YOUR FIELD · CORE DOMAIN

Chemical Engineering and Process Systems Engineering (CE&PSE)

Discusses the transformative impact of Generative AI on chemical and process systems engineering, highlighting challenges, opportunities, and applications.

Christophe Pennetier, André-Benoît De Jaegere

Generative Engineering · Volume I · written November 2024 to February 2025, two years after ChatGPT; its text is kept as published · https://generative-engineering.aquia.studio/fields/chemical/

CONTENTS

Overview 1
The Applications of GenAI in Chemical Engineering 2
How GenAI is Transforming Chemical Engineering 4
Case Studies of AI in Chemical Process Optimization 6
AI and Precision Solution Chemistry 8
GenAI in Continuous Industrial Process Engineering 10
GenAI for Fault Detection and Diagnosis of Industrial Process Systems 13
Challenges and Limitations 15
Future Directions for Gen AI in Chemical and Process Engineering 18
Startups and Projects Utilizing Explainable AI (XAI) in Chemical Engineering (CE) and Process Systems Engineering (PSE) 24
Conclusion 26
Citations for Section on Chemical Engineering 26
General References for Section on Chemical Engineering 29

Overview

Generative Artificial Intelligence (GenAI), driven by advancements in machine learning models such as transformers and diffusion models, has emerged as a transformative technology across multiple domains. In the fields of Chemical Engineering (CE) and Process Systems Engineering (PSE), GenAI has shown remarkable potential for automating traditionally complex tasks, optimizing workflows, and innovating processes, products, and materials. This literature review synthesizes recent research and applications to outline the impact, challenges, and opportunities of GenAI in these domains.

The Applications of GenAI in Chemical Engineering

Generative AI (GenAI) is increasingly being applied in chemical engineering to address challenges in molecular design, process optimization, and catalyst development. We will review:

1.

GenAI fields of applications in molecular and material design, process optimization and catalyst development

2.

How Generative AI is transforming chemical engineering by optimizing processes across multiple scales, from molecular to enterprise levels

3.

Case Studies of GenAI in Chemical Process Optimization

4.

AI and the Precision Solution Chemistry

Molecular and Material Design

GenAI is revolutionizing molecular and material design by enabling the creation of novel molecules and materials with specific desired properties. One of the most impactful applications of GenAI in chemical engineering is its role in molecular and material design. Generative models, such as Variational Autoencoders (VAEs) and Generative Adversarial Networks (GANs), have been extensively utilized to generate novel molecular structures with desired properties. These models, trained on chemical databases like PubChem or ChEMBL, can design molecules that optimize trade-offs between competing properties such as reactivity, stability, and environmental impact.

Example:

A study utilizing transformer-based models like MolGPT demonstrated the generation of drug-like molecules with superior pharmacokinetics. By integrating multi-objective optimization frameworks, researchers have also generated molecules aimed at decarbonization, such as CO2-absorbing materials.

Key advancements include:

Generative Models for Molecular Design: Techniques like Variational Autoencoders (VAEs) and reinforcement learning are being used to generate valid molecular structures that meet predefined property thresholds. For instance, a study combined deep learning with Self-Referencing Embedded Strings (SELFIES) to design surfactants with targeted attributes, showcasing the potential for tailored molecule creation[1].

Multimodal Generative Frameworks: Advanced models like X-LoRA-Gemma integrate multimodal generative AI to analyze, design, and test molecules. These frameworks use AI-AI and human-AI collaboration to identify molecular targets and optimize properties such as dipole moments and polarizability[4].

Startups in Material Discovery: Companies like CuspAI leverage GenAI to generate materials on demand, such as those for carbon capture[5]. Their platform combines molecular simulation and deep learning to streamline material discovery processes.

Process Optimization

GenAI is being applied to optimize chemical engineering processes, improving efficiency and reducing costs:

Reinforcement Learning for Process Automation: GenAI models powered by reinforcement learning dynamically adapt to complex production environments. These models autonomously fine-tune parameters, anticipate maintenance needs, and minimize downtime in manufacturing systems[2].

Business Process Management (BPM): Generative AI enhances operational efficiency by automating repetitive tasks, detecting bottlenecks, and optimizing workflows across industries. This has implications for chemical engineering processes such as supply chain management and resource allocation[6].

Catalyst Development

Catalyst design is another area where GenAI is making significant contributions:

Generative Models for Catalyst Design: AI-driven generative models propose new molecular structures with targeted catalytic properties.

These models are particularly suited for exploratory optimization problems beyond known molecular spaces[3].

High-Throughput Screening: Machine learning (ML) techniques integrated with generative models streamline catalyst discovery by predicting reaction mechanisms, constructing phase diagrams, and optimizing synthesis pathways[3].

Industrial Applications: Improved catalytic reactions designed using GenAI can enhance the efficiency of industrial chemical processes, such as those involved in sustainable energy production or emissions reduction.

Examples of Industry Adoption

CuspAI
Focuses on AI-designed materials for sustainability challenges like carbon capture. Their platform combines GenAI with experimental validation to ensure practical applicability[5].

S&P Global
Recently acquired ProntoNLP to expand its generative AI capabilities for data analysis. While not directly chemical-specific, this highlights broader adoption of GenAI tools by major corporations[9].

Pharmaceutical Companies
Firms like Bristol-Myers Squibb are leveraging generative AI for drug discovery, which shares parallels with catalyst development in terms of molecular design techniques[6.7-7][10].

How GenAI is Transforming Chemical Engineering

This section describes how GenAI enhances chemical processes from real-time optimization to energy efficiency and multiscale applications.

Real-Time Process Optimization

Dynamic Parameter Adjustment: GenAI models analyze real-time sensor data to optimize variables such as temperature, pressure, and flow rates. This allows for continuous fine-tuning of chemical processes, improving efficiency and product quality[11][12].

Use Case: In distillation columns, GenAI can adjust parameters like reflux ratio and feed temperature to enhance product purity without increasing energy consumption[11].

Process Flow Design and Simulation

Flowsheet Autocompletion: GenAI assists in designing process flows by autocompleting flowsheets and generating Process and Instrumentation Diagrams (P&IDs). This accelerates the workflow for engineers and reduces errors[17].

Simulation Integration: By interacting with simulation environments, GenAI generates models, runs simulations, and uses the results to optimize processes. This hybrid approach combines mechanistic knowledge with data-driven insights[15].

Predictive Maintenance

Proactive System Monitoring: GenAI predicts equipment failures by analyzing historical data, enabling timely maintenance. This minimizes downtime and ensures safety in high-stakes chemical environments[12][18].

Energy Efficiency and Sustainability

Reducing Energy Consumption: By optimizing process conditions, GenAI significantly reduces energy usage. For example, it can dynamically adjust heating zones in glass production to maintain uniform viscosity while reducing defects[11][14].

Sustainable Practices: GenAI enables process designs that minimize waste generation and optimize resource utilization, contributing to greener chemical engineering practices[13][16].

Autonomous Experimental Platforms

Self-Driven Laboratories: Platforms like GPT-Lab leverage GenAI to autonomously plan and execute experiments. These systems mine literature for experimental parameters, validate outcomes through high-throughput synthesis, and iterate designs for optimal results[13].

Multiscale Optimization

Macro-Level Integration: At the enterprise scale, GenAI optimizes supply chains and plant-wide operations by integrating diverse data types (e.g., textual, visual, experimental) into decision-making frameworks[13].

Quantum-Level Insights: At the molecular level, GenAI enhances understanding of fundamental phenomena, aiding in the design of efficient chemical processes from the ground up[15].

Key Tools and Frameworks

Aspen Plus and ChemCAD:** These tools integrate AI for process simulation and optimization.

Valispace: Offers AI-assisted systems engineering tools for multidisciplinary projects in chemical engineering[14].

Custom Platforms: Companies like AIZOTH provide tailored AI solutions that combine multiple techniques with frameworks like PDCA (Plan-Do-Check-Act) to refine processes with minimal data input[16].

Generative AI is reshaping chemical engineering by enabling real-time optimization, automating complex workflows, enhancing sustainability, and fostering innovation through autonomous experimentation. Its ability to integrate data-driven insights with simulation environments positions it as a transformative tool for achieving efficiency and innovation in chemical processes.

Case Studies of AI in Chemical Process Optimization

There are several successful case studies of AI optimizing chemical processes in the industry. These examples demonstrate how AI technologies have been applied to improve efficiency, reduce costs, enhance product quality, and accelerate innovation in chemical production.

Mesalazine Synthesis Optimization

AI was used to optimize the multistep synthesis of mesalazine, a pharmaceutical compound. Techniques like partial least squares (PLS) regression and artificial neural networks (ANNs) were employed to monitor and control reaction conditions in real-time. This approach improved the precision of intermediate and impurity quantification, enabling better reaction control and higher yields. Additionally, simulated training spectra accelerated the training process for AI models, making advanced data processing more accessible and cost-effective[19].

Catalyst Design at Johnson Matthey

Johnson Matthey utilized the AlchemiteAI platform to optimize catalyst design and related processes. This led to a 4% increase in yield for a key process and significant cost and energy savings during scale-up. The company also achieved efficiency improvements of 50-80% in experimental programs, showcasing the potential of AI in accelerating R&D while reducing resource consumption[22].

Distillation Column Optimization

AI was applied to optimize distillation columns by adjusting parameters such as reflux ratio, feed temperature, and pressure. This resulted in improved benzene purity by 0.5% without increasing energy consumption. The system continuously adjusted these variables in real-time, ensuring optimal performance and product quality[11].

Ink Formulation at Domino Printing Sciences

Domino Printing Sciences used AI to guide testing and optimize ink formulations. The application of machine learning reduced time-to-market, identified new candidate formulations, and enabled reformulation in response to market or regulatory changes. This approach significantly streamlined the development process[22].

Predictive Maintenance and Factory Throughput

A major chemical company leveraged AI for predictive maintenance, which improved factory throughput by nearly 10%. By analyzing historical data, AI models predicted equipment failures before they occurred, reducing downtime and enhancing operational efficiency[23].

Glass Production Quality Control

AI was employed to control temperature profiles along forehearths used in glass production. Machine learning models dynamically adjusted heating zones to maintain uniform viscosity in molten glass, reducing defects such as bubbles by 3% and ensuring higher-quality products[11].

Reaction Condition Optimization

AI-driven models have been used to identify optimal reaction conditions (e.g., temperature and pressure) for various chemical processes. These models reduced reaction times and improved yields while minimizing resource use[20].

Lubricant Property Prediction

In a study on lubricants, sparse experimental data was combined with molecular dynamics simulations using Alchemitesoftware. AI exploited property-property correlations to predict physical properties of known and new alkanes, aiding in material discovery with fewer experiments[22].

Key Benefits of AI in Chemical Industry Applications

- **Efficiency Gains:** Real-time monitoring and dynamic adjustments improve process efficiency.
- **Cost Reduction:** Optimized resource use reduces waste and energy consumption.
- **Innovation Acceleration:** Faster discovery of new materials or formulations.
- **Improved Quality:** Enhanced product consistency through precise control.
- **Sustainability:** Reduced environmental impact through waste minimization.

These case studies highlight how AI is transforming the chemical industry by addressing complex challenges with innovative solutions across research, manufacturing, and quality control processes.

AI and Precision Solution Chemistry

AI plays a transformative role in the customization of chemical products by enabling precision, efficiency, and innovation across various stages of product development and manufacturing.

Customization of Chemical Products

Tailored Formulations

AI analyzes customer data, market trends, and application-specific requirements to design chemical formulations that meet precise specifications. This enables companies to create products optimized for unique customer needs, enhancing satisfaction and market differentiation[25][30].

Predictive Modeling and Simulation

AI-powered predictive models simulate chemical reactions and predict outcomes, allowing researchers to test various combinations virtually before physical trials. This reduces trial-and-error experimentation,

accelerates product development, and ensures formulations are fine-tuned for specific applications[24][28].

Localized Production Customization
By integrating AI into small-scale reactors and production units, manufacturers can adapt processes to meet localized demand while minimizing waste and environmental impact. This ensures that products are customized for regional or niche markets efficiently[27].

Optimization of Material Properties
Machine learning algorithms analyze the relationship between material structures and their properties, suggesting modifications to enhance performance. This is particularly useful in industries like pharmaceuticals, coatings, and advanced materials where specific characteristics are critical[24][26].

Efficiency and Speed in Customization

Accelerated Research and Development (R&D):
AI shortens R&D timelines by identifying optimal chemical combinations faster than traditional methods. For instance, tools like IBM's RXN for Chemistry can forecast reaction outcomes, reducing time-to-market by up to 30%[25][23].

Automated Synthesis:
Robotic platforms guided by AI automate the synthesis of chemicals, ensuring consistent quality while enabling rapid prototyping of customized products[28].

Real-Time Process Adjustments:
AI systems dynamically adjust production parameters based on real-time data, ensuring that customized products meet exact specifications without delays or errors[27][31].

Sustainability and Cost-Effectiveness

Resource Optimization:
AI minimizes raw material usage by optimizing reaction pathways and reducing waste during production. This not only lowers costs but also aligns with sustainability goals[25][31].

Eco-Friendly Product Design:
AI-driven simulations help discover environmentally friendly materials and processes, reducing reliance on harmful substances while maintaining the desired product performance[27][23].

Reduction in Waste:
Predictive analytics ensure that manufacturing processes are efficient, minimizing byproducts and waste generation during the customization process[25][31].

Market Responsiveness

Trend Prediction:
AI leverages historical data to forecast market trends and customer preferences, enabling companies to anticipate demand for specific chemical properties or applications[26][30].

Enhanced Product-Market Fit:
By analyzing large datasets, AI ensures that customized products align closely with market needs, improving commercial success rates[23][30].

AI's integration into the chemical industry has revolutionized how companies approach product customization. By combining advanced analytics with machine learning models, businesses can deliver innovative solutions tailored to specific customer requirements while achieving efficiency, sustainability, and cost-effectiveness. This positions AI as a cornerstone for competitive advantage in the chemical sector.

GenAI in Continuous Industrial Process Engineering

The application of Artificial Intelligence (AI) in continuous industrial processes—such as cement production, steelmaking, sugar refining, and bioreactor/fermentation systems—has led to remarkable advancements. These technologies are driving operational efficiency, sustainability, and innovation across industries. We will review:

1. How GenAI is Revolutionizing Continuous Industrial Processes
2. How Generative AI Optimizes Bioreactors and Fermentation Systems

Revolutionizing Continuous Industrial Processes with AI

Cement Industry: Driving Efficiency and Sustainability
AI has become a cornerstone in modernizing the cement industry, addressing challenges like energy consumption, emissions reduction, and process inefficiencies.

Digital Twins for Optimization:
Companies like Basetwo and Ripik.AI use digital twins to simulate cement production processes. These tools enable real-time adjustments to parameters such as kiln temperature and fuel usage, leading to energy savings of up to 20% and significant reductions in CO_2 emissions[32][36].

Predictive Maintenance:
AI systems monitor rotary kilns and other critical equipment using thermal imaging and sensor data. For example, Vision AI detects hotspots in kiln linings early, allowing timely maintenance and preventing costly downtimes. Ripik.AI's solutions enhance rotary kiln performance by detecting anomalies like overheating or refractory damage[40][42].

Sustainable Cement Formulations:
AI supports the development of greener cement by simulating blends with alternative materials like fly ash or slag. Carbon Re's Delta Zero platform has been instrumental in reducing CO_2 emissions by over 50 kilotonnes annually per plant[43].

These advancements are enabling the cement industry to transition toward smarter, more sustainable operations while meeting global decarbonization goals.

Steel and Foundry Industries: Enhancing Process Control
The steelmaking process involves complex operations such as melting, casting, and refining. AI is addressing variability in these processes to improve efficiency and product quality.

Real-Time Process Optimization:
AI models analyze incoming material properties to dynamically adjust process parameters. Tools like PHOSA integrate with SCADA systems to optimize operations in real time, improving yields while reducing waste[48].

Anomaly Detection:
Machine learning algorithms identify deviations in casting or rolling processes. For example, Generative Adversarial Networks (GANs) generate realistic process data for training diagnostic systems, ensuring consistent product quality[33][44].

Predictive Maintenance:
AI-driven models predict equipment servicing needs for blast furnaces and mills, reducing costs and unexpected disruptions[37].

Energy Efficiency:
AI-powered tools optimize energy consumption during smelting and refining stages. This has significantly reduced costs and environmental impact across steel plants.

Sugar Refining: Real-Time Optimization
Sugar refining involves dynamic processes that require precise control for optimal performance. AI has introduced transformative solutions for this industry:

Digital Twins for Sugar Mills:
Unified data models integrate lab results, process historians, and instrumentation diagrams to create a comprehensive digital representation of each mill. These models provide actionable insights for operators to maximize yield while minimizing energy costs[34][45].

Enhanced Operator Performance:
AI-enhanced decision-making tools enable operators to perform at expert levels by analyzing historical trends alongside real-time conditions[34].

Precision Agriculture:
Platforms like GAMAYA's CanaSight use AI to predict sugarcane yield and sugar content with high accuracy. This improves resource allocation during harvesting, enhancing profitability while reducing waste[38].

These innovations are helping sugar producers achieve higher efficiency and profitability while addressing sustainability concerns.

Bioreactors and Fermentation Systems

Bioreactor Optimization through AI:
Bioreactors play a critical role in industries such as pharmaceuticals, bioengineering, food production, and more. AI is revolutionizing fermentation processes by optimizing environmental conditions for microorganisms:

Continuous Bioprocessing: Companies like Pow.bio leverage AI to split fermentation into growth and production phases, eliminating contamination risks while maximizing productivity[35].

Strain Optimization: Iterative machine learning platforms like TeselaGen accelerate the development of genetically engineered microorganisms for fermentation. This approach enables higher yields in pharmaceuticals or bio-based products[46][49].

Precision Fermentation: Researchers at Imperial College London use AI to genetically engineer yeast strains for food production or pharmaceutical applications. The iterative process tests thousands of genome edits per run, significantly improving output quality[39].

These innovations are enabling bioreactors to achieve higher productivity while maintaining consistency across large-scale operations.

Addressing Challenges in Fermentation:
AI also addresses common challenges faced by fermentation systems:

Data Integration: Advanced platforms unify data from sensors and lab records to create actionable insights for process control.

Problem Prevention: AI helps prevent issues such as shear stress or foam formation by continuously monitoring critical parameters like oxygen levels and substrate feed rates[41][49].

By leveraging these capabilities, bioreactor systems are achieving unprecedented levels of efficiency while reducing operational risks. AI has emerged as a transformative force across non-manufacturing industrial continuous processes. From optimizing cement kilns to streamlining sugar refining operations and enhancing bioreactor performance, these technologies are setting new benchmarks for efficiency, sustainability, and innovation. With ongoing advancements in machine learning and data analytics, the potential for further breakthroughs remains vast—paving the way for smarter industries that align with global sustainability goals.

GenAI for Fault Detection and Diagnosis of Industrial Process Systems

Fault detection is a critical aspect of maintaining safe and efficient operations in continuous process systems. Generative AI (GenAI) is transforming fault detection and diagnosis (FDD) in industrial process systems by leveraging advanced machine learning techniques to identify, predict, and explain faults in real time. This technology addresses challenges such as limited fault data, high-dimensional datasets, and evolving system behaviors, making it a game-changer for industries reliant on continuous processes.

We will review:

1. Anomaly detection
2. Predictive maintenance
3. Enhanced Fault Diagnosis with Explainability
4. Industry Use Cases

Anomaly Detection

Generative Adversarial Networks (GANs) are used to detect anomalies in complex systems like rotating machinery or power grids. These models can generate synthetic data to train diagnostic algorithms, improving accuracy even with limited real-world fault data[50][51].

CycleGAN for Rotating Machinery:
CycleGANs have been used to generate synthetic vibrational data for rotating machinery, effectively augmenting datasets with realistic fault scenarios. These synthetic datasets improve the training of diagnostic algorithms, achieving high accuracy even under varying operating conditions[52].

ACGAN-SDAE for Small Sample Sizes:
Auxiliary Classifier GAN (ACGAN) combined with Stacked Denoising Autoencoders (SDAE) generates high-quality labeled samples, enhancing fault diagnosis accuracy in scenarios with limited data availability. This approach has shown superior performance in diagnosing rolling bearing faults under noisy conditions[50].

GenAI-powered models like ACGAN-SDAE have demonstrated robust performance across varying load conditions and noisy environments, making them highly adaptable for diverse industrial applications[50][55].

Predictive Maintenance

By simulating failure scenarios and generating synthetic datasets, GenAI enables predictive maintenance strategies that preempt equipment failures. This reduces downtime and operational costs while enhancing system reliability[56].

Enhanced Fault Diagnosis with Explainability

Explainable AI (XAI) techniques integrated with GenAI improve the interpretability of fault diagnosis models:

Interpretable Models:
Tools like SHAP (Shapley Additive Explanations) are used to explain model predictions, making them more transparent for stakeholders. For example,

SHAP was applied to diagnose faults in HVAC systems, improving trust in AI-driven decisions[57].

Natural Language Explanations:
Large Language Models (LLMs) combined with diagnostic tools provide human-readable explanations of faults. For instance, Argonne National Laboratory's system explains sensor anomalies in nuclear power plants, helping operators understand complex issues[58].

Industry Use Cases

Chemical Plants:
LSTM-based networks achieve over 95% accuracy in diagnosing faults in the Tennessee Eastman Process (TEP), a benchmark chemical plant simulation[59].

Chillers and Rotating Machinery:
GAN-based methods enhance diagnostics by generating realistic failure patterns for chillers and bearings under varying load conditions[60].

Generative AI is revolutionizing fault detection and diagnosis by addressing long-standing challenges in process systems engineering. From creating synthetic datasets to enabling real-time anomaly detection and predictive maintenance, GenAI enhances operational efficiency, safety, and reliability across industries. As adoption grows, its ability to integrate explainable models and adaptive solutions will further solidify its role as a cornerstone technology in industrial automation.

Challenges and Limitations

Generative AI (GenAI) applications in Chemical Engineering and Industrial Process System Engineering face several challenges and limitations across four key domains: Data Quality and Availability, Interpretability, Computational Costs, Ethical Considerations and examples from the industry.

Data Quality and Availability

High-quality data is foundational for training effective GenAI models. However, in CE and IPSE, acquiring such data poses significant challenges:

Data Scarcity:
Industrial processes often involve proprietary or sensitive data, limiting access to comprehensive datasets[61][68].

Data Complexity:
Industrial datasets are highly heterogeneous, incorporating time-series data, sensor outputs, and operational logs that require extensive preprocessing[70].

Synthetic Data:
GenAI can generate synthetic datasets to augment training data, but this introduces risks of inaccuracies or bias if the synthetic data does not adequately reflect real-world conditions[61][73].

Solution paths:

Leveraging GenAI for data cleaning, anomaly detection, and filling missing values can improve data quality[70].

Collaboration between industries and academia to establish shared repositories of anonymized industrial data can help address availability issues[68].

Interpretability

The "black-box" nature of GenAI models limits their adoption in safety-critical fields like Industrial process engineering:

Opaque Decision-Making:
Understanding how a model generates outputs is critical in industrial contexts where errors could lead to catastrophic failures[71][77].

Trust Issues:
Lack of interpretability hinders trust among engineers and stakeholders[64][71].

Innovation paths:

Techniques like attention mechanisms, saliency maps, and Layer-wise Relevance Propagation (LRP) are being developed to enhance interpretability[75].

Startups like Goodfire AI are pioneering mechanistic interpretability methods to provide granular insights into model behavior[64].

Computational Costs

GenAI models are computationally intensive, making their deployment challenging:

Energy Consumption:
Training large models requires significant computational power, driving up energy costs and environmental impact[63][63].

Infrastructure Costs:
Maintaining the hardware and cloud infrastructure for running GenAI models is expensive, particularly for small enterprises[78][79].

Mitigation Strategies:

Hybrid cloud architectures and optimized coding practices can reduce computational costs by up to 50%[63][63].

Smaller, task-specific models trained on high-quality data can be more efficient than general-purpose large models[63].

Ethical Considerations

The ethical implications of GenAI applications in this engineering domain include:

Bias and Fairness:
Training data may embed biases that lead to unfair or unsafe outcomes in industrial settings[76][74].

Sustainability:
The high energy demands of GenAI raise concerns about environmental sustainability[65][74].

Misuse Risks:
Generative capabilities could be exploited for malicious purposes, such as creating unsafe chemical formulations or spreading misinformation about industrial processes[76].

Robust ethical frameworks emphasizing transparency, accountability, and sustainability are essential for responsible deployment[62][65]. Regular audits of AI systems can help identify biases or unintended consequences early on[80].

Examples from Industry

Startups like Unit8 are developing scalable GenAI solutions tailored for industrial applications while prioritizing transparency and cost-efficiency[69].

Hyperscalers such as Google Cloud leverage GenAI to optimize resource allocation in industrial processes, achieving faster training times and reduced costs[66].

S&P 500 Companies like Chevron Phillips Chemical use GenAI for IoT analytics in industrial settings, showcasing its potential despite high initial costs[67].

In conclusion, while GenAI offers transformative potential in Chemical Engineering and Industrial process engineering, addressing these challenges through collaborative efforts, technological innovation, and ethical governance is critical for its successful integration into these fields.

Future Directions for Gen AI in Chemical and Process Engineering

Four directions shape the future of GenAI in CE and PSE: Hybrid Models, Self-Supervised Learning (SSL), Decarbonization and Green Chemistry, and Model Interpretability.

Hybrid Models

Hybrid models improve the reliability of simulations in chemical engineering by combining the strengths of physics-based (mechanistic) models and data-driven machine learning (ML) approaches. This synergy addresses the limitations of each method individually, leading to more accurate, efficient, and robust simulations. Here's how hybrid models enhance reliability:

Improved Prediction Accuracy:
Hybrid models leverage the precision of first-principles models while incorporating the flexibility of ML to account for unknown or complex phenomena that are difficult to model using physics alone. For instance, direct hybrid models combine outputs from mechanistic and ML models in series or parallel configurations. This approach has been shown to improve predictions for chemical processes like polymerization and fermentation by correcting biases in standalone models[81][84]. In biopharmaceutical applications, hybrid models have demonstrated up to 40% better accuracy compared to traditional mechanistic models, particularly in scenarios with limited data[84].

Enhanced Extrapolation Capabilities:
Physics-based models are adept at extrapolating beyond experimental conditions due to their grounding in scientific principles. When combined with ML, hybrid models can extend this capability by refining predictions with data-driven corrections. For example, hybrid residual models quantify and adjust for prediction errors in mechanistic simulations, enabling more reliable extrapolations in batch processes and reactor systems[81][84].

Reduced Data Requirements:
Purely data-driven models often require large datasets to generalize effectively. Hybrid models mitigate this need by embedding physical laws into the ML framework, reducing dependence on extensive data while maintaining accuracy[83][84]. This is particularly valuable in chemical engineering applications where data collection can be expensive or time-consuming.

Robust Process Optimization:
Hybrid models facilitate process optimization by integrating mechanistic insights with real-time data analysis. This dual approach enables better control over dynamic systems, such as reactors or distillation columns, and improves decision-making under uncertainty[81][82]. For instance, hybrid modeling has been applied to optimize polymer manufacturing processes by combining kinetic equations with neural networks for precise parameter estimation[81].

Quantification of Uncertainty:
Hybrid approaches can quantify uncertainties in both mechanistic and ML components, providing a more comprehensive understanding of model reliability. This is critical for high-stakes applications like safety-critical process control or scale-up operations[81][86].

Applicability Across Scales:
Hybrid models are versatile and applicable across various scales of chemical engineering, from molecular-level simulations (e.g., catalyst design) to plant-wide operations (e.g., digital twins for smart operation) [82][85].

While hybrid models offer significant advantages, challenges such as computational complexity, integration of diverse data sources, and interpretability remain. Future research is focused on:

- Developing standardized frameworks for hybrid modeling.
- Enhancing interpretability through explainable AI techniques.

- Expanding applications in decarbonization and sustainable process design.

By addressing these challenges, hybrid modeling will continue to play a pivotal role in advancing chemical engineering simulations and process systems engineering.

Self-Supervised Learning

Self-supervised learning (SSL) helps overcome data scarcity in chemical engineering (CE) and process systems engineering (PSE) by leveraging large quantities of unlabeled data to pre-train models, which can then be fine-tuned for specific tasks using limited labeled data. This approach addresses the challenges posed by the high cost and time required to generate labeled datasets.

Leveraging Unlabeled Data:
SSL uses unlabeled data, which is typically abundant in industrial processes (e.g., sensor readings, operational logs), to learn meaningful representations of the underlying system. For example, SSL has been applied to fatigue damage prognostics by pre-training models on synthetic strain gauge data, enabling accurate predictions of remaining useful life (RUL) with minimal labeled examples[87][90]. In materials science, SSL frameworks like Crystal Twins have been used to predict material properties from large databases of unlabeled structural and compositional data, bypassing the need for expensive simulations or experiments[89].

Pretext Tasks for Representation Learning:
SSL employs "pretext tasks"—auxiliary learning tasks designed to extract useful features from unlabeled data. These tasks might include predicting missing sensor values, identifying anomalies, or reconstructing process states. For instance, in chemical process optimization, SSL could train a model to predict missing operational parameters or reconstruct reaction pathways, creating robust feature representations that can be fine-tuned for downstream tasks like fault detection or yield prediction.

Reducing Labeling Costs:
Traditional supervised learning requires extensive labeled datasets, which are costly and labor-intensive to generate in CE and PSE due to the need for domain expertise (e.g., annotating experimental results or process failures). SSL minimizes this dependency by generating pseudo-labels from unlabeled data[88][89]. This reduction in labeling requirements makes it feasible to apply advanced machine learning techniques even in resource-constrained environments.

Enhancing Model Generalization:
By pre-training on diverse and unlabeled datasets, SSL models learn generalizable features that can adapt to various tasks with minimal additional training. This is particularly valuable in CE and PSE, where processes often involve complex dynamics and variability across systems. For example, pre-trained SSL models can be fine-tuned for specific applications like optimizing distillation columns or predicting catalyst performance with limited task-specific data.

Applications in CE and PSE:

Process Monitoring: SSL can analyze sensor data to detect anomalies or predict equipment failures without requiring extensive failure labels.

Materials Discovery: In materials science, SSL has been used to predict properties of crystalline materials using large-scale databases of atomic structures[89].

Energy Systems: SSL frameworks have supported sustainable energy innovations by enabling efficient modeling of CO2 capture materials with minimal labeled experimental data[91].

Computational Efficiency:
Pre-trained SSL models often require less computational effort during fine-tuning compared to training a supervised model from scratch. This efficiency is critical in CE and PSE applications where computational resources may be constrained[87].

While SSL offers significant advantages for addressing data scarcity, challenges such as model interpretability, computational complexity during pre-training, and ensuring high-quality unlabeled datasets remain. Future research should focus on:

Developing domain-specific pretext tasks tailored to CE and PSE.

Integrating SSL with hybrid modeling approaches that combine physics-based simulations with machine learning.

Expanding the use of SSL for decarbonization efforts and green chemistry innovations.

By addressing these challenges, SSL can further enable robust AI-driven solutions even in data-scarce scenarios.

Decarbonization and Green Chemistry

GenAI can play a pivotal role in advancing sustainable practices in CE and PSE.

Applications in Decarbonization:

Startups like Entalpic use GenAI to design catalysts for energy-efficient chemical processes, contributing to decarbonization efforts in sectors like fertilizers, energy storage, and pollution control[92].

Machine learning models are being developed to optimize multiscale systems for sustainable pathways, such as reducing greenhouse gas emissions in petrochemical processes[93][94].

GenAI models can simulate chemical reactions and optimize process parameters to improve energy efficiency and reduce emissions. By identifying optimal configurations, it minimizes waste and energy use during production[95][96].

Applications in Green Chemistry:

Generative AI plays a transformative role in advancing decarbonization and green chemistry initiatives by enabling more efficient, sustainable, and innovative practices across process industries. Its ability to analyze, simulate, and optimize processes at various scales makes it a powerful tool for reducing environmental impact while driving technological progress. AI-driven tools are being used to design eco-friendly solvents, catalysts, and reaction pathways that align with green chemistry principles[97].

Designing Sustainable Chemicals and Materials:

GenAI accelerates the discovery of environmentally friendly chemicals by predicting molecular structures with desired properties. It proposes alternatives that reduce toxicity, improve biodegradability, and lower carbon footprints[98][27].

For instance, generative models have been used to design catalysts that enhance reaction rates while minimizing environmental impact[100][27].

Lifecycle Analysis:

Predictive models powered by GenAI evaluate the environmental impact of chemical processes throughout their lifecycle—from raw material extraction to disposal—helping industries adopt greener practices[100].

Optimizing Reaction Pathways:

GenAI simulates reaction conditions to identify the most sustainable methods for chemical production. This reduces reliance on hazardous solvents and reagents while maximizing yield[96][100].

A notable example is the use of AI-driven models to optimize catalytic processes for converting biomass into biofuels, enhancing yield while reducing waste[100].

Autonomous Experimental Platforms:

GenAI integrates with robotic systems to automate chemical synthesis tasks. By planning and executing synthesis routes with minimal human intervention, it accelerates the development of sustainable materials while reducing resource consumption[13][27].

Future Opportunities:

GenAI could help industries transition to circular economies by optimizing waste recycling processes or designing biodegradable materials. In the chemical industry, it supports scaling lab-scale processes to industrial levels while maintaining sustainability goals[103]. It can also assist policymakers by modeling the long-term environmental impacts of various industrial strategies.

Model Interpretability

Explainable AI (XAI) is a transformative tool for CE and PSE, addressing the "black-box" nature of traditional AI models while enhancing trust among stakeholders. By fostering transparency, reliability, and collaboration, XAI can drive the adoption of advanced AI-driven workflows in these critical industries.

Explainable AI (XAI) Techniques:

XAI techniques are essential for making AI models more transparent and interpretable, particularly in critical industrial applications such as fault detection, optimization, and process control. XAI is being used to optimize chemical processes by integrating machine learning with optimization algorithms[104].

For instance, hybrid models combining neural networks with genetic algorithms have reduced computational time significantly while achieving high accuracy in process optimization tasks like distillation. In product design, XAI enhances the prediction of physical properties like critical temperature or pressure using deep learning models. This helps in designing safer and more sustainable materials. Advanced XAI techniques

are being applied to simulate complex chemical processes, enabling engineers to understand model predictions better, especially for nonlinear systems with multiple variables. Companies like NVIDIA and Alphabet are leveraging AI to optimize industrial processes while maintaining transparency through explainable frameworks[107].

The complexity of chemical processes often requires domain-specific adaptations of general XAI techniques[105][106].

Startups and Projects Utilizing Explainable AI (XAI) in Chemical Engineering (CE) and Process Systems Engineering (PSE)

Discusses innovative startups and projects that leverage XAI for chemical and process systems engineering.

AiChemist by CHELONIA

AiChemist leverages XAI for molecular design and optimization in chemical R&D :

Optimizing biological activity and physico-chemical properties of compounds while minimizing toxicity.

Enhancing environmental chemistry by identifying harmful compounds to ensure safer and sustainable practices.

Bridging the explainability gap by translating AI insights into actionable information for chemists and regulatory bodies.

The project integrates advanced machine learning methods with XAI to make AI outputs interpretable, aiding both researchers and regulatory agencies.

CHAI Project by imec

Developing hybrid, explainable AI algorithms to optimize chemical processes :

Supporting less experienced process operators by providing actionable insights.

Assisting experienced engineers with contextualized dashboards that explain AI-driven predictions and suggestions.

Scaling expertise in chemical production through hybrid AI that combines data-driven learning with expert knowledge.

The project emphasizes dynamic visualization and feedback loops to continuously improve decision-making in process optimization[109].

Basetwo

A low-code platform for process engineers to build digital twins of manufacturing processes :

Predictive maintenance, fault detection, and process optimization using digital twins.

Providing interpretable insights into process behavior through XAI tools.

Basetwo's platform simplifies the integration of XAI into PSE workflows, enabling engineers to simulate, monitor, and optimize processes effectively[110].

XenonStack

Explainable AI in manufacturing, including applications relevant to PSE :

Predictive maintenance using sensor data to forecast equipment failures before they occur.

Real-time quality monitoring and anomaly detection in industrial processes.

XenonStack uses XAI to provide transparent justifications for AI predictions, fostering trust among users while improving operational efficiency[108].

SchNet4AIM

SchNet4AIM is a SchNet-based architecture designed for Explainable Chemical Artificial Intelligence (XCAI):

Predicting molecular properties such as atomic charges, delocalization indices, and interaction energies.

Assisting in material design and reaction engineering by providing interpretable chemical insights.

Unlike traditional black-box models, SchNet4AIM inherently integrates explainability by linking predictions to physically rigorous atomic or pairwise terms. This approach has enabled researchers to interpret complex chemical simulations, accelerating material discovery and improving the reliability of AI-driven insights[111].

Autonomous Simulation Agents

These agents are used for real-time data analysis and predictive modeling in polymer science:

- Suggesting new polymer structures with specific properties like flexibility or strength.
- Optimizing synthesis processes dynamically based on ongoing results.

By integrating XAI, these agents provide clear explanations of their predictions, helping researchers understand the rationale behind suggested designs or process adjustments. The use of XAI has improved material design workflows, reducing development time and enhancing safety in polymer applications[112].

These startups and projects demonstrate how XAI is being applied to enhance transparency, trust, and efficiency in CE and PSE workflows. They focus on bridging the gap between complex AI models and human interpretability, ensuring safer and more effective industrial operations.

Conclusion

Generative AI represents a paradigm shift in chemical engineering and process systems engineering. By automating complex tasks, optimizing processes, and driving innovation, it has the potential to redefine industry standards. However, challenges such as data limitations, interpretability, and computational costs must be addressed to fully harness its capabilities. As research advances, GenAI is poised to play a central role in the evolution of CE&PSE, particularly in the context of sustainability and decarbonization.

Citations for Section on Chemical Engineering

[1] ChemRxiv - Article Details

[2] Example Article 2
[3] Catalyst Design Narrative Study
[4] ChemRxiv - Article Details
[5] SiliconANGLE - CuspAI Raises $30M
[6] EOS Globe - Generative AI in BPM
[7] Yahoo Finance - Pharma Stocks Harnessing AI
[8] Santander Private Banking - The Potential of AI
[9] S&P Global Press Release
[10] S&P Global - AI in Pharmaceuticals
[11] Intelecy - AI in Chemical Industry Process Optimization
[12] Allied Academies - Applications of AI in Chemical Engineering
[13] Frontiers in Chemical Engineering - Article
[14] Valispace - AI in Chemical Engineering
[15] PI Research - Generative AI Project
[16] Aizoth - Smart Solutions in Chemical Engineering
[17] YouTube - IPPL6_j8MvE
[18] JCHR Article Download
[19] Thiem Connect - PDF Article
[20] Restack - AI in Chemical Engineering
[22] Intellegens - Chemicals Applications
[23] Unit8 - AI in Discovering New Chemicals
[24] Intimedia - AI in Chemical R&D
[25] DigitalDefynd - AI in Chemical Industry
[26] CEFIC - Sustainable Innovation in Chemical Industry
[27] Business Chemistry - AI in the Chemical Industry
[28] MDPI - Article
[30] AISuperior - AI Companies for Chemicals
[31] Spyro-Soft - EU AI Act in Chemical Industry
[32] CemAI - December 2022 LR
[33] AISTech - Details
[34] C3.ai - Optimize Sugar Production with AI
[35] SynBioBeta - AI and Automation in Bioreactors
[36] Wizata - Cement Industry
[37] Gunung Capital - AI in Steel Industry
[38] Gamaya - Patent for AI-driven Sugarcane Yield Prediction
[39] ISAAA - Crop Biotech Update
[40] Ripik.AI - AI in Cement Industry
[41] Isomerase - Fermentation Optimization
[42] DigitalDefynd - AI in Cement Industry
[43] Alliance LCCC - Carbon-Re AI Powered Decarbonisation
[44] IMIS AIST - Abstracts PDF
[45] SugarIndustry.info - Paper 32754
[46] Teselagen - Fermentation Optimization
[48] BCG - Value of AI in Steel Industry
[49] Isomerase UK - Fermentation Optimisation Article
[50] PMC - Article PMC7924884
[51] TechScience - IASC Article
[52] MDPI - Article 12458
[55] TechScience - CMES Article
[56] LinkedIn Pulse - Generative AI in Predictive Maintenance
[57] GitHub - Fault Detection and Diagnosis with XA
[58] TechXplore - Smart Diagnostics with GenAI
[59] arXiv - Article 2401.10266v1

[60] ResearchGate - GAN for Fault Detection in Chillers
[61] XenonStack - Generative AI for Data Quality
[62] IOS Press eBook - FAIA231182
[63] IBM Think - AI Economics Compute Cost
[64] SiliconANGLE - Goodfire AI Raises $7M
[65] Devoteam - Ethics of Generative AI
[66] Itsoli.ai - Unlocking Hyperscaler Speed
[67] Databricks - Data AI Use Cases
[68] Data Science Blog - Generative AI and Data Quality
[69] Unit8 - GenAI Hyperscaler
[70] Express Analytics - Generative AI Data Quality Management
[71] XCubeLabs - Explainability in GenAI Systems
[73] BigDataWire - Data Quality and GenAI
[74] TechTarget - Generative AI Ethics
[75] HTC Inc - Harnessing GenAI
[76] Ingosa.ai - Ethical Challenges in GenAI
[77] LarkSuite - Interpretability in AI
[78] LinkedIn - Making GenAI Affordable
[79] ORF Online - AI Economies of Hyper Scale
[80] Lumen Blog - Ethical AI Governance
[81] arXiv - Paper 2112.01475
[82] Frontiers in Chemical Engineering - Article
[83] PI Research - Hybird Modeling
[84] Frontiers in Chemical Engineering - Article 1157889
[85] AspenTech - Hybrid Models
[86] ResearchGate - Hybrid Science-Guided ML
[87] arXiv - Paper 2301.08441
[88] IBM Think - Self-Supervised Learning
[89] Nature - Article s41524-022-00921-5
[90] HAL Science - Document
[91] NETL - 23FECM_19_Yin.pdf
[92] Cathay Capital - Entalpic Raises $8.5M
[93] ACS Publications - Article on Decarbonizing Chemicals
[94] S&P Global - Decarbonizing Chemicals
[95] PharmiWeb - AI in Chemical Market
[96] KPMG - GenAI Enrc Sep26
[97] Tandfonline - Article 2312848
[98] Accenture - Making Generative AI Green
[100] Restack - AI in Green Chemistry
[103] IBM - GenAI Document
[104] Elmoukrie - AI in CE for My Site v2
[105] ResearchGate - Explainable AI in Process Engineering
[106] CEUR WS - Paper26
[107] SyntaxData - Quantifying AI Exposure in S&P 500
[108] XenonStack - Explainable AI in Manufacturing
[109] imec - CHAI Project
[110] Chemical Times - AI Startups in Chemical Industry
[111] Nature - Explainable Chemical AI
[112] Chelonia - Explainable AI for Molecules

General References for Section on Chemical Engineering

[G1] Computational Chemical Engineering Journal, various issues

[G2] AIChE Journal, special issues on AI in Chemical Engineering

[G3] Key insights synthesized from the uploaded documents on GenAI applications in CE&PSE

[G4] F6S - Companies with Explainable AI (XAI) technology

[G5] Reddit - Premier Companies for Chemical Engineering

[G6] MDPI - Article on Explainable AI in Engineering

[G7] ResearchGate - Explainable AI for Engineering Design

[G8] IEEE Xplore - Explainable AI in Systems Engineering

[G9] Harvard Data Science Review - Explainable AI

[G10] LinkedIn - New Directions in Chemical Engineering

[G11] ResearchGate - Process Systems Engineering

[G12] JMIR - Explainable AI in Healthcare

[G13] MDPI - Explainable AI in Chemical Engineering

[G14] IChemE - Chemical Engineering Matters

[G15] EDA - Handbook on Process Systems Engineering

[G16] Reddit - Opening a Chemical Engineering Firm

[G17] Google Cloud - Generative AI Use Cases

[G18] IBM - Case Study on Engineering

[G19] ResearchGate - Reloading Process Systems Engineering

[G20] Bohrium - Explainable AI in Engineering

[G21] Strathprints - Case Study on Explainable AI

[G22] Restack - AI in Chemical Engineering Case Studies

[G23] Restack - Innovative AI Applications in Chemical Engineering

[G24] OMP - Understanding AI through XAI

[G25] MDPI - Explainable AI in Engineering

[G26] Restack - Generative AI Applications in Chemical Engineering

[G27] Restack - AI Tools on GitHub

[G28] Restack - Generative AI in Chemical Processes

[G29] GreyB - AI in the Chemical Industry

[G30] Restack - AI Case Studies in Chemical Engineering

[G31] Postindustria - AI Use Cases in Chemical Industry

[G32] Restack - AI Applications and Predictions in Chemical Engineering

[G33] Markovate - AI Use Cases in Chemical Industry

[G34] Chinimandi - AI in Sugar Industry

[G35] ABB - AI in Cement Industry

[G36] IJAMTES - AI in Mechanical Engineering

[G37] SM Specialties - AI in Metal Casting

[G38] Harvard - Evaluating Generative Models

[G39] Sybershel - AI Investments in S&P 500

[G40] Nasdaq - AI Stocks in S&P 500

[G41] arXiv - Explainable AI Research

[G42] Taylor & Francis - Explainable AI in Engineering

[G43] Aalpha - Cost to Build Generative AI

[G44] Enterprise League - Explainable AI Startups

[G45] Remotely - Ethics in Generative AI
[G46] Itrex - Cost of Generative AI
[G47] Computer Weekly - AI Ethics
[G48] Kalray - Data Management for Generative AI
[G49] TechTarget - Explainable AI Systems
[G50] Encore - Ethics of Generative AI
[G51] Forbes - Best AI Stocks
[G52] S&P Global - AI Governance Challenge
[G53] LinkedIn - AI Activity in S&P 500
[G54] US Funds - AI Gains in S&P 500
[G55] Publicis Sapient - Ethical AI Implementation
[G56] Tech Informed - AI Risks in Fortune 500
[G57] Harvard Law - AI as a Risk Multiplier
[G58] Workplace Privacy Report - AI Risks in SEC Filings
[G59] ACS Publications - AI in Chemical Engineering
[G60] TU Delft - Hybrid Modeling Methodologies
[G61] Nature - AI in Chemical Engineering
[G62] MDPI - AI in Chemical Engineering
[G63] Core - AI in Chemical Engineering
[G64] PMC - AI in Chemical Engineering
[G65] PMC - AI in Chemical Engineering
[G66] News Medical - XAI in Drug Discovery
[G67] Siemens - Whitepaper on XAI