperature	7TI508
		Pressure	7PI504
	Flows	LPG In	7FIC501
P6	Variables	Temperature	47TI1565
		Pressure	47PIC1507
	Flows	LPG In	47FIC1545
P7	Variables	Temperature	147TI1081
		Pressure	147PIC1033
	Flows	LPG In	147FIC1058

However, it is important to recognize that the concentration values obtained in our simulation displayed minor deviations, approximately ±5%, compared to previous laboratory measurements taken from the tank. This variation is expected, given that the DT model covers only a part of the entire plant and accounts for the time difference between the simulation and laboratory measurements.

Despite the variations observed, the DT model provides an accurate representation of the unit's general performance, as shown in Table 3. This allows us to estimate the overall operational state of the plant, as explained in Section III-C.

As per our initial CE hypothesis outlined in Section III-B, it is crucial that the levels of C2 and C5 in the storage tank (O1) closely match the values illustrated in Fig. 6. This hypothesis remains valid even when we make adjustments to

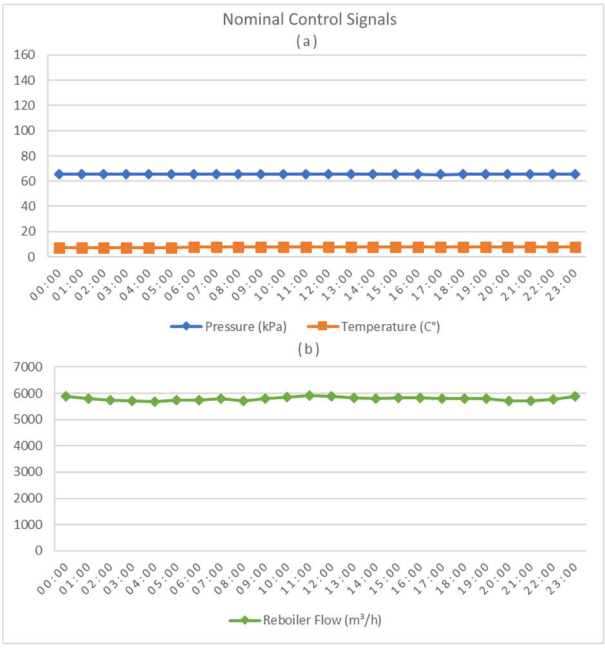

FIGURE 5. Nominal time-series data for pressure, temperature (a) , and reboiler (b) variables from the system ICS sensors.

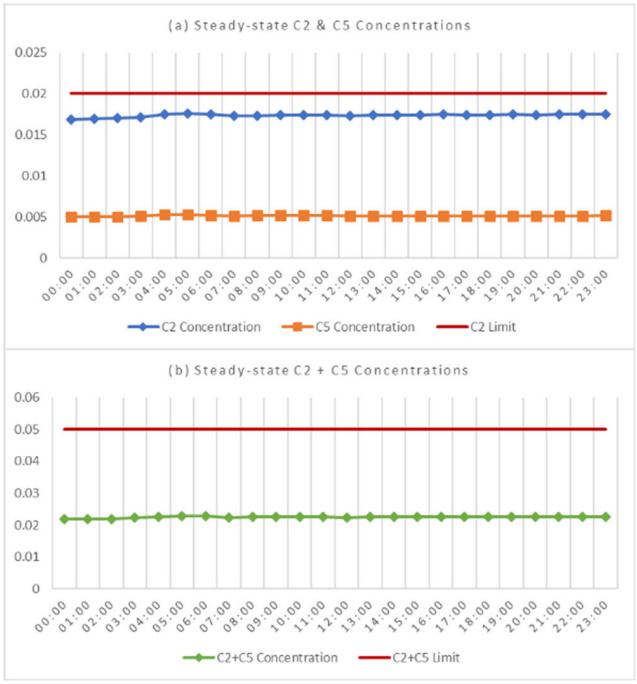

FIGURE 6. DT model steady-state individual (a) and cumulative (b) concentration values of C2 and C5 in the storage tank (O1) over time. The red line indicates the LPG product specifications for C2 and C5.

the ICS control inputs for any process variable within their shutdown limits.

C. STEP 4: VARYING ICS PROCESS EVENTS

In this step, we are replicating real-world events that can impact the stability of the system and examining how they affect the levels of C2 and C5 in the storage tank. To achieve this, we adjusted the settings for temperature, pressure, and reflow related to the P1-P5 processes (as detailed in Section III-C). We initiated these adjustments at 06:00 (t_0 = 6) and continued them for 12 hours. You can see the graphical representation of these adjustments for Debutanizer-3 (P5) in Figures 7a and 7b. Similar adjustments in temperature, pressure, and reflow settings were made for all the other processes we are studying (namely, P1-P5). It's important to note that the rest of the processes and their settings were not affected by these changes.

We repeated the simulation of the LPG purification unit, this time using the adjusted control signal values in our DT model. This simulation helped us recalculate the concentrations of C2 and C5 in the storage tank (O1). Since these concentrations were determined with the adjusted control signal, they reveal an unstable state. You can see the updated C2 and C5 concentrations in the storage tank (O1) over time in Figures 8a and 8b.

The results indicate that when we manipulated the control signals, there was a significant impact on the concentrations of C2 and C5 in the storage tank, as shown in Figure 8. Specifically, the concentration of C2 decreased notably between 06:00 and 11:00, followed by an increase from 11:00 to 17:00, and then a gradual decline until the end of the experiment. Conversely, the concentration of C5 steadily increased from 07:00 to 17:00 and then gradually decreased until the end of the experiment. Additionally, the cumulative concentrations of C2 and C5 followed a similar pattern, increasing from 11:00 to 17:00 and then decreasing thereafter.

It's important to note that the C2 concentration exceeded the specified limits, as depicted in Figure 8a. However, the combined concentrations of C2 and C5 remained within acceptable ranges, as shown in Figure 8b. *This observation confirms that manipulating the control signals in the system disrupts the operation and performance of the LPG purification unit.*

D. RESULTS

We utilized the C2 and C5 concentrations obtained from both the stable and unstable system states to compute the ICD indicator, which serves as a measure of the system's departure from its stable state (see Section III-B for details). The ICD indicator provided us with a means to assess and evaluate the system's ability to withstand operational-level alterations. The resilience curve, as illustrated in Figure 9, was generated using the ICD indicator. This curve illustrates how the operational level evolves over time as the system transitions through various phases: commencing from a stable state, moving towards an unstable state, and ultimately recovering.

Based on our findings, we observed that the ICD indicator remained stable, nearly zero, from 00:00 to 06:00, indicating minimal variation in impurity concentration and system stability during this period. However, from 06:00 to 23:00, the ICD indicator became negative, reaching a minimum of

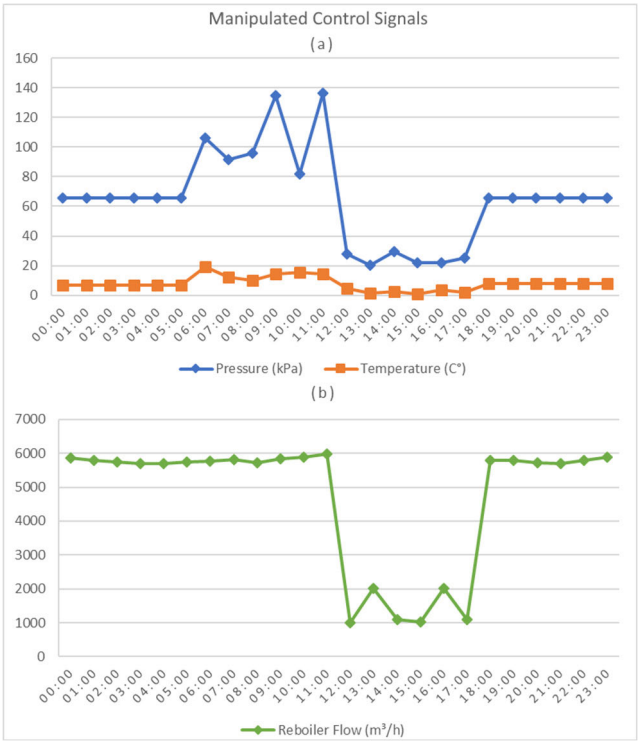

FIGURE 7. Time series data of manipulated temperature, pressure (a) , and reboiler (b) control signals for Debutanizer-3 (P5). Signal manipulation started at 06:00 ($t_0 = 6$) and lasted for 12 hours.

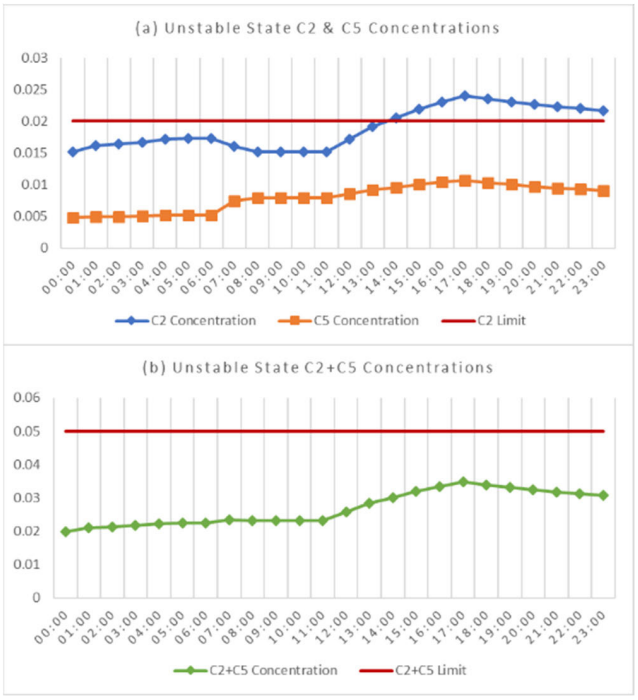

FIGURE 8. DT model unstable state individual (a) and cumulative (b) concentration values of C2 and C5 in the storage tank (O1) over time. The red line indicates the LPG product specifications regarding its impurities (i.e., C2 and C5).

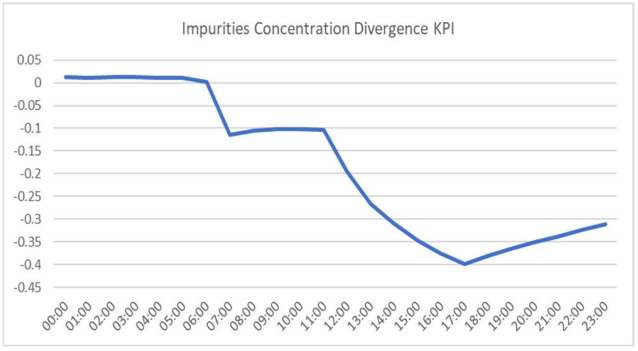

FIGURE 9. The systems resilience curve as produced by the ICD indicator.

−0.41. This marked departure from the stable state signified a substantial deviation from the expected impurity concentration. It's essential to note that lower ICD values represent more significant observed deviations. Thus, during this timeframe, impurity concentration diverged from the stable state, contradicting our initial CE hypothesis (see Section III-B). *This finding emphasizes the critical need to consider broader system effects, even when individual processes remain within their operational limits*. Crossing certain thresholds can lead to a situation where the system's production and operational state deviate from expected output.

The slope of the ICD indicator curve between 06:00 and 17:00 reflects the system's ability to withstand adverse events. A sharp decline in the ICD indicator between 06:00 and 07:00 and between 11:00 and 17:00 indicates adverse events during these periods negatively affecting impurity concentration. Despite the ICD indicator reaching a minimum of −0.12, *relatively stable ICD values recorded from 07:00 to 11:00 indicate effective system resistance to adverse events.*

The slope of the ICD indicator curve between 17:00 and 23:00 illustrates the system's ability to recover from adverse events. The increasing ICD indicator during this period suggests gradual recovery from earlier adverse events. However, within the chosen experimental timeframe, full recovery may not occur. Nevertheless, with undisturbed operation, impurity levels will eventually return to normal, and the system will recover, returning to its stable state. The ICD indicator enables experts to monitor impurity concentration deviations over time, assessing the system's capacity to resist and recover from adverse events.

Analyzing the critical flows, we noticed a recurring pattern where certain network nodes, specifically O1, appeared multiple times as targets. This pattern suggests that these nodes play a crucial role in the system's resilience but are susceptible to disruption during CE experiments. Additionally, flows involving C2 exhibited the most substantial deviations from the steady state. An attack on processes responsible for removing C2 could lead to significant disturbances in the system's operation, causing the concentration of C2 in storage tank (O1) to exceed specified limits. This would render the stored LPG unsellable due to noncompliance with required specifications (see Fig. 8a). This observation highlights the

TABLE 7. Top five critical flows from the resilience analysis step. Peak deviation refers to the maximum difference between the steady-state flow (normal operating condition) and the flow observed during the experiment.

Source	Target	Resource	Peak Deviation from steady state (m^3/h)
P5	P8	C2	0.19
P8	O1	C2	0.19
P1	P2	C2	0.17
P2	P6	C2	0.17
P6	O1	C2	0.17

heightened sensitivity and potential impact of these flows on overall system performance and stability. Experts can use this information to identify critical system flows above a threshold value, with a particular focus on flows originating from nodes P5 and P1, considered high-priority targets for mitigation measures.

Additionally, applying the ICD indicator at the process level *allows us to identify significant deviations in system processes*. This analysis prioritizes vulnerable processes based on their highest deviation from the steady state. Negative ICD values indicate notable deviations from the steady-state process, signaling potential vulnerabilities and areas for improvement. Table 8 lists the top three critical processes based on this approach.

Based on our findings, the most critical process is Debutanizer-3 (P5), which has a significant impact on the flow of both C2 and C5 resources. The ICD value for this process is -0.25, indicating a substantial deviation from the steady-state flow conditions. Debutanizer-1 (P1) is the second-ranked critical process, primarily affecting the C2 output flow. It exhibits an ICD value of -0.17, signifying a significant departure from normal operating conditions. Debutanizer-2 (P4) is the third-ranked critical process, also influencing the C2 output flow, but to a lesser extent, with an ICD value of -0.10. These findings emphasize the vulnerability of these processes in terms of their impact on system output flows. Experts can use this information to identify crucial system processes surpassing a predefined threshold. Based on our analysis, we prioritize addressing the Debutanizer-3 (P5) process as the most critical node, deserving high-priority mitigation measures. This conclusion aligns with the significant impact this specific process has and corroborates our critical flow analysis.

Our approach offers a streamlined way for experts to efficiently pinpoint and prioritize the most vulnerable aspects of the system, facilitating precise assessments. As a result, experts can concentrate their efforts on devising targeted mitigation strategies and contingency plans. Implementing these measures can lead to substantial enhancements in the system's resilience, ensuring smooth and secure operation.

TABLE 8. Top three critical processes from the resilience analysis step. The ICD values are computed by assessing the impact on output flow resources for each process. The ICD reflects the maximum divergence between the steady-state flows (normal operating conditions) and the flows observed during the experiment.

Process	Output flow(s)	Top ICD value
P5	C2 & C5	-0.25
P1	C2	-0.17
P4	C2	-0.10

E. LESSONS LEARNED

The results obtained from our experiment illustrate that a Cyber-Physical System's (CPS) ability to withstand deviations hinges on its capacity to absorb them effectively. In our case, the Programmable Logic Controller (PLC) monitors signals from sensors overseeing the distillation columns in the LPG purification unit. When there's a deviation in a process variable like temperature, pressure, or reflow, the PLC communicates with actuators to initiate adjustments at the column. However, adverse events can sometimes circumvent fail-safe mechanisms. For example, Denial of Service (DoS) attacks can overload the Industrial Control System (ICS) network, leading to congestion failures in control components like routers [55]. Consequently, ICS sensor values may not be transmitted to the control center on time, and the system might not detect these disconnections. Another scenario involves false data injection (FDI) attacks, where intruders manipulate sensor readings in a way that introduces errors into state variables and value calculations while evading bad data detection (BDD) mechanisms [56], [57]. In both cases, the distillation column's temperature, pressure, and reflow variables experience significant deviations from their expected values, leading to deviations in the system's output flows. This aligns with our experiment's outcomes, where the C2 and C5 output flows entering the storage tank reached unacceptable levels at 17:00 (see Fig. 9). Given the above, *if deviations in control signals caused by adverse events fall within the shutdown operating limits, the system cannot withstand the event without experiencing significant disruptions in its production output*.

Our critical flow analysis revealed that flows involving C2 were highly vulnerable. Importantly, this analysis underscores the storage tank (O1) as a critical node for system resilience. *The* system's *recovery rate and overall recovery depend on the impurity level in the storage tank (O1)* and how quickly impurities are reduced after the end of the adverse event (i.e., signal manipulation). Furthermore, *the storage tank (O1) influences the system's ability to resist adverse conditions*. The time it takes for the system to reach a critical state, characterized by out-of-spec impurities in the storage tank, is directly affected by the initial concentration of impurities before signal manipulation. Our experiment confirmed that a low impurity concentration in the storage tank leads to increased resistance, resulting in an extended time needed for the system to reach a critical state. The presence of tank(s)

acts as a buffer, effectively extending the response time available to operators to address unexpected events. This extended response time can significantly mitigate the impact of adverse conditions on the overall system.

Our critical process analysis identified Debutanizer-3 (P5) as the most critical process. Since the system comprises sequentially connected processes (Fig. 4), introducing redundancy for the identified critical process(es) could enhance system resilience [58]. However, the feasibility of redundancy depends on the specific industrial process because industrial purification plants often consist of sequentially connected processes (e.g., oil refineries, water and wastewater treatment plants). Typically, one process relies on the operation and efficiency of the previous process for its proper functioning. In our case study, the sulfur removal process, which is not directly dependent on the performance of the C2 removal process for its primary operation (i.e., sulfur removal), depends on the existence of the LPG flow, i.e., the operation of the preceding process.

The insights gained through our approach, including identifying and prioritizing critical flows and processes, form the basis for enhancing the system's resilience. Therefore, *engineers and experts should conduct periodic Cyber-Physical Experiments (CE) to assess and validate the effectiveness of redundancy and other mitigation measures on the system's resilience continually*. By implementing an effective resilience enhancement strategy that includes regular CE experiments, we can continuously validate and improve the effectiveness of mitigation measures and contingency plans.

V. DISCUSSION AND CONCLUSION

In this study, we introduce an innovative framework that combines Decision Trees (DT) using Multi-layer Feedforward Networks (MFN) with a Chaos Engineering (CE) methodology to quantitatively evaluate a system's resilience in its operational stage using real-time data against both deliberate malicious attacks and unintentional threats. The combination of DT, CE, and MFN enabled nondisruptive real-time testing and proactive observation of a system's behavior and performance.

The validation of our approach on an actual LPG purification process demonstrates its effectiveness in improving system resilience. We conducted Chaos Engineering experiments on a DT of a real-world LPG purification process within an oil refinery. These experiments provided strong evidence of notable operational performance deviations from the steady state, highlighting the system's vulnerability to adverse events, such as injection attacks and state deviation that can occur from integrity violations. Results detected various process-aware vulnerabilities that could derail the LPG purification process. Fixing these issues significantly enhanced system resilience, improving its ability to withstand adverse events and recover gracefully from them, as evident from the observed divergence curve slope.

To the best of our knowledge, this work represents the first study that combines DT and CE using an MFN to evaluate and improve CPS resilience. Our findings contribute to advancing the concepts of security and safety by design in critical infrastructure protection.

A. RESTRICTIONS AND FUTURE WORK

The accuracy of our analysis highly depends on the level of detail and quality of the modeled system inside the DT. Therefore, to improve the accuracy of our approach, we must ensure that the modeled system is as representative as possible of the actual system behavior. If the modeled unit is not representative of the overall behavior of the system, then the overall operating condition cannot be accurately reflected.

Second, experimental events in DT simulations face high-time complexity due to the utilization of nonlinear equations to model the physical processes; an issue that can strain resources if processes are to scale into larger systems and processes. Addressing this limitation requires exploring more efficient algorithms and computational techniques to reduce time complexity while maintaining the required level of accuracy.

Third, it is important to note that our current study primarily focuses on investigating the effects of disruptions within a system. However, in industrial CPS, interconnections between different systems can have a significant impact on the overall performance. For instance, disruptions in one system can propagate and affect the performance of the entire interconnected system (i.e., an oil refinery connected to a crude oil distribution pipeline network, communication, and electricity grid). Plant-wide CE experiments should work towards modeling multiple processes into a single DT to be able to capture subliminal vulnerabilities that may exist due to these component interdependencies.

Lastly, the generalization of the approach should be explored by assessing its applicability and effectiveness in other CI sectors. For instance, evaluating its effectiveness in Electric Power Grid Systems, where generators and batteries can be analogous to distillation units and storage tanks, would provide valuable insights and broaden its applicability.

APPENDIX
DISTILLATION COLUMN MODELLING

In this Appendix, we describe the MESH equations that model the equilibrium stages in a single one-feed two-product distillation column with a reboiler and condenser (Fig. 10a).

The MESH variables are referred to as state variables. For a general j^{th} stage and i^{th} component (Fig. 10b) these are: Stage temperatures, T_j (°C); Stage pressures, P_j $\left(\frac{kg}{cm^2}\right)$; Internal total vapor and liquid rates, V_j $\left(\frac{kgmol}{h}\right)$ and L_j $\left(\frac{kgmol}{h}\right)$; Stage compositions, $y_{j,i}$ and $x_{j,i}$, or instead, component vapor and liquid rates, $v_{j,i}$ $\left(\frac{kgmol}{h}\right)$ and $l_{j,i}$ $\left(\frac{kgmol}{h}\right)$.

The component material balance equations for the simple stage j (no feed), are given by:

$$L_{j-1}x_{j-1,i} + V_{j+1}y_{j+1,i} - L_jx_{j,i} - V_jy_{j,i} = 0 \quad (6)$$

$$l_{j-1,i} + v_{j+1,i} - l_{j,i} - v_{j,i} = 0 \quad (7)$$

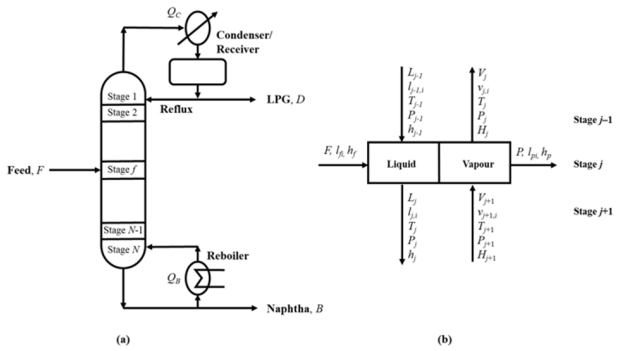

FIGURE 10. (a) Process flow diagram of a single-feed debutanizer column with reboiler and condenser. (b) Model of the separation stage.

The component balance for the feed stage, f, adds the liquid portion of the feed, l_{fi}, and for the product stage (D, B) the material withdrawn l_{pi}, is subtracted from the component material balance. By convention, the material leaving a tray has a negative value, and the material entering a tray has a positive value. The total material balance for the simple stage is given by:

$$V_{j+1} + L_{j+1} - V_j - l_j = 0 \tag{8}$$

The same convention applies to feed and product trays, where the total flow rate, F, is added, and the total product flow rate, P, is subtracted.

The flow compositions leaving the stage are in equilibrium. Therefore, the mole fractions of component i in the liquid and vapor flows leaving stage j are related by the equilibrium relation shown in (9) and (10), as follows:

$$y_{j,i} = K_{j,i} x_{j,i} \tag{9}$$

$$\frac{v_{j,i}}{V_j} = K_{j,i} \frac{l_{ji},}{L_j} \tag{10}$$

The equilibrium constant or K-value, $K_{j,i}$, can be a complex function, dependent on the flow composition, temperature and pressure. The operating line of the enriching section is given by (11) where R is the reflux/reflow ratio $\left(=\frac{L_0}{D}\right)$. Equal molar flow rates for liquid and vapor are assumed, that is $L_1 = L_2 = \cdots = L_n$ and $V_1 = V_2 = \cdots = V_{n+1}$.

$$y_{j,i+1} = \frac{R}{R+1} x_{j,i} + \frac{x_{j,D}}{R+1} \tag{11}$$

The summation equation or composition constraints simply states that the sum of the mole fractions on each stage is equal to unity for the liquid and vapor phases:

$$\sum_{i=1}^{C} x_{i,j} = 1 \tag{12}$$

$$\sum_{i=1}^{C} y_{i,j} = 1 \tag{13}$$

The energy balance for the simple stage, $j = 1, \ldots, N_t$ is given by:

$$L_{j-1} h_{j-1} + V_{j+1} H_{j+1} - L_j h_j - V_j H_j = 0 \tag{14}$$

The liquid enthalpy, h_j, and the vapor enthalpy, H_j (energy per mole or m^3) for each stage are functions of temperature, T_j, pressure, P_j, and composition: $H_j = H_j\left(T_j, P_j, y_{j,i}\right)$ and $h_j = h_j\left(T_j, P_j, x_{j,i}\right)$. For feed stage, the term Fh_f is added to the energy balance. The energy balance for the reboiler is calculated based on the reboiler duty $Q_R\left(\frac{kj}{h}\right)$ and the total bottom rate $B\left(\frac{kgmol}{h}\right)$ using (15), while the energy balance for the condenser is calculated based on the condenser duty $Q_c\left(\frac{kj}{h}\right)$ and the total distillate vapor rate $D\left(\frac{kgmol}{h}\right)$ using (16).

$$L_{N-1} h_{N-1} - V_N H_N - B h_N + Q_R = 0 \tag{15}$$

$$L_0 h_0 - V_1 H_1 - D h_0 - Q_C = 0 \tag{16}$$

The overall energy balance is given by:

$$F h_f - B h_N - D h_0 + Q_R - Q_C = 0 \tag{17}$$

REFERENCES

[1] S. U. Rehman, M. Ceglia, S. Siddiqui, and V. Gruhn, "Towards an importance of security for cyber-physical systems/Internet-of-Things," in *Proc. 8th Int. Conf. Softw. Inf. Eng.*, Apr. 2019, pp. 151–155, doi: 10.1145/3328833.3328855.

[2] M. Moghaddam and A. Deshmukh, "Resilience of cyber-physical manufacturing control systems," *Manuf. Lett.*, vol. 20, pp. 40–44, Apr. 2019, doi: 10.1016/j.mfglet.2019.05.002.

[3] U. Bodkhe, D. Mehta, S. Tanwar, P. Bhattacharya, P. K. Singh, and W.-C. Hong, "A survey on decentralized consensus mechanisms for cyber physical systems," *IEEE Access*, vol. 8, pp. 54371–54401, 2020, doi: 10.1109/ACCESS.2020.2981415.

[4] G. Stergiopoulos, D. A. Gritzalis, and E. Limnaios, "Cyber-attacks on the oil & gas sector: A survey on incident assessment and attack patterns," *IEEE Access*, vol. 8, pp. 128440–128475, 2020, doi: 10.1109/ACCESS.2020.3007960.

[5] T. Tsvetanov and S. Slaria, "The effect of the colonial pipeline shutdown on gasoline prices," *Econ. Lett.*, vol. 209, Dec. 2021, Art. no. 110122, doi: 10.1016/j.econlet.2021.110122.

[6] J. Cervini, A. Rubin, and L. Watkins, "Don't drink the cyber: Extrapolating the possibilities of oldsmar's water treatment cyberattack," *Int. Conf. Cyber Warfare Secur.*, vol. 17, no. 1, pp. 19–25, Mar. 2022, doi: 10.34190/iccws.17.1.29.

[7] X. Xu, T. Zhang, H. Huang, Z. Luo, and C. Zhang, "Research progress of CPS resilience," in *Proc. IEEE Int. Conf. Unmanned Syst. (ICUS)*, Guangzhou, China, Oct. 2022, pp. 1101–1106, doi: 10.1109/ICUS55513.2022.9986856.

[8] M. Segovia, J. Rubio-Hernan, A. R. Cavalli, and J. Garcia-Alfaro, "Cyber-resilience evaluation of cyber-physical systems," in *Proc. IEEE 19th Int. Symp. Netw. Comput. Appl. (NCA)*, Cambridge, MA, USA, Nov. 2020, pp. 1–8, doi: 10.1109/NCA51143.2020.9306741.

[9] F. Yu, Y. Hu, T. Zhang, and Y. Jin, "Special issue: Resilient distributed estimator with information consensus for CPS security," in *Proc. IEEE 38th Int. Conf. Comput. Design (ICCD)*, Hartford, CT, USA, Oct. 2020, pp. 41–44, doi: 10.1109/ICCD50377.2020.00023.

[10] C. Konstantinou, M. Sazos, A. S. Musleh, A. Keliris, A. Al-Durra, and M. Maniatakos, "GPS spoofing effect on phase angle monitoring and control in a real-time digital simulator-based hardware-in-the-loop environment," *IET Cyber-Phys. Syst., Theory Appl.*, vol. 2, no. 4, pp. 180–187, Dec. 2017, doi: 10.1049/iet-cps.2017.0033.

[11] S. Ouchani, K. Khebbeb, and M. Hafsi, "Towards enhancing security and resilience in CPS: A coq-maude based approach," in *Proc. IEEE/ACS 17th Int. Conf. Comput. Syst. Appl. (AICCSA)*, Antalya, Turkey, Nov. 2020, pp. 1–6, doi: 10.1109/AICCSA50499.2020.9316535.

[12] C. H. Fleming, C. Elks, G. Bakirtzis, S. Adams, B. Carter, P. Beling, and B. Horowitz, "Cyberphysical security through resiliency: A systems-centric approach," *Computer*, vol. 54, no. 6, pp. 36–45, Jun. 2021, doi: 10.1109/MC.2020.3039491.

[13] D. McFarlane, R. Srinivasan, A. Puchkova, A. Thorne, and A. Brintrup, "A maturity framework for operational resilience and its application to production control," in *Service Orientation in Holonic and Multi-Agent Manufacturing* (Studies in Computational Intelligence), vol. 762, T. Borangiu, D. Trentesaux, A. Thomas, and O. Cardin, Eds. Cham, Switzerland: Springer, 2018, pp. 51–62, doi: 10.1007/978-3-319-73751-5_5.

[14] C. Konstantinou, G. Stergiopoulos, M. Parvania, and P. Esteves-Verissimo, "Chaos engineering for enhanced resilience of cyber-physical systems," 2021, *arXiv:2106.14962*, doi: 10.48550/ARXIV.2106.14962.

[15] A. Basiri, N. Behnam, R. de Rooij, L. Hochstein, L. Kosewski, J. Reynolds, and C. Rosenthal, "Chaos engineering," *IEEE Softw.*, vol. 33, no. 3, pp. 35–41, May 2016, doi: 10.1109/MS.2016.60.

[16] C. Rosenthal and N. Jones, *Chaos Engineering: System Resiliency in Practice*, 1st ed. Beijing, China: O'Reilly Media, 2020.

[17] Y. Jiang, S. Yin, K. Li, H. Luo, and O. Kaynak, "Industrial applications of digital twins," *Phil. Trans. Roy. Soc. A, Math., Phys. Eng. Sci.*, vol. 379, no. 2207, Oct. 2021, Art. no. 20200360, doi: 10.1098/rsta.2020.0360.

[18] V. Wohlgemuth, B. Page, and W. Kreutzer, "Combining discrete event simulation and material flow analysis in a component-based approach to industrial environmental protection," *Environ. Model. Softw.*, vol. 21, no. 11, pp. 1607–1617, Nov. 2006, doi: 10.1016/j.envsoft.2006.05.015.

[19] G. Arampatzis, A. Angelis-Dimakis, M. Blind, and D. Assimacopoulos, "A web-based toolbox to support the systemic eco-efficiency assessment in water use systems," *J. Cleaner Prod.*, vol. 138, pp. 181–194, Dec. 2016, doi: 10.1016/j.jclepro.2016.02.065.

[20] L. Xu, Q. Guo, Y. Sheng, S. M. Muyeen, and H. Sun, "On the resilience of modern power systems: A comprehensive review from the cyber-physical perspective," *Renew. Sustain. Energy Rev.*, vol. 152, Dec. 2021, Art. no. 111642, doi: 10.1016/j.rser.2021.111642.

[21] A. Clark and S. Zonouz, "Cyber-physical resilience: Definition and assessment metric," *IEEE Trans. Smart Grid*, vol. 10, no. 2, pp. 1671–1684, Mar. 2019, doi: 10.1109/TSG.2017.2776279.

[22] A. Rahiminejad, M. Ghafouri, R. Atallah, A. Mohammadi, and M. Debbabi, "A cyber-physical resilience-based survivability metric against topological cyberattacks," in *Proc. IEEE Power Energy Soc. Innov. Smart Grid Technol. Conf. (ISGT)*, New Orleans, LA, USA, Apr. 2022, pp. 1–5, doi: 10.1109/ISGT50606.2022.9817513.

[23] G. Stergiopoulos, P. Dedousis, and D. Gritzalis, "Automatic network restructuring and risk mitigation through business process asset dependency analysis," *Comput. Secur.*, vol. 96, Sep. 2020, Art. no. 101869, doi: 10.1016/j.cose.2020.101869.

[24] N. Pelissero, P. M. Laso, and J. Puentes, "Impact assessment of anomaly propagation in a naval water distribution cyber-physical system," in *Proc. IEEE Int. Conf. Cyber Secur. Resilience (CSR)*, Rhodes, Greece, Jul. 2021, pp. 518–523, doi: 10.1109/CSR51186.2021.9527952.

[25] E. Guo and B. Bu, "CBTC systems resilience evaluation based on resource state model under DoS attacks," in *Proc. 7th Annu. Int. Conf. Netw. Inf. Syst. for Comput. (ICNISC)*, Guiyang, China, Jul. 2021, pp. 451–456, doi: 10.1109/ICNISC54316.2021.00087.

[26] A. S. Jin, L. Hogewood, S. Fries, J. H. Lambert, L. Fiondella, A. Strelzoff, J. Boone, K. Fleckner, and I. Linkov, "Resilience of cyber-physical systems: Role of AI, digital twins, and edge computing," *IEEE Eng. Manag. Rev.*, vol. 50, no. 2, pp. 195–203, 2nd Quart., 2022, doi: 10.1109/EMR.2022.3172649.

[27] S. Yoginath, V. Tansakul, S. Chinthavali, C. Taylor, J. Hambrick, P. Irminger, and K. Perumalla, "On the effectiveness of recurrent neural networks for live modeling of cyber-physical systems," in *Proc. IEEE Int. Conf. Ind. Internet (ICII)*, Orlando, FL, USA, Nov. 2019, pp. 309–317, doi: 10.1109/ICII.2019.00062.

[28] P. Eirinakis, S. Lounis, S. Plitsos, G. Arampatzis, K. Kalaboukas, K. Kenda, J. Lu, J. M. Rožanec, and N. Stojanovic, "Cognitive digital twins for resilience in production: A conceptual framework," *Information*, vol. 13, no. 1, p. 33, Jan. 2022, doi: 10.3390/info13010033.

[29] S. Guo, M. Wu, and C. Wang, "Symbolic execution of programmable logic controller code," in *Proc. 11th Joint Meeting Found. Softw. Eng.*, Paderborn Germany, Aug. 2017, pp. 326–336, doi: 10.1145/3106237.3106245.

[30] Z. Luo, F. Zuo, Y. Jiang, J. Gao, X. Jiao, and J. Sun, "Polar: Function code aware fuzz testing of ICS protocol," *ACM Trans. Embedded Comput. Syst.*, vol. 18, no. 5s, pp. 1–22, Oct. 2019, doi: 10.1145/3358227.

[31] X. Cai, K. Shi, K. She, S. Zhong, Y. C. Soh, and Y. Yu, "Performance error estimation and elastic integral event triggering mechanism design for T–S fuzzy networked control system under DoS attacks," *IEEE Trans. Fuzzy Syst.*, vol. 31, no. 4, pp. 1327–1339, Apr. 2023, doi: 10.1109/TFUZZ.2022.3199817.

[32] X. Cai, K. Shi, K. She, S. Zhong, and Y. Tang, "Quantized sampled-data control tactic for T–S fuzzy NCS under stochastic cyber-attacks and its application to truck-trailer system," *IEEE Trans. Veh. Technol.*, vol. 71, no. 7, pp. 7023–7032, Jul. 2022, doi: 10.1109/TVT.2022.3169349.

[33] Y. Chen, C. M. Poskitt, and J. Sun, "Learning from mutants: Using code mutation to learn and monitor invariants of a cyber-physical system," in *Proc. IEEE Symp. Secur. Privacy (SP)*, San Francisco, CA, USA, May 2018, pp. 648–660, doi: 10.1109/SP.2018.00016.

[34] A. Basiri, L. Hochstein, N. Jones, and H. Tucker, "Automating chaos experiments in production," 2019, *arXiv:1905.04648*, doi: 10.48550/ARXIV.1905.04648.

[35] M. A. Chang, B. Tschaen, T. Benson, and L. Vanbever, "Chaos monkey: Increasing SDN reliability through systematic network destruction," *ACM SIGCOMM Comput. Commun. Rev.*, vol. 45, no. 4, pp. 371–372, Sep. 2015, doi: 10.1145/2829988.2790038.

[36] K. A. Torkura, M. I. H. Sukmana, F. Cheng, and C. Meinel, "Security chaos engineering for cloud services: Work in progress," in *Proc. IEEE 18th Int. Symp. Netw. Comput. Appl. (NCA)*, Cambridge, MA, USA, Sep. 2019, pp. 1–3, doi: 10.1109/NCA.2019.8935046.

[37] K. A. Torkura, M. I. H. Sukmana, F. Cheng, and C. Meinel, "CloudStrike: Chaos engineering for security and resiliency in cloud infrastructure," *IEEE Access*, vol. 8, pp. 123044–123060, 2020, doi: 10.1109/ACCESS.2020.3007338.

[38] F. Poltronieri, M. Tortonesi, and C. Stefanelli, "A chaos engineering approach for improving the resiliency of IT services configurations," in *Proc. IEEE/IFIP Netw. Oper. Manage. Symp. (NOMS)*, Budapest, Hungary, Apr. 2022, pp. 1–6, doi: 10.1109/NOMS54207.2022.9789887.

[39] P. Dedousis, G. Stergiopoulos, G. Arampatzis, and D. Gritzalis, "A security-aware framework for designing industrial engineering processes," *IEEE Access*, vol. 9, pp. 163065–163085, 2021, doi: 10.1109/ACCESS.2021.3134759.

[40] D. Zhang, C. Li, H. H. Goh, T. Ahmad, H. Zhu, H. Liu, and T. Wu, "A comprehensive overview of modeling approaches and optimal control strategies for cyber-physical resilience in power systems," *Renew. Energy*, vol. 189, pp. 1383–1406, Apr. 2022, doi: 10.1016/j.renene.2022.03.096.

[41] J. Wang and G. Henke, "Tridiagonal matrix for distillation," *Hydrocarbon Process.*, vol. 45, no. 5, p. 155, 1966.

[42] P. Ge, F. Teng, C. Konstantinou, and S. Hu, "A resilience-oriented centralised-to-decentralised framework for networked microgrids management," *Appl. Energy*, vol. 308, Feb. 2022, Art. no. 118234, doi: 10.1016/j.apenergy.2021.118234.

[43] C. Tang. (Aug. 2017). *Key Performance Indicators for Process Control System Cybersecurity Performance Analysis*. Accessed: Jun. 6, 2022. [Online]. Available: https://www.nist.gov/publications/key-performance-indicators-process-control-system-cybersecurity-performance-analysis

[44] Z. Hau, J. H. Castellanos, and J. Zhou, "Evaluating cascading impact of attacks on resilience of industrial control systems: A design-centric modeling approach," in *Proc. 6th ACM Cyber-Phys. Syst. Secur. Workshop*, Taipei, Taiwan, Oct. 2020, pp. 42–53, doi: 10.1145/3384941.3409587.

[45] S. Boschert and R. Rosen, "Digital twin-the simulation aspect," in *Mechatronic Futures*, P. Hehenberger and D. Bradley, Eds. Cham, Switzerland: Springer, 2016, pp. 59–74, doi: 10.1007/978-3-319-32156-1_5.

[46] T. D. West and M. Blackburn, "Is digital thread/digital twin affordable? A systemic assessment of the cost of DoD's latest Manhattan project," *Proc. Comput. Sci.*, vol. 114, pp. 47–56, Jan. 2017, doi: 10.1016/j.procs.2017.09.003.

[47] A. El Saddik, "Digital twins: The convergence of multimedia technologies," *IEEE MultimediaMag.*, vol. 25, no. 2, pp. 87–92, Apr. 2018, doi: 10.1109/MMUL.2018.023121167.

[48] L. Fortuna, S. Graziani, and M. G. Xibilia, "Soft sensors for product quality monitoring in debutanizer distillation columns," *Control Eng. Pract.*, vol. 13, no. 4, pp. 499–508, Apr. 2005, doi: 10.1016/j.conengprac.2004.04.013.

[49] D. W. Hanson, T. Becker, and M. R. Resetarits, "FCC debutanizer revamp for flexibility and additional capacity," in *Proc. AIChE Spring Nat. Meeting Distillation Symp.* Texas, Apr. 2001.

[50] M. R. Resetarits and M. J. Lockett, "Distillation," in *Encyclopedia of Physical Science and Technology*. Amsterdam, The Netherlands: Elsevier, 2003, pp. 547–559, doi: 10.1016/B0-12-227410-5/00182-4.
[51] A. Bahadori, *Natural Gas Processing: Technology and Engineering Design*. Amsterdam, The Netherlands: Elsevier, 2014.
[52] X. Zhou, X. Xu, W. Liang, Z. Zeng, S. Shimizu, L. T. Yang, and Q. Jin, "Intelligent small object detection for digital twin in smart manufacturing with industrial cyber-physical systems," *IEEE Trans. Ind. Informat.*, vol. 18, no. 2, pp. 1377–1386, Feb. 2022, doi: 10.1109/TII.2021.3061419.
[53] M. Liu, S. Fang, H. Dong, and C. Xu, "Review of digital twin about concepts, technologies, and industrial applications," *J. Manuf. Syst.*, vol. 58, pp. 346–361, Jan. 2021, doi: 10.1016/j.jmsy.2020.06.017.
[54] J. Hensen and R. Lamberts, Eds., *Building Performance Simulation for Design and Operation*, 2nd ed. New York, NY, USA: Routledge, 2019.
[55] Y. Liu, P. Ning, and M. K. Reiter, "False data injection attacks against state estimation in electric power grids," *ACM Trans. Inf. Syst. Secur.*, vol. 14, no. 1, pp. 1–33, May 2011, doi: 10.1145/1952982.1952995.
[56] M. Doostmohammadian and U. A. Khanc, "Vulnerability of CPS inference to DoS attacks," in *Proc. 48th Asilomar Conf. Signals, Syst. Comput.*, Nov. 2014, pp. 2015–2018, doi: 10.1109/ACSSC.2014.7094825.
[57] H. Guo, J. Sun, and Z.-H. Pang, "Stealthy false data injection attacks with resource constraints against multi-sensor estimation systems," *ISA Trans.*, vol. 127, pp. 32–40, Aug. 2022, doi: 10.1016/j.isatra.2022.02.045.
[58] M. Barbeau, F. Cuppens, N. Cuppens, R. Dagnas, and J. Garcia-Alfaro, "Resilience estimation of cyber-physical systems via quantitative metrics," *IEEE Access*, vol. 9, pp. 46462–46475, 2021, doi: 10.1109/ACCESS.2021.3066108.

PANAGIOTIS DEDOUSIS received the B.Sc. degree in informatics from the University of Piraeus, Greece, and the M.Sc. degree in information systems from the Athens University of Economics and Business, Greece, where he is currently pursuing the Ph.D. degree in information security and critical infrastructure protection. He is also a Researcher with the Information Security and Critical Infrastructure Protection (INFOSEC) Research Group, Department of Informatics, Athens University of Economics and Business. His current research interests include information security, critical infrastructure protection, and risk assessment.

GEORGE STERGIOPOULOS received the B.Sc. degree in informatics from the University of Piraeus, Greece, and the M.Sc. degree in information systems and the Ph.D. degree in critical infrastructure protection at software and information interdependency levels from the Athens University of Economics and Business (AUEB), Greece. He is currently an Assistant Professor of cybersecurity with the Department of Information and Communication Systems Engineering, University of the Aegean, Samos, Greece. He has been a Principal Investigator with the INFOSEC Laboratory, AUEB, in many funded research projects in critical infrastructure protection, software security, malware, and network security. He has published several papers in peer-reviewed journals and international conferences.

GEORGE ARAMPATZIS received the Diploma degree in chemical engineering and the Ph.D. degree from the National Technical University of Athens, in 1991 and 2000, respectively. He is currently an Assistant Professor with the School of Production Engineering and Management (PEM), Technical University of Crete (TUC), and a Visiting Professor with the Mediterranean Agronomic Institute of Chania. Since 2018, he has been a member of the Academic Community, TUC. He is also the Coordinator of the Industrial and Digital Innovations Research Group (indigo). He is the author of six academic books, three book chapters, and more than 70 scientific publications in international peer reviewed journals and conference proceedings. He has participated in over 30 research projects at European and national level as a technical/scientific coordinator, a work package leader, and a senior researcher. His research interests include smart ICT technologies, process, system and service engineering, energy systems management, environmental systems management, water resources management, and decision making.

DIMITRIS GRITZALIS received the B.Sc. degree in mathematics from the University of Patras, Greece, the M.Sc. degree in computer science from the City University of New York, USA, and the Ph.D. degree in information systems security from the University of the Aegean, Greece. He is currently a Professor of cybersecurity with the Department of Informatics, Athens University of Economics and Business (AUEB), Greece. He was the Director of the M.Sc. Program in Information Systems Development and Security and the INFOSEC Research Group. He was an Associate Rector of Research and Financial Affairs and the President of the Life-Long Education Center, AUEB. He has published extensively in peer-reviewed journals and conferences. His current research interests include cybersecurity, risk assessment, critical infrastructure protection, and malware. He served as the President for the Greek Computer Society and an Associate Data Protection Commissioner of Greece. He serves as an Academic Editor for *Computers and Security* (Elsevier) and a Scientific Editor for the *International Journal of Critical Infrastructure Protection* (Elsevier).