Title
Multistage Infrastructure System Design: An Integrated Biofuel Supply Chain against Feedstock Seasonality and Uncertainty

Authors
Fan, Yueyue
Huang, Yongxi
Chen, Chien-Wei

Publication Date
2012-09-01

Research Report – UCD-ITS-RR-12-25

Multistage Infrastructure System Design: An Integrated Biofuel Supply Chain against Feedstock Seasonality and Uncertainty

September 2012

Yueyue Fan
Yongxi Huang
Chien-Wei Chen

Institute of Transportation Studies ◦ University of California, Davis
One Shields Avenue ◦ Davis, California 95616
PHONE (530) 752-6548 ◦ FAX (530) 752-6572
www.its.ucdavis.edu

Final Product

You are required to submit a final product for your project. This product will usually be in the form of a research report published through the ITS on-line database. Peer-review publications can be posted in lieu of a research report, provided they provide equivalent documentation of the project. In either case, the official STC acknowledgement must be included, and research reports should also include the official disclaimer, both provided below. For some projects, other types of products (e.g. CD-ROM, website) might be appropriate.

Research reports should give a complete description of the problem, approach, methodology, findings, conclusions, recommendations, etc., developed in the project and completely documents all data gathered, analyses performed, and results achieved.

Please use a front cover that looks like the template on the following page (feel free to copy-paste within bracketed text.

Please include a table of contents, list of tables, and list of figures showing corresponding page numbers.

<u>All reports and papers should include the following acknowledgement:</u>

> **ACKNOWLEDGEMENT**
> The research was supported by a grant from the Sustainable Transportation Center at the University of California Davis, which receives funding from the U.S. Department of Transportation and Caltrans, the California Department of Transportation, through the University Transportation Centers program.

<u>The inside of the front cover of the research report must show a disclaimer including the following:</u>

> **DISCLAIMER**
> The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the information presented herein. This document is disseminated under the sponsorship of the Department of Transportation University Transportation Centers Program, in the interest of information exchange. The U.S. Government assumes no liability for the contents or use thereof.

STC Final Research Report [R04-2]

Multistage Infrastructure System Design: An Integrated Biofuel Supply Chain against Feedstock Seasonality and Uncertainty

September, 2012

Yueyue Fan
Yongxi Huang
Chien-Wei Chen

Department of Civil and Environmental Engineering
University of California, Davis
One Shields Ave., Davis, CA 9561

Abstract

A biofuel supply chain consists of various interdependent components from feedstock resources all the way to energy demand sites. This study focuses on the design of an efficient biofuel supply chain system against seasonal variations and uncertainties of feedstock supply in an integrative manner. By integrating planning and operational decisions in a stochastic programming framework, we aim at finding an effective design strategy for biofuel supply chain that is economically viable and hedges well against a wide range of future uncertainties. A solution algorithm based on scenario decomposition is designed to overcome computational challenges involved in large-scale applications. A California case study is implemented to demonstrate the applicability of the proposed methods in evaluating the economic potential, the infrastructure needs, and the risk of wastes-based bioethanol production.

Keywords: biofuel supply chain, seasonality, uncertainty, stochastic programming, decomposition

1. Introduction

The goal of this paper is to establish an efficient biofuel supply chain system against potential feedstock supply uncertainty and seasonality, by integrating the planning and operation of an entire supply chain. Our focus is on biofuels that can be converted from cellulosic biomass such as biowastes and dedicated energy crops. Comparing to conventional liquid fuels and corn-based biofuel, cellulosic biofuels have better performances in terms of reducing greenhouse gas emissions (European Parlinament and Council, 2003; U.S. Congress, 2007), diversifying transportation fuels, and providing a solution to food versus energy debate (De La Torre et al., 2000; McCarl and Schneider, 2001; Perlack et al., 2005). To facilitate the success of cellulosic biofuel industry, an efficient supply chain system that ensures strong cost competiveness and reliability is crucial.

Studies have shown the interdependence of various components of biofuel supply chains and the importance of planning the system as a whole (Delucchi, 2006; Farrell and Sperling, 2007; Kim and Dale, 2005; Turner and Plevin, 2007; Unnasch and Pont, 2007; Zah et al., 2007). However, most existing studies on biofuel infrastructure system planning only considered part of the systems. For example, Freppaz et al. (2004) considered a production system from biomass to thermal and electricity generation; Tembo et al. (2003) considered a system including feedstock production, delivery, and processing. Only a limited number of studies (Parker et al., 2008) have adopted a "supply chain" concept (also referred as "well-to-wheel" approach in energy literature) in biofuel system planning. These studies were based on deterministic approaches, which assume perfect foresight of model input parameters.

The concept of supply chain, through better integration and coordination of various components of a supply system (such as procurement, production, storage, and marketing), can greatly improve system efficiency. On the other hand, reduced redundancy and buffer, which improve system efficiency under normal conditions, may present higher risk under unexpected events such as supply shortage, demand spike, technological failure, or attacks and disasters. Following the definition by Tang (2006), supply chain risks are categorized into operational risks and disruption risks. An operational risk refers to those recurrent risks such as supply and demand fluctuations that are inherent in supply chains. A disruption risk usually refers to external disruptions caused by natural and man-made disasters. By this definition, the risk we are addressing, feedstock supply uncertainty, belongs to the category of recurrent operational risks.

Despite of the importance of addressing uncertainties in biofuel supply system planning as identified in (Ekşioğlu et al., 2009; International Energy Agency, 2006), to our knowledge there are only three stochastic models in biofuel supply chain literature. Cundiff et al. (1997) focused only on storage facilities for herbaceous biomass; Dal-Mas et al. (2011) designed a multi-year corn-ethanol supply chain considering uncertain prices of corn, bioethanol, and Dried Distillers Grains with Solubles (DDGS); Chen and Fan (2012) established a stochastic biofuel supply chain model under uncertainties of feedstock supply and fuel demand, and proposed a decomposition method for solving large-scale problems. All these studies considered aggregated yearly operations and did not consider feedstock seasonality that inevitably introduces system interdependence across temporal dimension.

Since most production and delivery infrastructures of the emerging cellulosic biofuel industry are not in place yet, there presents an opportunity for incorporating risk and seasonality directly into the strategic planning of its supply chain systems. Strategic supply chain management aims at finding the best supply chain configuration, including location setup, procurement, production, storage, and distribution, to support efficient operations of the whole supply chain (Cordeau et al., 2006). A comprehensive reviews of recent progress in strategic supply chain management are given by Synder (2006) and Melo et al. (2007).

The key feature distinguishing this study from most existing work on biofuel supply chain is the integration of physical design and operational management as a whole in seeking mitigation strategies against feedstock uncertainty and seasonality. Facility spatiality, time variation of feedstock yields, and uncertainty are integrated into a stochastic programming framework. Optimal strategies on feedstock procurement, biofuel production, feedstock and fuel storage, and delivery are sought simultaneously to achieve the least expected total system cost. The proposed model is used to evaluate the economic potential and system effectiveness of converting corn stover and forest residues to ethanol in California. The real-world case study provides a realistic model incorporating both system dynamics and uncertainties. An efficient solution algorithm based on scenario decomposition is developed to overcome computational challenges involved in solving large-scale mix-integer stochastic programming problems. As identified in a recent comprehensive review on supply chain management (Melo et al., 2007), large-scale implementation of realistic models considering both dynamics and uncertainty is lacking in existing literature.

The paper is organized as follows. Model description and formulation are given in Section 2. A case study of converting corn stover and forest residues to ethanol in California, together with numerical implementation and sensitivity analysis, is detailed in Section 3. Conclusions and discussions are presented in Section 4.

2. Model Formulation

2.1 Modeling Background

A supply chain representation: Figure 1 represents a biofuel supply chain system from waste resources to end users, including feedstock storage, fuel production, and fuel storage in between. The arrows in Figure 1 denote flow (feedstock or fuel) directions. Note that the supply chain ends at city gates and that further fuel dispensing to individual refueling stations is omitted in this study. Strategic planning of this supply chain includes designing of the physical configuration of the supply chain system such as locations and the sizes of the production and storage facilities, as well as making corresponding operational decisions such as procurement strategy of the feedstock, production and storage amount, and flow transported between different layers of the supply chain.

Spatial and temporal dimensions: This problem spans over both spatial and temporal dimensions. Spatial dimension comes from the geographical distribution of the feedstock supply, facility locations, and demand sites. Temporal dimension is mainly caused by seasonality of the feedstock supply. Design for such a complex system is not trivial due to several tradeoffs in the system. For example, a centralized facility takes the advantage of economy of scale, but may result in higher transport cost. Storage of feedstock and fuel may impose an extra cost to the system, but may lower the risk of future supply shortage. By integrating the physical design of infrastructure systems and the operation management, this study aims to capture the system interdependence and balance the tradeoffs in both temporal and spatial dimensions.

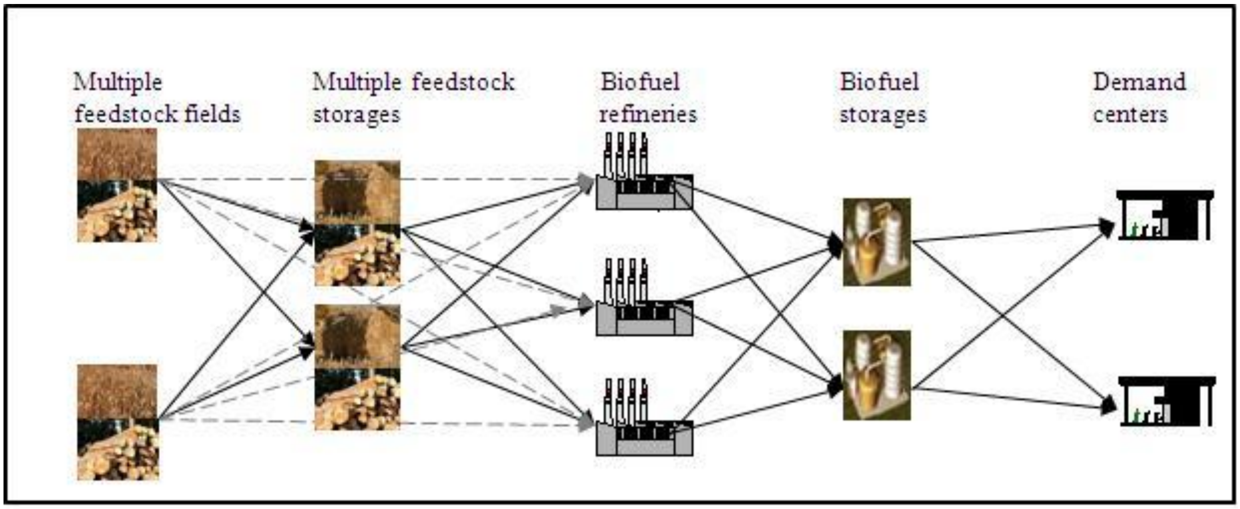


Figure 1. A Complete Biofuel Pathway

Planning vs. operational decisions under uncertainty: In addition to the system interdependence and the supply seasonality, handling uncertainty imposes another modeling challenge. In this study, we focus on the uncertainty of feedstock availability, which may be caused by climate (e.g., flood or drought) and natural disasters (e.g., wild fires). Planning decisions such as the locations and sizes of facilities (i.e., feedstock storages, refineries, and fuel storages) are usually made before the uncertain supply is known. Once these decisions are implemented, they cannot be easily modified. Operational decisions such as the production and storage quantities can be adjusted based on the actual realization of the uncertain supply. This feature fits well in a stochastic programming framework (Birge and Louveaux, 1997; Louveaux, 1986), which recognizes the non-anticipativity of planning decisions while allowing recourse for operational decisions.

Modeling framework: In this study, feedstock supply is assumed to take a discrete set of possible scenarios with associated probabilities. A mixed integer stochastic programming model is developed with a goal of minimizing the expected total system cost across all possible scenarios. To reflect the temporal dimension of the problem, all recourse decision variables are time (season) dependent. The decision variables to be determined by the model include:

- locations and sizes of refineries and fuel storage facilities (feedstock storage facilities have negligibly low capital cost),
- feedstock procurements,
- feedstock storages and deliveries, and
- ethanol productions, storages and distributions.

Complete notations for decision variables and model parameters are defined in Table 1.

Table 1. Notation Table

Parameters
T: index t, set of time phases (seasons);
Ω: index ω, set of uncertain scenarios;
L: index l, set of feedstock types;
I_l: index i_l, set of feedstock fields of type l;
H_l: index h_l, set of feedstock storage facilities of type l; it is assumed that feedstock of type l can only be stored at h_l;
S: index s, set of biofuel storage sizes - large, median, and small;
J: index j, set of ethanol refineries;
K: index k, set of ethanol storage facilities (terminals);
M: index m, set of demand cities;
$proc_l$: inelastic procurement cost (\$/dry ton) of feedstock type l;
$prod$: biofuel production cost (\$/gallon);
v: truck average traveling speed (mile/hr);
$fstor_l$: unit cost of feedstock storage (\$/dry ton);
$fcap_j$: refinery annualized fixed capital cost (\$) at location j;
$vcap_j$: refinery variable capital cost (\$/gallon) at location j;
$mcap_j$: the maximum allowable refinery capacity (gallon) at location j;
$fscap_k$: receiving facility and blending system cost (\$) at location k;
$fsvcap_k^s$: capital cost (\$) of fuel tank with size s;
$estor$: unit cost of ethanol storage (\$/gallon);
$mfscap^s$: the fuel storage capacity (gallon) of tank size s;
η_l : conversion rate (gallon/dry ton), measuring amount of biofuel converted from one unit of feedstock of type l;
MC_l: feedstock moisture content of type l;
d_{ij}: distance between node i and j;
$tdisb$:distance dependent transportation cost (\$/mile/truckload) of bulk solids, i.e., the cost of traveling one mile per truckload, including truck fuel, insurance, maintenance, and permitting expenses;
$ttimb$: travel time dependent transportation cost (\$/hr/truckload) of bulk solids, i.e., the cost of traveling one hour per truckload, including labor and capital costs;
$tdislq$:distance dependent transportation cost (\$/mile/truckload) of liquids;
$ttimlq$: travel time dependent transportation cost (\$/hr/truckload) of liquids;
$trcapb$: truck capacity (wet ton) of bulk solids, which varies with different feedstocks due to the moisture content;
$trcaplq$: truck capacity (gallon) of liquids;
$trLUb$: truck loading and unloading cost of bulk solids (\$/wet ton);
$trLUlq$:truck loading and unloading cost of liquids (\$/gallon);
$yield_{i_l}^t(\omega)$: feedstock yields of type l at field (dry ton) i_l during time t under scenario ω;
D_m^t : biofuel demand at city m during time t;

de_l^t : feedstock deterioration rate of type l during time t;

pl: penalty cost - unit cost of importing fuels ($/gallon)

Decision Variables
$zplant_j$:= 1 if refinery is opened at location j; =0 otherwise;
$zfuel_k^s$:= 1 if an biofuel storage of size s is placed at location k; =0 otherwise;
$zcap_j$: the size of refinery (gallon) at location j;
$Y_{i_l}^t(\omega)$: the amount of harvested feedstock of type l at i_l during time t under scenario ω;
$x_{ij}^t(\omega)$: the amount of feedstock/fuel transported from node i to node j during time t under scenario ω;
$FSQ_{h_l}^t(\omega)$: the amount of storage of feedstock type l at facility h_l at the beginning of time t under scenario ω;
$FQ_k^t(\omega)$: the quantity of biofuel available in the fuel storage k at the beginning of time t under scenario ω;
$yin_{lj}^t(\omega)$: the total supply to refinery at location j by feedstock type l during time t under scenario ω;
$pr_j^t(\omega)$: the amount of biofuel produced at refinery j during time t under scenario ω;
$q_m^t(\omega)$: the quantity of imported fuels for city m during time t under scenario ω.

2.2 Mathematical Formulation

The following multistage stochastic programming model is established.

Minimize:

$$\sum_{j\in J}(fcap_j\, zplant_j + vcap_j\, zcap_j) + \sum_{s\in S}\sum_{k\in K}(fscap_k + fsvcap_k^s) zfuel_k^s +$$

$$E_\omega \sum_{t\in T}\left(T_{feedstock}(\omega,t) + T_{fuel}(\omega,t) + \sum_{l\in L}\sum_{h_l\in I_l} fstor_l\, FSQ_{h_l}^t(\omega) + \sum_{j\in J} prod * pr_j^t(\omega) + \sum_{k\in K} estor * FQ_k^t(\omega) + \sum_{l\in L}\sum_{i\in I_l} proc_l\, Y_{i_l}^t(\omega) + \sum_{m\in M} pl \times q_m^t(\omega) \right) \quad (1)$$

$$T_{feedstock}(\omega,t) =$$

$$\sum_{l\in L}\sum_{i_l\in I_l}\sum_{j\in J}\left(\frac{(tdisb + \frac{ttimb}{v})d_{i_l j}}{trcapb} + trLUb \right)\frac{x_{i_l j}^t(\omega)}{MC_l} + \sum_{l\in L}\sum_{i_l\in I_l}\sum_{h_l\in I_l}\left(\frac{(tdisb + \frac{ttimb}{v})d_{i_l h_l}}{trcapb} + trLUb \right)\frac{x_{i_l h_l}^t(\omega)}{MC_l} + \sum_{l\in L}\sum_{h_l\in I_l}\sum_{j\in J}\left(\frac{(tdisb + \frac{ttimb}{v})d_{h_l j}}{trcapb} + trLUb \right)\frac{x_{h_l j}^t(\omega)}{MC_l} \quad (1.a)$$

$$T_{fuel}(\omega,t) =$$

$$\sum_{j\in J}\sum_{k\in K}\left(\frac{(tdislq + \frac{ttimlq}{v})d_{jk}}{trcaplq} + trLUlq \right) x_{jk}^t(\omega) + \sum_{k\in K}\sum_{m\in M}\left(\frac{(tdislq + \frac{ttimlq}{v})d_{km}}{trcaplq} + trLUlq \right) x_{km}^t(\omega) \quad (1.b)$$

The objective function (1) minimizes the expected total system cost, including the costs associated with planning and operational decisions. The planning-stage cost is the sum of facility capital costs. Since planning decisions are non-distinguishable across all scenarios, their cost is deterministic. The operational decisions are scenario dependent, so are the costs involved with feedstock procurement, storage, and delivery ($T_{feedstock}$), as well as fuel production and distribution (T_{fuel}). The delivery costs for feedstock and fuel, $T_{feedstock}$ and T_{fuel}, have similar cost structures. Let us use $T_{feedstock}$ in (1.a) as an example to explain the cost structure. It contains costs from three possible delivery trips - between field and refinery (denoted as *ij*), between field and storage (denoted as *ih*), and between storage and refinery (denoted as *hj*). For each trip, the transport cost is estimated by distance (*tdisb*)- and time (*ttimb*)- dependent costs, both of which are divided by truck capacity (*trcapb*) to convert the delivery quantity to number of truckloads. Truck loading/unloading (*trLUb*) cost is also considered and is linear to the delivery quantity. The moisture content (*MC*) is used to convert the feedstock dry ton to wet ton, for which the truck capacity is based on. Expression (1.b) can be explained similarly. The last term in (1) involving $q_m^t(\omega)$ denotes the penalty cost imposed on not producing enough fuel to satisfy demand. The penalty cost may be interpreted as the cost of importing fuels from other states or regions.

Subject to:

Constraints on biofuel refineries:

$$\sum_{i_l \in I_l} x_{i_l j}^t(\omega) + \sum_{h_l \in I_l} x_{h_l j}^t(\omega) = yin_{lj}^t(\omega) \quad \forall j \in J, t \in T, l \in L, \omega \in \Omega \tag{2}$$

$$\sum_{k \in K} x_{jk}^t(\omega) = pr_j^t(\omega) \quad \forall j \in J, t \in T, \omega \in \Omega \tag{3}$$

$$zcap_j = \max_{\omega}\{\sum_{t \in T} pr_j^t(\omega)\} \quad \forall j \in J \tag{4}$$

$$zcap_j \le mcap_j \quad \forall j \in J \tag{5}$$

$$\sum_{l \in L} yin_{lj}^t(\omega) \times \eta_l = pr_j^t(\omega) \quad \forall j \in J, t \in T, \omega \in \Omega \tag{6}$$

$$zcap_j \le \overline{M} z_j \quad \forall j \in J \tag{7}$$

Equation (2) computes the total amount of feedstock available at refinery *j* during time *t* under scenario ω. Equation (3) assures flow conservation constraints on refineries, stating that all produced biofuel is transported to fuel storages. The refinery size ($zcap_j$) is defined in equation (4), which equals the maximum annual fuel production in all scenarios. Constraint (5) limits the size by its maximum allowable capacity. Equation (6) calculates feedstock-to-fuel conversions. Constraint (7) is a logic constraint, meaning that there is no fuel production unless a refinery is operating at *j*.

Constraints on feedstock sites:

$$Y_{i_l}^t(\omega) \le yield_{i_l}^t(\omega) \quad \forall i_l \in I_l, t \in T, l \in L, \omega \in \Omega \tag{8}$$

$$Y_{i_l}^t(\omega) = \sum_{j \in J} x_{i_l j}^t(\omega) + \sum_{h_l \in I_l} x_{i_l h_l}^t(\omega) \quad \forall i_l \in I_l, t \in T, l \in L, \omega \in \Omega \tag{9}$$

Constraint (8) ensures that the total harvested feedstock cannot exceed its availability in each season. Note that the feedstock availability varies with its type *l* and time *t*. Similar to constraint (3), equation (9) is the feedstock flow conservation constraint.

Constraints on feedstock storages:

$$FSQ_{h_l}^{t+1}(\omega) = de_l^t FSQ_{h_l}^t(\omega) + \sum_{h_l \in I_l} x_{i_l h_l}^t(\omega) - \sum_{j \in J} x_{h_l j}^t(\omega) \quad \forall h_l \in I_l, t \in T, l \in L, \omega \in \Omega \quad (10)$$

Equation (10) defines the feedstock storage ($FSQ_{h_l}^{t+1}$) at the beginning of season *t*+1, which consists of discounted feedstock inventory during season *t* ($FSQ_{h_l}^t$) due to feedstock loss (accounted by de_l^t), and net feedstock inflow ($\sum_{h_l \in I_l} x_{i_l h_l}^t(\omega) - \sum_{j \in J} x_{h_l j}^t(\omega)$) in season *t*.

Constraints on biofuel storages:

$$FQ_k^{t+1}(\omega) = FQ_k^t(\omega) + \sum_{j \in J} x_{jk}^t(\omega) - \sum_{m \in M} x_{km}^t(\omega) \quad \forall k \in K, t \in T, \omega \in \Omega \quad (11)$$

$$FQ_k^t(\omega) \le \sum_{s \in S} mfscap^s zfuel_k^s \quad \forall k \in K, t \in T, \omega \in \Omega \quad (12)$$

$$\sum_{s \in S} zfuel_k^s \le 1 \quad \forall k \in K \quad (13)$$

Constraint (11) is the flow conservation constraint on fuel storage, which can be similarly explained as for constraint (10) except that there is no deterioration discount on fuels over time. Fuel storage is limited by its total capacity in constraint (12), where fuel storage size is chosen from a set of discrete values. Constraint (13) is to exclude the possibility of having more than one size of fuel tanks at a single site.

Constraints on satisfying the demands:

$$\sum_{k \in K} x_{km}^t(\omega) + q_m^t(\omega) = D_m^t \quad \forall m \in M, t \in T, \omega \in \Omega \quad (14)$$

Equation (14) requires the total demand to be satisfied by instate production and/or imports.

Integer and nonnegativity constraints:

$$zplant_j \in \{0,1\} \quad \forall j \in J \quad (15)$$

$$zfuel_k \in \{0,1\} \quad \forall k \in K \quad (16)$$

$$zcap_j \ge 0 \quad \forall j \in J \quad (17)$$

$$x_{ij}^t(\omega) \ge 0 \quad \forall i \in I, j \in J, t \in T, \omega \in \Omega \quad (18)$$

$$FSQ_{h_l}^t(\omega) \ge 0 \quad \forall h_l \in I_l, l \in L, t \in T, \omega \in \Omega \quad (19)$$

$$FQ_k^t(\omega) \ge 0 \quad \forall k \in K, t \in T, \omega \in \Omega \quad (20)$$

$$yin_{lj}^t(\omega) \ge 0 \quad \forall j \in J, l \in L, t \in T, \omega \in \Omega \quad (21)$$

$$pr_j^t(\omega) \ge 0 \quad \forall j \in J, t \in T, \omega \in \Omega \quad (22)$$

$\overline{M}$ is a big positive number.

3. Case Study: waste-based bioethanol production in California

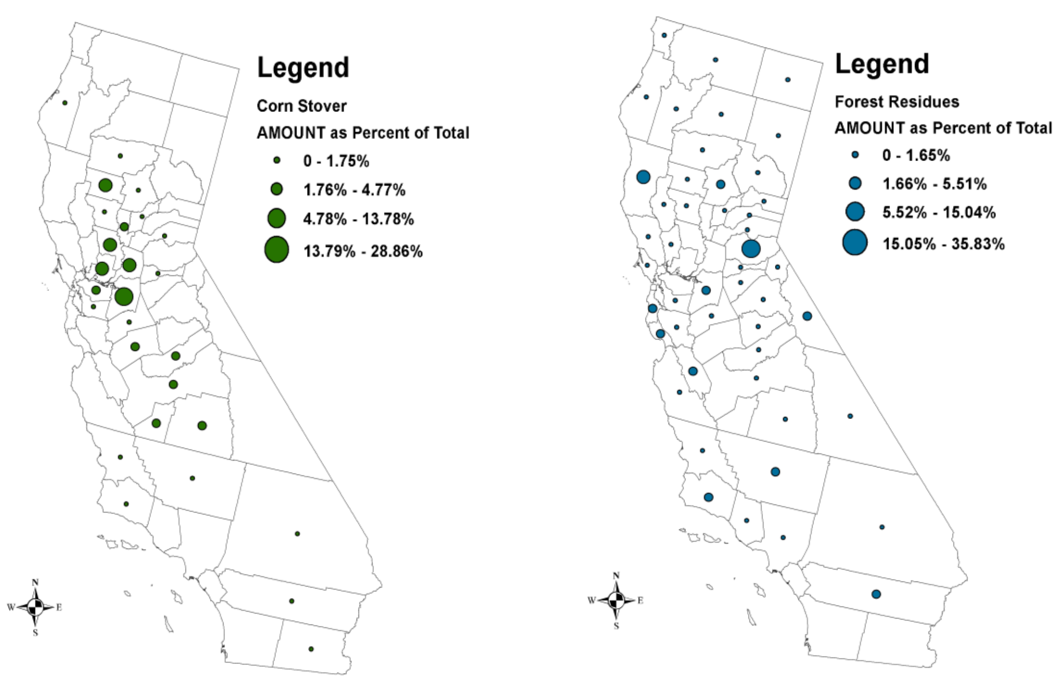
California is of particular interest for biofuel study due to its aggressive environmental policies promoting low-carbon fuels (including AB32, AB1493, Low Carbon Fuel Standard, etc). California has proposed an aggressive goal in the Bioenergy Action Plan targeting instate ethanol production at 20% of the total state's biofuel consumption by 2010, 40% by 2020 and 75% by 2050 (California Bioenergy Interagency Working Group, 2006). Advanced biofuel conversion technologies that use lignocellulosic biomass are anticipative to be ready for commercialization by 2020 (Parker et al., 2007). Ethanol demand by year 2020 is therefore set as the demand target in the study, which is projected to be 350Million Gallons per Year (MGY), given the current ethanol-to-gasoline blend rate E5.7 (Jenkins et al., 2007).

3.1 Model input

3.1.1 Potential Feedstock Resources

Research conducted by Idaho National Laboratory (Bioenergy Program, 2008) has demonstrated near-term feasibility of mixing multiple types of feedstock for biofuel production with single conversion technology through advanced uniform-feedstock preprocessing. California has a diverse set of biowaste feedstock resources (at inexpensive procurement cost) for ethanol production. Two types of biowaste resources – corn stover and forest residues, are considered in this study, both are abundant in California.

The feedstock yields vary significantly across the state. A thorough assessment of feedstock resources has been conducted by Western Governors' Association and the details are available in the report (Parker et al., 2007). The annual feedstock yields and locations are aggregated at county or city levels in geographic information system (GIS). To integrate feedstock resource data with transportation network data, it is assumed that feedstock produced in a county or city is available at the centroid node of that zone. The geographic distribution of corn stover and forest residue is plotted in Figure 2, in which the size of each dot is proportional to the feedstock quantity. Corn stover is mainly clustered in central valley region. Forest residue is widely distributed across the state with higher concentration in the northern part.



(a) corn stover (b) forest residue

Figure 2. Geographical Distributions of Feedstock Availability

The feedstock parameters, including total availability, conversion rates, moisture contents, deterioration factor, average procurement costs, and storage costs, are given in Table 2. The conversion rate is in the unit of gallons of ethanol converted from one dry ton of feedstock. Deterioration rate is the percentage loss in feedstock inventory over one season. In this study, a 10% corn stover loss in storage is assumed, which is within a range of 3.3%-18.1% loss reported in (Shinners et al., 2007). The deterioration rate 12% for forest residue is adopted from (Ashton et al., 2007). The deterioration rates are assumed to be constant all year around and identical across all geographical locations. The feedstock procurement cost is defined as the expense of transporting feedstock from fields to the roadside in a transportable form (Parker et al., 2007). The corn stover storage cost is $8/dry ton, given that they are stored in uncovered or tarped stacks (Sokhansanj et al., 2002). The cost for logger to stack forest residue on site is $2/dry ton (Petrolia, 2006), which is considered as the storage cost in the study.

Table 2. Feedstock Parameters

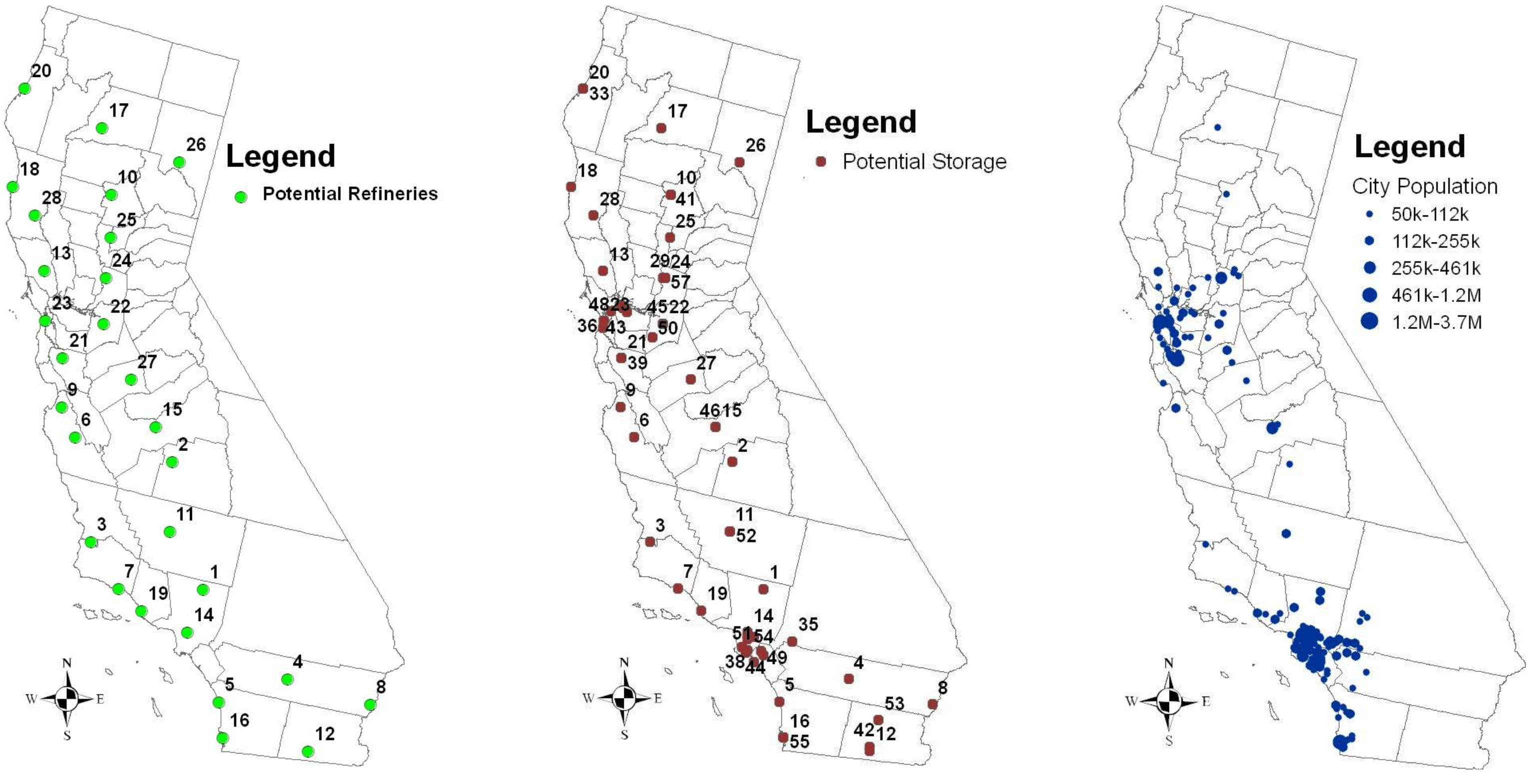
Feedstock types	Availability (thousand dry ton)	# of nodes[1]	Conversion rate	Moisture content (% weight)	Seasonal deterioration rate[2]	Avg procurement cost ($/dry ton)	Storage cost ($/dry ton)
Corn stover	563	27	80.6	15	10%	35	8
Forest residue	4,268	47	90.2	50	12%	30	2

Notes: 1. the total number of locations (centroid nodes) of each feedstock type.
2. All data in this table are adopted from (Parker et al., 2007) except seasonal deterioration rates and storage costs.

The two types of feedstock resources are not available all year around. Corn can only be harvested in fall season. Forest residue has wider harvesting window excluding winter season. The uncertainty of corn stover availability is mainly caused by corn production, whose historical data is reported by the USDA National Agricultural Statistics Service (USDA, 2001). Forestry fires are considered to be the major cause of forest residue yield fluctuations. The annual wildfires information is available from California Department of Forestry and Fire Protection's Fire and Resources Assessment Program (2011). We assumed the forestry loss is proportional to the size of fires. In this study, a set of ten scenarios has been generated for both feedstock types, based on 1999 to 2008 historical data.

3.1.2 Potential refineries, fuel storages, and demand clusters

A total of 28 sites were chosen as the candidate refinery locations (see Figure 3(a)) based on a set of criteria considering the accessibility to water and transportation infrastructures and zoning requirements (Parker et al., 2007). A total of 57 locations, including all potential refinery locations and existing city-gate fuel terminals, are candidates for siting ethanol storage facilities (Figure 3(b)). Cities with a population more than 50,000 are considered as demand centers. Figure 3(c) shows the geographic distribution of 143 demand centers in this study. The total annual ethanol demand from the selected demand centers is set to be 272MGY which is estimated based on the state-wide demand of 350 MGY (Jenkins et al., 2007) proportional to the population. Compared to the feedstock seasonality, the demands are relatively stable, and they vary with only 2% over seasons (U.S. Energy Information Administration, 2011).



(a) Ethanol refinery candidate locations (b) Ethanol storage candidate locations (c) Demand centers

Figure 3. Geographical Distributions of Demand Centers and Candidate Locations for Refineries and Storages

3.1.3 Bioethanol Production

An advanced biofuel conversion technology - LignoCellulosics Ethanol (LCE) via hydrolysis and fermentation conversion technology with specific Dilute Acid pretreatment process, is considered. Different from conventional technologies that consume corn and grain mostly, the new technology uses lignocellulosic biomass. It also features low cellulose enzyme cost and reasonably high ethanol yields. According to (Office of the Biomass Program, 2009), the mid-term projection of bioethanol production cost is $0.92 per gallon, which includes pretreatment, production, and distillation and solid recovery costs.

Refinery costs: The refinery cost includes fixed capital cost (facility setup cost) and variable capital cost (facility size-dependent cost). The fixed capital cost was annualized assuming a real discount rate of 10% and lifetime of 20 years, based on the 2020-year technology performance. The fixed capital cost is $6.157m and the variable capital cost is $0.314 per gallon. The size of the refinery is determined by the model subject to the constraint of maximum refinery capacity as 100MGY (Parker et al., 2007).

Fuel storage costs: Three different sizes of tank are considered for fuel storage: 25, 50, and 100 thousand barrels (1 barrel = 42 gallons). The capital costs associated with the three sizes are $450k, $765k, and $1.26m, respectively. Note that this cost already covers fees incurred in fuels storage and materials consumption so that there is no separate storage operational cost (i.e., $estor = 0$). In addition, receiving product by delivery trucks costs another $10k, and including gasoline blending systems adds additional $300k (Downstream Alternatives Inc., 2000).

Transportation costs: In order to estimate the costs of transporting feedstock and fuels in the entire supply chain system, a GIS-based transportation network was introduced. This network contains local, rural, urban roads and major highways. The shortest distances between feedstock fields, refineries, storages, and demand cities were calculated based on this network. Since only in-state production and delivery are considered, we consider trucking as the only transportation mode, assuming maximum loads of 25 tons for transporting bulk solids and 8,000 gallons for transporting liquid, and an average travel speed at 40 mile/hr. Transportation costs include three components: loading/unloading cost, time dependent travel cost, and distance dependent travel cost as summarized in Table 3. Time dependent cost includes labor and capital cost of trucks, while distance dependent cost includes fuel, insurance, maintenance, and permitting cost. Finally, the truck is fueled by diesel with cost of $2.50 per gallon (Parker et al., 2007).

Table 3. Trucking Cost

	Liquids	Bulk solids
Loading/unloading	$0.02/gallon	$5/wet ton
Time dependent	$32/hr/truckload	$29/hr/truckload
Distance dependent	$1.30/mile/truckload	$1.20/mile/truckload
Truck Capacity	8,000 gallons	25 wet tons

Source: (Parker et al., 2007)

3.2 Results and Sensitivity Analysis

In this section, we present results from the case study described above. The mixed integer SP model was solved by a solution algorithm based on scenario decomposition. The numerical implementation of this algorithm is described in Section 3.3.

3.2.1 Optimal System Results

The least expensive biofuel supply chain system is presented in Figure 4. The penalty cost is set high at $5/gallon to mandate the required level of in-state ethanol production. The breakdown of total system cost is presented in Figure 5 and feedstock supply strategies are summarized in Table 4, respectively.

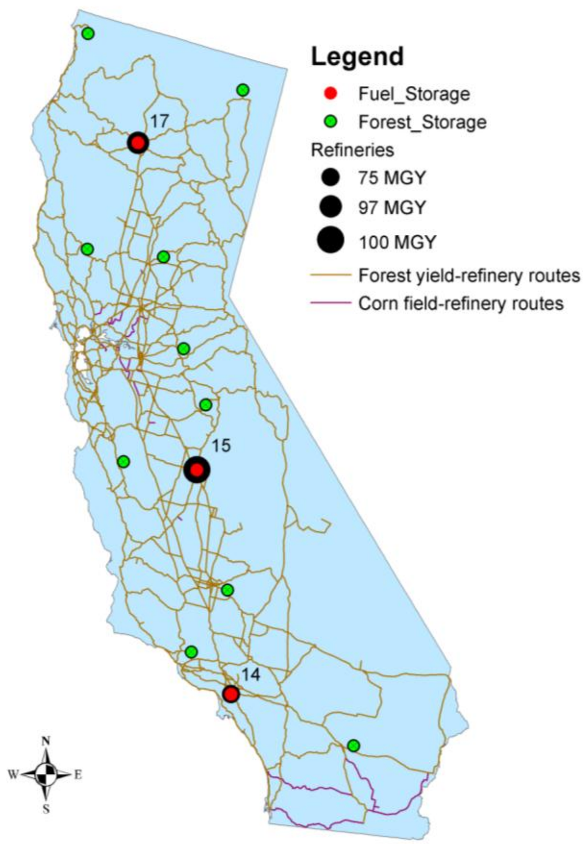


Figure 4. Optimal System Planning

Table 4. Annual Feedstock Supply Strategies (million dry ton)

feedstock	Refinery ID 14	Refinery ID 15	Refinery ID 17	Total feedstock supply
Corn stover	0.01	0.36	0.08	0.45
Forest residue	0.83	0.79	1.00	2.61
Total inflow to refinery	0.84	1.15	1.08	

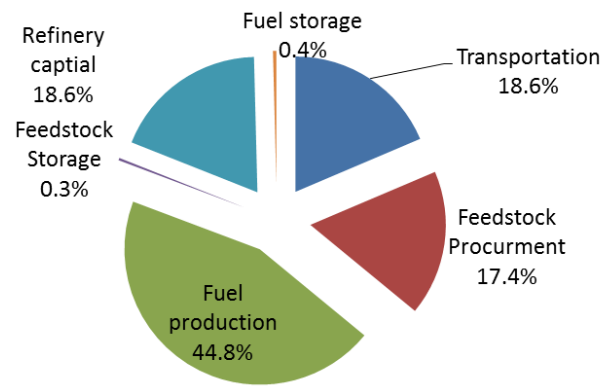


Figure 5. Breakdown of Total System Cost

<u>System planning strategy:</u> Three refineries are selected out of 28 candidate sites. Refinery #17 (with a size of 97MGY) and #14 (with a size of 75MGY) are designated to serve the two major demand clusters, Bay Area and Southern California (mainly concentrated in Los Angeles area), respectively. Refinery #15, centrally located and in proximity to both feedstock types, is built at a capacity of 100MGY to serve both demand clusters. Three small-size fuel storages are placed in co-existence with refineries to save transportation cost. Ten forest residue feedstock storages are included in the system as well, which are widely distributed in the state. The optimal supply chain does not require corn stover storage, because it is less economical (see Table 2) than forest residues.

Feedstock supply strategy: Table 4 presents the amount of feedstock supplied to each refinery. Overall, forest residues play a more important role as feedstock than corn stover; 86.7% of bioethanol is converted from forest residue and 13.3% is from corn stover. In general, the choice of consuming one type of feedstock over the other is a result of tradeoffs among various parameters, including feedstock availability, geographical distribution, feedstock-to-fuel conversion rate, procurement cost, deterioration rate, and moisture content. In this case, corn stover has lower availability and poorer performance than forest residue in terms of conversion rate, procurement cost, and unit storage cost.

Cost breakdowns: The expected cost of delivered bioethanol is $2.05 per gallon, which is highly competitive in comparison with historical California Fuel Ethanol Terminal Market Prices between $1.95 and $2.7 per gallon (California Energy Commission, 2011). In Figure 5, fuel production cost is identified as the major cost driver, accounting for almost half of the delivered cost. Transportation cost takes 18.6%, which suggests the importance of considering the spatiality of the problem and a need for efficient logistics and supply chain planning.

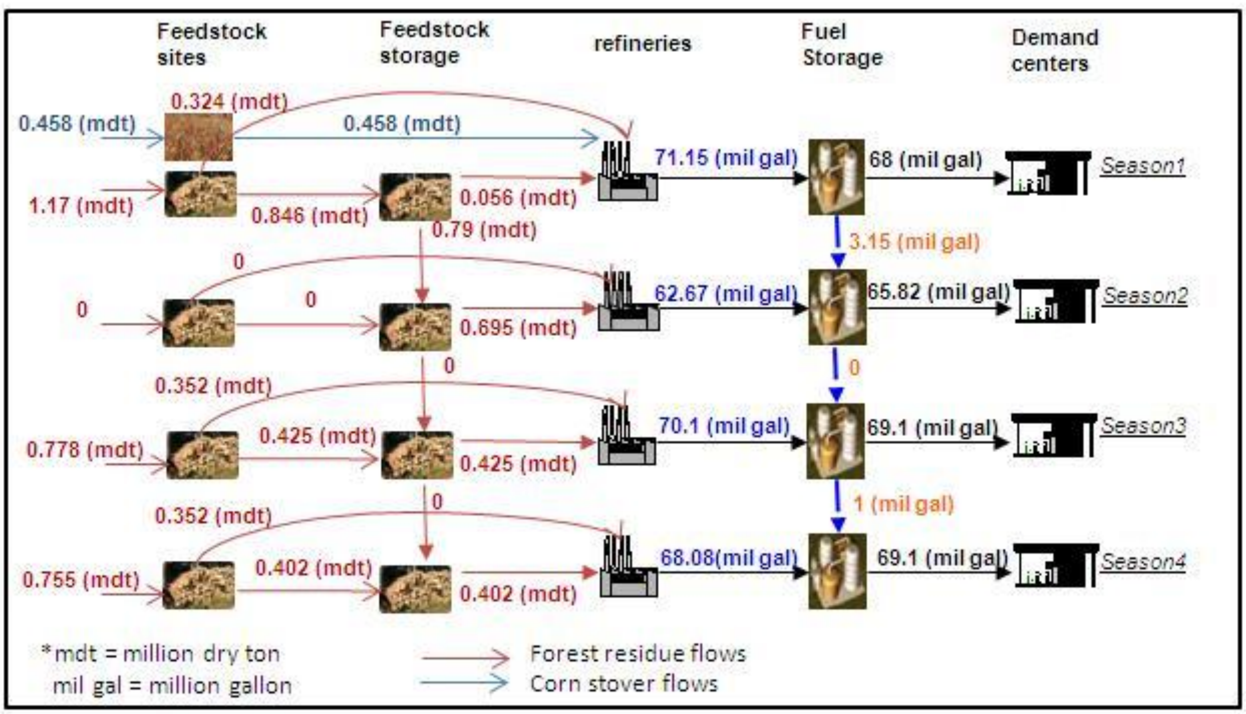


Figure 6. Flows in Supply Chain under One Scenario

Operational strategies: Operational decisions are allowed to have recourse - can be adjusted based on the actual realization of uncertain parameters, thus are scenario dependent. As an example, Figure 6 demonstrates how feedstock and fuel flow between different layers of the supply chain in one scenario. Corn stover is only available in the first season (beginning in September), in which all procured corn stover is transported to refineries directly, avoiding its high storage cost. In season 2, refineries are operated based on stored forest residues.

Role of storage: Even though feedstock and fuel storage accounts for less than 1% of the total system cost, it plays some important roles:

- *storage* function provides "buffer" for the system to adjust supply-demand discrepancy. For example, feedstock storage helps stabilize the refinery workload throughout the year and reduce the requirement of fuel storage, which is more capital intensive than feedstock storage. If no feedstock storage were placed in the system, the total system cost would rise by 2.5% and 16 large-size fuel storages would be needed.
- *redistribution* function over time and space increases system efficiency via consolidation, and improves the self-reorganization ability of the system hedging against potential supply disruption. As an example, in season 3, about half of the procured forest residue is consolidated at two of the storage sites, resulting in lower transportation cost.

Stochastic vs. deterministic solutions: In handling multiple possibilities of random events, a common engineering approach is to aggregate all scenarios to a single scenario by using expected value and then solve the corresponding deterministic problem. We call a solution from this approach the expected-value solution. The expected-value solution suggests a different system layout: three refineries located at 15, 27, and 28 with capacities of 100, 95, and 77MGY, respectively. They are proximate to feedstock fields. Three small-size fuel storage facilities are selected, but at different locations #15, #19 and #28.

Performance evaluation of stochastic and deterministic solutions: In this section, we evaluate the performance of stochastic and deterministic solutions in 100 randomly generated scenarios. For each one of the ten original scenarios, ten scenarios are generated following a normal distribution with three standard deviation (SD) levels at 0.1, 0.2 and 0.25. Note that the scenarios used for model evaluation are intentionally chosen to be slightly different from the model input, reflecting imperfect prediction of random parameter distribution. The comparison results are plotted in Figure 7.

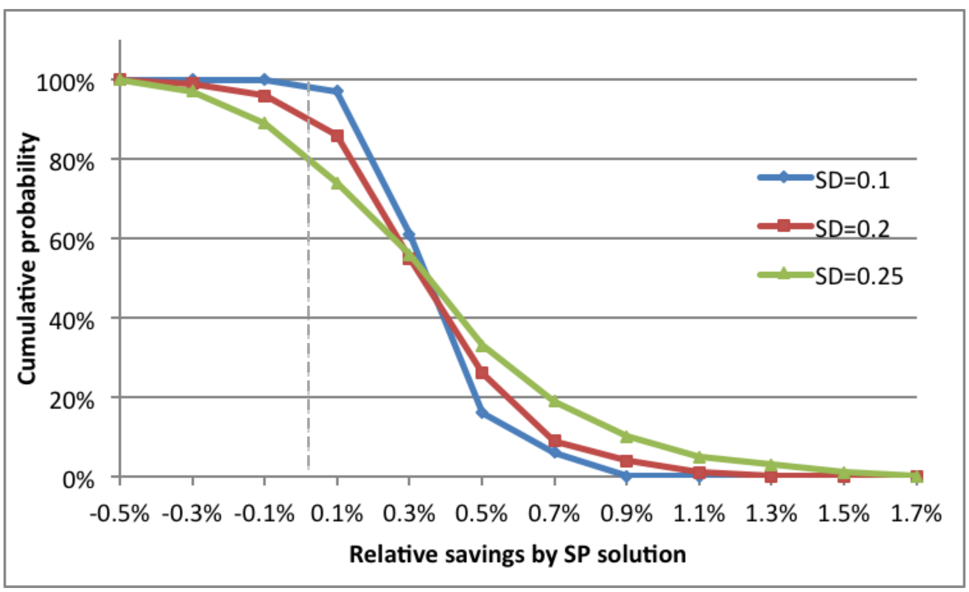


Figure 7. Performance Evaluation of Stochastic and Deterministic Solutions

The horizontal axis in Figure 7 represents the relative saving in total system cost (%) by the stochastic programming (SP) solution compared to the deterministic solution. The vertical axis represents the cumulative probabilities of having a saving large than or equal to a given threshold. For example, the chance of the SP solution over performs the deterministic solution (i.e., relative saving is larger than or equal to 0) across all tested scenarios is about 99%, 90%, and 80% for dataset of SD = 0.1, 0.2, 0.25, respectively. Comparing the distributions of the three lines corresponding to different levels of SD, it is observed that as the quality of input data degrades (reflected by increasing SD), the benefit of SP solution spans over a larger range, meaning the difference made by the two solutions becomes more noticeable. For dataset of SD=0.25, the saving by SP solution is between [-0.5%, 1.7%]. Note that such saving is achieved from adjusting the transportation, storage, and procurement costs which in total only accounts for 36.6% of the total system cost, since the total capital and production cost of refineries stays unchanged between the stochastic and deterministic solutions.

3.2.2 Sensitivity Analysis

Impact of the cost of imported fuels: The cost of imported biofuels is reflected by the penalty cost of demand unsatisfied by in-state production. A sensitivity analysis is conducted to learn how penalty cost impacts the optimal level of in-state production. Figure 8 shows the expected unit cost and the percentage of imported fuel under penalty cost between \$2 and \$2.3 per gallon. The horizontal axis represents the penalty cost. The vertical axis on the left corresponds to the unit cost of delivered fuel, and the one on the right represents the percentage of demand unsatisfied by in-state production. It is observed that the average delivered biofuel cost is stabilized around \$2/gallon when under a penalty cost at \$2/gallon or higher. The bars in the figure show that \$2 and \$2.3 are two critical points for penalty cost setting. When the penalty cost is set at \$2 or below, no in-state production is needed; all demand should be satisfied by imported fuel. As the penalty cost increases, it becomes more economical to satisfy the fuel demand by in-state production. Using geographic information systems (GIS) tool, we observed that most unsatisfied demand occurs in the southern part of the state. This suggests that the biofuel pathways considered in this study provide a better economic potential for the northern part of the state. When the penalty cost is set to be \$2.3 per gallon or high, it is efficient to satisfy demand all by in-state production.

Impact of demand growth: The state-wide demand of 272MGY in the baseline study is based on 5.7 blend rate. For E10 as most other states practice now, the total demand increases to 477MGY. A sensitivity analysis on demand between 272MGY and 477MGY is conducted, in which the penalty cost is set as \$5. The results are plotted in Figure 9. The horizontal axis corresponds to various fuel demands. The two vertical axes are defined the same as in Figure 8. The curve in Figure 9 shows that 350MGY is a critical demand point for the state. As the demand increases beyond 350MGY, the average unit cost of delivered fuel increases significantly. The higher cost is mainly due to increased cost on feedstock procurement and transportation and the penalty of unsatisfied demand. As shown in Figure 9 (bars), the portion of imported fuel increases as the demand grows, although the total amount of corn stover and forest residue in the state is sufficient to produce up to 435MGY ethanol. This suggests that to achieve the overall system efficiency, a certain portion of demand (mainly located in the southern part of the state) may not be served by in-state production.

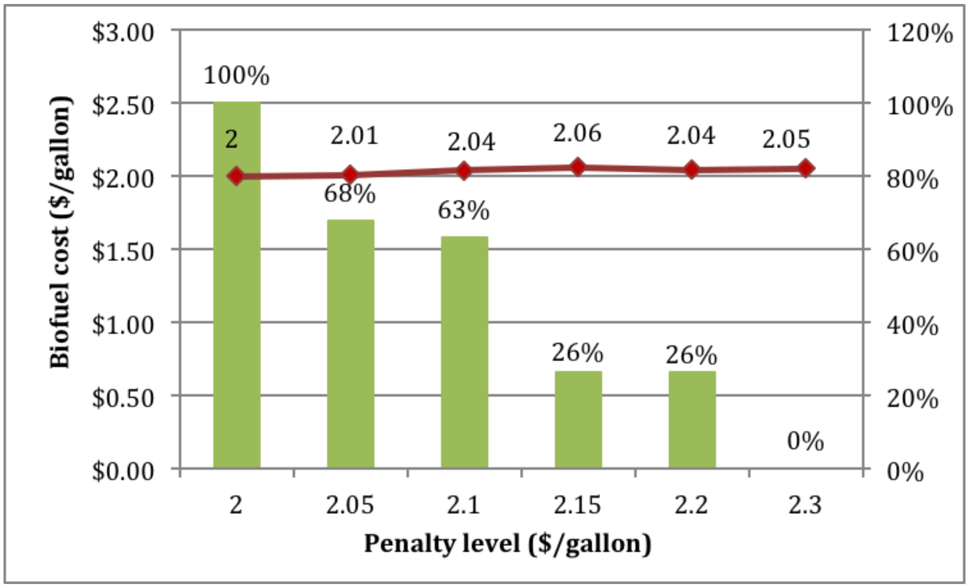


Figure 8. Impact of Penalty Level on Unit Biofuel Cost and Import Quantity

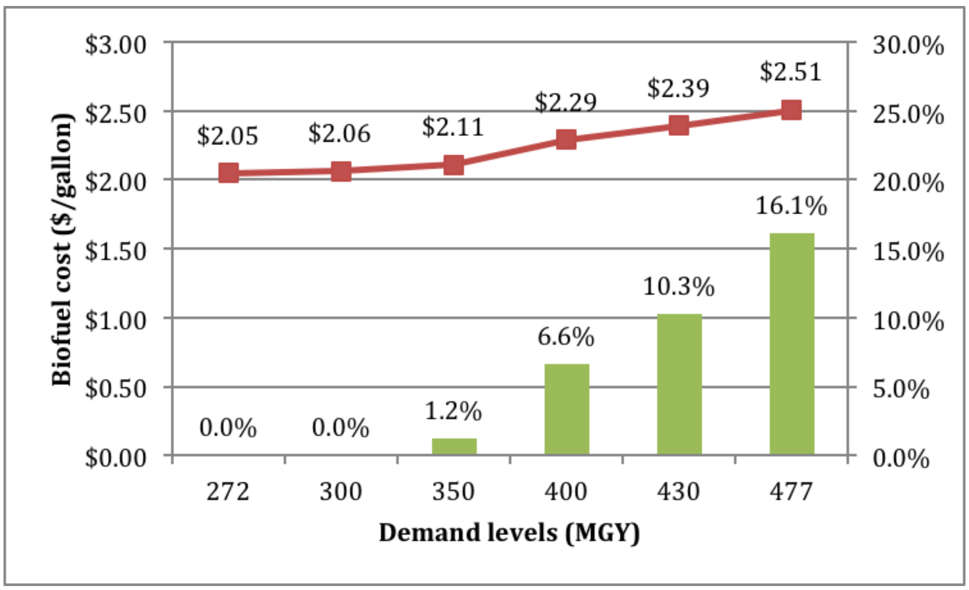


Figure 9. Impact of Demand on Unit Biofuel Cost and Import Quantity

3.3 Numerical Implementation

Although stochastic modeling approaches provide more reliable results, they often come with heavier computational burden for problems of non-trivial size. In some numerical experiments, we were not able to directly use commercial solvers to solve the problem within a reasonable amount of time. Therefore, decomposition methods were exploited to overcome the numerical difficulties.

There are a handful of decomposition methods in handling large-scale stochastic programming problems (Ruszczynski, 1997). Some well-documented and widely implemented methods include L-shaped method (also called vertical decomposition) (Van Slyke and Wets,

1969) and Progressive Hedging (PH) method (also called horizontal decomposition) (Rockafellar and Wets, 1991). Based on our previous research experience and numerical experiments, we found that the PH method is well suitable for solving this problem. PH method decomposes a stochastic problem across scenarios by partitioning the original problem into manageable scenario sub-problems. Recent successful applications of PH method in solving stochastic mix-integer problems can be found in (Chen and Fan, 2012; Fan and Liu, 2010; Watson and Woodruff, 2011).

The performance of PH algorithm is examined by comparing its computing time and the solution quality with those of CPLEX, the most widely used commercial solver for mix-integer linear programming. Figure 10 shows the execution time (CPU minute) of running the stochastic model with different number of scenarios. The “CPLEX” curve represents the running time from directly solving the problem using AMPL-CPLEX. All the experiments described in this paper were run using a Dell workstation with 12 GB RAM and Dual-Xeon 2.40 GHz processor under Windows 7 environment. In Table 5, the objective values obtained by the PH algorithm and AMPL-CPLEX are compared. The relative difference between the two solution methods is reported as the “gap” of numerical accuracy. In all experiments, the gap is relatively small (around 1% - 2%), indicating a good solution quality of the PH algorithm.

As reported by several researchers (Løkketangen and Woodruff, 1996; Mulvey and Vladimirou, 1991; Mulvey and Vladimirou, 1992; Watson and Woodruff, 2011), the convergence of the PH algorithm is largely influenced by the setting of parameter γ (see Appendix for a summary of the PH method). Previous research has suggested that an effective γ value should be close in magnitude to the sensitivity of the objective value with respect to the first-stage decision variable (Watson and Woodruff, 2010). In this study, the three different γ values were set as: γ_{plant} = 6.157 for refinery location variables, γ_{size} = 0.314 for refinery size variables, and γ_{fuel} = 1.57 for fuel storage location variables.

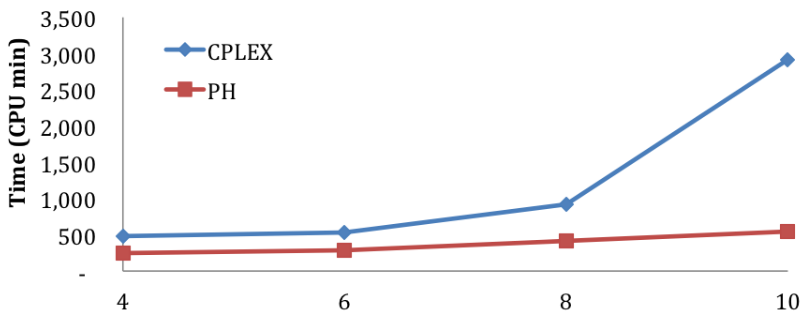


Figure 10. Computing Time by PH and CPLEX

Table 5. Solution Quality by PH and CPLEX

# of scenario	CPLEX ($M)	PH algorithm ($M)	Gap (%)
4	569.75	575.18	0.95
6	543.95	555.36	2.10
8	545.88	555.91	1.84
10	553.07	561.26	1.48

4. Conclusions and Discussions

This research is a new endeavor in biofuel supply chain planning, addressing geographical interdependence, system dynamics, and uncertainty in a single modeling framework. A mixed-integer multistage stochastic programming model that combines strategic and tactical system decision makings is proposed, with a goal of achieving the least total expected system cost. The model has been used to assess the economic potential of converting biowastes (i.e., corn stover and forest residue) to ethanol in California. The overall delivered bioethanol has a competitive cost at about $2.05 per gallon. In bioenergy research community, feedstock seasonality and uncertainty has been a major concern in bioenergy supply system planning. We found that when the entire biofuel supply chain is considered and when diversified feedstock resources are used, the system is able to mitigate the risk brought by feedstock variation. Feedstock and fuel storage, though accounting for an insignificant portion of the total cost, operated together with the rest of the supply chain, may provide critical storage and redistribution functions that help dissipate the supply-demand discrepancies. The real world case study demonstrates that the presented modeling and computing approach are suitable for biofuel infrastructure system planning under uncertainty.

Several directions for future extension may be pursued. From computing perspective, each scenario sub-problem resulted from the PH algorithm is still a large-scale mix-integer problem. Integration of effective mix-integer solver could largely improve the overall computational efficiency. How to organize evolving information describing random parameters presents another research challenge for large-scale multistage stochastic programming problems. One possible direction is to explore the concept of dynamic scenario tree generation, which does not require scenario data to be generated all at once thus reducing storage burden. From modeling perspective, non-recurrent disruptions (natural disasters and/or human made attacks) featuring low probability but severe consequences on energy supply chain, should be incorporated. However, due to different nature of non-recurrent disruptions, different modeling framework including risk quantification might be more suitable. This direction calls for an integration of system resilience research with bioenergy supply chain planning. It is one of our ongoing research topics.

Acknowledgements

The authors would like to deliver our sincere gratitude to researchers at Sustainable Transportation Energy Pathways (STEPS) program of University of California, Davis for their generous data support and helpful technical discussions. The research was supported by a grant from the Sustainable Transportation Center at the University of California Davis, which receives funding from the U.S. Department of Transportation and Caltrans, the California Department of Transportation, through the University Transportation Centers program. The contents of this paper reflect the views of the authors, who are responsible for the facts and the accuracy of the information presented herein.

References

Ashton, S., Jackson, B., Schroeder, R., 2007. Storing Woody Biomass, Fact Sheet 4.6. Southern Forest Research Partnership, Inc., Athens, GA, pp. 149-152.

Bioenergy Program, 2008. Uniform-Format Feedstock Supply System Design for Lignocellulosic Biomass. Idaho National Laboratory.

Birge, J.R., Louveaux, F., 1997. *Introduction to Stochastic Programming*, 1st ed. Springer, New York.

California Bioenergy Interagency Working Group, 2006. Bioenergy Action Plan for California.

California Department of Forestry and Fire Protection, 2011. FRAP Data.

California Energy Commission, 2011. Fuel Ethanol Terminal Market Price - 10 Year History, http://energyalmanac.ca.gov/transportation/ethanol_graphs/ethanol_10-year.html.

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.

Cordeau, J.-F., Pasin, F., Solomon, M., 2006. An integrated model for logistics network design. *Annals of Operations Research* 144(1), 59-82.

Cundiff, J.S., Dias, N., Sherali, H.D., 1997. A linear programming approach for designing a herbaceous biomass delivery system. *Bioresource Technology* 59(1), 47-55.

Dal-Mas, M., Giarola, S., Zamboni, A., Bezzo, F., Strategic design and investment capacity planning of the ethanol supply chain under price uncertainty, *Biomass & Bioenergy* 35(5), 2059-2071.

De La Torre, U., G., D., Ray, D.E., 2000. Biomass and bioenergy applications of the POLYSYS modeling framework. *Biomass and Bioenergy* 18(4), 291-308.

Delucchi, M., 2006. Lifecycle analyses of biofuels: Draft manuscript. Institute of Transportaion Studies, University of California, Davis, Davis, CA.

Downstream Alternatives Inc., 2000. The Current Fuel Ethanol Industry Transportation, Marketing, Distribution, and Technical Considerations. Oak Ridge National Laboratory Ethanol Project.

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.

European Parlinament and Council, 2003. Directive 2003/30/EC on the promotion and use of biofuels or other renewable fuels for transport, Brussels.

Fan, Y., Liu, C., 2010. Solving Stochastic Transportation Network Protection Problems Using the Progressive Hedging-based Method. *Networks and Spatial Economics* 10(2), 193-208.

Farrell, A.E., Sperling, D., 2007. A low-carbon fuel standard for california, part 1: Technical analysis. Institute of Transportation Studies, University of California, Davis, Davis, CA.

Freppaz, D., Minciardi, R., Robba, M., Rovatti, M., Sacile, R., Taramasso, A., 2004. Optimizing forest biomass exploitation for energy supply at a regional level. *Biomass and Bioenergy* 26(1), 15-25.

International Energy Agency, 2006. World Energy Outlook 2006.

Jenkins, B., Dempster, P., Gildart, M., Kaffka, S., 2007. California Biomass and Biofuels Production Potential (Draft). California Energy Commission.

Kim, S., Dale, B.E., 2005. Life cycle assessment of various cropping systems utilized for producing biofuels: Bioethanol and biodiesel. *Biomass and Bioenergy* 29(6), 426-439.

Løkketangen, A., Woodruff, D.L., 1996. Progressive Hedging and Tabu Search Applied to Mixed Integer (0,1) Multi-stage Stochastic Programming. *Journal of Heuristics* 2, 111-128.

Louveaux, F.V., 1986. Discrete stochastic location models *Annals of Operations Research* 6, 21-34.

McCarl, B.A., Schneider, U.A., 2001. CLIMATE CHANGE: Greenhouse Gas Mitigation in U.S. Agriculture and Forestry. *Science* 294(5551), 2481-2482.

Melo, T., Nickel, S., Gama, F.S.d., 2007. Facility Location and Supply Chain Management – A comprehensive review. *Institut Techno-und Wirtschaftsmathematik*.

Mulvey, J., Vladimirou, H., 1991. Applying the progressive hedging algorithm to stochastic generalized networks. *Annals of Operations Research* 31(1), 399-424.

Mulvey, J.M., Vladimirou, H., 1992. Stochastic network programming for financial planning problems. *Management Science* 38(11), 1642-1664.

Office of the Biomass Program, 2009. Biomass Multi-Year Program Plan, Enery Efficiency and Renewable Energy, U.S. Department of Energy.

Parker, N., Tittmann, P., Hart, Q., Lay, M., Cunningham, J., Jenkins, B., 2007. Strategic Development of Bioenergy in the Western States Development of Supply Scenarios Linked to Policy Recommendations, Task 3: Spatial Analysis and Supply Curve Development. Western Governors' Association.

Parker, N.C., Ogden, J.M., Fan, Y., 2008. The role of biomass in California's hydrogen economy. *Energy Policy* 36(10), 3925-3939.

Perlack, R.D., Wright, L.L., Turhollow, A.F., Graham, R.L., Stokes, B.J., Erbach, D.C., 2005. Biomass as Feedstock for a Bioenergy and Bioproducts Industry: The Technical Feasibility of a Billion-Ton Annual Supply. U.S. Department of Agriculture and Oak Ridge National Laboratory, Oak Ridge, TN.

Petrolia, D.R., 2006. Ethanol from Biomass: Economic and Environmental Potential of Converting Corn Stover and Hardwood Forest Residue in Minnesota, *American Agricultural Economics Association (New Name 2008: Agricultural and Applied Economics Association)*, Long Beach, CA.

Rockafellar, R.T., Wets, R.J.-B., 1991. Scenarios and Policy Aggregation in Optimization Under Uncertainty. *Mathematics of Operations Research* 16(1), 119-147.

Ruszczynski, A., 1997. Decomposition methods in stochastic programming. *Mathematical Programming* 79, 333-353.

Shinners, K.J., Binversie, B.N., Muck, R.E., Weimer, P.J., 2007. Comparison of wet and dry corn stover harvest and storage. *Biomass and Bioenergy* 31(4), 211-221.

Snyder, L.V., 2006. Facility location under uncertainty: a review. *IIE Transactions* 38(7), 547-554.

Sokhansanj, S., Turhollow, A., Perlack, R., 2002. Stochastic Modeling of Costs of Corn Stover Costs Delivered to an Intermediate Storage Facility, *ASAE Annual International Meeting/CIGR XVth World Congree*, Chicago, Illinois, USA.

Tang, C.S., 2006. Perspectives in supply chain risk management. *International Journal of Production Economics* 103(2), 451-488.

Tembo, G., Epplin, F.M., Huhnke, R.L., 2003. Integrative Investment Appraisal of a Lignocellulosic Biomass-to-Ethanol Industry. *Journal of Agricultural and Resource Economics* 28(3), 611-633.

Turner, B.T., Plevin, R.J., 2007. Creating markets for green biofuels: Measuring and improving environmental performance. Institute of Transportation Studies, University of California, Berkeley, Berkeley, CA.

U.S. Congress, 2007. Energy Policy and Security Act of 2007.

U.S. Energy Information Administration, 2011. U.S. Weekly Product Supply.

Unnasch, S., Pont, J., 2007. Fuel cycle assessment: Well-to-tank energy inputs, emissions and water impacts - draft consultant report. California Energy Commission, Sacramento, CA.

USDA, 2001. County Agricultural Commissioners' Data.

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

Watson, J.-P., Woodruff, D., 2011. Progressive hedging innovations for a class of stochastic mixed-integer resource allocation problems. *Computational Management Science* 8(4), 355-370.

Zah, R., Boni, H., Gauch, M., Hischier, R., Lehmann, M., Wager, P., 2007. Life cycle assessment f energy products: Environmental assessment of biofuels. Empa, Technology and Soceity lab, St. Gallen, Switzerland.

Appendix: A summary of the progressive hedging (PH) method

For the convenience of the readers, the basic scheme of PH method is briefly described below by using a generic mathematical model described in (1) and (2):

$$\text{Minimize } (c \cdot x) + \sum_{s \in S} P_s (t_s \cdot y_s) \qquad (1)$$

$$\text{Subject to: } (x, y_s) \in F_s \qquad \forall s \in S \qquad (2)$$

where S is the set of possible scenarios for random event, s ($s \in S$) denotes an individual scenario, x denotes the first-stage decisions with a cost coefficient vectors c, and y_s represents the second-stage decisions with associated cost coefficient vectors t_s. For each scenario $s \in S$, we denote the probability of the occurrence as P_s. The objective is to minimize the total cost of the first and second stages as described in (1). The decisions are subject to the constraints defined by the feasibility set F_s for each scenario s as described in Constraint (2).

The model defined by equation (1) and (2) can be simply separated into many scenario sub-problems. Solving the scenario sub-problems defined in all s ($s \in S$) will give us different s-dependent first-stage solutions, denoted as x_s for each $s \in S$. However, these solutions cannot be directly implemented, because at the time when the location decision solutions are implemented, one does not know yet which scenario is going to happen. In order to consolidate the s-dependent solutions to an *implementable* solution, we must impose the following condition:

$$x_s = x'_s \qquad \forall s \in S, \forall s' \in S, s' \neq s, \qquad (3)$$

or equivalently

$$x_s - z = 0 \qquad \forall s \in S \qquad (4)$$

where z is a vector of free variables. This condition is called a *non-anticipativity* constraint defined by Rockafellar and Wets (1991) which states that a reasonable policy should not require different actions relative to different scenarios if the scenarios are not distinguishable at the time when the actions are taken. Therefore, the overall stochastic program can be formulated as:

$$\text{Minimize } \sum_{s \in S} P_s Q_s(x_s, y_s) \qquad (5)$$

$$\text{Subject to: } (x_s, y_s) \in F_s \qquad \forall s \in S \qquad (6)$$

$$x_s - z = 0 \qquad \forall s \in S$$

Function $Q_s(x_s, y_s)$ is the total first- and second-stage cost in a given scenario s, which depends on the decisions x_s and y_s.

The PH method decomposes a stochastic problem across scenarios and partitions the problem into manageable sub-problems. Define

$$L_r(X, Y, z, W) = \sum_{s \in S} P_s Q_s(x_s, y_s) + (w_s)' \cdot (x_s - z) + \frac{1}{2}\gamma \|x_s - z\|^2 \qquad (7)$$

as the augmented Lagrangian, where W is the vector of dual variables for the constraints in (4), and $\gamma > 0$ is a penalty parameter associated with violation of the non-anticipativity constraints. Therefore, the augmented Lagrangian integrates the non-anticipativity constraints with the original objective function. The stochastic problem becomes

$$\text{Minimize } L_r(X, Y, z, W) \text{ over all } (x_s, y_s) \in F_s. \qquad (8)$$

Due to the nonseparable penalty term $\frac{1}{2}\gamma\left\|x_s - z\right\|^2$ in expression (7), the problem cannot be decomposed directly. The PH method achieves decomposition by alternately fixing the scenario solutions (x, y) and the implementable solution z in (7). The detailed procedures are described below.

PH algorithm procedure

Step 1

Set the iteration index $k = 0$.

Solve for each scenario sub-problem and then obtain $(x_s^{(0)}, y_s^{(0)}), \forall s \in S$.

Initialize $z^{(0)} := \sum_{s\in S} P_s x_s^{(0)}$ and $w_s^{(0)} := \gamma(x_s^{(0)} - z^{(0)})$

If $x_s^{(0)} = z^{(0)}, \forall s \in S$ then the optimal solution is found; otherwise continue with *step 2*.

Step 2

$k = k+1$

Solve for each scenario $\forall s \in S$

$x_s^{(k)} := \operatorname{argmin}_x (Q(x_s, y_s) + w_s^{k-1} x_s + \frac{\gamma}{2}\left\|x_s - z^{k-1}\right\|^2) : (x_s, y_s) \in F_s$

Update $z^{(k)} := \sum_{s\in S} P_s x_s^{(k)}$ and $w_s^{(k)} := w_s^{(k-1)} + \gamma(x_s^{(k)} - z^{(k)}), \forall s \in S$

Step 3

Check whether the termination criterion $\varepsilon = [\left\|z^{(k)} - z^{(k-1)}\right\|^2 + \sum_{s\in S} P_s \left\|x^{(k)} - z^{(k)}\right\|^2]^{1/2} \approx 0$ is reached; if yes, an optimal solution is found, otherwise, go to *step 2* and continue the iterations.