Computer-Aided Civil and Infrastructure Engineering **28** (2013) 247–259

Integrated Planning of Supply Chain Networks and Multimodal Transportation Infrastructure Expansion: Model Development and Application to the Biofuel Industry

Leila Hajibabai & Yanfeng Ouyang*

Department of Civil and Environmental Engineering, University of Illinois at Urbana-Champaign, IL, USA

Abstract: *As the biofuel industry continues to expand, the construction of new biorefinery facilities induces a huge amount of biomass feedstock shipment from supply points to the refineries and biofuel shipment to the consumption locations, which increases traffic demand in the transportation network and contributes to additional congestion (especially in the neighborhood of the refineries). Hence, it is beneficial to form public-private partnerships to simultaneously consider transportation network expansion and biofuel supply chain design to mitigate congestion. This article presents an integrated mathematical model for biofuel supply chain design where the near-optimum number and location of biorefinery facilities, the near-optimal routing of biomass and biofuel shipments, and possible highway/railroad capacity expansion are determined. The objective is to minimize the total cost for biorefinery construction, transportation infrastructure expansion, and transportation delay (for both biomass/biofuel shipment and public travel) under congestion. A genetic algorithm framework (with embedded Lagrangian relaxation and traffic assignment algorithms) is developed to solve the optimization model, and an empirical case study for the state of Illinois is conducted with realistic biofuel production data. The computational results show that the proposed solution approach is able to solve the problem efficiently. Various managerial insights are also drawn. It shall be noted that although this article focuses on the booming biofuel industry, the model and solution techniques are suitable for a number of application contexts that simultaneously involve network traffic equilibrium, infrastructure expansion, and facility location choices (which determine the origin/destination of multi-commodity flow).*

*To whom correspondence should be addressed. E-mail: *yfouyang@illinois.edu*.

DOI: 10.1111/j.1467-8667.2012.00791.x

1 INTRODUCTION

Biofuels, such as bioethanol, can be derived from crops (e.g., corn and sugarcane) and other cellulosic organic sources (such as wood and grass). Bioethanol can be used as an oxygenate in gasoline at various mixture ratios (Energy Information Administration, 2008). It has become one of the main renewable energy sources in the United States, contributing to half of the nation's consumed gasoline (Kang et al., 2010). Worldwide, biofuels support 1.8% of the total transport fuel needs as of 2008. Investment in biofuels production capacity exceeded $4 billion worldwide in 2007, and it continues to grow due to the ever-increasing demand for fuel additives and biofuel mandates. Other important factors, such as shortage of fossil fuel, need for energy security, and concerns over air quality, have also attracted enormous attention to biofuel production and development.

Production of biofuel imposes challenges to a variety of infrastructure systems that are associated with all stages of biofuel industry operations (including biomass production, harvesting, storage, transportation, processing into biofuel, and finally shipment of biofuel to consumption points) at regional, national, or even international levels. Biofuel supply chain planning should simultaneously consider the number and location of biofuel refineries, the required distribution networks, and transportation plans for bioproducts. Compared to traditional fossil fuels, the corn and cellulosic biomass contain a relatively lower density of energy content. Therefore, to satisfy the biofuel demand, biomass materials must be collected across large agricultural production areas, and there is a trade-off between using a few large centralized facilities (to benefit from production economies of scale) versus using a more decentralized network systems (to save transportation among supply, production, and consumption locations).

Economists have long studied the financial practicability of "biomass-related" energy supply chains (e.g., using corn residues as supplementary fuel in "coal-fired power plants") by considering farm production, transportation, and processing costs (English et al., 1981; Nienow et al., 2000). However, only after the biofuel industry boomed in recent years were logistics models proposed for integrated analysis on biomass collection, storage, and shipment (Sokhansanj et al., 2006). Kaylen et al. (2000) showed that producing ethanol from cellulosic products is cost-effective, while Tembo et al. (2003) developed multi-period, multi-region models that considered technical alternatives for biomass production, transportation, and processing. Mapemba (2005) estimated shipment cost of biomass to a biorefinery under various agricultural and production factors (e.g., biomass harvest days, frequency, and refinery capacity). Searcy et al. (2007) assessed a variety of possible modes for biomass and ethanol shipments. Trucking generally provides better service for short hauls, and hence turns out to be the most cost-effective mode for biomass transportation (Mahmudi and Flynn, 2006; Brown et al., 2007). Nevertheless, for large volumes and longer distances, or when roadway congestion prevents efficient truck-based shipments, rail shipment might become favorable (Kang et al., 2010). Eathington and Swenson (2007) used geospatial information systems (GIS) applications to facilitate decisions on refinery site, size, and technology under various scenarios regarding energy demand level, industry growth, and impacts on the job market. Tursun et al. (2008) and Kang et al. (2010) proposed multiyear supply chain models for production of corn and cellulosic based ethanol and byproducts. Bai et al. (2011) further considered traffic congestion and public travel delay as a result of increased feedstock and bioproducts shipments, and incorporated such effects into the planning of biorefinery location and transportation routing. A fixed-charge facility location model was jointly formulated with a traffic assignment model, and the problem was solved effectively by Lagrangian relaxation (Fisher, 1981) with an embedded convex combination method (Frank and Wolfe, 1956; Sheffi, 1985).

Biorefinery location and bioproduct shipment should be environmentally, technically, and financially sustainable. When biorefinery facilities are constructed, shipment of biomass and biofuel induces significantly higher shipment demand that originates or ends at the refineries. Such shipment demand could cause additional congestion delay in the transportation network particularly on roadway links (bottlenecks) where background (i.e., public) traffic demand is near or has reached roadway capacity. This not only increases travel time of the general public but also has a significant impact on the operational efficiency of the biofuel supply chain itself, which, in turn, shall affect the biorefinery location decisions.

Unfortunately, these endogenous relationships have largely been ignored in the supply chain planning literature. While there have been strategic transportation planning models that integrate sustainability issues into facility location design, network analysis, and spatial and regional economic analysis (e.g., Lopez and Monzon, 2010), the most relevant study in this direction is probably Bai et al. (2011), which shows that integrating refinery location and shipment routing decisions could help mitigate congestion impacts. However, the effectiveness of such demand-side strategies (e.g., routing of shipments) shall diminish when roadway capacity is already reached in the entire network. In this situation, the options on the capacity supply side (e.g., adding lanes or railroad segments) should be considered as an integral part of biofuel supply chain planning.

In general, capital investment and land use restrictions tend to prohibit extensive roadway expansion as an alternative to mitigate congestion (MTP, 2006). Hence, highway investment decisions under dynamic responsive traffic demand, have largely been focusing on long-term infrastructure maintenance planning and rehabilitation (e.g., Ouyang, 2007; Ng et al., 2009; Gao et al., 2012) or highway investment alternatives under a certain budget (e.g., Li et al., 2010). However, in the context of biofuel industry expansion, building additional capacity in the highway or railroad networks might be much more feasible. The capital investment associated with transportation infrastructure expansion may be considered as part of the biofuel industry's investment plan, and potential public-private partnership may be established to facilitate the investment (Unnikrishnan et al., 2009).

This article proposes an integrated mathematical model for biofuel supply chain design that encompasses biorefinery location, biomass and ethanol transportation, and possible infrastructure capacity expansion in the transportation network, as illustrated in Figure 1, while the objective is to minimize the total cost including the transportation costs (for bioproduct shipment and for the traveling public) and the infrastructure investments in both biorefineries and network capacity. The biomass and ethanol routing part of the problem can be modeled as a traffic assignment problem, and the refinery location part of the problem generally can be modeled as a fixed-charge facility location problem. However, the integration of facility location, shipment routing, and infrastructure expansion makes the problem very difficult to solve. We propose a genetic

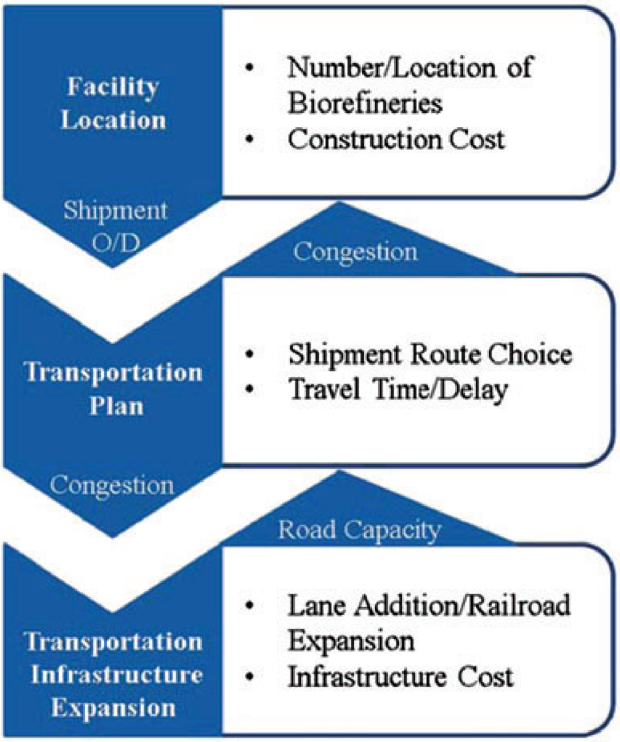

Fig. 1. Interactions among transportation planning, facility location, and infrastructure capacity expansion.

algorithm to handle the facility location and infrastructure capacity expansion part of the problem, while using embedded Lagrangian relaxation and convex combination algorithms to solve the routing decisions. The proposed methodology is applied to an empirical case study and a series of sensitivity analyses are conducted. Numerical results show that the proposed solution algorithms effectively solve the proposed problem, and it is also shown that the integration of biorefinery location, biomass shipment routing, and infrastructure capacity expansion decisions have a significant socioeconomic impact.

It shall be noted that although this article focuses on a specific biofuel supply chain design problem due to its importance, the modeling framework can be applied to a wide range of application contexts (e.g., urban land use development). Similar problems would arise as long as the origins, destinations, and paths of multiple types of commodity flow are determined simultaneously, while the capacity of network links could be expanded to mitigate congestion.

The exposition of this article is as follows. Section 2 will focus on the mathematical formulation and the notations used in the model. Section 3 will introduce the solution approach that has been used to solve the problem. Section 4 will present the empirical case study for the state of Illinois. Section 5 will elaborate the results and conduct a series of sensitivity analyses. Finally, Section 6 will summarize the study by providing conclusions and future trends.

2 MODEL FORMULATION

This section presents a mixed integer nonlinear program (MINLP) that simultaneously addresses biorefinery location, shipment routing, and infrastructure expansion decisions under traffic congestion. The objective of our model is to minimize the total costs for the entire supply chain including the investments in refinery construction and network capacity expansion, and transportation cost (for biomass and biofuel shipment, and public travel delay).

We let I^s and I^d represent the sets of biomass supply and biofuel demand locations, respectively. Region $i \in I^s$ produces biomass supply h_i^s, and region $i \in I^d$ has biofuel demand h_i^d. The units of both h_i^s and h_i^d are hourly passenger car equivalents, which are estimated based on full-truck load in volume and hourly passenger car flow equivalent for trucks (HCM, 2000). To simplify our calculations, for train shipments we assumed 125-ton freight cars (with approximately 26 tons maximum capacity), and converted the required shipment volumes into the number of rail cars needed. The amount of train traffic is then computed by the average number of rail cars per train (e.g., 100 cars per train).

Let J denote the set of candidate locations for biorefineries. The construction of a biorefinery at location $j \in J$ involves a fixed cost of m_j and generates a production capacity of C_j. In this study, we simply assume a fixed biorefinery capacity for each candidate location, as in Kang et al. (2010) and Bai et al. (2011). Such capacities are generally set to be very large in order to exploit the economy of scale in biofuel production. Sometimes, the refinery capacity at one given location could be selected from a discrete set. In such cases, the location part of the model should follow the multi-type facility location formulation (Hakan et al., 2012; Klose and Drexl, 2005; Jen et al., 1968) by introducing additional decision variables that indicate the production capacity (i.e., facility type) at each candidate location.

The selection of locations for biorefineries is determined by decision variables $\{Y_j\}$ as follows:

$$Y_j = \begin{cases} 1 & \text{if a biorefinery is built at } j \in J \\ 0 & \text{otherwise.} \end{cases}$$

Suppose that biomass and biofuel shipments go through a multimodal transportation network containing a set of highway links A, and a set of railroad links

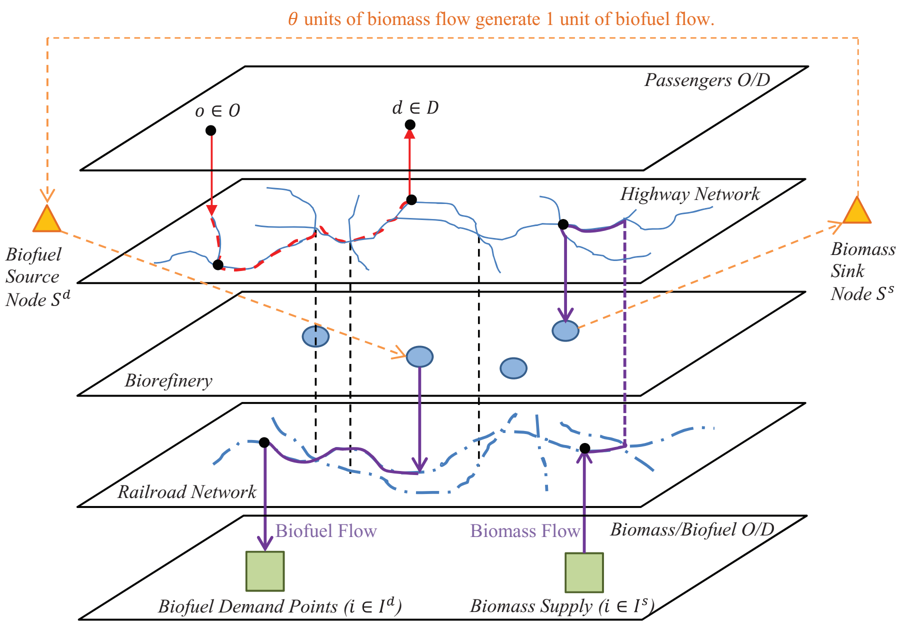

Fig. 2. Biorefinery candidate location, biomass production and biofuel demand points, and multimodal transportation networks.

B. Intermodal transshipments are possible at junctions and yards. Similar to Bai et al. (2011), two imaginary nodes, S^s and S^d, are added respectively as a sink node for the biomass transportation, and a source node for the biofuel shipments which are connected by a set V of virtual links to candidate location $j \in J$ if there is an open facility (i.e., $Y_j = 1$); see Figure 2. This can be interpreted as considering the set of biomass supply locations I^s as the origins of the biomass shipment with flow of h_i^s to node S^s as the only destination; meanwhile, node S^d is the only origin of biofuel shipments to the set of demand points I^d as destinations. This ensures that all the flows from the source node or to the sink node will pass through at least one open biorefinery. Figure 2 schematically illustrates the biofuel supply chain system, which includes biorefinery location, biomass production and biofuel demand points, the multimodal transportation networks, and the virtual links. Without losing generality, we include in B those "transshipment links" between the highway nodes and the railroad nodes (i.e., those vertical links between the highway layer and the railroad layer in Figure 2).

Let $K^{s,i}$ and $K^{d,i}$, respectively denote the set of intermodal paths from a biomass supply point $i \in I^s$ to the sink node S^s, and those from the source node S^d to a demand point $i \in I^d$. Biomass flow $f_k^{s,i}$ travels from $i \in I^s$ to S^s on a possible path $k \in K^{s,i}$, and biofuel flow $f_k^{d,i}$ travels from S^d to $i \in I^d$ on a path $k \in K^{d,i}$. We assume that under current technology, θ units (in terms of volume) of biomass generate one unit of biofuel. We further introduce link-arc incidence parameters $\delta_{a,k}^{s,i}$ and $\delta_{a,k}^{d,i}$ (Sheffi, 1985) as follows:

$$\delta_{a,k}^{\cdot,i} = \begin{cases} 1 & \text{if path } k \in K^{\cdot,i} \text{ includes link } a \in A \cup B \\ 0 & \text{otherwise} \end{cases}, \cdot \in \{s,d\}.$$

Similarly, we define parameters $\Delta_{j,k}^{s,i}$ and $\Delta_{j,k}^{d,i}$ for the virtual links connected to roadway network:

$$\Delta_{j,k}^{\cdot,i} = \begin{cases} 1 & \text{if path } k \in K^{\cdot,i} \text{ includes node } j \in J \\ 0 & \text{otherwise} \end{cases}, \cdot \in \{s,d\}.$$

Link flows on these virtual links, v_j^s and v_j^d represent biorefinery throughput at location $j \in J$. They can be expressed as the following:

$$v_j^s = \sum_{i \in I^s} \sum_{k \in K^{s,i}} f_k^{s,i} \Delta_{j,k}^{s,i},\ v_j^d = \sum_{i \in I^d} \sum_{k \in K^{d,i}} f_k^{d,i} \Delta_{j,k}^{d,i}.$$

In addition to the biomass/biofuel flows, we assume that there are passenger traffic flows on the highway transportation network from a set of origins, O to a set of destinations, D. Let $K^{o,d}$ denote the set of paths connecting a passenger flow origin $o \in O$ and destination $d \in D$ through the highway network, and let $f_k^{o,d}$ represent the flow on path $k \in K^{o,d}$. On each highway link $a \in A$, the total passenger flow is $\sum_{o \in O} \sum_{d \in D} \sum_{k \in K^{o,d}} f_k^{o,d} \delta_{a,k}^{o,d}$, where $\delta_{a,k}^{o,d} = 1$ if path $k \in K^{o,d}$ contains link $a \in A$, and 0 otherwise.

On the other hand, the traffic in existing railroad networks is usually less sensitive to the addition of biomass/biofuel traffic. For simplicity, we assume that there is a fixed background traffic flow on the railroad links, $r_a, a \in B$, in addition to the biomass/biofuel rail freight flows.

In summary, the total link flow x_a for all $a \in A \cup B$ is the summation of the background (passenger or rail) flow and the biomass/biofuel shipment flows, that is,

$$x_a = \begin{cases} \sum_{o \in O^h} \sum_{d \in D^h} \sum_{k \in K^{o,d}} f_k^{o,d} \delta_{a,k}^{o,d} + \sum_{i \in I^s} \sum_{k \in K^{s,i}} f_k^{s,i} \delta_{a,k}^{s,i} \\ \quad + \sum_{i \in I^d} \sum_{k \in K^{d,i}} f_k^{d,i} \delta_{a,k}^{d,i} & \forall a \in A \\ r_a + \sum_{i \in I^s} \sum_{k \in K^{s,i}} f_k^{s,i} \delta_{a,k}^{s,i} + \sum_{i \in I^d} \sum_{k \in K^{d,i}} f_k^{d,i} \delta_{a,k}^{d,i} & \forall a \in B. \end{cases} \tag{1}$$

In support of the supply chain design, we consider the option of expanding highway link capacity; let decision variable $Z_a \in \{0, 1, 2, \ldots\}, \forall a \in A$ be the number of lanes added to link a, and each additional lane yields a known extra capacity of q_a. The capacity of link a after the capacity expansion is the summation of the original link capacity Q_a and the additional capacity; that is, $Q_a + Z_a q_a$. The travel time on link $a \in A$, denoted by $t_a(x_a, Z_a)$, is assumed to take the following BPR function form based on the traffic volume and expanded link capacity:

$$t_a(x_a, Z_a) = t_0 \left(1 + \alpha \left(\frac{x_a}{Q_a + Z_a . q_a} \right)^{\beta} \right), \quad \forall a \in A,$$

where constant parameters $\alpha = 0.15$ and $\beta = 4$ (Bureau of Public Roads, 1970). By assuming the BPR function for link travel times, our model does not address possible correlation of delay in sequential links (e.g., due to queue spillover). Other traffic delay models may be used to handle such issues.

We assume that if a biorefinery is built at location j, a new railroad link may be optionally built to connect the biorefinery to the main railway network such that inflow to and outflow from the biofuel refinery can use either local roadways or this railroad connector. Such new railroad segments will be dedicated to biomass/biofuel transportation; that is, the background traffic on them is zero. We let a_j denote the candidate railroad connector link for candidate refinery location $j \in J$. Since such rail segments often contain a single track, we use decision variable $Z'_{aj} \in \{0, 1\}, \ \forall j \in J$ to represent the number of tracks added for refinery j. The simulation results in Lai and Barkan (2009) suggest that the travel time on an existing railroad link depends on the traffic volume, as follows:

$$t'_a(x_a) = t'_{a0} + E_o e^{F_o x_a}, \quad \forall a \in B,$$

where t'_{a0} presents the free flow travel time on the link, and x_a is the train traffic volume per day. Constant parameters E_o and F_o depend on the traffic and operating conditions of a railroad subdivision. For candidate connector a_j, the link travel time is simply:

$$t'_{a_j}(x_{a_j}, Z'_{a_j}) = (E_o e^{F_o x_{a_j}} + t'_{a_j 0}) / Z'_{a_j}, \quad \forall j \in J.$$

Obviously, when $Z'_{a_j} = 0$, no connector is built and the travel time goes to infinity. For those transshipment links, the travel cost/time per unit flow is constant. The value is converted from the handling cost/delay for one transshipment.

The cost for highway link capacity expansion, $c_a(Z_a)$, $\forall a \in A$, can be expressed as the product of link length l_a, the additional capacity $q_a Z_a$, and a cost coefficient w (Unnikrishnan et al., 2009), that is,

$$c_a(Z_a) = w l_a q_a Z_a, \quad \forall a \in A.$$

Similarly, the investment for adding a railroad connector can be expressed as follows:

$$c'_{a_j}(Z'_{a_j}) = w' l'_{a_j} Z'_{a_j}, \quad \forall j \in J,$$

where w' denotes the cost coefficient and l'_{a_j} is the length of the railroad connector. The mathematical optimization model that integrates location, routing, and network design decisions can be expressed as follows:

$$\begin{aligned} \min & \sum_{j \in J} (m_j Y_j + c'_{a_j}(Z'_{a_j})) + \sum_{a \in A} c_a(Z_a) \\ & + \rho \Bigg(\sum_{a \in A} x_a t_a(x_a, Z_a) + \sum_{j \in J} x_{a_j} t'_{a_j}(x_{a_j}, Z'_{a_j}) \\ & + \sum_{a \in B} x_a t'_a(x_a) \Bigg) \end{aligned} \tag{2}$$

subject to (1), and

$$v_j^{\cdot} = \sum_{i \in I^{\cdot}} \sum_{k \in K^{\cdot,i}} f_k^{\cdot,i} \Delta_{j,k}^{\cdot,i} \quad \forall j \in J, \cdot \in \{s, d\} \tag{3}$$

$$h_i^{\cdot} = \sum_{k \in K^{\cdot,i}} f_k^{\cdot,i} \quad \forall i \in I^{\cdot}, \cdot \in \{s, d\} \tag{4}$$

$$v_j^s \le C_j Y_j \quad \forall j \in J \tag{5}$$

$$\theta v_j^s = v_j^d \quad \forall j \in J \tag{6}$$

$$Z'_{a_j} \le Y_j \quad \forall j \in J \tag{7}$$

$$\sum_{i \in I^s} h_i^s \le \sum_{j \in J} C_j Y_j \quad \forall i \in I^s, \forall j \in J \tag{8}$$

$$Y_j \in \{0, 1\}; Z_a \geq 0, \quad \text{integer} \quad \forall j \in J, \forall a \in A \quad (9)$$

$$Z'_{a_j} \in \{0, 1\} \quad \forall j \in J \quad (10)$$

$$f_k^{\cdot,i} \geq 0 \quad \forall i \in I^{\cdot}, k \in K^{\cdot,i}, \quad \cdot \in \{s, d\} \quad (11)$$

$$f_k^{o,d} \geq 0 \quad \forall o \in O, \forall d \in D, k \in K^{o,d} \quad (12)$$

The objective function (2) minimizes the total system cost which includes facility construction investment, infrastructure capacity expansion cost, costs for biomass and biofuel transportation, and public travel cost, respectively. Parameter ρ converts link travel time to travel cost and reflects a relative weight of total travel cost against the construction costs. This system optimal objective may be more suitable for centrally controlled shipment trucks than for the public traffic (e.g., Abdul Aziz and Ukkusuri, 2012). An alternative formulation with a user equilibrium objective for the public traffic is shown in the Appendix. Constraints (1) indicate that the traffic flow on each highway network link is the sum of the background traffic and the passenger car equivalent flow for biomass and biofuel shipment on highway and railroad networks. Constraints (3) ensure that the flow on each virtual link is the sum of the biomass or biofuel flow from all paths which contain node $j \in J$. Constraints (4) show that the sum of all biomass flows out of a biomass production location should be equal to the supply at that location and all biofuel shipment flows into a demand point should be equal to the demand at that point. Constraints (5) ensure that the throughput v_j^s can be any nonnegative value no greater than the capacity of the refinery at candidate location $j \in J$ (if there is a biorefinery at that node). Constraints (6) show the flow conservation at the refinery, indicating that the inbound biomass flow converts to the equivalent outbound biofuel flow at each biorefinery. Constraints (7) guarantee that railroad expansion may occur only when there is a facility at $j \in J$. Constraints (8) ensure that the total capacity of biorefineries should exceed the total biomass supply. Finally, constraints (9)–(12) define the binary and nonnegative variables.

3 SOLUTION APPROACH

The integrated mathematical model (1)–(12) involves nonlinearity and mixed integer variables, and hence it is very difficult to solve this model to exact optimality. To overcome this challenge, we propose a hybrid solution approach that integrates the genetic algorithm (GA) (Adeli and Hung, 1995; Adeli and Kumar, 1995; Adeli and Cheng, 1994; Kang et al., 2009), Lagrangian relaxation (Fisher, 1981), and traffic assignment algorithms (Frank and Wolfe, 1956; Sheffi, 1985).

3.1 The GA framework

The GA has been used to effectively solve transportation network design problems (e.g., Ukkusuri et al., 2006; Putha et al., 2012). It begins with a population of individuals that develops through generations based on the principle of "survival of the fittest." Interested readers are referred to Adeli and Hung (1995) for more details on GA.

The general framework for our GA approach is depicted in Figure 3. First, the basic parameter setting such as the size of population (n) and the probabilities of crossover and mutation must be initialized. We define a chromosome to be a binary vector representation of the integer variables $\{Y_j\}$, $\{Z_a\}$, and $\{Z'_{a_j}\}$, such that the length of the chromosome depends on $|J|$, $|A|$, the maximum number of lanes we may add to a roadway link, and the maximum number of tracks we may add to the railroad network in the neighborhood of biorefineries. If we only allow at most one lane addition to each link (i.e., $Z_a \in \{0, 1\}$, $\forall a \in A$), then the length of the chromosome is simply $2|J| + |A|$.

In the initialization step, chromosomes are randomly generated for the first population, and each chromosome should satisfy constraints (8) to ensure feasibility of the solution; that is, the total biomass supply should not exceed the total capacity of the biorefinery facilities. Each chromosome in the population contains information on the facility location and road expansion decisions. Based on this information we perform traffic assignment to determine near-optimal link traffic volumes on roadway network $\{x_a\}$. The detail of this step is explained in the following section. The fitness function that we use to evaluate and rank each chromosome is the inverse of the objective value.

GA creates chromosomes for new populations through a series of operations including selection, crossover, and mutation. The tournament selection technique is used to choose the chromosomes for later perturbations in crossover and mutation operations. The crossover uses a multipoint technique where cross points are randomly selected for each part of the chromosome (e.g., those corresponding to location decisions $\{Y_j\}$ and capacity expansion decisions $\{Z_a\}$). Then, a bitwise mutation is used, that is, each cell of the chromosome (i.e., gene) is randomly flipped according to the probability of mutation. All parts of the chromosome are mutated in the same way but the cells representing the location decision are never switched with the cells representing the capacity expansion decision. Again, any newly generated chromosome should

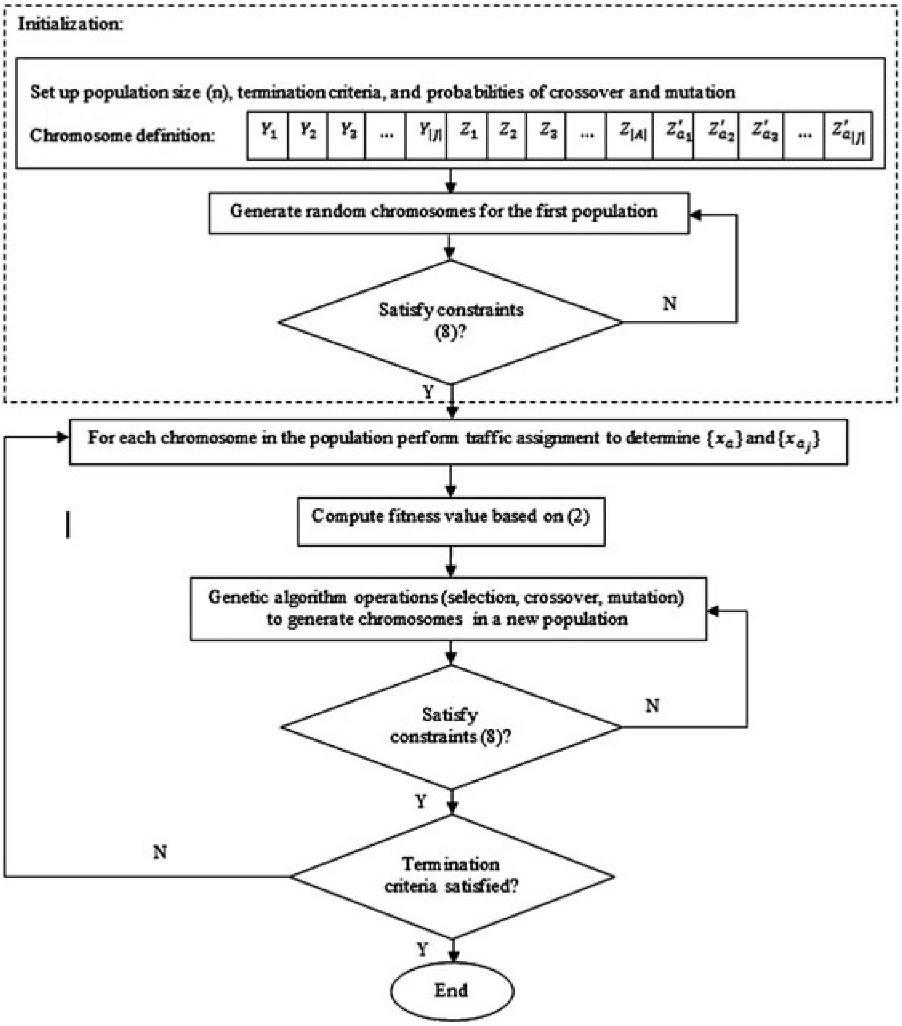

Fig. 3. General framework of the GA approach.

satisfy constraints (8) or otherwise it will be discarded. Finally, the algorithm terminates either when the predetermined maximum number of generations is reached, or if the best solution has not been improved over a certain number of consecutive generations. The best chromosome over all generations is recorded as the solution to the problem.

3.2 Routing decisions for a given chromosome

As discussed above, the facility location and roadway capacity expansion decisions are determined by the chromosomes in the GA framework. We use an embedded Lagrangian relaxation and convex combination algorithm to solve the remaining routing decisions.

For any given chromosome, the sets of decision variables $\{Y_j\}$, $\{Z_a\}$, and $\{Z'_a\}$ are known, therefore model (1)–(12) reduces to a simpler nonlinear program. Similar to Bai et al. (2011), we relax constraints (6) and the remaining problem can be expressed as follows:

$$\begin{aligned}
&\min \sum_{j\in J}(m_j Y_j + c'_{a_j}(Z'_{a_j})) + \sum_{a\in A} c_a(Z_a) \\
&\quad + \rho\left(\sum_{a\in A} x_a t_a(x_a, Z_a) + \sum_{j\in J} x_{a_j} t'_{a_j}(x_{a_j}, Z'_{a_j}) \right.\\
&\quad \left. + \sum_{a\in B} x_a t'_a(x_a)\right) + \sum_{j\in J}\mu_j\left(\theta v_j^s - v_j^d\right) \\
&= \min \sum_{j\in J}(m_j Y_j + c'_{a_j}(Z'_{a_j})) + \sum_{a\in A} c_a(Z_a) \\
&\quad + \sum_{a\in A} x_a[\rho t_a(x_a, Z_a)] + \sum_{j\in J} x_{a_j}\left[\rho t'_{a_j}(x_{a_j}, Z'_{a_j})\right]
\end{aligned}$$

$$+\sum_{a\in B} x_a[\rho t_a'(x_a)] + \sum_{j\in J}(\theta\mu_j)v_j^s$$
$$+\sum_{j\in J}(-\mu_j)v_j^d$$

subject to (1), (3)–(5), (7), and (9)–(12).

The first two terms in the objective function are constants (for a given chromosome). This relaxed problem becomes a traffic assignment problem for biomass and biofuel shipment flows, if the Lagrangian multiplier μ_j is interpreted as constant travel times on the virtual links (j, S^s) and (j, S^d). Then, the convex combination method can be applied within a Lagrangian relaxation framework to solve the routing problem, as shown in Bai et al. (2011).

4 CASE STUDY

The state of Illinois is expected to produce a large proportion of the nation's biofuel mandate for years 2007–2022. In our case study, each county in Illinois is considered as a biomass supply point and a biofuel consumption location. The Illinois dataset contains 98 nodes, which are mainly the centroid of the counties or most important intersection of interstate highways, which can serve both as the origins and the destinations of biomass and biofuel transportation. In addition, the network contains 374 links that are interstate highways and some local arterials. There are 20 candidate locations for biorefinery construction, selected based on socioeconomic factors such as land price, access to major transportation facilities, water availability, and agricultural activities. The supply of biomass (i.e., corn) and demand of ethanol by county are respectively calculated as a percentage share of the projected Illinois total supply and demand in year 2022 (Kang et al., 2010). For simplicity we have assumed that the total biomass supply equals the total demand. In case of unbalanced biomass supply and biofuel demand, a virtual node can be added to process the excessive supply or demand. Similar to Bai et al. (2011), the amounts of biomass and ethanol that a truck/railcar can carry are used to convert the shipment volume into the number of trucks/trains needed per year (and then per hour).

Furthermore, the maximum capacity of refineries in the Illinois case study is considered to be 300 million gallon per year and the annual fixed cost of constructing a biorefinery of this size will be \$ 27,000,000 (Kang et al., 2010). The fixed refinery construction cost is converted to the cost per hour assuming 20 years for serviceability and 260 days per year and 9 hours per day as effective working times for biomass and biofuel transportation (Bai et al., 2011). The transportation mode is both full-truck-load (Kang et al., 2010) and freight train shipping, and the traffic flow is converted to passenger car equivalent (PCE) per hour.

For comparison with the results in Bai et al. (2011), we only consider the routing of biomass and biofuel shipments in the multimodal network. Information on the hourly passenger traffic flow and the existing capacity of the 374 highway links are obtained from the National Transportation Atlas Database (RITA, 2008) and Illinois Department of Transportation (IDOT, 2007). The capacities of the interstate highways and local arterials are assumed to be 2,200 and 1,700 (pcphpl), respectively (HCM, 2000). We set $\rho = 20\,(\$/\text{hr} - \text{PCE})$ (Bai et al., 2011) and the cost coefficient factor for capacity expansion is $w = 1{,}000\,(\$/\text{lane} - \text{mile})$ (Unnikrishnan et al., 2009). Furthermore, we assumed that, realistically, at most one highway lane will be added to each link if road capacity expansion is required ($Z_a = 1$).

The Illinois railroad network contains 150 links, including those from all the class-I railroads. Since the agricultural supply infrastructure has been quite established in Illinois, we assume that the biomass production points already have access to either local tracks or highway transportation; and, there is no need to consider railway expansion in the neighborhood of biomass production farms. The information on background traffic volume per day and the railway network as well as the free flow speed on each segment are acquired from the railroad industry (Source: Norfolk Southern, Illinois Division, Timetable Number 1). The cost rate of constructing railroad segments $w' = 78{,}000\,(\$/\text{track} - \text{mile})$ (Source: http://tacnet.missouri.org/history/railroads/rrcosts.html).

4.1 Results and discussion

This proposed algorithm is coded in Visual C++ and run on a desktop computer with 2.67 GHz CPU and 2.00 GB memory. In the GA framework, the selection pressure is chosen to be 20, the population size 100, probability of crossover 0.8, probability of mutation 0.1, chromosome length 414, and random seed value 0.025. The program terminates when the best fitness value does not improve across a number of generations. For all numerical cases, the GA converges within 150 generations, taking less than 30 minutes of CPU time.

We consider the integrated biofuel supply chain design in different scenarios: (1) highway shipment only with no road expansion, (2) highway shipment and highway expansion only, (3) multimodal shipments with highway expansion only, (4) multimodal shipments with highway/railroad expansions, and finally (5) multimodal shipments with railroad expansion only. Since a 20-year life cycle is assumed for both refineries

Table 1
Comparison of near-optimal solutions in different scenarios

	Scenario 1 (Benchmark case)	*Scenario 2*	*Scenario 3*	*Scenario 4*	*Scenario 5*
Number of refineries	10	10	7	7	8
Location of refineries	2, 3 , 4, 8, 10, 12, 13, 14, 16, 18	2, 3, 4, 8, 10, 12, 14, 15, 16, 18	2, 4, 8, 9, 11, 12, 14	2, 4, 8, 9, 11, 12, 14	2, 7, 10, 11, 14, 15, 16, 17
Total cost for refineries ($)	5.400×10^9	5.400×10^9	3.780×10^9	3.780×10^9	4.320×10^9
Number of added lanes	–	74	46	28	–
Total cost for highway expansion ($)	–	5.884×10^9	3.726×10^9	2.281×10^9	–
Number of added railroad segments	–	–	–	4	–
Total cost for railroad expansion ($)	–	–	–	9.360×10^5	–
Total transportation cost ($)	6.711×10^{11}	3.994×10^{11}	3.805×10^{11}	3.781×10^{11}	3.893×10^{11}
Total system cost ($)	6.765×10^{11}	4.106×10^{11}	3.880×10^{11}	3.842×10^{11}	3.936×10^{11}

and roads, all costs reported here are 20-year totals. The first scenario is essentially the same problem in Bai et al. (2011), which only considers facility location and traffic assignment. The algorithm finds an objective value very consistent with the near-optimal results in Bai et al. (2011), see Table 1. This suggests that the proposed algorithm works very effectively and the solution quality is acceptable.

The second scenario considers additional capacity expansion decisions which turn out to influence the transportation cost significantly. It implies that highway capacity expansion is very effective in mitigating congestion, probably due to the high background traffic volume on the network. The third scenario takes advantage of intermodal shipments as well as highway capacity expansion, and the fourth scenario adds the option of railway expansion. The option of railroad shipments reduces the transportation cost to a great extent. Figure 4 depicts the refinery locations and the roadway expansions in this scenario. Finally, the fifth scenario considers multimodal transportation and expanding railroad segments only. It shows the benefit of using the cheaper transportation mode when the highway congestion increases, and therefore avoiding huge highway investments to facilitate the shipments.

In summary, it is observed that the capacity expansion in scenario 2 has contributed to significant cost reduction, as compared with scenario 1. It can also be seen from Table 1 that although road expansion activities might require additional infrastructure investments, the total system cost has been significantly reduced in scenario 2 (mainly due to the reduction in transportation cost). In addition, incorporation of railroad shipments (scenario 5), even without highway capacity expansion, has decreased the total cost considerably compared to the first scenario. Incorporation of railroad shipments has decreased the excessive number of lane additions in the highway network (compared to the ones in scenario 2). Scenario 4 also benefits from significant cost reduction due to railroad network expansion.

4.2 Sensitivity analysis

We also perform a series of sensitivity analyses to see how the values of the roadway expansion cost coefficient $w \in \{500, 1000, 1500, 2000\}$ and traveler time value $\rho \in \{0, 10, 20, 40\}$ influences the near-optimal solution. The results are shown in Figure 5.

A few interesting observations can be made. When w is zero, the expansion of roadway capacity bears no cost, and therefore, most of the roadway links tend to be expanded to enhance the traffic condition. Thus, the transportation cost in this case is much smaller than those in other cases. Increases in w make capacity expansion less

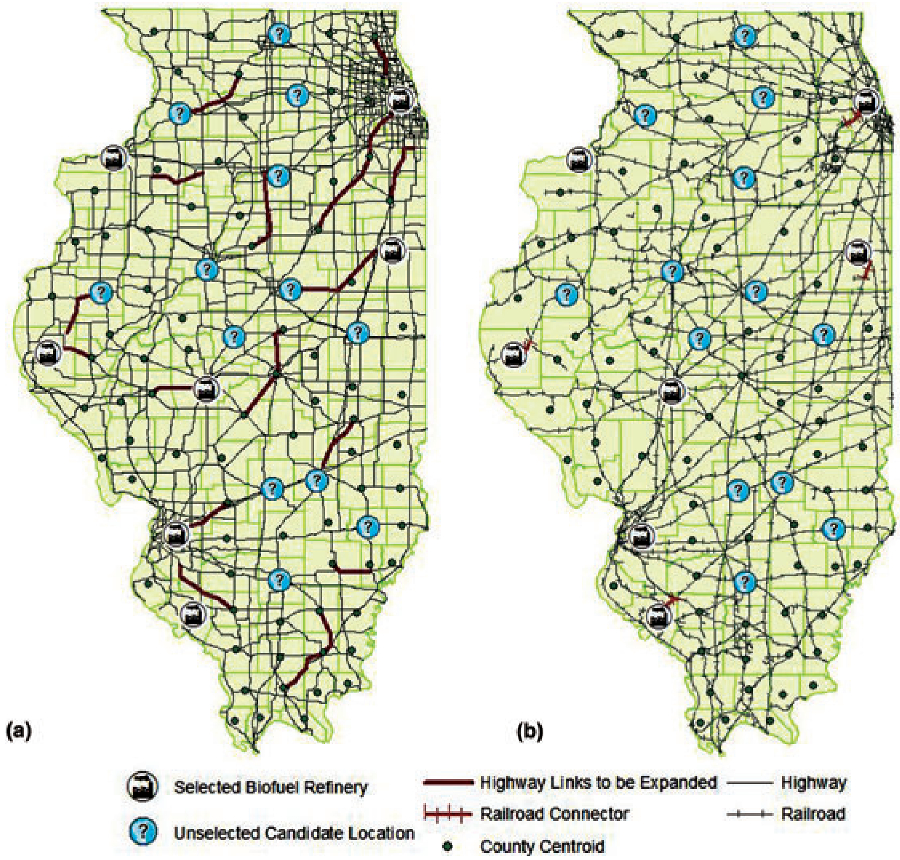

Fig. 4. Biorefinery locations and roadway capacity expansion (scenario 4): (a) highway expansions and (b) railroad expansions.

favorable, and this leads to an increase in transportation cost accordingly.

For a constant w, when ρ is small, the transportation cost is low and hence the number of biorefinery facilities is small (because shipping cost over long distance is cheap), and there is little incentive intention to perform road expansion. When the value of ρ increases, the transportation cost increases significantly, leading to more biorefineries in order to reduce the shipment distances.

Finally, sensitivity analysis is conducted to study how the solution quality and computation time depend on the choice of five GA parameters (i.e., random seeds, population size, mutation probability, crossover probability, selection pressure). One parameter value is perturbed at a time, where random seed value $\in\{0.025, 0.2, 0.5, 0.9\}$, population size $\in\{20, 40\}$, mutation probability $\in\{0.01, 0.03\}$, crossover probability $\in\{0.55, 0.8\}$, and selection pressure $\in\{2, 4, 6\}$. For each tested parameter combination, the best objective value and the computation time needed to achieve that solution are compared with those of the benchmark case (as reported in Table 1). Our computational results show that while the GA parameter values do slightly affect the number of iterations needed to achieve convergence (all within 30 minutes of CPU time), they virtually have no impact on the final objective value at convergence (all within 0.13% difference from the benchmark). Hence, we do not need to be very selective on the GA parameters.

5 CONCLUSION

This study presented an integrated mathematical model for biofuel supply chain design where the near-optimum number and location of biorefinery facilities, the near-optimal multimodal routing of biomass and biofuel shipment, and possible roadway capacity expansion are determined. The objective was to minimize the total cost for facility construction, roadway capacity expansion (including highway links and railway segments), and transportation (including both biomass/biofuel shipment and public travel). The congestion pattern and total transportation costs are determined based on traffic equilibrium flows on shipment routes under

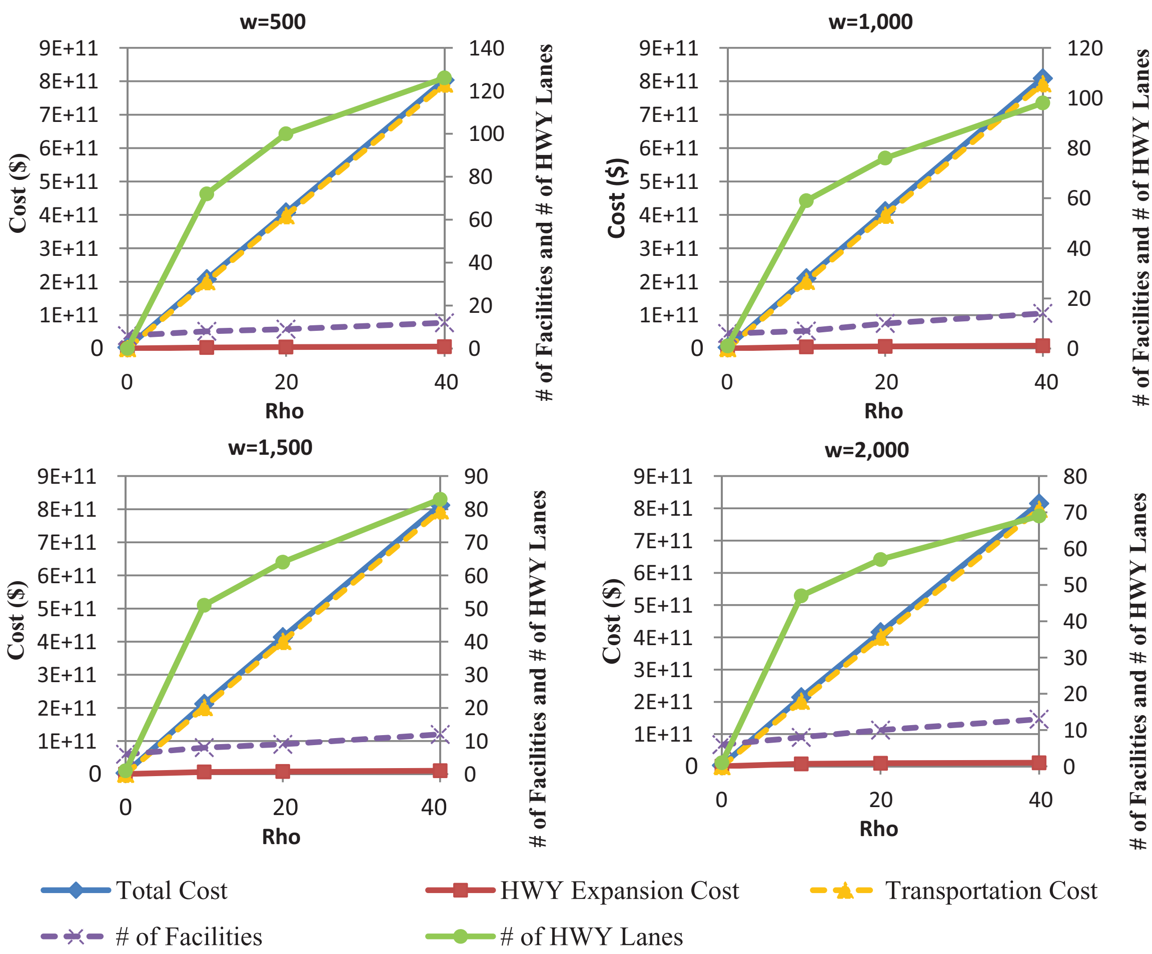

Fig. 5. Sensitivity analysis with regard to ρ and w.

expanded capacity. While the study has been motivated by the rapidly expanding biofuel supply chains, we shall note that the proposed mathematical model and solution techniques can be applied to other multicommodity flow optimization problems that simultaneously involve network traffic equilibrium, infrastructure expansion, and location choices (which determine the origin/destination of multicommodity flow).

To find the near-optimum solution to the proposed model, we have developed a hybrid GA framework that incorporates Lagrangian relaxation and convex combination algorithms. A real-world case study for the state of Illinois is conducted. The computational results show that the proposed algorithm is able to solve the problem efficiently. In addition, a series of sensitivity analyses are performed to draw managerial insights into the effects of roadway expansion cost coefficient (w) and traveler time value (ρ) on near-optimal supply chain design. Sensitivity analysis also shows that the solution at convergence is quite independent of the choice of GA parameters.

This research has assumed that the capital investment associated with roadway expansion is considered as part of the biofuel industry's investment plan. This is possible under potential public–private partnerships (Unnikrishnan et al., 2009). It may be interesting to expand our current framework to allow multiple stakeholders (e.g., public agency, biofuel industry, railroad company, and public travelers) to have independent or conflicting objectives. This will probably require multilevel programs with equilibrium constraints. In addition, as we observe a significant impact of roadway capacity expansion on the total system cost (see Table 1), it would be interesting to study in the future how roadway expansion budget would affect the optimal solution. Furthermore, by assuming the BPR function for link travel times, our current model does not address possible correlation of delay in sequential links (e.g., due to queue spillover). Other traffic delay models (e.g., cell transmission) may be used to handle such issues. Currently, our genetic algorithm randomly conducts chromosome operations (e.g., selecting some links to be widened), which may not be efficient because such operations may not take full advantage of the underlying network structure. It would be interesting to address possible correlation of sequential links in the future by including more chromosome operation rules (e.g., selecting a block of links) in our GA structure. The alternative formulation with a user equilibrium objective for the general road users (e.g., the one in the Appendix) leads to additional model complexity; efforts are being made to develop efficient algorithms to solve such problems.

ACKNOWLEDGMENTS

This research was supported in part by the U.S. National Science Foundation through Grants EFRI-RESIN #0835982 and CMMI #0748067. The first author would like to thank Yun Bai (University of Illinois at Urbana-Champaign) for providing the Illinois dataset and for helping with the traffic assignment model.

REFERENCES

Abdul Aziz, H. M. & Ukkusuri, S. V. (2012), Integration of environmental objectives in a system optimal dynamic traffic assignment model, *Computer-Aided Civil and Infrastructure Engineering*, **27**(7), 494–511.

Adeli, H. & Cheng, N.-T. (1994), Augmented Lagrangian genetic algorithm for structural optimization, *Journal of Aerospace Engineering, ASCE*, **7**(1), 104–18.

Adeli, H. & Hung, S. L. (1995), *Machine Learning–Neural Networks, Genetic Algorithms, and Fuzzy Sets*, John Wiley and Sons, New York.

Adeli, H. & Kumar, S. (1995), Distributed genetic algorithms for structural optimization, *Journal of Aerospace Engineering*, **8**(3), 156–63.

Bai, Y., Hwang, T., Kang, S. & Ouyang, Y. (2011), Biofuel refinery location and supply chain planning under traffic congestion. *Transportation Research Part B*, **45**(1), 162–75.

Brown, R., Orwig, E., Nemeth, J. & Subietta Rocha, C. (2007), *Economic Potential for Ethanol Expansion in Illinois*, Illinois Institute for Rural Affairs at Western Illinois University, Macomb, IL.

Bureau of Public Roads (1970), Federal Highway administration, U.S. D.O.T, Urban Transportation Planning, General Information and Introduction to System 360.

Eathington, L. & Swenson, D. A. (2007), *Dude, Where's My Corn? Constraints on the Location of Ethanol Production in the Corn Belt*. Presented at the 46th Annual Meeting of the Southern Regional Science, Association, Charleston, SC. March 29–31, 2007.

English, B. C., Short, C. & Heady, E. O. (1981), The economic feasibility of crop residues as an auxiliary fuel in coal-fired power plants, *American Journal of Agricultural Economics*, **63**(4), 136–44.

Fisher, M. L. (1981), The Lagrangian relaxation method for solving integer programming problems. *Management Science*, **27**(1), 1–18.

Frank, M. & Wolfe, P. (1956), An algorithm for quadratic programming. *Naval Research Logistics Quarterly*, **3**, 95–110.

Gao, L., Xie, C., Zhang, Z. & Waller, S. T. (2012), Network-level road pavement maintenance and rehabilitation scheduling for optimal performance improvement and budget utilization, *Computer-Aided Civil and Infrastructure Engineering*, **27**(4), 276–87.

Hakan, A. M., Temel, O. & Kuban, A. (2012), Efficient approximate solution methods for the multi-commodity capacitated multi-facility Weber problem, *Computers and Operations Research*, **39**(2), 225–37.

Highway Capacity Manual (HCM) (2000), *Transportation Research Board*. Washington, DC, ISBN: 0–309-06681–6.

IDOT, 2007. Illinois Department of Transportation (IDOT) Statistical Maps of Illinois-Average Annual Daily Traffic Maps. Available at: http://www.dot.il.gov/trafficmaps/table.htm, accessed August 8, 2009.

Jen, F. C., Pegels, C. C., & Dupuis, T. M. (1968), Optimal capacities of production facilities, *Management Science*, **14**(10), B-573–B-580.

Kang, M. W., Schonfeld, P. & Yang, N. (2009), Prescreening and repairing in a genetic algorithm for highway alignment optimization, *Computer-Aided Civil and Infrastructure Engineering*, **24**(2), 109–19.

Kang, S., Önal, H., Ouyang, Y., Scheffran, J. & Tursun, D. (2010), Optimizing the biofuels infrastructure: transportation networks and biorefinery locations in Illinois, in M. Khanna et al. (eds.), *Handbook of Bioenergy Economics and Policy*, *Series: Natural Resource Management and Policy*, Vol. 33, Springer Academic Publishers, New York.

Klose, A. & Drexl, A. (2005), Facility location models for distribution system design, *European Journal of Operational Research*, **162**(1), 4–29.

Lai, Y. C. & Barkan, C. P. L. (2009), An enhanced parametric railway capacity evaluation tool (RCET), Transportation Research Record, Transportation Research Board of the National Academies, Washington DC, *Journal of the Transportation Research Board*, **2117**, 33–40.

Li, Z., Madanu, S., Wang, Y., Abbas, M. & Zhou, B. (2010), A heuristic approach for selecting highway investment alternatives, *Computer-Aided Civil and Infrastructure Engineering*, **25**(6), 427–39.

Lin, D. Y. (2011), A dual variable approximation-based descent method for a bi-level continuous dynamic network design problem, *Computer-Aided Civil and Infrastructure Engineering*, **26**(8), 581–94.

López, E. & Monzón, A. (2010), Integration of sustainability issues in strategic transportation planning: a multi-criteria model for the assessment of transport infrastructure plans, *Computer-Aided Civil and Infrastructure Engineering*, **25**(6), 440–51.

Mahmudi, H. & Flynn, P. C. (2006), Rail vs. truck transport of biomas. *Applied Biochemistry and Biotechnology*, **129**(1–3), 88–103.

Mapemba, L. D. (2005), Cost to deliver lignocellulosic biomass to biorefinery. Ph.D. dissertation, The Oklahoma State University, Stillwater, OK.

Metropolitan Transportation Plan (MTP 2035) (2006), Road expansion, Sacramento region, Blueprint for Sustainable Communities 2006. Available at: http://www.sacog.org/2035, accessed June 7, 2010.

Ng, M. W., Lin, D. Y. & Waller, S. T. (2009), Optimal long-term infrastructure maintenance planning accounting for

traffic dynamics, *Computer-Aided Civil and Infrastructure Engineering*, **24**(7), 459–69.
Nie, Y. & Zhang, H. M. (2008), A variational inequality formulation for inferring dynamic origin-destination travel demands. *Transportation Research Part B*, **42**(2008), 635–62.
Nienow, S., McNamara, K. T. & Gillespie, A. R. (2000), Assessing plantation biomass for co-firing with coal in northern Indiana: a linear programming approach. *Biomass Bioenergy*, **18**, 125–35.
Ouyang, Y. (2007), Pavement resurfacing planning on highway networks: a parametric policy iteration approach. *ASCE Journal of Infrastructure Systems*, **13**(1), 65–71.
Putha, R., Quadrifoglio, L. & Zechman, E. (2012), Comparing ant colony optimization and genetic algorithm approaches for solving traffic signal coordination under oversaturation conditions, *Computer-Aided Civil and Infrastructure Engineering*, **27**(1), 14–28.
RITA (Research and Innovative Technology Administration) (2008), National Transportation Atlas Database 2008. Available at: http://www.bts.gov/, accessed October 7, 2009.
Searcy, E., Flynn, P., Ghafoori, E. & Kumar, A. (2007), The relative cost of biomass energy transport. *Applied Biochemistry and Biotechnology*, **137–140**(1–12), 639–52.
Sheffi, Y. (1985), *Urban Transportation Networks: Equilibrium Analysis with Mathematical Programming Methods*. Prentice Hall, Englewood Cliffs, NJ.
Sokhansanj, S., Kumar, A. & Turhollow, A. F. (2006), Development and implementation of integrated biomass supply analysis and logistics model (IBSAL), *Biomass and Bioenergy*, **30**(10), 838–47.
Srimantula, R. (2004), Study on capacity of railroad network and airport terminals for upper midwest freight corridor, Master thesis, Civil Engineering, University of Toledo, OH.
Tembo, G., Epplin, F. M. & Huhnke, R. L. (2003), Integrative investment appraisal of a lignocellulosic biomass-to-ethanol industry. *Journal of Agriculture and Resource Economics*, **28**, 611–33.
Tursun, D., Kang, S., Onal, H., Ouyang, Y., Scheffran, J. (2008), Optimum biorefinery locations and transportation network for the future biofuels industry in Illinois, in M. Khanna, (ed.), *Transition to a Bioeconomy: Environmental and Rural Development Impacts, Proceedings of Farm Foundation/USDA Conference*, St. Louis, MI, October 15–16, 2008, Farm Foundation.
Unnikrishnan, A. & Lin, D. Y. (2012), User equilibrium with recourse: continuous network design problem, *Computer-Aided Civil and Infrastructure Engineering*, **27**(7), 512–24.
Unnikrishnan, A., Valsaraj, V. & Damnjanovic, I. (2009), Design and management strategies for mixed public private transportation networks: a meta-heuristic approach. *Computer Aided Civil and Infrastructure Engineering*, **24**, 266–79.
Ukkusuri, S. V., Mathew, T. V., & Waller, S. T. (2006), Robust transportation network design under demand uncertainty, *Computer Aided Civil and Infrastructure Engineering*, **22**(1), 6–18.

APPENDIX: ALTERNATIVE FORMULATION WITH USER EQUILIBRIUM FOR BACKGROUND TRAFFIC

In this article, the congestion pattern and total transportation costs have been computed based on system optimal flows on shipment routes. In reality, user equilibrium flows could also be incorporated into the formulation to address alternative route choices of the background traffic b_a. To this end, formulation (1)–(12) can be rewritten into a mathematical program with equilibrium constraints (MPEC), as follows.

$$\min \sum_{j\in J}(m_j Y_j + c'_{a_j}(Z'_{a_j})) + \sum_{a\in A} c_a(Z_a) + \rho \left(\sum_{a\in A} x_a t_a(x_a, Z_a) + \sum_{j\in J} x_{a_j} t'_{a_j}(x_{a_j}, Z'_{a_j}) + \sum_{a\in B} x_a t'_a(x_a)\right)$$

subject to (1), (3)–(11), and

$$x'_a = \sum_{i\in I^s}\sum_{k\in K^{s,i}} f_k^{s,i}\delta_{a,k}^{s,i} + \sum_{i\in I^d}\sum_{k\in K^{d,i}} f_k^{d,i}\delta_{a,k}^{d,i} \quad \forall a \in A$$

and

$$b_a \in \operatorname{argmin}_{b_a} \sum_{a\in A}\int_{x'_a}^{x'_a+b_a} t_a(\omega, Z_a)\, d\omega$$

subject to (12) and

$$b_a = \sum_{o\in O^h}\sum_{d\in D^h}\sum_{k\in K^{o,d}} f_k^{o,d}\delta_{a,k}^{o,d} \quad \forall a \in A$$

$$\sum_{k\in K^{o,d}} f_k^{o,d} = q^{od} \quad \forall o \in O, \forall d \in D$$

The above MPEC model is generally difficult to solve. However, the literature has repeatedly shown that such MPEC models can often be effectively tackled by the GA approach (e.g., Unnikrishnan and Lin, 2012; Lin, 2011). In such a framework, once the $\{Y_j\}$, $\{Z_a\}$, and $\{Z'_a\}$ variables are fixed, the remaining problem becomes a multi-class mixed traffic assignment problem, which can be solved effectively. Interested readers are referred to (Nie and Zhang, 2008) for more detailed information.