Pima County Regional Dynamic Traffic Assignment Model Mid-Term Report

Prepared by:

Yi-Chang Chiu, Melrose Pan, Raymond Huang

DynusT Lab

The Depart of Civil and Architectural Engineering and Mechanics

The University of Arizona

02/01/2020

Table of Contents

1 Executive Summary 6
2 DynusT Model Data Collection and Initial Processing 7
1.1 Data Collection 8
1.2 DynusT Network Buildout Process 10
1.2.1 Network Clean-up 11
1.2.2 Importing and Converting the DynusT Network 15
1.2.3 DynusT Network Buildout 17
1.3 DynusT's Time-Dependent Demand Matrices 21
1.3.1 Descriptions of the StreetLight Data 21
1.3.2 StreetLight Data Rationality Check 24
1.3.3 StreetLight OD Data Analysis 27
1.3.4 Converting Streetlight Data to the DynusT Format 30
2 DynusT Model Validation and Calibration 31
2.1 DynusT Network Connection Testing 32
2.2 General Simulation Outputs and Statistics 33
2.3 OD Calibration and Validation 35
3 Concluding Remarks 39

List of Figures

Figure 2-1: Data requirements of DynusT. 7

Figure 2-2: DynusT Network Buildout Process. 8

Figure 2-3: An Illustration of the Link Shapefile. 9

Figure 2-4: Creating the Node Shapefile. 9

Figure 2-5: An Illustration of the Zone Shapefile. 10

Figure 2-6: An Illustration of Trimming the Network. 12

Figure 2-7: An Illustration of Short Links. 12

Figure 2-8: Relationship Between Zones and Generation Links/Destination Nodes ... 13

Figure 2-9: An Illustration of Adding Generation Links and Destination Nodes 14

Figure 2-10: An Illustration of Importing a Shape File. 15

Figure 2-11: An Illustration of Converting an ARC File to an Editable Network. 16

Figure 2-12: An Illustration of the Convert Arclink to Network Tool. 16

Figure 2-13: An Illustration of a Complete Network Conversion. 17

Figure 2-14: An Illustration of the Generation Links and Destination Nodes. 18

Figure 2-15: An Illustration of Generation Links in the Network. 18

Figure 2-16: An Illustration of Destination Nodes in the Network. 18

Figure 2-17: An Illustration of the Covert Centroids Tool. 19

Figure 2-18: Movement Direction Calculation Tool. 19

Figure 2-19: DynusT's Anisotropic Greenshield Family Speed-Density Curve Formulation. 20

Figure 2-20: An Illustration of Defining the Traffic Model 21

Figure 2-21: Traffic Analysis Zones. 22

Figure 2-22: StreetLight Calibration Links. 22

Figure 2-23: Important Links Missing Rate Distribution. 24

Figure 2-24: Visualization of the OD pairs with a missing rate greater than 0.6 on the map. 25

Figure 2-25: Locations of the Calibration Links (Orange). 26

Figure 2-26: StreetLight OD Underestimated Traffic Volume. 26

Figure 2-27: Heatmap Indicating the Number of Vehicles at 8 AM and 5 PM During the School Season. Missing Zones are Colored Gray. 27

Figure 2-28: Heatmap Indicating the Number of Vehicles at 8 AM and 5 PM During the Summer Break. Missing Zones are Colored Gray. 28

Figure 2-29: Heatmap Indicating the Number of Vehicles at 8 AM and 5 PM During the Snowbird Season. Missing Zones are Colored Gray. 28

Figure 2-30: Peak-hour (7 AM) and Off-peak Volume (1 AM) Comparison. 29

Figure 2-31: Departure Time Distributions. 30

Figure 2-32: An Illustration of the OD Demand Import Tool. 31

Figure 2-1: DynusT Model Validation and Calibration Process. 32

Figure 2-2: An Illustration of the DTA Simulator. 33

Figure 2-3: Departure, Arrival, Trip Duration (Travel Time), and Delay Profiles for the Baseline Run. 34

Figure 2-4: An Illustration of the Analysis Tool. 34

Figure 2-5: The Calibration Process. 35

Figure 2-6: Simulation Counts Compared with Screen Line Counts After Iterations 0, 6, and 12. 36

Figure 2-7: Differences between Simulation Counts and Screen Line Counts After Each Iteration 37

List of Tables

Table 2-1: Initial Link Type Statistics 11

Table 2-2: Network Statistics. 15

Table 2-3: Sample of Raw OD Data. 23

Table 2-4: Number of OD Pairs. 23

Table 2-5: Processed Data Example. 30

Table 2-6: DynusT Demand Format. 31

Table 2-1: DynusT Simulation Output Statistics. 33

Table 2-2: Summary of Counts and Errors on 30 Calibration Links. 38

1 EXECUTIVE SUMMARY

The University of Arizona DynusT research team has completed the Pima County regional Dynamic Traffic Assignment (DTA) model creation, validation, and calibration tasks. The details of these tasks are documented in this mid-term report.

The entire effort is generally composed of several sub-tasks: network built-out process, model calibration, and validation.

For the network creation, network and demand data were collected and converted to DynusT input file format. The link, node, and TAZ shapefiles were first preprocessed in QGIS. Then, the preprocessed shapefiles were converted to the format required by DynusT. Origin-destination (OD) data were collected from StreetLight, including school, summer and snowbirds scenarios. After going through some cleaning process, the temporal and spatial analysis was conducted, and some major issues of using StreetLight data such as underestimation were discovered. The OD data were converted to the format required by DynusT and used as the initial demand input.

Once the base model was created, the OD demand data were calibrated using the DynusT calibration tool and the AADT counts provided by the County. The purpose of OD calibration is to adjust the OD matrices such as the re-simulated and re-assigned link volumes at the provided permanent stations matches with the actual counts as possible. From the process, we discovered a significant underestimation of StreetLight data compared to the final results. The base simulation results were recorded, and volume, departure time and travel time profile were shown. Specifically. the OD demand was calibrated for 13 iterations, and an improved decreasing error between simulation and actual volume was observed 35% to 16.8%.

2 DYNUST MODEL DATA COLLECTION AND INITIAL PROCESSING

This section describes the process of developing a DynusT model through the DynuStudio platform, including the data requirements and the procedure for data collection and processing. DynuStudio is a data management platform designed explicitly for DynusT. As shown in Figure 2-1, network data (nodes and links) and origin and destination (O/D) demand are necessary for the model. A DynusT network is composed of zones, typically conventional Traffic Analysis Zones (TAZ) like in most transportation planning models, but can be any type of zone. Attributes of the zones include boundaries and locations. Attributes of the nodes include node ID's, coordinates, and associated TAZ's. Attributes of the links include the from/to nodes, length, number of lanes, road type, speed limit, and grade. Attributes of the time-dependent O/D records include the O/D zone, departure time, and volume. All of this necessary data has to be converted to the format required by DynusT.

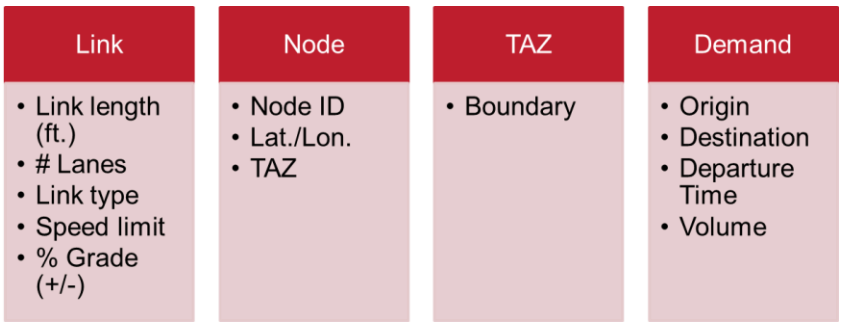


Figure 2-1: Data requirements of DynusT.

After converting the data, several steps are needed in order to complete the network buildout. Two of these steps are to tag the generation links and destination nodes in each zone. The generation links are the locations where vehicles leaving a zone are loaded. The destination nodes are the locations where existing vehicles are heading towards.

Figure 2-2 shows the general process for building a DynusT network:

- Mapping the terminal node of a link to a zone: This step assigns a terminal link to a zone (TAZ) centroid. Each zone is treated as an imaginary centroid connected to all destination nodes and generation links, but this node does not actually exist and is instead assigned by DynuStudio.
- Identifying generation links and destination nodes for each zone: This step assigns destination nodes to all existing nodes that are connected to a centroid

and assigns generation links to all outbound links from the destination nodes identified in the previous step.

- Calculating movement directions: Movements are required link attributes used to represent network connectivity and determine how vehicles can traverse from one link to another. This step uses the relative positions of connected adjacent nodes to determine the left-turn, right-turn, or straight movement.
- Completing sweeping network checks, cleanups, and adjustments: This step checks if the link connections, generation links, and destination nodes are in the correct zone. It also double-checks if the movement directions of all links are correct.
- Importing OD demand as a DAT file: This step converts the OD demand to a DAT file and imports the O/D demand function per DynusT's required format.

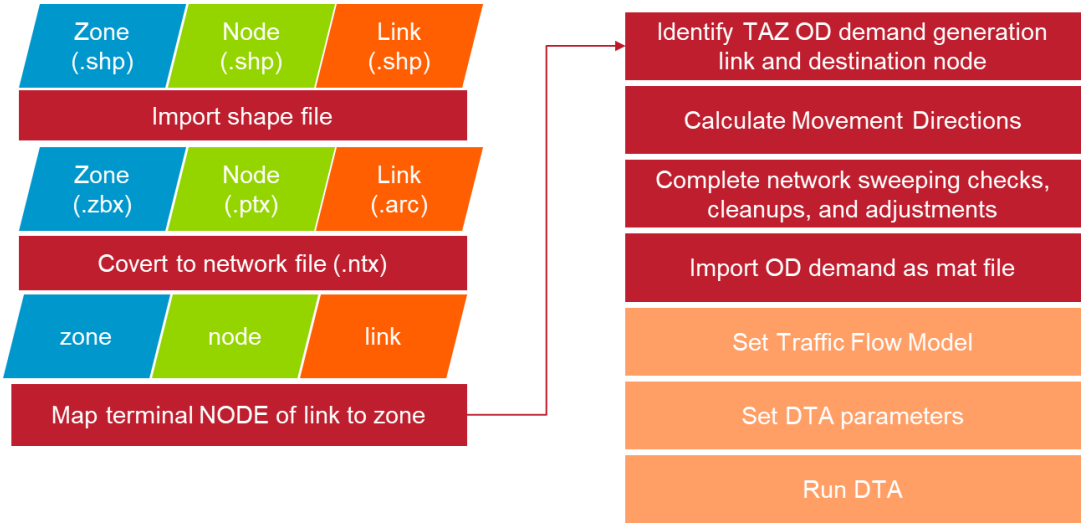


Figure 2-2: DynusT Network Buildout Process.

Other steps, such as setting up traffic flow models, setting up dynamic traffic assignment (DTA) parameters, and running DTA, are conducted after all the required data are imported into DynusT.

The following sections provide details regarding network data collection and process specific to this project.

1.1 Data Collection

GIS raw data (including the Link and Zone shapefiles) were provided by the Pima County Department of Transportation, and the majority of the required attributes related to links and zones (e.g., link name, link length, link type, number of lanes, speed limit, zone

ID) were contained in the shapefiles. A Node shapefile was not provided but was simply created by extracting the vertices from the Link shapefile, as indicated by Figure 2-4.

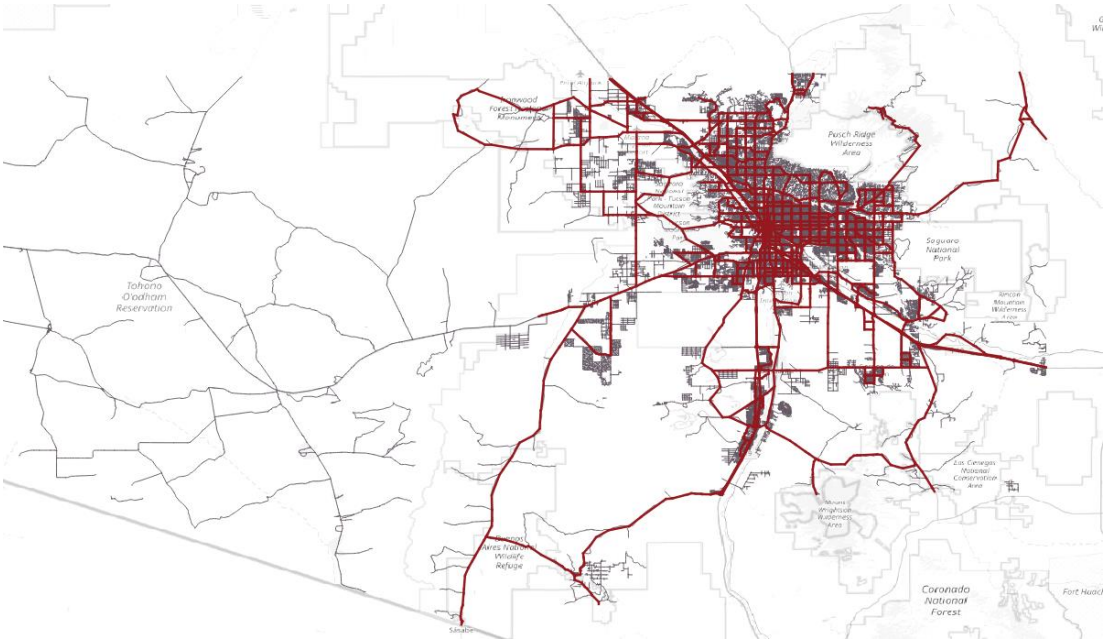


Figure 2-3: An Illustration of the Link Shapefile.

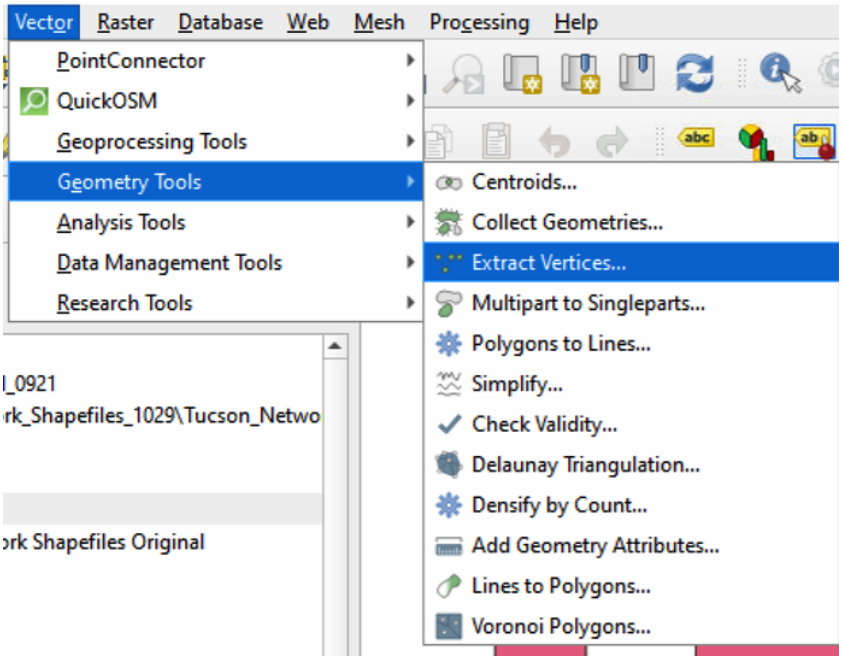


Figure 2-4: Creating the Node Shapefile.

The TAZ definitions in the DynusT model were consistent with the zones provided in the Zone shapefile, as shown in Figure 2-5. There were 1,104 TAZ's in total for the Tucson area, and they were indexed from 1 to 1,104.

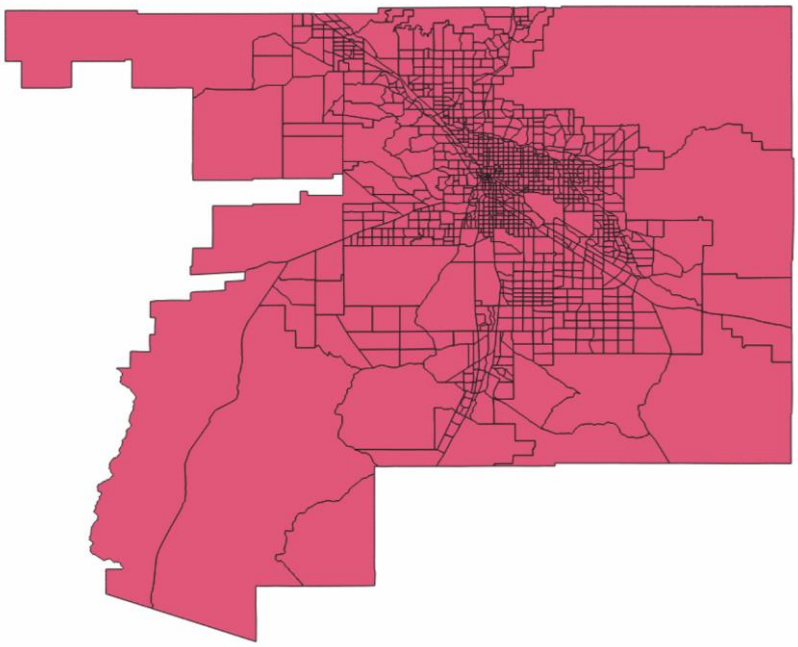


Figure 2-5: An Illustration of the Zone Shapefile.

With these zone definitions, the Origin-Destination (O/D) data was then processed. The time-dependent demand information was derived through StreetLight data, and more details are provided in Section 2.3. It is noteworthy that the TAZ definitions in the Zone shapefile were consistent with the ones in the StreetLight data. Thus, after the OD data generated by StreetLight were converted into the DynusT format, they directly inputted into the DynusT model without further adjustments.

Traffic data is necessary for calibrating the traffic flow model and the OD matrix in DynusT. Open data, including AADT, travel times, and travel speeds from StreetLight and Google Maps, were utilized as benchmarks.

1.2 DynusT Network Buildout Process

As previously discussed, the GIS-based network used for this study was provided by the Pima County Department of Transportation. The network provided desirable fundamental GIS raw data but needed additional processing such as filtering and compiling to prepare for a DynusT model. The following sections describe the processes of network clean up, the DynusT network import and conversion, and the DynusT network buildout.

1.2.1 Network Clean-up

For DTA simulation and modeling purposes, the GIS raw network needed to be cleaned up and processed. Several tasks were undertaken by the Metropia team, including:

- Trimming the Network

From Figure 2-3 we can see that the Link shapefile consisted of a large number of links of differing link types, but some of them were not within our area of focus. Therefore, the links that did not intersect with any TAZ's were filtered out first. Additionally, minor local roads, which accounted for 86.45% of the network, were trimmed off because they mainly facilitated the departure and arrival of vehicles and had little impact on the network traffic dynamics as a whole. Minor local roads primarily introduced vehicles into the network and took vehicles out of the network, and can be replaced with generation links and destination nodes defined in the DynusT model. Lastly, Indian routes were filtered out from the network since they were typically used by indigenous people and involved little to no vehicles.

After trimming the network, the scope of the network included the interstate highway, state routes, frontage roads, major local roads, the interstate ramps, state route ramps, and major road ramps (Figure 2-6).

Table 2-1: Initial Link Type Statistics

Road Category	Number of Links	%
Interstate highway	181	0.3
State route	475	0.788
Frontage road	248	0.411
Major local road	6223	10.32
Minor local road	52121	86.45
Indian route	765	1.269
Interstate ramp	220	0.365
State route ramp	23	0.038
Major road ramp	34	0.056

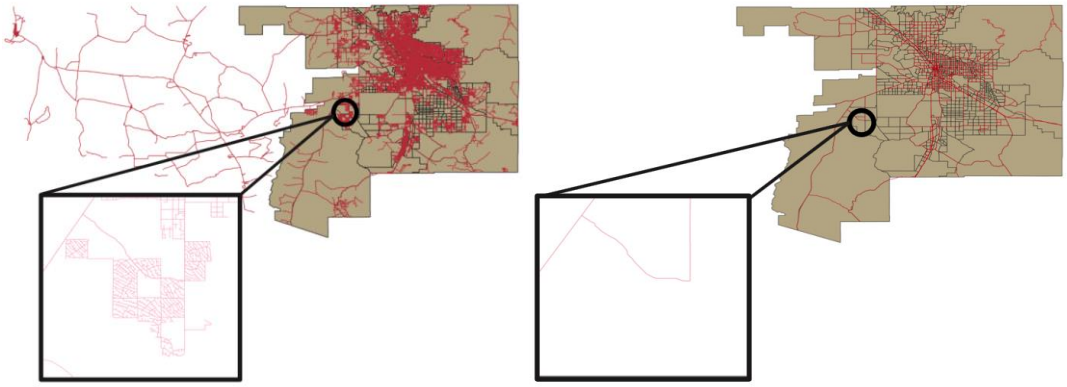


Figure 2-6: An Illustration of Trimming the Network.

(Left: Before Right: After)

- Merging Short Links

In order to represent roadway curve geometries, some facilities such as clover-leaf interchange links were divided into several smaller segments, as illustrated in Figure 2-7. All simulation models require a minimum link length, which is 300 ft for DynusT, as short segments can impact the DynusT model's realism. Furthermore, traffic dynamics on extremely short links give biased and misleading results in evaluating network performance. As such, we found it necessary to merge smaller links into longer links. We merged consecutive links whose lengths added up to 0.25 miles such that no links shorter than 0.25 miles remained in the network. This task was performed using an external Python script instead of the embedded function in QGIS.

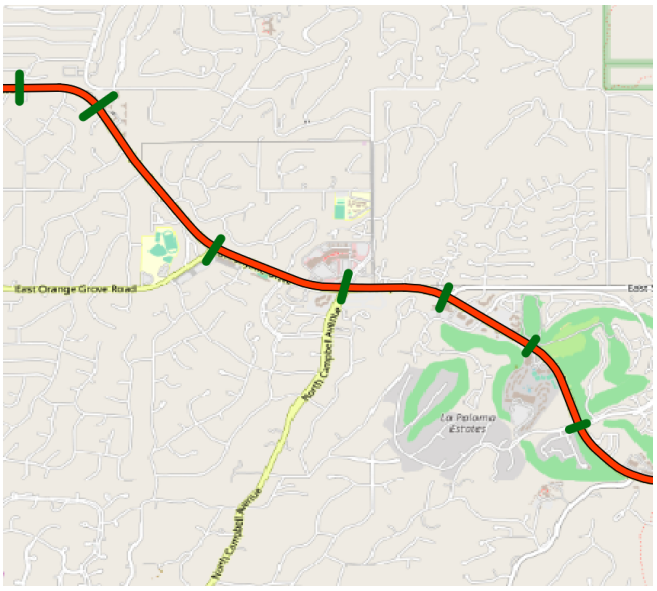


Figure 2-7: An Illustration of Short Links.

- Checking Network Connectivity

This step checked if all the links were connected because all GIS datasets have some random errors within. Although these errors may be invisible or trivial in GIS, they can lead to erroneous simulated traffic flows. We used the Disconnected Islands tool in the QGIS[1] software to check connectivity. Additionally, the shortest path algorithm was used to check if all freeway interchanges could be reached from the rest of the network. All interchanges were checked using this shortest path method. The results indicated that the Tucson network was fully connected, and every node was accessible from any other node in the network.

- Adding Generation Links and Destination Nodes

TAZ's are required in order to map the links and nodes to a zone such that DynusT can load generated vehicles emanating from a zone to associated generation links. Figure 2-8 shows the relationships between zones and generation links/destination nodes. Among the nodes within a specific TAZ, the four nodes closest to the TAZ centroid (north, south, east, and west of the centroid) were defined as the destination nodes for the TAZ, and the links connecting the TAZ centroid to these four nodes were defined as the generation links, as shown in Figure 2-9. According to this definition, the vehicles departing from a TAZ entered the network from these generation links, and the vehicles arriving at a TAZ left the network from the destination nodes.

Figure 2-8: Relationship Between Zones and Generation Links/Destination Nodes

1 https://qgis.org/en/site/

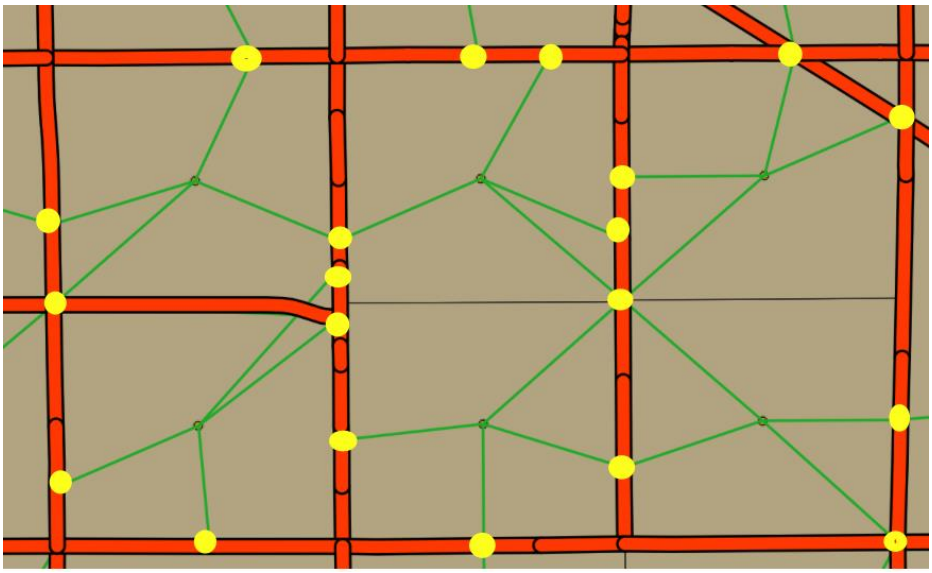


Figure 2-9: An Illustration of Adding Generation Links and Destination Nodes

(Green Line Segments: Generation Links Yellow Dots: Destination Nodes)

- Adding DynusT's Required Attributes

Since the shapefiles provided by the Pima County Department of Transportation contained most of the information required by DynusT, only the "link type" attribute for the Link shapefile, and the "node ID" and "node type" attributes for the Node shapefile were added to make the shapefiles complete.

- Link Type

The Link shapefile contained all the links in the network, including the actual network links and the virtual generation links that don't exist in the real network. Link Type is an attribute for differentiating them: the link type for generation links is 0, and the link type for actual links is ranging from 1 to 10.

- Node ID

The Node shapefile contained all the nodes, including the actual link nodes and TAZ centroid nodes. Node ID is the index used to differentiate between nodes. There were 1,104 TAZs in total, and numbers ranging from 1 to 1,104 were assigned to the centroids for TAZ #1 to TAZ #1,104.

- Node Type

The Node Type attribute was utilized to depict the control type of the node if the node represented an intersection. Since we didn't have signal control data, we set the Node Type of the nodes that represented intersections as 5, indicating that the intersections were under "Actuated Signal Control."

Overall, the statistics of the processed network are listed in Table 2-2.

Table 2-2: Network Statistics.

Number of Zones	1104
Number of Links	9247
Number of Nodes	5330

1.2.2 Importing and Converting the DynusT Network

Based on the previous step, the network shapefile was cleaned up, and all necessary attributes were prepared. The network was then imported into DynuStudio. The shapefile import tool imported the shapefiles (including the Link, Node, and Zone shapefiles) as an ARC file into DynuStudio (as shown in Figure 2-10).

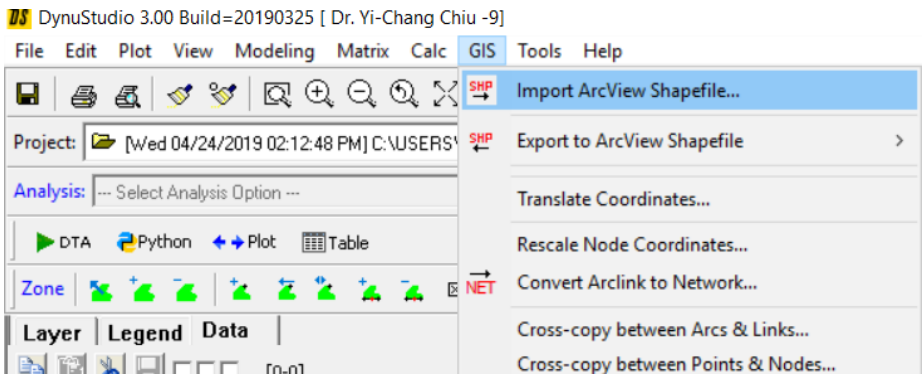


Figure 2-10: An Illustration of Importing a Shape File.

After importing the shapefile, the ARC file appeared in the Project Configuration window. Since the ARC file was not editable, we converted it into an NTX file to edit it in DynuStudio. The "Convert Arclink to network tool" conducted the conversion process (as shown in Figure 2-11)

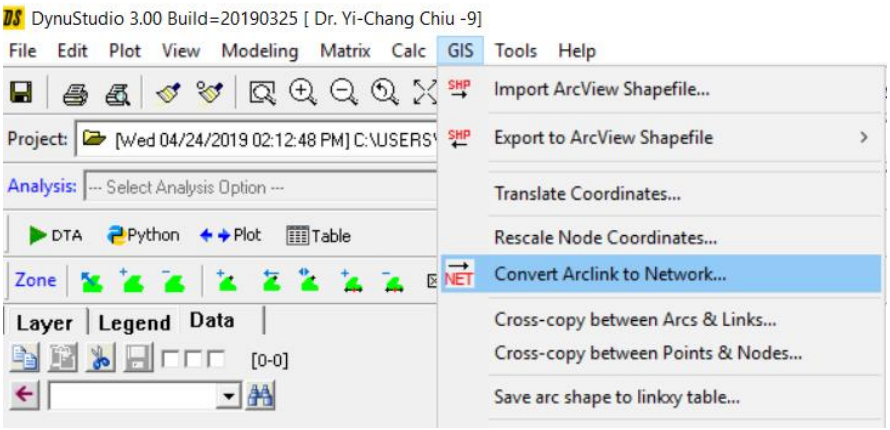


Figure 2-11: An Illustration of Converting an ARC File to an Editable Network.

Before starting the conversion, the direction, nodes, link length, link type, number of lanes, and the speed limit had to be set (as shown in Figure 2-12).

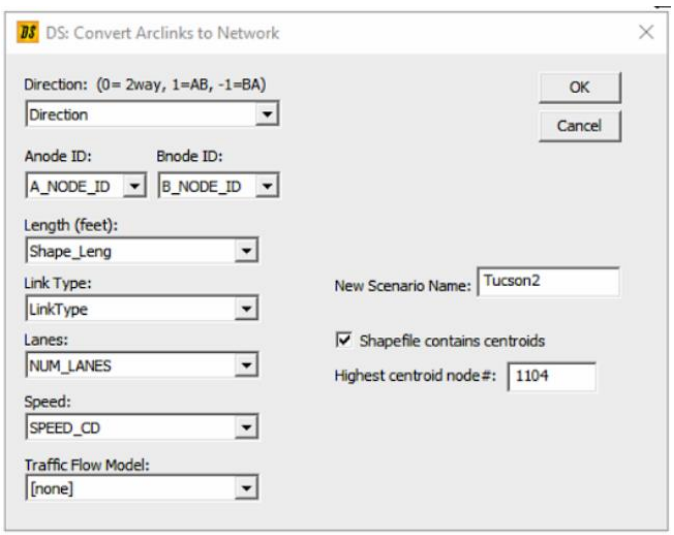


Figure 2-12: An Illustration of the Convert Arclink to Network Tool.

Once the tool completed the conversion, the network appeared on the plot area (as shown in Figure 2-13).

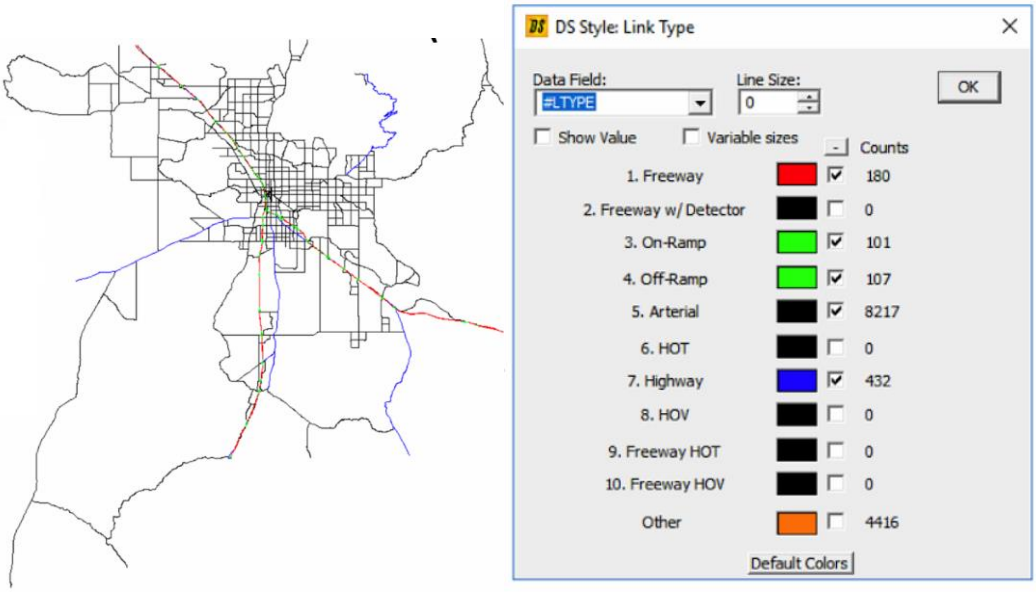


Figure 2-13: An Illustration of a Complete Network Conversion.

1.2.3 DynusT Network Buildout

The network import and conversion processes provided the necessary network data for DynusT. Three steps were then needed to build out the entire simulation-ready network in DynusT:

- Generating Generation Links and Destination Nodes

After the network conversion, TAZ information was added to the DynusT model. For each zone, we defined the corresponding generation links and destination nodes above. The covert centroids tool was used in DynusT (as shown in Figure 2-17). This tool performed two tasks: (1) it connected all the destination nodes to the imaginary centroid for that zone, and (2) it connected all generation links to the centroid. The conversion script calculated the loading weights based on the proportion of each link's lane miles to the total lane miles over all the generation links in the same TAZ. The generation links and destination nodes in the DynusT model are shown in Figure 2-15 and Figure 2-16.

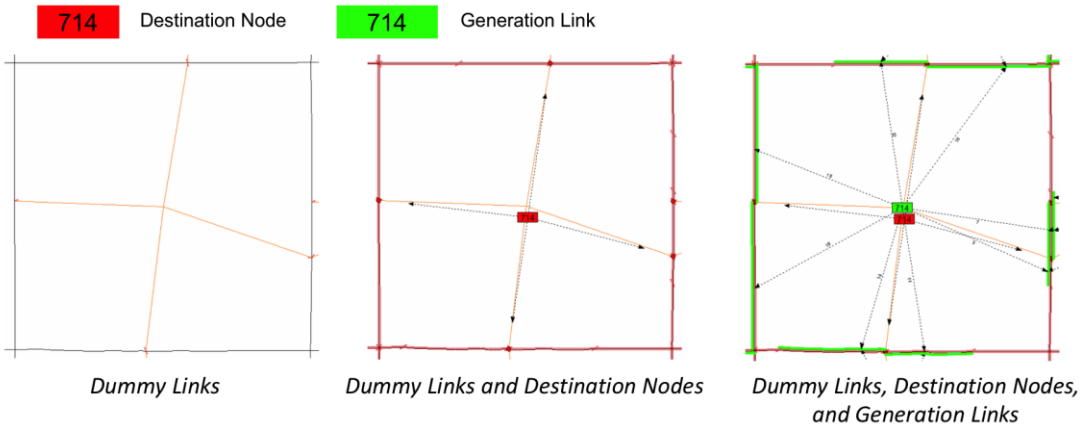


Figure 2-14: An Illustration of the Generation Links and Destination Nodes.

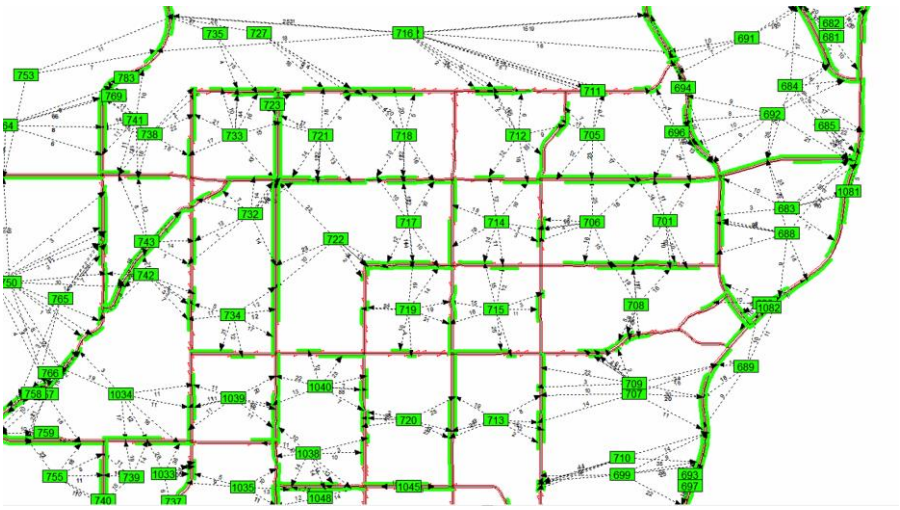


Figure 2-15: An Illustration of Generation Links in the Network.

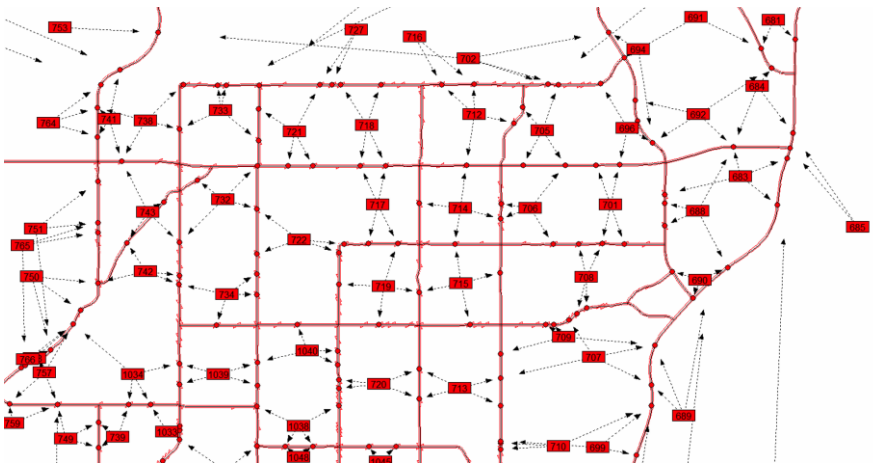


Figure 2-16: An Illustration of Destination Nodes in the Network.

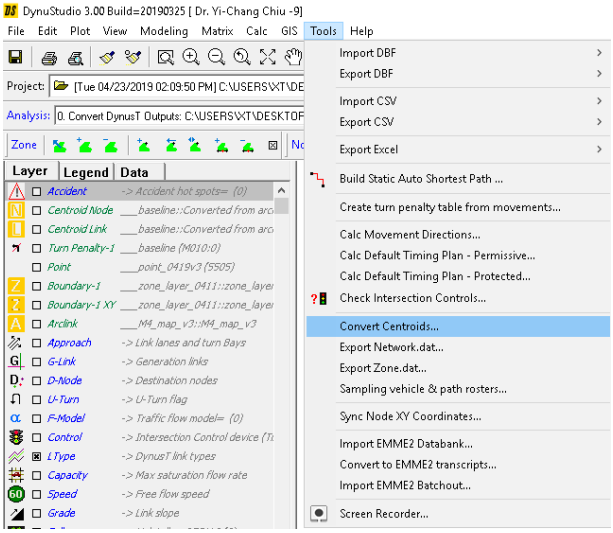


Figure 2-17: An Illustration of the Covert Centroids Tool.

- Calculating Movement Directions

Movements are required link attributes used to control how vehicles traverse from one link to the next. Movements are based on the connectivity of the adjacent outbound links to their downstream nodes and the relative orientations of these outbound nodes. These factors are used to determine which nodes can be reached via a left-turn, straight, or right-turn movement. The Calculate Movement Directions tool in DynuStudio generated movements automatically, but this automatic process is not 100% accurate (as many network editing tools are), so additional manual checks and revision were needed for nodes with unusual geometries.

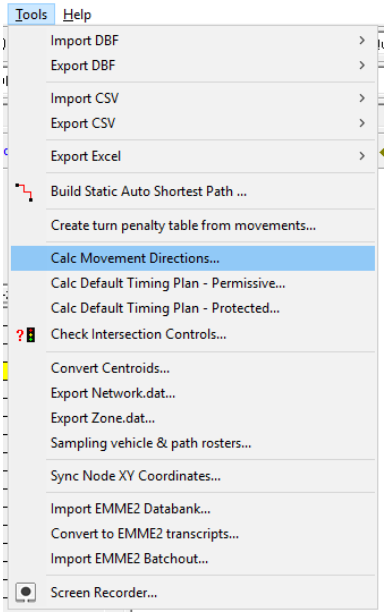


Figure 2-18: Movement Direction Calculation Tool.

The final model was visualized by loading Google Maps in the background to allow the modeler to reference the network with actual terrain on a map quickly.

- Defining the Link Traffic Flow Model

Each link was assigned with a speed-density (Greenshield family model) traffic flow model to govern its traffic flow properties during the simulation. The model simulated the capacity of each link since the Anisotropic Mesoscopic Simulation (AMS) model employed by DynusT obeys the traffic flow property $q = kv$, as shown in Figure 2-19. DynuStudio provides a tool for setting different traffic flow models. The user can set up to 10 traffic flow models, and each link can be assigned with one of the traffic flow models with the link attribute (#TFMODEL). DynuStudio sets traffic flow model 1 as a default value.

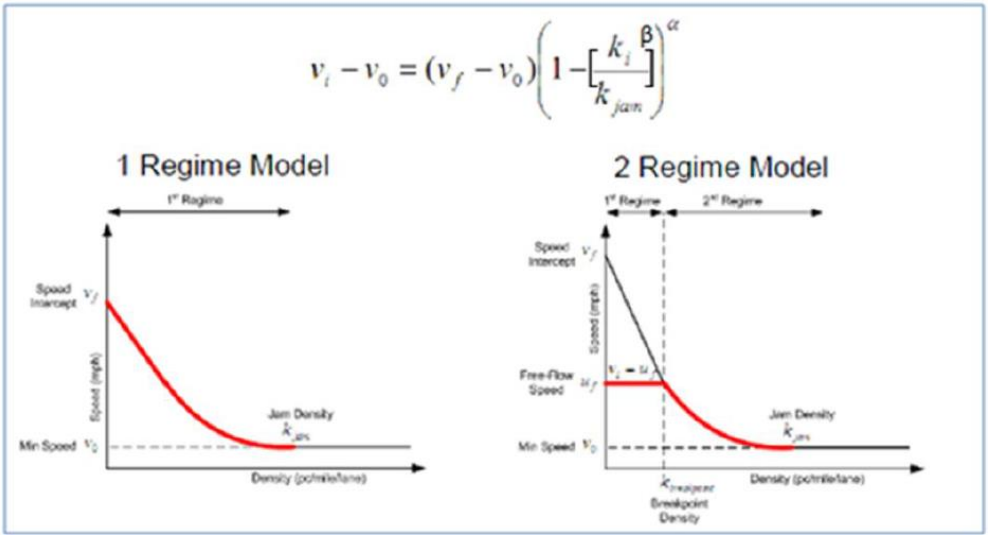


Figure 2-19: DynusT's Anisotropic Greenshield Family Speed-Density Curve Formulation.

Real traffic data was used to calibrate the model parameters, such as curve formulation, shape terms (alpha, beta), minimal speed, and jam density. Based on the calibration results, the traffic flow model was created and assigned to related links (Figure 2-20).

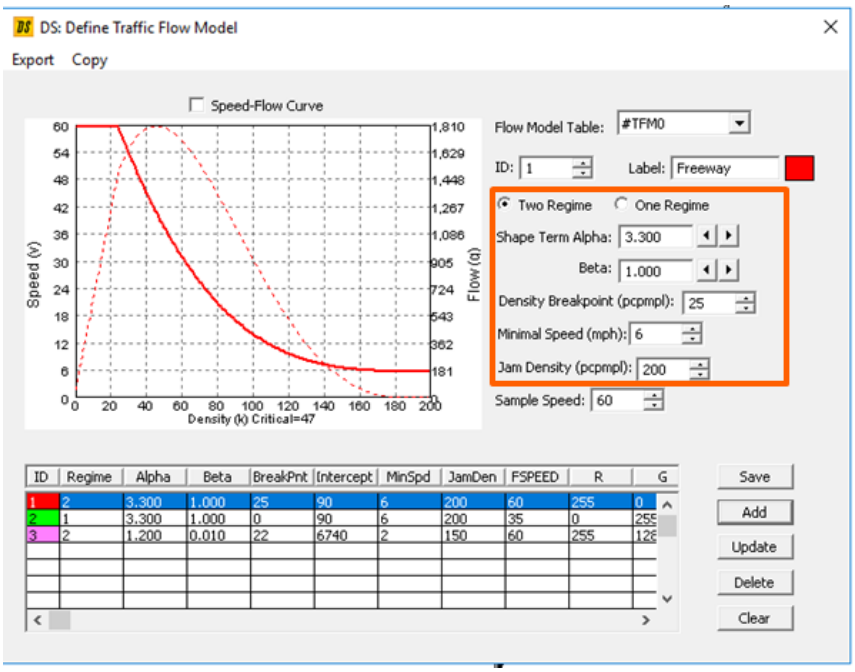


Figure 2-20: An Illustration of Defining the Traffic Model

1.3 DynusT's Time-Dependent Demand Matrices

1.3.1 Descriptions of the StreetLight Data[2]

StreetLight (STL) Data is a company that uses geospatial data points to measure how pedestrians, bikes, and vehicles interact. It provides demand input for DynusT. In order to generate the data, we needed to go through the following steps:

- Setting Traffic Analysis Zones (TAZ's)

First, we needed to set our area of interest to ensure that all the OD pairs were within this region. There were 1104 zones in total.

2 https://www.streetlightdata.com/

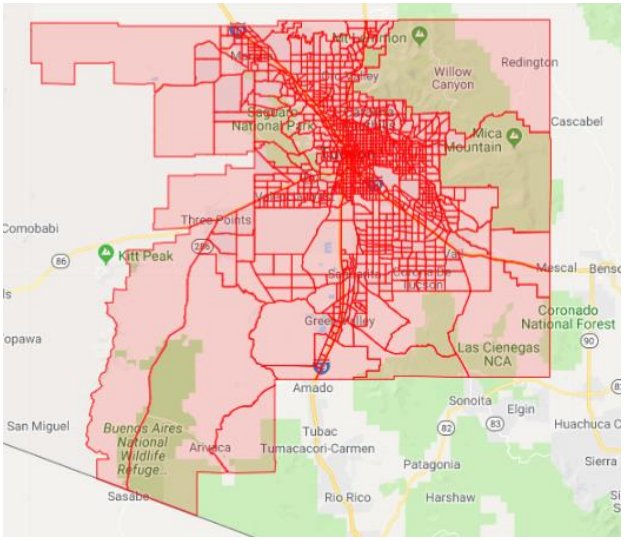


Figure 2-21: Traffic Analysis Zones.

- Setting Calibration Zones

STL used GPS data as a sample for their AADT estimates in order to calibrate their sample trip count to the estimated counts. Single-factor calibration was utilized to create a single normalization factor that was an average of the StreetLight Sample Trip Count compared to the StreetLight AADT estimates. This factor was then applied to the StreetLight Index, which scaled it to estimated counts.

We used 10 links in the TAZ's to get an unbiased estimate of vehicle counts. These links were on major roads and highways and had relatively higher AADT.

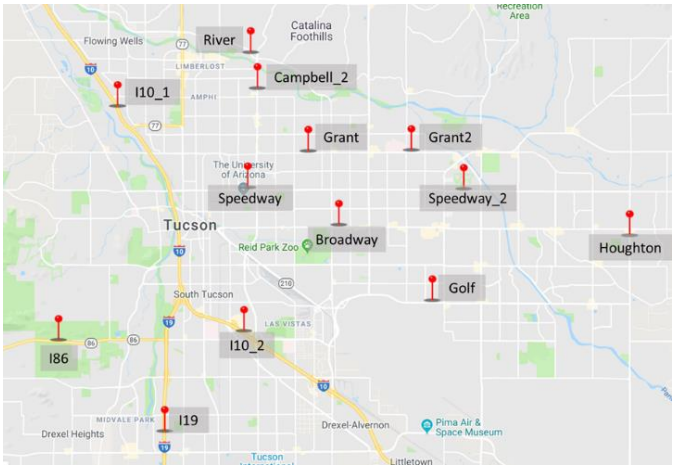


Figure 2-22: StreetLight Calibration Links.

- Setting the Time and Day to Generate Different Scenarios

We wanted to generate 24 hours of OD data from Monday to Thursday (weekdays) from STL. In order to compare the traffic patterns at different times, we generated demand information for three different scenarios. One was during the school season and included traffic data from March, April, September, and October. One was a school break that included traffic data in June, July, and August. Days after August 20th were excluded since most students were already back at school after that time. The last scenario was the snowbird season, where many people come to Tucson to enjoy a warm winter, and included data from November to January.

An example of the OD data we received from STL is shown in Table 2-3. We only needed weekday data, so we eliminated the rows that contained Day Type 2 and 0. OD Traffic (STL Index) represented the uncalibrated data that were collected directly by STL. OD-Traffic (Calibrated Index) contains vehicle counts calibrated by AADT. Zone traffic is the number of vehicles in the origin or destination zones.

Table 2-3: Sample of Raw OD Data.

Origin Zone ID	Destination Zone ID	Day Type	Day Part	O-D Traffic (StL Index)	**O-D Traffic (Calibrated Index)**	Origin Zone Traffic (StL Index)	Origin Zone Traffic (Calibrated Index)	Destination Zone Traffic (StL Index)	Destination Zone Traffic (Calibrated Index)
501	307	1: Average Weekday (M-Th)	20: 7pm (7pm-8pm)	1	1	97	164	138	233
633	1010	2: Average Weekend Day (Sa-Su)	12: 11am (11am-12noon)	7	11	114	191	143	240
263	57	0: Average Day (M-Su)	19: 6pm (6pm-7pm)	1	1	329	553	244	410
692	686	0: Average Day (M-Su)	20: 7pm (7pm-8pm)	2	4	74	124	67	112
68	247	1: Average Weekday (M-Th)	11: 10am (10am-11am)	2	3	205	345	322	541
154	499	0: Average Day (M-Su)	14: 1pm (1pm-2pm)	1	1	282	475	102	171

There were 1104 zones, theoretically giving us 1218816 sets of OD data for each time interval, with 24 time intervals in total. However, if the calibrated traffic count for an OD pair for a specific time period (e.g., Early AM) was below StreetLight's significance threshold, no results were shown in the files. The number of OD pairs provided by STL is summarized in Table 2-4.

Table 2-4: Number of OD Pairs.

Scenario	Months	# of OD per day
School Season	Mar, Apr; Sep, Oct	2990225
Summer (School Break)	June, Jul, Aug	2754810
Snowbird Season	Nov, Dec, Jan	2879107

Other than OD data, STL also provided Average Travel Time, Average Travel Speed, and trip length. Some trips had extra pings at the start and/or end, which caused their average travel times to be longer than expected and average travel speed to be lower than expected. Therefore, travel time and speed data were not used for future validation.

1.3.2 StreetLight Data Rationality Check

- Missing Values

In this section, we wanted to examine the raw dataset and see if it was reasonable enough to be used as the input demand data. As we mentioned above, there were some missing data in the raw dataset. They were not recorded in the dataset because the confidence level of the calibration did not meet the criteria of Streetlight. Therefore, we also needed to check if the dataset contained enough important information in order to run the simulation model.

While we should have 1218816 OD pairs (1104 × 1104 × 24), we were not interested in all of them, as some OD's had almost zero vehicles passing or were too early or late in terms of time. Therefore, we selected the most important links by their volume (hourly volume larger than 5 vehicles). This took up 0.7% of the total number of OD pairs. The missing rate distribution is shown in Figure 2-23. Most of these links had no missing values from 8 AM to 5 PM. The missing rate was a little higher if we considered all 24 hours because almost no vehicles passed through early in the morning or late at night. The average day time missing rate was 30%. This is enough for the input demand.

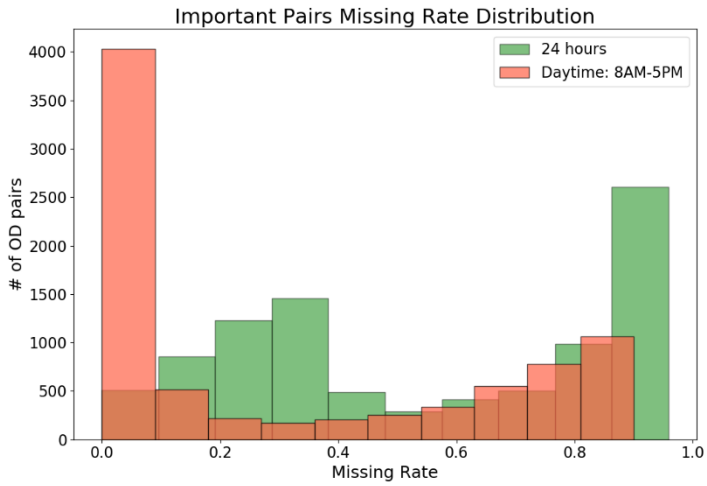


Figure 2-23: Important Links Missing Rate Distribution.

We were also interested in where the ODs had the highest missing value. Therefore, we visualized some of the OD pairs on the map. The width of the lines represented the

missing rate. The selected links all had a missing rate higher than 60%. As can be seen in Figure 2-24, longer OD pairs tended to have more missing values.

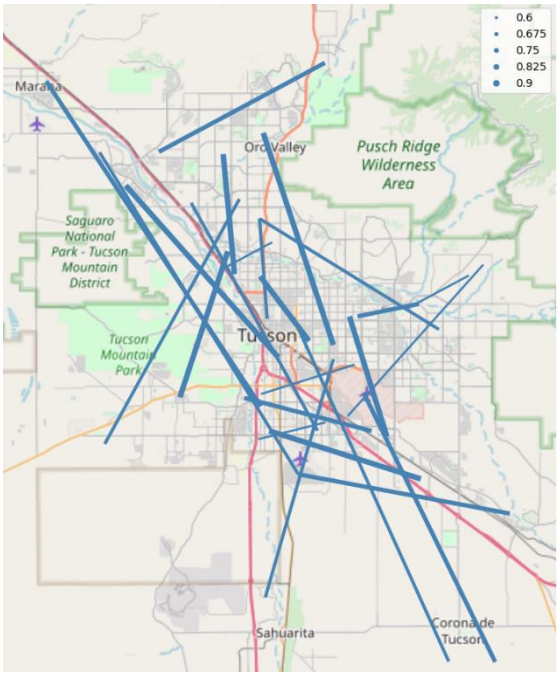


Figure 2-24: Visualization of the OD pairs with a missing rate greater than 0.6 on the map.

- Comparing with Permanent Counts

In order to see if the data overestimated or underestimated traffic conditions, AADT collected by fixed sensors were also utilized for comparison. By aggregating the OD demand, the volume of calibration links was calculated. The locations of the 30 links are shown in Figure 2-25.

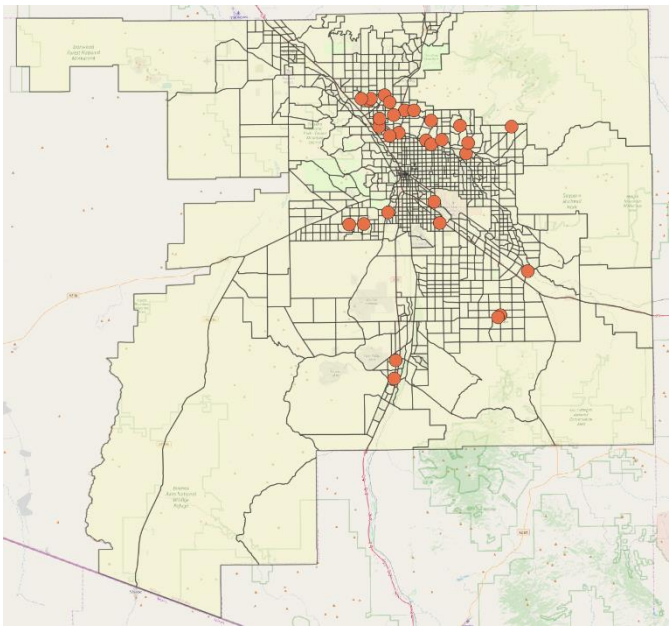


Figure 2-25: Locations of the Calibration Links (Orange).

The screen line counts represented the AADT collected from fixed sensors. The DynusT simulation volume was the output of DynusT fed with the StreetLight OD data. The linear fit shown with the 45-degree line is represented by the dashed yellow line. From Figure 2-26, we can see that all the data points coded in orange were under the 45-degree line, indicating that the StreetLight OD data underestimated the traffic conditions. This led to biased results in the DynusT simulation.

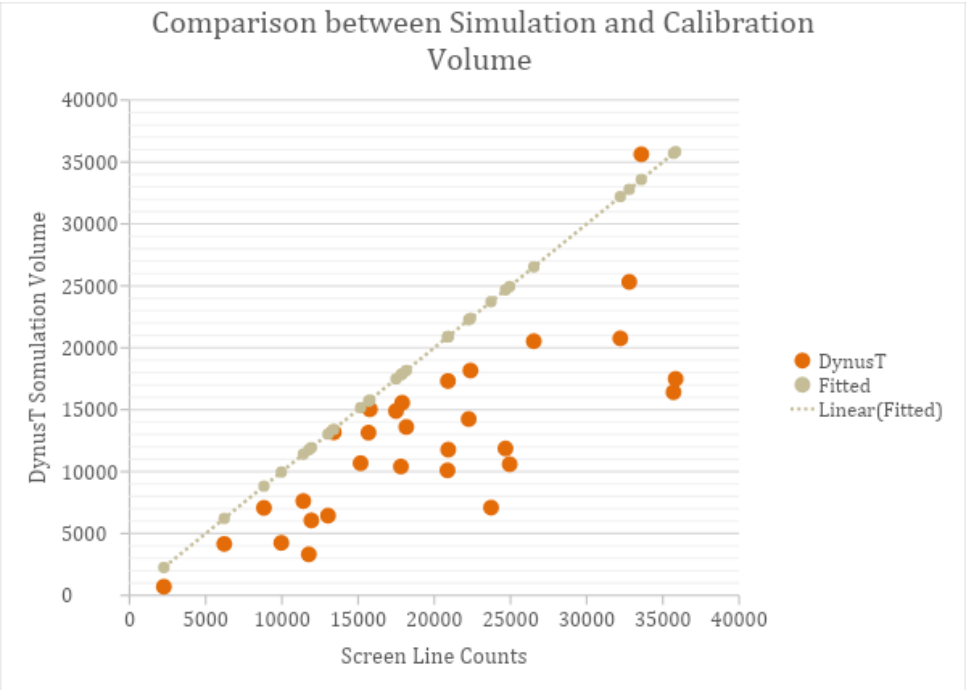


Figure 2-26: StreetLight OD Underestimated Traffic Volume.

To summarize, StreetLight OD data represented general traffic patterns and could be used as input for DynusT because the missing rate was not high. However, OD data were underestimated, leading to biased simulation results. Therefore, the demand data needed calibration with the help of permanent counts such that the simulation results reflected real traffic patterns. The next section describes a few observations of the StreetLight data.

1.3.3 StreetLight OD Data Analysis

- Spatial Distribution

We visualized the hourly origin and destination vehicle counts on the map in order to see a general traffic pattern for the three scenarios. In the school scenario (Figure 2-27), the hot spots were at Davis–Monthan Air Force Base and the University of Arizona during the morning peak. The hot spots were more spread out near Sahuarita and Cortaro during the afternoon peak.

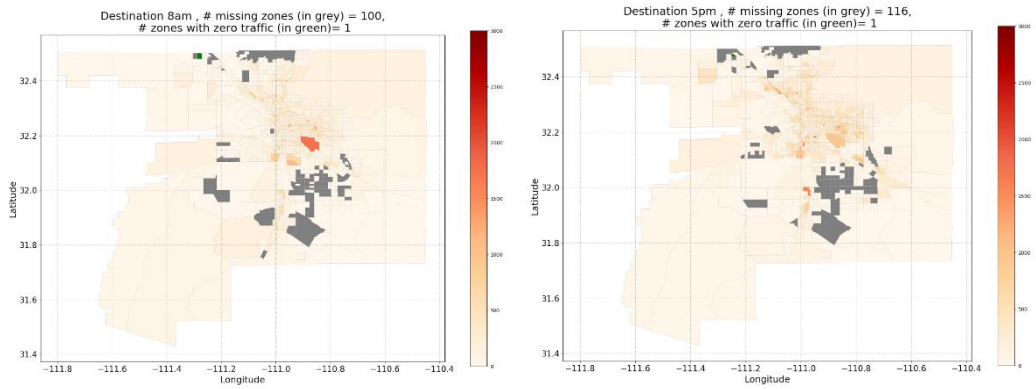


Figure 2-27: Heatmap Indicating the Number of Vehicles at 8 AM and 5 PM During the School Season. Missing Zones are Colored Gray.

In the summer (school break) scenario (Figure 2-28), the hot spot was mostly at Davis–Monthan Air Force Base. Compared with the school scenario, fewer vehicles drove into the school area during the morning peak. The hot spots were similar to those of the school scenario during the afternoon peak.

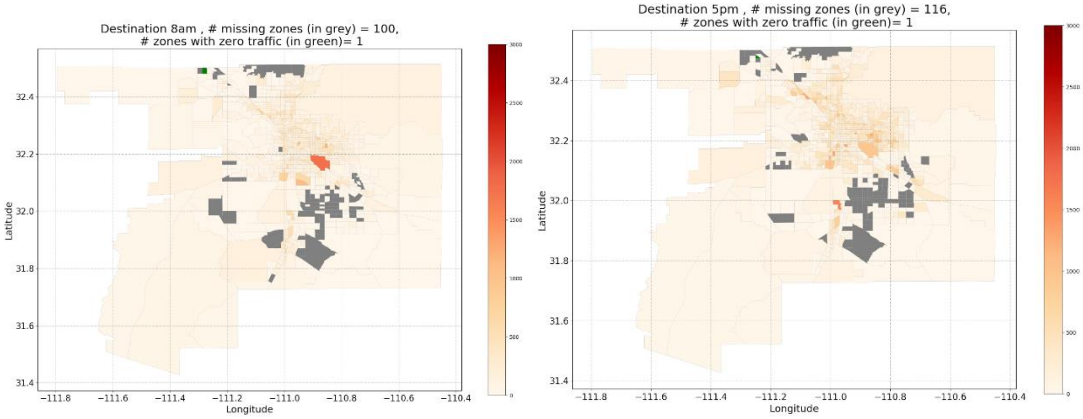


Figure 2-28: Heatmap Indicating the Number of Vehicles at 8 AM and 5 PM During the Summer Break. Missing Zones are Colored Gray.

From the snowbird scenario (Figure 2-29), the hot spot was still at Davis–Monthan Air Force Base. Compared with the school scenario, fewer vehicles drove into the school area during the morning peak. The hot spots were similar to those of the school and summer scenarios during the afternoon peak.

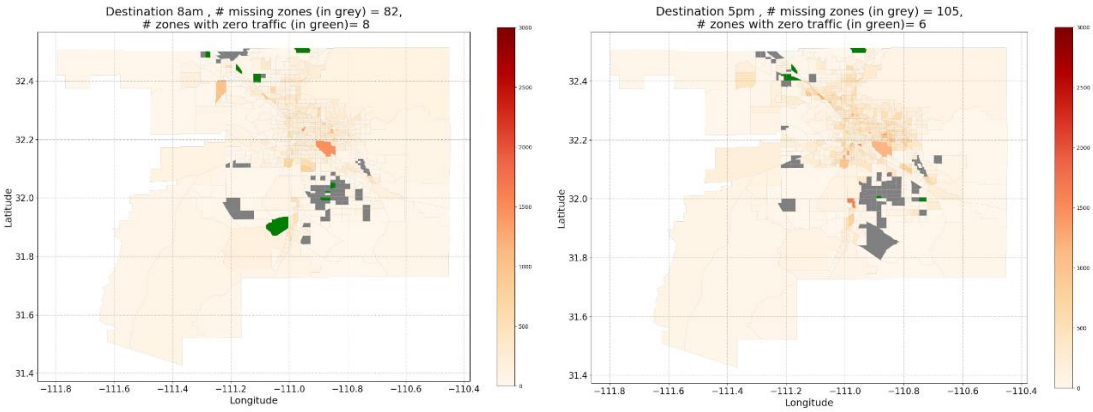


Figure 2-29: Heatmap Indicating the Number of Vehicles at 8 AM and 5 PM During the Snowbird Season. Missing Zones are Colored Gray.

Besides, we also compared the distributions of the missing zones during peak and non-peak hours (Figure 2-30). Zones colored in green represent zones with zero volume during the morning and afternoon peak. We found that there were more zero-traffic zones and zones with missing counts during off-peak hours, like 1 AM. Most of the missing zones during peak hours were actually mountain areas, where not many vehicles are generated or attracted to.

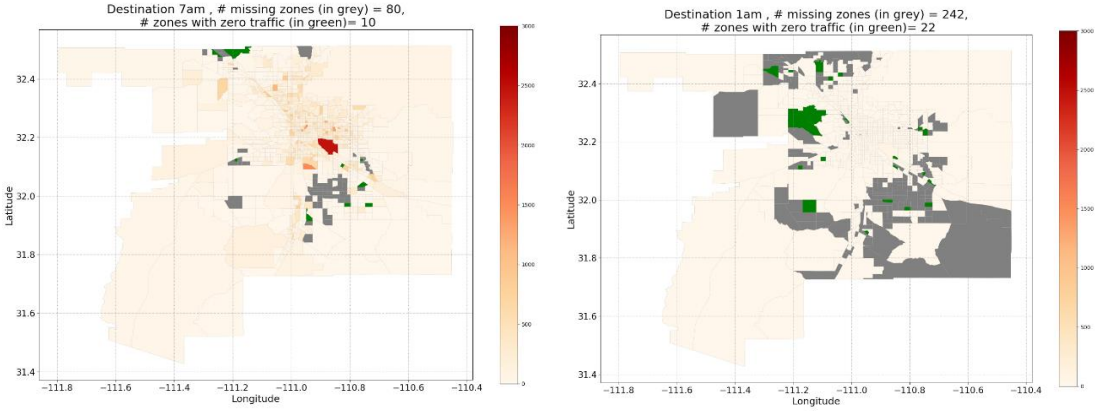
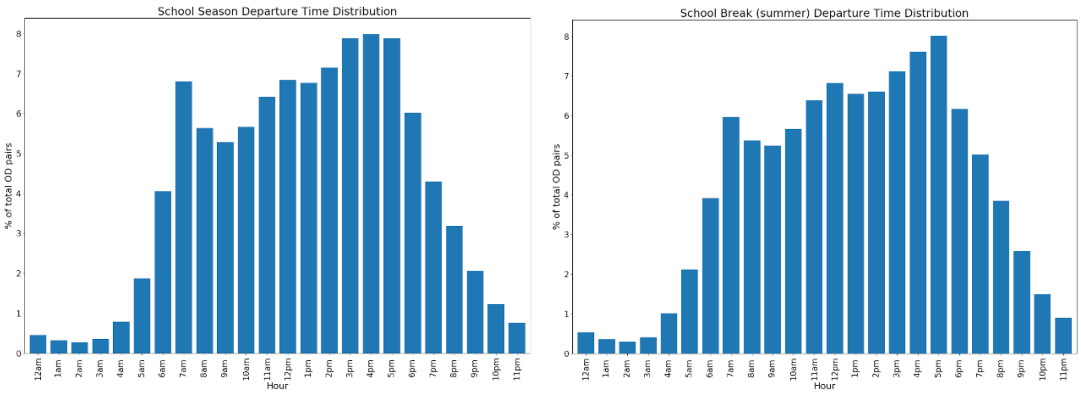


Figure 2-30: Peak-hour (7 AM) and Off-peak Volume (1 AM) Comparison.

To summarize, StreetLight OD data represented the true values of traffic and could be used as input for DynusT if we wanted to analyze peak-hour traffic patterns. In addition, however, volume collected by loop detectors or fixed-point sensors might be needed in order to observe off-peak traffic patterns.

- Temporal Patterns

Departure time distributions of the three scenarios were visualized from the StreetLight data in Figure 2-31. The highest morning volume was observed during the school season, where the traffic might be mainly generated by students. Afternoon peaks were higher and longer during the snowbird season, where the traffic might be generated by visitors. This is in line with our perceptions.



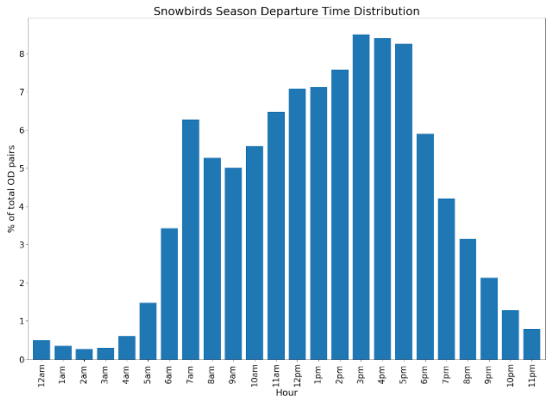


Figure 2-31: Departure Time Distributions.

1.3.4 Converting Streetlight Data to the DynusT Format

A few cleaning steps were done before converting the StreetLight spreadsheet into the DAT file required by DynusT. First, only weekday OD data were conserved since we only wanted to analyze the weekday patterns. Second, missing data were imputed as N/As in order to make the whole dataset complete such that the zone ID was continuous. Third, the data were sorted based on start hour, origin ID, and destination ID. The dataset after the cleaning steps looked like Table 2-5.

Table 2-5: Processed Data Example.

Origin Zone ID	Destination Zone ID	O-D Traffic (Calibrated Index)	Start Hour (min)
1	1	0	0.0
1	2	1	0.0
	⬅Continuous: 1 - 1104➡		
1	3	0	0.0
...	...	...	...
1	1	2	60.0
1	2	1	60.0
...	...	...	...
1104	1	3	0.0
1104	2	5	0.0
...	...		
1104	1104	0	1440.0

Next, the data were converted to the DynusT format. The process was done in Python.

1104*1104*24 time-dependent OD (TDOD) matrices were created in the DynusT format. Table 2-6 shows the TDOD in the DynusT format. The file includes (1) the number of demand intervals, (2) the demand loading factor, and (3) the starting time for each OD matrix

Table 2-6: DynusT Demand Format.

```
   24 1.000  # start time intervals, growth factor
      0.0   60.0   120.0   180.0   240.0   300.0   360.0   420.0   480.0
540.0   600.0   660.0   720.0   780.0   840.0   900.0   960.0   1020.0
1080.0   1140.0   1200.0   1260.0   1320.0   1380.0   1440.0  Start time
Start Time =     0.0  Each start time
    0.0000     0.0000     0.0000     0.0000     0.0000     0.0000
    0.0000     0.0000     0.0000     0.0000     0.0000     0.0000
    0.0000     0.0000     0.0000     0.0000     0.0000     0.0000
    0.0000     0.0000     0.0000     0.0000     0.0000     0.0000
    0.0000     0.0000     0.0000     0.0000     0.0000     0.0000
    0.0000     0.0000     0.0000     0.0000     0.0000     0.0000
    0.0000     0.0000     0.0000     0.0000     0.0000     0.0000
    0.0000     0.0000     0.0000     0.0000     0.0000     0.0000
    0.0000     0.0000     0.0000     0.0000     0.0000     0.0000
    0.0000     0.0000     0.0000     0.0000     0.0000     0.0000
    0.0000     0.0000     0.0000     0.0000     0.0000     0.0000
    0.0000     0.0000     0.0000     0.0000     0.0000     0.0000
    0.0000     0.0000     0.0000     0.0000     0.0000     0.0000
    0.0000     0.0000     0.0000     0.0000     0.0000     0.0000
```

6 values per line; new origin ID also starts in a new line

This format was saved as a DAT file and was imported into DynuStudio using the O/D demand function import tool.

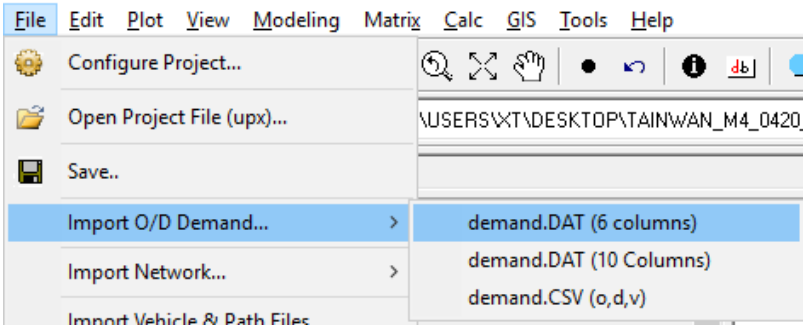


Figure 2-32: An Illustration of the OD Demand Import Tool.

2 DYNUST MODEL VALIDATION AND CALIBRATION

The baseline model needed to be reasonably calibrated and validated in order to use it for later comparative analysis. Figure 2-1 shows the general steps for DynusT model validation and calibration. We collected traffic flow, travel time, and travel speed data using real-world observations. Then, DynusT simulated the traffic flow based on the given network and demand, which we prepared beforehand. After the simulation, DynusT provided route travel time, travel speed, and volume. These Measures of Effectiveness (MOE) were then compared with real-world traffic conditions. The traffic flow model was then modified and adjusted according to the comparison results, and traffic was then re-simulated. This was done in an iterative process, by performing a comparison (validation), parameter adjustments (calibration), and re-simulation. The following sections describe the process of validation and calibration.

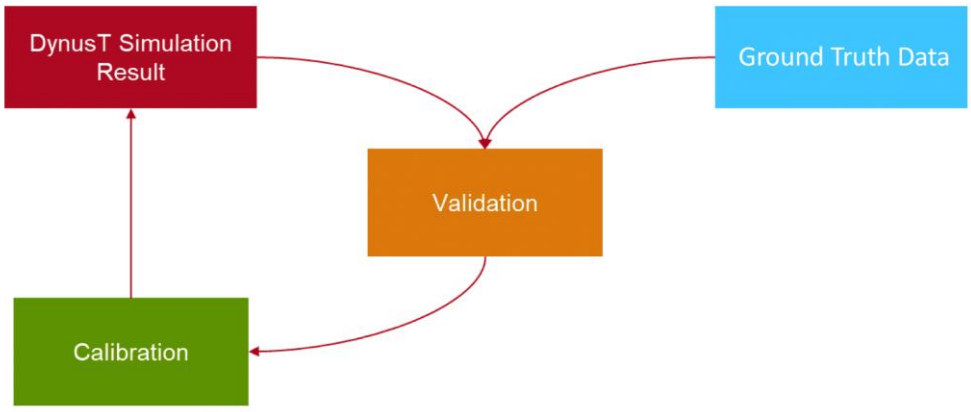


Figure 2-1: DynusT Model Validation and Calibration Process.

2.1 DynusT Network Connection Testing

After importing all the required data, we ran the DTA simulator (as shown in Figure 2-2). In the Run Window in DynuStudio, a user can set the traffic assignment method (one-shot vs. iterative equilibrium), set demand input methods (OD-based or trip roster-based), vehicle type rate, and other parameters. The simulator checks if all the input files and parameters are correctly prepared and, if so, starts to run the simulation-based traffic assignment. Before inputting the real-world demand into the model, we needed to test the connectivity of the network in DynuStudio. The test demand for each pair of origins and destinations was set as 1, and then the simulation was fired up. No errors popped up, let us say that the network is fully connected to the DynusT model.

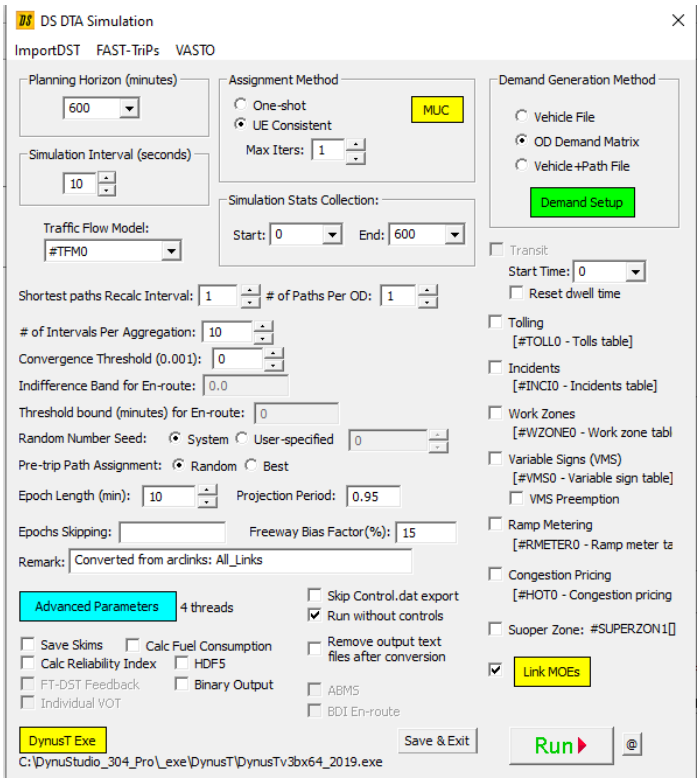


Figure 2-2: An Illustration of the DTA Simulator.

2.2 General Simulation Outputs and Statistics

A 1440-min (24 hours) simulation horizon run took about 30 minutes and 12GB memory on a Windows server with E5-2680 v2 2.8Ghz CPU. The primary output statistics for the simulation are listed below in Table 2-1. A total of 2.9 million vehicles were generated. The average travel time was 16.72 minutes, of which 1.72 minutes were in a stop-and-go situation. The average travel distance was 5.88 miles. The average speed was 21.09 mph, but this included both peak and off-peak hour statistics. The actual speed for vehicles during peak hours may be much lower.

Table 2-1: DynusT Simulation Output Statistics.

```
For All Vehicles in the Network
    Number of Vehicles              =        2929718
    Total Travel Time w/o queueing =    48985868.000
    Total Trip Distances            =    17218770.000
    Total STOP Time                 =     5047829.500
    Average Travel Time             =          16.720
    Average Trip Distance           =           5.877
    Average STOP Time               =           1.723
    Average travel Speed            =          21.090
```

From the departure (red) and arrival (green) profiles shown in Figure 2-3, one can see the clear demand spike starting at 6 AM in the simulation. The travel time (blue) profile shows that those who have long trips preferred departing early.

Figure 2-3: Departure, Arrival, Trip Duration (Travel Time), and Delay Profiles for the Baseline Run.

DynuStudio provides multiple analysis tools for analyzing simulation results at both the link and corridor levels for one or two scenario comparisons (as shown in Figure 2-4).

Figure 2-4: An Illustration of the Analysis Tool.

Simulation animations were also provided in the same analysis tool list. Several selected simulation animation videos are available in the playlist below.

https://youtu.be/ijHFOTCYFog

https://youtu.be/rct7PfICveE

https://youtu.be/rct7PfICveE

2.3 OD Calibration and Validation

In order to make the simulation results closer to real-life data, we utilized the DynusT calibration tool to adjust the input demand. The AADT of 30 links in Pima County collected by 30 fixed sensors were employed for calibration. The locations of the sensors are shown in Figure 2-5.

The calibration process can be summarized in the following diagram. One norm linear minimization problem was formed to reduce the difference between the calibration and simulation volume. The demand file was then rebuilt and fed into DynusT again. The volume calculated from the simulation was compared with the calibration data again to see if the difference was small enough.

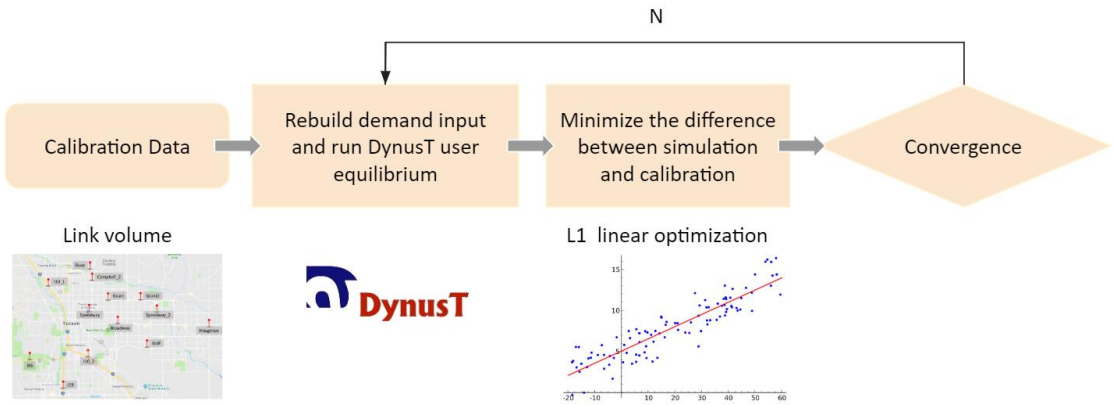


Figure 2-5: The Calibration Process.

After 13 iterations of calibration, the results converged. The calibrated counts compared with the screen line counts are shown in Figure 2-6. The gray dashed line is the 45-degree line. We can see that the calibrated counts were closer to the gray line, indicating that the errors between the simulation and real demand were reduced over iterations.

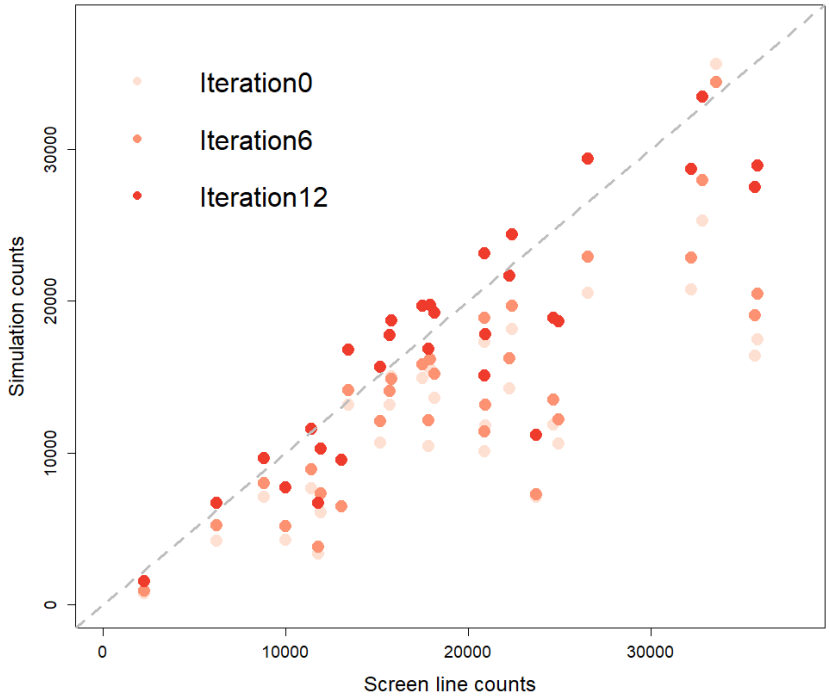


Figure 2-6: Simulation Counts Compared with Screen Line Counts After Iterations 0, 6, and 12.

The errors between the simulation counts and the screen line counts of each link grew smaller over various iterations and can be seen in Figure 2-7. During the first 8 iterations, the calibration tool was only allowed to change 20% of the total demand, so the error between the simulation counts and screen line counts did not decrease much. Because StreetLight data underestimated the OD much, the model was allowed to increase to a higher percentage of total demand during the next few iterations. The allowed change was then increased to 100%, and the error decreased dramatically, especially for links whose volume had relatively high deviations from screen line counts, which changed from red to green immediately on the figure.

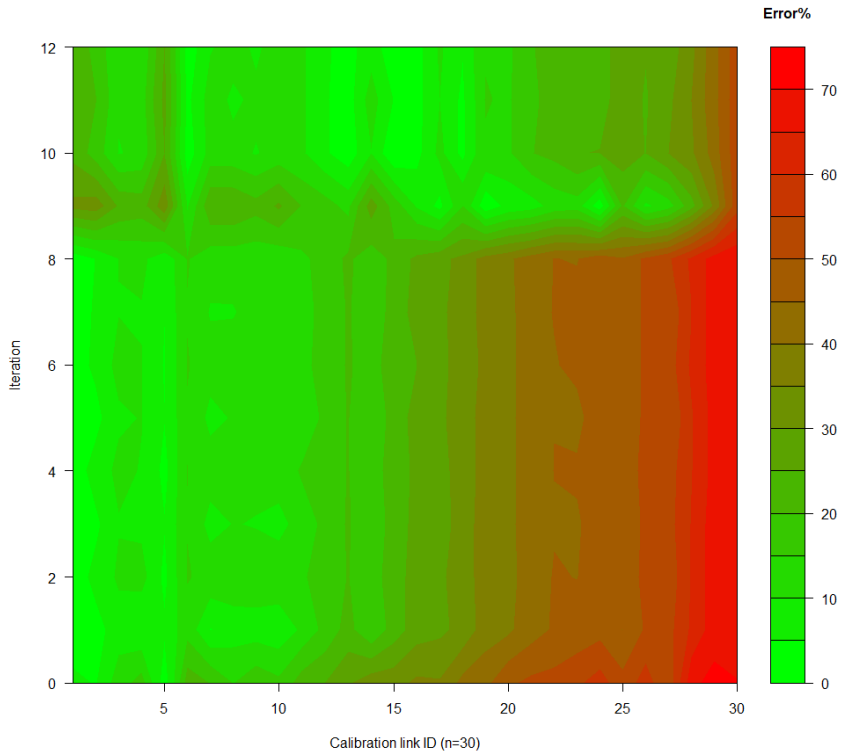


Figure 2-7: Differences between Simulation Counts and Screen Line Counts After Each Iteration

After the calibration, we reran DynusT, and the simulation counts were compared with the screen line counts for validation. We denoted the screen line counts of the ith link as SL_i and the DynusT simulation counts of the ith link as DC_i. The weighted mean absolute percentage error (WMAPE) was calculated using the following formula:

$$WMAPE = \sum_{i=1}^{30} \frac{SL_i}{\sum_{i=1}^{30} SL_i} \cdot \frac{|SL_i - DC_i|}{SL_i}$$

Links with higher volume had a higher weight when calculating the error. Comparisons between the simulation counts and the screen line counts are shown in Table 2-2. The error was reduced from 35.06% to 16.76%, indicating that the calibration tool corrected for the demand in the right direction and amount.

Table 2-2: Summary of Counts and Errors on 30 Calibration Links.

	Before Calibration	After Calibration
Total Simulation Counts (veh)	383412	549002
WMAPE	35.06%	16.76%

3 CONCLUDING REMARKS

Based on our years of experience in DTA modeling, the model established by the UArizona team has exhibited several innovative and exciting features and outcomes.

- Network preprocessing was done in QGIS and coded in Python, including generation links and destination node generation and network trimming. Since the processes were all pipelined, they can easily be extended or applied to other data.
- To the best of our knowledge, this is the first time the StreetLight data origin-destination data were used to serve as the primary data source for the development of the countywide large-scale time-dependent OD matrices. The detailed data management steps discussed in the previous sections clearly highlight the novelty in the modeling process. We also found that the incomplete and underestimation problems existed in the StreetLight data, which might cause bias in the simulation results. This finding can be used as a reference in the future use of StreetLight data. However, we came up with the solution in the Model Calibration section.
- In order to address the issues of underestimated OD data, we carried out a calibration scheme in the DynusT simulation process by solving a linear constrained minimization problem. The innovative utilization of the calibration tool successfully reduced the bias caused by StreetLight data, and the error of the simulation model was acceptable (16.7%) after calibration. The calibration tool can also be applied in other scenarios where the demand data were biased.