A Quantile-Based Scenario Analysis Approach to Biomass Supply Chain Optimization under Uncertainty

David S. Zamar[a], Bhushan Gopaluni[a], Shahab Sokhansanj[a,b], Nathaniel K. Newlands[c]

[a] *Department of Chemical and Biological Engineering, University of British Columbia, Vancouver, BC V6T 1Z3, Canada (e-mail: zamar.david@gmail.com, bhushan.gopaluni@ubc.ca, shahabs@chbe.ubc.ca)*

[b] *Resource and Engineering Systems Group, Environmental Sciences Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831-6422*

[c] *Science and Technology Branch, Agriculture and Agri-Food Canada, Lethbridge Research Centre, Lethbridge, AB, Canada (e-mail: nathaniel.newlands@agr.gc.ca)*

Abstract

Supply chain optimization for biomass-based power plants is an important research area due to greater emphasis on renewable power energy sources. Biomass supply chain design and operational planning models are often formulated and studied using deterministic mathematical models. While these models are beneficial for making decisions, their applicability to real world problems may be limited because they do not capture all the complexities in the supply chain, including uncertainties in the parameters. This paper develops a statistically robust quantile-based approach for stochastic optimization under uncertainty, which builds upon scenario analysis. We apply and evaluate the performance of our approach to address the problem of analyzing competing biomass supply chains subject to stochastic demand and supply.

Keywords: renewable energy systems, biomass, optimization, uncertainty, scenario analysis, stochastic approximation, robust estimation

1. Introduction

Presently, fossil fuels such as oil, coal and natural gas are the prime energy sources of the world. However, it is anticipated that these sources of energy will be depleted within the next 50 to 100 years [8, 9, 17]. The world is consuming more fossil fuel energy then is being discovered and the reserves of energy that can be cheaply mined have reached peak production [8, 9, 17]. Moreover, the expected environmental damages, such as global warming, acid rain and urban smog due to the production of emissions from the combustion of fossil fuels have compelled the world to reduce carbon emissions and shift towards utilizing sustainable and renewable energy sources [17]. Biomass has been recognized as a promising alternative energy source, since it is both renewable and CO_2 neutral [12]. However, renewable energy production from biomass faces many challenges due to uncertainty of its demand and continuous supply [6, 22]. The purpose of this study, is to develop and apply a stochastic optimization model to analyze the impact of inter-power plant competition for the available feedstock on cost structures.

Conventional optimization based approaches for the design of biomass supply chain networks assume that the operational characteristics, and hence the design parameters are deterministic [1, 11]. However, critical parameters, such as customer demand, prices, and resource capacities are usually uncertain. Sources of uncertainty in decision outcomes are typically due to three different conditions: (1) lack of knowledge, such as the quality characteristics of available biomass feedstocks; (2) noise, such as measurement errors or incomplete data; and (3) events that have not yet occurred, such as future energy demand or feedstock supply shortages [13]. Uncertainty creates a range of concerns regarding the volatility of one decision versus another.

Common approaches to solving constrained optimization problems under uncertainty assign a probability distribution to the unknown model parameters and maximize the expected value of an objective function subject to constraints, which are satisfied on average or with high probability [16, 21, 5]. The expected value of a random variable, X, written $E(X)$, is equal to the weighted average of the outcomes of the random variable, where the weights are given by the probabilities of those outcomes. Use of the expected value for selecting the best solution to a stochastic optimization problem explicitly assumes that the decision maker is primarily interested in the average behavior of the performance metric and is not concerned with other features of its distribution, such as quantiles or variance [2].

Dembo [5], proposed an extension of stochastic optimization, called scenario analysis (SA), in which uncertainties are modeled using discrete probability scenarios that approximate the true probability distributions. The realism of the scenarios chosen and the accurate estimation of the corresponding probabilities enhance the model performance. However, the SA approach generally leads to improved performance of the solution even if there are errors in the estimations of both the actual scenario outcomes and their probabilities [3, 24].

In this paper, we describe a modified scenario analysis approach that is applicable to a wide range of constrained optimization problems and is easily implementable. A preliminary variant of this methodology was motivated by a renewable energy research problem on biomass supply chain optimization [23]. Rather than optimizing the average performance of a solution, we optimize its corresponding quantiles across a set of sampled scenarios. The solutions obtained are statistically robust because they are not easily influenced by extreme scenarios that have a low probability of occurring. An advantage of our approach is that uncertainties do not need to be expressed in terms of a finite number of scenarios declared in advance. Instead, scenarios are sampled according to their estimated probability distribution. Another feature of our approach is the ability to handle large complex stochastic problems by compiling them into a set of similar deterministic constraint problems, which can be efficiently solved in parallel. This compilation allows us to use existing constraint solvers without any modifications, as well as call upon the power of hybrid solvers for non-conventional optimization problems.

The remainder of this paper if organized as follows. In Section 2, we provide a brief discussion of scenario analysis and potential advantages over the more familiar mean-based stochastic optimization approach. In Section 3, we present our quantile-based scenario analysis (QSA) approach for stochastic optimization. In Section 4, the motivating problem of this paper is formulated as a QSA problem. In Section 5, results are presented and compared to those obtained using standard alternative optimization techniques. Finally, Section 5 provides a summary and concluding remarks.

2. Constrained Optimization under Uncertainty

This section provides a brief introduction to stochastic optimization and scenario analysis. For more details, the reader may consult the work of Birge and Louveaux in [4] and references therein.

2.1. Stochastic Optimization

Most systems that need to be controlled or analyzed usually involve some level of uncertainty about the values to assign to the model parameters, if not about the model structure itself. In many situations not so much is lost by selecting a suitable model to represent the system and assigning reasonable values to the model parameters. In other circumstances, ignoring uncertainty may invalidate implications one may wish to draw from the analysis. Uncertainty can be incorporated into the problem in many ways. One that has been used successfully in a wide variety of situations is to model the uncertain quantities as random variables and assign them appropriate probability distributions. In the case of decision making under uncertainty, this leads us to stochastic optimization. Given a probability space $(\Omega,\, p)$, where Ω is the set of possible realizations of the random vector $\boldsymbol{\omega}$ and p is the associated probability density function, the stochastic optimization problem may be expressed in the standard inequality form of Equation (1).

$$\begin{aligned} \underset{\boldsymbol{x}}{\text{maximize}} \quad & E\{J(\boldsymbol{x},\boldsymbol{\omega})\} = \int_{\Omega} J(\boldsymbol{x},\boldsymbol{\omega})p(\boldsymbol{\omega})d\boldsymbol{\omega} \\ \text{subject to} \quad & E\{C_k(\boldsymbol{x},\boldsymbol{\omega})\} \leq 0,\ k=1,2,\ldots,m \end{aligned} \tag{1}$$

where the objective function, $J(\boldsymbol{x},\boldsymbol{\omega})$, and constraint functions, $C_k(\boldsymbol{x},\boldsymbol{\omega})$, depend on the optimization variables $\boldsymbol{x}$ and the values of the random parameters $\boldsymbol{\omega} \in \Omega$. Thus, $J(\boldsymbol{x},\boldsymbol{\omega})$ and $C_k(\boldsymbol{x},\boldsymbol{\omega})$, $k = 1, 2, \ldots, m$, are explicit functions of the optimization variables, for a given realization of the random parameters $\boldsymbol{\omega}$ with probability function $p(\boldsymbol{\omega})$. Here, m represents the total number of constraints, some of which may be deterministic and thus do not depend on $\boldsymbol{\omega}$. The goal is to choose $\boldsymbol{x}$, such that the constraints are satisfied on average and the objective is large on average over the probability space Ω. A common stochastic optimization approach, known as the sample average approximation (SAA) [18], is to pick sufficiently many independent samples from Ω and use them to approximate $p(\boldsymbol{\omega})$ and correspondingly the expected value functions shown in Equation (1). The SAA approach provides a solution that is not optimal for any given scenario, but optimal on average. Thus, it does not provide any assurance on how well the solution will perform (in terms of maximizing the objective function and satisfying the constraints) for different realizations of the random parameters $\boldsymbol{\omega}$. The scenario analysis approach described below addresses this issue.

2.2. Scenario Analysis

As described in the seminal paper by Dembo [5], the multi-scenario stochastic optimization problem is given by:

$$\begin{aligned} \underset{\boldsymbol{x}}{\text{maximize}} \quad & J(\boldsymbol{x}, \boldsymbol{\omega}) \text{ (uncertain objective)} \\ \text{subject to} \quad & C_u(\boldsymbol{x}, \boldsymbol{\omega}) \leq \mathbf{o}, \text{ (uncertain constraints)} \\ & C_d(\boldsymbol{x}) \leq \mathbf{o} \text{ (deterministic constraints)}, \end{aligned} \tag{2}$$

where a scenario is defined as a particular realization, $\boldsymbol{\omega_i}$, of an uncertain scenario $\boldsymbol{\omega} \in \Omega$, where $i \in I$, and I is an index set whose elements label (or index) specific members of the set Ω. Note that for each realized scenario, $\boldsymbol{\omega_i} \in \Omega$, the above problem reduces to a deterministic subproblem. It is evident that the solution of a single scenario poses no difficulty. On the other hand, solving each subproblem does not provide a definite way of determining what a reasonable solution to the original stochastic problem should be.

A fundamental issue in scenario analysis is how to combine the solutions from different scenarios to form a single reasonable solution to the underlying stochastic problem. As described by Dembo [5], a common coordination model that combines the scenario solutions into a single feasible solution is given by:

$$\begin{aligned} \hat{x} = \underset{\boldsymbol{x}}{\text{argmax}} \quad & \sum_{i \in I} ||J(\boldsymbol{x}, \boldsymbol{\omega_i}) - v_i||^2 p(\boldsymbol{\omega_i}) + \sum_{i \in I} ||C_u(\boldsymbol{x}, \boldsymbol{\omega_i}) - \boldsymbol{c_i}||^2 p(\boldsymbol{\omega_i}) \\ \text{subject to} \quad & C_d(\boldsymbol{x}) \leq \mathbf{o}, \end{aligned} \tag{3}$$

where $v_i = J(\boldsymbol{x_i}, \boldsymbol{\omega_i})$, $\boldsymbol{c_i} = C_u(\boldsymbol{x_i}, \boldsymbol{\omega_i})$ and $\boldsymbol{x_i}$ is the optimal solution of scenario $\boldsymbol{\omega_i}$. The coordinating model incorporates the random constraints into the objective function as a penalty. Associated with each scenario, $\boldsymbol{\omega_i}$, is a probability $p(\boldsymbol{\omega_i})$. This model attempts to track the scenario solutions as closely as possible while still maintaining feasibility. We may think of the solution, $\hat{\boldsymbol{x}}$, as a centroid of the various scenario-solutions, or simply as the average solution. In short, the scenario optimization approach to stochastic optimization proceeds in two stages:

(a) Compute a solution to the deterministic problem for each scenario.

(b) Solve a coordinating model to find a single feasible policy.

The problem referred to in stage (a) could be a linear, nonlinear or mixed-integer programming problem [5]. Alternatively, it could consist of a system of equations with stochastic coefficients or be any function dependent on stochastic parameters [5]. Nevertheless, for any assumed scenario, this problem is deterministic and can be solved using known algorithms [5].

The scenario analysis approach assumes that the number of scenarios is finite and that their corresponding probabilities are known. However, in practice this may not be the case, because the scenarios may come from a continuous distribution or their distribution (discrete or continuous) may not be known. Furthermore, although the scenario analysis approach attempts to find a solution that performs well across the scenarios, a more accurate and comprehensive comparison of the performance distribution associated with each candidate solution is desirable.

3. QSA Algorithm Description and Mathematical Formulation

3.1. QSA Algorithm

Our approach examines the performance of a solution $\boldsymbol{x_i}$, for a given scenario, $\boldsymbol{\omega_i} \in \Omega$, across other scenarios in Ω. Specifically, we evaluate the performance of $\boldsymbol{x_i}$ at a scenario $\boldsymbol{\omega_j}$, $i, j \in I$, based on two criteria: (1) its objective function value $J(\boldsymbol{x_i}, \boldsymbol{\omega_j})$, and (2) its constraint performance functions defined as

$$\psi_k(\boldsymbol{x_i}, \boldsymbol{\omega_j}) = C_k(\boldsymbol{x_i}, \boldsymbol{\omega_j})^+,\ k = 1, 2, \ldots, m. \tag{4}$$

The exponent $(+)$ indicates that only positive values are considered. The constraint performance functions express the magnitude of the deficiency of each unsatisfied constraint.

Our approach uses the concept of quantile of a distribution. A quantile, q_α of a distribution is a value which divides the distribution such that there is a given proportion, α of values below the quantile. For example, the median is the 0.5 quantile representing the central value of the distribution, such that half the values are less than or equal to it. The proposed approach is as follows. For each scenario solution $\boldsymbol{x_i}$, $i \in I$, we define the random variable $J(\boldsymbol{x_i}, \boldsymbol{\omega})$ with a cumulative distribution function (CDF) denoted by $F_{\boldsymbol{x_i}}$, and the random variables $\psi_k(\boldsymbol{x_i}, \boldsymbol{\omega})$, with CDFs denoted by $G_{\boldsymbol{x_i}k}$, $k = 1, 2, \ldots, m$. The CDFs $F_{\boldsymbol{x_i}}$ and $G_{\boldsymbol{x_i}k}$ convey the overall objective and constraint performance of the solution $\boldsymbol{x_i}$ across scenarios, respectively. We consider coordination models of the following form

$$\begin{aligned} \underset{i \in I}{\text{maximize}} \quad & f(F^{-1}_{\boldsymbol{x_i}}(t)), \\ \text{subject to} \quad & g_k(G^{-1}_{\boldsymbol{x_i}k}(t)) \leq 0,\ k = 1, 2, \ldots, m \end{aligned} \tag{5}$$

that are based on different functionals of the quantile functions $F^{-1}_{\boldsymbol{x_i}}$ and $G^{-1}_{\boldsymbol{x_i}k}$. For example, if we are interested in maximizing a weighted average of the quantiles of the objective and constraint distributions, then we may wish to solve the following coordination model

$$\begin{aligned} \underset{i \in I}{\text{maximize}} \quad & \int_0^1 F^{-1}_{\boldsymbol{x_i}}(t) \cdot \phi_0(t) dt \\ \text{subject to} \quad & \int_0^1 G^{-1}_{\boldsymbol{x_i}k}(t) \cdot \phi_k(t) dt \leq \gamma_k,\ \ k = 1, 2, \ldots, m, \end{aligned} \tag{6}$$

where the functions $\phi_0(t)$ and $\phi_1(t)$ are weighting functions, which must be positive and integrate to unity over the range 0 to 1 and γ_k are appropriate tolerances. This formulation of the problem attempts to maximize a weighted average of the quantiles of the objective function, subject to satisfying a weighted average of the quantiles of the constraint performance functions, also called a risk spectrum [19]. In practice, the CDFs

$F_{\boldsymbol{x_i}}$ and $G_{\boldsymbol{x_i}k}$ are unknown and may be estimated using Monte Carlo simulation. Assuming that we are able to sample from the probability space, (Ω, P), of scenarios, then $F_{\boldsymbol{x_i}}$ and $G_{\boldsymbol{x_i}k}$ may be estimated using the algorithm outlined in Algorithm 1. A visual representation of the procedure used to obtain the distribution of the objective function, $F_{\boldsymbol{x_i}}$, is shown in Figure 1. A similar procedure is applied to obtain the distribution of the constraints, $G_{\boldsymbol{x_i}k}$.

Once the deterministic solutions to the sampled scenario problems have been calculated, the QSA method can be used to obtain solutions with distinct probabilistic characteristics without having to perform any additional optimization.

Algorithm 1: QSA Algorithm

for *i in 1 to n* **do**

  // sample a scenario from $\boldsymbol{\Omega}$

  $\boldsymbol{\omega}_i \in \boldsymbol{\Omega}$

  // solve the deterministic subproblem

  $\boldsymbol{x_i} = \underset{\boldsymbol{x}}{\operatorname{argmax}}\, J(\boldsymbol{x}, \boldsymbol{\omega_i})$ subject to $C_k(\boldsymbol{x}, \boldsymbol{\omega_i}), \quad k = 1, 2, \ldots, m$

end

// evaluate each scenario solution across all sampled scenarios

for *i in 1 to n* **do**

  for *j in 1 to n* **do**

    $h_{ij} = J(\boldsymbol{x_i}, \boldsymbol{\omega_j})$

    for *k in 1 to m* **do**

      $c_{ijk} = \psi_k(\boldsymbol{x_i}, \boldsymbol{\omega_j})$

    end

  end

end

// estimate the objective and constraint CDFs for each solution

for *i in 1 to n* **do**

  $\hat{F}_{\boldsymbol{x_i}}(t) = \frac{1}{n} \sum_{j=1}^{n} I(h_{ij} \leq t)$

  for *k in 1 to m* **do**

    $\hat{G}_{\boldsymbol{x_i}k}(t) = \frac{1}{n} \sum_{j=1}^{n} I(c_{ijk} \leq t)$

  end

end

// obtain the QSA solution, for example

$\boldsymbol{x}^* = \underset{i \in \{1,2,\ldots,n\}}{\operatorname{argmax}} \int_0^1 F_{\boldsymbol{x_i}}^{-1}(t) \cdot \phi_0(t) dt$

subject to

$\int_0^1 G_{\boldsymbol{x_i}k}^{-1}(t) \cdot \phi_k(t) dt \leq \gamma_k, \ k = 1, 2, \ldots, m$

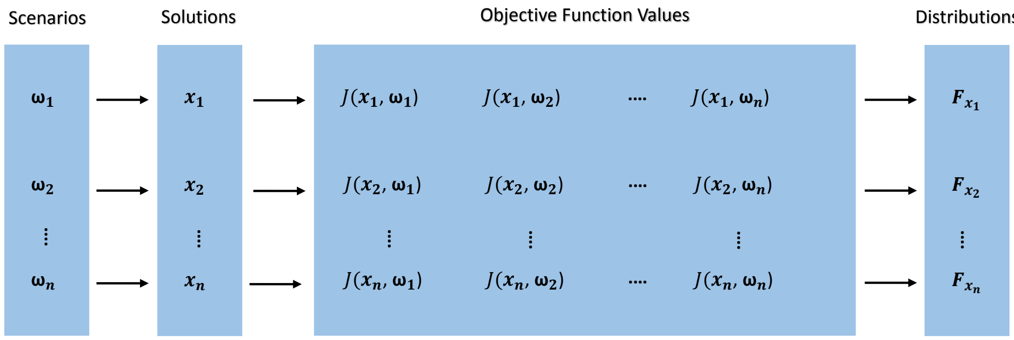

Figure 1: Diagram of the QSA Approach

3.2. Stopping Rule

A drawback of most metaheuristics is the absence of an effective stopping criteria [14]. Most methods stop after a given number of iterations or a maximum number of consecutive iterations is performed without improvement in the best known solution. In this section, we present a probabilistic stopping rule for the QSA algorithm when solving problems of the form specified in Equation (5).

Let $\hat{\boldsymbol{x}}$ be the QSA solution obtained using a set, $\Omega_1 \subset \Omega$, of n_1 realized scenarios. Let ρ be the probability that a randomly chosen scenario from the set Ω, whose deterministic solution, $\tilde{\boldsymbol{x}}$, satisfies the constraints $g_k(G^{-1}_{\tilde{\boldsymbol{x}}k}(q)) \leq 0$, $k = 1, 2, \ldots, m$, in Equation (5) and has a better performance over the set Ω_1 than that of $\hat{\boldsymbol{x}}$. We propose a stopping rule based on statistical evidence that $\rho < \rho_0$, for some specified value ρ_0. We sample an additional set of $n_2 \ll n_1$ independently and identically distributed scenarios, $\Omega_2 \subset \Omega$. Let $0 \leq N \leq n_2$ be the number of sampled scenarios in Ω_2 with solutions that have a better performance over the set Ω_1 than $\hat{\boldsymbol{x}}$. Suppose that the observed value of N is k. Our stopping rule tests the hypothesis $H_0 : \rho \geq \rho_0$ vs. $H_1 : \rho < \rho_0$ as follows. We reject H_0 at level α if $P_{H_0}(N \leq k) \leq \alpha$, for $\rho \geq \rho_0$. Under, $H_0 : \rho \geq \rho_0$, N follows the Binomial distribution, $Bin(n_2, \rho)$ with $\rho \geq \rho_0$. It can be shown [10] that

$$\sup_{\rho \geq \rho_0} P\left(Bin(n_2, \rho) \leq k\right) = P\left(Bin(n_2, \rho) \leq k\right) = \sum_{i=0}^{k} \binom{n_2}{i} \rho_0^i (1 - \rho_0)^{n_2 - i} \tag{7}$$

and we reject $H_0 : \rho \geq \rho_0$ in favour of $H_1 : \rho < \rho_0$ at level α if the RHS of Equation (7) is less than α, where α is a specified small probability, such as 0.05 corresponding to a 95% confidence level.

4. QSA Formulation of the Competing Biomass Supply Chains Problem

To investigate the performance of the QSA approach for energy supply chain design and optimization under uncertainty, we discuss a supply chain network design problem for supplying forest biomass feedstock to three competing power plants within a study region. We assume that the power plant locations are given, and that they utilize only one type of feedstock, namely forest harvest residue (FHR), which includes tops and branches and unmerchantable wood left after stand harvesting. Each power plant is associated with an annual energy demand expressed in terms of giga.

The area of the study region is 400 km^2 and is divided into 625 forest cells representing 16 km^2 each. The grid cells have an annual technical availability of FHR that is computed as a product of the harvesting factor and the theoretical availability of FHR at each cell. The harvesting factor is a value between 0 and 1, which evaluates the harvesting capability of a grid cell. The FHR available at each grid cell is also associated with a wet basis moisture content. The wet basis moisture content is used to describe the water content of biomass and is defined as the percentage equivalent of the ratio of the weight of water to the total weight of the biomass. When it comes to biomass, the higher the moisture content the lower the net energy content. The energy content of a wet ton of FHR was calculated using Equation (8), which expresses the relationship between lower heating value, l, and wet basis moisture content, m, as an average for wood in terms of BTU per ton [15].

$$l(m) = (8660 - 9712m) \times 2240 \tag{8}$$

The constants 8660 and -9712 in Equation (8) are in units of BTU per lb and represent the estimated average higher heating value of wood (across several species), and the effect of the moisture content on the lower heating value of wood, respectively. The multiplication factor 2240 is the conversion factor from lb to long ton (t).

The cost of procurement of a green ton of biomass was calculated based on processing and transportation costs. The processing cost includes harvesting, grinding, chipping and piling of FHR. The transportation cost from a forest cell to a power plant was assumed proportional to the distance between them. The vehicle considered for transporting forest biomass feedstock is a tractor-trailer with a 53-foot semitrailer and a payload of 40.55 tons. A cost of \$5.24 per green ton (gt) was included to account for loading and unloading operations. The charge-out rate for a biomass truck with operator was fixed at \$85 per hour (h). An average driving speed of 50 km/h was assumed. With this information, the cost of transporting forest biomass from each road cell of the study region to the three power plants was established. The remaining parameters are subject to random variability: biomass availability per forest cell; harvesting factor per forest cell; processing cost per forest cell; average FHR moisture content per forest cell; energy demand at each power plant.

The random parameters are assumed to be independent and normally distributed with mean and standard deviations given in Table 1. A Gaussian random field with an exponential correlation function was used to model the spatial dependence in FHR availability and the moisture content between grid cells, as shown in Figures 2 and 3 with respect to the location of the three biomass plants. Realizations of the values of the random parameters represent specific scenarios. For a single scenario, the parameters, indices and variables used for developing the optimization model are listed in Table 1, with random parameters being assigned normal distributions.

The objective function, which identifies the FHR catchment area of each power plant that minimizes the overall procurement cost, is specified in Equation (9) below. The constraints used in the model are specified

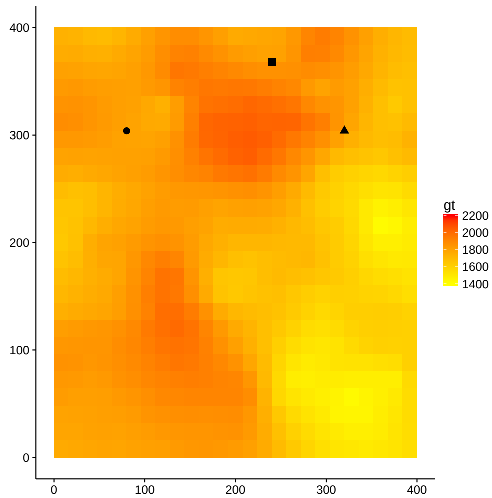

Figure 2: FHR Availability (gt)

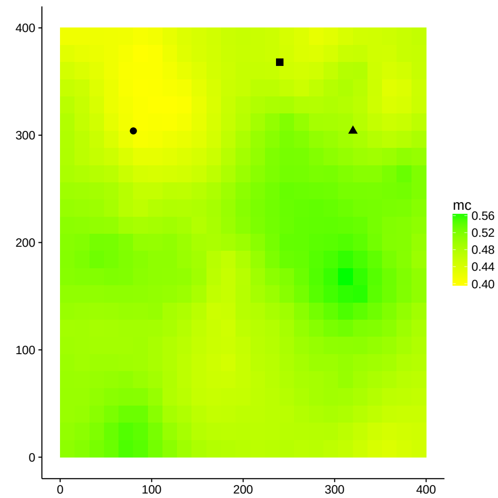

Figure 3: FHR Moisture Content (% Wet Basis)

Table 1: Optimization Model Notation

	Symbol	Description	Units	Value/Distribution
Sets				
	M:	set of power plants	–	$\{1, 2, 3\}$
	U:	set of forest cells	–	$\{1, 2, \ldots, 160000\}$
Indices				
	m:	power plant	–	$m \in M$
	u:	forest cells	–	$u \in U$
Parameters				
	λ :	load/unload overhead	$\$ \cdot \text{gt}^{-1}$	5.24
	α:	harvesting factor	%	$N(61, 5)$
	β:	processing cost at roadside	$\$ \cdot \text{gt}^{-1}$	$N(26, 3)$
	ϵ:	average driving speed of truck	$\text{km} \cdot \text{hr}^{-1}$	50
	ν :	truck capacity	gt	40.55
	δ_m:	annual energy demand of each power plant	MMBTU	$N(850000, 16000)$ $N(1100000, 20000)$ $N(1300000, 24000)$
	θ_u:	availability of FHR in the uth forest cell	gt	$N(4900, 280)$
	η_u:	wet basis moisture content of FHR in the uth forest cell	%	$N(50, 3)$
	d_{mu}:	distance from the uth forest cell to the mth power plant	km	Euclidean distance
	τ_{mu}:	charge out rate for a truck driver	$\$ \cdot \text{hr}^{-1}$	85
Variables				
	x_{mu} :	amount of annual FHR harvested from the uth forest cell for the mth power plant	gt	–

in Equations (10) to (12). To use a more compact notation, let $\boldsymbol{x} = (x_{11}, x_{12}, \ldots, x_{mu})$ represent the vector of decision variables and $\boldsymbol{\omega}$ be the vector of model parameters.

$$\underset{\boldsymbol{x}}{\text{minimize}} \quad J(\boldsymbol{x};\boldsymbol{\omega}) = \sum_{m\in M}\sum_{u\in U} x_{mu} \times \left(\beta + \lambda + \frac{2d_{mu} \times \tau_{mu}}{\nu}\right) \tag{9}$$

$$\text{subject to} \quad \sum_{u\in U} l(\eta_u) \times x_{mu} \geq \delta_u \quad \forall m \in M \tag{10}$$

$$\sum_{m\in M} x_{mu} \leq \vartheta_{us} \quad \forall u \in U \tag{11}$$

$$x_{mu} \geq 0 \quad \forall (m,u) \in M \times U \tag{12}$$

The R package "Rglpk", which is an R interface to the GNU linear programming Kit (GLPK), was used to solve the resulting deterministic linear optimization scenario subproblems [7]. GLPK is open source software for solving large-scale linear programming (LP), mixed integer linear programming (MILP) and other related problems.

In accordance with the QSA coordination model described in Equation (6), we seek a solution that minimizes the overall FHR procurement cost, subject to satisfying at least $(1 - \gamma_1) \times 100\%$ of the energy demand. We give an exponential increase in importance to the upper quantiles of the stochastic cost function and define the objective weight function, ϕ_0, in Equation (1) as

$$\phi_0(t) = \begin{cases} \frac{e^t}{(e^t-1)} & \text{if } \ 0 \leq x \leq 1 \\ 0 & \text{othewise.} \end{cases} \tag{13}$$

This resembles a minimax optimization, but instead of minimizing the worst-scenario performance we minimize a weighted average of the performances across all scenarios and assign larger penalties to less desirable performances. We define the constraint weight function, ϕ_1, in Equation (1) as

$$\phi_1(t) = \begin{cases} 1 & \text{if } \ t = q_1 \\ 0 & \text{othewise,} \end{cases} \tag{14}$$

where the parameter $0 < q_1 < 1$ represents the probability of satisfying the demand constraint. For this problem, the QSA penalty function, used in evaluating the i^{th} solution in the j^{th} scenario, is defined as follows:

$$C_1\left(\boldsymbol{x_i}, \boldsymbol{\omega_j}\right) = \frac{1}{3}\sum_{m=1}^{3}\left(\frac{\delta_{mi} - y_{mij}}{\delta_{mj}}\right)^{+}, \tag{15}$$

where δ_{mi} is the energy demand of the m^{th} power plant in the j^{th} scenario, and y_{mij} is the energy produced by the m^{th} power plant when applying the i^{th} solution to the j^{th} scenario. Here, the exponent $(+)$ indicates that only positive values of the parenthetical expression are considered. In other words, we do not penalize for exceedance of the demand. Equation (15) calculates the overall percentage of unfilled demand across the three power plants for a given solution and scenario.

Notice that a larger demand will typically incur a greater procurement cost. Hence, it would be erroneous to compare solutions in terms of the total procurement cost, because the demand is not fixed between scenarios. A larger demand will typically incur a greater procurement cost. If FHR is to be used as a fuel, its value lies in its energy content and not in its weight. In order to make valid comparisons between scenario solutions, the procurement cost should be made relative to the amount of energy produced. To this end, the total procurement cost of each solution and scenario combination is divided by the corresponding amount of energy obtained from the procured FHR. Thus, for the proper comparison of solutions evaluated at different scenarios we replace Equation (9) with

$$\tilde{J}(\boldsymbol{x};\boldsymbol{\omega}) = \frac{J(\boldsymbol{x};\boldsymbol{\omega})}{\sum\limits_{m\in M}\sum\limits_{u\in U} l(\eta_u)\times x_{mu}}, \tag{16}$$

which is in units of dollar per MMBTU.

5. Results

In this section, we evaluate the performance of the QSA approach for solving the stochastic optimization problem described in Section 4. In particular, we want to examine (1) its ability to identify optimal solutions (2) its precision and accuracy and (3) its computational complexity and scalability. To this effect, we compare the solution obtained by the QSA approach to that obtained by scenario analysis (SA) as described by Dembo [5].

The QSA approach for this problem minimizes a weighted average of the quantiles of the procurement cost, obtained across scenarios, according to the weight function given by Equation (13). We restrict attention to solutions that have a 90% probability of satisfying at least 95% of the demand across scenarios. Thus, $q_1 = 0.05$ in Equation (14). In contrast, the SA approach minimizes the coordination model given in Equation (3), where $J(\boldsymbol{x_i},\boldsymbol{\omega_i}) = \tilde{J}(\boldsymbol{x_i},\boldsymbol{w_j})$, $C_u(\boldsymbol{x},\boldsymbol{\omega_i}) = C_1\left(\boldsymbol{x_i},\boldsymbol{\omega_j}\right)$, and $C_d(\boldsymbol{x})$ is the deterministic constraint given in Equation 12. Minimization of the The SA coordination model was done using the spectral projected gradient method for large-scale optimization with simple constraints implemented in the“BB” package in R [20].

In order to asses the performance of the QSA approach, we compare its solution to that obtained using the standard SA approach. It was found that even for this modest stochastic optimization problem, the computational time of the SA method was so significant that it is almost time prohibitive to compute for a large number of scenarios. Given this limitation of the SA approach, a total of $n = 500$ scenarios were sampled from their underlying Gaussian distribution. A computation time of 119.7 and 1.3 minutes was required to solve this problem running on a 3.6 GHz Intel Xeon processor for the QSA and SA methods, respectively. This tremendous savings in computational time is an enormous advantage of the QSA method. To assess if the solution converged, we sampled an additional 200 scenarios and used the stopping rule of

Equation (7) with $\rho_0 = 0.05$. All of the 200 additionally sampled scenarios yielded solutions with lower performance compared to that of the QSA solution, resulting in a p-value of 3.51×10^{-5}. Therefore, we have strong evidence that the solution converged and reject the null hypothesis, $H_0 : \rho \geq 0.05$ in favor of $H_1 : \rho < 0.05$. In other words, we are confident that the QSA solution identified from the one thousand sampled scenarios is among the top 5% of possible scenario solutions for this problem.

The optimal catchment areas identified by the QSA and SA solutions are shown in Figures 4 and 5, respectively. The catchment area for the QSA solution is much more compact than that of the SA solution. The SA solution attempts to perform well on average. As a result, it tends to over collect in order to compensate for extreme scenarios with very poor FHR availability. This phenomenon is illustrated very clearly in Figure 5. Poor yields were simulated in several scenarios across the middle belt region of the study region, which is why the SA solution tries to skip it over. These outlier scenarios had no influence on the QSA solution, since it optimizes the quantiles of the objective and constraint distributions, which are much less sensitive to outliers than the mean. The CDF of the procurement cost obtained by the QSA and SA solutions are shown in Figure 6. The x-axis of Figure 6 represents the cost of procurement in dollar per MMBTU. Likewise, the empirical CDF of the constraint on the unfilled demand is shown in Figure 7. The x-axis of Figure 7 corresponds to the overall percentage of unfilled demand, which was calculated using Eq. (15). The y-axis of Figure 7 describes the probability that the overall percentage of unfilled demand will be less than or equal to the quantity on the x-axis. The distribution of the cost function of the QSA solution is strictly less than that of the SA solution. Thus, the QSA solution consistently outperforms the SA solution in terms of minimizing the cost in each scenario. The 0.90 quantile of unfilled demand constraint distribution is 4.76%, which satisfies the requirement that it be less than 5.0%. The SA solution on average collects 22.7% more FHR than the QSA solution to meet the same energy demand. Table 2 compares several quantiles of the cost distribution of the QSA and SA solutions.

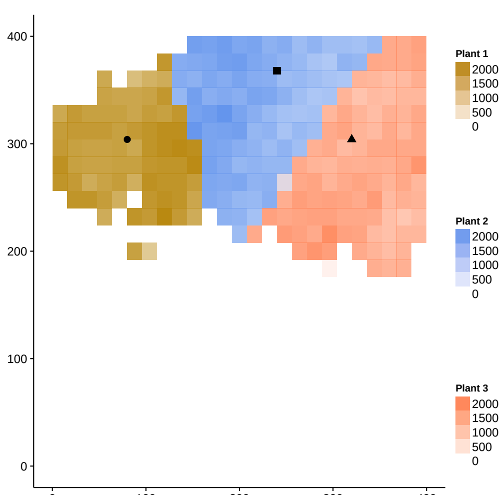

Figure 4: Catchment Area of the QSA Solution (gt)

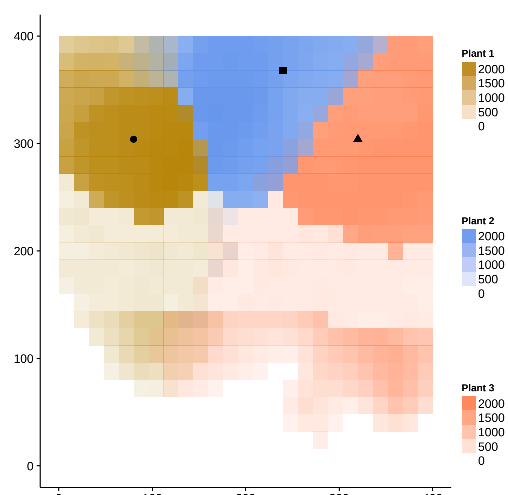

Figure 5: Catchment Area of the SA Solution (gt)

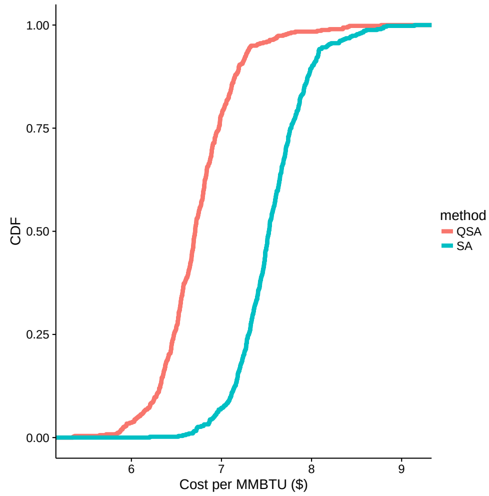

Figure 6: CDF of the Cost ($\$ \cdot \mathrm{MMBTU}^{-1}$)

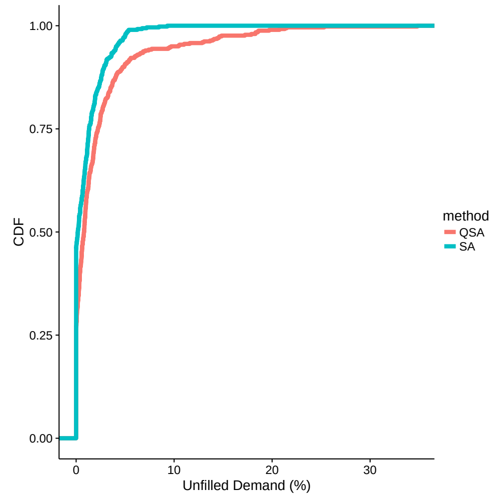

Figure 7: CDF of the Unfilled Demand (gt)

Table 2: Quantiles of the Cost Distribution of the QSA and SA Solutions

Quantile	QSA	SA
0.10	6.28	7.11
0.25	6.47	7.30
0.50	6.71	7.53
0.75	6.91	7.77
0.90	7.19	8.01

6. Conclusion

Standard approaches for optimization under uncertainty define the best solution as the one with the best expected value of the performance metric, subject to satisfying a set of constraints. However, when other statistics or attributes of the distribution of the objective function are more appropriate for comparing solutions (e.g. upper or lower quantiles, inter quartile range, overall dispersion, etc.), then the expected value based approaches are not able to accommodate the selection needs.

We presented a robust quantile-based scenario analysis procedure (QSA) for optimization under uncertainty. We applied our approach to address the problem of analyzing a system of biomass supply chains, which are competing for the same feedstock, subject to stochastic demand and supply in addition to numerous constraints. In such circumstances, it is important to identify solutions that will, with high probability, meet the energy demand and satisfy the problem constraints, for the majority of scenarios. The energy demand and feedstock supply were considered uncertain, but with known distributions. Our method was able to provide a solution that minimizes a weighted average of the procurement cost distribution, while satisfying at least 90% of the energy demand with a probability of 95%.

We compared the performance of the QSA approach to that of the classical SA approach for stochastic optimization. Results show that the QSA approach is the best performing method in terms of computational complexity and the percentage of scenarios solved optimally while satisfying the stochastic constraints.

References

[1] Md Bedarul Alam, Reino Pulkki, Chander Shahi, and Thakur Upadhyay. Modeling woody biomass procurement for bioenergy production at the atikokan generating station in northwestern ontario, canada. *Energies*, 5(12):5065–5085, 2012.

[2] Demet Batur and F Choobineh. A quantile-based approach to system selection. *European Journal of Operational Research*, 202(3):764–772, 2010.

[3] A.R. Bell, C. Brooks, and M. Prokopczuk. *Handbook of Research Methods and Applications in Empirical Finance*. Handbooks of Research Methods and Applications Series. Edward Elgar Publishing, 2013. ISBN 9780857936097. URL `https://books.google.ca/books?id=3DiJw7OyEpEC`.

[4] John R Birge and Francois Louveaux. *Introduction to stochastic programming*. Springer, 2011.

[5] Ron S Dembo. Scenario optimization. *Annals of Operations Research*, 30(1):63–80, 1991.

[6] Mahmood Ebadian. Design and scheduling of agricultural biomass supply chain for a cellulosic ethanol plant. 2013.

[7] Kurt Hornik, Stefan Theussl, and Maintainer Stefan Theussl. Package 'rglpk'. 2012.

[8] Larry Hughes and Jacinda Rudolph. Future world oil production: growth, plateau, or peak? *Current Opinion in Environmental Sustainability*, 3(4):225–234, 2011.

[9] Richard A Kerr. Technology is turning us oil around but not the world's. *Science*, 335(6068):522–523, 2012.

[10] Achim Klenke, Lutz Mattner, et al. Stochastic ordering of classical discrete distributions. *Advances in Applied probability*, 42(2):392–410, 2010.

[11] Tao Lin, Luis F Rodríguez, Yogendra N Shastri, Alan C Hansen, and KC Ting. Integrated strategic and tactical biomass–biofuel supply chain optimization. *Bioresource technology*, 156:256–266, 2014.

[12] Peter Rauch and Manfred Gronalt. The terminal location problem in the forest fuels supply network. *International Journal of Forest Engineering*, 21(2):32–40, 2010.

[13] A Ravi Ravindran. *Operations research and management science handbook*. Crc Press, 2007.

[14] Celso C Ribeiro, Isabel Rosseti, and Reinaldo C Souza. Effective probabilistic stopping rules for randomized metaheuristics: Grasp implementations. In *Learning and Intelligent Optimization*, pages 146–160. Springer, 2011.

[15] Joseph P Roise, Glenn Catts, Dennis Hazel, Alex Hobbs, and Chris Hopkins. Balancing biomass harvesting and drying tactics with delivered payment practice. Technical report, Technical Report-Not held in TRLN member libraries,) Greenville, SC: US Endowment for Forestry and Community, 2013.

[16] Nikolaos V Sahinidis. Optimization under uncertainty: state-of-the-art and opportunities. *Computers & Chemical Engineering*, 28(6):971–983, 2004.

[17] R Saidur, EA Abdelaziz, A Demirbas, MS Hossain, and S Mekhilef. A review on biomass as a fuel for boilers. *Renewable and Sustainable Energy Reviews*, 15(5):2262–2289, 2011.

[18] Alexander Shapiro. Sample average approximation. In *Encyclopedia of Operations Research and Management Science*, pages 1350–1355. Springer, 2013.

[19] Stoyan V Stoyanov. A moment expansion of downside risk measures. *Available at SSRN 1721227*, 2010.

[20] Ravi Varadhan and Paul Gilbert. BB: An R package for solving a large system of nonlinear equations and for optimizing a high-dimensional nonlinear objective function. *Journal of Statistical Software*, 32 (4):1–26, 2009. URL `http://www.jstatsoft.org/v32/i04/`.

[21] Roger JB Wets. The aggregation principle in scenario analysis and stochastic optimization. In *Algorithms and model formulations in mathematical programming*, pages 91–113. Springer, 1989.

[22] Dajun Yue, Fengqi You, and Seth W Snyder. Biomass-to-bioenergy and biofuel supply chain optimization: Overview, key issues and challenges. *Computers & Chemical Engineering*, 66:36–56, 2014.

[23] David S. Zamar, Bhushan Gopaluni, Shahab Sokhansanj, and Nathaniel K. Newlands. Robust Optimization of Competing Biomass Supply Chains Under Feedstock Uncertainty. In *9th International Symposium on Advanced Control of Chemical Processes*, pages 1223–1228, Whistler, British Columbia, 2015. The International Federation of Automatic Control. URL `https://www.google.com/search?q=zamar+robust+optimization&ie=utf-8&oe=utf-8#q=zamar+robust+optimization+stochastic`.

[24] Rachel ES Ziemba and William T Ziemba. *Scenarios for risk management and global investment strategies*. John Wiley & Sons, 2008.