Using Mathematical Modeling to Estimate Unified Origin-Destination Tables from Multiple Datasets

Jared Gearhart

Dan Pless

Linda Nozick

Dean Jones

Sandia National Labs

U.S. DEPARTMENT OF ENERGY

NNSA
National Nuclear Security Administration

Sandia National Laboratories

Exceptional service in the national interest

Overview

- Motivation
  - Transported goods often use multiple shipping modes and may pass through several countries
  - Knowing the ultimate origins and destinations of goods is essential in many analyses
  - No single data source is available
- Outline
  - Optimization methodology to estimate origin-destination tables
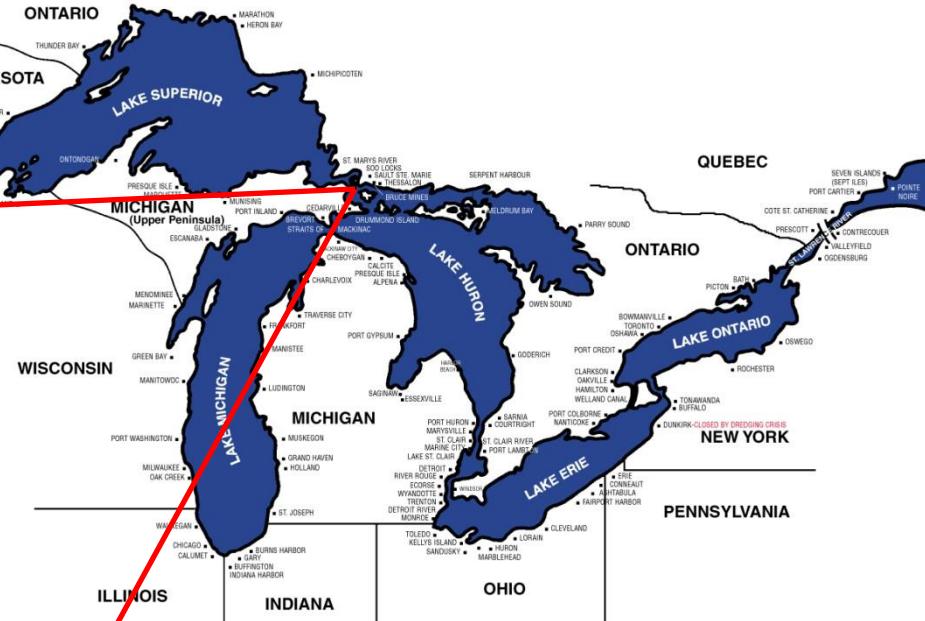 - Case study of disruption to the Soo Locks

Soo Locks Overview

Set of four locks administered by the US Army Corps of Engineers

Lake Carriers' Association
Great Lakes and St. Lawrence Seaway Ports
Listed West to East by State and Province

Capable of handling 1,000-foot "lakers"

Disruption

- Objective: determine the impact to the transportation and manufacturing industries if the Soo Locks are temporarily unavailable

- Knowing what is interrupted is insufficient

- Need to know origins and destinations

2010 Soo Locks Tonnage by Commodity

Commodity	Tonnage	Percent	Primary Direction
Iron Ore	40,570,000	54.4%	Down-bound
Coal	18,068,000	24.2%	Down-bound
Wheat	6,684,000	9.0%	Down-bound
Limestone	4,438,000	6.0%	Up-bound
Other	4,746,000	6.4%	N/A
Total	74,527,000	100.0%	N/A

US Army Corps of Engineers 2010 Annual Report for the Saint Mary's Falls Canal

Shipping Patterns

- Locks used by the US, Canada and rest of the world

- Primarily dry bulk goods moved through Soo Locks

- Goods transported to and from ports
  - Production and consumption occurs away from ports
  - Rail is the major transportation mode

- Multiple data sets required
  - US and Canada
  - Domestic and international
  - Rail and water

Data Sets

- Five data sets of interest

- US Army Corps of Engineers
    - Domestic Waterborne Data
    - International Waterborne Data

- Statistics Canada
    - Domestic Waterborne Data
    - International Waterborne Data

- Surface Transportation Board
    - US Rail Waybill

- No source for Canadian rail data identified

Data Example: Coal

Tons of Coal

Canada International Water Data

US Rail Data

US Domestic Water Data

5,000
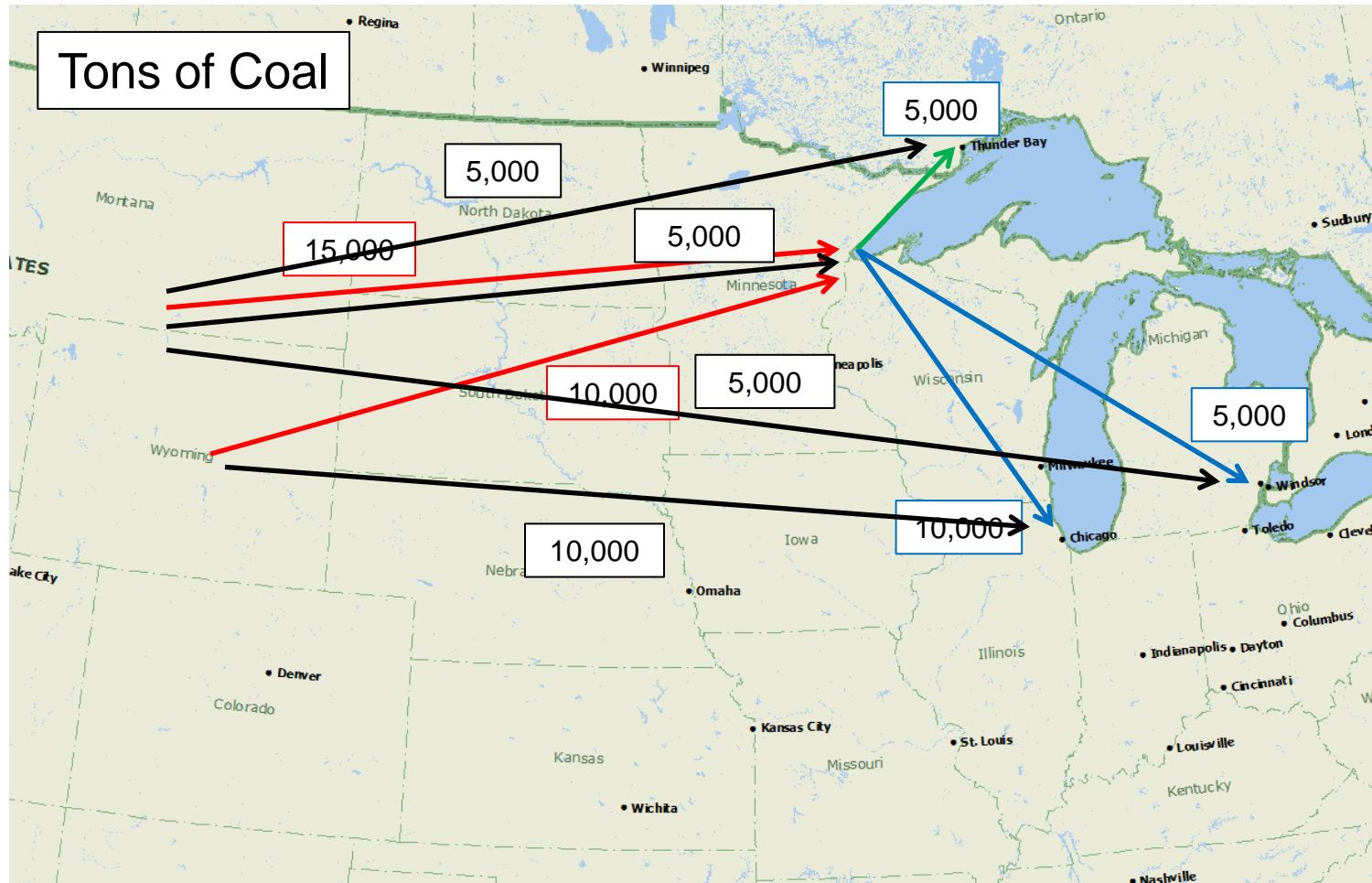15,000
10,000
5,000
10,000
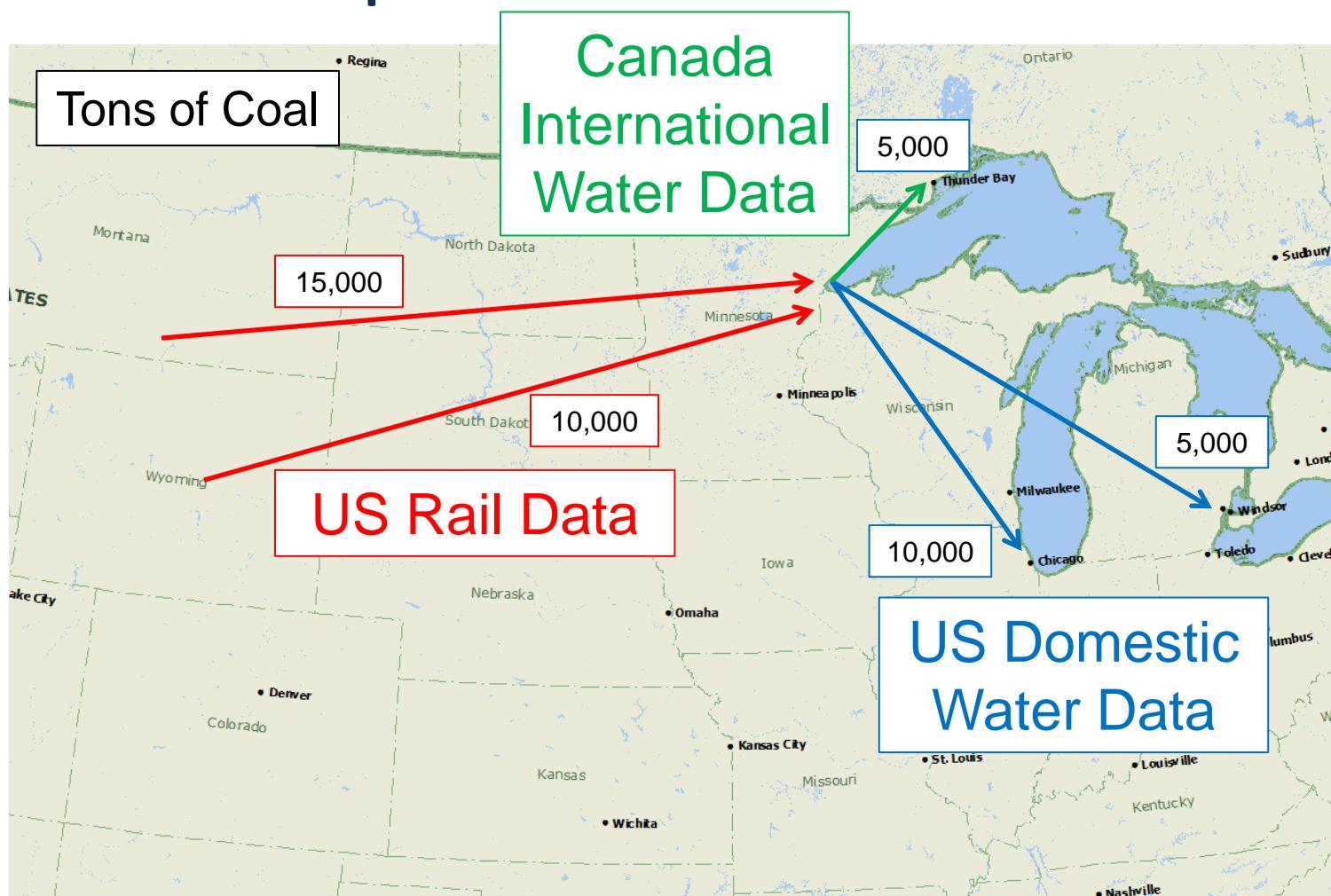
Each data point only provides a piece of the entire trip

OD Estimation

- Use optimization to combine each data set into a unified origin-destination table

- Objective: Create a multi-modal origin-destination table that minimizes the deviations across all data sets

- Model solved for each commodity
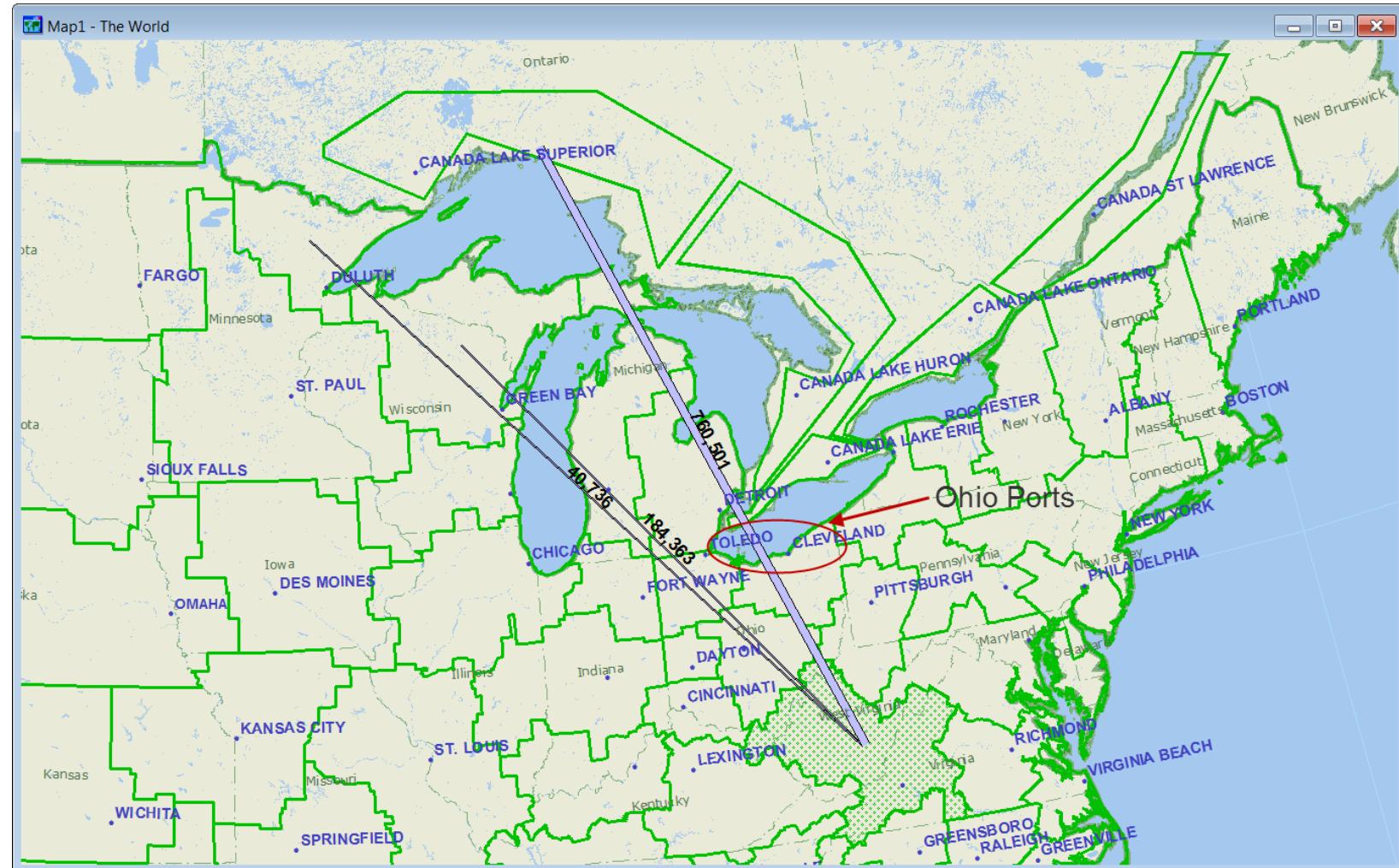
- Linear program

OD Estimation

Tons of Coal

Optimization connects each piece into a single path
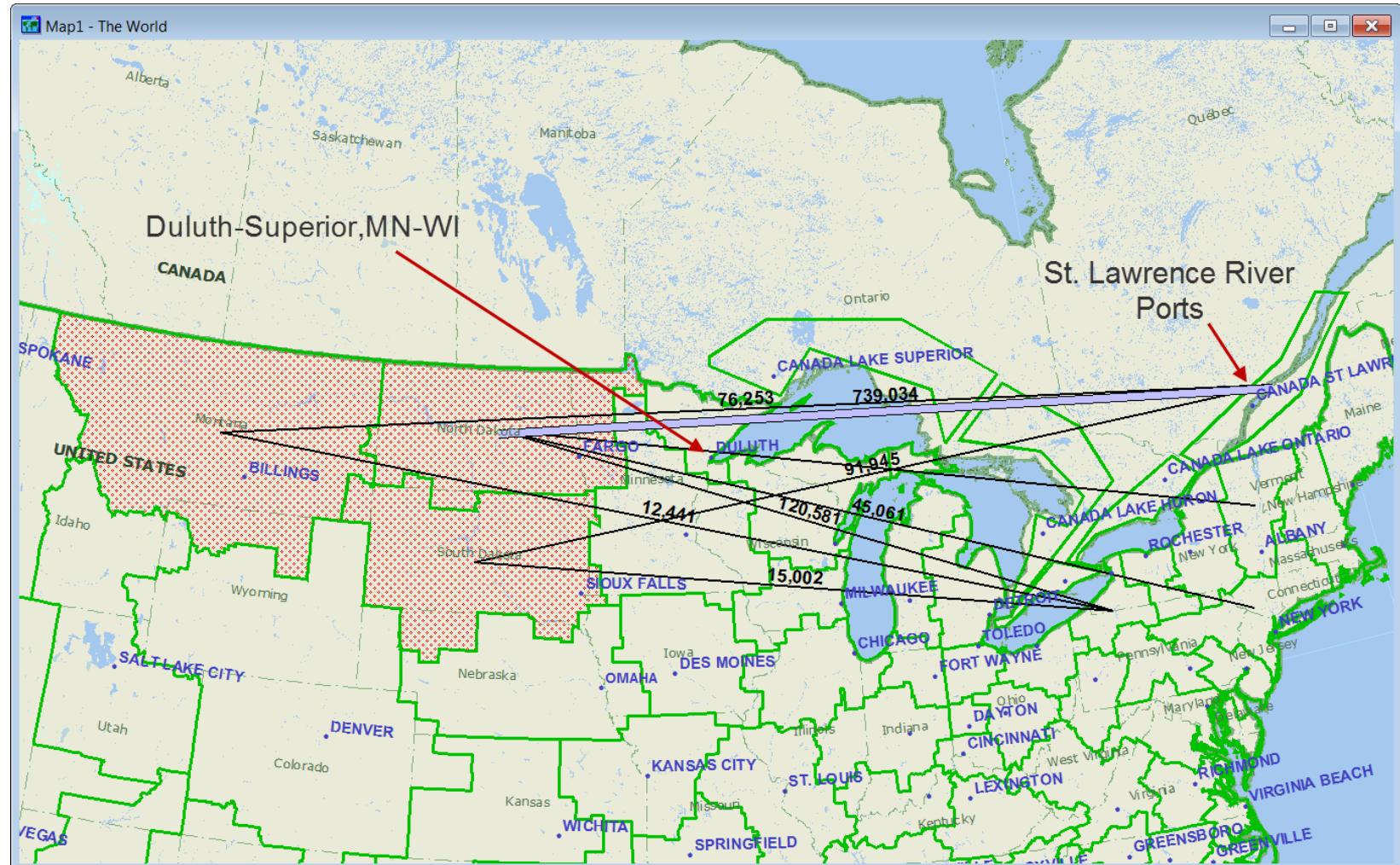
Results: Coal from Western US
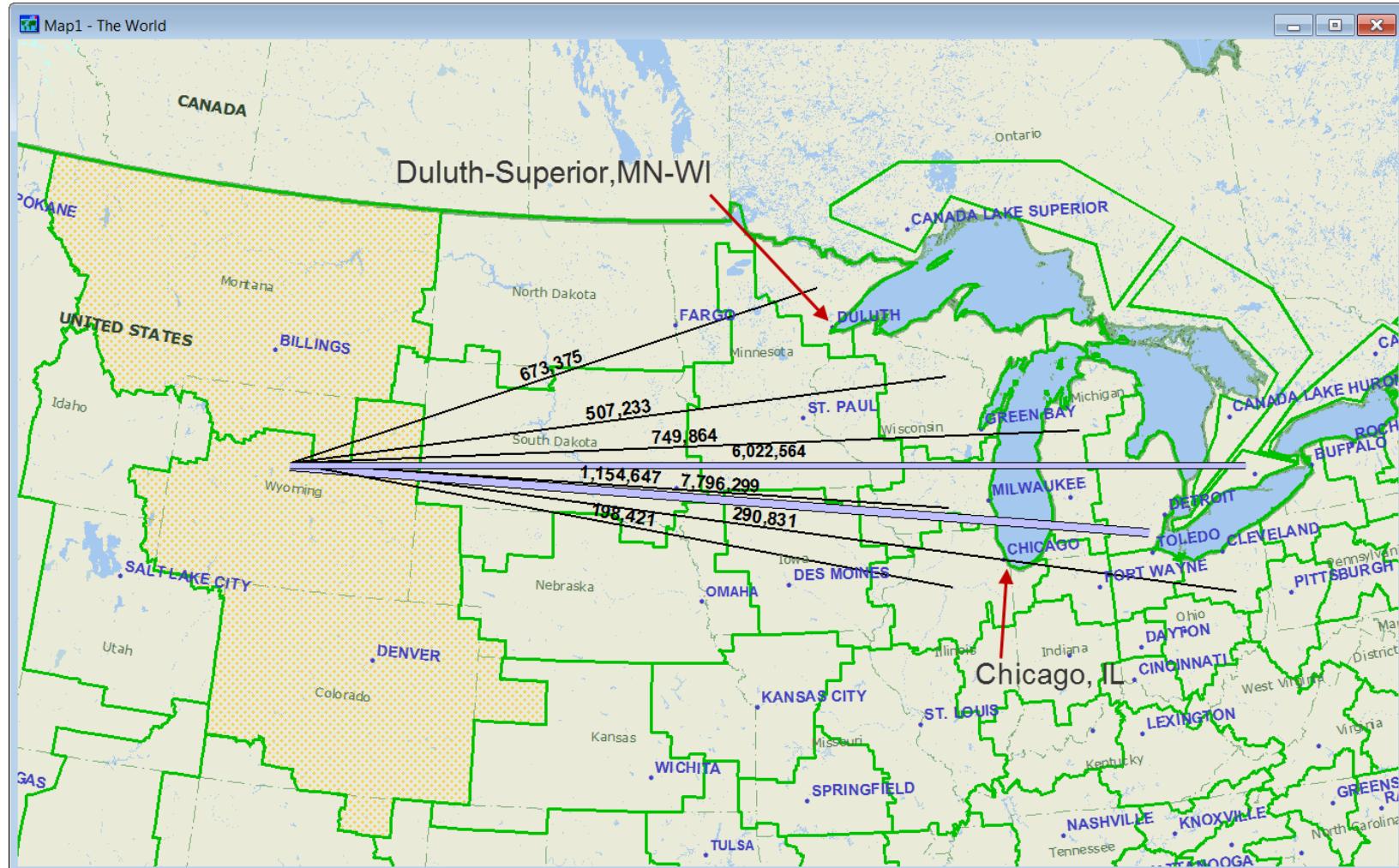
84% (17.5M Tons) of Soo Locks Coal

Results: Coal from Eastern US

5% (1.0M Tons) of Soo Locks Coal

Results: Grain from US Great Plains

The vast majority of US grain is shipped to ports in Quebec

Challenges

- Inconsistent categories across data sets
  - Many commodity systems
  - Countries use difference port coding systems
- Need to create connections between rail and water networks
- Overlapping data sets don't agree
  - Optimization resolves these issues

Conclusion

- Data is generally unavailable for multi-modal, multi-national systems

- Optimization techniques provide a rigorous method for synthesizing nonexistent data

- Resulting output permits detailed quantitative follow-on analysis