1

Introduction

*Peng Jiang and Jiahua Zhu**

State Key Laboratory of Materials-oriented Chemical Engineering, College of Chemical Engineering, Nanjing Tech University, Nanjing, China

*Corresponding author: jhzhu@njtech.edu.cn

Biomass has emerged as a highly attractive resource of renewable energy, chemicals, and materials due to its global abundance, affordability, and carbon-neutral features [1, 2]. Global biomass generation has been reported to surpass 10 billion tons annually on a dry-weight basis, representing an energy potential equivalent to more than 2 billion tons of standard coal [3]. Accordingly, biomass-derived carbon resources has been widely recognized as a promising alternative to conventional fossil feedstocks, offering significant potential for lowering fossil resource consumption and alleviating CO_2 emissions [4].

To achieve economic parity with conventional fossil-derived routes, a range of technological pathways and development strategies for biomass valorization have been put forward. As an illustrative case, the US Department of Energy designated a list of "Top 12 High-Value Biobased Chemicals" as key priority products for future deployment [5]. Likewise, the International Energy Agency (IEA) published the Technology Roadmap for Delivering Sustainable Bioenergy, which presents strategies to promote the sustainable use of biomass [6]. China possesses abundant biomass resources, yielding an annual output of 3.5 billion tons. The *3060 Zero-Carbon Biomass Potential Blue Book* indicates that these biomass resources are equivalent to about 460 million tons of standard coal and have the potential to reduce CO_2 emissions by about 10%, making a substantial contribution toward China's carbon neutrality objectives [7]. Recently, the fast advancement of electrification technologies has pushed new innovations in biomass conversion technologies. For instance, Tian et al. [8] developed a conceptual framework enabling electrochemically favored conversion pathways for biomass-derived molecules. Dossow et al. [9] carried out a comprehensive evaluation of the biomass-to-X (BtX) electrification roadmap. Their findings suggested that integrating electrification into BtX pathways could substantially advance the decarbonization of both the chemical industry and the transportation sector, while promoting the shift away from fossil-based resources.

In recent years, the utilization of biomass has progressed beyond heat and electricity generation, expanding toward the production of bulk fuels and chemicals, including methane, methanol, ethanol, and syngas. Research efforts have increasingly emphasized the production of high-value-added outputs, including green hydrogen [10, 11], sustainable aviation fuels [12], advanced functional materials [13–15], and platform chemicals [16, 17]. Despite decades of investigation, only a limited number of industrial-scale implementations have successfully satisfied sustainability requirements in terms of technical feasibility, economic viability, and environmental performance. Compared with conventional fossil resources such as coal and petroleum, several critical gaps in biomass utilization remain evident. First, comprehensive and reliable databases describing biomass composition and thermodynamic properties are still insufficient [18, 19]. Second, the establishment of robust first-principles models for biomass transformation remains challenging, especially for highly complex routes such as pyrolysis [20, 21], due to pronounced feedstock variability and the presence of multifaceted reaction mechanisms [22]. Third, system-level optimization continues to present major difficulties, as biomass valorization spans the full life cycle, covering biomass supply chains (BSCs), including harvesting, storage, and transportation [23], downstream conversion steps such as pretreatment, conversion, and product separation [24], and integrated optimization objectives that simultaneously consider carbon utilization, energy efficiency, economic feasibility, and environmental impacts [25]. Collectively, the strong coupling among heterogeneous data layers, modeling approaches, and system-level coordination poses a significant barrier to the practical implementation of biomass engineering technologies.

Data-driven methodologies, with machine learning (ML) as a representative example, have emerged as powerful tools for describing and predicting the behavior of highly complex systems by learning statistical relationships between input variables and corresponding outputs. Accordingly, a broad range of ML techniques has been applied to biomass utilization research, encompassing the estimation of biomass property descriptors [26], the development of process-level models [27, 28], and multi-objective optimization strategies [29]. Among the various biomass characteristics, the higher heating value (HHV) is considered a key parameter for evaluating feedstocks and is commonly employed in techno-economic assessments, computational modeling, and the engineering design of advanced conversion processes [30, 31]. For instance, Xing et al. [32] utilized ML models to estimate HHV based on ultimate and proximate composition data. The study demonstrated that ML-driven predictions $(R^2 > 0.94)$ considerably outperformed traditional empirical correlations $(R^2 < 0.70)$.

In biorefinery systems, process modeling serves as a fundamental tool for evaluating system performance, optimizing operating conditions, and estimating product distributions. Conventional models for biomass gasification are commonly constructed using thermodynamic and kinetic frameworks to describe pyrolysis behavior [33–35]. However, these models often require intensive numerical calculations and tend to exhibit limited predictive robustness when applied to heterogeneous biomass feedstocks. To overcome such challenges, data-driven modeling approaches

have gained increasing attention. For example, Baruah et al. [36] applied an artificial neural network (ANN) to fixed-bed downdraft gasifiers for predicting syngas compositions, including CH_4, CO, CO_2, and H_2, and reported high prediction accuracy $(R^2 = 0.986–0.993)$ with low error levels (RMSE = 0.052–0.092). Nevertheless, although ML-based techniques generally show strong fitting performance, their transferability and generalization are frequently constrained by limited datasets and the intrinsic complexity of biomass conversion systems. To solve the above problems, researchers put forward hybrid modeling solutions integrating mechanistic modeling with data-driven learning. These methods deliver enhanced physical interpretability, superior generalization and more stable predictive performance. Physics-informed neural networks, as a representative application, achieve an effective balance between prediction accuracy, model interpretability and generalization in biomass gasification studies [37]. With the rapid advancement of ML-assisted biomass research, this introductory chapter is structured around three key themes: (i) biomass valorization technologies, (ii) ML methodologies, and (iii) strategies for embedding ML into biomass conversion and utilization pathways.

1.1 Biomass Valorization Technologies

1.1.1 Traditional Conversion Technologies

Biomass valorization pathways are generally divided into two major categories: thermochemical routes and biochemical routes (Figure 1.1). Biomass, including

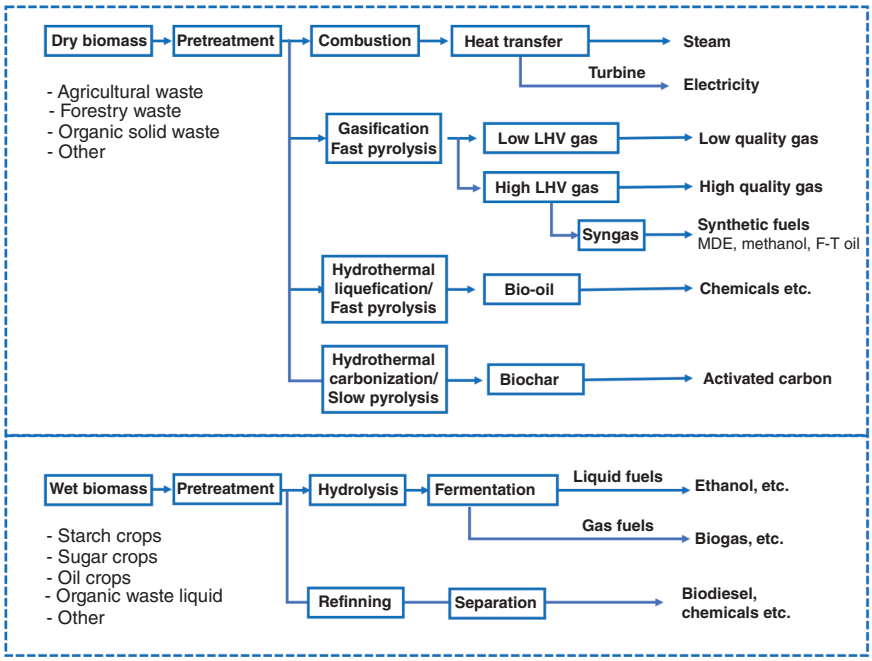

Figure 1.1 Traditional biomass conversion technologies.

agricultural and forestry residues, is primarily converted via thermochemical techniques. These processes enable the transformation of biomass into diverse energy carriers and valuable products, including electricity, biochar, bio-oil, and syngas. Owing to their relatively high conversion efficiency, technological maturity, and scalability, thermochemical approaches can be readily integrated into existing industrial systems, thereby providing practical substitution options for fossil-based sectors, including coal-fired power generation, cement production, and iron and steel manufacturing.

Wet biomass, such as algae and manure, is mainly transformed through biochemical pathways, which encompass aerobic treatments [38], anaerobic digestion [39], enzymatic transformations [40], and fermentation techniques [41]. These approaches allow for the targeted decomposition or transformation of biomass into various fuels and chemicals. In the case of lipid-rich biomass, additional processing steps are applied to produce diesel and ester-based compounds [42]. Despite the relative simplicity of conventional biomass conversion methods in terms of equipment and operational demands, the resulting outputs typically have low economic value, encompassing heat, electricity, bio-oil, and syngas [43].

1.1.2 High-value Conversion Technologies

Recently, researchers have put forward multiple strategies including biomolecular engineering, catalytic conversion and biomass poly-generation. These techniques can be applied to synthesize bio-based functional materials, platform chemicals and high-energy-density fuels like biohydrogen and sustainable aviation fuels (Figure 1.2).

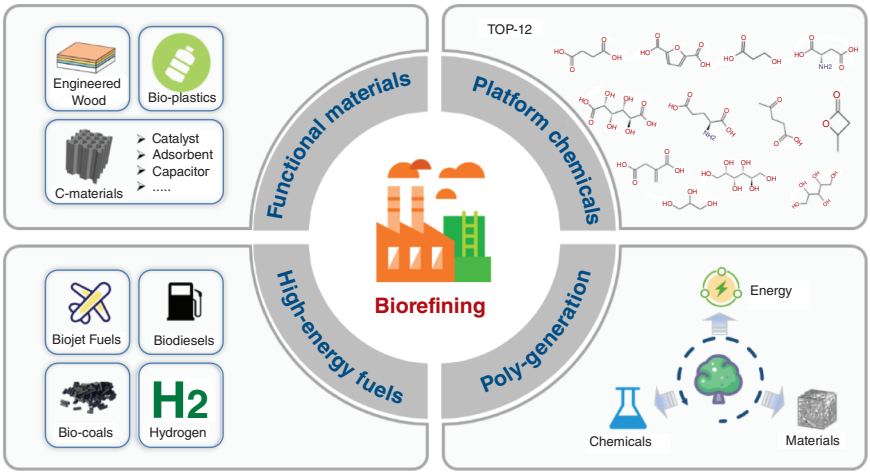

Figure 1.2 Technologies for converting biomass into high-value products. *Source*: [74] / with Permission of Elsevier.

1. **Bio-based functional materials.** The compositional and structural characteristics of biomass have been harnessed to develop advanced functional materials. For example, Hu et al. [45] prepared high-performance eutectogel materials using lignin as the feedstock. This material exhibits excellent mechanical toughness, freeze resistance, underwater adhesion, self-healing ability, and antibacterial activity. A biomass-derived porous carbon material derived from sugarcane bagasse also demonstrates outstanding gas adsorption and separation performance, with a CO_2 working capacity of 12.3 mmol/g and a CO_2/H_2 selectivity as high as 47.8 under the conditions of 273 K and 10 bar [46]. In addition, the self-organization of chitin and cellulose generates fiber networks with excellent structural integrity and adsorption capacity, achieving nearly complete removal of microplastics from diverse water samples [47]. In addition, by optimizing pyrolysis and activation processes, our group has transformed the inherent structures of biomass into hierarchical porous carbon architectures [24, 48–51]. Based on these insights, a variety of functional hierarchical porous carbons were fabricated and used as catalyst supports [52–55] and adsorbents for volatile organic compounds [56].
2. **Bio-based platform chemicals.** The National Energy Technology Laboratory (NETL) highlighted in its bioenergy report [5] 12 molecules considered the most promising substitutes for conventional petroleum-derived building blocks. Among these, 5-hydroxymethylfurfural (HMF) has emerged as a key platform chemical. Our group further demonstrated that microwave-assisted processing markedly improves the fructose-to-HMF transformation, enhancing energy efficiency by 10-fold relative to conventional heating methods [57–59]. Additionally, the photocatalytic conversion of biomass-derived HMF into 5-formyl-2-furancarboxylic acid and 2,5-diformylfuran has been investigated [60]. Recently, research attention has shifted toward producing platform chemicals directly from biomass while preserving its macromolecular structure. However, most studies still rely on model compounds rather than actual biomass constituents, and reaction conditions are often simplified, limiting the practical scalability of these approaches.
3. **Bio-based high-energy fuels.** Sun et al. [61] developed an integrated continuous catalytic strategy that couples selective hydrogenation, aldol condensation, and hydrodeoxygenation, enabling the efficient conversion of isophorone and furfural derived from lignocellulosic biomass into C9–C14 cycloalkanes. The target products were obtained with a yield of 88.32% and an overall carbon yield of 99.6%, while both catalytic efficiency and selectivity significantly outperformed those of conventional processes. In addition, Kumar et al. [62] proposed a multistep catalytic pathway for the conversion of mixed sugars from corn stover, in which bicyclic frameworks were selectively constructed via acid-catalyzed dehydration followed by hydrogenative ring rearrangement, producing jet fuel-range bicycloalkanes with high energy density. When blended with conventional Jet A fuel at a volume fraction of 37%, the resulting fuel exhibited markedly improved volumetric heat value and freezing point characteristics, demonstrating its suitability for direct application in aviation propulsion systems.

Fan and Zeng [63, 64] demonstrated the production of high-purity hydrogen from biomass while simultaneously capturing CO_2 in situ through chemical looping technology. However, a major challenge in the production of bio-based high-energy fuels is the efficient removal of oxygen, and its removal typically involves energy-intensive and cost-intensive processes. Therefore, developing more efficient oxygen management strategies during the production of high-energy fuels remains a critical area requiring in-depth research.

4. **Biorefining poly-generation.** Recently, diverse biorefinery strategies have been proposed to simultaneously enhance the economic performance of biomass utilization routes and strengthen their potential for achieving net-negative carbon emissions. For example, Khoshgoftar Manesh et al. [65] developed an innovative solar-biomass-wind poly-generation system. By integrating core subsystems such as gasification, ammonia synthesis, and supercritical CO_2 cycle, the system efficiently converts municipal solid waste (as biomass feedstock) into three high-value products: ammonia, process steam, and electricity. The payback period of the system is further shortened to 2.99 years, demonstrating excellent techno-economic performance and application potential. Studies on poly-generation indicate that combining advanced fractionation and conversion techniques can enhance product yields, lower carbon emissions, and strengthen the economic feasibility of biomass-based biorefineries. Despite its potential, poly-generation remains a highly intricate process. Scaling up this technology to demonstration-level operations is challenging, requiring substantial experimental validation and being hindered by limited databases and process models.

1.2 Machine Learning Methods

ML models refer to computational algorithms that enhance performance on specific tasks by learning from experience, as evaluated by established performance metrics [66]. This definition emphasizes the fundamental principle of ML: leveraging data to progressively improve task execution. Figure 1.3a illustrates the hierarchical relationships of artificial intelligence (AI), ML, and deep learning (DL). By uncovering complex, often hidden patterns in datasets, ML enables the extraction of interconnections and the accurate prediction of target properties. As shown in Figure 1.3b, ML methods are typically divided into three main types: supervised, semi-supervised, and unsupervised learning.

Supervised regression learning is currently the most widely used approach in biomass conversion (Figure 1.3b). ML encompasses a wide range of algorithms commonly used to analyze complex data and develop predictive models. These methods include classical techniques such as linear regression and generalized linear models, as well as probabilistic approaches like Gaussian processes. Ensemble-based algorithms, including random forest (RF), XGBoost, AdaBoost, and gradient-boosted decision trees (GBDTs), are widely used for their robust performance. In addition,

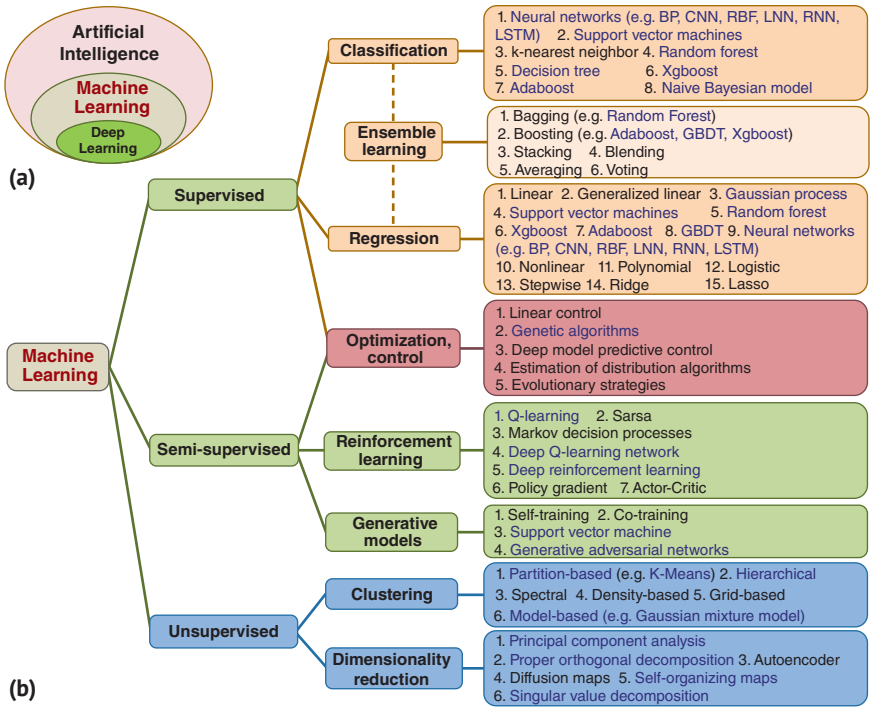

Figure 1.3 (a) The relationships among AI, ML, and DL, and (b) a classification of ML algorithms. *Source*: [66] / with permission of American Chemical Society.

a variety of neural network architectures have been developed to handle nonlinear and high-dimensional datasets, including backpropagation (BP) networks, recurrent neural networks (RNNs), layered neural networks (LNNs), convolutional neural networks (CNNs), radial basis function networks (RBFs), and long short-term memory (LSTM) networks.

A variety of ML libraries and frameworks are available for implementing algorithms, among which Scikit-learn remains the most widely adopted, alongside TensorFlow, PyTorch, and Keras. Each algorithm relies on a set of hyperparameters that require careful adjustment to achieve optimal predictive performance for specific datasets and problem contexts. Hyperparameter tuning may be performed either manually, employing techniques such as grid search or random search, or automatically through the use of optimization algorithms. Given the substantial computational cost and time inefficiency of manual tuning, automated optimization approaches are generally preferred for identifying the most effective hyperparameter configurations. The principles and applications of hyperparameter optimization methods have been extensively discussed in the literature [67, 68]. Frequently used algorithms for hyperparameter optimization include particle swarm optimization (PSO), simulated annealing (SA), genetic algorithms (GA), and Bayesian optimization (BO). While neural network (NN) algorithms tend to be more complex and computationally intensive than traditional ML approaches such as tree-based

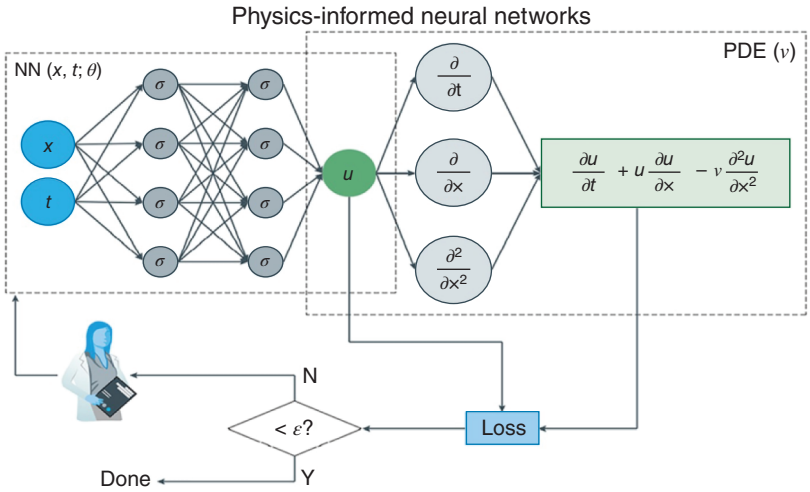

Figure 1.4 Schematic diagram of a physics-informed neural network (PINN). *Source*: [69] / with permission of Springer Nature.

models, they provide greater flexibility and have the potential to achieve superior predictive accuracy.

Additionally, several challenges remain for both simple and ensemble ML models. First, limited data availability and the absence of critical input features such as biomass particle size, reaction characteristics, and reactor configuration can significantly hinder model accuracy. Second, most data-driven models suffer from poor interpretability, which limits their generalization and effectiveness in process optimization. Moreover, the lack of integration of kinetic information makes it difficult to establish quantitative relationships between parameters, such as residence time and reactor dimensions, which are typically well captured by mechanism-based models. Recently, hybrid models that integrate mechanistic knowledge with data-driven approaches have attracted increasing attention in process modeling. For instance, PINNs embed governing equations into the loss function, thereby constraining model training to physical laws and boundary conditions (Figure 1.4). This approach enhances prediction accuracy while maintaining physical consistency, particularly in systems with limited or noisy data.

1.3 AI/ML-integrated Biomass Valorization

After retrieval and rigorous screening from the Web of Science database, a total of 898 relevant publications themed on "biomass + ML/AI" from 2014 to 2024 were obtained. This book employed CiteSpace to conduct bibliometric analysis [70]. Figure 1.5 illustrates the keyword co-occurrence network for the same period, comprising 418 nodes and 1190 links. The top 60 most frequently appearing keywords are summarized in Table 1.1.

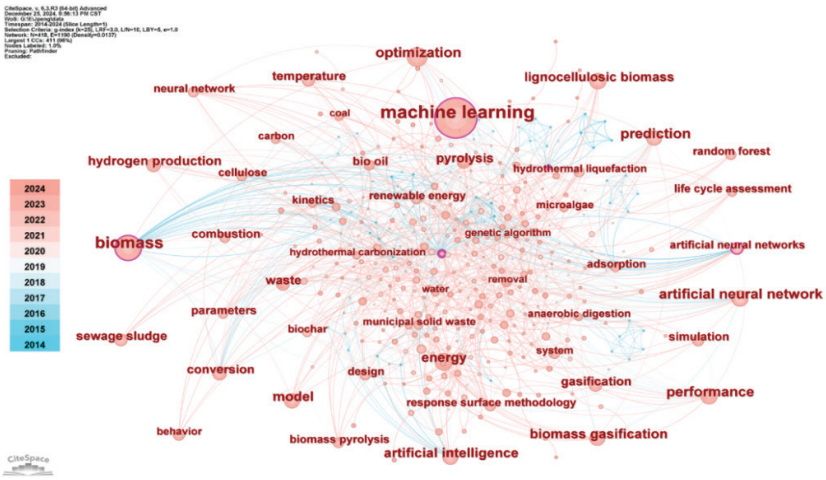

Figure 1.5 Analysis of keyword co-occurrence from 2014 to 2024.

Table 1.1 Overview of keyword clusters ranked by frequency of appearance.

No.	Frequency	Centrality	Keyword	No.	Frequency	Centrality	Keyword
1	359	0.11	Machine learning	31	28	0.03	Biomass pyrolysis
2	166	0.15	Biomass	32	28	0.08	Adsorption
3	88	0.05	Optimization	33	28	0.08	Kinetics
4	80	0.08	Performance	34	27	0.01	Hydrothermal liquefaction
5	79	0.08	Prediction	35	27	0.02	Behavior
6	77	0.08	Artificial neural network	36	27	0.03	Biochar
7	73	0.10	Artificial intelligence	37	26	0.02	Microalgae
8	71	0.03	Energy	38	26	0.01	Municipal solid waste
9	65	0.05	Biomass gasification	39	25	0.01	Carbon
10	62	0.02	Model	40	25	0	System
11	59	0.05	Pyrolysis	41	23	0.07	Anaerobic digestion
12	53	0.03	Hydrogen production	42	22	0.02	Hydrothermal carbonization
13	53	0.02	Lignocellulosic biomass	43	22	0	Removal
14	48	0.05	Waste	44	20	0.01	Water
15	46	0.06	Conversion	45	20	0.01	Coal

(Continued)

Table 1.1 (Continued)

No.	Frequency	Centrality	Keyword	No.	Frequency	Centrality	Keyword
16	46	0.03	Temperature	46	20	0.07	Genetic algorithm
17	45	0.01	Sewage sludge	47	19	0	Oil
18	44	0.02	Gasification	48	19	0.08	Classification
19	40	0.07	Parameters	49	19	0.01	Steam gasification
20	37	0.04	Combustion	50	19	0.04	Yield
21	35	0.02	Bio-oil	51	18	0	Gas
22	35	0.01	Simulation	52	18	0.01	Higher heating value
23	33	0.12	Artificial neural networks	53	18	0.07	Models
24	29	0.05	Design	54	17	0.08	Wood
25	29	0.04	Life cycle assessment	55	17	0.04	Pretreatment
26	29	0.04	Response surface methodology	56	17	0.21	Neural networks
27	29	0.04	Cellulose	57	17	0.03	Activated carbon
28	28	0.05	Renewable energy	58	17	0.03	Exergy analysis
29	28	0.02	Random forest	59	16	0.01	Regression
30	28	0.06	Neural network	60	16	0.02	Co pyrolysis

Co-occurrence analysis of keywords identified multiple research focuses, highlighting the central themes of biomass properties, conversion technologies, products, process modeling, and system-level analysis. Notably, hydrogen production emerged as a prominent topic. Based on these bibliometric insights, the following sections review recent advances in interpretable ML applied to biomass valorization, summarized into three core research areas: ML for data generation, ML for process modeling, and ML for system optimization (Figure 1.6).

1.3.1 AI-assisted Data Acquisition

A central difficulty in biomass valorization is the precise modeling of its conversion processes [19]. Unlike conventional chemical substances, biomass displays substantial complexity, with diverse chemical and elemental compositions and intricate

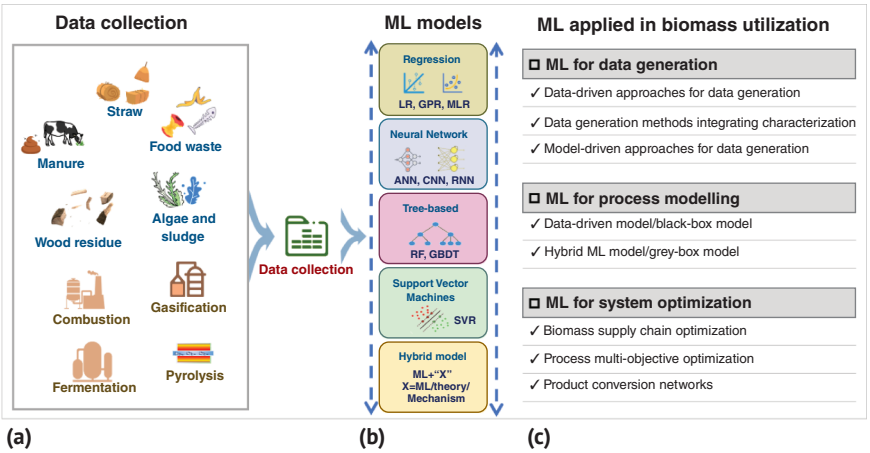

Figure 1.6 Advances in research combining biomass with machine learning, showing (a) data acquisition, (b) development of ML models, and (c) implementation of these models in biomass valorization processes.

thermal characteristics that are crucial for accurately predicting conversion behavior [71]. According to authoritative databases like Phyllis and NIST, a complete biomass dataset must encompass proximate analysis, ultimate analysis, chemical composition, ash components and thermodynamic properties [72, 73]. Nevertheless, acquiring these basic data via standard experiments and thermodynamic calculations is still laborious and time-consuming, and the accuracy is generally not assured [74].

ML techniques provide an effective framework for data generation, enhancement, preprocessing, and classification. Jiang et al. [44] provided a comprehensive review of current ML-assisted data generation strategies, which can be broadly categorized into three major categories. The "data-to-data" approach refers to purely data-driven methods for predicting target properties from readily available experimental data. For instance, HHVs can be predicted using ultimate and proximate analyses of biomass instead of direct calorimetric measurements. In contrast, the "characterization-to-data" paradigm combines ML algorithms with conventional analytical methods to extract or generate high-value information. Representative studies have integrated thermogravimetric analysis with ML models to estimate biochar yields, reconstruct TG/DTG profiles, and predict thermodynamic and kinetic parameters associated with biomass pyrolysis [75]. The "model-to-data" strategy utilizes established mathematical, statistical, or mechanistic models as data sources for expansion and refinement. In this context, generative adversarial networks (GAN) have shown considerable promise under data-scarce conditions. Wei et al. [76] demonstrated that GAN-based augmentation coupled with RF modeling significantly improved predictive accuracy in a bio-polymerization system. Furthermore, Jiang et al. [28] proposed a mechanistic pyrolysis framework grounded in Gibbs free energy minimization and combined it with ML techniques to expand datasets by preferentially retaining samples exhibiting lower deviations from experimental observations.

1.3.2 AI-assisted Process Modeling

Process modeling has become an essential approach for evaluating, optimizing, and predicting product yields in biomass valorization and biorefinery systems. Conventional modeling frameworks primarily depend on thermodynamic principles and reaction kinetics to represent the behavior of biomass pyrolysis [33–35]. As a representative example, Yang et al. [77] developed a detailed Aspen Plus simulation incorporating 21 kinetic rate expressions distributed across 14 sequential reactor units to describe the pyrolysis process. Similarly, Peters et al. [78, 79] developed a more comprehensive kinetic scheme comprising 149 equations to describe the volatilization, thermal decomposition, and depolymerization mechanisms during poplar pyrolysis. These efforts underscore the importance of integrating comprehensive reaction kinetics into process simulators to improve the predictive accuracy of biomass pyrolysis modeling. However, pyrolysis kinetic models still face significant challenges when applied to new biomass types due to complex reaction networks and substantial biomass compositional variability [80, 81].

In contrast to kinetic models, equilibrium models are preferred for their simplicity and broader extrapolation capability [82]. These modeling approaches are grounded in first-principles calculations, where mass and energy balance equations are systematically integrated with Gibbs free energy minimization techniques. Such a framework enables the reliable prediction of equilibrium product distributions under specified temperature and pressure conditions. Extensive research efforts have been devoted to use equilibrium model for developing biomass combustion, gasification, and pyrolysis processes [83, 84]. The results indicated that the equilibrium-based framework offers advantages in terms of computational simplicity, robust convergence, and good scalability. However, this approach inherently assumes that all reactions proceed to chemical equilibrium, an assumption that deviates from the actual behavior of biomass pyrolysis systems [82]. Consequently, such idealized conditions may introduce noticeable deviations between model outputs and experimentally observed process performance.

In summary, purely mechanistic representations of biorefining systems, including equilibrium-based and kinetic formulations, often fail to deliver reliable predictive performance because they depend on comprehensive thermodynamic and kinetic parameters that are difficult to acquire under realistic operating conditions. By comparison, data-driven methods have emerged as a powerful tool for representing the behavior of highly nonlinear and complex processes [85]. As depicted in Figure 1.7, ML-based process modeling can be broadly divided into two categories: fully data-driven ML frameworks (black-box models) and hybrid modeling strategies (gray-box models) that integrate physical knowledge with learning algorithms.

1.3.3 AI-assisted System Optimization

Effective system-level optimization in biomass utilization depends on the coordinated consideration of upstream resource characteristics, such as spatial distribution, feedstock categories, and availability, and downstream utilization routes, including value-added conversion pathways and substitution of fossil energy in

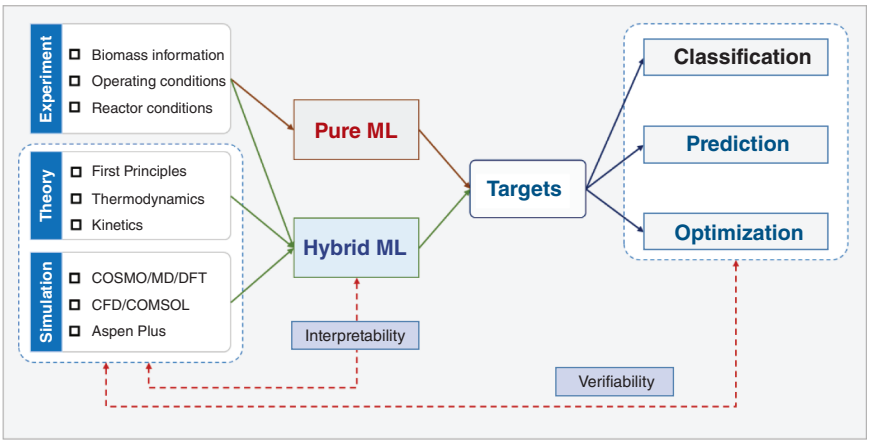

Figure 1.7 Application of ML in process modeling, illustrating the categorization of modeling strategies into data-driven ML approaches (black-box models) and integrated ML-mechanism frameworks (gray-box models).

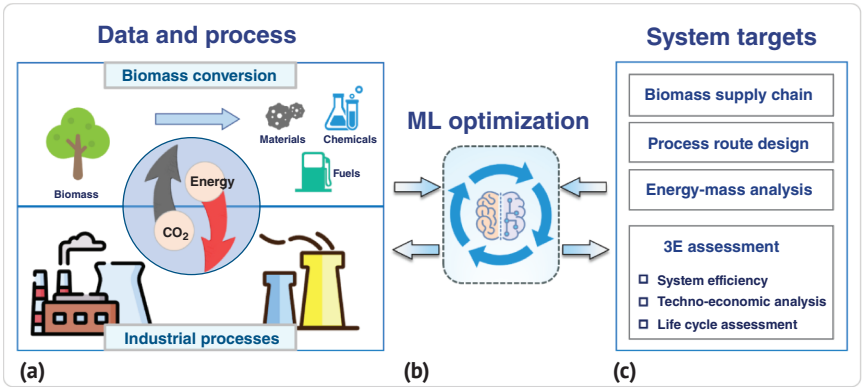

Figure 1.8 (a) Data generation associated with biomass conversion processes or fossil-fuel substitution pathways; (b) implementation of ML-based optimization strategies; (c) definition and evaluation of system-level objectives.

industrial systems. The realization of high-performance biorefining operations therefore requires a comprehensive trade-off among economic feasibility, environmental performance, and energy efficiency. However, the strong coupling between complex reaction kinetics and thermodynamic behaviors significantly constrains the effectiveness of traditional optimization techniques. These limitations become more pronounced when multiple and often competing objectives must be addressed simultaneously, such as cost reduction, emission mitigation, and energy performance across heterogeneous biomass resources and conversion technologies [86]. To overcome these challenges, ML-based optimization frameworks have been increasingly embedded into the construction of multiscale biomass databases and process modeling platforms, enabling system-wide optimization from perspectives of efficiency, techno-economic performance, and carbon mitigation (Figure 1.8).

At present, such ML-driven optimization strategies have been applied to a broad range of domains, including biomass supply chain planning [87], multi-objective process optimization [88], and the structural design of biomass conversion networks [89].

References

1 Savage, N. (2011). Fuel options: the ideal biofuel. *Nature* **474**: S9–S11.
2 Bartling, A. W., Stone, M. L., Hanes, R. J., Bhatt, A., Zhang, Y., Biddy, M. J., Davis, R., Kruger, J. S., Thornburg, N. E., and Luterbacher, J. S. (2021). Techno-economic analysis and life cycle assessment of a biorefinery utilizing reductive catalytic fractionation. *Energy & Environmental Science* **14**: 4147–4168.
3 Liu, W.-J., Li, W.-W., Jiang, H., and Yu, H.-Q. (2017). Fates of chemical elements in biomass during its pyrolysis. *Chemical Reviews* **117**: 6367–6398.
4 Liao, Y., Koelewijn, S.-F., Van den Bossche, G., Van Aelst, J., Van den Bosch, S., Renders, T., Navare, K., Nicolaï, T., Van Aelst, K., Maesen, M., Matsushima, H., Thevelein, J. M., Van Acker, K., Lagrain, B., Verboekend, D., and Sels, B. F. (2020). A sustainable wood biorefinery for low-carbon footprint chemicals production. *Science* **367**: 1385–1390.
5 Werpy, T. A., Petersen, G. E., Aden, A., Bozell, J. J., Holladay, J. E., White, J. F., Manheim, A., Eliot, D., Lasure, L. L., and Jones, S. (2004). *Top Value Added Chemicals from Biomass.* Volume 1 - *Results of Screening for Potential Candidates from Sugars and Synthesis Gas*. National Renewable Energy Laboratory.
6 IEA. (2017). *Technology Roadmap Delivering Sustainable Bioenergy*. Paris, France: International Energy Agency.
7 Liu, N., Zhou, J., Wang, R., Zhang, J., Yang, Z., Tian, F., Hu, Y., and Wu, Z. (2025). Green and efficient depolymerization and recycling of polyester/cotton blended fabrics by ethanolamine. *Sustainable Chemistry and Pharmacy* **45**: 102011.
8 Tian, C., Dorakhan, R., Wicks, J., Chen, Z., Choi, K.-S., Singh, N., Schaidle, J. A., Holewinski, A., Vojvodic, A., Vlachos, D. G., Broadbelt, L. J., and Sargent, E. H. (2024). Progress and roadmap for electro-privileged transformations of bio-derived molecules. *Nature Catalysis* **7**: 350–360.
9 Dossow, M., Klüh, D., Umeki, K., Gaderer, M., Spliethoff, H., and Fendt, S. (2024). Electrification of gasification-based biomass-to-X processes – a critical review and in-depth assessment. *Energy & Environmental Science* **17**: 925–973.
10 Turner, J. A. (2004). Sustainable hydrogen production. *Science* **305**: 972–974.
11 Liu, L., Jiang, P., Qian, H., Mu, L., Lu, X., and Zhu, J. (2022). CO_2-negative biomass conversion: an economic route with co-production of green hydrogen and highly porous carbon. *Applied Energy* **311**: 118685.
12 Wang, M., Dewil, R., Maniatis, K., Wheeldon, J., Tan, T., Baeyens, J., and Fang, Y. (2019). Biomass-derived aviation fuels: challenges and perspective. *Progress in Energy and Combustion Science* **74**: 31–49.

13 Upton, B. M. and Kasko, A. M. (2016). Strategies for the conversion of lignin to high-value polymeric materials: review and perspective. *Chemical Reviews* **116**: 2275–2306.

14 Ding, Y., Pang, Z., Lan, K., Yao, Y., Panzarasa, G., Xu, L., Lo Ricco, M., Rammer, D. R., Zhu, J., and Hu, M. J. C. R. (2022). Emerging engineered wood for building applications. *Chemical Reviews* **123**: 1843–1888.

15 Shi, X., Wang, Z., Liu, S., Xia, Q., Liu, Y., Chen, W., Yu, H., and Zhang, K. (2024). Scalable production of carboxylated cellulose nanofibres using a green and recyclable solvent. *Nature Sustainability* **7**: 315–325.

16 Qiu, Z. and Li, C.-J. (2020). Transformations of less-activated phenols and phenol derivatives via C–O cleavage. *Chemical Reviews* **120**: 10454–10515.

17 Mika, L. T., Cséfalvay, E., and Németh, Á. (2018). Catalytic conversion of carbohydrates to initial platform chemicals: chemistry and sustainability. *Chemical Reviews* **118**: 505–613.

18 Vassilev, S. V., Baxter, D., Andersen, L. K., and Vassileva, C. G. (2010). An overview of the chemical composition of biomass. *Fuel* **89**: 913–933.

19 Siewert, R., Zaitsau, D. H., Emel'yanenko, V. N., and Verevkin, S. P. (2019). Biomass valorization: thermodynamics of the Guerbet condensation reaction. *Journal of Chemical & Engineering Data* **64**: 4904–4914.

20 Wang, S., Dai, G., Yang, H., and Luo, Z. (2017). Lignocellulosic biomass pyrolysis mechanism: a state-of-the-art review. *Progress in Energy and Combustion Science* **62**: 33–86.

21 Dente, M., Bozzano, G., Faravelli, T., Marongiu, A., Pierucci, S., and Ranzi, E. (2007). Kinetic modelling of pyrolysis processes in gas and condensed phase. *Advances in Chemical Engineering* **32**: 51–166.

22 Mettler, M. S., Vlachos, D. G., and Dauenhauer, P. J. (2012). Top ten fundamental challenges of biomass pyrolysis for biofuels. *Energy & Environmental Science* **5**: 7797–7809.

23 Lo, S. L. Y., How, B. S., Leong, W. D., Teng, S. Y., Rhamdhani, M. A., and Sunarso, J. (2021). Techno-economic analysis for biomass supply chain: a state-of-the-art review. *Renewable and Sustainable Energy Reviews* **135**: 110164.

24 Jiang, G., Cai, L., Wang, S., Laaksonen, A., Feng, X., Mu, L., Lu, X., and Zhu, J. (2020). Critical role of carbonized cellulose in the evolution of highly porous biocarbon: seeing the structural and compositional changes of spent mushroom substrate by deconvoluted thermogravimetric analysis. *Industrial & Engineering Chemistry Research* **59**: 22541–22548.

25 Yang, Q., Zhou, H., Bartocci, P., Fantozzi, F., Masek, O., Agblevor, F. A., Wei, Z., Yang, H., Chen, H., Lu, X., Chen, G., Zheng, C., Nielsen, C. P., and McElroy, M. B. (2021). Prospective contributions of biomass pyrolysis to China's 2050 carbon reduction and renewable energy goals. *Nature Communications* **12**: 1698.

26 Jiang, P., Li, L., Lin, H., Ji, T., Mu, L., Ji, Y., Lu, X., and Zhu, J. (2024). Establishing a generalized model for accurate prediction of higher heating values of substances with large ash fractions. *Green Chemical Engineering* **6**: 372–379.

27 Jiang, P., Wang, C., Fan, J., Ji, T., Mu, L., Lu, X., and Zhu, J. (2024). Hybrid residual modelling of biomass pyrolysis. *Chemical Engineering Science* **293**: 120096.

28 Jiang, P., Fan, J., Li, L., Wang, C., Tao, S., Ji, T., Mu, L., Lu, X., and Zhu, J. (2024). A hybrid approach combining mechanism-guided data augmentation and machine learning for biomass pyrolysis. *Chemical Engineering Science* **296**: 120227.

29 Li, Y., Jiang, P., Li, L., Ji, T., Mu, L., Lu, X., and Zhu, J. (2025). Integrating machine learning for multi-objective optimization of biomass conversion to hydrogen. *Journal of Cleaner Production* **494**: 144948.

30 Qian, H., Guo, X., Fan, S., Hagos, K., Lu, X., Liu, C., and Huang, D. (2016). A simple prediction model for higher heat value of biomass. *Journal of Chemical & Engineering Data* **61**: 4039–4045.

31 Su, G., Jiang, P., Ong, H. C., Zhu, J., Amin, N. A. S., Zulkifli, N. W. M., and Ibrahim, S. (2023). Co-production of biochar and electricity from oil palm wastes for carbon dioxide mitigation in Malaysia. *Journal of Cleaner Production* **423**: 138749.

32 Xing, J., Luo, K., Wang, H., Gao, Z., and Fan, J. (2019). A comprehensive study on estimating higher heating value of biomass from proximate and ultimate analysis with machine learning approaches. *Energy* **188**: 116077.

33 Lu, L., Gao, X., Dietiker, J.-F., Shahnam, M., and Rogers, W. A. (2022). MFiX based multi-scale CFD simulations of biomass fast pyrolysis: a review. *Chemical Engineering Science* **248**: 117131.

34 Sharma, A., Wang, S., Pareek, V., Yang, H., and Zhang, D. (2015). Multi-fluid reactive modeling of fluidized bed pyrolysis process. *Chemical Engineering Science* **123**: 311–321.

35 Xiong, Q., Kong, S.-C., and Passalacqua, A. (2013). Development of a generalized numerical framework for simulating biomass fast pyrolysis in fluidized-bed reactors. *Chemical Engineering Science* **99**: 305–313.

36 Baruah, D., Baruah, D. C., and Hazarika, M. K. (2017). Artificial neural network based modeling of biomass gasification in fixed bed downdraft gasifiers. *Biomass and Bioenergy* **98**: 264–271.

37 Ren, S., Wu, S., and Weng, Q. (2023). Physics-informed machine learning methods for biomass gasification modeling by considering monotonic relationships. *Bioresource Technology* **369**: 128472.

38 Valentino, F., Lorini, L., Gottardo, M., Pavan, P., and Majone, M. (2020). Effect of the temperature in a mixed culture pilot scale aerobic process for food waste and sewage sludge conversion into polyhydroxyalkanoates. *Journal of Biotechnology* **323**: 54–61.

39 Kunatsa, T. and Xia, X. (2022). A review on anaerobic digestion with focus on the role of biomass co-digestion, modelling and optimisation on biogas production and enhancement. *Bioresource Technology* **344**: 126311.

40 Østby, H., Hansen, L. D., Horn, S. J., Eijsink, V. G. H., and Várnai, A. (2020). Enzymatic processing of lignocellulosic biomass: principles, recent advances and perspectives. *Journal of Industrial Microbiology & Biotechnology* **47**: 623–657.

41 Yin, Y., Hu, Y., and Wang, J. (2022). Co-fermentation of sewage sludge and lignocellulosic biomass for production of medium-chain fatty acids. *Bioresource Technology* **361**: 127665.

42 Luque, R. and Clark, J. H. (2011). Biodiesel-like biofuels from simultaneous transesterification/esterification of waste oils with a biomass-derived solid acid catalyst. *ChemCatChem* **3**: 594–597.

43 Lin, X., Wang, P., Hong, R., Zhu, X., Liu, Y., Pan, X., Qiu, X., and Qin, Y. (2022). Fully lignocellulosic biomass-based double-layered porous hydrogel for efficient solar steam generation. *Advanced Functional Materials* **32**: 2209262.

44 Jiang, P., Wang, C., She, W., Ye, W., Li, Y., Ji, T., Mu, L., Lu, X., and Zhu, J. (2026). Machine learning accelerated data generation, process modelling and system optimization for biomass conversion and valorization. *Renewable and Sustainable Energy Reviews* **226**: 116439.

45 Hu, J., Zhao, Y., Song, T., Pang, B., Wang, T., Zhou, S., Du, J., Lv, Y., Tao, Y., Lu, J., Fu, C., and Wang, H. (2026). Green and tough eutectogel from demethylated lignin: bridging renewable biomass and harsh-environment sensing. *Green Chemistry* **28**: 2770–2780.

46 Sun, X., Li, R., Zhang, B., Wang, H., Cheng, Y., Guan, J., Liu, Q., and Chen, X. (2025). Function–structure–synthesis: machine learning enabled closed-loop design of biomass-derived porous carbon materials. *ACS Sustainable Chemistry & Engineering* **13**: 7698–7709.

47 Wu, Y., Chen, S., Wu, J., Liu, F., Chen, C., Ding, B., Zhou, X., and Deng, H. (2024). Revivable self-assembled supramolecular biomass fibrous framework for efficient microplastic removal. *Science Advances* **10**: eadn8662.

48 Chen, L., Ji, T., Brisbin, L., and Zhu, J. (2015). Hierarchical porous and high surface area tubular carbon as dye adsorbent and capacitor electrode. *ACS Applied Materials & Interfaces* **7**: 12230–12237.

49 Chen, L., Ji, T., Yuan, R., Mu, L., Brisbin, L., and Zhu, J. (2015). Unveiling mesopore evolution in carbonized wood: interfacial separation, migration, and degradation of lignin phase. *ACS Sustainable Chemistry & Engineering* **3**: 2489–2495.

50 Jiang, G., Liu, L., Xiong, J., Luo, Y., Cai, L., Qian, Y., Wang, H., Mu, L., Feng, X., and Lu, X. J. M. B. (2022). Advanced material-oriented biomass precise reconstruction: a review on porous carbon with inherited natural structure and created artificial structure by post-treatment. *Macromolecular Bioscience* **22**: 2100479.

51 Xing, H., Jiang, G., Xiong, J., Zhu, Z., Zhang, L., Chen, M., and Zhu, J. (2023). The structure evolution from intrinsic micro-nano skeleton of natural wood to hierarchical porous carbon. *Microporous and Mesoporous Materials* **359**: 112653.

52 Wang, S., Jiang, G., Yang, Z., Mu, L., Ji, T., Lu, X., and Zhu, J. (2022). Boosted H_2O_2 productivity by sulfur-doped Pd/carbon catalysts via direct synthesis from H_2 and O_2 at atmospheric pressure. *ACS Sustainable Chemistry & Engineering* **10**: 13750–13758.

53 Jiang, G., Mou, L., Wang, Z., Zhang, L., Ji, T., Mu, L., Jiang, J., Lu, X., and Zhu, J. (2024). Manganese–palladium dual-atom catalyst boosts direct H_2O_2 synthesis beyond 2 wt % at atmospheric conditions. *ACS Catalysis* **14**: 17055–17064.

54 Qiu, Y., Jiang, P., Ye, W., Hu, J., Zhang, B., Ji, T., Mu, L., Feng, X., Lu, X., and Zhu, J. (2024). A spatially integrated electrochemical–thermal tandem reaction for continuous mild synthesis of propylene oxide. *Green Chemistry* **26**: 11206–11215.

55 Xiong, J., Deng, Z., Zhu, Z., Jiang, G., Qiu, Y., Chen, M., Mu, L., Lu, X., and Zhu, J. (2024). Nitrogen-doped porous carbon with regulated pore structure and surface wettability as flow electrodes for efficient capacitive deionization. *Journal of Electroanalytical Chemistry* **965**: 118356.

56 Cao, D., Ji, Y., Liu, L., Li, L., Li, L., Feng, X., Zhu, J., Lu, X., and Mu, L. J. F. (2022). A facile and green strategy to synthesize N/P co-doped bio-char as VOCs adsorbent: through efficient biogas slurry treatment and struvite transform. *Fuel* **322**: 124156.

57 Ji, T., Tu, R., Li, L., Mu, L., Liu, C., Lu, X., and Zhu, J. (2018). Localizing microwave heat by surface polarization of titanate nanostructures for enhanced catalytic reaction efficiency. *Applied Catalysis B: Environmental* **227**: 266–275.

58 Ji, T., Tu, R., Mu, L., Lu, X., and Zhu, J. (2018). Structurally tuning microwave absorption of core/shell structured CNT/polyaniline catalysts for energy efficient saccharide-HMF conversion. *Applied Catalysis B: Environmental* **220**: 581–588.

59 Ji, T., Yang, Z., Song, S., Zhou, T., Mu, L., Lu, X., and Zhu, J. (2022). Intense microwave heating at strongly polarized solid acid/water interface for energy-efficient platform chemical production. *Chemical Engineering Science* **262**: 118035.

60 Hong, L., Zhang, H., Hu, L., Xiao, R., and Chu, S. (2024). Near-infrared light–driven biomass conversion. *Science Advances* **10**: eadn9441.

61 Sun, M., Zhang, X., Yan, J., Zhang, H., Li, Z., Huang, L., Jin, S., Wang, W., and Li, N. (2026). High-density aviation fuel or diesel-range naphthenes are synthesized from biomass-derived isophorone and furfural. *Catalysts* **16**: 83.

62 Kumar, A., Katahira, R., Yang, Z., Bayles, A., Kumar, A., Ruddy, D., Bell, D. C., Heyne, J. S., Johnson, D. K., and Mittal, A. (2025). Advancing sustainable aviation fuel with high-energy-density bicycloalkanes production from corn stover mixed sugars. *Cell Reports Physical Science* **6**: 102692.

63 Fan, L.-S., Zeng, L., Wang, W., and Luo, S. (2012). Chemical looping processes for CO_2 capture and carbonaceous fuel conversion – prospect and opportunity. *Energy & Environmental Science* **5**: 7254.

64 Zeng, L., Cheng, Z., Fan, J. A., Fan, L.-S., and Gong, J. (2018). Metal oxide redox chemistry for chemical looping processes. *Nature Reviews Chemistry* **2**: 349–364.

65 Manesh, M. H. K., Davadgaran, S., Rabeti, S. A. M., and Blanco-Marigorta, A. M. (2025). Feasibility study of green ammonia and electricity production via an innovative wind-solar-biomass polygeneration system. *Applied Energy* **384**: 125467.

66 Venkatasubramanian, V. (2019). The promise of artificial intelligence in chemical engineering: is it here, finally? *AIChE Journal* **65**: 466.

67 Zhu, L.-T., Chen, X.-Z., Ouyang, B., Yan, W.-C., Lei, H., Chen, Z., and Luo, Z.-H. (2022). Review of machine learning for hydrodynamics, transport, and reactions

in multiphase flows and reactors. *Industrial & Engineering Chemistry Research* **61**: 9901–9949.

68 Yang, L. and Shami, A. (2020). On hyperparameter optimization of machine learning algorithms: theory and practice. *Neurocomputing* **415**: 295–316.

69 Karniadakis, G. E., Kevrekidis, I. G., Lu, L., Perdikaris, P., Wang, S., and Yang, L. (2021). Physics-informed machine learning. *Nature Reviews Physics* **3**: 422–440.

70 Ma, J., Zheng, L., and Yu, F. (2024). Current status and future prospects of biochar application in electrochemical energy storage devices: a bibliometric review. *Desalination* **581**: 117597.

71 Peduzzi, E., Boissonnet, G., and Maréchal, F. (2016). Biomass modelling: estimating thermodynamic properties from the elemental composition. *Fuel* **181**: 207–217.

72 Zhao, G., Jiang, P., Zhang, H., Li, L., Ji, T., Mu, L., Lu, X., and Zhu, J. (2024). Mapping out the regional low-carbon and economic biomass supply chain by aligning geographic information systems and life cycle assessment models. *Applied Energy* **369**: 123599.

73 Domalski, E. S., Jobe Jr, T. L., and Milne, T. A. (1986). *Thermodynamic Data for Biomass Conversion and Waste Incineration*. Washington, DC: National Bureau of Standards; Solar Energy Research.

74 Escamez, S., Terryn, C., Gandla, M. L., Yassin, Z., Scheepers, G., Näsholm, T., Sundman, O., Jönsson, L. J., Lundberg-Felten, J., Tuominen, H., Niittylä, T., and Paës, G. (2021). Fluorescence lifetime imaging as an in situ and label-free readout for the chemical composition of lignin. *ACS Sustainable Chemistry & Engineering* **9**: 17381–17392.

75 Jiang, P., She, W., Zhang, H., Li, L., Lin, H., Ji, T., Mu, L., Lu, X., and Zhu, J. (2025). Unlocking the synergistic effects in biomass-plastic co-pyrolysis using a hybrid machine learning approach with blended rule descriptors. *Biomass and Bioenergy* **201**: 108130.

76 Wei, S., Chen, Z., Arumugasamy, S. K., and Chew, I. M. L. (2022). Data augmentation and machine learning techniques for control strategy development in bio-polymerization process. *Environmental Science and Ecotechnology* **11**: 100172.

77 Xu, D., Yang, S., Su, Y., Shi, L., Zhang, S., and Xiong, Y. (2021). Simultaneous production of aromatics-rich bio-oil and carbon nanomaterials from catalytic co-pyrolysis of biomass/plastic wastes and in-line catalytic upgrading of pyrolysis gas. *Waste Management* **121**: 95–104.

78 Peters, J. F., Banks, S. W., Bridgwater, A. V., and Dufour, J. (2017). A kinetic reaction model for biomass pyrolysis processes in Aspen Plus. *Applied Energy* **188**: 595–603.

79 Peters, J. F., Iribarren, D., and Dufour, J. (2015). Simulation and life cycle assessment of biofuel production via fast pyrolysis and hydroupgrading. *Fuel* **139**: 441–456.

80 Sharma, N. and Liu, Y. A. (2022). A hybrid science-guided machine learning approach for modeling chemical processes: a review. *AIChE Journal* **68**: e17609.

81 Jiang, P., Zhao, G., Liu, L., Zhang, H., Mu, L., Lu, X., and Zhu, J. (2022). A negative-carbon footprint process with mixed biomass feedstock maximizes conversion efficiency, product value and CO_2 mitigation. *Bioresource Technology* **351**: 127004.

82 Lestinsky, P. and Palit, A. (2016). Wood pyrolysis using aspen plus simulation and industrially applicable model. *GeoScience Engineering* **62**: 11.

83 Shahbaz, M., AlNouss, A., Parthasarathy, P., Abdelaal, A. H., Mackey, H., McKay, G., and Al-Ansari, T. (2020). Investigation of biomass components on the slow pyrolysis products yield using Aspen Plus for techno-economic analysis. *Biomass Conversion and Biorefinery* **12**: 669–681.

84 Safarian, S., Rydén, M., and Janssen, M. (2022). Development and comparison of thermodynamic equilibrium and kinetic approaches for biomass pyrolysis modeling. *Energies* **15**: 3999.

85 Wang, H., Fu, T., Du, Y., Gao, W., Huang, K., Liu, Z., Chandak, P., Liu, S., Van Katwyk, P., and Deac, A. (2023). Scientific discovery in the age of artificial intelligence. *Nature* **620**: 47–60.

86 Liu, Z., Cui, Z., Wang, M., Liu, B., and Tian, W. (2024). A machine learning proxy based multi-objective optimization method for low-carbon hydrogen production. *Journal of Cleaner Production* **445**: 141377.

87 Mokarram, M., Ronizi, S. R. A., and Negahban, S. (2024). Optimizing biomass energy production in the southern region of Iran: a deterministic MCDM and machine learning approach in GIS. *Energy Policy* **195**: 114350.

88 Hiloidhari, M., Sharno, M. A., Baruah, D. C., and Bezbaruah, A. N. (2023). Green and sustainable biomass supply chain for environmental, social and economic benefits. *Biomass and Bioenergy* **175**: 106893.

89 Gong, J., Garcia, D. J., and You, F. (2016). Unraveling optimal biomass processing routes from bioconversion product and process networks under uncertainty: an adaptive robust optimization approach. *ACS Sustainable Chemistry & Engineering* **4**: 3160–3173.