Optimization models for biorefinery supply chain network design under uncertainty

Narges Kazemzadeh\textsuperscript{a)} and Guiping Hu\textsuperscript{b)}

Department of Industrial and Manufacturing Systems Engineering, Iowa State University, Ames, Iowa 50011, USA

(Received 5 May 2013; accepted 10 September 2013; published online 24 September 2013)

Biofuel industry has attracted much attention due to its potential to reduce dependency on fossil fuels and contribute to the renewable energy. The high levels of uncertainty in feedstock yield, market prices, production costs, and many other parameters are among the major challenges in this industry. This challenge has created an ongoing interest on studies considering different aspects of uncertainty in investment decisions of the biofuel industry. This study aims to determine the optimal design of supply chain for biofuel refineries in order to maximize annual profit considering uncertainties in fuel market price, feedstock yield, and logistic costs. In order to deal with the stochastic nature of parameters in the biofuel supply chain, we develop two-stage stochastic programming models in which Conditional Value at Risk (CVaR) is utilized as a risk measure to control the amount of shortage in demand zones. Two different approaches including the expected value and CVaR of the profit are considered as the objective function. We apply these models and compare the results for a case study of the biomass supply chain network in the state of Iowa to demonstrate the applicability and efficiency of the presented models.

\textsuperscript{c)}Copyright 2013 AIP Publishing LLC. [http://dx.doi.org/10.1063/1.4822255]

I. INTRODUCTION

Biofuel, as an important source of energy, has created increasing interest in the past few years due to its environmental and economic benefits. One of the most significant advantages of biofuel is its potential to reduce dependency on fossil fuel. Moreover, second generation biofuel provides the benefit of avoiding competition with food production and promotes rural development in agricultural regions by using lignocellulosic biomass as feedstock.\textsuperscript{10}

U.S. Environmental Protection Agency (EPA) regulations affect the production and use of biofuel across the biofuel supply chain. EPA has proposed rules in a revised Renewable Fuel Standard (RFS-2) that govern how biofuels are produced and used in the U.S. RFS-2 has set a goal of producing $36 \times 10^9$ gallons of biofuels in 2022 as shown in Figure 1.

One of the most important aspects of biofuel production planning is the design of biomass supply chain networks. Thus far, numerous studies have been conducted on supply chain design of biorefineries.\textsuperscript{11,21,33} However, the biofuel industry has been challenged by the significant uncertainties along the biofuel supply chain such as the available feedstock supply, logistic costs, and consumer demands. Therefore, it is of great importance to consider the impacts of uncertainties to the biofuel supply chain network design.

There is a rich literature on supply chain design. An \textit{et al.}\textsuperscript{3} reviewed previous research on biofuel and petroleum-based fuel supply chain. Shah\textsuperscript{29} discussed the advantages and challenges of the process industry supply chain optimization. The author reviewed the studies in infrastructure design, modeling, analysis, planning, and scheduling together with some industrial examples. Bowling \textit{et al.}\textsuperscript{7} presented an optimization model with the objective of maximizing net
profit considering overall sales and the costs for the feedstock, transportation costs, capital costs for the facilities, and the operational costs for the facilities. Eksioglu et al.\textsuperscript{11} proposed a mathematical model to design the supply chain of biorefineries needed to produce biofuel and applied the model for the state of Mississippi in a case study. Gan\textsuperscript{12} developed an analytical framework for supply of biomass considering feedstock production, energy conversion, and environmental benefits/costs to minimize the total cost of both feedstock and electricity production and to determine the optimal power plant size. Sultana and Kumar\textsuperscript{12} used Geographic Information Systems to determine the optimal locations, sizes, and number of bio-energy facilities in Alberta, Canada while optimizing the transportation cost. An integrated mathematical model to determine the optimal comprehensive supply chain/logistics decisions to minimize the total cost is proposed by Zhang \textit{et al.}\textsuperscript{37} They showed the application of this model with a case study in the state of North Dakota. Vera \textit{et al.}\textsuperscript{34} developed a framework for finding the optimum location and capacity of a power plant fed with residues from olive oil producing areas.

An optimization model for the strategic design of a hybrid first/second generation ethanol supply chain is developed by Akgul \textit{et al.}\textsuperscript{1} This model addresses sustainability issues such as the use of food crops, land use requirements of second generation crops, and competition for biomass with other sectors. In another work,\textsuperscript{2} they proposed a multi-objective optimization model of hybrid first/second generation biofuel supply chains to analyze the trade-off between the economic and environmental objectives as well as the impact of carbon tax on the economic and environmental performance of the biofuel supply chain. The authors demonstrated the applicability of the model with a case study of bioethanol production in the UK. Kim \textit{et al.}\textsuperscript{20} developed a mixed integer linear programming model to determine the fuel conversion technologies, capacities, biomass locations, and the logistics of transportation from the locations of forestry resources to the conversion sites and then to the markets. Judd \textit{et al.}\textsuperscript{17} considered the impact of biomass crop yield, harvest method, and economies of scale in biorefinery capacities on the total cost. The problem of finding the best location for a biorefinery plant considering the local availability of biomass and geographical distribution of customers has been studied by Leduc \textit{et al.}\textsuperscript{22}

A number of studies considered dynamic models planning over multiple periods. Huang \textit{et al.}\textsuperscript{16} proposed a mathematical model that integrates spatial and temporal dimensions for strategic planning of future bioethanol supply chain systems. As a case study, the authors applied the model to investigate the economic potential and infrastructure requirements for bioethanol production from eight waste biomass resources in California. Sokhansanj \textit{et al.}\textsuperscript{31} developed a dynamic integrated biomass supply analysis and logistics model to simulate the collection, storage, and transport operations of supplying agricultural biomass to a biorefinery. A dynamic

![FIG. 1. Renewable Fuel Standards (RFS-2) goal (Source: EPA, EIA).](image-url)
nonlinear mixed integer programming model is developed by Shabani and Sowlati\textsuperscript{28} to maximize the overall value of the supply chain of forest biomass.

The majority of the literature on biofuel supply chain design assumes all the parameters in the system are known a priori. In biofuel supply chain, however, there is a high level of uncertainty that can be encountered in practice. Hence, it is important to develop approaches to deal with the uncertainties associated with the biofuel supply chain design.\textsuperscript{14,27}

A number of recent studies in this field have considered the uncertainties associated with the supply chain. Awudu and Zhang\textsuperscript{5} discussed uncertainties in biofuel supply chain management and reviewed related works. A dynamic mixed integer linear programming for strategic design and planning of a supply chain in a period of 10 yr was developed by Dal-Mas \textit{et al.}\textsuperscript{8} while considering uncertainty on biomass production cost and product selling price. The objective of their model was to minimize the expected net present value related to each scenario deriving from the combination of corn purchase costs and fuel ethanol market price. This model was used for the corn-to-ethanol production supply chain in Northern Italy as a test case. Sodhi and Tang\textsuperscript{30} introduced a two-stage stochastic model for supply chain management under uncertainty by applying a stochastic mixed integer non-linear method. Decisions such as the production topology, plant sizing, product selection, product allocation are considered. Kim \textit{et al.}\textsuperscript{19} proposed a two-stage mixed integer stochastic model to determine the size and location of the biorefineries. To design the problem in a manageable size, they considered only the bounds of the parameters. Marvin \textit{et al.}\textsuperscript{23} considered a mixed integer linear programming to determine optimal locations and capacities of biorefineries with biomass harvest and distribution. They also performed sensitivity analysis to verify the impact of price uncertainty on the decisions. Giarola \textit{et al.}\textsuperscript{13} general mixed integer linear programming modeling framework is developed to assess the design and planning of a multi-period and multi-echelon bioethanol upstream supply chain under market uncertainty considering economic and environmental (global warming potential) performance. Awudu and Zhang\textsuperscript{5} proposed a stochastic linear programming model for a biofuel supply chain under demand and price uncertainties within a single-period planning framework to maximize the expected profit. The decisions are to determine the amount of raw materials purchased, the amount of raw materials consumed, and the amount of products produced. A simulation model is another useful tool for supply chain management in biofuel industry due to the complexity and degree of uncertainty in such problems.\textsuperscript{15,18,24,31,36}

While it has been demonstrated that biofuel industry is more vulnerable to risk compared to many other industries,\textsuperscript{8} there are only a few studies dealing with the uncertainty in the biofuel supply chain design. The literature reviewed in this paper considered the uncertain parameters while maximizing the profit or minimizing the costs. One of the challenges, however, is to quantify the adverse impact of the uncertain parameters on demands satisfaction as well as the economic objectives. Feedstock supply is a main source of uncertainty in the biofuel supply chain, because it is highly dependent on the weather and can be negatively affected by pests or diseases. For instance, fluctuation of feedstock supply has a large impact on the level of satisfied biofuel demands. As a consequence, the system may not be able to meet all the demands, or there might be excesses of the supply. In addition, the uncertainty on the selling price of the biofuel, and logistic costs including transportation and operation costs related to the feedstock preparation at the field will directly impact the supply chain system.

In this study, we aim to develop a mathematical modeling framework to design a biorefinery supply chain considering uncertainties in fuel market price, feedstock supply, and logistic costs including transportation and operation costs. Mixed integer programming models with a two-stage stochastic programming approach were applied to address the uncertainties. The first-stage makes the capital investment decisions including the locations and capacities of the biorefineries. Once the first-stage decisions are determined, the second-stage determines the biomass and gasoline flows. The objective function is to maximize the annual profit which is revenue minus costs. Two different types of objectives were considered: expected value of profit, $E(\text{Profit})$, and conditional value at risk of profit, $\text{CVaR}(\text{Profit})$. The proposed models also illustrate the impact of incorporating CVaR in constraints on satisfying demand and controlling the amount of shortage in demand zones.
The rest of the paper is organized as follows: in Sec. II, the problem statement for biofuel supply chain is presented. Then, we discuss the stochastic programming models for this problem in Sec. III. In order to highlight the efficiency and applicability of the presented models, a case study in the state of Iowa and the results are presented in Sec. IV. Finally, we conclude the paper in Sec. V with summary of findings.

II. PROBLEM STATEMENT

The goal of this study is to develop a mathematical modeling framework to design a supply chain network for biofuel considering uncertainty in the system. The biofuel supply chain network consists of biomass production, harvesting, transportation, conversion, and fuel distribution. Figure 2 shows a schematic structure of the biofuel supply chain. In order to design the supply chain network, we developed two optimization models with different objective functions. These models determine the best locations of the biorefineries to maximize the profit while reducing the risk of biofuel shortages in demand centers. They also specify the amount of biomass transported from harvesting sites to biorefineries as well as the amount of gasoline shipped to the demand nodes.

The parameters used in the problem are defined as follows:

- Set of biomass feedstock harvesting sites.
- Feedstock availability at each harvesting site with the potential fluctuation of yield due to seasonality and weather conditions.
- Sustainability factor for each feedstock harvesting site.
- Feedstock collection and loading cost with a known probability distribution.
- Feedstock transportation cost with a known probability distribution.
- The distance between nodes of the supply chain network based on great circle distance.
- Set of potential biorefineries locations along with the possible set of capacity levels of each one.
- Set of demand zones with the amount of associated demand.
- Biofuel transportation costs.

Several assumptions are made in the presented models. We assume that the feedstock supply and the logistic costs (including transportation, collection, and loading costs) are uncertain due to high impacts of these parameters on the efficiency of the network. In these models, each biorefinery can be provided by more than one feedstock harvesting site, and each demand can be satisfied by more than one biorefinery. In addition, each harvesting site can serve more than one biorefinery and also each biorefinery can supply more than one demand zone.

The models in this paper are developed to design a biofuel supply chain network to maximize the profit and minimize the costs while controlling the biofuel shortage in demand centers. The objective function of the models is to maximize the total profit (revenue from selling biofuel deducted by total cost including collecting, transporting, and operational costs). The aim is to determine the locations and capacities of biorefineries, and the quantities of biomass feedstock shipped between harvesting sites and biorefineries, as well as the quantities of biofuel transported between biorefineries and demand zones.

FIG. 2. Structure of the biofuel supply chain.
III. MODEL FORMULATION

We formulate two stochastic programming models to maximize the profit in a biofuel supply chain network. The uncertainties in the models are defined with a set of uncertain parameters described by discrete distributions. Scenarios are generated based on the combination of the uncertain parameters. A two-stage stochastic programming approach was incorporated to investigate the decision making under the uncertainties. The fundamental idea behind two-stage stochastic programming is the concept of recourse, which is the ability to take corrective action after a realization of a scenario. The first-stage decisions involve variables that have to be decided before the actual value of uncertainties are realized. After the first-stage, the uncertainties are revealed, and the decision maker must choose an action that optimizes the objectives according to the realization of the scenario. In this problem, the first-stage decision is for the capital investment including the locations and capacities of the biorefineries. The second-stage variables are those that can be determined after the realization of the uncertain parameters. Once the uncertainties of available feedstock is resolved, the second-stage decisions are made, which include the flows of the biomass from harvesting sites to biorefineries and the flows of biofuel to demand zones.

We adopt the concept of Conditional value at Risk (CVaR) in the second objective function and in the constraints as a risk measure to incorporate the uncertainties design setting. As a consequence of uncertainties, there may be biofuel shortage for the demand zones. However, it is not desirable to have a large amount of shortage in a single demand node. Hence, CVaR is employed as a risk measure to control the shortage in each demand zone. The concept of CVaR is also employed in the objective function formulation. Uncertainties in biofuel market price and logistic costs are considered. We consider two different types of objectives: expected value of profit, \( E(\text{Profit}) \), and conditional value at risk of profit, \( \text{CVaR}(\text{Profit}) \). Although CVaR has its original meaning as a function typically used for loss distribution, it can be also defined for a profit distribution to decrease the risk associated with profit. In the remainder of this section, we will first explain the concept of CVaR for the loss distribution and the profit distribution. Then, we will elucidate the constraints in the models, and finally, the objective functions applied in the models are discussed.

A. Value at risk and conditional value at risk

A common way to incorporate risk-aversion concept into an optimization model is the use of Value at Risk (VaR) constraints. VaR is a popular measure for its comprehensibility; however, because of the conceptual and computational limitations, it is preferred to use CVaR constraints.\(^6,25,26\)

In this study, we used CVaR constraints to model the risk and uncertainty for the shortage of demand. In the definition of VaR and CVaR of a loss function, usually the tail on the right side of a probability density function is considered, so in this problem we also use the definition of CVaR for the tail on the right side of a probability density function of shortage.

The \( \text{VaR}_{1-\alpha} \) of a random variable of \( X \) is the lowest value of \( t \) such that, with probability \( \alpha \), the loss will not be more than \( t \), whereas the \( \text{CVaR}_{1-\alpha} \) is the conditional expectation of loss above that amount \( t \) (Ref. 26), that is,

\[
\text{VaR}_{1-\alpha}(X) = \inf \{ t : Pr(X \leq t) \geq 1 - \alpha \}, \]
\[
\text{CVaR}_{1-\alpha}(X) = E[X|X \geq \text{VaR}_{1-\alpha}].
\]

Figure 3 depicts the concept of VaR and CVaR of loss or shortage associated with \( \alpha \) percentile for a continuous distribution. Since the stochastic parameters in this study are assumed to be discrete distributed, the shortages on demand are defined in a discrete distribution as well. Another representation of CVaR\(_{(1-\alpha)}\) for a discrete distribution is

\[
\text{CVaR}_{1-\alpha}(X) = \inf \left\{ t + \frac{1}{\alpha} E[(X - t)_+] \right\},
\]

where \( (a)_+ = \max\{0, a\} \).\(^9\)
In the biofuel supply chain design, CVaR of loss (shortage of fuel demand in this study) is chosen as a criterion to control the risk of fuel shortage in demand areas. A constraint which limits the upper bound of the CVaR of shortage is incorporated in the model.

Although CVaR is typically defined for an adverse distribution in literature of finance, it can be defined for a favorable distribution such as the distribution of profit. In this study, CVaR is also utilized to incorporate the uncertainty for the profit. For a distribution of the profit, the definition of VaR and CVaR is considered for the tail on the left side of a probability density function.

The VaR$_{1-\beta}$ of a random variable of $X$ is the highest value of $t$ such that, with probability $\beta$, the profit will not be less than $t$, whereas the CVaR$_{1-\beta}$ is the conditional expectation of profit below that amount $t$, as follows:

$$\text{VaR}_{1-\beta}(X) = \sup \{t : Pr(X \geq t) \geq 1 - \beta\},$$

$$\text{CVaR}_{1-\beta}(X) = E[X | X \leq \text{VaR}_{1-\beta}].$$

Figure 4 shows VaR and CVaR of profit associated with $\beta$ percentile. For a discrete distribution, another representation of CVaR$_{1-\beta}$ is

$$\text{CVaR}_{1-\beta}(X) = \sup \left\{ t \in \mathbb{R} : Pr(t \leq X) \geq 1 - \beta \right\}. \tag{2}$$

B. Constraints in the model

In this section, we present a two-stage stochastic programming formulation for biofuel supply chain network design where locations for biorefineries are assumed to be centroid of the counties and demand nodes are based on Metropolitan Statistical Areas (MSAs). We assume that the available feedstock, the price, collection and loading costs, and biomass transportation costs have discrete distribution. Table I describes the notations used in the model.

The first-stage constraints of the model enforce the selection of biorefinery locations. A set of binary variables, $\delta_{lj}$, is defined to determine whether a biorefinery with capacity level of $l$ is located in a candidate location $j$. To ensure that the cost of building biorefineries does not exceed the available budget $B$, the following constraint is used:

$$\sum_j \sum_l c_l^j \delta_{lj} \leq B. \tag{3}$$
In each candidate location, only one biorefinery can be built, which is specified by the following constraints:

$$\sum_l \delta_{lj} \leq 1, \quad \forall j \in N.$$  (4)

The rest of the constraints refer to the second-stage decisions which specify the amount of feedstock and biofuel flows among the nodes of the supply chain network based on which scenario happens considering supplies and demands, respectively.

In our models, the biomass supply is assumed to be uncertain with a known distribution from which we take samples, called scenarios and represented by $S$. Given the set of counties, $N$, that produce biomass feedstock, each county $i \in N$ has $A_{is}$ tons per year of corn stover in scenario $s$ available. A sustainability factor of the corn stover, $S_i$, must remain in the field to provide winter cover and prevent soil erosion. Therefore, each county can provide at most $(1 - S_i)A_{is}$ tons of corn stover per year in scenario $s$.

It is assumed that transport distances within one county are negligible in feedstock transportation costs. Each county, $j \in N$, can be a candidate for a biorefinery facility with the capacity of $U_{lj}$. The flow of the feedstock from biorefinery $i$ to the biorefinery facility $j$ in
scenario \( s \) is denoted by \( f_{ij} \). The total quantity of feedstock transported from county \( i \) cannot exceed the amount of feedstock available at the county in each scenario, which is satisfied by

\[
\sum_j f_{ij} \leq (1 - S_i)A_{is}, \quad \forall i \in N, \quad \forall s \in S. \tag{5}
\]

Capacity constraints are also incorporated in the model. The total flow of feedstock into the biorefinery facility is \( \sum_i f_{ij} \). The material loss factor \( e_j \in [0, 1] \) accounts for possible losses during loading, transportation, and unloading at county \( j \). Factor \( e_j \in [0, 1] \) is feedstock dependent. Therefore, the amount of feedstock that can be processed to biofuel at a facility is less than or equal to the capacity, \( U_{lj} \), in county \( j \) in each scenario, which is denoted by

\[
(1 - e_j) \sum_i f_{ij} \leq \sum_j U_{lj} \delta_{lj}, \quad \forall j \in N, \quad \forall s \in S. \tag{6}
\]

The biorefineries convert the biomass feedstock into biofuel which will be shipped to the MSAs. Decision variable \( q_{jks} \) represents the quantity of biofuel shipped from biorefinery \( j \) to the MSA \( k \) under the scenario \( s \). In a scenario \( s \), biofuel shipped from biorefineries to a certain MSA \( k \) may not satisfy its demand \( G_k \). The shortage is represented by \( sh_{ks} \), as shown in constraint (7)

\[
\sum_j q_{jks} + sh_{ks} = G_k, \quad \forall k \in M, \quad \forall s \in S. \tag{7}
\]

It is assumed that all the biomass shipped to a biorefinery is converted to biofuel, where \( Y \) is a conversion factor associated to the production yield. This is represented by

\[
(1 - e_j) \sum_i f_{ij}Y = \sum_k q_{jks}, \quad \forall j \in N, \quad \forall s \in S. \tag{8}
\]

As discussed earlier, the feedstock available to convert to biofuel may not be enough to satisfy all the demands; therefore, there may be shortages in MSAs. To manage the amount of shortages in demand zones, CVaR is employed as a risk measure. The decision makers have the flexibility to determine the limits on the CVaR of shortage which is denoted by \( H \). Based on the definition of CVaR for a discrete distribution, to enforce a limit on CVaR of shortage associated with \( z \)-quantile, i.e., \( \text{CVaR}_{1-z}(sh) \leq H \), constraints (9)–(11) are included

\[
\eta + \frac{1}{z} \sum_s w_s r_s \leq H, \quad \forall s \in S. \tag{9}
\]

\[
r_s \geq sh_{ks} - \eta, \quad \forall k \in M, \quad \forall s \in S, \quad \forall s \notin S. \tag{10}
\]

\[
r_s \geq 0, \quad \forall s \in S. \tag{11}
\]

Note that these constraints are based on linearization of Eq. (1) by introducing auxiliary variables \( r_s \) and \( \eta \).

According to constraints (3), we can derive valid inequalities formulated as the following:

\[
\sum_j \delta_{lj} \leq \lfloor B/C^b_l \rfloor, \quad \forall l \in L. \tag{12}
\]

As these constraints make the feasible region tighter, this will facilitate problem solution process, making it more efficiently. This is employed in the case study section.

C. Objective function

The objective of the models is to maximize the expected profit which is defined as the revenue from selling the biofuel subtracted by the total cost. Various types of costs are incurred in the biofuel supply chain network. The first one is the unit cost of collection and loading of feedstock shipped and delivered to the biorefinery facilities, which is denoted by \( C_{is} \). The other one
is $C_{s}^{ST}$ which refers to the unit transportation cost for biomass feedstock. The collection and loading cost and transportation cost are highly dependent on the economic/market conditions, and thus $C_{s}^{SC}$ and $C_{s}^{ST}$ are based on the expected value of the costs. Assuming the distance between county $i$ and $j$ as $D_{ij}$, the total expected cost of loading, collection, and transportation of biomass feedstock is $\sum_{s} (C_{s}^{SC} + \tau D_{ij} C_{s}^{ST}) w_{dij}$. Here $\tau$ is a tortuosity factor that accounts for the actual distance that must be traveled due to the available geography and transportation infrastructure.

$C_{s}^{CGC}$ is a unit conversion cost to produce a gallon of biofuel at the biorefinery. The total conversion cost is thus $\sum_{s} D_{ij} C_{s}^{CGC} w_{ij}$. Biofuel is shipped to the MSA by pipelines at a unit cost of $C_{s}^{CGT}$, so the total biofuel transportation cost equals $\sum_{s} Q_{ij} C_{s}^{CGT} w_{ij}$. Total capital cost to build the biorefineries is $\sum_{l} C_{l}^{F} \delta_{lj}$. Since we consider the annual profit in the objective function, we adopt the amortized capital investment concept. Therefore, the annual payments for a period of $t = 30$ years with interest rate of $z = 8\%$ is

$$\text{PMT(Investment)} = \frac{z (1 + z)^t}{(1 + z)^t - 1}.$$

To compute the profit, we need to calculate the revenue. The expected price biofuel sold at in MSA $k$ is denoted by $P_{k}$. Therefore, the total revenue obtained by selling the product is $\sum_{j,k,s} P_{k} w_{djs}$. The total profit can be defined as the total revenue subtracted by the total costs.

To maximize the total profit, two modeling approaches are considered. The first is to maximize the expected value of the total profit which is referred to as $E(\text{Profit})$ in the rest of the paper. The model with objective of $E(\text{Profit})$ is formulated as follows:

$$\begin{align*}
\max & \quad \sum_{j,k,s} P_{k} w_{djs} - \sum_{i,j,s} (C_{s}^{SC} + \tau D_{ij} C_{s}^{ST}) w_{dij} - \sum_{j,k,s} (C_{s}^{CGC} + D_{ij} C_{s}^{CGT}) w_{ij} \delta_{lj} - \text{PMT} \left( \sum_{l} C_{l}^{F} \delta_{lj} \right), \\
\text{s.t.} & \quad \text{Constraints (3) - (11),} \\
& \quad f_{ij} \geq 0, \quad \forall i, j \in N, \quad \forall s \in S, \\
& \quad q_{djs} \geq 0, \quad \forall k \in M, \quad \forall s \in S, \\
& \quad sh_{djs} \geq 0, \quad \forall k \in M, \quad \forall s \in S, \\
& \quad \delta_{lj} \in \{0, 1\}, \quad \forall j \in N, \quad \forall l \in L.
\end{align*}$$

It should be noted that risks associated with profit are not explicitly considered in the first approach with objective of $E(\text{Profit})$. Therefore, in the second approach, we adopt the CVaR of profit for objective function to maximize the profit in the cases of unfavorable scenarios.

The goal of the second approach is to maximize the CVaR of the total profit which is referred to as $CVaR(\text{Profit})$ in the rest of the paper. In other word, the objective function can be viewed as maximization of the expected value of $\beta$-percentile of the worst case of the total profit. The notation related to the new assumptions is updated in Table II. Variables $\zeta$ and $\nu_{s}$ are applied to formulate and linearize CVaR of the profit according to the definition of CVaR for the discrete distribution.

The model with the objective of $CVaR(\text{Profit})$ associated with $\beta$-percentile is formulated as follows. The objective function used in this model is a linearization of Eq. (2) by introducing auxiliary variables $\nu_{s}$ and $\zeta$.

<table>
<thead>
<tr>
<th>TABLE II. Updated parameters for the stochastic model with objective of $CVaR(\text{Profit})$.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Profit,</td>
</tr>
<tr>
<td>Revenue,</td>
</tr>
<tr>
<td>Cost,</td>
</tr>
<tr>
<td>$\zeta$, $\nu_{s}$</td>
</tr>
</tbody>
</table>
IV. CASE STUDY

The stochastic mixed integer linear models proposed in this study are aimed to design a biorefinery supply chain with the consideration of uncertainties. The problem is formulated in two mathematical models with two different objective functions: \( E(\text{Profit}) \) and \( CVaR(\text{Profit}) \). The models consider the uncertainties in the fuel market price, feedstock supply, and logistic costs. A novelty in the proposed models is to consider the control of the shortage of biofuel for demand zones based on the CVaR of shortage.

In this case study, we examine the supply chain network design for conversion of biomass into biofuel in the state of Iowa. Biomass can be harvested and collected in every county in the state. The feedstock is then transported from county centroid to the biorefineries for conversion to biofuel. The biofuel is transported to the demand areas, which are based on the MSAs in Iowa. It is assumed that the transportation distance within the county has a negligible effect on feedstock transportation costs. The goal is to determine the optimal biorefineries locations and capacities with the objective of maximizing the annual profit while controlling the risk of the biofuel shortages at the MSAs.

In this section, we first explain the data used in this case study. Then, we analyze and discuss the model output and draw managerial insights for biofuel supply chain network design.

A. Data sources for the case study

In the state of Iowa, there are 99 counties which are potential biomass harvesting locations. Each county is also considered as a candidate location to build a biorefinery with capacity level of 1000, 1500, or 2000 ton per day for the conversion to biofuel. The maximum available budget assigned to this project is $3,000,000,000. We consider 21 MSAs in Iowa as the demand areas. Biofuel demand at each MSA is estimated as a percent of the state-level gasoline consumption as provided by Energy Information Administration (EIA). This percent is based on the ratio of the population within the MSA and the total population of the state.

The confidence levels to define the CVaR of shortage \( \alpha \) and CVaR of profit \( \beta \) are both assumed to be 20% in this case study. The impacts of different confidence levels are not within the scope of this study. The upper bound for biofuel shortage at MSAs is assumed to be \( H = 200,000,000 \) gallons per year.

Material loss factor \( e \), which accounts for possible losses during loading, transportation, and unloading, is assumed to be 0.05. Tortuosity factor \( \tau \) is considered 1.29, which is
multiplied by distances and shows the actual distances that must be traveled according to the geographical infrastructure. Based on the experimental data, the biorefinery process yield of feedstock, \( Y \), is assumed to be 0.218. The sustainability factors, \( S_i \), to be 0.718 at all counties.35

In this case study, the scenarios are generated based on the average values of the parameters and their deviation according to the historical records. We considered 16 scenarios for available feedstock, 3 scenarios for price of gasoline, 2 scenarios for feedstock collection and loading costs, and 2 scenarios for transportation cost. In our data, we used the price of gasoline as an estimate for the price of biofuel. Tables III list possible scenarios and their probabilities considered for each parameter. The combination of these scenarios constructs 192 scenarios in total for this problem.

B. Results analysis and discussion

The proposed models aim to determine capital investment decisions on the location and capacities of the biorefineries, the feedstock transportation, and biofuel delivery decisions. The first-stage decisions have to be made before the uncertainties are realized, and the second-stage decisions are made after the realization of the system parameters. In this study, the first-stage decisions include the capital investment decisions (the location and capacities of the biorefineries). Once the uncertainties are realized, the second-stage decisions are made which include the flows of the biomass from harvesting sites to biorefineries and the flows of biofuel to demand areas. The uncertainties considered in this problem consist of feedstock supply, fuel market price, and logistic costs. Two modeling approaches are adopted in the objective function formulation: expected value and CVaR of profit. In the first approach, the objective function is to maximize the expected value of profit. The profit of the project is an important performance

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Available feedstock</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scenario 1</td>
<td>( A - 8% A )</td>
<td>1/16</td>
</tr>
<tr>
<td>Scenario 2</td>
<td>( A - 7% A )</td>
<td>1/16</td>
</tr>
<tr>
<td>Scenario 3</td>
<td>( A - 6% A )</td>
<td>1/16</td>
</tr>
<tr>
<td>Scenario 4</td>
<td>( A - 5% A )</td>
<td>1/16</td>
</tr>
<tr>
<td>Scenario 5</td>
<td>( A - 4% A )</td>
<td>1/16</td>
</tr>
<tr>
<td>Scenario 6</td>
<td>( A - 3% A )</td>
<td>1/16</td>
</tr>
<tr>
<td>Scenario 7</td>
<td>( A - 2% A )</td>
<td>1/16</td>
</tr>
<tr>
<td>Scenario 8</td>
<td>( A - 1% A )</td>
<td>1/16</td>
</tr>
<tr>
<td>Scenario 9</td>
<td>( A + 1% A )</td>
<td>1/16</td>
</tr>
<tr>
<td>Scenario 10</td>
<td>( A + 2% A )</td>
<td>1/16</td>
</tr>
<tr>
<td>Scenario 11</td>
<td>( A + 3% A )</td>
<td>1/16</td>
</tr>
<tr>
<td>Scenario 12</td>
<td>( A + 4% A )</td>
<td>1/16</td>
</tr>
<tr>
<td>Scenario 13</td>
<td>( A + 5% A )</td>
<td>1/16</td>
</tr>
<tr>
<td>Scenario 14</td>
<td>( A + 6% A )</td>
<td>1/16</td>
</tr>
<tr>
<td>Scenario 15</td>
<td>( A + 7% A )</td>
<td>1/16</td>
</tr>
<tr>
<td>Scenario 16</td>
<td>( A + 8% A )</td>
<td>1/16</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Price of gasoline</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scenario 1</td>
<td>( P - 10%P )</td>
<td>1/3</td>
</tr>
<tr>
<td>Scenario 2</td>
<td>( P )</td>
<td>1/3</td>
</tr>
<tr>
<td>Scenario 3</td>
<td>( P + 10%P )</td>
<td>1/3</td>
</tr>
</tbody>
</table>
measure to evaluate the effectiveness of the decision. However, the expected profit approach does not explicitly address the risk of decision making under the unfavorable events. In order to manage the system risks, we adopted CVaR of profit as the second approach in the objective function.

It should be noted that it is of great importance to control the shortage of demand in the system. One of the challenges in this model is incurring a large amount of shortage in a single demand MSA. We design a risk measure in the constraints to level the shortage and decrease the probability of larger shortage occurring in the network. We consider CVaR of the shortage and set an upper bound on that to control the risk of shortage through demand areas.

In the case study, the state of Iowa is selected due to the data availability to demonstrate the effectiveness and applicability of the proposed modeling framework. There are 99 counties in Iowa with biomass feedstock supply, each of which is considered as a candidate location for biorefinery. The demand zones are 21 MSAs located in the state. We implemented the proposed models with different assumptions in the case study to compare and analyze the results: the model with objective function of $E(\text{Profit})$ with and without the CVaR constraints on the shortage, and also the model with objective function of $\text{CVaR}(\text{Profit})$ with and without the CVaR constraints on the shortage.

We implement two proposed models in this case study and compare them to the models with the same assumption but without considering CVaR constraints on shortages. Model (A) refers to the model with the objective of $E(\text{Profit})$, and Model (B) is the model with the objective of $\text{CVaR}(\text{Profit})$. These models are implemented in CPLEX Python API version 12.2.

- In Model (A), the objective is to maximize $E(\text{Profit})$. At first, we implement this model while there is no control on the shortage of demand. The version of model (A) without considering the constraints on shortages is as follows:

$$
\max \sum_{j,k,s} P_{ks} w_s d_{jks} - \sum_{i,j,s} (C_{is}^{SC} + \tau D_{ij}C_{si}^{ST}) w_{ijs} - \sum_{j,k,s} (C_{ij}^{GC} + D_{jk}C_{jT}^{GT}) w_s d_{jks} - \text{PMT} \left( \sum_{f} C_{fi}^T \delta_{fj} \right),
$$

s.t. Constraints (3) – (8), Constraints (12),

$$
f_{ijs} \geq 0, \quad \forall i, j \in N, \quad \forall s \in S,
q_{jks} \geq 0, \quad \forall k \in M, \quad \forall s \in S,
sh_{ks} \geq 0, \quad \forall k \in M, \quad \forall s \in S,
\delta_{ij} \in \{0, 1\}, \quad \forall j \in N, \quad \forall l \in L.
$$

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Feedstock collection and loading cost</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scenario 1</td>
<td>$C_{is}^{SC} - 10% C_{is}^{SC}$</td>
<td>1/2</td>
</tr>
<tr>
<td>Scenario 2</td>
<td>$C_{is}^{SC} + 10% C_{is}^{SC}$</td>
<td>1/2</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Transportation cost</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scenario 1</td>
<td>$C_{ij}^{ST} - 10% C_{ij}^{ST}$</td>
<td>1/2</td>
</tr>
<tr>
<td>Scenario 2</td>
<td>$C_{ij}^{ST} + 10% C_{ij}^{ST}$</td>
<td>1/2</td>
</tr>
</tbody>
</table>
Figure 5 shows the results from model (A) without considering shortage constraints. As shown in Figure 5, there is a large amount shortage in one MSA, which is about 528,000 gallons per year. This motivated the use of a risk measure to control the shortage through MSAs.

- In this case we consider Model A which include constraints (9)–(11) as the CVaR constraints on the shortage. Figure 6 shows that when we add CVaR constraints on the shortage to the model with the objective of $E(Profit)$, the shortages are split in a more reasonable way, such that the system will not incur that large amount of shortage in any single MSA. It should be noted that the number of MSAs with biofuel shortage is increased. In other words, after incorporating the CVaR constraints on the shortage, the system shortage is more dispersed in the system which mitigates the system risks. In addition, after incorporating constraints (9)–(11) the total amount of shortage decreases in this model. This is due to the limit forced on the shortage. In this model, the expected value of profit decreased about 4% due to the additional constraints added to the model.

- Model (B) considers the objective of $CVaR(Profit)$. The following formulation refers to this model while there is no control on the biofuel shortages.

![Figure 5](image1.png) Biorefineries locations for the model with the objective of $E(Profit)$ (Model A) without considering CVaR constraints on shortage of demand.

![Figure 6](image2.png) Biorefineries locations for the model with the objective of $E(Profit)$ (Model A) with considering CVaR constraints on shortage of demand.
According to Figure 7, the results show that the amounts of shortage are very large in three MSAs. The total amount of shortages is more than total shortages in Model (A). The reason is that Model (A) tries to maximize the expected profit without the risk control of unfavorable events in the objective function; however, Model (B) attempts to maximize the profit in the averse conditions which is associated with the system risks.

- Now we consider Model (B) while enforcing an upper bound on CVaR of shortage in order to avoid concentrated biofuel shortages for the MSAs by adding constraints (9)–(11) to the model.

The results from the model with objective of \( \text{CVaR(Profit)} \) with the CVaR of shortage constraints is shown in Figure 8. When we add the CVaR of shortage constraints, the amount of shortage in a single MSA is dispersed which is similar to Model (A). As shown in Figure 8, although we have more MSAs with shortage, we do not have any concentrated shortage in a single MSA as we had from Model (B) without CVaR constraints. Moreover, the total shortage is less than the same model without considering CVaR constraints on the shortage.
After applying the constraints on the shortage in this model, the expected value of profit increased about 8% although the objective value (i.e., CVaR of profit) decreased due to the additional constraints. However, model B resulted in smaller profit compared to model A. This is because that model B tries to improve the profit in the worst cases, while model A aims to maximize the expected value of profit.

The observations from both models indicate that using CVaR constraints is a reasonable approach to address the risk of the shortage. It can be applied in the system in which the risk of occurring large amounts of shortage in a single MSA is expensive. The reason is that the constraints of the model make the inevitable shortage to be split through all the MSAs according to parameter $\alpha$ in the CVaR, and therefore it is not allowed to have a large amount of shortage in a single MSA. In addition, comparison of Model A and B, regardless of CVaR constraints, shows that model B is more appropriate for more conservative decision makers because of the property of risk-aversion embedded in its objective function. This risk aversion property can be set according to the decision maker preference by changing parameter $\beta$ in the CVaR, in the objective function. As stated before, studying changes in parameter $\alpha$ and $\beta$ was not included in the scope of this study.

In summary, comparisons between models with two different objective functions indicates that unsurprisingly the models with the objective function of $E(\text{Profit})$ provide smaller shortages, whereas the models with the objective function of $C\text{VaR}(\text{Profit})$ yield larger shortages. In addition, models without the CVaR constraints on the shortage result in a larger concentrated amount of shortages in the MSAs which is due to that there is no upper bound on the amount of shortage in a specific demand area. However, the models with CVaR constraints on the shortage result in more MSAs with shortages, but the amount of shortage in each MSA is reduced. In other words, enforcing an upper bound on the CVaR of the shortage prevent the occurrence of a large amount of shortage in a single MSA. This result is as expected, as the CVaR constraints set a limit on the amount of shortage in a single MSA.

V. CONCLUSION

Biofuels play an important role in providing clean and secure energy and promoting economic growth. One of the most important and challenging issues of biofuel production is biofuel supply chain network design. The general structure of biofuel supply chain consists of biomass production, harvesting, transportation, conversion, and fuel distribution. The biomass is harvested at the farms and shipped to the biorefineries. At biorefineries, the feedstock is converted to biofuel and then transported to demand areas. In the research arena of biofuel supply
chain network design, one of the biggest challenges is to deal with uncertainties along the supply chain.

The goal of this study is to explore the design of a biofuel supply chain network under uncertainty. We proposed a mathematical programming framework with the approach of two-stage stochastic programming to determine capital investment decisions on the location and capacities of the biorefineries, the feedstock transportation and biofuel delivery decisions. The uncertainties considered in this problem consist of feedstock supply, fuel market price, and logistic costs. Two modeling approaches are adopted in the objective function formulation: expected value and CVaR of profit.

To sum up, this study provided a mathematical modeling framework to the biofuel supply chain network design under uncertainty. Two types of objective functions: expected value of profit and CVaR of profit were considered. The first approach focuses on maximize the expected profit where the latter approach is more on the mitigation of system risk under adverse conditions. The impacts of incorporating the stochastic shortage control are also investigated by incorporating the CVaR of shortage as a constraint in the model.

We conclude the paper by pointing out two future research directions. Biofuel supply chain network design depends on many parameters and factors. However, the proposed method only provides a basic framework to study the biofuel supply chain under uncertainty. It is suggested to extend these models to consider additional operational assumptions in future studies. In addition, the larger the number of scenarios, the more accurate the decisions would be. Consequently, the computational complexity would substantially increase. Therefore, exploring more efficient algorithms to solve the problem could be another direction for future work in this area.


