

Article

Consolidated Bioprocessing of Lignocellulosic Biomass: A Review of Experimental Advances and Modeling Approaches

Mark Korang Yeboah *  and Dirk Söffker 

Chair of Dynamics and Control, University of Duisburg-Essen, Lotharstraße, 47057 Duisburg, Germany

* Correspondence: mark.korangyeboah@uni-due.de

Abstract

Growing global energy demand and concerns over climate change and fossil fuel depletion have increased interest in sustainable bioproducts such as ethanol. Unlike first-generation (1G) ethanol derived from food crops (e.g., corn), second-generation (2G) ethanol is produced from lignocellulosic biomass, an abundant non-food resource that addresses key sustainability concerns. Consolidated bioprocessing (CBP) integrates enzyme production, hydrolysis, and fermentation into a single step, using either microbial consortia or engineered microorganisms, thereby simplifying the process and potentially reducing costs compared with separate hydrolysis and fermentation (SHF) and simultaneous saccharification and fermentation (SSF). However, CBP systems are complex due to dynamic interactions among microbial communities, metabolic pathways, and process conditions. Addressing this complexity requires modeling approaches that capture nonlinear relationships and support robust process optimization. Machine learning (ML)-based models offer data-driven tools to represent complex bioprocess dynamics, improve predictive accuracy, and optimize bioproduct formation, thereby supporting progress toward commercial viability. Although CBP can be applied to a range of bioproducts, this review primarily focuses on lignocellulosic ethanol and closely related biofuels. The review provides a comprehensive overview of key CBP processes, the current state of CBP modeling, major limitations, and the emerging role of ML in addressing modeling challenges. It summarizes recent modeling techniques for CBP, including polynomial models and response surface methodologies, and discusses regression and neural network approaches in detail. Both first-principles and data-driven modeling strategies are considered, highlighting advances that can improve the scalability and efficiency of CBP for bioproduction. Overall, this review offers perspectives on modeling-enabled pathways for utilizing low-cost lignocellulosic biomass in sustainable bioprocessing.

Keywords: consolidated bioprocessing; lignocellulosic biomass; mechanistic modeling; machine learning; hybrid modeling; digital twins; model predictive control

Academic Editor: Ronald Charles
Sims

Received: 31 October 2025

Revised: 24 February 2026

Accepted: 28 February 2026

Published: 5 March 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

Rising global energy demand, alongside concerns about climate change and fossil fuel depletion, has intensified the search for renewable and sustainable alternatives [1,2]. Among these, biofuels, particularly ethanol, have attracted substantial attention because they can reduce greenhouse gas emissions and dependence on petroleum-based fuels [3]. However, first-generation ethanol derived from food crops such as corn and sugarcane raises ethical and environmental concerns related to food security and land use [4,5]. These concerns have encouraged a shift toward second-generation ethanol produced from

lignocellulosic biomass, an abundant non-food resource from agricultural residues, forest waste, and municipal solid waste [6,7]. Although this 1G/2G transition has driven much of the interest in lignocellulosic ethanol, consolidated bioprocessing (CBP) has also been investigated for starch-derived and other carbohydrate-rich feedstocks within broader biorefinery contexts.

Lignocellulosic biomass nevertheless poses substantial processing challenges due to its complex and recalcitrant structure. The dense cellulose, hemicellulose, and lignin matrix resists enzymatic breakdown, typically requiring pretreatment to release fermentable sugars [8–10]. Conventional bioconversion is commonly performed in multiple steps (pretreatment, enzymatic hydrolysis, and fermentation), each requiring distinct conditions and adding to costs and inefficiencies [7,11]. Consolidated bioprocessing addresses these drawbacks by integrating enzyme production, hydrolysis, and fermentation into a single step, thereby simplifying the process and reducing time and cost. It can employ engineered microorganisms or microbial consortia capable of performing the required functions, providing a promising route toward more economical bioconversion [3,10,12]. Although CBP can also be applied to starch-based substrates and other agro-industrial residues, this review focuses on lignocellulosic biomass because of its abundance and minimal competition with food.

Industrial adoption of CBP remains limited by technical barriers, including low product yields, limited strain robustness, inconsistent performance at scale, and difficulties in optimizing tightly coupled process stages [13,14]. Overcoming these challenges requires improved understanding and control of processes. Traditional empirical approaches (e.g., response surface methodology, RSM) have supported optimization but often struggle to represent the nonlinear and dynamic behavior of CBP systems [12,15]. Advanced computational approaches, particularly machine learning (ML) and artificial intelligence (AI), offer complementary tools for capturing complex variable interactions, supporting optimization, and improving the prediction of system behavior [14,15].

This review provides an overview of CBP of lignocellulosic biomass, emphasizing key processes, the current state of modeling, and limitations that restrict broader application. It covers established and emerging modeling strategies, including polynomial and response surface methods, as well as regression-based and neural network approaches. We also discuss the integration of first-principles and data-driven models, highlighting how recent advances, especially in ML, can support improved prediction, optimization, and scalability. By synthesizing these developments, the review offers perspectives on how modeling can accelerate efficient and sustainable conversion of low-cost lignocellulosic biomass into valuable bioproducts.

This review synthesizes peer-reviewed studies on consolidated bioprocessing of lignocellulosic biomass, with emphasis on experimental advances and modeling (first-principles, data-driven, and hybrid). Literature was identified using keyword-based searches (e.g., “consolidated bioprocessing”, “lignocellulosic ethanol”, “hybrid modeling”, “machine learning”, “digital twin”) in major scholarly databases, complemented by backward/forward citation screening. Selection and reporting were guided by established systematic-review reporting recommendations [16].

2. Research Advances in CBP for Bioproduction

Consolidated Bioprocessing presents a fully integrated bioprocess in which the generation of enzymes, enzymatic hydrolysis of the released cellulose, and fermentation of the glucose to ethanol all take place in a single process step (Figure 1) [6,13,14]. This process can be accomplished in two different ways: either by cultivating together a community of microbes in efficient reactor systems that have niches for each community member and

divide the labor of hydrolysis and fermentation among different species or by using a genetically modified organism (GMO) in conventional bioreactors that can both hydrolyze cellulose and ferment sugars to ethanol [9,13]. This approach improves the economic viability of second-generation bioproducts by shortening the fermentation period and reducing capital and operating expenses. It presents the highest level of process integration for bioprocessing [6,9]. Due to ongoing fermentation, this mechanism also prevents the accumulation of inhibitory quantities of reducing sugars (glucose), because the released glucose is continuously consumed by the fermenting microorganisms, which improves saccharification and fermentation to produce bioproducts [3,12,15].

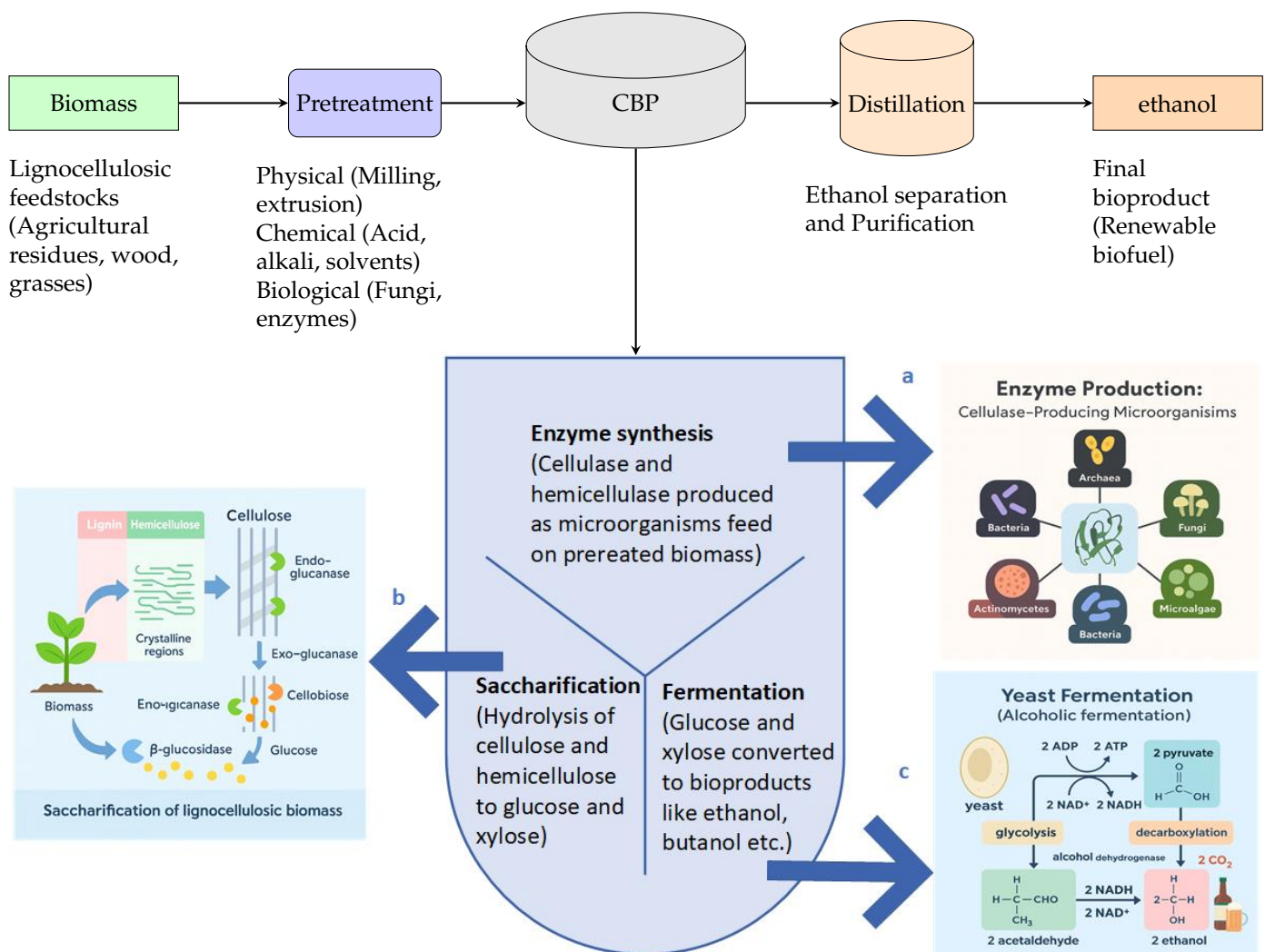


Figure 1. A schematic diagram of consolidated bioprocessing for the production of ethanol. Redrawn from: a [17], b [18], and c [19].

2.1. Enzyme Synthesis

In CBP, enzyme synthesis is tightly linked to the type of microorganism used, its natural habitat, and the cultivation strategy employed. Microbes such as bacteria, yeasts, and filamentous fungi can convert organic materials (e.g., lignocellulosic residues, agro-industrial wastes, food-processing byproducts, and algal biomass) into various biofuels, including ethanol, biogas, and biohydrogen, or into lipid-rich intermediates for biodiesel. For the synthesis of hydrolytic enzymes required in CBP (e.g., cellulases, hemicellulases, and lignin-degrading enzymes), two main cultivation strategies are employed: submerged fermentation (SmF), where microorganisms grow in a liquid medium, and solid-state

fermentation (SSF), where growth occurs on moist solid particles with minimal free water [20–22]. In SSF, non-photosynthetic microbes such as bacteria and filamentous fungi grow in thin water films and gas-filled pores between particles, obtaining carbon and other nutrients from the solid substrate and dissolved compounds [3,20]. Owing to simple equipment, lower capital costs, greater yields per reactor volume, reduced contamination, low effluent generation, and minimal aeration and agitation requirements, SSF has in some cases replaced SmF for producing selected chemicals and enzymes from lignocellulosic substrates [12,20]. Representative microorganisms and enzyme systems produced via SmF and SSF that are relevant for CBP-oriented biofuel production are summarized in Table 1.

Cellulose can be broken down by many bacteria, actinomycetes, and filamentous fungi, but filamentous fungi are widely used in commercial enzyme production because the amount and diversity of enzymes generated are typically greater than those produced by yeasts or bacteria [9,22,23]. SSF, which mimics the natural environment of these fungi, is particularly suitable for producing cellulolytic and hemicellulolytic enzymes [24], and most industrial cellulases and hemicellulases are secreted by filamentous fungi as enzyme complexes [13,21,25,26]. In parallel, cellulolytic bacteria such as *Ruminococcus* and *Clostridium* inhabit ruminant and insect guts [23,27], where raw biomass is directly digested in multi-chambered stomachs [28,29]. This “natural CBP” provides a blueprint for engineered CBP, demonstrating how complex biomass can be converted by coordinated microbial communities [25,29,30]. Among bacterial candidates, the obligate anaerobe *Clostridium thermocellum* is the most extensively studied CBP-enabling organism and exhibits one of the fastest growth rates on crystalline cellulose [20,23,31]. Synthetic co-cultivation systems with specialized roles and mixed consortia including bacteria, yeasts, and fungi are increasingly considered attractive CBP agents because enzymatic tasks can be distributed across multiple populations [25,32–34].

At the level of cellulose hydrolysis, cellulase-mediated conversion can be broadly described in three phases: adsorption of cellulase onto the cellulose surface, biotransformation of cellulose into fermentable sugars, and subsequent desorption of the enzyme [34]. A suite of enzymes acts synergistically: exo- and endo- β -glucanases attack the cellulose chain from the reducing and non-reducing ends and exhibit strong synergism [21,35], while β -glucosidase converts cellobiose into glucose, preventing its accumulation and alleviating catabolic repression. These hydrolytic processes are controlled by substrate concentration, enzyme dosage, and reaction conditions [36]. For example, *Trichoderma* species typically produce higher amounts of endo- and exo- β -glucanases and lower amounts of β -glucosidase [36,37], whereas *Aspergillus* species generate more endo- β -glucanase and β -glucosidase and relatively less exo- β -glucanase [18,22,34]. Reducing sugar yield and reaction rates increase at low substrate concentrations but decrease at high loadings due to end-product inhibition when released sugars accumulate faster than they are fermented. In well-coupled CBP configurations, simultaneous fermentation reduces sugar accumulation and thus alleviates end-product inhibition [37,38]. While increasing enzyme dosage enhances reducing sugar output, it also raises processing costs, motivating optimization of temperature, pH, incubation time, and enzyme loading. In addition, lignin negatively affects cellulases through non-productive adsorption and irreversible binding, reducing cellulose accessibility and overall hydrolysis efficiency [24,30,35,39–41].

Genetic engineering plays a vital role in enhancing enzyme synthesis in CBP systems [42]. Through synthetic biology and metabolic engineering, heterologous enzyme genes have been introduced into industrial strains to broaden their catalytic capabilities, and promoters, signal peptides, and secretion pathways have been optimized to increase enzyme yield and activity under fermentation conditions [33,43]. Parallel efforts aim to design microbial consortia or co-culture systems in which different organisms provide com-

plementary enzymatic activities, further boosting biomass conversion [32,33]. However, challenges remain in balancing enzyme production with microbial growth and product synthesis: overproduction of enzymes can impose metabolic burdens that reduce viability and productivity [44], and the complex regulation of enzyme gene expression in response to diverse and often recalcitrant biomass feedstocks demands finely tuned control systems [38]. Despite these challenges, advances in systems biology, omics technologies, and bioprocess optimization are steadily improving the feasibility and economic viability of CBP, supporting sustainable large-scale biofuel production [43].

Table 1. Functions of lignocellulose-degrading enzymes and their associated microorganisms.

Enzymes	Specific Types	Function	Microorganisms (Bacterial and Fungal Species)	References
Cellulases	Endoglucanase (EG)	Breaks internal bonds in cellulose, creating new chain ends.	Bacterial: <i>Clostridium</i> sp., <i>Cellulomonas</i> sp., <i>Thermomonospora</i> sp., <i>Bacillus</i> sp., <i>Streptomyces</i> sp., <i>R. flavefaciens</i> , <i>Pedobacter</i> sp., <i>F. succinogenes</i> , <i>R. albus</i> , <i>Mucilaginibacter</i> sp. Fungal: <i>T. reesei</i> , <i>T. viride</i> , <i>A. niger</i> , <i>P. helicum</i> , <i>P. betulinus</i> , <i>A. nidulans</i> , <i>A. fumigatus</i> , <i>A. oryzae</i> , <i>M. grisea</i> , <i>N. crassa</i> , <i>F. gramineum</i>	[18,24,26,37,45]
	β -Glucosidase (BG)	Hydrolyzes cellobiose to glucose.		
	Exoglucanase (CBH)	Releases cellobiose from cellulose chain ends.		
Hemicellulases	Xylanases	Cleave β -1,4-xylosidic bonds in xylan.	Bacterial: <i>Bacillus</i> sp., <i>P. bryantii</i> , <i>P. xylanivorans</i> , <i>F. succinogenes</i> , <i>R. albus</i> , <i>Pedobacter</i> sp., <i>Mucilaginibacter</i> sp. Fungal: <i>A. niger</i> , <i>P. betulinus</i> , <i>R. flavefaciens</i> , <i>B. cinerea</i> , <i>A. nidulans</i> , <i>A. fumigatus</i> , <i>A. oryzae</i> , <i>M. grisea</i> , <i>F. gramineum</i>	[9,21,34,35]
	β -Xylosidase	Converts xylooligomers into xylose.		
	α -Galactosidase	Removes galactose side chains.		
	Acetyl esterase	Removes acetyl groups from xylan.		
	Mannanase	Hydrolyzes mannans.		
Lignases	Laccase (LaC)	Oxidizes lignin via radical generation.	Bacterial: <i>A. lipoferum</i> , <i>B. subtilis</i> , <i>C. basiliensis</i> , <i>R. ornithinolytica</i> , <i>Prevotella</i> sp., <i>Pseudomonas</i> sp., <i>Pseudobutyrvivibrio</i> sp. Fungal: <i>D. squalens</i> , <i>G. applanatum</i> , <i>T. reesei</i> , <i>T. longibrachiatum</i> , <i>M. tremellosus</i> , <i>P. chrysosporium</i> , <i>C. subvermispora</i> , <i>P. cinnabarinus</i> , <i>Pleurotus</i> sp., <i>P. rivulosus</i>	[31,32,39,40]
	Lignin peroxidase (LiP)	Degrades lignin using H_2O_2 .		
	Manganese peroxidase (MnP)	Degrades lignin using Mn^{3+} radicals.		
	Versatile peroxidase (VP)	Combines catalytic features of LiP and MnP.		

2.2. Glucose Production (Hydrolysis)

Most cellulose fibers consist of parallel, 1,4-linked glucose polymers that are linear and rigid due to extensive hydrogen bonding; the chains are unbranched and resist bending because of strong intra- and intermolecular interactions between glucose units [46,47]. These polymers form a water-resistant crystalline array, and because cellulose is resistant to hydrolysis by weak acids or alkalis, conversion to fermentable glucose typically requires strong acids and/or high temperatures [41,46]. Such methods are not widely used because acids must be recovered, and dilute-acid treatment at high temperatures can cause browning reactions and generate inhibitory compounds that reduce glucose yields and hinder subsequent ethanol fermentation [47]. Consequently, enzymatic hydrolysis using mixtures of cellulases and accessory enzymes is generally preferred [31,38]. Since lignocellulosic biomass contains both cellulose and hemicellulose, pretreatment followed by enzymatic hydrolysis releases a mixture of sugars: cellulose-derived glucose is often the dominant hexose, while hemicellulose contributes substantial pentoses, particularly xylose and arabinose. Xylose can constitute a significant fraction of the total sugar pool in many agricultural residues and is therefore critical for high overall carbon utilization in CBP [23,48].

Cellulases primarily act on accessible regions of cellulose, including loose ends of amorphous cellulose at fiber surfaces (exo- β -1,4-glucanases) and flexible internal segments (endo- β -1,4-glucanases). Both enzyme types release cellobiose, which must be further converted to glucose. Depending on the preceding delignification strategy, severe treatments may increase the number of amorphous sites or loose ends available for cellulase attack, yet overall hydrolysis can remain slow [25,49,50]. Enzyme adsorption to a finite number of target sites is often a rate-limiting step; therefore, increasing enzyme concentration does not necessarily increase hydrolysis rates as in typical homogeneous enzyme reactions. As a result, cellulose hydrolysis can remain sluggish, requiring long residence times and, in some cases, multiple reactors. In CBP, however, fermentation, enzymatic hydrolysis, and enzyme synthesis occur simultaneously in the same broth, which can reduce cellulase product inhibition by continuously consuming released glucose [41,48].

During CBP, microorganisms produce biomass-degrading enzymes to generate monomeric sugars for growth and metabolism. In natural settings, these enzymes are deployed mainly via two systems: cellulosomal complexes and free-enzyme systems [49,51]. In the free-enzyme system, secreted enzymes act individually on biomass substrates and are widely used in biorefineries because they are comparatively straightforward to implement. Given the structural complexity of cellulose and hemicellulose, a diverse enzyme set is needed to cleave different linkages during lignocellulose hydrolysis; these “molecular scissors” release specific monosaccharides from complex polysaccharides [30,52,53].

2.3. Microbial Fermentation

Using yeasts or fermentative bacteria, microbial fermentation converts sugars released from lignocellulosic biomass into biofuels or biochemicals [49,54,55]. This can be carried out alongside enzymatic hydrolysis (simultaneous saccharification and fermentation, SSF), separately from hydrolysis (separate hydrolysis and fermentation, SHF), or by integrating both functions within microorganisms (consolidated bioprocessing, CBP) [50–52]. Glucose and xylose may be fermented individually or simultaneously (co-fermentation) [49,53], and from a cost perspective, CBP is often regarded as more economical than other process configurations.

To enable efficient xylose fermentation, extensive metabolic engineering of *Saccharomyces cerevisiae* and *Zymomonas mobilis* has demonstrated xylose utilization and ethanol production [54,56]. Nonetheless, further improvements are required to achieve industrially competitive performance, particularly in rapid glucose–xylose co-fermentation, inhibitor tolerance, and strain robustness. Xylose is a key component of lignocellulosic hydrolysates, yet many industrial yeast strains (e.g., wild-type *S. cerevisiae*) cannot naturally ferment xylose, and even engineered strains often show lower rates and yields on xylose than on glucose. Alternative xylose-utilizing organisms, including native pentose-fermenting yeasts (e.g., *Scheffersomyces stipitis*, *Pichia stipitis*) and certain bacteria, have therefore been explored for CBP. However, efficient co-fermentation remains challenging due to glucose repression, redox imbalances in heterologous xylose pathways, and inhibitory compounds generated during pretreatment [30,45].

Microorganisms such as *Clostridium thermocellum* and *Clostridium phytofermentans* can directly ferment lignocellulosic material into biofuels. The anaerobic thermophilic bacterium *C. thermocellum* typically ferments at around 60 °C and produces a cellulosome, an enzyme complex that can degrade cellulose more effectively than free enzymes in pure-substrate studies [48,52,54]. Its main fermentation products include ethanol and acetic acid; despite its promise as a CBP organism, *C. thermocellum* is constrained by limited tolerance to fermentation products [6,37,51]. Another active research focus is the development of CBP yeast, where cellulase and hemicellulase genes are introduced into biofuel-producing yeast

strains [26,36,53,55]. While ethanol remains the primary target in most CBP studies, mixed sugar streams containing glucose and xylose have also been converted to other products (e.g., ABE solvents or organic acids) using suitable microbial consortia, although the experimental and modeling literature remains largely ethanol-focused.

2.4. Challenges in Sugar Utilization and Bioproduct Formation

One of the major challenges in CBP is the efficient and balanced utilization of mixed sugars derived from lignocellulosic biomass [57]. Lignocellulose contains both hexoses (e.g., glucose) and pentoses (e.g., xylose and arabinose), but many native CBP microorganisms cannot effectively metabolize pentoses. Even in engineered strains, glucose repression, where glucose inhibits the uptake and metabolism of other sugars, can lead to inefficient sugar utilization [42,58]. This sequential consumption reduces overall fermentation efficiency and prolongs processing time, negatively affecting productivity and yield [59,60].

Currently, separate hydrolysis and fermentation (SHF) remains a common configuration for lignocellulosic biofuel production [61,62]. An alternative is simultaneous saccharification and co-fermentation (SSCF), where co-fermenting microorganisms convert hydrolysates containing both glucose and xylose [50–52]. However, xylose fermentation is often significantly slower than glucose fermentation. In *S. cerevisiae*, xylose uptake largely relies on native glucose transporters due to the lack of dedicated xylose transporters [38,54]. Because these transporters have a higher affinity for glucose, xylose is typically consumed after glucose depletion in *S. cerevisiae* (and many other microbes) [36]. By that stage, accumulated ethanol and other fermentation metabolites can further inhibit xylose metabolism [48,55]. Nevertheless, some microorganisms can occasionally consume glucose and xylose simultaneously (i.e., without strong catabolite repression) under specific conditions (e.g., *Thamnidium elegans* under aerobic conditions) [45,53,57].

The toxicity of sugar-degradation intermediates (e.g., furfural and hydroxymethylfurfural, HMF) and fermentation byproducts (e.g., acetic acid) is another major limitation. These compounds, generated during biomass pretreatment, can inhibit microbial metabolism [8,63,64] by disrupting sugar transport, enzyme activity, and cellular respiration, thereby impairing both sugar utilization and bioproduct formation [64,65]. Mitigation strategies include developing inhibitor-tolerant strains and engineering detoxification pathways that neutralize or bypass these inhibitors [66]. In addition, CBP imposes a metabolic burden because the host must simultaneously support enzyme production, sugar catabolism, and product synthesis, often creating resource competition between growth, enzyme expression, and product-formation pathways [63,66]. This load can reduce the efficiency of one or more process components, leading to suboptimal yields [50]. Approaches such as dynamic pathway regulation, compartmentalization, and co-culture systems are being explored to alleviate this burden and improve overall process balance [60,67].

Achieving high titers, volumetric productivities, and yields remains a major obstacle in CBP development. Many promising microorganisms exhibit low productivity under industrial conditions due to suboptimal enzyme secretion, limited precursor availability, or feedback inhibition in product pathways. Moreover, the heterogeneity and complexity of lignocellulosic substrates can introduce variability in sugar release and fermentation performance. Addressing these challenges requires integrated efforts combining metabolic engineering, process optimization, and systems biology to enhance robustness and efficiency for commercial-scale bioproduction.

2.5. Experimental Approaches for Optimizing CBP Systems

Optimizing CBP systems requires a strategic combination of experimental and statistical approaches to enhance enzyme production, substrate utilization, and overall process

efficiency. Various optimization techniques, including classical one-factor-at-a-time (OFAT) methods and advanced statistical models such as RSM, are employed to identify key process variables and their interactions [68]. Factorial designs, such as the Plackett–Burman design (PBD) for screening significant parameters and the Central Composite Design (CCD) for response optimization, allow for precise control over fermentation conditions, leading to improved yields and cost-effective process optimization [69]. Additionally, bioreactor scaling studies, conducted under controlled and uncontrolled conditions, provide insights into the kinetic behavior of microbial strains and the feasibility of large-scale production [70]. By integrating these approaches, CBP systems can be fine-tuned to maximize microbial efficiency and enzyme productivity, paving the way for industrial applications in biofuel and bioproduct manufacturing [59,71].

A co-immobilized cultivation system incorporating *Trichoderma reesei*, *Aspergillus niger*, and *Zymomonas mobilis* has been shown to improve enzymatic hydrolysis and fermentation efficiency for ethanol production [72]. The synergistic action of *T. reesei* and *A. niger* increased saccharification enzyme activity, resulting in a cellulose conversion rate of 46.27% and a reducing sugar concentration of 2.57 g/L, while immobilized *Z. mobilis* in alginate beads enhanced fermentation, achieving an ethanol concentration of 0.56 g/L under the reported conditions [72]. Co-immobilized cultivation outperformed suspension cultivation in terms of saccharification enzyme activity and overall ethanol production, highlighting the importance of spatial organization and carrier choice as experimental levers for CBP optimization [71]. Although this configuration was optimized empirically, it also illustrates key variables, such as mass-transfer limitations within beads and the coupling between hydrolysis and fermentation, that could be more systematically explored using mechanistic or computational models [73].

Metabolic engineering also plays a crucial role in improving CBP-based ethanol production. A co-culture of *Clostridium thermocellum* and *Thermoanaerobacterium saccharolyticum* achieved an ethanol titer of 38 g/L from 92 g/L of avicel, while reducing organic acid byproducts that typically inhibit fermentation [74]. This study indicated that metabolic constraints, rather than cellulolytic efficiency, were the primary limiting factors in ethanol production. With approximately 90% avicel hydrolysis and ethanol titers reaching about 80% of theoretical yields, these findings support the potential of genetic engineering in optimizing CBP systems [58,74]. Further improvements have been observed using membrane-based bioreactors (MBMs) [74]. In one configuration, the combination of *Trichoderma reesei*, *Saccharomyces cerevisiae*, and *Scheffersomyces stipitis* enabled high ethanol yields from undetoxified dilute-acid-pretreated wheat straw, reaching 7.2 g/L ethanol, with 79% of the ethanol diffusing into the gas phase for in situ removal. This reduced toxicity and improved downstream processing efficiency, highlighting the advantages of co-culturing cellulolytic and fermentative microorganisms for industrial ethanol production [71,74]. Further optimization of strain selection and media composition could enhance the applicability of MBM technology in large-scale operations.

A stable artificial symbiotic system was developed for ABE (acetone–butanol–ethanol) fermentation through CBP using *Clostridium cellulovorans* and *Clostridium beijerinckii* [75]. This mixed-culture approach facilitated simultaneous saccharification and fermentation of alkali-extracted deshelled corn cobs (AECC) without the need for external enzyme addition. The system degraded 68.6 g/L of AECC and produced 11.8 g/L of solvents (2.64 g/L acetone, 8.30 g/L butanol, and 0.87 g/L ethanol) within 80 h, demonstrating a substantial improvement over initial conditions. Real-time polymerase chain reaction (PCR) analysis of 16S ribosomal RNA (rRNA) gene sequences provided insights into microbial interactions, revealing cooperative and competitive dynamics between *C. cellulovorans* and *C. beijerinckii* that contributed to improved biobutanol production [75]. The co-culturing approach

leveraged the metabolic complementarities of *C. cellulovorans* for cellulose degradation and *C. beijerinckii* for solventogenesis, eliminating the need for externally added butyrate. However, the molecular mechanisms governing interspecies interactions remain unclear, necessitating further studies to elucidate the metabolic pathways involved [76].

A summary of different experimental approaches for CBP optimization is shown in Table 2. Although empirical optimization can yield strong results, integrating advanced modeling techniques, such as machine learning (e.g., artificial neural networks (ANNs)) and hybrid models, can further improve process predictability and scalability. Overall, these findings underscore the importance of microbial synergy in biofuel research and provide a foundation for optimizing CBP strategies to achieve cost-effective, sustainable ethanol and biobutanol production [73,75].

Table 2. Summary of selected experimental (Trial-and-error/OFAT) approaches for CBP optimization.

Microbial Consortia	Substrate	Bioproduct	Yield/Productivity ^a	Reference
Co-culture of <i>Clostridium beijerinckii</i> ^b and <i>Clostridium cellulovorans</i> ^b	Alkali-extracted deshelled corn cobs	Acetone, butanol, ethanol (ABE)	2.64 g/L acetone, 8.30 g/L butanol, 0.87 g/L ethanol; Productivity = 11.8 g/L of ABE solvents in less than 80 h	[75]
<i>T. reesei</i> ^f BCRC 31863, <i>A. niger</i> ^f BCRC 3113, <i>Z. mobilis</i> ^b BCRC 10809	Carboxymethyl-cellulose	Ethanol	Productivity = 0.56 g/L and reducing sugar conversion = 11.2 % in 24 h	[72]
<i>Clostridium thermocellum</i> ^b and <i>Thermoanaerobacterium saccharolyticum</i> ^b	Avicel	Ethanol, acetate, lactate	Productivity = 38 g/L of ethanol in 146 h	[58]
<i>Trichoderma reesei</i> ^f , <i>Saccharomyces cerevisiae</i> ^f , and <i>Scheffersomyces stipitis</i> ^f	Wheat straw	Ethanol	Yield = 67 %	[74]
<i>Saccharomyces cerevisiae</i> ^f and <i>C. phytofermentans</i> ^b	α -cellulose	Ethanol	Yield = 22 g/L ethanol from 100 g/L of α -cellulose	[77]
<i>Trichoderma reesei</i> ^f and <i>Candida molischiana</i> ^f	α -cellulose	Ethanol	Yield = 15 %	[78]
<i>Clostridium thermocellum</i> ^b and <i>Clostridium thermolacticum</i> ^b	Micro-crystallized cellulose (MCC)	Ethanol	Yield = 75 %	[79]
<i>Phlebia radiata</i> ^f and <i>Saccharomyces cerevisiae</i> ^f	Waste lignocellulose material	Ethanol	Productivity = 32.4 g/L in 30 days	[55]
<i>Acremonium cellulolyticus</i> ^f and <i>Saccharomyces cerevisiae</i> ^f	Solka-Floc (SF)	Ethanol	Concentration = 8.7–46.3 g/L	[80]
<i>Acetivibrio thermocellus</i> ^b and <i>Thermoclostridium stercorarium</i> ^b	Mixture of cellulose and xylan	Ethanol	Concentration = 40.4 mM	[81]

^a Values are reported as given in the original references; units and fermentation times therefore differ among studies and are not intended for direct quantitative comparison. ^b Bacterial species; ^f fungal species (including yeasts).

3. Review of Recent Modeling Approaches for CBP

Modeling in the context of CBP refers to the formulation of computational or mathematical representations that capture biological, chemical, and physical dynamics within a CBP system. Such models support understanding system behavior, optimizing process performance, and predicting outcomes under different operating conditions. Traditionally, bioreactor models have been classified as structured or unstructured and segregated or unsegregated, depending on whether intracellular composition and population heterogeneity are explicitly represented [82]. Building on these foundations and enabled by advances in computational power and data availability, recent CBP studies are often grouped more pragmatically into two broad categories: mechanistic models [83] and data-driven models [84], with hybrid approaches increasingly combining elements of both. In this contribution, mechanistic modeling refers to models that represent CBP through explicit causal process descriptions (e.g., mass balances, reaction/transport steps, growth and inhibition mecha-

nisms), whereas first-principles-based modeling denotes the subset derived predominantly from fundamental laws (conservation relations, thermodynamics, and transport) with minimal empirical “shortcut” fitting; thus, the terms overlap but are not strictly identical. Mechanistic models rely on established biological, chemical, or physical laws to represent microbial metabolism, enzyme kinetics, and reactor dynamics, including deterministic models (e.g., Monod kinetics, structured models, computational fluid dynamics (CFD)) and stochastic models for simulating system variability [83,85]. In contrast, data-driven models, including machine learning (ML) and statistical approaches such as RSM, infer relationships directly from experimental data without requiring detailed prior knowledge of system mechanisms [86]. Response surface methodology, though sometimes described as a modeling tool, is more accurately a statistical technique for process optimization that generates empirical (often polynomial) response approximations [87]. A third emerging category involves hybrid models that integrate first-principles structure with data-driven components to leverage the strengths of both approaches [88]. This framework guides the discussion below, highlighting applications, strengths, and limitations for biofuel process design and optimization.

Current CBP modeling efforts focus on improving system efficiency or yield [76,78] and predicting performance under varying process conditions [89]. Deterministic and stochastic models have been used to simulate microbial growth, enzyme production dynamics, substrate degradation, and product formation. Kinetic models (e.g., Monod-type and structured models) describe metabolic interactions and resource allocation of CBP-relevant microorganisms, including *Clostridium thermocellum* and engineered yeasts [90,91]. CFD and optimization techniques have also been applied to assess mixing, reactor performance, and process bottlenecks [73,85]. To improve predictive capability, multi-scale and hybrid approaches increasingly combine genome-scale metabolic (GEM) network reconstructions with process-level simulations to examine metabolic fluxes and energy trade-offs in CBP systems [1,2,76]. ML algorithms are also being explored to analyze experimental datasets and support in silico optimization of fermentation conditions, helping identify key parameters that affect yield and productivity and enabling targeted genetic and process engineering strategies [92,93].

Mathematical modeling has also emerged as a useful tool for optimizing microbial consortia in CBP. A co-culture of *C. phytofermentans*, *Saccharomyces cerevisiae* cdt-1, and *Candida molischiana* demonstrated long-term stability under controlled oxygen diffusion, producing 22 g/L ethanol from 100 g/L α -cellulose [77]. The addition of exogenous cellulases increased ethanol production, underscoring the importance of regulating enzyme activity. Computational models incorporating ecological theory and metabolic flux analysis could further clarify microbial interactions and support the design of stable and efficient CBP consortia [77,94]. Modeling can also support optimization of environmental conditions. In a study using *Trichoderma reesei* and *Candida molischiana*, temperature shifts from 30 °C to 50 °C after 36 h increased reducing sugar and glucose production by 95% and 70%, respectively [78]. Computational simulations could help refine such dynamic parameter adjustments by identifying optimal temperature and pH profiles for maximizing conversion of lignocellulosic substrates to ethanol [78,94].

The integration of thermophilic and cellulolytic bacteria, such as *Acetivibrio thermocellus* and *Thermoclostridium stercorarium*, has been shown to enhance cellulose degradation and ethanol fermentation [81]. Combined cellulase systems can yield nearly twice the ethanol of single-strain fermentations, and models incorporating metabolic pathways and enzyme expression dynamics can help optimize these interactions for improved efficiency and scalability [81]. Although empirical approaches (e.g., one-factor-at-a-time (OFAT) and trial-and-error experimentation) remain common, they often overlook interactions among

multiple variables and can lead to suboptimal outcomes [94]. Computational tools such as RSM and ANN provide more efficient alternatives by capturing interactions and optimizing process parameters with fewer experiments, supporting the transition from laboratory studies to industrial production [73,81,93].

Overall, CBP modeling has broadened beyond traditional kinetic and stoichiometric descriptions toward more integrated, multi-scale frameworks. While foundational models (e.g., Monod-based kinetics and structured growth simulations) remain important for describing microbial metabolism and enzyme dynamics, recent work increasingly emphasizes hybrid modeling that couples genome-scale reconstructions with process-level simulations [95,96]. Modeling of microbial consortia has likewise expanded, applying ecological theory and metabolic flux analysis to improve co-culture stability and productivity [97]. Advanced computational tools, including ML, RSM, and ANN, are being leveraged to optimize fermentation parameters and predict outcomes under complex, dynamic conditions [94]. Despite progress, challenges remain in accurately simulating lignocellulose degradation and multi-organism interactions at scale; nevertheless, tighter integration of modeling with experimental design continues to drive advances toward scalable, efficient CBP systems for industrial biofuel production [92].

3.1. Polynomial Models

Polynomial models are among the simplest mathematical approaches for describing bioprocess systems, including CBP. These models utilize polynomial equations to approximate relationships between variables such as substrate concentration, enzyme activity, microbial growth, and product formation. By fitting experimental data to polynomial functions, researchers can derive empirical models that capture general trends in the system behavior [98,99]. Polynomial models are particularly useful for their ease of implementation, low computational demands, and ability to represent nonlinear relationships over limited ranges of process parameters. They are often employed in the initial stages of process modeling or optimization, where a quick and approximate understanding of system behavior is sufficient [100].

Despite their simplicity, polynomial models have significant limitations when applied to complex bioprocesses like CBP. These models are inherently empirical and lack a mechanistic basis for representing the biological interactions within the system. As a result, they are often inadequate for accurately predicting process performance outside the range of experimental data used for model calibration [98,99]. For instance, polynomial models may fail to capture enzyme–substrate dynamics, metabolic shifts, or inhibition effects that are critical for CBP processes. Moreover, as the polynomial degree increases to improve fit, models become more susceptible to overfitting, reducing reliability and robustness for practical applications [93,101].

The study in [101] illustrates these limitations in a CBP context: higher-order polynomials failed to capture complex, nonlinear interactions, as reflected in negative R^2 values and F-statistics of 0 (p -values of 1), indicating overfitting and a lack of model significance. In that case, polynomial models offered limited predictive power and provided little insight into relationships among variables [101]. While linear models showed some explanatory capability, the overall performance suggests that alternative approaches may be required for CBP systems with strong nonlinearities and inhibition effects.

Accordingly, polynomial models are best viewed as preliminary tools for exploring trends rather than as reliable predictors under dynamic or industrially relevant conditions [100,101]. Their limited extrapolation capability and lack of mechanistic interpretability constrain their use for scenario analysis, optimization, and scale-up. Recent methodological development has therefore shifted toward more sophisticated frameworks, including

kinetic models [70,102], hybrid approaches [96,103], and multi-scale models [71,88], as well as machine learning methods that can better capture nonlinear, high-dimensional relationships [94,104].

3.2. Response Surface Methodology

Response surface methodology is a statistical and mathematical technique widely used for process optimization and empirical modeling of bioprocess systems, including CBP. Response surface methodology explores relationships between multiple independent variables and one or more response variables, providing a structured framework to optimize operating conditions [68,105]. By fitting experimental data to second-order polynomial models, it generates response surfaces that describe how parameters such as temperature, pH, substrate concentration, and inoculum size influence responses, including enzyme activity, microbial growth, and product yield. This approach is particularly useful for identifying operating optima while reducing experimental effort through systematic design of experiments (DoE) such as Box–Behnken or Central Composite Designs [87,106]. A key strength is its ability to evaluate interacting variables simultaneously, which is valuable for CBP, where enzyme synthesis, saccharification, and fermentation occur in one step [68,69]. Response surfaces and contour plots further support visualization of variable effects, enabling targeted process adjustments [100,106].

An RSM-based optimization of saccharification variables for ionic-liquid (triethylammonium hydrogen sulfate [TEA][HSO₄]) pretreated *Saccharum spontaneum* under one-pot consolidated bioprocessing (OPCB) showed substantial improvements in sugar yield and cellulose conversion [107]. Enzyme load, temperature, time, and pH were optimized, with temperature identified as the most influential factor; the study reported a 2.7-fold increase in sugar yield (up to 531.0 mg/g biomass) and 93.7% cellulose conversion, with good agreement between predicted and experimental values [69,107]. Response surface methodology was also applied to model and optimize CBP for bacterial alkaline phosphatase (ALP) production [105]. Using Plackett–Burman screening followed by a rotatable central composite design (RCCD), the effects of molasses, ammonium nitrate (NH₄)₂NO₃, and potassium chloride (KCl) were evaluated; the fitted quadratic model supported selection of optimal factor levels and led to a reported 94-fold increase in ALP production relative to the basal medium [105].

Using RSM, Selvakumar et al. [108] optimized ethanol production from *Manihot esculenta* Crantz YTP1 stems via CBP with *Cellulomonas fimi* and *Zymomonas mobilis*. Acid pretreatment (acetic, nitric, and mixed acids) was optimized for delignification, achieving 85% lignin removal and 87.45% cellulose release. Response surface methodology (in combination with ANN) was used to optimize pH, temperature, agitation, and time, resulting in maximum cellulase activity (11.63 IU/mL) and ethanol production (9.39 g/L). ANOVA indicated significant quadratic models with high R² values (0.9679 for cellulase and 0.9691 for ethanol), and optimal conditions (pH 5, 31 °C, 150 rpm, 24 h) were experimentally validated [108].

Despite these advantages, RSM remains an empirical, polynomial-based approach and may not capture complex biological mechanisms, nonlinear feedback, or strongly dynamic behavior outside the experimental design space [69,106]. This limits extrapolation and scenario analysis for CBP systems with time-varying inhibition, multi-organism interactions, or scale-dependent effects. Consequently, RSM is often most effective as an early-stage optimization tool or as a complementary component within more advanced kinetic/first-principles and machine-learning or hybrid modeling frameworks aimed at improving robustness and predictive accuracy [87,102,104].

3.3. Machine Learning-Based Modeling of CBP

Machine learning (ML)-based modeling has emerged as a powerful, data-driven approach for representing CBP dynamics. Techniques such as artificial neural networks (ANNs), decision trees, and support vector machines (SVMs) can analyze experimental datasets to identify patterns and relationships among process variables [84,109,110]. Unlike polynomial or kinetic models, many ML methods do not require explicit assumptions about system behavior, making them well-suited for the nonlinear and dynamic characteristics of CBP [102,111]. For example, ML models can predict biofuel yields, enzyme production rates, and substrate consumption from inputs such as temperature, pH, substrate concentration, and microbial growth conditions, often improving predictive performance over purely empirical approaches [112,113].

A key advantage of ML is its ability to handle high-dimensional, multivariate data and capture complex interactions among variables [109,113]. In CBP systems, where enzyme production, saccharification, and fermentation occur simultaneously, ML algorithms can integrate information from experimental data, omics measurements, and process monitoring [86,114]. Deep learning and ensemble methods can further improve performance as additional data become available [111,113,115]. However, ML models typically require large, high-quality datasets for training and validation, and limited interpretability (particularly for deep learning) can hinder mechanistic insight. To address these limitations, ML is increasingly combined with first-principles-based models to improve both prediction accuracy and process understanding, supporting more reliable optimization and control of CBP systems [109,112].

In CBP applications, the choice of learning paradigm influences both predictive performance and reliance on domain expertise. Supervised learning (e.g., SVMs and neural networks) is widely used to predict key outcomes (e.g., ethanol yield, enzyme activity) but depends on the availability and quality of labeled experimental data, which can be limited in CBP research. Unsupervised methods (e.g., clustering and dimensionality reduction such as PCA and t-SNE) are useful for exploring high-dimensional omics or process datasets without labels and for identifying latent structure that can guide hypothesis generation. Semi-supervised strategies provide a compromise when labeled data are scarce by leveraging larger pools of unlabeled data to improve robustness. Thus, selecting supervised, unsupervised, or semi-supervised learning reflects practical trade-offs between data availability and the degree of expert knowledge needed to develop, interpret, and validate models for CBP [114,116,117].

3.3.1. Regression Models

Regression models are a core subset of machine learning approaches used to predict CBP outcomes by relating input variables (e.g., substrate concentration, enzyme activity, pH, temperature) to responses such as biofuel yield, microbial growth, or product formation. Linear regression is often used for initial analyses, but because CBP behavior is typically nonlinear, more advanced formulations (e.g., polynomial, ridge, and lasso regression) are frequently applied to improve accuracy and reduce overfitting [70,109,118]. Beyond these, methods such as support vector regression (SVR), decision tree regression, and random forest/gradient-boosted regression have proven effective for capturing complex CBP relationships [103,110]. SVR uses kernel functions to represent nonlinear dependencies in higher-dimensional spaces [118], while tree-based methods can accommodate multicollinearity and variable interactions and often retain some interpretability. Collectively, these regression approaches can support the prediction of CBP performance and the identification of influential parameters for improving yields and efficiency [86].

In regression-based CBP modeling, data preprocessing is critical. Removing inconsistent measurements and outliers (e.g., guided by residual analysis) can narrow residual ranges, reduce RMSE/MSE, and improve R^2 values, thereby stabilizing fitted models [89,119]. Recent applications of Gaussian process regression (GPR) to ethanol production data indicate that, with appropriate kernel choices (e.g., Matérn-type or exponential kernels), GPR can outperform simple polynomial surrogates and other regression techniques, particularly for nonlinear responses across varying CBP conditions [111]. These studies also emphasize the importance of rigorous training/validation/test separation to obtain unbiased performance estimates on unseen CBP scenarios [89,119].

Despite strong in-range performance, regression models can degrade when extrapolating to untested operating conditions, and small or noisy datasets can lead to underfitting or overfitting. To improve generalization and practical usefulness, current developments increasingly integrate regression with complementary ML methods (e.g., neural networks) or embed regression components in hybrid frameworks that incorporate mechanistic constraints and domain knowledge [95,103,120]. Going forward, emphasis should be placed on improving robustness, scalability, and interpretability, particularly for deployment in broader operational scenarios, through better feature engineering, uncertainty-aware prediction, and mechanistically informed hybridization [112,120,121].

3.3.2. Neural Network Models

Neural network models have attracted significant attention for their ability to represent complex, nonlinear CBP behavior. Inspired by biological neural networks, these models consist of interconnected layers of neurons that transform input data into predictions [95,111]. Artificial neural networks are well-suited to CBP because they can learn relationships between inputs (e.g., enzyme activity, substrate concentration, microbial growth, and fermentation conditions) and outputs such as product yield and process efficiency [100]. When trained on experimental datasets, ANNs can capture nonlinearities that are difficult to describe with traditional regression or polynomial models [95,100]. More advanced architectures, including deep neural networks (DNNs), can incorporate many interacting variables and integrate large-scale inputs such as process monitoring data, omics data (e.g., genomics and metabolomics), and environmental conditions to improve predictive performance [95,115,122]. Training typically relies on backpropagation and related optimization algorithms to minimize prediction error [95], and approaches such as transfer learning can support model updates when new data become available.

Artificial neural network modeling has been applied to optimize cellulase activity and ethanol yield by selecting an appropriate network architecture [108]. In that study, multiple topologies were evaluated, and the best-performing model was a 4-8-1 topology using a hyperbolic tangent (Tanh) hidden-layer activation and a linear output layer. High predictive accuracy was reported, with R^2 values of 0.9950, 0.9960, and 0.9851 for training, validation, and test datasets (cellulase activity), and 0.9797, 0.9907, and 0.9804 for ethanol yield; overall R^2 values were 0.9908 (cellulase) and 0.9794 (ethanol) [108]. Compared with RSM, the ANN achieved lower RMSE and absolute average deviation (AAD%) values, and the optimal conditions corresponded to cellulase activity of 11.63 ± 1.23 IU/mL and ethanol yield of 9.39 ± 0.33 g/L [108].

Model performance is strongly influenced by dataset quality and encoding. Several studies report that carefully curated experimental datasets can outperform larger but weakly curated or synthetic datasets, as residual outliers and poorly represented variability can degrade generalization on independent test sets [89,111,119]. This emphasizes the importance of preprocessing, feature design, and rigorous separation of training, validation, and test data. Key limitations include sensitivity to noisy or limited data and limited

mechanistic interpretability (the “black-box” issue) [95,114]. Consequently, current efforts increasingly combine ANNs with first-principles structure and/or explainable AI (XAI) to improve transparency while retaining predictive accuracy [123,124]. Despite extensive experimental work on CBP, relatively few studies report explicit quantitative models (e.g., RSM, ANN, GPR), and these are summarized in Table 3.

Table 3. Modeling approaches applied to CBP systems ^a.

Modeling Approach	Microorganisms	Substrate	Bioproduct	Performance Metrics	Reference
RSM	<i>Hangateiclostridium thermocellum</i> KSMK1203 and consortium of <i>Cellulomonas fimi</i> MTCC 24 and <i>Zymomonas mobilis</i> MTCC 92	Pre-treated <i>Allium ascalonicum</i> leaves	Ethanol	$R^2 = 0.9397$	[35]
	<i>Cellulomonas fimi</i> MTCC 24 and <i>Zymomonas mobilis</i> MTCC 92	Thermo-chemo pretreated <i>Manihot esculenta</i> Crantz YTP1 stem	Cellulase	$R^2 = 0.9679$ RMSE = 0.7943	[108]
			Ethanol	$R^2 = 0.9691$ RMSE = 1.0526	
ANN	<i>Cellulomonas fimi</i> MTCC 24 and <i>Zymomonas mobilis</i> MTCC 92	Thermo-chemo pretreated <i>Manihot esculenta</i> Crantz YTP1 stem	Cellulase	$R^2 = 0.990$ RMSE = 0.5151	[108]
			Ethanol	$R^2 = 0.979$ RMSE = 0.6575	
	Mixed microbial consortia	Experimental and synthetic datasets with different lignocellulosic substrates	Ethanol	R = 0.918, MSE = 0.186 (val.; $n = 1615$) ^a R = 0.784, MSE = 0.568 (test; $n = 15$) ^a	[89]
GPR	Mixed microbial consortia	Experimental and synthetic datasets with different lignocellulosic substrates	Ethanol	R^2 up to 0.97; RMSE as low as 0.24 (best model) performance depends on kernel choice	[89]

^a n denotes the number of data points in the corresponding split: $n = 1615$ for the validation set and $n = 15$ for the test set.

3.4. Summary of the State of the Art in First-Principles and Data-Driven Modeling of CBP

State-of-the-art CBP modeling can be grouped into two main approaches: mechanistic models (including first-principles-based formulations) and data-driven models. Mechanistic models represent CBP through explicit causal descriptions of the underlying biological, chemical, and physical processes and are often implemented as deterministic mass-balance and kinetic frameworks. They incorporate representations of enzyme–substrate interactions, microbial growth, and metabolic fluxes to predict system behavior [70,96,103]. Common examples include Monod-type growth kinetics [83], Michaelis–Menten enzyme models [125], and genome-scale metabolic network reconstructions [126]. Because of their mechanistic structure, these models can provide interpretable insight into enzyme activity, substrate hydrolysis, and product formation and support parameter optimization grounded in system dynamics [92,127]. However, their application is often constrained by complexity and extensive parameterization requirements: accuracy depends on reliable kinetic constants, reaction pathways, and system constraints that can be difficult to obtain for CBP involving multiple microbial species and variable environmental factors [127]. In addition, mechanistic formulations may struggle to represent stochasticity and strongly time-varying behavior, particularly at industrial scales where feedstock composition and operating conditions can change substantially. These limitations have contributed to growing interest in data-driven approaches [96,103].

Data-driven models, including regression methods, neural networks, support vector machines, and ensemble learning techniques, have shown strong potential for representing and optimizing CBP systems. They are well-suited to high-dimensional, nonlinear, and multivariate datasets commonly encountered in CBP research [89]. Machine learning

methods can predict biofuel yields, enzyme production rates, and substrate utilization from historical data, making them useful for optimization and scenario analysis [84,86]. Advances in deep learning and hybrid frameworks also enable the integration of omics data, process-monitoring information, and experimental results to improve predictive accuracy. In this way, data-driven approaches can complement first-principles modeling, particularly under variable operating conditions [96,115].

A persistent challenge for bioprocess modeling is the systematic collection and harmonization of heterogeneous experimental data, as highlighted by [84]. Consolidated bioprocessing studies often employ highly variable conditions, limiting the integration of results for broader analysis. More detailed and standardized reporting is therefore needed to improve data management and predictive performance. For example, microbial consortia are frequently encoded only coarsely, without specifying whether strains are wild-type or engineered or reporting the relative abundance of consortium members. Such omissions can increase prediction error because strain identity and strain ratios can strongly affect ethanol productivity across substrates. Similarly, pretreatment strategies, which influence lignin removal, cellulose accessibility, and ultimately ethanol yield, are often recorded qualitatively or not encoded in modeling datasets. Including richer microbial descriptors (e.g., strain type and consortium composition) together with quantitative pretreatment information will be essential to improve the accuracy, interpretability, and transferability of CBP models.

Despite their advantages, data-driven methods still require large, high-quality datasets for training, and the black-box nature of some algorithms can limit interpretability. Recent progress in explainable AI and hybrid modeling is helping bridge this gap by linking machine learning predictions to mechanistic understanding [96,128]. By combining predictive power with mechanistic interpretability, hybrid models are emerging as a promising direction for CBP, supporting accurate prediction while enabling biological insight, diagnosing inefficiencies, and optimizing process conditions [86,127].

A comparative table summarizing the advantages and disadvantages of first-principles and data-driven modeling approaches in CBP, using the qualitative scoring system (++ very strong, + strong, 0 neutral, – weak, – – very weak) is shown in Table 4.

Table 4. Comparison of first-principles vs. data-driven modeling approaches in CBP.

Criteria	First-Principles-Based Models	Data-Driven Models
Interpretability and mechanistic insight	++	–
Performance under limited data availability	+	– –
Predictive accuracy under known conditions	++	+
Ability to update with new experimental results	–	++
Computational complexity	–	0
Handling multivariate interactions	0	++
Suitability for early-stage research	++	0
Need for system understanding	++	–
Ease of implementation	–	+

Interpretability is a central distinction between first-principles-based and data-driven modeling in CBP. First-principles models are typically transparent because they are derived from established biochemical and physical laws, providing direct insight into how variables such as enzyme kinetics and substrate concentrations shape system outcomes [129,130]. This transparency makes them useful for diagnosing performance issues and supporting rational process optimization [131]. By contrast, many data-driven methods, especially deep learning, are often treated as “black boxes” because their internal decision logic is difficult to explain. Although predictive accuracy can be high, limited interpretability may hinder biological insight and obscure the root causes of observed behavior [119,131]. Another difference concerns reliance on prior knowledge and adaptability. Mechanistic

models require detailed inputs (e.g., kinetic rates, metabolic pathways, stoichiometry), which can be difficult to obtain for complex or poorly characterized CBP systems. Data-driven approaches reduce this dependence by learning patterns directly from empirical data, making them well-suited for exploratory analyses when mechanistic insight is incomplete [129,130,132]. They are also easier to update: retraining on new data can improve performance over time, whereas first-principles formulations often need reparameterization (and sometimes reformulation) as conditions change [133]. Combined with their capacity to represent nonlinear, high-dimensional interactions, this adaptability can be advantageous in dynamic and heterogeneous CBP environments [134].

The two approaches also differ in data requirements and sensitivity to data quality. First-principles-based models can produce reliable predictions from relatively small datasets when key mechanistic parameters are well defined [129], and their grounding in biological principles can help maintain performance in data-sparse settings [119,132]. In contrast, data-driven models usually benefit from large, high-quality datasets; when data are limited, imbalanced, or noisy, they are more prone to underfitting, overfitting, and reduced generalization unless careful preprocessing, feature selection, and regularization are applied [135]. For complex multivariate inputs, such as omics data and time-varying process variables, data-driven architectures are often better suited because they can manage high-dimensional spaces and uncover hidden structure in large-scale datasets, supporting CBP applications that integrate genomics, proteomics, and environmental monitoring [119,134]. While mechanistic models can incorporate such information, this typically requires extensive customization and simplifying assumptions that may reduce fidelity. Parameter tuning can also be more labor-intensive in first-principles-based models due to manual estimation of many interdependent variables [133,136], whereas data-driven approaches rely more on automated optimization routines (often at higher computational cost). Both approaches, however, can be affected by missing or incomplete data: data-driven models may require imputation strategies, while mechanistic models are constrained by the completeness of available mechanistic knowledge [136].

Beyond prediction, first-principles-based models are valuable for elucidating biological mechanisms and cause-and-effect relationships, supporting hypothesis-driven research [137]. Their scalability to industrial conditions, however, can be limited when underlying assumptions break down in large-scale and highly variable environments [131,135]. Data-driven models are often more amenable to industrial deployment because they can adapt to operational data streams and improve as new information is collected, albeit with higher computational demands, particularly for deep learning, during training and validation [132,137]. In comparison, mechanistic models are often less resource-intensive, especially when simplified kinetics or steady-state assumptions are used. A summary of different modeling approaches applicable to CBP research, highlighting their potential applications, is shown in Table 5.

Table 5. Characterization of modeling approaches relevant to CBP research with potential applications.

Model Type	Description and Potential Application in CBP
Deterministic models (including kinetic/structured models)	Use ordinary differential equations (ODEs) to simulate microbial growth, enzyme production, substrate degradation, and product formation. This includes Monod-type and structured kinetic formulations, and can be extended to represent co-cultures and substrate competition in CBP systems.
Stochastic models	Incorporating random variables to account for biological noise and fluctuations in microbial behavior. Useful for microbial consortia, feedstock composition variability, and uncertain process conditions.
Computational Fluid Dynamics (CFD)	Simulation of reactor hydrodynamics, mixing patterns, mass transfer, and heat exchange. Can be used to optimize large-scale CBP bioreactors and reduce process bottlenecks.

Table 5. Cont.

Model Type	Description and Potential Application in CBP
Multi-scale modeling	Integration of genome-scale metabolic models with process-level dynamics to understand intracellular fluxes and system behavior at different scales. Potentially useful to link metabolic engineering with reactor performance in CBP.
Hybrid models	Combination of mechanistic models with data-driven approaches like support vector machines or random forests to improve prediction accuracy and interpretability. Hybrid models are useful for predicting CBP outcomes with novel feedstocks.
Reinforcement learning models	Utilizing reward-based algorithms to optimize process parameters dynamically. Can be applied to adaptive control of CBP processes, e.g., feeding strategies or environmental adjustments.
Digital twins/soft sensors/Model Predictive Control (MPC)	Integration of mechanistic and data-driven models with online measurements for state estimation and real-time optimization/control of CBP (e.g., inhibition mitigation, drift detection, high-solids operation).
Evolutionary algorithms	Optimization techniques inspired by natural selection. Can be used to optimize multi-objective CBP process parameters, microbial community composition, or pathway design.

Building on this complementarity, recent work increasingly emphasizes deployment-oriented hybrid frameworks for real-time monitoring and control, often framed as digital twins.

3.4.1. Digital Twins and Bioprocessing 4.0

Recent “Bioprocessing 4.0” efforts emphasize digital twins that couple mechanistic balances with data-driven components to enable real-time monitoring, optimization, and control. In this context, digital twins typically integrate first-principles structure (stoichiometry, transport, and constraints) with soft sensors and learned models calibrated on process data, improving adaptability under variable conditions [133]. Hybrid-modeling perspectives increasingly position these twins as the deployment target for advanced monitoring and model-predictive control, particularly when paired with online analytics (e.g., spectroscopy) and uncertainty-aware predictions [95]. Such frameworks are relevant to CBP because they can support operation at high-solids loadings, mitigate inhibition effects, and detect consortium drift through continuous model updating and state estimation [138].

Within this digital-twin context, physics-informed ML, particularly physics-informed neural networks (PINNs), is gaining traction as a compact way to strengthen hybrid CBP models by embedding first-principles structure (e.g., mass balances and known kinetics) directly into the learning objective. By constraining neural networks with mechanistic relationships, PINNs can improve generalization and yield physically consistent state estimates useful for real-time optimization and control. Recent work has demonstrated PINN-enabled model-predictive control (MPC) in bioprocess settings, highlighting how physics-informed models can support closed-loop operation under uncertainty and changing process conditions [139]. In parallel, hybrid modeling has evolved from shallow grey-box formulations toward deep, physics-regularized architectures explicitly designed for monitoring and advanced control applications [95,133]. More broadly, panoramic reviews of hybrid fusion for process monitoring emphasize that combining mechanistic constraints with data-driven observers can improve robustness in noisy, industrial environments, which is directly relevant to CBP scale-up and soft-sensing deployment [140].

3.4.2. Uncertainty Quantification (UQ)

For CBP modeling and scale-up, uncertainty-aware predictions are essential because parameter variability, feedstock heterogeneity, and measurement noise can materially change optimal operating decisions. Uncertainty quantification can be introduced through confidence/credible intervals from Bayesian calibration and inference, which quantify

uncertainty in parameters and propagate it to model outputs [141]. Gaussian process regression is particularly attractive because it provides both a mean prediction and an associated predictive variance, enabling uncertainty bounds for soft sensors and decision support [142]. In control, these uncertainty estimates can be embedded in (stochastic/robust) model-predictive control to tighten constraints, reduce violation risk, and maintain performance under disturbances and model mismatch [142,143]. Overall, integrating UQ improves the reliability of optimization and MPC policies and supports safer translation of CBP models from laboratory conditions to variable industrial operation [141,143].

3.4.3. Key Points

Across CBP modeling, simple polynomial models are most useful for rapid, low-cost trend fitting within narrow experimental ranges, but they generalize poorly and offer limited physical insight. Response surface methodology provides an efficient design-of-experiments framework to identify local optima and variable interactions, yet its second-order surrogate form can miss strongly nonlinear or time-varying behavior outside the design space. Regression and tree-based ML models can capture nonlinear mappings and rank influential factors, but their reliability typically depends on dataset quality and coverage, particularly when extrapolating to new feedstocks or operating windows. Neural networks further improve flexibility for high-dimensional and highly coupled CBP data, although interpretability and robustness under sparse or shifting data remain challenges. First-principles models remain essential for mechanistic understanding and scale-aware reasoning but may under-represent complex inhibition, consortium dynamics, and multi-scale couplings without additional structure. Taken together, these trade-offs motivate hybrid frameworks that embed conservation laws and physicochemical constraints while learning residual kinetics from data, especially when paired with online sensing and uncertainty-aware control to enable adaptive operation under industrially realistic variability. The overall modeling workflow is summarized in Figure 2, and the main takeaways are listed below.

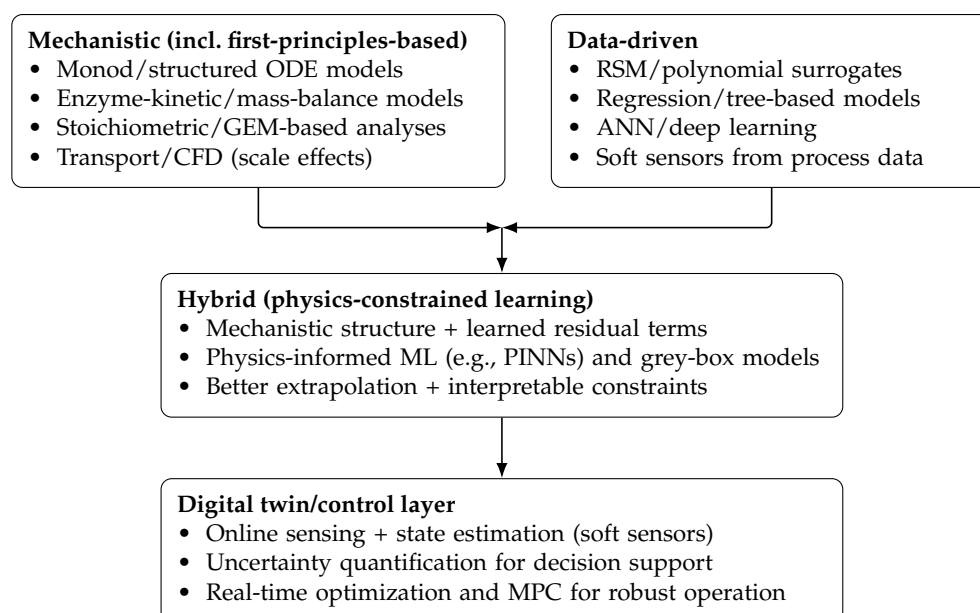


Figure 2. Modeling taxonomy for CBP: mechanistic and data-driven approaches motivate hybrid frameworks, which enable digital-twin deployment for monitoring, optimization, and control.

4. Summary and Conclusions

Consolidated bioprocessing offers a cost-effective and sustainable route to valorize lignocellulosic biomass by integrating enzyme production, hydrolysis, and fermentation into a single operation, thereby reducing unit operations compared to SHF/SSF. This review focused on lignocellulosic CBP for ethanol and related products and surveyed modeling approaches spanning first-principles methods (kinetic and stoichiometric balances) and data-driven or surrogate models (e.g., RSM-based response surfaces, regularized regression, and neural networks). Across enzyme synthesis, hydrolysis, and fermentation, persistent bottlenecks include incomplete pentose co-utilization, product/solids inhibition at high loadings, scale-dependent mass-transfer limitations, and limited strain robustness under industrial stressors. Mechanistic models support interpretability and scale-aware reasoning but can struggle with strongly nonlinear, time-varying couplings among hydrolysis, transport, and metabolism. Data-driven models capture complex mappings and support soft sensing, yet often lack physical guarantees and robust extrapolation across feedstocks and operating windows. This complementarity motivates hybrid modeling, where mechanistic structure enforces stoichiometry, thermodynamic feasibility, and transport constraints, while learned components capture residual kinetics, inhibition cross-terms, and context-specific regulation with uncertainty estimates suitable for control and scale-up. Reported case studies indicate improved predictive accuracy, tighter operation, and faster design cycles relative to single-paradigm baselines.

Recent work has advanced bioprocess modeling by augmenting kinetic/stoichiometric reconstructions with machine-learning surrogates, transfer learning, and active learning to accelerate parameter estimation, soft sensing, and decision support. However, key gaps remain: many studies are laboratory-bound, multi-scale data (molecular to bioreactor) are rarely integrated coherently, and multi-omics (genomics, transcriptomics, proteomics, metabolomics) is underutilized for parameterization and validation. Future progress will benefit from combining real-time spectroscopy and soft sensors (for sugars, inhibitors, and redox/metabolic state) with robust model-predictive control to mitigate catabolite repression, product/solids inhibition, and consortium drift. Priority directions include physics-constrained hybrid frameworks trained on curated multi-omics and time-series datasets (supported by standardized protocols and open, well-annotated data spanning lab-to-pilot scales for reproducible benchmarking) and integration of digital twins with techno-economic and life-cycle assessments to target cost- and risk-dominant levers (e.g., solids handling, uptime, in situ product removal). Ultimately, coupling mechanistic interpretability with data-driven predictive performance, and closing the loop through online sensing and control, will be pivotal for translating CBP from proof-of-concept to reliable, large-scale production of ethanol and higher-value bioproducts.

Author Contributions: Conceptualization, M.K.Y. and D.S.; formal analysis, M.K.Y. and D.S.; investigation, M.K.Y.; supervision, D.S.; original draft writing, M.K.Y.; review—writing and revision, M.K.Y. and D.S.; final review, D.S. All authors have read and agreed to the published version of the manuscript.

Funding: This work is partly supported through a scholarship awarded to the first author by the German Academic Exchange Service (DAAD) for his Ph.D. study at the Chair of Dynamics and Control, UDE, Germany.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding author.

Acknowledgments: Support by the Open Access Publication Fund of the University of Duisburg-Essen and German Academic Exchange Service (DAAD) is gratefully acknowledged.

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

References

1. Sebestyén, V. Renewable and Sustainable Energy Reviews: Environmental impact networks of renewable energy power plants. *Renew. Sustain. Energy Rev.* **2021**, *151*, 111626. [\[CrossRef\]](#)
2. Nastasi, B.; Markovska, N.; Puksec, T.; Duić, N.; Foley, A. Renewable and sustainable energy challenges to face for the achievement of Sustainable Development Goals. *Renew. Sustain. Energy Rev.* **2022**, *157*, 112071. [\[CrossRef\]](#)
3. Li, Z.; Waghmare, P.R.; Dijkhuizen, L.; Meng, X.; Liu, W. Research advances on the consolidated bioprocessing of lignocellulosic biomass. *Eng. Microbiol.* **2024**, *4*, 100139. [\[CrossRef\]](#) [\[PubMed\]](#)
4. Bertrand, E.; Dussap, C.G. First Generation Bioethanol: Fundamentals—Definition, History, Global Production, Evolution. In *Liquid Biofuels: Bioethanol*; Springer: Berlin/Heidelberg, Germany, 2022; pp. 1–12.
5. De Almeida, M.A.; Colombo, R. Production chain of first-generation sugarcane bioethanol: characterization and value-added application of wastes. *BioEnergy Res.* **2023**, *16*, 924–939. [\[CrossRef\]](#)
6. Althuri, A.; Gujjala, L.K.S.; Banerjee, R. Partially consolidated bioprocessing of mixed lignocellulosic feedstocks for ethanol production. *Bioresour. Technol.* **2017**, *245*, 530–539. [\[CrossRef\]](#)
7. Roukas, T.; Kotzekidou, P. From food industry wastes to second generation bioethanol: A review. *Rev. Environ. Sci. Bio/Technol.* **2022**, *21*, 299–329. [\[CrossRef\]](#)
8. Wei Kit Chin, D.; Lim, S.; Pang, Y.L.; Lam, M.K. Fundamental review of organosolv pretreatment and its challenges in emerging consolidated bioprocessing. *Biofuels Bioprod. Biorefining* **2020**, *14*, 808–829. [\[CrossRef\]](#)
9. Olguin-Maciel, E.; Singh, A.; Chable-Villacis, R.; Tapia-Tussell, R.; Ruiz, H.A. Consolidated Bioprocessing, an Innovative Strategy towards Sustainability for Biofuels Production from Crop Residues: An Overview. *Agronomy* **2020**, *10*, 1834. [\[CrossRef\]](#)
10. Guo, J.; Liu, D.; Xu, Y. Perspectives and advances in consolidated bioprocessing strategies for lignin valorization. *Sustain. Energy Fuels* **2024**, *8*, 1153–1184. [\[CrossRef\]](#)
11. Moonsamy, T.A.; Mandegari, M.; Farzad, S.; Görgens, J.F. A new insight into integrated first and second-generation bioethanol production from sugarcane. *Ind. Crop. Prod.* **2022**, *188*, 115675. [\[CrossRef\]](#)
12. Singhanian, R.R.; Patel, A.K.; Singh, A.; Haldar, D.; Soam, S.; Chen, C.W.; Tsai, M.L.; Dong, C.D. Consolidated bioprocessing of lignocellulosic biomass: Technological advances and challenges. *Bioresour. Technol.* **2022**, *354*, 127153. [\[CrossRef\]](#) [\[PubMed\]](#)
13. Periyasamy, S.; Beula Isabel, J.; Kavitha, S.; Karthik, V.; Mohamed, B.A.; Gizaw, D.G.; Sivashanmugam, P.; Aminabhavi, T.M. Recent advances in consolidated bioprocessing for conversion of lignocellulosic biomass into bioethanol—A review. *Chem. Eng. J.* **2023**, *453*, 139783. [\[CrossRef\]](#)
14. Re, A.; Mazzoli, R. Current progress on engineering microbial strains and consortia for production of cellulosic butanol through consolidated bioprocessing. *Microb. Biotechnol.* **2023**, *16*, 238–261. [\[CrossRef\]](#) [\[PubMed\]](#)
15. Gupte, A.P.; Di Vita, N.; Myburgh, M.W.; Cripwell, R.A.; Basaglia, M.; van Zyl, W.H.; Viljoen-Bloom, M.; Casella, S.; Favaro, L. Consolidated bioprocessing of the organic fraction of municipal solid waste into bioethanol. *Energy Convers. Manag.* **2024**, *302*, 118105. [\[CrossRef\]](#)
16. Page, M.J.; McKenzie, J.E.; Bossuyt, P.M.; Boutron, I.; Hoffmann, T.C.; Mulrow, C.D.; Shamseer, L.; Tetzlaff, J.M.; Akl, E.A.; Brennan, S.E.; et al. The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ* **2021**, *372*, n71. [\[CrossRef\]](#)
17. Korsá, G.; Konwarh, R.; Masi, C.; Ayele, A.; Haile, S. Microbial cellulase production and its potential application for textile industries. *Ann. Microbiol.* **2023**, *73*, 13. [\[CrossRef\]](#)
18. Ma, X.; Li, S.; Tong, X.; Liu, K. An overview on the current status and future prospects in *Aspergillus* cellulase production. *Environ. Res.* **2024**, *244*, 117866. [\[CrossRef\]](#)
19. CareerPower. Anaerobic Respiration: Definition, Equation and Examples. 2024. Available online: <https://www.careerpower.in/school/biology/anaerobic-respiration> (accessed on 26 September 2025).
20. Kumar, V.; Ahluwalia, V.; Saran, S.; Kumar, J.; Patel, A.K.; Singhanian, R.R. Recent developments on solid-state fermentation for production of microbial secondary metabolites: Challenges and solutions. *Bioresour. Technol.* **2021**, *323*, 124566. [\[CrossRef\]](#)
21. Althuri, A.; Mohan, S.V. Sequential and consolidated bioprocessing of biogenic municipal solid waste: A strategic pairing of thermophilic anaerobe and mesophilic microaerobe for ethanol production. *Bioresour. Technol.* **2020**, *308*, 123260. [\[CrossRef\]](#)
22. Ramos, M.D.; Sandri, J.P.; Claes, A.; Carvalho, B.T.; Thevelein, J.M.; Zangirolami, T.C.; Milessi, T.S. Effective application of immobilized second generation industrial *Saccharomyces cerevisiae* strain on consolidated bioprocessing. *New Biotechnol.* **2023**, *78*, 153–161. [\[CrossRef\]](#)

23. Wen, Z.; Ledesma-Amaro, R.; Lu, M.; Jin, M.; Yang, S. Metabolic engineering of *Clostridium cellulovorans* to improve butanol production by consolidated bioprocessing. *ACS Synth. Biol.* **2020**, *9*, 304–315. [\[CrossRef\]](#)
24. Bhardwaj, N.; Kumar, B.; Agrawal, K.; Verma, P. Current perspective on production and applications of microbial cellulases: A review. *Bioresour. Bioprocess.* **2021**, *8*, 95. [\[CrossRef\]](#) [\[PubMed\]](#)
25. Minnaar, L.; den Haan, R. Engineering natural isolates of *Saccharomyces cerevisiae* for consolidated bioprocessing of cellulosic feedstocks. *Appl. Microbiol. Biotechnol.* **2023**, *107*, 7013–7028. [\[CrossRef\]](#) [\[PubMed\]](#)
26. Bhati, N.; Shreya; Sharma, A.K. Cost-effective cellulase production, improvement strategies, and future challenges. *J. Food Process Eng.* **2021**, *44*, e13623. [\[CrossRef\]](#)
27. Weimer, P.J. Degradation of cellulose and hemicellulose by ruminal microorganisms. *Microorganisms* **2022**, *10*, 2345. [\[CrossRef\]](#)
28. Balla, A.; Silini, A.; Cherif-Silini, H.; Bouket, A.C.; Boudechicha, A.; Luptakova, L.; Alenezi, F.N.; Belbahri, L. Screening of cellulolytic bacteria from various ecosystems and their cellulases production under multi-stress conditions. *Catalysts* **2022**, *12*, 769. [\[CrossRef\]](#)
29. Fu, R.; Han, L.; Li, Q.; Li, Z.; Dai, Y.; Leng, J. Studies on the concerted interaction of microbes in the gastrointestinal tract of ruminants on lignocellulose and its degradation mechanism. *Front. Microbiol.* **2025**, *16*, 1554271. [\[CrossRef\]](#)
30. Wen, Z.; Li, Q.; Liu, J.; Jin, M.; Yang, S. Consolidated bioprocessing for butanol production of cellulolytic *Clostridia*: Development and optimization. *Microb. Biotechnol.* **2020**, *13*, 410–422. [\[CrossRef\]](#)
31. Singh, L.K.; Chaudhary, G. *Advances in Biofeedstocks and Biofuels, Biofeedstocks and Their Processing*; John Wiley & Sons: Hoboken, NJ, USA, 2016.
32. Jiang, Y.; Lv, Y.; Wu, R.; Lu, J.; Dong, W.; Zhou, J.; Zhang, W.; Xin, F.; Jiang, M. Consolidated bioprocessing performance of a two-species microbial consortium for butanol production from lignocellulosic biomass. *Biotechnol. Bioeng.* **2020**, *117*, 2985–2995. [\[CrossRef\]](#)
33. Gao, H.; Li, M.; Yang, L.; Jiang, Y.; Jiang, W.; Yu, Z.; Zhang, W.; Xin, F.; Jiang, M. Spatial niche construction of a consortium-based consolidated bioprocessing system. *Green Chem.* **2022**, *24*, 7941–7950. [\[CrossRef\]](#)
34. Srivastava, M.; Srivastava, N.; Ramteke, P.W.; Mishra, P.K. *Approaches to Enhance Industrial Production of Fungal Cellulases*; Springer: Berlin/Heidelberg, Germany, 2019; p. 200.
35. Kavitha, S.; Gajendran, T.; Saranya, K.; Selvakumar, P.; Manivasagan, V.; Jeevitha, S. An insight—A statistical investigation of consolidated bioprocessing of *Allium ascalonicum* leaves to ethanol using *Hangateiclostridium thermocellum* KSMK1203 and synthetic consortium. *Renew. Energy* **2022**, *187*, 403–416. [\[CrossRef\]](#)
36. Fang, H.; Deng, Y.; Pan, Y.; Li, C.; Yu, L. Distributive and collaborative push-and-pull in an artificial microbial consortium for improved consolidated bioprocessing. *AIChE J.* **2022**, *68*, e17844. [\[CrossRef\]](#)
37. Shahab, R.L.; Luterbacher, J.S.; Brethauer, S.; Studer, M.H. Consolidated bioprocessing of lignocellulosic biomass to lactic acid by a synthetic fungal-bacterial consortium. *Biotechnol. Bioeng.* **2018**, *115*, 1207–1215. [\[CrossRef\]](#)
38. Sharma, J.; Kumar, V.; Prasad, R.; Gaur, N.A. Engineering of *Saccharomyces cerevisiae* as a consolidated bioprocessing host to produce cellulosic ethanol: Recent advancements and current challenges. *Biotechnol. Adv.* **2022**, *56*, 107925. [\[CrossRef\]](#) [\[PubMed\]](#)
39. Cagide, C.; Castro-Sowinski, S. Technological and biochemical features of lignin-degrading enzymes: A brief review. *Environ. Sustain.* **2020**, *3*, 371–389. [\[CrossRef\]](#)
40. Atiwesh, G.; Parrish, C.C.; Banoub, J.; Le, T.A.T. Lignin degradation by microorganisms: A review. *Biotechnol. Prog.* **2022**, *38*, e3226. [\[CrossRef\]](#) [\[PubMed\]](#)
41. Schlembach, I.; Hosseinpour Tehrani, H.; Blank, L.M.; Büchs, J.; Wierckx, N.; Regestein, L.; Rosenbaum, M.A. Consolidated bioprocessing of cellulose to itaconic acid by a co-culture of *Trichoderma reesei* and *Ustilago maydis*. *Biotechnol. Biofuels* **2020**, *13*, 207. [\[CrossRef\]](#) [\[PubMed\]](#)
42. Cunha, J.T.; Romani, A.; Inokuma, K.; Johansson, B.; Hasunuma, T.; Kondo, A.; Domingues, L. Consolidated bioprocessing of corn cob-derived hemicellulose: engineered industrial *Saccharomyces cerevisiae* as efficient whole cell biocatalysts. *Biotechnol. Biofuels* **2020**, *13*, 138. [\[CrossRef\]](#)
43. Yan, Q.; Fong, S.S. Challenges and advances for genetic engineering of non-model bacteria and uses in consolidated bioprocessing. *Front. Microbiol.* **2017**, *8*, 2060. [\[CrossRef\]](#)
44. Banner, A.; Toogood, H.S.; Scrutton, N.S. Consolidated bioprocessing: Synthetic biology routes to fuels and fine chemicals. *Microorganisms* **2021**, *9*, 1079. [\[CrossRef\]](#)
45. Amoah, J.; Ishizue, N.; Ishizaki, M.; Yasuda, M.; Takahashi, K.; Ninomiya, K.; Yamada, R.; Kondo, A.; Ogino, C. Development and evaluation of consolidated bioprocessing yeast for ethanol production from ionic liquid-pretreated bagasse. *Bioresour. Technol.* **2017**, *245*, 1413–1420. [\[CrossRef\]](#)
46. Zeng, M.; Pan, X. Insights into solid acid catalysts for efficient cellulose hydrolysis to glucose: Progress, challenges, and future opportunities. *Catal. Rev.* **2022**, *64*, 445–490. [\[CrossRef\]](#)
47. Sari, N.K.; Ernawati, D.; Sari, K.N. Optimization of Glucose from *Saccharomyces cerevisiae* Liquid Waste Using the Acid Hydrolysis Process. In *Nusantara Science and Technology Proceedings*; Galaxy Science: Malang, Indonesia, 2024; pp. 13–17.

48. Lu, J.; Lv, Y.; Jiang, Y.; Wu, M.; Xu, B.; Zhang, W.; Zhou, J.; Dong, W.; Xin, F.; Jiang, M. Consolidated Bioprocessing of Hemicellulose-Enriched Lignocellulose to Succinic Acid through a Microbial Cocultivation System. *Acs Sustain. Chem. Eng.* **2019**, *8*, 9035–9045. [\[CrossRef\]](#)
49. Liu, Y.J.; Li, B.; Feng, Y.; Cui, Q. Consolidated bio-saccharification: Leading lignocellulose bioconversion into the real world. *Biotechnol. Adv.* **2020**, *40*, 107535. [\[CrossRef\]](#)
50. Li, J.; Chen, B.; Gu, S.; Zhao, Z.; Liu, Q.; Sun, T.; Zhang, Y.; Wu, T.; Liu, D.; Sun, W.; et al. Coordination of consolidated bioprocessing technology and carbon dioxide fixation to produce malic acid directly from plant biomass in *Myceliophthora thermophila*. *Biotechnol. Biofuels* **2021**, *14*, 186. [\[CrossRef\]](#) [\[PubMed\]](#)
51. Dempfle, D.; Kröcher, O.; Studer, M.H.P. Techno-economic assessment of bioethanol production from lignocellulose by consortium-based consolidated bioprocessing at industrial scale. *New Biotechnol.* **2021**, *65*, 53–60. [\[CrossRef\]](#)
52. He, F.; Chen, J.; Gong, Z.; Xu, Q.; Yue, W.; Xie, H. Dissolution pretreatment of cellulose by using levulinic acid-based protic ionic liquids towards enhanced enzymatic hydrolysis. *Carbohydr. Polym.* **2021**, *269*, 118271. [\[CrossRef\]](#)
53. Kumar, V.; Fox, B.G.; Takasuka, T.E. Consolidated bioprocessing of plant biomass to polyhydroxyalkanoate by co-culture of *Streptomyces* sp. SirexAA-E and *Priestia megaterium*. *Bioresour. Technol.* **2023**, *376*, 128934. [\[CrossRef\]](#)
54. Tsai, S.L.; Sun, Q.; Chen, W. Advances in consolidated bioprocessing using synthetic cellulosomes. *Curr. Opin. Biotechnol.* **2022**, *78*, 102840. [\[CrossRef\]](#)
55. Mattila, H.; Kačar, D.; Mali, T.; Lundell, T. Lignocellulose bioconversion to ethanol by a fungal single-step consolidated method tested with waste substrates and co-culture experiments. *AIMS Energy* **2018**, *6*, 866–879. [\[CrossRef\]](#)
56. Zhang, Y.W.; Yang, J.J.; Qian, F.H.; Sutton, K.B.; Hjort, C.; Wu, W.P.; Jiang, Y.; Yang, S. Engineering a xylose fermenting yeast for lignocellulosic ethanol production. *Nat. Chem. Biol.* **2025**, *21*, 443–450. [\[CrossRef\]](#)
57. Mazzoli, R. Current progress in production of building-block organic acids by consolidated bioprocessing of lignocellulose. *Fermentation* **2021**, *7*, 248. [\[CrossRef\]](#)
58. Argyros, D.A.; Tripathi, S.A.; Barrett, T.F.; Rogers, S.R.; Feinberg, L.F.; Olson, D.G.; Foden, J.M.; Miller, B.B.; Lynd, L.R.; Hogsett, D.A.; et al. High Ethanol Titters from Cellulose by Using Metabolically Engineered Thermophilic, Anaerobic Microbes. *Appl. Environ. Microbiol.* **2011**, *77*, 8288–8294. [\[CrossRef\]](#) [\[PubMed\]](#)
59. Sharma, A.; Saini, H.; Thakur, B.; Soni, R.; Soni, S.K. Consolidated bioprocessing of biodegradable municipal solid waste for transformation into biofertilizer formulations. *Biomass Convers. Biorefinery* **2024**, *14*, 20923–20937. [\[CrossRef\]](#)
60. Qiu, Y.; Wu, M.; Bao, H.; Liu, W.; Shen, Y. Engineering of *Saccharomyces cerevisiae* for co-fermentation of glucose and xylose: Current state and perspectives. *Eng. Microbiol.* **2023**, *3*, 100084. [\[CrossRef\]](#) [\[PubMed\]](#)
61. Beluhan, S.; Mihajlovski, K.; Šantek, B.; Ivančić Šantek, M. The Production of Bioethanol from Lignocellulosic Biomass: Pretreatment Methods, Fermentation, and Downstream Processing. *Energies* **2023**, *16*, 7003. [\[CrossRef\]](#)
62. Dey, P.; Pal, P.; Kevin, J.D.; Das, D.B. Lignocellulosic bioethanol production: Prospects of emerging membrane technologies to improve the process—A critical review. *Rev. Chem. Eng.* **2020**, *36*, 333–367. [\[CrossRef\]](#)
63. Panagopoulos, V.; Boura, K.; Dima, A.; Karabagias, I.K.; Bosnea, L.; Nigam, P.S.; Kanellaki, M.; Koutinas, A.A. Consolidated bioprocessing of lactose into lactic acid and ethanol using non-engineered cell factories. *Bioresour. Technol.* **2022**, *345*, 126464. [\[CrossRef\]](#)
64. Li, B.; Wang, L.; Wu, Y.J.; Xia, Z.Y.; Yang, B.X.; Tang, Y.Q. Improving acetic acid and furfural resistance of xylose-fermenting *Saccharomyces cerevisiae* strains by regulating novel transcription factors revealed via comparative transcriptomic analysis. *Appl. Environ. Microbiol.* **2021**, *87*, e00158–21. [\[CrossRef\]](#)
65. Bhavana, B.K.; Mudliar, S.N.; Bokade, V.; Debnath, S. Effect of furfural, acetic acid and 5-hydroxymethylfurfural on yeast growth and xylitol fermentation using *Pichia stipitis* NCIM 3497. *Biomass Convers. Biorefinery* **2024**, *14*, 4909–4923.
66. Singh, N.; Gupta, A.; Mathur, A.S.; Barrow, C.; Puri, M. Integrated consolidated bioprocessing for simultaneous production of Omega-3 fatty acids and bioethanol. *Biomass Bioenergy* **2020**, *137*, 105555. [\[CrossRef\]](#)
67. Bar-Peled, L.; Kory, N. Principles and functions of metabolic compartmentalization. *Nat. Metab.* **2022**, *4*, 1232–1244. [\[CrossRef\]](#) [\[PubMed\]](#)
68. Upadhyay, A.; Kovalev, A.A.; Zhuravleva, E.A.; Pareek, N.; Vivekanand, V. Enhanced production of acetic acid through bioprocess optimization employing response surface methodology and artificial neural network. *Bioresour. Technol.* **2023**, *376*, 128930. [\[CrossRef\]](#) [\[PubMed\]](#)
69. Breig, S.J.M.; Luti, K.J.K. Response surface methodology: A review on its applications and challenges in microbial cultures. *Mater. Today Proc.* **2021**, *42*, 2277–2284. [\[CrossRef\]](#)
70. Forster, T.; Vázquez, D.; Müller, C.; Guillén-Gosálbez, G. Machine learning uncovers analytical kinetic models of bioprocesses. *Chem. Eng. Sci.* **2024**, *300*, 120606. [\[CrossRef\]](#)
71. Maitra, S.; Singh, V. A consolidated bioprocess design to produce multiple high-value platform chemicals from lignocellulosic biomass and its technoeconomic feasibility. *J. Clean. Prod.* **2022**, *377*, 134383. [\[CrossRef\]](#)

72. Liu, Y.K.; Yang, C.A.; Chen, W.C.; Wei, Y.H. Producing bioethanol from cellulosic hydrolyzate via co-immobilized cultivation strategy. *J. Biosci. Bioeng.* **2012**, *114*, 198–203. [\[CrossRef\]](#)
73. Olson, D.G.; McBride, J.E.; Shaw, A.J.; Lynd, L.R. Recent progress in consolidated bioprocessing. *Curr. Opin. Biotechnol.* **2012**, *23*, 396–405. [\[CrossRef\]](#)
74. Brethauer, S.; Studer, M.H. Consolidated bioprocessing of lignocellulose by a microbial consortium. *Energy Environ. Sci.* **2014**, *7*, 1446–1453. [\[CrossRef\]](#)
75. Wen, Z.; Wu, M.; Lin, Y.; Yang, L.; Lin, J.; Cen, P. Artificial symbiosis for acetone-butanol-ethanol (ABE) fermentation from alkali extracted deshelled corn cobs by co-culture of *Clostridium beijerinckii* and *Clostridium cellulovorans*. *Microb. Cell Factories* **2014**, *13*, 92. [\[CrossRef\]](#)
76. Mbaneme-Smith, V.; Chinn, M.S. Consolidated bioprocessing for biofuel production: Recent advances. *Energy Emiss. Control Technol.* **2015**, *3*, 23–44. [\[CrossRef\]](#)
77. Zuroff, T.R.; Xiques, S.B.; Curtis, W.R. Consortia-mediated bioprocessing of cellulose to ethanol with a symbiotic *Clostridium phytofermentans*/yeast co-culture. *Biotechnol. Biofuels* **2013**, *6*, 59. [\[CrossRef\]](#) [\[PubMed\]](#)
78. Bu, Y.; Alkotaini, B.; Salunke, B.K.; Deshmukh, A.R.; Saha, P.; Kim, B.S. Direct ethanol production from cellulose by consortium of *Trichoderma reesei* and *Candida molischiana*. *Green Process. Synth.* **2019**, *8*, 416–420. [\[CrossRef\]](#)
79. Xu, L.; Tschirner, U. Improved ethanol production from various carbohydrates through anaerobic thermophilic co-culture. *Bioresour. Technol.* **2011**, *102*, 10065–10071. [\[CrossRef\]](#)
80. Park, E.Y.; Naruse, K.; Kato, T. One-pot bioethanol production from cellulose by co-culture of *Acremonium cellulolyticus* and *Saccharomyces cerevisiae*. *Biotechnol. Biofuels* **2012**, *5*, 64. [\[CrossRef\]](#)
81. Wang, N.; Yan, Z.; Liu, N.; Zhang, X.; Xu, C. Synergy of Cellulase Systems between *Acetivibrio thermocellus* and *Thermoclostridium stercorarium* in Consolidated-Bioprocessing for Cellulosic Ethanol. *Microorganisms* **2022**, *10*, 502. [\[CrossRef\]](#)
82. Pinto, J.; Antunes, J.; Ramos, J.; Costa, R.S.; Oliveira, R. Modeling and optimization of bioreactor processes. In *Current Developments in Biotechnology and Bioengineering*; Elsevier: Amsterdam, The Netherlands, 2022; pp. 89–115.
83. Jin, Q.; Wu, Q.; Shapiro, B.M.; McKernan, S.E. Limited mechanistic link between the Monod equation and methanogen growth: A perspective from metabolic modeling. *Microbiol. Spectr.* **2022**, *10*, e02259-21. [\[CrossRef\]](#)
84. Cheng, Y.; Bi, X.; Xu, Y.; Liu, Y.; Li, J.; Du, G.; Lv, X.; Liu, L. Artificial intelligence technologies in bioprocess: Opportunities and challenges. *Bioresour. Technol.* **2023**, *369*, 128451. [\[CrossRef\]](#)
85. Sadino-Riquelme, M.; Donoso-Bravo, A.; Zorrilla, F.; Valdebenito-Rolack, E.; Gómez, D.; Hansen, F. Computational fluid dynamics (CFD) modeling applied to biological wastewater treatment systems: An overview of strategies for the kinetics integration. *Chem. Eng. J.* **2023**, *466*, 143180. [\[CrossRef\]](#)
86. Khanal, S.K.; Tarafdar, A.; You, S. Artificial intelligence and machine learning for smart bioprocesses. *Bioresour. Technol.* **2023**, *375*, 128826. [\[CrossRef\]](#)
87. Madhuvanthi, S.; Jayanthi, S.; Suresh, S.; Pugazhendhi, A. Optimization of consolidated bioprocessing by response surface methodology in the conversion of corn stover to bioethanol by thermophilic *Geobacillus thermoglucosidasius*. *Chemosphere* **2022**, *304*, 135242. [\[CrossRef\]](#) [\[PubMed\]](#)
88. Gopalakrishnan, S.; Johnson, W.; Valderrama-Gomez, M.A.; Icten, E.; Tat, J.; Ingram, M.; Shek, C.F.; Chan, P.K.; Schlegel, F.; Rolandi, P.; et al. COSMIC-dFBA: A novel multi-scale hybrid framework for bioprocess modeling. *Metab. Eng.* **2024**, *82*, 183–192. [\[CrossRef\]](#) [\[PubMed\]](#)
89. Yeboah, M.K.; Asiedu, N.Y.; Dogbe, S.; Addo, A. Performance of Machine Learning Based-Modelling Approach in Consolidated Bioprocessing with Microbial Consortium for Bioethanol Production. *Ind. Biotechnol.* **2024**, *20*, 77–97. [\[CrossRef\]](#)
90. Salakkam, A.; Phukoetphim, N.; Laopaiboon, P.; Laopaiboon, L. Mathematical modeling of bioethanol production from sweet sorghum juice under high gravity fermentation: Applicability of Monod-based, logistic, modified Gompertz and Weibull models. *Electron. J. Biotechnol.* **2023**, *64*, 18–26. [\[CrossRef\]](#)
91. Sen, R.; Roy, S. *Biofuel Production: Biological Technologies and Methodologies*; CRC Press: Boca Raton, FL, USA, 2022.
92. Dempfle, D.B. Model-Based Scale-Up of a Continuously Operated Consolidated Bioprocess Based on a Microbial Consortium for the Production of Ethanol. Ph.D. Thesis, École Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland, 2022. Available online: <https://infoscience.epfl.ch/entities/publication/3d8bd464-8b40-4939-b7d0-59372ff14f47> (accessed on 20 November 2024).
93. Rendón-Castrillón, L.; Ramírez-Carmona, M.; Ocampo-López, C.; Gómez-Arroyave, L. Mathematical model for scaling up bioprocesses using experiment design combined with Buckingham Pi theorem. *Appl. Sci.* **2021**, *11*, 11338. [\[CrossRef\]](#)
94. Mangipudi, S.; Reddy, D.G.V.; Ranganathan, P. Computational tools in bioprocessing. In *Current Developments in Biotechnology and Bioengineering*; Elsevier: Amsterdam, The Netherlands, 2022; pp. 211–231.
95. Agharafeie, R.; Ramos, J.R.C.; Mendes, J.M.; Oliveira, R. From Shallow to Deep Bioprocess Hybrid Modeling: Advances and Future Perspectives. *Fermentation* **2023**, *9*, 922. [\[CrossRef\]](#)

96. Mahanty, B. Hybrid modeling in bioprocess dynamics: Structural variabilities, implementation strategies, and practical challenges. *Biotechnol. Bioeng.* **2023**, *120*, 2072–2091. [\[CrossRef\]](#)
97. Liang, Y.; Ma, A.; Zhuang, G. Construction of environmental synthetic microbial consortia: Based on engineering and ecological principles. *Front. Microbiol.* **2022**, *13*, 829717. [\[CrossRef\]](#)
98. Abdelgalil, S.; Soliman, N.; Abo-Zaid, G.; Abdel-Fattah, Y. Bioprocessing strategies for cost-effective large-scale production of bacterial laccase from *Lysinibacillus macroides* LSO using bio-waste. *Int. J. Environ. Sci. Technol.* **2021**, *19*, 1633–1652. [\[CrossRef\]](#)
99. Sriraman, V.; Johnrajan, J.; Yazhini, K.; Rathinasabapathi, P. Bioprocess engineering essentials: Cultivation strategies and mathematical modeling techniques. In *Industrial microbiology and Biotechnology: A New Horizon of the Microbial World*; Springer: Berlin/Heidelberg, Germany, 2024; pp. 247–276.
100. de Andrade Bustamante, R.; de Oliveira, J.S.; Dos Santos, B.F. Modeling biosurfactant production from agroindustrial residues by neural networks and polynomial models adjusted by particle swarm optimization. *Environ. Sci. Pollut. Res.* **2023**, *30*, 6466–6491. [\[CrossRef\]](#)
101. do Nascimento, J.F.C.; dos Reis, B.D.; de Baptista Neto, Á.; Lerin, L.A.; de Oliveira, J.V.; de Paula, A.V.; Remonato, D. Comparing a polynomial DOE model and an ANN model for enhanced geranyl cinnamate biosynthesis with Novozym® 435 lipase. *Biocatal. Agric. Biotechnol.* **2024**, *58*, 103240. [\[CrossRef\]](#)
102. Monteiro, M.; Fadda, S.; Kontoravdi, C. Towards advanced bioprocess optimization: A multiscale modelling approach. *Comput. Struct. Biotechnol. J.* **2023**, *21*, 3639–3655. [\[CrossRef\]](#) [\[PubMed\]](#)
103. O'Brien, C.M.; Zhang, Q.; Daoutidis, P.; Hu, W.S. A hybrid mechanistic-empirical model for in silico mammalian cell bioprocess simulation. *Metab. Eng.* **2021**, *66*, 31–40. [\[CrossRef\]](#)
104. Tsafrakidou, P.; Manthos, G.; Zagklis, D.; Mema, J.; Kornaros, M. Assessment of substrate load and process pH for bioethanol production—Development of a kinetic model. *Fuel* **2022**, *313*, 123007. [\[CrossRef\]](#)
105. Abdelgalil, S.A.; Soliman, N.A.; Abo-Zaid, G.A.; Abdel-Fattah, Y.R. Dynamic consolidated bioprocessing for innovative lab-scale production of bacterial alkaline phosphatase from *Bacillus paralicheniformis* strain APSO. *Sci. Rep.* **2021**, *11*, 6071. [\[CrossRef\]](#) [\[PubMed\]](#)
106. Djimtoingar, S.S.; Derkyi, N.S.A.; Kuranchie, F.A.; Yankyera, J.K. A review of response surface methodology for biogas process optimization. *Cogent Eng.* **2022**, *9*, 2115283. [\[CrossRef\]](#)
107. Vaid, S.; Sharma, S.; Dutt, H.C.; Mahajan, R.; Bajaj, B.K. One pot consolidated bioprocess for conversion of *Saccharum spontaneum* biomass to ethanol-biofuel. *Energy Convers. Manag.* **2021**, *250*, 114880. [\[CrossRef\]](#)
108. Selvakumar, P.; Kavitha, S.; Sivashanmugam, P. Optimization of process parameters for efficient bioconversion of thermo-chemo pretreated *Manihot esculenta* Crantz YTP1 stem to ethanol. *Waste Biomass Valorizat.* **2019**, *10*, 2177–2191. [\[CrossRef\]](#)
109. Mondal, P.P.; Galodha, A.; Verma, V.K.; Singh, V.; Show, P.L.; Awasthi, M.K.; Lall, B.; Anees, S.; Pollmann, K.; Jain, R. Review on machine learning-based bioprocess optimization, monitoring, and control systems. *Bioresour. Technol.* **2023**, *370*, 128523. [\[CrossRef\]](#)
110. Razzak, S.A.; Alam, M.S.; Hossain, S.Z.; Rahman, S.M. Tree-based machine learning for predicting *Neochloris oleoabundans* biomass growth and biological nutrient removal from tertiary municipal wastewater. *Chem. Eng. Res. Des.* **2024**, *210*, 614–624. [\[CrossRef\]](#)
111. Serrano-Bermúdez, L.M. Bioprocesses in the Era of Artificial Intelligence. *Rev. Colomb. Biotechnol.* **2025**, *27*, 1–4. [\[CrossRef\]](#)
112. Duong-Trung, N.; Born, S.; Kim, J.W.; Schermeyer, M.T.; Paulick, K.; Borisyak, M.; Cruz-Bournazou, M.N.; Werner, T.; Scholz, R.; Schmidt-Thieme, L.; et al. When bioprocess engineering meets machine learning: A survey from the perspective of automated bioprocess development. *Biochem. Eng. J.* **2023**, *190*, 108764. [\[CrossRef\]](#)
113. Helleckes, L.M.; Hemmerich, J.; Wiechert, W.; von Lieres, E.; Grünberger, A. Machine learning in bioprocess development: From promise to practice. *Trends Biotechnol.* **2023**, *41*, 817–835. [\[CrossRef\]](#) [\[PubMed\]](#)
114. Baako, T.M.D.; Kulkarni, S.K.; McClendon, J.L.; Harcum, S.W.; Gilmore, J. Machine learning and deep learning strategies for Chinese hamster ovary cell bioprocess optimization. *Fermentation* **2024**, *10*, 234. [\[CrossRef\]](#)
115. Krishna, V.V.; Pappa, N.; Rani, S.J.V. Deep learning based soft sensor for bioprocess application. In *Proceedings of the 2021 IEEE Second International Conference on Control, Measurement and Instrumentation (CMI), Kolkata, India, 8–10 January 2021*; IEEE: Piscataway, NJ, USA, 2021; pp. 155–159.
116. Imamoglu, E. Artificial Intelligence and/or Machine Learning Algorithms in Microalgae Bioprocesses. *Bioengineering* **2024**, *11*, 1143. [\[CrossRef\]](#)
117. Singh, A.; Singhal, B. Role of machine learning in bioprocess engineering: Current perspectives and future directions. In *Design and Applications of Nature Inspired Optimization: Contribution of Women Leaders in the Field*; Springer: Berlin/Heidelberg, Germany, 2023; pp. 39–54.
118. Luo, Y.; Kurian, V.; Ogunnaike, B.A. Bioprocess systems analysis, modeling, estimation, and control. *Curr. Opin. Chem. Eng.* **2021**, *33*, 100705. [\[CrossRef\]](#)
119. Linardatos, P.; Papastefanopoulos, V.; Kotsiantis, S. Explainable AI: A review of machine learning interpretability methods. *Entropy* **2020**, *23*, 18. [\[CrossRef\]](#)

120. Rogers, A.W.; Song, Z.; Ramon, F.V.; Jing, K.; Zhang, D. Investigating ‘greyness’ of hybrid model for bioprocess predictive modelling. *Biochem. Eng. J.* **2023**, *190*, 108761. [\[CrossRef\]](#)
121. Agharafeie, R.; Oliveira, R.; Ramos, J.R.C.; Mendes, J.M. Application of hybrid neural models to bioprocesses: A systematic literature review. *Authorea Prepr.* **2023**. [\[CrossRef\]](#)
122. Yang, C.T.; Kristiani, E.; Leong, Y.K.; Chang, J.S. Big data and machine learning driven bioprocessing—recent trends and critical analysis. *Bioresour. Technol.* **2023**, *372*, 128625. [\[CrossRef\]](#)
123. Angelov, P.P.; Soares, E.A.; Jiang, R.; Arnold, N.I.; Atkinson, P.M. Explainable artificial intelligence: An analytical review. *Wiley Interdiscip. Rev. Data Min. Knowl. Discov.* **2021**, *11*, e1424. [\[CrossRef\]](#)
124. Das, A.; Rad, P. Opportunities and challenges in explainable artificial intelligence (xai): A survey. *arXiv* **2020**, arXiv:2006.11371. [\[CrossRef\]](#)
125. Srinivasan, B. A guide to the Michaelis–Menten equation: Steady state and beyond. *FEBS J.* **2022**, *289*, 6086–6098. [\[CrossRef\]](#) [\[PubMed\]](#)
126. Clark, T.J.; Guo, L.; Morgan, J.; Schwender, J. Modeling plant metabolism: From network reconstruction to mechanistic models. *Annu. Rev. Plant Biol.* **2020**, *71*, 303–326. [\[CrossRef\]](#) [\[PubMed\]](#)
127. Narayanan, H.; Luna, M.F.; von Stosch, M.; Cruz Bournazou, M.N.; Polotti, G.; Morbidelli, M.; Butté, A.; Sokolov, M. Bioprocessing in the digital age: The role of process models. *Biotechnol. J.* **2020**, *15*, 1900172. [\[CrossRef\]](#)
128. Rozov, S. Machine Learning and Deep Learning methods for predictive modelling from Raman spectra in bioprocessing. *arXiv* **2020**, arXiv:2005.02935. [\[CrossRef\]](#)
129. Nazemzadeh, N.; Malanca, A.A.; Nielsen, R.F.; Gernaey, K.V.; Andersson, M.P.; Mansouri, S.S. Integration of first-principle models and machine learning in a modeling framework: An application to flocculation. *Chem. Eng. Sci.* **2021**, *245*, 116864. [\[CrossRef\]](#)
130. Stiglic, G.; Kocbek, P.; Fijacko, N.; Zitnik, M.; Verbert, K.; Cilar, L. Interpretability of machine learning-based prediction models in healthcare. *Wiley Interdiscip. Rev. Data Min. Knowl. Discov.* **2020**, *10*, e1379. [\[CrossRef\]](#)
131. Krishnan, M. Against interpretability: A critical examination of the interpretability problem in machine learning. *Philos. Technol.* **2020**, *33*, 487–502. [\[CrossRef\]](#)
132. Zhang, Z.; Zhang, H. Application of Big Data Analysis in Intelligent Industrial Design Using Scalable Computational Model. *Scalable Comput. Pract. Exp.* **2025**, *26*, 1180–1195. [\[CrossRef\]](#)
133. Isoko, K.; Cordiner, J.L.; Kis, Z.; Moghadam, P.Z. Bioprocessing 4.0: A pragmatic review and future perspectives. *Digit. Discov.* **2024**, *3*, 1662–1681. [\[CrossRef\]](#)
134. Yang, S.; Wan, M.P.; Chen, W.; Ng, B.F.; Dubey, S. Model predictive control with adaptive machine-learning-based model for building energy efficiency and comfort optimization. *Appl. Energy* **2020**, *271*, 115147. [\[CrossRef\]](#)
135. Habib, M.K.; Ayankoso, S.A.; Nagata, F. Data-driven modeling: Concept, techniques, challenges and a case study. In *Proceedings of the 2021 IEEE International Conference on Mechatronics and Automation (ICMA), Takamatsu, Japan, 8–11 August 2021*; IEEE: Piscataway, NJ, USA, 2021; pp. 1000–1007.
136. Han, Z.; Gao, C.; Liu, J.; Zhang, J.; Zhang, S.Q. Parameter-efficient fine-tuning for large models: A comprehensive survey. *arXiv* **2024**, arXiv:2403.14608.
137. Penloglou, G.; Kiparissides, A. Advanced Modeling of Biomanufacturing Processes. *Processes* **2024**, *12*, 387. [\[CrossRef\]](#)
138. Herrera-Ruiz, J.F.; Fontalvo, J.; Prado-Rubio, O.A. Hybrid Modeling for Bioprocesses: Architectures, Applications, and Perspectives. *Eng. Rep.* **2025**, *7*, e70502. [\[CrossRef\]](#)
139. Catalão, M.; Pinto, J.; Torres, C.A.; Freitas, F.; Reis, M.A.; Costa, R.S.; Oliveira, R. Bioprocess model-predictive control with physics-informed neural networks: Driving microbiome evolution toward high polyhydroxyalkanoates production capacity. *J. Process Control* **2025**, *156*, 103594. [\[CrossRef\]](#)
140. Ali, H.; Safdar, R.; Liu, J.; Binti Abd Manan, T.S.; Hu, G.; Rasool, M.H.; Yao, Y.; Gao, F. Hybrid fusion paradigm in advanced process monitoring: A panoramic review and future perspectives. *Ind. Eng. Chem. Res.* **2025**, *64*, 22465–22514. [\[CrossRef\]](#)
141. Helleckes, L.M.; Osthege, M.; Wiechert, W.; von Lieres, E.; Oldiges, M. Bayesian calibration, process modeling and uncertainty quantification in biotechnology. *PLoS Comput. Biol.* **2022**, *18*, e1009223. [\[CrossRef\]](#)
142. Morabito, B.; Pohlodek, J.; Kranert, L.; Espinel-Ríos, S.; Findeisen, R. Efficient and simple Gaussian process supported stochastic model predictive control for bioreactors using HILO-MPC. *IFAC-PapersOnLine* **2022**, *55*, 922–927. [\[CrossRef\]](#)
143. Rashedi, M.; Rafiei, M.; Demers, M.; Khodabandehlou, H.; Wang, T.; Tulsyan, A.; Undey, C.; Garvin, C. Machine learning-based model predictive controller design for cell culture processes. *Biotechnol. Bioeng.* **2023**, *120*, 2144–2159. [\[CrossRef\]](#)

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.