HOW TPB STAFF DEVELOPS TRANSIT NETWORKS USED BY THE REGIONAL TRAVEL MODEL

Mark S. Moran
Manager, Model Development, COG/TPB staff

Regional Public Transportation Subcommittee
February 27, 2018

Overview

- TPB travel demand forecasting model
- Transportation networks used by the model
- Transit schedule data used to code networks
  - Machine-processible data: GTFS
  - Non-machine-processible data
- Use of GTFS data to code transit networks
  - Benefits
  - Issues
  - Potential solutions
- Next steps

Image credit: Mark Moran

TPB Travel Demand Forecasting Model: Overview

- Aggregate, trip-based model, a.k.a. four-step model
- Trip generation: Predict the no. of trip ends generated in each zone
- Trip distribution: Predict where trips are going, i.e., connecting trip ends into trips
- Mode choice: Predict the share of trips made by each travel mode
- Trip assignment

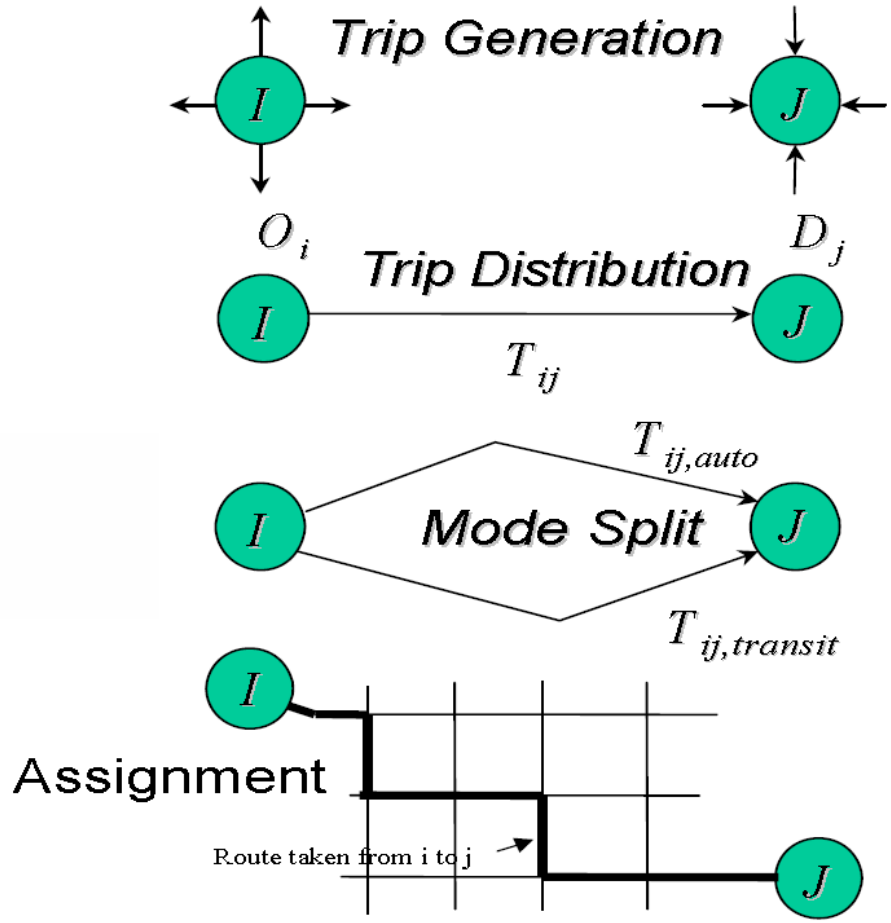
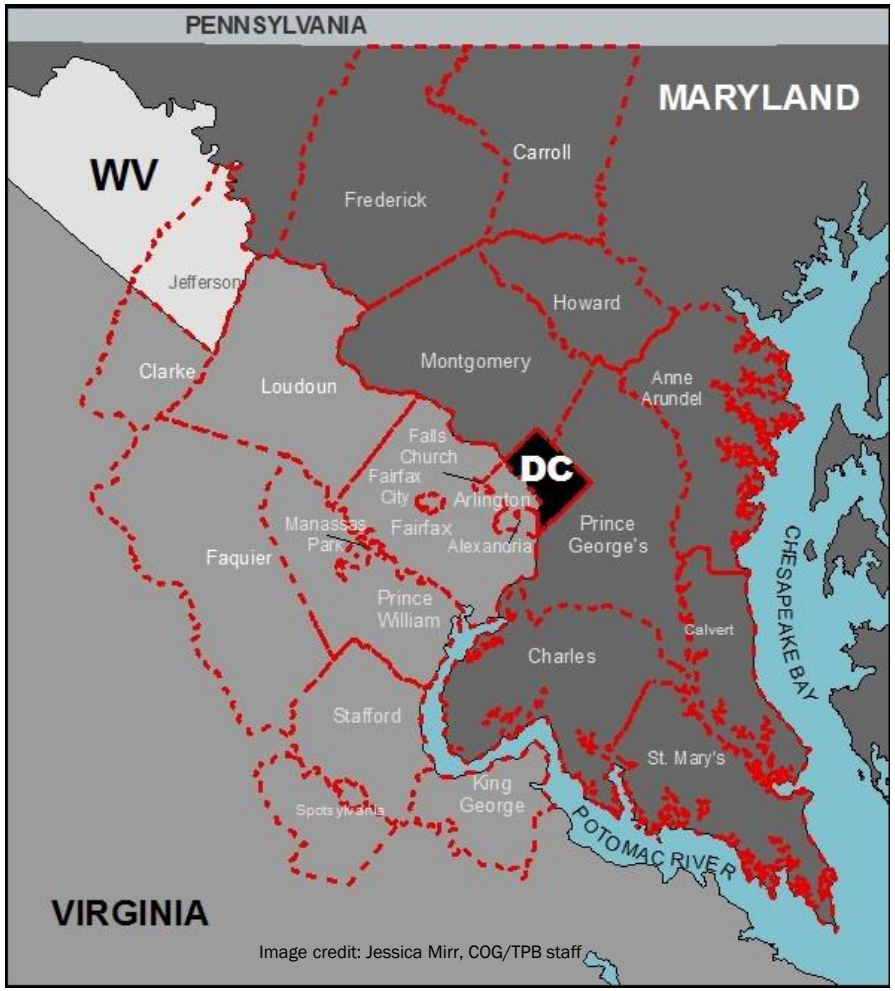


Image credit: Meyer, Michael D., and Eric J. Miller. Urban Transportation Planning: A Decision-Oriented Approach. McGraw-Hill Higher Education, 2001. p. 272.

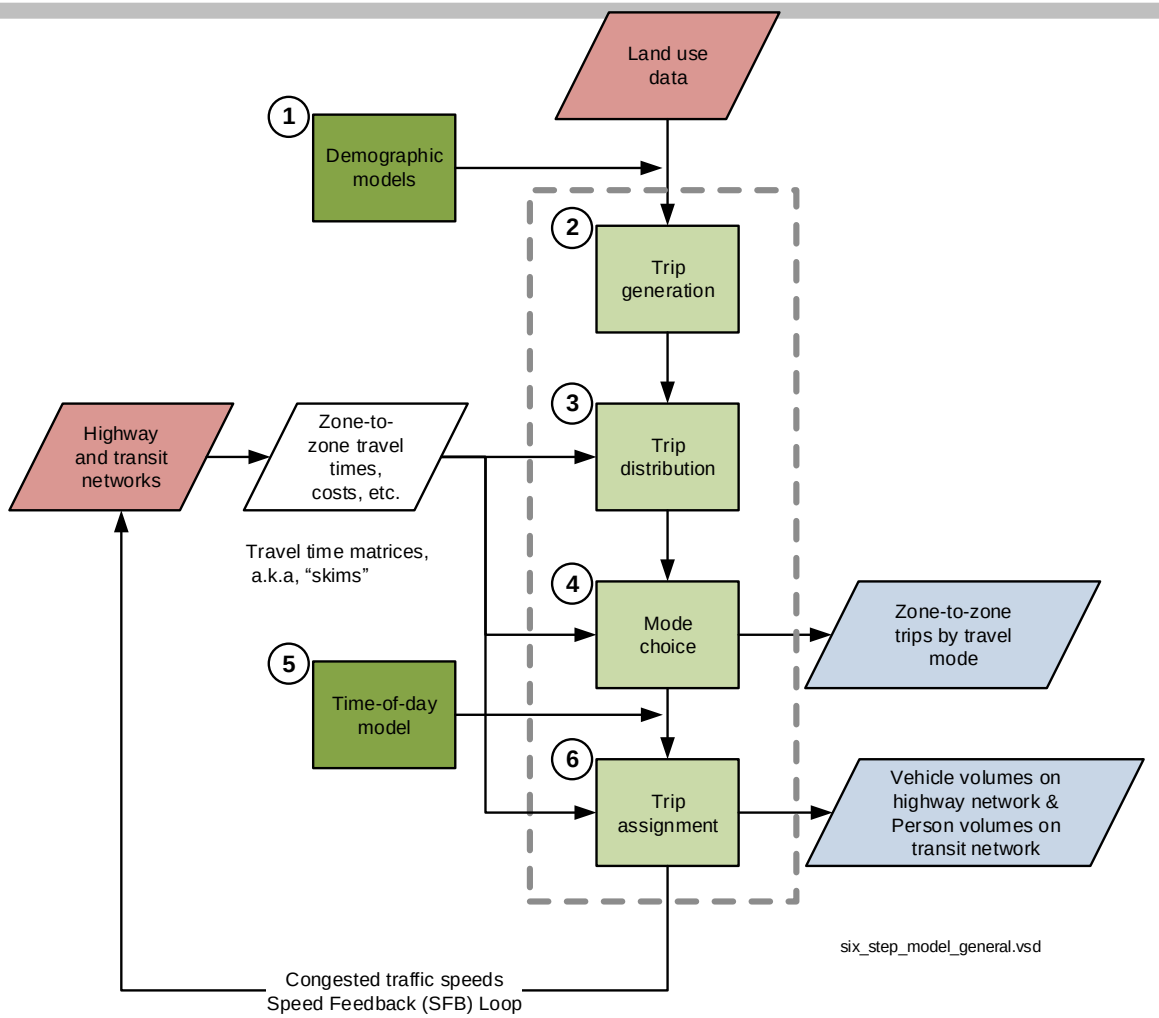
TPB Travel Demand Forecasting Model: Modeled Area & Calibration

- 6,800 sq. mi.
- 22 jurisdictions: DC, MD, VA, & 1 county in WV
- Modeled area is larger than either the MSA (4,000 sq. mi.) or TPB planning area
- 7 million people, 4 million jobs
- Model
  - Calibrated to year 2007
  - Validated to year 2010
  - Documentation on website



TPB Travel Demand Forecasting Model: Inputs and Outputs

- Major inputs
  - Land use data
  - Transportation networks
- Major outputs
  - Zone-to-zone trips by mode
  - Vehicle volumes on highway network
  - Person volumes on transit network



Transportation networks used in the regional travel demand model

- Highway network
  - Aggregate-level (not all streets)
  - Coverage: Freeways (all), Arterials (all), Collectors (some), Local (few)
- Transit network
  - Built on top of the highway network
  - Additional features:
    - Transit-only infrastructure (stations, rail links, PNR lots, access links)
    - Transit service
      - Two times of day: Peak period & off-peak period

Transportation networks used in the regional travel demand model (2)

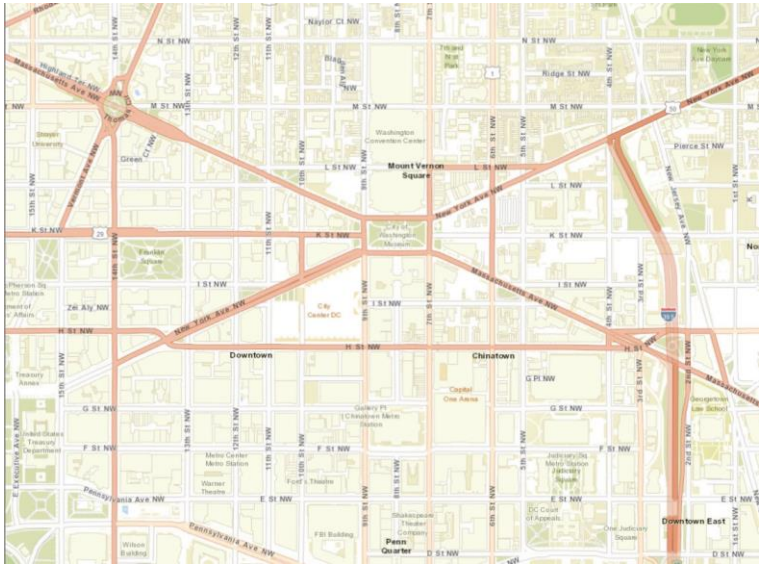
- Source of data for transit networks: Transit agencies (you!)
- Data needed:
  - Transit routes
  - Avg. run time (speed) for each route
  - Avg. headway for each route
- Data formats
  - Machine readable
    - Primarily General Transit Feed Specification (GTFS)
  - Non-machine readable
    - Paper schedules, PDF files, websites

Highway Network

- Links represent road segments
- Nodes represent intersections, merge/diverge points, & zone centroids
- Simplistic depiction of roadway connectivity and capacity
- Used to model vehicle flows between TAZs (not within TAZs)

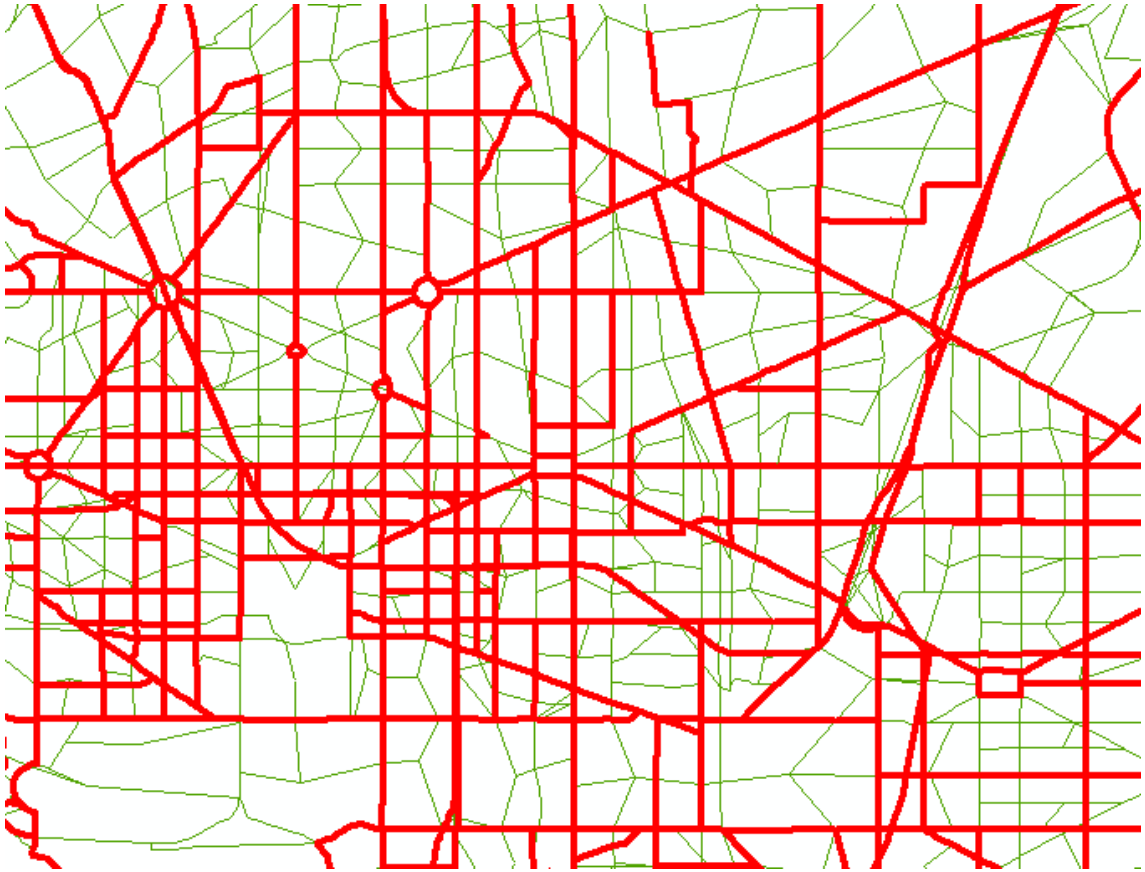
COG/TPB highway network (blue)

All-streets network



Transit Network

- Transit network is built on top of COG's highway network
- Two time periods
  - Peak (AM peak, 7:00 - 7:59 AM)
  - Off Peak (midday, 10:00 AM – 2:59 PM)
- We calculate
  - Avg. headway
  - Avg. run time



Transit Network (2)

- Transit schedule data
  - Machine-readable format: GTFS
  - Paper schedules, PDF files, websites
- 85% of our transit schedule data is in GTFS format (goal: 100%)
- Transit agencies providing GTFS data (on the right):

WMATA - Metrobus/rail

Arlington County – ART Bus

City of Alexandria – DASH Bus

DC Circulator

Fairfax City – CUE Bus

Fairfax County – Fairfax Connector

Falls Church

Frederick County - TransIT

Howard County – Howard Transit

City of Laurel - Connect-a-Ride

Lee Coaches Commuter Bus

MARC Train

Maryland MTA

Montgomery County – Ride-On Bus

Prince George's County – The Bus

Prince William County - OMNI Link/Ride

GTFS

- In 2005, Google and Portland's TriMet transit agency developed an electronic data format for incorporating transit data into online maps
- Initially known as "Google Transit Feed Specification" (GTFS) but was later renamed to "General Transit Feed Specification" (also GTFS)
- Became the default format for sharing public transit scheduling information.

Image credit: Mark Moran

GTFS (2)

- GTFS data is a set of text files that represent a snapshot of a transit agency's scheduled service:
  - agency.txt
  - routes.txt
  - trips.txt
  - stops.txt
  - stop_times.txt
  - calendar.txt
  - …
- Zipped into a file and uploaded to the GTFS Data Exchange website (http://transit.land/ or http://transitfeeds.com/).

Transit Network (3)

- Example GTFS post-processed data table

Agency	TLD	DTPTLD	ORI	DEST	RunTime	Headway	Operation
ART Bus	ART41	ART41N#	Columbia Pike	Court House Metro	29	60	AM
ART Bus	ART41	ART41S	Court House Metro	S Dinwiddie Street	38	15	AM
ART Bus	ART41	ART41S	Court House Metro	S Dinwiddie Street	38	15	OP
ART Bus	ART41	ART41N	S Dinwiddie Street	Court House Metro	26	15	AM
ART Bus	ART41	ART41N	S Dinwiddie Street	Court House Metro	29	15	OP

Transit Network (4)

- Example of non-machine processible data from the website
- Run time and headway will be calculated manually

Weekday Northbound

	Columbia Pike & Dinwiddie	Glebe & Columbia Pike	Ballston Common Mall	Court House Metro	Notes
41	5:30 A	5:37 A	5:45 A	5:54 A	Trip starts at Columbia Pike & Greenbrier 5 minutes earlier.
41	5:50 A	5:57 A	6:05 A	6:14 A	Trip starts at Columbia Pike & Greenbrier 5 minutes earlier.
41	6:10 A	6:17 A	6:25 A	6:34 A	Trip starts at Columbia Pike & Greenbrier 5 minutes earlier.
41	6:30 A	6:37 A	6:45 A	6:54 A	Trip starts at Columbia Pike & Greenbrier 5 minutes earlier.
41	6:40 A	6:49 A	6:58 A	7:09 A	
41	7:00 A	7:07 A	7:15 A	7:24 A	
41	7:15 A	7:22 A	7:30 A	7:39 A	
41	7:30 A	7:37 A	7:45 A	7:54 A	Trip starts at Columbia Pike & Greenbrier 5 minutes earlier.
41	7:40 A	7:49 A	7:58 A	8:09 A	
41	7:55 A	8:04 A	8:13 A	8:24 A	
41	8:10 A	8:19 A	8:28 A	8:39 A	
41	8:25 A	8:34 A	8:43 A	8:54 A	
41	8:40 A	8:49 A	8:58 A	9:09 A	
41	8:55 A	9:04 A	9:13 A	9:24 A	
41	9:10 A	9:19 A	9:28 A	9:39 A	
41	9:25 A	9:34 A	9:43 A	9:54 A	
41	9:40 A	9:49 A	9:58 A	10:09 A	

Use of GTFS data to code transit networks

- GTFS data has greatly increased the efficiency of our annual network updates and minimizes errors
  - Means the model is informed by the very latest information
- Because GTFS is both machine processible and uses a standard format, one can process it using computer programs, which reduces manual processing and staff time
- However, there are still some issues regarding the regional GTFS data

Use of GTFS data to code transit networks (2)

- Many transit agencies have made their own minor changes to the GTFS standard, which complicates computer processing
  - For example the stops.txt data file

    It can be recorded as :

    stop_id, stop_name, stop_lat, stop_lon, location_type

    4000002, S Washington St and Church St, 38.794674, -77.04908, 0

    Or

    stop_id, stop_code, stop_name, stop_lat, stop_lon

    83, 51001, "Ballston Metro G, Fairfax Dr, EB @ N Stafford, NS", 38.882092, -77.110876

Use of GTFS data to code transit networks (3)

- Example, stops.txt

stops.txt

File: **Required**

Field Name	Required	Details
stop_id	**Required**	The **stop_id** field contains an ID that uniquely identifies a stop, station, or station entrance. Multiple routes may use the same stop. The **stop_id** is used by systems as an internal identifier of this record (e.g., primary key in database), and therefore the **stop_id** must be dataset unique.
stop_code	Optional	The **stop_code** field contains short text or a number that uniquely identifies the stop for passengers. Stop codes are often used in phone-based transit information systems or printed on stop signage to make it easier for riders to get a stop schedule or real-time arrival information for a particular stop. The **stop_code** field contains short text or a number that uniquely identifies the stop for passengers. The **stop_code** can be the same as **stop_id** if it is passenger-facing. This field should be left blank for stops without a code presented to passengers.
stop_name	**Required**	The **stop_name** field contains the name of a stop, station, or station entrance. Please use a name that people will understand in the local and tourist vernacular.
stop_desc	Optional	The **stop_desc** field contains a description of a stop. Please provide useful, quality information. Do not simply duplicate the name of the stop.
stop_lat	**Required**	The **stop_lat** field contains the latitude of a stop, station, or station entrance. The field value must be a valid WGS 84 latitude.

Image credit: Google Transit APIs: https://developers.google.com/transit/gtfs/reference/#stopstxt

Use of GTFS data to code transit networks (4)

- Different agencies interpret the “standard” in different ways
  - e.g., In WMATA GTFS data, the arrival_time in stop_times.txt file could be 25:02:03 (other agencies: 1:02:03)
- Same agencies change the format they use in GTFS data from year to year
  - COG uses Transit Route Name and OD Name to match GTFS data records to the transit network
  - Inconsistent GTFS data requires us to manually match GTFS to the COG transit network, negating some of the benefits of using machine-processible data

Potential Solutions

- Encourage all transit agencies to publish schedule data in GTFS format
  - Google provides free software for agencies to publish schedule in GTFS format
- Encourage all transit agencies in the region to agree on one standard format
- Specify data type (e.g., numeric, character) and domain for each attribute of the GTFS files

Next steps

- Further streamline processing of GTFS data
- Leverage geometric information within GTFS for transit network validation and editing

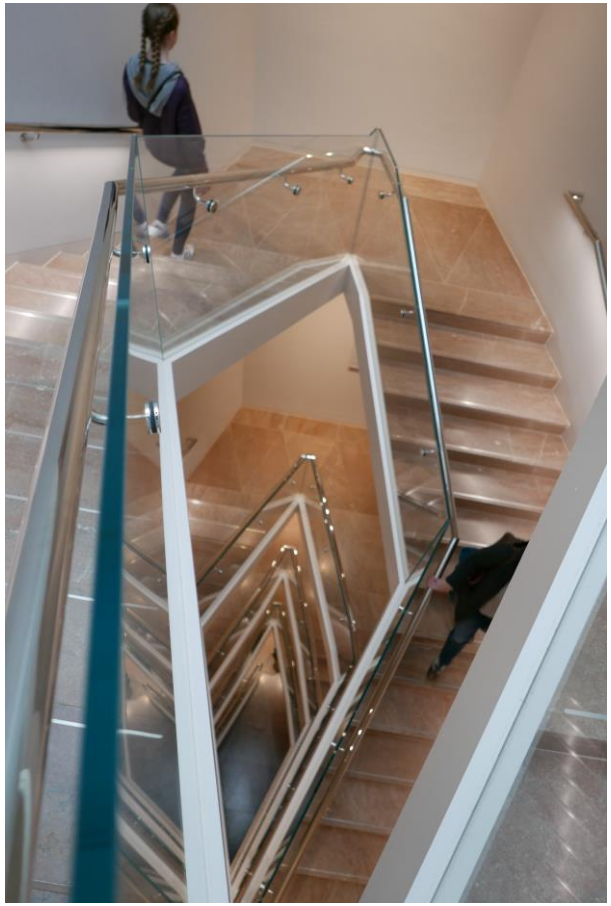


Image credit: Mark Moran

Acknowledgements

- Network development group (Jim Yin, Wanda Owens, Joe Davis) and Meseret Seifu

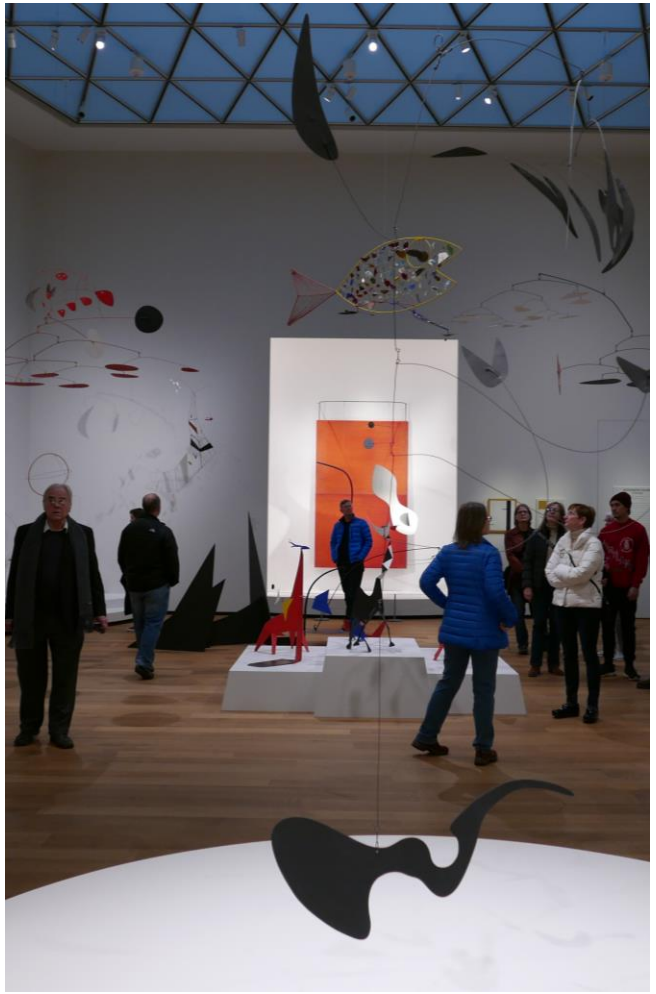


Image credit: Mark Moran

Mark S. Moran

Manger, Model Development
(202) 962-3392
mmoran@mwcog.org

Jim Yin

Principal Transportation Engineer
(202) 962-3361
jyin@mwcog.org

mwcog.org/TPB

Metropolitan Washington Council of Governments
777 North Capitol Street NE, Suite 300
Washington, DC 20002