Accepted Manuscript

Supplier selection and operation planning in biomass supply chains with supply uncertainty

Duc Huy Nguyen, Haoxun Chen

PII: S0098-1354(18)30739-7
DOI: 10.1016/j.compchemeng.2018.07.012
Reference: CACE 6166

To appear in: *Computers and Chemical Engineering*

Received date: 21 December 2017
Revised date: 15 June 2018
Accepted date: 16 July 2018

Please cite this article as: Duc Huy Nguyen, Haoxun Chen, Supplier selection and operation planning in biomass supply chains with supply uncertainty, *Computers and Chemical Engineering* (2018), doi: 10.1016/j.compchemeng.2018.07.012

Highlights

- Stabilization of feedstock supply in biomass supply chains considered
- Stochastic programming model for biomass supplier selection and operation planning
- An enhanced and regularized decomposition algorithm proposed to solve the model
- Performance of the algorithm evaluated by numerical experiments
- Sensitivity analysis conducted to evaluate the impacts of key parameters

Supplier selection and operation planning in biomass supply chains with supply uncertainty

Duc Huy NGUYEN[a], Haoxun CHEN[a,*]

[a]*Laboratory of Industrial Systems Optimization, Charles Delaunay Institute, University of technology of Troyes, France*

Abstract

Bioenergy is considered a potential solution to reduce carbon footprint and fight against global warming. However, uncertainty in the harvest of biomass could lead to the instability of feedstock supply that has a significant impact on the sustainability of biomass supply chain. In this paper, we present a two-stage stochastic programming model dealing with supplier selection to stabilize feedstock supply of a biomass supply chain in uncertain environments. The model involves the first stage decisions for the supplier selection and the second-stage decisions for planning transportation, inventory and production operations. To reduce the computational burden for large instances, we propose an enhanced and regularized L-shaped decomposition algorithm to solve the model. The applicability of this model and the performance of the solution method are evaluated by numerical studies. Sensitivity analysis shows that the values of some parameters related to suppliers have significant impacts on the optimal expected cost and supplier selection.

Keywords: Biomass supply chain, Supplier selection, Uncertainty, Stochastic programming

1. Introduction

Nowadays, humanity is facing enormous challenges of global warming and the exhaustion of fossil resources according to Shafiee and Topal (2009). The use of fossil fuel is one of the major causes of global warming that affects severely human life and increases natural disaster in recent decades. Therefore, renewable energy will play a significant role in the energy transition and appears to be a very promising solution to sustainable development in the future. The various renewable resources require different technologies to capture energy such as solar, hydro-power, wind, geothermal, marine and biological resources. They are clean energy sources and do not contribute to global warming, greenhouse effect and air pollution. However, the exploitation of renewable energy sources requires much time, high costs of investment and incentive policies from the government.

*Corresponding author

Email addresses: duc_huy.nguyen@utt.fr; nguyenduchuyvn@gmail.com (Duc Huy NGUYEN), haoxun.chen@utt.fr (Haoxun CHEN)

Bioenergy is generated from sources such as wood, agricultural products, animal and plant wastes. In contrast to wind and solar power, biomass can be stored in existing infrastructures in response to various demand according to Kaut et al. (2015). Nowadays, bioenergy industry has to deal with challenges related to low energy density, high harvest, and transportation costs of biomass to compete with fossil energy. The research in Ekşioğlu et al. (2009) shows that an efficient supply chain design and management could reduce the biofuel price significantly.

The sustainability of a biomass supply chain depends on the supply network design and uncertainty management considering feedstock supply, bioenergy demand, price. To stabilize the feedstock supply, the study in Li et al. (2011) concludes that lignocellulosic biomass (corn, wheat, and switchgrass...) can be used as feedstock and is interchangeable in bioenergy production because their compositions are very similar. The multi-biomass approach could be a potential solution, which uses a mix of feedstocks (corn, wheat, and switchgrass) harvested/collected in different time periods. According to Andersson et al. (2002), biomass storage facilities in a biorefinery site could reduce the dependency of feedstock seasonality. Such storage facilities allow us to collect enough biomass feedstock during low demand periods to meet the peak demand. An alternative is to increase biofuel production during low-cost periods as a buffer storage for peak demand periods.

Although biomass is a cheap raw material, costs associated with transportation, storage, and preprocessing might lead to a significant increase in total cost. Moreover, the reliability of feedstock supply considerably affects the efficiency of a biomass supply chain, so building a stable relationship with biomass suppliers is very important to assure a steady supply of high-quality biomass at low prices. The partnership established with suppliers should be integrated with long-term planning to cope with uncertain environments.

To the best of our knowledge, no previous work studied a multi-period biomass supply chain model coping with supplier selection and the planning of transportation and production operations in uncertain environments. For this reason, we present a stochastic programming model that can be used to capture most activities in a biomass supply chain and the uncertainty in feedstock supply. The main contributions of this paper include:

Firstly, we propose a two-stage stochastic programming model for biomass supply chain management. In the model, the first stage decision is to select suppliers who provide long/mid term feedstock supply. The second-stage decisions determine the amount of biomass to purchase, to transport, to pre-process and the volume of bioethanol to produce in each period.

Secondly, we develop an enhanced and regularized L-shaped decomposition method to solve the stochastic programming model with a given number of scenarios. The Monte Carlo sampling approach is used to find the required number of scenarios for ensuring that the exact value of an output of the model is located within a confidence interval with a given probability.

Thirdly, we conduct a numerical study to evaluate the performance of the proposed algorithm. The numerical results show that the algorithm can find an optimal solution in a reasonable computation time while a commercial solver cannot for large instances. Sensitivity analysis is performed to investigate the effect of some critical parameters on the optimal expected cost and suppliers selection.

This paper is organized as follows: Section 2 describes the main characteristics of biomass supply chains and gives a literature review. In Section 3, the problem studied and its model are introduced. We describe the solution method for the model in Section 4. A numerical study is presented, and its results are analyzed in Section 5. Section 6 concludes this paper with some remarks for future research.

2. Literature review

In this section, we will review relevant problems and studies that arise in the design and management of biomass supply chain, such as uncertain in biomass availability and supplier selection. Moreover, an overview of optimization models and solution approaches for biomass supply chain design and operation planning is presented.

2.1. Biomass supply chain

In comparison with traditional supply chains, biomass supply chains have many different features regarding their structures and characteristics. Their structures could be naturally classified according to logistics activities required by supplying biomass from farmers to production sites and delivering biofuel from the production sites to service stations (gas stations). These activities include ground preparation and plantation, cultivation, harvesting, handling, storage, conversion, transportation and utilization of the biofuel at service stations. Unlike fossil energy, biomass decays during storage, at a rate estimated to be 1-2% of the material stocked per month under the ambient storage, according to Rentizelas et al. (2009).

Different types of uncertainty present in almost all logistics activities of a biomass supply chain, not only arise from external environments, but also from the supply chain itself Bot et al. (2015). Firstly, the feedstock supply is always seasonal and depends on weather conditions and harvest time. In addition, a plant's location has a significant impact on the sustainability of feedstock supply thanks to geographical dispersion. Due to the heterogeneous and low-density characteristics, pretreatment operations are essential to homogenize and increase biomass density that may help us to reduce transportation costs of biomass significantly. Moreover, the fluctuation of purchase price and biofuel demand affects the efficiency of a supply chain network eventually. In some cases, it is hard to find an optimal network design because of uncertainty (fluctuation) n price and demand. The efficiency of a biomass supply chain is also affected by long-term contracts with suppliers, transportation and local distribution infrastructure, conversion technology and policies of the government according to Ba et al. (2016); Gold and Seuring (2011).

2.2. Supplier selection

Companies need to consider every essential factor during their decision-making process to increase competitiveness and respond to the requirements of markets rapidly.Scott et al. (2013) pointed out that the incoherent supplier selection could cause an unsustainable supply chain system such as unreliable operations of the bioenergy plant, shortage of fuel supply and greenhouse gas emissions. According to Govindan et al. (2013), supplier selection plays one of the most vital roles in the sustainability of a supply chain and in achieving its social, environmental, and economic goals.

Besides, uncertainty is a critical factor in the design and management of bioenergy supply chains according to Shabani and Sowlati (2016a). The study of Ekşioğlu et al. (2009) shows that transportation costs and biomass availability are two primary factors influencing the biomass supply chain design decisions. Their results also indicate that biomass collection costs and biomass processing costs have no significant impact on the decisions.

In a biomass supply chain, the supplier selection becomes more complicated and high challenging due to the presence of uncertainty. Different from traditional supply chains, biomass feedstock supply, is always uncertain because its availability and yield depend on seasonality (weather condition, temperature) and harvest time (Sharma et al., 2013; Yue et al., 2014; Ba et al., 2016). In summary, biomass availability is a critical factor that must be accounted for because it has a substantial impact on the sustainability of a biomass supply chain. Therefore, this uncertainty should be considered in supplier selection.

Although there are some papers dealing with supplier selection in other types of supply chain such as food supply chains (Amorim et al., 2016; Hammami et al., 2014), the study of supplier selection in a biomass supply chain was rare in the literature. There is no work considering strategical supplier selection along with tactical planning under uncertain environments although the two types de decisions are interdependent

In this paper, we propose an integrated model for supplier selection and operation planning in a biomass supply chain, which differs from the models mentioned above because of additional side constraints imposed by the biomass application. Moreover, we develop an enhanced and regularized L-shaped decomposition method for solving our model optimally. This method is different from the one proposed by Amorim et al. (2016), which applies Bender cuts to a Branch and Cut scheme, and the one proposed by Hammami et al. (2014), which solves the deterministic mixed integer programming equivalent of a stochastic programming model by using a commercial MILP solver CPLEX.

2.3. Optimization models for biomass supply chain design and operations

Mathematical programming and optimization techniques have been employed to design and manage biomass supply chains considering their different components and various decisions at strategic, tactical and operational levels according to Shabani et al. (2013). These mathematical models can be classified into two categories: deterministic and stochastic. Most studies ignored uncertainty by only considering deterministic models. A bibliographic analysis in Pérez et al. (2017) shows that 75% of studies which focus on economic and environmental criteria adopt deterministic models.

2.3.1. Deterministic models

A deterministic model is useful for ”what-if” scenarios where decision-makers can observe the outcomes of their decisions in different inputs/conditions. Earlier works on the design and management of biomass supply chains usually consider long-term decisions such as selecting production technologies, biomass preprocessing techniques, determining biomass storage and processing facilities' capacities and locations.

Samsatli et al. (2015) proposed a biomass supply chain model to evaluate the economic and environmental impacts under different scenarios and constraints concerning optimal resources and technologies selected. In Hombach et al. (2016), a mixed-integer linear programming (MILP) model was proposed to design a second generation biofuel supply chain in a German region under European biofuel regulations. Woo et al. (2016) proposed a MILP model for designing a supply chain network against fluctuations of biomass availability and hydrogen demand. The impact of biomass quality-related costs on operational costs in a biofuel supply chain was investigated in Castillo-Villar et al. (2016).

Few studies integrated activities in a biomass supply chain at the tactical decision level which involves transportation mode, harvest equipment and supplier selection such as Morales-Chvez et al. (2016); Ba et al. (2015); Duc-Huy et al. (2017). The impact of environmental uncertainty was ignored in their models. There are few mathematical models cope with the operational activities such as harvest schedule, production schedule, inventory control and vehicle routing problems. In Li et al. (2017), a new concept, called ”distance potential” was proposed, which could be used to optimize biomass logistics networks. A transportation planning problem in a forest biomass supply chain under work-time constraints and predetermined orders was presented in Han and Murphy (2012).

Based only on the expected value, deterministic models are not able to well capture events in the tails of the distributions of random variables. This disadvantage could lead to a negative impact on the performance of the models. For example, some scenarios with a low probability of occurrence might have a high impact on the total cost/profit of a system. So, stochastic models may be more appropriate.

2.3.2. Stochastic models

According to Sahinidis (2004), there are three principal approaches to cope with uncertainty: stochastic programming (recourse models, robust stochastic programming, and probabilistic models), fuzzy programming (flexible and possibilistic programming), and stochastic dynamic programming. In the literature for biomass supply chains, most mathematical models are deterministic and only few stochastic models exist. These stochastic models are usually based on two-stage stochastic programming with recourse to design supply chains at the strategic level (selecting production technologies, determining the location and capacity of each facility).

Awudu and Zhang (2013) studied a biofuel supply chain for biorefineries already placed and producing biofuels under demand and price uncertainties in North Dakota. They applied Bender decomposition algorithm and Monte Carlo method to approximate the value of second-stage based on a set of scenarios. Osmani and Zhang (2013) consider a similar problem but with objective to determine a location of biorefineries under more uncertainties (supply, demand, and prices). The solution technique is identical to those previously mentioned.

Kim et al. (2011) also study a similar problem but with objective to determine location and size of two types of conversion facilities with the high level of uncertainty in supply amounts, market demands, market prices, and processing technologies. This model is implemented on the commercial software GAMS and uses the CPLEX solver. Chen and Fan (2012) studied the same problem but consider separately two major sources of uncertainties, feedstock supply and fuel demand. They applied Progressive hedging algorithm to solve this model optimize a biomass supply chain under uncertainties in biofuel demand and feedstock supply.

Similar to Osmani and Zhang (2013), Kaut et al. (2015) want to determine the location of facility and flow of biomass in a supply chain with uncertainty in demand, supply and prices. They proposed a hybrid stochastic programming-robust optimization model and solved this MLIP model by using FICO Xpress Optimizer.

Most works only consider road transportation, except for the study of Marufuzzaman et al. (2014) focused on selecting transportation mode (rented truck, facility-owned trucks, and pipelines) and facility location of a biodiesel supply chain with uncertainties in feedstock and technology development. They investigate the impact of carbon regulatory mechanisms on the design and management of a supply chain. They proposed an algorithm combining Lagrangian relaxation and L-shaped method to solve this model.

Most researchers study supply chains for biofuel production whereas Shabani and Sowlati (2016b) investigate a forest-based biomass supply chain for a power-plant with uncertainty in biomass quality and biomass availability. This model is solved using the AIMMS software and CPLEX solver.

Almost works mentioned above consider biomass supply chains at only one period such as Chen and Fan (2012); Awudu and Zhang (2013); Kim et al. (2011); Marufuzzaman et al. (2014). In our model, we consider a biomass supply chain in multi-period and also dealing with uncertainty in feedstock supply such as Chen and Fan (2012). Moreover, we consider a new feature ”suppliers selection” in our models while these above papers deal with facilities' capacities and locations. That makes our paper different from these above papers.

There exist different methods to solve a two-stage stochastic programming model such as L-shaped algorithm in Awudu and Zhang (2013); Marufuzzaman et al. (2014), Progressive Hedging algorithm in Chen and Fan (2012). Some authors used a commercial solver to find an optimal solution of a deterministic equivalent linear program model such as Shabani et al. (2014); Shabani and Sowlati (2016b); Kim et al. (2011); Osmani and Zhang (2013); Kaut et al. (2015). In this paper, we develop an enhanced and regularized L-shaped decomposition method to tackle computational burden of the problem we study.

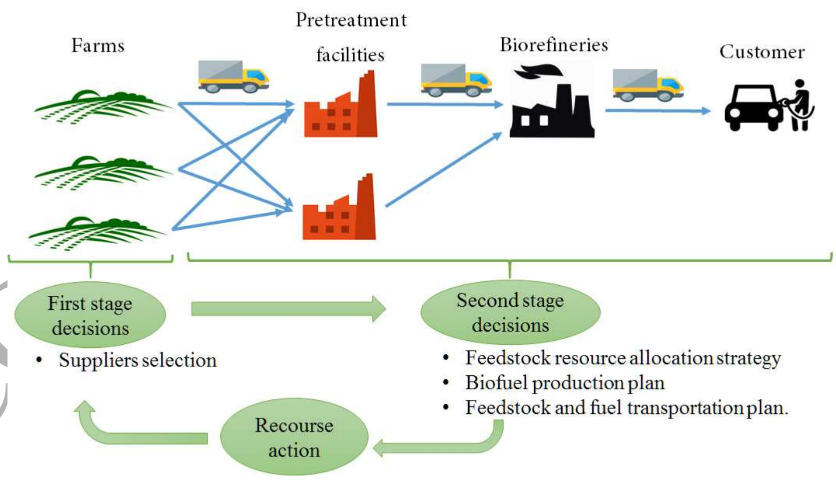

Figure 1: A two stage stochastic programming model with recourse for a biomass supply chain

3. Problem description and model formulation

3.1. Problem description

This paper studies a biomass supply chain under uncertainty in supply as shown in Figure 1. Biorefineries seek to establish a long-term relationship with certain local/regional suppliers to stabilize its feedstock supply. Under a mid/long-term contract, the unit purchase price and the minimum supply quantity are predetermined. The logistics activities occurred in each period are presented as follows:

- Biorefineries can buy biomass from international/national markets (if available) to meet part of biofuel demand only at a reasonable price of biomass.
- In each period, biomass is transported from contracted suppliers to preprocessing facilities, then to biorefineries. In each conversion facility, the supplied biomass feedstock is preprocessed and then converted into biofuel or be kept as inventory at the biorefinery.
- After satisfying the biofuel demand in each period, any excess volume of biofuel is kept as inventory. In case of shortfall in biofuel requirement, a relatively high shortage cost is charged for any unmet biofuel demand.
- Biomass preprocessed/biofuel produced is limited by biomass preprocessing/processing capacity and also by storage capacity in each facility. During storage, biomass has a ratio of loss about 1-2% per month.

We assume that the production infrastructure such as pretreatment facilities and biorefineries have been located. The unit transportation costs from local/regional suppliers to biorefineries are pre-calculated based on their distances and fixed costs are associated.

In the following, we introduce a two-stage stochastic programming model to tackle the supplier selection and operation planning in a biomass supply chain. The general framework of the two-stage stochastic programming model is given in the next subsection.

3.2. Two-stage stochastic programming model

In a two-stage stochastic programming model with recourse, the decision variables are divided into those of two different stages. In the first stage, the decision maker chooses decisions x associated with a cost $c^T x$ before knowing the actual realization of random parameters ξ. In the case of a discrete distribution, the random vector ξ has a finite number of possible realizations, called scenarios $s \in S$. In the second stage, an action associated with vector y_s (recourse action) can be taken to correct the adverse effects that result from the first stage decision x. The joint objective function of two-stage recourse model is to minimize the first stage cost and the expected cost of the second stage. In most applications, the first stage is often related to strategic decisions such as facilities' capacities and locations whereas the second stage usually refers to operational or tactical decisions such as production schedule, inventory control, transportation planning.

The conventional formulation of a two-stage stochastic programming model is presented as follows:

$$\min_{x} \quad \left\{ c^T x + \sum_{s\in S} p_s Q(x,s) \mid Ax = b, \quad x \geq 0 \right\} \tag{1}$$

$$\text{where:} \quad Q(x,s) = \min_{y_s} \{ q_s y_s \mid W_s y = h_s - T_s x \quad , y_s \geq 0 \}$$

In this model, the vector x is the first-stage decision variables and y_s is the second-stage decision variables for each scenario s. The objective of this problem is to minimize the cost of the first stage $(c^T x)$ and the expected cost over all scenarios of the second stage $\sum_{s\in S} p_s Q(x,s)$. Each scenario s is associated with a probability of occurrence p_s and the sum of the probabilities for all scenarios equal to 1. T_s is the matrix related to the first stage decision variables x. W_s is a recourse matrix, q_s and h_s are two vectors associated with scenario s in the second stage model.

3.3. Mathematical formulation of the problem

In this section, we formulate a two-stage stochastic programming model that considers the selection of suppliers with a mid/long-term contract, inventory management and transportation planning in a biomass supply chain. The first-stage decision variables include only binary variables which determine the choice of suppliers with long/mid-term contracts. The second-stage variables include the amount of biomass to be purchased from suppliers, the amount of biomass feedstock to be transported to biorefinery, the amount of biomass to be used to transform into bioethanol, the volume of unmet bioethanol, the inventory of biomass and bioethanol in each period (Figure 1).

We consider a planning horizon of T periods, each period corresponds to a month. Let I be a set of biomass suppliers, P be a set of pretreatment facilities, B be a set of biorefineries, S be a set of stochastic scenarios and A be the set of arcs in the biomass supply chain A. An arc $(i, j) \in A$ stands for a path with biomass flow from node i to node j in the supply chain network considered. A node may be a supplier, a pretreatment facility or a biorefinery. Let $pre(n)$, $suc(n)$ denote, respectively, the set of predecessor and the set of successors of a given node n. For example, $suc(i), i \in I$ indicates a set of pretreatment facilities to which biomass feedstock is delivered from farm i; $pre(p), p \in P$ indicates a set of farms that supply biomass feedstock to pretreatment facility p; $suc(p), p \in P$ indicates a set of biorefineries to which biomass is delivered from pretreatment facility p.

The first-stage and second stage model are presented as follows:

$$\textbf{(SP):} \quad \min_{\varphi} \left\{ mcost^T \varphi + \sum_{s\in S} p_s Q(\varphi, s) \mid \varphi \in \{0,1\} \right\} \tag{2}$$

$$\text{where:} \quad Q(\varphi, s) = \min_{y_s} \left\{ f(\varphi, s, y_s) \mid \text{subject to } (5)-(20) \right\} \tag{3}$$

where $\varphi = \{\varphi(i), i \in I\}$ are the first stage decision variables related to mid/longterm contracts between the biorefineries and suppliers, $\varphi(i) = 1$ if a long term contract is established with supplier i, 0 otherwise. The vector $mcost = \{mcost(i), i \in I\}$ includes fixed cost $mcost(i)$ for establishing

Table 1: List of parameters and variables

Set	
T	Set of periods t
I	Set of suppliers i
P	Set of preprocessing facilities p
B	Set of biorefineries b
M	Set of markets m
A	Set of all arcs in the biomass supply chain $A = (I \times P) \cap (P \times B) \cap (M \times B)$
S	Set of scenarios s
Deterministic parameters	
	Local/Regional suppliers
$mcost(i)$	Fixed cost for establishing a contract with supplier i
$costA^I(i)$	Unit purchase price of biomass from supplier i (\$/ton)
$q_{min}(i)$	Minimum supply quantity of biomass (tons) under contract with supplier i during planning horizon T
	International/national market
$avail^M(m,t)$	Amount of biomass (tons) available at market m in period t
$costA^M(m)$	Unit purchase price of biomass from market m (\$/ton)
	Preprocessing site
$\gamma_p^{max}(p)$	Maximum amount of biomass (tons) that could be preprocessed at preprocessing facility p
$\gamma_p^{min}(p)$	Minimum amount of biomass (tons) that could be preprocessed at preprocessing facility p
$w^{max}(p)$	Maximum biomass storage capacity (tons) at preprocessing facility p in each period
$w^{min}(p)$	Minimum biomass storage capacity (tons) at preprocessing facility p in each period
$w^{init}(p)$	Initial stock of biomass (tons) at preprocessing facility p
$hcostP(p)$	Unit holding cost at preprocessing facility p (\$/ton)
$tcostP(p)$	Unit biomass preprocessing cost of preprocessing facility p (\$/ton)
$\lambda^P(p)$	Ratio of biomass storage loss at preprocessing facility p in each period
$\eta^P(p)$	Preprocessing rate at preprocessing facility p in each period
	Biorefinery site
$\gamma_b^{max}(b)$	Maximum amount of biomass (tons) that could be processed by biorefinery b in each period
$\gamma_b^{min}(b)$	Minimum amount of biomass (tons) that could be processed by biorefinery b in each period
$V_{av}^{init}(b)$	Initial stock of biomass (tons) of biorefinery b
$V_{av}^{max}(b)$	Maximum biomass storage capacity (tons) of biorefinery b in each period
$V_{av}^{min}(b)$	Minimum biomass storage capacity(tons)of biorefinery b in each period
$V_{ap}^{init}(b)$	Initial stock of biofuel (litres) of biorefinery b
$V_{ap}^{max}(b)$	Maximum biofuel storage capacity(litres) of biorefinery b in each period
$V_{ap}^{min}(b)$	Minimum biofuel storage capacity (litres) of biorefinery b in each period
$pcost(b)$	Unit shortage cost for unmet biofuel demand at biorefinery b (\$/tons)
$tcostB(b)$	Unit biofuel production cost at biorefinery b (\$/tons)
$hcostB_{av}(b)$	Unit inventory holding cost of biomass at biorefinery site b in each period (\$ / ton)
$hcost_{ap}^B(b)$	Unit inventory holding cost of biofuel at biorefinery b in each period (\$/ ton)
$\eta^B(b)$	Biofuel yield rate of biomass for biorefinery b (litre/ton)
$\lambda^B(b)$	Ratio of biomass storage loss at biorefinery b in each period
$demand(b,t)$	Demand of biofuel(litre) at biorefinery b in period t
	Transportation:
$Vcost(i,j)$	Unit transportation cost from node i to node j (\$ /tons)
Stochastic parameters	
$avail^I(i,t,s)$	Amount of biomass (tons) available at supplier i in period t under scenario s
Decision variables	*Continuous variables*
$x(i,j,t,s)$	A flow of biomass on arc $(i,j) \in A$ in period t under scenario s, originally from node i to node j
$w(p,t,s)$	Biomass inventory level (tons) at preprocessing facility p in period t under scenario s
$V_{av}(b,t,s)$	Inventory level of biomass at biorefinery b in period t under scenario s
$V_{ap}(b,t,s)$	Inventory level of biofuel at biorefinery b in period t under scenario s
$q^I(i,t,s)$	Total amount of biomass transported from supplier i in period t under scenario s
$q^P(p,t,s)$	Total amount of biomass transported to preprocessing facility p in period t under scenario s
$z(b,t,s)$	Amount of biomass required (tons) to produce biofuel at refinery b in period t under scenario s
$short(b,t,s)$	Shortage of biofuel (litres) at biorefinery b in period t under scenario s
	Binary variables
$\varphi(i)$	Equal to 1 if a mid/long-term contract between the biorefineries and supplier i is established, 0 otherwise
	Dual variables
$\pi_{(.)}$	Dual variables associated with the constraints (5)-(20) for problem **(P2)**. For readability, each dual variable is written in bracket at the right hand of the corresponding constraint in model **(P2)**

a contract with supplier i, which is independent of scenario s. The function $Q(\varphi, s)$ refers to the cost function of the second stage problem for each scenario s, in which the first stage decisions, φ are already made.

For convenience, all notations used in the model formulation are summarized in Table 1. All continuous decision variables are non-negative. The constraints (5 - 20) will be presented in the next part of this section. The different components of the second stage total cost $f(\varphi, s, y_s)$ under scenario s are given below; where y_s contains all second-stage decision variables, $y_s = (x, w, V_{av}, V_{ap}, q^P, z, short)$.

$$\begin{aligned} f(\varphi,s,y_s) = &\sum_{t\in T}\sum_{(i,j)\in A} x(i,j,t,s) * Vcost(i,j) \\ &+ \sum_{t\in T}\sum_{p\in P} \eta^P * q^P(p,t,s) * tcostP(p) \\ &+ \sum_{t\in T}\sum_{b\in B} short(b,t,s) * pcost(b) \\ &+ \sum_{t\in T}\sum_{b\in B} \eta^B * z(b,t,s) * tcostB(B) \\ &+ \sum_{t\in T}\sum_{p\in P} hcost^P(p) w(p,t,s) \\ &+ \sum_{t\in T}\sum_{b\in B} hcost^B_{av}(b) V_{av}(b,t,s) \\ &+ \sum_{t\in T}\sum_{b\in B} hcost^B_{ap}(b) V_{ap}(b,t,s) \\ &+ \sum_{t\in T}\sum_{i\in I} q^I(i,t,s) * costA^I(i) \\ &+ \sum_{t\in T}\sum_{m\in M}\sum_{b\in B} x(m,b,t,s) * costA^M(m) \end{aligned} \tag{4}$$

where $x(i,j,t,s)$ is the flow of biomass on arc $(i,j) \in A$ in period t under scenario s. Such an arc (i,j) could be arc (i,p) or (p,b) that refers to an arc from supplier $i \in I$ to preprocessing facility $p \in P$ or from preprocessing facility $p \in P$ to biorefinery b. $w(p,t,s)$ is the inventory level at preprocessing facility p in period t under scenario s.

The variables $V_{av}(b,t,s)$, $V_{ap}(b,t,s)$ are the inventory level of biomass and the inventory level of biofuel at biorefinery b in period t under scenario s, respectively. $z(b,t,s)$, $short(b,t,s)$ are the amount of biomass used to produce biofuel and the shortage of biofuel at biorefinery b in period t under scenario s, respectively. $q^I(i,t,s)$ is the total amount of biomass transported from supplier i to the pretreatment facilities in period t under scenario s. $q^P(p,t,s)$ is the total amount of biomass transported to pretreatment facility p in period t under scenario s. Note that the decision variables $q^I(i,t,s)$ and $q^P(p,t,s)$ are redundant but they are introduced for the readability of the model. $costA^I(i)$, $costA^M(m)$ are unit purchase price of biomass from supplier i and from market m, respectively; $tcostP(p)$ is unit biomass preprocessing cost at pretreatment facility p; $hcostB_{av}(b)$, $hcost^B_{ap}(b)$ are unit inventory holding costs of biomass and biofuel at biorefinery site b in each period, respectively; $\eta^B(b)$ is the biofuel yield rate and $\lambda^B(b)$ is the ratio of biomass storage loss at biorefinery b in each period.

The objective function is to minimize the total cost of the supply chain which includes the first stage costs and the expected second-stage costs. The first stage costs comprise the total costs for establishing contracts with local/regional suppliers. Equation (4) gives the second stage costs $Q(\varphi, s)$ for each scenario s, which comprise nine costs: transportation costs (1st term), pretreatment costs (2nd term), shortage costs (3rd term), biofuel conversion costs (4th term), holding costs (5th-7th terms), and purchase costs (9th-10th terms).

Constraints on suppliers

$$q^I(i,t,s) = \sum_{p \in suc(i)} x(i,p,t,s) \qquad \forall t \in T, i \in I, s \in S \qquad [\pi^1_{i,t,s}] \qquad (5)$$

$$q^I(i,t,s) \leq avail^I(i,t,s).\varphi(i) \qquad \forall t \in T, i \in I, s \in S \qquad [\pi^2_{i,t,s}] \qquad (6)$$

$$\sum_{t \in T} q^I(i,t,s) \geq q_{min}(i).\varphi(i) \qquad \forall i \in I, s \in S \qquad [\pi^3_{i,s}] \qquad (7)$$

Equation (5) defines the total amount of feedstock transported from supplier i to the pretreatment facilities in each period under scenario s.

Inequality (6) states that there is no feedstock supply from supplier i unless there is a mid/long-term contract between the supplier and biorefinery b. This inequality also ensures that the total purchase amount does not exceed the available amount of the supplier in each period under scenario s.

Inequality (7) specifies the minimum supply quantity under a contract between a supplier i and the biorefinery b during planning horizon T.

Constraints on market

$$\sum_{b \in B} x(m,b,t,s) \leq avail^M(m,t) \qquad \forall t \in T, m \in M \qquad [\pi^4_{m,t,s}] \qquad (8)$$

Inequality (8) ensures that the purchase quantity of biomass is also limited by the market's supply availability in each scenario.

Constraints on preprocessing sites

$$q^P(p,t,s) = \sum_{i \in pre(p)} x(i,p,t,s) \qquad \forall t \in T, p \in P, s \in S \qquad [\pi^5_{p,t,s}] \qquad (9)$$

$$w(p,t,s) = (1-\lambda^P(p))w(p,t-1,s) + \eta^P q^P(p,t,s) - \sum_{b \in suc(p)} x(p,b,t,s) \qquad \forall t \geq 2, p \in P, s \in S \qquad [\pi^6_{p,t,s}] \qquad (10)$$

$$w(p,1,s) = (1-\lambda^P(p))w^{init}(p) + \eta^P q^P(p,1,s) - \sum_{b \in suc(p)} x(p,b,1,s) \qquad \forall p \in P, s \in S \qquad [\pi^7_{p,s}] \qquad (11)$$

$$\gamma^{min}_p(p) \leq q^P(p,t,s) \leq \gamma^{max}_p(p) \qquad \forall t \in T, p \in P, s \in S \qquad [\pi^{8a/b}_{p,t,s}] \qquad (12)$$

$$w^{min}(p) \leq w(p,t,s) \leq w^{max}(p) \qquad \forall t \in T, p \in P, s \in S \qquad [\pi^{9a/b}_{p,t,s}] \qquad (13)$$

Equation (9) defines the total amount of biomass transported from suppliers to pretreatment facility p in each period under scenario s. Equation (10) and (11) ensure the inventory balance for each stock at preprocessing facility under scenario s. These constraints ensure that no more biomass is delivered from or processed at a location than its available amount in stock. The capacity of each preprocessing facility and its inventory capacity are given by constraints (12) and (13), respectively.

Constrains on refinery sites

$$V_{av}(b,t,s) = (1-\lambda^B(b))V_{av}(b,t-1,s) - z(b,t,s) + \sum_{p \in pre(b)} x(p,b,t,s) + \sum_{m \in M} x(m,b,t,s) \qquad \forall t \geq 2, b \in B, s \in S \qquad [\pi^{10}_{b,t,s}] \tag{14}$$

$$V_{av}(b,1,s) = (1-\lambda^B(b))V^{init}_{av}(b) - z(b,1,s) + \sum_{p \in pre(b)} x(p,b,1,s) + \sum_{m \in M} x(m,b,1,s) \qquad \forall b \in B, s \in S \qquad [\pi^{11}_{b,s}] \tag{15}$$

$$V_{ap}(b,t,s) = V_{ap}(b,t-1,s) + \eta^B z(b,t,s) - demand(b,t) + short(b,t,s) \qquad \forall t \geq 2, b \in B, s \in S \qquad [\pi^{12}_{b,t,s}] \tag{16}$$

$$V_{ap}(b,1,s) = V^{init}_{ap}(b) + \eta^B z(b,1,s) - demand(b,1) + short(b,1,s) \qquad \forall b \in B, s \in S \qquad [\pi^{13}_{b,s}] \tag{17}$$

$$\gamma^{min}_b(b) \leq z(b,t,s) \leq \gamma^{max}_b(b) \qquad \forall t \in T, b \in B, s \in S \qquad [\pi^{14a/b}_{b,t,s}] \tag{18}$$

$$V^{min}_{av}(b) \leq V_{av}(b,t,s) \leq V^{max}_{av}(b) \qquad \forall t \in T, b \in B, s \in S \qquad [\pi^{15a/b}_{b,t,s}] \tag{19}$$

$$V^{min}_{ap}(b) \leq V_{ap}(b,t,s) \leq V^{max}_{ap}(b) \qquad \forall t \in T, b \in B, s \in S \qquad [\pi^{16a/b}_{b,t,s}] \tag{20}$$

Equations (14 -17) are the inventory balance constraints for each biomass or biofuel stock at biorefinery b, respectively. These constraints ensure that no more biomass or biofuel is delivered from or processed at a location than its available amount in stock. They also ensure that biofuel demand is satisfied as much as possible, and any unsatisfied demand is lost. The shortage cost is charged for any unit lost.

Equation (19) and (20) are inventory capacity constraints for biomass and biofuel stock at biorefinery b. The conversion capacity of each refinery facility is given by constraints (18).

4. Solution approach

In this section, the stochastic programming model (SP) is transformed into a deterministic equivalent model (DEP) by applying the scenario approach. The second model can be solved by a commercial solver like CPLEX ou GUROBI. However, if the number of scenarios is large, the resolution of the model requires an enormous memory and computational effort. Therefore, we propose in this section an enhanced and regularized decomposition (ERD) method for the model to overcome the computational difficulties.

4.1. Deterministic equivalent programming model

We reformulate the model **(SP)** as a deterministic mixed integer linear programming (MILP) model as follows:

$$\textbf{(DEP)} \quad \min_{x,y_s} \quad \left(mcost^T \varphi + \sum_{s \in S} p_s f(\varphi, s, y_s) \right)$$
$$s.t. \quad \varphi(i) \in \{0,1\} \quad \forall i \in I$$
$$y_s \quad \text{subject to constraints (5)-(20)}$$

This (DEP) model could be solved by a commercial solver like CPLEX or GUROBI. However, for an instance considering a large number of scenarios, the size of the model is too large to be solved by such a solver due to the memory requirement of a computer. For example, in case of an instance with $I = 60$, $P = 6$, $T = 12$ and $S =$1000 scenarios,, the coefficient matrix of the DEP model has 1.680.000 rows and 1.716.060 columns. A computer uses 12 bytes to store each non-zero element of this matrix. Thus, it requires at least 1.716.060x1.680.000 $\times 12/10^9 \approx 240$ GB to store the matrix. So, to solve this instance, the PC should have at least 240GB of RAM. Such high memory requirement could not be met by most PCs. For this reason, we develop a stochastic decomposition method to solve the model in a reasonable time with much less memory requirement.

4.2. Multi-cut L-shaped algorithm

The scenario-based stochastic programming approach allows to capture the uncertainty in an approximate way with the precision depending on the number of scenarios considered. The size of the model can grow dramatically due to the consideration of a large number of scenarios. By extended a Bender decomposition in stochastic programming, Van Slyke and Wets (1969) introduced an iterative method, namely the L-shaped algorithm. To speed up this algorithm, Birge and Louveaux (1988) showed that the problem of calculating the expected value in the second stage problem could be decomposed by scenarios s and multiple cuts as many as the number of scenarios can be generated. These cuts allow to reduce the solution space of the first stage variables in each iteration. The multi-cut L-shaped algorithm takes a fewer number of iterations to reach an optimal solution in comparison with the corresponding L-shaped algorithm, but each iteration may take a longer computational time in solving a large number of scenario subproblems.

We now introduce a new variable vector $\theta = (\theta_s, s \in S)$ that provides an link between the first-stage problem **(P1)** and the scenario subproblem **(P2)**. θ_s can be interpreted as an approximation of $Q(\varphi, s)$. The first-stage problem **(P1)** is thus given by:

$$\textbf{(P1)} \quad \min_{\varphi,\theta} \quad mcost^T \varphi + \sum_{s \in S} p_s \theta_s$$
$$s.t. \quad \theta_s \geq d_s^k \varphi + e_s^k \quad \forall s \in S, \ k = 1,2...K \tag{21}$$
$$\varphi(i) \in \{0,1\} \quad \forall i \in I$$

where k denotes the k^{th} iteration, K is the number of iterations so far. In the first iteration, the first-stage problem is solved without considering the second stage variables and their constraints.

Then, the subproblem for each scenario $s \in S$ is formulated as follows and is solved with the values of the first-stage variables, φ given by the solution of the first-stage problem **(P1)**.

$$\textbf{(P2)} \quad \min_{y_s} \quad f(\varphi, s, y_s)$$
$$s.t. \quad \text{constraints} \quad (5)-(20)$$

In each iteration, the first-stage problem and the subproblem for each scenario s are linked through optimality cuts. These cuts (21) are generated from the dual of the subproblem **(P2)**. The coefficients d_s^k and e_s^k of the optimality cut are given as follows:

$$d_s^k = \pi_{k,s}^T T_s \quad \forall s \in S, \;\; k = 1,2...K$$
$$e_s^k = \pi_{k,s}^T h_s \quad \forall s \in S, \;\; k = 1,2...K \tag{22}$$

where $\pi_s = \{\pi_s^i\} \quad (\forall i = 1...16)$ is the optimal dual vector of the subproblem **(P2)** for each scenario $s \in S$. h_s is the vector associated with scenario s and T_s is the matrix related to the first stage decision variables φ in the corresponding model (1).

By solving the dual of the subproblem (P2), we obtain the optimal dual variable π_s for each scenario and then compute the coefficient d_s^k and e_s^k as follows:

$$\begin{aligned}
d_s^k &= \sum_{t\in T}\sum_{i\in I} avail^I(i,t,s)\pi_{i,t,s}^2 + \sum_{i\in I} q_{min}(i)\pi_{i,s}^3 \\
e_s^k &= \sum_{t\in T}\sum_{m\in M} avail^M(m,t,s)\pi_{m,t,s}^4 + \sum_{p\in P}(1-\lambda^P(p))w^{init}(p)\pi_{p,s}^7 \\
&+ \sum_{t\in T}\sum_{p\in P}\gamma_p^{max}(p)\pi_{p,t,s}^{8a} + \sum_{t\in T}\sum_{p\in P}\gamma_p^{min}(p)\pi_{p,t,s}^{8b} \\
&+ \sum_{t\in T}\sum_{p\in P} w^{max}(p)\pi_{p,t,s}^{9a} + \sum_{t\in T}\sum_{p\in P} w^{min}(p)\pi_{p,t,s}^{9b} \\
&+ \sum_{b\in B}(1-\lambda^B(b))V_{av}^{init}(B)\pi_{b,s}^{11} - \sum_{t\geq 2\in T}\sum_{b\in B} demand(b,t)\pi_{b,t,s}^{12} \\
&+ \sum_{b\in B}\left[V_{ap}^{init}(B) - demand(b,1)\right]\pi_{b,s}^{13} + \sum_{t\in T}\sum_{b\in B}\gamma_b^{max}(b)\pi_{b,t,s}^{14a} \\
&+ \sum_{t\in T}\sum_{b\in B}\gamma_b^{min}(b)\pi_{b,t,s}^{14b} + \sum_{t\in T}\sum_{b\in B} V_{av}^{max}(b)\pi_{b,t,s}^{15a} \\
&+ \sum_{t\in T}\sum_{b\in B} V_{av}^{min}(b)\pi_{b,t,s}^{15b} + \sum_{t\in T}\sum_{b\in B} V_{ap}^{max}(b)\pi_{b,t,s}^{16a} \\
&+ \sum_{t\in T}\sum_{b\in B} V_{ap}^{min}(b)\pi_{b,t,s}^{16b}
\end{aligned}$$

This multi-cut L-shaped algorithm is presented in pseudo-code in Algorithm 1. It is outlined as follows:

i) In each iteration, the first-stage problem is solved and the dual of subproblem (**P2**) is solved as many times as the number of scenarios $s \in S$

Algorithm 1 : Multi-cut L-shaped algorithm

Initialization: Set $k = 1$, $LB = -\infty$, $UB = \infty$.

Step 1: In each iteration k, solve the first stage problem **(P1)** with all the optimality cuts generated at all previous iterations. Denote the optimal objective value of the problem as α^k and the optimal solution of the first stage decision variables φ as φ^k.

If $\alpha^k \geq LB$, set $LB = \alpha^k$

Step 2: Solve all the dual of subproblems of **(P2)** with the values of first stage decision variables fixed as $\varphi = \varphi^k$ and obtain the optimal dual vector for each scenario $s \in S$ and .

Set $\beta^k = \sum_{i\in I} \varphi^k(i) * mcost(i) + \sum_{s\in S} p_s\theta_s$.

If $\beta^k \leq UB$ then update $UB = \beta^k$

Step 3: If $(UB - LB)/UB \leq \varepsilon$, stop and get an optimal solution. Otherwise, go to Step 4.

Step 4: Add the optimality cuts to the first-stage problem. Set $k = k+1$, and go to Step 1.

ii) By solving the first-stage problem, the values of the first stage decision variables φ are determined and the value of function $Q(\varphi, s)$ at each scenario s in the second stage is approximated by θ_s.

iii) By solving the dual of subproblem **(P2)**, the expected value of the second stage problem for all scenarios can be found for the given value of the first stage variables found in the previous step.

iv) In each iteration, after solving the dual of the scenario subproblem, multiple cuts of type (21) with the number equal to that of scenarios are added to the first-stage problem.

4.3. Enhanced and regularized decomposition approach

4.3.1. Regularized decomposition

In general, the classical L-shaped method has two drawbacks: (1) At initial iterations; cuts are often inefficient; (2) At final iterations, cuts become degenerately. To avoid these drawbacks, Ruszczyński and Świtanowski (1997) introduced a regularized decomposition method that combines a multi-cut approach with adding a quadratic regularized term in the objective function. We have adapted this approach to our problem. The regularized decomposition approach is given in Algorithm 2.

In addition, we have also applied several acceleration techniques to improve the convergence of the regularized decomposition algorithm. The valid inequalities presented in Section 4.3.2-4.3.4 are added to the regularized first-stage problem **(RD)** involved. These valid inequalities are described in the next section.

4.3.2. Valid inequalities

The first-stage problem only includes the binary constraints of φ and a few cuts. In some initial iteration, that may lead to the fact that very few suppliers are selected and that correspond to a small objective function value (lower bound). In addition, for such small number of suppliers selected, there is a little demand can be met in the scenario sub-problems. Consequently, it causes high shortage costs and leads to a high value of upper bound.

Algorithm 2 : Regularized decomposition algorithm

Initialization: Set $k = 0$, $LB = -\infty$ and $UB = \infty$. Select a^1 as a feasible solution.

Step 1: Set $k = k+1$. Solve the regularized first-stage problem

$$\begin{aligned} \textbf{(RD)} \quad & \min_{\varphi,\theta} \quad mcost^T \varphi + \sum_{s\in S} p_s\theta_s + \frac{1}{2}\|\varphi - a^k\|^2 \\ s.t. \quad & \theta_s \geq e_s^k \varphi + d_s^k \quad \forall s \in S, \ k = 1, 2...K \\ & \varphi(i) \in \{0,1\} \quad \forall i \in I \end{aligned}$$

Let (φ^k, θ^k) be an optimal solution of the (RD) problem at iteration k, where $\varphi^k = (\varphi^k(i), i \in T)$ is a vector of φ.

If $(UB - LB)/UB \leq \varepsilon$, stop and get an optimal solution. Otherwise, go to Step 2

Step 2: Solve the dual of subproblem (P2) for all $s \in S$. Compute an optimality cut (21) as before.

Step 3: If $mcost^T \varphi^k + \sum_{s\in S} p_s Q(\varphi, s) \leq mcost^T a^k + \sum_{s\in S} p_s Q(a^k, s)$

then update $UB = mcost^T \varphi^k + \sum_{s\in S} p_s Q(\varphi, s)$ and set $a^{k+1} = \varphi^k$, go to Step 1.

Else, $a^{k+1} = a^k$, go to Step 1.

To avoid these inefficiencies in the initial iterations, it is necessary to integrate more information obtained from subproblems into the first stage problem to improve more significantly the quality of lower and upper bound in each iteration. From the nature of the biomass supply chain problem, valid inequalities may be added to the first-stage problem to improve the convergence rate and produce high-quality solutions.

From inequalities (5), (7) and (12), we obtain the following valid inequalities:

$$\begin{aligned} & \sum_{i\in I} q_{min}(i)\varphi(i) \leq \sum_{t\in T}\sum_{i\in I} q^I(i,t,s) = \sum_{t\in T}\sum_{p\in P} q^P(p,t,s) \leq \sum_{t\in T}\sum_{p\in P} \gamma_p^{max}(p) \\ & \Longrightarrow \sum_{i\in I} q_{min}(i)\varphi(i) \leq T \sum_{p\in P} \gamma_p^{max}(p) \quad \forall s \in S \end{aligned} \tag{23}$$

These valid inequalities (23) ensure the total minimum supply quantity from suppliers selected that are not surpassed a maximal amount biomass processed at preprocessing facility.

We also have:

$$\begin{aligned} & \sum_{t\in T}\sum_{i\in I} avail^I(i,t,s)\varphi(i) \geq \sum_{t\in T}\sum_{i\in I} q^I(i,t,s) \geq \sum_{t\in T}\sum_{p\in P} \gamma_p^{max}(p) \geq \sum_{t\in T}\sum_{p\in P} \gamma_p^{min}(p) \\ & \Longrightarrow \sum_{t\in T}\sum_{i\in I} avail^I(i,t,s)\varphi(i) \geq T \sum_{p\in P} \gamma_p^{min}(p) \quad \forall s \in S \end{aligned} \tag{24}$$

These valid inequalities (24) ensure that the suppliers selected are capable of satisfying at least a minimum amount biomass processed at preprocessing facility.

Besides, adding these valid inequalities (23) (24) not only reduce the feasible region of the first-stage problem but also avoids the infeasibility of the scenario subproblems in each iteration.

4.3.3. Knapsack inequalities

Let UB be the current upper bound in the iteration considered. By combining the optimal cut: $\theta_s \geq e_s^k \varphi + d_s^k$ and $UB \geq mcost^T \varphi + \sum_{p \in P} p_s \theta_s$, we have:

$$UB \geq mcost^T \varphi + \sum_{p \in P} p_s \theta_s \geq mcost^T \varphi + \sum_{s \in S} p_s (e_s^k \varphi + d_s^k)$$
$$\Longrightarrow \big(mcost + \sum_{s \in S} p_s e_s^k\big)^T \varphi \leq UB - \sum_{s \in S} p_s d_s^k$$

As φ are binary variables, we can rounded down two sides of the above inequality and obtain a following valid inequality follows:

$$\lfloor \big(mcost + \sum_{s \in S} p_s e_s^k\big)^T \rfloor \varphi \leq \lfloor UB - \sum_{s \in S} p_s d_s^k \rfloor \tag{25}$$

Where $\lfloor a \rfloor$ is a number rounded down from a given number a. When an upper bound ub_s is available, adding a knapsack inequality (25) along with the optimality cut (21) could tighten the first-stage problem and lead to a good quality solution in each iteration.

4.3.4. Specific optimality cut

Laporte and Louveaux (1993) defined an optimality cut for the general L-shaped algorithm with the binary first-stage decision variable. This specific optimality cut can be added to accelerate the L-shaped algorithm when the first-stage variables are binary variables. We reformulated the specific optimality cut for the multi-cut L-shaped algorithm as follows:

$$\sum_{s \in S} p_s \theta_s \geq (q_s - L) \left(\sum_{i \in \tau} \varphi(i) - \sum_{i \notin \tau} \varphi(i) - \mid \tau \mid +1 \right) + L \tag{26}$$

where: $q_s = \sum_{s \in S} p_s Q(\varphi, s)$ is the corresponding recourse function value in each iteration. L is a lower bound satisfying: $L \leq min_{\varphi} \{\sum_{s \in S} p_s Q(\varphi, s)\} \; \forall \varphi$. The set τ is defined as the set of all selected suppliers i: $\tau = \{i \mid \varphi(i) = 1\}$ and $\mid \tau \mid$ denote the cardinality of set τ.

By adding the valid inequalities presented in Section 4.3.2- 4.3.4 into the regularized first-stage problem **(RD)**, we could enhance the performance of the solution method, called (ERD) method.

4.4. Determination of the number of scenarios by Monte Carlo Sampling

To approximate the distribution of a random event, we use a Monte Carlo sampling approach to generate the scenarios which allow to reduce the model size according to Linderoth et al. (2006); Shapiro (2000). Each scenario is assigned with the same probability, and the sum of the probabilities of all scenarios is equal to 1. Using a statistical method in Shapiro and Homem-de Mello (1998), we are able to determine the minimum number of scenarios required to obtain a solution within a confidence interval for a given level of confidence α. The idea of this approach is based on the theory of probability to find a relationship between reducing the number of scenarios and the reliability of the obtained solution. This method is based on the measure of the confidence interval of the expected total cost. The Monte Carlo sampling variance estimator $\sigma(n)$ is given as follows:

$$\sigma(n) = \sqrt{\frac{\sum_{s=1}^{n}(\mathbb{E}[totalcost] - totalcost_s)^2}{n-1}} \tag{27}$$

where n is the number of scenarios used to estimate the Monte Carlo sampling variance, and $totalcost_s$ is the total system cost under scenario s. The minimum number of scenarios N required can be calculated by:

$$N = \left(\frac{z_{\alpha/2}\sigma(n)}{H}\right)^2 \tag{28}$$

where H is a given confidence interval and α is a given level of confidence. The value $z_{\alpha/2}$ is determined by $Pr(z \leq z_{\alpha/2}) = 1 - \alpha/2$, where $z \sim N(0,1)$. In summary, the procedure for determining the minimum number of scenarios is given as follows:

i) Solve the stochastic programming model with a small number of scenarios n (E.g n= 10-100).
ii) Estimate the value of sampling estimator $\sigma(n)$.
iii) Determine the number of scenarios N required for a given confidence interval H and level of confidence α.

5. Numerical study

In this section, numerical studies are conducted to evaluate the performance of the proposed algorithm and demonstrate the effectiveness of our approach. The algorithm is implemented in Python 3.5 on an HP computer with Intel Core i5-4210M CPU 2.66 GHz and 8.0 GB RAM. All linear programming (LP) problems are solved by using commercial solver GUROBI 6.5. According to the Benchmarks of Optimization Software (Mittelmann, 2018), GUROBI performs better than CPLEX on the MIPLIB2010 benchmark set. That is why we chose GUROBI as MILP solver in our numerical study.

5.1. Data generation

In the numerical study, we use data from Osmani and Zhang (2013); Zhang et al. (2012); Osmani and Zhang (2014) to make our tested instances more realistic. The dataset is summarized in Appendix A.6. We consider a biomass supply chain over one year-horizon divided into 12 periods. Miscanthus is used as biomass raw material, and it is harvested from March to April. In our study, the first period, $t = 1$ corresponds to March. The model data assumptions are presented as follows:

a. One period in this study corresponds to one month. The biorefinery seeks to sign a contract with several local feedstock suppliers. The initial inventory quantity required to commence operations is 25 000 tons of biomass for a biofuel facility.

Table 2: Problem size of the deterministic equivalent model

I	P	B	T	No. Of scenarios	No. Of binary variables	No. Of continuous variables
40	4	1	12	50	50	58 200
40	4	1	12	100	100	116 400
40	4	1	12	1000	100	1 164 000
50	5	1	12	50	50	72 000
50	5	1	12	100	100	144 000
50	5	1	12	1000	100	1 440 000
60	6	1	12	50	50	85 800
60	6	1	12	100	100	171 600
60	6	1	12	1000	100	1 716 000

b. The installed biorefinery has a production capacity of 190 MLPY with a unit production cost of 0.2$/ liter. The bioethanol yield rate of lignocellulosic biomass is about 313 liters/ton. The unit biofuel storage cost and unit biomass storage cost at biorefinery are 0.227 $/liter and 0.9 $/ton , respectively. The unit shortage cost is 1.06 $/liter.
c. All 50 potential suppliers are located within 100km radius of the biorefinery.
d. The amount of biomass available in each supplier is estimated from land availability Osmani and Zhang (2013, 2014) and miscanthus yield (from 10 to 14 ton/hecta according to Clifton-Brown et al. (2007)). For each supplier, its miscanthus yield is generated from uniform distribution $U[10,14]$.
e. The pretreatment facility has a production capacity of 2 Mton/year with unit production cost of 13.94 $/ton and its transformation yield rate is 0.75. Unit biomass storage cost at pretreatment facility is 1.125 $/ton.
f. The ratio of biomass storage loss at pretreatment and at biorefinery are 3% and 1% respectively.
g. Biofuel demand is generated from Normal probability distribution $N(2133,4138)$ and truncated in the interval $[2006,2280]$ based on historical data (1981-2005) in Osmani and Zhang (2013). The monthly biofuel demand is also adjusted to the production capacity of the biorefinery.
h. Contracts with suppliers are characterized by fixed cost, purchase price and minimum quantity to supply. Purchase price varies from 42 to 45 $/ton. The minimum supply quantity varies from 20 000 to 30 000 tons. Fixed cost for establishing a contract varies from $20 000 to $30 000.
i. Biomass in market has a purchase price 55 $/ton. However, the distance from a market to the biorefinery varies from 150 to 200km.

The main focus of our work is to study the impact of uncertain supply on the strategic decision made. This uncertainty is incorporated into the model by considering a set of possible scenarios. The number of scenarios should be large enough to capture the entire range of the probability distribution of feedstock supply.

5.2. Performance evaluation of the solution algorithm

In this section, we describe our computational experiments on the algorithm proposed in the previous section when solving instances of different sizes. The computation time for solving the model depends on the number of periods, suppliers, preprocessing facilities, biorefineries, and scenarios. We tested this model for instances with I = 40, 50, 60; P = 4, 5, 6; T = 12 and number of scenarios S = 50, 100, 1000. The dimension of the deterministic equivalent problem for these instances is presented in Table 2. It shows that the size of the stochastic programming model increases rapidly as a function of the number of scenarios, suppliers and pretreatment sites.

Table 3 summarizes the size, the number of scenarios and the execution time of the solver GUROBI, the multi-cut L-Shaped method (MLS), the regularized decomposition method (RD) and the enhanced & regularized decomposition L-shaped method (ERD) for these instances.

We fixed the stopping condition of the algorithm as follows: the optimality gap $\varepsilon \leq 0.01(\%)$ or CPU time $\geq$ $time_{max}$=1800s or the number of iterations $\geq iteration_{max} = 1000$, where the optimality gap is calculated by: $\varepsilon(\%) = 100 * (UB - LB/UB)$. We also generated 10 samples of biofuel demand for each instance, denoted by d_i ($\forall i = 1...10$). As the experimental results indicate, the (ERD) method and GUROBI constantly outperforms by a factor of 2 the standard multi-cut L-Shaped method (MLS). On average, the (ERD) method is 1.01, 1.21, 1.16, 1.20, 1.20 and 1.32 times faster than GUROBI for all instances studied, respectively (6 sub-tables in the 1st and 2nd lines of Table 3). Therefore, the (ERD) method can be used to solve the considered problem because it can find an optimal solution in a shorter computational time.

For large instances, the size of the model is too large to be solved by solver GUROBI because the computer requires at least 240GB of memory. Three sub-tables in the third line of Table 3 show that the (ERD) algorithm still performed well for large instances because the second stage problem is decomposed into multiple subproblems with reasonable size.

By comparing the last two columns, we observe that the added valid inequalities allow improving the convergence speed of the solution method significantly. In fact, adding the valid inequalities presented in Section 4.3.2- 4.3.4 could reduce 8.8-30.3% of the computation time in comparison with the RD method (the classic regularization approach without adding valid inequalities). This result proves the relevance of adding these inequalities in the solution method.

5.3. Analysis of the solution

This section analyzes our numerical results. Figure 2 shows the cost distribution of the base case for the biomass supply chain over planning horizon T= 12. The optimal value of the total expected cost of the supply chain is \$114 271 335 when the biomass supply chain includes 1 biorefinery, 1 pretreatment site, 50 suppliers over 12 time periods.

The biofuel production cost and feedstock purchase cost are the primary cost drivers that account for nearly 39.5 % and 34.8% of the total system costs respectively. The transportation cost has an essential role in the biomass supply chain with approximatively 14.7% of total system cost. This result shows the importance of transportation operation in biomass supply chains. The holding cost is only 2.9%, so in this case, the storage of a significant amount of biomass at biorefinery seems not critical.

Table 3: Comparison of solution approaches

Problem size						CPU time			
I	**P**	**B**	**T**	**S**	**D**	**Gurobi**	**MLS**	**RD**	**ERD**
40	4	1	12	50	d_1	2.24	4.47	2.46	2.09
40	4	1	12	50	d_2	2.29	5.53	4.38	3.51
40	4	1	12	50	d_3	2.27	4.35	2.56	2.21
40	4	1	12	50	d_4	2.32	4.38	2.49	2.07
40	4	1	12	50	d_5	2.42	6.16	4.38	3.61
40	4	1	12	50	d_6	2.31	4.36	2.26	1.79
40	4	1	12	50	d_7	2.43	4.37	2.27	1.85
40	4	1	12	50	d_8	2.44	4.39	2.43	2.01
40	4	1	12	50	d_9	2.30	6.67	2.34	2.07
40	4	1	12	50	d_{10}	2.31	4.83	2.27	1.88
Average						**2.33**	**4.95**	**2.78**	**2.31**

Problem size						CPU time			
I	**P**	**B**	**T**	**S**	**D**	**Gurobi**	**MLS**	**RD**	**ERD**
50	5	1	12	50	d_1	2.99	7.18	3.52	2.83
50	5	1	12	50	d_2	3.04	7.75	5.08	4.11
50	5	1	12	50	d_3	3.18	7.67	5.66	4.17
50	5	1	12	50	d_4	3.21	5.19	3.24	2.29
50	5	1	12	50	d_5	3.22	5.22	2.71	2.49
50	5	1	12	50	d_6	3.29	7.78	2.68	2.27
50	5	1	12	50	d_7	2.97	5.24	2.77	2.29
50	5	1	12	50	d_8	3.08	5.13	2.74	2.16
50	5	1	12	50	d_9	3.11	5.10	2.63	2.10
50	5	1	12	50	d_{10}	3.15	5.13	2.65	2.11
Average						**3.12**	**6.14**	**3.37**	**2.68**

Problem size						CPU time			
I	**P**	**B**	**T**	**S**	**D**	**Gurobi**	**MLS**	**RD**	**ERD**
60	6	1	12	50	d_1	3.77	9.96	3.94	3.32
60	6	1	12	50	d_2	3.86	6.45	3.68	3.27
60	6	1	12	50	d_3	3.82	9.52	4.22	3.55
60	6	1	12	50	d_4	3.87	6.53	3.46	2.90
60	6	1	12	50	d_5	3.93	6.51	3.91	3.22
60	6	1	12	50	d_6	3.85	6.55	3.83	3.30
60	6	1	12	50	d_7	3.80	6.66	3.57	3.05
60	6	1	12	50	d_8	3.77	6.48	3.90	3.23
60	6	1	12	50	d_9	3.75	6.63	3.60	2.96
60	6	1	12	50	d_{10}	3.74	6.53	3.44	2.90
Average						**3.82**	**7.18**	**3.76**	**3.17**

I	**P**	**B**	**T**	**S**	**D**	**Gurobi**	**MLS**	**RD**	**ERD**
40	4	1	12	100	d_1	4.75	8.82	4.75	4.05
40	4	1	12	100	d_2	4.84	8.84	4.92	4.43
40	4	1	12	100	d_3	4.91	8.86	4.52	3.79
40	4	1	12	100	d_4	5.10	8.76	4.61	4.24
40	4	1	12	100	d_5	4.89	8.82	4.63	3.88
40	4	1	12	100	d_6	4.99	8.85	4.52	3.90
40	4	1	12	100	d_7	4.86	8.77	4.57	4.11
40	4	1	12	100	d_8	4.76	8.93	4.72	3.94
40	4	1	12	100	d_9	4.82	8.81	4.50	3.83
40	4	1	12	100	d_{10}	5.08	9.03	5.02	4.35
Average						**4.90**	**8.85**	**4.68**	**4.05**

I	**P**	**B**	**T**	**S**	**D**	**Gurobi**	**MLS**	**RD**	**ERD**
50	5	1	12	100	d_1	6.56	15.58	9.76	7.56
50	5	1	12	100	d_2	6.57	10.20	5.11	4.13
50	5	1	12	100	d_3	6.88	15.66	6.19	5.03
50	5	1	12	100	d_4	6.64	10.43	5.47	4.50
50	5	1	12	100	d_5	6.42	15.38	9.93	8.17
50	5	1	12	100	d_6	6.58	10.38	5.21	4.27
50	5	1	12	100	d_7	6.53	15.44	9.85	8.06
50	5	1	12	100	d_8	6.70	10.40	5.24	4.39
50	5	1	12	100	d_9	6.66	10.24	5.13	4.32
50	5	1	12	100	d_{10}	6.27	10.34	5.35	4.44
Average						**6.58**	**12.41**	**6.72**	**5.49**

I	**P**	**B**	**T**	**S**	**D**	**Gurobi**	**MLS**	**RD**	**ERD**
60	6	1	12	100	d_1	7.86	13.04	7.16	6.09
60	6	1	12	100	d_2	7.97	13.05	7.87	6.36
60	6	1	12	100	d_3	8.20	12.88	7.13	6.03
60	6	1	12	100	d_4	8.12	12.83	6.75	5.44
60	6	1	12	100	d_5	8.09	13.06	6.83	5.97
60	6	1	12	100	d_6	8.17	13.04	6.52	5.49
60	6	1	12	100	d_7	7.98	13.07	6.70	5.60
60	6	1	12	100	d_8	8.17	13.05	10.48	8.47
60	6	1	12	100	d_9	7.97	20.10	6.63	5.91
60	6	1	12	100	d_{10}	8.14	12.94	6.77	5.74
Average						**8.07**	**13.71**	**7.28**	**6.11**

I	**P**	**B**	**T**	**S**	**D**	**Gurobi**	**MLS**	**RD**	**ERD**
40	4	1	12	1000	d_1	OOM	88.34	51.23	50.21
40	4	1	12	1000	d_2	OOM	90.66	54.76	48.46
40	4	1	12	1000	d_3	OOM	89.24	60.40	43.27
40	4	1	12	1000	d_4	OOM	152.32	67.64	51.28
40	4	1	12	1000	d_5	OOM	103.24	54.74	67.24
40	4	1	12	1000	d_6	OOM	102.47	57.34	55.69
40	4	1	12	1000	d_7	OOM	104.47	54.90	45.58
40	4	1	12	1000	d_8	OOM	98.17	53.38	48.48
40	4	1	12	1000	d_9	OOM	93.81	60.51	44.55
40	4	1	12	1000	$d_1$0	OOM	90.04	53.19	46.48
Average							**101.28**	**56.81**	**50.12**

I	**P**	**B**	**T**	**S**	**D**	**Gurobi**	**MLS**	**RD**	**ERD**
50	5	1	12	1000	d_1	OOM	126.97	65.53	62.81
50	5	1	12	1000	d_2	OOM	190.58	77.51	77.30
50	5	1	12	1000	d_3	OOM	189.96	118.39	103.34
50	5	1	12	1000	d_4	OOM	183.64	72.89	69.58
50	5	1	12	1000	d_5	OOM	126.61	72.30	68.57
50	5	1	12	1000	d_6	OOM	191.83	86.76	72.63
50	5	1	12	1000	d_7	OOM	190.03	116.46	108.27
50	5	1	12	1000	d_8	OOM	188.73	115.86	110.90
50	5	1	12	1000	d_9	OOM	190.80	67.58	55.73
50	5	1	12	1000	d_{10}	OOM	128.39	70.18	64.31
Average							**170.75**	**86.35**	**79.34**

I	**P**	**B**	**T**	**S**	**D**	**Gurobi**	**MLS**	**RD**	**ERD**
60	6	1	12	1000	d_1	OOM	227.08	150.31	111.68
60	6	1	12	1000	d_2	OOM	229.65	159.33	110.65
60	6	1	12	1000	d_3	OOM	231.56	109.85	95.94
60	6	1	12	1000	d_4	OOM	322.03	200.92	123.39
60	6	1	12	1000	d_5	OOM	243.87	162.12	123.67
60	6	1	12	1000	d_6	OOM	240.04	162.35	116.21
60	6	1	12	1000	d_7	OOM	231.45	161.97	126.01
60	6	1	12	1000	d_8	OOM	323.20	168.89	127.85
60	6	1	12	1000	d_9	OOM	247.60	80.47	84.95
60	6	1	12	1000	d_{10}	OOM	227.30	89.19	88.38
Average							**252.38**	**144.54**	**110.87**

(*)OOM: out of memory

Figure 3 shows the vital role of biomass storage facility in a biorefinery site. Due to seasonality and short harvest time of biomass, we must collect enough biomass feedstock during the harvest season and store it later in order to cope with supply variability and to satisfy biofuel demand over the whole planning horizon. That explains why the inventory level of biomass during the harvest season (period $t = 1,2,3$) is higher than that in other periods. The biomass storage facility provides a buffer for the biomass supply chain to protect against the fluctuations of supply and demand.

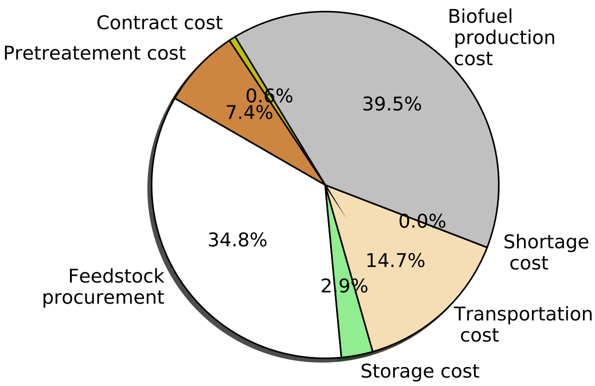

Figure 2: Cost distribution for the entire supply chain

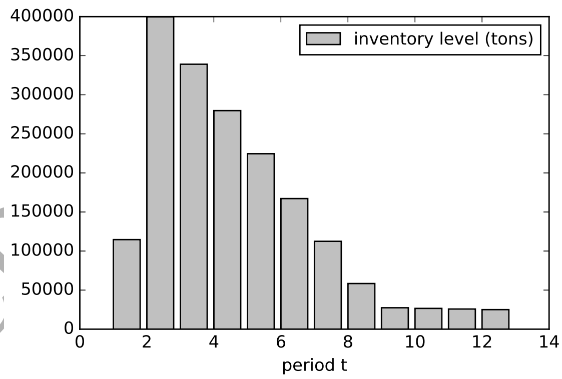

Figure 3: Evolution of biomass inventory level at biorefinery

5.4. Sensitivity analysis

We investigated the impacts of some critical parameters such as purchase price from a market, minimum supply quantity and fixed contracted cost on the expected total system cost since they have significant impacts on feedstock supply. A sensitivity analysis is conducted to learn how purchase price, minimum quantity to supply and fixed contracted cost impact the optimal expected total cost of biomass supply chain and the total number of contracted suppliers.

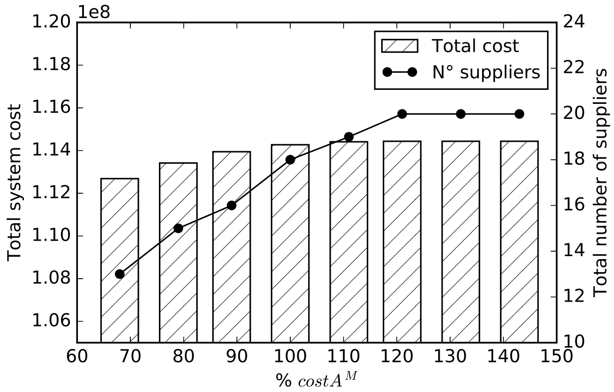

Figure 4: Impact of unit purchase price of biomass from market

5.4.1. Effect of purchase price

Figure 4 shows the total system cost and the number of contracted suppliers under different values of the purchase price which ranges between 70% and 140 % of its base value. The horizontal axis represents the various purchase price of biomass from an international/regional market. The bar graphs with the height measured by the vertical axis on the left correspond to the total costs of this supply chain under different values of the purchase price. The dotted line with the value of each point measured by the vertical axis on the right indicates the number of contracted suppliers at each possible value of the purchase price considered. As the market price increases, the optimal total system cost increases slightly from \$ 112 685 030 to \$ 114 410 480 (less than 2%), and then it remains constant at \$ 114 435 526 when the market price is set at 120% or higher than its basic level. This confirms that when the market price increases too high, decision-makers have less interest in purchasing an additional quantity of biomass from markets to meet biofuel demand. In this case, they would intend to establish more long-term supply contracts to ensure the stability of feedstock supply. This explains why the number of contracted suppliers increases when the market price grows. This result suggests that to achieve the overall system efficiency, a balance between purchase price under long-term contracts and the one from market supply need to be considered. In the long term, the quantity and price of biomass supplied from suppliers would be guaranteed by long-term contracts whereas the price and availability of biomass from markets vary highly due to external conditions such as weather.

5.4.2. Effect of fixed cost

Factors such as volatile market price, uncertain yield and limited land availability would contribute the difficulty in estimating a fixed purchase price and the minimum supply quantity in long-term contracts. This explains the motivation we conduct a study of the effect of fixed cost on the biomass supply chain. In Figure 5, the horizontal axis represents the fixed cost varying from 50% to 400% in comparison to its base value. The bar graphs with the height measured by the vertical axis on the left hand present the total expected system costs under different values of the fixed cost. The dotted line with the value of each point measured by the vertical axis on the right hand indicates the number of contracted suppliers at each possible value of the fixed cost. The

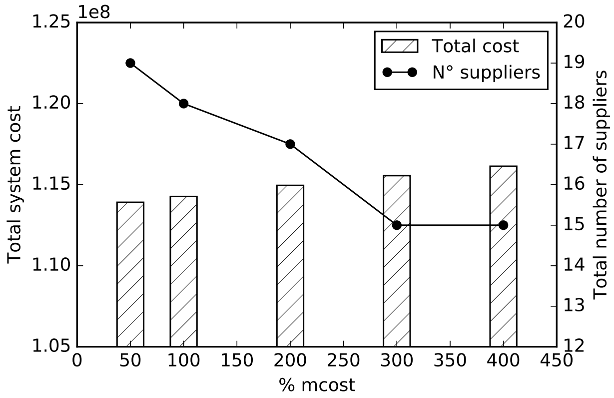

Figure 5: Impact of fixed cost for establish contracts

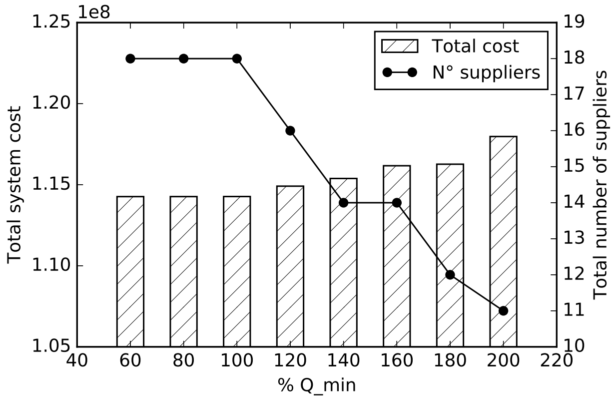

Figure 6: Impact of minimum contracted quantity of biomass

result indicates that the increase of fixed cost slightly affects the total system cost and the total number of contracted suppliers. This also reveals the fact that the increase in fixed cost could lead to higher total system cost and might hinder the establishment of long-term contracts.

5.4.3. Effect of the minimum supply quantity of suppliers

Figure 6 presents the impacts of a contracted supply quantity on the total system cost of biomass supply chain and the number of contracted suppliers. The horizontal axis represents the various contracted quantity delivered from 60% to 200% in comparison with its base value. The bar graphs with the height measured by the vertical axis on the left hand show the total expected system costs under different values of the minimum supply quantity. The dotted line with the value of each point measured by the vertical axis on the right hand indicates the number of contracted suppliers at each possible value of the minimum supply quantity.

This result shows that with the increase in minimum supply quantity, the total system cost will decrease; however, this decrease is not as significant as the increase in supply quantity. On the other hand, the total number of contracted suppliers decreases as the supply quantity increases.

In this case, decision-maker would require additional quantity from market to satisfy the final demand but only at a reasonable market price.

5.5. Value of the stochastic solution

In a deterministic approach, the decision-maker may assume expected supply availability based on average yield (12 tons/hectare) and seek for optimal contracts with suppliers (farmer) according to these supply availability. In other words, we formulate a deterministic model by using expected values of biomass supply availability. Based on the solution of the model found, we fix the values of the first-stage variables in formulation **(SP)** and then solve the model.

Table 4: Comparison of results from stochastic program

Solution	**Total costs ($)**
Deterministic	124 088 777
Stochastic	114 271 335
VSS	9 817 442
VSS (%)	7.9%

This approach represents the expected value solution as we mentioned before, but it might lead to unfavorable consequences. Here, as shown in Table 4, using the expected value solution could provide in the planning horizon the annual total cost of $ 124 088 777. This total cost is 7.9% higher than the one resulting from the corresponding stochastic programming model, $ 114 271 335. The choice of suppliers would affect the expected total cost strongly under the stochastic environments. We can consider that the value, $VSS = 124088777 - 114271335 = \$9817442$, represents the possible gain by taking account of the random variations of biomass supply. This value is called the value of the stochastic solution (VSS) according to Birge and Louveaux (2011).

5.6. Impact of integrated supplier selection and operation planning

Previously, most researchers considered two types of decisions, supplier selection and operation planning, separately. In this paper, we consider operation planning when making supplier selection, i.e., aggregating these two types of decisions in an integrated model. In this section, we provide both an intuitive reasoning and quantitative evidences for our choice.

For this purpose, we formulate a mathematical model (see Appendix B) in which the two types of decisions are made sequentially. In this model, supplier selection and operation planning are made sequentially with the following two phases:

(i) In the first phase, we only consider the selection of suppliers subject to constraints on supply availability (across a set of scenarios), biomass demand and the minimum supply quantity under each contract between a supplier and the bio-refinery. The biomass demand is estimated from the biofuel demand at the bio-refinery in each period.

(ii) In the second phase, operation planning under different scenarios is made with the suppliers selected in the first phase.

To evaluate the impact of the integrated decision-making, we tested some instances with $I = 30$; $P = 3$; $T = 12$ and number of scenarios $S = 50$, to compare two approaches: our integrated supplier selection and operation planning approach and the sequential planning approach. Table 5 shows

Table 5: Impact of integrated supplier selection and operation planning on number of suppliers selected: Baseline case

Instance	I_1	I_2	I_3	I_4	I_5	I_6	I_7	I_8	I_9	I_10
Integrated planning approach	11	11	12	11	11	11	11	10	12	13
Sequential approach	12	12	16	12	13	10	11	12	10	14

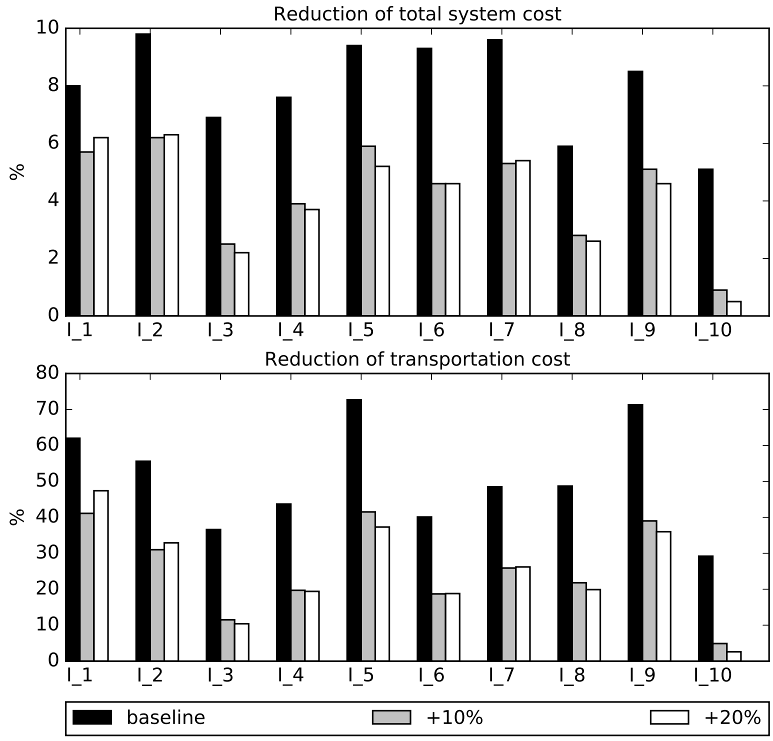

Figure 7: Impact of integrated supplier selection and operation planning on total system cost and transportation cost

the solution structure will change in terms of the number of supplier selected if supplier selection and operation planning are considered in an integrated way. The results depicted in Figure 7 show that the integrated planning approach can reduce 5.1% to 9.8% of the total supply chain cost in comparison with the sequential approach. Note that the reduction of transportation costs is also important, from 29% to 72% for the baseline case.

To evaluate the impact of biomass availability on the total system cost and supplier selection, we vary the quantity of biomass availability for each supplier from 10% to 20%. As the results show, the reduction of total cost realized by the integrated approach seems to decrease when the biomass availability increases. This can be explained by the fact that as the feedstock supply becomes abundant, a small number of suppliers could satisfy the biomass demand. The smaller number also leads to the reduction of fixed costs for establishing contracts with the suppliers and transportation costs.

Intuitively, without considering operating planning, in each scenario the refinery will not know exactly how many tons of biomass it requires to produce biofuel in each period, because the biofuel to produce in each period may be different from its demand. Consequently, it will not know exactly the quantity of biomass to be transported in each period from each contracted supplier or the market to it (via a pretreatment facility). In other words, the transportation quantities and the transportation costs considered in the supplier selection model in the first phase of the sequential approach are only approximate (in each scenario), this will lead to a significant increase of transportation costs as shown in Figure 7. On the other hand, the integrated planning approach can correct the adverse effects created by making strategic level decision (supplier selection) without considering related operational level decisions (production and transportation planning). In fact, the integrated planning approach can improve the relevance of suppliers selection by considering transportation costs and the availability of biomass resources as well as the deterioration of biomass during storage. Eventually, this causes the difference of supplier selection decisions between the two approaches and leads to that the sequential approach has higher transportation costs than the integrated approach at the operational level. That may explain why our approach has a significant lower transportation cost and a lower total system cost than the sequential planning approach.

The biomass availability seems to be a critical factor affecting the choice between the two approaches because the sequential planning approach could provide a good solution in a situation when the biomass yield is high, and the amount of biomass available at suppliers is abundant. However, this situation is too optimistic to account for because biomass feedstock supply is uncertain and always depends on weather condition and temperature. Thus, the quantitative results presented above reveal that it would be better to consider operation planning when making supplier selection decisions. It should be noted that the integration of strategical and tactical decisions is also adopted by problems in other domains such as location-routing problems.

6. Conclusions

We have proposed a mathematical model to tackle the supplier selection and operation planning problem in biomass supply chains to help decision-makers facing uncertainty of biomass feedstock supply. The objective is to minimize the total system cost of a biomass supply chain.

We have applied an enhanced and regularized L-shaped method to solve the two-stage stochastic programming model. This technique allows us to decompose a high dimensional stochastic model into subproblems of reasonable size. These sub-problems could be solved on a personal computer with limited memory. Moreover, our proposed method could find an optimal solution faster than the standard L-shaped decomposition method and commercial MILP solver Gurobi. Besides, we have analyzed the impacts of critical parameters on the optimal expected cost of the system and supplier selection.

Several directions for future research may be pursued as considering other uncertainties (price, quality, external demand, conversion technology). Another possibility is to integrate a more detailed transportation planning with the number of truck trips as a decision variable in each period. For this extension, the second stage problem may become a MILP model which imposes a high computational challenge. More effective algorithms may be developed to solve the complex and high dimensional model. Our ultimate goal is to develop a decision-making support tool for biomass supply chain management.

Appendix A. Data of model parameters

Table A.6: Data of model parameters

Input parameters	Value
Bioethanol yield (litre/tonne)	313
Unit shortage cost for ummet bioethanol demand ($/litre)	1.06
Unit bioethanol production cost ($/litre)	0,200
Unit biomass storage cost at biorefinery ($/litre)	0.9
Unit biofuel storage cost ($/ton)	0.228
Annual production capacity of biorefinery(MLPY)	190
Unit preprocessing cost at preprocessing facility ($/tonne)	13.94
Unit biomass storage cost at pretreatment facility ($/tonne)	1.125
Transformation yield rate at pretreatment facility i	75
Ratio of biomass storage loss in time period (%)	1-3%
Annual production capacity of pretraitement facility (Mton)	2.00
Unit biomass purchase price from market ($/ton)	55
Unit biomass purchase price from suppliers ($/ton)	42-45
Minimum quantitity to supply from each supplier (tons/year)	20 000-30 000
Fixed cost for establishing a contract ($)	20 000-30 000
Demand of biofuel (MLPY)	$N(2133,4138)$ on $[2016,2280]$
Yield biomass feedstocks (tons/hecta)	Uniform[10, 14]

Appendix B. The sequential approach

In this model, supplier selection and operation planning are made sequentially with the following two phases. In the first phase, we only consider the selection of suppliers subject to constraints on supply availability across a set of scenarios (B.1-B.2), biomass demand (B.3) and the minimum supply quantity under contract (B.4). The biomass demand is estimated from the biofuel demand at the biorefinery in each period. The objective is to minimize the total cost including fixed costs for establishing a contract and purchasing costs from suppliers and markets. The mathematical model for the first phase is given as follows:

$$\min \quad \sum_{i\in I} mcost(i)\varphi(i) + \sum_{t\in T}\sum_{s\in S} p_s \left[\sum_{i\in I} x(i,t,s)costA^I(i) + \sum_{m\in M} x(m,t,s)costA^M(m)\right]$$

$$\text{s.t.} \quad x(i,t,s) \le avail^I(i,t,s).\varphi(i) \quad \forall t\in T, i\in I, s\in S \tag{B.1}$$

$$x(m,t,s) \le avail^M(m,t,s) \quad \forall t\in T, m\in M, s\in S \tag{B.2}$$

$$\sum_{i\in I} x(i,t,s) + \sum_{m\in M} x(m,t,s) = \bar{d}_t \quad \forall t\in T, s\in S \tag{B.3}$$

$$\sum_{t\in T} x(i,t,s) \ge q_{min}(i)\varphi(i) \forall t\in T, s\in S \tag{B.4}$$

where $x(i,t,s)$ and $x(m,t,s)$ are the quantity of biomass purchased from supplier $i \in I$ and market $m \in M$ in period t under scenario s, respectively. $\bar{d_t}$ is the biomass demand in each period considered. It is estimated from the biofuel demand $demand(b,t)$, the biofuel rate $\eta^B(b)$ and the preprocessing rate η^P, as follows: $\bar{d_t} = \sum_{b \in B} demand(b,t)/(\eta^B(b) \times \eta^P)$.

(ii) In the second phase, operation planning under different scenarios is made with the set of suppliers selected in the first phase (i.e the value of vector φ is known in this phase). The mathematical model for the second phase is given as follows:

$$\min_{y_s} \quad \sum_{s \in S} p_s f(\varphi, s, y_s)$$
$$s.t. \quad \text{constraints} \quad (5)-(20) \qquad \forall s \in S$$

It is necessary to recompute the total system cost by combining the operational costs obtained in the second phase with the costs for establishing biomass purchase contracts in the first phase.

References

Amorim, P., Curcio, E., Almada-Lobo, B., Barbosa-Póvoa, A. P., Grossmann, I. E., 2016. Supplier selection in the processed food industry under uncertainty. European Journal of Operational Research 252 (3), 801–814.

Andersson, G., Asikainen, A., Björheden, R., Hall, P., Hudson, J., Jirjis, R., Mead, D., Nurmi, J., Weetman, G., 2002. Production of forest energy. Bioenergy from Sustainable Forestry, 49–123.

Awudu, I., Zhang, J., 2013. Stochastic production planning for a biofuel supply chain under demand and price uncertainties. Applied Energy 103, 189–196.

Ba, B. H., , C., Prodhon, C., 2016. Models for optimization and performance evaluation of biomass supply chains: An operations research perspective. Renewable Energy 87, 977–989.

Ba, B. H., Prins, C., Prodhon, C., 2015. A new tactical optimization model for bioenergy supply chain. In: International Conference on Applied Mathematics and OR. Vol. 2.

Birge, J. R., Louveaux, F., 2011. Introduction to stochastic programming. Springer Science & Business Media.

Birge, J. R., Louveaux, F. V., 1988. A multicut algorithm for two-stage stochastic linear programs. European Journal of Operational Research 34 (3), 384–392.

Bot, P., van Donk, D. P., Pennink, B., Simatupang, T. M., 2015. Uncertainties in the bidirectional biodiesel supply chain. Journal of Cleaner Production 95, 174–183.

Castillo-Villar, K. K., Minor-Popocatl, H., Webb, E., 2016. Quantifying the impact of feedstock quality on the design of bioenergy supply chain networks. Energies 9 (3), 203.

Chen, C.-W., Fan, Y., 2012. Bioethanol supply chain system planning under supply and demand uncertainties. Transportation Research Part E: Logistics and Transportation Review 48 (1), 150–164.

Clifton-Brown, J. C., Breuer, J., Jones, M. B., 2007. Carbon mitigation by the energy crop, miscanthus. Global Change Biology 13 (11), 2296–2307.

Duc-Huy, N., Haoxun, C., Nengmin, W., 2017. Modeling and optimization of biomass supply chain with two types of feedstock suppliers. In: Proceedings of IESM 2017 7th International Conference on Industrial Engineering and Systems Management. Saarbrcken, Germany, pp. 610–615.

Ekşioğlu, S. D., Acharya, A., Leightley, L. E., Arora, S., 2009. Analyzing the design and management of biomass-to-biorefinery supply chain. Computers & Industrial Engineering 57 (4), 1342–1352.

Gold, S., Seuring, S., 2011. Supply chain and logistics issues of bio-energy production. Journal of Cleaner Production 19 (1), 32–42.

Govindan, K., Khodaverdi, R., Jafarian, A., 2013. A fuzzy multi criteria approach for measuring sustainability performance of a supplier based on triple bottom line approach. Journal of Cleaner Production 47, 345–354.

Hammami, R., Temponi, C., Frein, Y., 2014. A scenario-based stochastic model for supplier selection in global context with multiple buyers, currency fluctuation uncertainties, and price discounts. European Journal of Operational Research 233 (1), 159–170.

Han, S.-K., Murphy, G. E., 2012. Solving a woody biomass truck scheduling problem for a transport company in western oregon, usa. Biomass and Bioenergy 44, 47–55.

Hombach, L. E., Cambero, C., Sowlati, T., Walther, G., 2016. Optimal design of supply chains for second generation biofuels incorporating european biofuel regulations. Journal of Cleaner Production 133, 565–575.

Kaut, M., Egging, R., Flatberg, T., Uggen, K. T., 2015. Blomstan optimization model for the bioenergy supply chain. In: Handbook of Bioenergy. Springer, pp. 37–66.

Kim, J., Realff, M. J., Lee, J. H., 2011. Optimal design and global sensitivity analysis of biomass supply chain networks for biofuels under uncertainty. Computers & Chemical Engineering 35 (9), 1738–1751.

Laporte, G., Louveaux, F. V., 1993. The integer l-shaped method for stochastic integer programs with complete recourse. Operations research letters 13 (3), 133–142.

Li, J.-M., Li, A.-H., Varbanov, P. S., Liu, Z.-Y., 2017. Distance potential concept and its applications to the design of regional biomass supply chains and solving vehicle routing problems. Journal of Cleaner Production.

Li, Z., Liu, Y., Liao, W., Chen, S., Zemetra, R. S., 2011. Bioethanol production using genetically modified and mutant wheat and barley straws. biomass and bioenergy 35 (1), 542–548.

Linderoth, J., Shapiro, A., Wright, S., 2006. The empirical behavior of sampling methods for stochastic programming. Annals of Operations Research 142 (1), 215–241.

Marufuzzaman, M., Eksioglu, S. D., Huang, Y. E., 2014. Two-stage stochastic programming supply chain model for biodiesel production via wastewater treatment. Computers & Operations Research 49, 1–17.

Mittelmann, H., 2018. Benchmarks for optimization software. Available as `http://plato.la.asu.edu/bench.html`.

Morales-Chvez, M. M., Jos A, S.-M., William, A. S., 2016. A mixed-integer linear programming model for harvesting, loading and transporting sugarcane. a case study in peru. DYNA 83 (195), 173–179.

Osmani, A., Zhang, J., 2013. Stochastic optimization of a multi-feedstock lignocellulosic-based bioethanol supply chain under multiple uncertainties. Energy 59, 157–172.

Osmani, A., Zhang, J., 2014. Economic and environmental optimization of a large scale sustainable dual feedstock lignocellulosic-based bioethanol supply chain in a stochastic environment. Applied Energy 114, 572–587.

Pérez, A. T. E., Camargo, M., Rincón, P. C. N., Marchant, M. A., 2017. Key challenges and requirements for sustainable and industrialized biorefinery supply chain design and management: A bibliographic analysis. Renewable and Sustainable Energy Reviews 69, 350–359.

Rentizelas, A. A., Tolis, A. J., Tatsiopoulos, I. P., 2009. Logistics issues of biomass: the storage problem and the multi-biomass supply chain. Renewable and Sustainable Energy Reviews 13 (4), 887–894.

Ruszczyński, A., Świtanowski, A., 1997. Accelerating the regularized decomposition method for two stage stochastic linear problems. European Journal of Operational Research 101 (2), 328–342.

Sahinidis, N. V., 2004. Optimization under uncertainty: state-of-the-art and opportunities. Computers & Chemical Engineering 28 (6-7), 971–983.

Samsatli, S., Samsatli, N. J., Shah, N., 2015. Bvcm: a comprehensive and flexible toolkit for whole system biomass value chain analysis and optimisation–mathematical formulation. Applied Energy 147, 131–160.

Scott, J. A., Ho, W., Dey, P. K., 2013. Strategic sourcing in the uk bioenergy industry. International journal of production economics 146 (2), 478–490.

Shabani, N., Akhtari, S., Sowlati, T., 2013. Value chain optimization of forest biomass for bioenergy production: a review. Renewable and Sustainable Energy Reviews 23, 299–311.

Shabani, N., Sowlati, T., 2016a. Evaluating the impact of uncertainty and variability on the value chain optimization of a forest biomass power plant using monte carlo simulation. International journal of green energy 13 (7), 631–641.

Shabani, N., Sowlati, T., 2016b. A hybrid multi-stage stochastic programming-robust optimization model for maximizing the supply chain of a forest-based biomass power plant considering uncertainties. Journal of Cleaner Production 112, 3285–3293.

Shabani, N., Sowlati, T., Ouhimmou, M., Rönnqvist, M., 2014. Tactical supply chain planning for a forest biomass power plant under supply uncertainty. Energy 78, 346–355.

Shafiee, S., Topal, E., 2009. When will fossil fuel reserves be diminished? Energy policy 37 (1), 181–189.

Shapiro, A., 2000. Stochastic programming by monte carlo simulation methods.

Shapiro, A., Homem-de Mello, T., 1998. A simulation-based approach to two-stage stochastic programming with recourse. Mathematical Programming 81 (3), 301–325.

Sharma, B., Ingalls, R. G., Jones, C. L., Khanchi, A., 2013. Biomass supply chain design and analysis: basis, overview, modeling, challenges, and future. Renewable and Sustainable Energy Reviews 24, 608–627.

Van Slyke, R. M., Wets, R., 1969. L-shaped linear programs with applications to optimal control and stochastic programming. SIAM Journal on Applied Mathematics 17 (4), 638–663.

Woo, Y.-b., Cho, S., Kim, J., Kim, B. S., 2016. Optimization-based approach for strategic design and operation of a biomass-to-hydrogen supply chain. International Journal of Hydrogen Energy 41 (12), 5405–5418.

Yue, D., You, F., Snyder, S. W., 2014. Biomass-to-bioenergy and biofuel supply chain optimization: overview, key issues and challenges. Computers & Chemical Engineering 66, 36–56.

Zhang, F., Johnson, D. M., Johnson, M. A., 2012. Development of a simulation model of biomass supply chain for biofuel production. Renewable Energy 44, 380–391.