GIS Based Two-stage Stochastic Facility Location Considering Planting Plan Uncertainty

Ou Sun

Systems and Industrial Engineering
University of Arizona

ExxonMobil On-site Interview
Jan 16, 2019

Outline

Outline

Network Optimization in Power Grid

(1) Subgraph connectivity by MIP
Many combinatorial optimization problems involve finding a connected subgraph embedded in a larger graph. We study the general problem as solving connectivity constraints of undetermined subgraphs by different MIP formulations.

(2) Multistage sensor placement
Sensor placement on a power grid in a multistage framework constrained by a budget limitation. To improve computational efficiency, the problem is converted to a multistage maximum flow problem.

(3) Probabilistic and reliable connected sensor networks
Each bus in the power system is observed with some reliability level (chance-constrained) under line/sensor contingencies. The reliability of the connectivity of the sensor network is also guaranteed.

Optimization in Other Networks

(4) Social networks

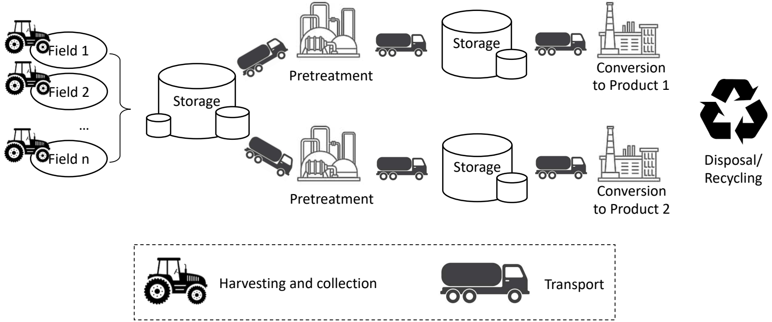
Socially influenced shortest path problem: Social networks are integrated into the traditional shortest path problem and used to model and predict personal decisions in daily life. The problem is modeled as a bilevel MIP.

(5) Transportation networks

GIS-based two-stage stochastic facility location problem: GIS analysis is applied to find potential facility location candidates. Supply uncertainty is also considered and modeled by stochastic programming.

Outline

Biomass Supply Chain



- A typical biomass supply chain consists of harvesting and collection, storage, transportation, pretreatment, and conversion

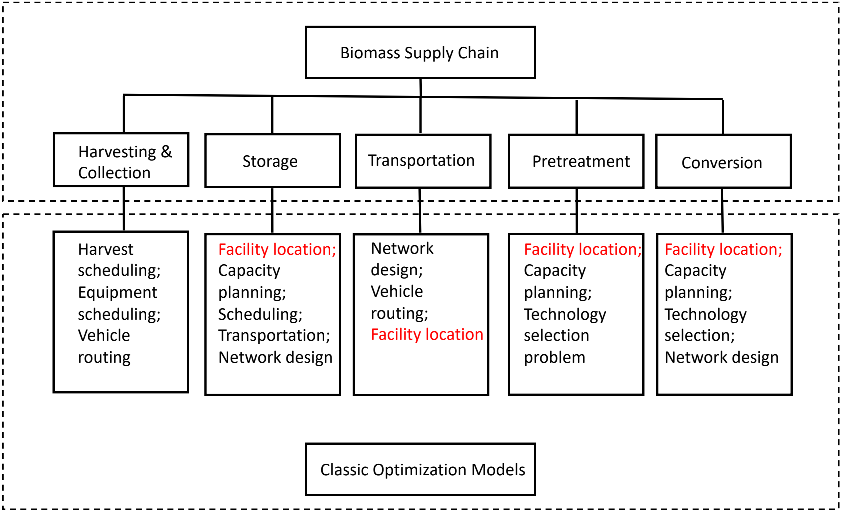
Biomass Supply Chain

THE UNIVERSITY OF ARIZONA

- Motivations of researches on biomass supply chain
  (1) the urgent desire to relieve the pressure of dependence on petroleum in energy industry
  (2) the growing demand for protection of our environment
  (3) biomass is a good replacement of fossil fuels

- Characteristics of a biomass supply chain
  - spacial fragmentation and dispersed distribution of biomass $\rightarrow$ complicated design and management
  - high moisture and low bulk density $\rightarrow$ high transportation and handling cost
  - uncertainties: biomass supply (seasonality, unpredictable climates) , demand, prices and costs.

Facility Location Problem

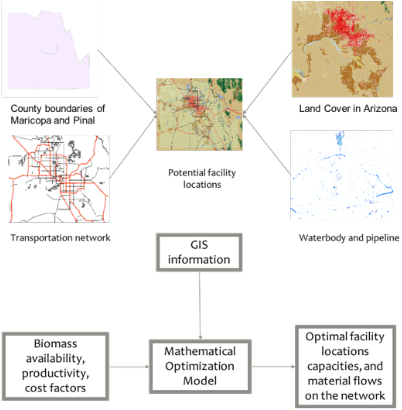


Facility Location Problem

- Importance of facility location problem
  - facility location decisions require large investment
  - vital to the feedstock logistics
  - long-term influence on accessibility to biomass and other materials
  - decisions affect not only costs but also the income
  - regarding costs, facility location affects a great variety of them:
    - land costs
    - labor costs
    - raw materials
    - transportation and logistics
- Location alternatives
  - expansion of an existent facility
  - start a new facility in a new area
  - shut down of a facility

Methodology Overview

THE UNIVERSITY OF ARIZONA



- A decision framework based on GIS information for facility location
  - GIS information as input for the mathematical model
  - two stage mixed integer stochastic programming
  - consider the uncertainty from the planting plan/ the availability of the biomass supply
  - output of the model is the decisions on optimal facility locations, the capacity for each facility, the material flows on the constructed network

Highlights of the Work

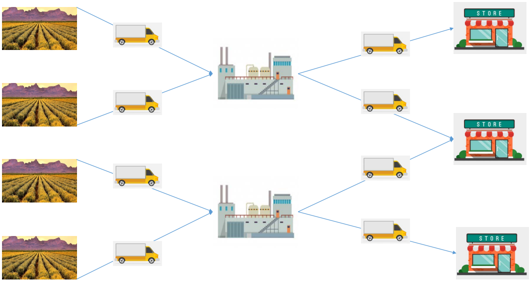
- The goal is to find conversion facility locations which connect suppliers and customers in order to minimize the system cost.
- GIS is integrated into a optimization model to identify potential candidates.
- A two-stage stochastic programming model is formulated to capture the uncertainty coming from supply.
- Different from uncertain sources of supply in references, we consider the planting plan uncertainty.

Outline

Variables

- Parameters
  - c^h (\$/metric tonne): unit harvesting/hauling cost of biomass
  - c^l(\$/metric tonne): unit loading and unloading cost of biomass
  - c^{t_1} (\$/metric tonne*km): unit transportation cost of biomass
  - $c^{t_2,q}$ (\$/metric tonne*kilometer): unit transportation cost of product type q
  - f_j (\$): fixed cost of building a facility at location $j \in J$
  - $Yield$ (metric tonne/acre): yield of biomass
  - A_i (acre): area of supply field $i \in I$
  - S_i (metric tonne): the productivity of each field i, e.g. $S_i = Yield \times A_i$
  - Min_j (metric tonne): the minimum capacity of the facility at location j
  - Max_j (metric tonne): the maximum capacity of the facility at location j
  - p: the planned maximum number of facilities
  - θ_q: the conversion rate from biomass (metric tonne) to product (metric tonne) $q \in Q$
  - D_k^q (metric tonne): the demand at location $k \in K$ for product q per year
  - d_{ij} (km): the distance between location i and location j

Variables



- Decision Variables
  - $x_j \in \{0, 1\}$: whether to locate a facility at location j or not
  - $cap_j \geq 0$: capacity of biomass supply for facility location j
  - $y_{ij} \geq 0$: commodity flow of biomass from supply i to facility j
  - $y_{jk}^q \geq 0$: commodity flow of product q from facility j to demand k

A deterministic Model

$$
\begin{aligned}
\min \; & \sum_{j\in J} f_j x_j + (c^h + c^l)\sum_{i\in I}\sum_{j\in J} y_{ij} + c^{t_1}\sum_{i\in I}\sum_{j\in J} d_{ij}y_{ij} + \sum_{q\in Q}\sum_{j\in J}\sum_{k\in K} c^{t_2,q} d_{jk} y_{jk}^q \\
s.t. \; & \sum_{j\in J} y_{ij} \le S_i, \forall i \in I \\
& \sum_{i\in I} y_{ij} = cap_j, \forall j \in J \\
& y_{ij} \le Mx_j, \forall i \in I, \forall j \in J \\
& y_{jk}^q \le Mx_j, \forall j \in J, \forall k \in K, \forall q \in Q \\
& Min_j x_j \le cap_j \le Max_j x_j, \forall j \in J \\
& \sum_{i\in I} y_{ij} \times \theta_q = \sum_{k\in K} y_{jk}^q, \forall j \in J, \forall q \in Q \\
& \sum_{j\in J} y_{jk}^q \ge D_k^q, \forall k \in K, \forall q \in Q \\
& \sum_{j\in J} x_j \le p \\
& x_j \in \{0,1\}, cap_j \ge 0, y_{ij}, y_{jk}^q \ge 0
\end{aligned}
$$

A Stochastic Model-First Stage

- In the first stage, strategic decisions of the facility locations and capacities are made.
- $\omega \in \Omega$: random vector of supplies in different scenarios
- We assume that there are finite number of outcomes with known probabilities of $\widetilde{\omega}$
- $p(\omega)$: the probability of occurrence for $\omega \in \Omega$

$$
\begin{aligned}
\min \ & \sum_{j \in J} f_j x_j + \sum_{\omega \in \Omega} p(\omega) Q_\omega(x) \\
s.t. \ & Min_j x_j \leq cap_j \leq Max_j x_j, \forall j \in J \\
& \sum_{j \in J} x_j \leq p \\
& x_j \in \{0,1\}, cap_j \geq 0
\end{aligned}
$$

Second Stage

- In the second stage, based on the first stage decisions, decide on the operational decisions of commodity flows
- For one realization of $\omega \in \Omega$

$$\mathbf{Q}_\omega(\mathbf{x}, \mathbf{Cap}) = \min (c^h + c^l) \sum_{i\in I} \sum_{j\in J} y_{ij}^{(\omega)} + c^{t_1} \sum_{i\in I} \sum_{j\in J} d_{ij} y_{ij}^{(\omega)} + \sum_{q\in Q} \sum_{j\in J} \sum_{k\in K} c^{t_2,q} d_{jk} y_{jk}^{q,(\omega)}$$

$$\begin{aligned}
s.t.\ & \sum_{j\in J} y_{ij}^{(\omega)} \le S_i^\omega, \forall i \in I \\
& \sum_{i\in I} y_{ij}^{(\omega)} = cap_j, \forall j \in J \\
& y_{ij}^{(\omega)} \le Mx_j, \forall i \in I, \forall j \in J \\
& y_{jk}^{q,(\omega)} \le Mx_j, \forall j \in J, \forall k \in K, \forall q \in Q \\
& \sum_{i\in I} y_{ij}^{(\omega)} \times \theta_q = \sum_{k\in K} y_{jk}^{q,(\omega)}, \forall j \in J, \forall q \in Q \\
& \sum_{j\in J} y_{jk}^{q,(\omega)} \ge D_k^q, \forall k \in K, \forall q \in Q \\
& y_{ij}^{(\omega)}, y_{jk}^{q,(\omega)} \ge 0
\end{aligned}$$

Outline

Benders Decomposition - Single Cut Version

- The binary variables only appear in the first stage, and the second stage has only continuous decision variables.
- The ordinary Benders Decomposition algorithm can be applied well in this problem.
- When applying Benders Decomposition, we assume the first stage problem as the master problem and the second stage problem as the subproblem.
- The second stage problem is a linear program and there is no duality gap of this problem.

Benders Decomposition - Single Cut Version

THE UNIVERSITY OF ARIZONA

The dual of the second stage problem is:

$$
\begin{aligned}
(\mathbf{DSP})\ \eta_\omega(\mathbf{x},\mathbf{Cap}) = \max\ & \sum_{i\in I} S_i^\omega \alpha_i^{(\omega)} + \sum_{j\in J} cap_j \beta_j^{(\omega)} + \sum_{i\in I}\sum_{j\in J} Mx_j \gamma_{ij}^{(\omega)} + \sum_{j\in J}\sum_{k\in K}\sum_{q\in Q} Mx_j \delta_{jk}^{q,(\omega)} + \sum_{k\in K}\sum_{q\in Q} D_k^q \mu_k^{q,(\omega)} \\
& \alpha_i^{(\omega)} + \beta_j^{(\omega)} + \gamma_{ij}^{(\omega)} + \sum_{q\in Q} \theta_q \lambda_j^{q,(\omega)} \le c^h + c^{t1} d_{ij}, \forall i \in I, j \in J \\
& \delta_{jk}^{q,(\omega)} - \lambda_j^{q,(\omega)} + \mu_k^{q,(\omega)} \le c^{t2,q} d_{jk}, \forall j \in J, k \in K, \forall q \in Q \\
& \alpha_i^{(\omega)}, \gamma_{ij}^{(\omega)}, \delta_{jk}^{q,(\omega)} \le 0, \beta_j^{(\omega)}, \lambda_j^{q,(\omega)} urs, \mu_k^{q,(\omega)} \ge 0
\end{aligned}
$$

The Benders reformulation is:

$$
\begin{aligned}
\min\ & \sum_{j\in J} f_j x_j + \eta \\
s.t.\ & Min_j x_j \le cap_j \le Max_j x_j, \forall j \in J \\
& \sum_{j\in J} x_j \le p \\
& \eta \ge \sum_{\omega\in\Omega} p(\omega)\Big(\sum_{i\in I} S_i^\omega \alpha_i^{\omega,m} + \sum_{j\in J} cap_j \beta_j^{\omega,m} + \sum_{i\in I}\sum_{j\in J} Mx_j \gamma_{ij}^{\omega,m} + \sum_{j\in J}\sum_{k\in K}\sum_{q\in Q} Mx_j \delta_{jk}^{q,\omega,m} + \sum_{k\in K}\sum_{q\in Q} D_k^q \mu_k^{q,\omega,m}\Big), \forall m \in M_\omega \\
& 0 \ge \sum_{i\in I} S_i^\omega \alpha_i^{\omega,n} + \sum_{j\in J} cap_j \beta_j^{\omega,n} + \sum_{i\in I}\sum_{j\in J} Mx_j \gamma_{ij}^{\omega,n} + \sum_{j\in J}\sum_{k\in K}\sum_{q\in Q} Mx_j \delta_{jk}^{q,\omega,n} + \sum_{k\in K}\sum_{q\in Q} D_k^q \mu_k^{q,\omega,n}, \forall n \in N_\omega, \omega \in \Omega \\
& x_j \in \{0,1\}, cap_j \ge 0
\end{aligned}
$$

Benders Decomposition - Single Cut Version

THE UNIVERSITY OF ARIZONA

Algorithm *Benders Decomposition-Single-Cut Version*

1: Initialization Step: Find an initial feasible solution $x_j(0), cap_j(0)$;
Set $LB = \sum_{j\in J} f_j x_j(0)$ and $UB = \infty$;
Let $t = 0, M_\omega = \emptyset, N_\omega = \emptyset$, and go to the main step.

2. Main Step:

2.1 Solve the **(DSP)** for each scenario ω

2.2 If any scenario **(DSP)** is unbounded, $n \leftarrow n+1$,
get extreme rays and add feasibility cut to **(RMP)**:
$0 \geq \sum_{i\in I} S_i^\omega \alpha_i^{\omega,n} + \sum_{j\in J} cap_j \beta_j^{\omega,n} + \sum_{i\in I}\sum_{j\in J} Mx_j\gamma_{ij}^{\omega,n}$
$\sum_{j\in J}\sum_{k\in K}\sum_{q\in Q} Mx_j \delta_{jk}^{q,\omega,n} + \sum_{k\in K}\sum_{q\in Q} D_k^q \mu_k^{q,\omega,n}$

2.3 Else, $m \leftarrow m+1$,
get extreme points and add optimality cut to **(RMP)**:
$\eta \geq \sum_{\omega\in\Omega} p(\omega)(\sum_{i\in I} S_i^\omega \alpha_i^{\omega,m} + \sum_{j\in J} cap_j \beta_j^{\omega,m} + \sum_{i\in I}\sum_{j\in J} Mx_j\gamma_{ij}^{\omega,m} +$
$\sum_{j\in J}\sum_{k\in K}\sum_{q\in Q} Mx_j \delta_{jk}^{q,\omega,m} + \sum_{k\in K}\sum_{q\in Q} D_k^q \mu_k^{q,\omega,m})$
Set $UB = f_j x_j(t) + \sum_{\omega\in\Omega} p(\omega) obj(DSP^\omega)$

2.4 If $UB - LB \leq \epsilon$, stop with optimal solution $\eta(t), x_j(t), Cap_j(t)$;
Else, set $t \leftarrow t+1$

2.5 Solve the **(RMP)** by adding feasibility cuts and optimality cuts,
set $LB = f_j x_j(t) + \eta(t)$, and go to main step

BD within a B&C framework

THE UNIVERSITY OF ARIZONA

- In a classic implementation of Benders decomposition, at each iteration the relaxed master problem which is a MIP is solved to optimality.
- Then given the optimal solution of this relaxed master problem, one checks whether any feasibility or optimality cut should be added.
- This procedure is repeated until no feasibility or optimality cuts are added to the relaxed master problem.
- However, repeatedly solving this MIP to optimality can be computationally time consuming.
- Instead, we deploy the Benders cuts within a branch-and-bound leading to B&C algorithm.
- We use callback features of CPLEX to implement our B&C algorithm.

BD within a B&C framework

THE UNIVERSITY OF ARIZONA

- We start the branch-and-bound tree search over the integer variables with the relaxed master problem at root node.
- Suppose a node is not fathomed due to infeasibility or by bound, then we execute the Benders cut generation algorithm and add the cuts to the current node.
- If the current node's LP relaxation solution satisfies integrality constraints, we use the LazyConstraintCallback feature of CPLEX to add lazy cuts violated by an integer solution of the current formulation.
- We check for Benders feasibility cuts and optimality cuts. If no such cut is found, the best found feasible solution is updated with the current solution, and the node is fathomed.
- If the LP relaxation is not integer feasible, the UserConstraintCallback feature of CPLEX is called to add user cuts cutting off non-integer solutions and if no violating cut is found, then branching will be done.

Outline

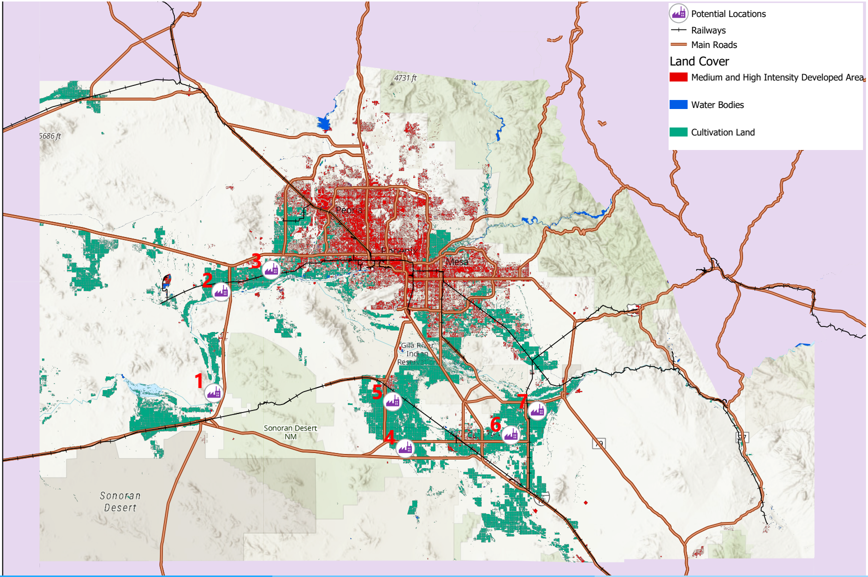
Potential Candidates

THE UNIVERSITY OF ARIZONA

- Objective
  - Find a processing location in AZ for converting guayule to rubber, resin, and bagasse.

- Study Region
  - Two counties in Arizona: Maricopa, Pinal
- Candidates selection Criteria
  - (a) accessibility to biomass supplies, close to feasible planting fields
  - (b) within 2-mile straight line distance to major roads
  - (c) within 2-mile straight line distance to railway
  - (d) within 2-mile straight line distance of the pipeline/major water bodies
  - (e) one candidate be surrounded by at least 100,000 acres
  - (f) Avoid developed area (high intensity and medium intensity)

Potential Candidates



Other Input Data

- Data inputs of parameters for optimization model

Inputs of Parameters for Optimization Model

Parameter	Number	References
harvest and haul of biomass ($/metric tonne)	37.33	[25, 26]
loading and unloading cost for biomass($/metric tonne)	5.7	[26, 27]
biomass variable mileage cost ($/metric tonne*km)	0.12	[27]
transportation cost for rubber and resin ($/metric tonne)	66.14	Industry partner
conversion rate for rubber	0.06091875	industry partner
conversion rate for resin	0.060309563	industry partner
yield (metric tonnes/acre)	10	industry partner
annually minimal capacity of processing center with respect to biomass (metric tonne)	405000	industry partner
annually maximal capacity of processing center with repect to biomass (metric tonne)	499500	industry partner

Other Input Data

THE UNIVERSITY OF ARIZONA

- Supply uncertainty
  - In the planning stage, the adoption rate for each farm from current planting crop to guayule is not known.
  - 154 farms in total
  - For each farm, the possible adoption rate ranges from 30% to 90%
- Demand

Customers and Demands

Location	Products	Demand of rubber (Metric Tonne)	Demand of resin (Metric Tonne)
Des Moines, Iowa	Agricultural Machinery Tires	5000	5000
Russellville, Arkansas	Tubes and Flaps	5000	5000
Bloomington, Illinois	Off-the-road Tires	5000	5000
La Vergne, Tennessee	Tires	5000	5000
Abilene, Texas	Retreading Materials	3000	3000
Muscatine, Iowa	Retreading Equipment	3000	3000

Numerical Experiments

THE UNIVERSITY OF ARIZONA

- Results for deterministic cases

Adoption Rate	30%	35%	40%	45%	50%	55%	60%	65%	70%	75%	80%	85%
Facility Location	5	5	5	5	5	5	5	5	5	5	5	5
Capacity(Metric Tonne)	431109	431109	431109	431109	431109	431109	431109	431109	431109	431109	431109	431109
Cost(10^7 \$)	2.2710	2.2624	2.2587	2.2561	2.2553	2.2551	2.2541	2.2544	2.2538	2.2535	2.253	2.2532

- Stochastic Test Cases – Case 1 & Case 2

	Scenario	1	2	3	4	5	6	7	8	9	10	11	12	13
	Adoption rate	0.3	0.35	0.4	0.45	0.5	0.55	0.6	0.65	0.7	0.75	0.8	0.85	0.9
Case1	Probability	0.3	0.2	0.1	0.1	0.05	0.05	0.05	0.05	0.02	0.02	0.02	0.02	0.02
Case2	Probability	0.02	0.02	0.02	0.02	0.02	0.05	0.05	0.05	0.05	0.1	0.1	0.2	0.3

- Results for stochastic cases

	Facility Location	Capacity (Metric Tonne)	Cost (10^7 \$)
Case 1	5	431109	2.26154
Case 2	5	431109	2.25415

Numerical Experiments

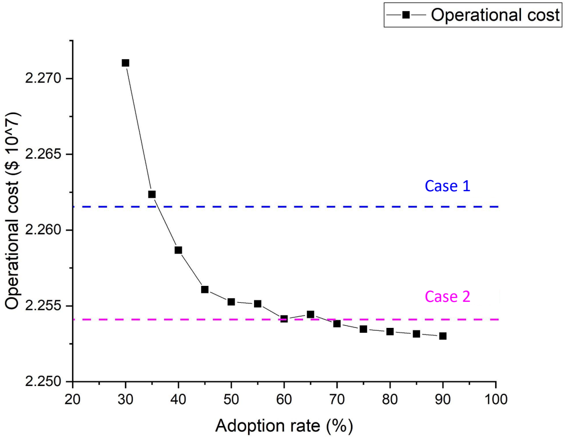
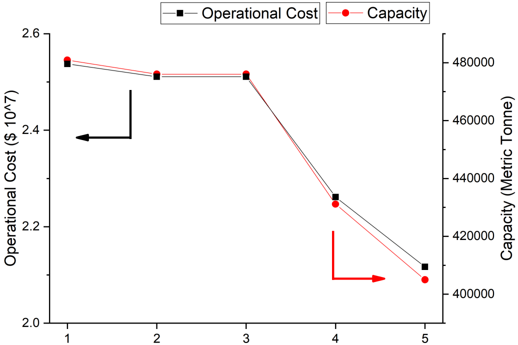


Figure: Operational Cost for Different Scenarios

Numerical Experiments

Sensitivity Analysis to Demand under Case 1

No.	rubber demand	resin demand	location	capacity (metric tonne)	operational cost(10^7\$)
1	29000	29000	5	480852	2.53751
2	29000	26000	5	476044	2.51083
3	29000	23000	5	476044	2.51083
4	26000	26000	5	431109	2.26154
5	24000	24000	5	405000	2.11668



Thank you!!

Acknowledgement. We appreaciate the support from SBAR. This material is based upon funding provided by the USDA-NIFA, Grant # 2017-68005-26867.