


Optimal Wholesale Facilities Location within the Fruit and Vegetables Supply Chain: A LP-MIP Heuristic Approach

Hamideh Etemadnia[1], Stephan Goetz[1,2], Patrick Canning[3]

[1]Agricultural Economics, Sociology and Education,
[2]The Northeast Regional Center for Rural Development,
Pennsylvania State University

[3]United States Department of Agriculture,
Economic Research Service

This research was supported by USDA-NIFA AFRI Grant #2011-68004-30057: Enhancing Food Security of Underserved Populations in the Northeast through Sustainable Regional Foods Systems, a project led by the Northeast Regional Center for Rural Development

PENNSTATE College of Agricultural Sciences Agricultural Economics and Rural Sociology

NE RCRD The Northeast Regional Center for Rural Development

USDA


RUNGIS

Outline

- Introduction and background
- Hub Location Problem
- Objective and Problem Formulation
- Experimental Results and Analysis
- Conclusions, and
- Future Work

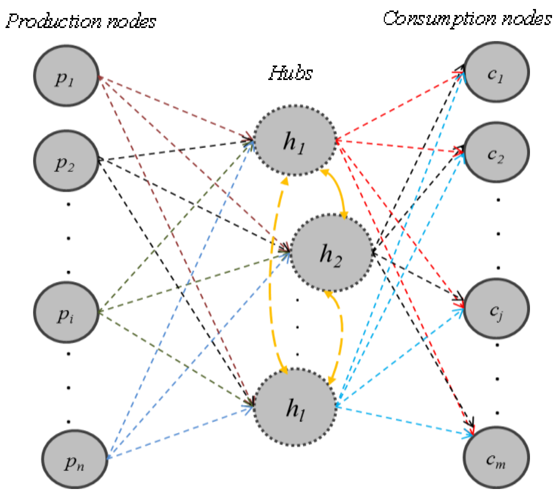
Introduction

- Population growth is posing a challenge to food availability and accessibility.
- To maintain the balance between supply and the growing demand for the food products, the number of production and consumption sites increase.
- The emergence of more production-consumption nodes also complicates food accessibility and availability.
- Perishability and freshness challenges
- Interest in locally produced food
- “Know Your Food, Know Your Farmer” (USDA)

Question : “What is a practical way of bringing food products to customers at reasonable cost by significantly increasing the role of locally produced foods in satisfying existing demand and consumers’ need?”

Hub Location Problem

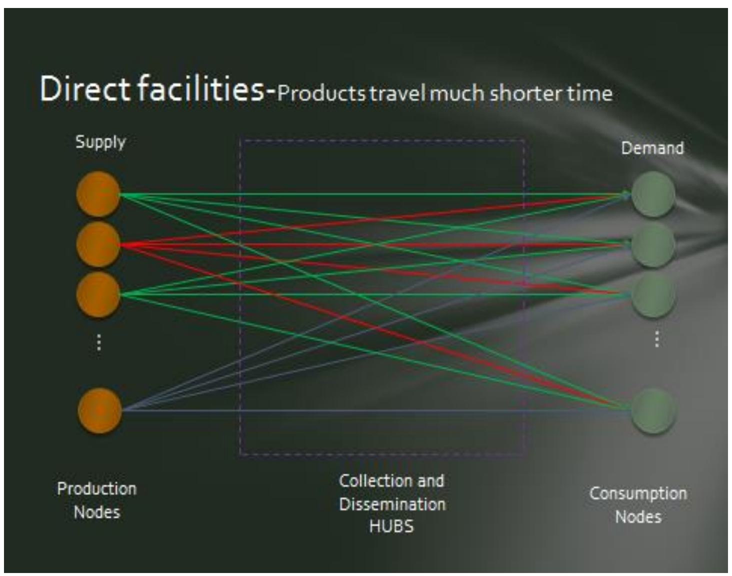
- The hub location problem arises when flow (travelers, airline passengers, cargos, farm products, mails, etc.) must be sent from an origin node to a destination node.
- A hub location is defined as existing wherever placing a direct link between each OD pair is either challenging or costly.

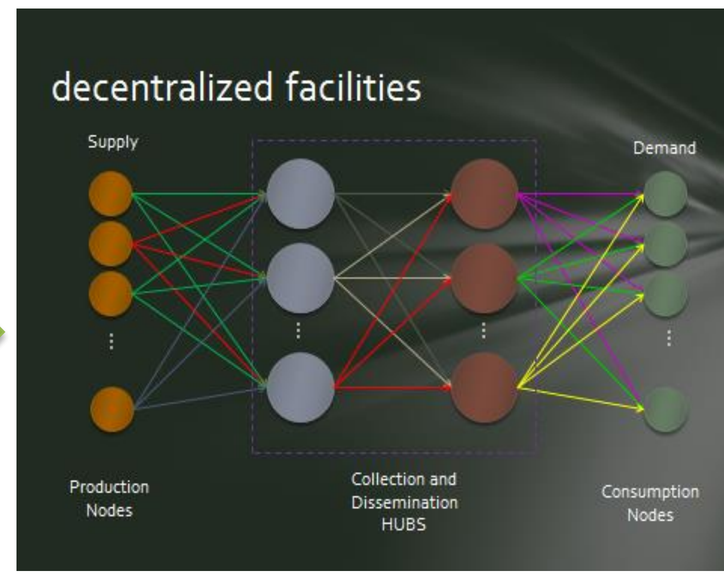


- Campbell (1994)and Campbell and O'Kelly (1994 – 2012) provide Comprehensive introduction, survey, and commentary review on hub location research. Formulations and solution approaches for the Capacitated Multiple Allocation Hub Location Problem (CMAHLP) are presented in (Ebery et al. 2000).
- GIS-based solutions are also proposed to solve the location problem by finding the optimal number and location of facilities in a supply-demand management network (Gu et al. 2009, Trubint et al. 2006, Large et al. 2004).

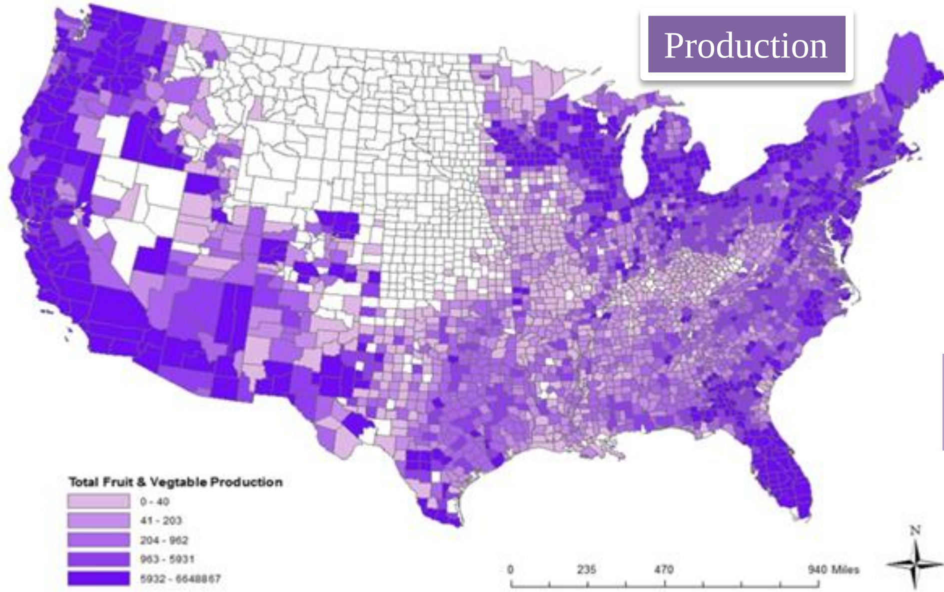
Food Distribution Hubs

- The hub network allows a large number of production and consumption nodes to be connected with fewer links.
- Reducing the number of links and their distances reduces food transportation costs and final product prices.





Total Fruit and Vegetable Production and Consumption Distribution (2007)

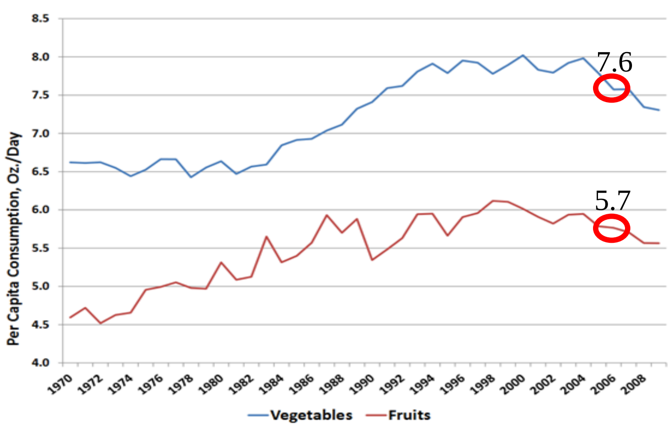


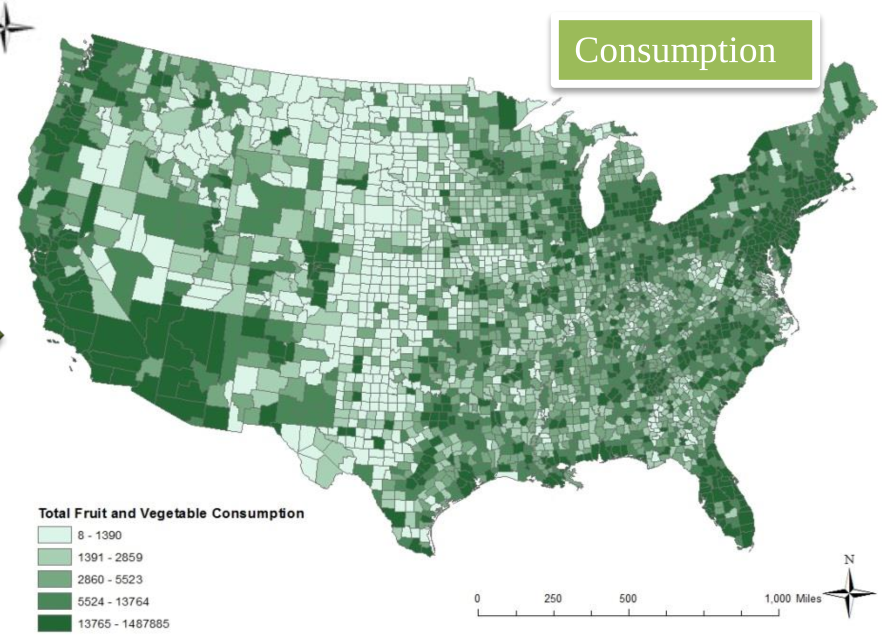
State FIPS	State
11	District of Columbia
31	Nebraska
38	North Dakota
46	South Dakota
56	Wyoming

No Data Available for these states in this study

Based on 2007 USDA/NASS data, a total of 2,522 counties in the U.S. contain fruit- or vegetable-producing farms.

US Per Capita Consumption of Fruits and Vegetables (Source: USDA ERS - USDA Rural Development, Executive Summary (2013))

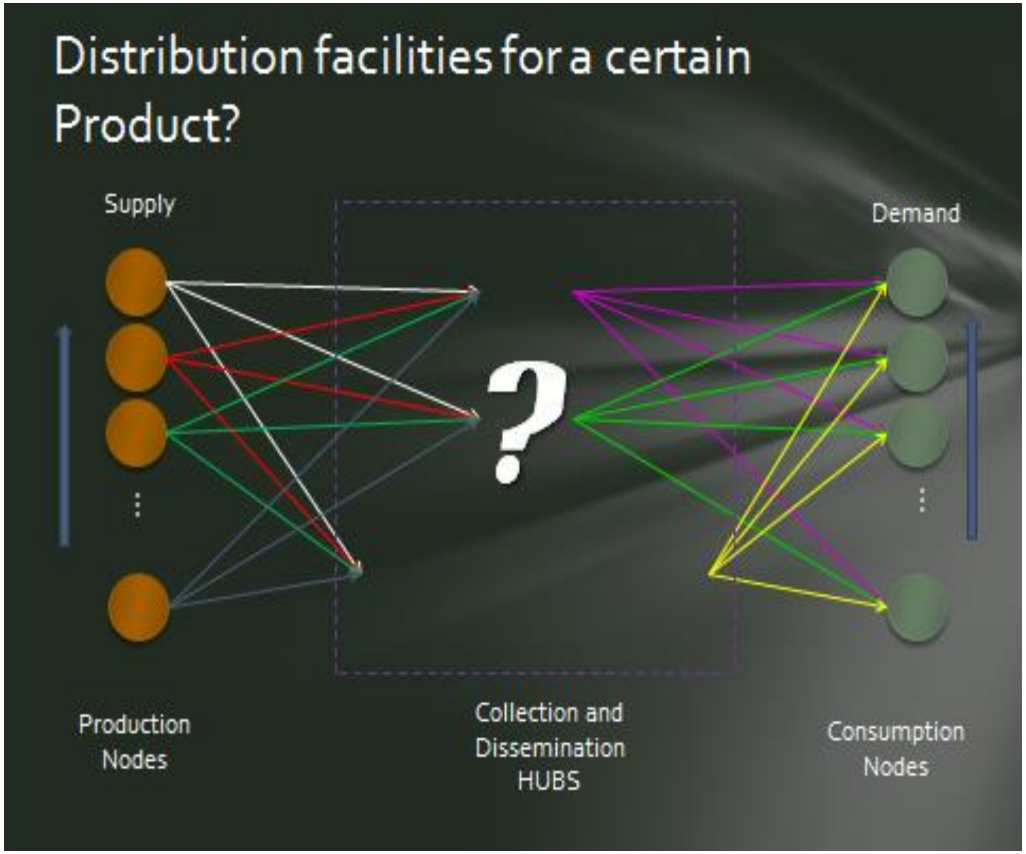




Objective

- Design and locate an optimal hub-based logistics network of wholesale markets within the food supply chain system through the followings:
  - Considering transportation impedance where the total travel cost between the processing and retail markets is minimized.
  - The product does not travel more than the maximum allowed predefined distance between the processing-wholesale hub and retail market (regional food access) for Land Transport.
  - Higher cost is associated with Air Transport.
  - Wholesale hubs are closer to the retail markets than to the processing facilities.
  - The optimal number of wholesale market hub locations is determined based on logistic performance, hub capacity and demand in the supply chain network.

Objective (Cont.)



Problem Formulation

Minimize

$$TC = \left(\sum_{i,h\in FS} ms_{ih} f(d_{ih}) + \sum_{h,j\in FD} md_{hj} f\left(d_{hj}\right)\right).CL + \left(\sum_{i,h\notin FS} ms_{ih} f(d_{ih}) + \sum_{h,j\notin FD} md_{hj} f\left(d_{hj}\right)\right).CA + \sum_h F_h Z_h \quad (1)$$

Subject to:

$$\sum_h ms_{ih} \leq p_i \quad \text{for all } i \quad (2)$$

$$\sum_h md_{hj} = c_j \quad \text{for all } j \quad (3)$$

$$\sum_i ms_{ih} = \sum_j md_{hj} \quad \text{for all } h \quad (4)$$

$$\sum_i ms_{ih} \leq Z_h.U_h \quad \text{for all } h \quad (5)$$

$$\sum_i ms_{ih} \geq Z_h.L_h \quad \text{for all } h \quad (6)$$

$$\sum_j md_{hj} \leq Z_h.U_h \quad \text{for all } h \quad (7)$$

$$\sum_j md_{hj} \geq Z_h.L_h \quad \text{for all } h \quad (8)$$

Where

$$Z_h = \begin{cases} 1 & if\ county\ node\ h\ is\ a\ hub \\ 0 & therewise, \end{cases}$$

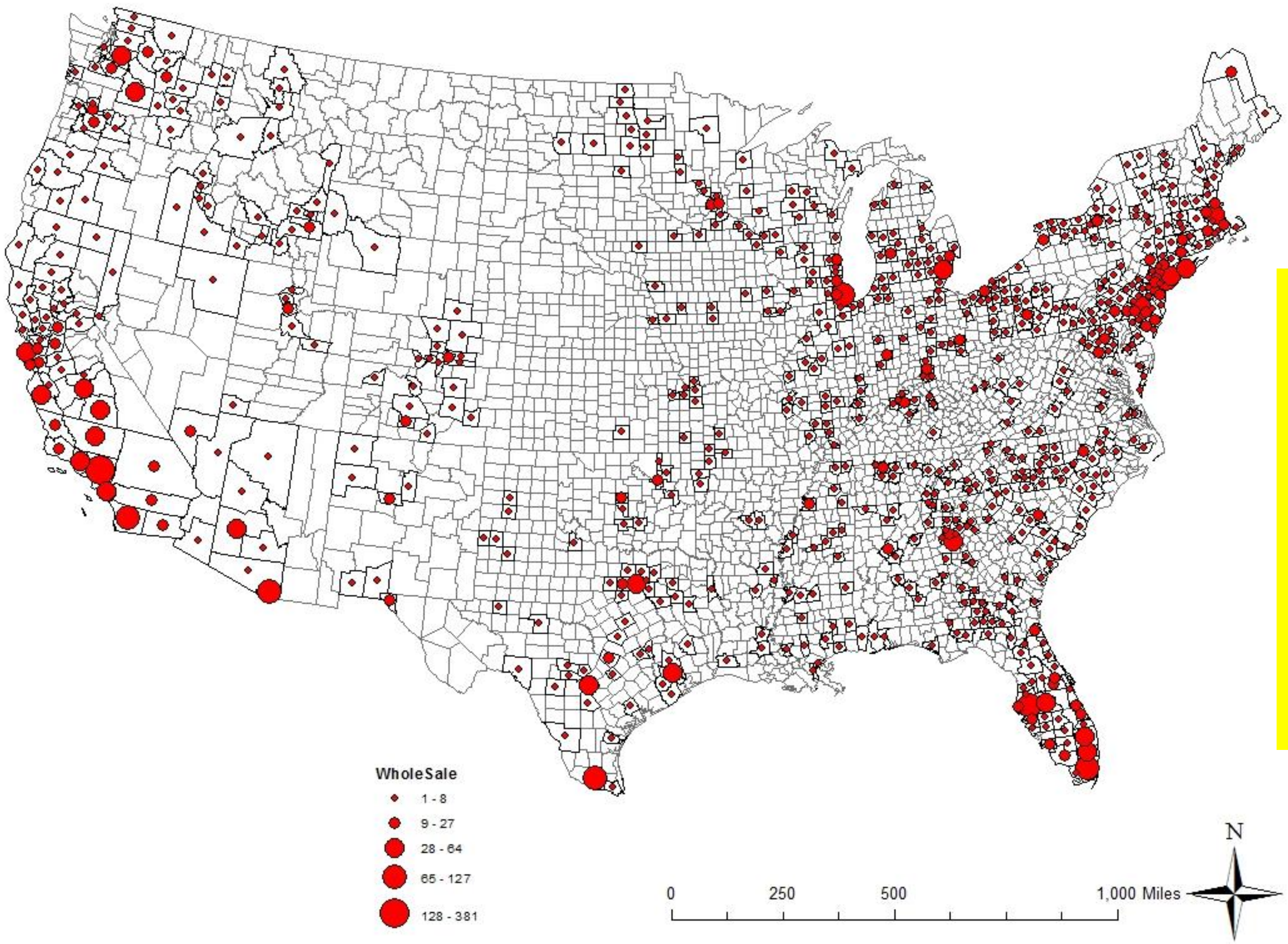
$$ms_{ih}, md_{ih} \geq 0$$

Variable Definition

Index (variable)	Definition
i	production location
j	consumption location
h	hub location
$f(d_{ij})$	impedance values as function of highway miles between any $i - j$ location pairs
CL	fixed cost (Land Transportation) per mile per ton value ($ per ton mile)
CA	fixed cost (Air Transportation) per mile per ton value ($ per ton mile)
F_h	fixed cost of locating and operating a hub in county h ($)
N $(\|N\| = n)$	a set of counties to be interconnected
H $(\|H\| = h)$	the estimated set of total hubs to be constructed
p_i	total supply in production location i (tons)
c_j	total demand in consumption location j (tons)
ms_{ih}	fraction of the quantities shipped from production location i to hub location h (tons)
md_{hj}	fraction of the quantities shipped from hub location h to consumption location j (tons)
Z_h	integer variable: $Z_h = 1$ if location h is a hub, and 0 otherwise
U_h	maximum capacity of hub facility in location h (tons)
L_h	minimum capacity of hub facility in location h should fulfilled (tons)
TC	total cost ($)
TP	threshold distance between production locations and hub locations (mile)
TM	threshold distance between hub locations to consumption locations (mile)
FS	subsets of distances between production regions to hub locations with respect to TP
FD	subsets of distances between hub locations to consumption locations with respect to TM

Fruit and Vegetable Industry

Vegetable and Fruit Wholesale Facilities (2007).



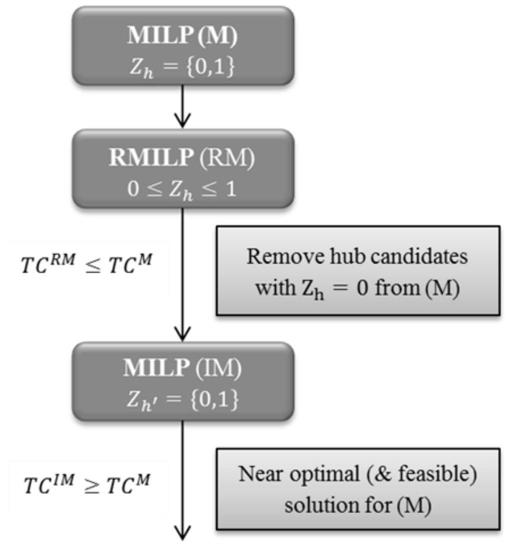
Goal:
To understand how optimal locations of the Wholesale Markets adjust over time with changing hub capacity constraints and products' travel distance.

Fresh Fruit and Vegetable Industry

- The network consists a total of **3080** counties
- Maximum **distance** of **3,637.3** miles is between Monroe, Florida and San Juan, Washington
- Maximum production is estimated to be **6,648,867** tons in **Fresno County** in the state of **California** (1,682,763 (tons) fruit and 4,966,104 (tons) vegetable).
- Maximum demand is estimated to be **1,487,885** tons in **Los Angeles** County in the state of **California**.
- Total Fruit + Vegetable **production** is **75,454,796** tons
- Demand in each county is estimated by multiplying US per capita consumption of fruits and vegetables by county population
- Total **demand** for Fruit + Vegetable is **45,409,579** tons

Model Implementation

- **A heuristic approach is developed to considerably reduce the computational time:**
  - Initially relax the integer variables ($0 \leq Z_h \leq 1$) of model solve a relaxed MILP (RMILP) as a linear programming problem.
  - The solution to this model would potentially contain some excluded nodes with $Z_h = 0$ and some potential hubs with $Z_h \neq 0$.
  - Omit those hub candidates from the model and build a smaller sized model called Intermediate Model (IM).



- **Naturally, the IM model eliminates Z_h, ms_{ih}, md_{hj} variables and all associated constraints related to $Z_h = 0$.**

Model Implementation (Cont.)

- We implemented our model and algorithm in GAMS 23.7.3 and used CPELX 12.4 as the MILP solver with the following CPLEX options:
  solvefinal=0 ;   threads=4;

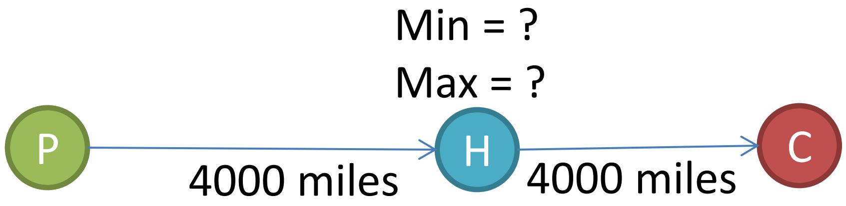
- All optimization studies were carried out on High Performance Computing (HPC) Systems, each case running on four processor cores with a memory allocation of 60 GB. Several simulations were conducted using the model to determine the optimal number of U.S. F&V hubs and their locations.

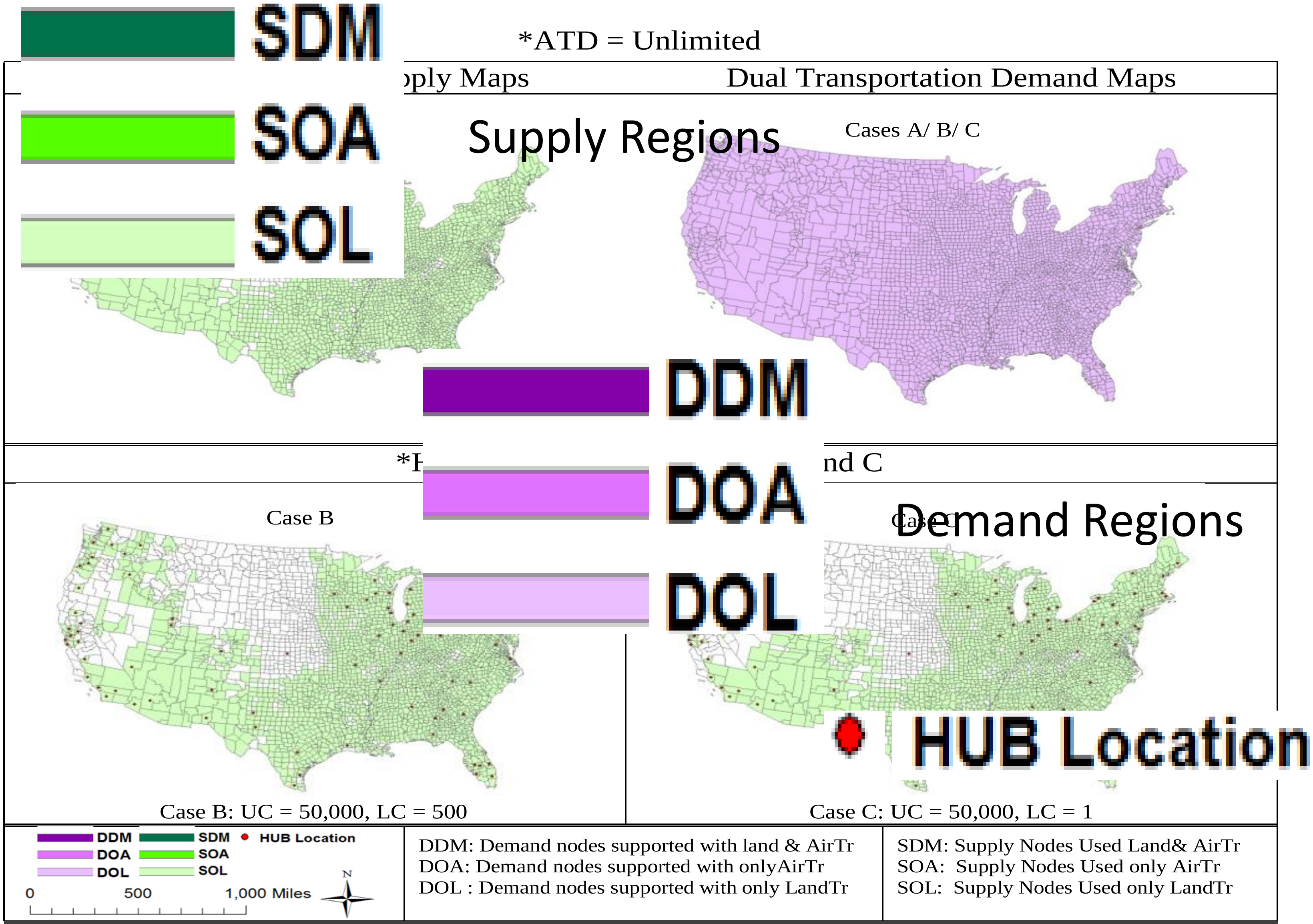
Experimental Results and Analysis(Cont.)

Result of number of hubs with *unlimited average travel distance and limited upper/lower bound of hub capacity

*distance greater than the maximum distance between the two farthest counties across the U.S., **unlimited = 500,000

Case No.	Hub(s) Max Capacity (Ton/100)	Hub(s) Min Capacity (Ton/100)	No. of Hub(s)	No. of Supply Nodes Used only LandTr	Objective Function $x10^6$	Relative GAP	Elapsed Time(h)
A	**unlimited	500	92	2411	729	0.029	8:20
B	50,000	500	91	2411	729	0.029	9:26
C	50,000	1	86	2411	730	0.029	13:45



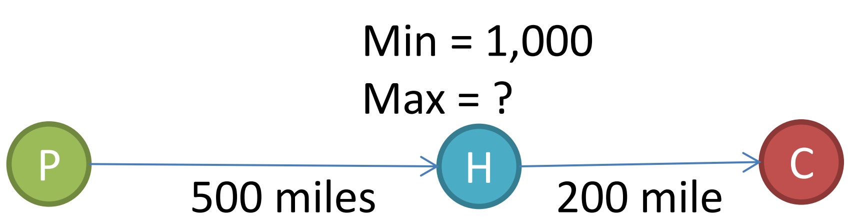


*ATD = Average Travel Distance, UC = Upper hub capacity limit, LC = Lower hub capacity limit

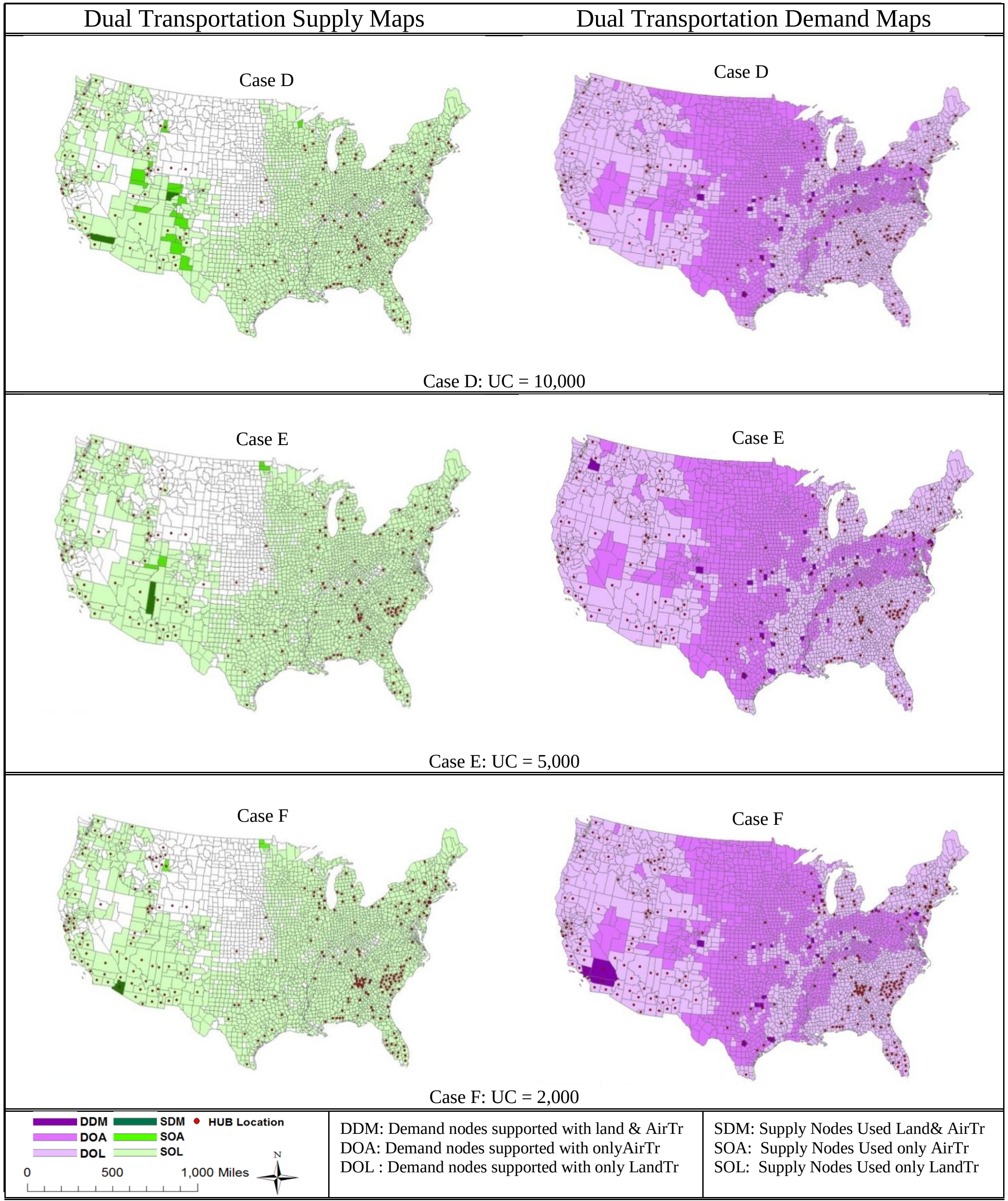
Experimental Results and Analysis(Cont.)

Result of number of hubs with fixed average travel distance and variable upper bound of hub capacity

Case No.	Hub(s) Max Capacity (Ton/100)	No. of Hub(s)	No. of Supply Nodes Used only LandTr	No. of Supply Nodes Used only AirTr	No. of Supply Nodes Used Land& AirTr	No. of Demand nodes supported with only LandTr	No. of Demand nodes supported with onlyAirTr	No. of Demand nodes supported with land & AirTr	Objective Function $x10^6$	Relative GAP	Elapsed time(h)
D	10,000	160	2418	17	2	1634	1422	24	14,451	0.024	2:30
E	5,000	170	2429	4	1	1645	1407	28	14,650	0.025	5:01
F	2,000	258	2435	3	1	1577	1405	28	15,757	0.021	3:17



*ATD_SH = 500, ATD_HD = 200, LC = 1,000

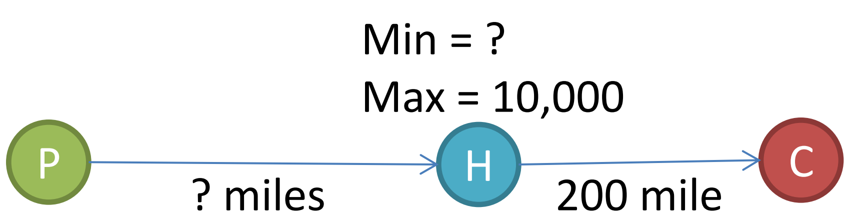


ATD_SH = Average Travel Distance from supply locations to hubs, ATD_HD = Average Travel Distance from hubs to demand locations, UC = Upper hub capacity limit, LC = Lower hub capacity limit

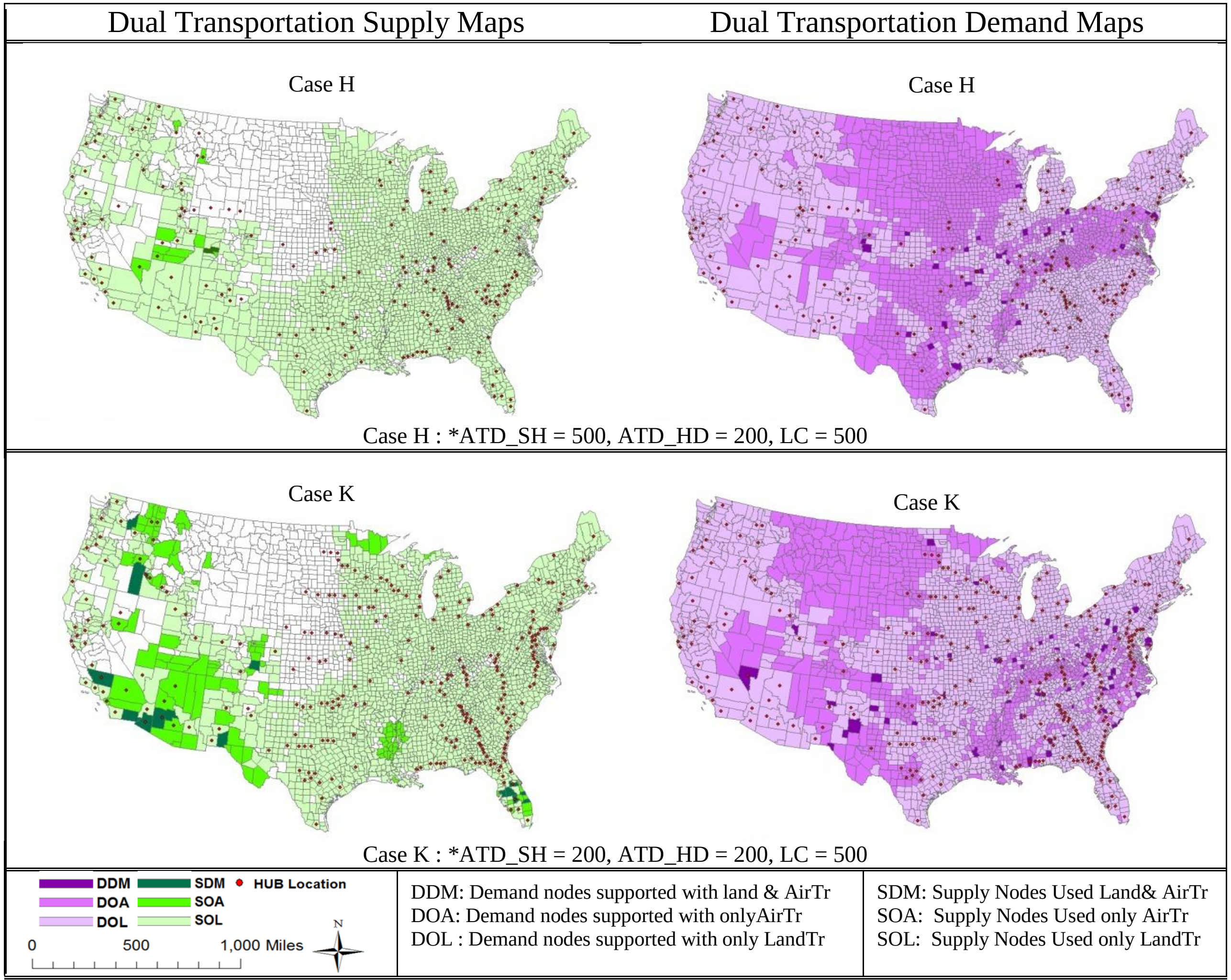
Experimental Results and Analysis(Cont.)

Result of number of hubs with variable average travel distance and lower bound of hub capacity

Case No.	Production-Hub Max Dist (Mile)	Hub(s) Min Capacity fulfilled (Ton/100)	No. of Hub	No. of Supply Nodes Used only LandTr	No. of Supply Nodes Used only AirTr	No. of Supply Nodes Used Land& AirTr	No. of Demand nodes supported with only LandTr	No. of Demand nodes supported with onlyAirTr	No. of Demand nodes supported with land & AirTr	Objective Function x10^6	Relative GAP	Elapsed Time(h)
G	500	1,000	160	2418	17	2	1634	1422	24	14,451	0.024	2:30
H	500	500	193	2424	10	1	1743	1302	35	14,286	0.012	2:31
I	500	1	288	2438	0	0	1860	1162	58	14,151	0.003	1:47
J	200	1,000	202	2142	276	16	1846	1200	34	27,779	0.018	2:48
K	200	500	280	2307	111	15	2179	843	58	27,513	0.009	2:19
L	200	1	514	2382	36	20	2415	616	49	27,340	0.002	2:03



*UC = 10,000



* ATD_SH = Average Travel Distance from supply locations to hubs, ATD_HD = Average Travel Distance from hubs to demand locations, UC = Upper hub capacity limit, LC = Lower hub capacity limit

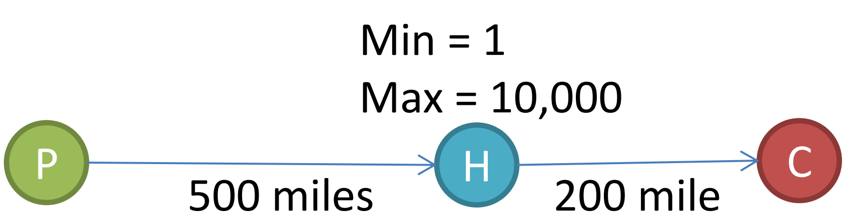
Experimental Results and Analysis(Cont.)

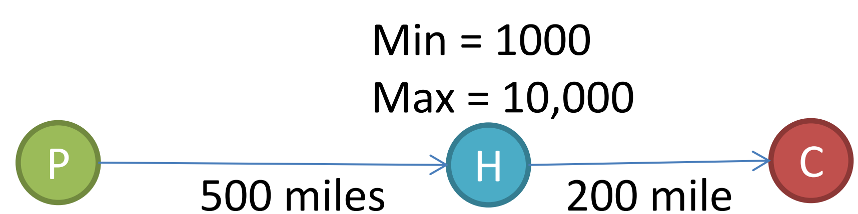
Result comparison of the MIP and LP-MIP solutions

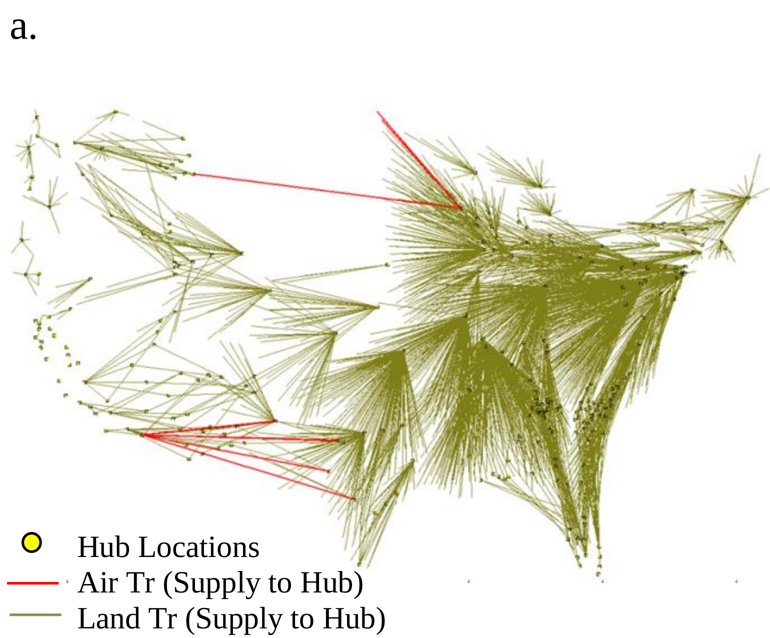
* (') Conventional MIP solution (CPLEX)

**LP-MIP Solution

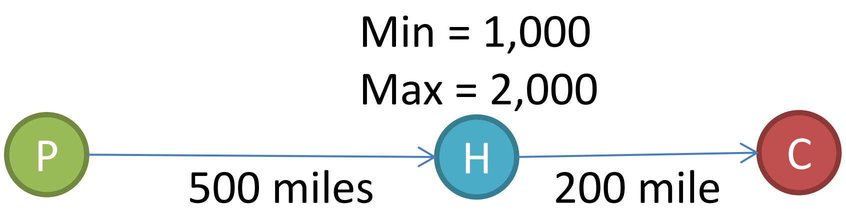
Case No.	Hub(s) Min Capacity fulfilled (Ton/100)	No. of Hub(s)	No. of Supply Nodes Used only LandTr	No. of Supply Nodes Used only AirTr	No. of Supply Nodes Used Land& AirTr	No. of Demand nodes supported with only LandTr	No. of Demand nodes supported with onlyAirTr	No. of Demand nodes supported with land & AirTr	Objective Function x10^6	Relative GAP	Elapsed Time(h)
*I'	500	283	2436	2	0	1863	1163	54	1,4151	0.000	56:01
**I	500	288	2438	0	0	1860	1162	58	14,151	0.003	1:47
G'	1,000										72:00
G	1,000	160	2418	17	2	1634	1422	24	14,451	0.024	2:30

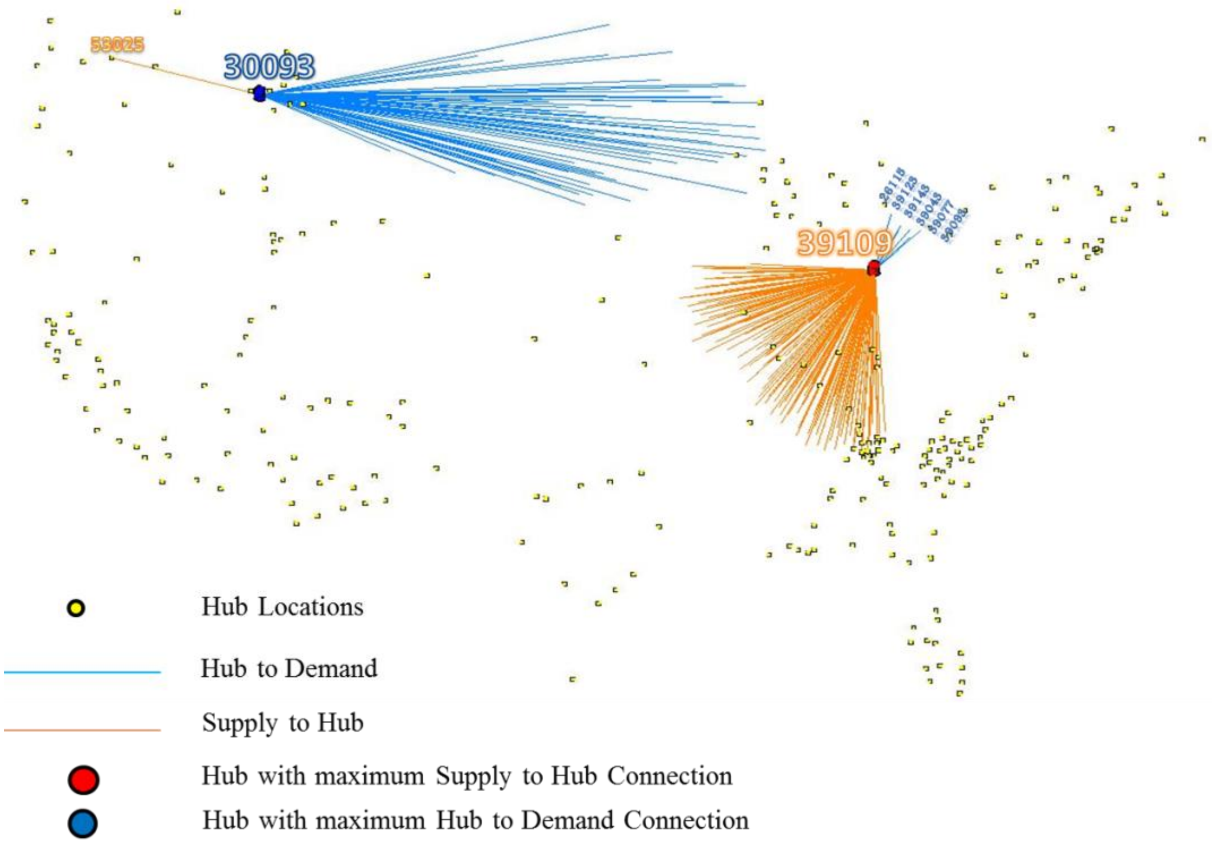






a) Land and air transport connections of supply-to-hub nodes
b) Land and air transport connections of hub-to-demand nodes





Maximum supply connection links the hub 39109 and maximum hub 30093 to demand connection

Conclusion

- This work presents a mathematical formulation of the food industry hub location problem and a LP-MIP heurist solution.
- Our findings allow us to draw several implications for supporting regional food systems by locating wholesale facilities to enable access to production and consumption sites.
- The model is sensitive to the distance over which commodities are allowed to travel (Land and Air Transport)
- The results show the effect of varying these parameters on the selection of hub locations.
- Our analysis is potentially useful for policymakers and private decision-makers from a number of perspectives. (social planner - private sector firms)

Future Work (Limitations)

- **Products are assumed to be undifferentiated**
  - Leads to unrealistic commodity mix outcomes
- **Cost to build hubs and ship products are not empirically estimated**
  - These costs drive model results so such data points should be empirically estimated and/or validated with market data
- **Consumer preferences are not represented**
  - Limits the models use for policy scenario analysis
- **Model does not distinguish shipping verses handling costs**
  - Transportation hubs do not exist
  - Scale economies in transportation are assumed away
- **Import and Export has not been considered**

Future Work (Integrated Modeling Approach)

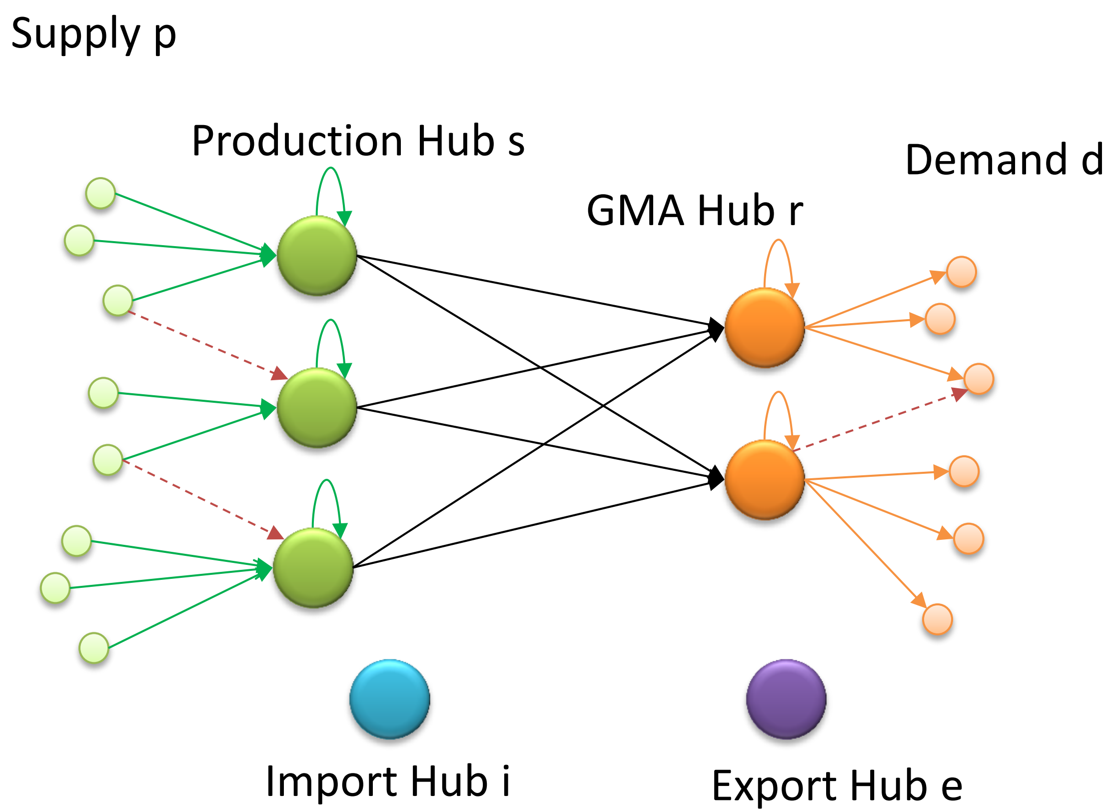
- Restate the mixed integer hub location model as a model of optimal investment
  - Hub construction and operation costs are empirically estimated
  - Economies of scale are empirically estimated

- Restate the model of intra-industry trade to include hub operations with economies of scale estimated in optimal investment model
  - The role of distribution infrastructure in determining marketability of regionally sourced produce is represented in the model
  - The two combined models complete the node-hub, hub-hub, and hub-note process

Future Work (Research Questions to be Addressed)

- What are the costs and impact of new hub infrastructure investments in the U.S.?
  - Where and how many hubs are optimal
  - How much of a gain in market share can regionally sourced produce achieve through a more cost competitive distribution network
- What role to farmland and labor supply constraints play in answers to above question?
- What impact to produce costs and availability occur with large energy price spikes?
  - Baseline impacts verses post hub investment impacts

Future Work (Cont.)



THANK YOU

?