Engineering 52 (2025) 60–67

Contents lists available at ScienceDirect

Engineering

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

Research
AI for Process Manufacturing—Perspective

A Perspective on Artificial Intelligence for Process Manufacturing

Vipul Mann [a], Jingyi Lu [b], Venkat Venkatasubramanian [a], Rafiqul Gani [b,c,d,e,*]

[a] Department of Chemical Engineering, Columbia University, New York City, NY 10027, USA
[b] Key Laboratory of Smart Manufacturing in Energy Chemical Processes, Ministry of Education, East China University of Science and Technology, Shanghai 200237, China
[c] PSE for SPEED Company, Charlottenlund 2920, Denmark
[d] Department of Applied Sustainability, Széchenyi István University, Győr 9026, Hungary
[e] Sustainable Energy and Environment Thrust, The Hong Kong University of Science and Technology (Guangzhou), Guangzhou 511453, China

ARTICLE INFO

Article history:
Received 7 October 2024
Revised 15 December 2024
Accepted 6 January 2025
Available online 13 February 2025

Keywords:
Artificial intelligence
Machine learning
Process systems engineering
Manufacturing

ABSTRACT

To achieve sustainable development goals and the requirements of a circular economy, a new class of intelligent computer-aided methods and tools is needed. Artificial intelligence (AI) techniques have been gaining much attention due to their ability to provide options to tackle the challenges we are currently facing. However, the successful application of AI to solve problems of current interest requires AI to be integrated with associated process systems engineering methods and tools that are already available or being developed. In this perspective paper, we highlight the use of a collection of process systems engineering methods and tools augmented by AI techniques to solve problems related to process manufacturing, with a focus on chemical product design, process synthesis and design, process control, and process safety and hazards.

1. Introduction

Process manufacturing (PM) is a core activity in chemical, biochemical, and related engineering disciplines, as it involves the conversion of selected raw materials into products needed by modern society [1]. PM encompasses various process engineering problems, such as continuous or batch operations, quality control, operability, safety, and hazards, for a wide spectrum of chemical, biochemical, and related industries. According to Sargent [2], process systems engineering (PSE) is about problem-solving using a systems approach; that is, "Some formulate their synthesis, design and/or control problem, or some useful simplification of it, in precise mathematical terms, and then seek to exploit the mathematical structure to obtain an effective algorithm, while others seek insight on the problem structure from physical intuition." Also, many of these process engineering problems are solved both online (or in real time) and offline. As pointed out by Pistikopoulos et al. [3], artificial intelligence (AI) and related techniques can help augment the core activities of engineers to identify better products that can be made from appropriate raw materials while satisfying the constraints of sustainable development goals [4]. While methods and tools from PSE are routinely used to solve many PM problems [5], there has recently been increased interest in combining them with AI techniques. For example (to mention a few), Decardi-Nelson et al. [6] highlighted the use of generative AI in PSE, Cao et al. [7] reported the use of machine learning (ML) for the group-contribution-based property prediction of ionic liquids, Baratsas et al. [8] proposed a hybrid stochastic and ML-based forecasting model for the energy sector, Ripla [9] emphasized the leveraging of AI in manufacturing planning and scheduling, and Daoutidis et al. [10] provided a perspective on the major challenges and directions in future process control research and its industrial implementation.

After being dormant for many decades, AI has recently emerged as a major breakthrough in many fields. Venkatasubramanian [11], quoting Rich [12], has stated that "Artificial intelligence is the study of how to make computers do things at which, at the moment, people are better." He points out that the implication here is that AI could eventually end up doing all the "things" that humans do and do them much better—that is, it could achieve super-human performance, as witnessed in systems such as DeepMind's AlphaZero for Chess and Go [13], the deep reinforcement learning (RL) approach used to defeat the fastest driver in Gran Turismo Sport [14], and ChatGPT [15]. Successful application of AI techniques also depends

* Corresponding author.
E-mail address: rgani2018@gmail.com (R. Gani).

https://doi.org/10.1016/j.eng.2025.01.014
2095-8099/

on the AI's ability to perform knowledge modeling. Here, ML—also referred to as numeric AI—plays an important role in various forms, such as ensemble modeling, stochastic inference, neural networks, and reinforced learning [16]. While a combination of data, knowledge models, and inference engines leads to the development of various AI techniques, in this paper, we mainly highlight the opportunities for *hybrid AI*, which is defined as the integration of ML methods with first-principles-based methods of symbolic AI (i.e., semantics- and logic-based methods to provide insights into meaning). As shown in Fig. 1, the integration of methods for numeric AI and symbolic AI with PSE tools results in hybrid AI tools for PSE. When combined with the domain knowledge of specific systems, these tools would be able to tackle various problems related to four focused topics within PM: chemical product design, process synthesis and design, process control and monitoring, and process safety and hazards. In principle, hybrid AI tools for PSE would allow better representations of the domain knowledge of the system, improved definitions of problems to be solved, flexible model architectures that would help to determine more efficient solution strategies based on the choice of numerical solvers, and many more advantages. That is, hybrid AI tools for PSE would be able to manage the complexity of the system-specific problems being solved. Mann et al. [17] gave an example of the development and use of hybrid AI tools for PSE with applications in process synthesis, design, and innovation. Other examples can be found in mechanistic model discovery [18] and reaction prediction [19].

This perspective paper aims to highlight—for a selected set of topics within PM—the concept of augmenting well-known PSE tools with appropriate AI techniques to find novel solutions. Here, we take the view that a perspective paper should take the form of a focused review that provides the reader with an overview of the subject and gives insights into advances and challenges the future may hold, while providing selective coverage rather than an in-depth review of an area [20]. Therefore, this perspective paper promotes the use of augmented PSE tools with AI techniques for selected topics within PM. Following the introduction, we define four selected topics within PM in Section 2; next, we highlight the current state of the art in the use of AI techniques to solve problems in the selected topics in Section 3. In Section 4, we provide perspectives on future developments by which well-known PSE methods and tools could be augmented through AI techniques to increase their application spectrum and yield new and innovative solutions. Finally, in Section 5, we provide concluding remarks.

2. Problem definitions

In Fig. 2, we illustrate examples of problems to be solved related to the four selected topics within PM. It should be noted that the figure does not describe all problems related to the four topics within PM and should be considered only for purposes of concept presentation. The solution to the product design problem answers questions related to what to make. The solution to the process synthesis and design problem verifies a process flowsheet and its design to manufacture a desired product. The solution to the control system design problem verifies the achievement of the objective of the control problem, which is to keep the process at a specified set-point, subject to, for example, disturbances in the process input variables or changes in the actuator variables. In this way, the process design and control are related because the control variables are the product specifications, while the actuators are the design variables. The objective of process safety and hazards analysis is to manage the risk related to, for example, the safe operation of the process. The release of chemicals from the system is related to managing safety, waste, and sustainability.

In Fig. 2, the shaded boxes indicate the hybrid AI tools for PSE, and the directional dotted lines indicate the connections between these tools. A product design problem, such as which product to make (y) and which raw material to use (f), is answered through the process, the optimizer, and their connections. A process synthesis and design problem, such as which process (flowsheet) to employ to convert the selected raw material into the desired product, is also answered through the process, the optimizer, and their connections, but using a different set of algorithms and tools. A process control and/or monitoring problem, such as verification of the controller performance, is answered through the process, the control, the optimizer, and their connections. The role of the optimizer in each case is to provide the optimal product, flowsheet, or controller by solving numerical optimization problems with or without the aid of numeric AI. Process safety and hazards are related, for example, to the streams coming into the system (e.g., the process) and going out from the system. That is, what environmental and health hazards would result from the release of the

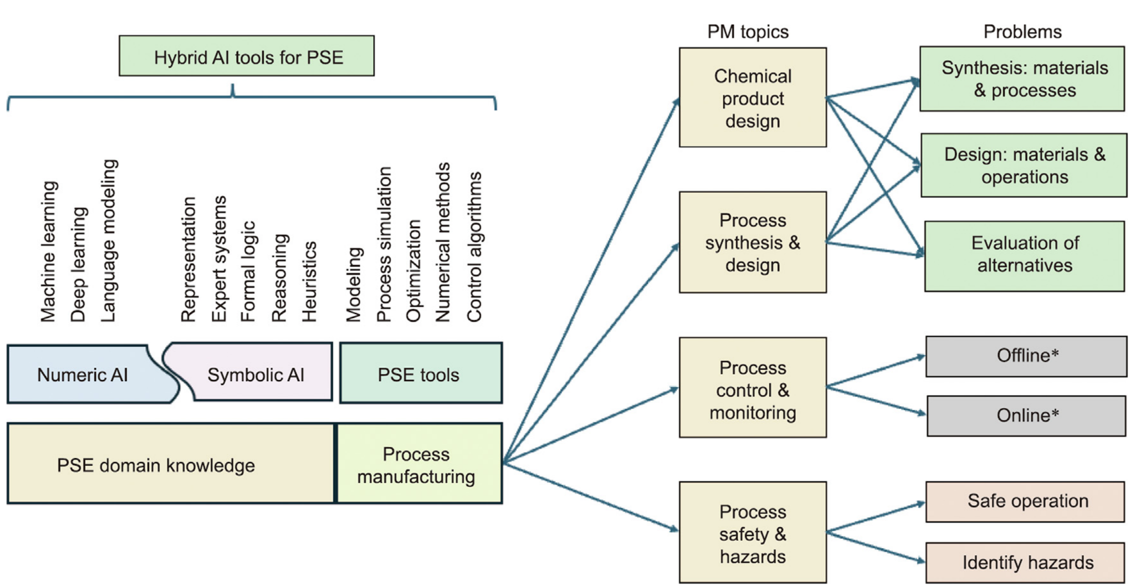

Fig. 1. The concept of hybrid AI in PM, highlighting the connections between the various components of hybrid AI tools for PSE (left), the four selected PM topics (middle), and examples of problems (right) (where * indicates "control and monitoring").

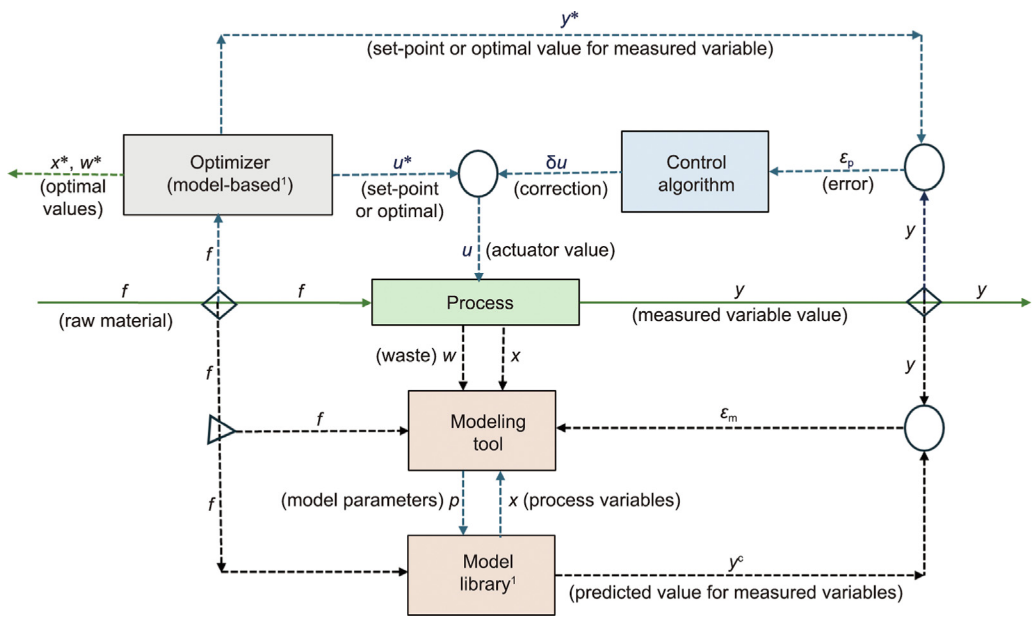

Fig. 2. Illustration of problems and their solution schemes within PM topics related to product design, process design, process control, and process analysis. In the figure, the text in parenthesis defines the corresponding symbol.

identified dangerous chemicals? It should be noted that the modeling tool and the model library perform the core function of supplying the various forms of models needed to solve all problems within the four PM topics. The efficiency of these hybrid AI tools for PSE (the shaded boxes in Fig. 2) can be optimized through the appropriate integration of selected AI techniques with the needed PSE tools.

3. Current state of the art

3.1. Chemical product design

In chemical product design, the objective is to synthesize (and design) chemicals-based products that match a set of desired product needs converted to a set of product property functions. The product could be a single species, such as a solvent; a blend of liquids, such as a liquid fuel; a functional product, such as a combination of chemicals and materials, as in a drug or cosmetic; or a device, where a chemical-based component aids a function of the device as in air purifier or inhaler. Computer-aided molecular or mixture design (CAMD) mainly involves the solution of four subproblems: representation of the product structure, generation of product alternatives, prediction of product functions, and evaluation and selection of promising alternatives. Alshehri et al. [21] have provided an overview of deep learning and knowledge-based methods for CAMD, highlighting the importance of benchmarking and empirical rigor in building deep learning models, together with a hybrid modeling paradigm that integrates deep learning approaches with an available wealth of knowledge-driven CAMD methods and tools. On the important topic of the representation of molecular structures in CAMD techniques, Mann et al. [22] proposed a Grammar2vec learning framework to generate property-agnostic molecular representations that embed structural information based on simplified molecular-input line-entry system (SMILES) grammar. This learning framework is combined with an interpretable ML framework for thermodynamic property estimation. Much work has also been done to develop more accurate property models; for example, Chen et al. [23] reported a scalable and integrated ML framework for molecular properties prediction. In the area of drug design, Che et al. [24] proposed a virtual screening framework based on binding site selectivity for small-molecule drug discovery. In the important area of chemistry and materials science related to chemical product design, Cheng et al. [25] provided an overview on the application of AI techniques.

Four examples are given in Table 1 [26–29] to highlight state-of-the-art developments in the selected areas of chemical product design. The first example is a tool for design of refrigerants and their performance verification in refrigeration processes, with the management of databases and property model libraries [26]. The second is a framework that combines chemistry rules with ML to select structural building blocks for *de novo* molecular design [27]. In the third example, a network-based approach is used to search for cost-effective, high-quality reaction pathways aimed at identifying synthesizable and viable routes, while considering cost-effectiveness, reaction yield, and diversity in the search process [28]. The last example provides a discussion on how AI and ML could assist in the development of key component materials used in cosmetics and formulated products including surfactants, polymers, fragrances, preservatives, and hydrogels [29].

Table 1
List of selected state-of-the-art developments in chemical product design.

Hybrid AI approach	Description	Ref.
Data management combined with CAMD	Highlights integrated refrigerant design and refrigeration cycle design through a database, CAMD, and modeling	[26]
Structural building blocks for *de novo* molecular design	Designs innovative new synthetic entities generated by optimization strategies	[27]
Network-based reaction pathway search	Searches for cost-effective, high-quality reaction pathways considering multiple factors	[28]
Use of AI and ML in the development of cosmetics and formulated products	Contributes to the development of materials such as surfactants, polymers, fragrances, preservatives, and hydrogels	[29]

3.2. Process synthesis and design

Process synthesis aims to identify the optimal processing route for converting a given raw material into a desired product. In contrast, process design aims to determine the optimal design of the operations within the selected processing route. From a sustainability perspective, the designed process must also satisfy performance criteria related to economics, operability, and sustainability. Hybrid AI approaches combining process knowledge and data-driven methods are increasingly being proposed for process synthesis, design, and innovation. Tula et al. [30] highlighted ProCAFD, an integrated computer-aided method with associated AI-augmented software that searches through the entire design space for sustainable process options while considering indicators for economic, life-cycle, sustainability, process safety, and operability aspects. Several PSE methods and tools augmented with AI techniques are employed at every step of the synthesis, design, and innovation workflow. Oeing et al. [31] presented an ML classifier that recommends appropriate separation systems. The classifier operates on a vector of process values, including relevant substance properties and characteristic physical quantities. The database is constructed using examples of processes from a flowsheet simulator to train and validate the model. Savage et al. [32] developed hybrid ML-based surrogates through a combination of Gaussian processes and artificial neural networks to efficiently model nonlinear systems, such as entire process flowsheets. Their study demonstrated that, when integrated with efficient stochastic algorithms, these data-driven models can identify high-quality optimal solutions while significantly reducing the computational time.

Table 2 [17,33–35] lists selected examples of state-of-the-art developments in the application of AI in process synthesis and design. The first example is an AI-based multi-level flowsheet representation and generation framework that incorporates text-based, hypergraph-based, and ontology-based methods with embedded model-based calculation tools [17], where flowsheets at each level, represented by the extended simplified flowsheet-input line-entry system (eSFILES), are connected with process generators, process simulators, and inferencing algorithms to solve various flowsheet-related interconnected problems. The second example is a surrogate modeling approach that uses Bayesian symbolic regression [33]. The third is the development of a hybrid framework for decomposition-based process design that combines RL with mathematical programming using a two-level agent structure. The higher level connects process sections to construct the overall process, while the lower level learns to build and solve individual process sections, facilitating modularity and flexibility [34]. The last example is the development of an evolutionary method for process synthesis (process synthesis by evolution, PSEvo) that utilizes redefined unit operations and their associated models as elementary building blocks [35]. It should be noted that, as Fig. 1 indicates, hybrid AI tools for PSE incorporate domain knowledge; therefore, although a great deal of development has been achieved with purely data-driven approaches based entirely on deep learning and/or RL for process synthesis and design, such approaches are not covered in this paper.

3.3. Process control and monitoring

In a tutorial review paper, Ren et al. [36] highlighted recent developments in time-series neural network modeling, along with its use in model predictive control (MPC). A nonlinear process example was introduced to demonstrate the application of different neural-network-based modeling approaches and to evaluate their performance in terms of closed-loop stability and prediction accuracy. These approaches often treat chemical processes as a black box. To address this issue, data-driven neural network models have been combined with mechanism models to reduce their dependency on the scale and quality of the data. Another approach in learning-based control focuses on directly learning the control policy without relying on a process model. Here, the control input is adjusted by modifying the control signal from the previous cycle based on the tracking error within the framework of iterative learning control (ILC). For continuous processes, the model can be replaced by a data-dependent representation of the system's open-loop and closed-loop dynamics [37]. This approach is based on a fundamental lemma, which states that a finite set of system trajectories can represent a linear system's entire set of trajectories, provided these trajectories arise from sufficiently excited dynamics. RL is another typical example of process control for both batch and continuous processes. It has been found to outperform MPC in real-time execution, since it only involves a forward run of the policy network, whereas MPC must solve several optimization problems sequentially. A comparison of RL and MPC was reported by Shin et al. [38], who emphasized that, instead of building a model of the process as MPC generally does, RL builds a model of the "closed-loop" objective function through trial and error. However, system safety and stability are crucial issues to consider, as RL learns the control policy via exploration and exploitation. The system may thus be driven into unknown operational regions during exploration without guaranteed safety and stability.

Table 3 [39–42] lists examples of selected state-of-the-art developments in the application of AI in process control and

Table 2
List of selected state-of-the-art developments in process synthesis and design.

Hybrid AI approach	Description	Ref.
eSFILES	Multi-level flowsheet representation and generation framework combining symbolic AI and process engineering knowledge	[17]
Bayesian symbolic regression	Surrogate modeling approach using Bayesian symbolic regression for improved performance in process optimization	[33]
RL-based process design	Hybrid framework combining RL with mathematical programming for decomposition-based process design	[34]
PSEvo	Evolutionary method for process synthesis using redefined unit operations and nonlinear programming	[35]

eSFILES: extended simplified flowsheet-input line-entry system; PSEvo: process synthesis by evolution.

Table 3
List of selected state-of-the-art developments in process control and monitoring.

Learning-based modeling and control	Description	Ref.
PINN-based modeling	Incorporates prior physical information, such as plant dynamics and reaction mechanisms, into the training phase of a neural network model	[39]
Iterative learning control	Presents a control strategy that directly adjusts the control input based on the input signal and tracking errors from the previous cycle in batch processes	[40]
Data-driven control	Designs a controller using data from system trajectories without relying on a predefined process model structure	[41]
Industrial AI	Provides perspectives on nonstationary process monitoring in the era of industrial AI	[42]

PINN: physics-informed neural network.

monitoring. The first example involves the use of process models identified via physics-informed neural networks (PINNs) in MPC to reduce online computation by avoiding the need to solve large-scale ordinary differential equations (ODEs) [39]. The second involves batch process control, where the controlled input is adjusted by modifying the control signal from the previous cycle based on a tracking error, within the framework of ILC [40]. The third example focuses on how to incorporate data-dependent stability and performance requirements into the control design process [41]. The last example relates to aspects of nonstationary process monitoring in the era of industrial AI [42].

3.4. Process safety and hazards

The flipside of real-time process control and fault diagnosis is process safety or hazard analysis. Disasters happen in different domains and facilities, are triggered by different events, and involve different factors. Nevertheless, there are certain common underlying patterns behind such systemic failures. There is an alarming similarity in major accidents, which underscores the important fundamental lessons that must be learned to prevent such events from recurring. To understand these patterns and learn from them, it is necessary to go beyond analyzing them as independent one-off accidents and examine them from the broader perspective of the potential fragility of all complex engineered systems. All these disasters must be studied from a common systems engineering perspective in order to thoroughly understand their commonalities and differences and thus design and control such systems better in the future. Furthermore, such studies must be carried out in concert with public policy experts in order to translate all the scientific and engineering lessons into effective policies and regulations. AI can assist us enormously in this challenging endeavor. As described by Venkatasubramanian et al. [43–45] and Zhao et al. [46,47], a hybrid AI system can reduce the time, effort, and expense involved in a process hazards analysis (PHA) review; make the review more thorough, detailed, and consistent; minimize human errors; and free the team to concentrate on the more complex aspects of the analysis, which are unique and challenging to automate.

Table 4 [48–51] lists selected examples of state-of-the-art developments in the application of AI in process safety and hazards. The first example focuses on the adoption of transfer learning to facilitate the training of RL [48], while the second covers safe transfer learning for accurate quantifying mismatching (between a real plant and a digital twin) or system uncertainty based on online data [49]. The third example involves the identification of abnormal operational conditions based on on-off data-driven fault detection, fault isolation, and fault identification methods using both online and offline process measurements [50]. In the last example, quantitative structure–use relationship (QSUR) models are used to screen for potential functional substitutes while incorporating hazard metrics [51].

Table 4
List of selected state-of-the-art developments in the application of AI in process safety and hazards.

Hybrid AI approach	Description	Ref.
Sim2Real transfer	Trains the control policy offline using a simulator, then applies it to real plants using transfer learning to avoid unsafe exploration	[48]
Uncertainty qualification	Provides a measure of system uncertainty so it can be explicitly considered in controller and process design	[49]
Data-driven fault diagnosis	Identifies abnormal operational conditions using process data	[50]
QSUR	Screens potential functional substitutes using a functional use database and hazard metrics	[51]

QSUR: quantitative structure–use relationship.

4. Perspectives and future directions

We argue that many engineering applications require the incorporation of first principles-based knowledge, which is often not explicitly considered in purely data-driven approaches such as large language models (LLMs). The success of LLMs is limited in highly scientific domains, as they cannot yet reason, plan, or explain due to their lack of in-depth mechanistic domain knowledge. This is a problem in domains such as chemical engineering, which are governed by fundamental laws of physics and chemistry (and biology), constitutive relations, and highly technical knowledge about materials, processes, and systems. Although purely data-driven ML has its immediate uses, the long-term success of AI in scientific and engineering domains would depend on the development of hybrid AI systems that effectively combine first principles and technical knowledge. We call these hybrid AI tools "large knowledge models" (LKMs) [52]. In this section, we outline the challenges and opportunities in developing LKMs for chemical product design, process synthesis and design, process control and monitoring, and process safety and hazards.

4.1. Chemical product design

While significant progress has been made toward developing frameworks for hybrid AI tools for chemical product design, there are several outstanding challenges that must be addressed to ensure the wider adoption of such frameworks. AI applications in chemistry and materials science will need to be embedded within the workflow of chemical product design [25]. Three additional perspectives and challenges are discussed below:

- **Better utilization of existing libraries of chemicals.** The available property databases need to be integrated into AI frameworks, such that, in addition to using them for model training, the application ranges of the property models and their predictive power and accuracy can be significantly improved. Here, a combination of fundamental principles of chemistry and thermodynamics, molecular structure versus property relations, analysis of available data, and so forth is essential. For newly synthesized molecules (or mixtures), uncertainties in property estimates could be minimized through AI-aided models that can safely extrapolate beyond the boundaries defined by the available data.
- **Computer power versus efficient computational algorithms.** Even though computation has become cheaper, and it is easier to train end-to-end models or foundational models for product property predictions, searching for new and/or optimal molecules with desired property functions for large molecules or polymers, such as in drugs or detergents, remains a challenge. Developing a computationally intensive data-driven modeling approach integrated with *a priori* knowledge-based intelligent problem solution algorithms could be a better option than aiming for bigger and faster computers.
- **Managing complexity with hybrid AI.** Computer-aided chemical product design often requires the matching of several properties, and it may not be possible for a single molecule to match them. In this case, a problem reformulation is necessary that may either change the target properties or look for a mixture instead of a single species. Although the generation of candidates (i.e., molecular structures or mixtures) is not the limiting step, as any combinatorial explosion for the generation step can

be handled through smart hybrid AI techniques, accurate evaluation of the generated candidates is still a challenge, and hybrid AI-based frameworks that facilitate multi-property optimization across multiple scales could make a useful contribution in this area.

4.2. Process synthesis and design

While progress has been achieved in the area of process synthesis and design, significant additional work is needed to enable a wider spectrum of applications. Current AI-augmented systems are mainly limited to chemical processes, handling the production of organic chemicals. Data and domain knowledge are limited for pharmaceuticals, food, and related industrial sectors. Consequently, models for process operations with embedded constitutive (property) models for the involved operations are not available in process synthesis, design, and innovation. The use of data-based surrogates could be an option if the measured data satisfies the essential thermodynamic principles (e.g., the conservation of mass and energy) that govern the operation of these processes [53]. Three additional perspectives and challenges are discussed below:

- **A unified, context-aware database of process flowsheets.** A unified database with appropriate flowsheet context and process knowledge would be able to facilitate the development of data-driven hybrid approaches for problems related to flowsheet synthesis and design. The limitations of AI-augmented approaches that rely purely on data-driven methods to generate process flow diagrams could be addressed by including the path of chemicals as they enter and leave the generated processing route. Without this, the flow diagram serves no purpose; notably, the chemical paths are important because they determine the need for specific operations within the processing route. An important question is how much *a priori* knowledge is needed to generate (or synthesize) the processing route and design the process operations.
- **Sustainability as a core component of flowsheet development.** Given the current challenges humanity is facing, it is no longer sufficient to design a process that is economically feasible and satisfies some process specifications. It is now also necessary for the process to be sustainable, safe, operable, and so forth, in addition to being economically feasible, minimizing waste, and reducing greenhouse gas emissions. Hybrid AI techniques can help to obtain the desired sustainable and novel alternatives.
- **The integration of optimization-based methods with hybrid AI.** While optimization-based frameworks offer advantages in terms of incorporating domain knowledge, a more versatile integration with ML and AI methods—that is, a hybrid AI-based modeling tool—could provide a more flexible framework with a built-in library of surrogates that could be employed to formulate and solve a wide range of process optimization problems.

4.3. Process control and monitoring

In the realm of process controller design and monitoring, data has become increasingly pivotal in refining process models, enhancing control policies, and accelerating online optimization. However, current learning-based control and monitoring schemes depend heavily on extensive measurements from manufacturing processes. There is a pressing need to extend these schemes to more complex manufacturing scenarios, such as changing operational conditions and scarce feedback signals [54]. Four additional perspectives and challenges are discussed below:

- **Adapting to changing operational conditions.** When operational conditions change, such as variations in raw material supply, the process dynamics also shift. It is crucial for the process models learned from data to quickly adapt to these new conditions. However, newly arriving data may have a different distribution from previously collected data. A key question is how to efficiently learn process dynamics under shifting data distributions and train ML models incrementally.
- **Handling limited feedback signals and ensuring safety.** In many manufacturing scenarios, operational procedures are highly complex, with limited feedback signals and strict safety requirements. In such cases, control policies learned from data may not perform as well as those based on human expertise. Learning control policies from expert demonstrations can mitigate operational risks. A fundamental challenge in this context lies in effectively addressing the scarcity and non-optimality of expert demonstrations. Integrating techniques such as inverse RL and few-shot learning into the design can be valuable.
- **Incorporating diverse measurement signals.** With advancements in measurement techniques, measurement signals have become diverse in modality. For instance, industrial cameras generate images, and acoustic signals can also be utilized in controller design. The challenge lies in efficiently incorporating multiple modalities and sources of signals into controller design. Techniques such as multi-modal information fusion and multi-fidelity learning can be explored to address this issue.
- **Practical implementation of AI-augmented control algorithms.** Although RL has made considerable advances in industrial process control applications, several critical challenges persist, notably concerning training sample efficiency and ensuring closed-loop stability in practical settings. Exploring the integration of RL and MPC either within a hierarchical architecture or through an online–offline combination approach warrants a thorough investigation.

4.4. Process safety and hazards

Although there has been recent progress in the application of AI for fault detection and diagnosis in pharmaceutical bioprocesses [55], most of the discussed techniques were demonstrated about 20–30 years back, as outlined in Refs. [56–58]. A major issue is the collection of data—such as a list of chemicals in use, their properties, their hazardous effects, and much more information. As measured data for chemicals are limited, predictive models are needed. However, can ML-based models provide the required predictive power? Significant efforts are needed to avoid human-made disasters. Three additional perspectives and challenges are discussed below:

- **A database of dangerous chemicals.** It is necessary to identify the most hazardous chemicals, regulate their use, and provide guidelines for actions if these chemicals are accidentally released. This will require a significant investment of time and resources, as measured data is available for only a fraction of the chemicals that have been identified. Also, missing data for a chemical does not necessarily indicate that it is safe.
- **Creation of better language models.** There is a need to create LKMs that relate products with their manufacturing processes, the list of chemicals involved, and their regulations in order to prevent disasters. This may require embedding property-model-based calculation tools within the problem-solving workflow using the LKMs.
- **Integration of hazardous and safety issues within process analysis.** Given the nature of current industrial pollution levels and the threat from hazardous chemicals, hazardous aspects

need to be more fundamentally ingrained in such model developments. Restricting the space of hazardous chemicals such that they are never generated or explored, rather than using an *ad hoc* filtering approach, should be considered. Inspiration from language models in which toxicity is penalized during training could be used to develop better guidelines.

5. Conclusions

The PM systems highlighted in Figs. 1 and 2 represent tasks that present unique challenges in modeling and execution, exceeding those encountered in typical process design, control, and safety [59]. These systems consist of social elements (i.e., humans) and technical components (i.e., the design of process operations such as pumps, valves, and reactors). They are particularly complex due to the integral role of human elements and the potential for major systemic failures. Designing such systems and their control mechanisms at various levels to ensure safe operation throughout their life cycles is a highly challenging task. The vast number of interconnected components with nonlinear interactions in these systems can result in "emergent" behavior, making it difficult to predict, optimize, and control. In addition, these systems do not operate in isolation; they interact with physical, market, and regulatory environments, leading to intricate feedback dynamics. Human decision-making and associated errors are key components of these feedback processes. Their nonlinearity, interconnectedness, and interactions with humans and the environment render these system-of-systems particularly fragile and prone to systemic failures. This also highlights the importance of product, process, control system, and system analysis integration.

While successful applications of AI-augmented PSE methods and tools can be found in specific manufacturing systems, such as image processing, fault diagnosis, and process control, a great deal of work is needed to develop AI-augmented PSE tools to solve a wide range of practical engineering problems, such as the integrated design and operation of a sustainable chemical process or the design of a novel material for currently infeasible reaction synthesis or low-energy-consumption separation processes. Instances of large-scale failures in various fields, such as the global financial crisis (2007–2008) [60], the BP Deepwater Horizon oil spill (2010) [61], and the Indian power outage (2012) [62], serve as continuous reminders of the vulnerability inherent in complex sociotechnical systems [63]. These instances underscore the urgent need for real-time operator advisory systems.

Finally, can hybrid AI help to make intelligent decisions *a priori* so that innovative sustainable PM can be performed with efficient human decision-making and/or avoid systemic failures? It is notable that systemic failures may occur when a complete system breaks down, typically involving a significant entity whose malfunction adversely affects many individuals and their environment, leading to substantial financial damage. Therefore, a new class of AI-augmented PSE tools is necessary so that the needed data can be efficiently transferred to model-based process simulation and/or optimization techniques for failure-free decision-making. However, such tools will require knowledge models to link with a wide range of PSE tools, such as chemical properties prediction, process operation verification, process synthesis and/or design methods, and many more. The goal is to solve problems such as the integrated synthesis, design, and manufacturing of large drug molecules or zero-emission sustainable chemical processes that manufacture products needed by modern society through augmented knowledge models. To achieve these goals, models, modeling, and model-based systems must continue to play a major role in the development and application of hybrid AI tools for PSE [53].

CRediT authorship contribution statement

Vipul Mann: Writing – original draft, Conceptualization. **Jingyi Lu:** Writing – original draft, Conceptualization. **Venkat Venkatasubramanian:** Writing – original draft, Conceptualization. **Rafiqul Gani:** Writing – review & editing, Writing – original draft, Conceptualization.

Declaration of competing interest

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

Acknowledgments

Jingyi Lu and Rafiqul Gani acknowledge the funding support from the National Natural Science Foundation of China (62394343) and the Program of Introducing Talents of Discipline to Universities (the 111 Project) (B17017).

References

[1] Gani R, Bałdyga J, Biscans B, Brunazzi E, Charpentier JC, Drioli E, et al. A multi-layered view of chemical and biochemical engineering. Chem Eng Res Des 2020;155:133–45.
[2] Sargent RWH. Advances in modelling and analysis of chemical process systems. Comput Chem Eng 1983;7(4):219–37.
[3] Pistikopoulos EN, Barbosa-Povoa A, Lee JH, Misener R, Mitsos A, Reklaitis GV, et al. Process systems engineering—the generation next? Comput Chem Eng 2021;147:107252.
[4] Sustainable development goals [Internet]. New York City: United Nations; [cited 2025 Jan 21]. Available from: https://www.un.org/sustainabledevelopment/.
[5] Grossmann IE, Harjunkoski I. Process systems engineering: academic and industrial perspectives. Comput Chem Eng 2019;126:474–84.
[6] Decardi-Nelson B, Alshehri AS, Ajagekar A, You F. Generative AI and process systems engineering: the next frontier. Comput Chem Eng 2024;187:108723.
[7] Cao P, Chen J, Chen G, Qi Z, Song Z. A critical methodological revisit on group-contribution based property prediction of ionic liquids with machine learning. Chem Eng Sci 2024;298:120395.
[8] Baratsas S, Iseri F, Pistikopoulos EN. A hybrid statistical and machine learning based forecasting framework for the energy sector. Comput Chem Eng 2024;188:108740.
[9] Ripla A. Leveraging AI in manufacturing planning and scheduling for manufacturing excellence [Internet]. Sunnyvale: LinkedIn Corporation; 2024 Apr 25 [cited 2025 Jan 21]. Available from: https://www.linkedin.com/pulse/leveraging-ai-manufacturing-planning-scheduling-andre-ripla-pgcert-1lobe/.
[10] Daoutidis P, Megan L, Tang W. The future of control of process systems. Comput Chem Eng 2023;178:108365.
[11] Venkatasubramanian V. The promise of artificial intelligence in chemical engineering: is it here, finally? AIChE J 2019;65(2):466–78.
[12] Rich E. Artificial intelligence. New York City: McGraw-Hill, Inc.; 1983.
[13] Silver D, Hubert T, Schrittwieser J, Hassabis D. AlphaZero: shedding new light on chess, shogi, and Go [Internet]. London: Google DeepMind; 2018 Dec 6 [cited 2025 Jan 21]. Available from: https://deepmind.google/discover/blog/alphazero-shedding-new-light-on-chess-shogi-and-go.
[14] Gran Turismo Sport [Internet]. San Mateo: Sony Interactive Entertainment Inc.; [cited 2025 Jan 21]. Available from: https://www.gran-turismo.com/us/gtsport/top/.
[15] openai.com/chatgpt [Internet]. San Francisco: OpenAI, LLC; [cited 2025 Jan 21]. Available from: https://openai.com/chatgpt/.
[16] Lee JH, Shin J, Realff MJ. Machine learning: overview of the recent progresses and implications for the process systems engineering field. Comput Chem Eng 2018;114:111–21.
[17] Mann V, Sales-Cruz M, Gani R, Venkatasubramanian V. eSFILES: intelligent process flowsheet synthesis using process knowledge, symbolic AI, and machine learning. Comput Chem Eng 2024;181:108505.
[18] Chakraborty A, Sivaram A, Venkatasubramanian V. AI-DARWIN: a first principles-based model discovery engine using machine learning. Comput Chem Eng 2021;154:107470.
[19] Mann V, Venkatasubramanian V. Predicting chemical reaction outcomes: a grammar ontology-based transformer framework. AIChE J 2021;67(3):e17190.
[20] Mann V, Gani R, Venkatasubramanian V. Group contribution-based property modeling for chemical product design: a perspective in the AI era. Fluid Phase Equilib 2023;568:113734.
[21] Alshehri AS, Gani R, You F. Deep learning and knowledge-based methods for computer-aided molecular design—toward a unified approach: state-of-the-art and future directions. Comput Chem Eng 2020;141:107005.

[22] Mann V, Brito K, Gani R, Venkatasubramanian V. Hybrid, interpretable machine learning for thermodynamic property estimation using Grammar2vec for molecular representation. Fluid Phase Equilib 2022;561:113531.
[23] Chen G, Song Z, Qi Z, Sundmacher K. A scalable and integrated machine learning framework for molecular properties prediction. AIChE J 2023;69(10):e18185.
[24] Che X, Liu Q, Yu F, Zhang L, Gani R. A virtual screening framework based on the binding site selectivity for small molecule drug discovery. Comput Chem Eng 2024;184:108626.
[25] Cheng AH, Ser CT, Skreta M, Guzmán-Cordero A, Thiede L, Burger A, et al. Spiers Memorial Lecture: how to do impactful research in artificial intelligence for chemistry and materials science. Faraday Discuss 2025;256:10–60.
[26] Kuprasertwong N, Padungwatanaroj O, Robin A, Udomwong K, Tula AK, Zhu L, et al. Computer-aided refrigerant design: new developments. Computer-Aided Chem Eng 2021;50:19–24.
[27] Button A, Merk D, Hiss JA, Schneider G. Automated *de novo* molecular design by hybrid machine intelligence and rule-driven chemical synthesis. Nat Mach Intell 2019;1(7):307–15.
[28] Badowski T, Molga K, Grzybowski BA. Selection of cost-effective yet chemically diverse pathways from the networks of computer-generated retrosynthetic plans. Chem Sci 2019;10(17):4640–51.
[29] Xin H, Virk AS, Virk SS, Akin-Ige F, Amin S. Applications of artificial intelligence and machine learning on critical materials used in cosmetics and personal care formulation design. Curr Opin Colloid Interface Sci 2024;73:101847.
[30] Tula AK, Eden MR, Gani R. Hybrid method and associated tools for synthesis of sustainable process flowsheets. Comput Chem Eng 2019;131:106572.
[31] Oeing J, Henke F, Kockmann N. Machine learning based suggestions of separation units for process synthesis in process simulation. Chemie Ingenieur Technik 2021;93(12):1930–6.
[32] Savage T, Almeida-Trasvina HF, del Río-Chanona EA, Smith R, Zhang D. An adaptive data-driven modelling and optimization framework for complex chemical process design. Computer-Aided Chem Eng 2020;48:73–8.
[33] Jog S, Vázquez D, Santos LF, Caballero JA, Guillén-Gosálbez G. Hybrid analytical surrogate-based process optimization via Bayesian symbolic regression. Comput Chem Eng 2024;182:108563.
[34] Khan AA, Lapkin AA. Designing the process designer: hierarchical reinforcement learning for optimisation-based process design. Chem Eng Process 2022;180:108885.
[35] Neveux T. *Ab-initio* process synthesis using evolutionary programming. Chem Eng Sci 2018;185:209–21.
[36] Ren YM, Alhajeri MS, Luo J, Chen S, Abdullah F, Wu Z, et al. A tutorial review of neural network modeling approaches for model predictive control. Comput Chem Eng 2022;165:107956.
[37] Willems JC, Rapisarda P, Markovsky I, De Moor BLM. A note on persistency of excitation. Syst Control Lett 2005;54(4):325–9.
[38] Shin J, Badgwell TA, Liu KH, Lee JH. Reinforcement learning—overview of recent progress and implications for process control. Comput Chem Eng 2019;127:282–94.
[39] Karniadakis GE, Kevrekidis IG, Lu L, Perdikaris P, Wang S, Yang L. Physics-informed machine learning. Nat Rev Phys 2021;3(6):422–40.
[40] Lu J, Cao Z, Zhao C, Gao F. 110th anniversary: an overview on learning-based model predictive control for batch processes. Ind Eng Chem Res 2019;58(37):17164–73.
[41] De Persis C, Tesi P. Formulas for data-driven control: stabilization, optimality, and robustness. IEEE Trans Automat Contr 2020;65(3):909–24.
[42] Zhao C. Perspectives on nonstationary process monitoring in the era of industrial artificial intelligence. J Process Contr 2022;116:255–72.
[43] Venkatasubramanian V, Zhao J, Viswanathan S. Intelligent systems for HAZOP analysis of complex process plants. Comput Chem Eng 2000;24(9–10):2291–302.
[44] Venkatasubramanian V, Zhang Z. TeCSMART: a hierarchical framework for modeling and analyzing systemic risk in sociotechnical systems. AIChE J 2016;62(9):3065–84.
[45] Venkatasubramanian V, Luo Y, Zhang Z. Control of complex sociotechnical systems: importance of causal models and game theory. Comput Chem Eng 2019;123:1–11.
[46] Zhao C, Bhushan M, Venkatasubramanian V. PHASuite: an automated HAZOP analysis tool for chemical processes: part I: knowledge engineering framework. Process Saf Environ Prot 2005;83(6):509–32.
[47] Zhao C, Bhushan M, Venkatasubramanian V. PHASuite: an automated HAZOP analysis tool for chemical processes: part II: implementation and case study. Process Saf Environ Prot 2005;83(6):533–48.
[48] Lin R, Chen J, Xie L, Su H, Huang B. Facilitating reinforcement learning for process control using transfer learning: perspectives. 2024. arXiv:2404.00247.
[49] Cheng S, Quilodrán-Casas C, Ouala S, Farchi A, Liu C, Tandeo P, et al. Machine learning with data assimilation and uncertainty quantification for dynamical systems: a review. IEEE/CAA J Autom Sin 2023;10(6):1361–87.
[50] Ding SX. Data-driven design of fault diagnosis and fault-tolerant control systems. London: Springer; 2015.
[51] Phillips KA, Wambaugh JF, Grulke CM, Dionisio KL, Isaacs KK. High-throughput screening of chemicals as functional substitutes using structure-based classification models. Green Chem 2017;19(4):1063–74.
[52] Venkatasubramanian V, Chakraborty A. *Quo vadis* ChatGPT? From large language models to large knowledge models. Comput Chem Eng 2025;192:108895.
[53] Mansouri SS, Sivaram A, Savoie CJ, Gani R. Models, modeling and model-based systems in the era of computers, machine learning and AI. Comput Chem Eng 2025;194:108957.
[54] Raven DB, Chikkula Y, Patel KM, Al Ghazal AH, Salloum HS, Bakhurji AS, et al. Machine learning & conventional approaches to process control & optimization: industrial applications & perspectives. Comput Chem Eng 2024;189:108789.
[55] Aghaee M, Mishra A, Krau S, Tamer IM, Budman H. Artificial intelligence applications for fault detection and diagnosis in pharmaceutical bioprocesses: a review. Curr Opin Chem Eng 2024;44:101025.
[56] Dash S, Venkatasubramanian V. Integrated framework for abnormal event management and process hazards analysis. AIChE J 2003;49(1):124–39.
[57] Venkatasubramanian V. Prognostic and diagnostic monitoring of complex systems for product life cycle management: challenges and opportunities. Comput Chem Eng 2005;29(6):1253–63.
[58] Hamdan IM, Reklaitis GV, Venkatasubramanian V. Real-time exceptional events management for a partial continuous dry granulation line. J Pharm Innov 2012;7(3–4):95–118.
[59] Venkatasubramanian V. Systemic failures: challenges and opportunities in risk management in complex systems. AIChE J 2011;57(1):2–9.
[60] Duignan B. Financial crisis of 2007–08: global economics [Internet]. Chicago: Britannica Money; 2019 Mar 13 [cited 2025 Jan 21]. Available from: https://www.britannica.com/money/financial-crisis-of-2007-2008.
[61] Pallardy R. Deepwater Horizon oil spill [Internet]. Chicago: Encyclopædia Britannica, Inc.; 2010 Jul 9 [cited 2025 Jan 21]. Available from: https://www.britannica.com/event/Deepwater-Horizon-oil-spill.
[62] Lai LL, Zhang HT, Lai CS, Xu FY, Mishra S. Investigation on July 2012 Indian blackout. In: Proceedings of the 2013 International Conference on Machine Learning and Cybernetics; 2013 Jul 14–17; Tianjin, China. Piscataway: IEEE; 2013. p. 92–7.
[63] Bookstaber R, Glasserman P, Iyengar G, Luo Y, Venkatasubramanian V, Zhang Z. Process systems engineering as a modeling paradigm for analyzing systemic risk in financial networks. J Invest 2015;24(2):147–62.