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
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.
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

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.

<table>
<thead>
<tr>
<th>OBJECTID *</th>
<th>OPERATOENA</th>
<th>LENGTH_KM</th>
<th>CN_IMP</th>
<th>CP_IMP</th>
<th>SHORT_IMP</th>
</tr>
</thead>
<tbody>
<tr>
<td>2394</td>
<td>Canadian National</td>
<td>0.3408</td>
<td>0.3408</td>
<td>8.5207</td>
<td>8.5206</td>
</tr>
<tr>
<td>2395</td>
<td>Canadian National</td>
<td>17.1895</td>
<td>17.1895</td>
<td>429.7365</td>
<td>429.7365</td>
</tr>
<tr>
<td>2396</td>
<td>Canadian National</td>
<td>2.4654</td>
<td>2.4654</td>
<td>61.6341</td>
<td>61.6338</td>
</tr>
<tr>
<td>277</td>
<td>Canadian Pacific</td>
<td>12.4158</td>
<td>310.3957</td>
<td>12.4158</td>
<td>310.3943</td>
</tr>
<tr>
<td>278</td>
<td>Canadian Pacific</td>
<td>1.8236</td>
<td>45.5911</td>
<td>1.8236</td>
<td>45.5909</td>
</tr>
<tr>
<td>85</td>
<td>Great Western Railway</td>
<td>42.762748</td>
<td>1069.06871</td>
<td>1069.06871</td>
<td>42.762748</td>
</tr>
<tr>
<td>86</td>
<td>Great Western Railway</td>
<td>59.577607</td>
<td>1489.44018</td>
<td>1489.44018</td>
<td>59.577607</td>
</tr>
<tr>
<td>87</td>
<td>Great Western Railway</td>
<td>5.30507</td>
<td>132.626755</td>
<td>132.626755</td>
<td>5.30507</td>
</tr>
</tbody>
</table>
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)
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
Methodology – Parse Route Segments from Rail Data

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

Routing Option 1
(1 segment)

Routing Option 2
(2 segments)

Routing Option 3
(2 segments)

Routing Option 4
(3 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
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
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
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:
  o Emergency response services
  o Population density
  o Sensitive infrastructure
  o 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.
Questions?
Lead Developer – David Hennessy

Methodology – David Hennessy, Priti Zachariah

Assistant Developers – Ethan Howieson, Ahmed El-Mokadem