Big Data Analysis to Measure Delays of Canadian Domestic and Cross-Border Truck Trips

Kevin Gingerich and Hanna Maoh
University of Windsor

Freight Day VI Symposium
Hosted by the University of Toronto Transportation Research Institute - Centre for Urban Freight Analysis
Wednesday March 1, 2017

Agenda

- **Introduction**
- **Data Overview**
  - GPS Data
  - Converting raw data into observable trips
- **Trip delays**
  - Measuring delay as a function of the total trip (instead of just single location events)
  - Identifying cross-border delays
  - Separating expected / unexpected delays

Definitions

- Big Data for transportation can be categorized as follows (ITF, 2015[1]):

- **Opportunistic sensing**
  - Collected for one purpose, used for another
  - Such as GPS data from trucks

Focus of this presentation on truck derived GPS data

- **Purposed sensing**
  - Collected for the purpose it is designed for
  - Such as fixed detectors, etc.

- **Crowd sensing**
  - Using content sharing platforms such as facebook, twitter, etc. to understand individuals and their travel patterns

[1]*International Transport Forum (ITF) 2015. - Big Data and Transport: Understanding and Assessing Options.*

Data

- In this presentation, we use a dataset of **GPS pings** observing truck movements across Canada and the continental U.S.

Data Statistics

For a one month period occurring in July, 2013

30,000 Canadian owned trucks

Owned by 580 carriers

- Each GPS ping provides the location of a truck at a particular point in time

Carrier ID	Truck ID	Latitude	Longitude	Day	Time
1042	554	48.47848	-114.14864	20130302	145845
1042	589	52.54987	-108.13242	20130309	224532
1165	1147	47.34894	-109.78547	20130328	062234

Data shown here is artificial

Stop Events

- As part of the data processing, the purpose of a stop event is estimated as either:

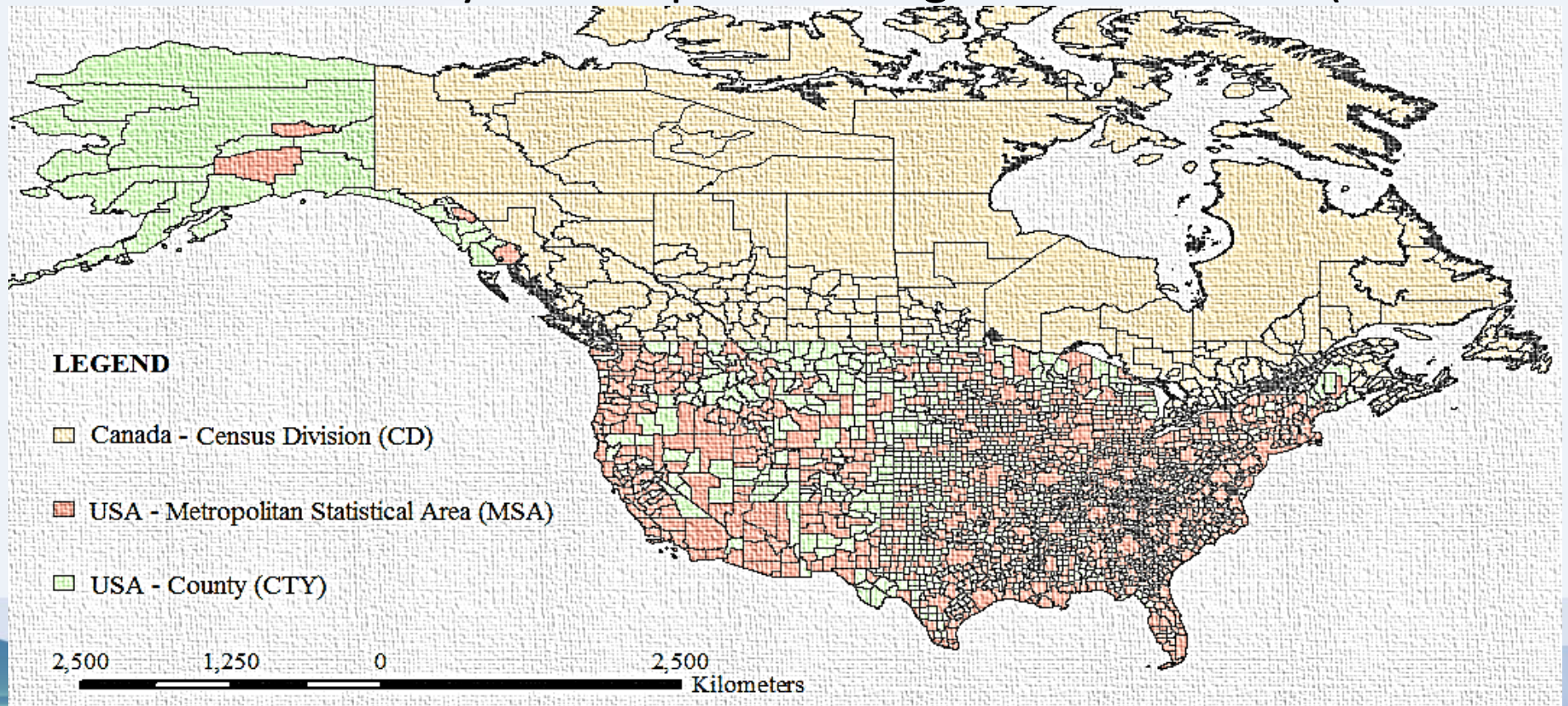
1. **Primary stop** – Vehicle is stopped to transfer goods
2. **Secondary stop** – Vehicle is stopped for other purposes such as fuel refills, driver break, etc.

- The **primary stops** are used as end points to define ***trips***
  - 221,800 trips derived for the one month period

Trip Zones

- The trips are aggregated (for both origins and destinations) by zones at the census division level (in Canada) and MSA or county level (In the U.S.)

- The focus of this study is on trips travelling between zones (inter-zonal)

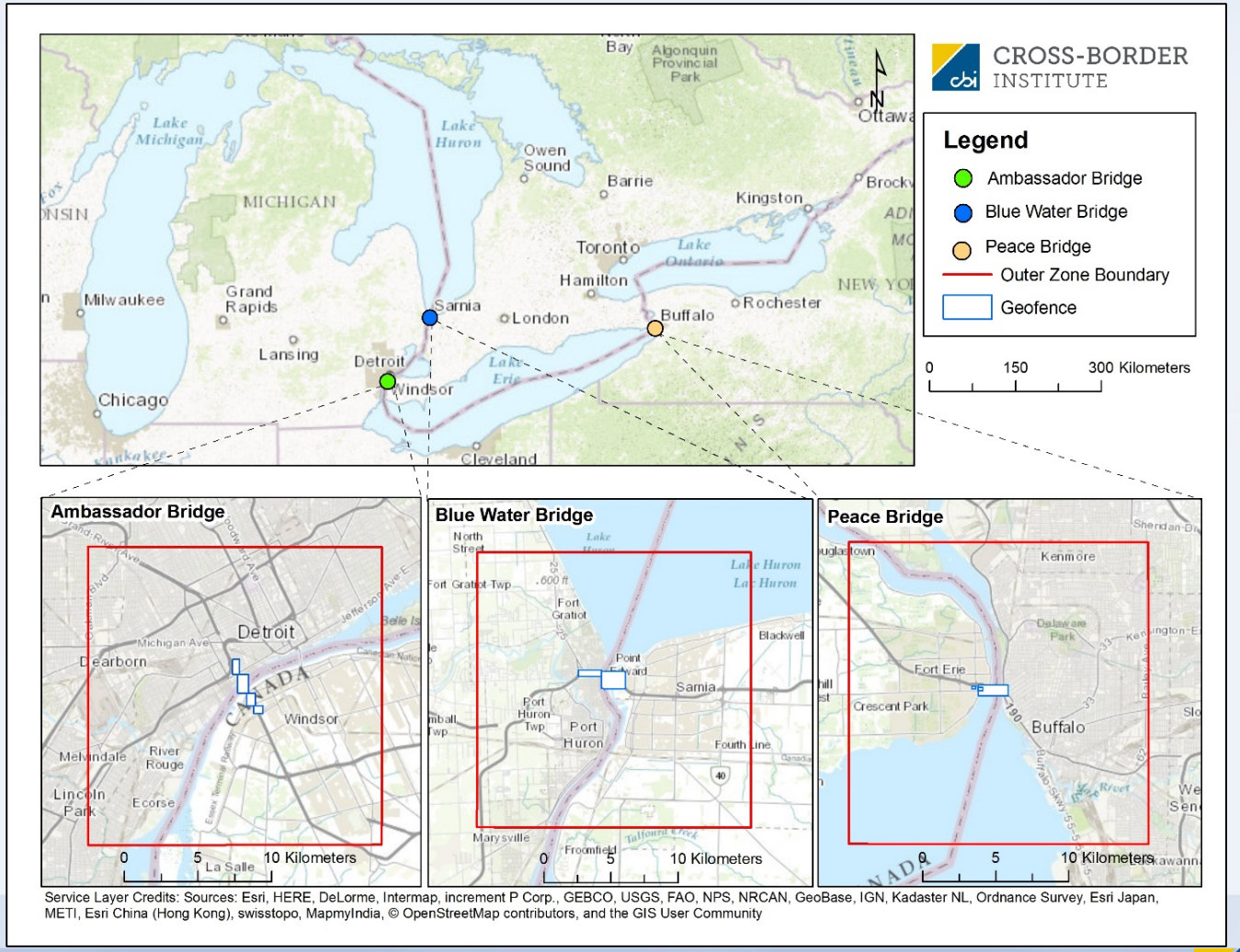


Trip Delays

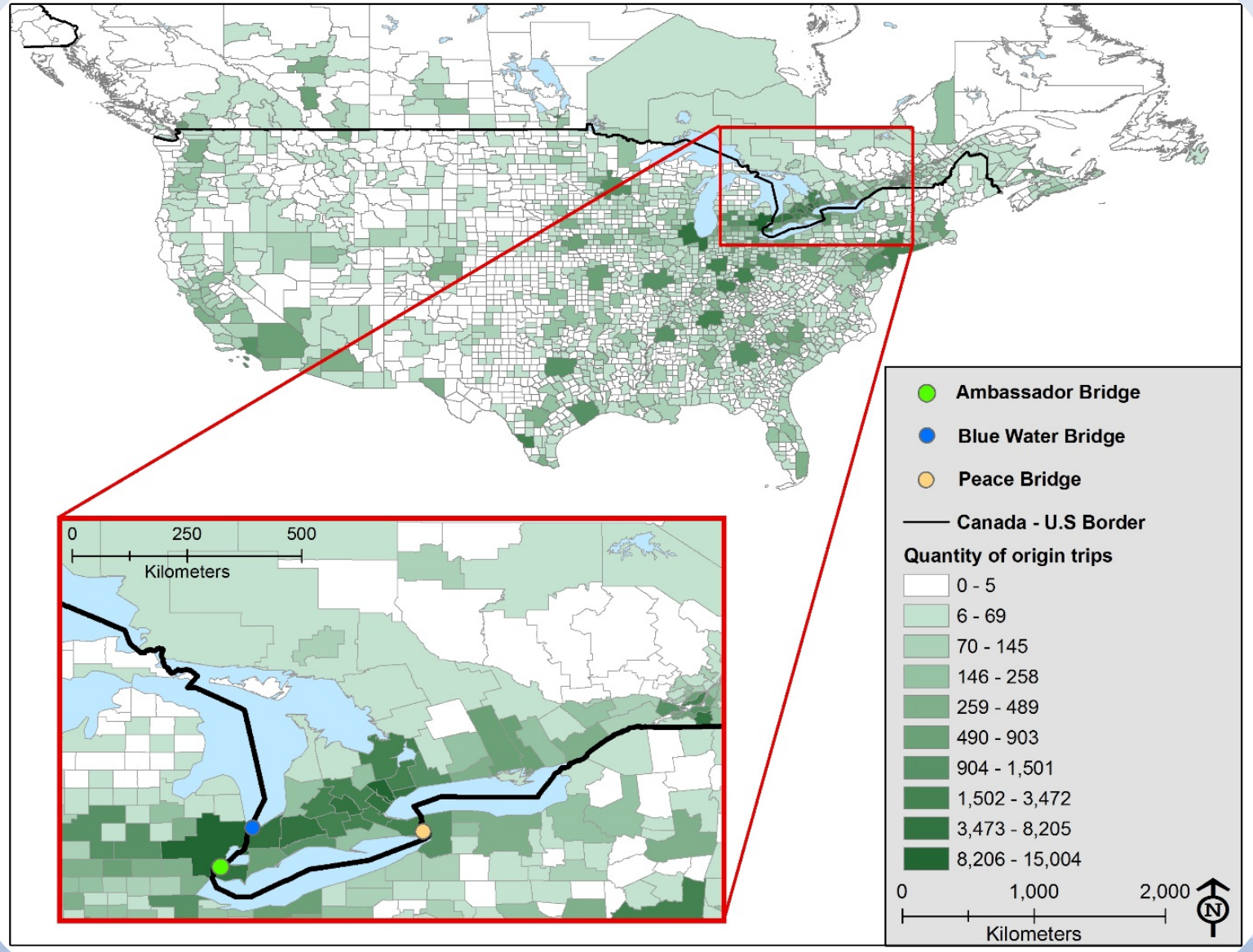
- **Part 1: Border delays**

Border Locations

- Border wait times are analyzed at three border crossings :
  - Blue Water Bridge
  - Ambassador Bridge
  - Peace Bridge



Border Crossing Trips by Origin Frequency

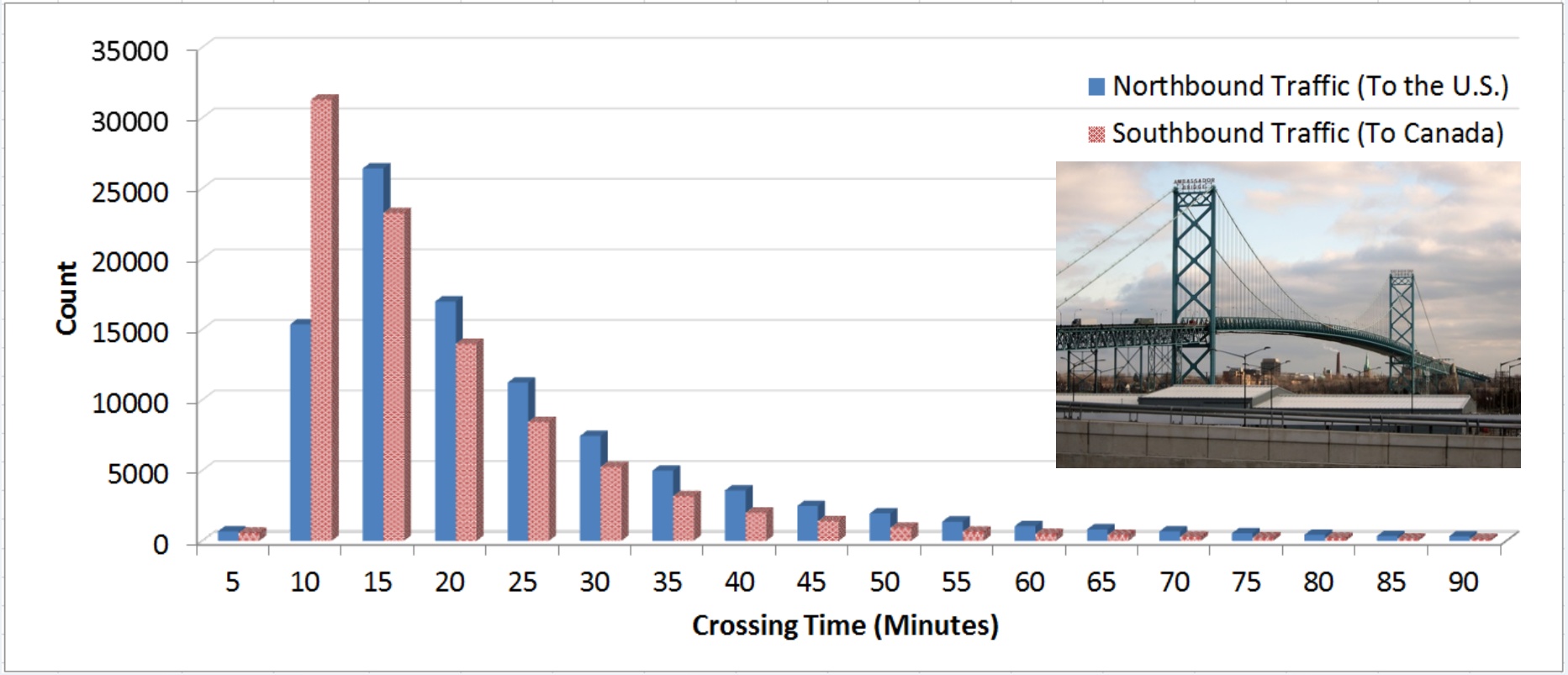


Border Crossing Statistics (minutes)

	Ambassador Bridge		Blue Water Bridge		Peace Bridge	
	CAN bound	USA bound	CAN bound	USA bound	CAN bound	USA bound
Minimum	3.35		4.61		2.58	
Median	13.6	14.3	11.3	13.1	12.3	22.0
Average	17.6	18.9	16.8	18.2	17.6	27.1
95 Percentile	42.3	48.3	48.7	48.0	47.6	69.2

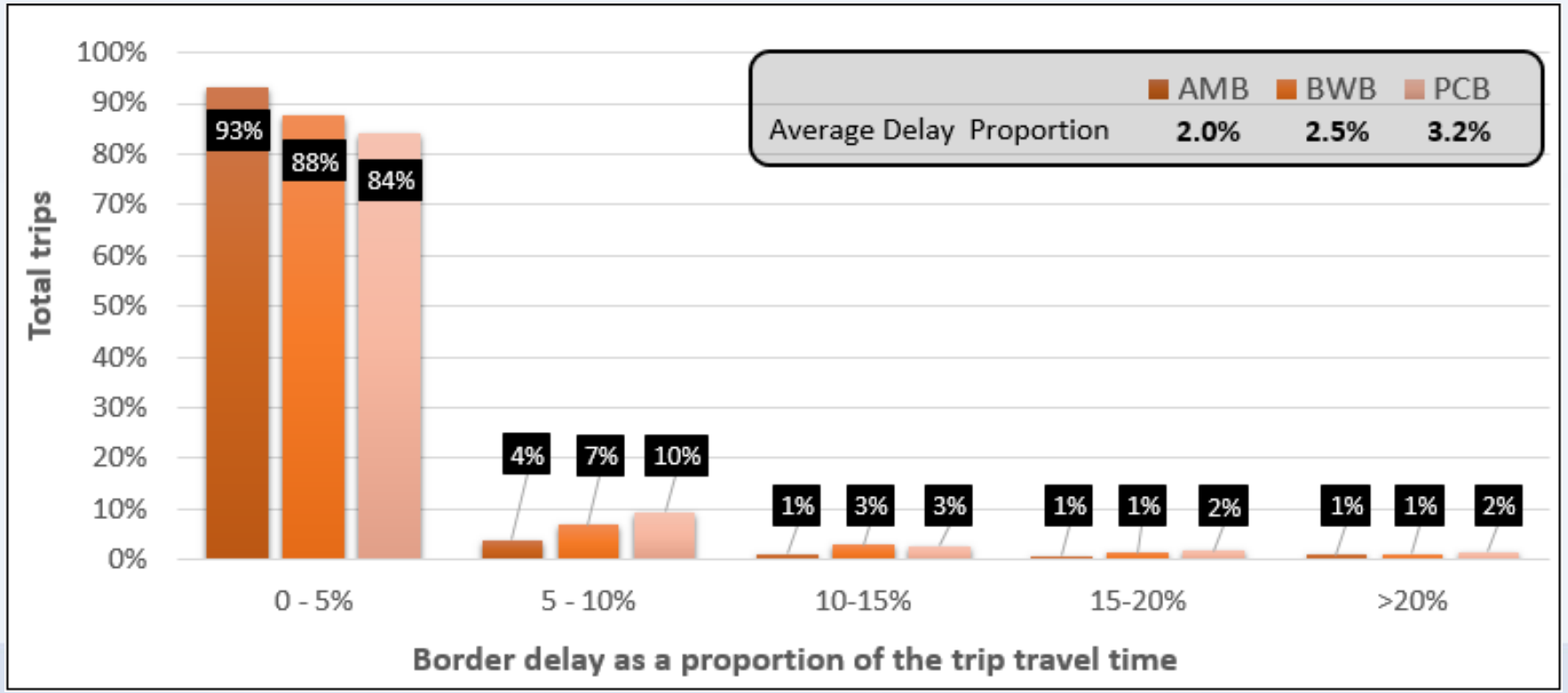
Border Crossing Distributions

- Crossing time distribution – Ambassador Bridge

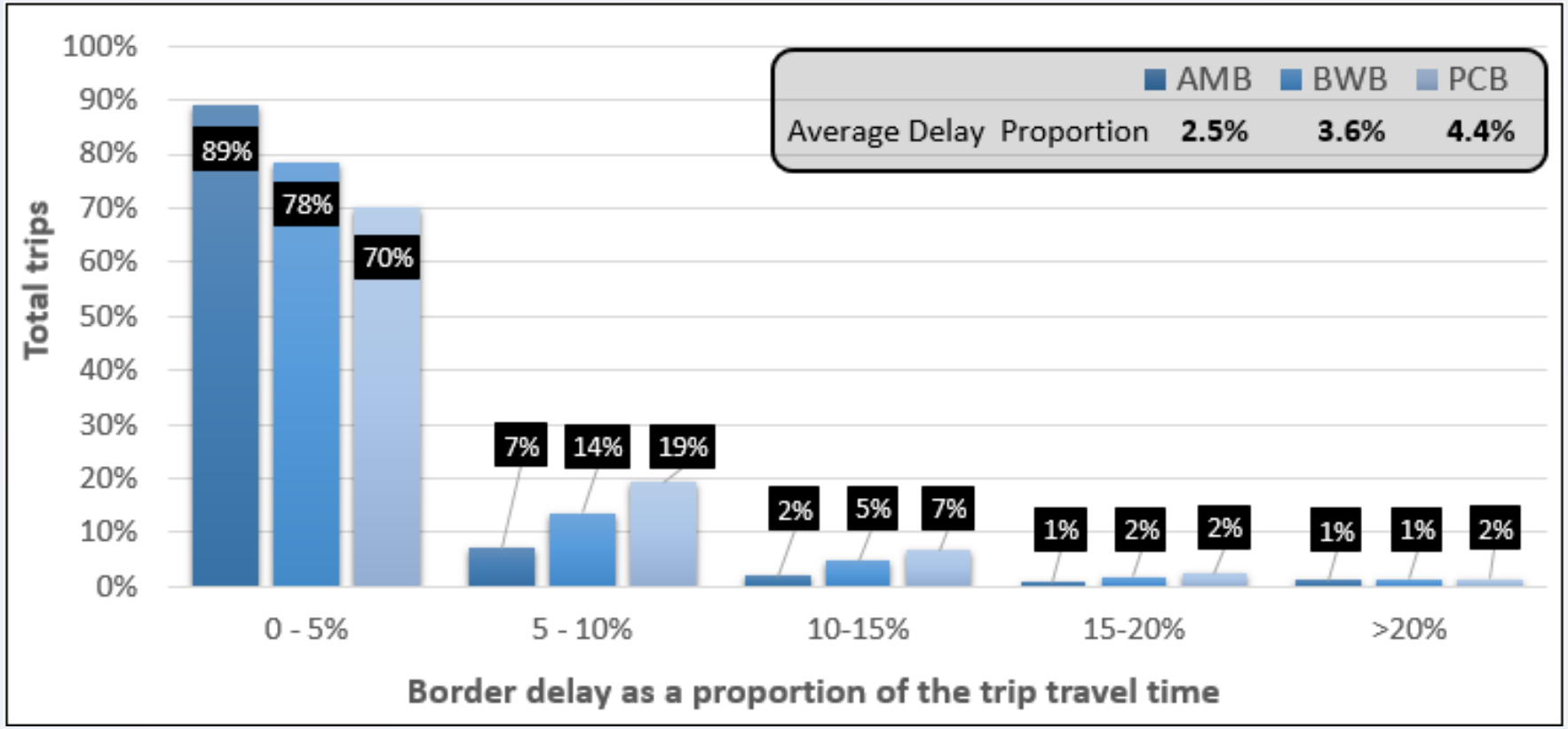


Canada Bound Border Delays

- We can determine the delay as a proportion of the total travel time since we know:
  1. the delay at the bridge
  2. the total travel time



U.S. Bound Border Delays



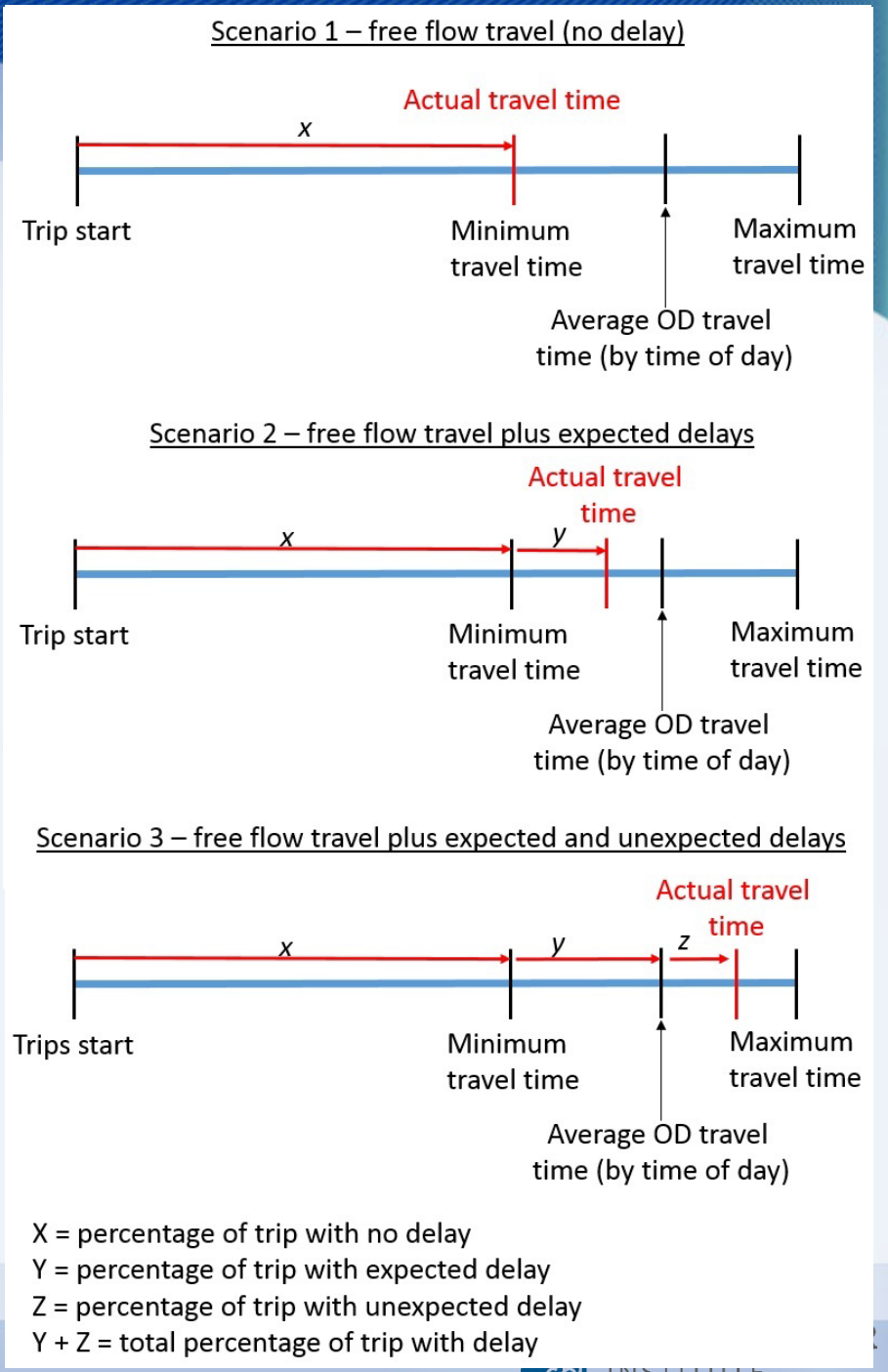
Trip Delays

- **Part 2: Expected / Unexpected Delays**

Expected / Unexpected Delays

- To further analyze delays, **all trips** (not just int'l trips) were broken down into proportions of:
  1. Free flow travel (no delay)
  2. Expected delay
  3. Unexpected delay

- Calculations require **the trip travel time** and the **minimum/average** travel times for any trip between the origin-destination zones and during the specified time of day



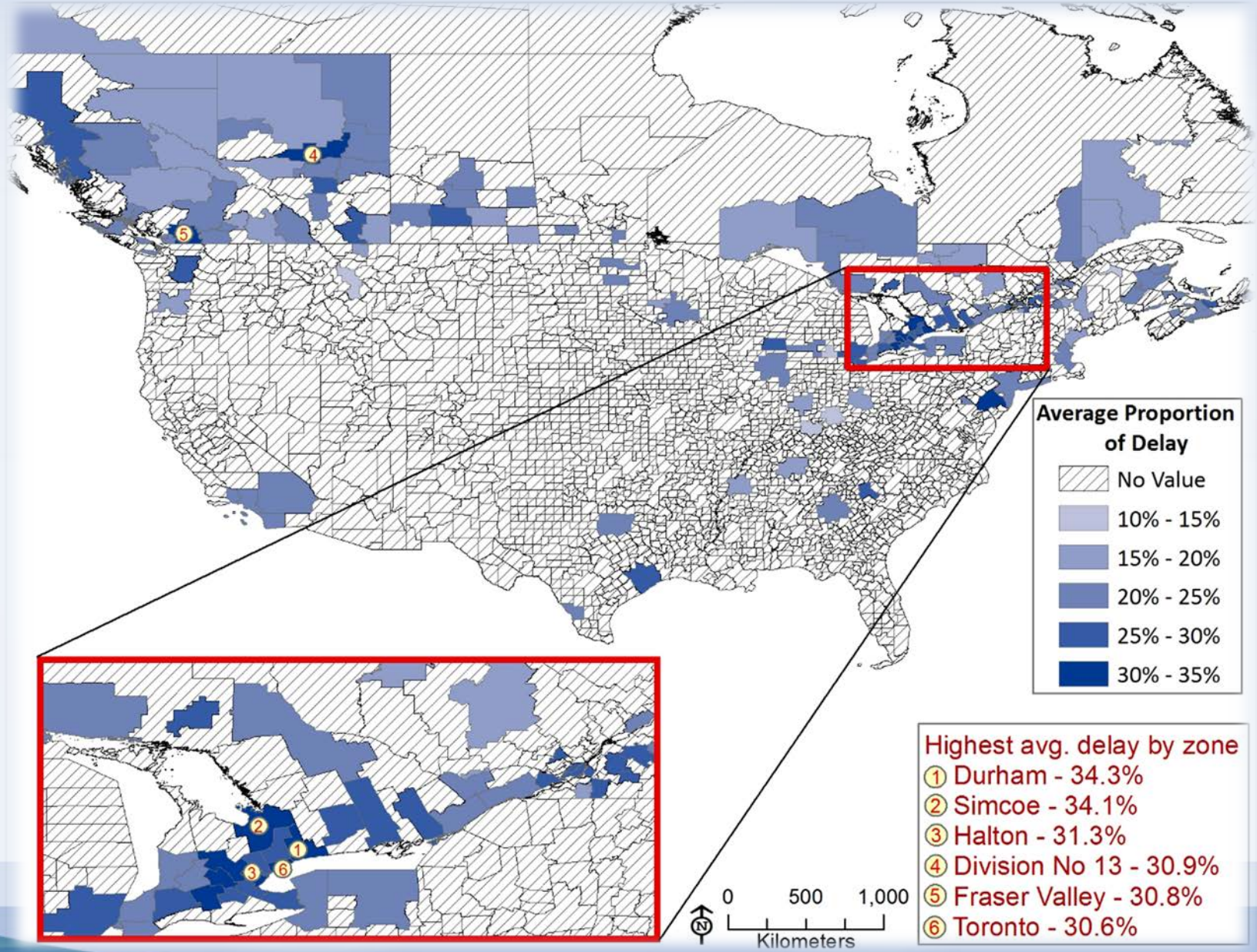
Trip Delay Results

- Only origin destination (OD) zonal pairs with at least **40 trips** were retained
  - Reducing the trip count from **221,807** to **83,654** trips belonging to **756** zonal pairs

- Average proportions include **75%** for the free flow travel time (no delay); **19%** for expected delays; **6%** for unexpected delays

- Average total delays for each origin/destination zone were also calculated
  - Range of values from **10% to 34%**

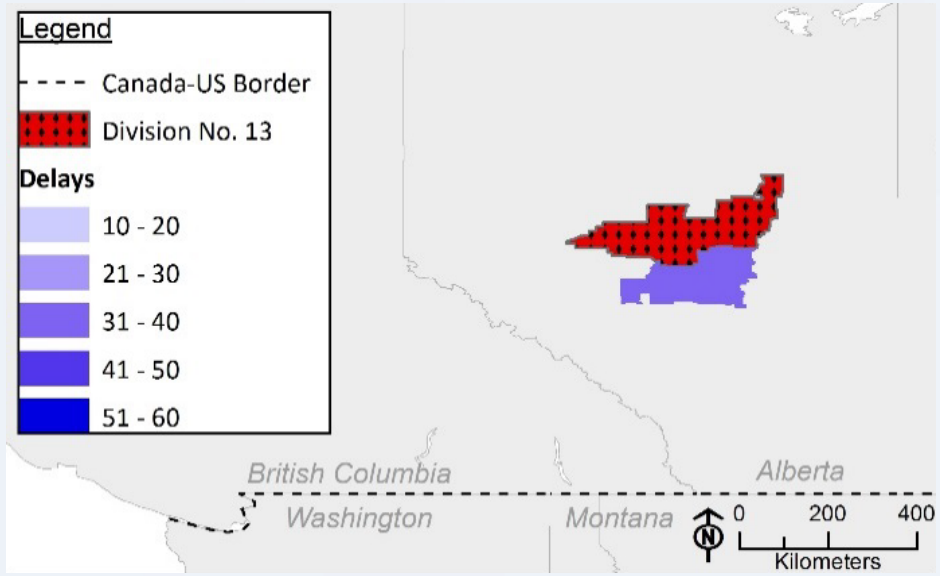
Average zone delays



Connected zones pairs

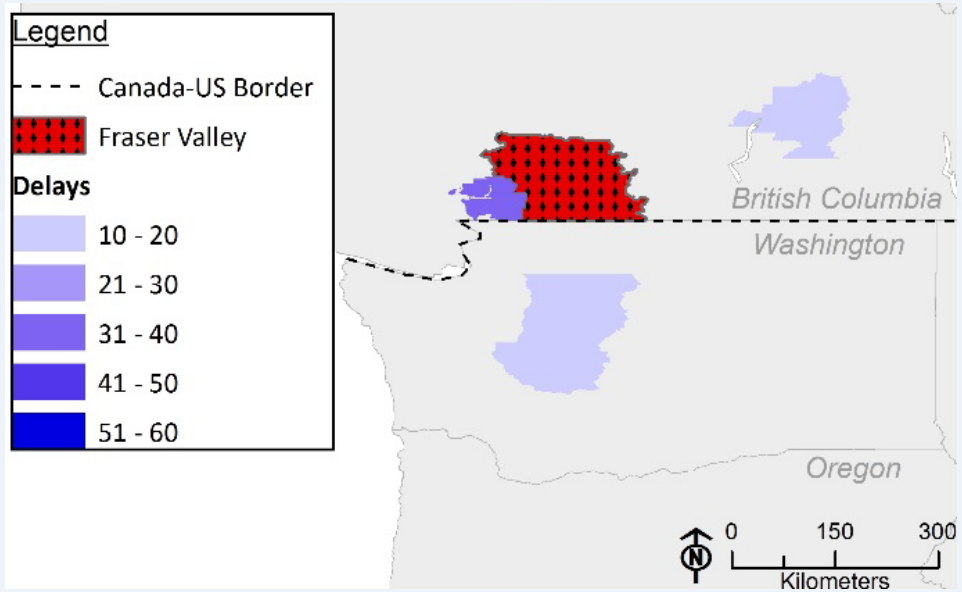
- While 2 western province locations exhibit high delays, they are based on a very small sample of origin-destination pairs

4



Only one suitable zone pair

5

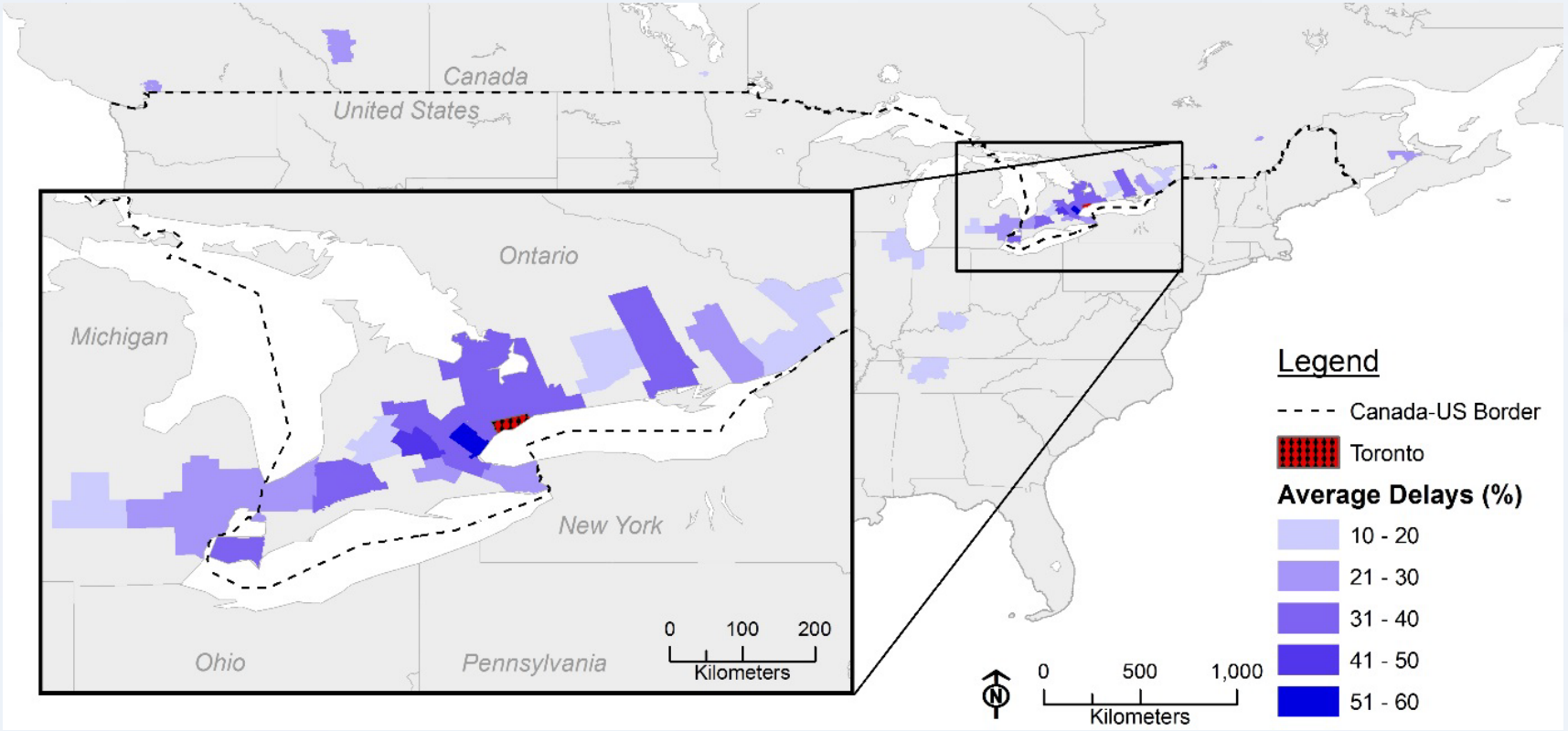


Only three suitable zone pairs

Connected zones pairs

- By contrast, the Toronto zone has a large number of zonal pairs (32)

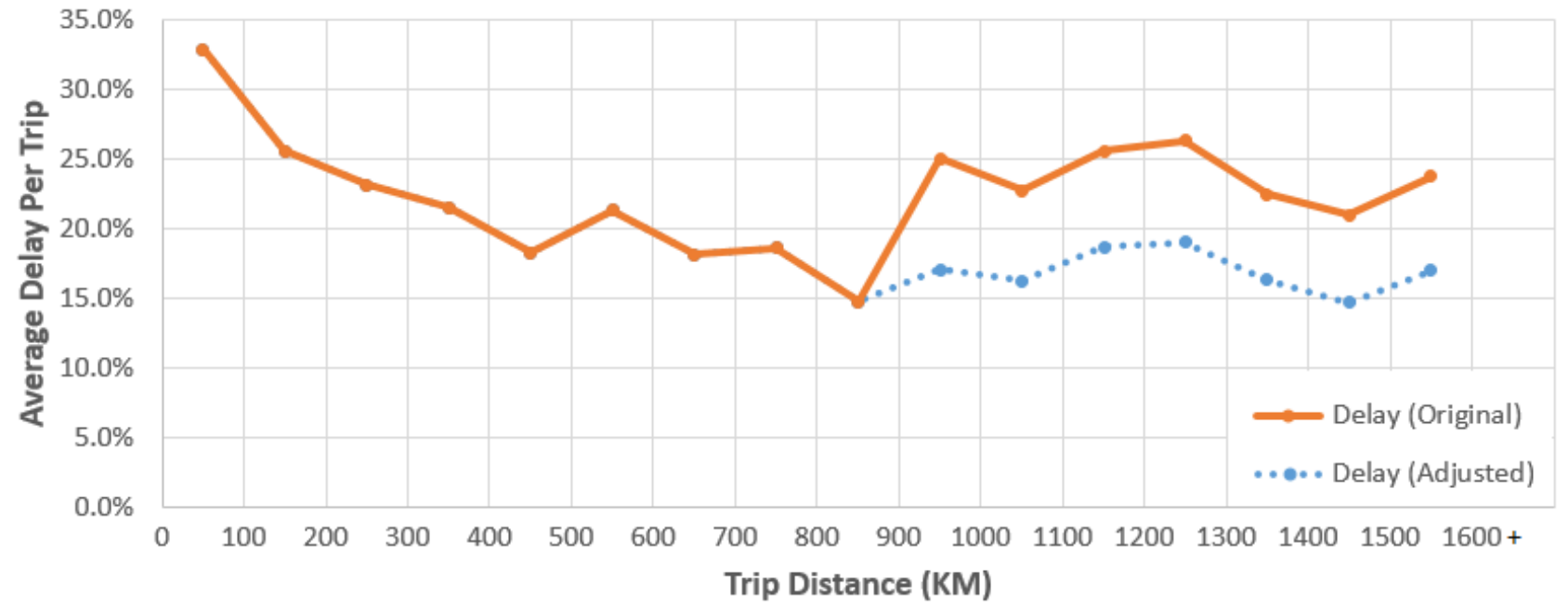
6



Trip delay and distance

- Delay and distance relationship suggests that short delays tend to have higher delays (as a proportion of the entire trip)

- The adjusted line is a better representation of the actual pattern
  - The increase in the original line is due to a different cut-off time for valid trips that are larger than 900 km



Acknowledgements

- We would like to thank **Transport Canada**, who loaned the GPS data used in this study
  - in addition to methodologic discussions provided by **Louis-Paul Tardif** and **Andrew Carter**

- We are also thankful to the financial contributions provided by NSERC

Thank you for watching!