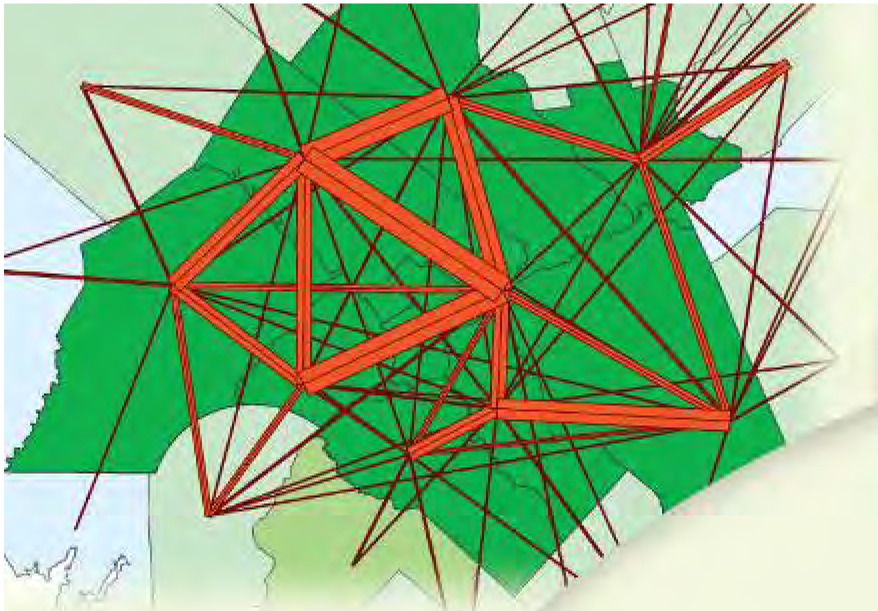
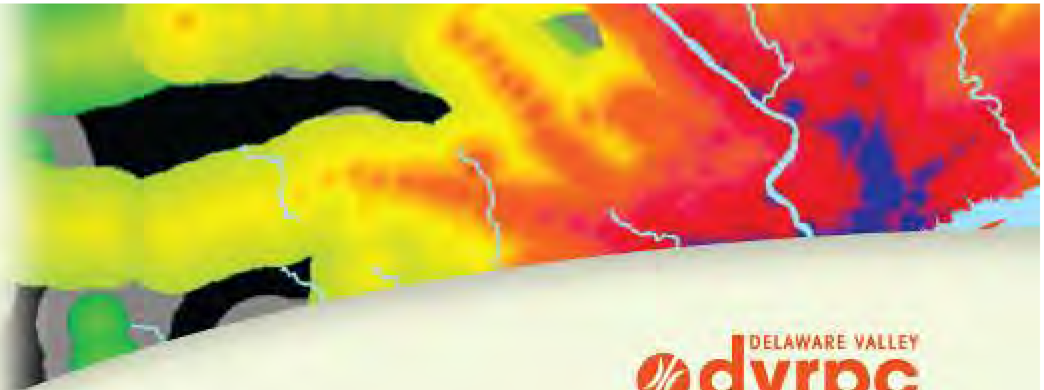




DELAWARE VALLEY REGIONAL PLANNING COMMISSION (dvrpc)

DVRPC's New Regional Network Model -

Using Google Transit Feeds and OpenStreetMap in Transportation Planning

March, 2011

IREG

Presented by:

Wolfgang Scherr

Office of Modeling and Analysis

Outline

- Travel forecasting at DVRPC
    - The regional travel model (TIM)
    - The network model
- New data sources
    - Google Transit feeds (GTFS)
    - OpenStreetMap (OSM)
    - Other
- Creation of the network model
- Applications
- Conclusions

What is dvrpc ?

The Delaware Valley Regional Planning Commission

- The MPO for the Greater Philadelphia region
- Population: 5.5 million
- 9 counties, 353 municipalities in PA and NJ
- Governed by a board of county and state representatives
- Planning for growth and development of the region

DVRPC's Travel Modeling

Office of Modeling

- 8 full time permanent staff
- Responsible for the majority of travel forecasts in the region
- Region-wide forecasts:
  - Air quality conformity analysis
  - Long range planning
- Detailed forecasts for subarea studies:
  - Highway alternatives analysis
  - Transit studies / FTA New Starts

The regional travel forecasting model TIM

- TIM 1.0 – Migration: TRANPLAN to VISUM
  - Since April 2010 used in all new projects
- TIM 2.0 – in the works …
  - New demand model
  - New network model
- TIM 3.0 – Future, activity-based model (ABM)

Purpose of the Network Model

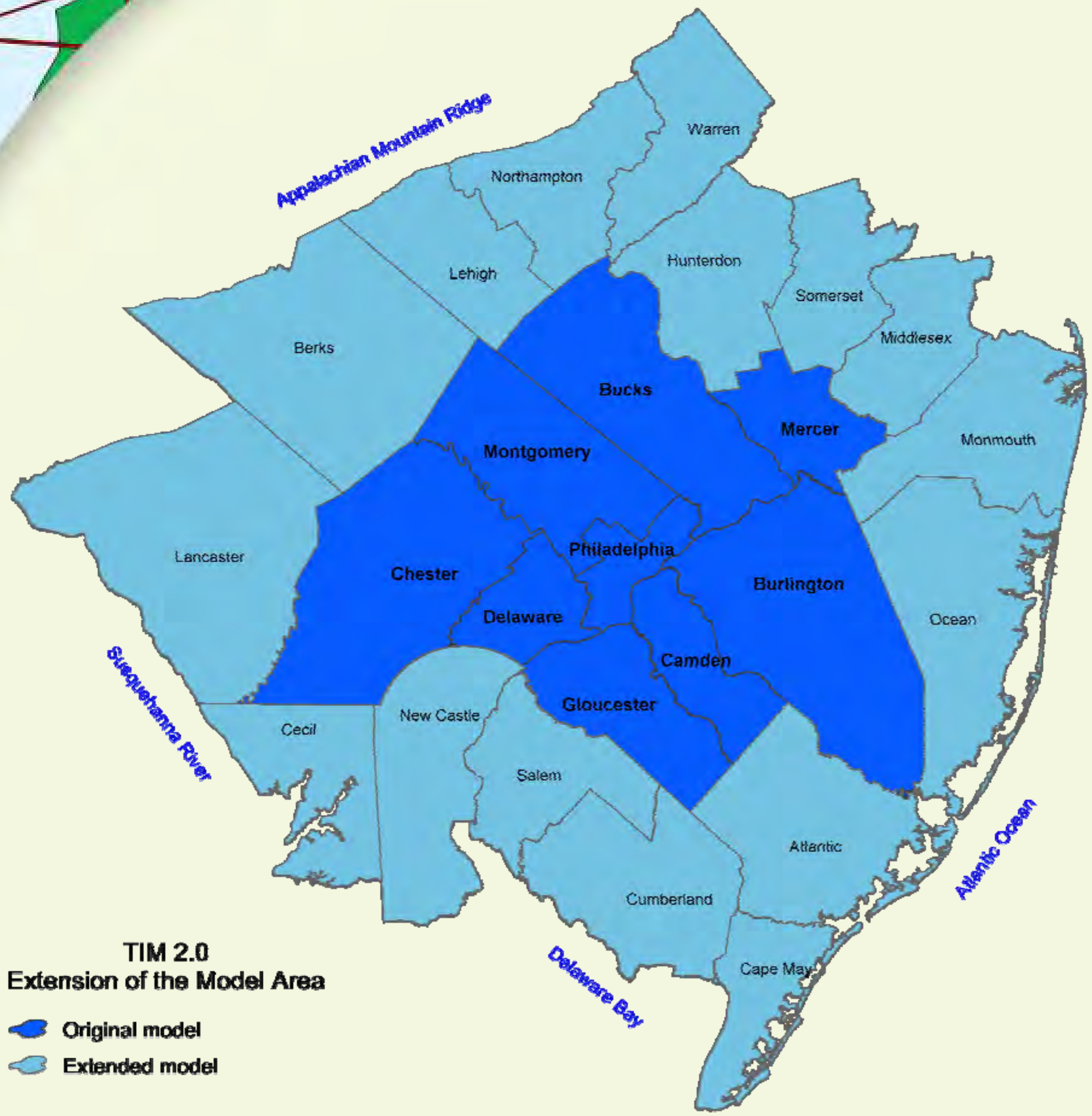
The network model

- represents transport supply in the forecasting model
- allows to compute for the entire region:
  - Travel times
  - Travel cost
  - Routes (paths) for highway vehicles
  - Routes (paths) for transit passengers

Requirements

- Geographically accurate
- Routable street data
- Reliable transit service data
- Automated updates if possible

Model Area Extension in TIM 2.0



Counties

- 9 member counties in TIM 1.0
- 25 counties in TIM 2.0

Benefits of the extension:

- 99% of journey to work shed
- better forecast at borders
- easier start-up
  of inter-MPO projects

Travel analysis zones in TIM 2.0

- internal 3150 (up from 1900)
- external 250
- total 3400

Data Sources for TIM 2.0 Network

When looking for data sources, we had choices:

- Government GIS
- Commercial street data: NAVTEQ, TeleAtlas
- Web 2.0 data

Web 2.0

- Everybody can contribute to the content
- “Mash-up”: data from various sources are integrated
- Examples: Wikipedia, Facebook etc.

Geography and transportation are emerging in web 2.0

- Voluntary geographic information (VGI)
  - Everybody can contribute to electronic maps
  - Everybody can use the data
- Examples: OpenStreetMap, Wikimapia, Google Earth

Data Sources for TIM 2.0 Network

Data sources chosen for TIM 2.0:

- OpenStreetMap (OSM) – main source streets
- Google transit feeds (GTFS) – main source transit
- Regional GIS
  - Street data for two counties (Philadelphia, Montgomery)
  - Boundaries, hydrography

Why web 2.0 data?

- OSM: Overall fair data quality, routability
- GTFS: Accurate schedules, regular updates, standardization
- Both: No copy right restrictions, no cost

The OpenStreetMap

The OSM foundation

- www.osm.org
- Non-profit, based in the U.K., founded in 2006
- Provides the organizational framework

Data distribution

- Can be used for any commercial or non-commercial purpose
- Free of charge
- Distributed under a “Creative Common” license agreement
  - As users of the data, we give reference to the copy right

Data content

- Routable street network plus other geography
- U.S. data derived from an import of the 2005 TIGER file
- In the U.K., the data start being used in commercial navigation services

Volunteers

- They generate the map
  - Uploading data from their private GPS devices
  - Editing directly on www.osm.org
- Currently no active user group in the Delaware Valley

Google Transit Feeds (GTFS)

The GTFS format

- = General Transit Feed Specification
- released by Google as open-source format in 2005
- has emerged as new industry standard for transit data
- Many applications use the data, not only Google Transit™

Data content

- Operations schedule with routes, patterns, vehicle/train trips
- Stop locations (GPS detail)
- Optional: fare information, route alignments

Data distribution

- Over 170 transit agencies in the U.S. and Canada publish schedules in GTFS
- In the DVRPC region:
  - SEPTA (www.septa.org/developer)
  - NJ Transit (njtransit.com/developer)
  - PATCO since 2010
  - AMTRAK, TMACC: working on their first feed

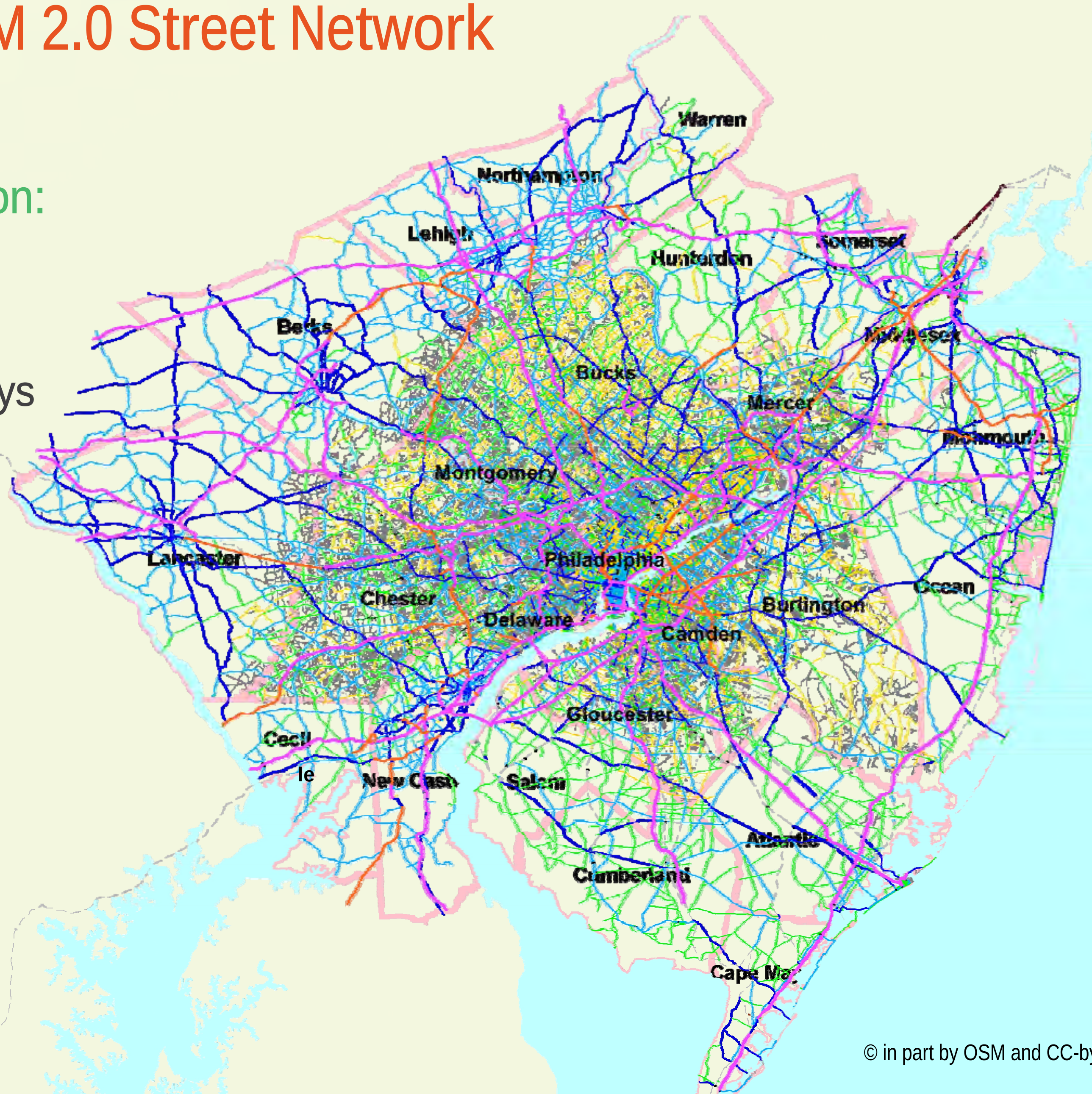
TIM 2.0 Street Network

Within DVRPC region:

- All (!) streets

Extended area:

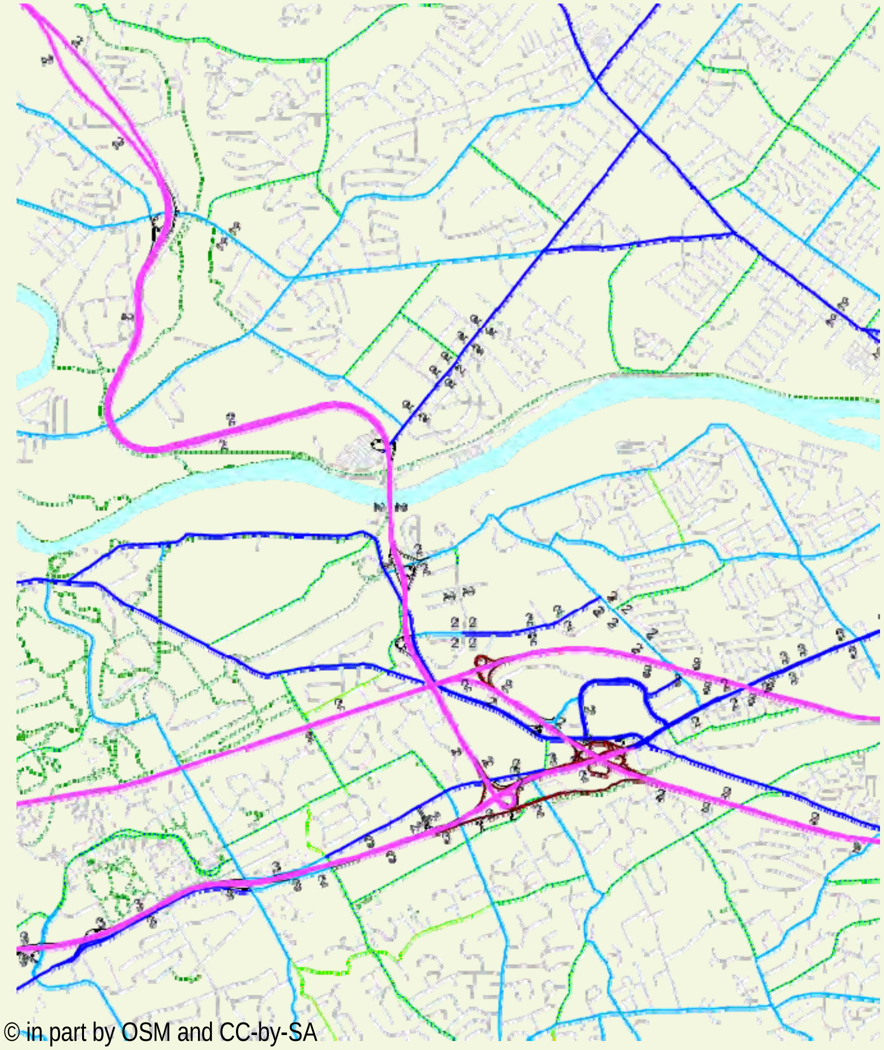
- Major highways



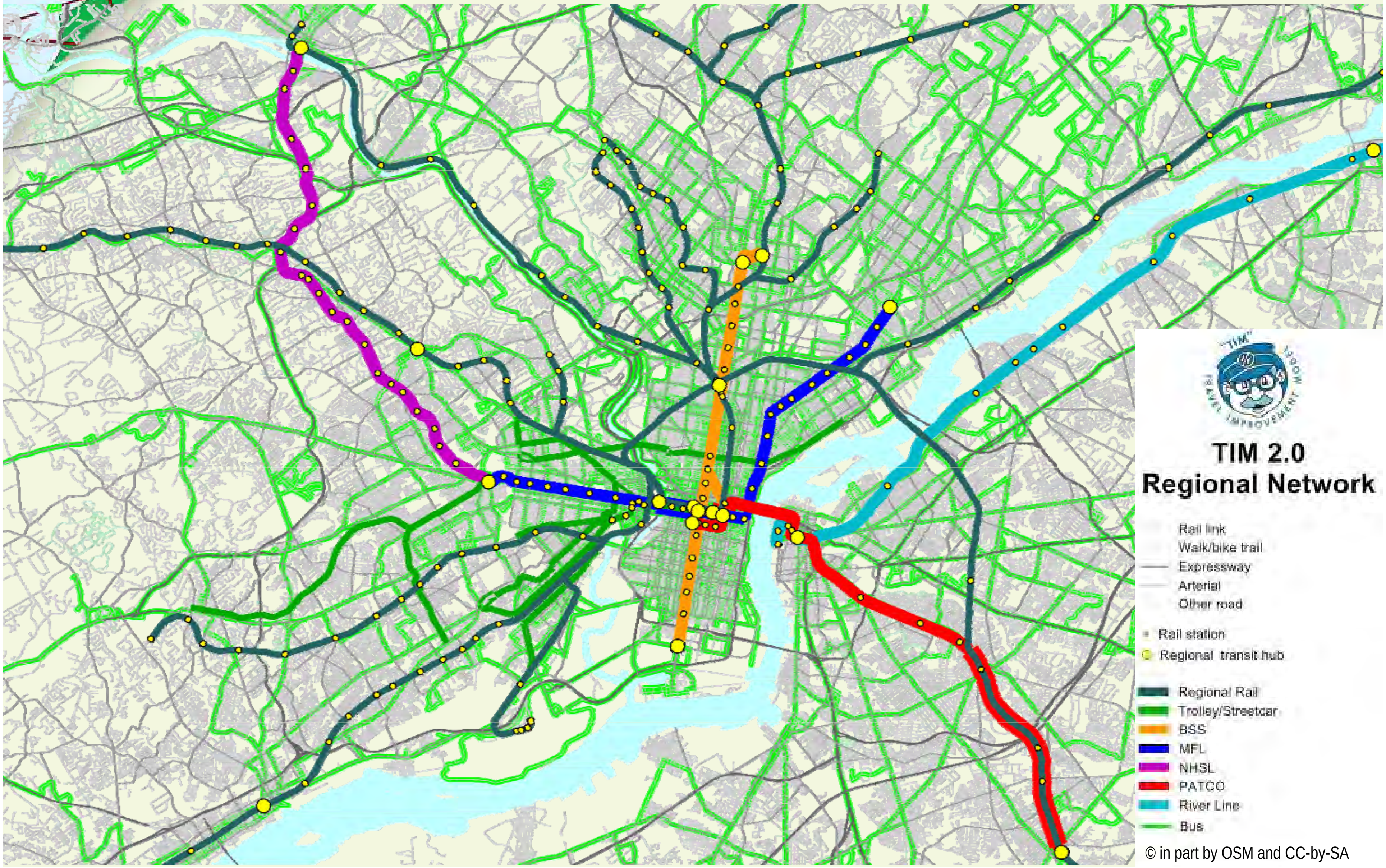
Accurate Street Topology in TIM 2.0

Legacy DVRPC network (TIM 1.0)

TIM 2.0 network model



Integrated Street & Transit Network



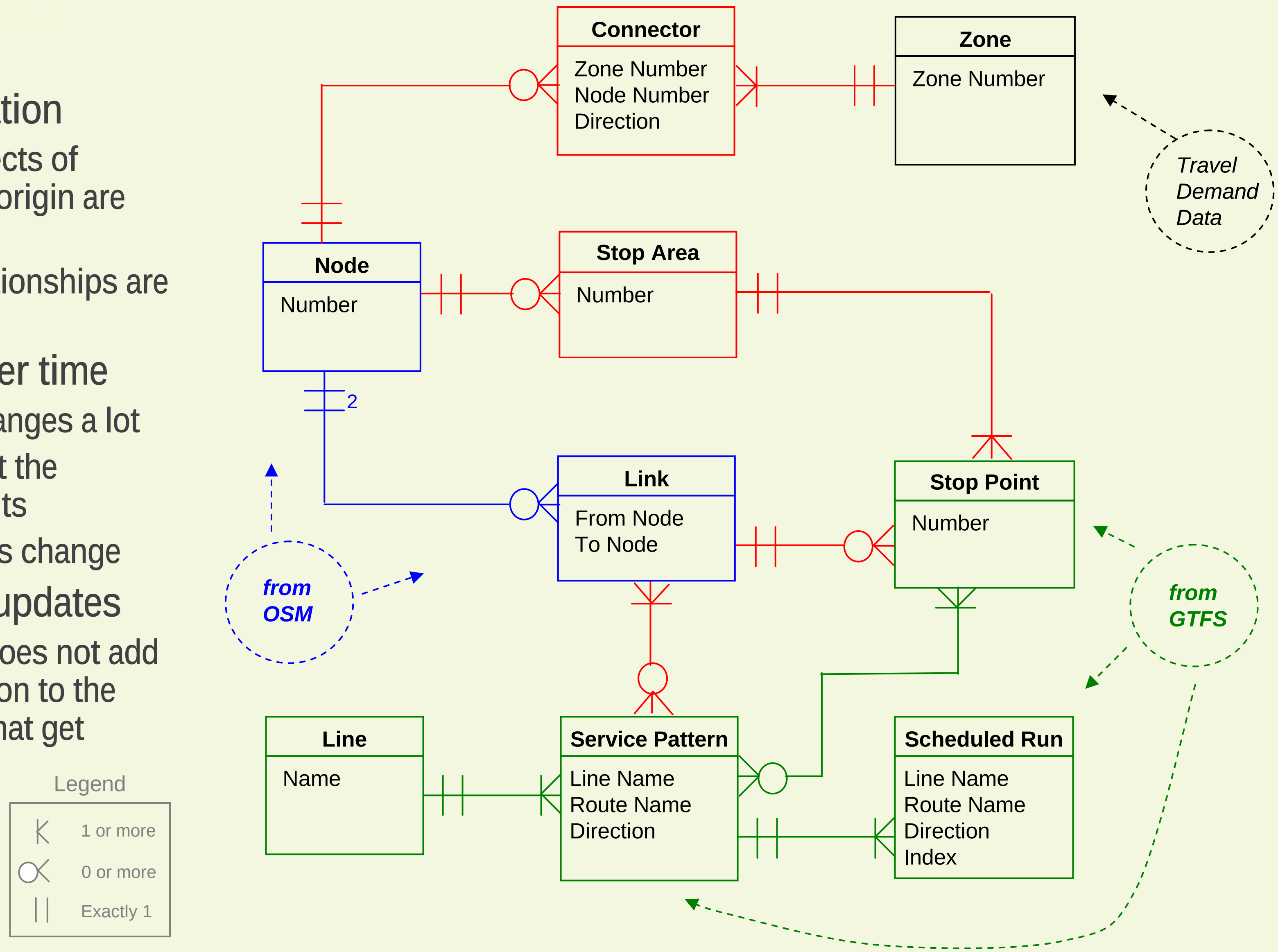
Data Integration and Enhancement

Process

- Data translation into VISUM format
  - OSM
  - GTFS
- Integration of street network with transit data
  - Snapping of GPS stops into the street network
  - Bus route alignment
- Street data enhancement
  - Link attribute coding (class, speed, lanes)
  - Correction of errors in routability (connectivity and directionality)
- Transit network enhancement
  - Stop organization (grouping of stop points)
  - Transfers
  - Access by walk and auto
  - Fare systems
- Calibration of Path Building and Path Flows

Simplified Entity-Relationship Diagram

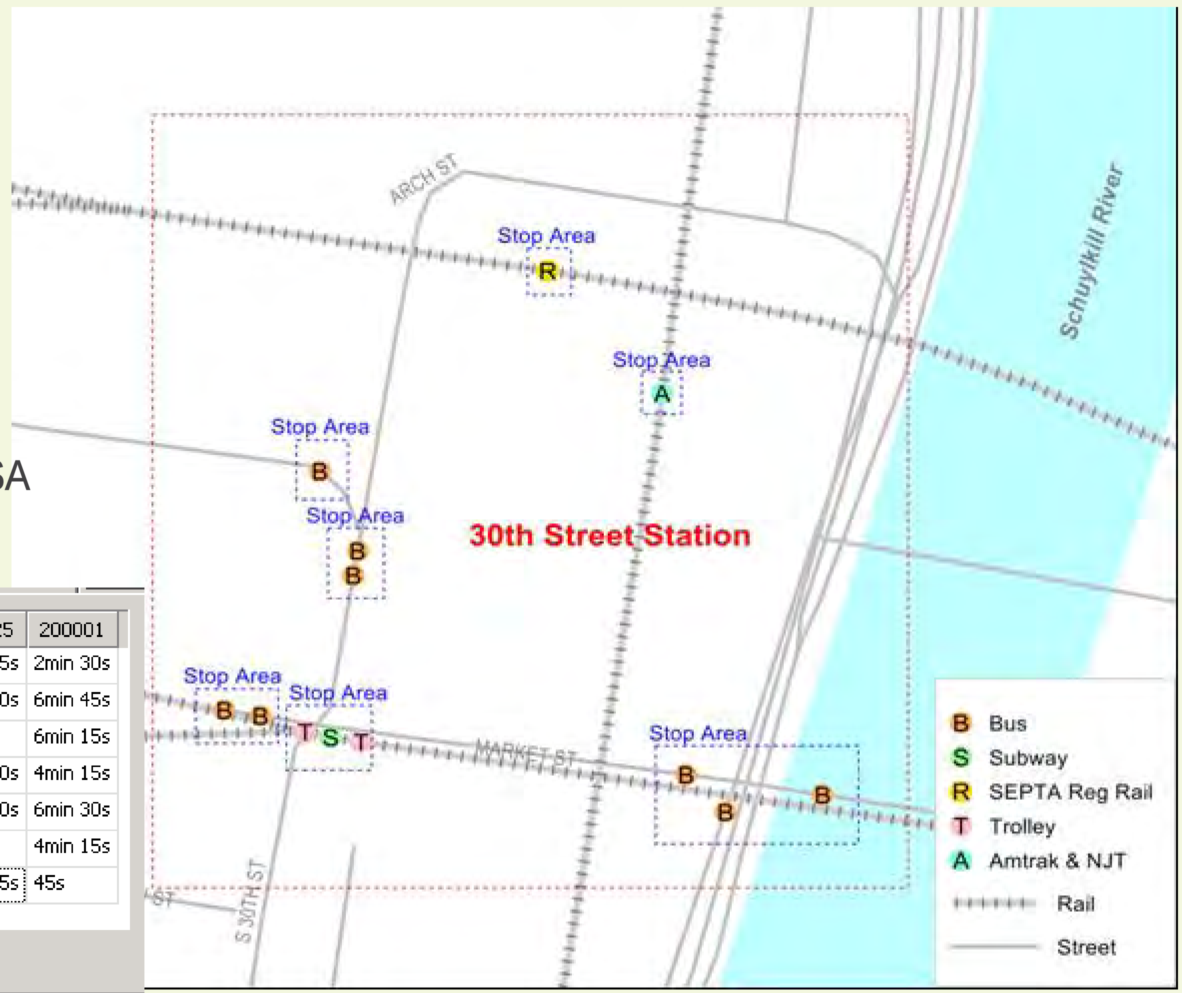
- Data integration
  - Data objects of different origin are merged
  - New relationships are created
- Changes over time
  - GTFS changes a lot
  - … but not the stop points
  - Our zones change
- Automated updates
  - DVRPC does not add information to the objects that get updated



Transit Stop and Transfer Modeling

Three levels of stops

- Stop points (SP)
  - GPS level
  - reported in GTFS
- Stop area (SA)
  - transfer connections
  - access times
- Stop
  - Groups many SP and SA

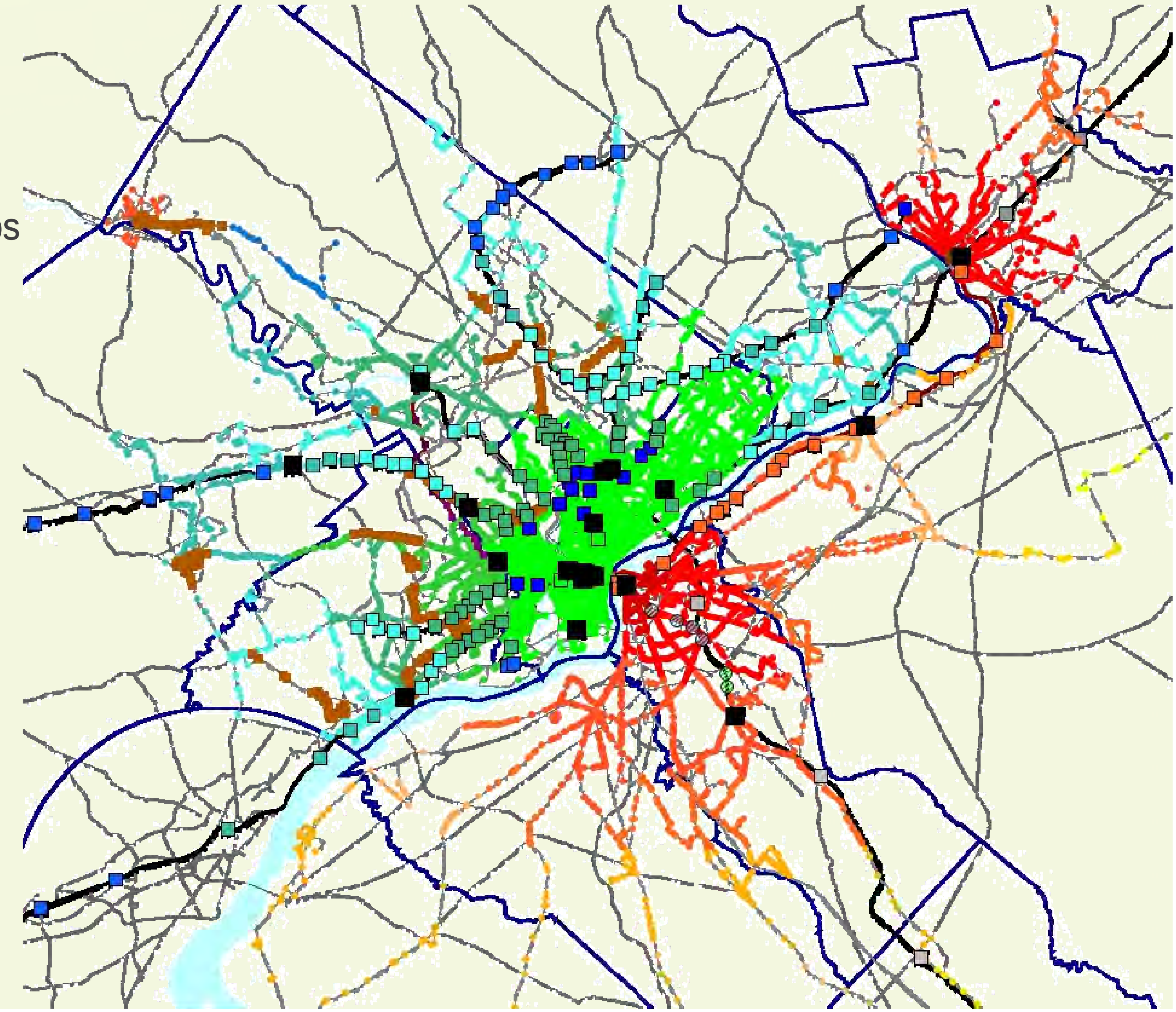


Transfer Times between SA:

	1102	119025	120643	121529	121532	128325	200001
1102	45s	6min 45s	6min 30s	6min 15s	6min 45s	4min 15s	2min 30s
119025	6min 45s	1min	1min 30s	4min 45s	1min 45s	3min 30s	6min 45s
120643	6min 30s	1min 30s	45s	4min	1min	3min	6min 15s
121529	6min 15s	4min 45s	4min	1min	4min 15s	6min 30s	4min 15s
121532	6min 45s	1min 45s	1min	4min 15s	1min 15s	3min 30s	6min 30s
128325	4min 15s	3min 30s	3min	6min 30s	3min 30s	1min	4min 15s
200001	2min 30s	6min 45s	6min 15s	4min 15s	6min 30s	4min 15s	45s

TWk Walk

Coding the Transit Fare System



Fare system:

- Membership of stops in fare zones
- Fare rules
- Prices

DVRCP region:

- Not yet available from SEPTA or NJ Transit in GTFS

The New Network in Numbers

Number of network objects	Legacy network (TIM 1.0)	New network (TIM 2.0)
Street segments	50,000	580,000
Transit stops (stop points)	5,000	18,000
Transit service patterns	2,000	6,000
TAZ (traffic analysis zones)	2,000	3,400

The Effort in Numbers (Feb 2011)

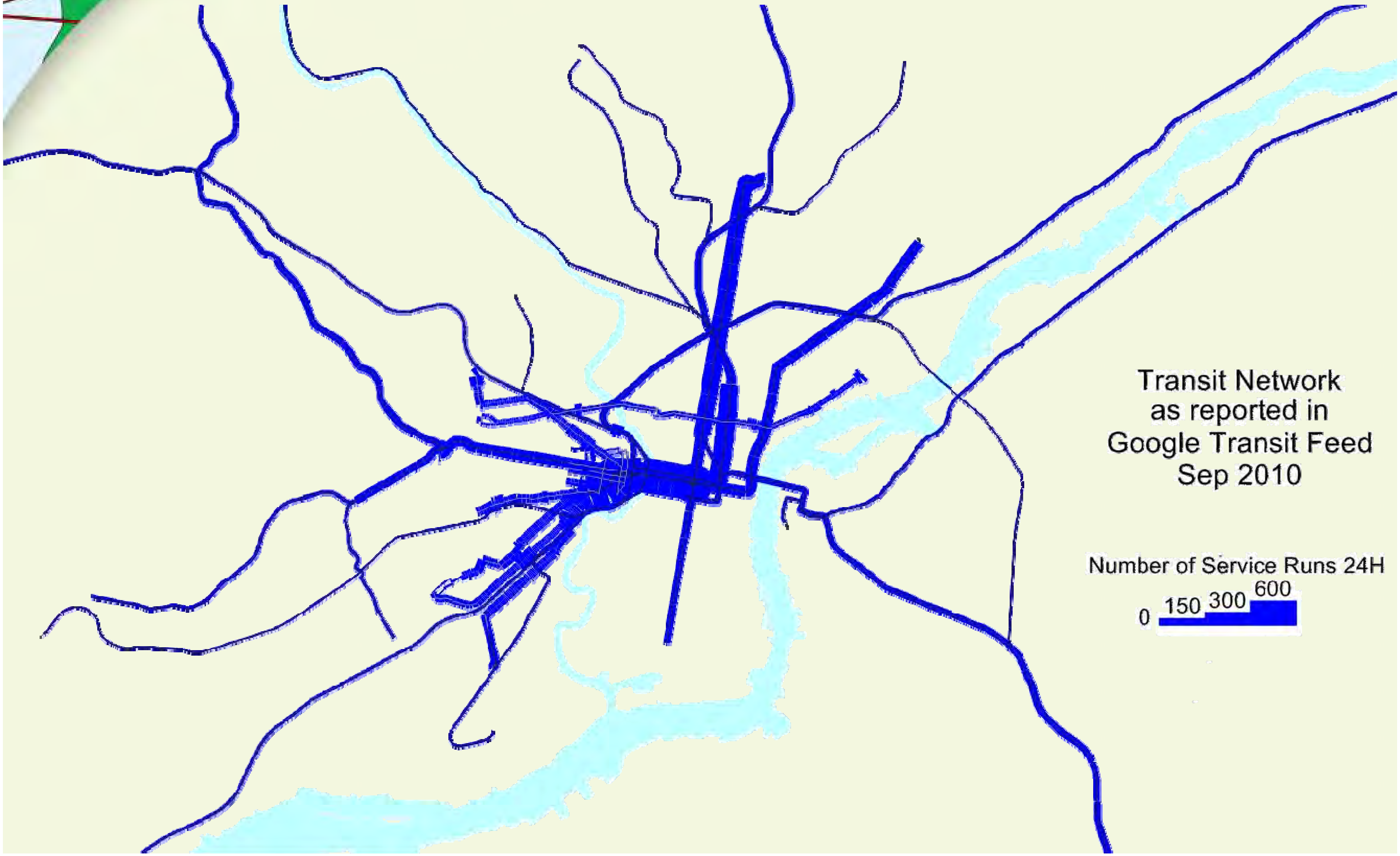
Task	Man-Months	% Full-time	% Intern
Process raw OSM&GTFS data	2.3	80%	20%
Coding of street attributes, parking	6.6	62%	38%
Integration hwy with transit	7.6	48%	52%
Transit transfers & fare system	5.0	82%	18%
Development, integration of TAZ	3.7	13%	88%
Calibration, hwy assignment	6.5	77%	23%
Calibration, transit assignment	5.6	86%	14%
TOTAL	37.3	64%	36%

Applications

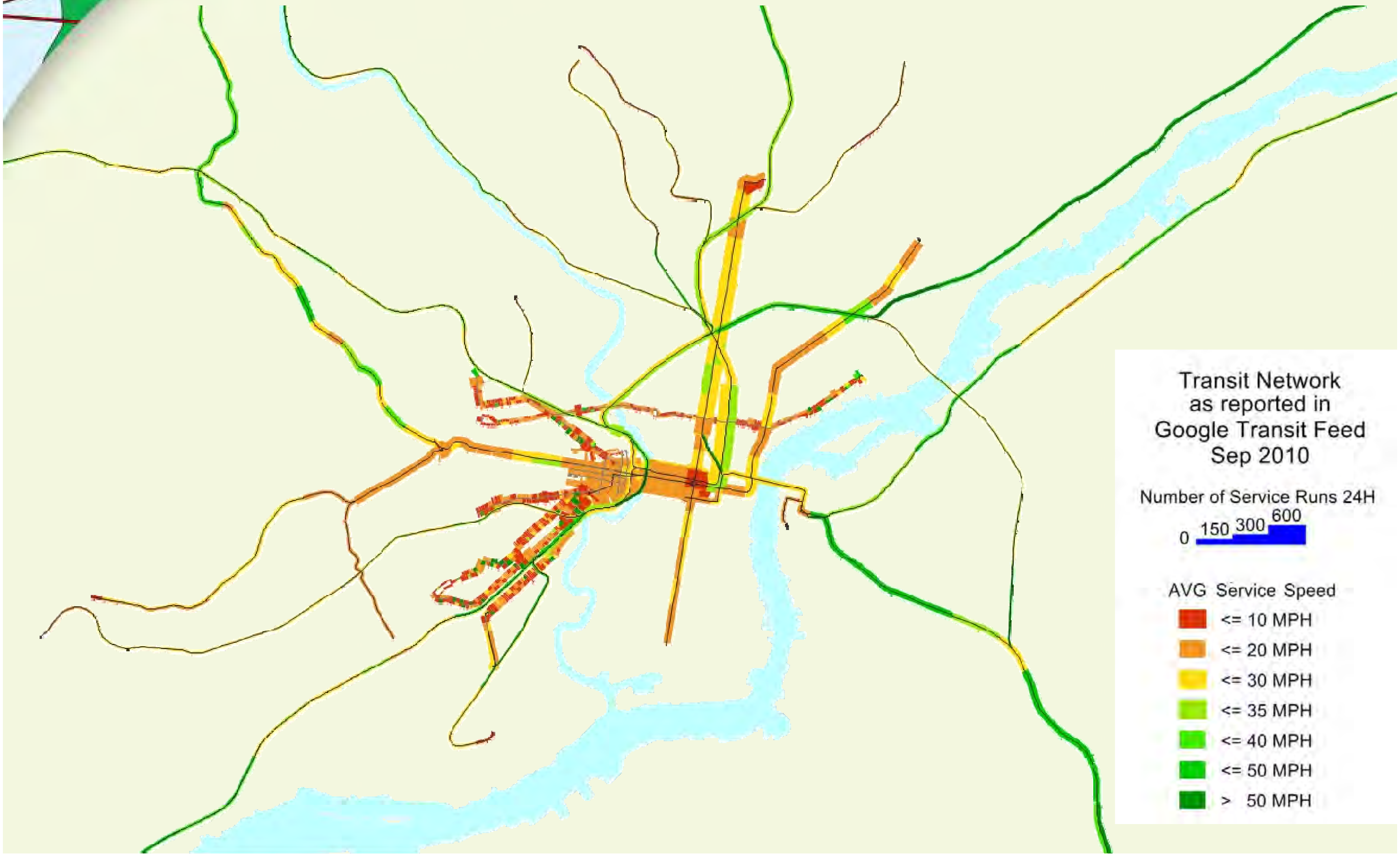
Examples on the following slides

- Transit service analysis (service frequency and speed)
- Network flow simulation (“assignment”)
- Travel time computation (“isochrones”)

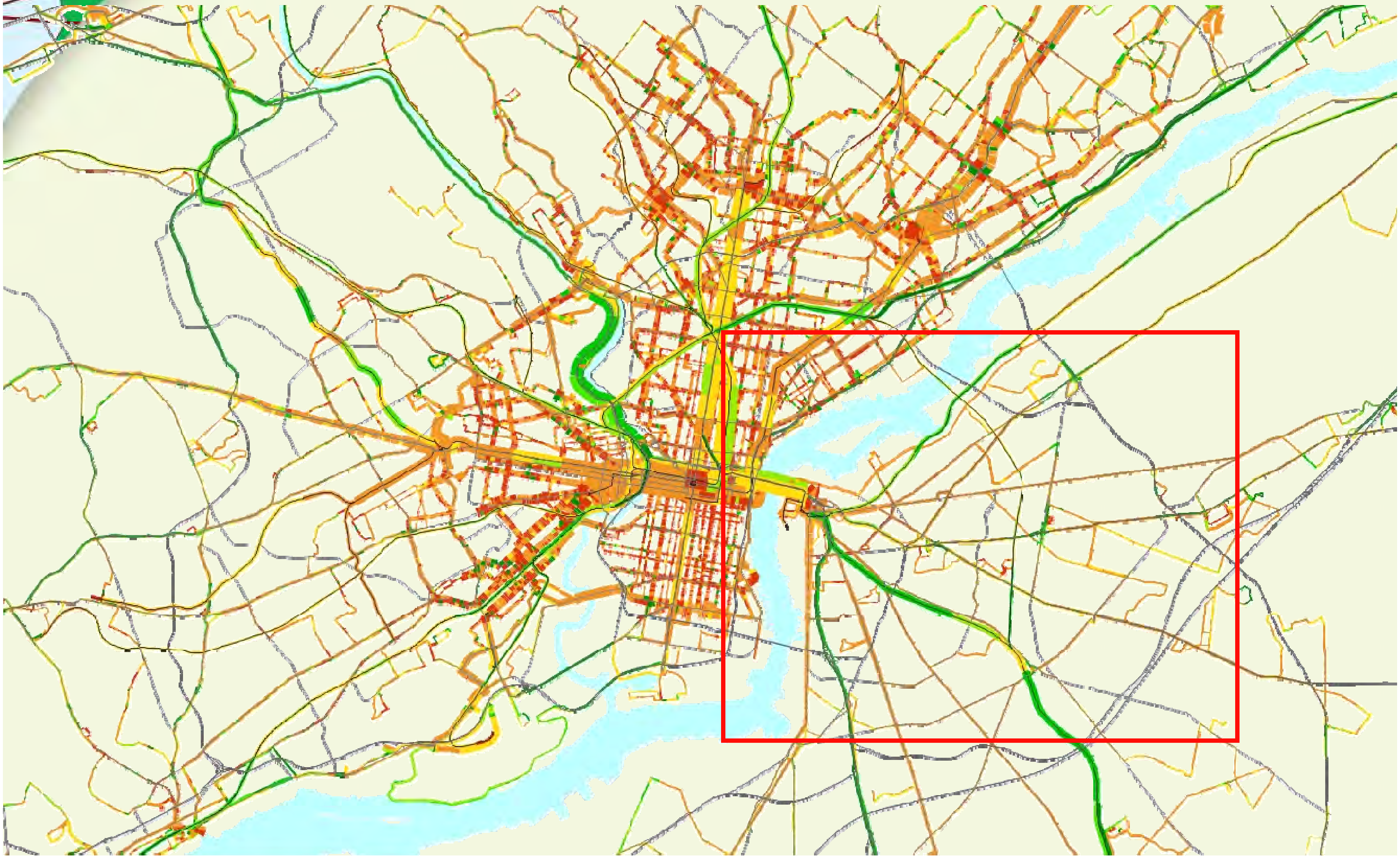
Transit Service Analysis (1)



Transit Service Analysis (2)

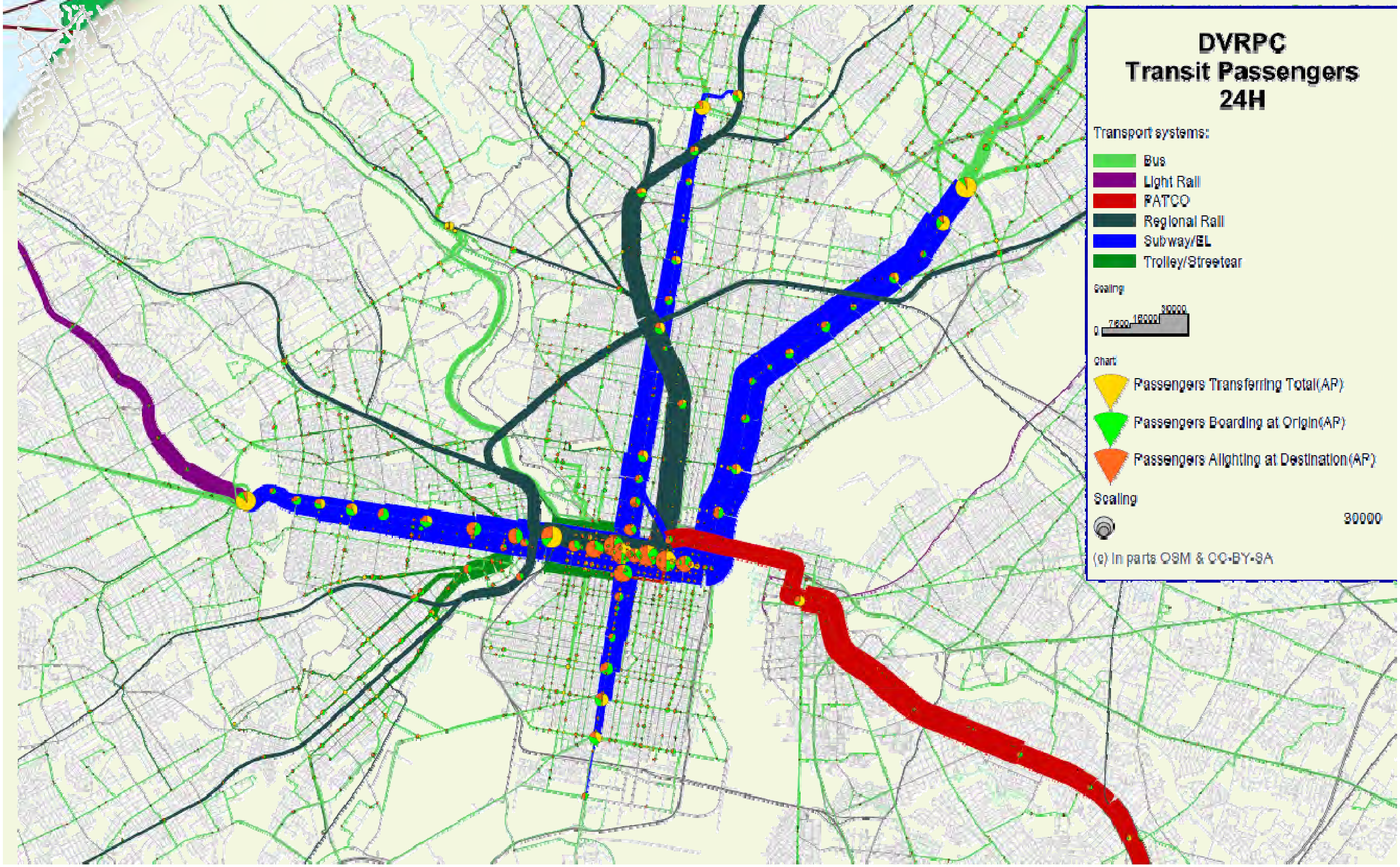


Transit Service Analysis (3)

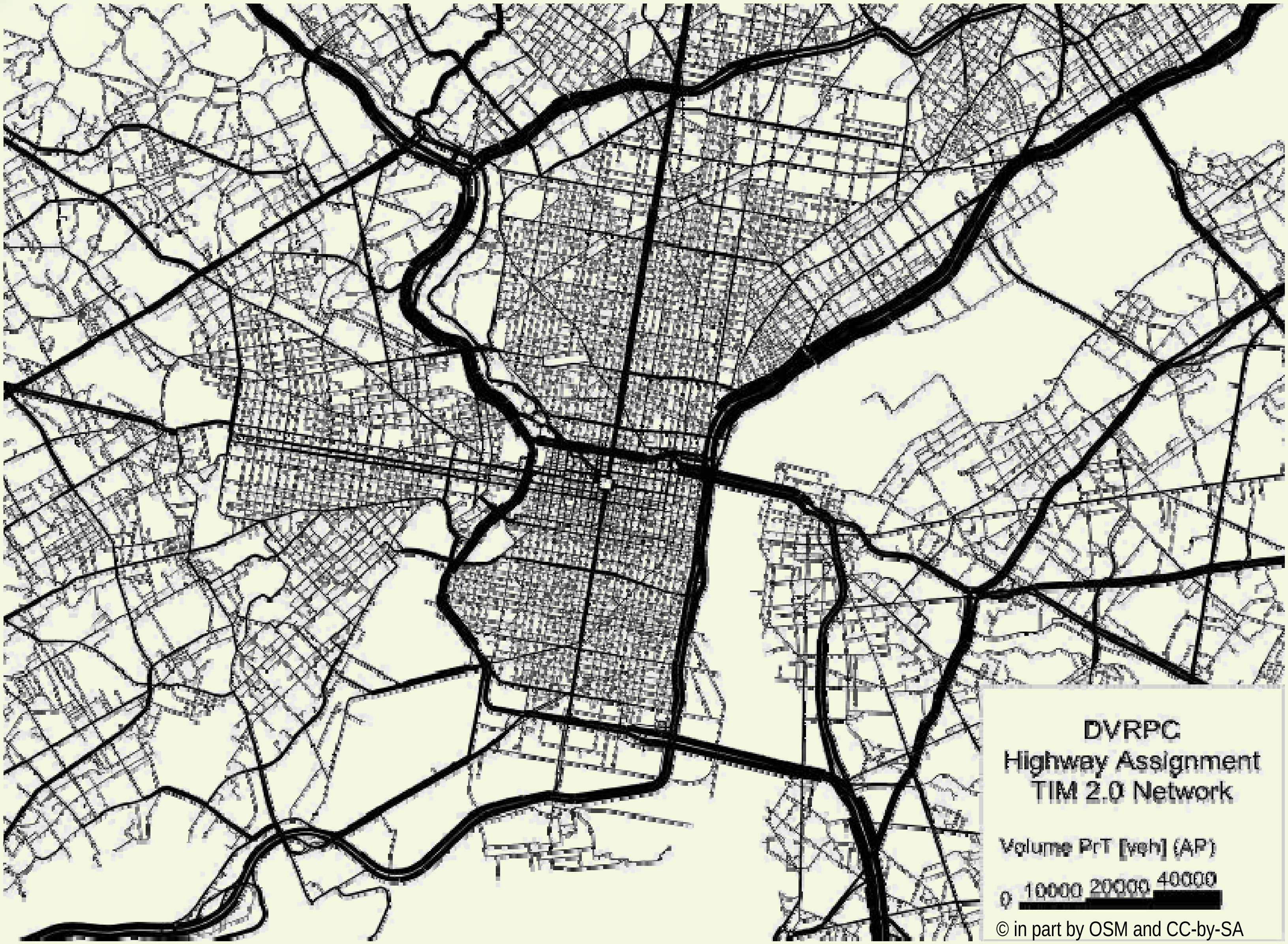


Transit Service Analysis (4)

Transit Ridership Volumes (“Assignment”)

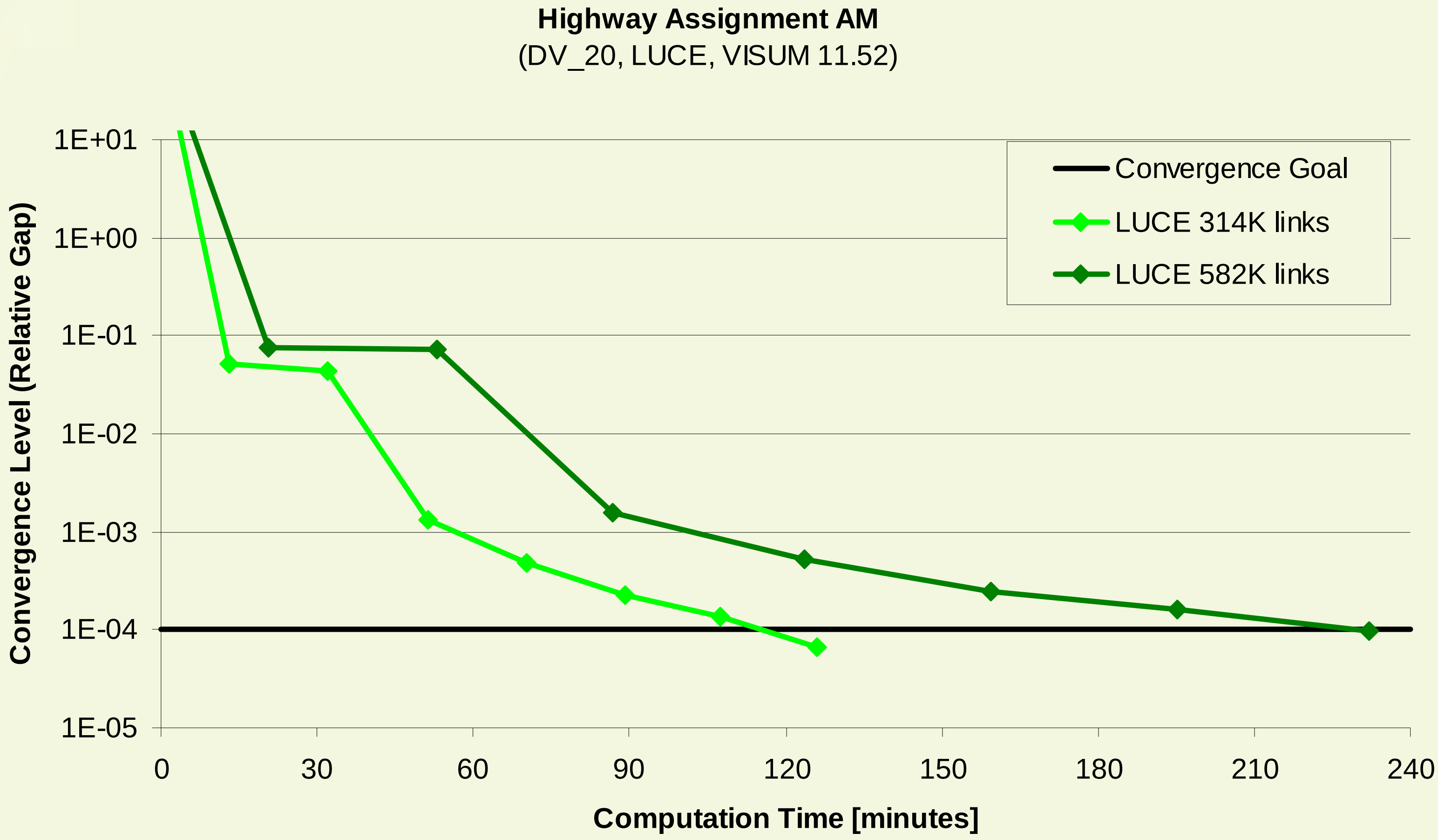


Highway Flow Volumes ("Assignment")

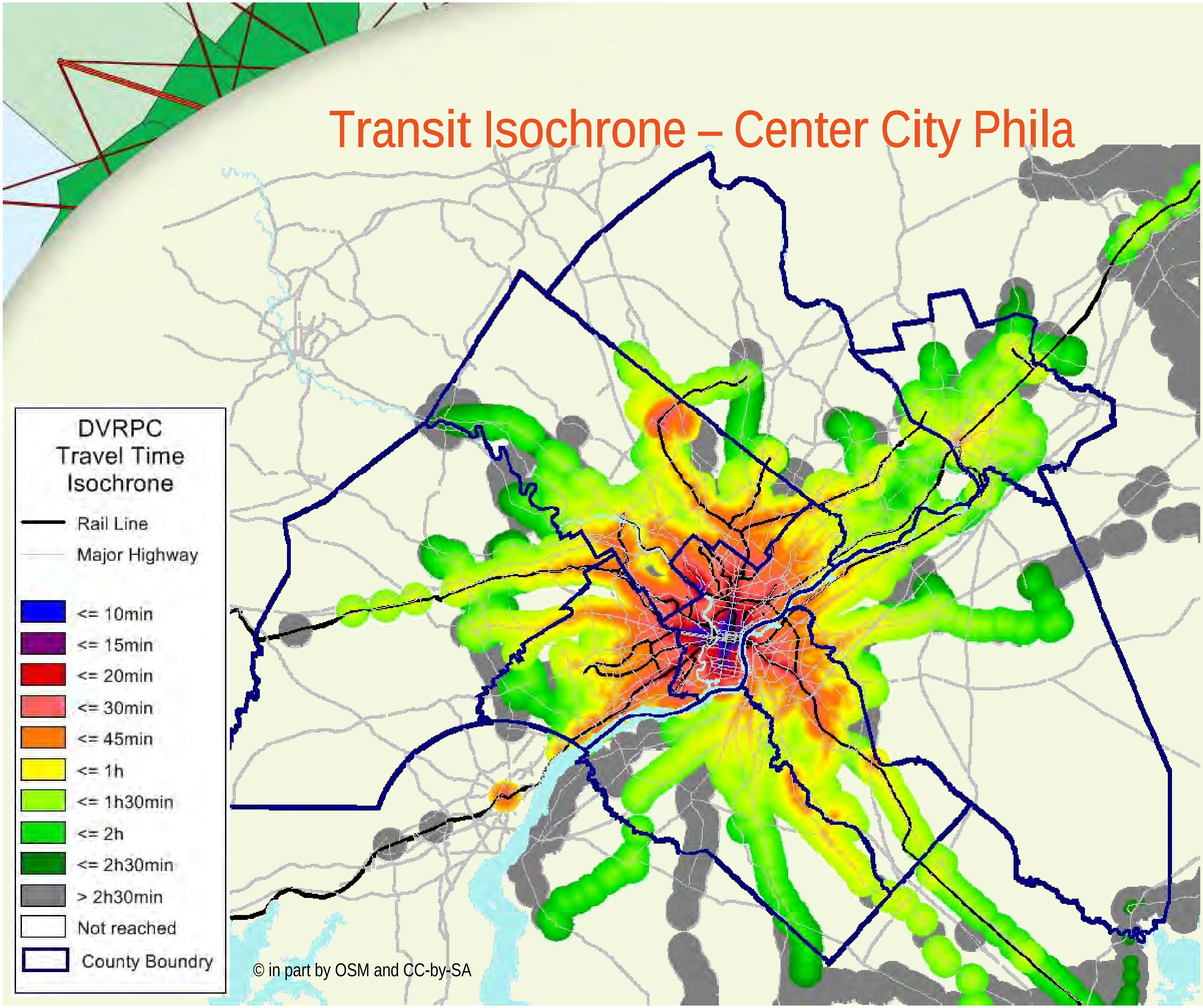


Computation Time

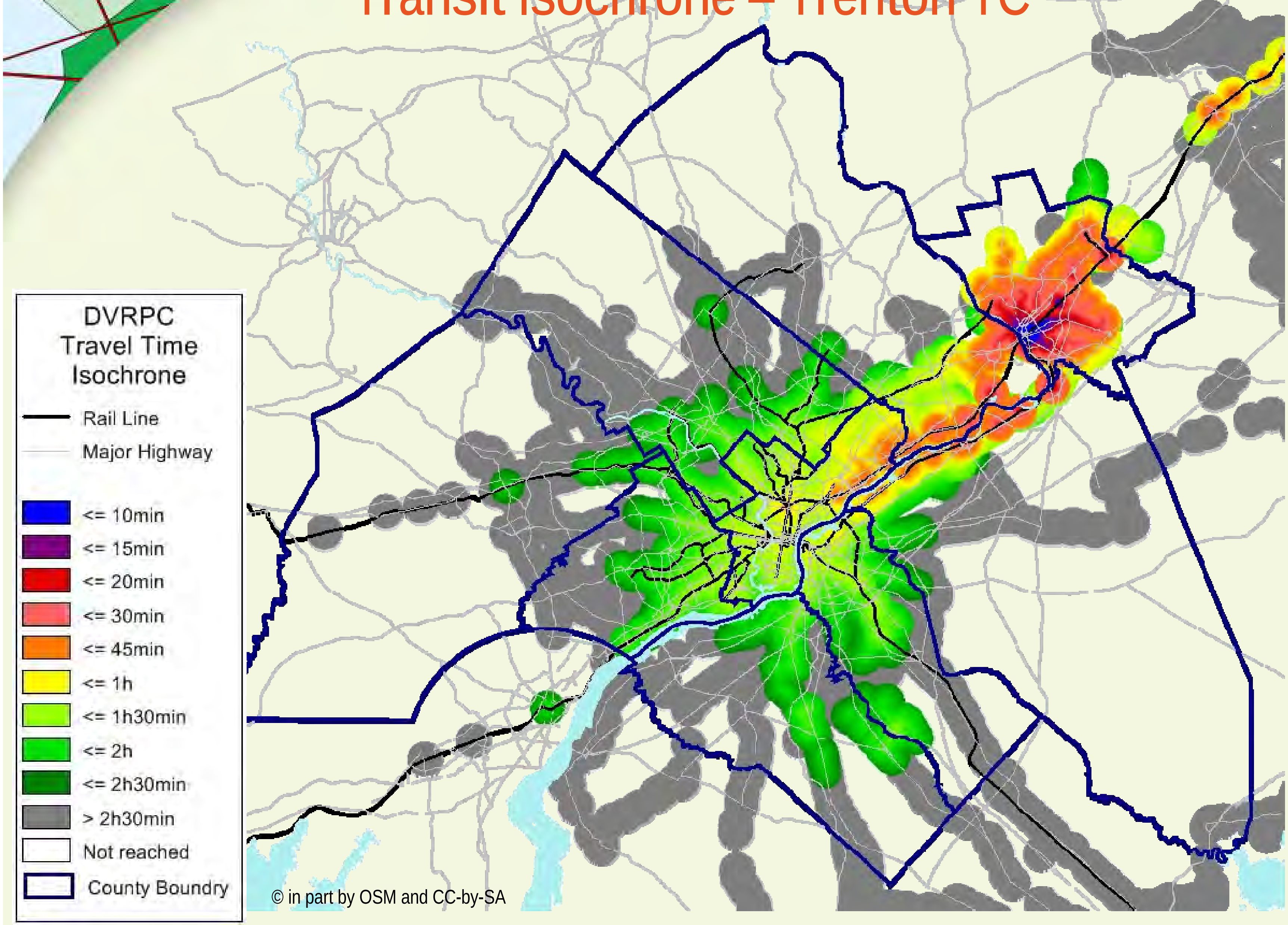
- The large size of the TIM 2.0 network file challenges our computer resources …



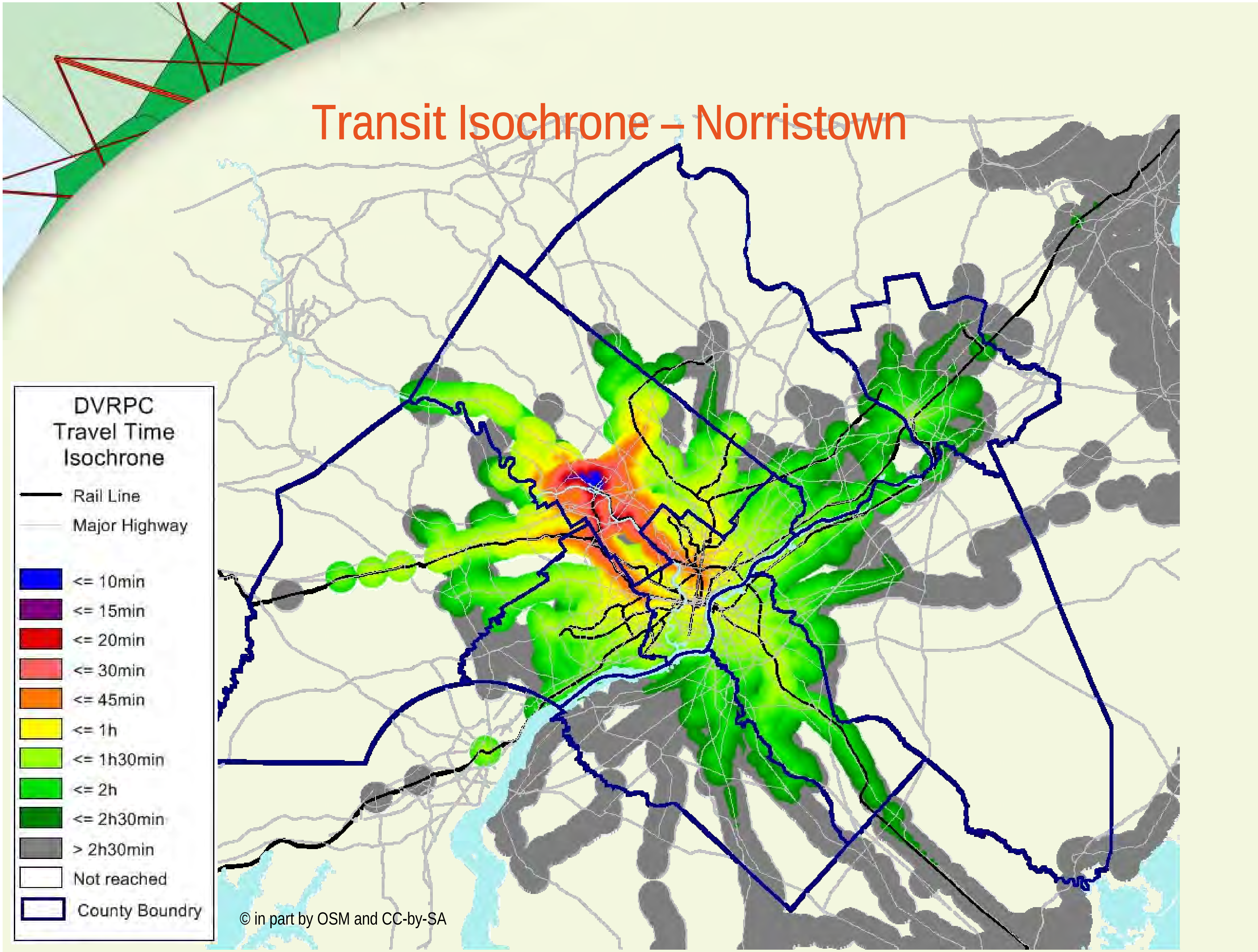
Transit Isochrone – Center City Phila



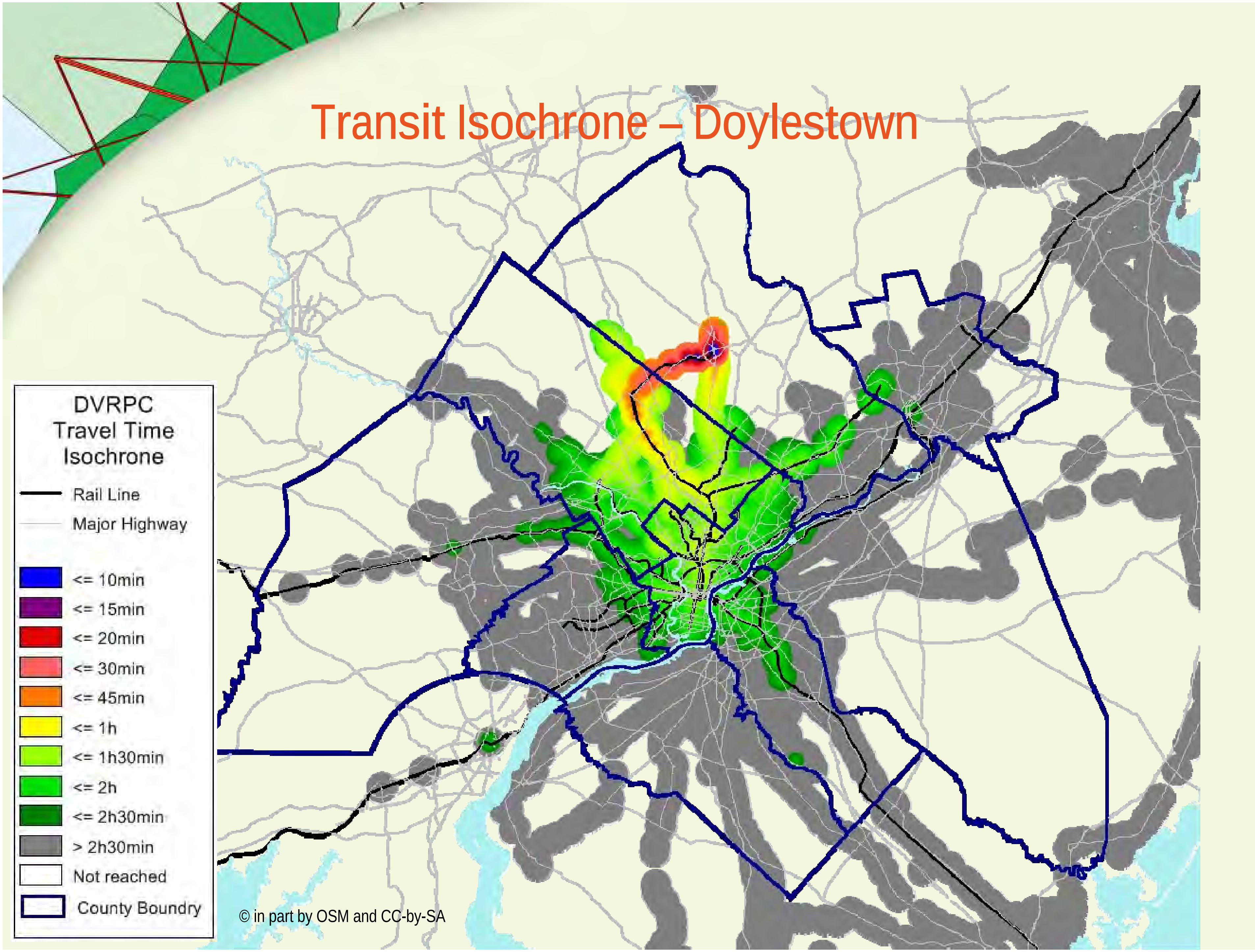
Transit Isochrone – Trenton TC



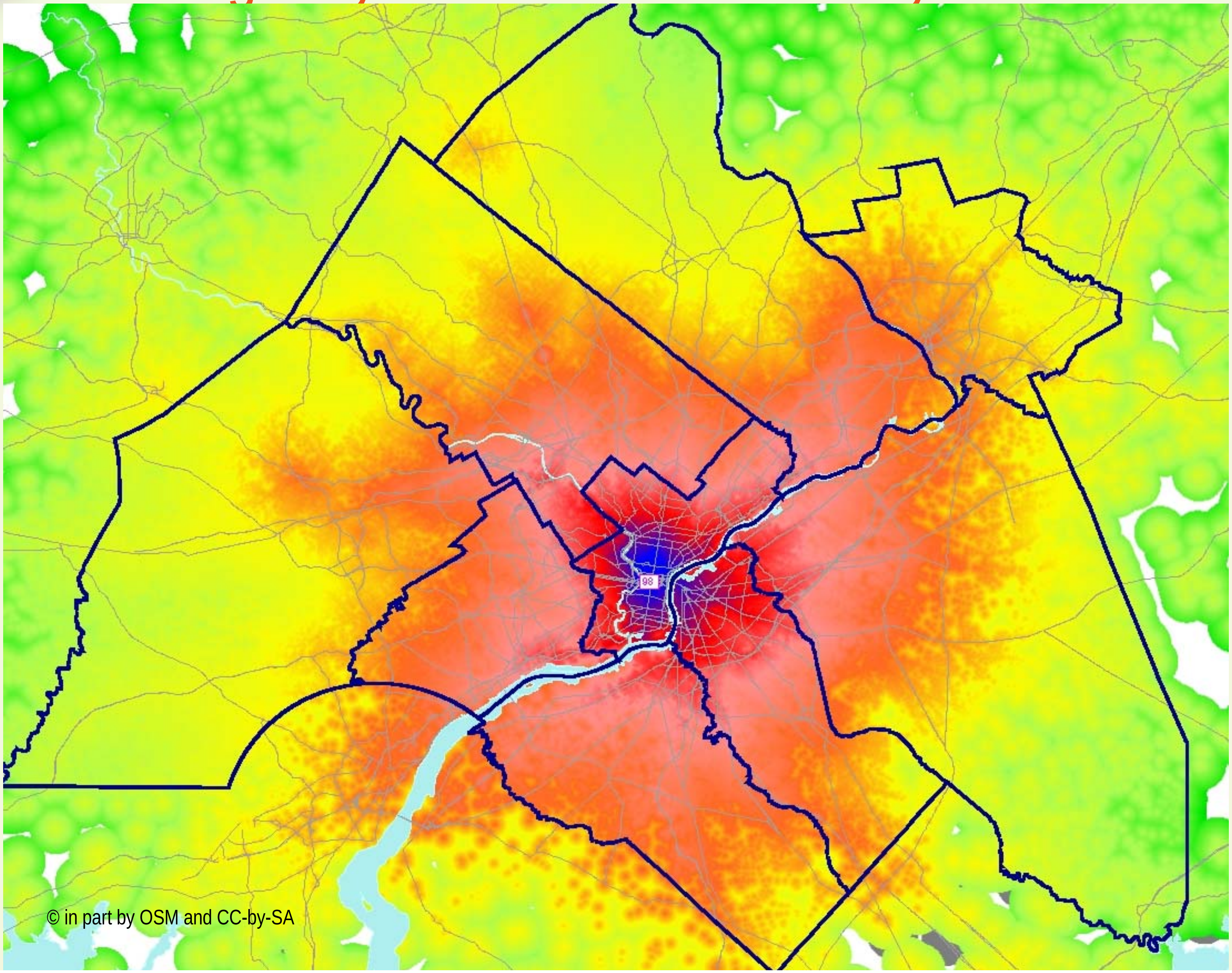
Transit Isochrone – Norristown



Transit Isochrone – Doylestown



Highway Isochrone – Center City



Summary and Conclusions

Use of web 2.0 network data in a forecasting model

- To our knowledge, first MPO in the U.S. to do so
- Benefit from the OSM
  - Provided routable street data without copyright restrictions
- Benefit from Google Transit feeds
  - Accurate representation of transit service
  - Ability for regular updates in a standard format
  - Ability to conduct operational analysis

Open question – future relationships with web 2.0

- OSM was a one-time, one-way data exchange
  - will someone feed our enhancements back?
- Established relationships with people behind the Google feeds (SEPTA, NJ Transit)

Other aspects of TIM 2.0 network

- High level of detail
  - All streets, GPS bus stops
  - Benefits mainly long term (activity-based model, non-motorized travel, operational studies)
  - It seems that we are able to deal with the computational challenges in the short run
- Interaction between the modes
  - Better modeling of P&R, effect of buses on street capacity
- Better representation of times of day
- Extended model area

ACS Data Import Tool

By Elizabeth He
ehe@dvrpc.org

ACS Data

- First ACS 5 year estimates released on Dec 14 2010.
- Data is available down to tracts and block groups level.
- Most of the data can be found from American Factfinder.

ACS Data

- Not all the detail data can be found from American Fact-Finder.
- There are 295 Detail tables only available as csv file on FTP site.
- Need Excel 2007 to extract Block Groups data.
- Individual staff download the data will slowdown network and also resulting in duplicate information

DVRPC ACS Database

- Centralize the data.
- Eliminate the duplicate downloads.

ACS Data on FTP site

- Data source on FTP site
- CSV data file
- Seq Excel file

Data Import

- How long it takes to import entire PA dataset into database manually?
- 5-10 min to create each table in oracle
- 5-10 min to create sqlloader scripts
- 1 min to run the import scripts
- Total 11 – 21 min to import one CSV file.
- 21 hr -41 hr to import the 117 CSV file for one state.

ACS Data Java Import Tool

- Read the seq file one at the time, and then create tables into oracle
- Generate the sqlloader script base the fields in the sql file.
- Run the sqlload commend on DOS to import the data into oracle.
- Generate the log files for each data import.

ACS Data Java Import Tool

- How long it take to import dataset into database using the Tool?
- Demo

ACS Data Java Import Tool

Dataset	State/County/ MCD	Track/ Block_Group
2009 ACS 5 yr est.	Yes	Yes
2009 ACS 3 yr est.	Yes	
2009 ACS 1 yr est.	Yes	

ACS Web Data Mining

- 28 Subject Area.
- 117 Data Files.
- 930 Tables.
- Demo

Questions