Transport Canada Transports Canada

Leveraging Network Analyst and Developing a Python Tool to Automate Production of Commodity Flows

Presented by Elise Pietroniro and Ethan Howieson
Safety, Research and Analysis, Transportation of Dangerous Goods, Transport Canada

July 12, 2018 – Esri User Conference San Diego, CA

Canada

Overview

- Organization
- Project Description
- Commodity Flows
- Required Data
- Getting Started
- Methodology
- Automation
- Benefits
- Next Steps

Organization

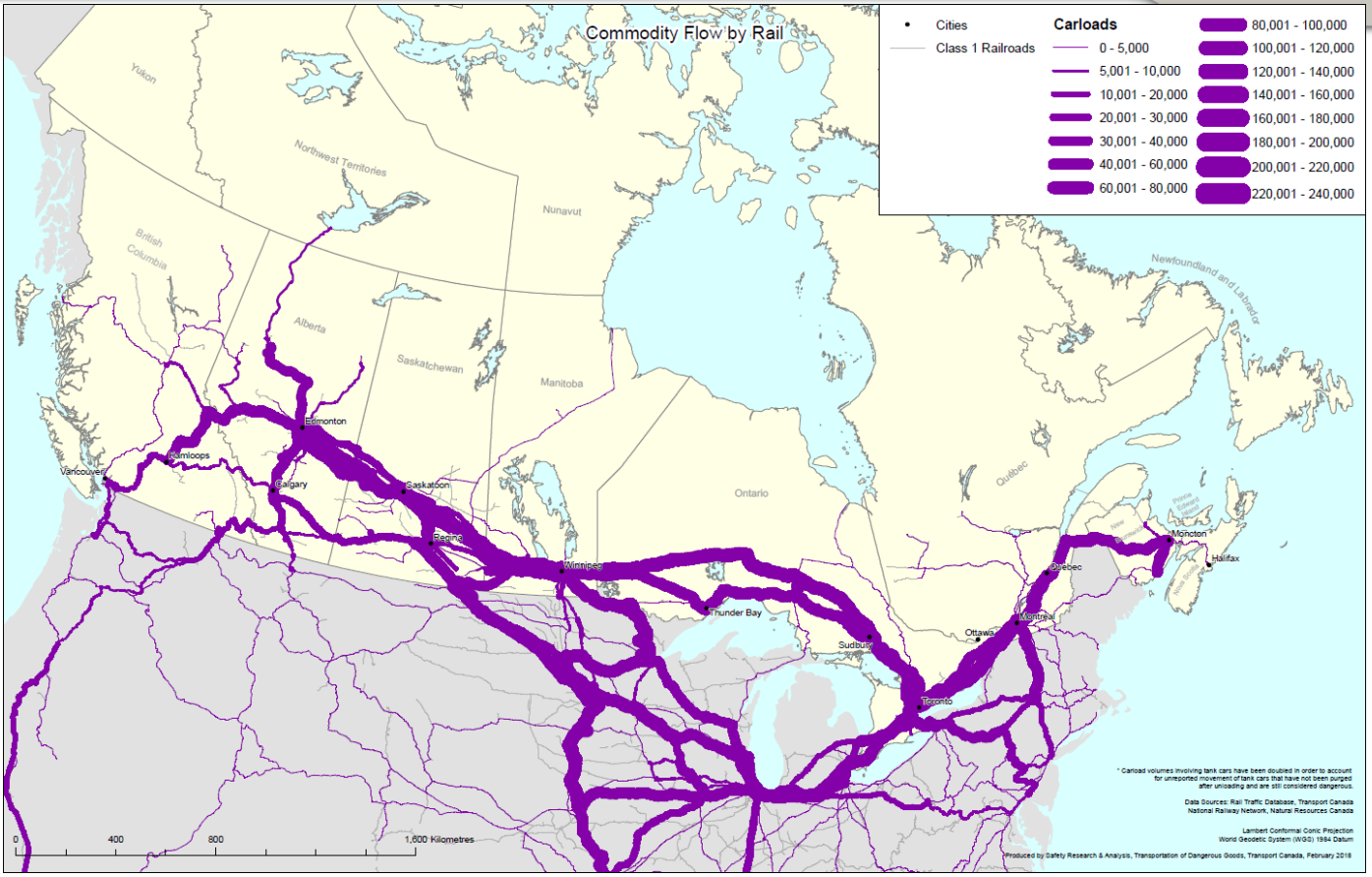
- Transport Canada (TC), based on risks, develops safety standards and regulations, provides oversight and gives expert advice on dangerous goods to promote public safety in the Transportation of Dangerous Goods Directorate (TDG) by all modes of transport in Canada.
- Located in Ottawa, Ontario.
- In 2014, the Safety Research and Analysis Branch (SRA) within TDG introduced a new team focused on GIS.

Project Description

- Project initiated to address formal risk assessment on flammable liquids
- Development of workflow and automated processes to generate commodity flows along rail corridors using MS Excel, ESRI's ArcGIS 10.6, Network Analyst and Python 2.7
- Requirement to automate workflow to increase productivity

Previous Project Methodology

- Previously, this methodology was time intensive. This included,
  - Manual parsing of large spreadsheets.
  - Labour intensive manipulation of routing logic (for all intents and purposes this was done by hand).
- Large datasets difficult to manipulate.
- Needed approximately two weeks to run a single commodity flow.
- Each map was designed individually, there was no continuity.


Commodity Flow by Rail
Cities
Class 1 Railroads
Carloads
0 - 5,000
5,001 - 10,000
10,001 - 20,000
20,001 - 30,000
30,001 - 40,000
40,001 - 60,000
60,001 - 80,000
80,001 - 100,000
100,001 - 120,000
120,001 - 140,000
140,001 - 160,000
160,001 - 180,000
180,001 - 200,000
200,001 - 220,000
220,001 - 240,000
Yukon
Northwest Territories
Nunavut
British Columbia
Alberta
Saskatchewan
Manitoba
Ontario
Quebec
Newfoundland and Labrador
Edmonton
Kamloops
Vancouver
Calgary
Saskatoon
Regina
Winnipeg
Thunder Bay
Sudbury
Ottawa
Toronto
Montreal
Quebec
Moncton
Halifax
0
400
800
1,600 Kilometres
* Carload volumes involving tank cars have been doubled in order to account for unreported movement of tank cars that have not been purged after unloading and are still considered dangerous.
Data Sources: Rail Traffic Database, Transport Canada
National Railway Network, Natural Resources Canada
Lambert Conformal Conic Projection
World Geodetic System (WGS) 1984 Datum
Produced by Safety Research & Analysis, Transportation of Dangerous Goods, Transport Canada, February 2018

What Are Commodity Flows?

Aggregation of carloads of commodities visualized along a route*

- Commodity flows for this project were calculated for rail routes
- Utilized reported data on dangerous goods from Canadian rail carriers for movements
- Modelled on North American rail network

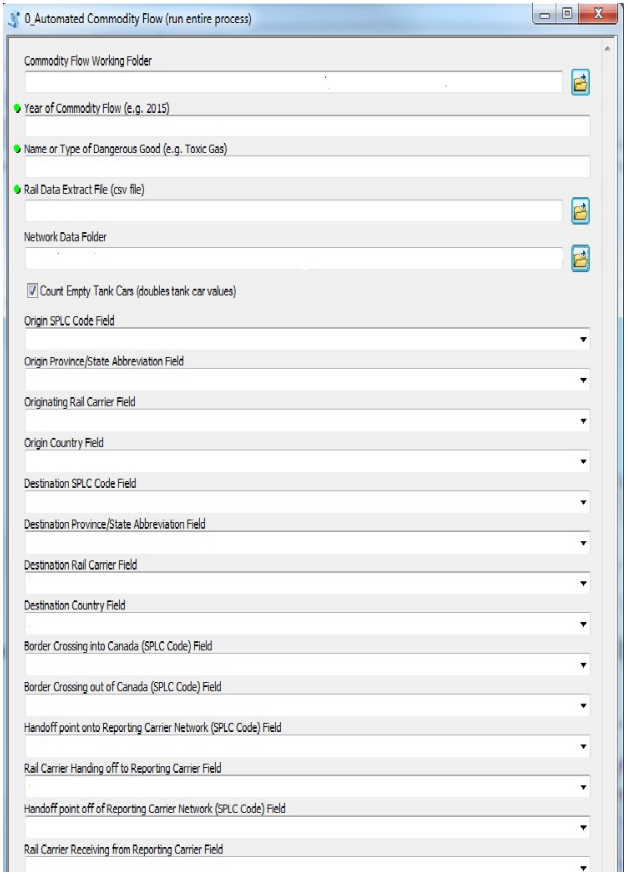
*Note: All data depicted in this presentation are synthetic

Required Data

- Commodity Movement (Origin-Destination)
  - *Canada Transportation Act* (Section 50 & 51) requires reporting on movement for class 1 rail carriers in Canada: Canadian National/Canadian Pacific
  - *Transportation Information Regulations (TIR)* – Details of what is shared
- Rail terminal locations
  - Standard Point Location Codes (SPLC) maintained by American Association of Railways (through RAILINC)
- Base rail network
  - National Railway Network dataset (including all North American rail) maintained by Natural Resources Canada and TC

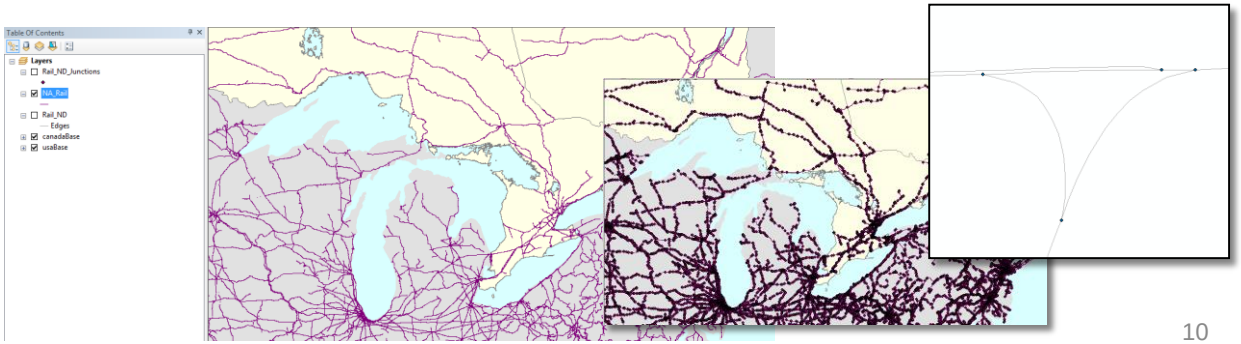
Rail Network sources. NWRN, Natural Resources Canada, US National Atlas, http://osav-usdot.opendata.arcgis.com/datasets/2553aa5e457349efb600502050bf9c3c_0, and Mexico

Required Data



Getting Started – Building A Network Dataset

- What is a network dataset?
  - Networks used by ArcGIS Network Analyst are stored as network datasets
  - Implements various parameters to direct how a rail car moves along the network
- Simple network to build – uses impedance values (no turns or elevations to model) for least impedance routing solution



Getting Started – Determining Cost Values

- Impedance Values linked to rail carriers (CN, CP and shortlines)
- Impedance calculated as ratio of length to ownership (1:1 for owned track, 1:25 for other carrier's track)
  - Ratios are arbitrary; however, trackage rights will vary depending on which rail carrier is moving on which rail line.

OBJECTID *	OPERATOENA	LENGTH_KM	CN_ IMP	CP_ IMP	SHORT_IMP
2394	Canadian National	0.3408	0.3408	8.5207	8.5206
2395	Canadian National	17.1895	17.1895	429.7384	429.7365
2396	Canadian National	2.4654	2.4654	61.6341	61.6338

OBJECTID *	OPERATOENA	LENGTH_KM	CN_ IMP	CP_ IMP	SHORT_ IMP
277	Canadian Pacific	12.4158	310.3957	12.4158	310.3943
278	Canadian Pacific	1.8236	45.5911	1.8236	45.5909
279	Canadian Pacific	9.6794	241.9849	9.6794	241.9838

OBJECTID *	OPERATOENA	LENGTH_KM	CN_IMP	CP_IMP	SHORT_ IMP
85	Great Western Railway	42.762748	1069.06871	1069.06871	42.762748
86	Great Western Railway	59.577607	1489.44018	1489.44018	59.577607
87	Great Western Railway	5.30507	132.626755	132.626755	5.30507

Methodology – Input Data

- Each rail data record indicates movement of a commodity and includes:
  - Quantity of commodity (carloads)
  - Rail carrier moving the commodity
  - Rail terminal locations (SPLC) for origin, destination, and station points along the route where commodity is loaded onto and off of the rail carrier's network (OnStat/OffStat)
  - Border crossings into or out of Canada, if applicable (OnCross/OffCross)

Year1	Month1	Carrier	Oncross	Offcross	Origin	Origarea	Origroad	Destinaton	Destinarea	Destroad	Onstat	Onroad	Offstat	Offroad	Cars
2017	2	123	99875	11224	1236	NY	555	3215	IL	697	99875	555	11224	657	65
2017	4	123		11224	4568	AB	123	4765	AL	567			11224	336	87
2017	7	321			4568	ON	687	6765	NS	321	4687	687			12
2017	8	321			9987	BC	321	4458	AB	321					5
2017	12	444	66578		1234	TX	233	4321	SK	444	66578	881			302

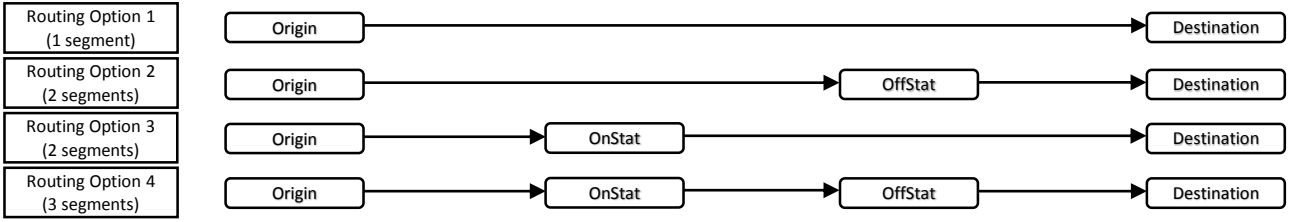
Methodology – Conditioning Input Data

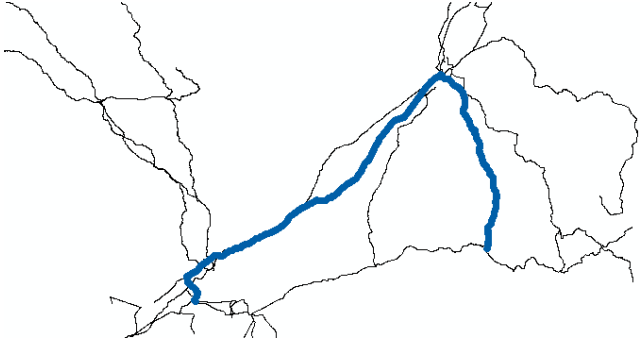
- Rail data organized into CSV (comma separated value) table and interpreted via Python script to isolate relevant information for network routing within Network Analyst

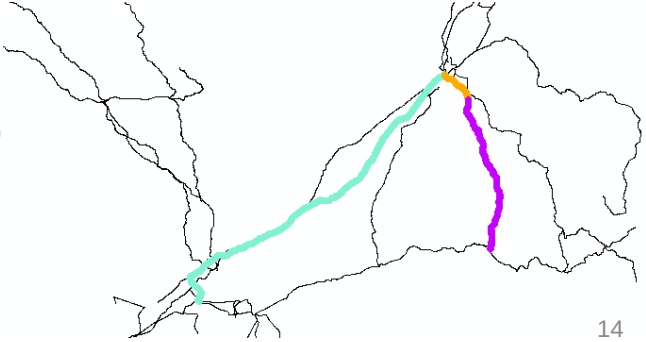
ONode	DNode	Cars	Carrier	Intermodal_Flag	TradeDir	Split	ONodeLon	ONodeLat	DNodeLon	DNodeLat	RailRecord
12345	54321	5	210	1	Can_Can	0-Orig to Dest	-12.3456	65.4321	-65.4321	12.3456	1
9875	5789	18	598	0	Can_border	1-OffStat to Dest	-98.7654	45.6789	-45.6789	88.7654	2
51325	54692	18	126	0	NoBorder	1-Orig to OffStat	-12.3456	65.4321	-65.4321	12.3456	2
55368	44893	32	235	1	Can_Can	0-Orig to Dest	-98.7654	45.6789	-45.6789	88.7654	3
6549	6489	65	888	0	border_Can	2-Orig to OnStat	-12.3456	65.4321	-65.4321	12.3456	4
63321	898736	23	301	0	Can_Can	2-OnStat to Dest	-98.7654	45.6789	-45.6789	88.7654	4

Methodology – Parse Route Segments from Rail Data

- Each rail data record parsed as origin-destination pairs to form route segments







Methodology – Parse Route Segments from Rail Data

- Script interprets records and assigns to each O-D segment:
  - a ***rail carrier*** value to calculate routing cost values for appropriate rail network (CN, CP or shortline);
  - a ***trade direction* value** (domestic, foreign or No Border) for routing in appropriate country;
  - a ***record ID*** value to link route segment to its original rail data extract record; and
  - ***lat/lon coordinates*** of corresponding SPLC (rail terminals) linked to origin and destination.
- Segments aggregated on sum of carloads per origin-destination pair

ONode	DNode	Cars	Carrier	Intermodal_Flag	TradeDir	Split	ONodeLon	ONodeLat	DNodeLon	DNodeLat	AggID
30363	33066	1	105	0	Can_Can	0-OffStat to Dest	-73.72598	45.58919	-71.2866	46.8007	1
93809	93768	2	103	0	Can_Can	0-Orig to OnStat	-122.94863	49.16206	-122.89014	49.22381	2
87465	87727	327	103	0	Can_Can	0-OffStat to Dest	-114.04736	55.16361	-117.49423	56.21361	3
49915	30196	152	103	0	Can_Can	2-Orig to OnStat	-82.43004	42.76774	-73.67929	45.47593	4
86306	64180	296	103	0	Can_Can	0-Orig to Dest	-113.18119	53.75839	-98.4909	49.979	5
51200	54500	74	103	0	Can_Can	0-OffStat to Dest	-79.45981	46.31385	-81.00498	49.0517[illegible]	[illegible]

Methodology – Logic and Validation

- Script interprets route segments with specific logic and validation checks to ensure sensible results:
  - Ignores invalid SPLC codes
    - avoids subsequent routing analysis errors
  - Detects apparent conflicts in rail carriers along route
    - Ex: Origin point with CN transfers to station point onto CN
  - Detects inconsistencies regarding border crossings
    - Ex: Origin and Destination countries are different countries, but border crossings indicate that commodity returns to same country
    - Ex: No border crossing indicated, but origin and destination countries clearly indicate cross-border movement

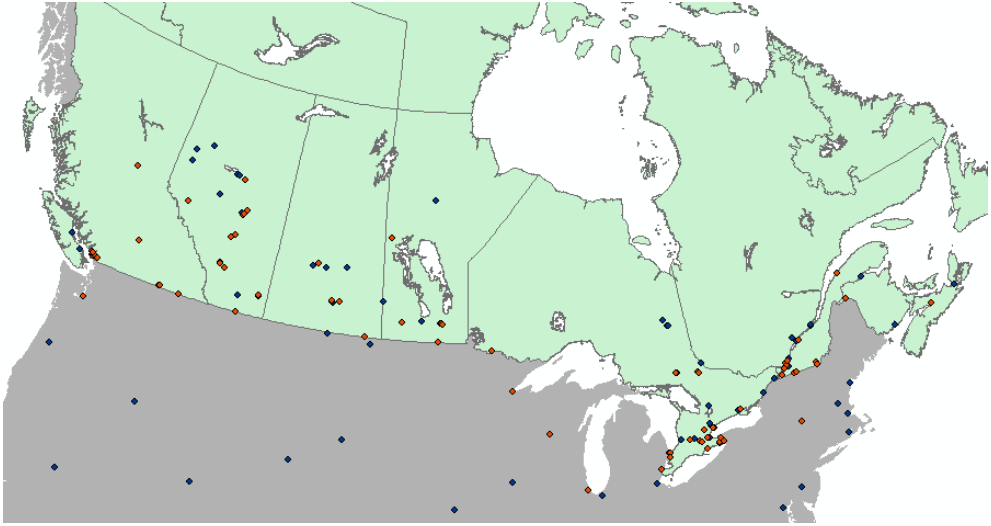
Methodology – Logic and Validation

- Records which fail logical interpretation are compiled as ‘rejected data’ for manual troubleshooting

ross	DGUNID_D	DEquiptype	OAreaNm	OCtryNm	DAreaNm	DCtryNm	Reason
	ANHYDROUS AMMONIA (UN1005)	Tank - NEED TO CONFIRM	Ontario	Canada	New York	United States	Invalid border routing - missing OnCross value'
	ANHYDROUS AMMONIA (UN1005)	Tank - NEED TO CONFIRM	Ontario	Canada	New York	United States	'Invalid border routing - missing OffCross value'
	ANHYDROUS AMMONIA (UN1005)	Tank - NEED TO CONFIRM	Ontario	Canada	New York	United States	'Invalid border routing - missing OffCross value'
	ANHYDROUS AMMONIA (UN1005)	Tank - NEED TO CONFIRM	Quebec	Canada	Maine	United States	'Invalid border routing - missing OffCross value'
	ANHYDROUS AMMONIA (UN1005)	Tank - NEED TO CONFIRM	Quebec	Canada	Maine	United States	'Invalid border routing - missing OffCross value'
	SULPHUR DIOXIDE (UN1079)	TANK CARS 12000 THROUGH 18000 GALS	Alabama	United States	Saskatchewan	Canada	Invalid SPLC: -01234'
	SULPHUR DIOXIDE (UN1079)	TANK CARS 12000 THROUGH 18000 GALS	Alberta	Canada	British Columbia	Canada	Invalid SPLC: -098765'
	SULPHUR DIOXIDE (UN1079)	TANK CARS 12000 THROUGH 18000 GALS	Alberta	Canada	British Columbia	Canada	Invalid SPLC: -065587'
	SULPHUR DIOXIDE (UN1079)	TANK CARS 12000 THROUGH 18000 GALS	Manitoba	Canada	Quebec	Canada	Invalid SPLC: -01324'
	SULPHUR DIOXIDE (UN1079)	TANK CARS 12000 THROUGH 18000 GALS	Manitoba	Canada	Quebec	Canada	Invalid SPLC: -01324'
	SULPHUR DIOXIDE (UN1079)	TANK CARS 12000 THROUGH 18000 GALS	British Columbia	Canada	Alberta	Canada	Invalid SPLC: -012324'
	SULPHUR DIOXIDE (UN1079)	TANK CARS 12000 THROUGH 18000 GALS	British Columbia	Canada	Nova Scotia	Canada	Invalid SPLC: -0112233'

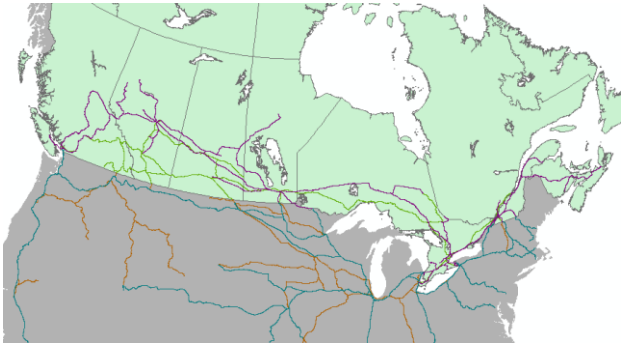
Methodology – Routing with Network Analyst

- Origin and destination points placed according to SPLC point coordinates



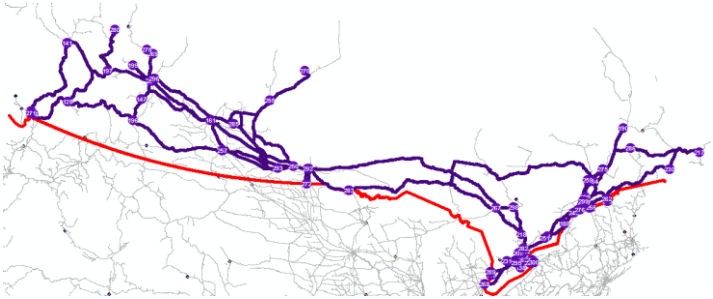
Methodology – Routing with Network Analyst

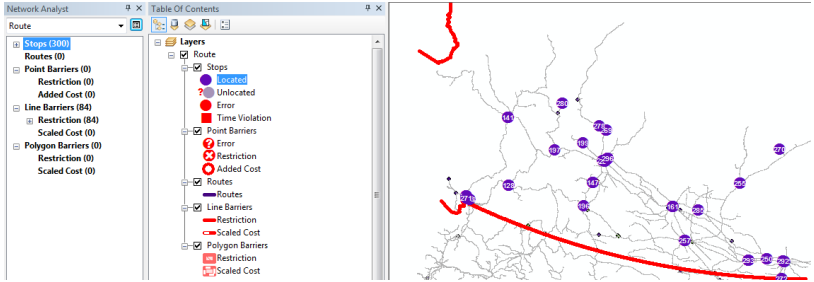
- Segments filtered by **rail carrier** (i.e. CN, CP or shortline) and **trade direction** (domestic, foreign or no border restriction)
  - Route analysis performed for each category (i.e. CN domestic, CN foreign, CN no border, CP domestic, etc.)
    - Rail carrier category dictates which cost value to be used on network (CN, CP or shortline cost fields)
    - Trade direction dictates which border restriction is implemented to force appropriate routing



Methodology – Input Route Analysis Parameters

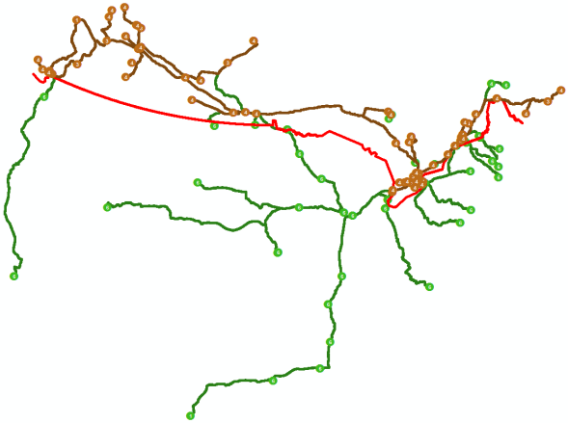
- Filter segments
- Add origin point locations to Stops
- Add destination point locations to Stops
- Add border restriction to Barriers
- Solve routing analysis





Methodology – Compile Results

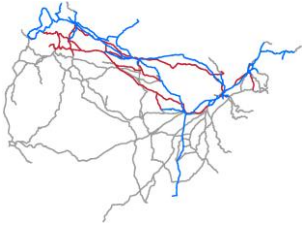
Series of geoprocessing steps are used to aggregate multiple analyses into total carload counts per segment



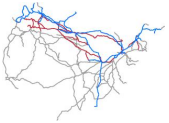
- Each result requires aggregation into single commodity flow – each segment a single record in the feature class

Methodology – Geoprocessing Results

- Merge – to merge CN/CP/shortline domestic/foreign/no border routes



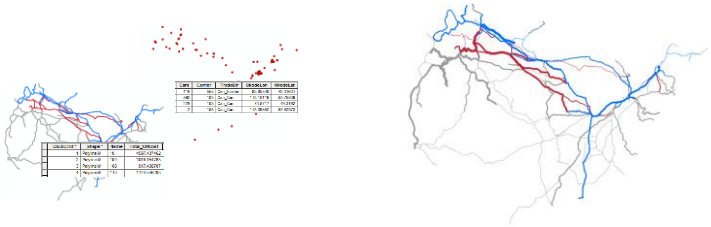
- Dissolve – to create single feature from merged rail features



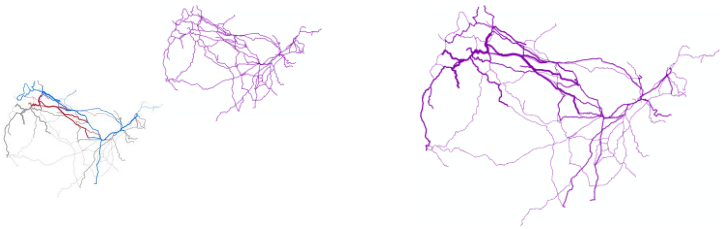
- Clip – to clip relevant portion of North American rail features with dissolved feature class

Methodology – Geoprocessing Results

- Table Join – to join carload count attribute to merge feature class



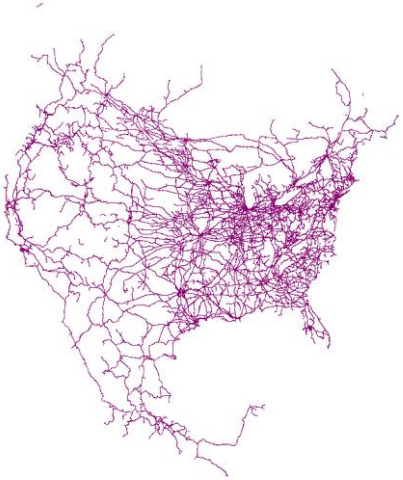
- Spatial Join – to join carload count attributes to clipped rail features



- Dissolve – to aggregate carload counts on spatially joined features

Methodology – Geoprocessing Hurdles

- Memory Issues
  - Rail network consists of millions of vertices, consistently causing memory errors when performing some of the previous geoprocessing steps even with best of breed computers.
    - Steps broken down into more manageable chunks
- Ongoing refinement of logical interpretation of rail records
  - Routing errors addressed manually on case-by-case basis, with improvements in logical interpretation included in parsing script for subsequent analyses



Methodology – Map Commodity Flow Results

- Route analysis results are applied to a map document template with standardized layout and symbology.
- Product exported as PDF for distribution.
- Future development of online commodity flows for internal use only are in the works.

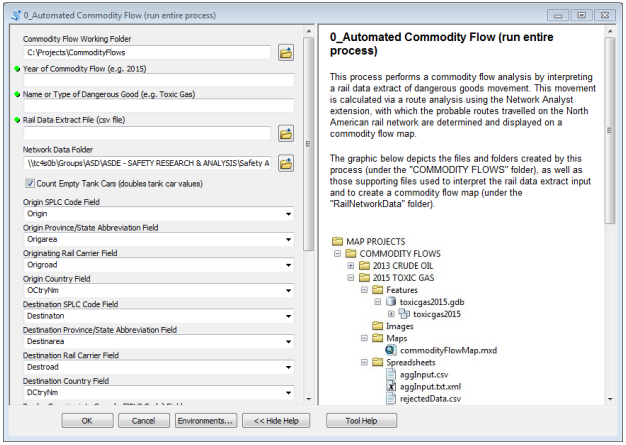
Automation

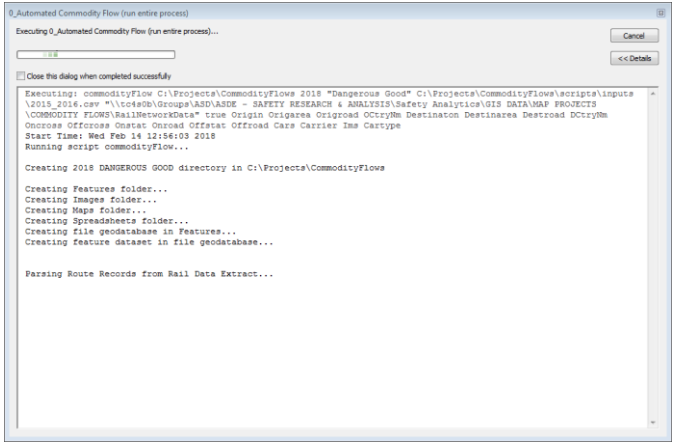
- Every step of the commodity flow analysis is automated for efficiency and repeatability
  - Gain in efficiency – two week manual process condensed to 15-45 minute automated process

CommodityFlow.pyt
- 0_Automated Commodity Flow (run entire process)
- 1_Create Folder Directory
- 2_Parse Route Records from Rail Data Extract
- 3_Create Point Feature Classes from Route Records
- 4_Calculate Route Layers
- 5_Aggregate Route Carload Values
- 6_Map Commodity Flow Results

Automation

- Python toolbox developed to execute commodity flow analysis using convenient interface for input

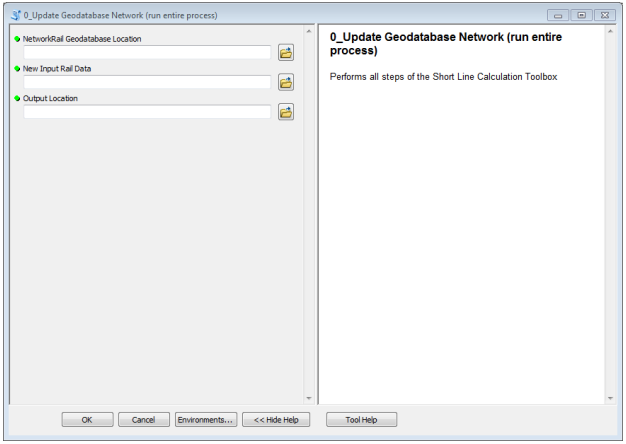




Automation

- Workflow to update data for Network Dataset creation with latest information also automated

ShortLineCalculation_Toolbox(FINAL).pyt
- 0_Update Geodatabase Network (run entire process)
- 1_Archive Geodatabase
- 2_Copy Features
- 3_Short Line Calculation

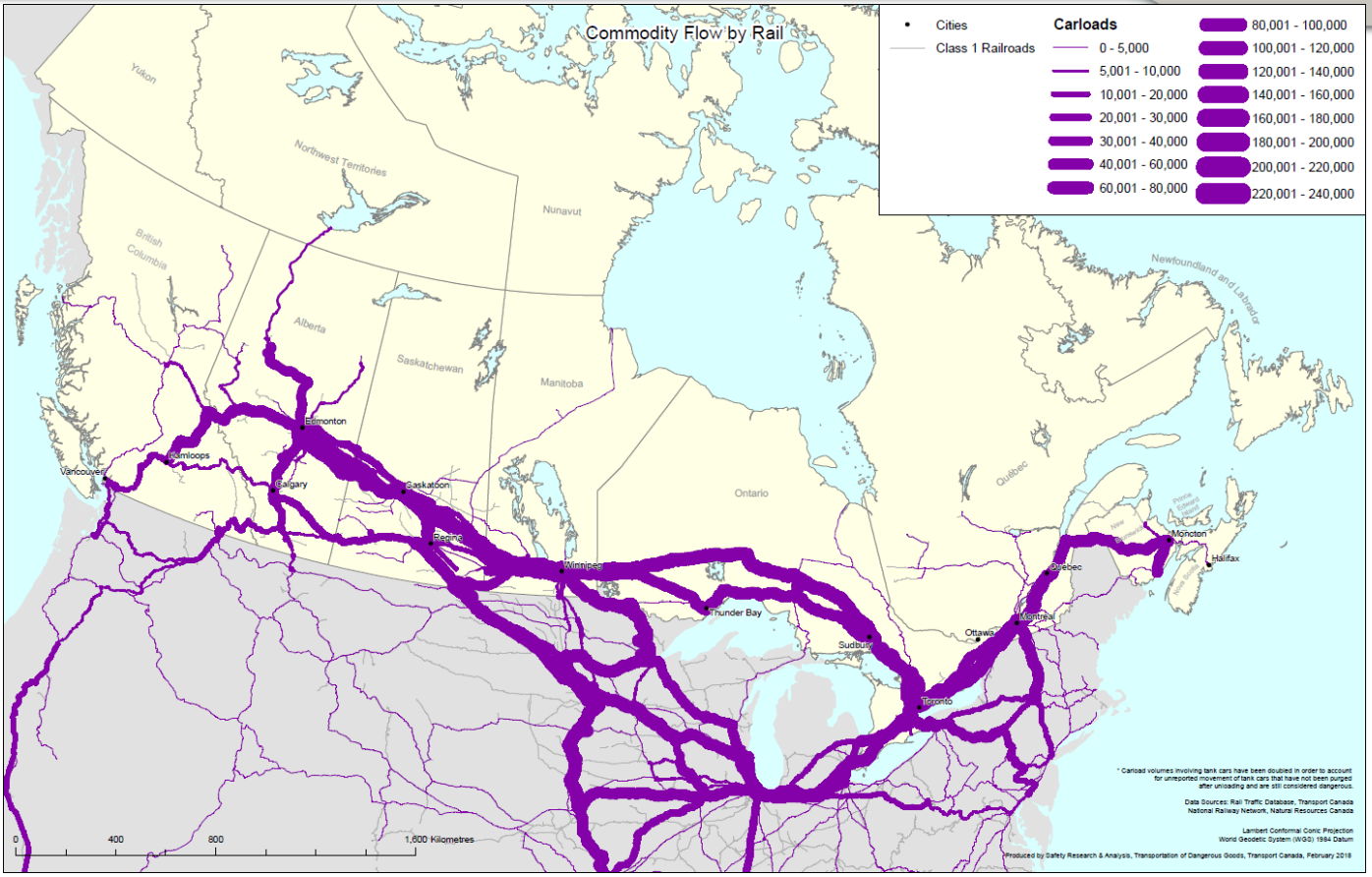


Benefits

- Automation of process greatly improved efficiency, allows for rapid turnaround of information regarding movement of dangerous commodities by train;
- We can improve on the logic to create a more accurate simulation of the flow of traffic. As new information becomes available, we can continue to improve on that accuracy.
- Commodity flows can be used to spatially analyze data across multiple variables in order to assess safety in transportation:
  - Emergency response services
  - Population density
  - Sensitive infrastructure
  - Topography

Next steps

- Continue to be proactive in risk mitigation.
- Increase evidence-based policy decisions.
- Commodity flows on modes other than rail, i.e. trucks.
- Continue to improve on logic to improve simulation.


Commodity Flow by Rail
Cities
Class 1 Railroads
Carloads
0 - 5,000
5,001 - 10,000
10,001 - 20,000
20,001 - 30,000
30,001 - 40,000
40,001 - 60,000
60,001 - 80,000
80,001 - 100,000
100,001 - 120,000
120,001 - 140,000
140,001 - 160,000
160,001 - 180,000
180,001 - 200,000
200,001 - 220,000
220,001 - 240,000
Yukon
Northwest Territories
Nunavut
British Columbia
Alberta
Saskatchewan
Manitoba
Ontario
Quebec
Newfoundland and Labrador
Edmonton
Kamloops
Vancouver
Calgary
Saskatoon
Regina
Winnipeg
Thunder Bay
Sudbury
Ottawa
Toronto
Montreal
Quebec
Moncton
Halifax
0
400
800
1,600 Kilometres
* Carload volumes involving tank cars have been doubled in order to account for unreported movement of tank cars that have not been purged after unloading and are still considered dangerous.
Data Sources: Rail Traffic Database, Transport Canada
National Railway Network, Natural Resources Canada
Lambert Conformal Conic Projection
World Geodetic System (WGS) 1984 Datum
Produced by Safety Research & Analysis, Transportation of Dangerous Goods, Transport Canada, February 2018

Questions?

Lead Developer – David Hennessy

Methodology – David Hennessy, Priti Zachariah

Assistant Developers – Ethan Howieson, Ahmed El-Mokadem

Crown copyright protects works created under the direction and control of Transport Canada.