Received: 27 January 2024 | Revised: 19 March 2024 | Accepted: 6 April 2024

DOI: 10.1002/aic.18458

RESEARCH ARTICLE

Process Systems Engineering

AIChE JOURNAL

Design and operation of modular biorefinery supply chain under uncertainty using generalized Benders decomposition

Yuqing Luo | Marianthi Ierapetritou

Department of Chemical and Biomolecular Engineering, University of Delaware, Newark, Delaware, USA

Correspondence
Marianthi Ierapetritou, Department of Chemical and Biomolecular Engineering, University of Delaware, 150 Academy Street, Newark, DE 19716 USA.
Email: mgi@udel.edu

Funding information
National Science Foundation, Grant/Award Numbers: 2217472, GCR CMMI 1934887, OIA - 2119754

Abstract

Biomass supply chain performance is heavily affected by uncertainties stemming from supply, demand, or unexpected disruptions. Unlike petrochemical plants that use crude oil, biorefineries often have to deal with the uneven spatial-temporal distribution of feedstock supply. The modular production strategy provides more flexibility in chemical manufacturing by allowing fast capacity expansion and unit movement. However, modeling and optimizing modular biomass supply chain under uncertainties becomes challenging due to high dimensionality and the existence of discrete decisions. This work optimizes the multiperiod biomass supply chain using the rolling horizon planning and two-stage stochastic programming framework. We then applied generalized Benders decomposition to reduce the computational complexity of the stochastic mixed integer nonlinear programming supply chain optimization. Furthermore, the solution of the stochastic programming could be used to quantitatively describe the life-cycle assessment uncertainties of the biomass supply chain performance, demonstrating seasonality and random variability.

KEYWORDS
biomass, generalized benders decomposition, life-cycle assessment, modular production, uncertainty

1 | INTRODUCTION

Using renewable biomass feedstock is an efficient way to decarbonize the chemical industry.[1] However, biomass supplies are spatially distributed and unstable due to various harvest times and unpredictable yields.[2,3] Thus, managing uncertainties throughout the biomass supply chain, including supply, demand, and disruptions, is essential to promote sustainable chemical production.[4,5]

Modular production with expandable or movable units is established as an effective strategy to improve the production facility performances and flexibility of the supply chain.[6,7] First of all, multiple standardized smaller modules, instead of large equipment, are chosen to meet the capacity or demand.[8] Therefore, the production sites can respond to the shifts of supply or demand centers and unexpected situations by installing or relocating units as needed.[9,10] Additional benefits of modular production include lower initial investment risk, more straightforward maintenance, and shorter time to market.[10–13] For instance, Shao et al. showed an up to three times investment risk reduction achieved by modular production in a biogas production and utilization process.[7] Shao et al. utilized the graph representation of the spatial superstructure to design plastic waste upcycling supply chain and demonstrates the trade-off between costs and modularity measure.[14,15] Allman and Zhang demonstrated movable modules' ability to handle demand center movement in the supply chain with module configuration considerations.[9,16] Allman et al. further showcased the modular biomass-to-energy supply chain's response to supply pattern

Abbreviations: BARON, branch and reduce optimization navigator; GWP, global warming potential; GBD, generalized benders decomposition; HMF, 5-hydroxymethylfurfural; LCA, life-cycle assessment; MILP, mixed integer linear programming; MINLP, mixed integer nonlinear programming; MSSP, multi-stage stochastic programming; NLP, nonlinear programming; MSH, molten salt hydrate; PSA, pressure-sensitive adhesive; RD – RCF, reactive distillation reductive catalytic fractionation; TSSP, two-stage stochastic programming.

changes in Minnesota and North Carolina.[16] Tan and Barton formulated a multi-period optimization model for gas monetization processes and established the profitability of mobile plants under uncertain prices, supply, and demand.[17] Bhosekar et al. introduced the surrogate models to replace feasibility constraints in the modular production systems and design the modular supply chain with "economy-of-numbers."[8,18,19]

Despite the benefits of modular production, the operation of modular supply chain is subject to multiple uncertainties.[20] Stochastic programming is a versatile framework to incorporate uncertainties as different scenarios into optimization problem.[20,21] For example, Gao and You applied a bilevel two-stage stochastic programming (TSSP) framework to a decentralized supply chain with shale well productivity and processing performance uncertainties.[22] Tong et al. included low-high scenarios for biomass availability, product demand, crude price and availabilities, and possible technology improvement in an integrated biofuel and petroleum supply chain.[23] The decisions in a multi-period supply chain planning problem are often taken sequentially as uncertainties unfold in each period.[24] Multi-stage stochastic programming (MSSP) provides a general framework for such multi-period problems by accurately reflecting the sequence and interactions of uncertainties and decisions.[25] However, the high computational complexity limits the application of MSSP models in large-scale supply chain optimization problems.[24,26] One way to simplify this multi-stage decision process under uncertainties is to take the rolling horizon approach, which plans only a fraction of the entire planning period (i.e., prediction interval) by TSSP at a time but iteratively re-solve the planning with updated information.[27,28] The rolling horizon method decomposes the time dimension of the planning problem and maintains small scenario tree sizes, resulting in a much shorter computational time.[29]

The stochastic rolling horizon planning problem may still contain a large number of scenarios when multiple sources of uncertainties need to be sampled. Most stochastic programming problems demonstrate decomposable block structures that Lagrangean or Benders decompositions can exploit.[30] Benders decomposition (L-shaped algorithm) is commonly adopted in stochastic programming to decompose the subproblems by scenarios and construct Benders cuts of the master problem by dual variables. However, integer variables in the second stage (recourse actions) prevent the direct Bender decomposition implementation, as dual multipliers are not available from the mixed-integer subproblems.[31] Some additional techniques have been incorporated to overcome this limitation, but integer recourse actions still pose computational challenges.[32] For example, Li and Grossmann addressed the integer recourse actions in stochastic programming by an improved L-shaped algorithm with strengthened Benders and Lagrangean cuts.[33] Decomposition-based branch-and-bound algorithms were also implemented to handle the discrete decisions in both stages of the TSSP.[32,34]

Uncertainty analysis also attracts increasing interest in LCA.[35] Due to the high measurement cost, obtaining multiple observations to understand the material or energy consumption fluctuation of a running process is expensive.[36] LCA practitioners typically build the life-cycle inventory and perform the analysis using nominal input-output values for all possibly uncertain parameters. Uncertainties are then propagated by assuming (mostly lognormal) distributions for each unit process flow based on source, completeness, and correlations of the parameters (e.g., Pedigree methods).[37] The weakness of the Pedigree method is that it relies only on experts' opinion of data quality, rather than using any actual observations (i.e., empirical data). Consequently, the standard deviations given by the Pedigree method do not carry any physical meaning nor quantitatively reflect the magnitudes of LCA result variability.[37,38]

In this work, we first formulated a modular biomass supply chain optimization with spatial-temporal distributed feedstocks, demand uncertainty, and disruptions. Rolling horizon planning and generalized Benders decomposition were then implemented to solve the proposed problem efficiently. Next, the stochastic supply chain optimization solutions were utilized to understand potential realizations of emissions—LCA uncertainties. This analysis framework utilizes available empirical information (e.g., demand, supply, and disruption) to quantify LCA uncertainties, which provide valuable and objective insights to guide better decisions than the traditional Pedigree approach.[39,40]

2 | PROBLEM FORMULATION

The modular supply chain network covers the locations of feedstock supply regions, production sites, and retail markets. Considering different biomass growth and harvest cycles, the optimization model aims to plan an entire year of the modular supply chain operation. The uncertainties influencing biomass conversion facilities are modeled in a TSSP framework with "here-and-now" and "wait-and-see" decisions.[22] The primary goal of the optimization model is to reduce the biomass supply chain cost with movable production modules. The planning decisions could be classified into two categories: (1) module-related, for example, new module installation and existing module movement; (2) operation-related, for example, supplier selection, chemical production, material transportation, and backorder.

For each planning interval, "here-and-now" decisions are made before the current month starts.[22] At the beginning of each time period, the biomass availability, facility disruption, and product demand in this current month are revealed, whereas those for the future months remain uncertain. It is assumed that forecasting models exist for uncertain feedstock supply, facility outage, and market demand. The imperfect prediction is captured by adding random variabilities to the forecasts due to the lack of historical market data.[19] This uncertain scenario assumption could be improved by adopting regression models, especially time series ones such as autoregressive integrated moving average, when large dataset is available in the future.[41] The prediction intervals of time series regressions will update when new data come in and will accurately reflect forecasting errors in real-time. After the uncertain parameters are known, the "wait-and-see" decisions are made for each scenario of the future months.[22] This setting allows the supply chain to respond to the uncertainties but also ensures future information is not leaked, which mimics how uncertainties affect decision-making.[26]

In this work, four products from the lignocellulosic biomass are produced in two steps: molten salt hydrate (MSH) hydrolysis and reactive-distillation reductive catalytic fractionation (RD-RCF).[42] As shown in Figure 1, the first MSH reaction breaks down cellulose and

hemicellulose components into sugars, and the following dehydration reaction converts them to furfural and 5-hydroxymethylfurfural (HMF), respectively.[43] The unreacted lignin is fractionated by RD-RCF to form phenolics, functionalized into monomers, and eventually polymerized into pressure-sensitive adhesives (PSA). The conversion coefficients are extracted from the Aspen Plus simulation (Aspen Tech, Burlington, MA) to reflect the feedstock chemical compositions and reaction yields.[43–46]

2.1 | Objective function

The optimization problem is formulated to minimize the total expected cost of the modular biomass supply chain (Equation (1)), including new module installation, existing module movement, disrupted module restoration, facility operation, material transportation, and backorder cost. In the stochastic programming formulation, sample average approximation was considered to convert continuous distributions of a large number of uncertain parameters into a smaller discrete scenario set ($\omega \in \Omega$).[20] Next, the costs from different scenarios are reconciled as one objective by taking the expected value.

$$\text{Total cost} = \mathbb{E}_{\omega}[\text{Install}_{\omega} + \text{Move}_{\omega} + \text{Restore}_{\omega} + \text{OPEX}_{\omega} + \text{Trans}_{\omega} + \text{Backorder}_{\omega}]. \quad (1)$$

Equation (2) calculates the total installation cost of each scenario ω by the sum of all newly added units $z_{j,m,t,\omega}$ of type m at production site j during time t.

$$\text{Install}_{\omega} = \sum_{t \in T} \sum_{m \in M} \sum_{j \in J} \left(q_m \cdot z_{j,m,t,\omega}\right). \quad (2)$$

The movement cost of units is proportional to number of units m moved from location j to j' during time t and the unit movement cost $r_{j,j',m}$. Here, $r_{j,j',m}$ is based on the distance between the two sites j and j' and the weight of unit m.

$$\text{Move}_{\omega} = \sum_{t \in T} \sum_{m \in M} \sum_{j \in J} \sum_{j' \in J} \left(r_{j,j',m} \cdot v_{j,j',m,t,\omega}\right). \quad (3)$$

After a disruption happens at site j and time t ($\text{dis}_{j,t,\omega} = 1$), modules in the affected location should be restored at a cost (res_m) to continue their normal operation (Equation (4)).

$$\text{Restore}_{\omega} = \sum_{t \in T} \sum_{m \in M} \sum_{j \in J} \left(\text{res}_m \cdot n_{j,m,t,\omega} \cdot \text{dis}_{j,t,\omega}\right). \quad (4)$$

The operating cost has two parts, the first includes utilities and resource consumptions. This term is thus proportional to the operating level (i.e., total flow rate into the operating units in one location, $\sum_{s \in S} \sum_{f \in F} Q_{s,j,f,t,\omega}$). The second part is quadratic in the flow rates, which is a typical term considered in the supply chain optimization to represent the busy operation or congestion.[47,48]

$$\text{Operate}_{\omega} = \sum_{t \in T} \sum_{j \in J} \left(o_{m1} \cdot \sum_{s \in S} \sum_{f \in F} Q_{s,j,f,t,\omega} + o_{m2} \cdot \left(\sum_{s \in S} \sum_{f \in F} Q_{s,j,f,t,\omega} \right)^2 \right). \quad (5)$$

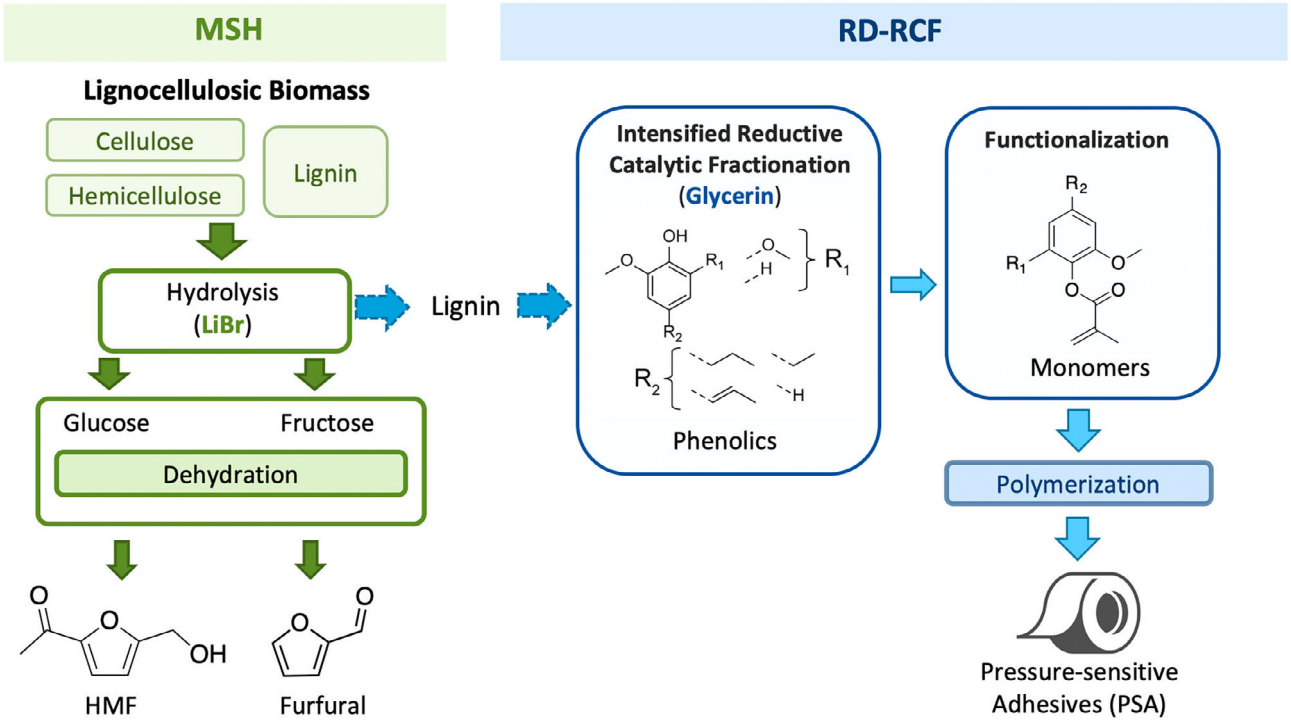

FIGURE 1 Lignocellulosic biomass conversion steps considered in modular biomass supply chain.

The material transportation cost is calculated by the distance and weight of feedstocks ($\sum_{t\in T}\sum_{f\in F} Q_{s,j,f,t,\omega}$) and products ($\sum_{t\in T}\sum_{p\in P} D_{j,r,p,t,\omega}$) shipped during every time t.

$$\begin{aligned}\text{Transport}_{\omega} = &\sum_{s\in S}\sum_{j\in J}\left(h_{s,j}\cdot\sum_{t\in T}\sum_{f\in F} Q_{s,j,f,t,\omega}\right)\\ &+\sum_{j\in J}\sum_{r\in R}\left(h_{j,r}\cdot\sum_{t\in T}\sum_{p\in P} D_{j,r,p,t,\omega}\right),\end{aligned} \tag{6}$$

where $h_{s,j}$, and $h_{j,r}$ are the transportation costs of feedstocks from supply region s to site j, product from site j to market r.

Although the biomass conversion facilities are expected to satisfy the demand as much as possible due to contract, unmet demands could still happen in extreme cases. Therefore, purchasing from other chemical producers, as backorder $B_{r,p,t,\omega}$ for product p at retailer r during time t, is permitted to fulfill the contract at a higher price (penalty $b_{r,p}$).

$$\text{Backorder}_{\omega} = \sum_{t\in T}\sum_{p\in P}\sum_{r\in R} b_{r,p}\cdot B_{r,p,t,\omega}. \tag{7}$$

The environmental impacts are evaluated by a "cradle-to-gate" LCA that includes carbon sequestration, manufacturing, transportation, and penalties for backorder purchases. Multiple products are often obtained from biomass conversion facilities.[8,44] In this work, the most valuable product is selected as the main product, and other by-products are treated as deductible emission credits based on the "avoided burden" approach.[49] More specifically, the emission of producing the same amount of by-product stream in the standalone reference processes is removed from the final results.[49] The total emissions of the biorefinery supply chain are given by Equation (8), which is tracked throughout the supply chain planning but not included in the objective function.

$$\begin{aligned}\text{Emission}_{\omega} = &\sum_{t\in T}\sum_{m\in M}\sum_{j\in J}\sum_{j'\in J'}\left(\beta^{\text{move}}_{j,j',m}\cdot v_{j,j',m,t,\omega}\right)\\ &+\sum_{t\in T}\sum_{f\in F}\sum_{s\in S}\sum_{j\in J}\left(\beta^{\text{transport}}_{s,j}\cdot Q_{s,j,f,t,\omega}\right)\\ &+\sum_{t\in T}\sum_{p\in P}\sum_{j\in J}\sum_{r\in R}\left(\beta^{\text{transport}}_{j,r}\cdot D_{j,r,p,t,\omega}\right)+\sum_{m\in M}\sum_{t\in T}\sum_{j\in J}\beta^{\text{operate}}_{m}\\ &\cdot\left(\sum_{s\in S}\sum_{f\in F} Q_{s,j,f,t,\omega}\right)+\sum_{t\in T}\sum_{f\in F}\sum_{s\in S}\sum_{j\in J}\left(\beta^{\text{seq}}_{f}\cdot Q_{s,j,f,t,\omega}\right)\\ &+\sum_{t\in T}\sum_{p\in P}\sum_{r\in R}\eta_{r,p}\cdot B_{r,p,t,\omega}-\sum_{t\in T}\sum_{p'\in P'}\sum_{r\in R}\eta_{r,p'}\cdot\delta_{r,p',t,\omega},\end{aligned} \tag{8}$$

where $\beta^{\text{move}}_{j,j',m}$ is the emission of moving module m from j to j'; β_m is the "gate-to-gate" emission of modular m occurring at the production stage; $\beta^{\text{transport}}_{s,j}$ and $\beta^{\text{transport}}_{j,r}$ are the emissions when transporting feedstock f or product p from supplier s to site j, and from site j to market r; β^{seq}_{f} is the carbon sequestration during the growth stage of feedstock f; $\eta_{r,p'}$ is the "cradle-to-gate" emission of the by-product p' that serves as the credit (substitution); $\delta_{r,p',t,\omega}$ is the uncertain demand of byproduct p' at market r during time t in scenario ω.

2.2 | Module-related constraints

The module expansion and relocation at each site are modeled by the following conservation Equation (9):

$$n_{j,m,t,\omega} = \begin{cases} n_{j,m,0} + z_{j,m,t,\omega} + \sum_{j'\in J'}\left(v_{j',j,m,t,\omega} - v_{j,j',m,t,\omega}\right) & \forall j,m,\omega,\ \text{if } t=1\\ n_{j,m,t-1,\omega} + z_{j,m,t,\omega} + \sum_{j'\in J'}\left(v_{j',j,m,t,\omega} - v_{j,j',m,t,\omega}\right) & \forall j,m,\omega,\ \text{if } t\geq 2\end{cases}, \tag{9}$$

where $n_{j,m,t,\omega}$ is the number of unit m at production site j during time period t; $n_{j,m,0}$ is the number of module m at production site j during time 0, $z_{j,m,t,\omega}$ is the newly purchased unit m at time t at the same site; $v_{j,j',m,t,\omega}$ is the number of modules moved from site j to j' at time $t-1$.

The total number of units moved away from site j at time t should not exceed the number of units already located there in the previous period (Equation (10)).

$$n_{j,m,t-1,\omega} \geq \sum_{j'\in J} v_{j,j',m,t,\omega}, \forall j,m,t,\omega. \tag{10}$$

One essential design consideration for a modular chemical plant is the maximum number of production units that it can accommodate. This capacity is limited by the size of the backbone facility, which provides the utilities and other necessary services to operate modular production lines.[11] The maximum number of each type of unit m allowed at each site j is limited by $N_{j,m,t}$ (Equation (11)).

$$n_{j,m,t,\omega} \leq N_{j,m,t} \quad ,\forall j,m,t,\omega. \tag{11}$$

The above constraints (Equations (9)–(11)) involve only integer variables that are related to the number of units. However, these discrete decisions lead to a mixed-integer nonlinear programming (MINLP) supply chain optimization problem with a large number of variables, given all combinations of time, location, scenario, material, and unit types.

2.3 | Operation-related constraints

At each production site, feedstock and product material flows follow the mass balance Equation (12). The conversion coefficients are based on the process simulation and literature data of feedstock compositions.[8,43–45]

$$\sum_{r\in R} D_{j,r,p,t,\omega} = \sum_{f\in F}\text{conv}_{f,p}\cdot\sum_{s\in S} Q_{s,j,f,t,\omega} \quad ,\forall j,p,t,\omega. \tag{12}$$

Additionally, nonlinear effects could also appear in the scale-ups and the conversion coefficients. For instance, reduced conversion as the flow rate increases has been introduced as a concave log function in the work of Kocis and Grossmann.[50] To capture this nonlinear

relationship, we introduce a similar logarithmic term to indicate the reduced conversion coefficients as the flow rate increases.

$$\sum_{r\in R} D_{j,r,p,t,\omega} \le \alpha \cdot \sum_{f\in F} \mathrm{conv}_{f,p} \cdot \sum_{s\in S} Q_{s,j,f,t,\omega} + (1-\alpha)\cdot \sum_{f\in F} \mathrm{conv}_{f,p} \cdot \ln\left(\sum_{s\in S} Q_{s,j,f,t,\omega} + 1\right), \forall j,p,t,\omega, \quad (13)$$

where $Q_{s,j,f,t,\omega}$ is the flowrate of feedstock f from supplier s to site j during time t; $\mathrm{conv}_{f,p}$ is the nominal conversion of product p using feedstock f; $D_{j,r,p,t,\omega}$ the flowrates of product p from site j to market r. The tuning parameter α between 0 and 1 adjusts how much the apparent reaction conversion coefficients deviate from the ideal linear relationship. It is shown that formulating the concave constraints as inequality preserves the convexity.[50] The current case study aims to reduce the total cost of the biomass supply chain while meeting biomass-based chemical demands. Thus, the constraints are almost always active because the system would want to produce as much as possible to meet the contract order and reduce the backorder.[50] Equation (13) becomes inactive only in the extreme cases where all retailers place very few orders on certain chemicals across the entire market. Hence, manufacturing sites choose to reduce the transportation cost by shipping fewer products to these contract retailers than what they can produce. In reality, this situation could lead the product managers to participate more in a spot-price market and sell the excess chemicals on a "free-on-board" basis, where the manufacturers are not responsible for shipping.[51] Moreover, if profit-maximization is selected as the goal of planning, the production sites will try to maximize what they send out, keeping Equation (13) active.

As an essential constraint that links the integer and continuous variables, Equation (14) requires that production activity cannot exceed the total installed capacity at site j. Here, unexpected disruption ($\mathrm{dis}_{j,t,\omega}$) influences the usable capacities on the right-hand side: if disruption happens, the plant site j cannot treat any feedstock; otherwise, the maximum feedstock flowrates are bounded by the capacity of unit m (c_m) and the number of units at that location during time t.

$$\sum_{s\in S}\sum_{f\in F} Q_{s,j,f,t,\omega} \le \sum_{m\in M} c_m \cdot n_{j,m,t,\omega} \cdot \left(1-\mathrm{dis}_{j,t,\omega}\right), \forall j,t,\omega. \quad (14)$$

Next, the backorder constraint is imposed by Equation (15) if the production cannot satisfy the uncertain demand. Producers could purchase the desired product from other suppliers (at an increased price) in the same market r to fulfill the contract.

$$B_{r,p,t,\omega} \ge \delta_{r,p,t,\omega} - \sum_{j\in J} D_{j,r,p,t,\omega}, \forall r,p,t,\omega, \quad (15)$$

where $\delta_{r,p,t,\omega}$ is the uncertain demand and $B_{r,p,t,\omega}$ is the backorder of product p at market r during t.

Finally, different feedstock availabilities ($\vartheta_{s,f,t,\omega}$) in each region s and time t are compiled based on the Biofuel Atlas database, which determines the upper bounds of the supply (Equation 16)

$$\sum_{j\in J} Q_{s,j,f,t,\omega} \le \vartheta_{s,f,t,\omega} \quad , \forall s,f,t,\omega, \quad (16)$$

2.4 | Non-anticipativity constraints (NACs)

The proposed modular supply chain model considers the uncertainties as possible scenarios, and the stochastic programming is expanded to its deterministic equivalent by creating copies of variables and constraints for each scenario.[32] Based on the TSSP model formulation, "here-and-now" decisions are those taken in the first month of the planning, and include the units placement, installation, movement, feedstock transportation, product shipment, and backorders. These decisions are planned before uncertain supply $\vartheta_{s,f,t,\omega}$, demand $\delta_{r,p,t,\omega}$, and disruption $\mathrm{dis}_{j,t,\omega}$ are revealed. Consequently, "here-and-now" variables in the first stage should be consistent among all scenarios because planners cannot predict the uncertainty realization to change their actions accordingly.[52] This restriction is referred to as the NACs and enforced by Equations (17)–(22).

$$n_{j,m,1,\omega} = n_{j,m,1,1} \quad , \forall j,m,\omega, \quad (17)$$

$$z_{j,m,1,\omega} = z_{j,m,1,1} \quad , \forall j,m,\omega, \quad (18)$$

$$v_{j',j,m,1,\omega} = v_{j',j,m,1,1} \quad , \forall j,j',m,\omega, \quad (19)$$

$$Q_{s,j,f,1,\omega} = Q_{s,j,f,1,1} \quad , \forall s,j,f,\omega, \quad (20)$$

$$D_{j,r,p,1,\omega} = D_{j,r,p,1,1} \quad , \forall j,r,p,\omega, \quad (21)$$

$$B_{r,p,1,\omega} = B_{r,p,1,1} \quad , \forall r,p,\omega. \quad (22)$$

3 | SOLUTION STRATEGY FOR THE MINLP MODULAR SUPPLY CHAIN PROBLEM

3.1 | Rolling horizon planning

The rolling horizon method can incorporate uncertain information into the supply chain planning problems and simplify the multi-stage decision-making.[27,28] Instead of looking at the entire planning horizon, a TSSP model is first developed at a truncated period, referred to as the prediction horizon (Figure 2A). The first step is to sample a scenario tree for uncertain parameters during the prediction horizon to formulate the optimization problem. Next, this smaller-size stochastic optimization is solved, and the "here-and-now" decisions (e.g., number of modules and flow rates) are implemented and passed to the following time period. For example, the number of units at each site in the first month is now treated as a parameter ($n_{j,m,0}$) of Equation (9) in the next planning problem. The time window is then moved forward, and the new scenario tree is updated using the forecasting model. A TSSP supply chain model starting the following month is constructed and solved again to continue the iterations until all time periods are planned.

Similar to NACs in the full-scale supply chain model, the parameters and decisions for the first month will be consistent among all sampled scenarios (Figure 2B).[19] However, decisions beyond the first month of each prediction horizon ("wait-and-see") are allowed to vary in different scenarios to respond to future situations. Overall, the rolling horizon approach maintains a small optimization model size (i.e., shorter time horizon and fewer uncertain scenarios) in each step by temporally decomposing the long planning horizon and solving each subperiod sequentially.[28] The prediction horizon length is determined based on the trade-off between solution quality and computational complexity. Although the modular strategy allows the supply chain to change its production capacities rapidly, a longer prediction horizon helps avoid myopic decisions.[53] For example, a short prediction horizon may overlook the long-term benefit of installing new units and favor the option of backorder when the demand stress is high (Figure S20). Rolling horizon planning also allows decision-makers to fully utilize updated uncertainty information and react to changes swiftly.[54] Therefore, this approach provides high-quality solutions for applications with low computational requirements.[55] For instance, rolling horizon with parametric programming has been shown to efficiently solve large-scale reactive scheduling problems of combined heat and power network.[28]

3.2 | Generalized Benders decomposition

Despite the significant problem size reduction achieved by rolling horizon planning, the resulting stochastic MINLP supply chain problem still contains a large copy of variables and constraints for each scenario. However, the stochastic supply chain problem is relatively sparse and has a unique block structure that could be exploited by decomposition algorithms.[33,56] Generalized Benders decomposition (GBD) splits the MINLP optimization problems into smaller mixed-integer linear programming (MILP) master problem and nonlinear programming (NLP) subproblems by fixing the complicating variables. GBD has been shown to efficiently solve large-scale structured optimization problems with a wide variety of applications.[57] For instance, Paules and Floudas applied the GBD algorithm to solve a MINLP process design problem with distillation column and heat integration.[58] Mitrai and Daoutidis implemented GBD with the hybrid multi-cut algorithm to efficiently solve the integrated planning, scheduling, and control problems.[56] Stochastic blockmodeling was also proposed to help select variables for Benders decomposition in complex problems.[56,59,60]

In rolling horizon planning, the only state variable that connects different time windows (i.e., consecutive prediction horizon) is the number of units ($n_{j,m,t,\omega}$). This variable also serves as the link between module-related and flowrate-related subsets of constraints that divide the overall supply chain planning task. This structure creates the opportunity for applying GBD to split the original MINLP problem into a pure-integer master problem and several continuous nonlinear subproblems (Figure 3).

The Benders master problem (MB^k) contains the integer variables in all scenarios, which deals with the installation, movement, and restoration of modules (Figure 4A). The last term of the objective function, $\mathbb{E}_{\omega}[\theta_{\omega}]$, is the value function of the average operation-related costs, which is approximated by a series of Benders optimality cuts (Equation (23)). This master problem provides a lower bound for the original problem because only a fraction of the constraints are included, and the new objective function is an underestimation of the actual cost. The original stochastic problem is a relatively complete recourse model such that any realization of the first-stage decision will lead to at least a feasible solution.[33] Even in the worst-case scenario, when disruptions or insufficient supply lead to very low

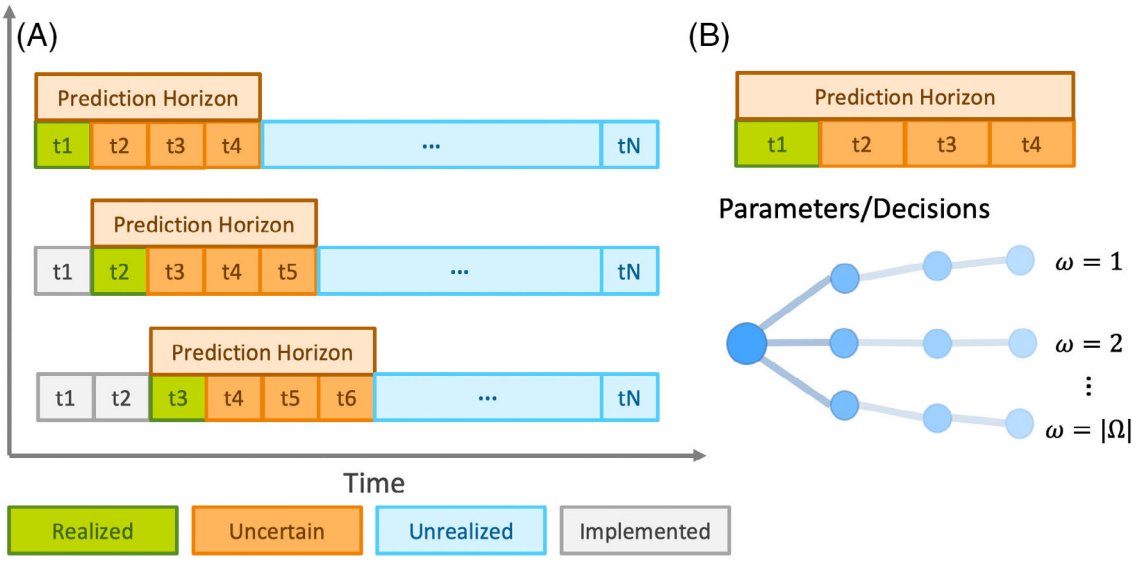

FIGURE 2 (A) Rolling horizon planning scheme and (B) Scenario tree obtained from rolling horizon.

production activities, all demands could still be satisfied with the backorder. Thus, no Benders feasibility cuts are introduced into the master problem. This work implements the multi-cut version of the Benders cut (Equation (23)) without cut selection, where $|\Omega|$ number of cuts are added in one iteration to approximate the value function individually. Different cut selection techniques could be applied to improve the efficiency of the Benders decomposition, especially when the master problem is very large.[61] However, the current biomass supply chain case study has a relatively simple Benders master problem, which is solved easily with commercial MILP solver even when all $K \cdot |\Omega|$ Benders cuts are included in the Kth iteration. This formulation has faster convergence than the standard single-cut version in many applications.[56,62] In Equation (23), $z^k_{SP,\omega}$, and $\lambda^k_{j,m,t,\omega}$ are the optimal solution and dual multipliers obtained from the subproblems; $\tilde{n}^k_{j,m,t,\omega}$ is the fixed first stage variables in the kth iteration.

$$\left(\mathrm{MB}^k\right): \quad \min_{n,z,v} \mathbb{E}_\omega[\mathrm{Install}_\omega + \mathrm{Move}_\omega + \mathrm{Restore}_\omega] + \mathbb{E}_\omega[\theta_\omega]$$

$$\text{s.t.} \quad \text{Equation}\,(9)-(11) \qquad \text{Unit balance and limits}$$

$$\text{Equation}\,(17)-(19) \qquad \text{Non}-\text{anticipativity constraints}$$

$$\theta_\omega \geq z^k_{SP,\omega} + \sum_{t\in T}\sum_{m\in M}\sum_{j\in J} \lambda^k_{j,m,t,\omega} \cdot \left(n_{j,m,t,\omega} - \tilde{n}^k_{j,m,t,\omega}\right)$$

$$n, z, v \in \mathbb{Z}^+. \tag{23}$$

The subproblems $\left(\mathrm{SSB}^k_\omega\right)$ are defined for each scenario (ω) once the history of units in all locations has been fixed. Each subproblem contains only continuous variables, shipment quantities, and backorders, which could be solved efficiently with commercial NLP solvers due to the reduced size (Figure 4B). The number of units obtained from the master problem ($\tilde{n}_{j,m,t,\omega}$) is fixed in the subproblem by Equation (24), which serves the crucial role of passing information from the master problem to the subproblems.[33] The dual multipliers of Equation (24) are extracted as λ^k_ω and used in Benders cuts (Equation (23)). Moreover, fixing the module-related integer variables will lead to a feasible solution to the original planning problem, which provides a valid upper bound.

	$v_{j,j',m,t,\omega}$	$z_{j,m,t,\omega}$	$n_{j,m,t,\omega}$	$Q_{s,j,f,t,\omega}$	$D_{j,r,p,t,\omega}$	$B_{r,p,t,\omega}$
Unit Balance	■	■	■			
Movement Limit	■		■			
Site Size Limit			■			
Production Capacity			■	■		
Biomass Conversion				■	■	
Supply Limit				■		
Backorder					■	■

FIGURE 3 Block structure of modular supply chain optimization problem.

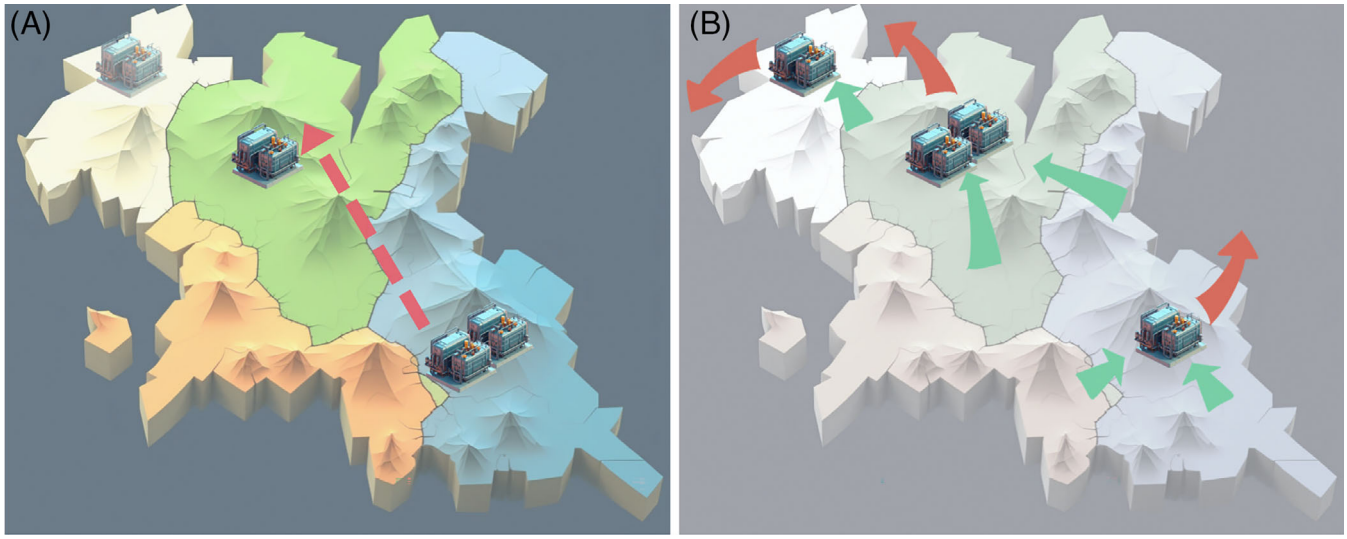

FIGURE 4 (A) GBD master problem involves module installation and movement. (B) GBD subproblem involves material transportation, production, and backorder.

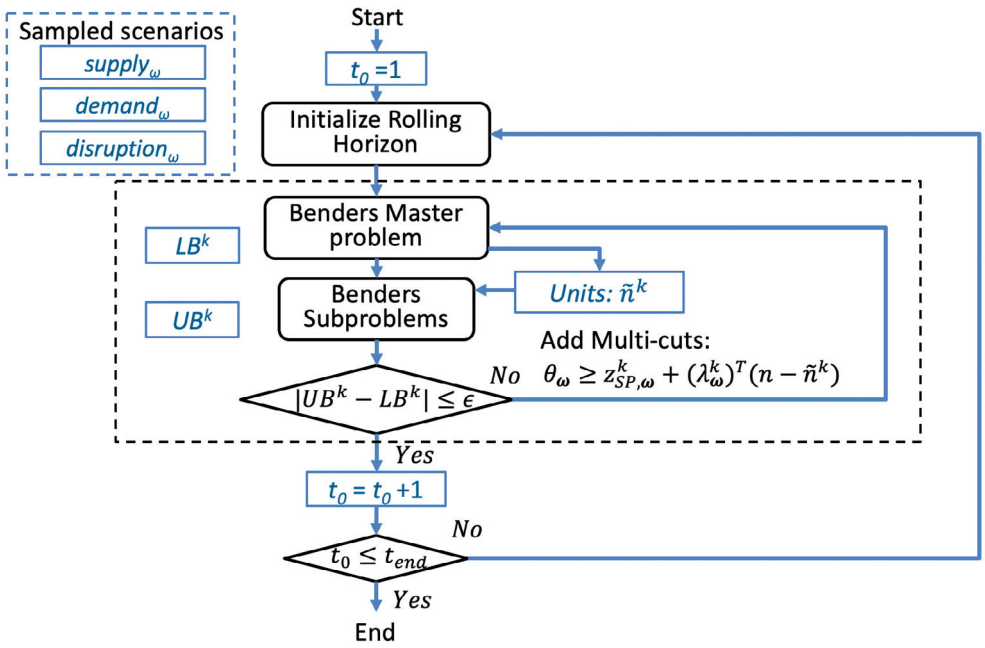

FIGURE 5 Proposed scheme for rolling-horizon and generalized Benders decomposition algorithm applied to modular biomass supply chain problem.

$$\left(\mathrm{SSB}^k_\omega\right): \quad \min_{n,Q,B} \left(\mathrm{transport}_\omega + \mathrm{operate}_\omega + \mathrm{backorder}_\omega\right)$$

$$\begin{aligned} s.t. \quad & \text{Equation (13)} && \text{Chemical production} \\ & \text{Equation (14)} && \text{Capacity limits} \\ & \text{Equation (15)} && \text{Backorder (demand)} \\ & \text{Equation (16)} && \text{Supply limits} \\ & n_{j,m,t,\omega} = \tilde{n}_{j,m,t,\omega} && \text{Fixing number of units} \\ & n, Q, B \in \mathbb{R}^+ && (24) \end{aligned}$$

Figure 5 illustrates the solution procedure for the modular biomass supply chain planning problem. In the outer loop, a rolling horizon is performed to focus on a moving and truncated time window of the entire planning period. After the uncertain scenario set is sampled, the inner Benders decomposition loop begins by initializing the upper and lower bounds to $\pm\infty$. The GBD algorithm then solves the master problem to obtain all integer decisions related to modules and updates the lower bound. Next, these integer variables are fixed in subproblems to optimize the best operating flow rates in each scenario. The upper bound and dual multipliers are obtained to construct new Benders cuts in the master problem. This iteration continues until the gap between upper and lower bounds is within the maximum allowed level.

4 | CASE STUDY: BIOMASS SUPPLY CHAIN WITH MODULAR PRODUCTION UNITS

The biomass supply chain network structure in this work is summarized in Figure 6A, where the feedstocks are harvested and transported to selected facility locations. Each supply region is represented by the geometric center of the county and the production facilities are chosen as the big cities within the region to ensure enough workforce and well-developed transportation infrastructure.[16] The production units could be moved around or installed to increase the capacity. Finally, the products are shipped to different markets to meet the demands.

4.1 | Modular biomass supply chain problem definition

The states of Missouri and Illinois are selected in this case study due to the abundant supply of lignocellulosic biomass, especially corn stover. However, corn stover is primarily available in Fall as residuals of the corn harvest.[16] In other seasons, willow and poplar are more readily available to fill the feedstock supply gap in biorefinery operations (Figure 6B). The potential supply of corn stover, poplar, and willow in this region is based on the National Renewable Energy Laboratory's (NREL) biomass supply data (Biofuel Atlas).[63] The Python package, GeoPandas, was utilized to calculate the pairwise distances between suppliers, sites, and markets.[64] We then incorporated temporal variation in biomass supply by further dividing the annual supply into each month. Random noises are added to the compiled biomass supply data to reflect the imperfect forecast and uncertain crop yields.[19] This forecasting error will increase when moving into the distant future, which is captured by larger variances of the Gaussian noises than feedstock supply in the near future.[19,65]

Demand was treated as time series data with trends, seasonality, and added random variability.[41] In this illustrative case study, it is assumed that a 2% monthly growth trend and sinusoidal seasonal patterns with varied magnitudes and frequencies exist for

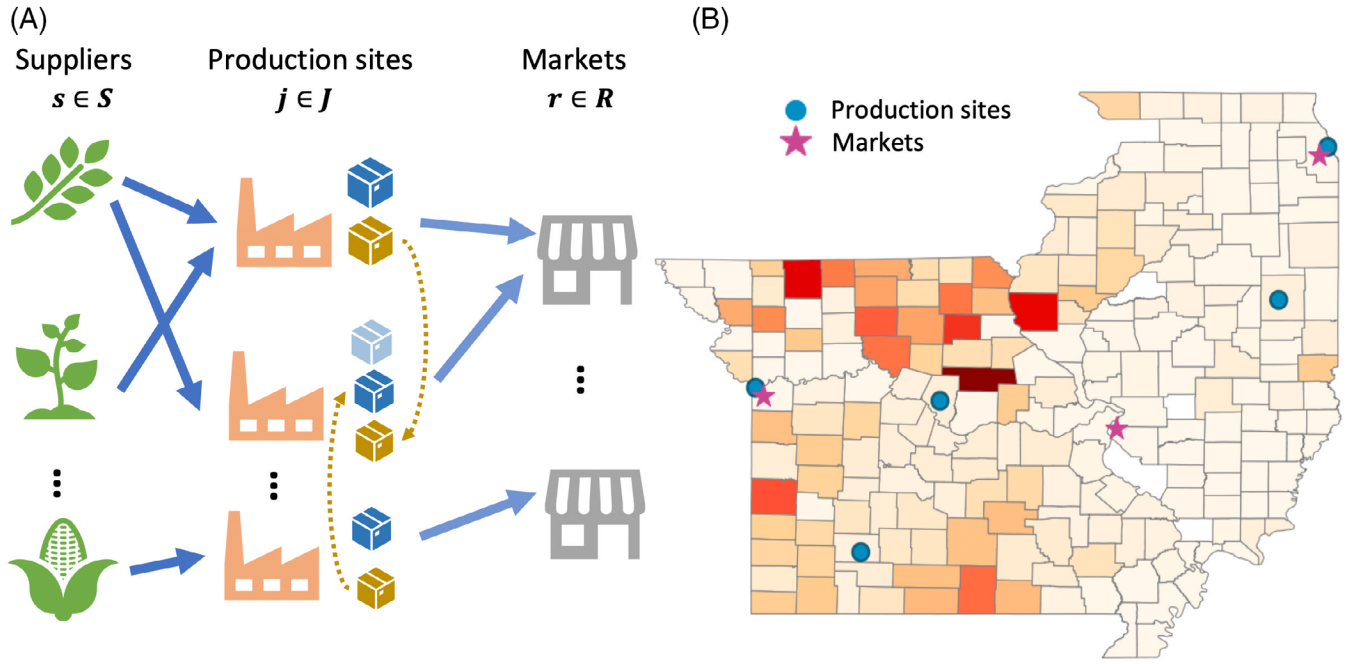

FIGURE 6 (A) Biomass supply chain network. (B) Poplar and willow distribution in the case study.

each product's demand. Similar to the biomass supply, Gaussian noises with more significant variances are added to later months of a prediction horizon. Chicago, Kansas City, and St. Louis are selected as the major trading centers for biomass-based chemical products. This forecasting assumption could be refined in the future by classic regression or machine learning models when sales records or market reports of biomass-based products are available.[66]

The capital and operating costs are based on Aspen Plus simulation and technoeconomic analysis for MSH and RD-RCF technologies.[8,43] Disruptions are assumed to happen randomly in each location with relatively low probabilities (Table S2). To ensure representative cases are included in the uncertain scenario set, 50 samples are taken in a 3-month prediction horizon. The modular biomass supply chain model contains 3 feedstock types, 210 supply regions, 5 production sites, a 12-month planning horizon, 3 markets, 2 conversion units, and 4 products. Other model parameters are listed in Supplemental Information.

4.2 | Performance of generalized Benders decomposition algorithm

All optimization models were implemented in Pyomo on a computer with Intel Xeon E-2274G CPU @ 4.00 GHz 32 GB RAM. The monolithic MINLP problem without decomposition was solved by the BARON solver. For generalized Benders decomposition, solvers for master problem and subproblems were CPLEX and CONOPT, respectively. CONOPT has been chosen as the NLP solver for subproblems due to the convex nature of the formulation and the absence of integer variables. The following two strategies are applied to accelerate the GBD performance.

4.2.1 | Demand satisfaction constraints

In the first few Benders decomposition iterations, not enough Benders cuts exist in the master problem to provide estimation for lower-level operation costs (θ_ω). The master problem dwells on lowering module-related costs without being concerned about the backorder penalties. Thus, GBD tends to explore low-quality solutions at the beginning—installing fewer units and producing insufficient products. In many cases, valid inequalities are typically derived to restrict Benders master problems to explore good initial points.[16,67] However, the proposed biomass supply chain optimization has relatively complete recourse, making it challenging to derive valid inequalities for the master problem.[68] Nevertheless, Equation (25) (by combining Equations (13) and (14)) serves similar functions as valid inequalities to restrict the master problem[67]:

$$\sum_{j \in J} \sum_{m \in M} \max_{f \in F} [\mathrm{conv}_{f,p}] \cdot c_m \cdot n_{j,m,t,\omega} \geq \eta \cdot \sum_{r \in R} \delta_{r,p,t,\omega} \quad \forall p,t,\omega. \tag{25}$$

The physical meaning of Equation (25) is that the total modules in the system should satisfy at least η % demand for time t and scenario ω. Here η is a tunable parameter reduced to zero in later GBD stages to lift this constraint and avoid affecting the solution quality. Due to the high backorder penalty, the biomass supply chain tries to use biomass feedstock to meet the contract if possible.

4.2.2 | Stabilization

Many GBD implementations display slow final convergence due to abrupt changes in solutions between two successive iterations.[69] This effect is common when the optimization exhibits solution degeneracy (i.e., several alternatives have similar objective functions), which is the case with the modular biomass supply chain. As a result, the oscillation of master problem solutions adds Benders cuts in regions far apart, rather than continuously improving the value function in one region of interest. Several stabilization methods have been applied to suppress the variation between iterations, including the level-bundle and trust-region methods.[70,71] This work applies the trust-region method by adding Equation (26) to the Benders master problem (MB^k), where the parameter M^{TR} determines the trust-region size

$$\sum_{j,m,t,\omega} | n_{j,m,t,\omega} - \tilde{n}^k_{j,m,t,\omega} | \le M^{TR}. \quad (26)$$

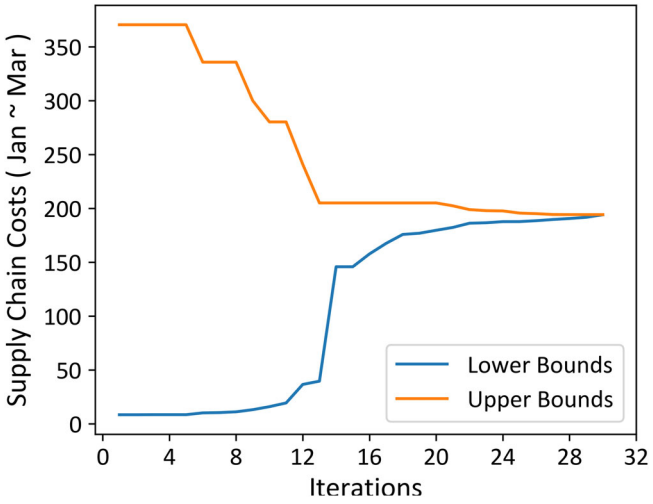

FIGURE 7 Upper and lower bounds of the GBD algorithm for the first prediction horizon.

Figure 7 shows the upper and lower bounds in the first prediction horizon of the biomass supply chain planning (January–March). The model details of the original optimization problem and GBD problems are summarized in Table 1.

Planning the entire 12-period biomass supply chain is optimized by rolling horizon, and GBD converges after 17.7 h. On average, each prediction horizon takes the GBD algorithm 1.5 h to reach less than a 0.5% optimality gap, where 23.4 and 288 s are spent on solving the master and subproblems in each iteration, respectively. The current GBD implementation solves 50 subproblems sequentially, which takes 92.4% of the time in each Benders iteration. However, all subproblems are independent of each other, and the computational time could be significantly reduced when parallel computing is adopted.[72] In contrast, the monolith problem still contains a large number of variables and constraints even after applying rolling horizon planning. The BARON solver could not find a feasible solution for the MINLP problem for a 3-month prediction horizon even after 20 h. Warm starts with two feasible solutions that have been implemented in BARON: (a) no unit has been installed in each location and all demands are satisfied by backorder, and (b) all production sites install the maximum number of units allowed. However, neither of the warm start strategies improves the MINLP model's convergence even after 20 h.

As a comparison, a deterministic supply chain model is formulated using mean parameter values of supply and demand. However, the deterministic model could not capture the effect of disruptions because this parameter takes only the value of 0 or 1 with a disruption probability. The deterministic model also underestimated the occurrence of backorder due to supply and demand variabilities. Overall, the stochastic supply chain planning model suggests more frequent movement of existing units rather than the installation of new units, which potentially reduces the units that are not fully operated because of disruptions and supply shortages (Figures S21 and S22). In this biomass modular supply chain case study, the values of stochastic solution (VSS) of each prediction horizon are between 3.5 and 41.7 M$ as a result of disruption and variations in demand or supply. This VSS result is around 1.3%–8.3% of the total cost and aligns with the typical ranges reported in the literature of similar modular production systems.[16,17] The modular production strategy enables the biomass supply chain to change capacities swiftly.[9,16] Moreover, the rolling horizon planning re-optimizes the supply chain using updated information, which allows the system to perform well even when sub-optimal decisions are made by the deterministic model in the previous step.[29] Additionally, the parameter choices (e.g., cost, supply, and demand) could potentially affect the VSS values in different supply chain cases, which should be further investigated in future work.

4.3 | Implications for LCA uncertainty analysis

In addition to the optimal supply chain design and operation, the proposed stochastic rolling horizon model keeps track of the LCA uncertainties of biomass-based chemical production. The functional unit of LCA was chosen as 1 kg of PSA supplied to the market because it is the main value-added product. Other byproducts (HMF, furfural, and lignin) were accounted for as the credits using the "avoided burden" method. Background emission data, such as transportation and utility generation, were collected from the Ecoinvent v3.3 database and literature results.[43,73] As one of the most used LCA metrics, global warming potential (GWP) was the environmental indicator used in this case study. When the plant grows, carbon dioxide is removed from the atmosphere and sequestrated in biomass species by photosynthesis.[74] This carbon sequestration effect is a primary benefit of using biomass feedstock, which is estimated by the carbon content of feedstocks through mass balance.[75,76]

TABLE 1 Optimization model characteristics of each prediction horizon.

		Generalized Benders decomposition	
	Monolith problem	Benders master problem	Benders subproblem
Model type	MINLP	MILP	NLP
Solvers	BARON	CPLEX	CONOPT
# of integer variables	10,500	10,500	0
# of continuous variables	483,300	0	9696
# of constraints	265,868	10,334[a]	2035
Average solution time (s)	–	23.4	5.75

[a]Every iteration includes additional 50 constraints (Benders multi-cuts) to the master problem.

The NACs (Equations (17)–(22)) in the stochastic programming render all decisions, costs, and emissions identical for the first month of each prediction horizon. The trajectory of the first-month GWP in each rolling horizon time window thus shows the nominal LCA performance of the modular production system from January to November, as shown by the solid line in Figure 8. Furthermore, the TSSP model allows decisions to vary under different uncertain demand/supply/disruption scenarios starting from the second month of the prediction horizon. Each set of decisions and, hence, emissions represent a possible behavior of the biomass supply chain that we may observe in the future. Therefore, the ensemble of all GWPs corresponding to different scenarios will give a quantitative representation of the actual variability of LCA results, which is reflected by the shaded area in Figure 8.

As expected, the nominal GWPs (with realized uncertainties) fall within the predicted uncertainty range from each sampled scenario. Additionally, the modular supply chain demonstrates two folds of GWP uncertainty relevant to LCA studies. First, the benefit of using biomass as the feedstock for chemical production varies substantially throughout the year, representing a seasonality. This seasonal pattern happens because the feedstock type, availability changes, and the byproduct credits are also affected by the varying demand over time. Unlike steady-state conditions in conventional petrochemical plants, the seasonality of biomass supply chain can create challenges in carbon accounting but also business opportunities.[74] For instance, if a downstream company plans to purchase PSA from this designed biomass supply chain, the emission associated with the production could range from −4.95 to 1.10 kg CO_2-eq/kg PSA throughout the year, and drastically reduced by 75% just between July and August. With increasing customer awareness of sustainable products, the vendors are more likely to succeed if they select appropriate timings for re-stocking and clearly label their product carbon footprint.

Second, even within the same month, there are considerable variabilities in the product's greenhouse gas emissions. For instance, the 5%–95% confidence interval covers a 0.55 kg CO_2-eq/kg PSA range in February; but this range expanded to 5.75 kg CO_2-eq/kg PSA in September. The confidence interval gives LCA practitioners an idea of how the results may differ from the nominal (average) values, which helps them make a more robust comparison between technology or product alternatives.[35] This static uncertainty behavior falls into the

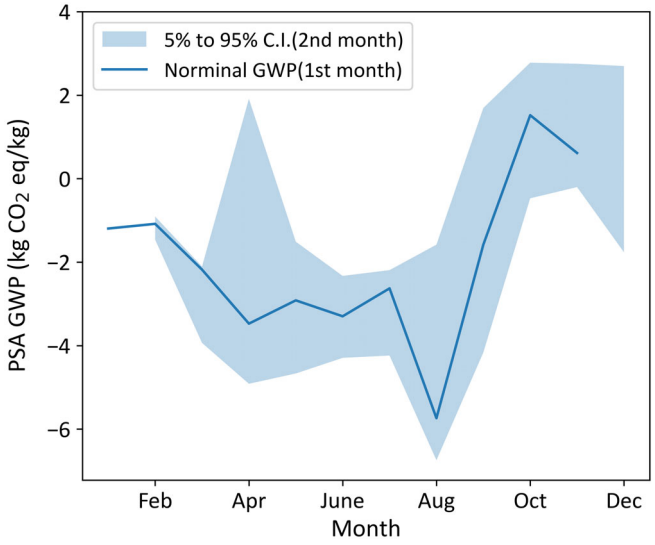

FIGURE 8 Tracked modular biomass supply chain performance in rolling horizon planning.

realm of the traditional LCA uncertainty discussion. Currently, most LCA uncertainty analyses rely on the Pedigree method to define parameter distributions in uncertainty propagation. The Pedigree method first asks practitioners to rate data quality on a scale of 1–5 from 6 perspectives: reliability (U_1), completeness (U_2), temporal correlation (U_3), geographical correlation (U_4), technological correlation (U_5), and sample size (U_6). These scores are then mapped to uncertainty factors and used in Equation (27) to calculate standard deviations for LCA process input-output relations (i.e., unit process flows).

$$\sigma^2 = \exp\left(\sqrt{(\ln U_1)^2 + (\ln U_2)^2 + (\ln U_3)^2 + (\ln U_4)^2 + (\ln U_5)^2 + (\ln U_6)^2}\right). \quad (27)$$

Yet, this procedure uses only qualitative judgment of data quality instead of actual data from the physical system, which does not quantitatively reflect the inherent variabilities in LCA results.[77]

The LCA uncertainty quantification in this work makes use of quantitative uncertain information from the actual supply chain system: spatial-temporal variations of supply and demand, forecasting models, and disruptions. Consequently, the stochastic programming-

based approach provides the empirical foundation and physical meanings for LCA uncertainties that the traditional Pedigree method lacks.[37]

5 | CONCLUSIONS

Distributed modular production strategy has the potential to improve the biomass supply chain's performance and flexibility. In this work, we utilized a rolling horizon planning to model the modular supply chain's design and operation under the spatial-temporal uncertainties of supply, demand, and disruption. A TSSP model was formulated to capture parameter uncertainties by scenario trees in each rolling-horizon time window. Integer variables in both stages (i.e., current and future months) are introduced for the movement and installation of production units.

We then applied a generalized Benders decomposition algorithm to the resulting MINLP based on its block structure. The Benders master problem is a MILP with integer variables involving the movement and expansion of modules in each site. The Benders subproblems are NLPs for planning the transportation, operation, and backorder actions under each uncertain scenario once the unit placement is decided.

The proposed supply chain optimization model with mobile modules was demonstrated in a biomass conversion case study of the Missouri-Illinois region, which was solved effectively by the combined rolling horizon and generalized Benders decomposition strategy. The main biomass sources in this region (i.e., corn stover, willow, and poplar) are grown unevenly and harvested during different seasons, encouraging the movement of the production modules. The option of having mobile units also alleviates the loss of unexpected production site disruptions. In the future, parallel computing[78] or scenario reduction techniques[79] can potentially further reduce the solution time because solving a large number of subproblems is the main time-consuming step.

In addition to reducing the costs of the modular biomass supply chain, the rolling-horizon stochastic programming framework is helpful in tracking and understanding the LCA uncertainties. This optimization-based uncertainty quantification method uses quantitative spatial-temporal information to demonstrate temporally-explicit LCA seasonality and inherent variance. These results and analysis framework can be extended to a variety of other LCA applications.

First, dynamic LCA studies establish more accurate emission accountings based on non-static emission factors and life cycle inventories.[80] The proposed multiperiod modular supply chain planning model provides such dynamic information to build the inventory, including material flow rates and production activity. Time-dependent dynamic LCA results are useful, especially when setting emission targets and evaluating process improvement over time.[81]

Second, market dynamics could be included in the current optimization framework to capture how demands are affected by prices. Although the demand-price elasticity functions are typically nonlinear, they will affect only the material flow subproblems that are compatible with the proposed generalized benders decomposition solution strategy.[82] Introducing the demand's response to price not only closely models the supply chain economics, but also determines how much petroleum-based commodities could be substituted by products from a biomass-based supply chain.[83] This substitution effect is an integral part of the consequential LCA, which evaluates the environmental impacts of decisions that change the business-as-usual system.[84]

Finally, prospective LCA analyzes emerging technologies in the early stage, which naturally involves predictive models and uncertainty information.[85] In the future, when biomass-based products have increased market share and ample market data are available, more detailed and reliable forecasting models could be constructed and incorporated into the stochastic programming framework.

ACKNOWLEDGMENTS

This work is financially supported by the National Science Foundation's Growing Convergence Research program (NSF GCR CMMI 1934887), Grant No. 2217472, and Grant No. OIA–2119754.

DATA AVAILABILITY STATEMENT

The spatial distribution of biomass availability is based on the National Renewable Energy Laboratory's (NREL) biomass supply data (Biofuel Atlas).[63] The model parameters used in the case study are tabulated in Tables S1–S4 of the Supporting Information. The numerical data shown in Figures 7 and 8 are tabulated in Tables S5–S9.

SETS/INDICES

$f \in F$	raw material feedstock
$p \in P$	products (e.g., lignin, HMF, furfural, and pressure-sensitive adhesive)
$s \in S$	supply regions
$j \in J$	production sites
$r \in R$	retailer (market)
$m \in M$	module types
$\omega \in \Omega$	scenarios
$t \in T$	time periods (month)
$k \in \mathrm{K}$	generalized Benders decomposition iteration

PARAMETERS

$\mathrm{conv}_{f,p}$	conversion coefficients of producing $p \in P$ from feedstock $f \in F$
q_m	capital cost of module $m \in M$
$r_{j,j',m}$	cost of moving module $m \in M$ from site $j \in J$ to $j' \in J$
o_{m1}	operating cost of module $m \in M$
o_{m2}	congestion cost of module $m \in M$, reflecting busy operation
res_m	restoration cost of module after disruption $m \in M$
$\mathrm{dis}_{j,t,\omega}$	indicator of whether disruption happens at site $j \in J$ during time $t \in T$ and scenario $\omega \in \Omega$
$h_{s,j}$	transportation costs of feedstocks from supply $s \in S$ to site $j \in J$
$h_{j,r}$	transportation costs of products from site $j \in J$ to market $r \in R$

$\delta_{r,p,t,\omega}$ uncertain demand of product p at market r during time t in scenario ω

$b_{r,p}$ the backorder cost for product $p \in P$ at market $r \in R$

$n_{j,m,0}$ number of modules $m \in M$ at production site $j \in J$ during time 0

$\tilde{n}^{k}_{j,m,t,\omega}$ fixed number of modules $m \in M$ at site $j \in J$ during time $t \in T$ based on the subproblem solution for scenario $\omega \in \Omega$ in k^{th} generalized Benders decomposition iteration

$\lambda^{k}_{j,m,t,\omega}$ dual values of subproblem constraints fixing the number of units in kth generalized Benders decomposition iteration

$z^{k}_{SP,\omega}$ optimal solution of subproblem $\omega \in \Omega$ in kth generalized Benders decomposition iteration

α tunable parameter representing the nonlinearity of reaction conversion constraints

η percentage of demand satisfied by biomass conversion operations

VARIABLES

$n_{j,m,t,\omega}$ number of modules $m \in M$ at production site $j \in J$ during time $t \in T$ in scenario $\omega \in \Omega$

$v_{j,j',m,t,\omega}$ number of module $m \in M$ moved from site $j \in J$ to site $j' \in J$ during time $t \in T$ in scenario $\omega \in \Omega$

$z_{j,m,t,\omega}$ number of newly purchased module $m \in M$ at production site $j \in J$ during time $t \in T$ in scenario $\omega \in \Omega$

$Q_{s,j,f,t,\omega}$ flowrate of feedstock from supplier $s \in S$ to $j \in J$ during time $t \in T$ in scenario $\omega \in \Omega$

$D_{j,r,p,t,\omega}$ flowrate of product $p \in P$ from site $j \in J$ to retailer $r \in R$ during time $t \in T$ in scenario $\omega \in \Omega$

$B_{r,p,t,\omega}$ unmet demand for product $p \in P$ at market $r \in R$ during time $t \in T$ in scenario $\omega \in \Omega$

θ_{ω} value function ($\omega \in \Omega$) approximated by Benders cut

ORCID

Yuqing Luo https://orcid.org/0000-0002-4174-946X

REFERENCES

1. Batchu SP, Hernandez B, Malhotra A, Fang H, Ierapetritou M, Vlachos DG. Accelerating manufacturing for biomass conversion via integrated process and bench digitalization: a perspective. *React Chem Eng*. 2022;7:813-832.
2. Breunig HM, Huntington T, Jin L, Robinson A, Scown CD. Temporal and geographic drivers of biomass residues in California. *Resour Conserv Recycl*. 2018;139:287-297.
3. O'Neill EG, Martinez-Feria RA, Basso B, Maravelias CT. Integrated spatially explicit landscape and cellulosic biofuel supply chain optimization under biomass yield uncertainty. *Comput Chem Eng*. 2022;160:107724.
4. Roni MS, Thompson DN, Hartley DS. Distributed biomass supply chain cost optimization to evaluate multiple feedstocks for a biorefinery. *Appl Energy*. 2019;254:113660.
5. Sharma P, Vlosky R, Romagnoli JA. Strategic value optimization and analysis of multi-product biomass refineries with multiple stakeholder considerations. *Comput Chem Eng*. 2013;50:105-129.
6. Bhosekar A, Ierapetritou M. A framework for supply chain optimization for modular manufacturing with production feasibility analysis. *Comput Chem Eng*. 2021;145:107175.
7. Shao Y, Hu Y, Zavala VM. Mitigating investment risk using modular technologies. *Comput Chem Eng*. 2021;153:107424.
8. Bhosekar A, Athaley A, Ierapetritou M. Multiobjective modular biorefinery configuration under uncertainty. *Ind Eng Chem Res*. 2021;60(35):12956-12969.
9. Allman A, Zhang Q. Dynamic location of modular manufacturing facilities with relocation of individual modules. *Eur J Oper Res*. 2020;286(2):494-507.
10. Baldea M, Edgar TF, Stanley BL, Kiss AA. Modular manufacturing processes: status, challenges, and opportunities. *AIChE J*. 2017;63(10):4262-4272.
11. Sievers S, Seifert T, Franzen M, Schembecker G, Bramsiepe C. Fixed capital investment estimation for modular production plants. *Chem Eng Sci*. 2017;158:395-410.
12. Weber RS, Snowden-Swan LJ. The economics of numbering up a chemical process enterprise. *J Adv Manuf Process*. 2019;1(1–2):e10011.
13. Luo Y, Selvam E, Vlachos DG, Ierapetritou M. Economic and environmental benefits of modular microwave-assisted polyethylene terephthalate depolymerization. *ACS Sustain Chem Eng*. 2023;11(10):4209-4218.
14. Shao Y, Ma J, Zavala VM. A spatial superstructure approach to the optimal design of modular processes and supply chains. *Comput Chem Eng*. 2023;170:108102.
15. Shao Y, Zavala VM. Modularity measures: concepts, computation, and applications to manufacturing systems. *AIChE J*. 2020;66:e16965.
16. Allman A, Lee C, Martín M, Zhang Q. Biomass waste-to-energy supply chain optimization with mobile production modules. *Comput Chem Eng*. 2021;150:107326.
17. Tan SH, Barton PI. Optimal dynamic allocation of mobile plants to monetize associated or stranded natural gas, part II: dealing with uncertainty. *Energy*. 2016;96:461-467.
18. Bhosekar A, Ierapetritou M. Modular design optimization using machine learning-based flexibility analysis. *J Process Control*. 2020;90:18-34.
19. Bhosekar A, Badejo O, Ierapetritou M. Modular supply chain optimization considering demand uncertainty to manage risk. *AIChE J*. 2021;67:e17367.
20. Santoso T, Ahmed S, Goetschalckx M, Shapiro A. A stochastic programming approach for supply chain network design under uncertainty. *Eur J Oper Res*. 2005;167(1):96-115.
21. Oliveira F, Grossmann IE, Hamacher S. Accelerating benders stochastic decomposition for the optimization under uncertainty of the petroleum product supply chain. *Comput Oper Res*. 2014;49:47-58.
22. Gao J, You F. A stochastic game theoretic framework for decentralized optimization of multi-stakeholder supply chains under uncertainty. *Comput Chem Eng*. 2019;122:31-46.
23. Tong K, Gong J, Yue D, You F. Stochastic programming approach to optimal design and operations of integrated hydrocarbon biofuel and petroleum supply chains. *ACS Sustain Chem Eng*. 2014;2(1):49-61.
24. Li C, Grossmann IE. A review of stochastic programming methods for optimization of process systems under uncertainty. *Front Chem Eng*. 2021;2:622241.
25. Apap RM, Grossmann IE. Models and computational strategies for multistage stochastic programming under endogenous and exogenous uncertainties. *Comput Chem Eng*. 2017;103:233-274.
26. Grossmann IE, Apap RM, Calfa BA, Garcia-Herreros P, Zhang Q. Mathematical programming techniques for optimization under uncertainty and their application in process systems engineering. *Theor Found Chem Eng*. 2017;51(6):893-909.
27. Zamarripa M, Marchetti PA, Grossmann IE, et al. Rolling horizon approach for production–distribution coordination of industrial gases supply chains. *Ind Eng Chem Res*. 2016;55(9):2646-2660.
28. Kopanos GM, Pistikopoulos EN. Reactive scheduling by a multiparametric programming rolling horizon framework: a case of a network

of combined heat and power units. *Ind Eng Chem Res*. 2014;53(11):4366-4386.
29. Lejarza F, Kelley MT, Baldea M. Feedback-based deterministic optimization is a robust approach for supply chain management under demand uncertainty. *Ind Eng Chem Res*. 2022;61(33):12153-12168.
30. Castro PM, Grossmann IE, Zhang Q. Expanding scope and computational challenges in process scheduling. *Comput Chem Eng*. 2018;114:14-42.
31. You F, Wassick JM, Grossmann IE. Risk management for a global supply chain planning under uncertainty: models and algorithms. *AIChE J*. 2009;55(4):931-946.
32. Li C, Grossmann IE. A generalized benders decomposition-based branch and cut algorithm for two-stage stochastic programs with nonconvex constraints and mixed-binary first and second stage variables. *J Glob Optim*. 2019;75(2):247-272.
33. Li C, Grossmann IE. An improved L-shaped method for two-stage convex 0–1 mixed integer nonlinear stochastic programs. *Comput Chem Eng*. 2018;112:165-179.
34. Li C, Grossmann IE. A finite epsilon-convergence algorithm for two-stage stochastic convex nonlinear programs with mixed-binary first and second-stage variables. *J Glob Optim*. 2019;75(4):921-947.
35. Igos E, Benetto E, Meyer R, Baustert P, Othoniel B. How to treat uncertainties in life cycle assessment studies? *Int J Life Cycle Assess*. 2019;24(4):794-807.
36. Piccinno F, Hischier R, Seeger S, Som C. From laboratory to industrial scale: a scale-up framework for chemical processes in life cycle assessment studies. *J Clean Prod*. 2016;135:1085-1097.
37. Ciroth A, Muller S, Weidema B, Lesage P. Empirically based uncertainty factors for the pedigree matrix in ecoinvent. *Int J Life Cycle Assess*. 2016;21(9):1338-1348.
38. Suh S, Qin Y. Pre-calculated LCIs with uncertainties revisited. *Int. J. Life Cycle Assess*. 2017;22(5):827-831.
39. Svanes E, Vold M, Hanssen OJ. Effect of different allocation methods on LCA results of products from wild-caught fish and on the use of such results. *Int J Life Cycle Assess*. 2011;16(6):512-521.
40. Ziyadi M, Al-Qadi IL. Model uncertainty analysis using data analytics for life-cycle assessment (LCA) applications. *Int J Life Cycle Assess*. 2019;24(5):945-959.
41. Hyndman RJ, Athanasopoulos G. *Forecasting: Principles and Practice*. 3rd ed. OTexts; 2021.
42. Luo Y, O'Dea RM, Gupta Y, et al. A life cycle greenhouse gas model of a yellow poplar Forest residue reductive catalytic fractionation biorefinery. *Environ Eng Sci*. 2022;39(10):821-833.
43. Athaley A, Annam P, Saha B, Ierapetritou M. Techno-economic and life cycle analysis of different types of hydrolysis process for the production of p-xylene. *Comput Chem Eng*. 2019;121:685-695.
44. Luo Y, Ierapetritou M. Multifeedstock and multiproduct process design using neural network surrogate flexibility constraints. *Ind Eng Chem Res*. 2023;62(5):2067-2079.
45. O'Dea RM, Pranda PA, Luo Y, et al. Ambient-pressure lignin valorization to high-performance polymers by intensified reductive catalytic deconstruction. *Sci Adv*. 2022;8(3):eabj7523.
46. Luo Y, Ierapetritou M. Comparison between different hybrid life cycle assessment methodologies: a review and case study of biomass-based p-xylene production. *Ind Eng Chem Res*. 2020;59(52):22313-22329.
47. Fischetti M, Ljubić I, Sinnl M. Benders decomposition without separability: a computational study for capacitated facility location problems. *Eur J Oper Res*. 2016;253(3):557-569.
48. Fischetti M, Ljubić I, Sinnl M. Redesigning benders decomposition for large-scale facility location. *Manage Sci*. 2016;63(7):2146-2162.
49. Anastasopoulou A, Keijzer R, Patil B, Lang J, van Rooij G, Hessel V. Environmental impact assessment of plasma-assisted and conventional ammonia synthesis routes. *J Ind Ecol*. 2020;24:1171-1185.
50. Kocis GR, Grossmann IE. Relaxation strategy for the structural optimization of process flow sheets. *Ind Eng Chem Res*. 1987;26(9):1869-1880.
51. Inderfurth K, Kelle P, Kleber R. Dual sourcing using capacity reservation and spot market: optimal procurement policy and heuristic parameter determination. *Eur J Oper Res*. 2013;225(2):298-309.
52. Escudero LF, Garín MA, Pérez G, Unzueta A. Lagrangian decomposition for large-scale two-stage stochastic mixed 0-1 problems. *TOP*. 2012;20(2):347-374.
53. Powell WB. *Approximate Dynamic Programming: Solving the Curses of Dimensionality*. 2nd ed. John Wiley & Sons, Inc.; 2011.
54. Perea-López E, Ydstie BE, Grossmann IE. A model predictive control strategy for supply chain optimization. *Comput Chem Eng*. 2003;27(8):1201-1218.
55. Wang L, Lu Z, Ren Y. A rolling horizon approach for production planning and condition-based maintenance under uncertain demand. *Proc Inst Mech Eng, Part O: J Risk Reliab*. 2019;233(6):1014-1028.
56. Mitrai I, Daoutidis P. A multicut generalized benders decomposition approach for the integration of process operations and dynamic optimization for continuous systems. *Comput Chem Eng*. 2022;164:107859.
57. Geoffrion AM. Generalized benders decomposition. *J Optim Theory Appl*. 1972;10(4):237-260.
58. Paules GE, Floudas CA. Stochastic programming in process synthesis: a two-stage model with MINLP recourse for multiperiod heat-integrated distillation sequences. *Comput Chem Eng*. 1992;16(3):189-210.
59. Mitrai I, Tang W, Daoutidis P. Stochastic blockmodeling for learning the structure of optimization problems. *AIChE J*. 2022;68(6):e17415.
60. Mitrai I, Daoutidis P. Efficient solution of Enterprise-wide optimization problems using nested stochastic blockmodeling. *Ind Eng Chem Res*. 2021;60(40):14476-14494.
61. Paterakis NG. Hybrid quantum-classical multi-cut benders approach with a power system application. *Comput Chem Eng*. 2023;172:108161.
62. Gebreslassie BH, Yao Y, You F. Design under uncertainty of hydrocarbon biorefinery supply chains: multiobjective stochastic programming models, decomposition algorithm, and a comparison between CVaR and downside risk. *AIChE J*. 2012;58(7):2155-2179.
63. The Biofuels Atlas. Innovative Data Energy Applications (IDEA). https://maps.nrel.gov/?da=biofuels-atlas Accessed December 28, 2023.
64. Jordahl K, Van den Bossche J, Wasserman J, et al. GeoPandas Version 0.14.3. 2019; https://geopandas.org/ Accessed December 28, 2023.
65. Nagy B, Farmer JD, Bui QM, Trancik JE. Statistical basis for predicting technological progress. *PLoS One*. 2013;8(2):e52669.
66. Blackburn R, Lurz K, Priese B, Göb R, Darkow I-L. A predictive analytics approach for demand forecasting in the process industry. *Int Tran Oper Res*. 2015;22(3):407-428.
67. Poudel SR, Marufuzzaman M, Bian L. Designing a reliable bio-fuel supply chain network considering link failure probabilities. *Comput Ind Eng*. 2016;91:85-99.
68. Saharidis GKD, Boile M, Theofanis S. Initialization of the Benders master problem using valid inequalities applied to fixed-charge network problems. *Expert Syst Appl*. 2011;38(6):6627-6636.
69. Sifuentes WS, Vargas A. Hydrothermal scheduling using benders decomposition: accelerating techniques. *IEEE Trans Power Syst*. 2007;22(3):1351-1359.
70. Colonetti B, Finardi EC. Combining Lagrangian relaxation, benders decomposition, and the level bundle method in the stochastic hydrothermal unit-commitment problem. *Int Trans Electr Energy Syst*. 2020;30(9):e12514.
71. Alshamsi A, Diabat A. Large-scale reverse supply chain network design: an accelerated benders decomposition algorithm. *Comput Ind Eng*. 2018;124:545-559.
72. Qi Y, Sen S. The ancestral Benders' cutting plane algorithm with multi-term disjunctions for mixed-integer recourse decisions in stochastic programming. *Math Program*. 2017;161(1):193-235.
73. Wernet G, Bauer C, Steubing B, Reinhard J, Moreno-Ruiz E, Weidema B. The ecoinvent database version 3 (part I): overview and methodology. *Int J Life Cycle Assess*. 2016;21(9):1218-1230.

74. Courchesne A, Bécaert V, Rosenbaum RK, Deschênes L, Samson R. Using the Lashof accounting methodology to assess carbon mitigation projects with life cycle assessment. *J Ind Ecol*. 2010;14(2):309-321.
75. Luo Y, Kuo MJ, Ye M, Lobo R, Ierapetritou M. Comparison of 4,4′-dimethylbiphenyl from biomass-derived furfural and oil-based resource: technoeconomic analysis and life-cycle assessment. *Ind Eng Chem Res*. 2022;61(25):8963-8972.
76. Liao Y, Koelewijn S-F, Van den Bossche G, et al. A sustainable wood biorefinery for low-carbon footprint chemicals production. *Science*. 2020;367(6484):1385-1390.
77. Henriksson PJG, Guinée JB, Heijungs R, de Koning A, Green DM. A protocol for horizontal averaging of unit process data—including estimates for uncertainty. *Int J Life Cycle Assess*. 2014;19(2):429-436.
78. Li X. Parallel nonconvex generalized Benders decomposition for natural gas production network planning under uncertainty. *Comput Chem Eng*. 2013;55:97-108.
79. Garcia-Herreros P, Wassick JM, Grossmann IE. Design of resilient supply chains with risk of facility disruptions. *Ind Eng Chem Res*. 2014;53(44):17240-17251.
80. Collinge WO, Landis AE, Jones AK, Schaefer LA, Bilec MM. Dynamic life cycle assessment: framework and application to an institutional building. *Int J Life Cycle Assess*. 2013;18(3):538-552.
81. Levasseur A, Lesage P, Margni M, Deschênes L, Samson R. Considering time in LCA: dynamic LCA and its application to global warming impact assessments. *Environ Sci Technol*. 2010;44(8):3169-3174.
82. Kaplan U, Türkay M, Karasözen B, Biegler LT. Optimization of supply chain systems with price elasticity of demand. *INFORMS J Comput*. 2011;23(4):557-568.
83. Zhao X, You F. Consequential life cycle assessment and optimization of high-density polyethylene plastic waste chemical recycling. *ACS Sustain Chem Eng*. 2021;9(36):12167-12184.
84. AzariJafari H, Yahia A, Amor B. Removing shadows from consequential LCA through a time-dependent modeling approach: policy-making in the road pavement sector. *Environ Sci Technol*. 2019;53(3):1087-1097.
85. Arvidsson R, Tillman A-M, Sandén BA, et al. Environmental assessment of emerging technologies: recommendations for prospective LCA. *J Ind Ecol*. 2018;22(6):1286-1294.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Luo Y, Ierapetritou M. Design and operation of modular biorefinery supply chain under uncertainty using generalized Benders decomposition. *AIChE J*. 2024;70(8):e18458. doi:10.1002/aic.18458