DRAFT REPORT

BIG DATA ANALYTICS FOR THE NORTHEAST INDIANA REGION

the science of insight

12.21.2017

PREPARED FOR:
NORTHEASTERN INDIANA REGIONAL COORDINATING COUNCIL

SUBMITTED BY:
RSG

2709 Washington Ave., Suite 9
Evansville, Indiana 47714
802.295.4999
www.rsginc.com

IN COOPERATION WITH:
CONVERGENCE PLANNING

PREPARED FOR:
NORTHEASTERN INDIANA REGIONAL COORDINATING COUNCIL

CONTENTS

List of Figures

List of Tables

1.0 INTRODUCTION

The Northeastern Indiana Regional Coordinating Council (NIRCC) has made important investments in passively collected big data to better understand the movements of both people and truck freight into, out of, through, and within Northeastern Indiana. This report documents the selection, processing, and analysis of this data and what it reveals about travel patterns in Northeast Indiana.

The study was organized into four component tasks:

- Data Selection
- Data Validation
- Data Expansion
- Data Analysis and Visualization

Each of these tasks is documented in the following sections of the report. The data selection section presents information on the various sources of big data available at the time of the study, the advantages and disadvantages of the various datasets, and rationale for the ultimate decision to purchase Streetlight data. The data validation section presents the results of data checks to ensure generally uniform coverage across the region, some general review of the data and comparison of the available datasets, and comparison of the data to traffic counts on the highway network in the region prior to expansion. The data expansion section documents the expansion of the passively collected big data to traffic counts to correct for systematic biases and ensure the expanded data is representative of all travel in the region. Finally, the data analysis and visualization section presents the results of analysis of the data including information on the amount travel between the various communities comprising the region and on auto and truck travel patterns for key corridors in the region.

2.0 DATA SELECTION

Passively collected big data on trip origins and destinations presents a valuable and powerful new source of data for travel modeling and forecasting. Passive origin-destination (OD) data include information from observations of millions of individual trips that can be harnessed for travel modeling and forecasting and simply understanding travel patterns in a region like Northeast Indiana. Moreover, passive data collection can provide OD data more cost effectively than traditional household travel surveys. However, the new data is not without its limitations, one key limitation being that passive OD data is typically aggregated (to protect privacy concerns and for data manageability) and anonymous (not including any traveler characteristics). Another important limitation is that passively collected data does not constitute a random sample and is not generally representative unless it is carefully expanded to correct for systematic biases.

While it will never be a full replacement for survey data, because passive data is by its nature anonymous and thus lacking in travelers' characteristics and purposes which are important for many types of forecasting (such as mode choice), passive OD data compliments traditional survey data extremely well, providing types of information that surveys cannot, or cannot without great cost. In particular, passive OD data can provide information on trucks and visitors, both of which are very costly to survey. Passive OD data can also collect much larger samples, which are important for less frequent phenomenon like longer distance trips, and for providing a detailed understanding of the OD patterns of simple daily resident trips. While surveys capture many important details of daily resident trips, particularly regarding purpose and mode, no cost-constrained survey can provide a picture of the OD trip matrix itself at the level of zones or even moderately disaggregate districts. Traditional surveys typically contain observations for 3% or less of the cells in the OD matrix. In contrast, passive OD data typically provides observations for a quarter to a third of the cells in the OD matrix. This data enables alternative data driven model frameworks which can produce more accurate results and a better ability to understand travel patterns in general.

Passively collected big data is a rapidly evolving subject area. As recently as 10 years ago, as of 2017, there were no commercial sources of passively collected OD data. However, over the past several years, many organizations and companies have begun to offer their data for use in transportation planning and analysis. As of 2017, there are four technologies (or types) of passive OD data in use (each of these types of data are described and discussed in some detail below):

1. Cellular Tower Signaling
2. LBS (Location Based Services)
3. GPS (Global Positioning Systems)
4. Bluetooth

Each technology requires its own equipment and has its own limitations. For instance, cellular phone tower triangulation has limited resolution based on the spacing of towers and relies on communications between devices and towers that is not optimized for transportation data needs. GPS devices can provide accurate locational data, but sometimes ID persistence is an issue that can limit data processing techniques. Bluetooth transceivers are required to detect Bluetooth-enabled devices and must be deployed on site to collect the data for a limited number of locations. Despite these limitations, these new technologies provide information on millions of trips to support a robust understanding of regional travel patterns.

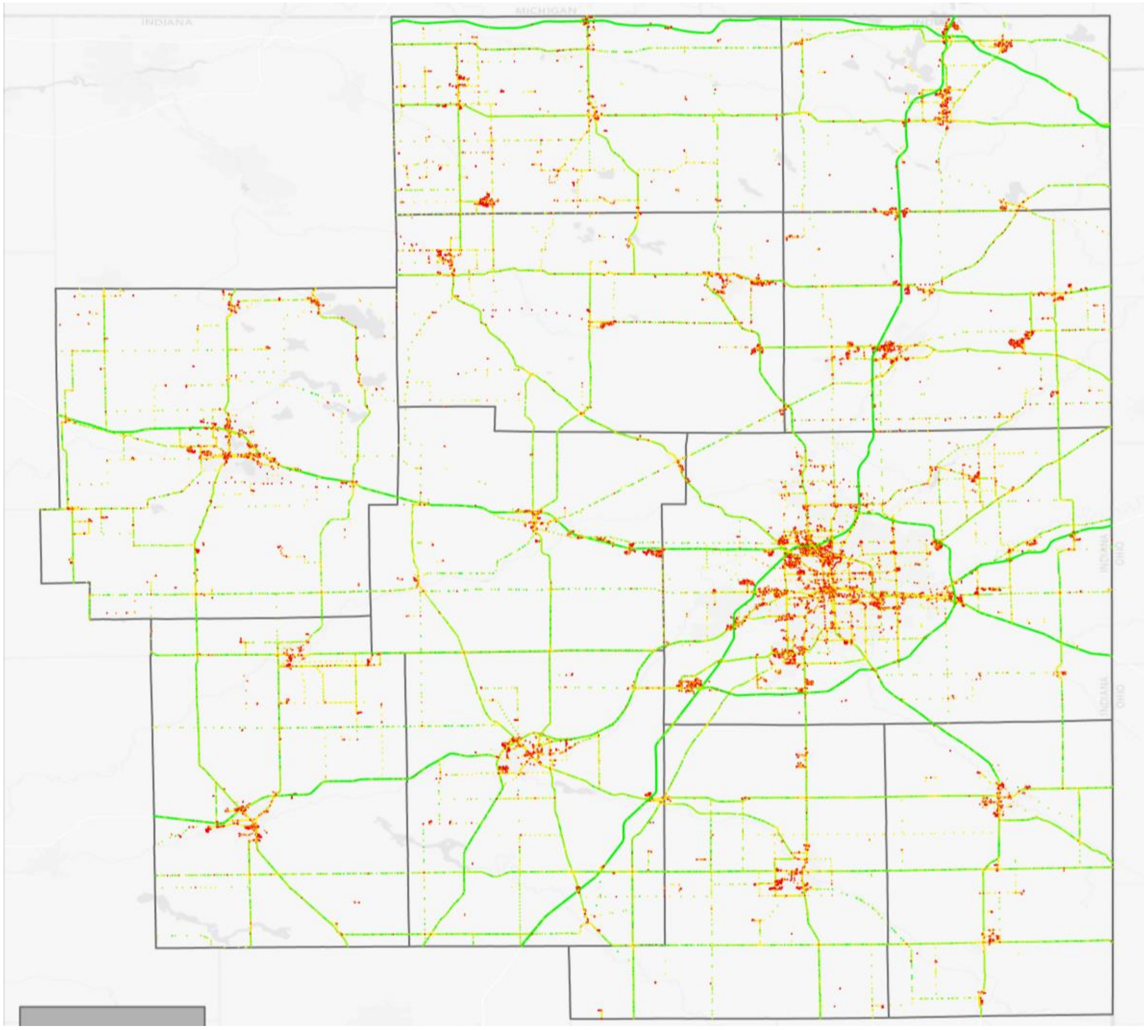
Although RSG conducts Bluetooth surveys, Bluetooth was not considered a candidate technology for this study for several reasons, most notably because it cannot provide information on trip origins and destinations, only on where trips pass, and because it cannot provide information separately for cars and trucks. Bluetooth is best for corridor OD studies and external cordon-line studies, but is not generally well-suited to multi-purpose regional analyses such as this.

NIRCC contacted and requested quotes for passive OD data from three data providers: AirSage, ATRI, and Streetlight Data. RSG assisted NIRCC in negotiating discounted pricing from AirSage and Streetlight, in part, contingent on finalizing the data purchase by the end of 2016. AirSage currently provides data only on total traffic based on cell tower signaling. ATRI provides GPS data only for heavy trucks. StreetLight provides two datasets, one based on LBS with total traffic and one from GPS which is broken out by cars and trucks. Since NIRCC was interested both in general travel patterns between communities in the region as well as truck-specific patterns associated with individual facilities, there were essentially two purchase options which could provide this information: NIRCC could purchase both AirSage and ATRI or just Streetlight.

NIRCC, with assistance and advice from RSG, also requested information from each data vendor to help determine their sample penetration. While all big data providers can boast observations of large numbers of trips over long time periods, the size of their samples varies (including somewhat by region, particularly for non-commercial travel) and the proportion of all trips (or all commercial trips) included in their sample is a key consideration in their value. Some providers have been more willing than others to provide sample data or penetration statistics, and often present statistics in the most favorable light and sometimes in ways that make comparisons with competitors difficult. It is therefore important to define a common metric to compare comparable data as much as possible across different vendors. Rather than focus on the portion of the population observed in the dataset, it is better to focus on the portion of trips observed on a particular corridor. This latter metric will always be lower than the former sort because many people may be observed at some point in the dataset but only a portion of their trips are observed, thus it is the portion of trips, rather than of people that gives a better indication of the amount of travel captured.

For the purposes of this study, each vendor was asked to report the average number of daily trips they observe on I-69 between US 6 and US 20. This can be compared to the counted AADT of 29,900 vpd to estimate sample penetration. This corridor was chosen because AirSage can only report facility level numbers with confidence for isolated rural corridors. The vendors were asked to provide data for a particular timeframe (the month in which the traffic count was taken), but for reasons of convenience, the vendors provided data from other timeframes and it was judged that this would not skew the results enough to have a significant impact on the evaluation.

FIGURE 1: ATRI TRUCK GPS POSITION POINTS FROM A SINGLE DAY



The two commercial vendors provided the requested average daily number of vehicles observed in the corridor. ATRI did not provide this data (in part because their key technical staff person was on vacation) but did share a screen shot of their truck GPS position points in the region from a single day (November 2, 2016). The screenshot, shown below in Figure 1, is generally consistent with expectations that ATRI would have a roughly 10% sample of all heavy trucks based on a number of studies around the country. AirSage provided data from January of 2015 which indicated that they had approximately an 8% sample in the corridor, see Figure 2. Streetlight shared their market penetration information in a webinar with NIRCC and the consultant team on December 19, 2016. Their data indicated that their LBS dataset had a sample penetration of 5.5 – 8.0%; their truck GPS data had a sample penetration of 11.5%, but their non-commercial GPS data had a sample penetration of roughly 0.4%. These results indicate that either AirSage and ATRI or Streetlight could provide adequate and generally comparable sample penetration, but that Streetlight's non-commercial GPS data may be of limited use due to its very low sample rate.

FIGURE 2: AIRSAGE SAMPLE PENETRATION

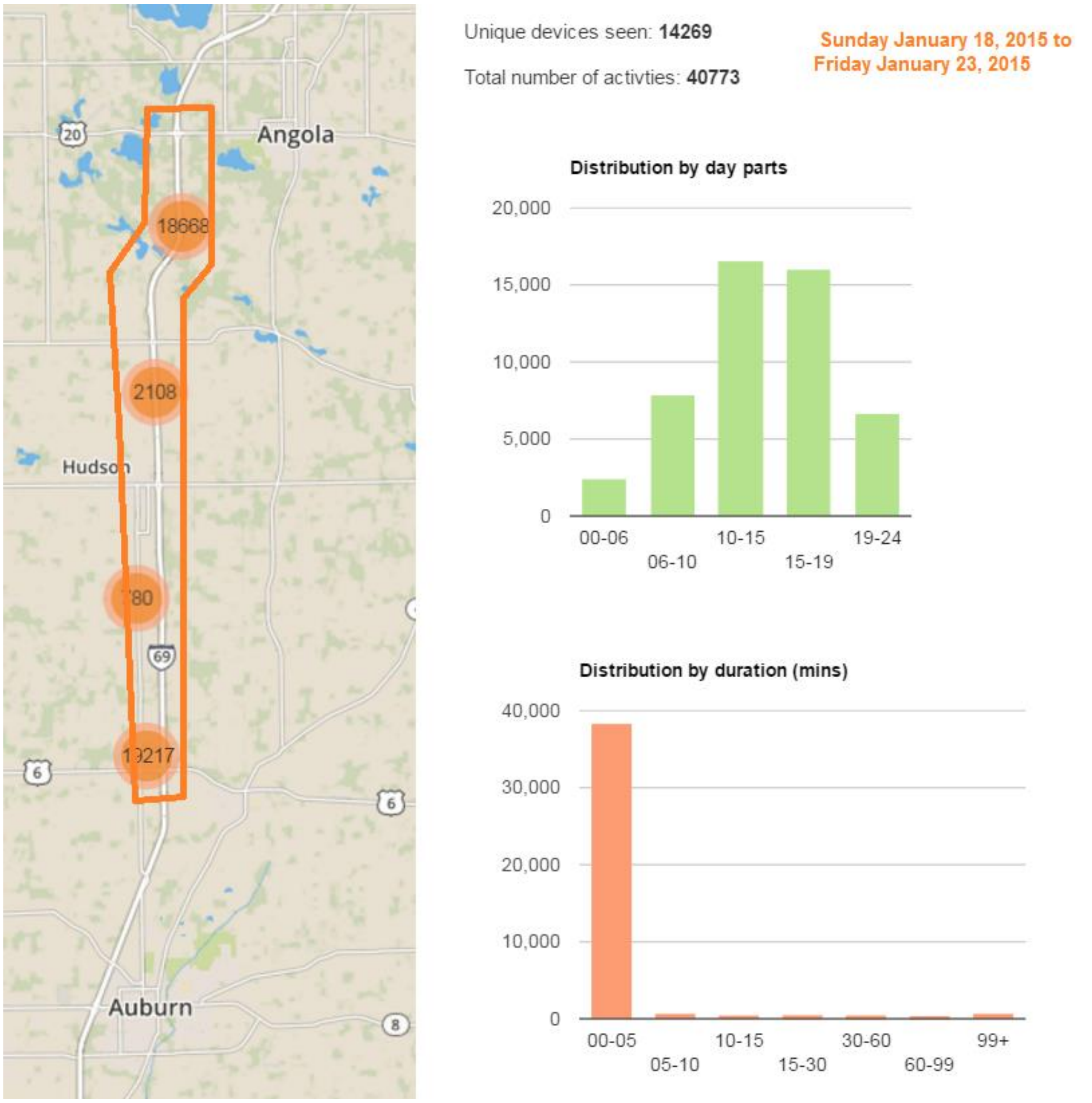


Table 1 summarizes and compares the advantages and disadvantages of the various datasets available for purchase by NIRCC in December of 2016. The main advantage of the AirSage and ATRI option was that it was less expensive than Streetlight Data. The key advantages of the Streetlight data were

- it provided up to two years of data (rather than a single month from AirSage and ATRI) which means that despite the similar sample penetration rates, it provided substantially more data
- the locational precision of Streetlight's LBS total traffic data (10-100 m) was substantially better than that of AirSage's cell-based data (~200-2000 m)
- Streetlight offered an unlimited number of zones / datasets for both its LBS and GPS data (with some limitations on very small zones that would violate privacy protections); whereas, AirSage's data was limited to single dataset of a fixed, limited number of zones (250-500)

- through its website Streetlight offered an unlimited number of direct select link analysis of truck OD patterns associated with a specific facility; whereas, a limited number of these could be produced using ATRI data and they would require substantially more effort to process
- Streetlight's truck GPS data could also provide information on medium-duty trucks, whereas ATRI could only provide data on heavy trucks

While AirSage's sample penetration was slightly higher than Streetlight's LBS data and ATRI could support the filtering of intermediate stops on long trips in a more robust way than Streetlight, these were deemed to be relatively minor issues for NIRCC's purposes. In the end, NIRCC selected and licensed Streetlight data despite the fact that it was more expensive because it was believed to offer better value overall.

TABLE 1: COMPARISON OF AVAILABLE BIG DATA FOR NIRCC

	AirSage			ATRI	Streetlight	
Price	**$12,307**	**$19,391**	$24,437	**$12,000**	**$41,836**	
Number of Zones	250	500		Unlimited	**Unlimited**	
Location Technology	Cell Tower Signaling			GPS	GPS	LBS
Locational Accuracy	~ 200 - 2000 meters			1 - 10 meters	1 - 10 meters	**10-100 meters**
Universe	All Travel			Heavy Trucks	Personal, Med. & Hvy. Truck	All Travel
Sample Penetration	8%			10%?	~0.4% personal, 11.5% truck	5.5-8.0%
Sampling Issues	Cell tower signaling somewhat dependent on phone use					
Coverage Issues	poor coverage in some rural areas					?
Select Link (Zone) Analysis	**Indirect only**			**Limited direct, indirect**	**Unlimited direct**	**Indirect only**
Filtering of Intermediate Stops on Long Trips	No (premium option)			Yes	No (manual removal of truck stops)	
Data Collection Time Period	1 month			1 month	**2 years**	**Multiple months**
Time Periods	Average Weekday and Average Weekend		Time-of-Day	User defined, requires more processing	Average Weekday or Average Day	
User Classes / Residency	None	Regional Residents, Non-Residents		None	**Heavy trucks / Medium Trucks**	**none/premium (not included)**
Purpose	No purpose imputation included			N/A	N/A	premium (not included)
Expansion	Residence market share based; must be adjusted			None; must be added	Single count-based factor; must be adjusted	
Non-disclosure	N/A			Restricted Use of Disaggregate Data	N/A	

3.0 DATA VALIDATION

Streetlight Data provides several datasets which can be analyzed separately or together. There are two primary data providers: INRIX and Cuebiq.

- Streetlight Data
  - INRIX (GSP)
    - Automobiles / Light Trucks (pickups/SUVs)
    - Medium Trucks
    - Heavy Trucks
  - Cuebiq (LBS)
    - All vehicles

INRIX provides navigational GPS data gathered passively primarily form in-vehicle navigational devices and secondarily from mobile device navigation applications. The INRIX GPS data is broken out by vehicle class into autos, medium trucks, and heavy trucks. Cuebiq provides location based services (LBS) data which constitutes the best locational information available to a mobile device at a given point in time. Thus, LBS data is a mix of underlying technologies including GPS, Bluetooth and WiFi beacons, and cell tower trilateration. Cuebiq's LBS data is currently only available for all vehicles.

The validation of NIRCC's Streetlight data comprised primarily of two efforts. First, the data was analyzed to identify any coverage issues. Second, the data was scaled to and compared with total traffic counts in the region. In this second effort, some comparisons were also made between the LBS and GPS datasets.

3.1 | SCREENING FOR POOR COVERAGE

Quick response trip generation methods were used with basic, freely available socioeconomic data to produce rough estimates of trip ends for each district and these were compared to the marginal sums of the OD data matrices to identify any likely "holes" where data coverage was missing or significantly limited.

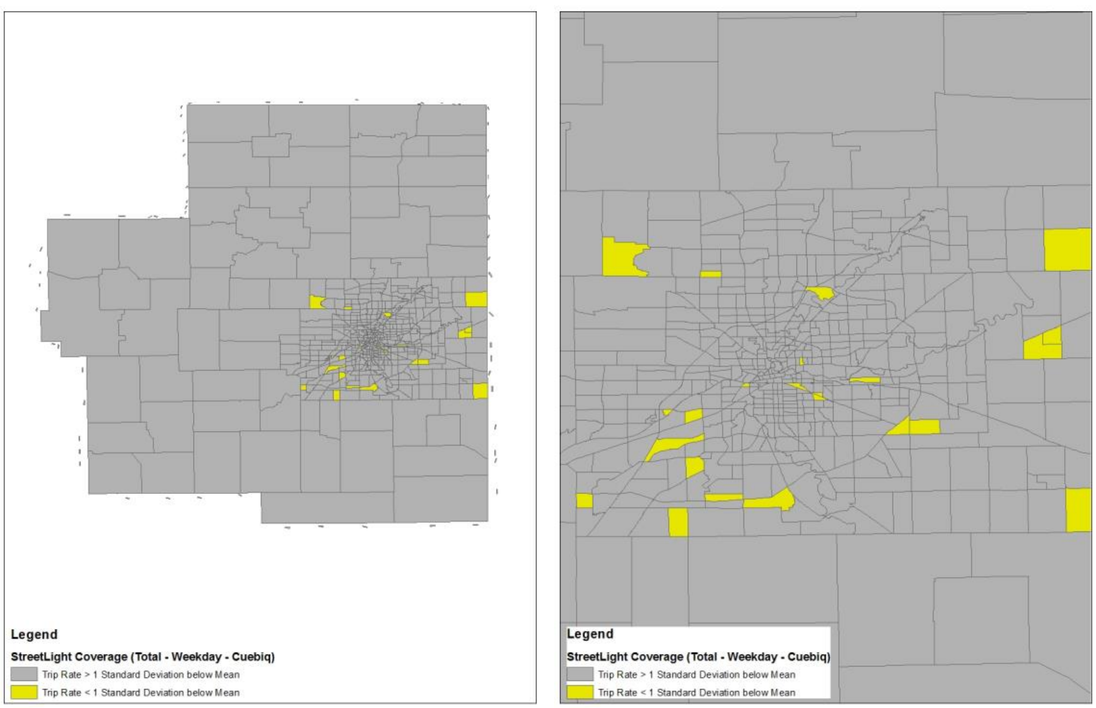
It should be mentioned that StreetLight reports trip index instead of trips; however, the trip indexes can be considered as scaled trips due to specific processes conducted by StreetLight. Rescaling to bring trips to their real values is a key element of this study. The StreetLight trip index, therefore, will be called StreetLight trips for the rest of this document.

The first screening for coverage problems was to examine the internal consistency of the passive data relative to the freely available socioeconomic data from the Census. While some variability in trip rates (person or employee) is real and to be expected, that variability has some reasonable bounds, so that by examining the data it is possible to identify cases where it is likely that data is missing.

A metric (trip rate) was defined as the ratio of total trip ends (both origins and destinations) to a socioeconomic attribute of the zone. In the first step, the marginal sums of the OD trip tables were calculated. The total trips of each zone were then divided by a zone socioeconomic characteristic to get a uniform measure for comparison. The zone characteristic chosen for each mode (auto, truck) should reflect the trips made by that mode. As a result, total employment was used as the socioeconomic attribute for truck trips while a combined measure was calculated for auto trips. This simple combined measure is equal to summation of number of households, retail employment, and total employment, to account for the fact that retail figures into trip attractions more prominently.

The mean and standard deviation of coverage measures were then separately calculated for auto and truck trips (for INRIX) or all trips (for Cuebiq) over all zones. Statistically, the outliers are defined as the zones with the coverage lower than mean coverage minus coverage standard deviation times 2. Fortunately, none of the zones in either dataset shows low coverage with this threshold; however, this observation does not mean zones are necessarily similar coverage-wise. As a result, a new threshold for each mode was defined as the mean coverage minus coverage standard deviation to recognize zones with lower trip coverage compared to others. This much stricter criteria does identify a small number of zones that may have lower coverage. All of these zones were relatively small and primarily rural zones in Allen County. Figure 3 illustrates the potentially low coverage zones in the Cuebiq dataset. The INRIX data was similar. More details documenting the analysis can be found in Appendix C.

FIGURE 3: CUEBIQ TRIP COVERAGE IN THE STUDY AREA AND ALLEN COUNTY



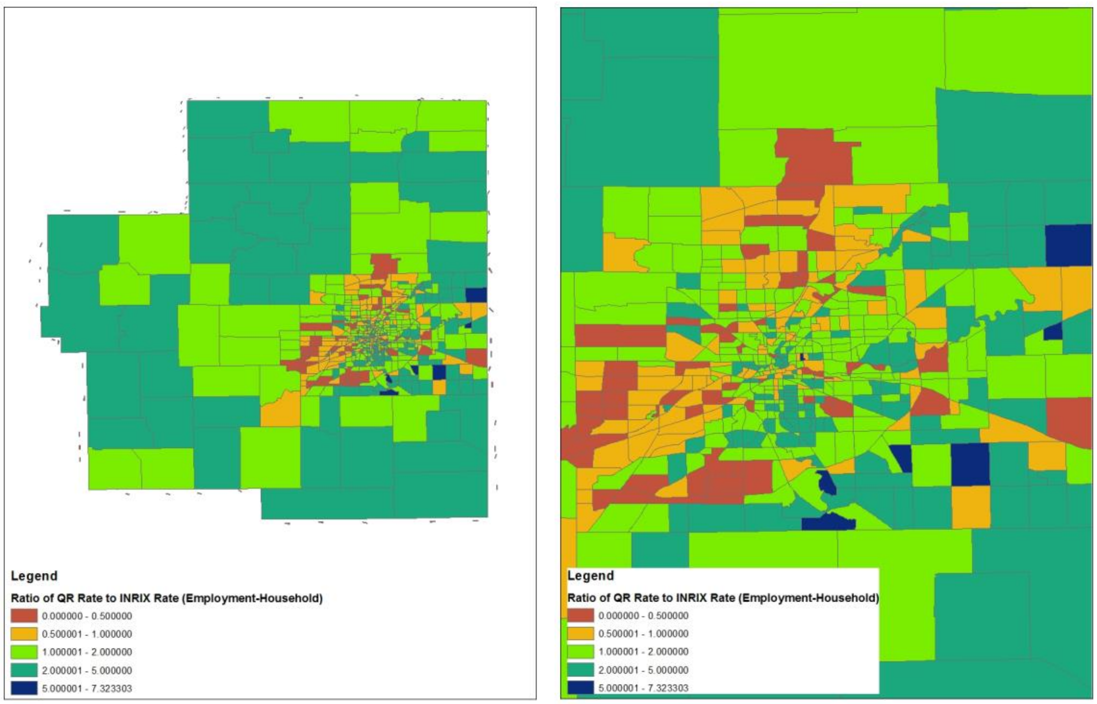
Although both datasets coverage is statistically acceptable based on this analysis, trip rates were also compared to trip rates estimated by NCHRP Quick Response (QR) Method. Trip production and attraction in NCHRP QR method are calculated as follows:

*Trip Production = 6.5 * Number of Households (based on NCHRP 716)*

*Trip Attraction = 17.2 * Retailers + 4.1 * Service Employment + 1.5 * The Remaining Employment + 1.9 * Number of Households + 1.45 * Total Employment (based on NCHRP 365).*

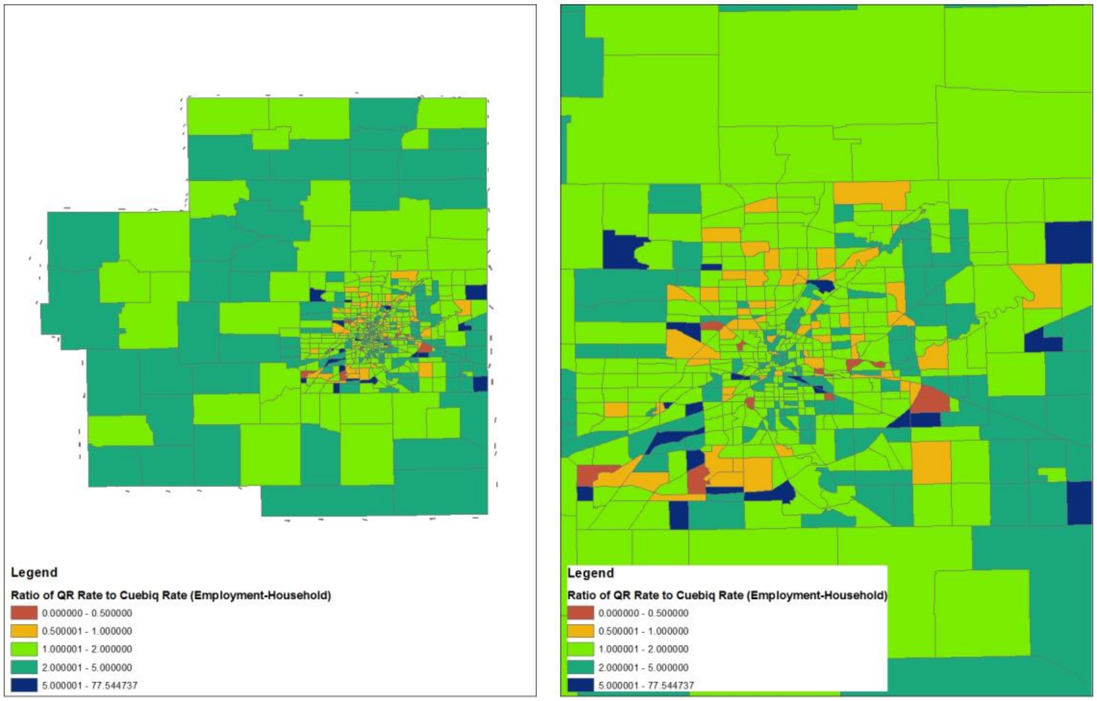
The summation of trip productions and attractions is then divided by the combined socioeconomic attribute of the zone (summation of number of households, retail employment, and total employment which is called "Employment" in the following maps) to get the QR trip rate which will be compared to Streetlight trip rates in the maps below and in Appendix C.

FIGURE 4: RATIO OF QR TO INRIX TRIP RATE IN THE STUDY AREA AND ALLEN COUNTY



The trip rates implied by the both passive data sets are a little bit low in rural areas relative to the QR rates. This could be due to poorer coverage / penetrations of the relevant technologies in rural areas but is equally likely due to the fact that the QR rates were developed for urban areas and rural areas are known to have lower trip rates. In any event, if there is deficient coverage in the rural areas, it is clearly of a magnitude that can be addressed and corrected for by the data expansion.

FIGURE 5: RATIO OF QR TO CUEBIQ TRIP RATE IN THE STUDY AREA AND ALLEN COUNTY



3.2 | SCALING AND TEST COMPARISON WITH TRAFFIC COUNT DATA

Traffic count data provide the best, unbiased information on the total amount of truck and passenger vehicle traffic on the road. Ultimately, the OD data was expanded on the basis of traffic count data as described in the following section. However, prior to this, as part of the data validation, the data was scaled to the overall level of traffic reflected in counts across the region and a preliminary comparison was made between traffic counts and the result of assigning the OD data to the roadway network. This initial comparison was included as part of the data validation task because it is not uncommon for the preliminary comparison with count data to reveal issues with the data which sometimes require coordination with the vendor to address or some processing to clean/correct.

RSG developed a regional network by combining the NIRCC's model network with the highway network from the Indiana Statewide Travel Demand Model outside the coverage area of NIRCC's model and creating centroid connectors to connect this regional network to the data OD districts.

In addition, RSG analyzed 2015 counts received from different sources such as county and state and tagged the links in the network with the corresponding counts. Figure 6 presenting the links with the 2015 AADT in the network indicates that the network has a very broad count coverage especially inside the Allen County. Table 2 also summarizes the counts by vehicle class and shows how many links out of 7,189 links have counts for each vehicle class.

FIGURE 6: LINKS WITH THE 2015 COUNTS IN THE NETWORK

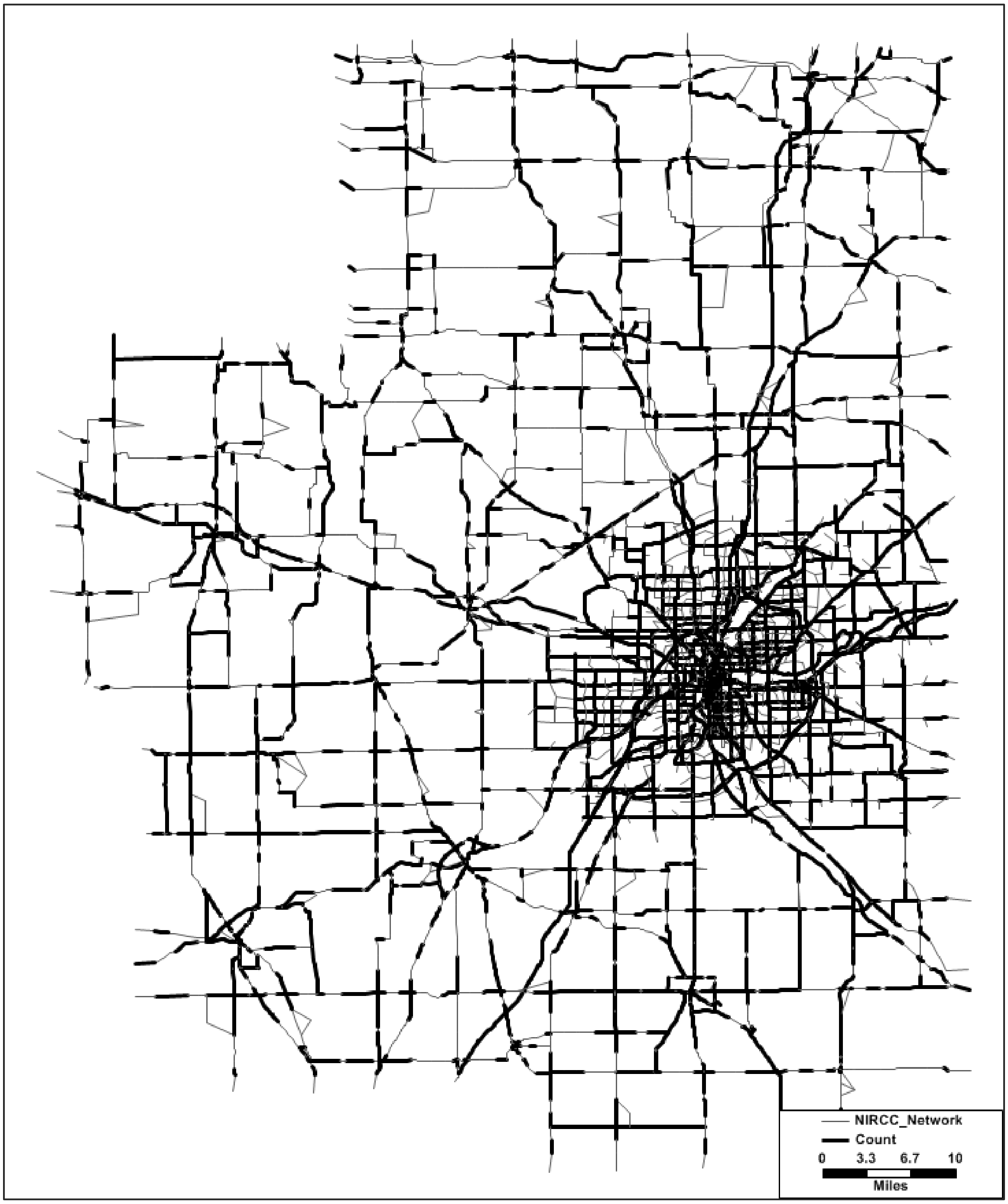


TABLE 2: NUMBER OF COUNT STATIONS BY VEHICLE CLASS

Vehicle Class	Total ADT	Auto	Truck	SUT	MUT
Number of Count Stations	3,719	2,452	2,452	1,346	1,346

As can be seen in Table 2, there are more links for total traffic than for specific classes. Therefore, validation and expansion of the data to total counts provides more confidence than results for specific vehicle classes. As mentioned in Section 3.1, INRIX data includes auto, medium, and heavy truck trip tables separately; however, CUEBIQ data includes one trip table containing all modes. Table 3 compares the original INRIX and CUEBIQ data before any scaling or expansion.

TABLE 3: ORIGINAL CUEBIQ AND INRIX TRIP TABLES

Trip	Total Trips	External-External	Intrazonal	Non-EE, Interzonal
CUEBIQ	24,608,896	19,388,997	4,629,565	590,334
INRIX – Total	7,027,388	2,675,002	3,879,095	473,291
– Personal	1,241,733	170,747	920,623	150,363
– Medium Trucks	1,599,207	371,987	1,024,831	202,389
– Heavy Trucks	4,186,448	2,132,268	1,933,641	120,539

As seen in Table 3, the CUEBIQ trip table captures many trips between external catchment zones that never enter the study area. These external-external (E-E) trips inflates the total number of trips, but once these are accounted for, CUEBIQ and INRIX report roughly similar numbers of total trips, within about 20% of each other. The INRIX data does not have the issue with trips outside the area since vehicles can be captured as they enter/exit the study area at external stations. Accounting for intrazonal trips, which can be difficult to measure accurately, further improves the agreement between the remaining trips. However, an examination of the breakout of trips by vehicle type in the INRIX data reveals that truck trips are significantly over-represented relative to personal vehicles.

Appendix C documents the details of the results of the initial assignments and scaling steps. The best preliminary results were produced using a combination of INRIX and Cuebiq trips and are presented below.

FIGURE 7: LOADING ERROR OF THE HYBRID TRIP TABLE

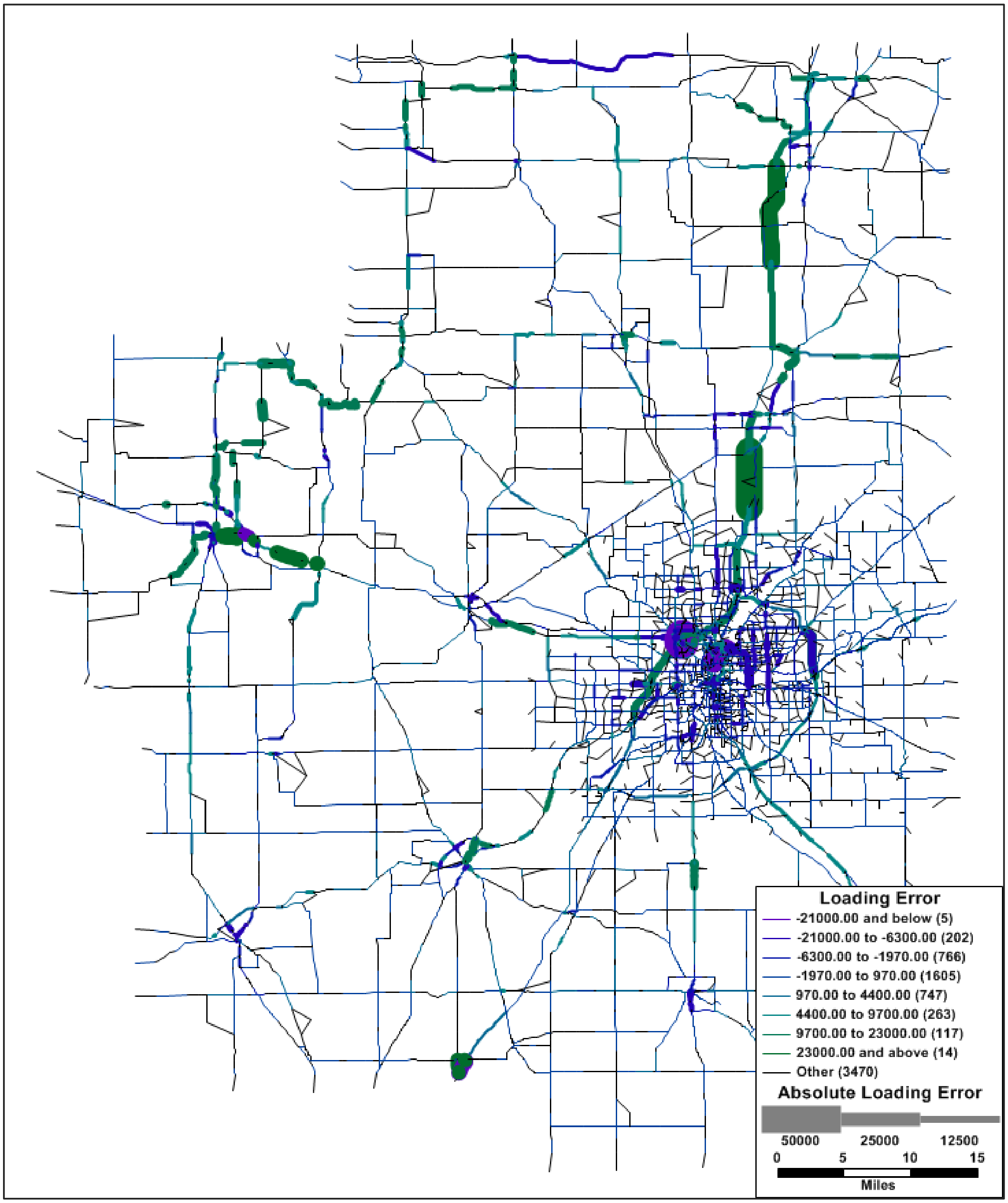


TABLE 4: ASSIGNMENT STATISTICS FOR HYBRID TRIP TABLE

Assignment Statistic	Average Error	Loading Error (%)	RMSE (%)	MAPE
All Facilities	1.78	0.03	71.91	155.08
Freeways	7,028.98	36.73	63.10	257.13
Arterials	-593.48	-5.78	51.25	154.04
Collectors	154.18	5.31	124.57	151.10
Locals	332.96	26.39	117.83	196.49
Urban Links	-1,135.10	-11.73	54.14	73.45
Rural Links	1,156.19	31.52	115.79	248.26

The project team carefully analyzed the assignment statistics of INRIX, CUEBIQ, and HYBRID trips after overall scaling to compare them with each other. The main results of this analysis can be summarized as follows:

1- Although the overall loading error is about zero, the global RMSE is high and almost in the same range in all three datasets.

2- Among INRIX trips, auto performs much better than trucks. In addition, medium trucks need more adjustment than heavy trucks. This issue might be due to errors in count stations; however, they are the only available observations.

3- The same pattern is observed in all assignment results which is underloading in urban areas and overloading in rural areas. The RMSE in rural areas is very high too, which confirms StreetLight has major issues in the detection of trips in suburban and rural areas.

4- The hybrid trip table has the best RMSE in rural areas while CUEBIQ trips are the best along freeways.

3.3 | COMPARISON WITH HOUSEHOLD SURVEY DATA

To investigate the underloading/overloading issue in urban/rural areas, the trip duration frequency distribution of INRIX and Cuebiq were also compared to the local survey data used to develop the NIRCC model. Figure 8 and Figure 9 present number of trips by trip duration. Overall, the expanded survey shows less trips than the other two sources. According to Figure 9, the survey shows a higher percentage of short trips (shorter than 15 minutes) while StreetLight trips have higher shares in trips between 20 to 55 minutes. This observation confirms the assignment statistics showing underloading in urban areas (shorter trips) and overloading in rural areas (longer trips) and tends to confirm the suspected trip length bias common in passive OD data. It is important to note that the survey did not cover all the rural areas covered by the passive data, and this may partly explain the difference, but likely not all of it.

FIGURE 8: TRIPS BY DURATION IN THE SURVEY, INRIX, AND CUEBIQ

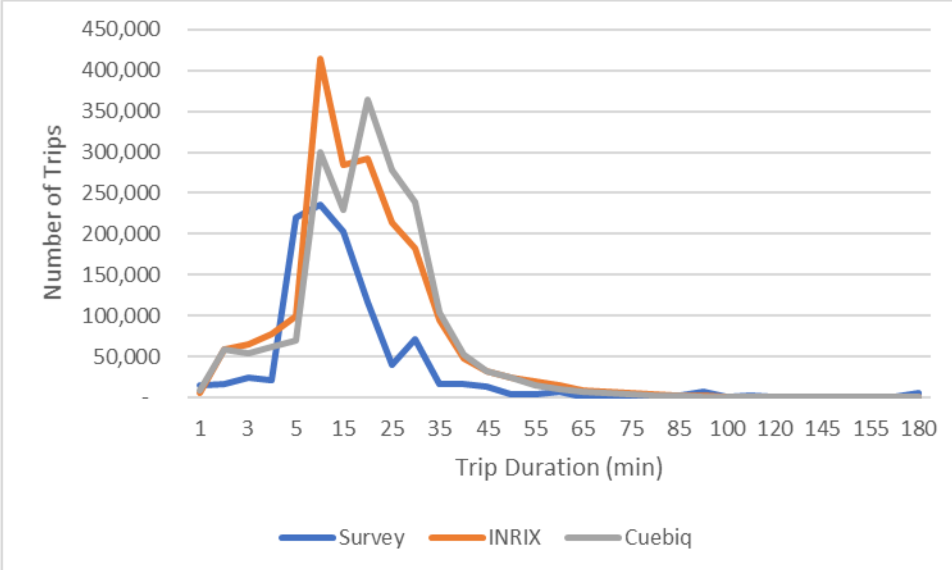
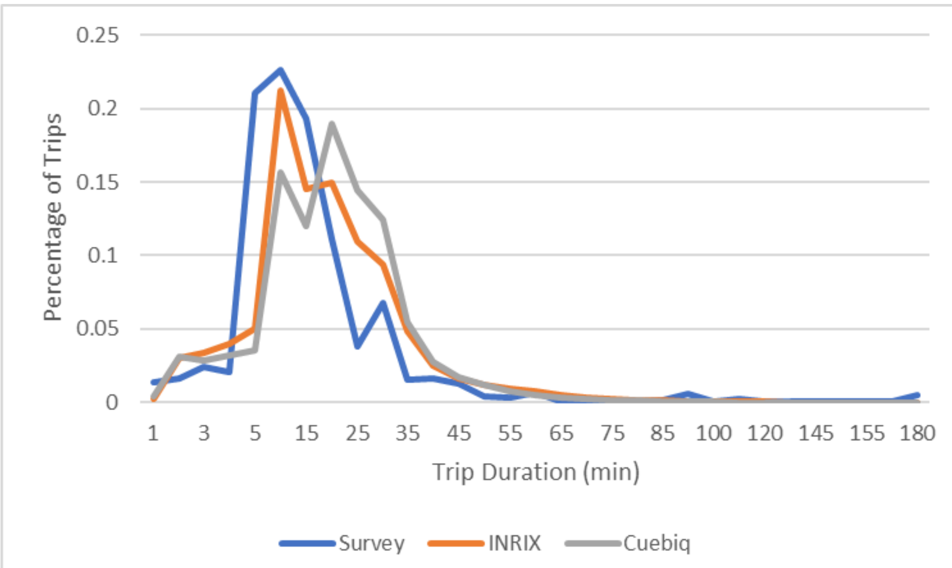


FIGURE 9: TRIP DURATION FREQUENCY DISTRIBUTION IN THE SURVEY, INRIX, AND CUEBIQ



In summary, data validation efforts indicate that both the INRIX and Cuebiq data sources represent biased travel patterns, skewed towards overrepresentation of longer trips. However, there are no critical areas of missing data or coverage holes in the data, so it was possible to expand the data to more accurately reflect travel in the region as described in the next section.

4.0 DATA EXPANSION

All existing commercially available passively collected OD data are based on incomplete sample frames. These commercially available datasets exclude travelers without mobile devices while they travel, and these datasets include only a select portion of travelers with mobile devices. Moreover, short-distance trips or short-duration activities are often under-represented in the data because they require more frequent observations of position. Travel to and from locations with poor coverage can also go un- or under-detected. Failure to account for such biases can lead to erroneous representations and faulty predictions of trip lengths, trip flows between origins and destinations, and present and future travel activity and traffic in general.

Traffic counts provide unbiased information on the spatial distribution of traffic. Traffic counts are currently the only data available to support expansion methods for passive OD data capable of correcting systematic biases related to coverage (rather than market penetration) and trip length or activity duration. The following section provides more information and illustrations of various methods by which traffic counts can be used to expand OD data.

Some analysts or modelers are reluctant to "mix" supply and demand data in this way. This reluctance may be rooted in the idea that traffic counts should provide independent validation of demand estimates developed solely from other sources. The development of travel demand models is sometimes presented in this way, but this is extremely misleading. In actual practice, demand estimates, whether based on "pure" synthetic models or directly observed data, are always adjusted to reconcile with or "validate to" traffic counts. The acknowledgement of this and the use of traffic counts in a well-defined process of expanding or adjusting demand estimates should be preferable to their use to adjust demand estimates in a series of ad hoc and often poorly documented manual adjustments.

4.1 | CANDIDATE METHODS FOR EXPANDING PASSIVE DATA

Multiple methods exist for expanding passively collected OD data. A taxonomy of expansion methods, focused on methods used in practice and particularly on methods using traffic counts, is presented below. In all, seven methods are presented, divided into various categories. At the highest level, these methods can be divided into two categories depending on whether they use traffic counts or whether they rely on another estimate of sample penetration. Two methods of expansion that do not use traffic counts are discussed because they are in common use. The remaining five methods use traffic counts in various ways to expand passive OD data. These can be divided first based on whether they use only a single or multiple expansion factors. The latter can then be divided based on whether they make use of a network assignment model. Those that do make use of a network assignment model can be further divided into parametric methods that do not use ODME algorithms and nonparametric methods that do. Finally, ODME-based methods can be divided into those that rely directly on an ODME algorithm and those that use ODME to develop a simplified set of expansion factors.

FIGURE 10: TAXONOMY OF PRACTICAL METHODS OF PASSIVE OD DATA EXPANSION

- Other Sample Penetration Methods
    - **Market Penetration**
    - **Trip Generation-Based**
- Traffic Count Methods
    - **Simple Scaling**
    - Variable Scaling
        - Matrix Partitioning
            - **Iterative Screenline Fitting**
        - Network Assignment-Based
            - **Parametric Scaling**
            - Nonparametric (ODME)
                - **Direct ODME**
                - **Indirect ODME**

The seven methods presented here have been included because they are known to be in use in practice, although level of use varies considerably. Although sometimes a single expansion method (most commonly simple scaling to traffic counts) is applied naively, but most well-conceived expansion efforts use several of these methods in combination.

The following discussion highlights both the advantages and limitations of various methods for NIRCC's analysis, recognizing that different methods or combination of methods may be appropriate for different applications.

MARKET PENETRATION

Market Penetration methods require information on the portion of the population included in the sample. To be meaningful and effective, this information must be available at a relatively fine level of geographic resolution. AirSage, for example, uses data on cell carrier market penetration at the block group level to develop expansion factors for its data. Although valuable in correcting for demographic biases (such as those related to income, which have been demonstrated, for instance, in GPS data for the Detroit metropolitan area), they do not correct for biases related to trip detection such as trip length bias or coverage biases. In any event, market penetration data is not available for INRIX or Cuebiq data at this time, so this expansion method is not a viable option for NIRCC's datasets.

TRIP-GENERATION-BASED

Trip Generation-Based methods develop expansion factors for the passive data based on zonal level comparisons of observed trips in the passive data with estimated trips from trip generation models such as the comparisons presented in the previous section as part of data validation. While using these comparisons to generate expansion factors may help address coverage issues, it could also introduce biases from the trip generation model used, and there is some reason to believe (such as the correlation of the INRIX and CUEBIQ trip rates) that at least some of the trip rate variation observed may be real. Moreover, while this method can help address coverage issues, it cannot address systematic biases related to frequency of observations and trip length. Thus, while comparisons with trip rates can be helpful and were used in exploring and validating the NIRCC datasets, they are not recommended as an expansion method.

SIMPLE SCALING TO COUNTS

Simple Scaling was applied as a first step in expansion of the NIRCC data as presented in the previous section. Independent scaling of the external and internal trips to traffic counts yielded substantial improvement in the validity of the data in comparison to traffic counts, but also was unable to address trip length bias and achieve satisfactory results. Thus, this preliminary scaling step is recommended as the first step in the expansion of the NIRCC datasets, but additional steps are also required.

ITERATIVE SCREENLINE FITTING

Iterative Screenline Fitting or **Matrix Partitioning** is unique in that makes use of traffic counts to produce a number of expansion factors which may be able to correct for systematic biases but without using a network assignment model. Avoiding the use of a network assignment model is an advantage since the use of any model can introduce error. Moreover, this approach typically can only make use of a subset of traffic counts in a region resulting in a holdout sample of counts which can still be used to provide independent validation of the passive OD Data.

The approach works by first identifying "screenlines" or "cutlines" such as are commonly used to validate travel models. Each screenline should partition the study region into two subareas and align with the zone system used to define ODs, and traffic counts should be available or taken everywhere the roadway network crosses the screenline. (For this reason it is helpful to choose screenlines which follow natural / physical barriers such as rivers, freeways, and railroads which have limited roadway crossings.) The sum of the traffic counts along each screenline can then be compared to the number of trips in the OD matrix which cross the screenline. This can comparison can be made without a network assignment model by partitioning or aggregating the OD matrix. Since each screenline partitions the region into two subareas, A and B, all origins and destinations can be identified as falling in either A or B. The OD matrix can then be aggregated into a matrix of just four cells: trips from A to A, trips from A to B, trips from B to B, and trips from B to A. The two off-diagonal cells (trips from A to B and from B to A) cross the screenline while the others do not. In this way, groups of OD trips can be compared against screenline counts without a network assignment model, and a preliminary expansion factor developed as the sum of the screenline counts divided by the sum of the off-diagonal elements of the aggregated matrix. The iterative screenline fitting process works by iterating or looping over the screenlines, factoring trips crossing each screenline to match the screenline counts. Although this factoring guarantees that the OD trips match the sum of counts for the current screenline, each factor has the potential to introduce disagreement between the OD trips and previous screenlines since individual OD pairs may cross several screenlines and thus have several differing factors applied. For this reason, the iteration is needed so that the expansion factors for individual OD pairs can stabilize to values that minimize errors versus all the screenline counts (but do not guarantee perfect agreement of the OD data with any individual screenline count).

This approach is not believed to be widespread but has seen application in several areas including Anchorage, Chattanooga, and San Diego. The relative value of the approach compared to simple scaling to total counts and its ability to address systematic biases in the passive data is largely a function of the number of screenlines which can be constructed for use in the procedure. A moderately large number of screenlines may be required in order to fully correct for trip length related biases as ODME-based methods can and it may be difficult to construct a large number of screenlines in some areas.

The NIRCC study area district/zone system and regional network with traffic counts was examined to determine how many screenlines could be constructed and how well they could isolate demand of interest. Count coverage in the region was good and allowed the definition of 82 screenlines, presented in Figure 11. Figure 12 illustrates the screenlines in Allen County in more detail. Of the total 82 screenlines, 66 are polygons surrounding a part of the region and defining the total demand entering and exiting it. The remaining 16 screenlines cut the entire region into two distinct subareas and define the demand crossing between them. Screenlines were built with respect to several facts including count station locations, zone borders, centroid connector locations, and natural\physical barriers such as rivers, freeways, and waterways. For instance, one major rule is that all centroid connectors of each zone must fall in the same partition created by screenlines. A total of 520 or just under 14% of the 3,719 links with AADT cross the screenlines, leaving a robust hold-out sample of 3,199 counts which would not be used by this method and could provide independent validation. Roughly half (56%) of the links used by the screenlines are in Allen County, which is reflective of the relative level of network and district detail – 2,191 or 59% of all counts are in Allen County.

Given the ability to form a very robust set of screenlines while still allowing for a large holdout sample, the iterative screenline fitting method is recommended as the second and hopefully primary method for expanding the passive data for the NIRCC region.

PARAMETRIC SCALING

Parametric Scaling is perhaps the most straightforward way of addressing and correcting for systematic trip length biases in passive OD data. In theory, this method may be able to be applied to GPS-based datasets without the use of a network assignment model, but this would require substantial data processing and in practice, to date, it is only known to have been applied through the use of a network assignment model. The approach is to estimate the parameters of a formula that produces expansion factors for trips as a function of their length or other attributes. The parameters are estimated using least squares error (LSE) versus traffic counts. The parameter estimation can be formulated as a bilevel programming problem but is particularly difficult (NP-Hard) as it involves an equilibrium constraint in the lower level traffic assignment problem. Hence, non-linear optimization metaheuristics such as genetic algorithms, etc., are typically used to solve for the parameters. (Although simpler line search methods can be used if the expansion factor is modeled as a simple linear function of trip length, this approach is not recommended as evidence points to a nonlinear relationship as well as the significance of other factors.)

This approach has the advantage of producing relatively easily understood expansion factor formulas and avoiding the ambiguities of ODME-based approaches. Moreover, it is firmly grounded in a robust statistical procedure, and can therefore, in theory for instance, be used to determine the statistical significance of systematic biases in the data. However, the involvement of a network assignment model and resulting need to employ metaheuristics to estimate the parameters of the expansion factor function make the process both mathematically complex and computationally intensive, ultimately making it a costly approach. Since iterative screenline fitting looks like a promising option for the NIRCC region, this method is not recommended at this time.

DIRECT ODME

Direct ODME is believed to be one of the most common approaches to expanding passive data in practice and is also widely documented in the literature. It is important to recognize that there are a variety of different ODME algorithms in use and that different algorithms can produce significantly different results and have different properties. ODME methods which use OD data only as a "seed" or starting point and produce a final adjusted OD matrix purely by minimizing errors versus traffic counts are not appropriate for expanding big OD data as they can significantly distort the observed data. However, methods which attempt find a solution and produce a final OD matrix which minimizes errors versus counts and versus the original OD data or only with appropriate constraints on adjustments to the original OD data can be powerful and appropriate methods for data expansion. These methods are capable of correcting systematic biases related to trip lengths as well as coverage "holes" (provided there is at least some observations in the "holes" to expand).

A proper understanding of ODME is grounded in two important facts. First, counts do provide real information about underlying OD patterns, and second, counts alone cannot be used to identify OD patterns. Both of these facts can be proved mathematically. The truth of the former is demonstrated, for instance, in the method of iterative screenline fitting. The truth of the latter is evident from the fact that the number of "known" traffic counts is always substantially smaller than the number of "unknown" OD flows so the problem is statistically under-determined and there is not a unique set of OD flows that correspond to a set of traffic counts on a network.

On the one hand, from the first fact that counts do provide information about the underlying OD patterns, it is clear that ODME has real potential to improve or correct OD matrices from big data. From the second fact, that counts alone cannot identify OD patterns, it should be clear that ODME methods focused solely on count data are ill-conceived. A balanced ODME approach recognizes the value of both traffic count data and Big OD Data and uses traffic counts to improve the representativeness of OD data being careful not to mangle or distort it. In fact, it is import to understand that mathematically, because the OD solution space dwarfs the network solution space, the OD data is more important than the count data in producing a good final solution. So long as an ODME method is used consistent with this fact, it can be an efficient and powerful tool for expanding Big OD Data.

Direct ODME has several practical advantages as a method to expand passive OD data. Since software implementations of ODME algorithms are widely available, direct ODME is one of the quickest and easiest ways of expanding OD data to traffic counts and correcting for systematic trip length biases. In fact, ODME can correct for a variety of different types of errors or biases in the

OD data without requiring complex methods or in-depth analysis. However, this is a two-edged sword the flip side of which is that ODME can over-correct and distort OD patterns to over-fit count data if not used carefully and appropriately. This danger and the distrust that it inspires in some professionals is the main drawback of the method together with its lack of transparency and the difficulty of understanding the underlying issues which the expansion adjustments are addressing.

Using ODME in combination with and secondary to other expansion methods such as iterative screenline fitting or parametric scaling can allow the imposition of tighter constraints on the ODME adjustments and greater confidence in the expansion while also allowing a tighter fit to traffic counts. If NIRCC desires a tighter fit to counts than is ultimately produced by iterative screenline fitting, then a final round of ODME expansion adjustments, within strict constraints, is recommended as a final, optional step in the expansion.

INDIRECT ODME

Indirect ODME involves analyzing the results of ODME to develop a simpler set of expansion factors. This approach can actually coincide with parametric scaling, but can also involve the development of non-parametric schemes of expansion factors based on trip length, districts, etc. A more limited set of expansion factors can be more readily understood and interpreted than a multitude of direct ODME-based expansion factors, and in this way inspire greater confidence in some cases. Moreover, this approach can help establish the amount of the ODME adjustments related to particular phenomenon such as trip length and confirm that these adjustments, such as changes in average trip lengths in themselves result in better agreement between the OD data and traffic counts independent of the details of the ODME adjustments.

The main advantage of this approach is its relatively higher level of transparency and interpretability of final results compared to ODME, its support of insights from ODME, and its modest level of effort and leveraging of widely available ODME algorithms. The level of effort associated with the approach can vary depending on the complexity of the expansion factors developed. Basic schemes to address trip length bias can be applied with only marginal additional effort compared to direct ODME, while complex schemes using multiple factors can require substantial effort. The additional increment of effort beyond direct ODME is one of the disadvantages of the approach together with the inability of the method to produce as good agreement with counts as direct ODME or correct for errors in the data that are more difficult to understand or identify.

Indirect ODME is not anticipated to offer substantial additional value for NIRCC beyond what would be achieved by iterative sceenline fitting and direct ODME, and can, in fact, be difficult to use when ODME is used as a secondary expansion method.

4.2 | EXPANSION APPROACH

RSG recommended a two or three stage approach to develop expansion factors for NIRCC's passive OD data. The expansion factors would start by building off of the simple scaling to counts performed as part of the data validation process and would continue with iterative screenline fitting. The results would be examined after iterative screenline fitting, and if a tighter fit to counts was still desired, direct ODME would be applied with careful constraints, to provide the final layer of expansion factors.

ITERATIVE SCREENLINE FITTING

Iterative Screenline Fitting (ISF) is a methodology to expand passively collected OD data to traffic counts along screenlines used as control data. This method is unique in that it makes use of traffic counts to produce expansion factors to correct for systematic biases—but without relying on a network assignment model. Although assignment is sometimes used as a convenience in the application algorithm, the results do not depend on it since whether a trip crosses a screenline is a function of its origin and destination, independent of its route. Avoiding reliance on a network assignment model is an advantage since the use of any model can introduce error. Moreover, this approach typically can only make use of a subset of traffic counts in a region (those along screenlines) resulting in a holdout sample of counts, which can still be used to provide independent validation of the expansion.

The approach works by first identifying "screenlines" or "cutlines"—which are commonly used to validate travel models. Each screenline should partition the study region into two subareas and align with the zone system used to define O-Ds. The sum of the traffic counts along each screenline can then be compared to the number of trips in the O-D matrix that cross the screenline. This comparison can be made without a network assignment model by partitioning or aggregating the O-D matrix. The ratio of the off-diagonal trips in the aggregated matrix to the counts on the screenline provides an expansion factor. The expansion factors from a number of screenlines can be combined in an iterative fashion.

The effectiveness and value of this method is a function of the number of screenlines that can be constructed for use in the method. The method may not achieve as good a fit to counts as alternative methods, but this stands to reason since it only uses the subset of the counts on screenlines. Thus, it is a natural complement to methods like ODME which can achieve better fit to counts, but rely heavily on network assignment.

4.1.1 | Screenlines and ISF Application for NIRCC

The first step of ISF, therefore, is screenline definition. Figure 11 presents the 93 screenlines constructed for the region. Figure 12 illustrates the screenlines in Allen County. 66 screenlines out of the total 93 are polygons surrounding a part of the region and the other 27 screenlines are lines crossing the region and partitioning the entire region to 2 subareas. Screenlines were built with respect to several facts including count station locations, zone borders, centroid connector locations, and natural\physical barriers such as rivers, freeways, and waterways. For instance, one major rule is that all centroid connectors of each zone must fall in the same partition created by screenlines. 902 links out of 3,738 links with AADT cross the screenlines. 508 links out of 902 links (56 percent) are in Allen County. As a result, the majority portion of counts can be used for validation (only 24 percent of counts will be used in the trip table adjustment). In addition, slightly over half of the count stations selected for trip table adjustment are in Allen county, which is consistent with the fact that this county includes 2,197 count stations out of 3,738 (59 percent of all count stations).

FIGURE 11: SELECTED SCREENLINES IN THE REGION

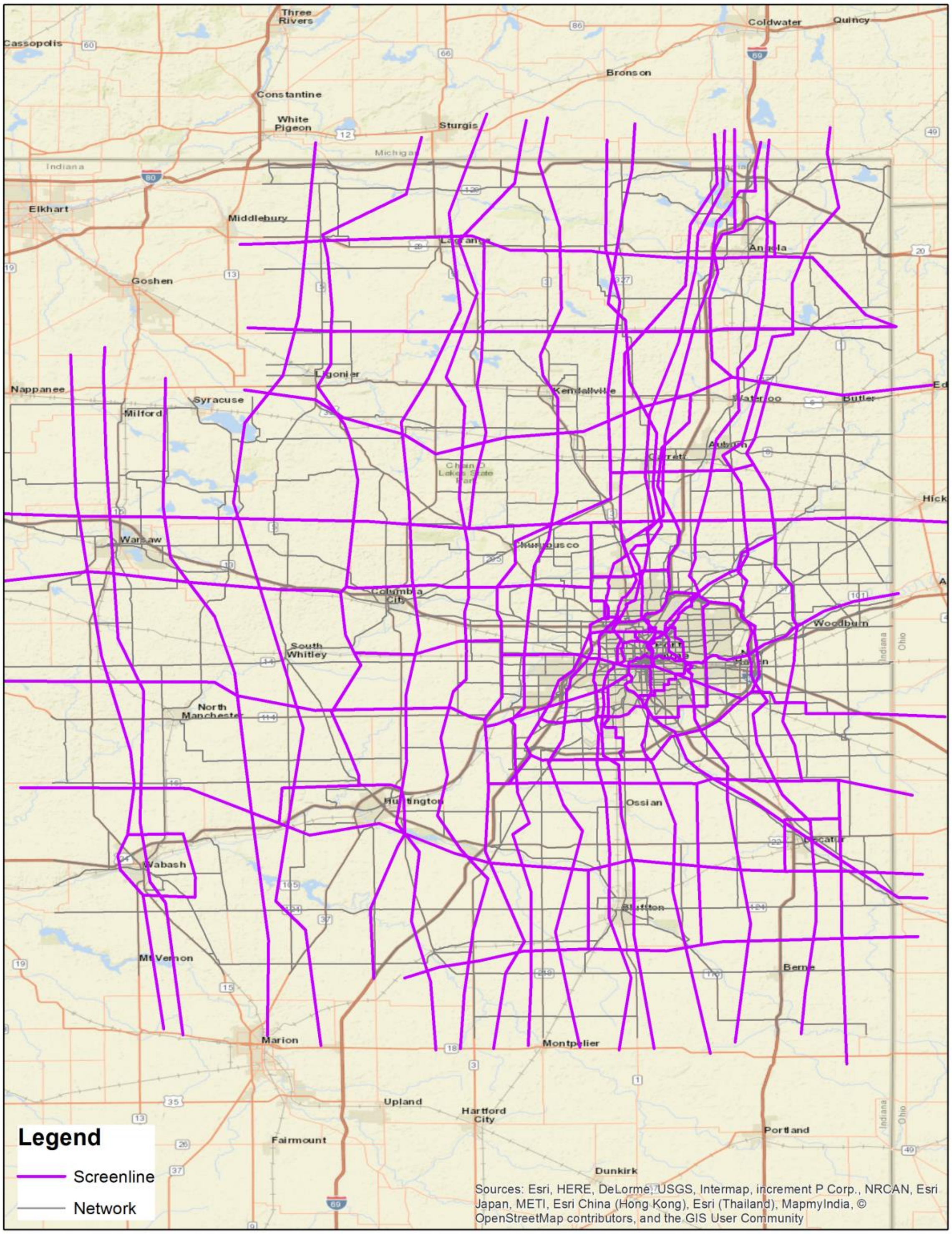
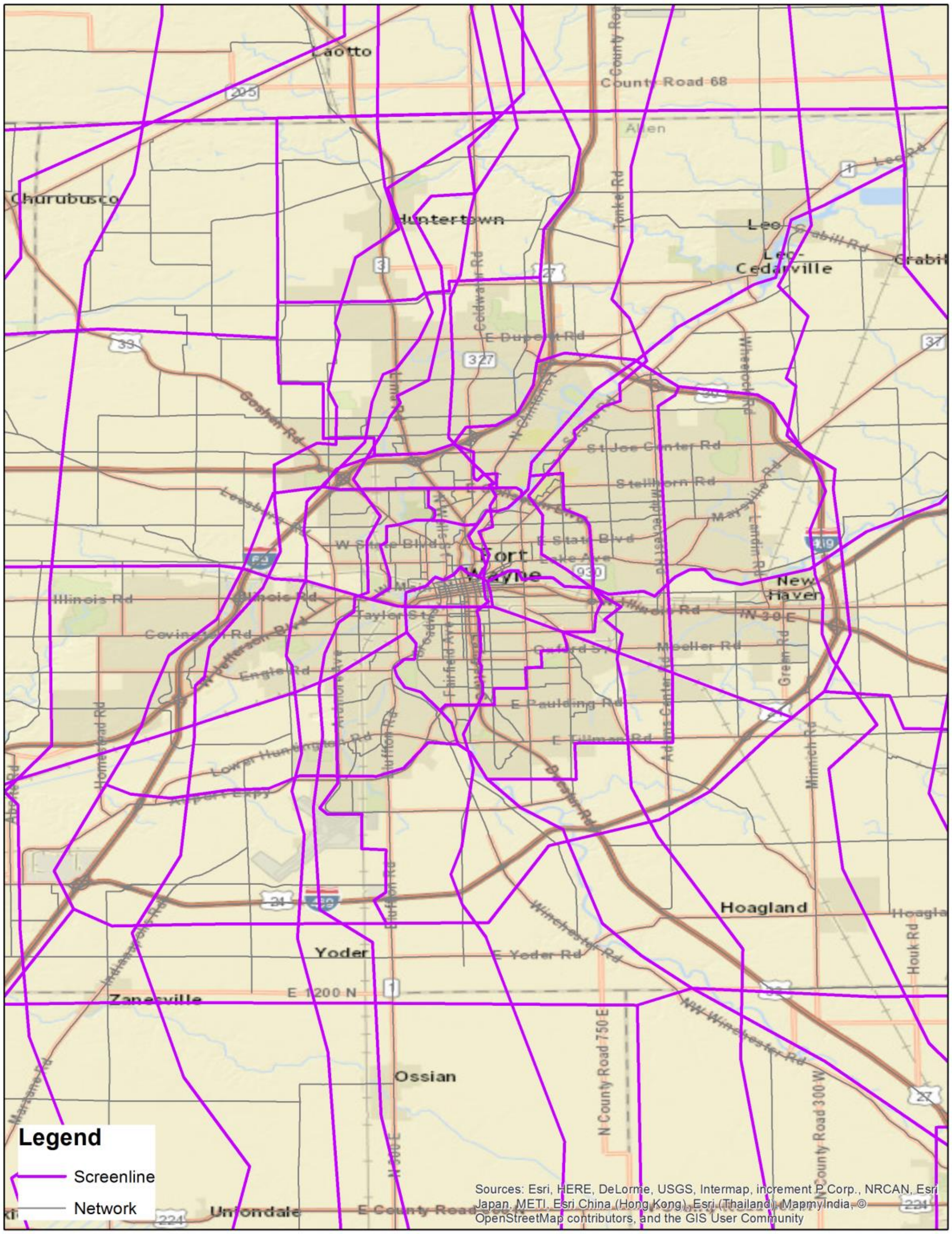


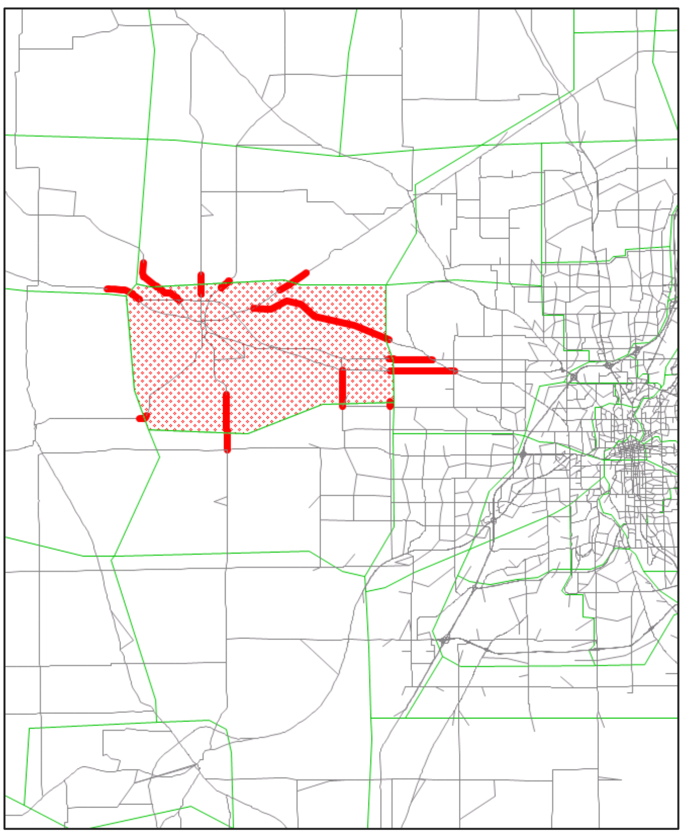
FIGURE 12: SELECTED SCREENLINES IN ALLEN COUNTY



Every polygon screenline divides the TAZs to two groups: TAZs inside the polygon and TAZs outside the polygon. Similarly, the crossing screenlines separate zones into two groups as each group includes zones on only one side of the line. As a result, each screenline divides the trip table into four quadrants. The diagonal quadrants cover trips for which both origin and destination are either inside or outside of the polygon (or on only one side of screenline). The non-diagonal quadrants include trips from zones inside the polygon to the zones outside the polygon and vice versa (or from one side to the other side of a screenline).

Figure 13 shows one polygon screenline and the network links crossed by the screenline. The ratio of weighted total count to weighted total model volume for the links is defined as the expansion factor for the corresponding screenline. The expansion factor is only applied to non-diagonal cells since the ratio is calculated based on the links going from one side to the other side of each screenline. Figure 13 illustrates that the selected links are not likely affected by trips within the polygon or between zones outside the polygon. This is the reason the expansion factors are not applied to diagonal cells in the trip table.

FIGURE 13: NETWORK LINKS CROSSED BY ONE POLYGON SCREENLINE



As mentioned previously, the weighted average ratio of counts to trips based on the network links crossing each screenline is used to calculate the expansion factor. Three separate weights are used in the expansion factor as follows:

1- Number of screenlines the network is included on

2- Functional class of the network link

3- Area type of the network link

As seen in Figure 11, there are several screenlines in the model area and there is a chance that a network link is crossed by several screenlines. Without weighting, the counts on these links would affect the final trip table more than other links crossing only one screenline. Especially if a link has a very high volume or count and crosses several screenlines, the trip table adjustment factor could be skewed toward the link's count-to-volume ratio. To avoid giving more priority to these links, therefore, a weight was defined for each network link which is equal to one over number of screenlines the link crosses over. The total volume and AADT of these links are therefore taken into account and all screenline counts are given equal weight as expansion factors are computed for all screenlines.

Since the counts on lower-rank facility types may not be as accurate as higher-rank facility types and also the assignment performance along higher-rank facility types is more important than other facility types, another weight was introduced to ISF which is equal to two for interstate highways and 1 everywhere else. Moreover, the area type weight is equal to two for any link in Allen County or Adam County and one for other links as the network has higher resolution and details in Allen county and Adam County and counts are believed to be more accurate here.

One matrix of expansion factors, taking weights into account, is assigned to each screenline (totally 93 matrix) and quadrants are separately defined for each matrix according to the corresponding screenline. All cells in diagonal quadrants are filled with one while non-quadrant cells are filled with the expansion factor of the corresponding screenline. The overall expansion factor for each cell for each iteration is then the average of all non-one expansion factors for that cell. The factors are applied at the end of each iteration and several iterations are required to achieve a converged result.

4.1.2 | Summary of ISF Results

Detailed results of the ISF process are presented in Appendix D. The Cuebiq expansion was relatively straightforward, using total ADT counts along the screenlines as the control data. Initially, the INRIX data was expanded using the class specific counts. However, this did not produce attractive results, particularly for the single-unit truck (SUT) class. Therefore, the project team modified the approach, expanding the SUT and MUT trips together using total truck counts as the control and then splitting the matrix back out to SUT and MUT based on the origin-destination specific proportions in the original data.

Application of ISF revealed issues with the reliability of the count data, particularly for low volume roads and counts outside of Adams and Wells counties. Effort was made to ensure that issues were addressed on counts with substantial volume, but fit to counts is presented here only for counts > 1000 ADT because the quality of counts with less volume than this could not be assured (however, statistics including those counts can be found in the appendix).

Figure 14 and

Table 5 present the result of expanding the Cuebiq data using the iterative screenline fitting technique.

FIGURE 14: ASSIGNMENT LOADING ERROR FOR CUEBIQ TRIPS AFTER ISF

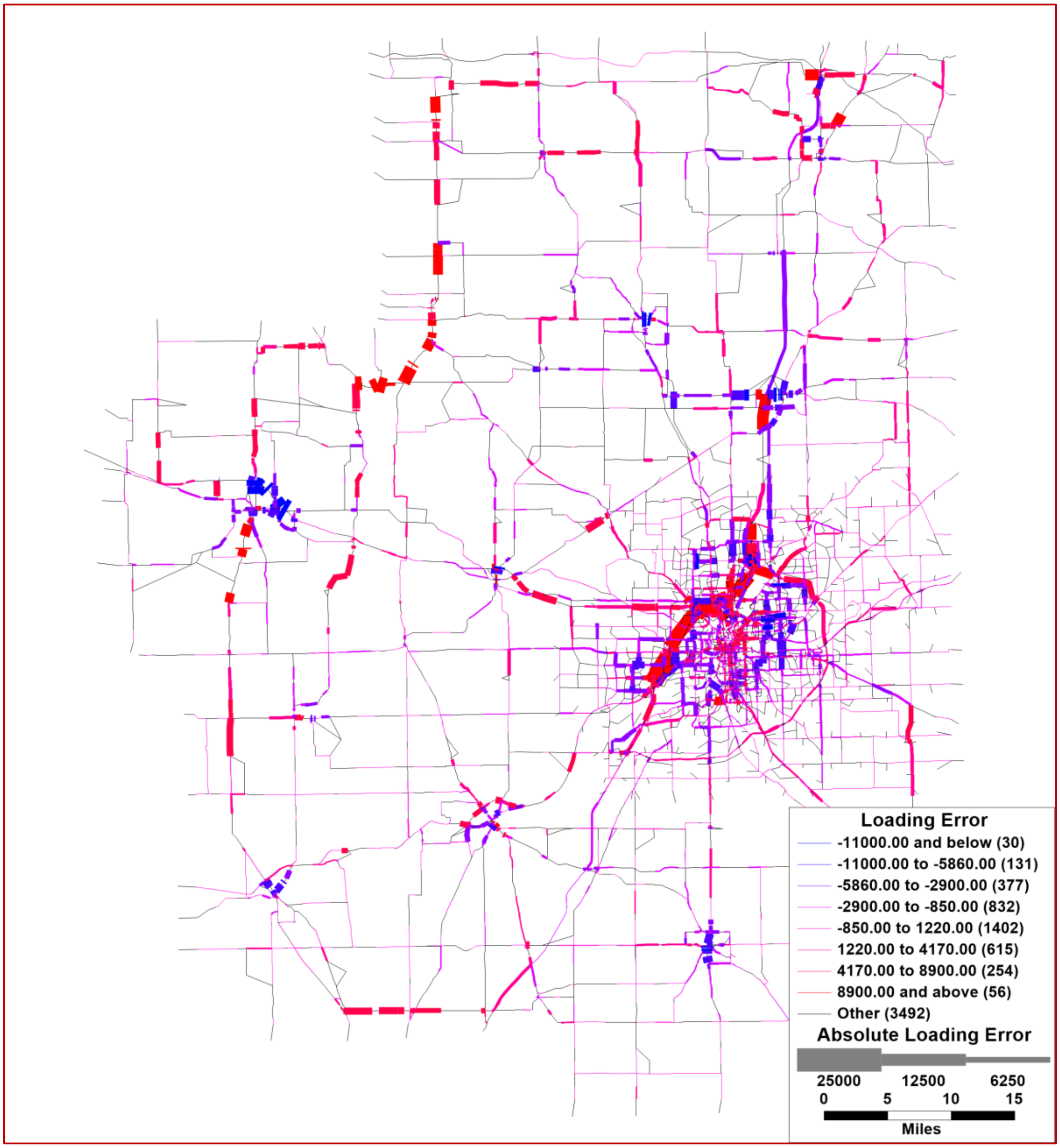


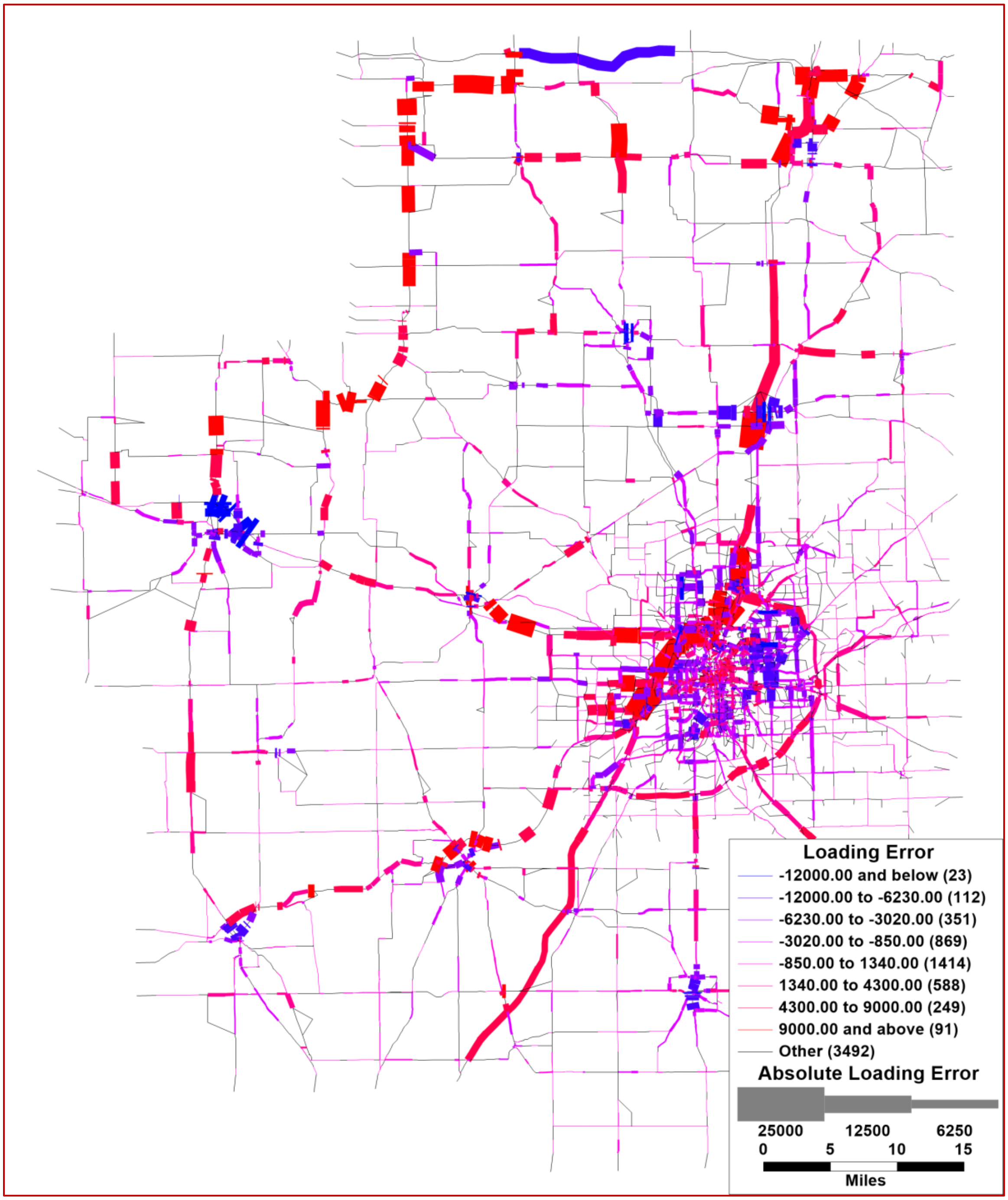
TABLE 5: ASSIGNMENT STATISTICS FOR EXPANDED CUEBIQ TRIPS FOR ANY LINK WITH AADT>1000

Assignment Statistic	Average Error	Loading Error (%)	RMSE (%)	MAPE
All Facilities	-137.1	-1.8	48.7	48.2
Freeways	4007.0	20.5	37.0	25.9
Arterials	-60.9	-0.6	38.0	35.3
Collectors	-664.9	-18.5	84.9	63.7
Locals	-1057.5	-45.4	97.4	56.4
Urban Links	-547.8	-5.6	42.7	42.4
Rural Links	243.1	5.1	61.8	54.4

TABLE 6: ASSIGNMENT STATISTICS FOR ALL EXPANDED INRIX TRIPS (AADT > 1000)

Assignment Statistic	Average Error	Loading Error (%)	RMSE (%)	MAPE
All Facilities	-126.3	-1.6	51.3	48.4
Freeways	6403.6	32.7	42.7	38.1
Arterials	-143.7	-1.4	40.4	37.3
Collectors	-643.3	-17.9	85.4	61.1
Locals	-1070.5	-45.9	97.3	53.6
Urban Links	-627.4	-6.4	44.0	42.7
Rural Links	409.6	8.6	69.9	54.5

FIGURE 15: ASSIGNMENT LOADING ERROR FOR INRIX TRIPS AFTER ISF



ISF significantly improved the assignment results for both datasets. The overall RMSE versus ADT counts decreased by nearly 20% for both datasets. There was improvement across all classes, but particularly for rural roadways. The loading error map also confirms that ISF has resolved some of the initial trip table issues though some issues still remain.

Comparison between Table 5 and Table 6 indicates that the model performance with expanded Cuebiq trips is slightly better than INRIX expanded trips. For instance, as can be seen in the loading error maps, the I-90 corridor in the study area is underloaded in the INRIX loaded network; however, it is in a good agreement with the counts in the Cuebiq loaded network.

ORIGIN-DESTINATION MATRIX ESTIMATION

Although ISF substantially improved the fit of the passively collected trips to traffic counts, ODME was ultimately employed to further reconcile the data and provide a better fit to counts. Origin-destination matrix estimation (ODME) is conceptually similar to ISF as both methods adjust trips according to count observations; however, ODME uses all counts and a network assignment model, whereas, ISF uses only the counts on screenlines and does not rely on network assignment, such as user equilibrium. There can be a danger of over-fitting trip data to counts using ODME. Therefore, to limit the modification of the trip table by ODME, upper and lower bounds were introduced to ODME adjustments. ODME could not increase any cell in the trip table by a factor greater than 3 or decrease it by a factor less than 0.5. The number of ODME iterations was also limited to 15, and the results were examined to see the amount of perturbation from the original data (documented in the next subsection).

For each type of passively collected trips explained in the previous sections (Cuebiq, INRIX, and Hybrid) ODME was run in two situations as follows:

1- Independent ODME: ODME was run on the trip table after fratar (using the same starting point as ISF), and

2- Sequential ODME: ODME was run on the trip table after ISF (taking the ISF results as the inputs to ODME).

Independent ODME was done primarily to check and evaluate ISF performance; while, sequential ODME was expected to obtain the best trip table since it should have a better starting point. The closer ISF results to ODME results, the better the performance of ISF in adjustment of trips. Independent ODME results confirm that ISF significantly improved trips using only a portion of traffic counts. In fact, ISF enhanced trips more than half of the amount ODME could while using many less traffic counts. Full results of ODME including the Independent ODME results and results for the Hybrid trip table are reported in Appendix D. The final results for ODME applied sequentially after ISF are presented below for the INRIX and Cueibiq datasets.

Figure 16 and Table 7 present the results for the Cuebiq data. The RMSE of sequential ODME is the best among all runs of expanded Cuebiq trips, including independent ODME. Thus, ISF and ODME run sequentially expanded the passively collected trips better than either of the approaches alone.

FIGURE 16: ASSIGNMENT LOADING ERROR FOR CUEBIQ TRIPS AFTER SEQUENTIAL ODME

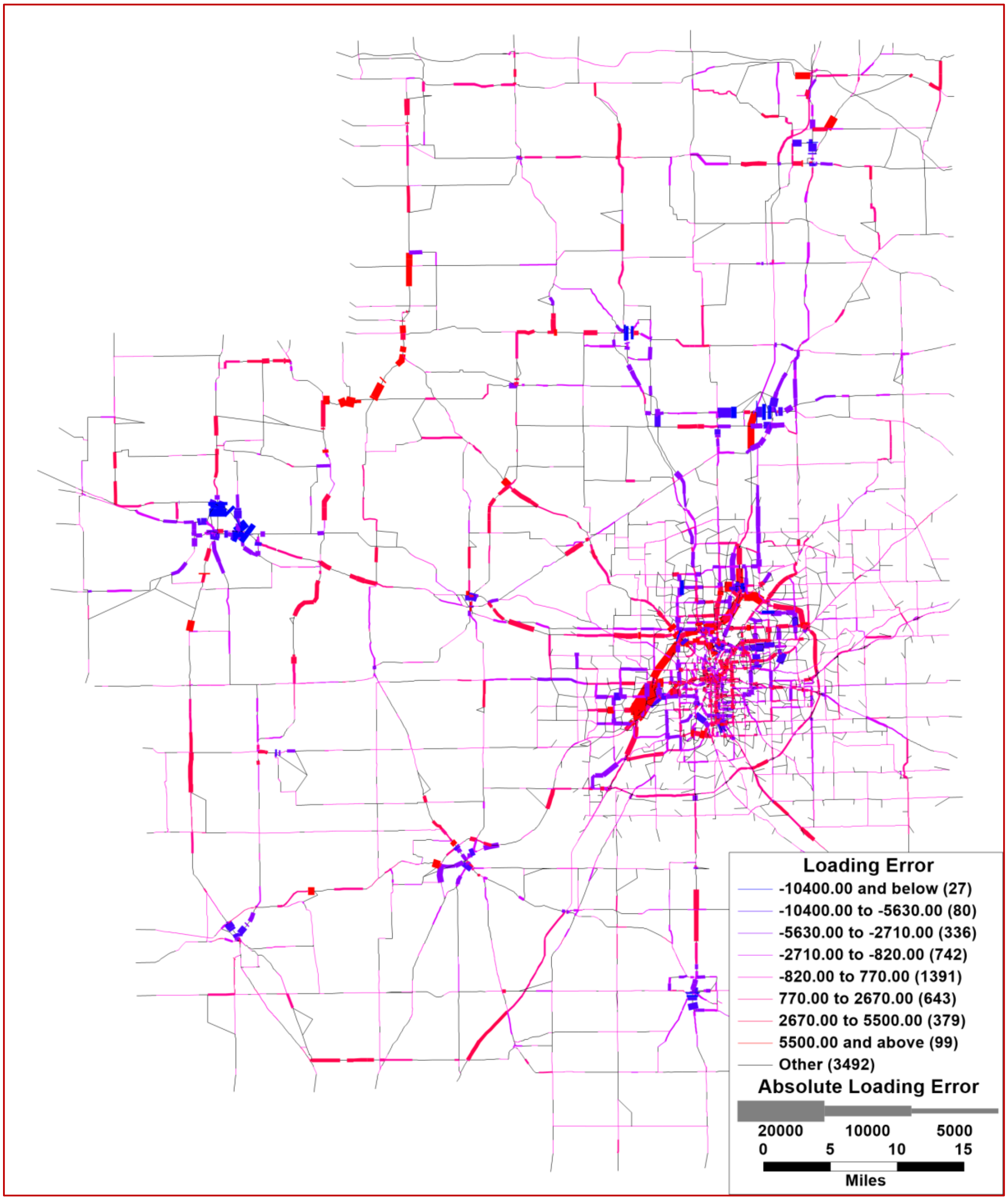


TABLE 7: ASSIGNMENT STATISTICS FOR CUEBIQ TRIPS AFTER SEQUENTIAL ODME (AADT > 1000)

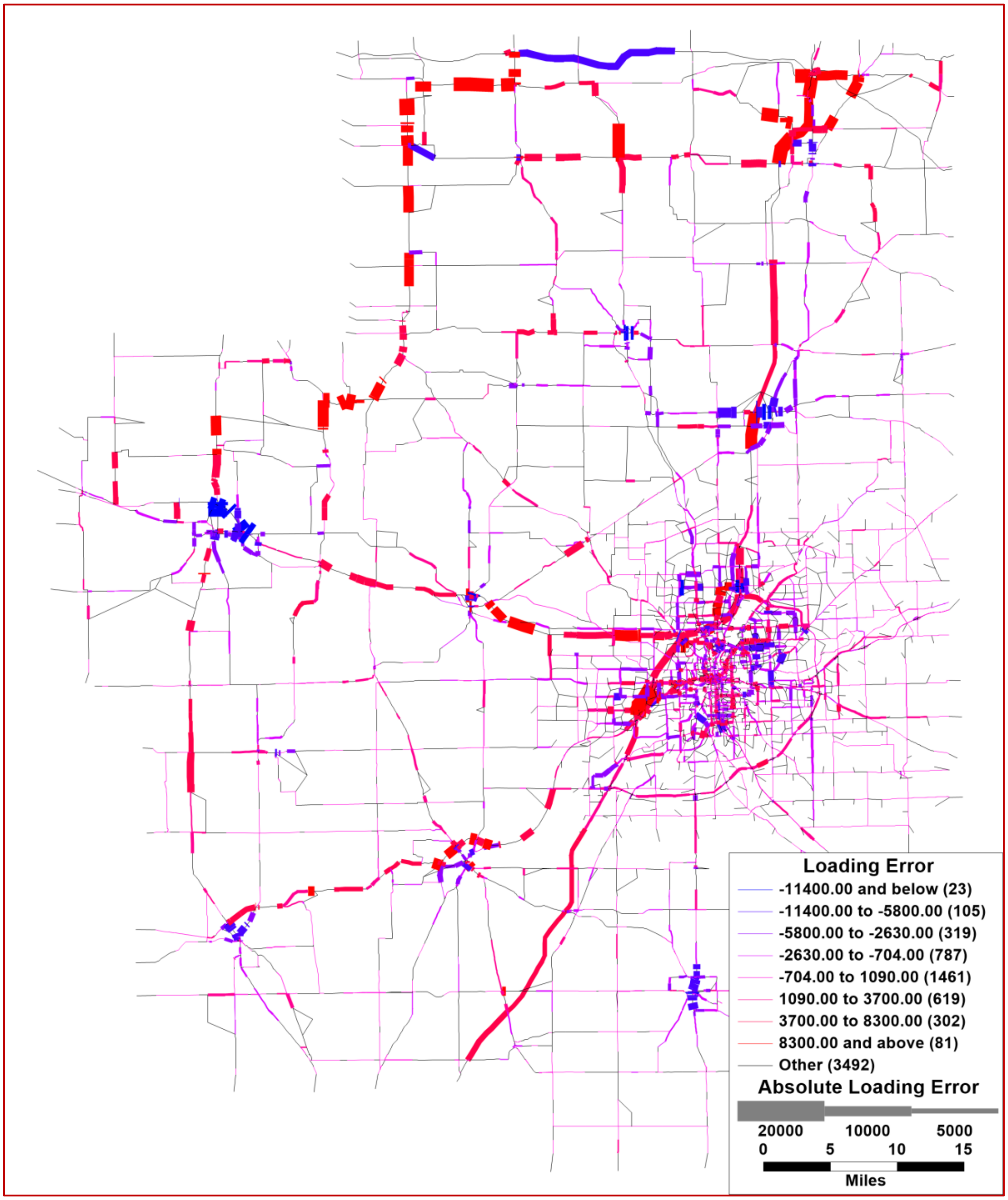
Assignment Statistic	Average Error	Loading Error (%)	RMSE (%)	MAPE
All Facilities	-167.5	-2.2	39.2	38.1
Freeways	2367.9	12.1	23.8	16.9
Arterials	121.6	1.2	32.4	30.4
Collectors	-823.8	-22.9	62.5	47.4
Locals	-1127.3	-48.4	97.2	57.2
Urban Links	-367.9	-3.7	35.4	35.9
Rural Links	-36.0	-0.8	45.4	39.7

Figure 17 and Table 8 present the results for the INRIX data. As with Cuebiq, ISF and ODME run sequentially expanded the passively collected INRIX trips better than either of the approaches alone although the advantage of the sequential process over independent ODME was less for INRIX.

TABLE 8: ASSIGNMENT STATISTICS FOR ALL INRIX TRIPS AFTER SEQUENTIAL ODME (AADT > 1000)

Assignment Statistic	Average Error	Loading Error (%)	RMSE (%)	MAPE
All Facilities	11.6	0.2	45.9	41.9
Freeways	3357.3	17.1	28.8	22.7
Arterials	163.8	1.6	37.1	32.6
Collectors	-498.3	-13.9	78.6	53.0
Locals	-910.5	-39.1	94.3	56.2
Urban Links	-370.1	-3.8	38.6	36.2
Rural Links	423.8	8.9	65.8	48.4

FIGURE 17: ASSIGNMENT LOADING ERROR FOR INRIX TRIPS AFTER SEQUENTIAL ODME



As can be seen from a comparison of Table 7 and Table 8, the overall RMSE for INRIX is higher than Cuebiq. The expanded Cuebiq data performs particularly better for rural areas and the imbalance between urban and rural errors is much less than originally or in INRIX. Although the Cuebiq data is only marginally better than INRIX in the urban areas, it is an order of magnitude more accurate in rural areas.

MATRIX COMPARISONS

When using ODME, it is important to ensure that the powerful ODME process does not overly perturb or mangle the original trip data in the process of matching counts. This section of the report therefore documents a comparison of how the trip data or OD matrix was changed by ODME both in comparison to the results of ISF and in comparison to the original data after it had only been fratared and scaled. The selected measures are Mean Absolute Error (MAE) and Mean Absolute Percentage Error (MAPE). Both MAE and MAPE report the difference between the input and output trips from ODME. Equations 1 and 2 shows the formulas for MAE and MAPE.

$$\text{MAE} = \frac{\sum_{i=1}^{n} |\, Y_i - X_i|}{n} \qquad \text{Eq. 1}$$

$$\text{MAPE} = 100 \times \frac{1}{n} \sum_{i=1}^{n} \frac{|Y_i - X_i|}{|Y_i|} \qquad \text{Eq. 2}$$

Independent ODME is based on trips after fratar. The measures for independent ODME, therefore, are calculated based on fratared trips. On the other hand, fratared trips are adjusted with ISF and the resulted trips are adjusted again in sequential ODME. As a result, sequential ODME measures are calculated for both fratared and ISF trips in order to understand the effect of ODME adjustments on the ISF results as well as the combined effect of ISF and ODME on the original fratared data (in part, for comparison with independent ODME).

TABLE 9: MATRIX COMPARISON BETWEEN SEQUENTIAL ODME AND ISF

Statistic	CUEBIQ	INRIX	HYBRID
MAE	2.7	2.1	0.8
MAPE	17.6%	28.7%	7.2%

TABLE 10: MATRIX COMPARISON BETWEEN SEQUENTIAL ODME AND ORIGINAL FRATAR

Statistic	CUEBIQ	INRIX	HYBRID
MAE	3.5	3.4	4.7
MAPE	20.7%	42.3%	19.4%

According to Table 9 and Table 10, on average, ODME did not change the trip tables dramatically. Even compared to the original data before ISF, the Cuebiq and Hybrid trip tables after ODME expansion are still typically within about 20% of their original values, well within reasonable assumptions of sampling error for these passive data sources. Even the INRIX dataset, which does exhibit more perturbation through both stages of expansion is still generally within 50% of original values, which means the implied expansion factors are not extreme. In summary, comparison of the expanded and original trip tables validates the reasonableness of the expansion factors and does not indicated over-fitting by ODME.

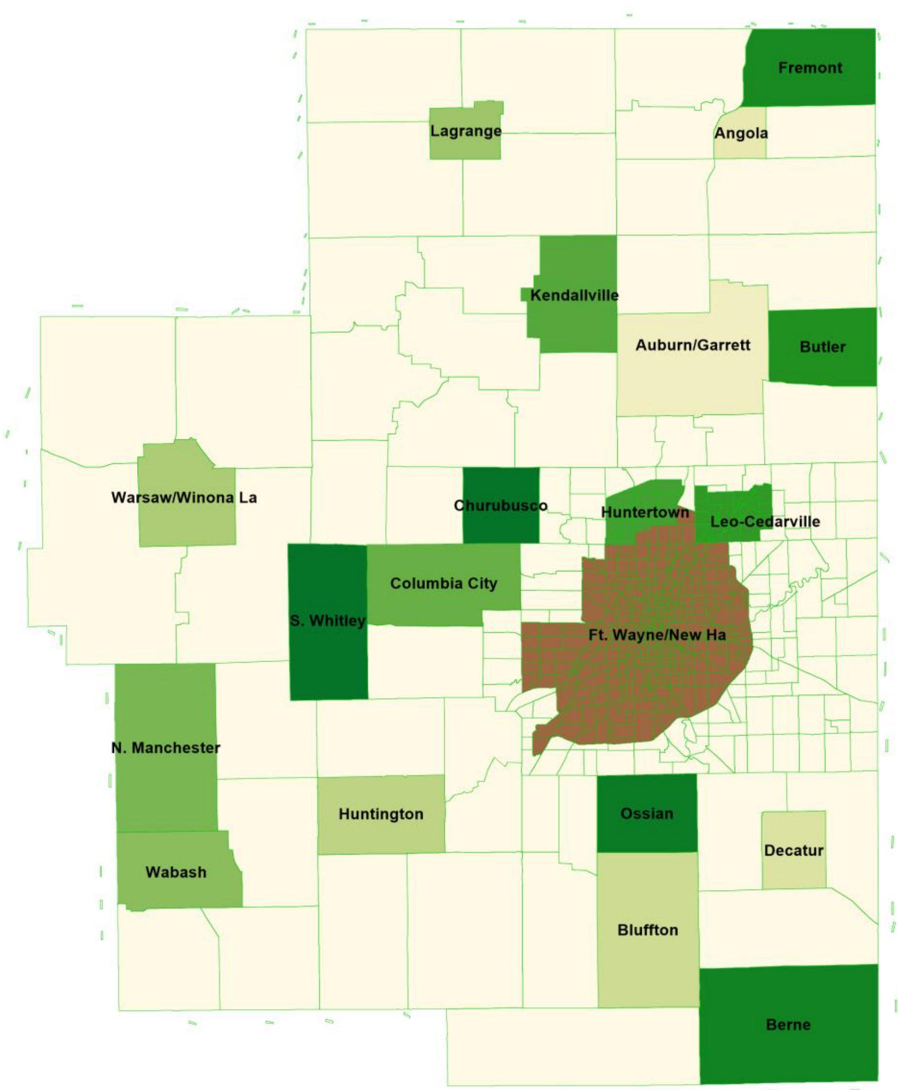
5.0 DATA ANALYSIS AND VISUALIZATION

5.1 | REGIONAL INTERACTIONS OF COMMUNITIES OF INTEREST

In this section of the report, a variety of visualizations and descriptive statistics are produced to develop insights and a deeper understanding of regional travel in Northeast Indiana. The main focus was two-fold, first, to understand the level of interaction among the various communities in the region, and second, to understand the demand patterns served by key facilities in the region, especially the proportion of trucks on specific facilities that are passing through the region versus interacting with it.

First, the daily trips between selected communities in the region are analyzed. NIRCC staff identified 12 primary communities and 8 secondary communities for this analysis as shown in Figure 18.

FIGURE 18: PRIMARY AND SECONDARY COMMUNITIES IN THE REGION



Selected primary communities are as follows:

1- Fort Wayne/New Haven in Allen County
2- Auburn/Garrett in DeKalb County
3- Angola in Steuben County
4- Decatur in Adams County
5- Bluffton in Wells County
6- Huntington in Huntington County
7- Warsaw/Winona Lake in Kosciusko County
8- Lagrange in Lagrange County
9- Wabash in Wabash County
10- North Manchester in Wabash County
11- Columbia City in Whitley County, and
12- Kendallville in Noble County

Selected secondary communities are as follows:

1- Huntertown in Allen County
2- Leo-Cedarville in Allen County
3- Butler in DeKalb County
4- Fremont in Steuben County
5- Berne in Adams County
6- Ossian in Wells County
7- Churubusco in Whitley County
8- South Whitley in Whitley County

The Cuebiq trips expanded by ISF and ODME were selected to visualize the interactions between communities since this dataset generated the best fit to counts among all expanded trip tables. The trip table was then aggregated to the level of defined communities to find regional interactions between them. Table 11 reports the daily interaction between communities based on expanded Cuebiq trips. To reduce the size of the table and make it readable all secondary communities were reported together. The rest of the region zones which are not part of communities were also grouped and labeled "Other" in the table. Since Fort Wayne/New Haven plays the main role in the study area, its interactions with other communities are highlighted in Table 11. According to the table, Auburn has the highest interaction with Fort Wayne/New Haven among all primary communities, followed by Columbia City, Bluffton, Warsaw, and Huntington. However, the table also clearly shows that the interaction between the communities is small compared to the trip-making within the communities.

TABLE 11: DAILY TRIPS BETWEEN COMMUNITIES

Community	Ft. Wayne	Auburn	Angola	Decature	Bluffton	Huntington	Warsaw	Lagrange	Wabash	N Manchester	Columbia City	Kendallville	Secondary	Other	Total
Ft. Wayne	1,202,846	4,544	470	1,334	2,863	1,980	2,143	360	791	281	3,082	1,055	50,118	70,339	1,342,205
Auburn	4,626	86,830	236	22	19	30	52	32	16	3	149	962	2,651	9,682	105,311
Angola	438	209	32,504	2	6	0	9	225	0	0	4	607	2,970	12,111	49,085
Decature	1,295	21	6	48,603	1,007	28	10	0	13	0	10	4	582	4,800	56,380
Bluffton	2,705	13	2	890	53,018	322	7	0	17	0	50	3	4,044	4,985	66,057
Huntington	1,811	22	1	46	383	73,057	45	6	359	36	212	6	376	7,321	83,680
Warsaw	1,951	49	11	4	4	52	127,113	2	92	128	572	12	350	14,495	144,835
Lagrange	316	47	278	2	0	0	6	8,279	0	0	33	558	81	7,318	16,917
Wabash	860	13	0	30	8	383	110	0	39,402	444	83	0	49	4,702	46,085
N Manchester	342	0	0	0	0	34	137	0	448	928	25	0	25	2,066	4,007
Columbia City	3,087	173	2	10	29	198	520	22	67	18	40,891	165	1,249	6,731	53,161
Kendallville	1,031	992	557	3	5	6	19	583	1	0	50	38,917	478	7,709	50,350
Secondary	50,192	2,583	3,035	590	3,986	409	497	68	49	25	1,156	548	66,070	21,049	150,257
Other	72,412	9,656	11,847	4,848	4,730	7,432	14,516	7,337	4,825	1,882	6,736	7,589	21,087	340,860	515,759
Total	1,343,914	105,153	48,948	56,384	66,056	83,932	145,184	16,914	46,081	3,745	53,054	50,427	150,129	514,169	2,684,090

To illustrate the overall travel pattern between communities, chord diagrams were also generated. Figure 19 presents the interactions between primary and secondary communities as well as rural areas in the region and external areas. Rural areas includes all internal zones which are not part of the defined communities. External zones were also grouped into the four following groups to represent the location of origins/destinations outside the region:

1- Michigan
2- Ohio
3- Indiana South, and
4- Indiana west

In Figure 19, color and degrees of arcs identify communities and bands of thickness representing magnitude of flow connecting them. According to Figure 19, travel within communities far outweighs travel between them within the region. Since, the chord diagram is dominated with Fort Wayne interactions, which resulted in very narrow bands between some communities, flows between communities were presented based on a logarithmic scale in Figure 20. Although interactions between communities with lower trips are more visible in Figure 20, it still confirms the overwhelming dominance of travel by intra-community trips.

FIGURE 19: DAILY INTERACTIONS BETWEEN COMMUNITIES, RURAL, AND EXTERNALS

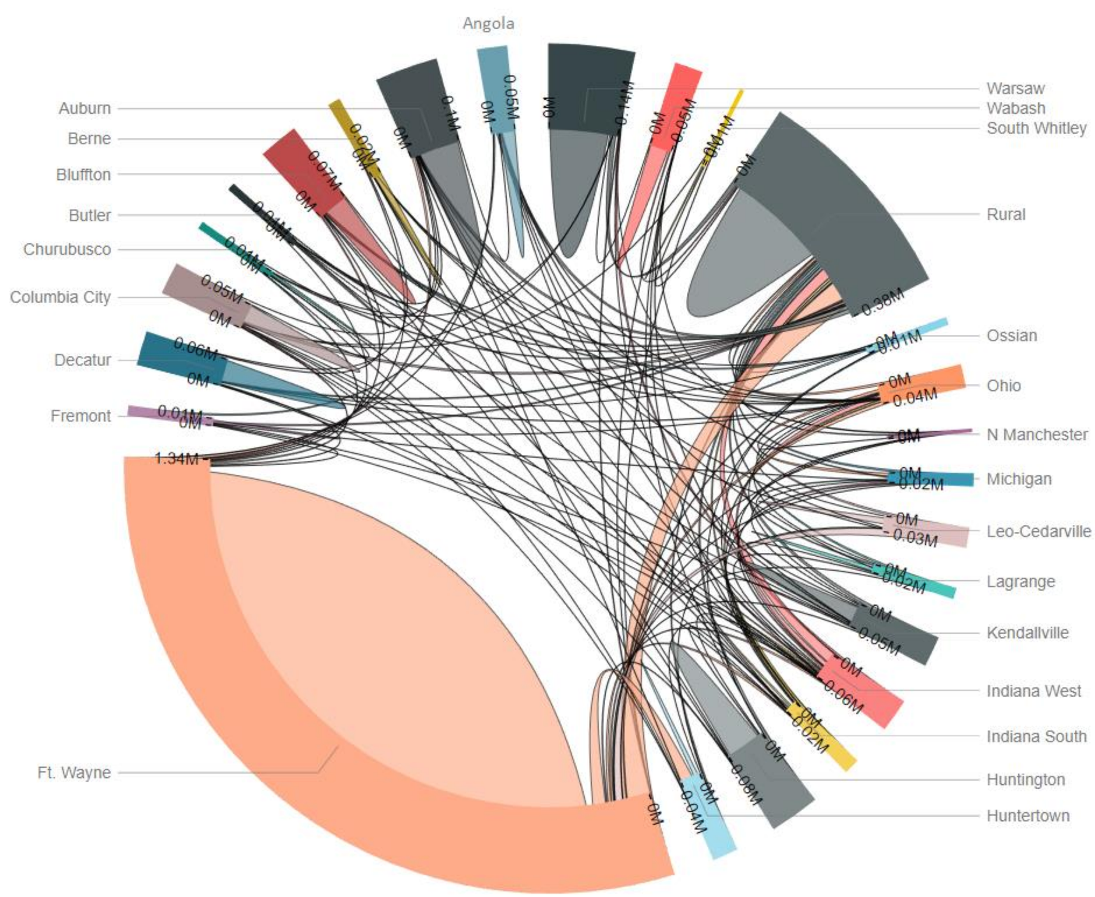
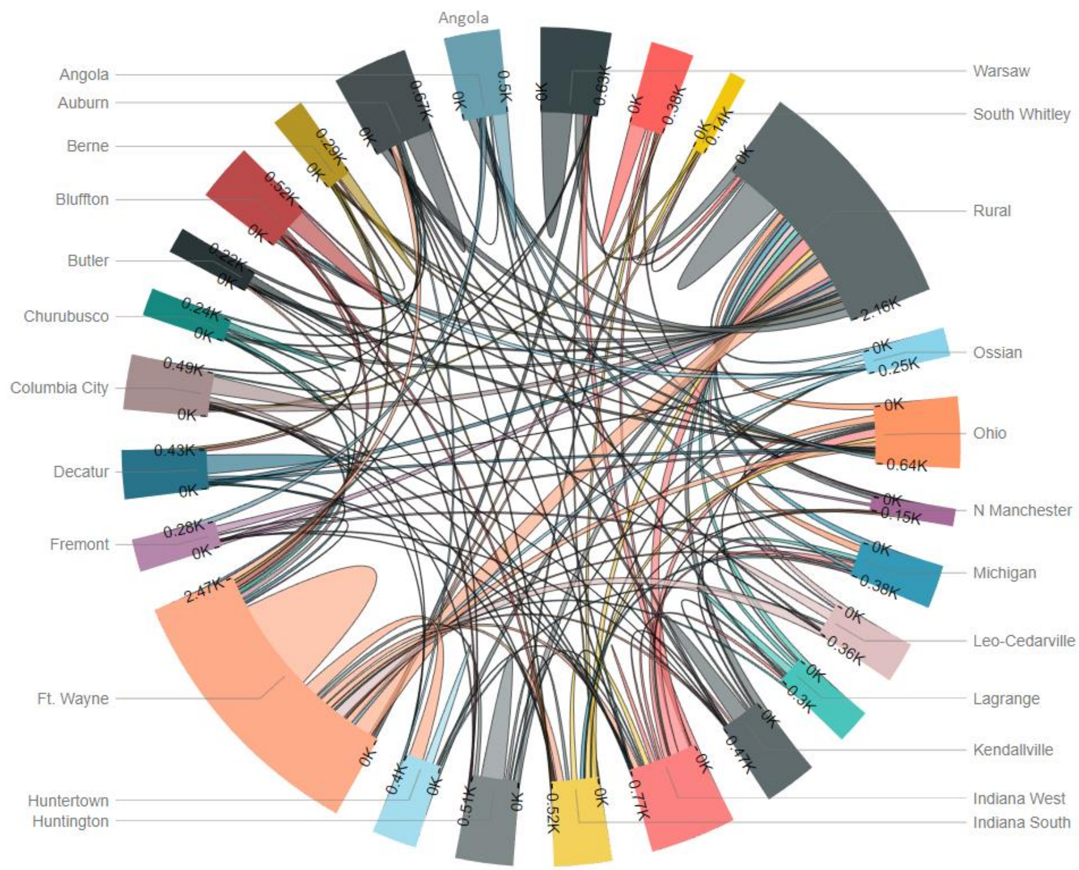
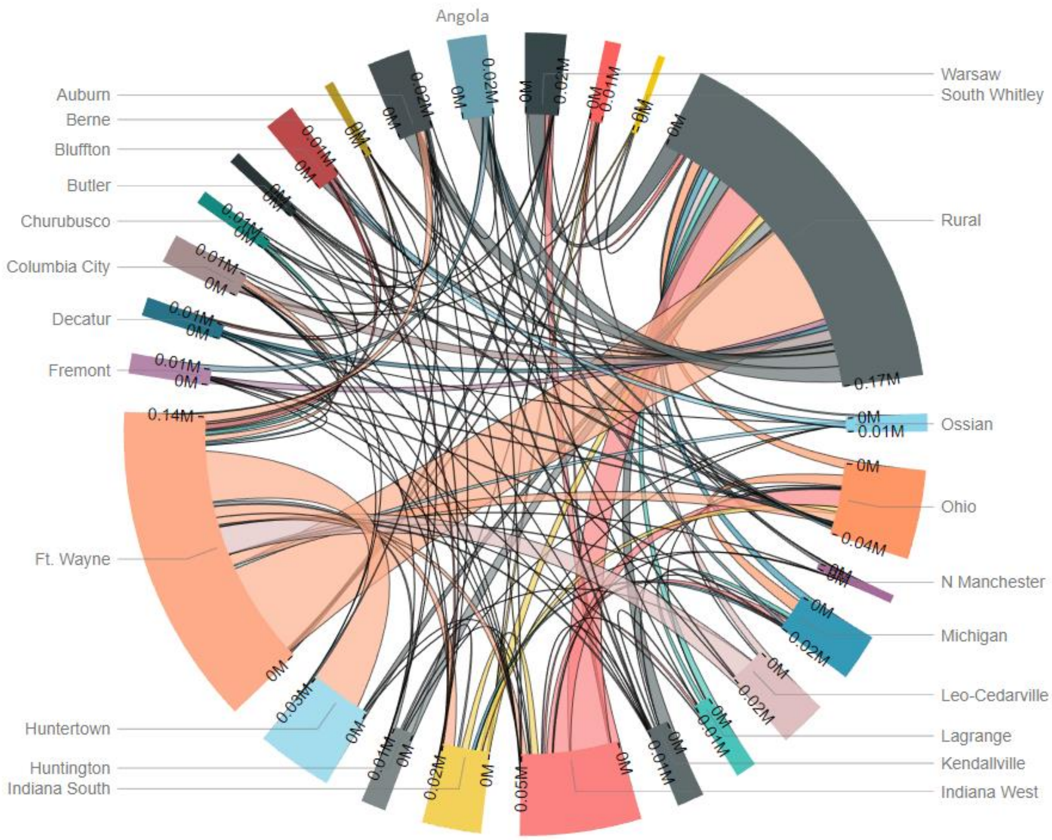


FIGURE 20: DAILY INTERACTIONS BETWEEN COMMUNITIES, RURAL, AND EXTERNALS IN LOGARITHMIC SCALE



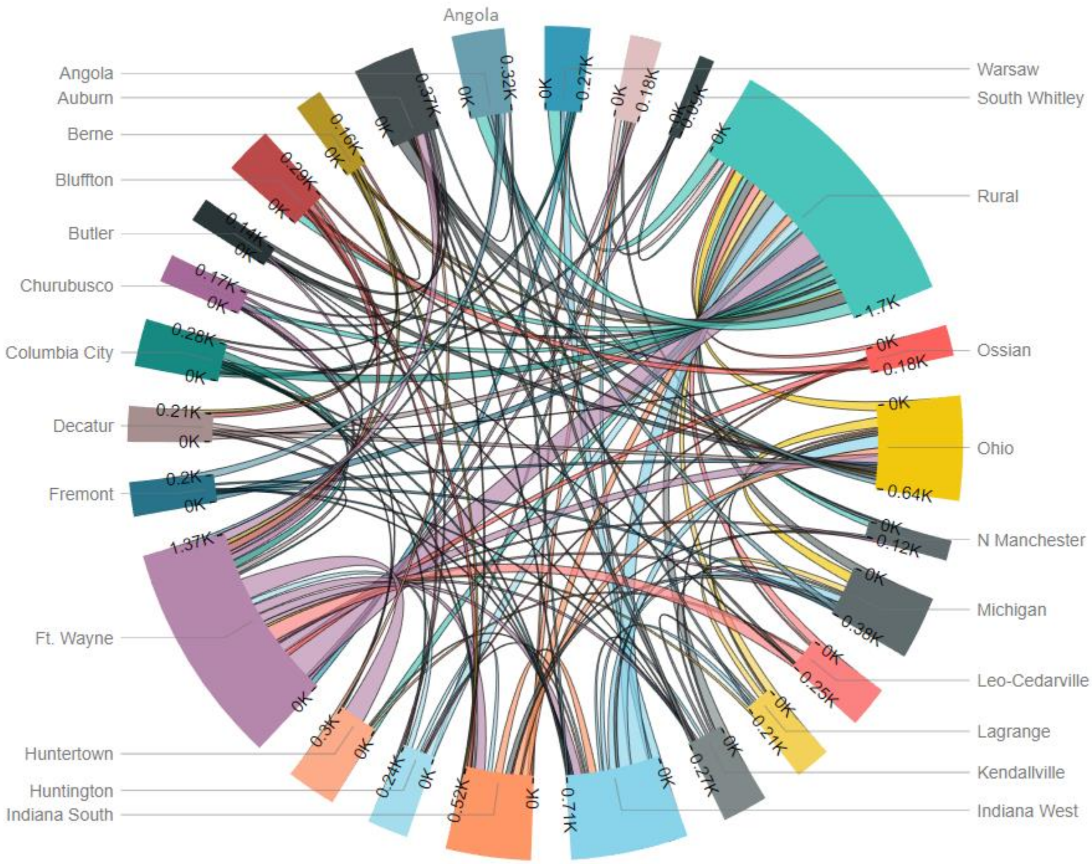
To better illustrate interactions between communities, chord diagrams were also generated without intra-district trips, which are shown in Figure 21 and Figure 22.

FIGURE 21: DAILY INTERACTIONS BETWEEN COMMUNITIES, RURAL, AND EXTERNALS WITHOUT INTRA-DISTRICT TRIPS



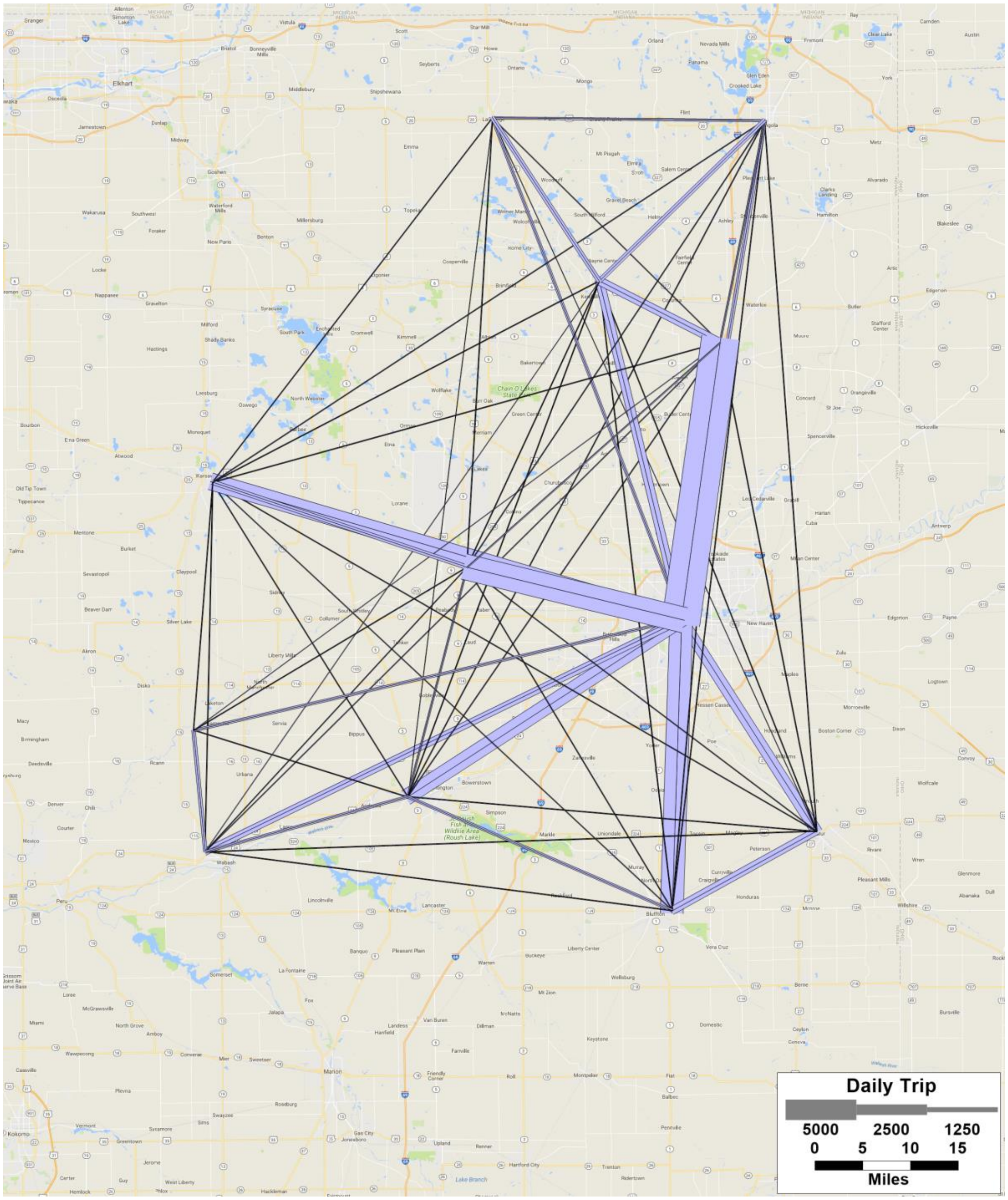
According to Figure 21 and Figure 22, travel to/from communities is dominated by travel into/out of the region (Indiana West, Indiana South, Ohio, Michigan) and by trips between communities and surrounding rural areas. Among external origins/destinations, the dominance of Indiana West is reasonable given both the proximity of the South Bend- Elkhart area and the size of the Chicago region.

FIGURE 22: DAILY INTERACTIONS BETWEEN COMMUNITIES, RURAL, AND EXTERNALS WITHOUT INTRA-INTRADISTRICT TRIPS IN LOGARITHMIC SCALE



Desire line maps are another way to illustrate the overall travel pattern using a line connecting each pair of communities and its thickness as the magnitude of the flow between communities. Figure 23 which presents the desire lines between primary communities, reconfirms that interactions with Fort Wayne/New Haven are dominant. It also shows that Auburn/Garrett, Columbia City, Bluffton, Warsaw, and Huntington are main partners of Fort Wayne/New Haven among primary communities. Bluffton-Ossian have the highest interaction between all primary community pairs not involving Fort Wayne/New Haven based on the Cuebiq data.

FIGURE 23: DESIRE LINES BETWEEN PRIMARY COMMUNITIES



Since Fort Wayne/New Haven is the main community in the region, its interactions with all primary and secondary communities are illustrated in Figure 24. According to Figure 24, interactions with

Huntertown and Leo-Cedarville are greater than other communities which is plausible given their close proximity.

FIGURE 24: DESIRE LINES FROM/TO FORT WAYNE/NEW HAVEN

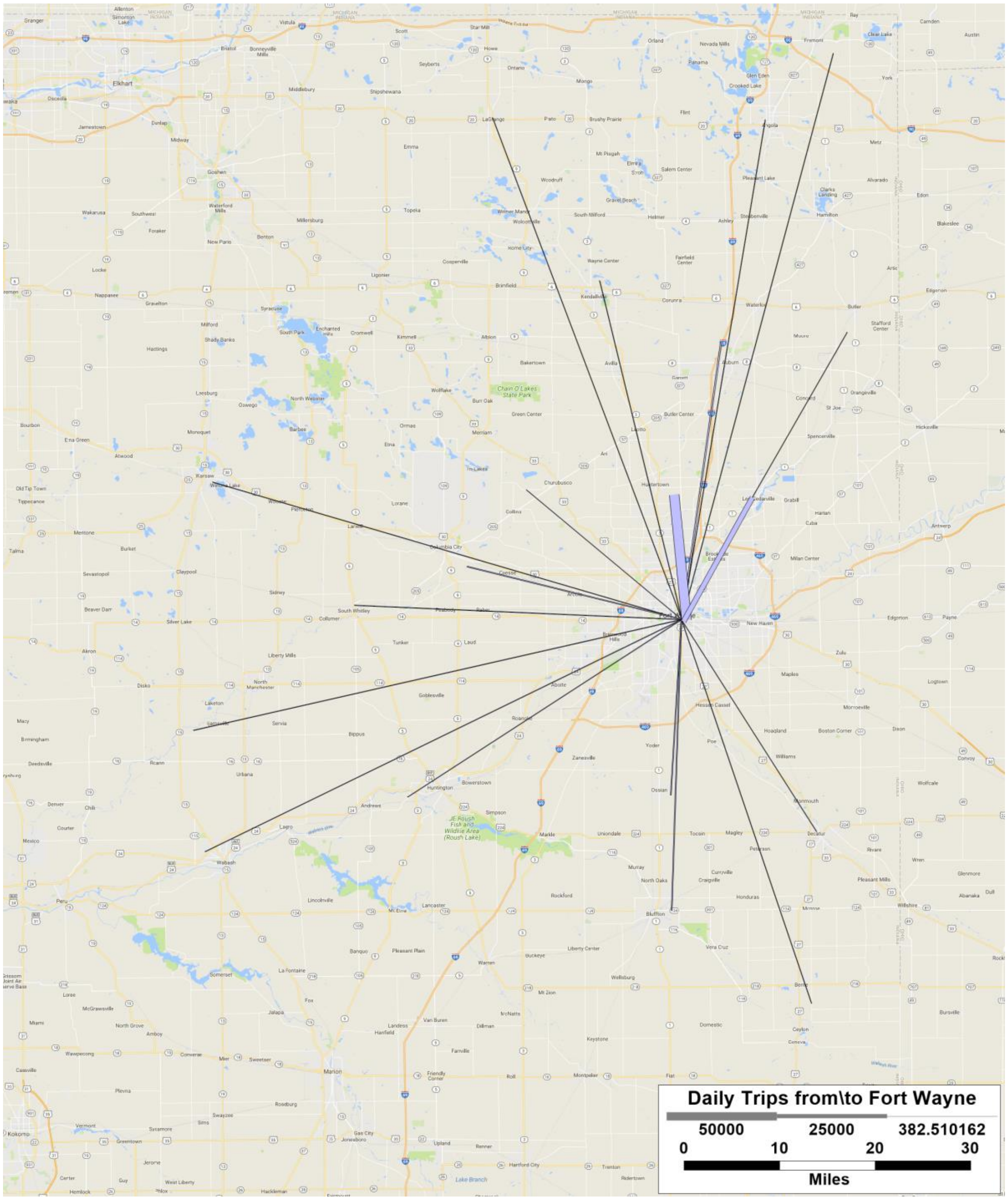
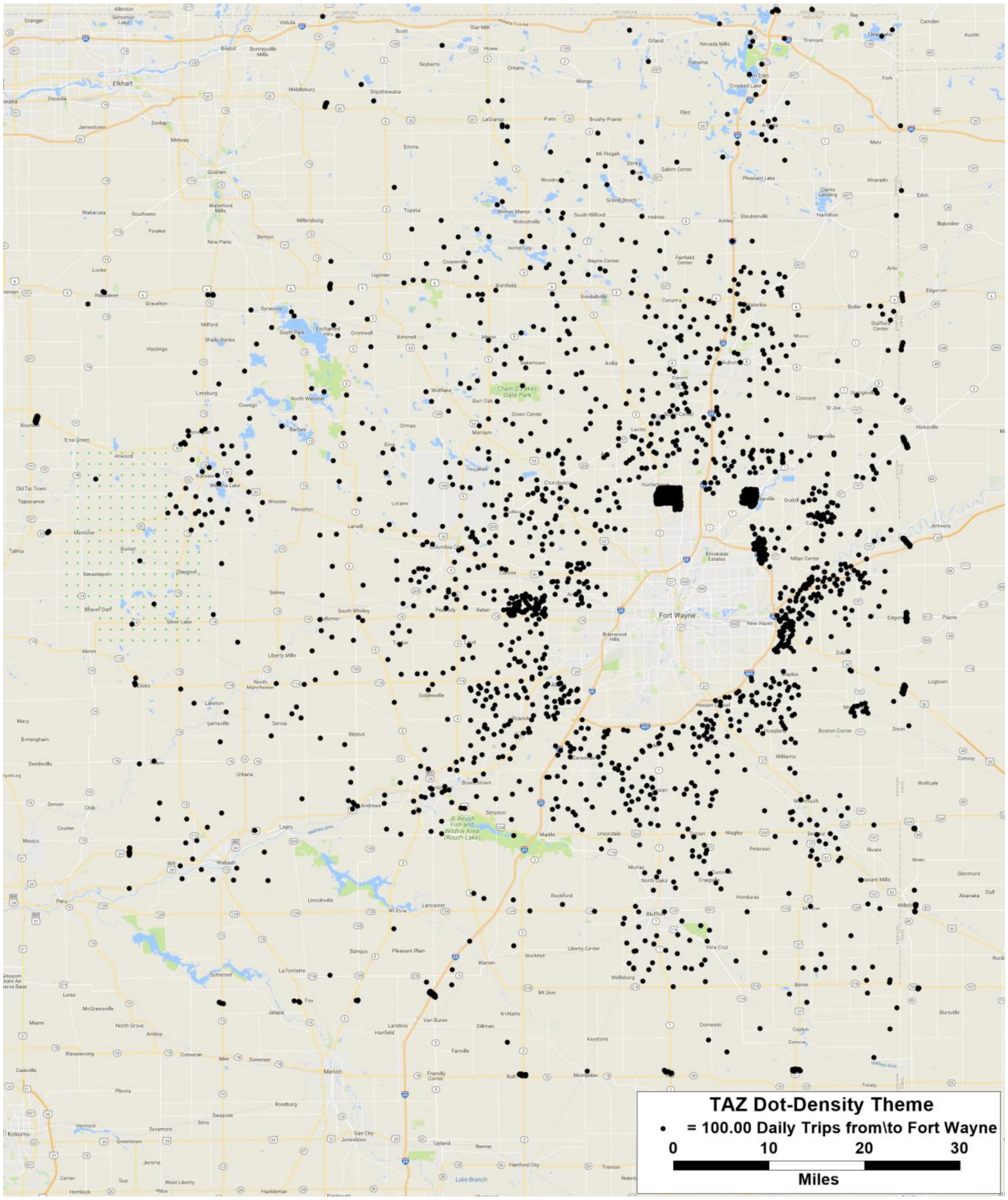


Figure 25 shows origins and destinations bound to/from Fort Wayne/New Haven taking all rural zones into consideration. Each dot in Figure 25 represents 100 daily trips to/from Fort Wayne/New

Haven. Similar to desire line map, Huntington and Leo-Cedarville shows the highest interaction with Fort Wayne/New Haven.

FIGURE 25: ORIGINS/DESTINATIONS BOUND TO/FROM FORT WAYNE/NEW HAVEN



More desire line maps for primary and secondary community can be found in Appendix A. Table 12 reports top 20 daily flows and the corresponding communities in the region. Fort Wayne/New Haven is in most of the top 20 daily flows as expected. Interaction between Ossian and Bluffton is also in Top 5 based on the Cuebiq data. School trips might be one of the main reasons of high traffic volume between these two communities.

Although agreement with traffic count data clearly indicates that the Cuebiq data is more accurate than INRIX, in order to verify the results shown in Table 12, the same analysis was also conducted using the INRIX trips expanded by ISF and ODME, with the results also reported in Table 12. According to Table 12, both Cubiq and INRIX ranks are in generally good agreement with each other. Although 5 pairs in Cuebiq top 20 are not in INRIX top 20, they are all in INRIX top 40 out of existing 380 community pairs. Moreover, 8 community pairs in the top 10 are common in both datasets with very close ranks. Although flow between Ossian and Bluffton is not in top 5 in INRIX, it is still in top 20, which confirms the importance of this interaction.

TABLE 12: TOP 20 DAILY FLOWS BETWEEN COMMUNITIES

CUEBIQ Rank	CUEBIQ Daily Flow	Community 1	Community 2	INRIX Rank	INRIX Daily Flow
1	56,479	Huntertown	Ft. Wayne/New Haven	1	54,320
2	30,473	Leo-Cedarville	Ft. Wayne/New Haven	2	29,001
3	9,170	Auburn/Garrett	Ft. Wayne/New Haven	4	8,188
4	7,139	Ossian	Bluffton	13	3,008
5	6,169	Columbia City	Ft. Wayne/New Haven	3	11,838
6	5,798	Fremont	Angola	5	8,176
7	5,568	Bluffton	Ft. Wayne/New Haven	7	5,279
8	4,500	Ossian	Ft. Wayne/New Haven	9	4,850
9	4,104	Churubusco	Ft. Wayne/New Haven	12	3,528
10	4,094	Warsaw/Winona Lake	Ft. Wayne/New Haven	10	4,011
11	3,791	Huntington	Ft. Wayne/New Haven	6	6,014
12	2,629	Decatur	Ft. Wayne/New Haven	14	2,827
13	2,354	Butler	Auburn/Garrett	16	2,224
14	2,086	Kendallville	Ft. Wayne/New Haven	11	3,845
15	1,955	Kendallville	Auburn/Garrett	23	1,285
16	1,953	Berne	Ft. Wayne/New Haven	30	889
17	1,896	Bluffton	Decatur	26	1,093
18	1,651	Wabash	Ft. Wayne/New Haven	19	1,908
19	1,508	S. Whitley	Ft. Wayne/New Haven	36	635
20	1,340	S. Whitley	Columbia City	38	612

Table 13 reports top 5 partner communities for each community. Fort Wayne/New Haven is the first partner for most of communities as expected and no lower than the third highest partner in for any community.

TABLE 13: TOP 5 PARTNER COMMUNITIES FOR EACH PRIMARY AND SECONDARY COMMUNITY

Community	First	Second	Third	Forth	Fifth
Ft. Wayne/New Haven	Huntertown	Leo-Cedarville	Auburn/Garrett	Columbia City	Bluffton
Auburn/Garrett	Ft. Wayne/New Haven	Butler	Kendallville	Huntertown	Leo-Cedarville
Angola	Fremont	Kendallville	Ft. Wayne/New Haven	Lagrange	Auburn/Garrett
Decatur	Ft. Wayne/New Haven	Bluffton	Berne	Ossian	Huntington
Bluffton	Ossian	Ft. Wayne/New Haven	Decatur	Berne	Huntington
Huntington	Ft. Wayne/New Haven	Wabash	Bluffton	Columbia City	Berne
Warsaw/Winona Lake	Ft. Wayne/New Haven	Columbia City	Churubusco	N. Manchester	Wabash
Lagrange	Kendallville	Ft. Wayne/New Haven	Angola	Fremont	Auburn/Garrett
Wabash	Ft. Wayne/New Haven	N. Manchester	Huntington	Warsaw/Winona Lake	Columbia City
N. Manchester	Wabash	Ft. Wayne/New Haven	Warsaw/Winona Lake	Huntington	Columbia City
Columbia City	Ft. Wayne/New Haven	S. Whitley	Warsaw/Winona Lake	Churubusco	Huntington
Kendallville	Ft. Wayne/New Haven	Auburn/Garrett	Angola	Lagrange	Fremont
Huntertown	Ft. Wayne/New Haven	Auburn/Garrett	Leo-Cedarville	Columbia City	Churubusco
Leo-Cedarville	Ft. Wayne/New Haven	Auburn/Garrett	Huntertown	Butler	Warsaw/Winona Lake
Butler	Auburn/Garrett	Ft. Wayne/New Haven	Kendallville	Leo-Cedarville	Angola
Fremont	Angola	Kendallville	Ft. Wayne/New Haven	Auburn/Garrett	Lagrange
Berne	Ft. Wayne/New Haven	Decatur	Bluffton	Ossian	Huntington
Ossian	Bluffton	Ft. Wayne/New Haven	Berne	Huntington	Decatur
Churubusco	Ft. Wayne/New Haven	Auburn/Garrett	Columbia City	Warsaw/Winona Lake	Huntertown
S. Whitley	Ft. Wayne/New Haven	Columbia City	Warsaw/Winona Lake	Huntington	N. Manchester

5.2 | ORIGINS AND DESTINATIONS FOR KEY FACILITIES

Flows on key facilities in the region were also analyzed to understand the origins and destinations they serve. A total of 60 gates on 9 major corridors were defined as reported in Table 14 and shown in Figure 26, and for each gate an INRIX trip table was downloaded using the middle filter option with a pass-through zone, since the Streetlight webtool did not allow this type of analysis with the Cuebiq data.

TABLE 14: LOCATION OF PRIMARY AND SECONDARY GATES

Primary Gate	Secondary Gate
I-69 North of I-80/I-90	I-69 South of I-80/I-90
I-69 South of US 6	I-69 North of US 6
I-69 South of US 30	I-69 between SR 1 & I-469N
I-69 South of I-469S	I-69 South of I-469N
I-69 North of SR 18	I-69 North of US 24
I-469 East of I-69 N	I-69 North of I-469S
I-469 between US 30 & US 24	I-469 North of US 24
I-469 East of US 27	I-469 South of US 30
I-469 East of I-69S	I-469 West of US 27
I-80/I-90 East SR 13	I-80/I-90 East SR 9
I-80/I-90 West of IN/OH State Line	I-80/I-90 West SR 9
US 20 East of SR 13	US 20 East of Lagrange
US 20 West of I-69	US 20 East of I-69
US 20 West of IN/OH State Line	US 6 West of Kendallville
US 6 West of SR 5	US 6 West of I-69
US 6 East of Kendallville	US 6 East of I-69
US 6 West of IN/OH State Line	US 33 West of Churubusco
US 33 East of SR 5	US 30 West of Warsaw
US 33 West of US 30	US 30 West of Columbia City
US 33 West of IN/OH State Line	US 30 West of US 33
US 30 West of SR 19	US 30 East of US 33
US 30 East of Warsaw	US 30 East of I-69
US 30 East of Columbia City	US 30 West of I-469
US 30 East of US 27 (Lima Rd)	US 30 East of I-469
US 30 West of IN/OH State Line	US 24 East of Wabash
US 24 West of Wabash	US 24 West of I-69
US 24 West of Huntington	US 24 East of I-469
US 24 South of CR E900N	US 27 South of SR 930
US 24 West of IN/OH State Line	US 27 North of I-469
US 27 South of Geneva	US 27 South of I-469

Trip tables were then expanded based on the INRIX expansion factors obtained by sequential ODME for the entire region. A congested travel time based all-or-nothing assignment was then run for each trip table to map truck and auto flows through each gate.

FIGURE 26: LOCATION OF PRIMARY AND SECONDARY GATES

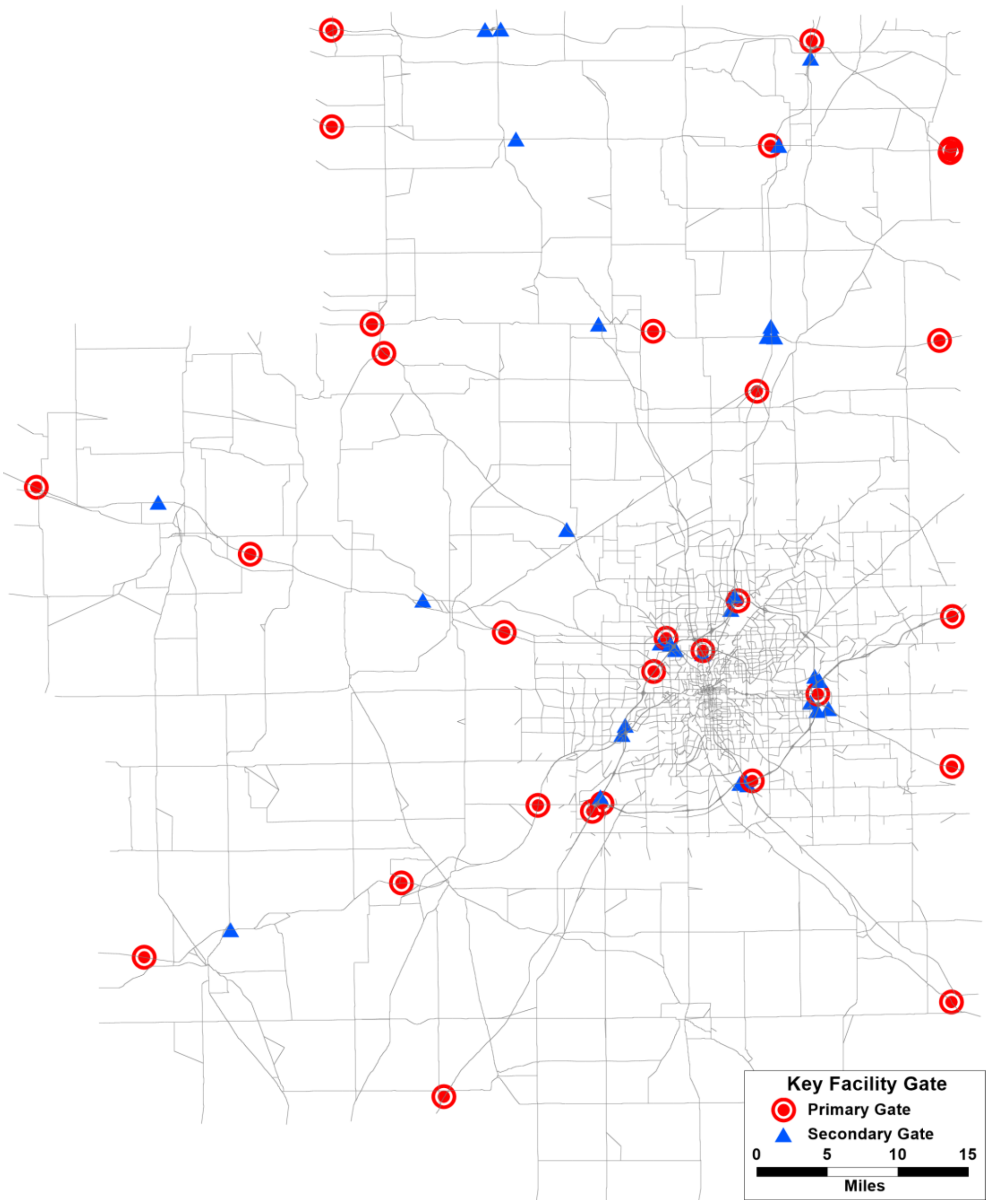
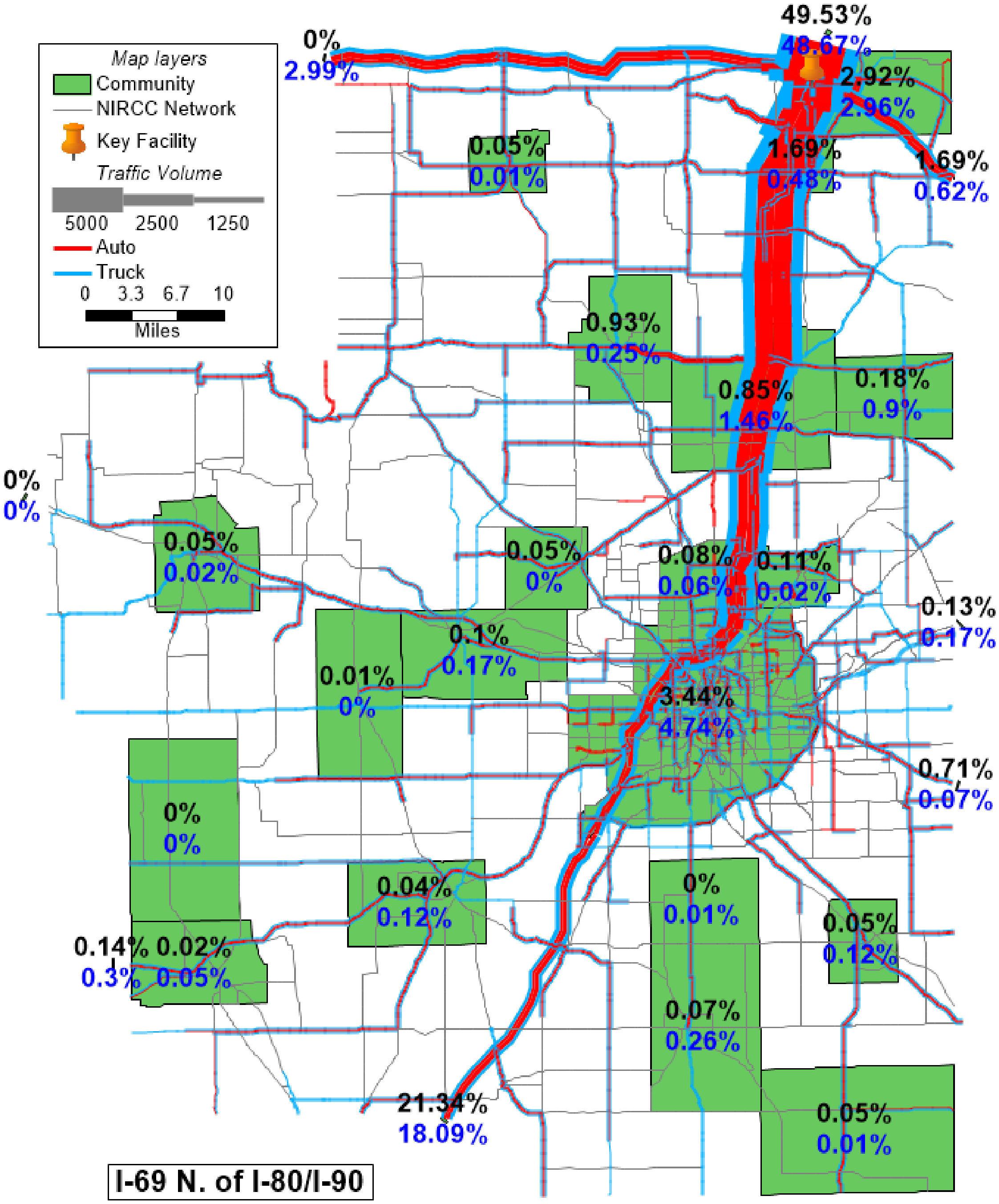


Figure 27 presents auto and truck volumes passing through the first primary gate which is on I-69 north of I-80/I-90 as well as the percentage of these flows to/from each community defined in Section 5.0 and major externals such as I-69 north and south, I-80/I-90 east and west, US30 east and west, and US24 east and west. The pin presents the location of the gate and auto and truck volumes/shares are shown with different colors.

FIGURE 27: LOADING VOLUME AND ORIGIN-DESTINATION OF A KEY FACILITY LOCATION BY VEHICLE CLASS



According to Figure 27, a plurality of trips (46%) are passing through the region via I-69 which is reasonable. It also shows that 3.4 percent of auto trips and 4.7 percent of truck trips passing through this gate are bound to/from Fort Wayne/New Haven – since half of the origins/destinations are at the external station, this means twice these numbers or about 7% of auto and 10% of truck trips to and from this external station involve Fort Wayne. Fort Wayne/New Haven has the highest share of auto and truck trips among communities in the region, consistent with common sense and the

Section 5.0 results. It should be mentioned that rural areas are also included in the community share calculation although their shares are not shown on the map. Maps for other primary and secondary gates were also generated which can be found in Appendix B.

To facilitate general understanding of trucking patterns, expanded truck trip table were also aggregated to the regions shown in Figure 28 with external zones assigned to five following categories:

1- Michigan
2- Ohio/I-80
3- Ohio/US-24
4- Indiana South
5- Indiana Southwest
6- Indiana West

The resulting aggregate trip table is reported in Table 15.

FIGURE 28: TRUCK REGION TO AGGREGATE EXPANDED TRUCK TRIPS

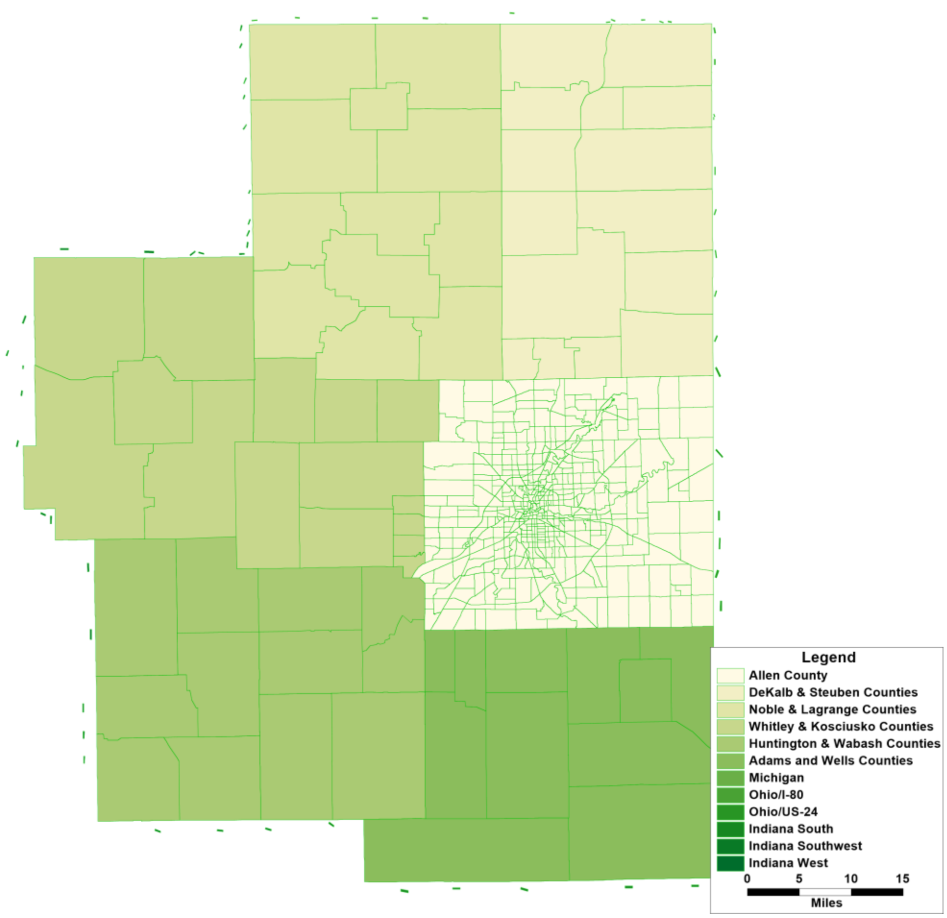


TABLE 15: AGGREGATED TRUCK TRIPS BASED ON EXPANDED INRIX DATA (SEQUENTIAL ODME)

County	Allen	DeKalb & Steuben	Noble & LaGrange	Whitley & Kosciusko	Huntington & Wabash	Adams & Wells	Michigan	Ohio / I-80	Ohio/ US-24	Indiana South	Indiana Southwest	Indiana West
Allen	20,593	1,531	464	1,839	2,134	1,533	507	63	3,167	1,402	210	1,514
DeKalb & Steuben	1,582	6,990	1,011	72	145	48	1,922	3,781	553	583	65	1,041
Noble & LaGrange	454	1,030	8,154	141	53	19	304	1,535	92	146	14	3,849
Whitley & Kosciusko	1,844	73	145	6,653	1,527	74	52	42	341	64	107	2,887
Huntington & Wabash	1,966	129	49	1,416	4,774	208	114	16	379	770	730	349
Adams & Wells	1,416	44	18	76	202	3,047	42	3	648	824	48	183
Michigan	474	1,819	297	51	105	39	0	97	35	1,414	127	983
Ohio/I-80	63	3,781	1,535	42	16	3	78	0	0	12	1	678
Ohio/US-24	3,167	553	92	341	379	648	30	0	0	1,316	192	812
Indiana South	1,516	632	159	69	829	843	1,556	11	1,350	0	7	683
Indiana Southwest	210	65	14	107	730	48	116	1	165	6	36	151
Indiana West	1,514	1,041	3,849	2,887	349	183	959	659	623	781	110	287

All expanded trip tables for all gates are also summarized in Table 16 and Table 17 reporting number of flows in the following five categories:

1- Trips passing through the region without stopping
2- Trips to and from the region (with one origin or destination in the region and the other outside the region)
3- Trips within Fort Wayne/New Haven
4- Trips between Fort Wayne/New Haven and other parts of the region
5- Trips within the region not involving Fort Wayne/New Haven

Table 16 summarizes primary gates' flows and Table 17 summarizes secondary gates' flows. The top 20 gates serving trips inside the region, trips to and from the region, and trips through the region were also reported in Table 18 to Table 20 and shown in Figure 29 to Figure 31.

The results show how different corridors and locations, within those corridors, function differently and are important to different travel markets. As expected, I-69 is clearly the most important facility for trips through the region. However, the data also reveals some potentially less obvious facts, such as that I-469 (between I-69 south and US 24) and US 24 (east of I-469) are together the second most important corridor for trips passing through the region – just slightly ahead of the I-80/90 Indiana Toll Road. I-69, US 30, and US 24 are also confirmed as the key facilities serving trips within the region, perhaps unsurprisingly, while a perhaps more surprisingly diverse list of facilities play an important role in serving trips to and from the region (although some of these may be serving short trips across the study area boundary). Further study of these results should be fruitful in support of regional planning.

TABLE 16: SUMMARY OF PRIMARY GATES' TRIPS

Gate Location	Through Trips	To/From the Region	Inside Ft. Wayne/New Haven	To/From Ft. Wayne/New Haven	Inside the Region without Ft. Wayne/New Haven
I-69 N. of I-80/I-90	11,083	9,522	4	14	120
I-69 S. of US 6	9,194	9,807	65	7,552	8,919
I-69 S. of US 30	9,947	9,548	36,770	17,596	3,752
I-69 S. of I-469S	14,643	11,433	184	4,403	1,586
I-69 N. of SR 18	14,722	13,503	18	45	68
I-469 E. of I-69 N	1,709	7,995	22,156	12,232	3,854
I-469 between US 30 & US 24	7,327	10,825	5,357	8,931	4,821
I-469 E. of US 27	5,676	6,979	2,287	5,203	2,572
I-469 E. of I-69S	5,489	5,275	1,800	8,301	3,045
I-80/I-90 E. SR 13	5,242	3,727	0	2	71
I-80/I-90 W. of IN/OH State Line	3,325	16,846	0	2	115
US 20 E. of SR 13	355	5,159	4	5	293
US 20 W. of I-69	619	2,144	5	395	4,464
US 20 W. of IN/OH State Line	341	2,868	2	1	38
US 6 W. of SR 5	1,555	8,594	10	253	3,094
US 6 E. of Kendallville	715	1,812	4	736	6,002
US 6 W. of IN/OH State Line	229	2,930	0	5	96
US 33 E. of SR 5	1,333	4,294	11	1,157	1,488
US 33 W. of US 30	1,174	2,824	7,241	7,893	1,195
US 33 W. of IN/OH State Line	324	3,153	0	49	150
US 30 W. of SR 19	1,446	8,572	6	10	52
US 30 E. of Warsaw	1,424	4,274	36	4,459	11,297
US 30 E. of Columbia City	1,299	4,114	138	13,270	16,966
US 30 E. of US 27	6	809	36,126	7,265	403
US 30 W. of IN/OH State Line	1,754	11,564	24	87	129
US 24 W. of Wabash	1,057	6,285	6	17	272
US 24 W. of Huntington	758	3,166	13	2,300	5,517
US 24 S. of CR E900N	626	2,592	117	9,829	4,437
US 24 W. of IN/OH State Line	5,766	6,595	21	39	64
US 27 S. of Geneva	172	4,486	4	20	49

TABLE 17: SUMMARY OF SECONDARY GATES' TRIPS

Gate Location	Through Trips	To/From the Region	Inside Ft. Wayne/New Haven	To/From Ft. Wayne/New Haven	Inside the Region without Ft. Wayne/New Haven
I-69 S. of I-80/I-90	9,331	11,120	4	457	3,200
I-69 N. of US 6	9,744	10,341	37	4,162	8,873
I-69 between SR 1 & I-469N	9,070	9,206	25,215	21,885	4,914
I-69 S. of I-469N	10,032	10,095	29,482	22,205	3,595
I-69 N. of US 24	9,788	9,403	22,862	14,658	3,716
I-69 N. of I-469S	9,844	9,687	5,902	10,598	2,533
I-469 N. of US 24	2,040	8,474	6,627	9,771	4,761
I-469 S. of US 30	5,534	5,399	2,726	4,955	3,038
I-469 W. of US 27	5,674	5,811	1,840	5,517	2,541
I-80/I-90 E. SR 9	4,713	5,250	0	30	995
I-80/I-90 W. SR 9	5,411	4,253	0	12	794
US 20 E. of Lagrange	557	1,854	5	237	6,690
US 20 E. of I-69	202	1,944	7	769	5,541
US 6 W. of Kendallville	844	3,613	9	2,209	8,889
US 6 W. of I-69	739	2,052	4	1,263	6,387
US 6 E. of I-69	440	2,220	23	1,887	5,790
US 33 W. of Churubusco	1,218	3,131	37	3,394	3,420
US 30 W. of Warsaw	1,268	8,001	7	392	3,763
US 30 W. of Columbia City	1,203	3,811	48	6,259	10,390
US 30 W. of US 33	1,414	4,398	10,409	18,313	3,526
US 30 E. of US 33	2,396	7,072	17,140	22,810	4,246
US 30 E. of I-69	130	3,933	25,353	19,373	683
US 30 W. of I-469	35	2,446	4,982	8,957	267
US 30 E. of I-469	1,580	10,285	80	4,204	1,802
US 24 E. of Wabash	758	3,300	9	1,794	4,315
US 24 W. of I-69	308	2,042	26,655	11,484	1,338
US 24 E. of I-469	5,756	5,699	200	6,128	1,890
US 27 S. of SR 930	64	649	26,637	5,964	239
US 27 N. of I-469	152	1,366	1,375	5,553	210
US 27 S. of I-469	432	4,405	232	6,285	1,776

TABLE 18: TOP 20 GATES SERVING TRIPS INSIDE THE REGION

Rank	Daily Trips	Facility
1	58,118	I-69 S. of US 30
2	55,282	I-69 S. of I-469N
3	52,013	I-69 between SR 1 & I-469N
4	45,409	US 30 E. of I-69
5	44,195	US 30 E. of US 33
6	43,794	US 30 E. of US 27 (Lima Rd)
7	41,237	I-69 N. of US 24
8	39,478	US 24 W. of I-69
9	38,241	I-469 E. of I-69 N
10	32,840	US 27 S. of SR 930
11	32,247	US 30 W. of US 33
12	30,373	US 30 E. of Columbia City
13	21,159	I-469 N. of US 24
14	19,110	I-469 between US 30 & US 24
15	19,033	I-69 N. of I-469S
16	16,697	US 30 W. of Columbia City
17	16,537	I-69 S. of US 6
18	16,329	US 33 W. of US 30
19	15,793	US 30 E. of Warsaw
20	14,383	US 24 S. of CR E900N

FIGURE 29: TOP 20 GATES SERVING TRIPS INSIDE THE REGION

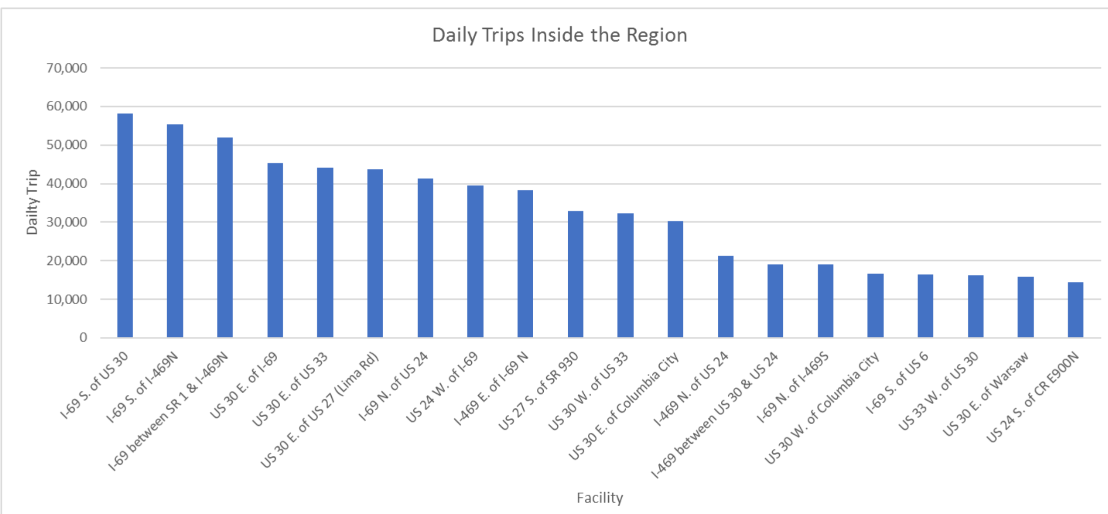


TABLE 19: TOP 20 GATES SERVING TRIPS TO AND FROM THE REGION

Rank	Daily Trips	Facility
1	16,846	I-80/I-90 W. of IN/OH State Line
2	13,503	I-69 N. of SR 18
3	11,564	US 30 W. of IN/OH State Line
4	11,433	I-69 S. of I-469S
5	11,120	I-69 S. of I-80/I-90
6	10,825	I-469 between US 30 & US 24
7	10,341	I-69 N. of US 6
8	10,285	US 30 E. of I-469
9	10,095	I-69 S. of I-469N
10	9,807	I-69 S. of US 6
11	9,687	I-69 N. of I-469S
12	9,548	I-69 S. of US 30
13	9,522	I-69 N. of I-80/I-90
14	9,403	I-69 N. of US 24
15	9,206	I-69 between SR 1 & I-469N
16	8,594	US 6 W. of SR 5
17	8,572	US 30 W. of SR 19
18	8,474	I-469 N. of US 24
19	8,001	US 30 W. of Warsaw
20	7,995	I-469 E. of I-69 N

FIGURE 30: TOP 20 GATES SERVING TRIPS TO AND FROM THE REGION

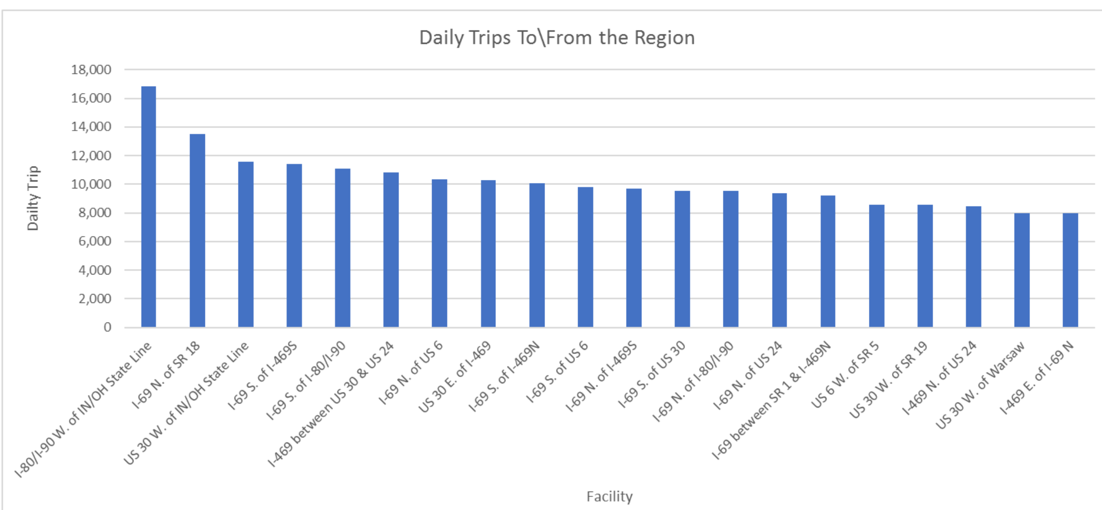
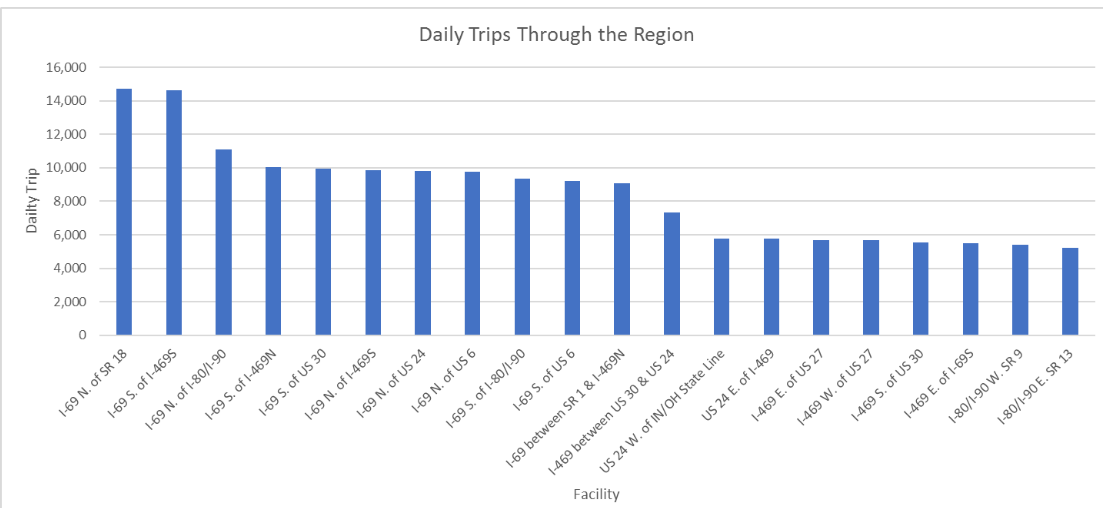


TABLE 20: TOP 20 GATES SERVING TRIPS THROUGH THE REGION

Rank	Daily Trips	Facility
1	14,722	I-69 N. of SR 18
2	14,643	I-69 S. of I-469S
3	11,083	I-69 N. of I-80/I-90
4	10,032	I-69 S. of I-469N
5	9,947	I-69 S. of US 30
6	9,844	I-69 N. of I-469S
7	9,788	I-69 N. of US 24
8	9,744	I-69 N. of US 6
9	9,331	I-69 S. of I-80/I-90
10	9,194	I-69 S. of US 6
11	9,070	I-69 between SR 1 & I-469N
12	7,327	I-469 between US 30 & US 24
13	5,766	US 24 W. of IN/OH State Line
14	5,756	US 24 E. of I-469
15	5,676	I-469 E. of US 27
16	5,674	I-469 W. of US 27
17	5,534	I-469 S. of US 30
18	5,489	I-469 E. of I-69S
19	5,411	I-80/I-90 W. SR 9
20	5,242	I-80/I-90 E. SR 13

FIGURE 31: TOP 20 GATES SERVING TRIPS THROUGH THE REGION



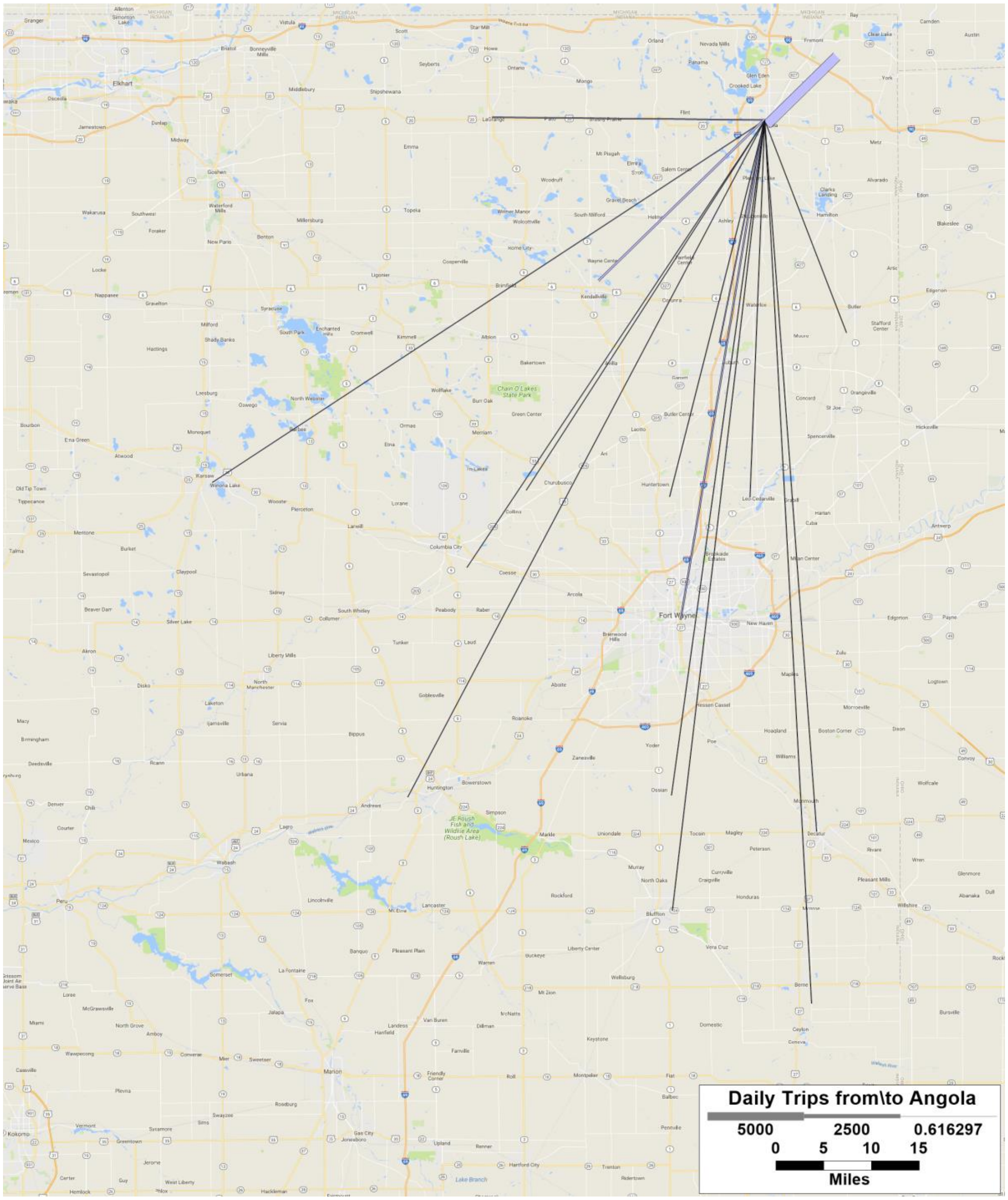


FIGURE 32. DESIRE LINES FOR ANGOLA

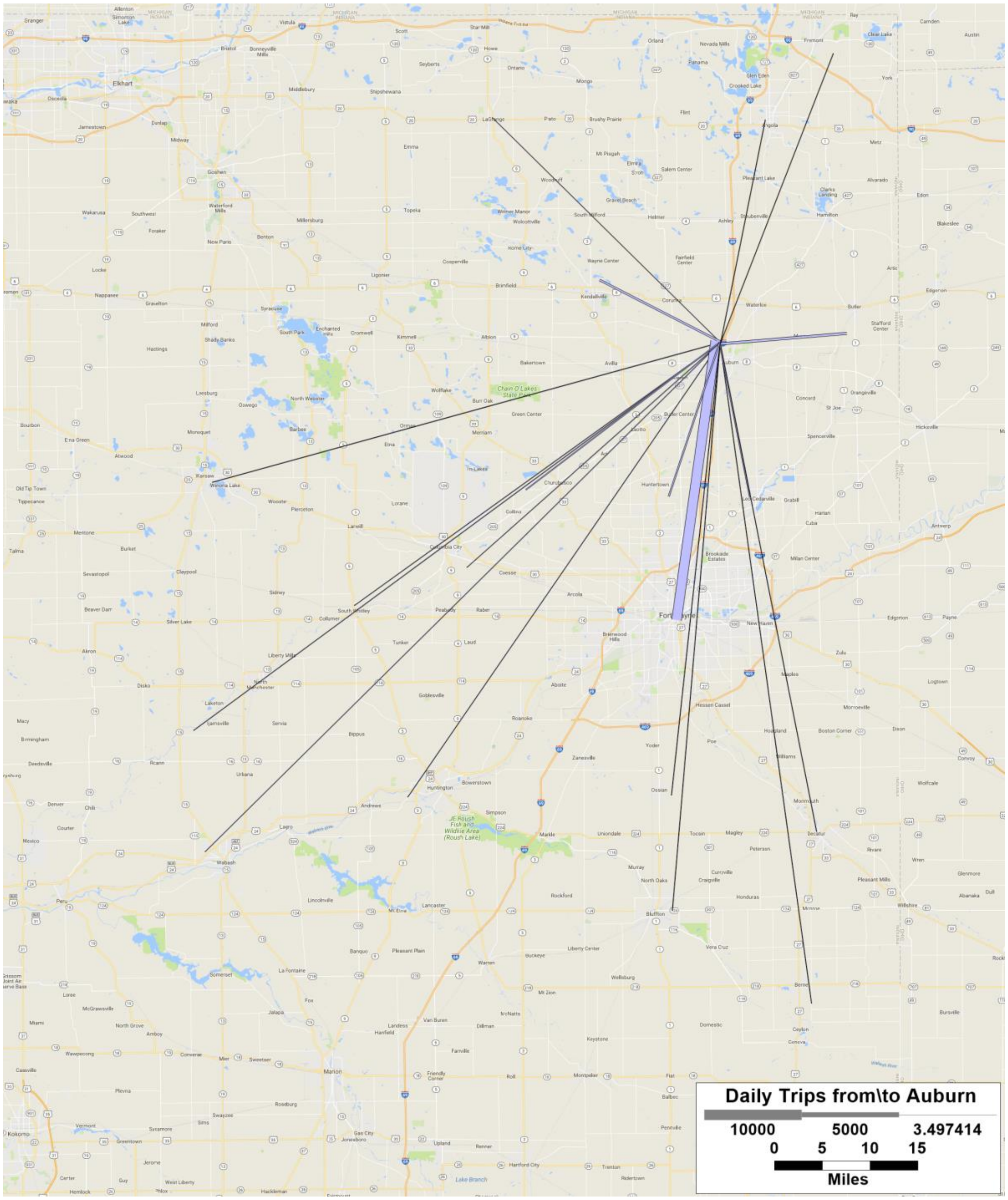


FIGURE 33. DESIRE LINES FOR AUBURN

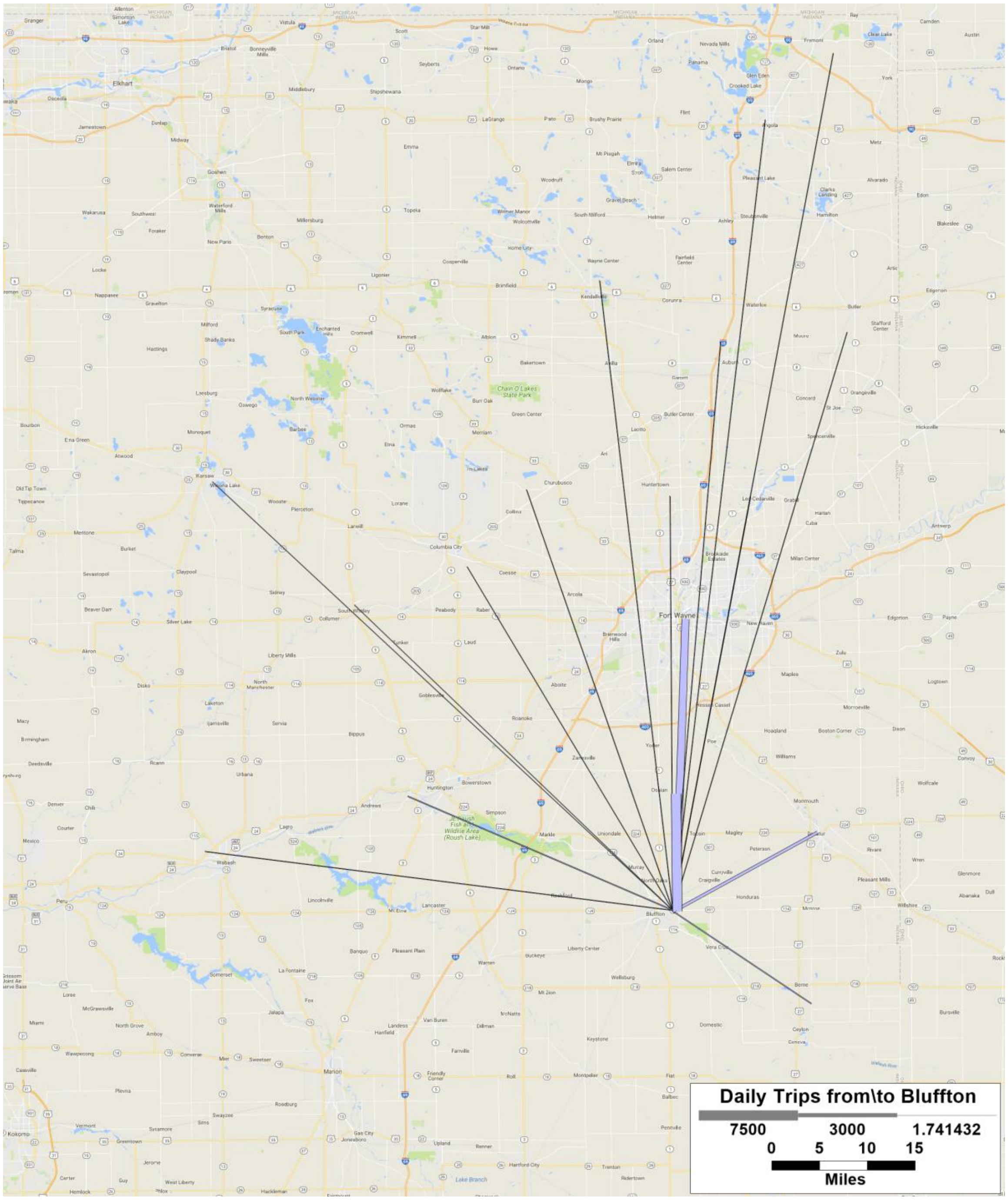


FIGURE 34. DESIRE LINES FOR BLUFFTON

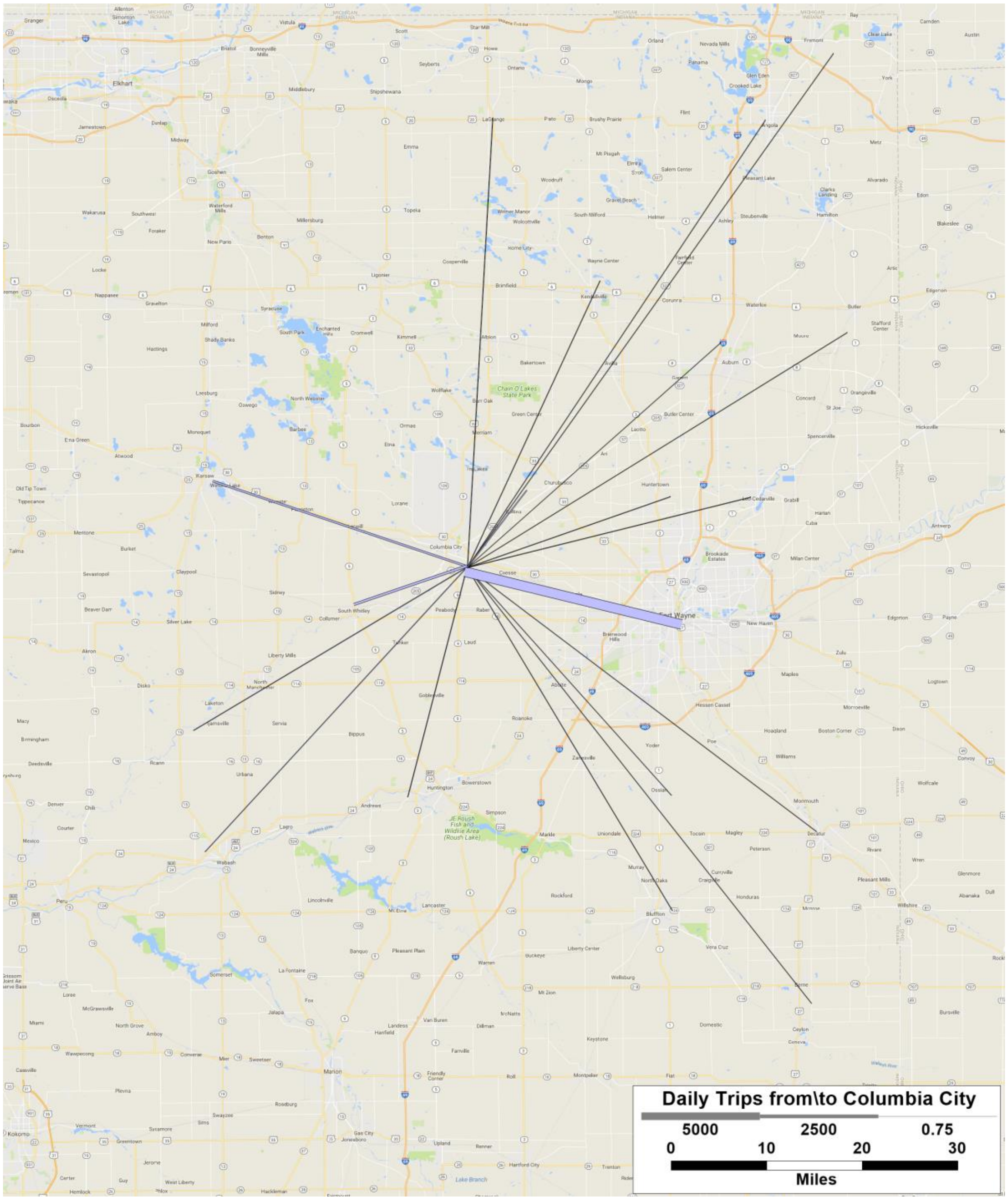


FIGURE 35. DESIRE LINES FOR COLUMBIA CITY

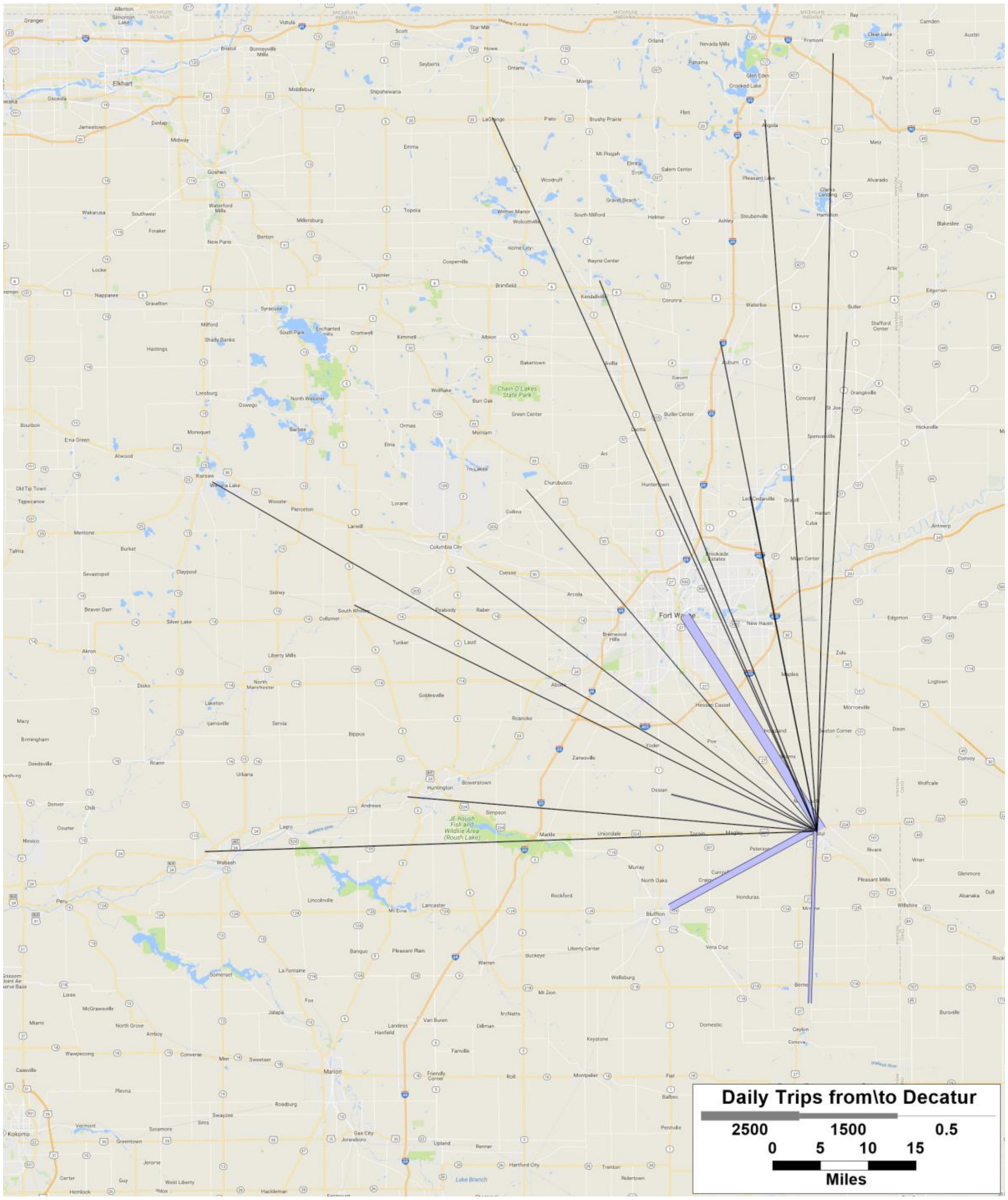


FIGURE 36. DESIRE LINES FOR DECATUR

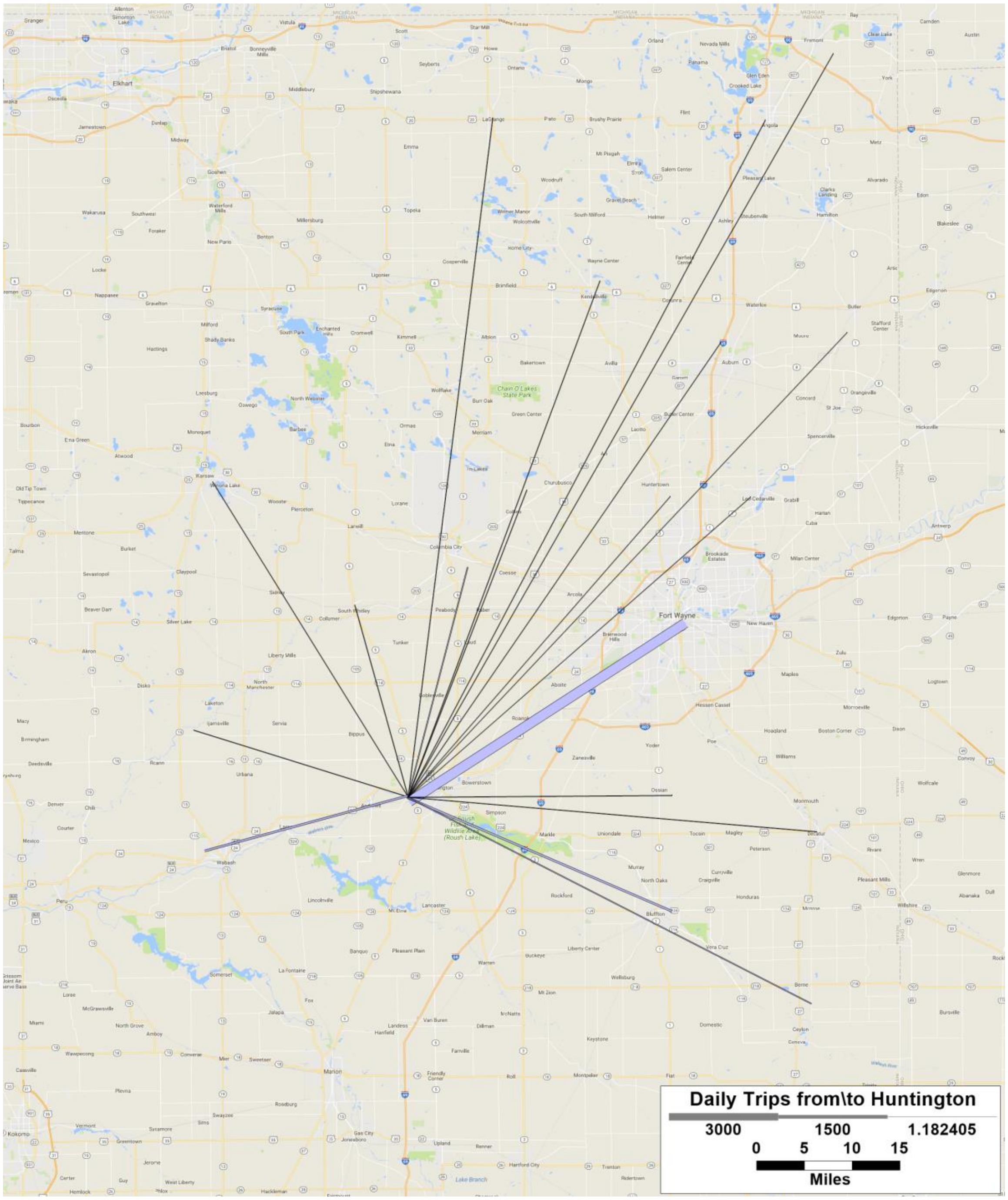


FIGURE 37. DESIRE LINES FOR HUNTINGTON

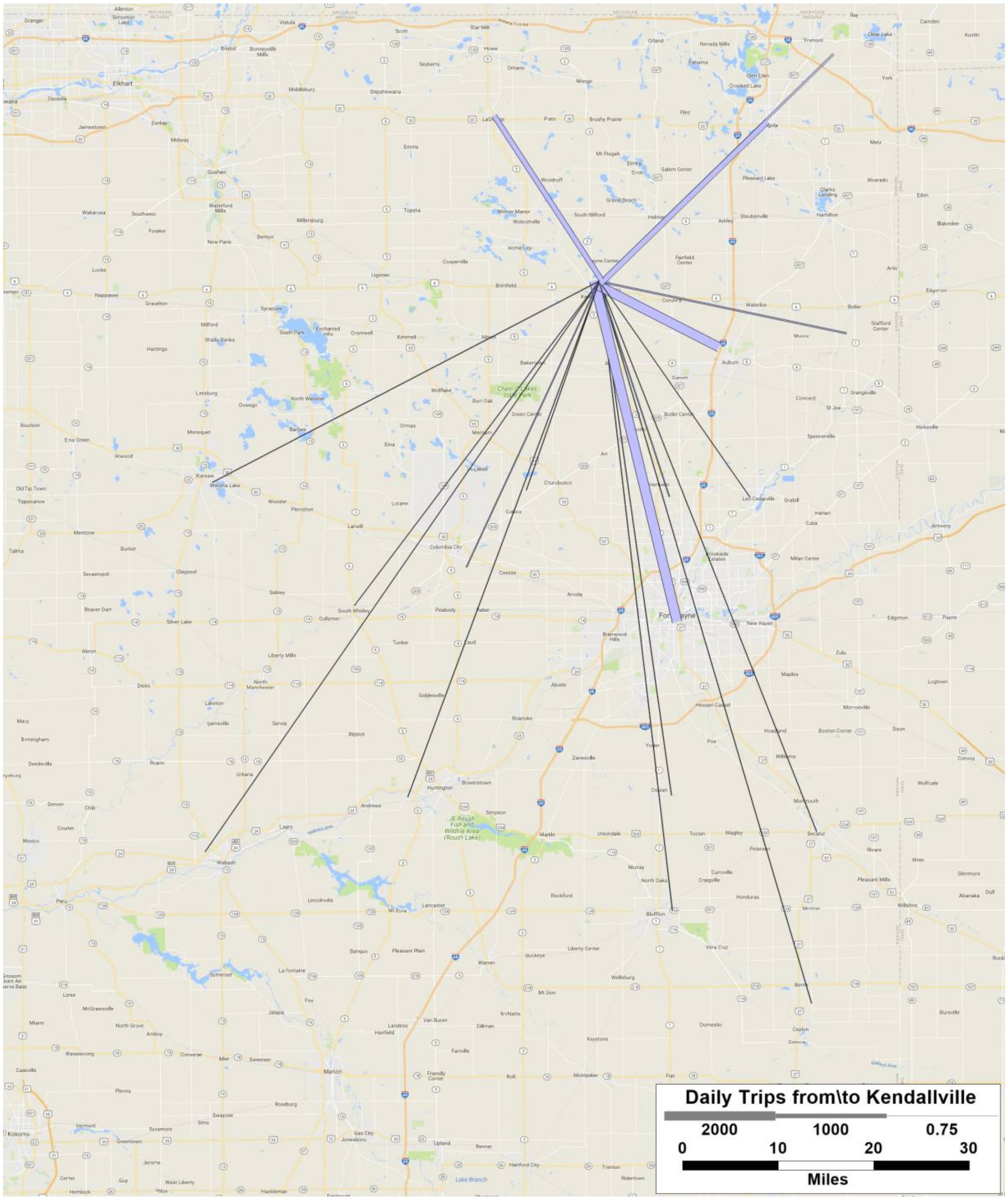


FIGURE 38. DESIRE LINES FOR KENDALLVILLE

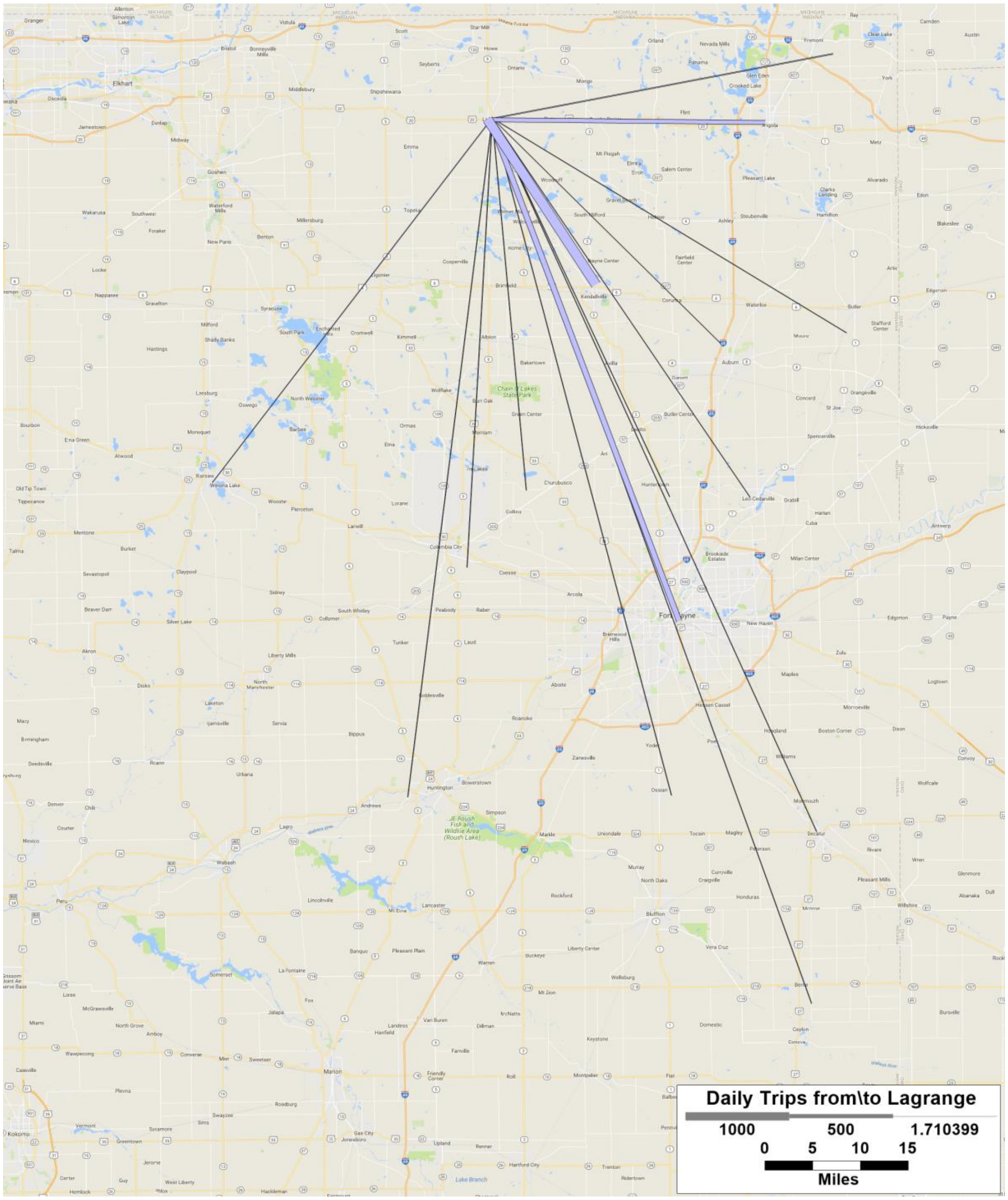


FIGURE 39. DESIRE LINES FOR LAGRANGE

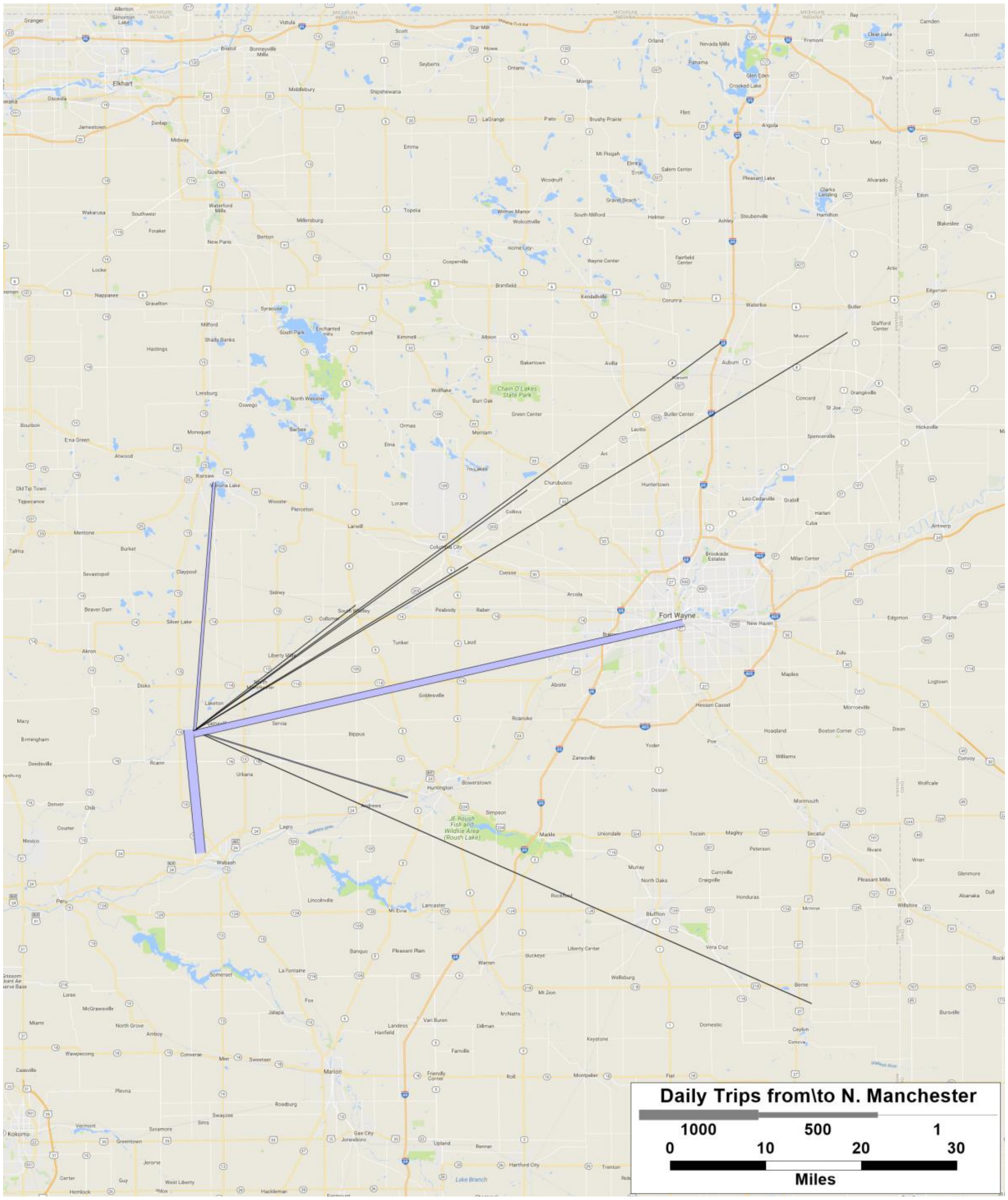


FIGURE 40. DESIRE LINES FOR N. MANCHESTER

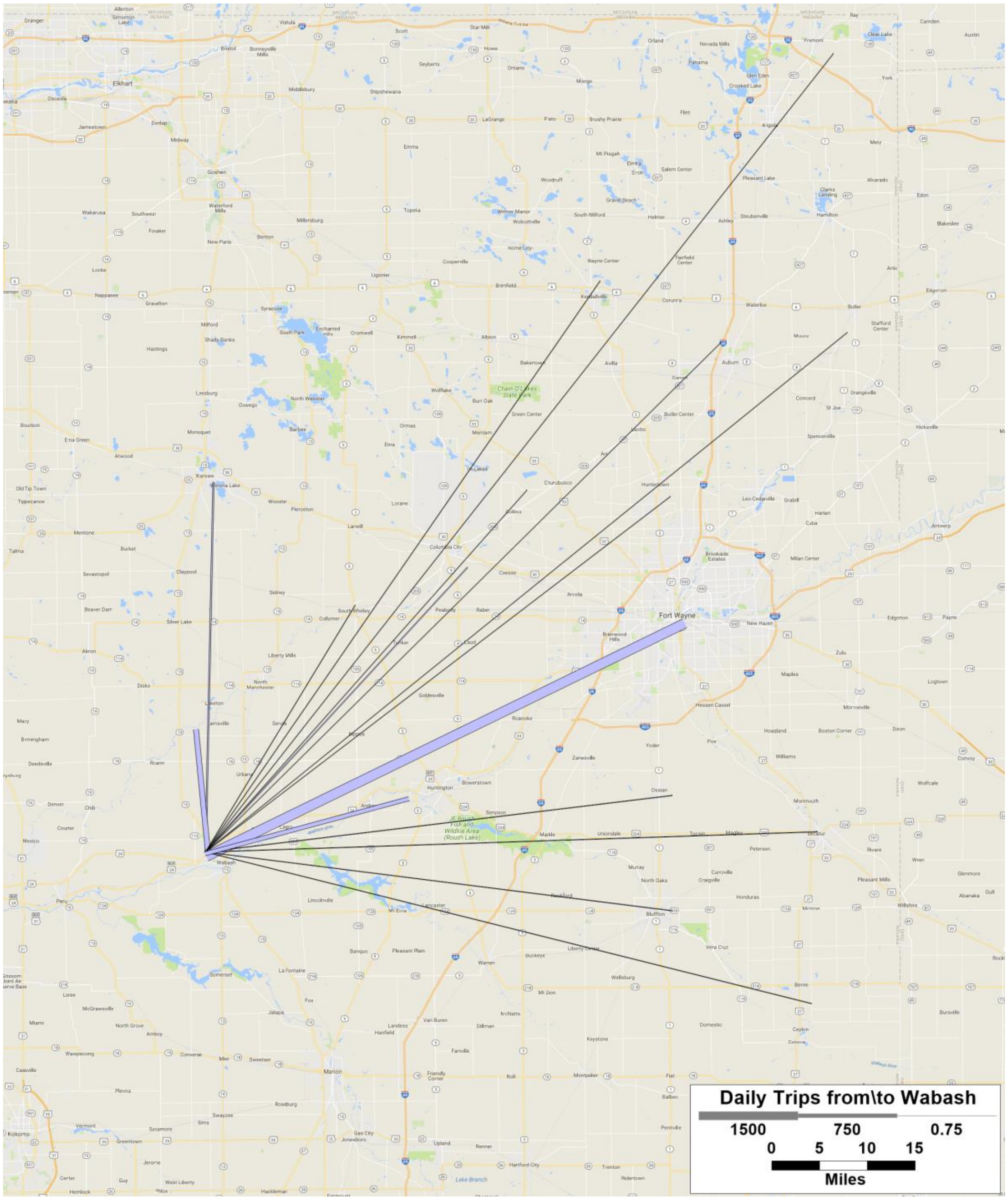


FIGURE 41. DESIRE LINES FOR WABASH

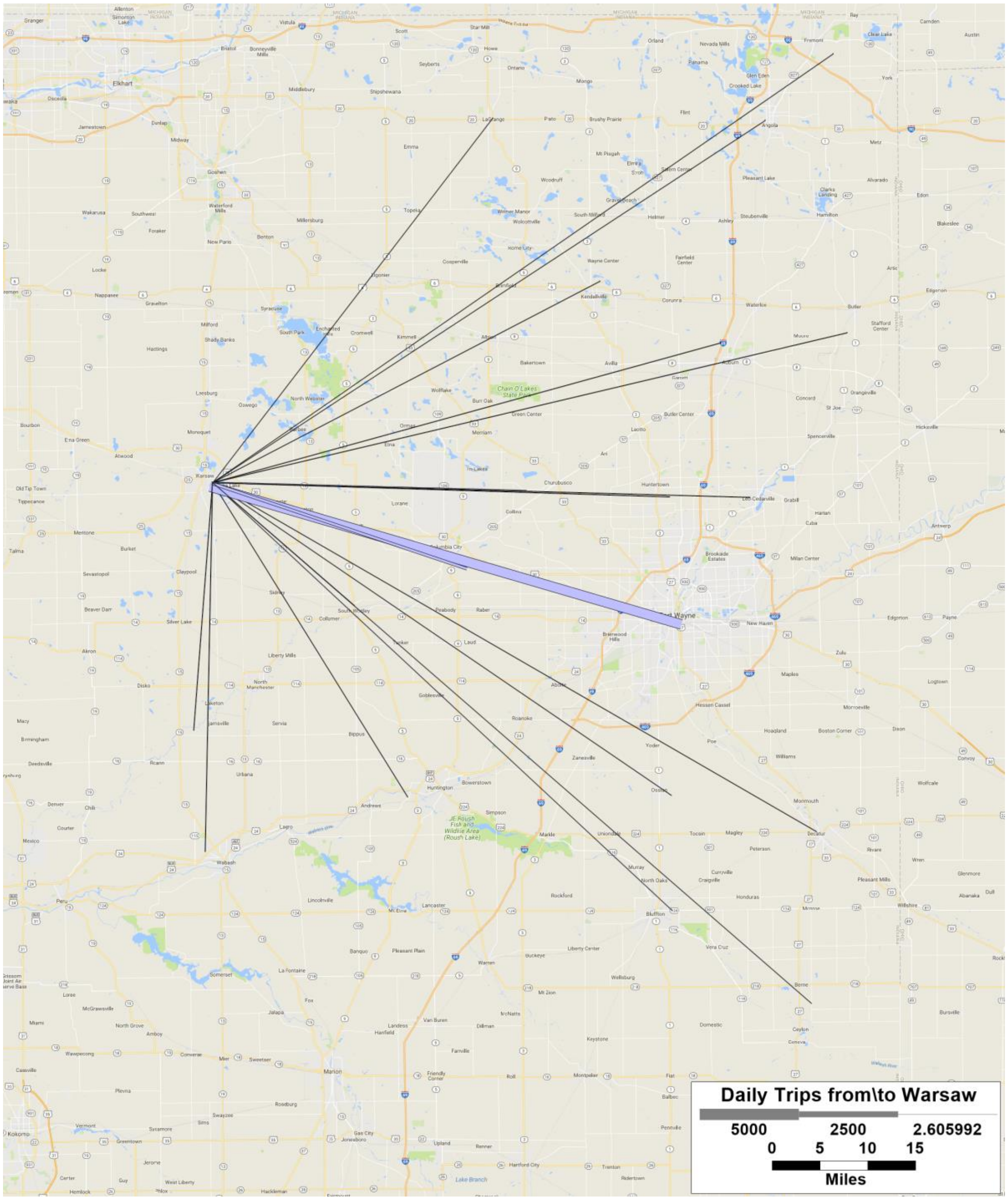


FIGURE 42. DESIRE LINES FOR WARSAW

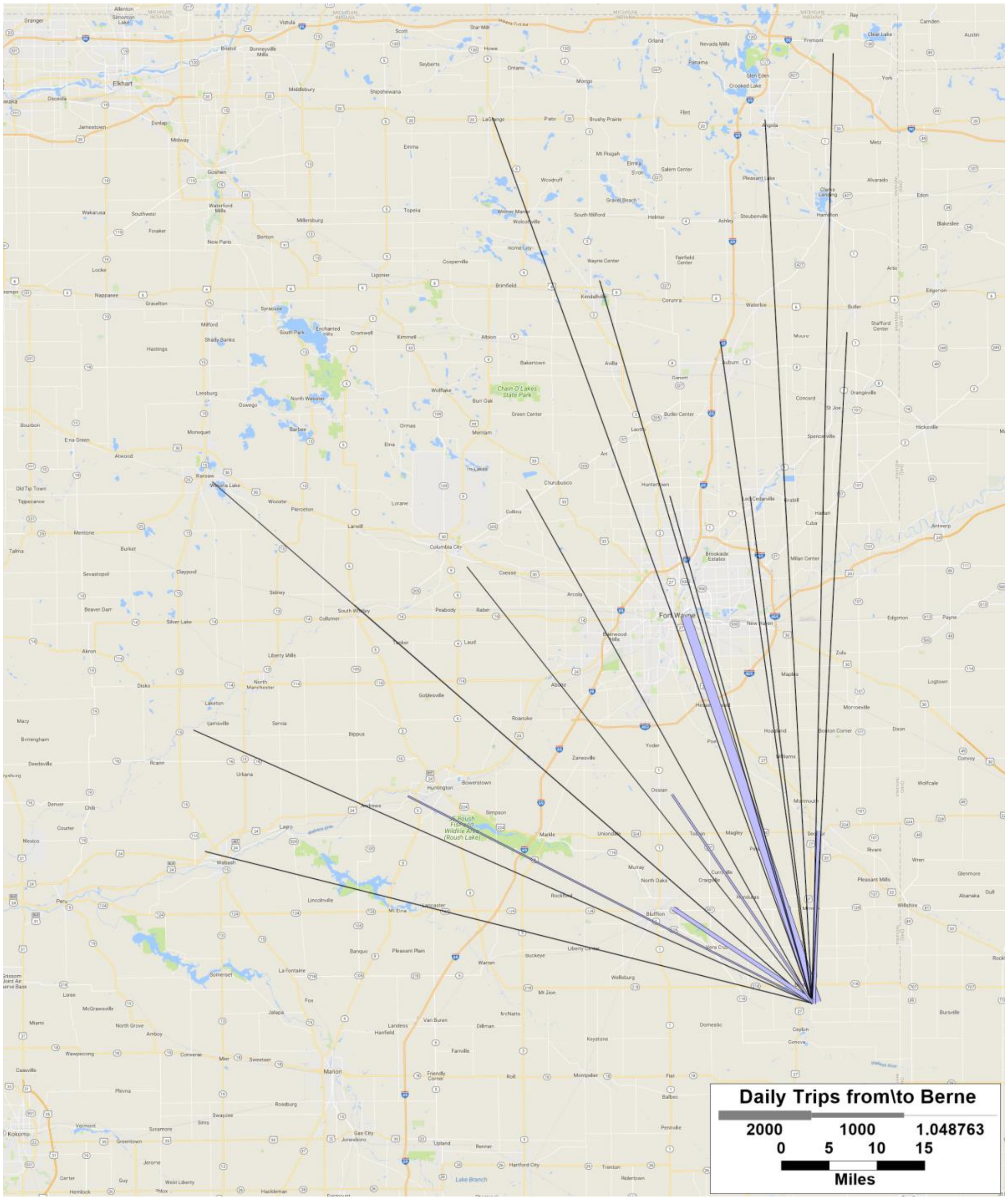


FIGURE 43. DESIRE LINES FOR BERNE

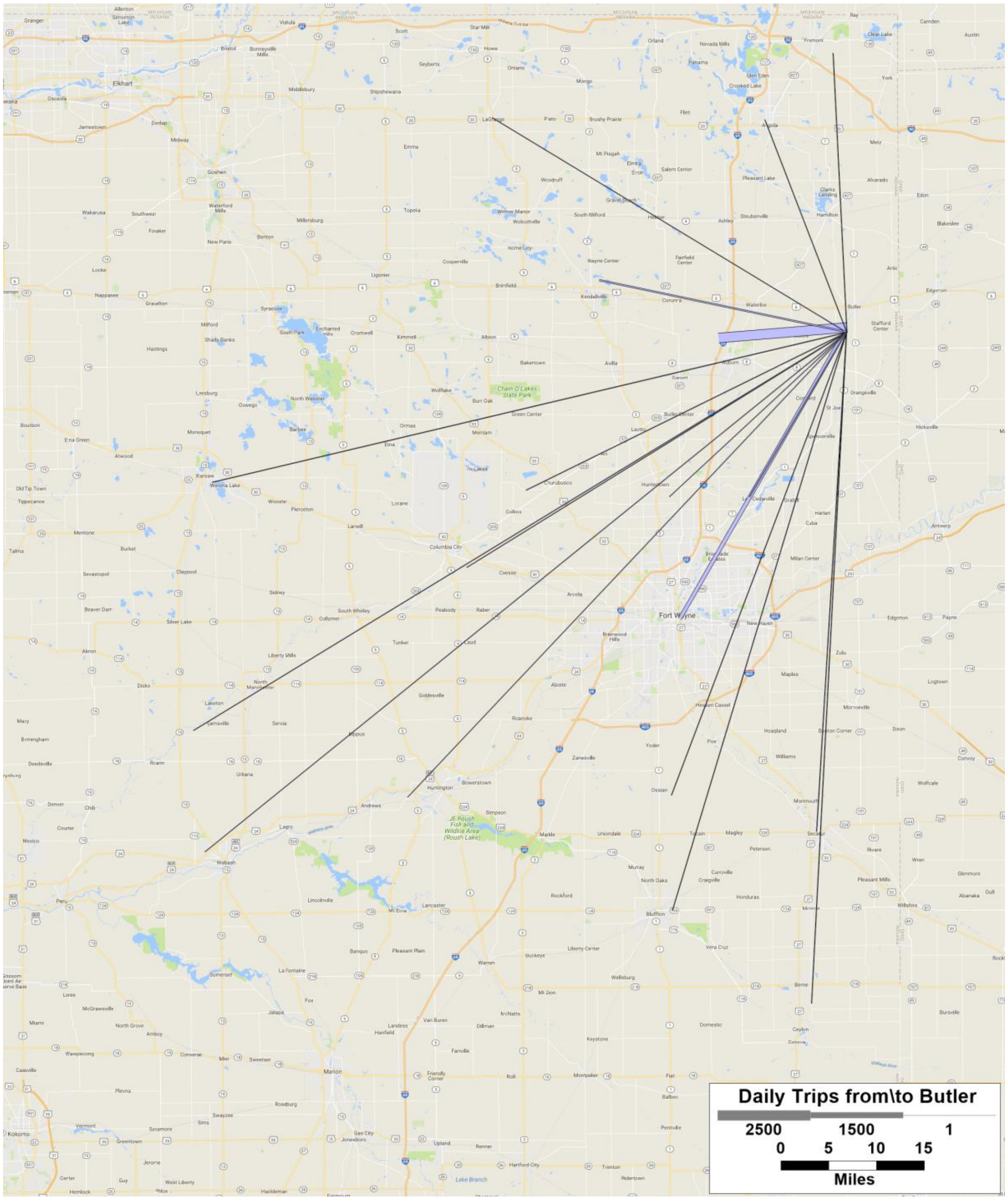


FIGURE 44. DESIRE LINES FOR BUTLER

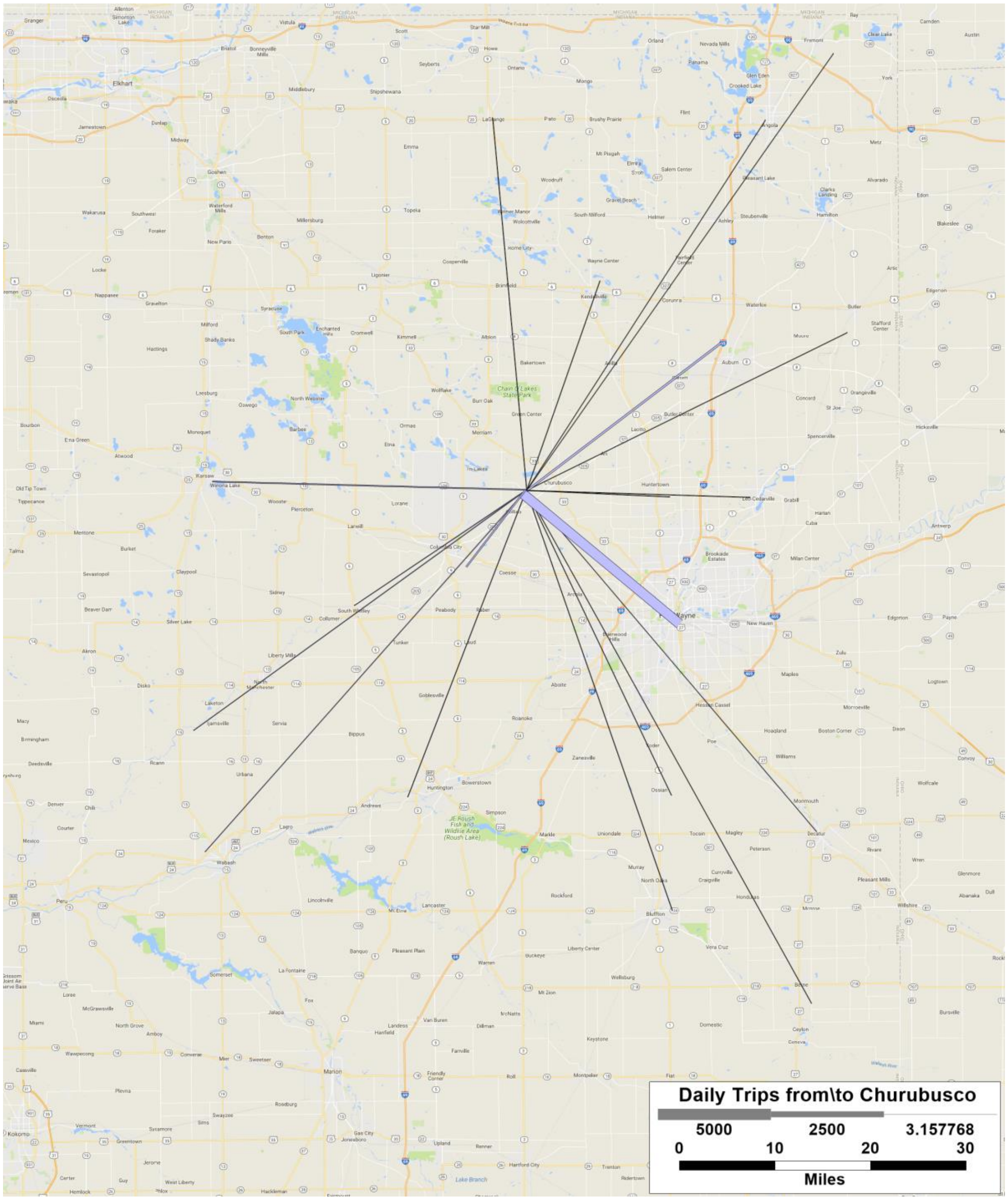


FIGURE 45. DESIRE LINES FOR CHURRUBUSCO

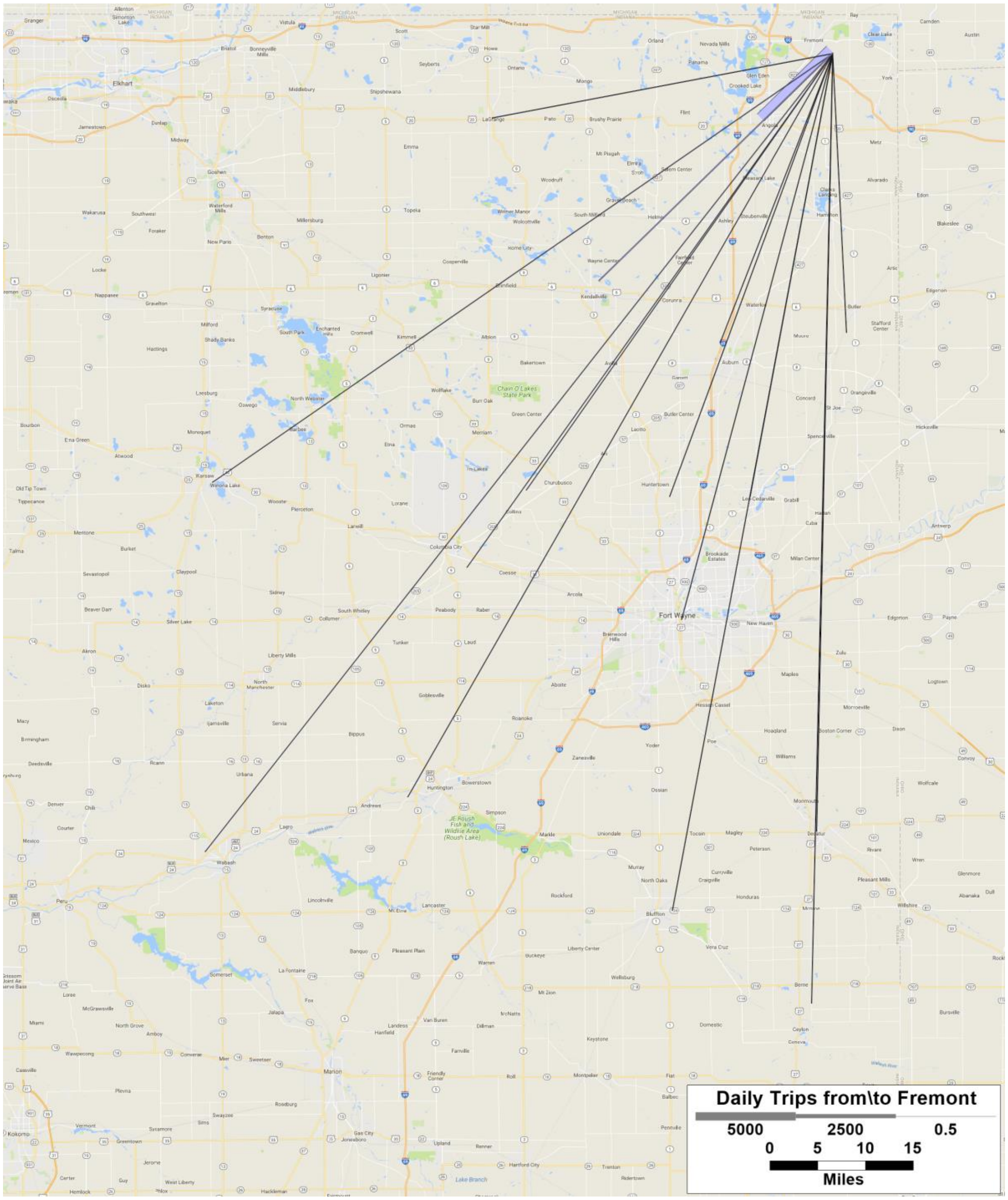


FIGURE 46. DESIRE LINES FOR FREMONT

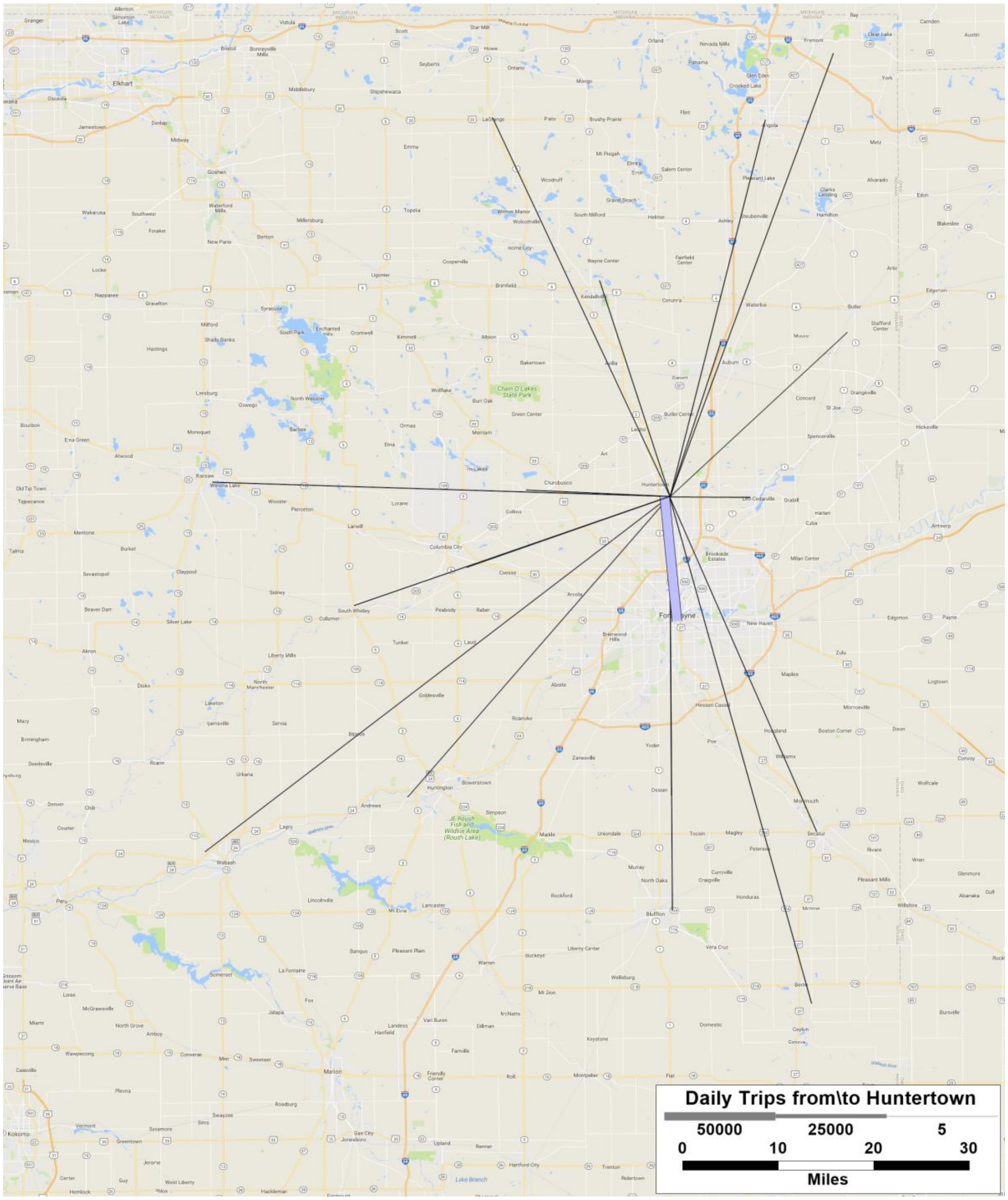


FIGURE 47. DESIRE LINES FOR HUNTERTOWN

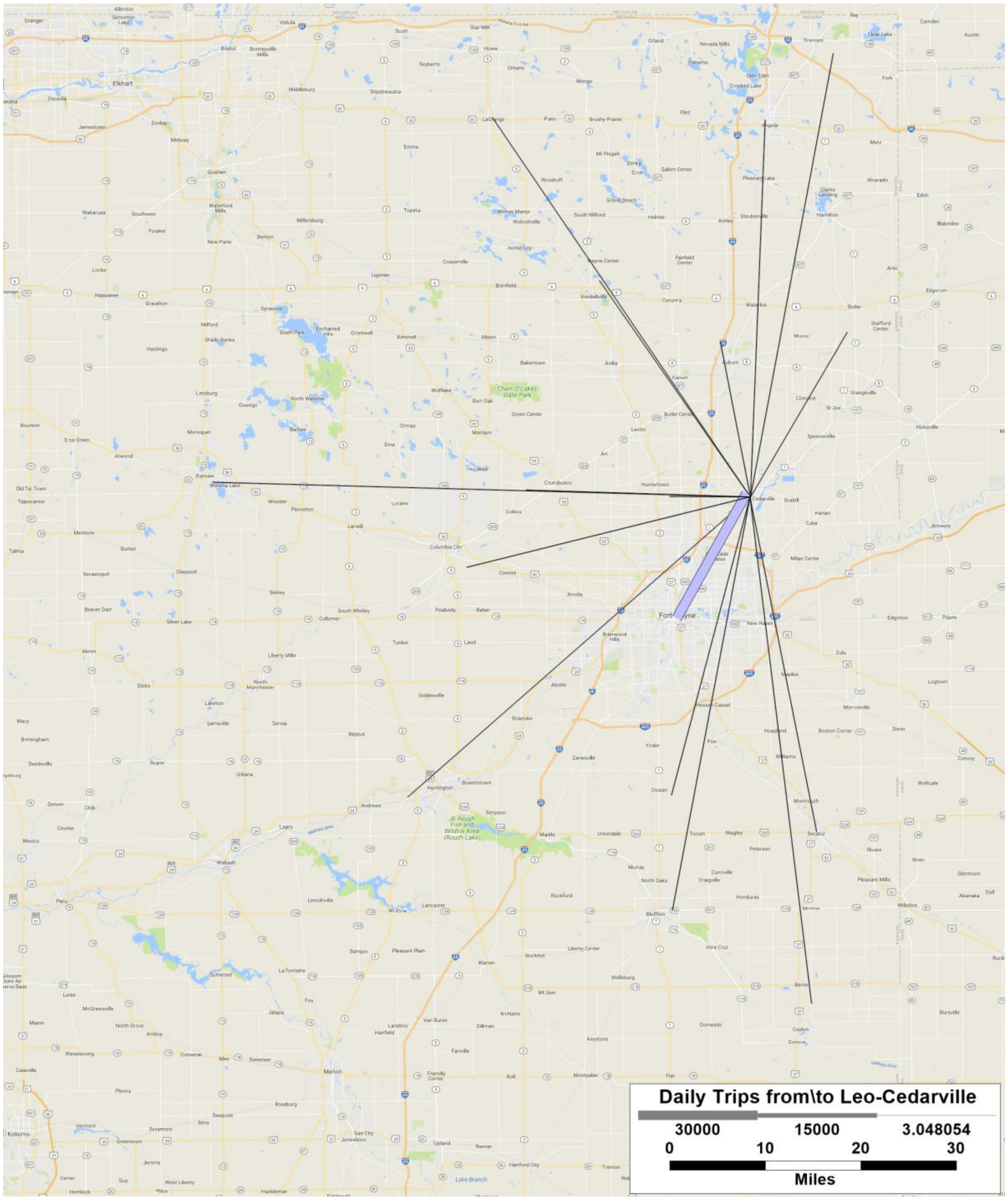


FIGURE 48. DESIRE LINES FOR LEO-CEDARVILLE

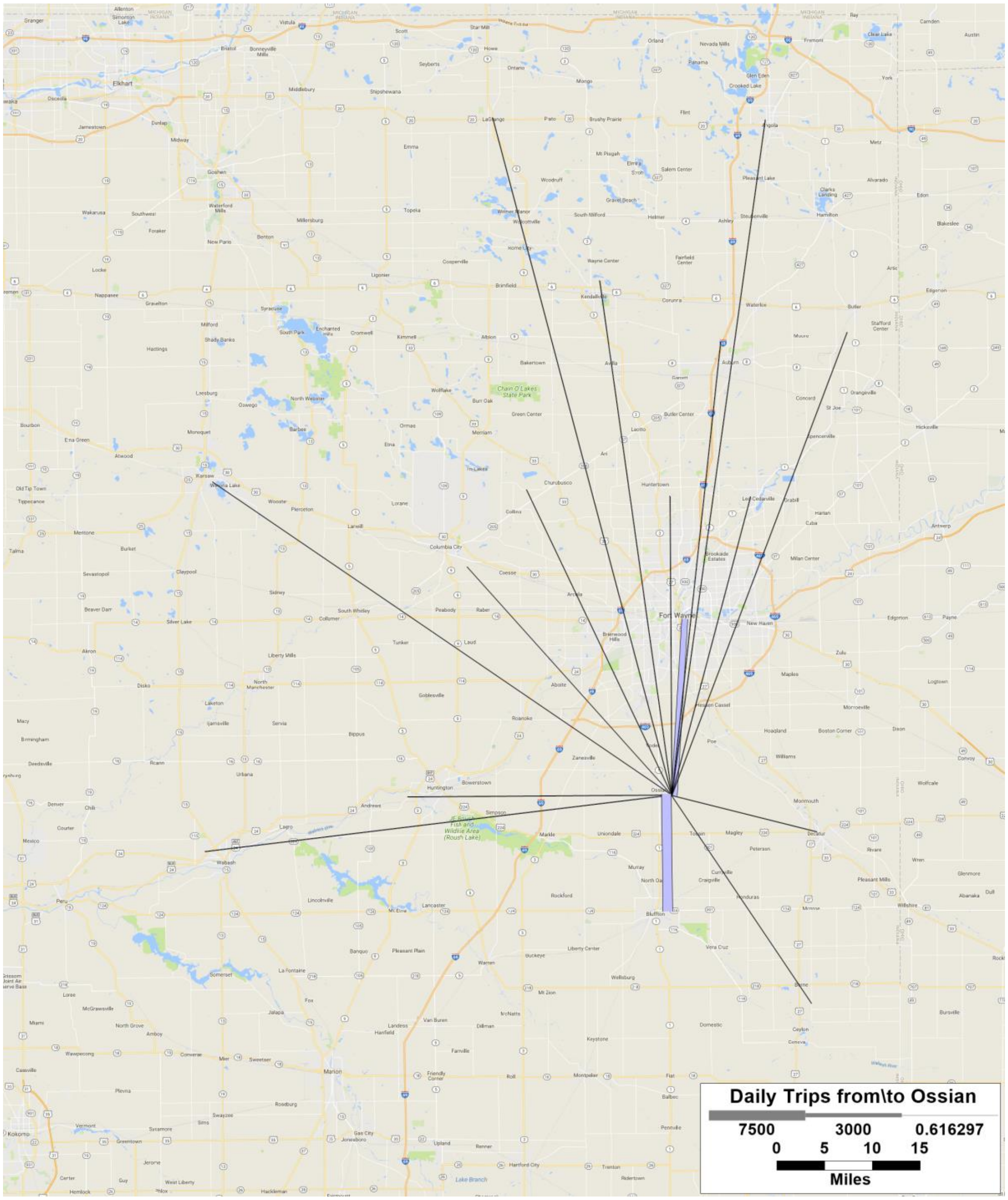


FIGURE 49. DESIRE LINES FOR OSSIAN

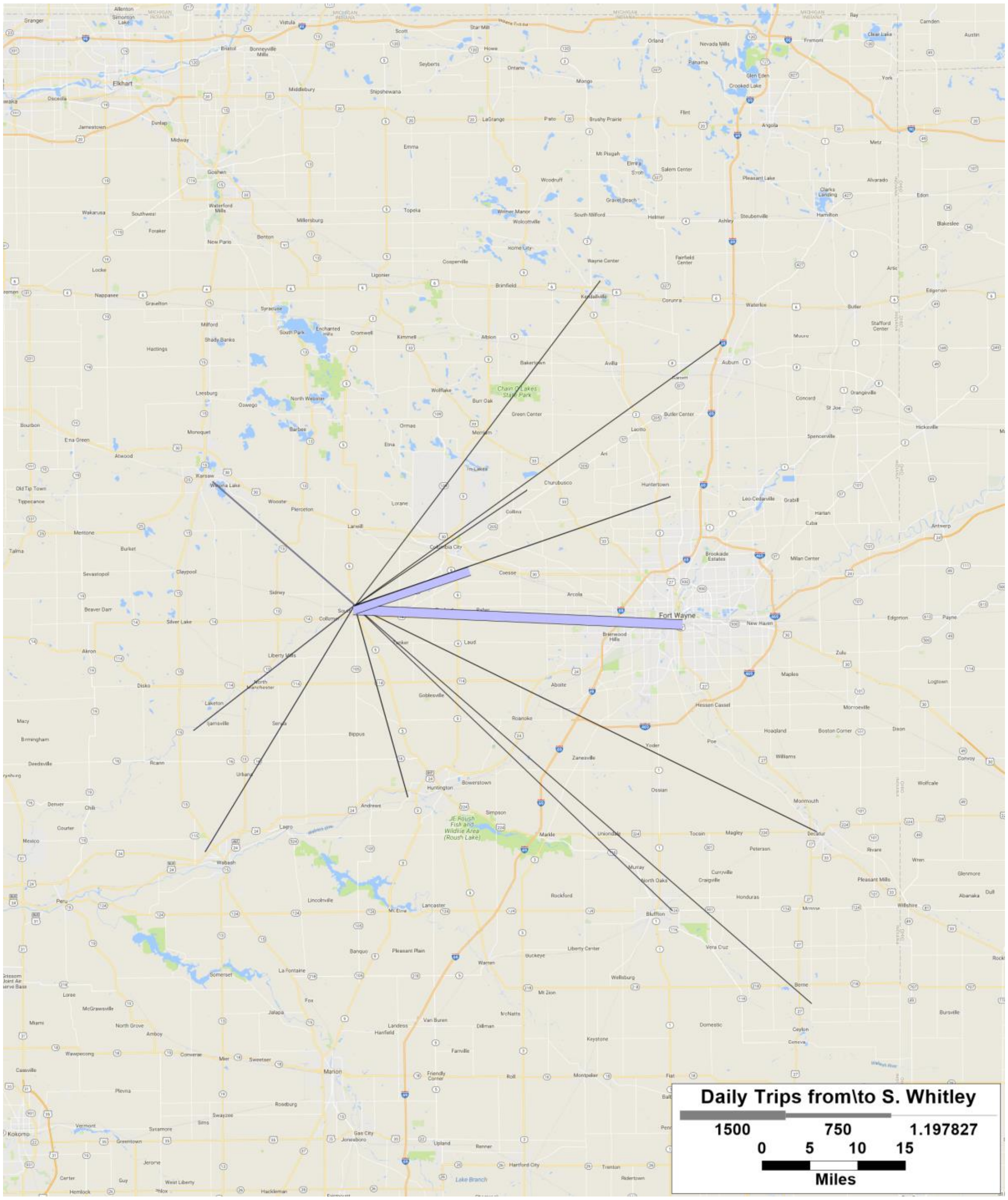


FIGURE 50. DESIRE LINES FOR S. WHITLEY

APPENDIX B. KEY FACILITY ANALYSIS MAPS

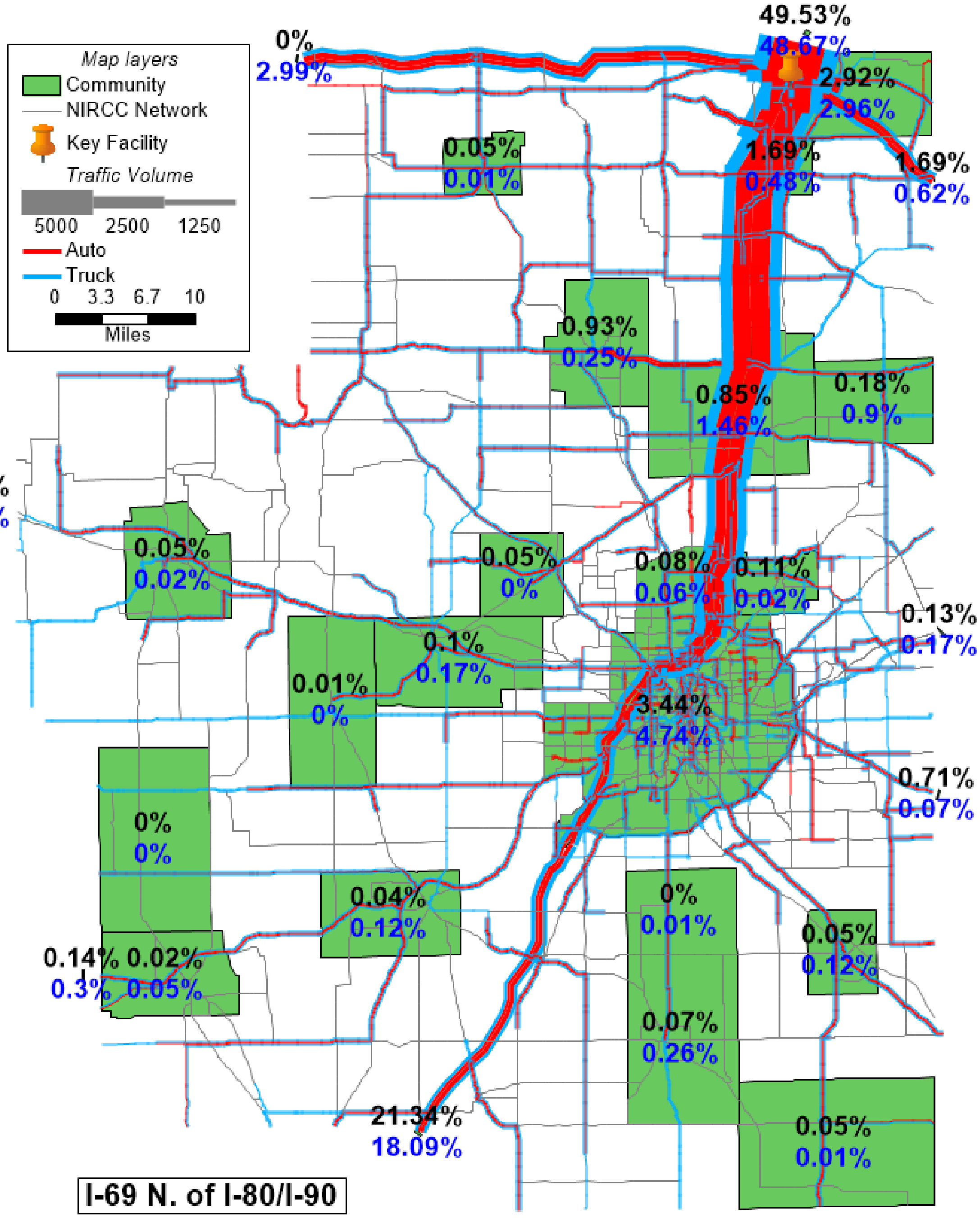
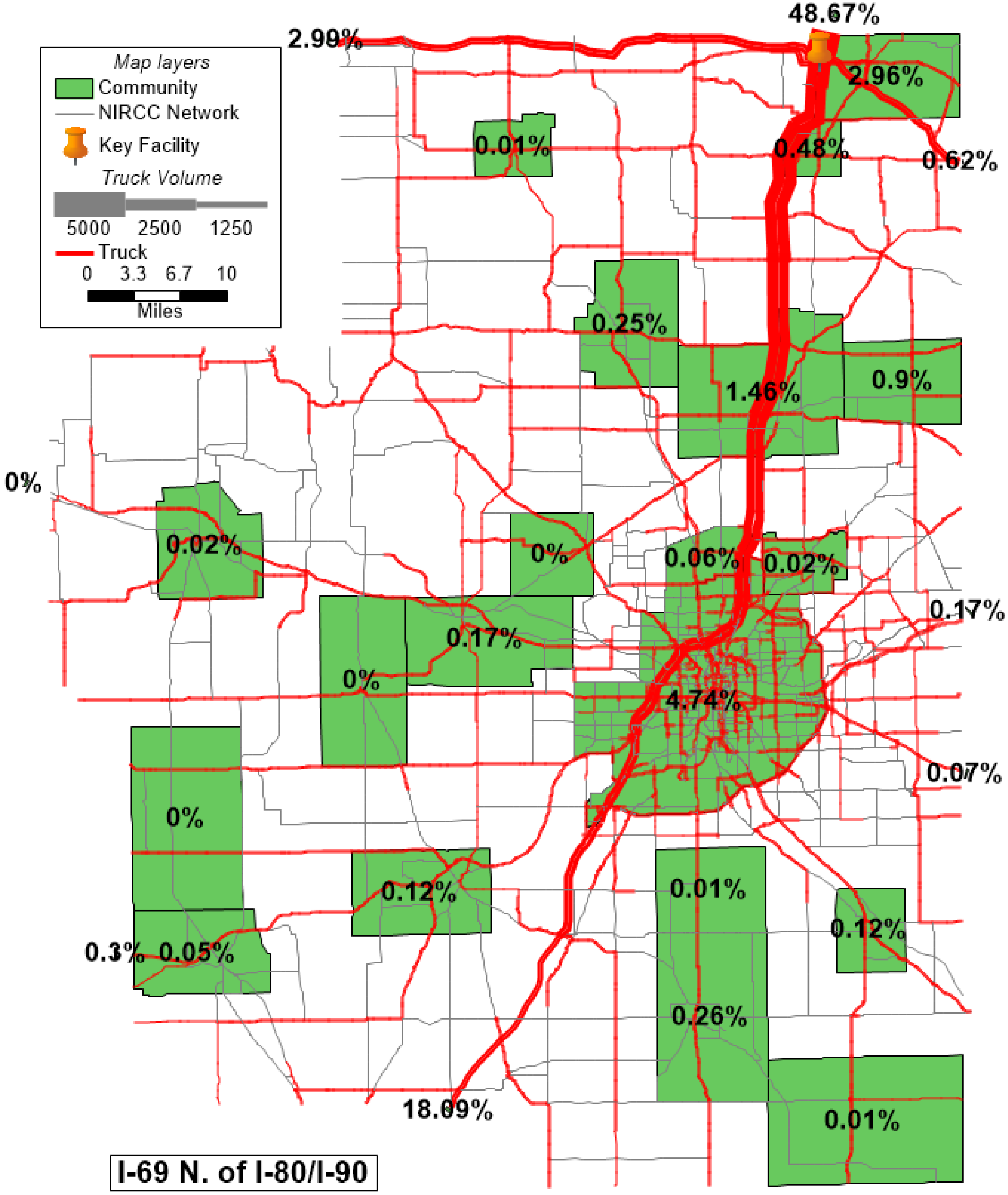


FIGURE 51. DEMAND AT I-69 N. OF I-80/I-90



I-69 N. of I-80/I-90

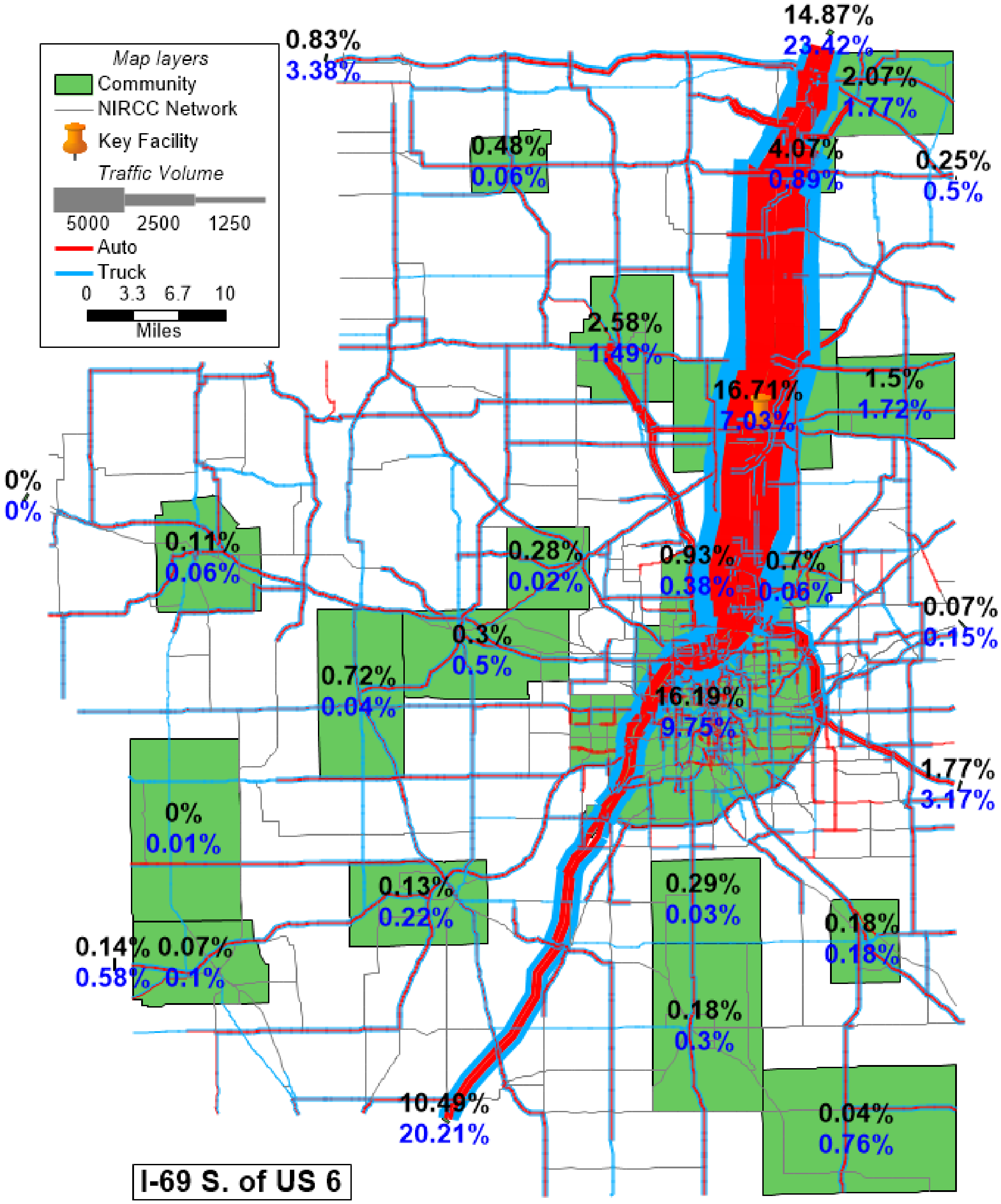
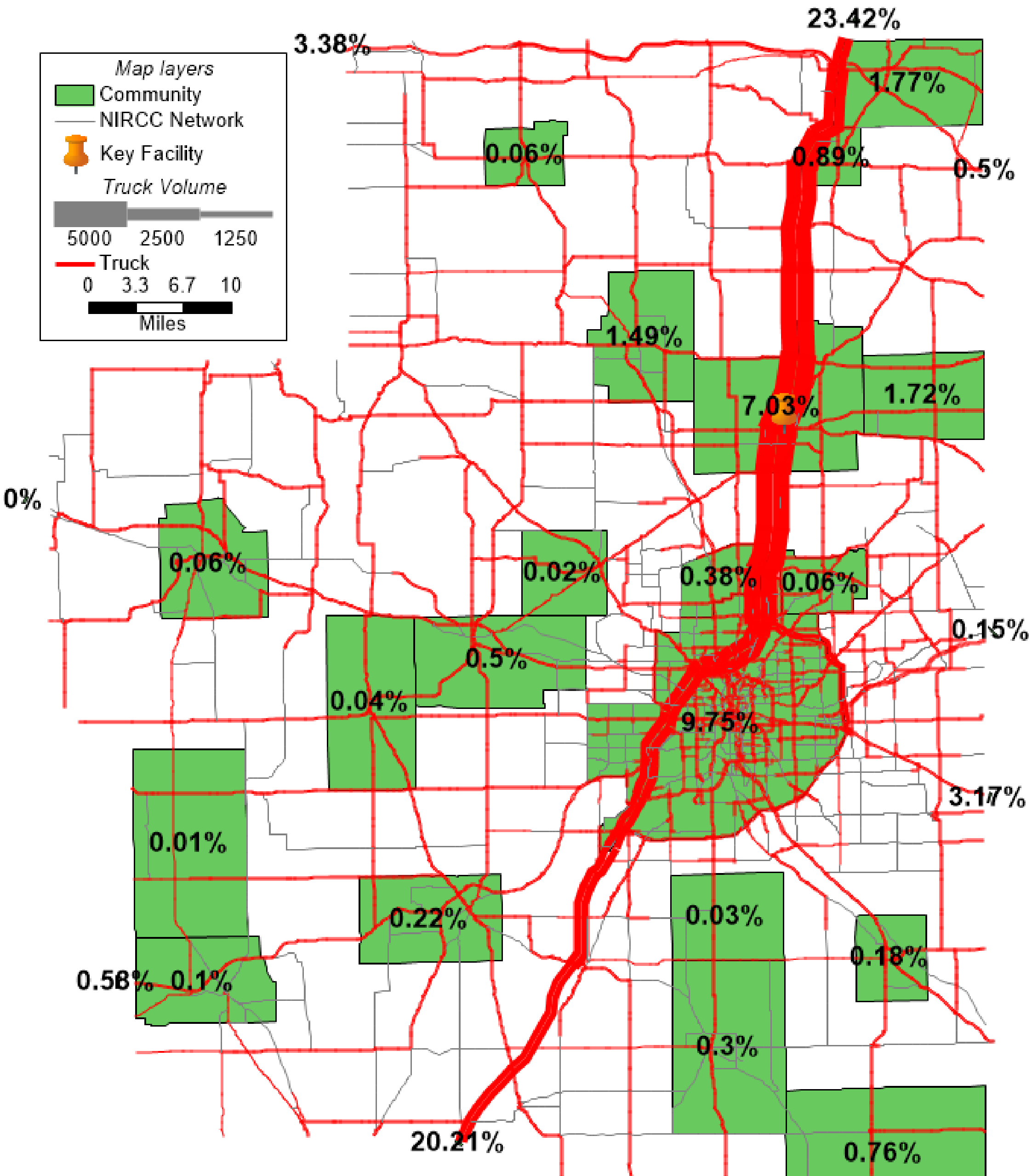


FIGURE 52. DEMAND AT I-69 S. OF US 6



Map layers
Community
NIRCC Network
Key Facility
Truck Volume
5000 2500 1250
Truck
0 3.3 6.7 10
Miles
23.42%
3.38%
1.77%
0.06%
0.89%
0.5%
1.49%
7.03%
1.72%
0%
0.06%
0.02%
0.38%
0.06%
0.15%
0.5%
0.04%
9.75%
3.17%
0.01%
0.22%
0.03%
0.18%
0.58%
0.1%
0.3%
20.21%
0.76%

I-69 S. of US 6

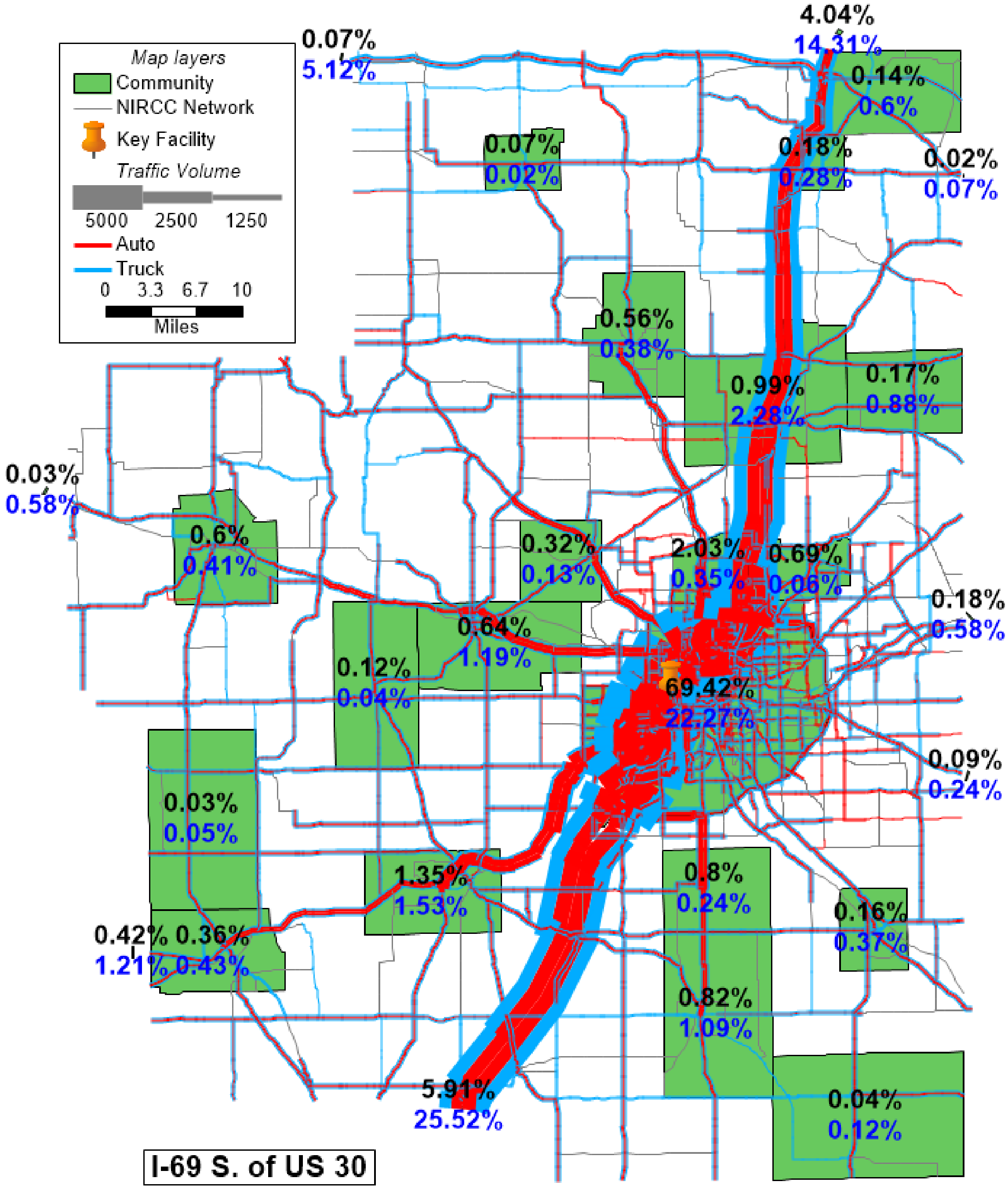
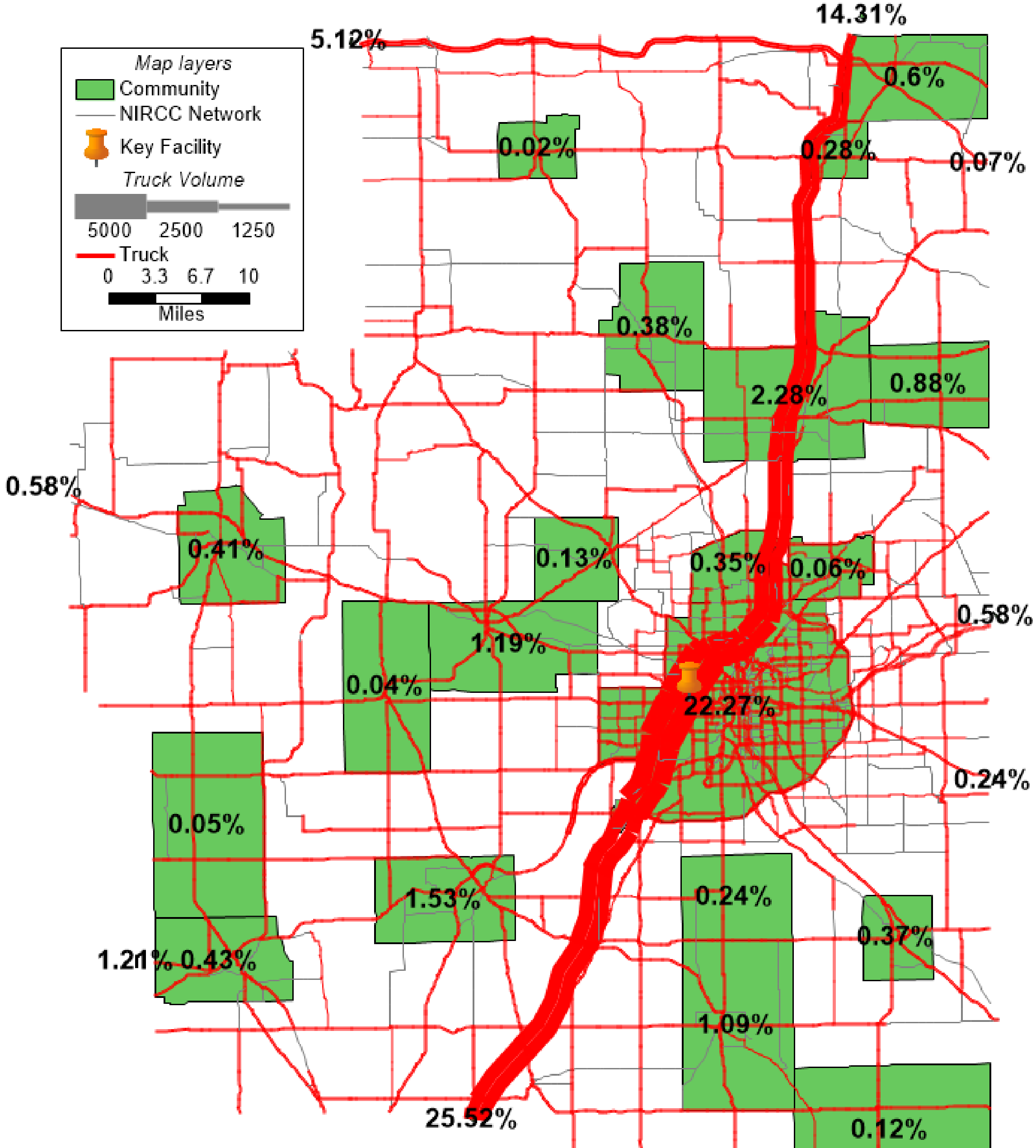


FIGURE 53. DEMAND AT I-69 S. OF US 30


Map layers
Community
NIRCC Network
Key Facility
Truck Volume
5000
2500
1250
Truck
0
3.3
6.7
10
Miles
5.12%
14.31%
0.6%
0.02%
0.28%
0.07%
0.38%
2.28%
0.88%
0.58%
0.41%
0.13%
0.35%
0.06%
0.58%
1.19%
0.04%
22.27%
0.24%
0.05%
1.53%
0.24%
0.37%
1.21%
0.43%
1.09%
25.52%
0.12%

I-69 S. of US 30

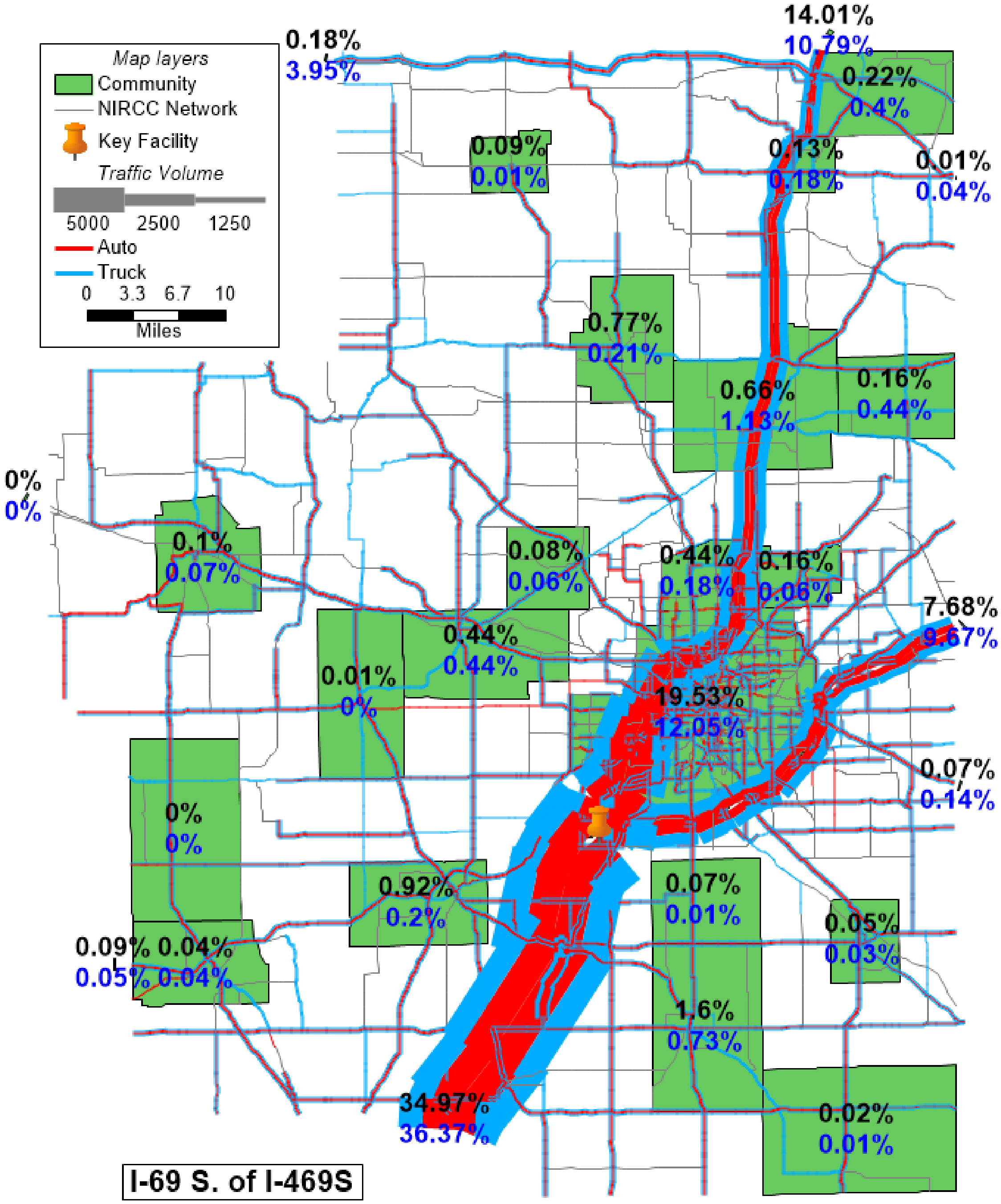
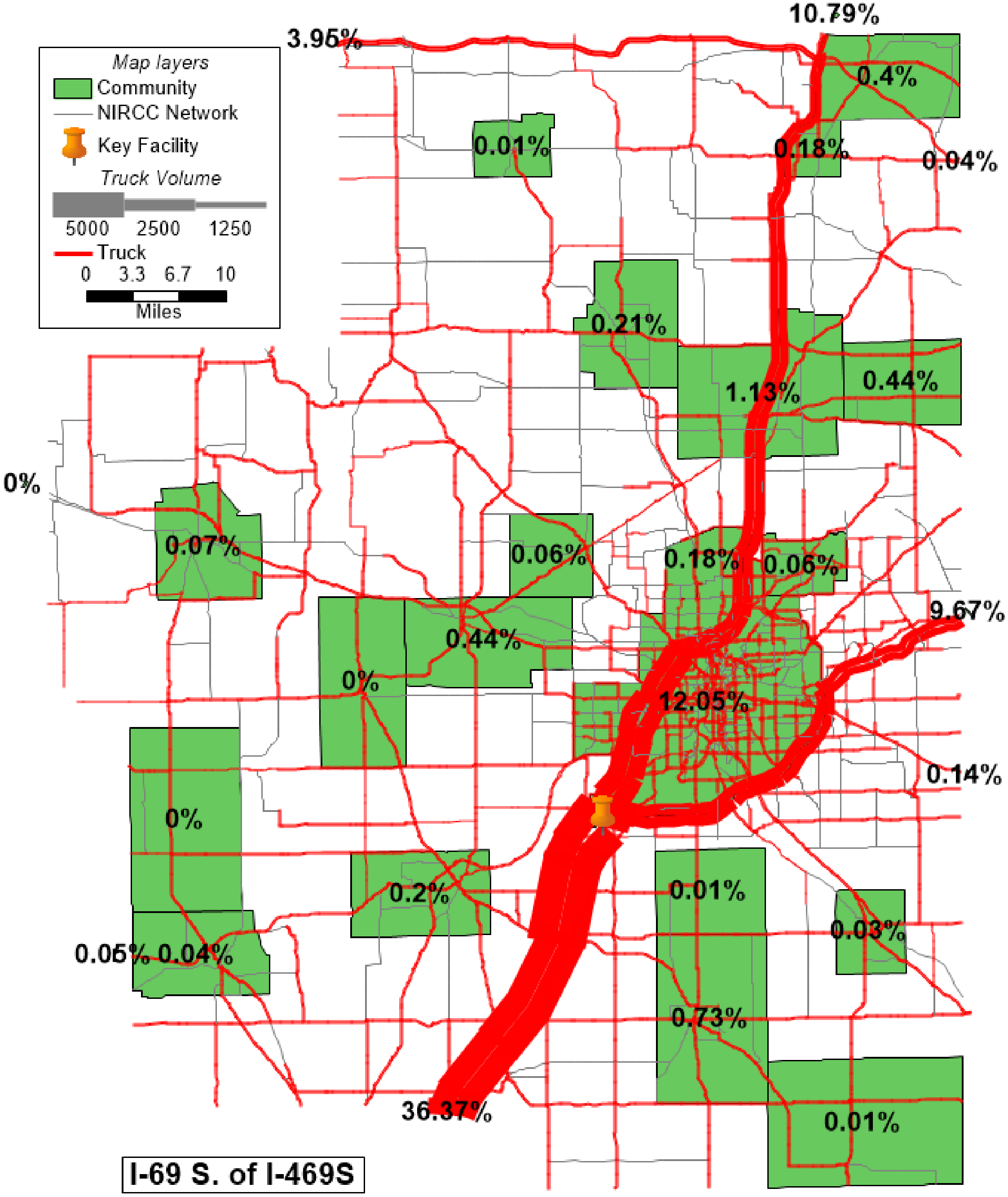


FIGURE 54. DEMAND AT I-69 S. OF I-469S



Map layers
Community
NIRCC Network
Key Facility
Truck Volume
5000 2500 1250
Truck
0 3.3 6.7 10
Miles
3.95%
10.79%
0.4%
0.01%
0.18%
0.04%
0.21%
1.13%
0.44%
0%
0.07%
0.06%
0.18%
0.06%
9.67%
0.44%
0%
12.05%
0.14%
0%
0.2%
0.01%
0.03%
0.05%
0.04%
0.73%
36.37%
0.01%
I-69 S. of I-469S

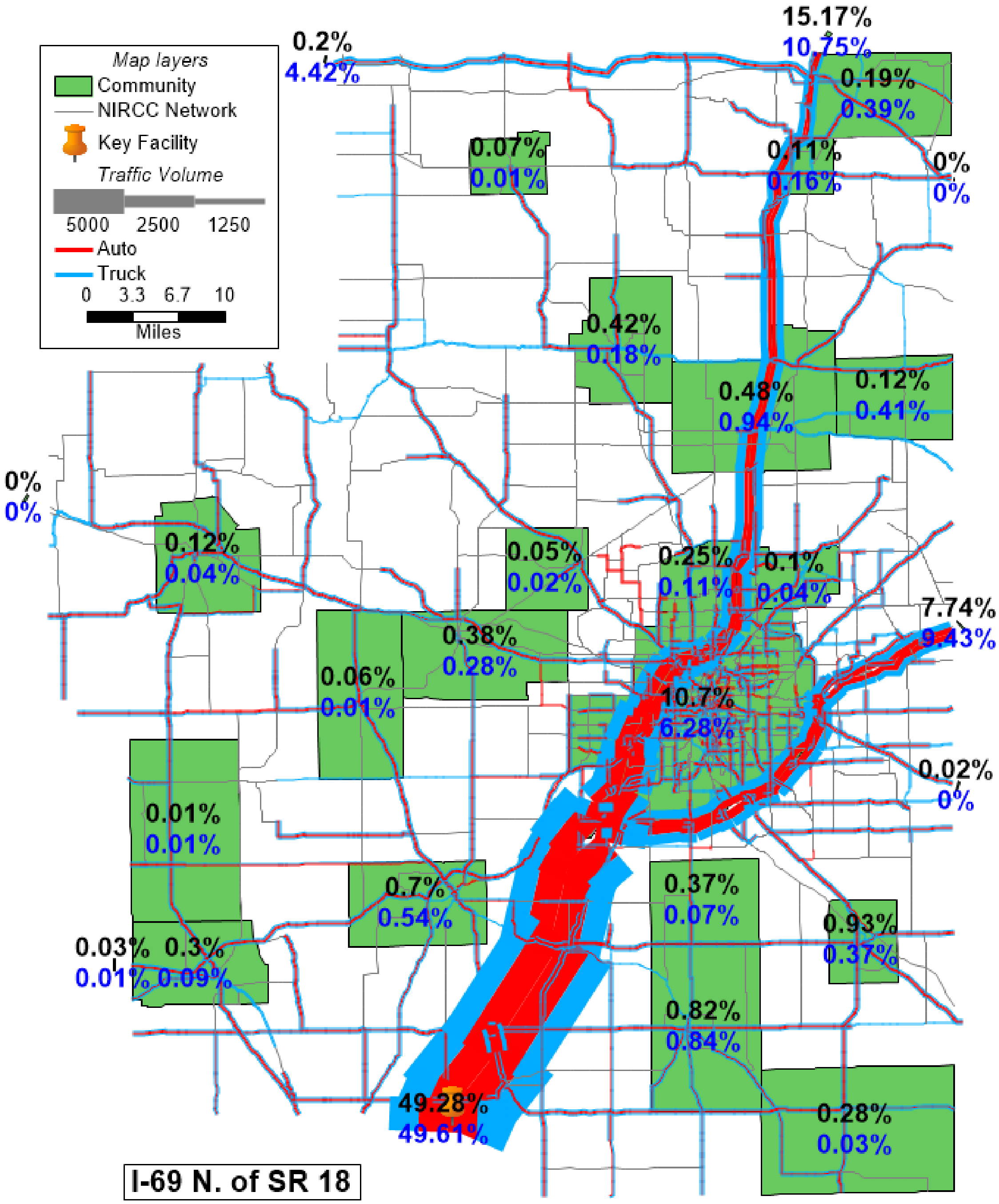
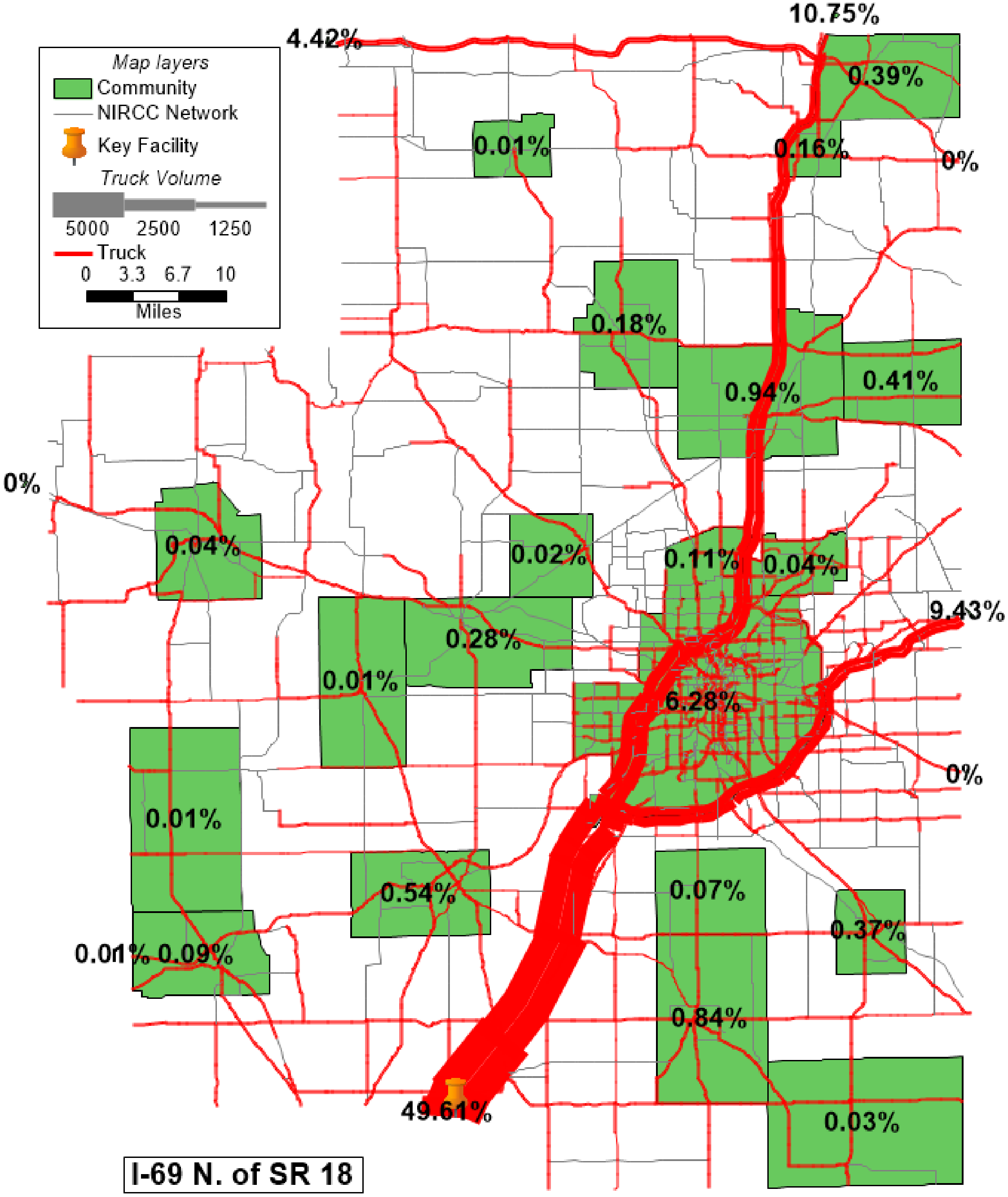


FIGURE 55. DEMAND AT I-69 N. OF SR 18


Map layers
Community
NIRCC Network
Key Facility
Truck Volume
5000
2500
1250
Truck
0
3.3
6.7
10
Miles
4.42%
10.75%
0.39%
0.01%
0.16%
0%
0.18%
0.94%
0.41%
0%
0.04%
0.02%
0.11%
0.04%
9.43%
0.28%
0.01%
6.28%
0%
0.01%
0.54%
0.07%
0.37%
0.01%
0.09%
0.84%
49.61%
0.03%

I-69 N. of SR 18

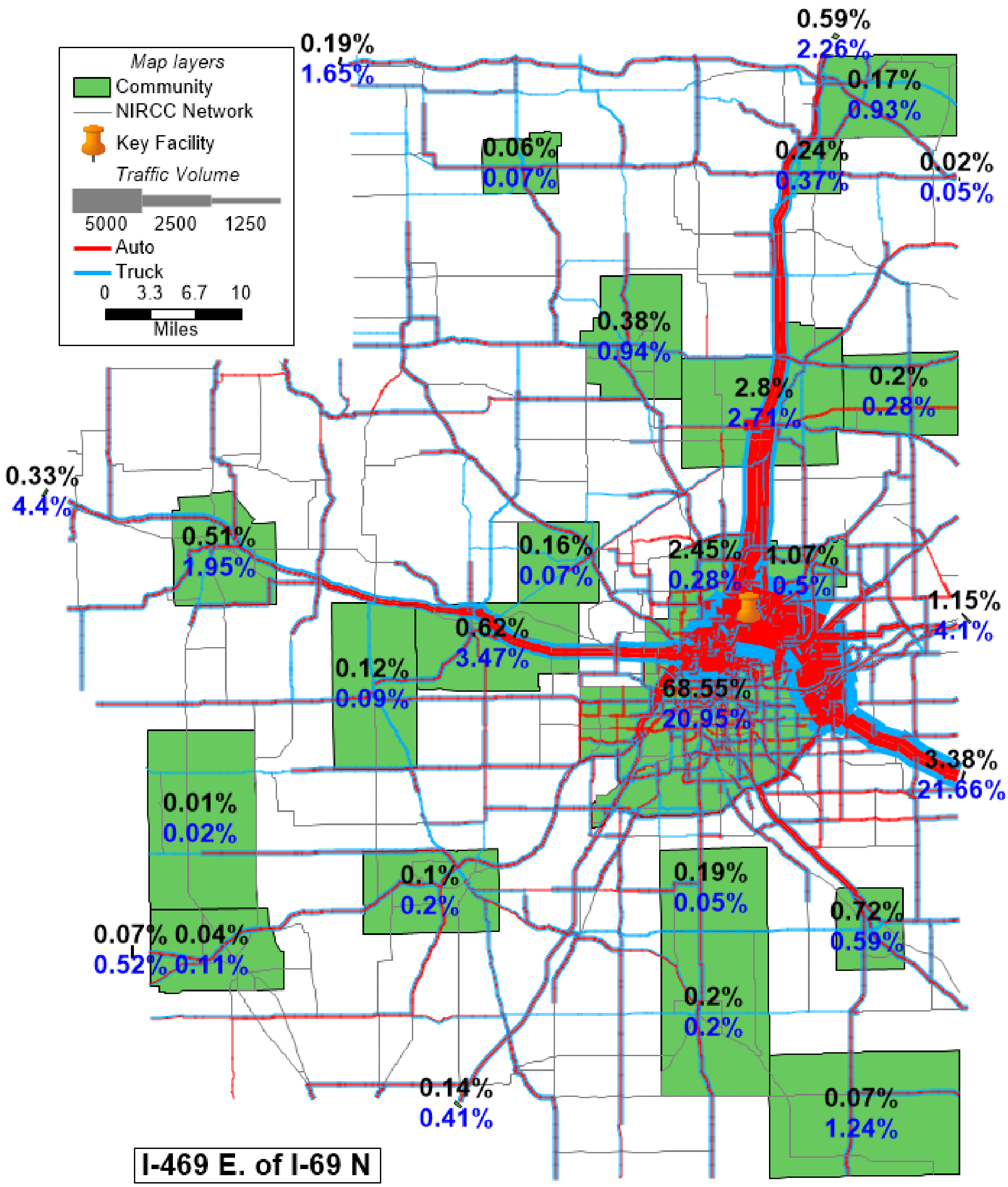
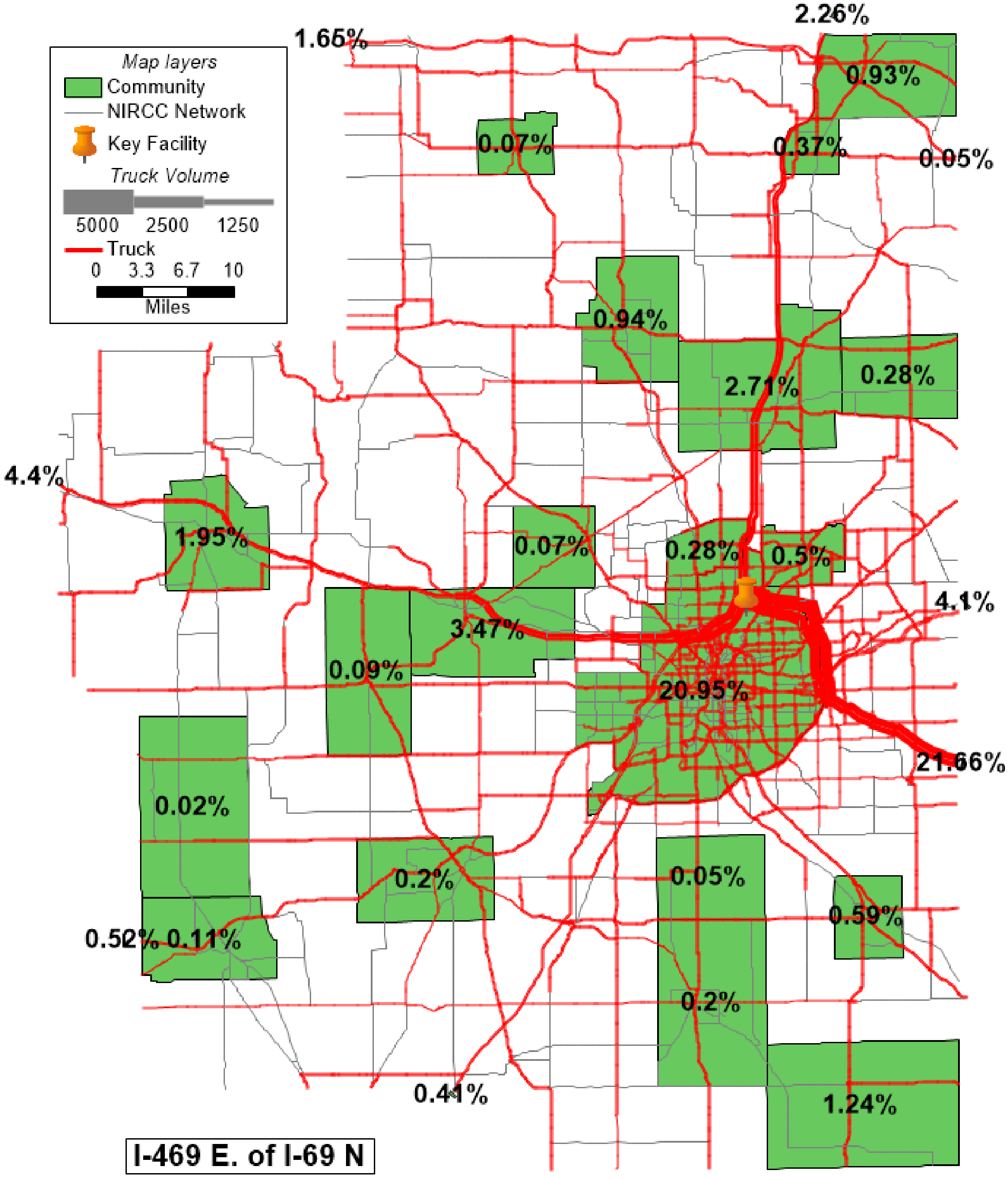


FIGURE 56. DEMAND AT I-469 E. OF I-69 N



I-469 E. of I-69 N

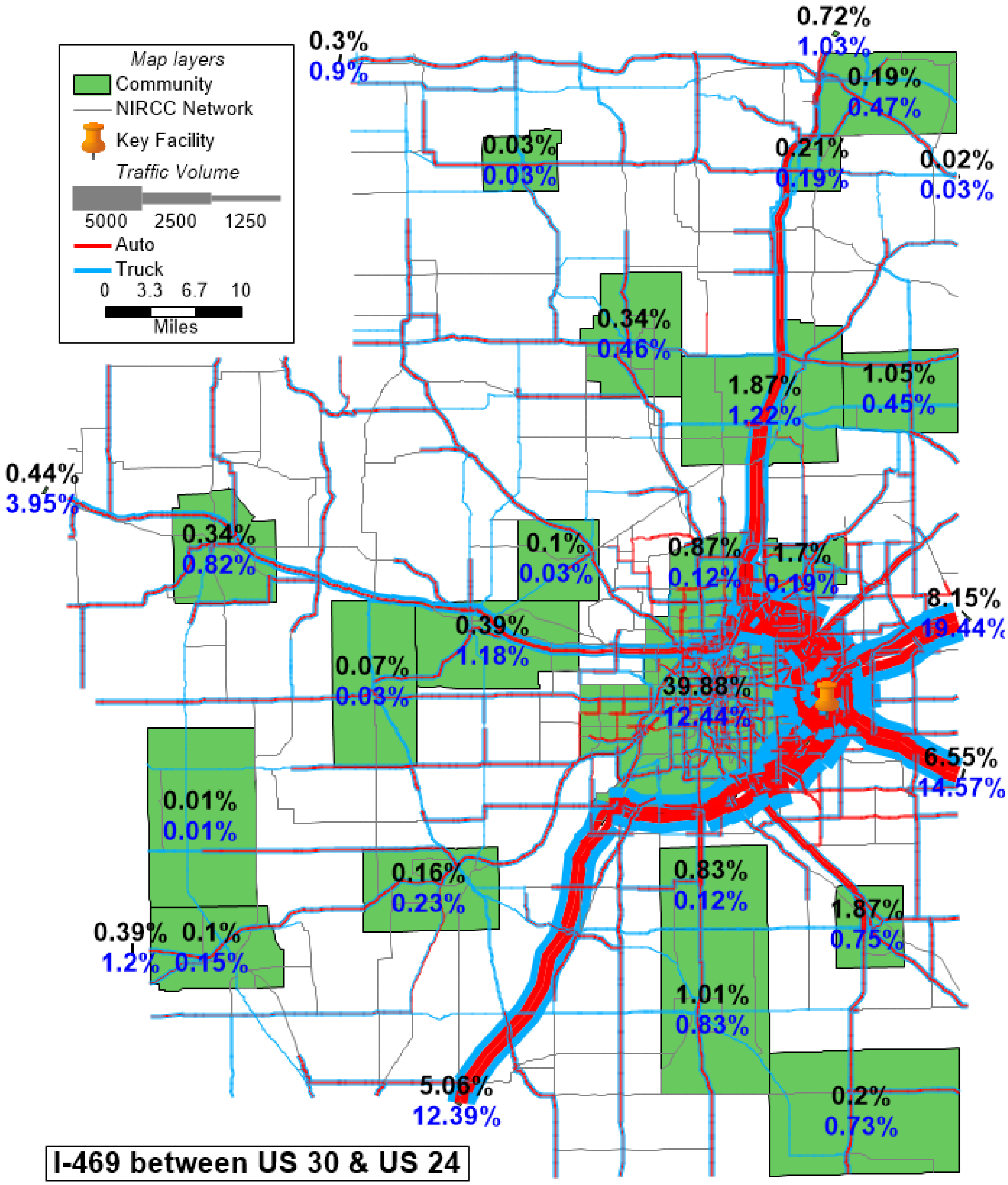
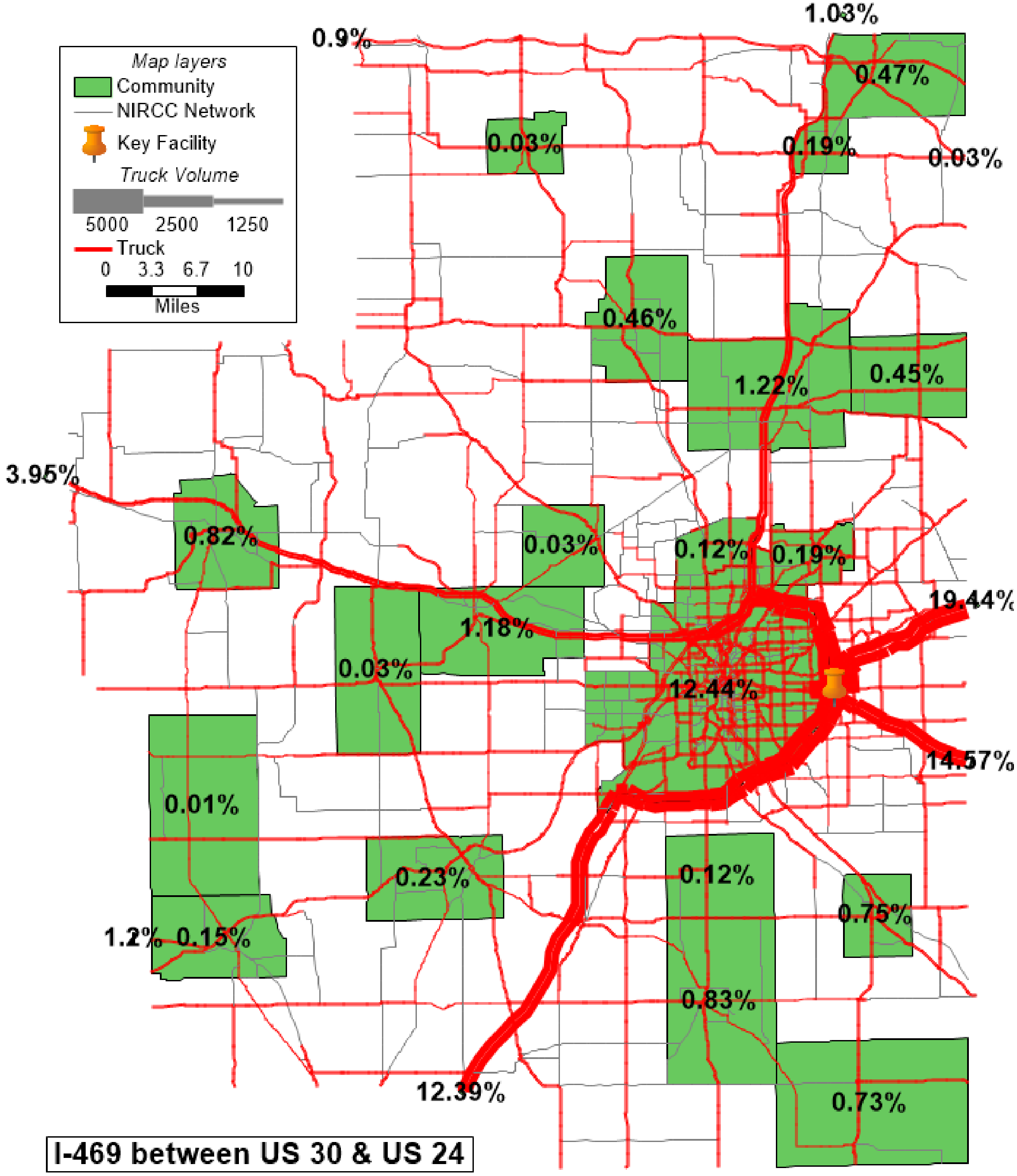


FIGURE 57. DEMAND AT I-469 BETWEEN US 30 & US 24



I-469 between US 30 & US 24

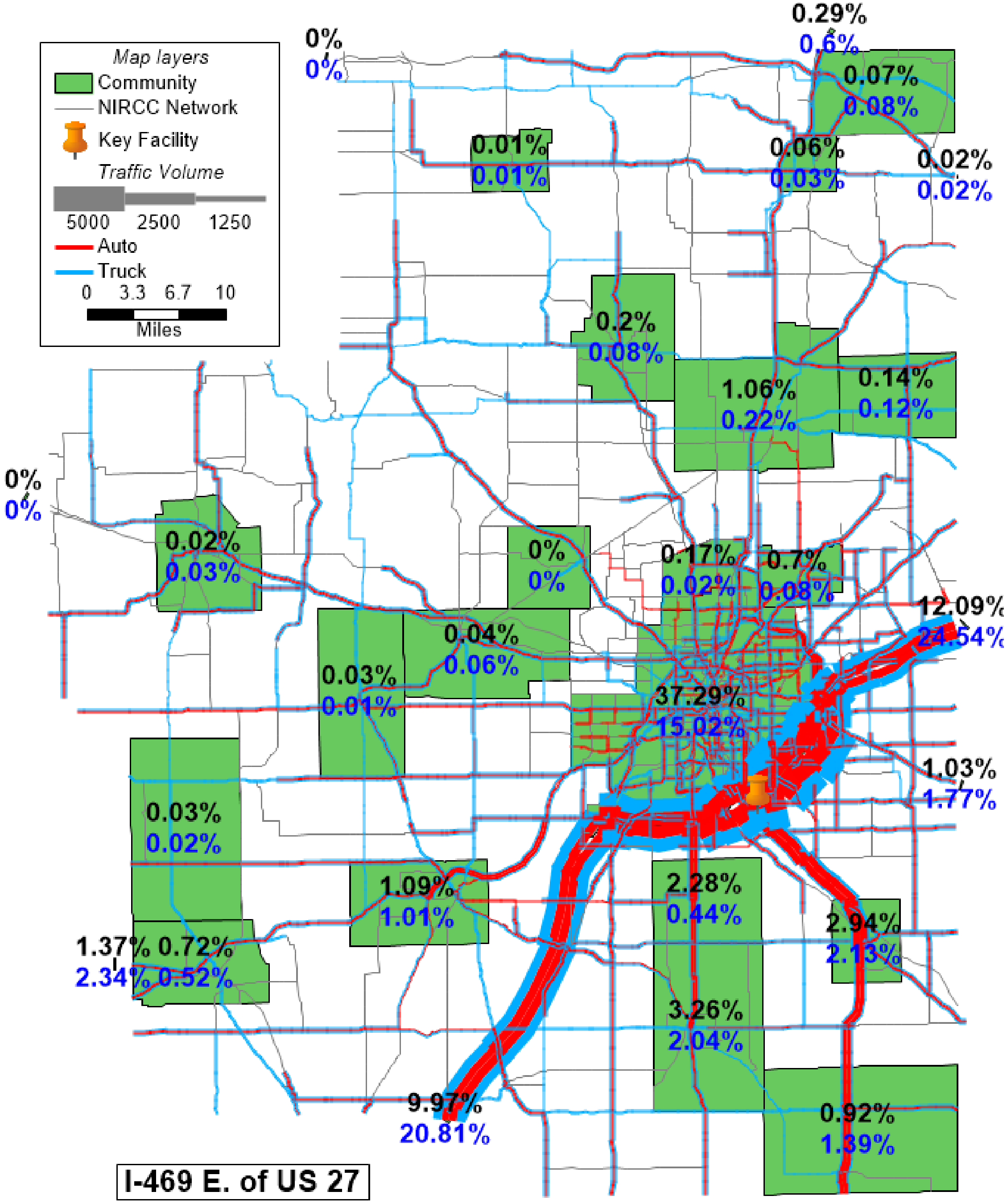
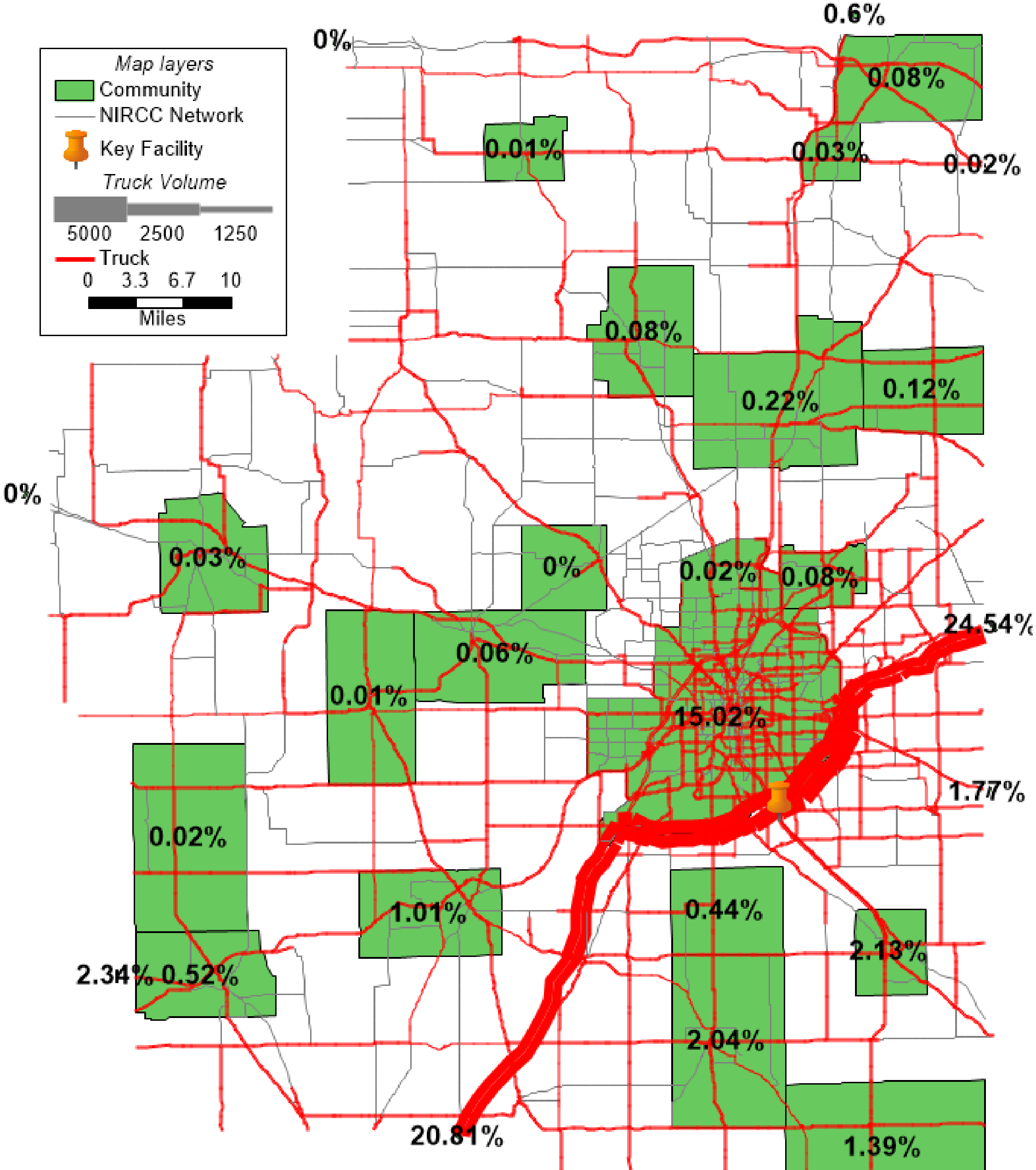


FIGURE 58. DEMAND AT I-469 E. OF US 27



I-469 E. of US 27

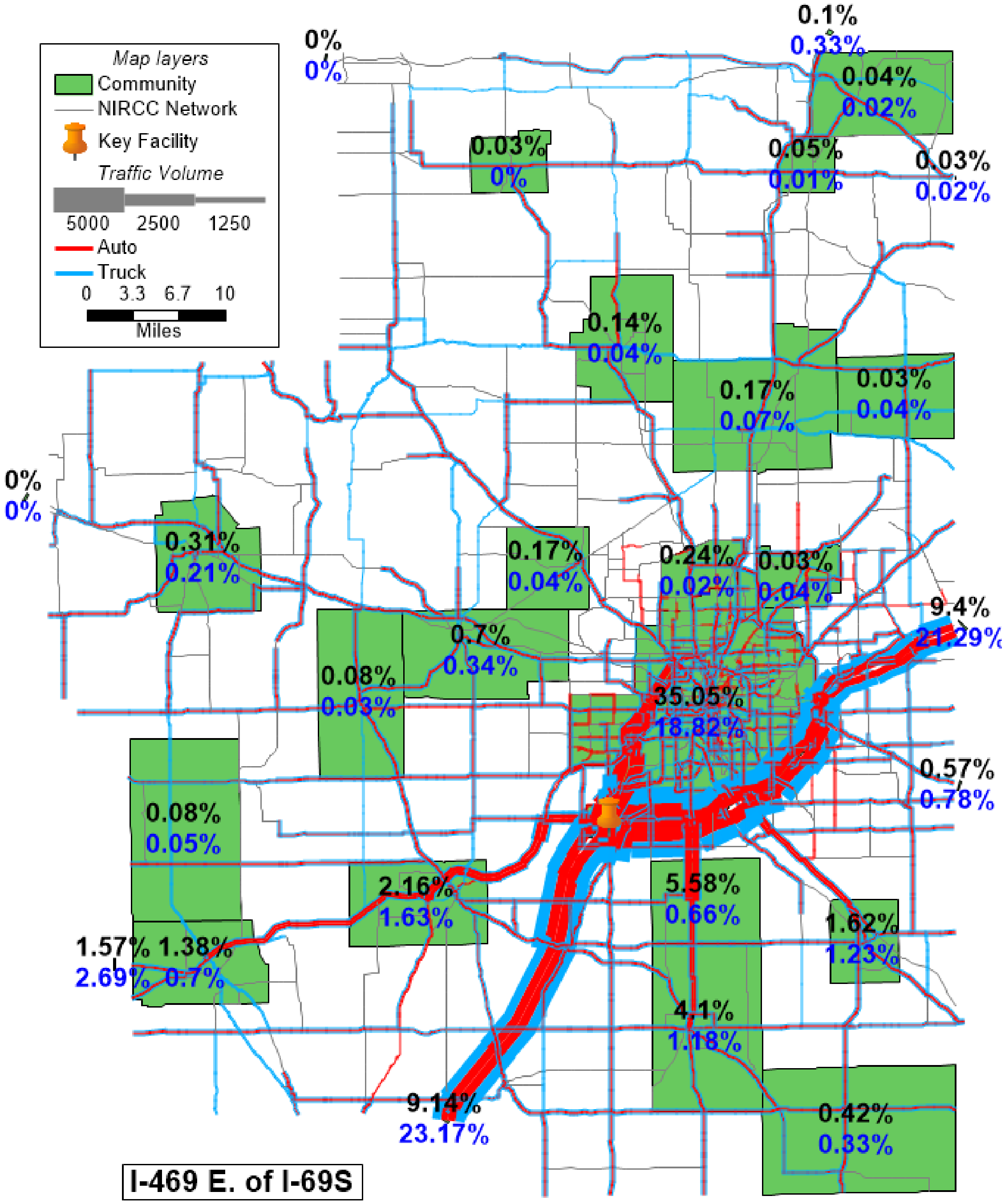
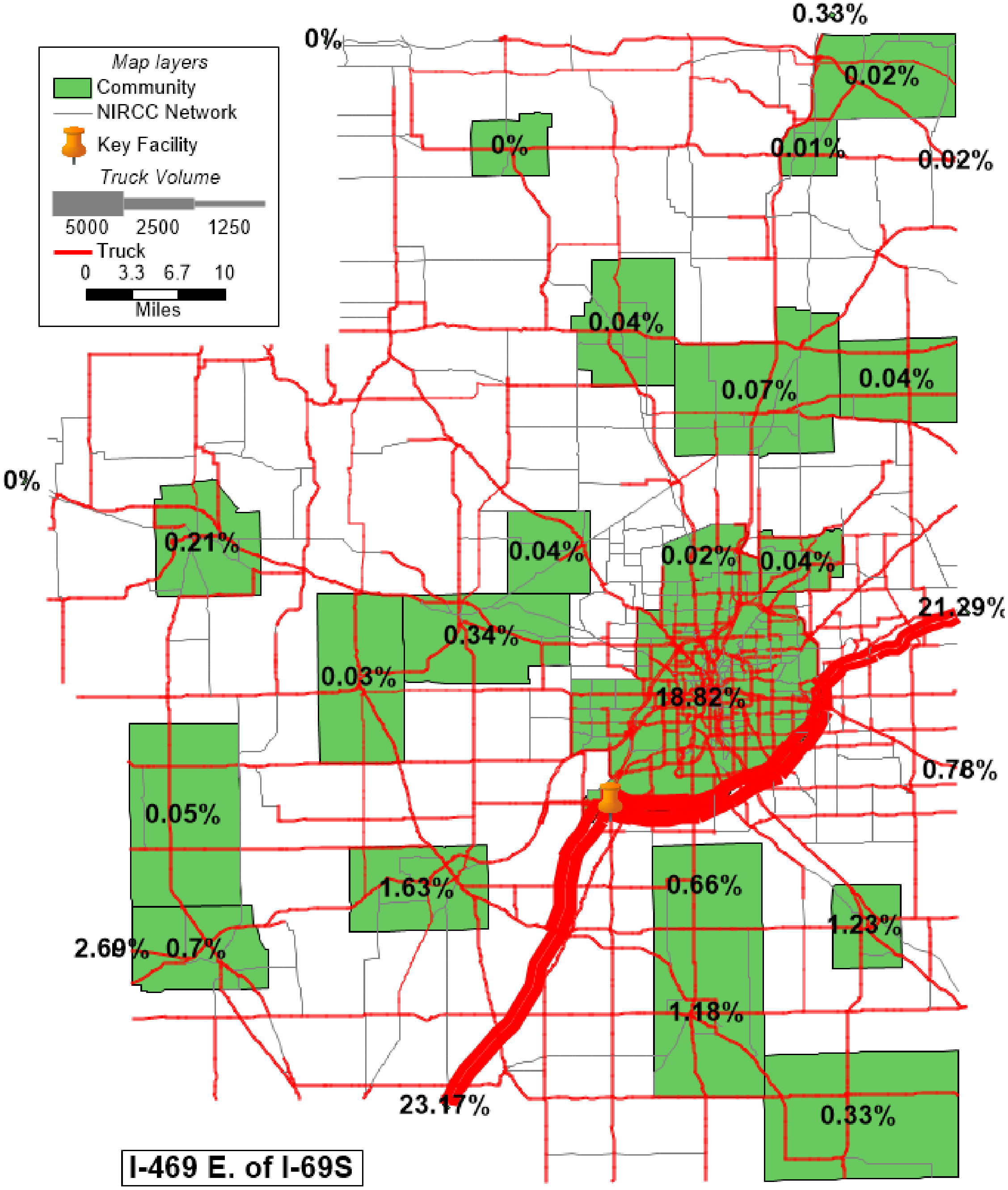


FIGURE 59. DEMAND AT I-469 E. OF I-69 S



I-469 E. of I-69S

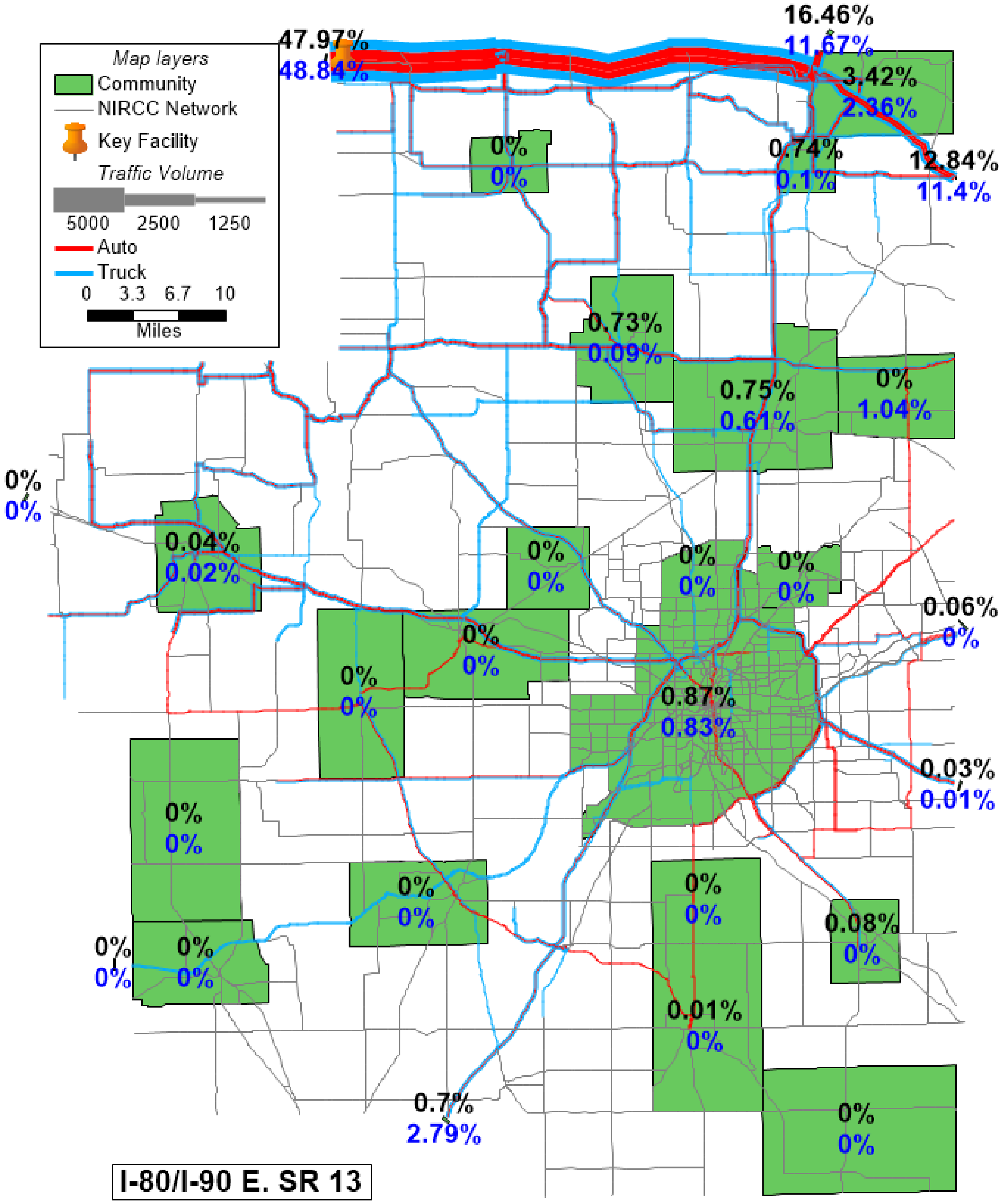
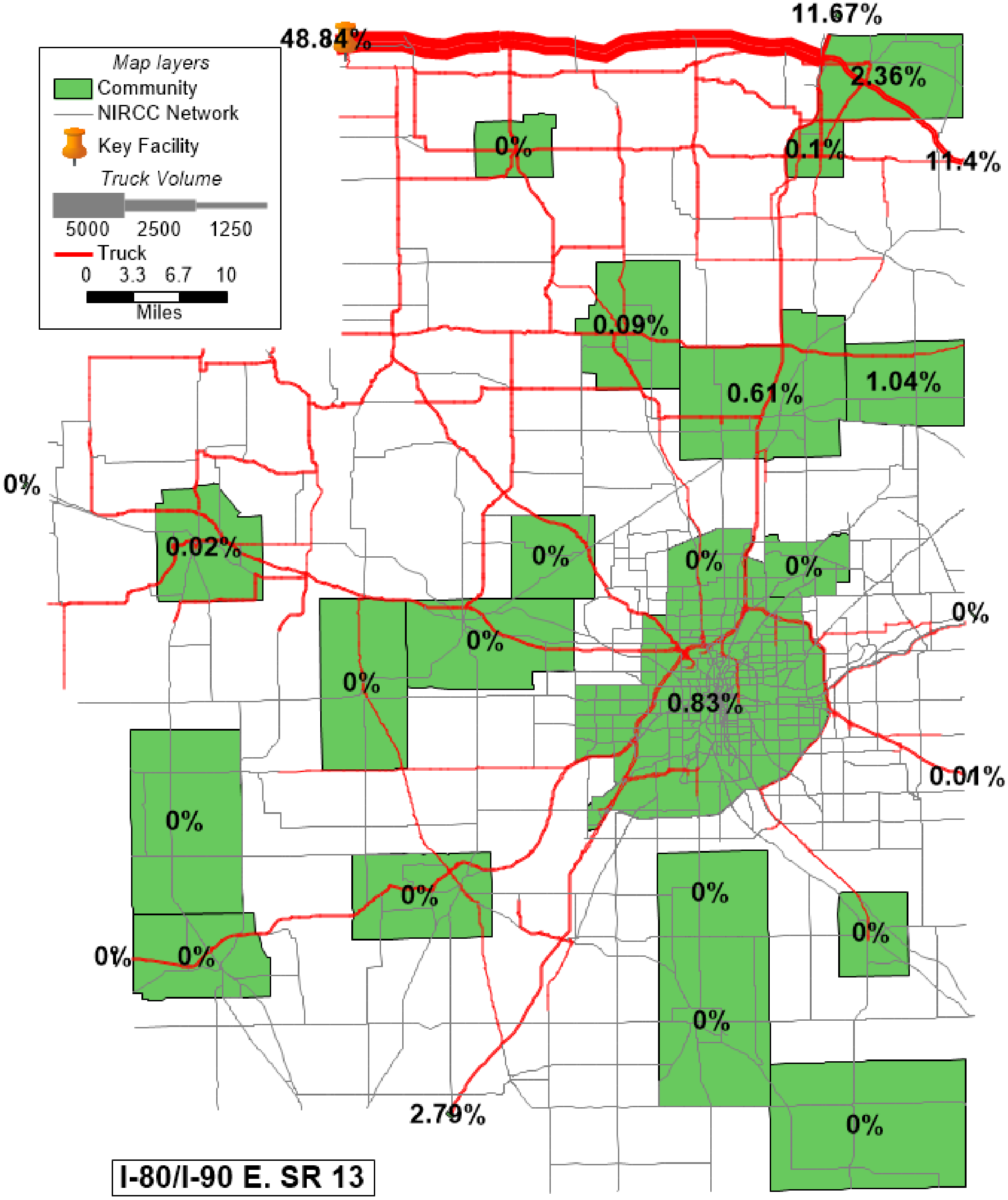


FIGURE 60. DEMAND AT I-80/90 E. OF SR 13



I-80/I-90 E. SR 13

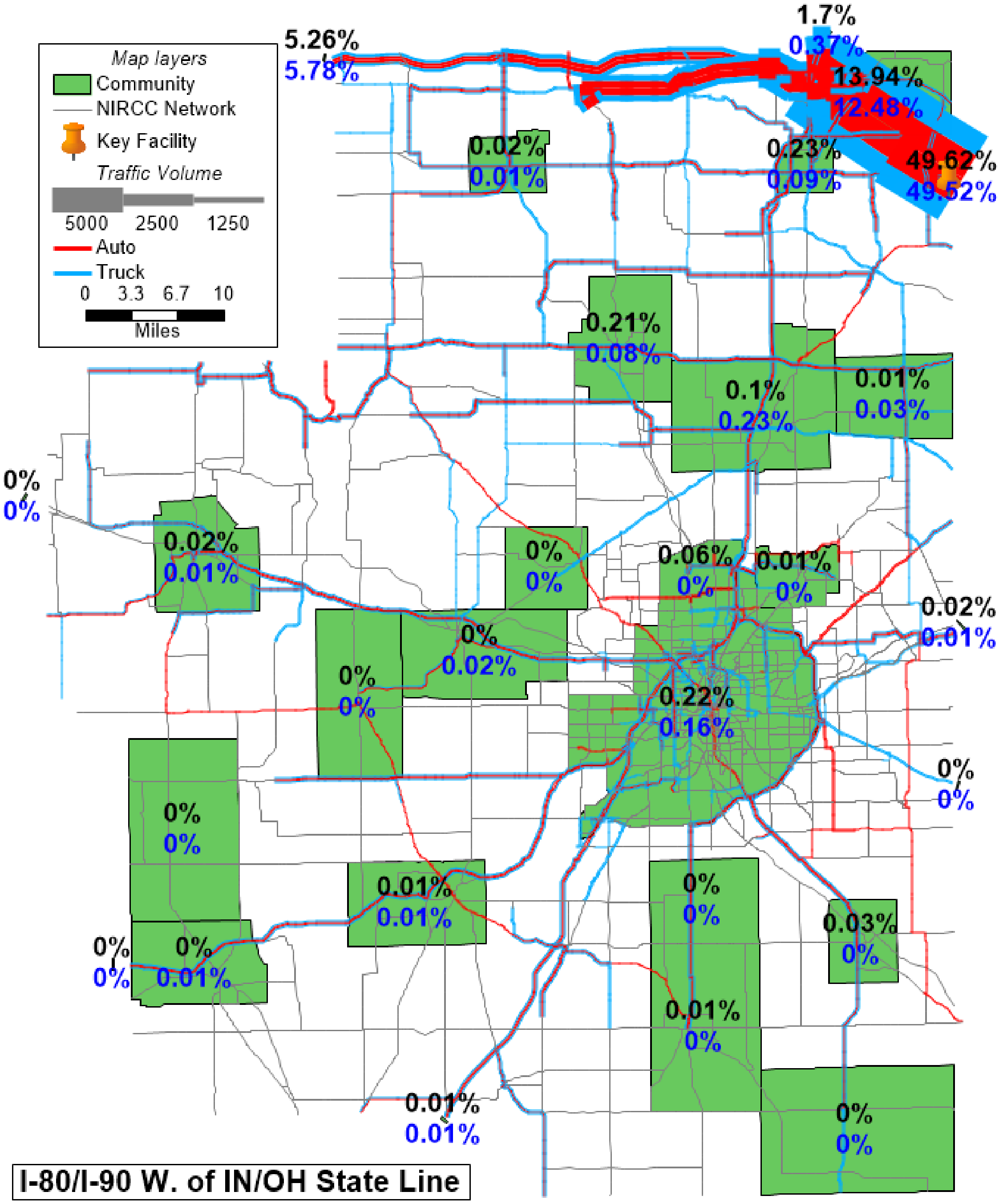
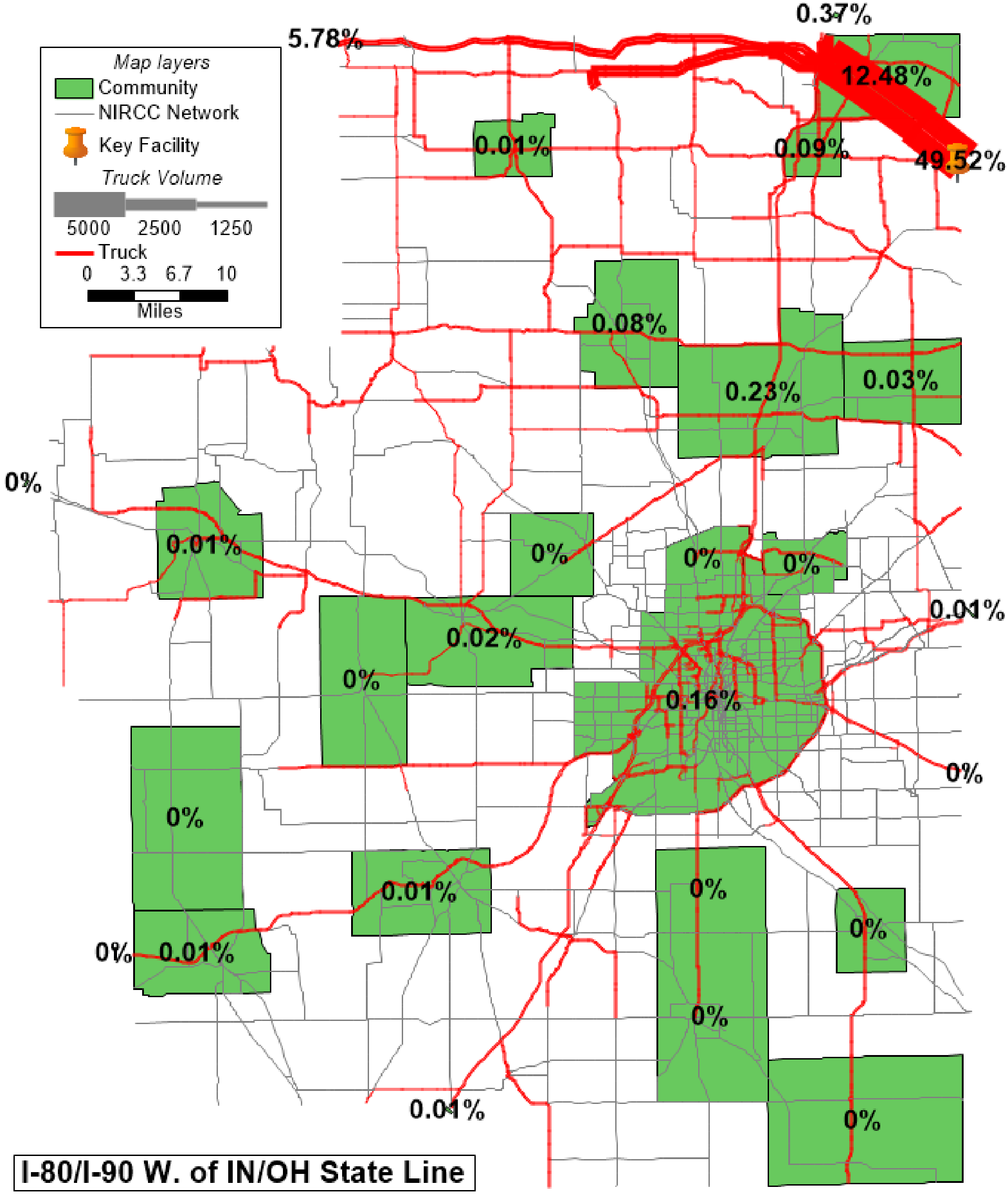


FIGURE 61. DEMAND AT I-80/90 W. OF IN/OH STATE LINE



I-80/I-90 W. of IN/OH State Line

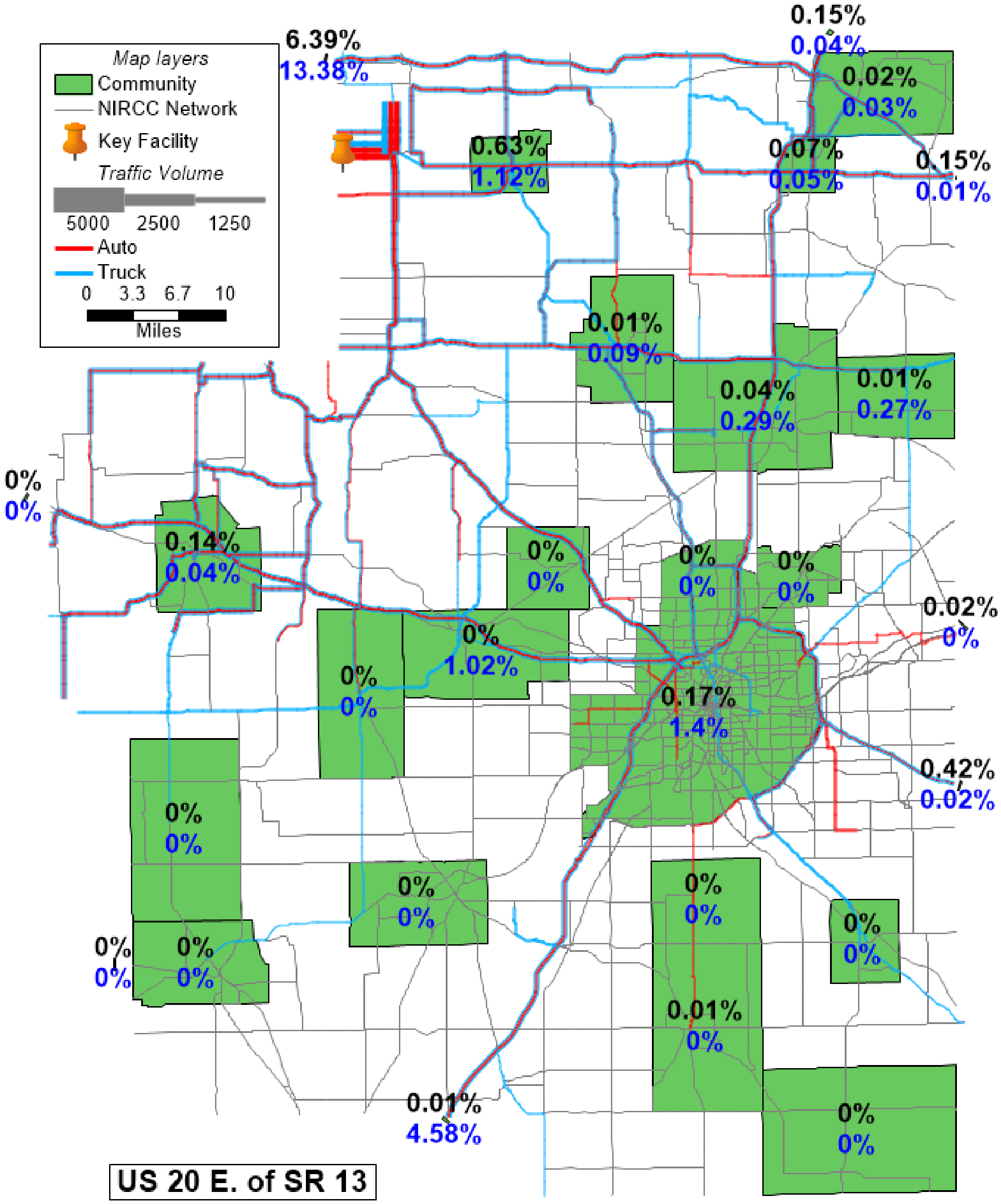
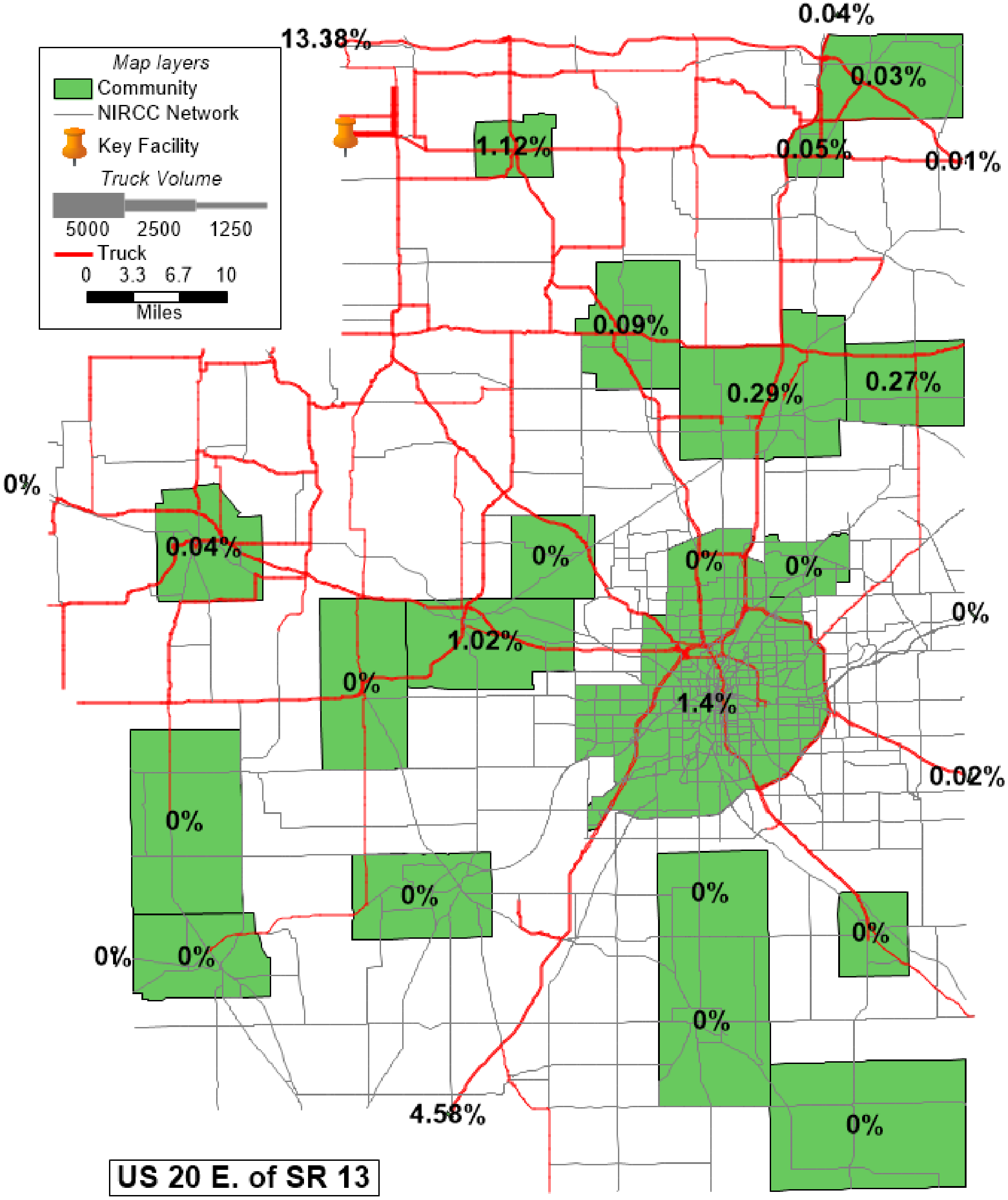


FIGURE 62. DEMAND AT US 20 E. OF SR 13



US 20 E. of SR 13

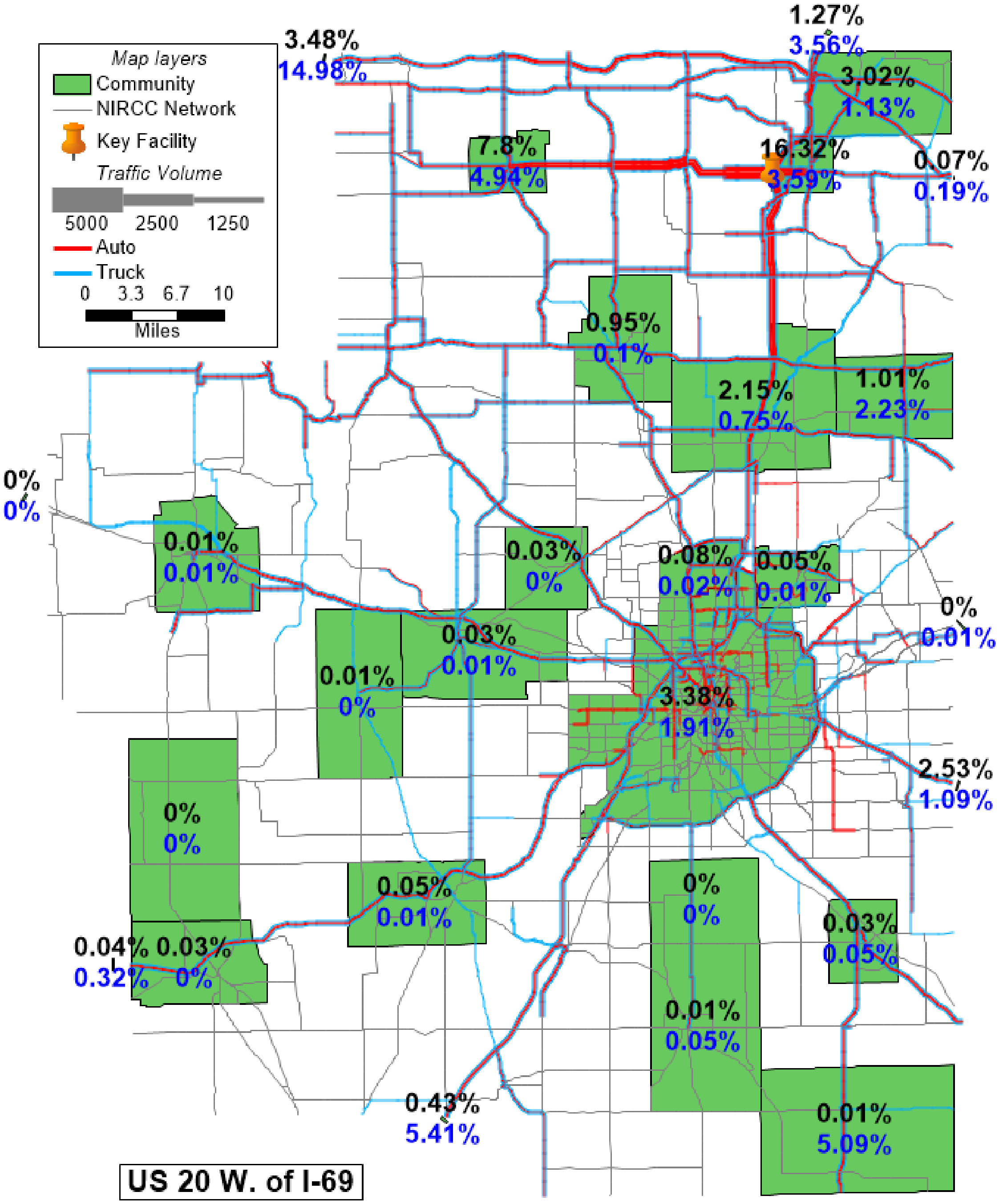
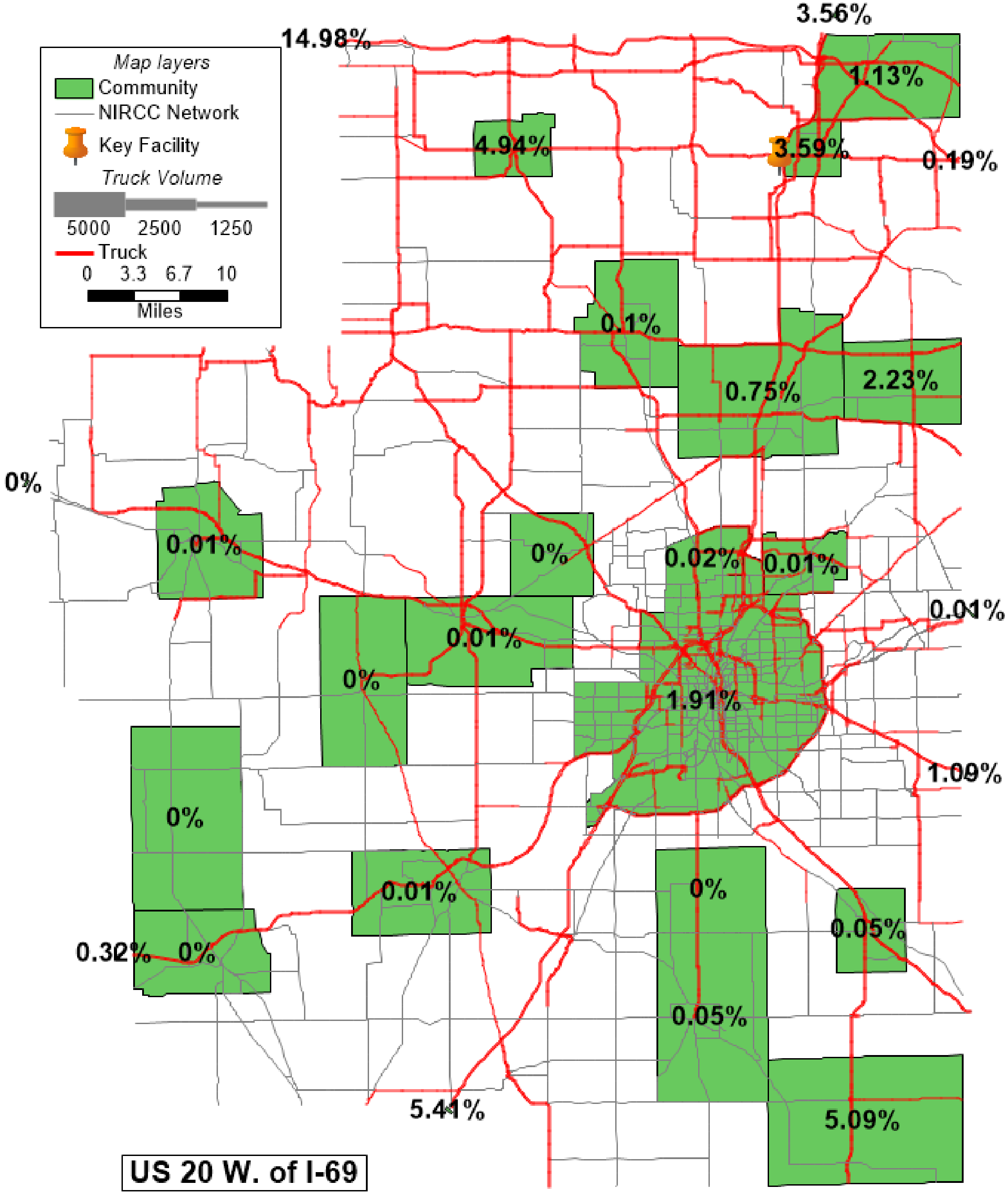


FIGURE 63. DEMAND AT US 20 W. OF I-69



Map layers
Community
NIRCC Network
Key Facility
Truck Volume
5000
2500
1250
Truck
0
3.3
6.7
10
Miles
14.98%
3.56%
1.13%
4.94%
3.59%
0.19%
0.1%
0.75%
2.23%
0%
0.01%
0%
0.02%
0.01%
0.01%
0.01%
0%
1.91%
1.09%
0%
0.01%
0%
0.05%
0.32%
0%
0.05%
5.41%
5.09%
US 20 W. of I-69

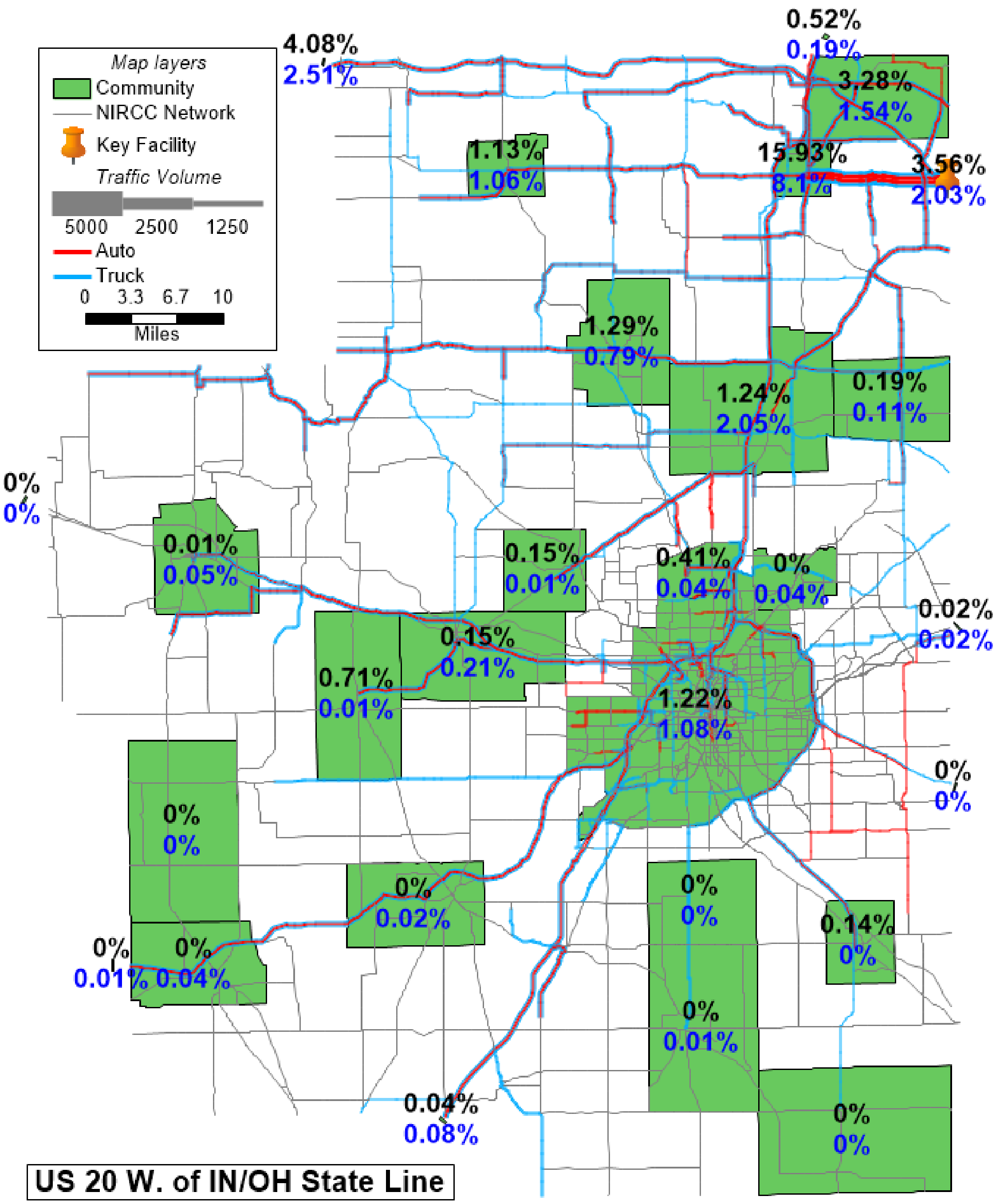
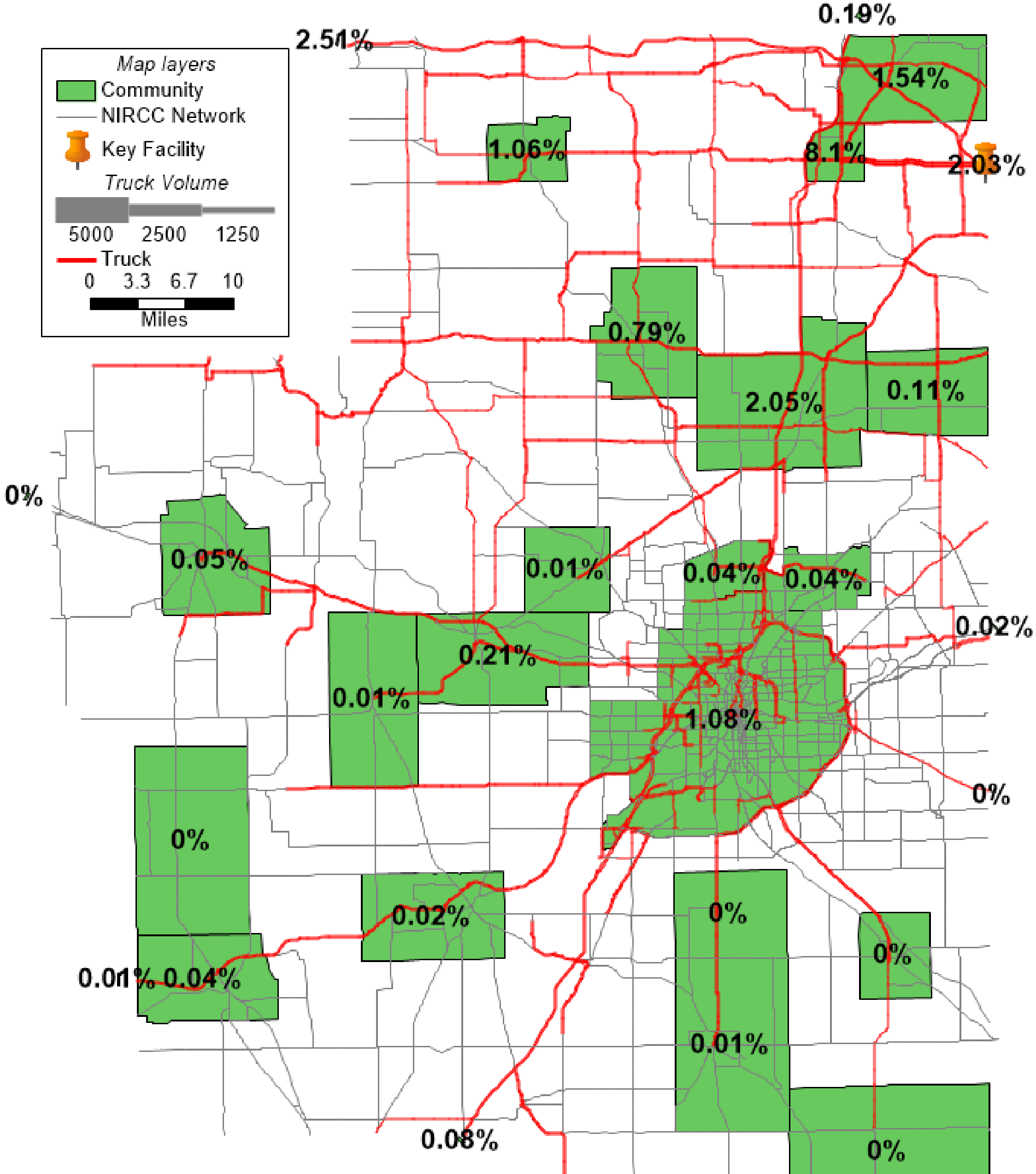


FIGURE 64. DEMAND AT US 20 W. OF IN/OH STATE LINE



US 20 W. of IN/OH State Line

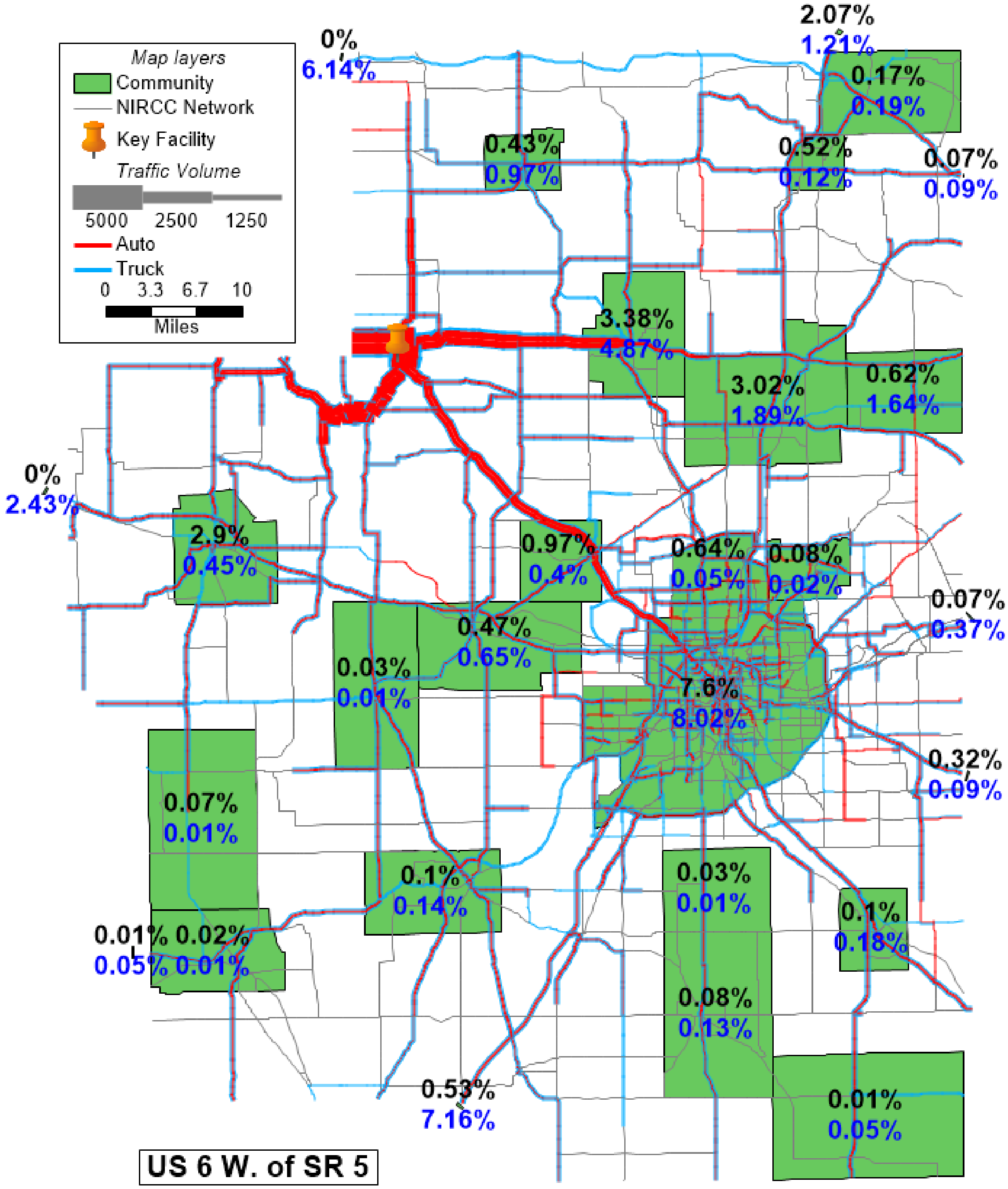
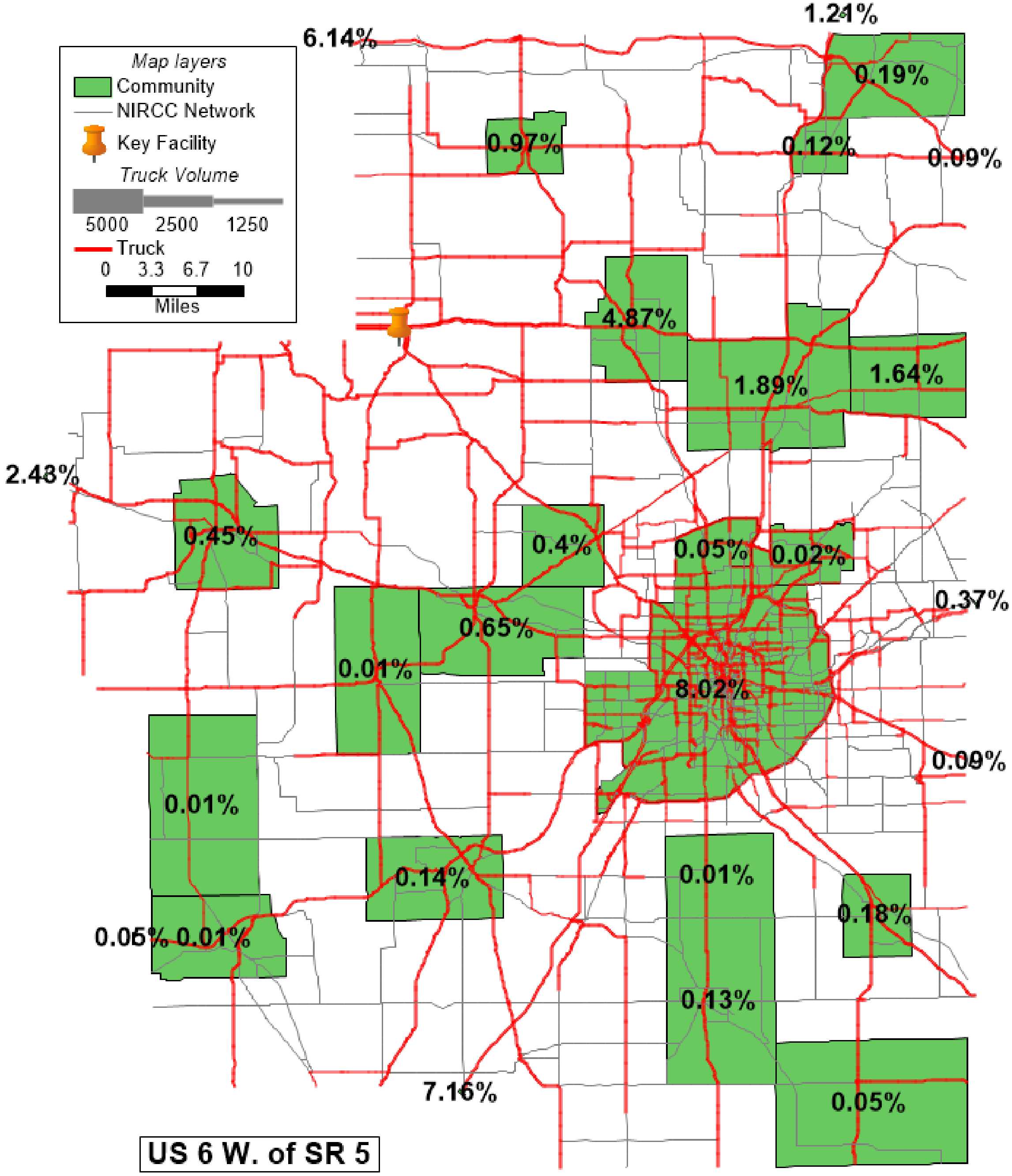


FIGURE 65. DEMAND AT US 6 W. OF SR 5


Map layers
Community
NIRCC Network
Key Facility
Truck Volume
5000 2500 1250
Truck
0 3.3 6.7 10
Miles
6.14%
1.24%
0.19%
0.97%
0.12%
0.09%
4.87%
1.89%
1.64%
2.48%
0.45%
0.4%
0.05%
0.02%
0.37%
0.65%
0.01%
8.02%
0.09%
0.01%
0.14%
0.01%
0.18%
0.06%
0.01%
0.13%
7.16%
0.05%
US 6 W. of SR 5

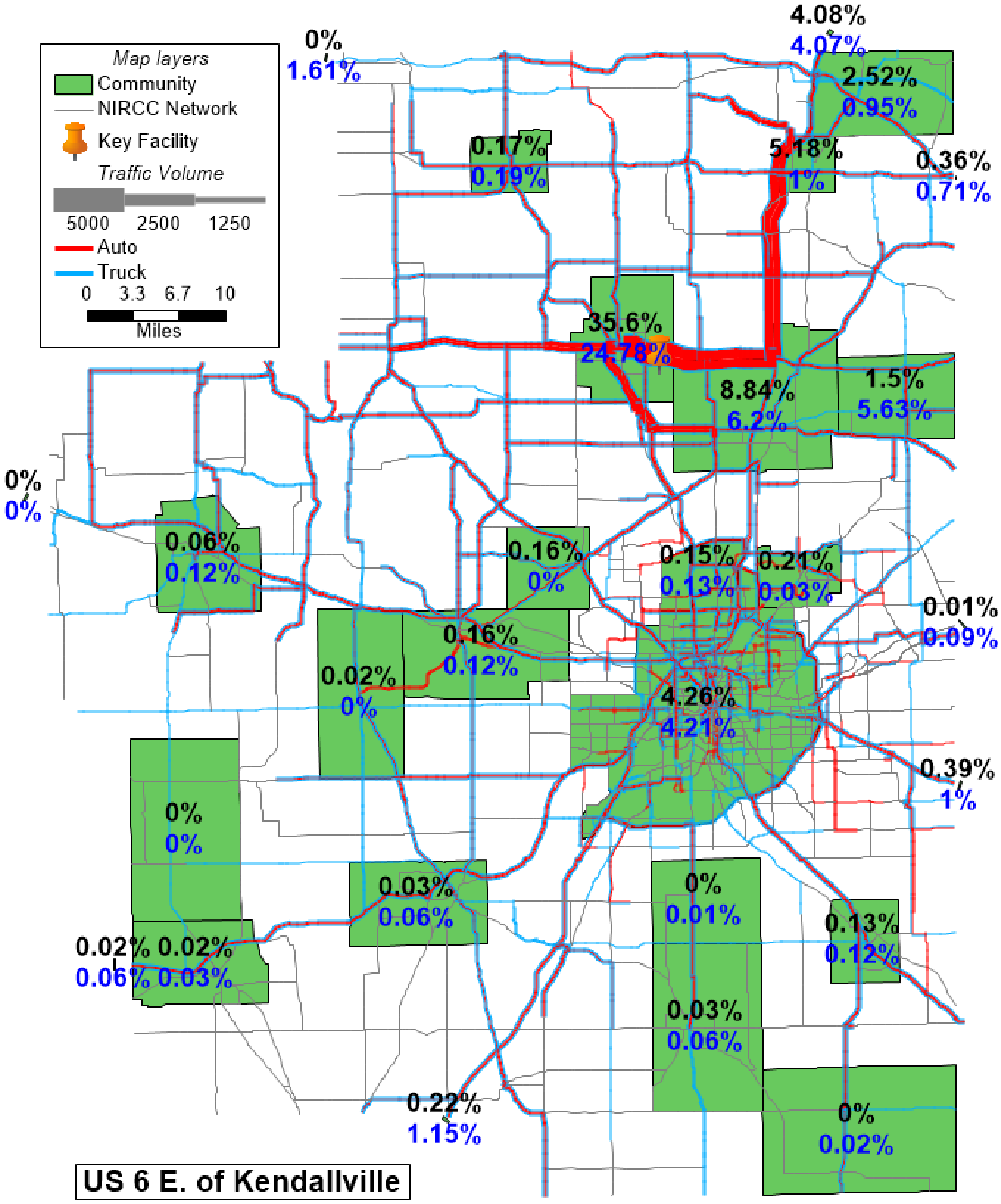
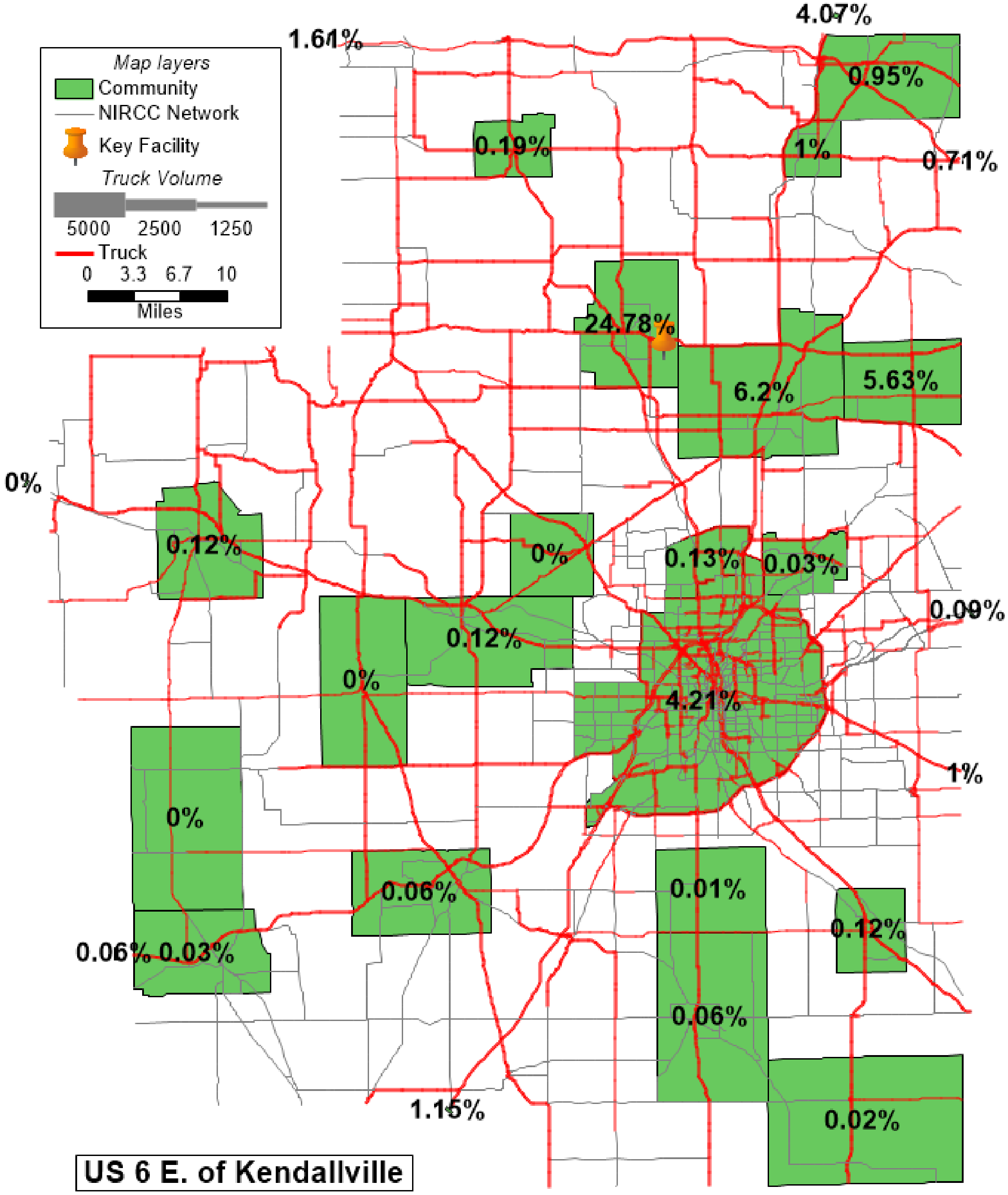


FIGURE 66. DEMAND AT US 6 E. OF KENDALLVILLE


Map layers
Community
NIRCC Network
Key Facility
Truck Volume
5000
2500
1250
Truck
0
3.3
6.7
10
Miles
4.07%
1.61%
0.95%
0.19%
1%
0.71%
24.78%
6.2%
5.63%
0%
0.12%
0%
0.13%
0.03%
0.09%
0.12%
0%
4.21%
1%
0%
0.06%
0.01%
0.12%
0.06%
0.03%
0.06%
1.15%
0.02%
US 6 E. of Kendallville

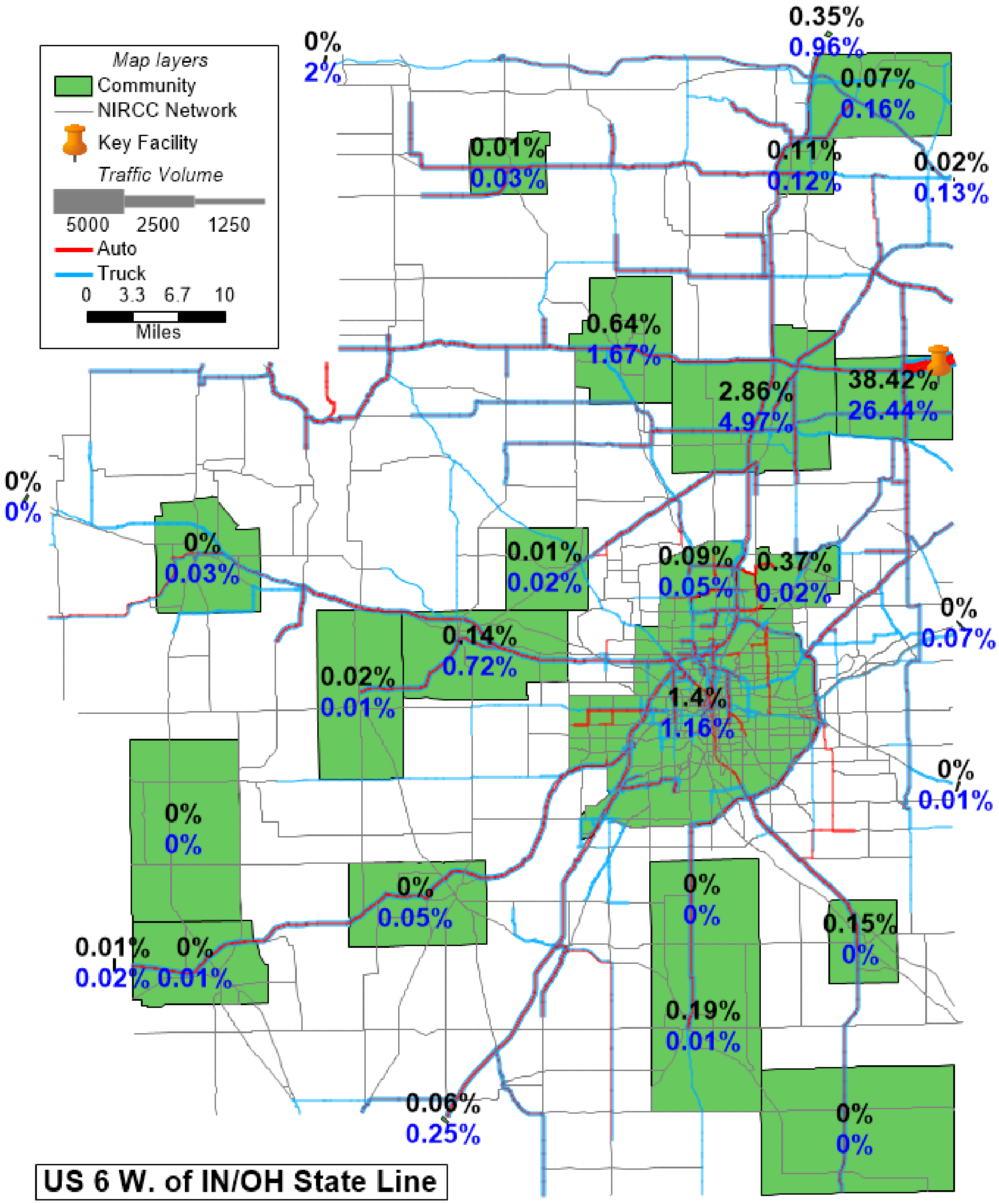
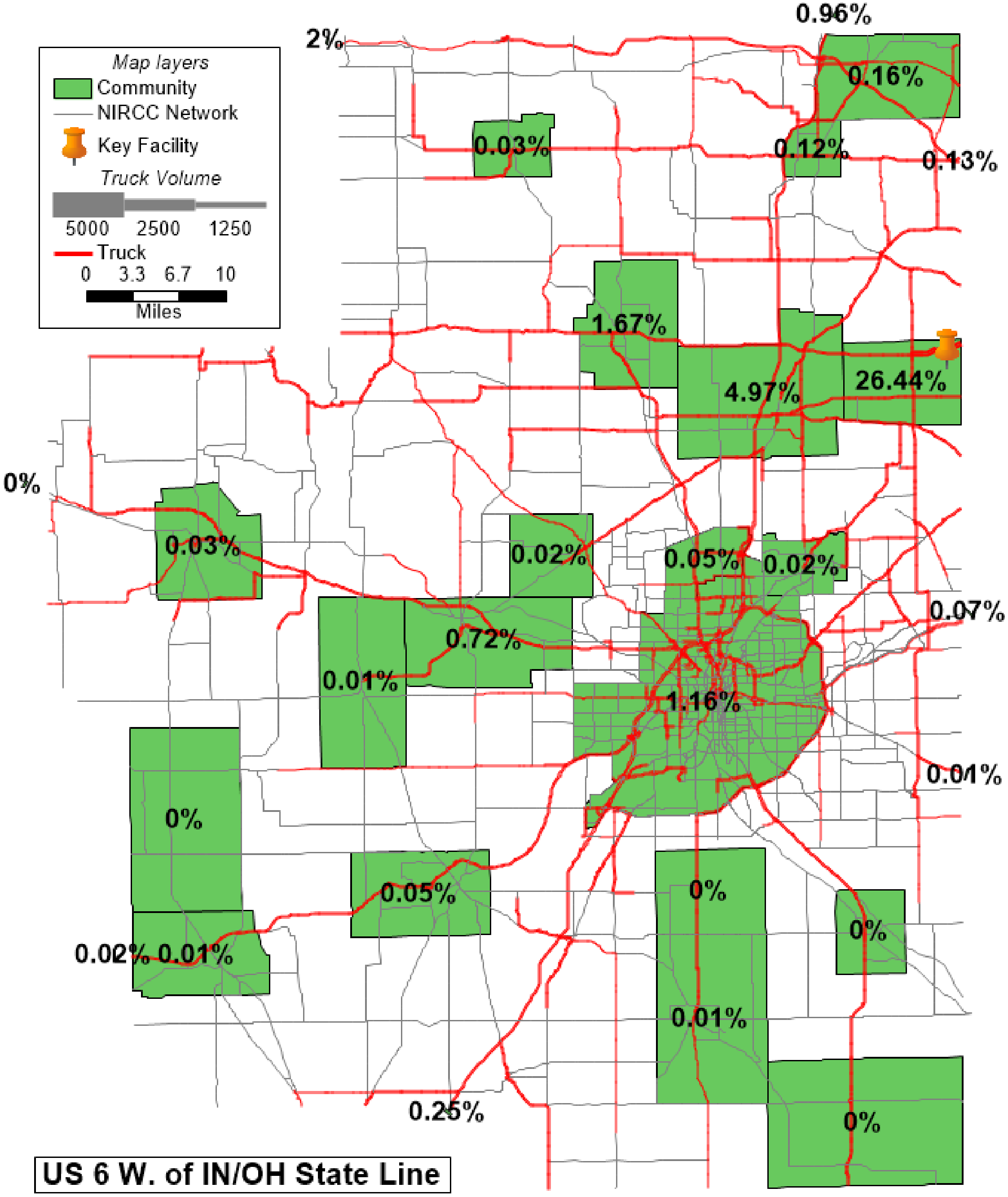


FIGURE 67. DEMAND AT US 6 W. OF IN/OH STATE LINE



US 6 W. of IN/OH State Line

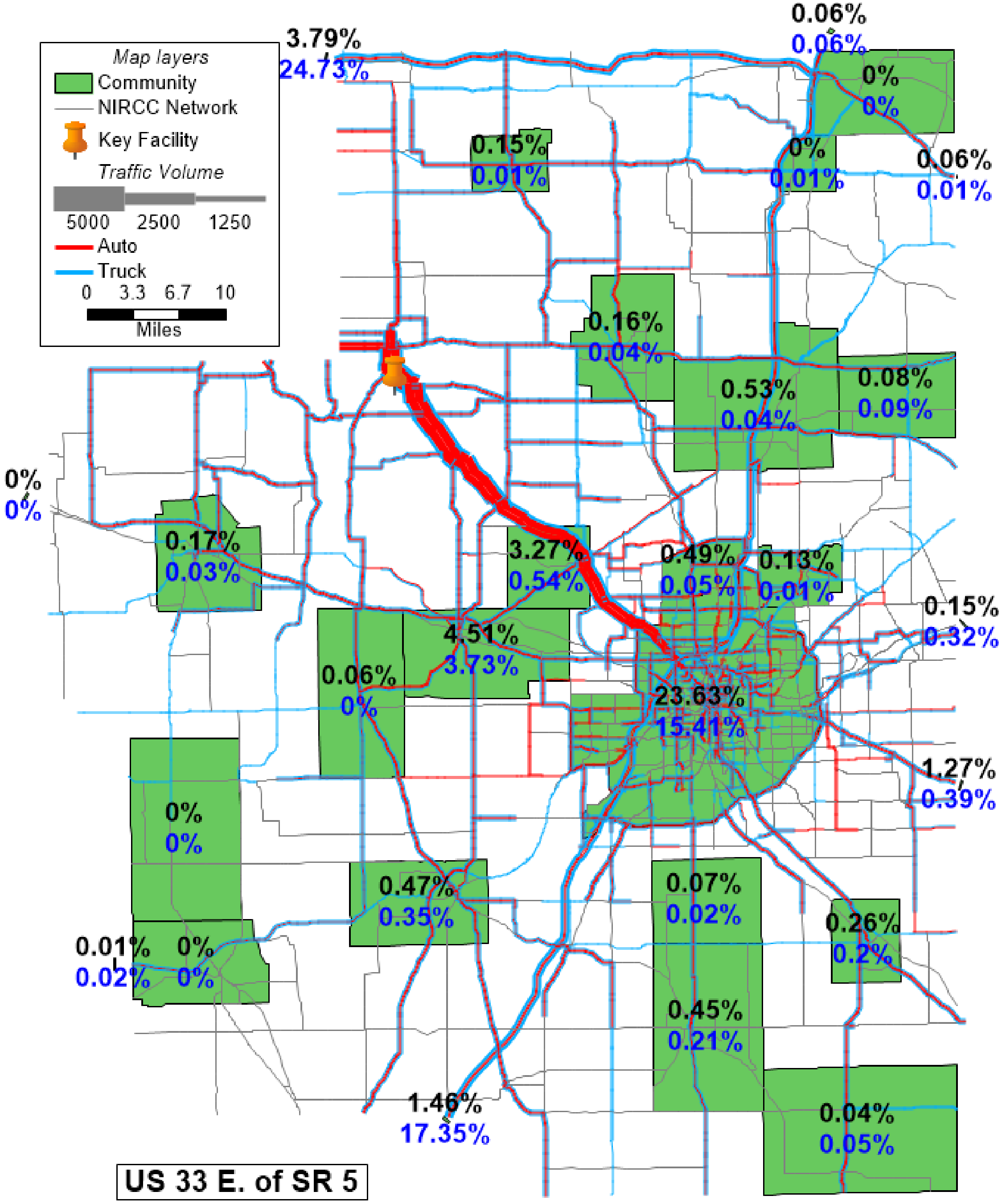
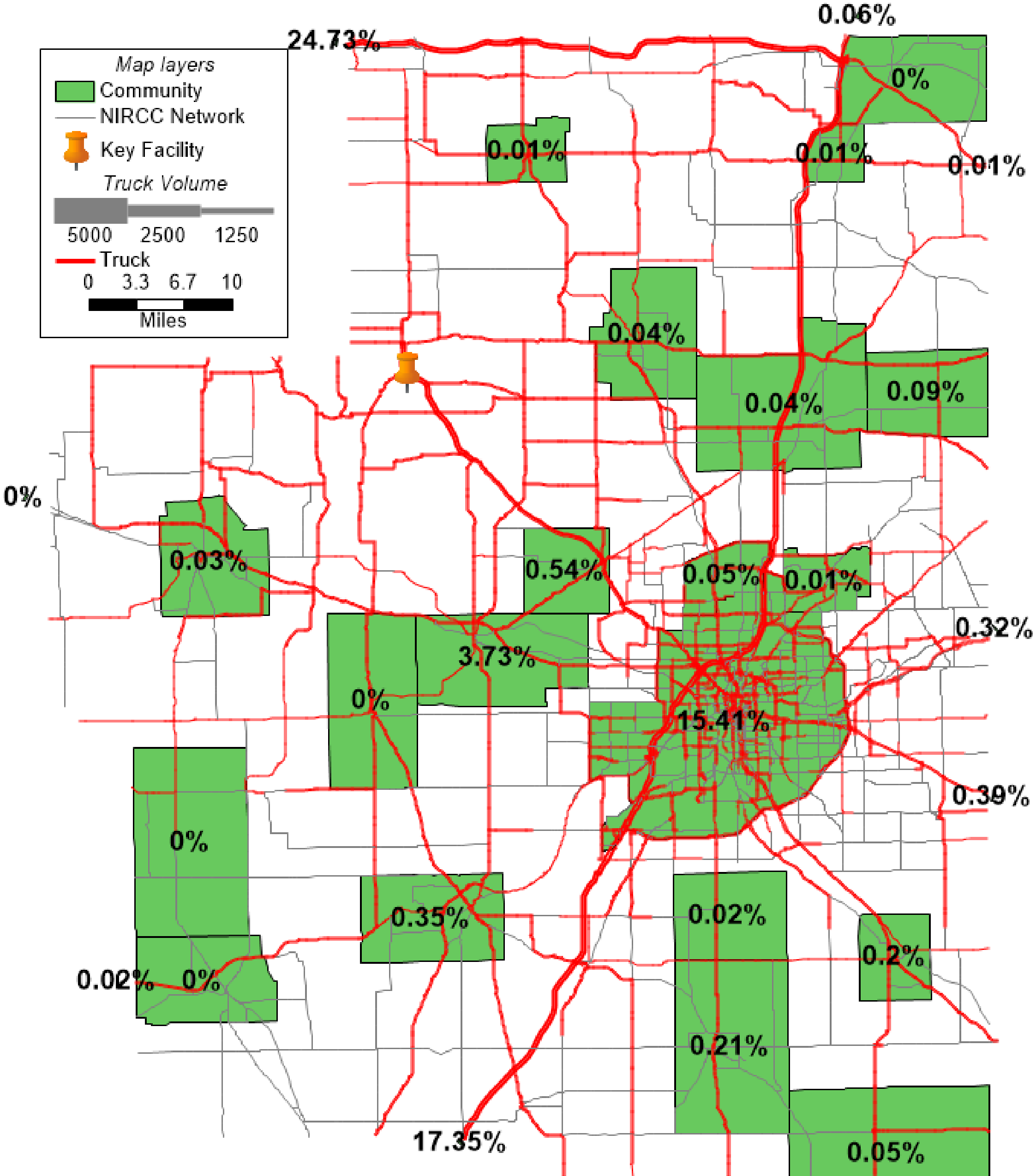


FIGURE 68. DEMAND AT US 33 E. OF SR 5


Map layers
Community
NIRCC Network
Key Facility
Truck Volume
5000
2500
1250
Truck
0
3.3
6.7
10
Miles
24.73%
0.06%
0%
0.01%
0.01%
0.01%
0.04%
0.04%
0.09%
0%
0.03%
0.54%
0.05%
0.01%
0.32%
3.73%
0%
15.41%
0.39%
0%
0.35%
0.02%
0.2%
0.02%
0%
0.21%
17.35%
0.05%

US 33 E. of SR 5

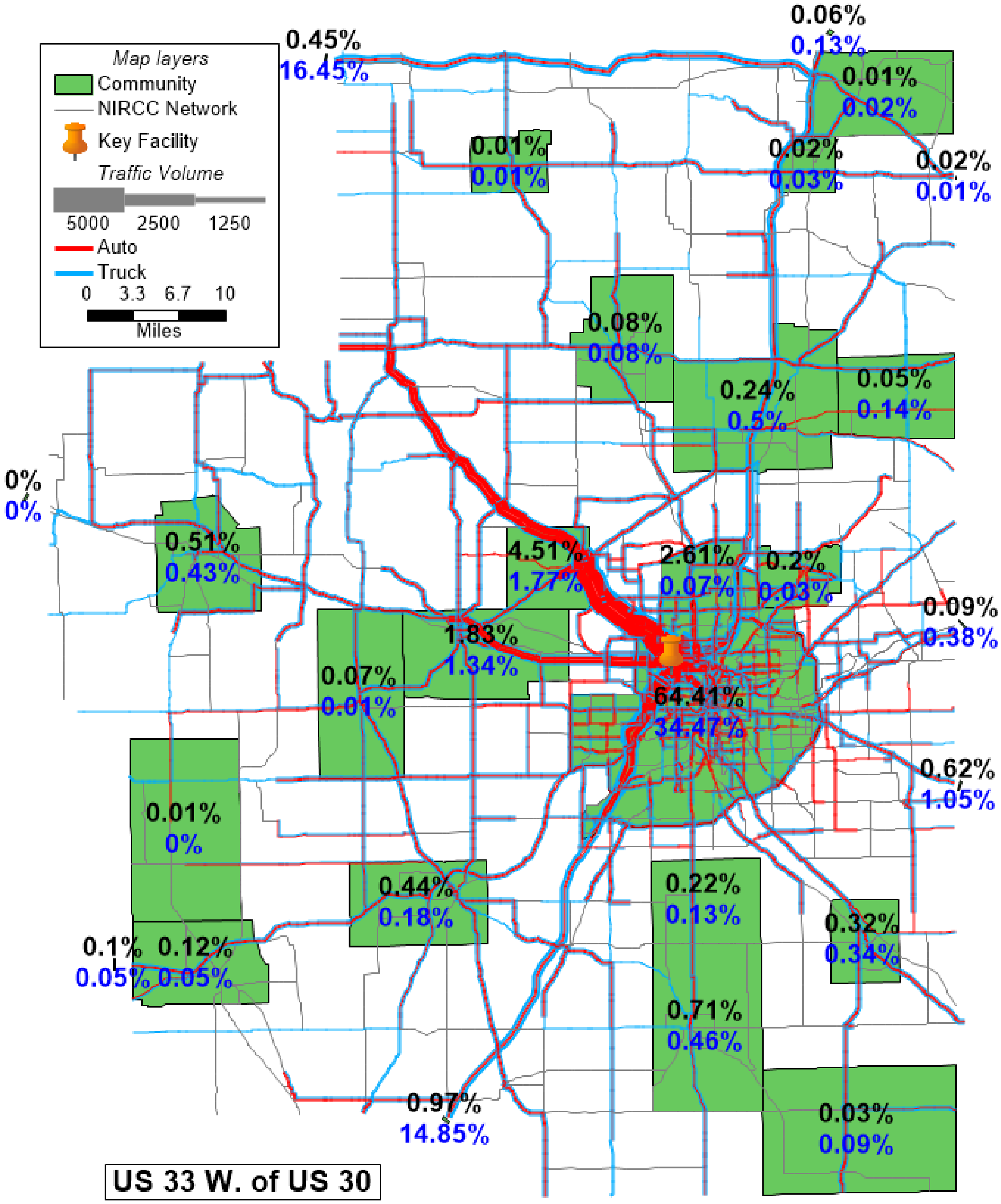
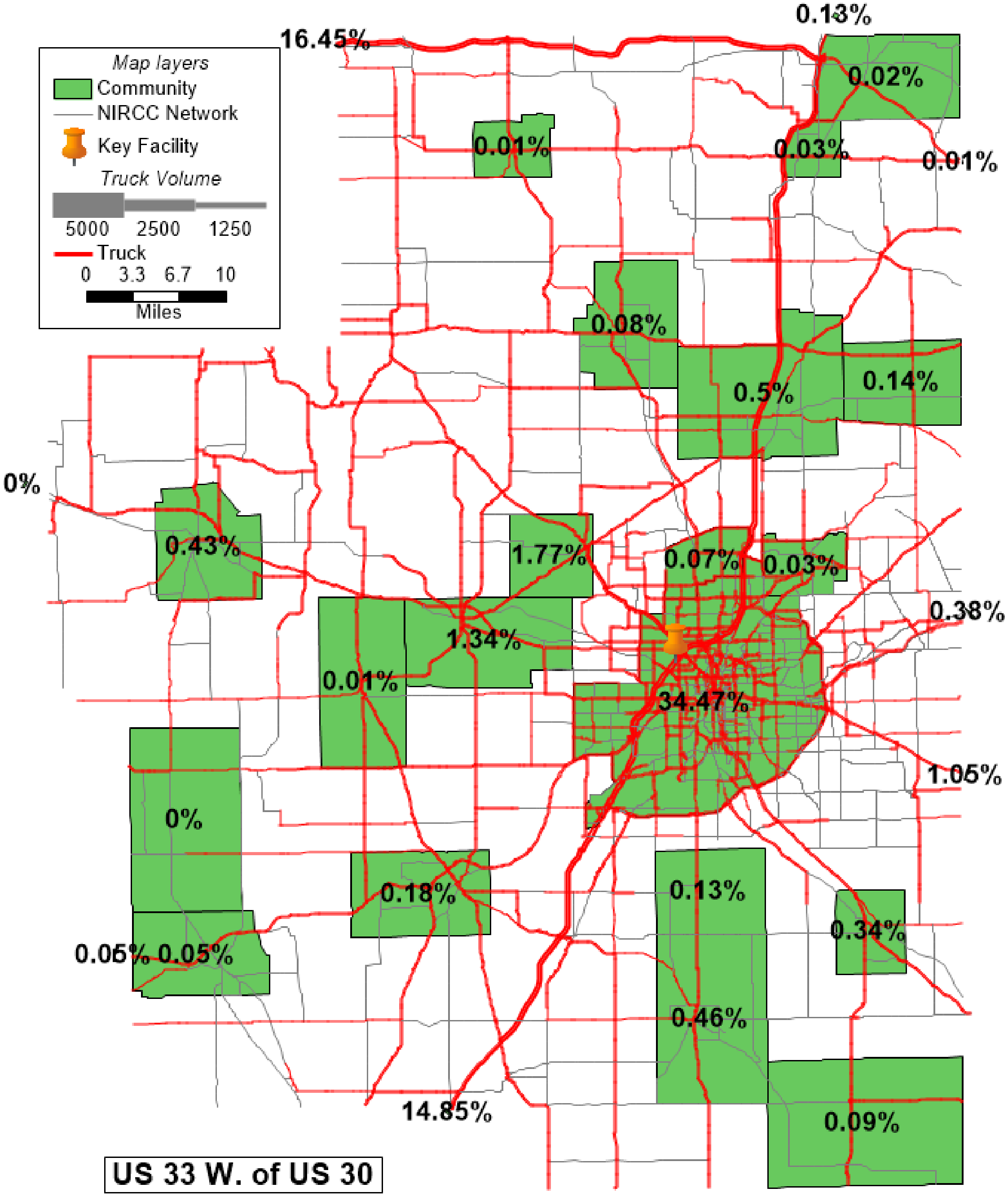


FIGURE 69. DEMAND AT US 33 W. OF US 30



US 33 W. of US 30

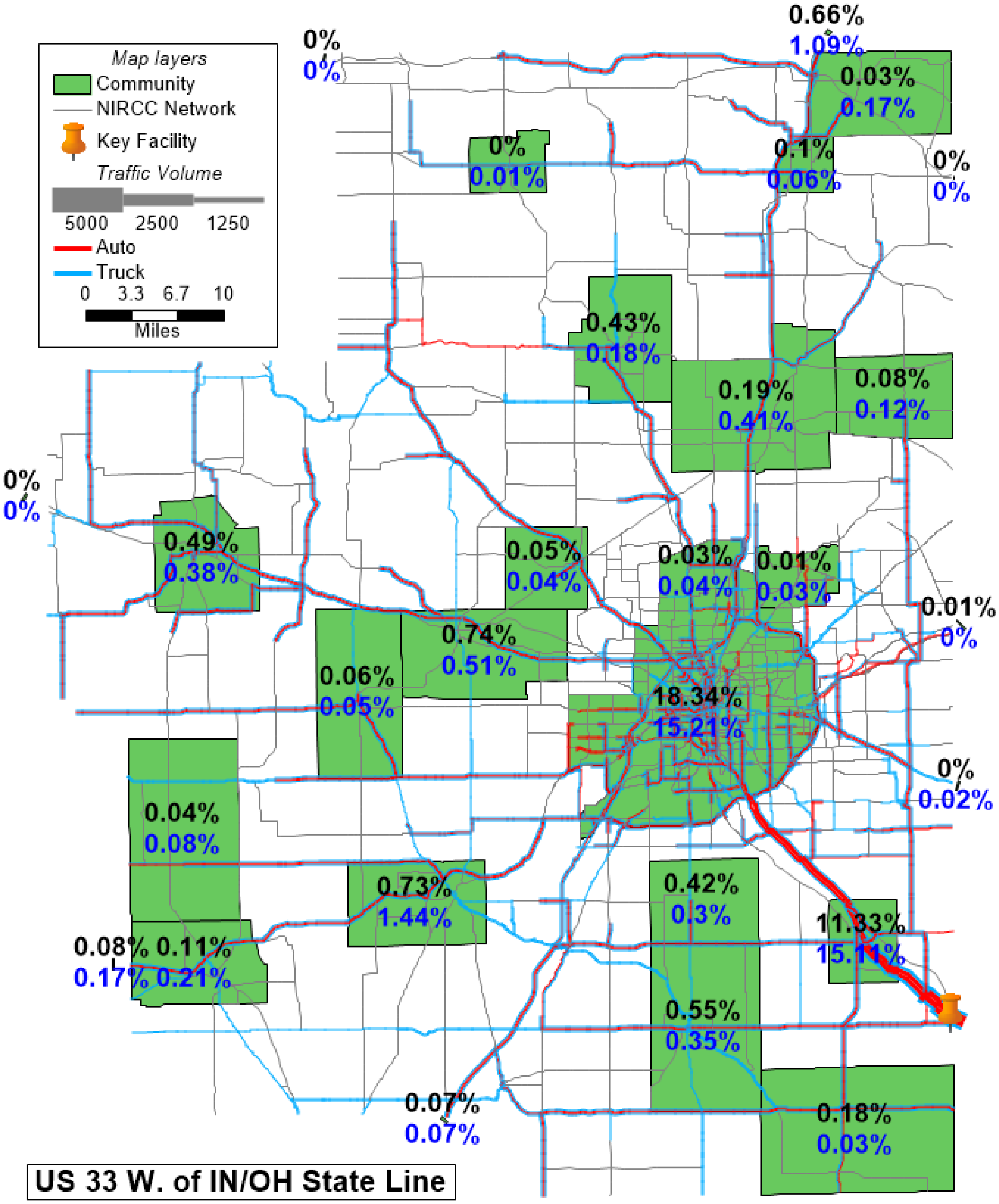
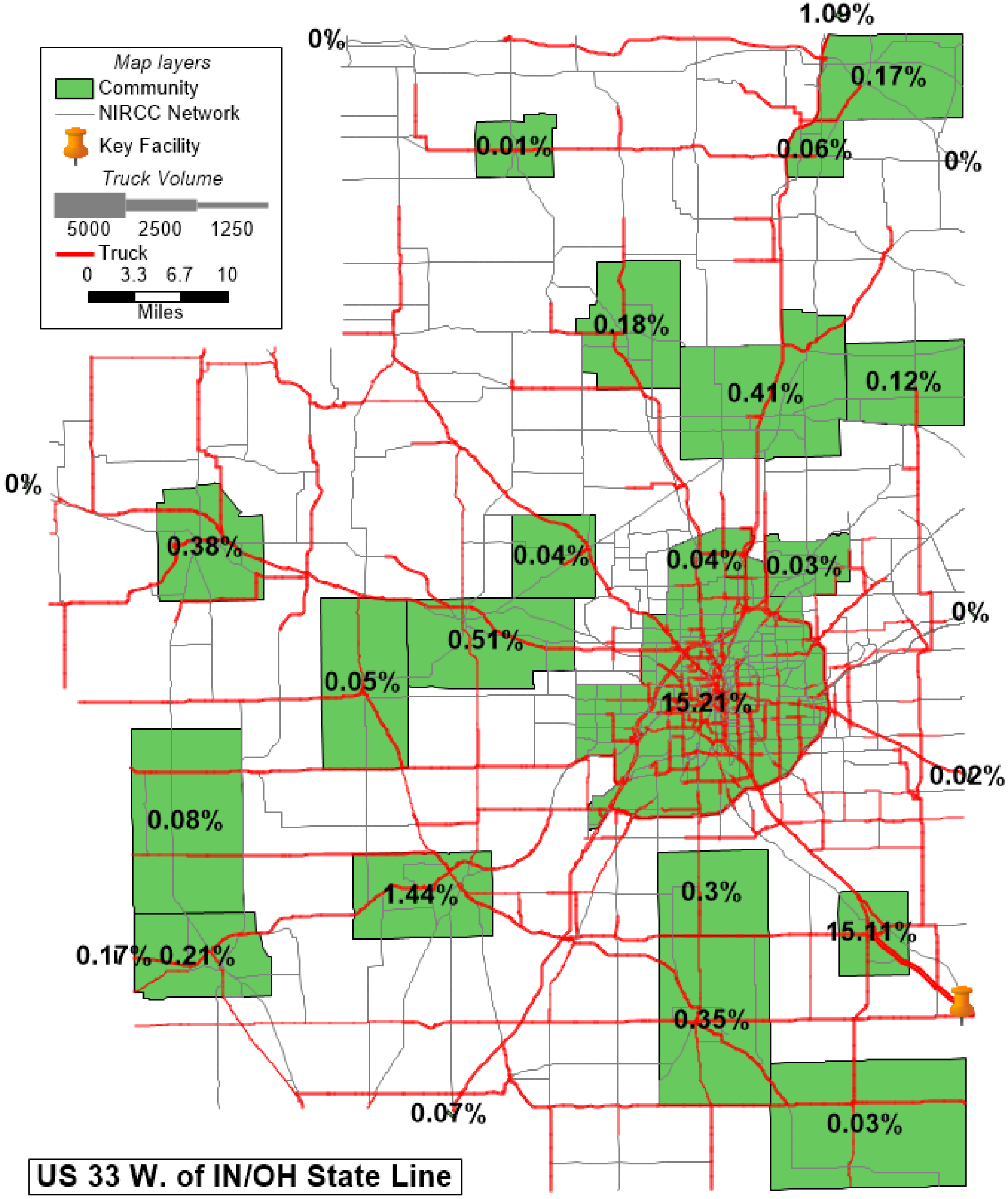


FIGURE 70. DEMAND AT US 33 W. OF IN/OH STATE LINE



US 33 W. of IN/OH State Line

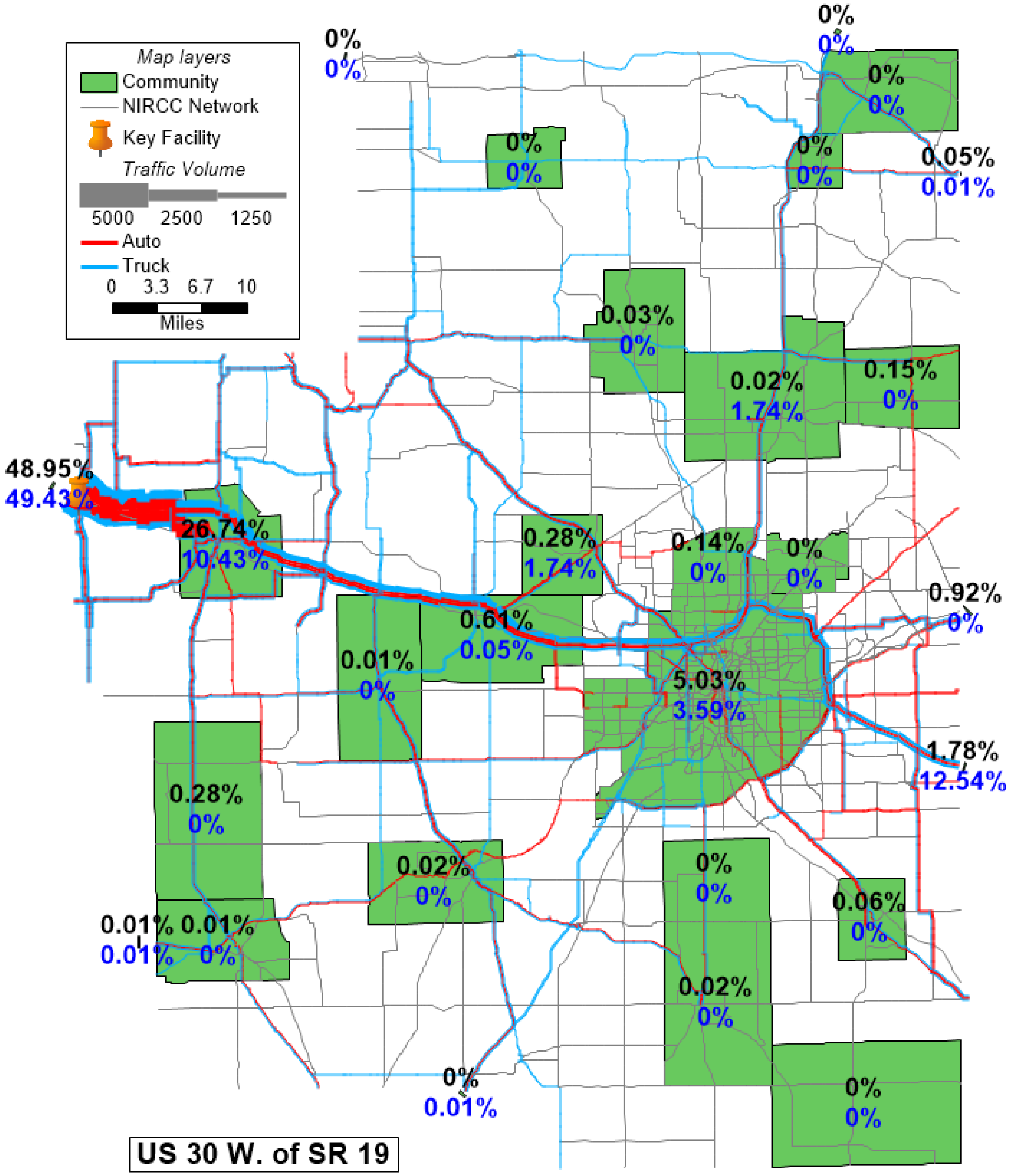
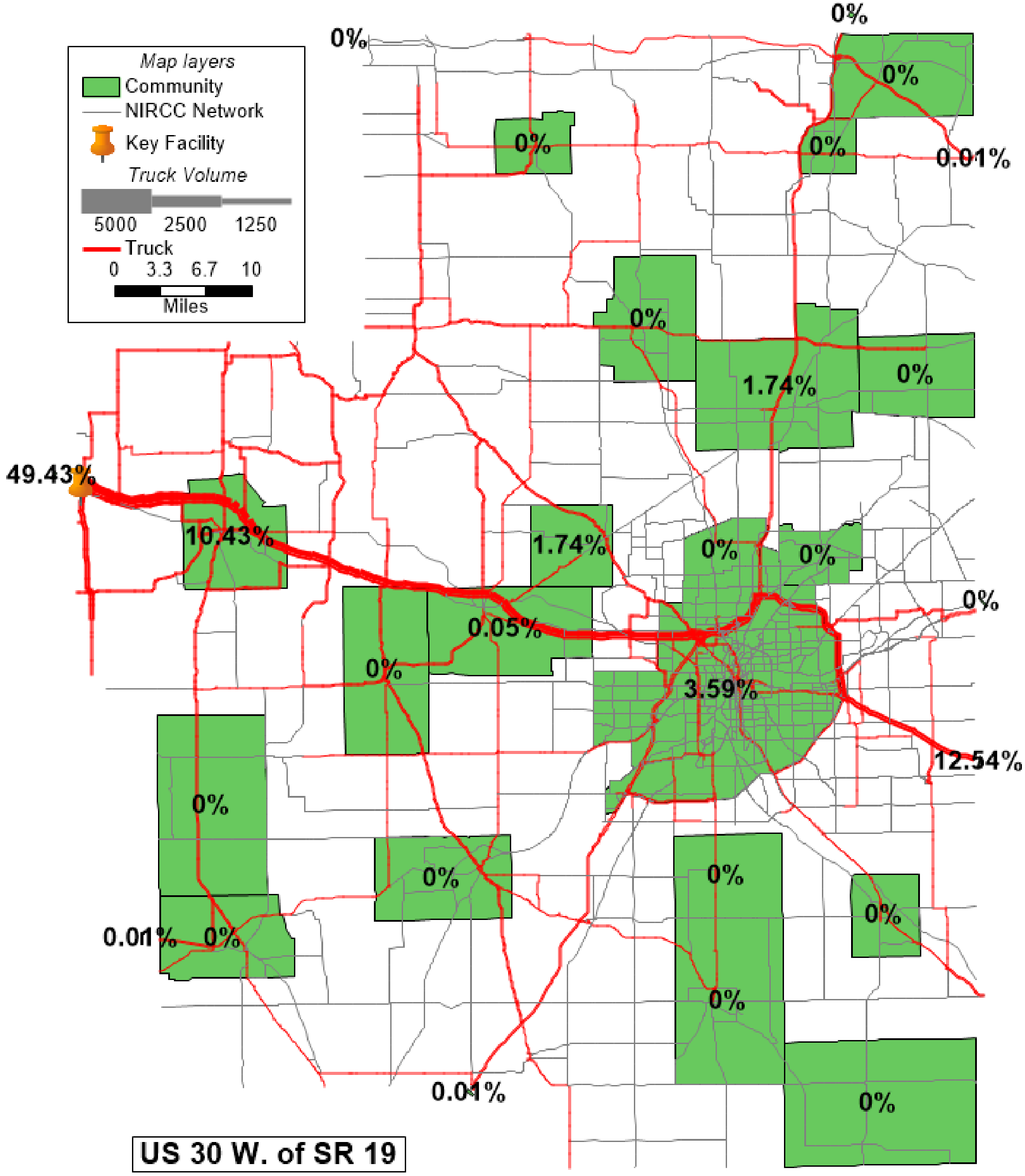


FIGURE 71. DEMAND AT US 30 W. OF SR 19



US 30 W. of SR 19

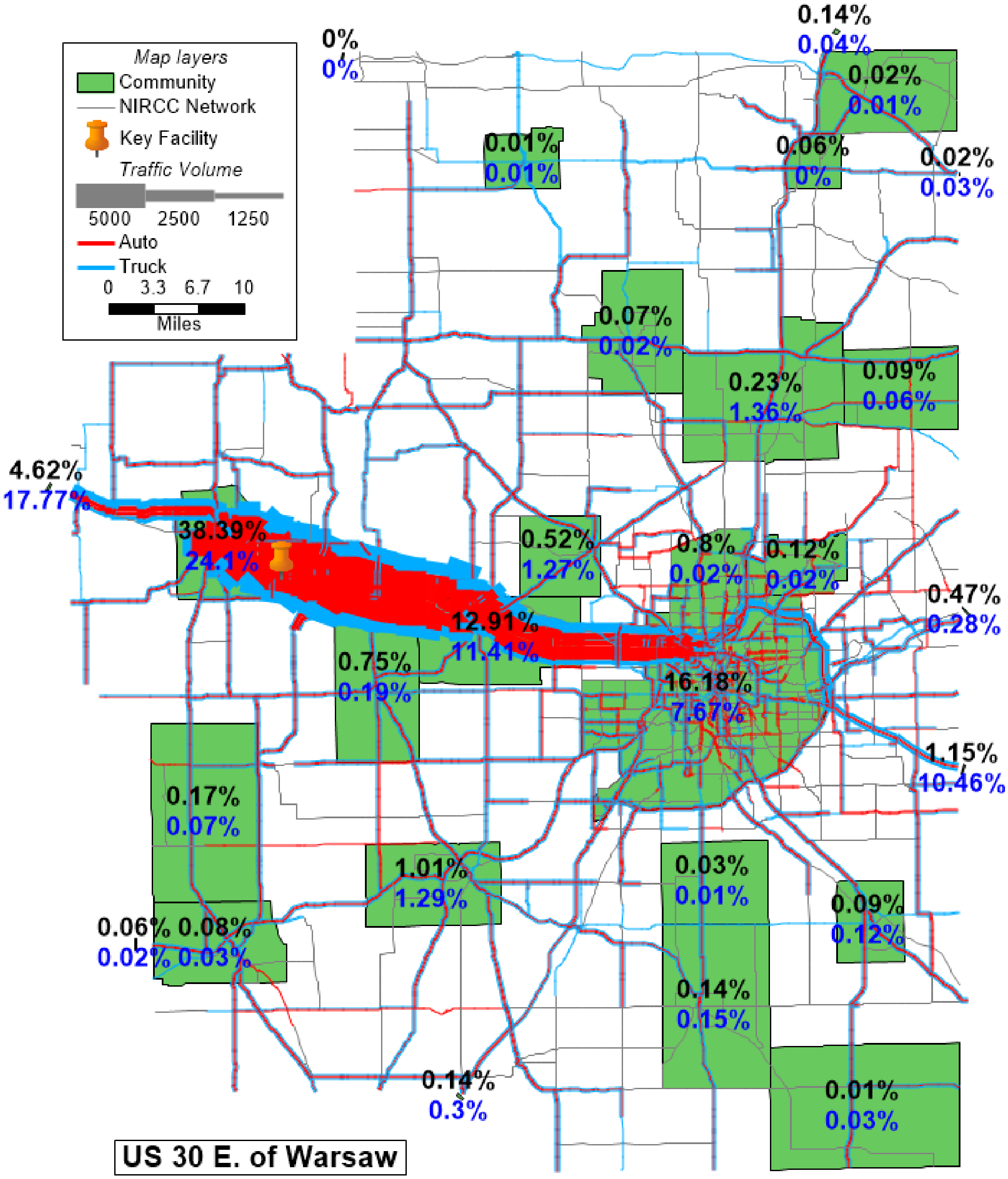
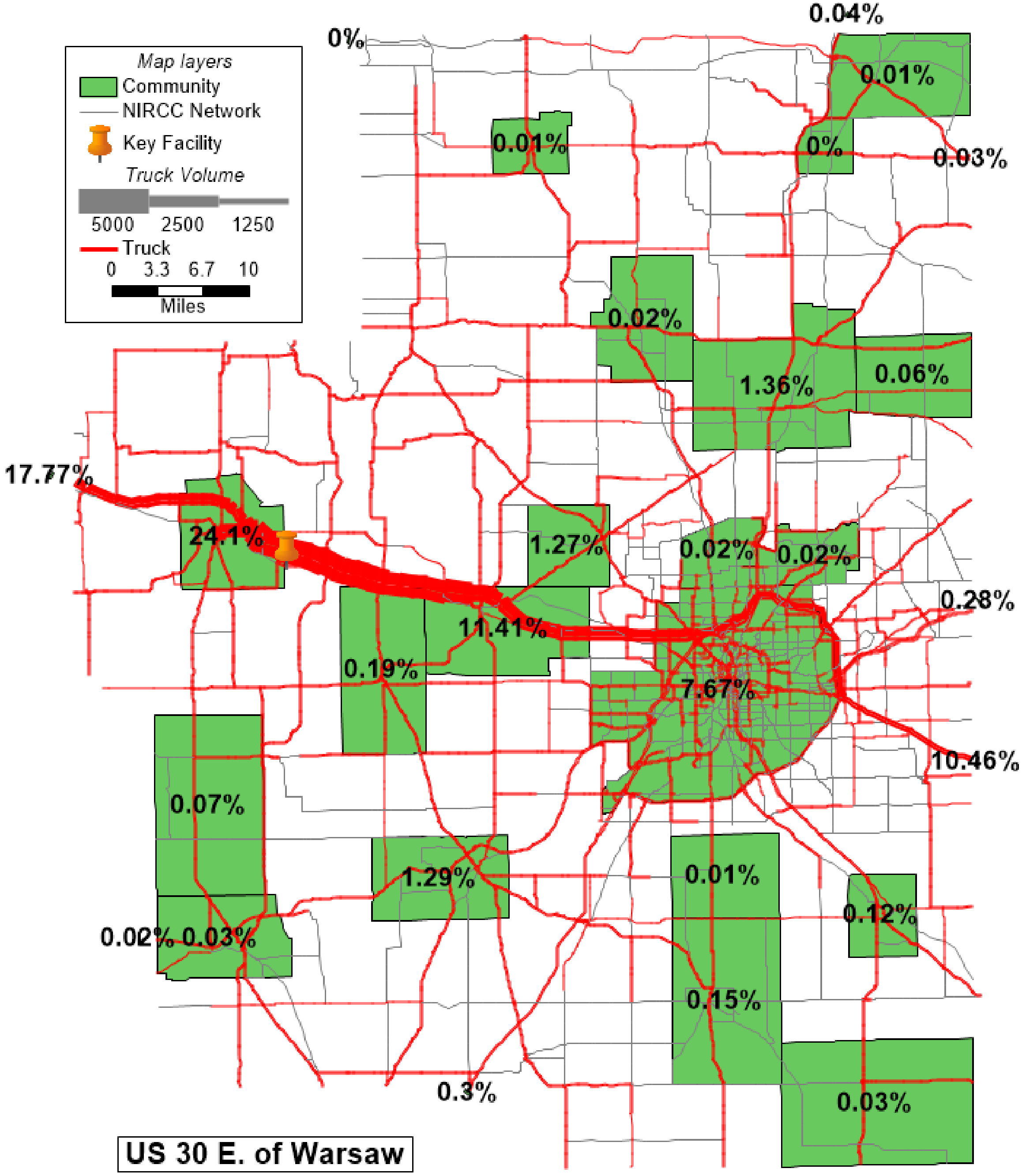


FIGURE 72. DEMAND AT US 30 E. OF WARSAW


Map layers
Community
NIRCC Network
Key Facility
Truck Volume
5000 2500 1250
Truck
0 3.3 6.7 10
Miles
0%
0.04%
0.01%
0.01%
0%
0.03%
0.02%
1.36%
0.06%
17.77%
24.1%
1.27%
0.02%
0.02%
0.28%
11.41%
0.19%
7.67%
10.46%
0.07%
1.29%
0.01%
0.12%
0.02%
0.03%
0.15%
0.3%
0.03%
US 30 E. of Warsaw

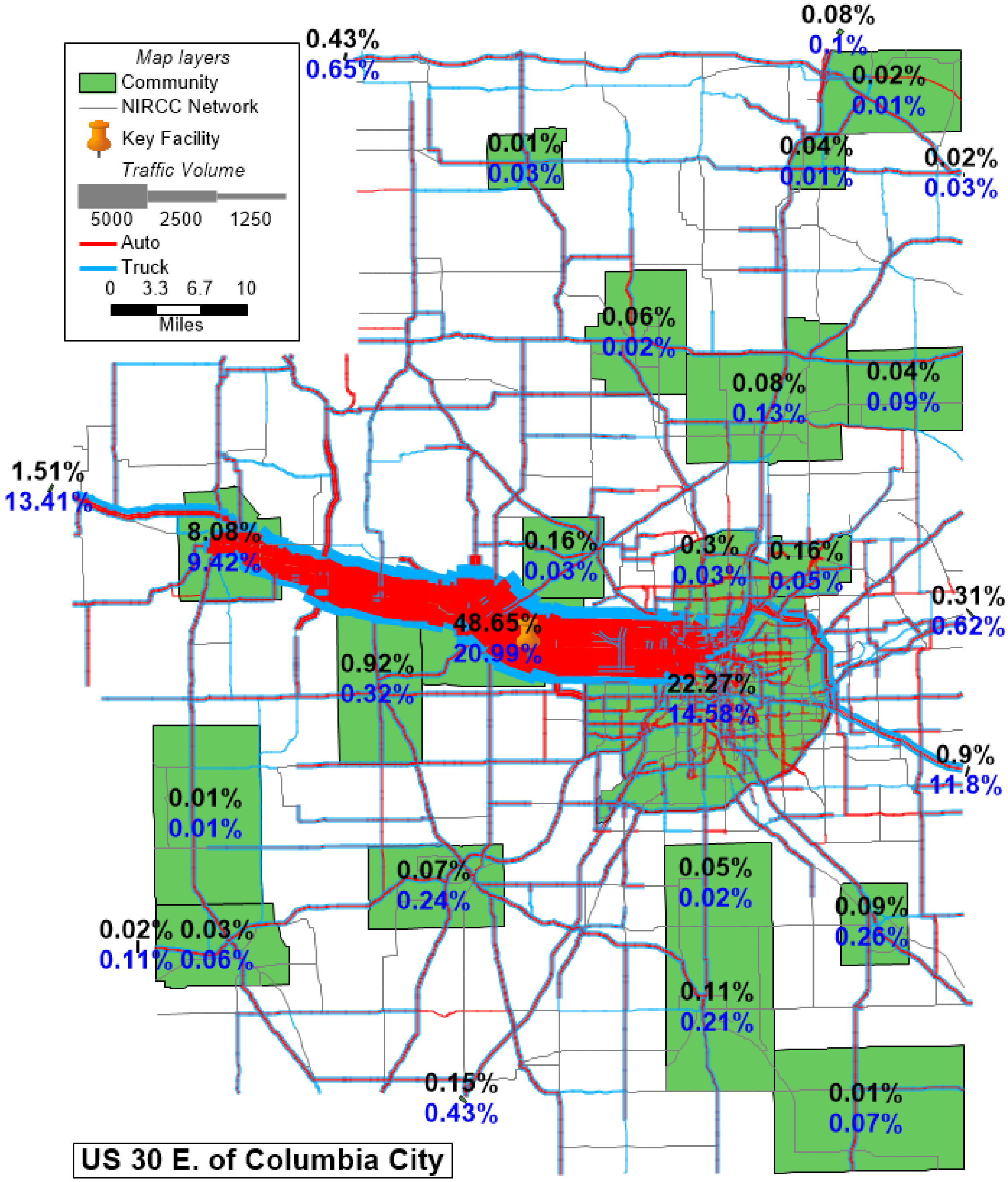
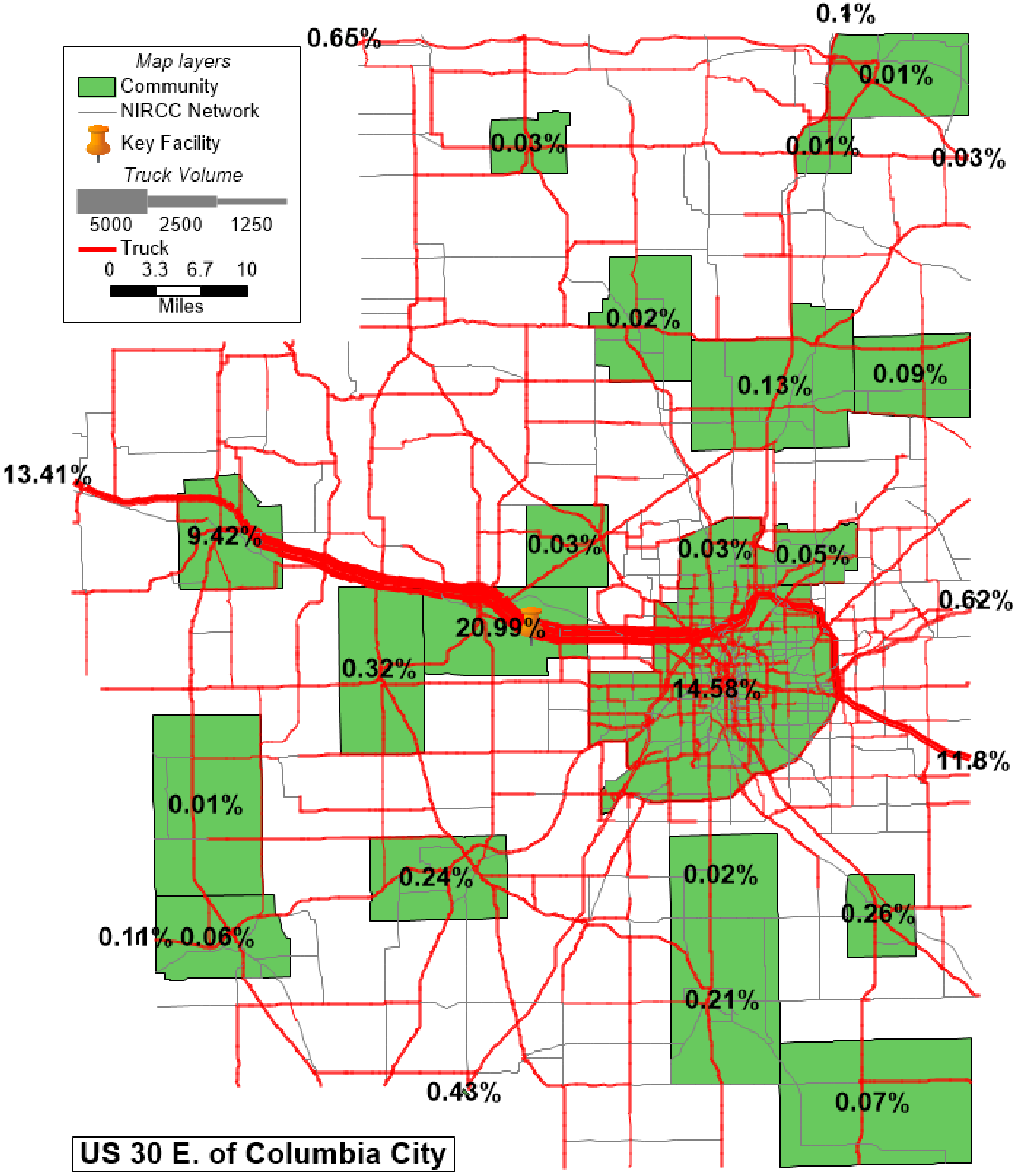


FIGURE 73. DEMAND AT US 30 E. OF COLUMBIA CITY



US 30 E. of Columbia City

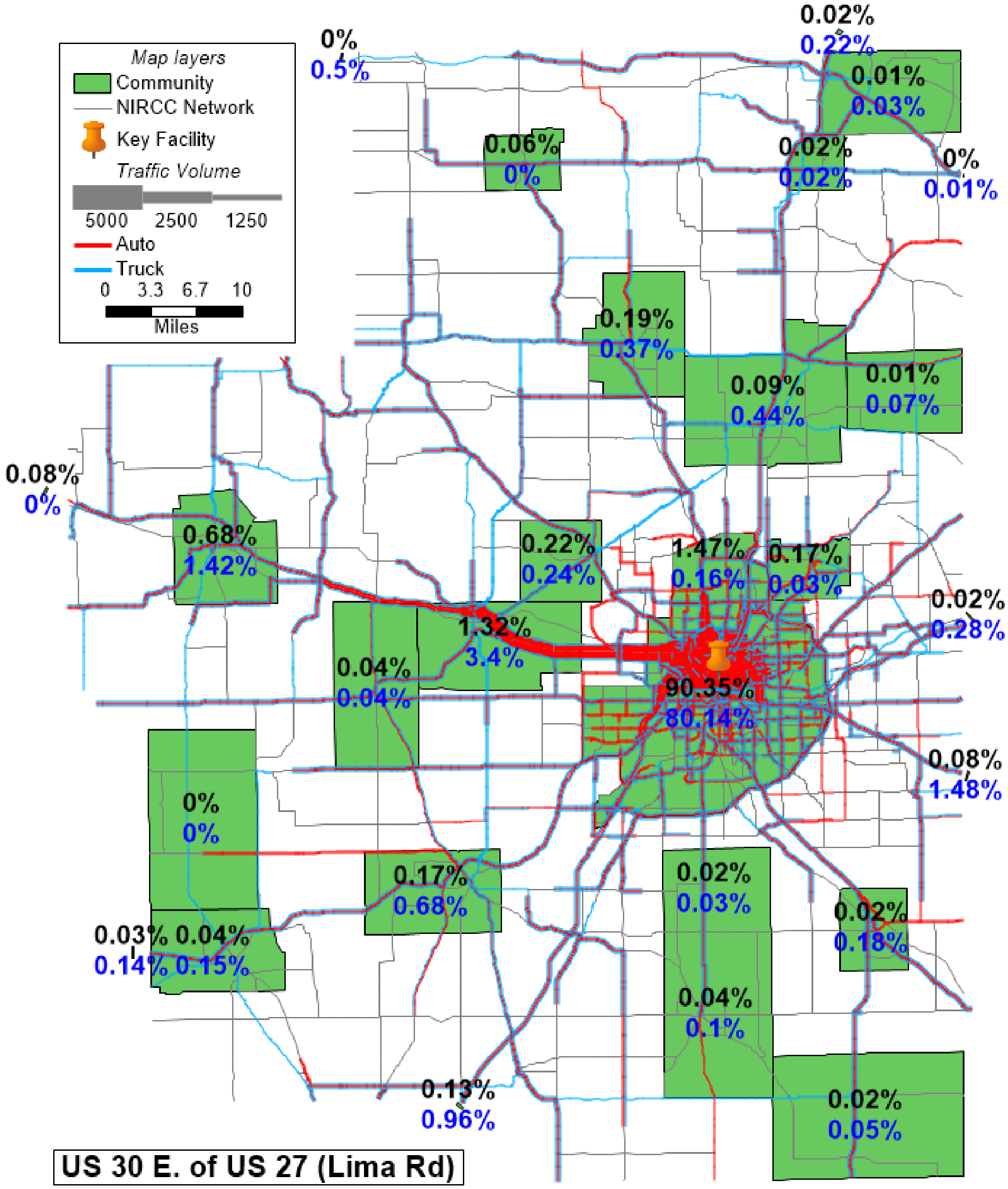
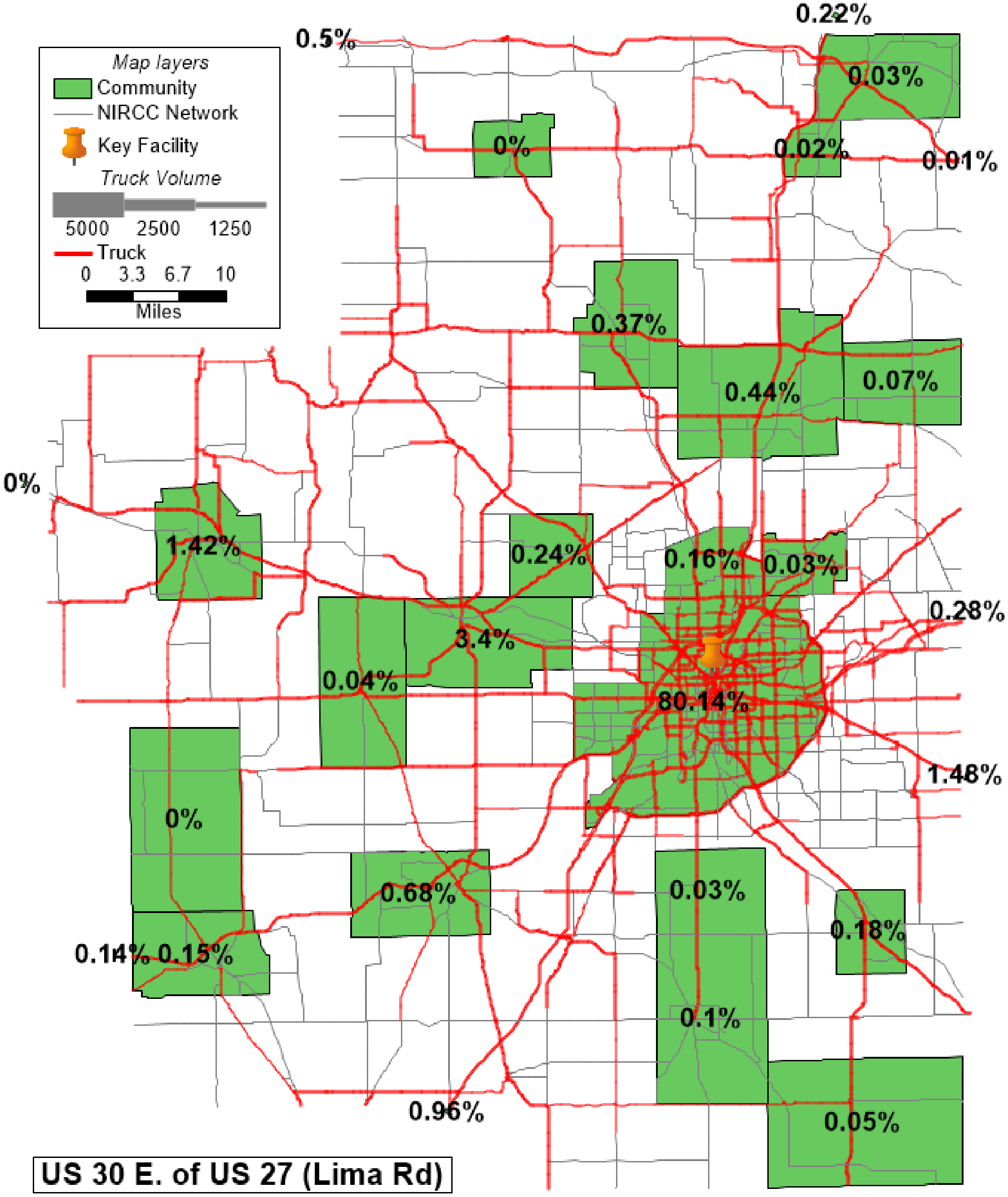


FIGURE 74. DEMAND AT US 30 E. OF US 27 (LIMA RD.)



US 30 E. of US 27 (Lima Rd)

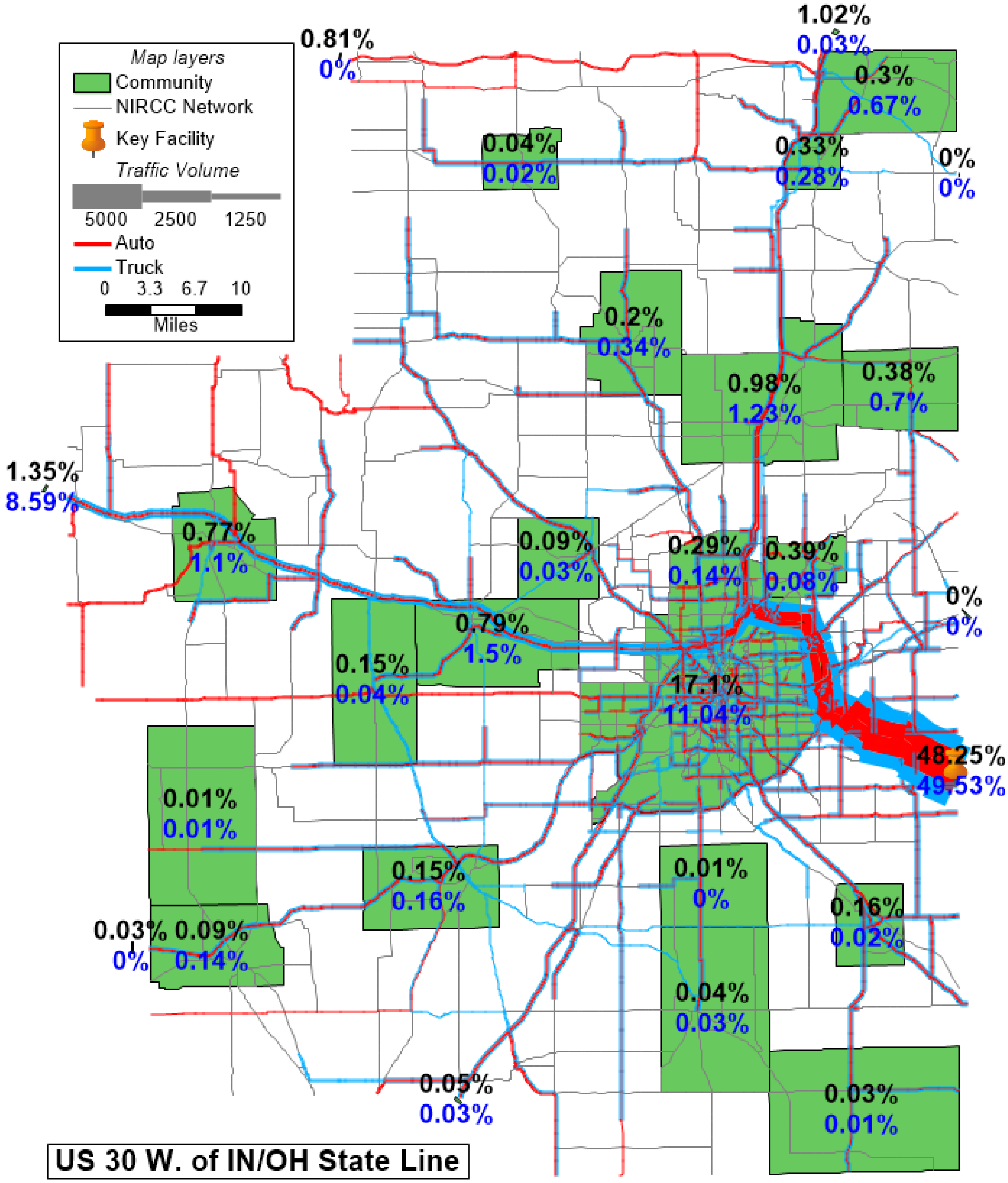
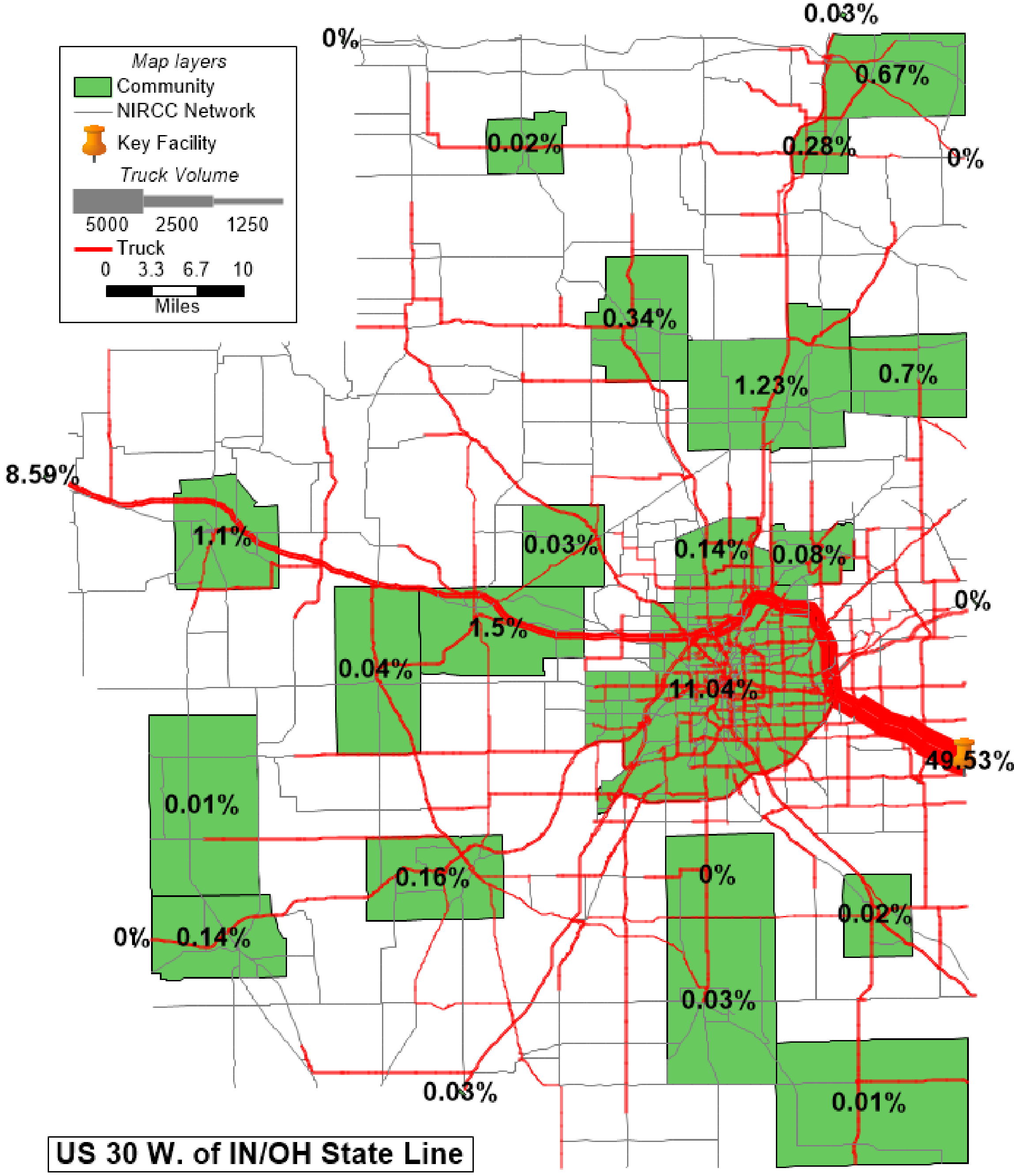


FIGURE 75. DEMAND AT US 30 W. OF IN/OH STATE LINE



US 30 W. of IN/OH State Line

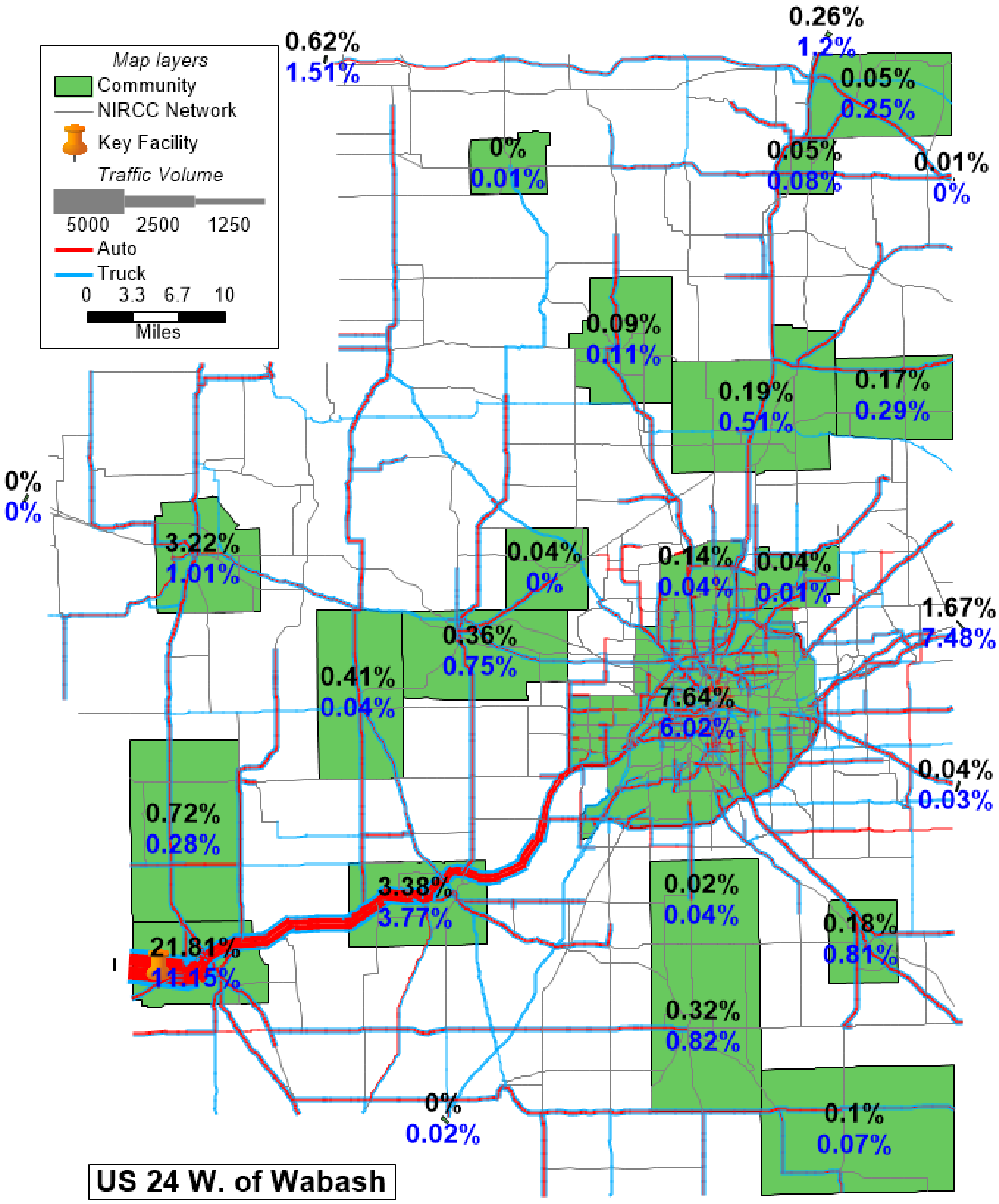
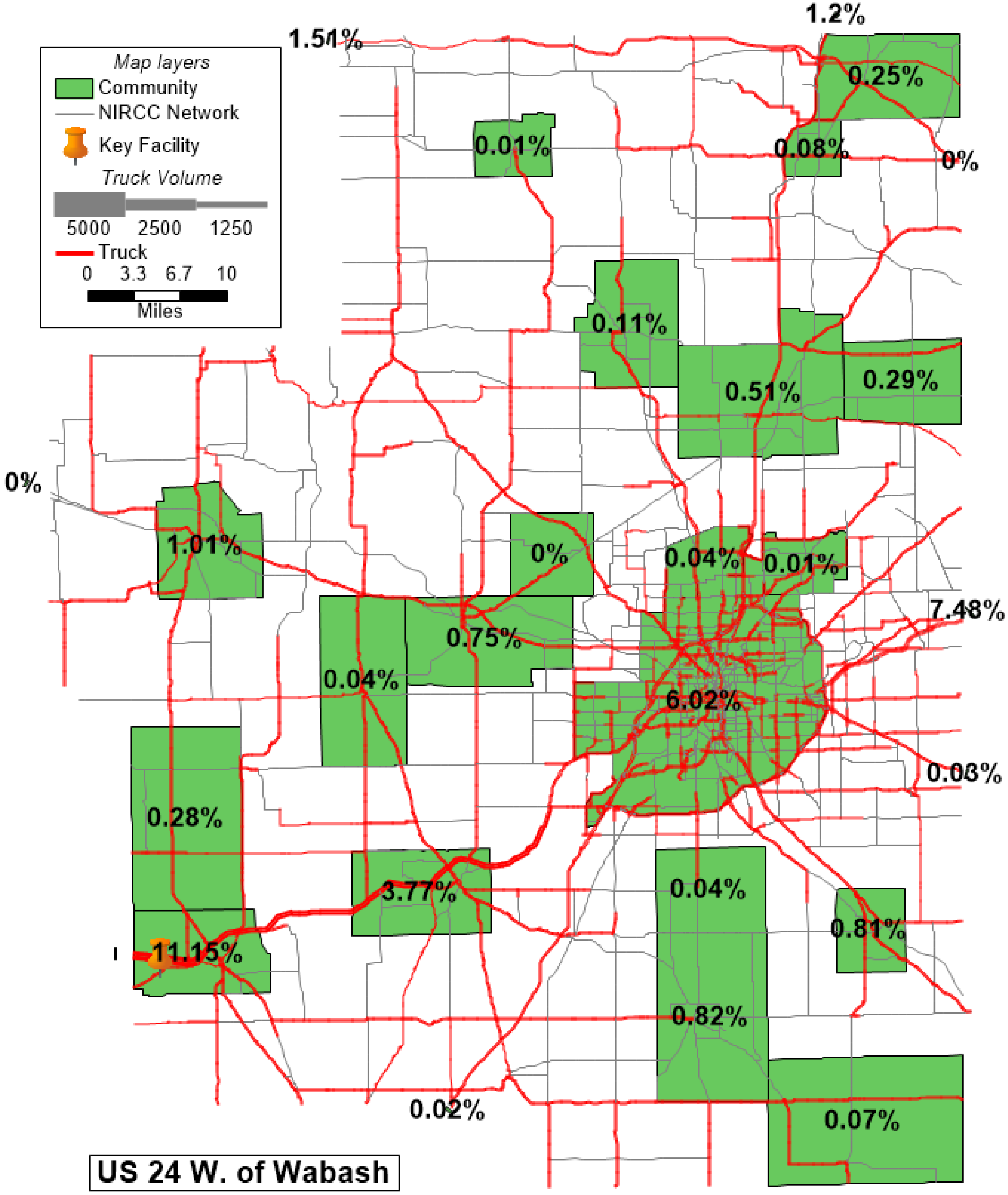


FIGURE 76. DEMAND AT US 24 W. OF WABASH


Map layers
Community
NIRCC Network
Key Facility
Truck Volume
5000
2500
1250
Truck
0
3.3
6.7
10
Miles
1.54%
1.2%
0.25%
0.01%
0.08%
0%
0.11%
0.51%
0.29%
0%
1.01%
0%
0.04%
0.01%
7.48%
0.75%
0.04%
6.02%
0.03%
0.28%
3.77%
0.04%
0.81%
11.15%
0.82%
0.02%
0.07%
US 24 W. of Wabash

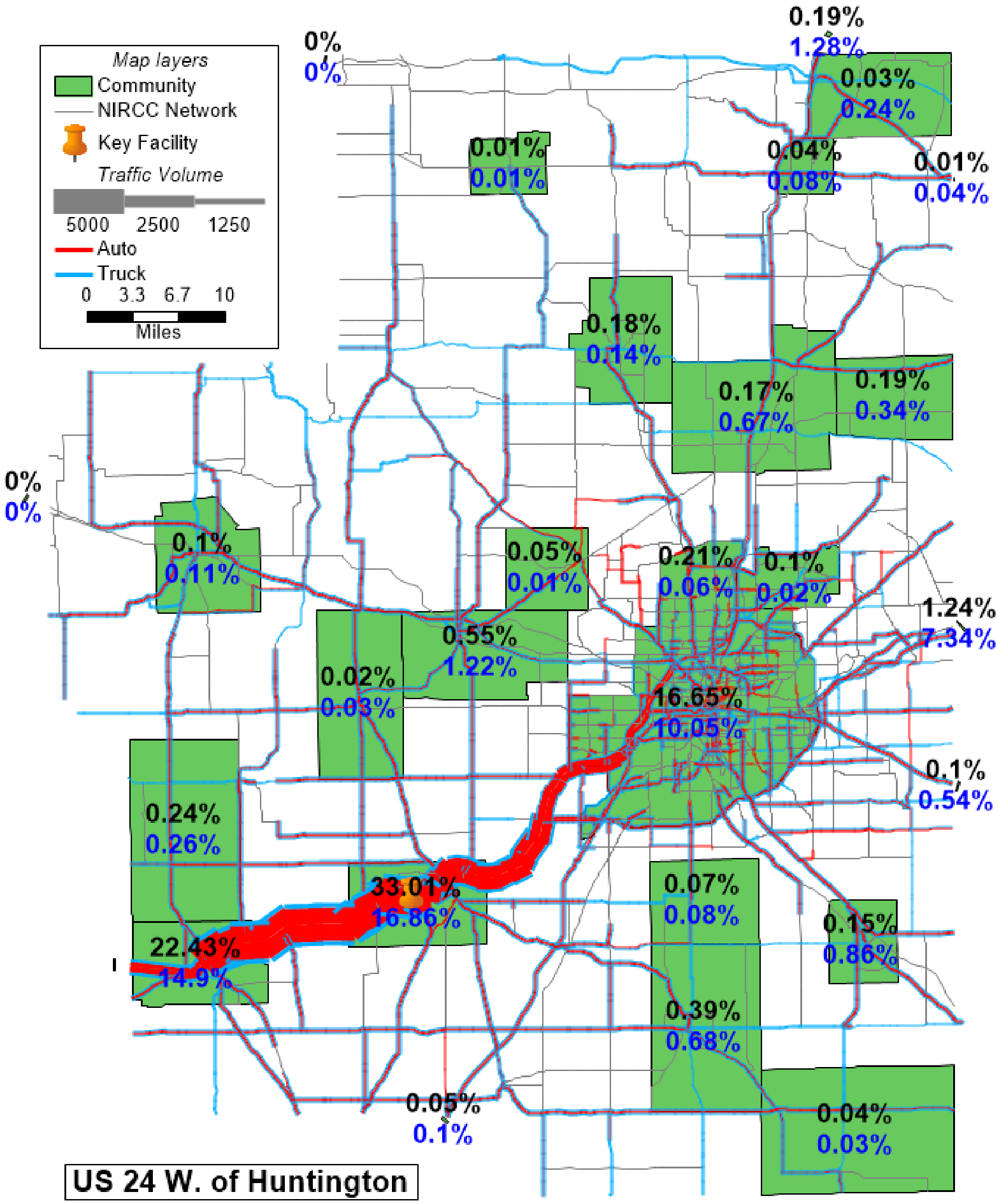
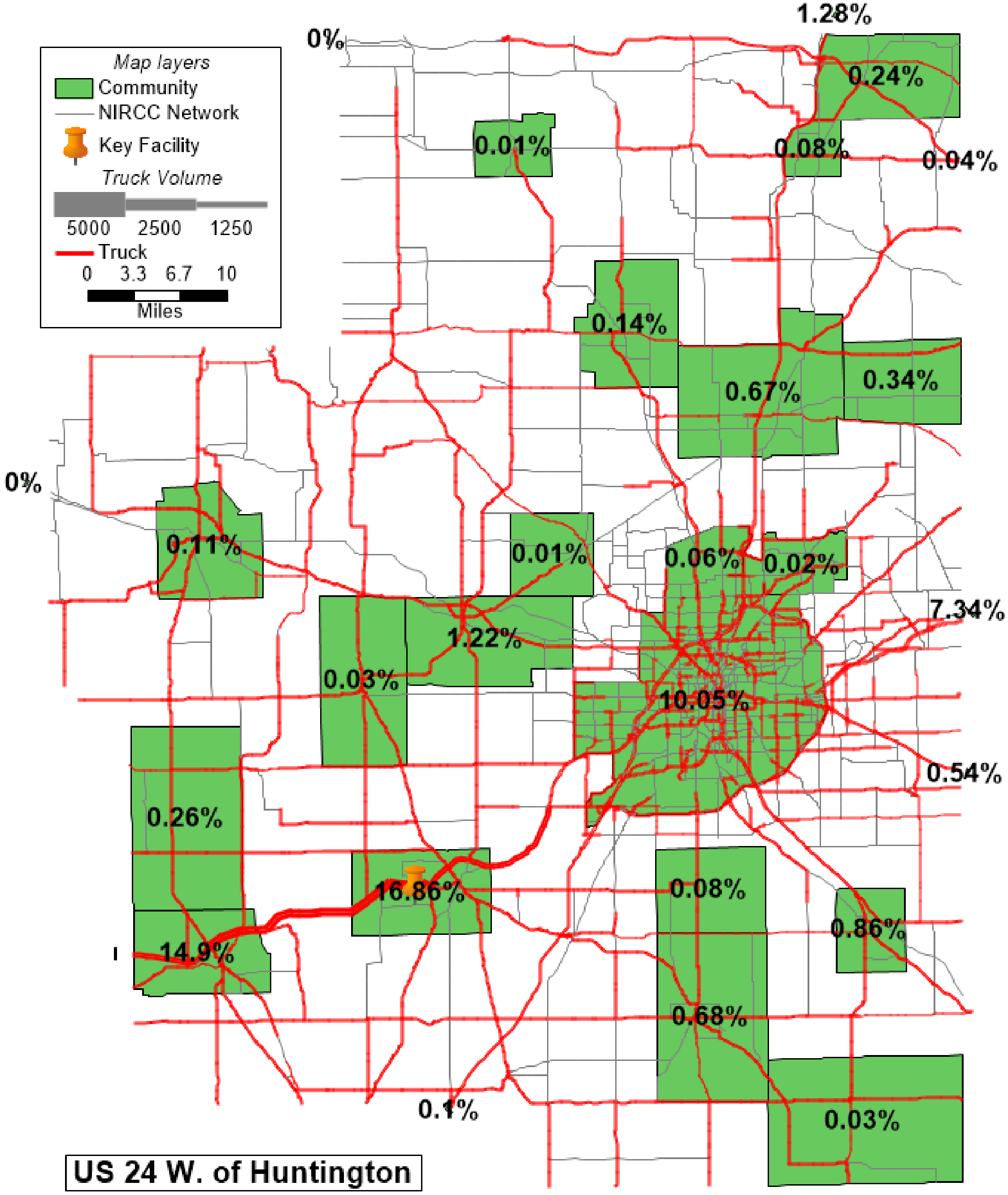


FIGURE 77. DEMAND AT US 24 W. OF HUNTINGTON



US 24 W. of Huntington

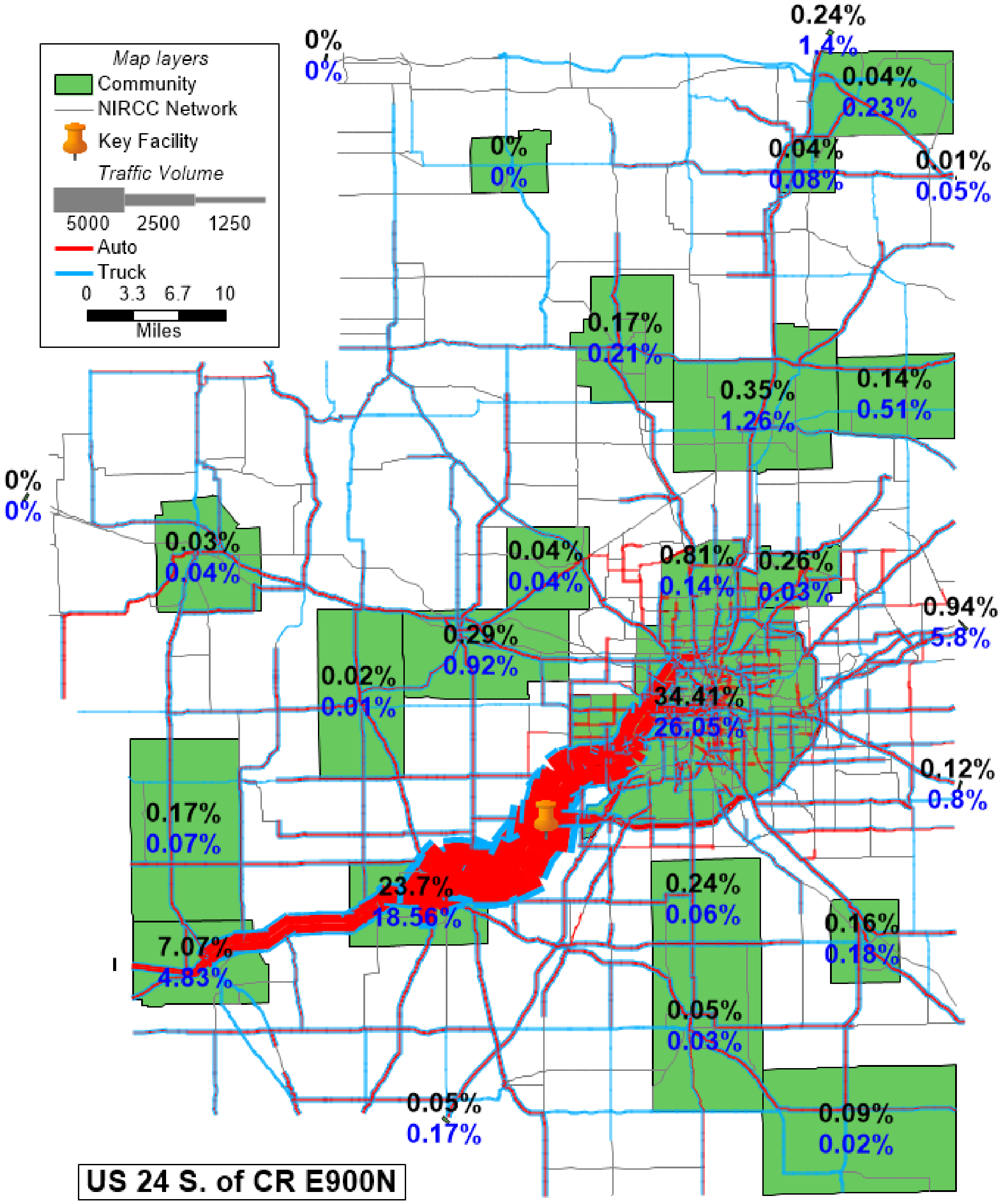
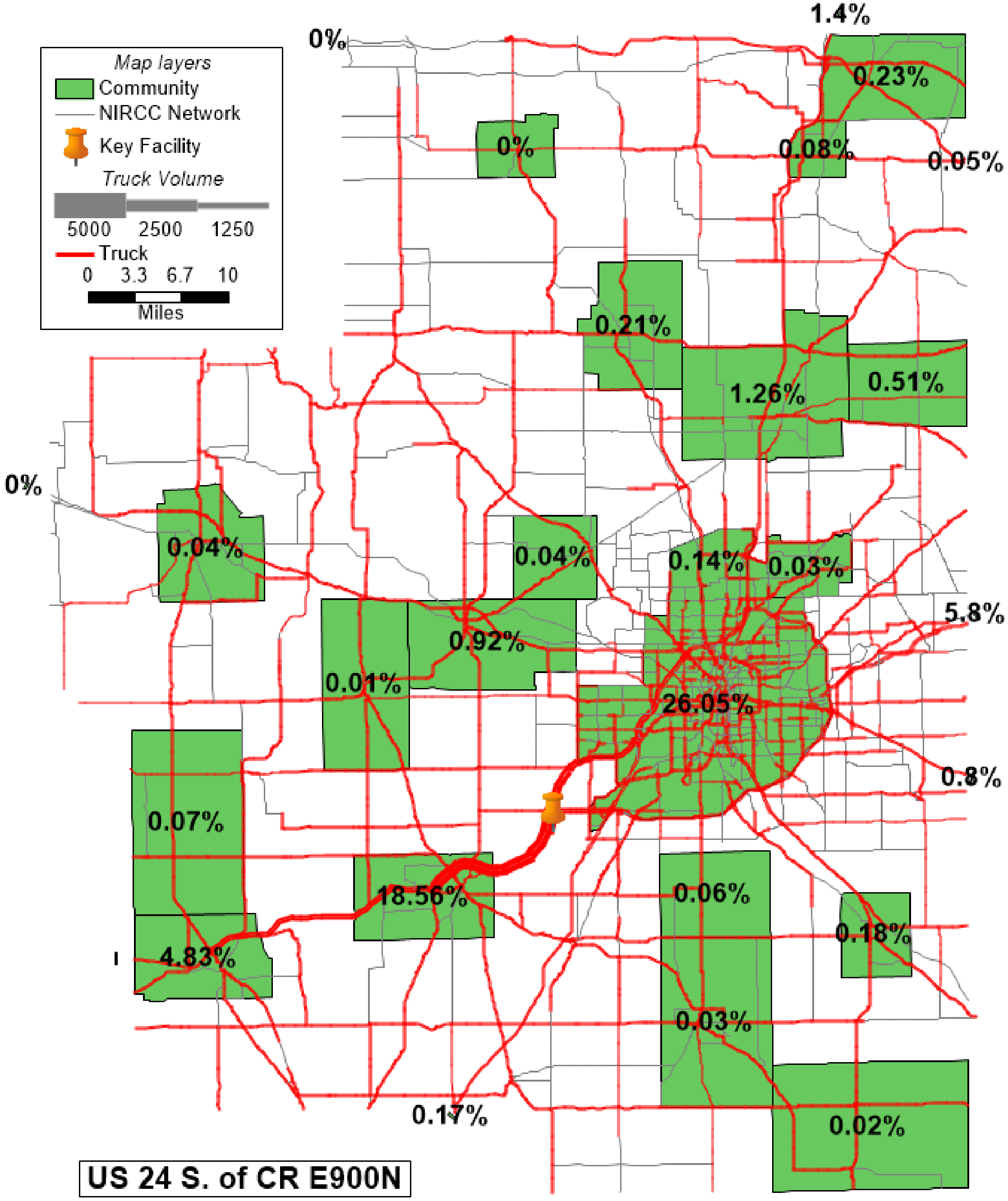


FIGURE 78. DEMAND AT US 24 S. OF CR E900N



Map layers
Community
NIRCC Network
Key Facility
Truck Volume
5000 2500 1250
Truck
0 3.3 6.7 10
Miles
0%
1.4%
0.23%
0%
0.08%
0.05%
0.21%
1.26%
0.51%
0%
0.04%
0.04%
0.14%
0.03%
5.8%
0.92%
0.01%
26.05%
0.8%
0.07%
18.56%
0.06%
0.18%
4.83%
0.03%
0.17%
0.02%
US 24 S. of CR E900N

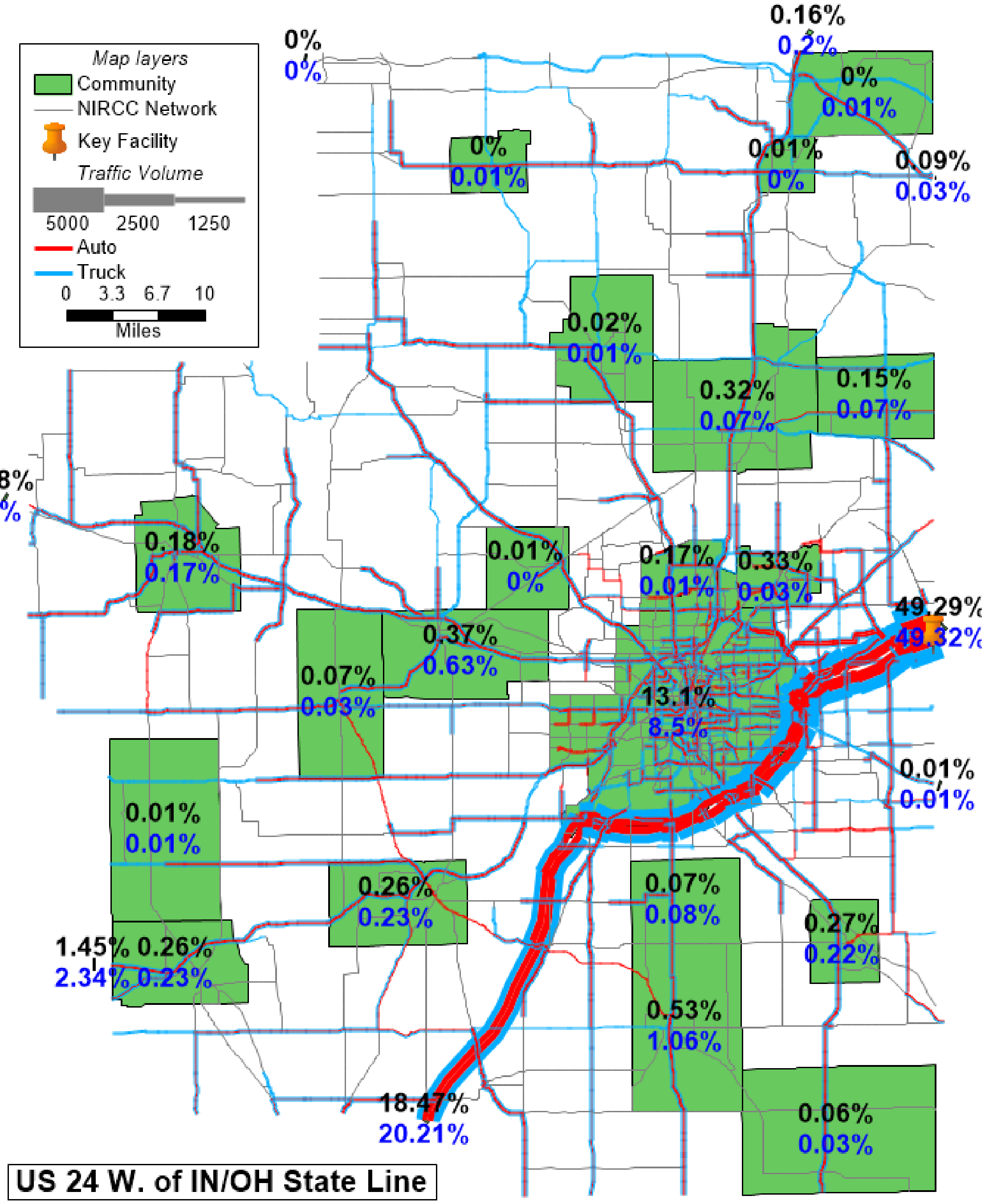
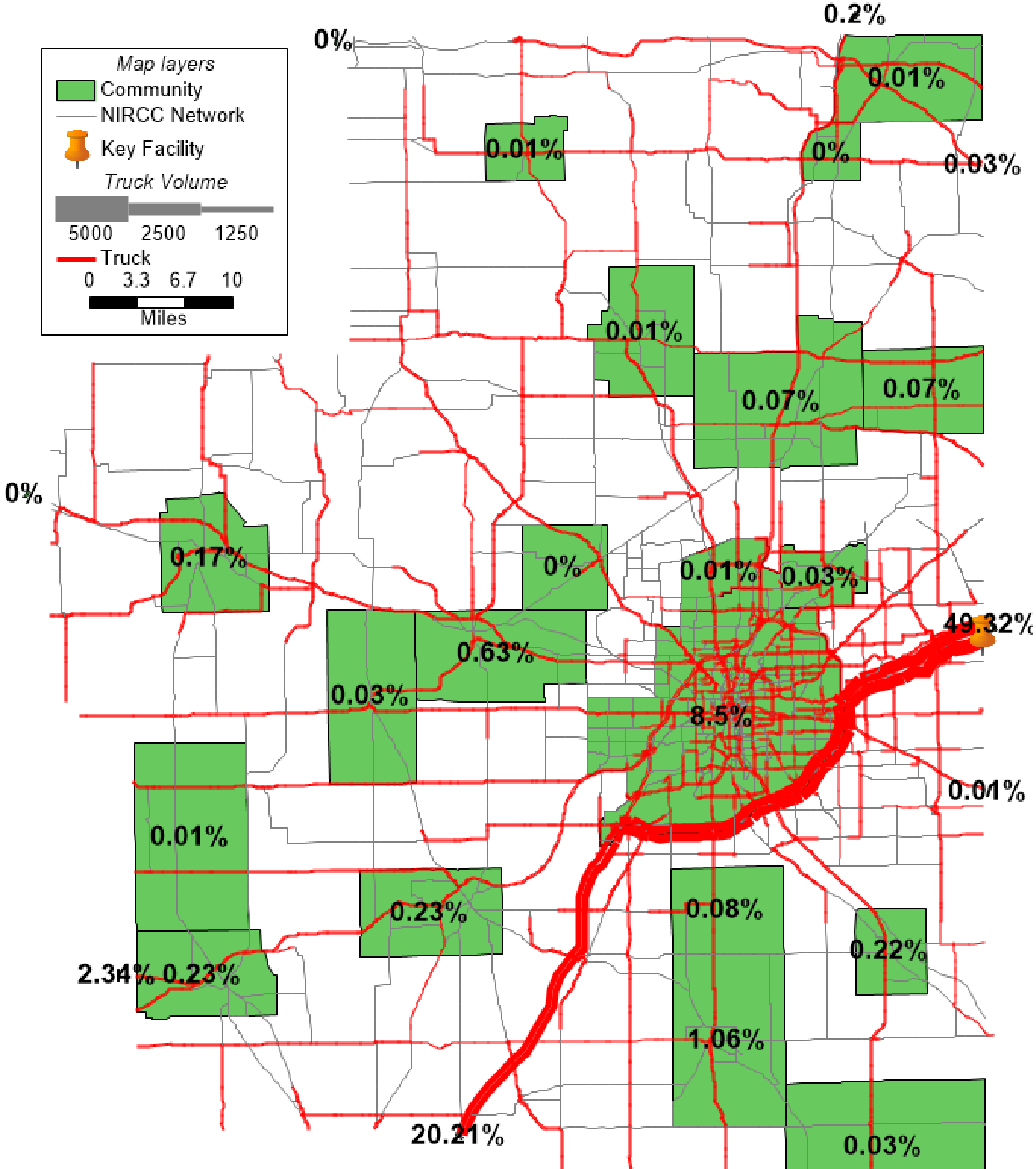


FIGURE 79. DEMAND AT US 24 W. OF IN/OH STATE LINE



US 24 W. of IN/OH State Line

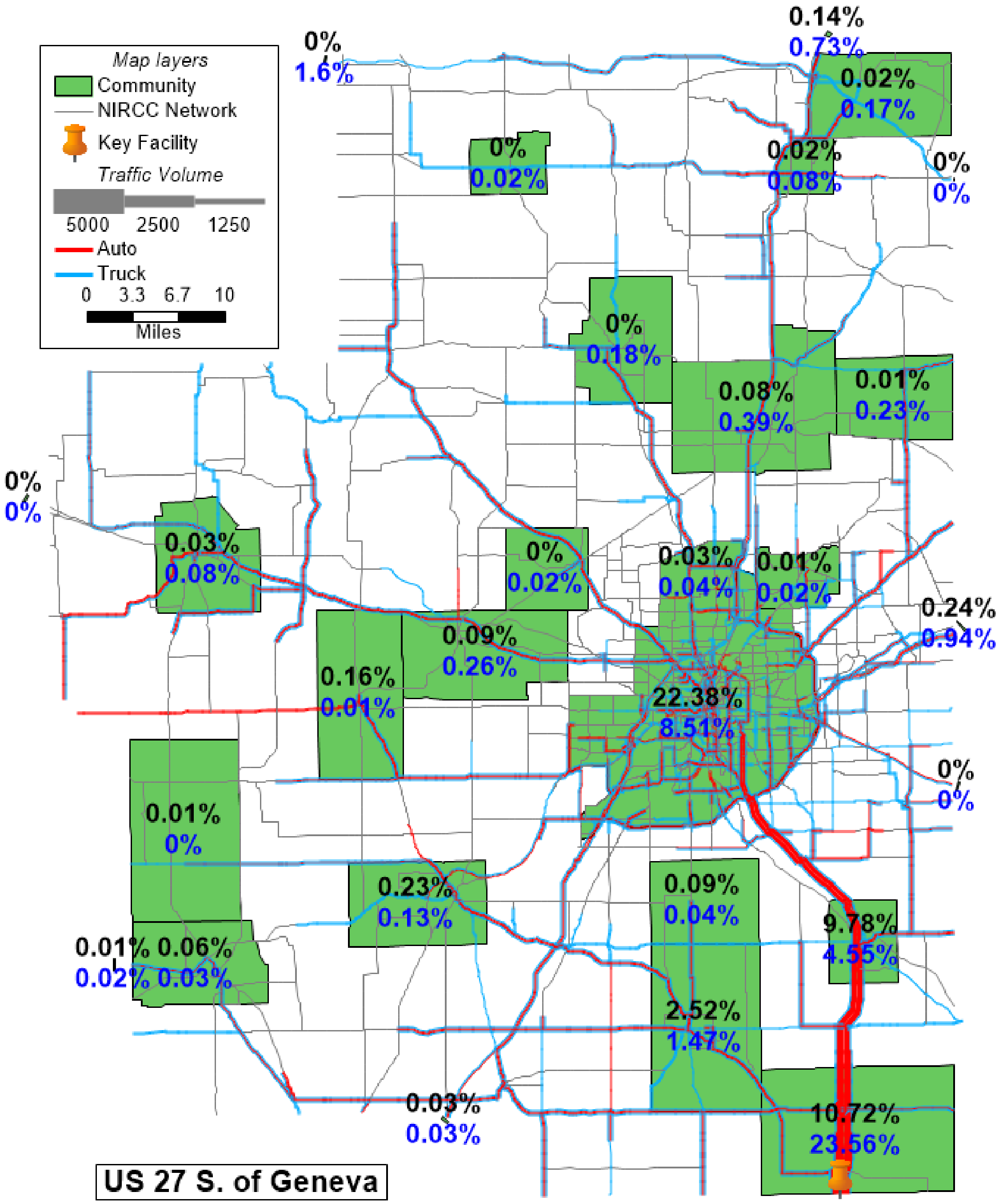
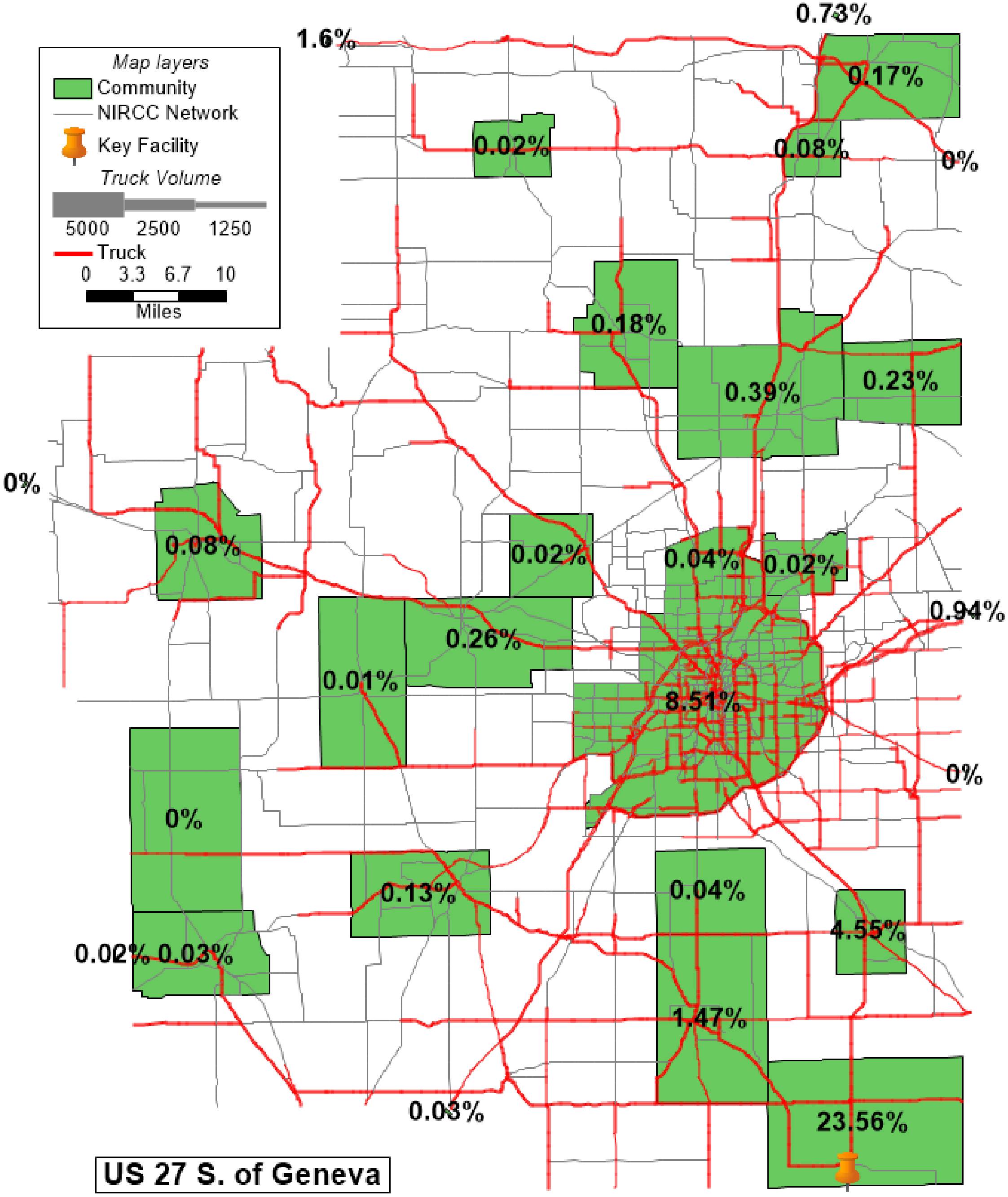


FIGURE 80. DEMAND AT US 27 S. OF GENEVA



US 27 S. of Geneva

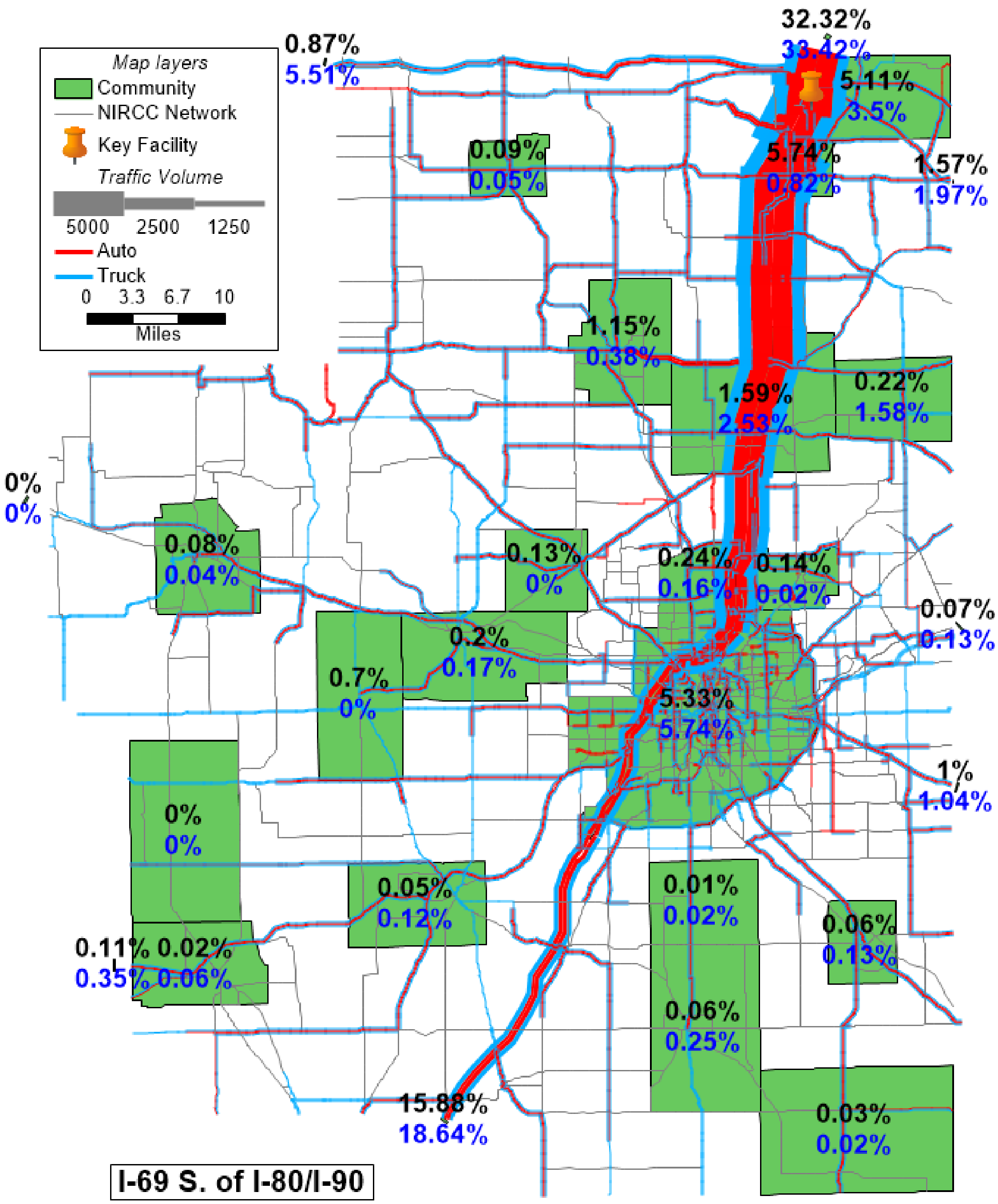
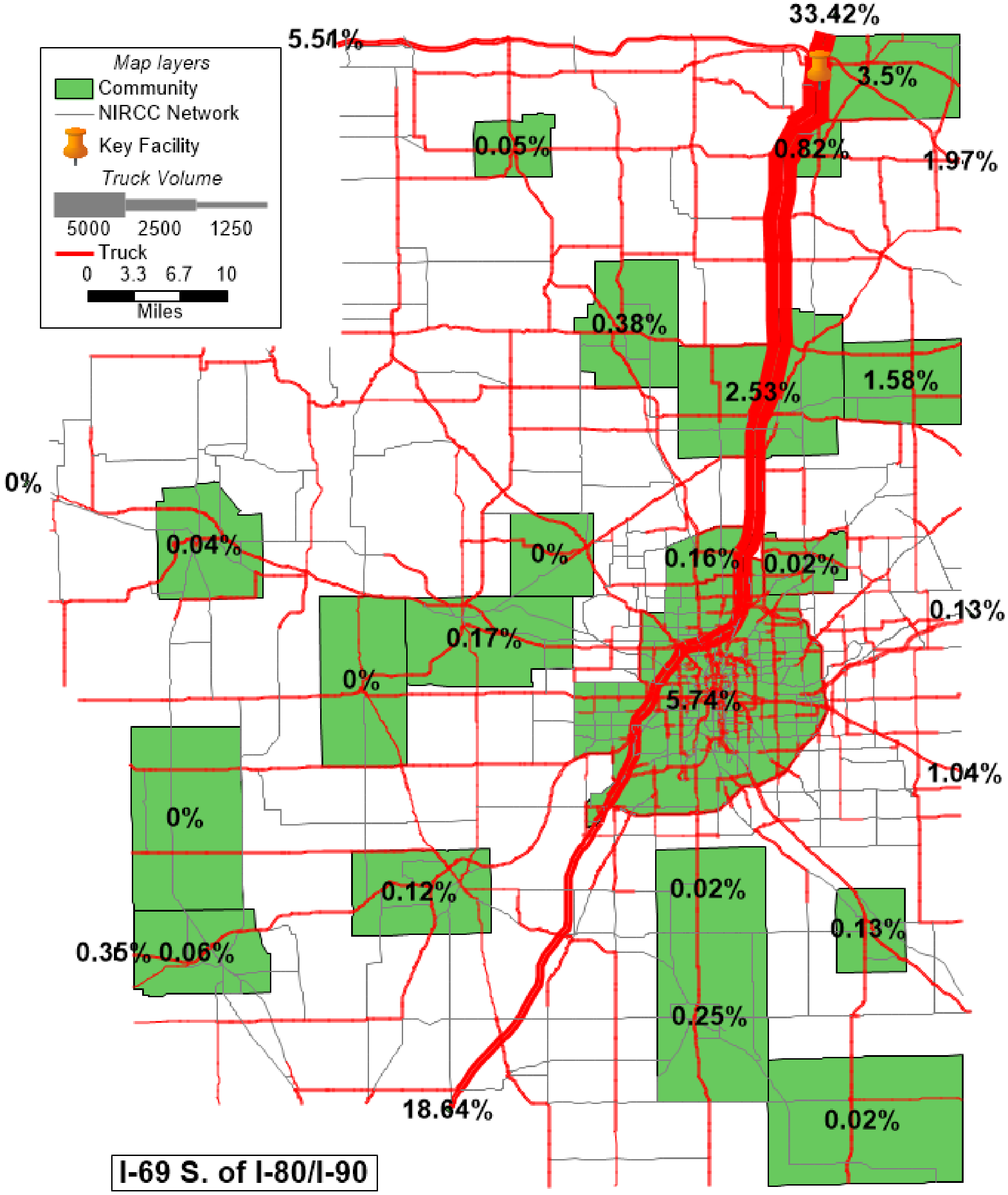


FIGURE 81. DEMAND AT I-69 S. OF I-80/90



I-69 S. of I-80/I-90

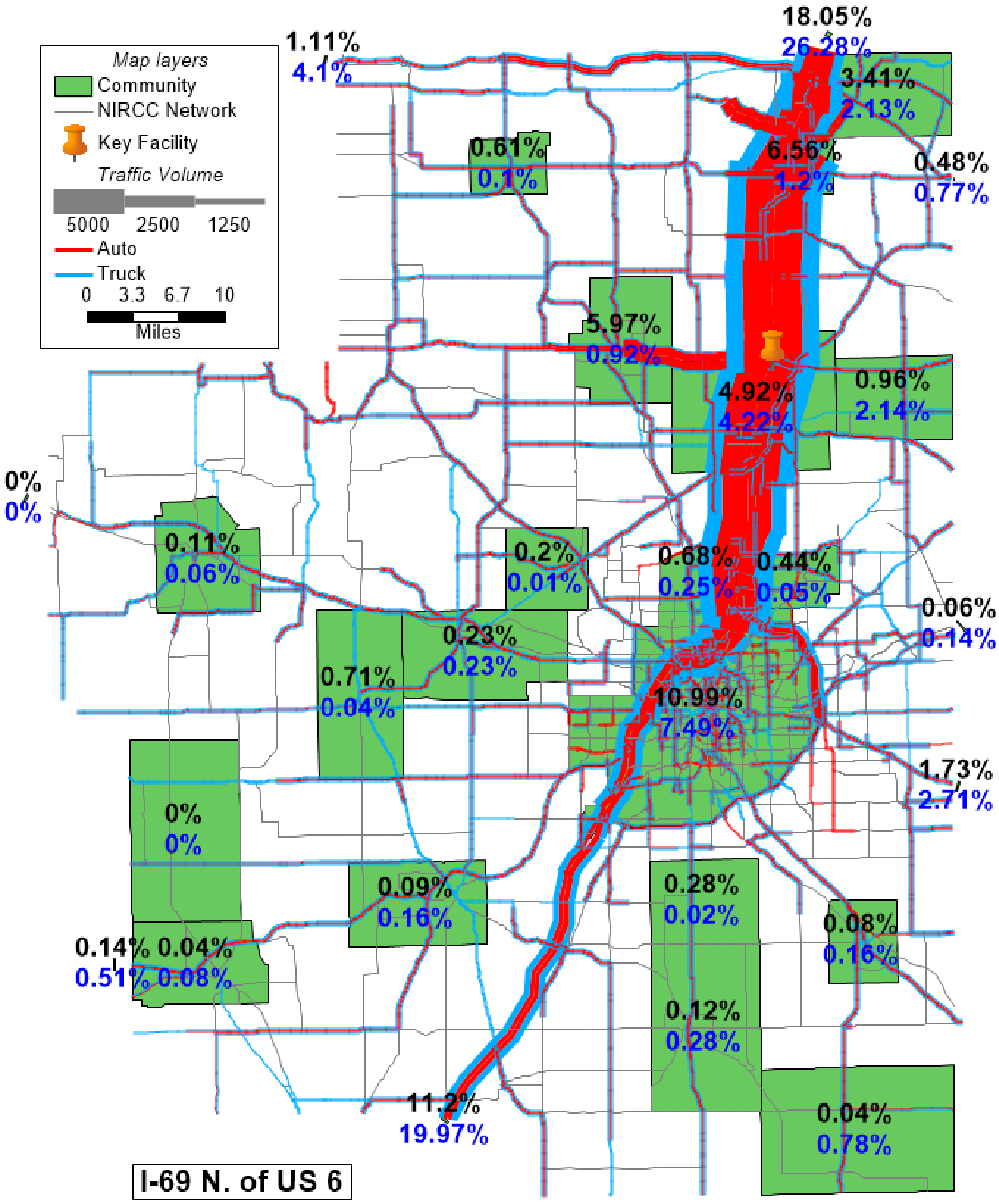
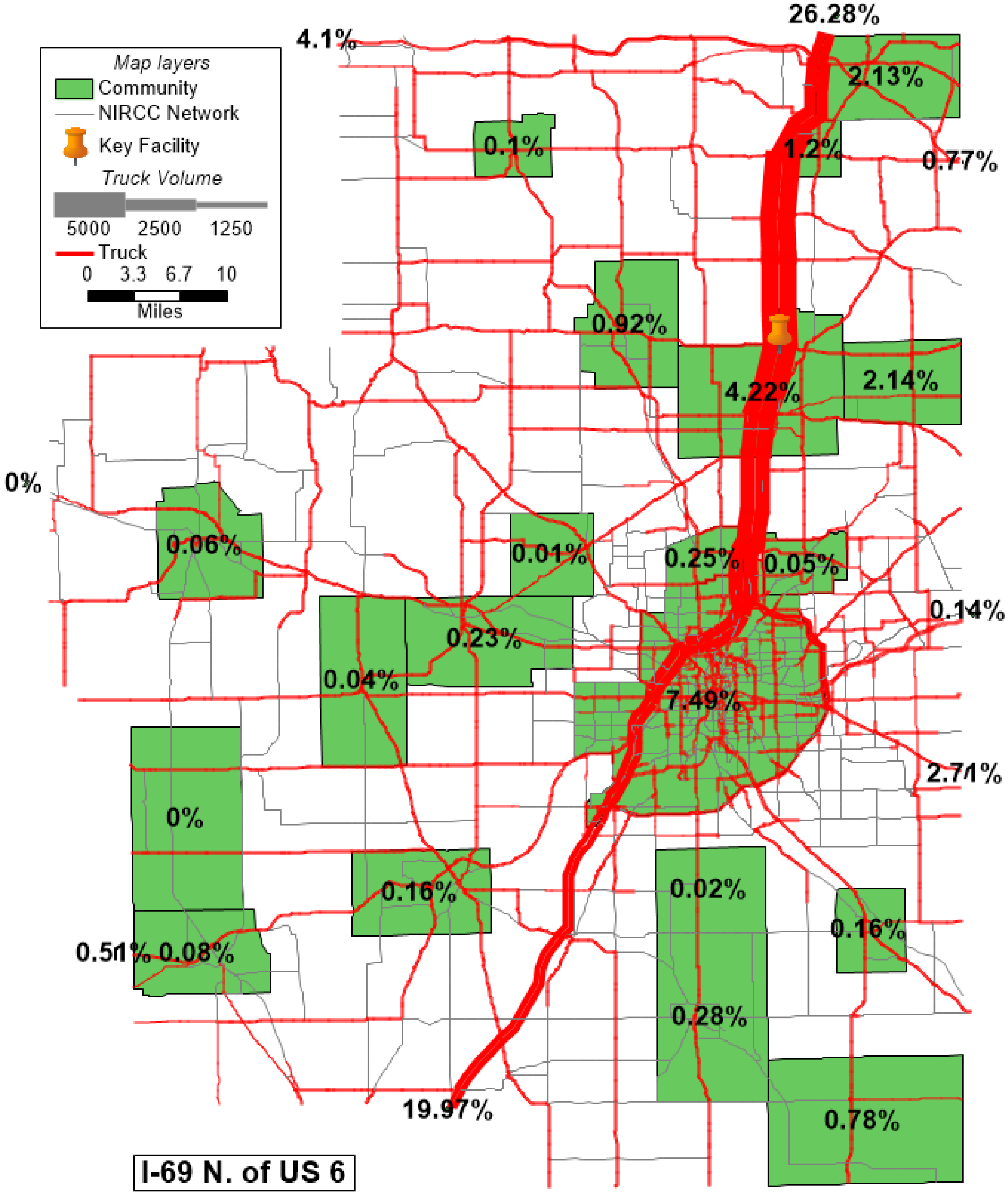


FIGURE 82. DEMAND AT I-69 N. OF US 6


Map layers
Community
NIRCC Network
Key Facility
Truck Volume
5000 2500 1250
Truck
0 3.3 6.7 10
Miles
26.28%
4.1%
2.13%
0.1%
1.2%
0.77%
0.92%
4.22%
2.14%
0%
0.06%
0.01%
0.25%
0.05%
0.14%
0.23%
0.04%
7.49%
2.71%
0%
0.16%
0.02%
0.16%
0.51%
0.08%
0.28%
19.97%
0.78%
I-69 N. of US 6

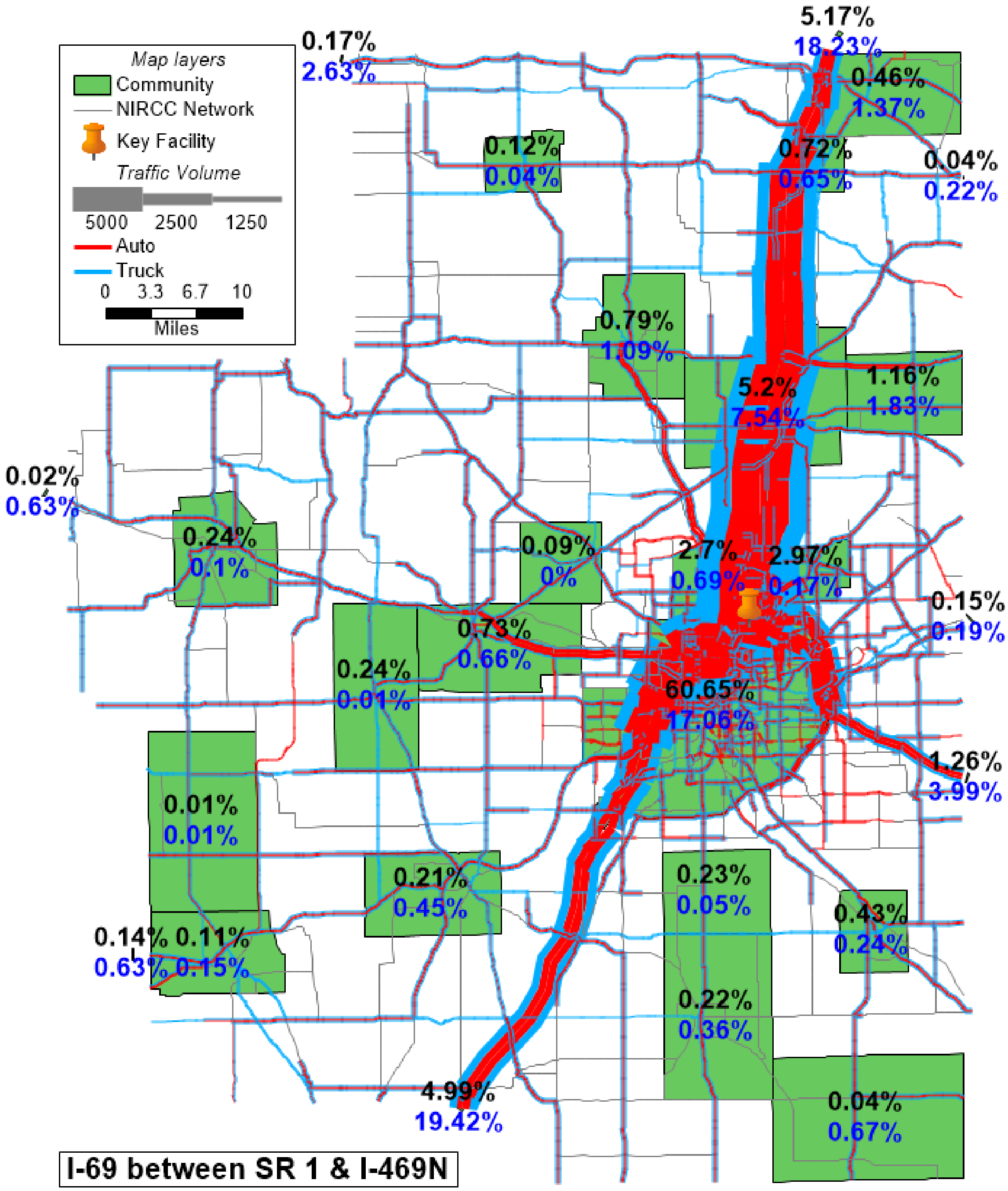
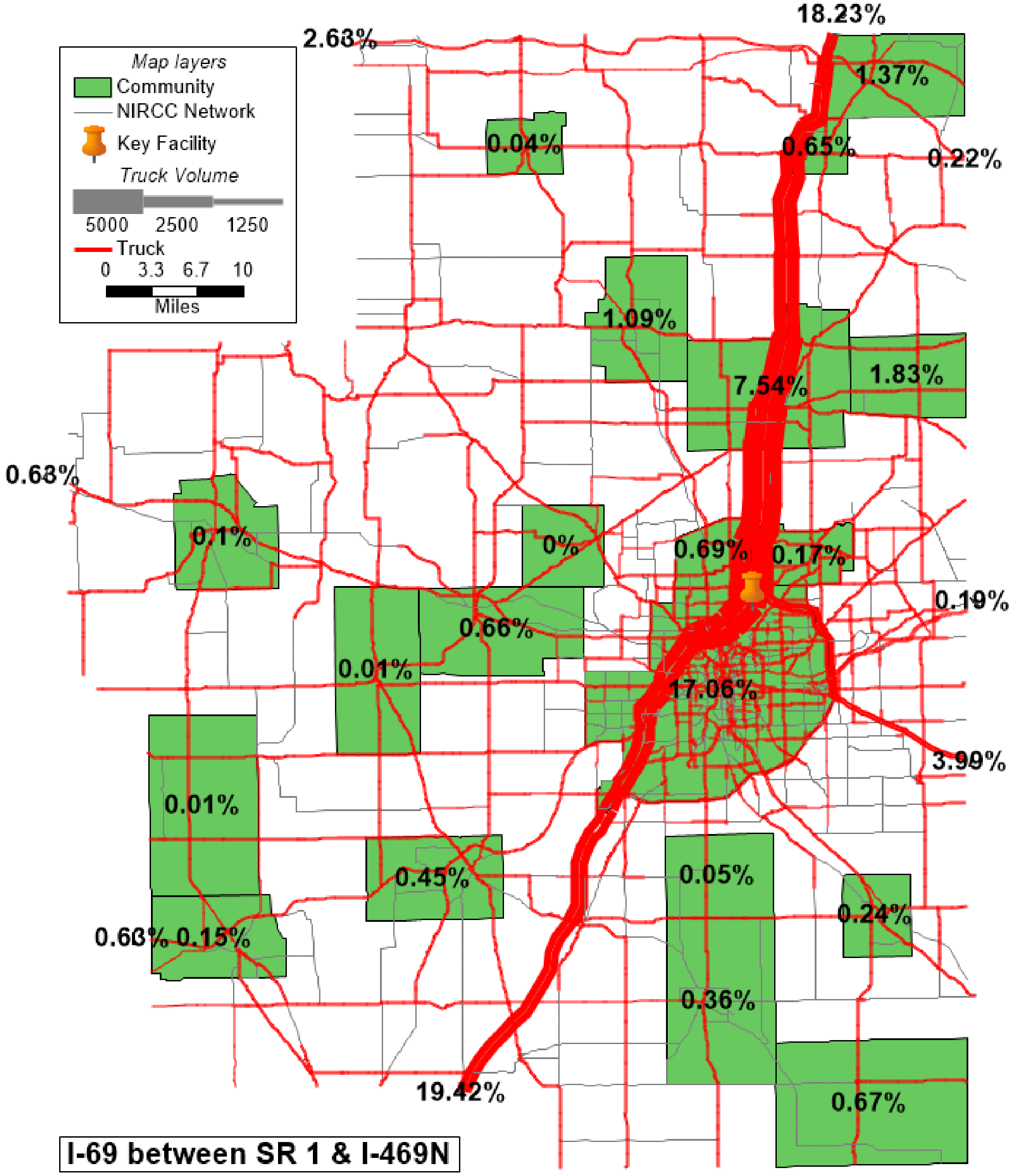


FIGURE 83. DEMAND AT I-69 BETWEEN SR 1 AND I-469 N



I-69 between SR 1 & I-469N

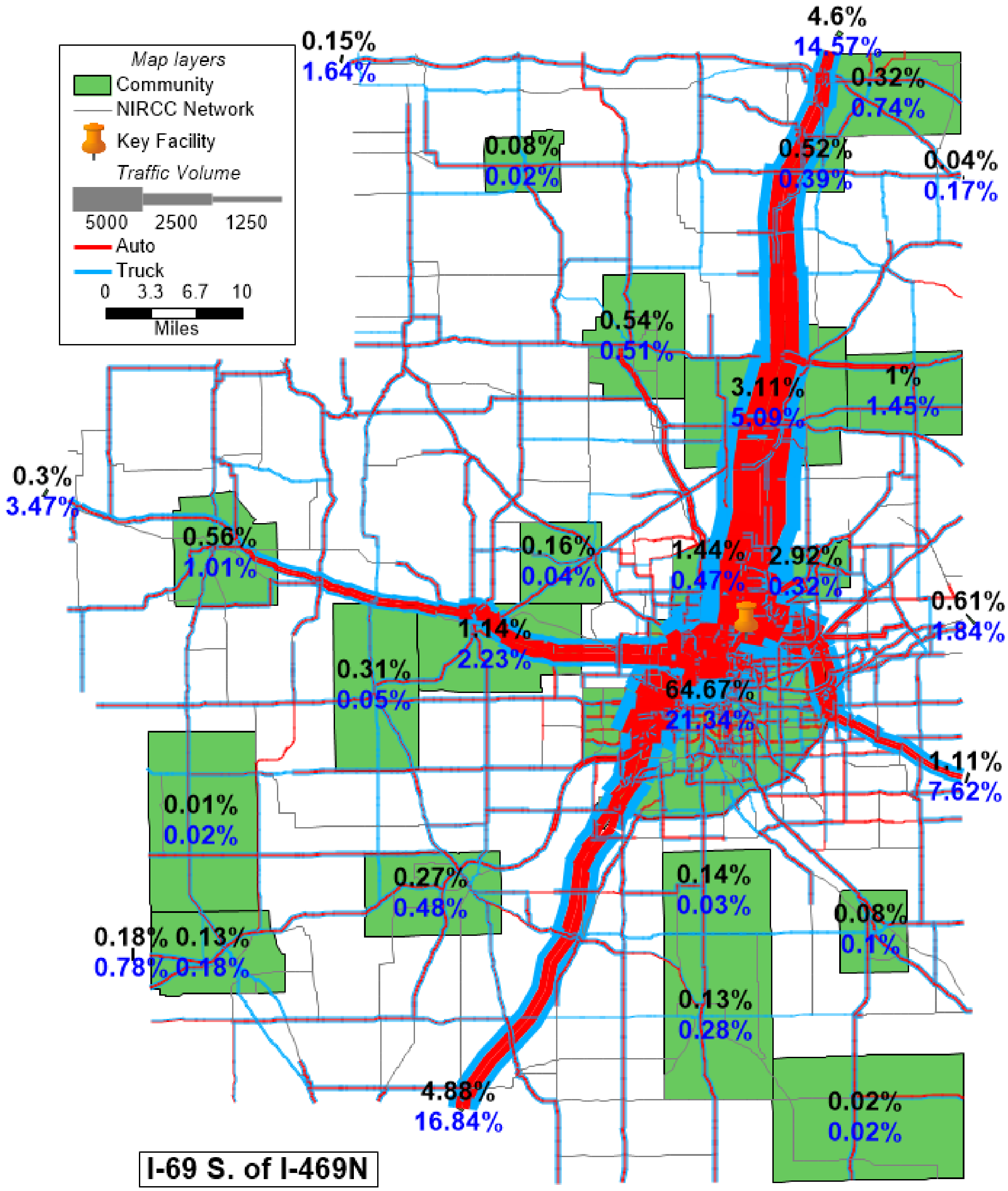
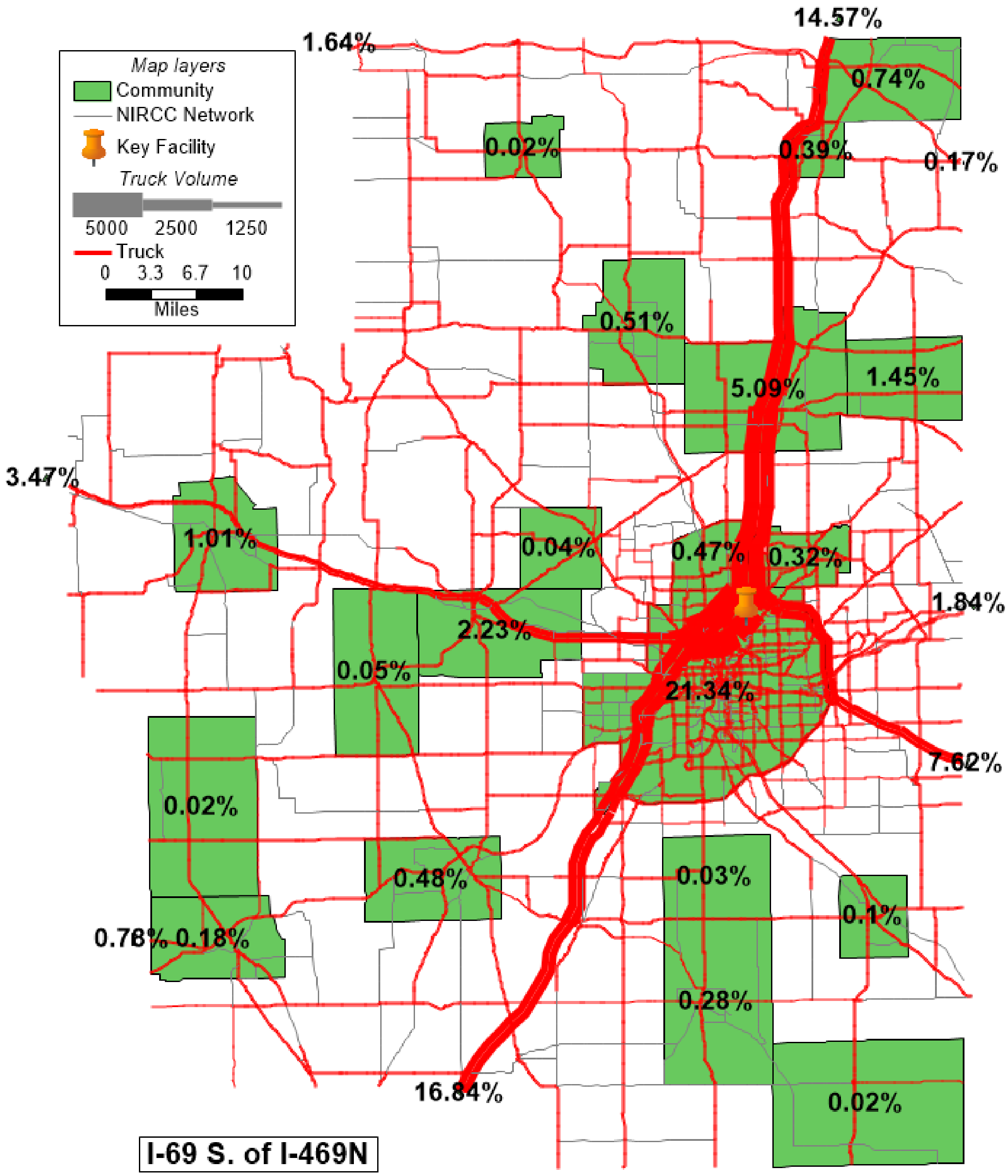


FIGURE 84. DEMAND AT I-69 S. OF I-469 N


Map layers
Community
NIRCC Network
Key Facility
Truck Volume
5000
2500
1250
Truck
0
3.3
6.7
10
Miles
14.57%
1.64%
0.74%
0.02%
0.39%
0.17%
0.51%
5.09%
1.45%
3.47%
1.01%
0.04%
0.47%
0.32%
1.84%
2.23%
0.05%
21.34%
7.62%
0.02%
0.48%
0.03%
0.1%
0.78%
0.18%
0.28%
16.84%
0.02%
I-69 S. of I-469N

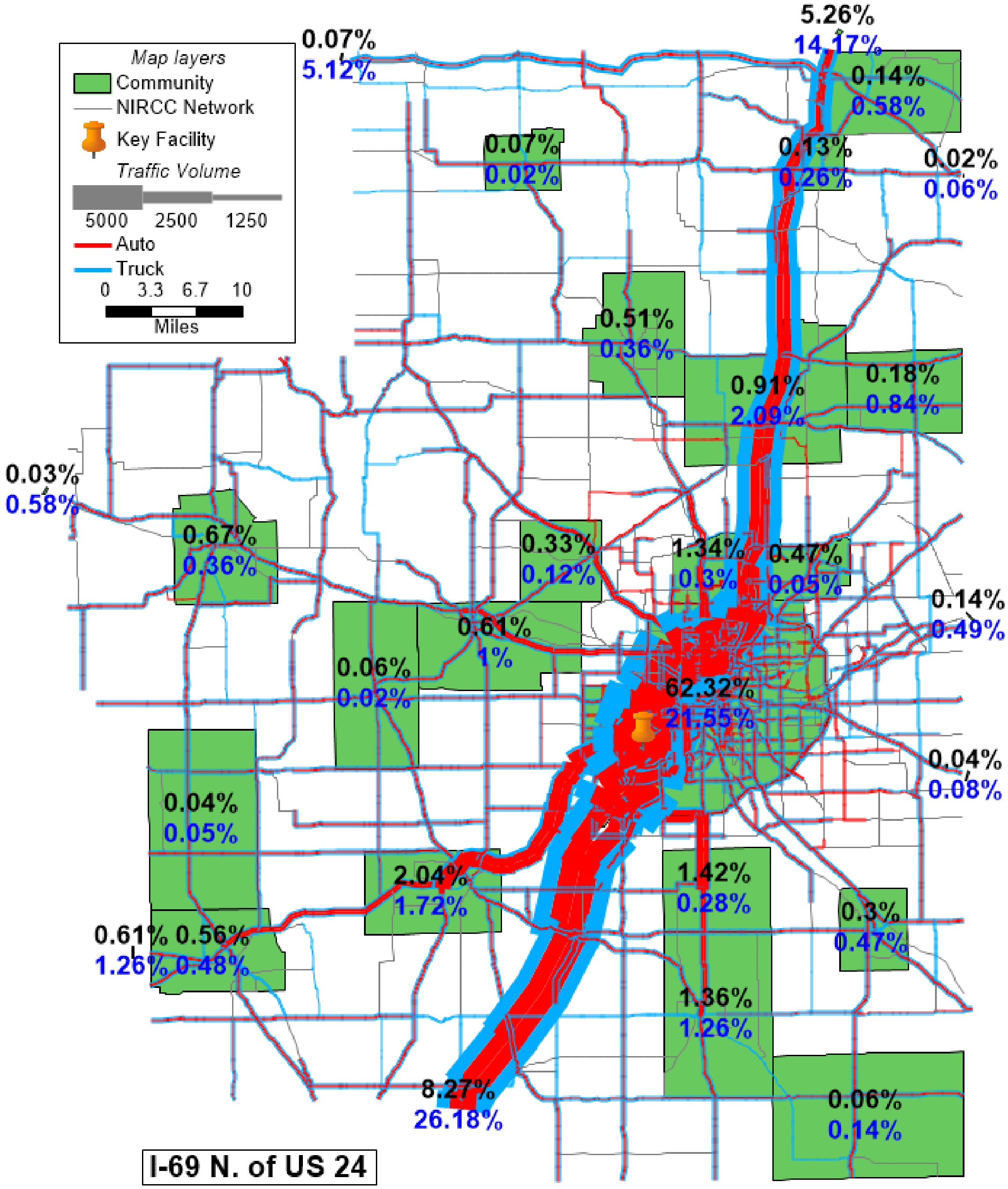
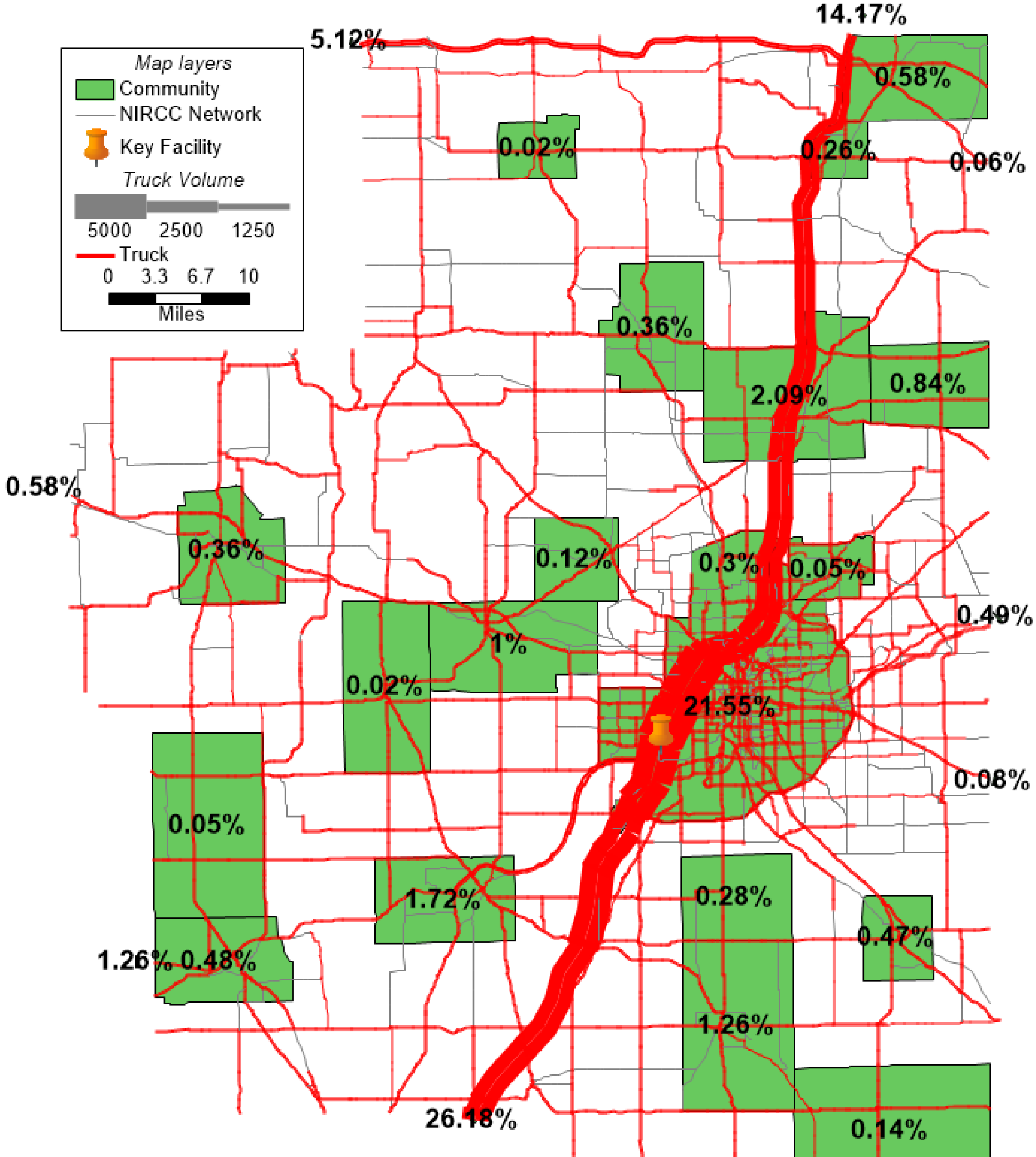


FIGURE 85. DEMAND AT I-69 N. OF US 24



Map layers
Community
NIRCC Network
Key Facility
Truck Volume
5000
2500
1250
Truck
0
3.3
6.7
10
Miles
14.17%
5.12%
0.58%
0.02%
0.26%
0.06%
0.36%
2.09%
0.84%
0.58%
0.36%
0.12%
0.3%
0.05%
0.49%
1%
0.02%
21.55%
0.08%
0.05%
1.72%
0.28%
0.47%
1.26%
0.48%
1.26%
26.18%
0.14%

I-69 N. of US 24

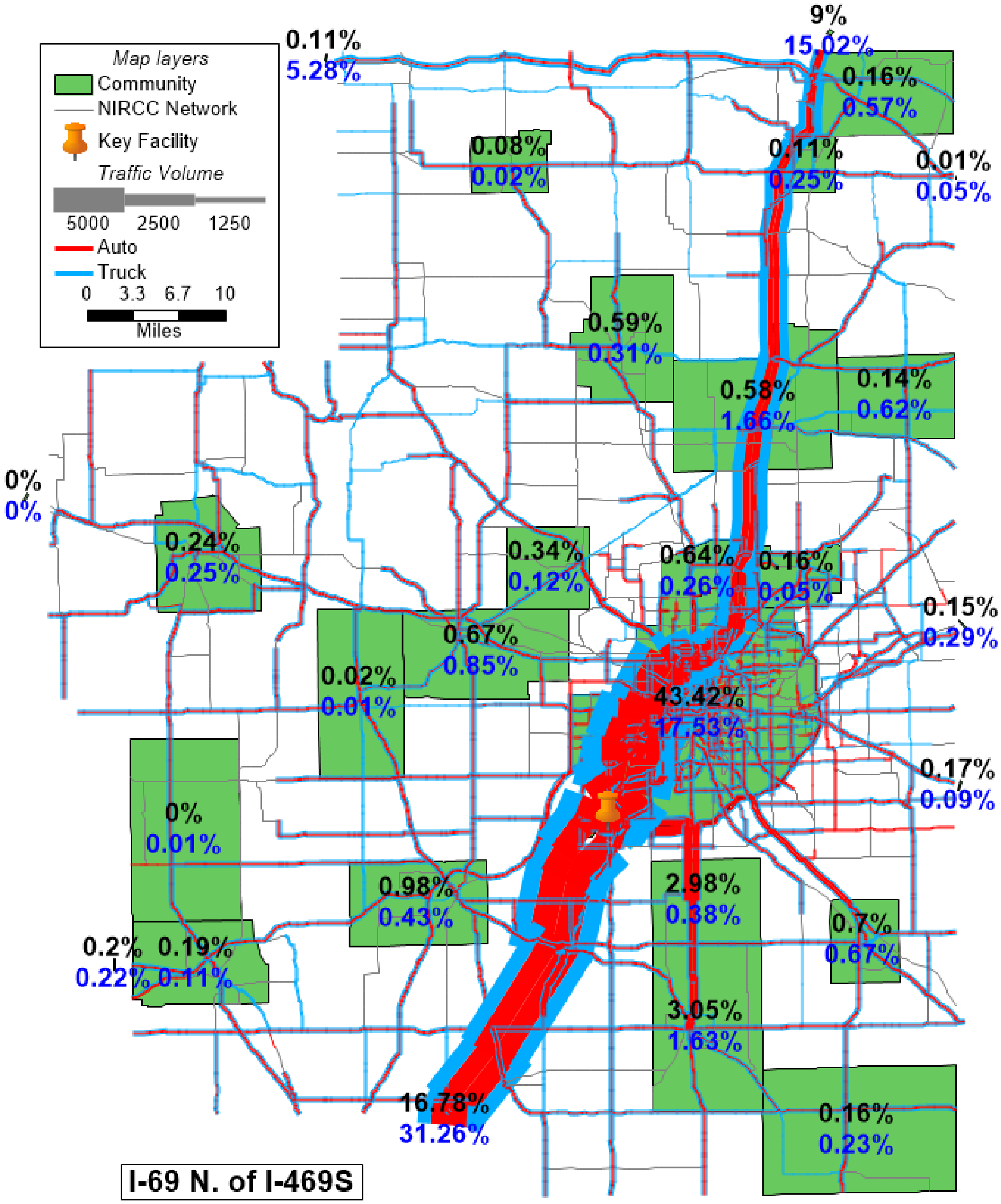
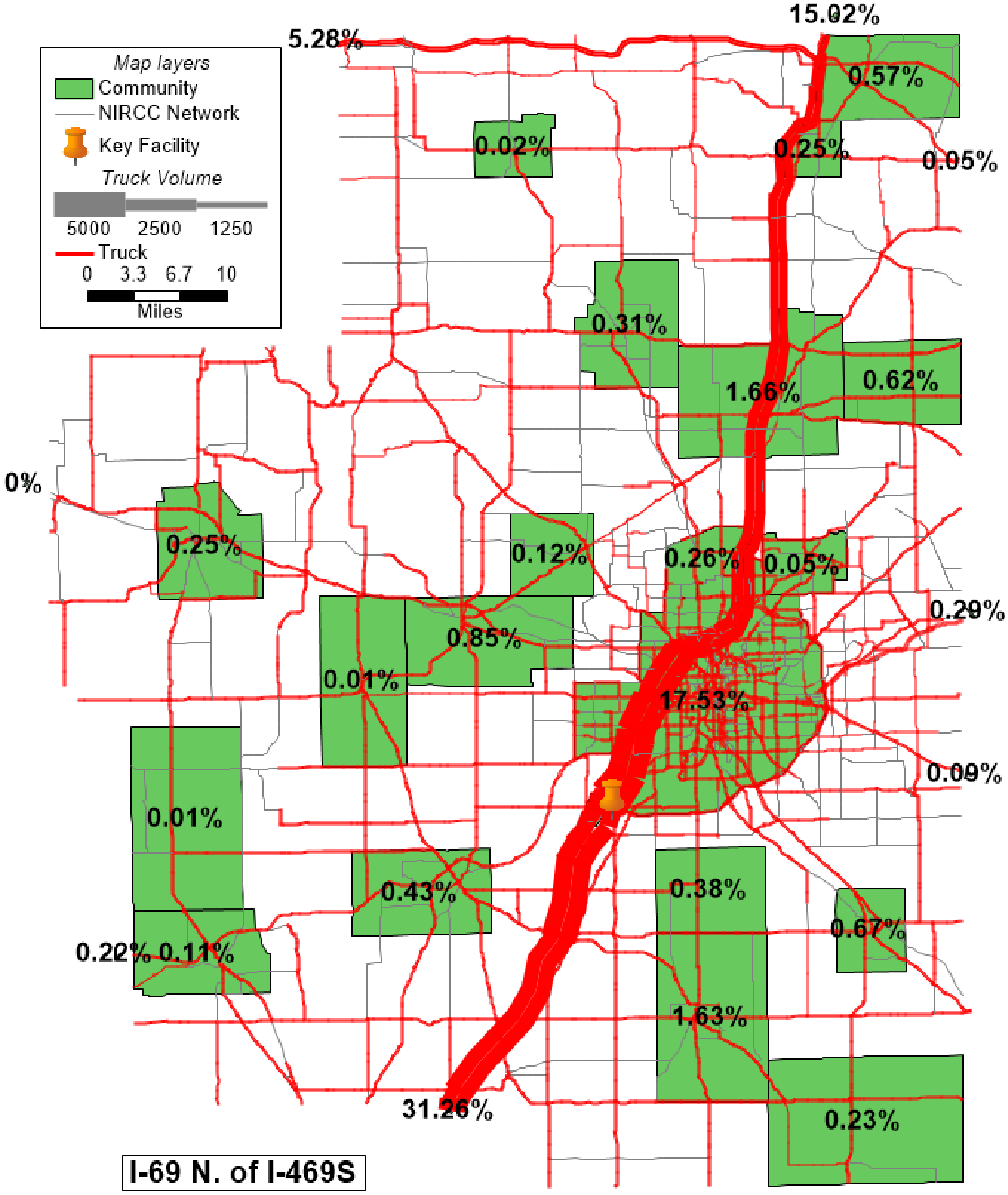


FIGURE 86. DEMAND AT I-69 N. OF I-469 S


Map layers
Community
NIRCC Network
Key Facility
Truck Volume
5000
2500
1250
Truck
0
3.3
6.7
10
Miles
15.02%
5.28%
0.57%
0.02%
0.25%
0.05%
0.31%
1.66%
0.62%
0%
0.25%
0.12%
0.26%
0.05%
0.29%
0.85%
0.01%
17.53%
0.09%
0.01%
0.43%
0.38%
0.67%
0.22%
0.11%
1.63%
31.26%
0.23%
I-69 N. of I-469S

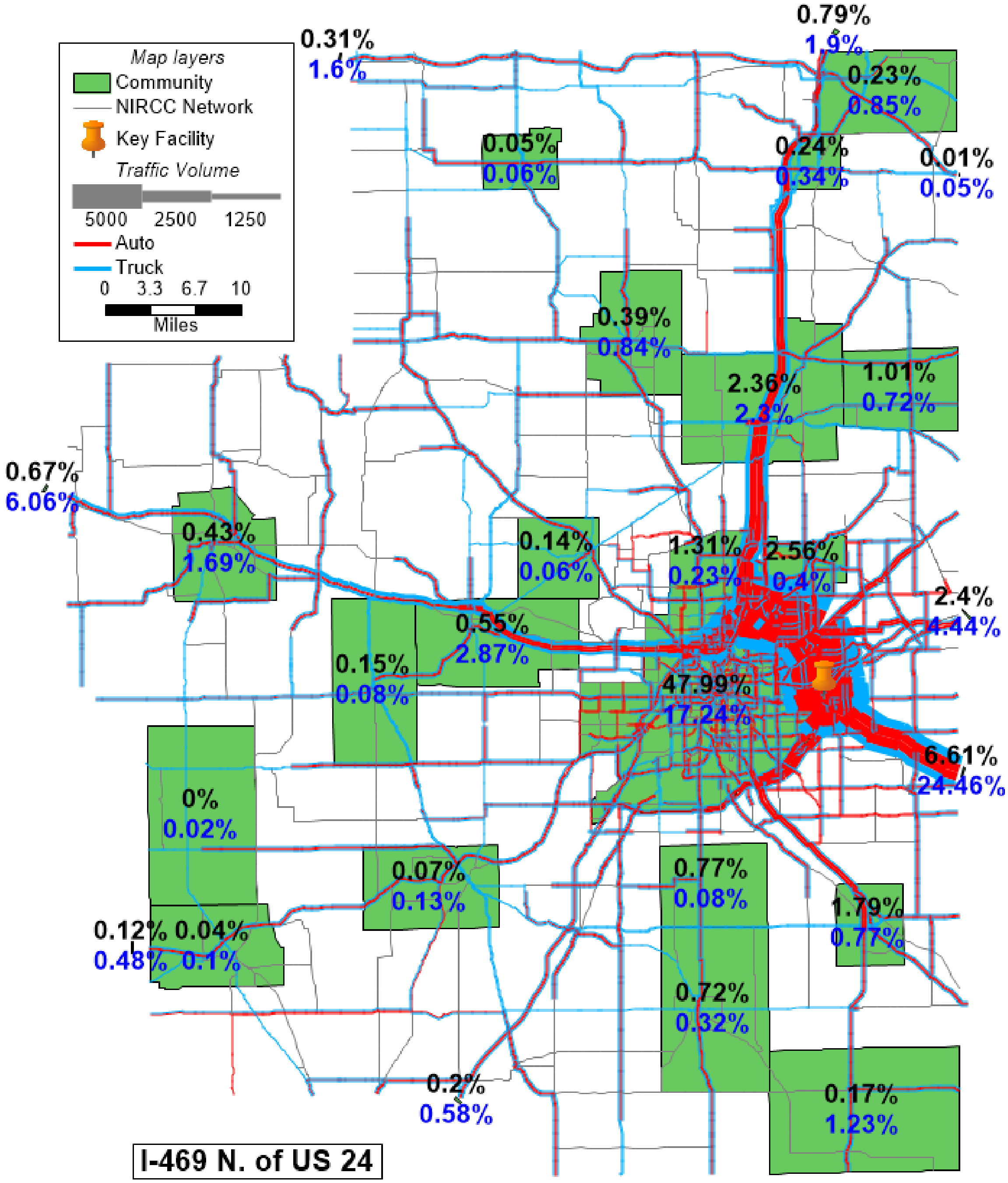
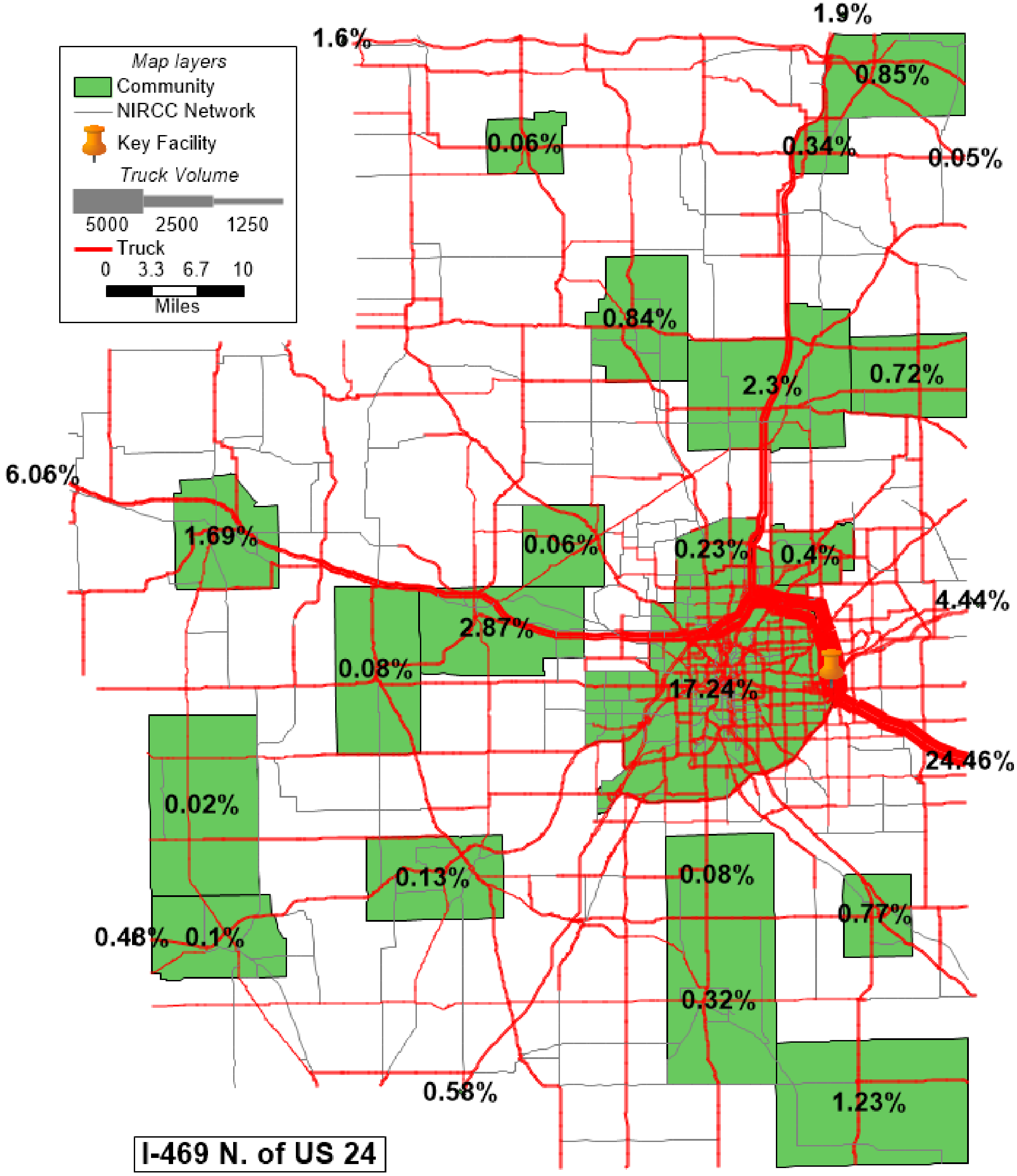


FIGURE 87. DEMAND AT I-469 N. OF US 24



I-469 N. of US 24

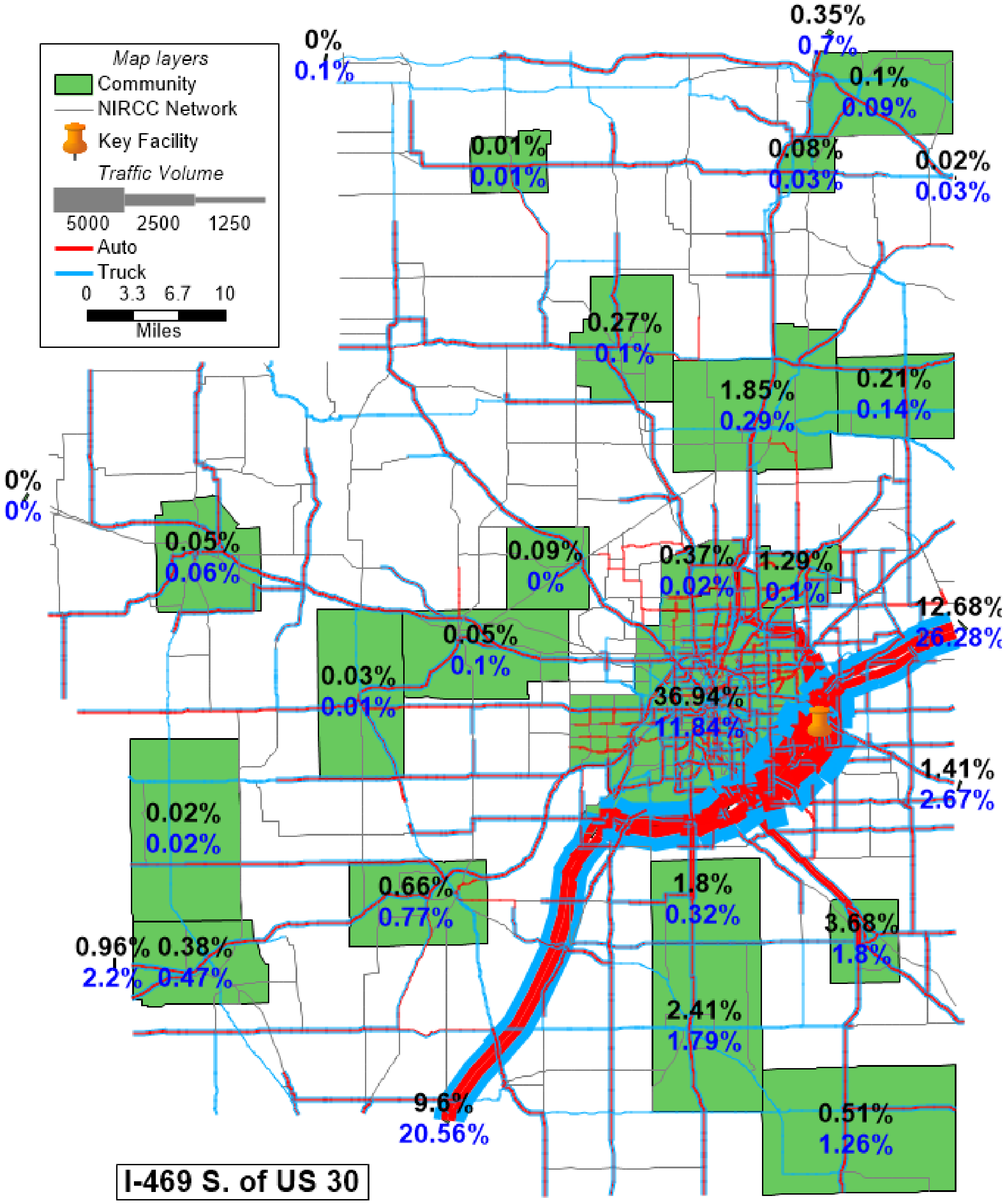
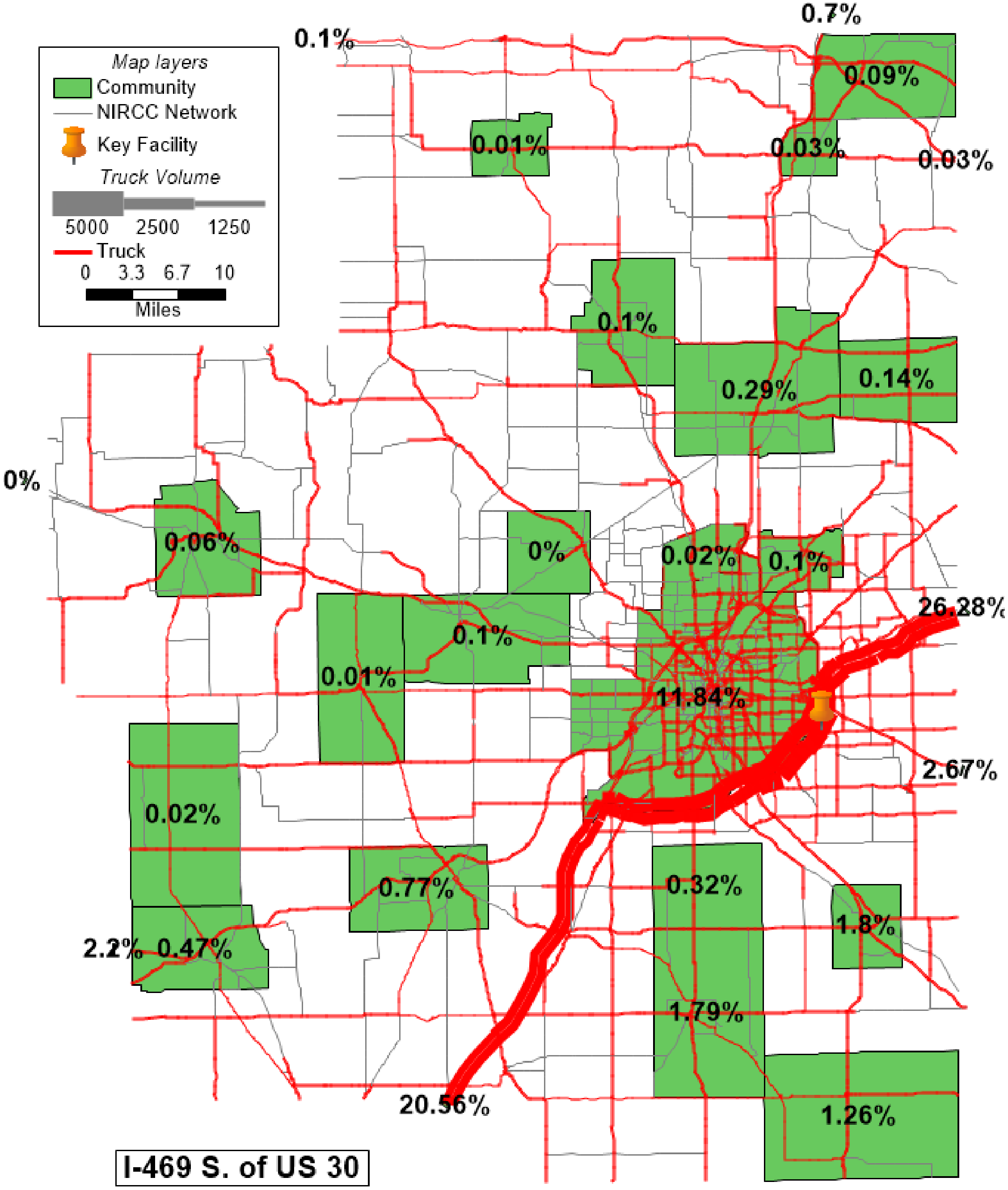


FIGURE 88. DEMAND AT I-469 S. OF US 30



I-469 S. of US 30

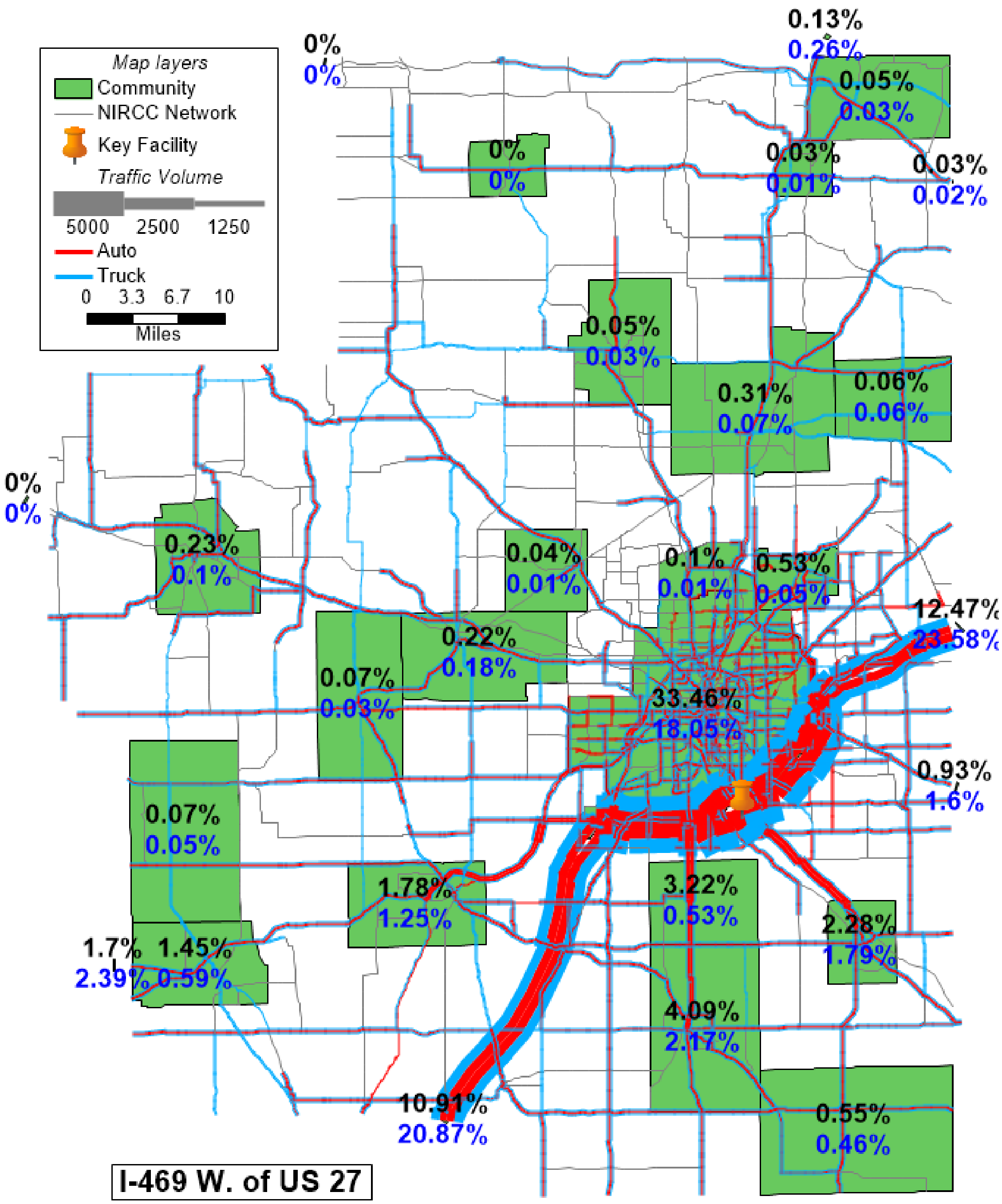
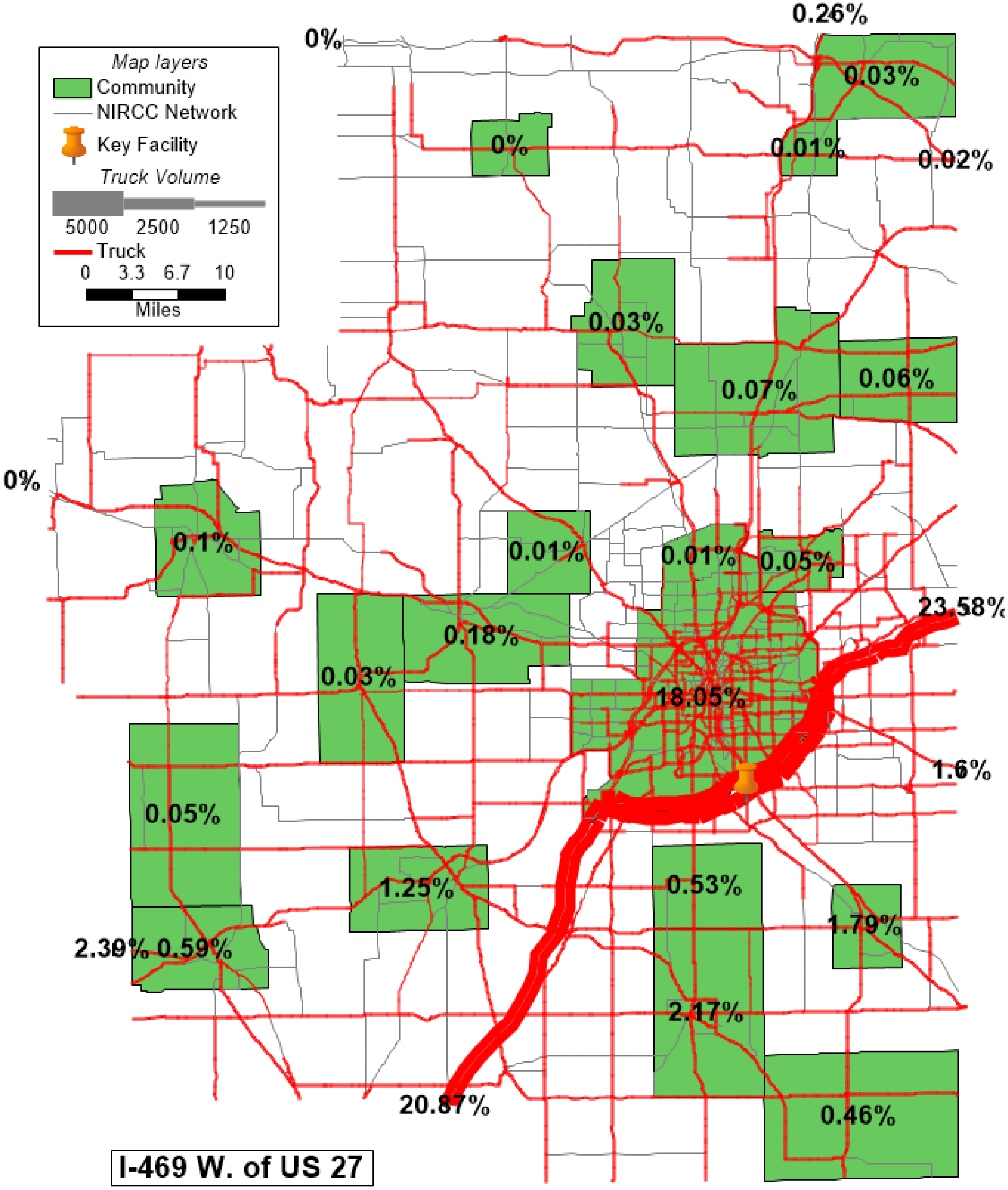


FIGURE 89. DEMAND AT I-469 W. OF US 27



I-469 W. of US 27

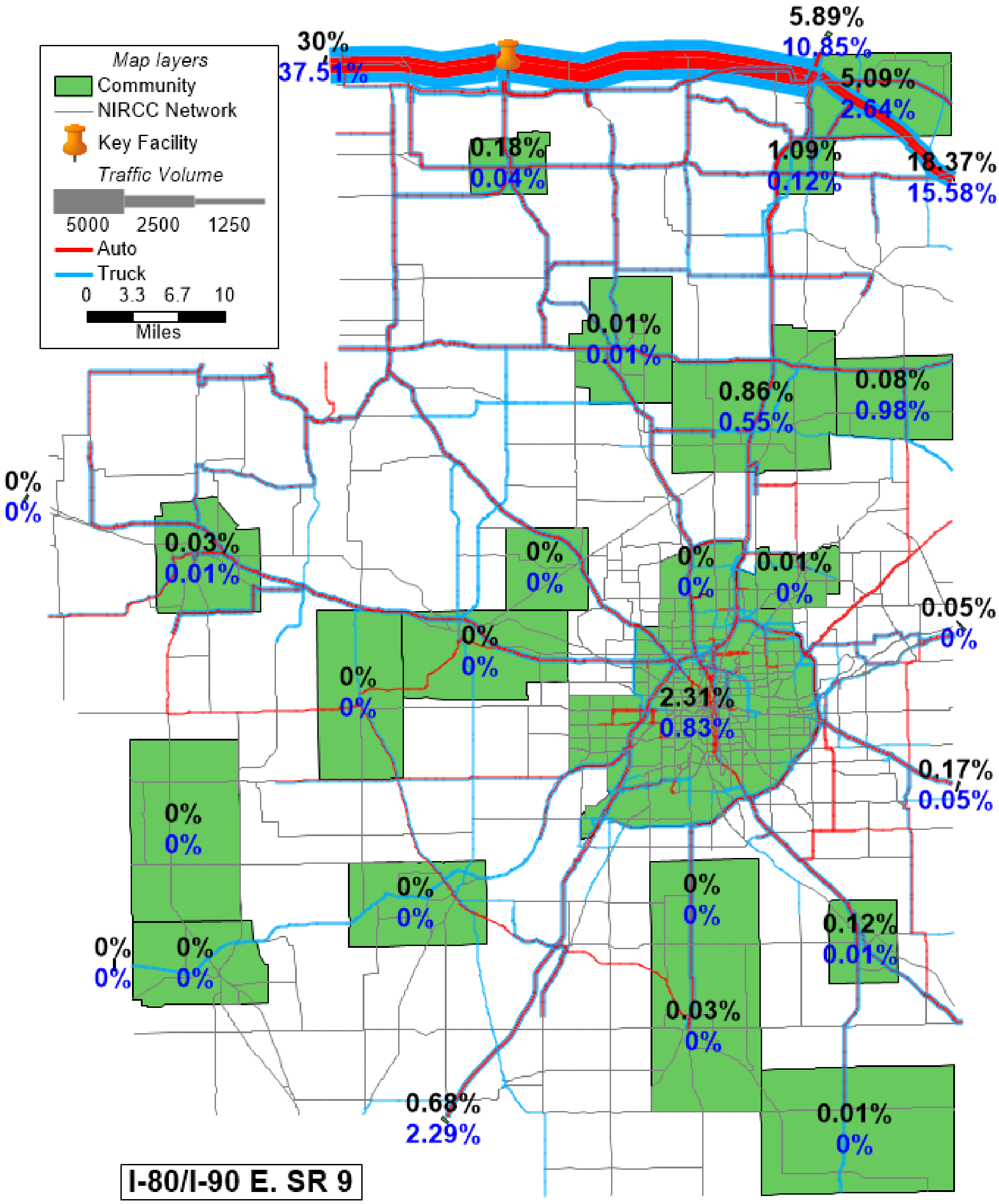
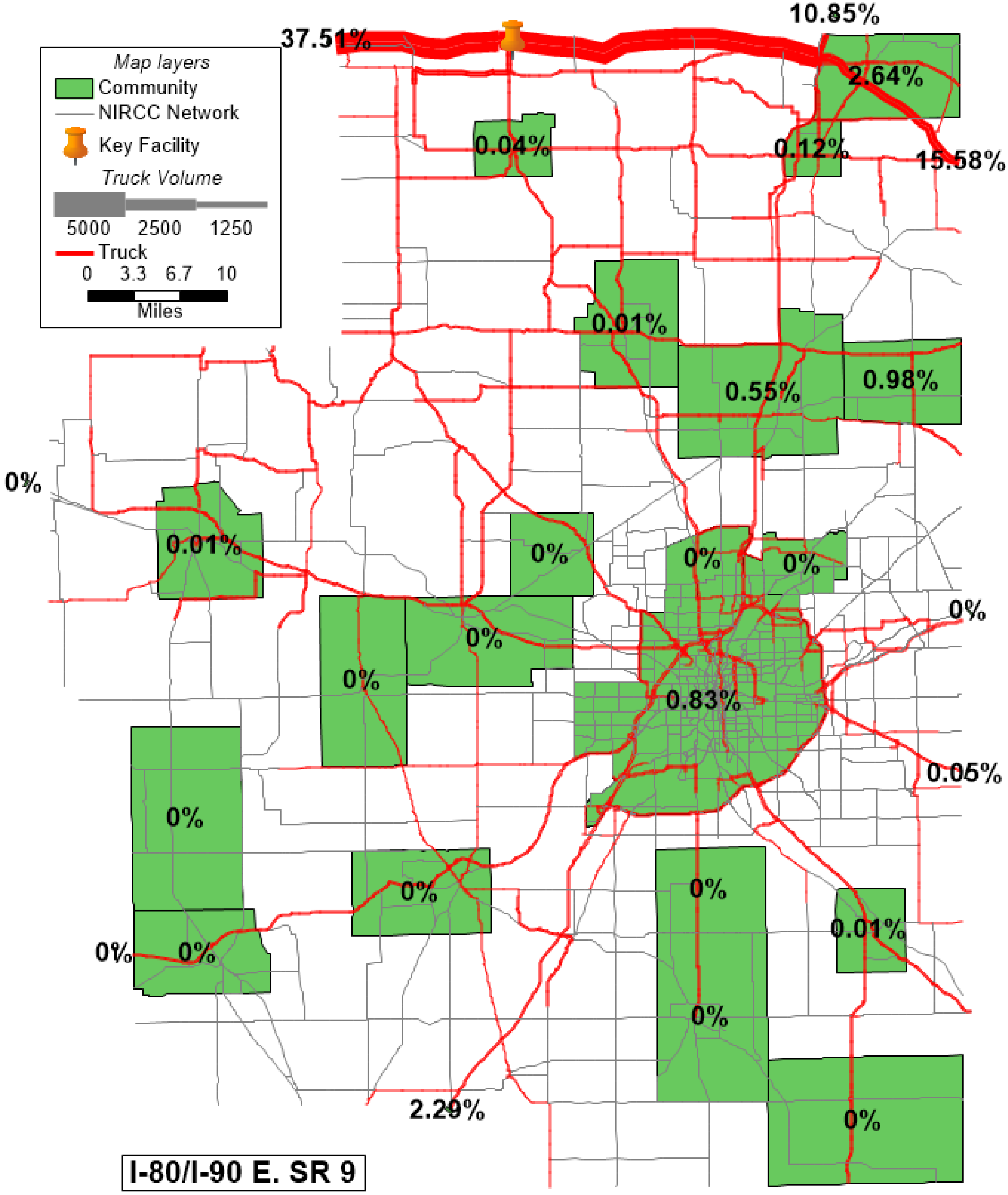


FIGURE 90. DEMAND AT I-80/90 E. OF SR 9


Map layers
Community
NIRCC Network
Key Facility
Truck Volume
5000
2500
1250
Truck
0
3.3
6.7
10
Miles
37.51%
10.85%
2.64%
0.04%
0.12%
15.58%
0.01%
0.55%
0.98%
0%
0.01%
0%
0%
0%
0%
0%
0%
0.83%
0.05%
0%
0%
0%
0.01%
0%
0%
0%
2.29%
0%
I-80/I-90 E. SR 9

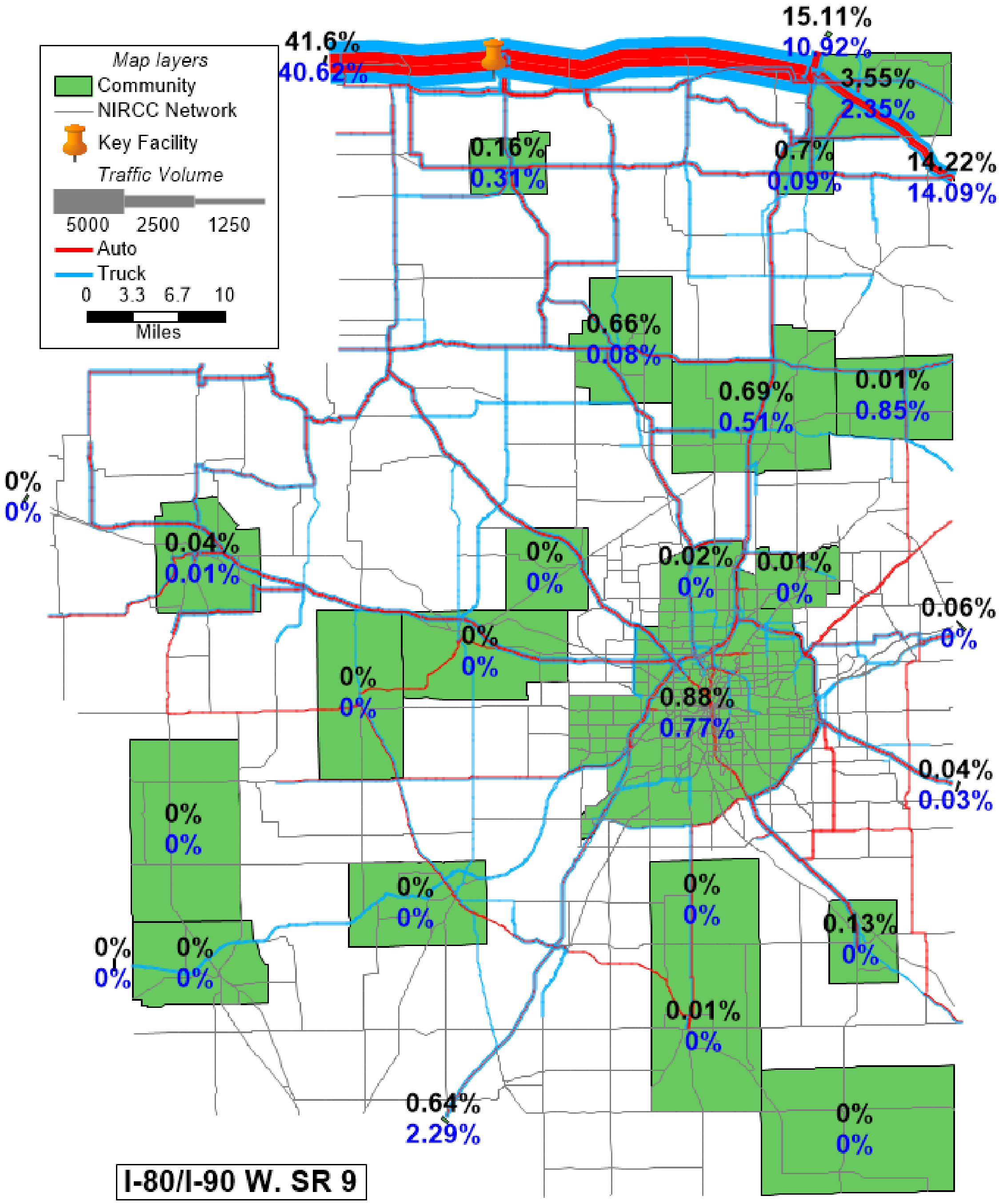
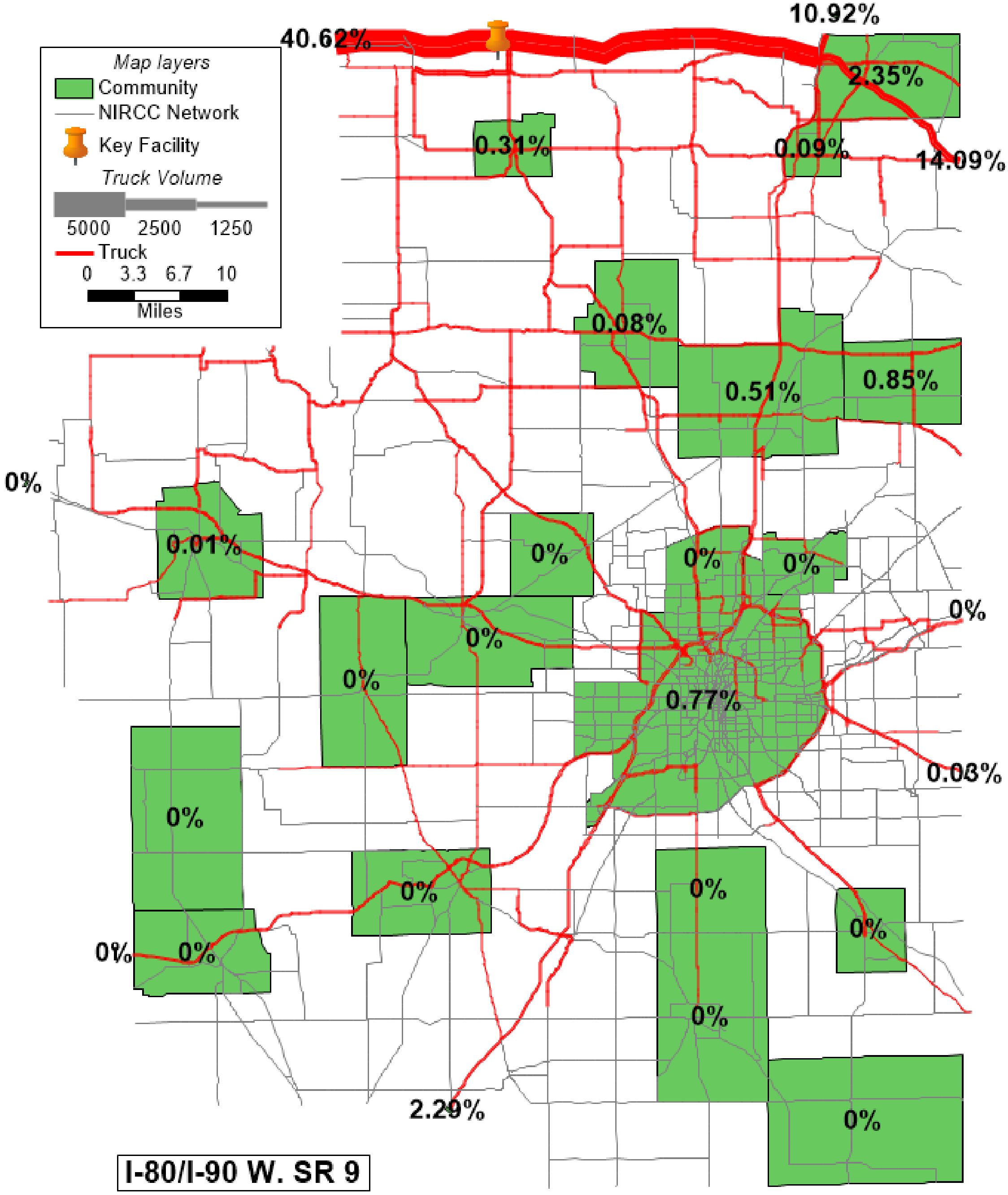


FIGURE 91. DEMAND AT I-80/90 W. OF SR 9


Map layers
Community
NIRCC Network
Key Facility
Truck Volume
5000
2500
1250
Truck
0
3.3
6.7
10
Miles
40.62%
10.92%
2.35%
0.31%
0.09%
14.09%
0.08%
0.51%
0.85%
0%
0.01%
0%
0%
0%
0%
0%
0%
0.77%
0.03%
0%
0%
0%
0%
0%
0%
0%
0%
2.29%
0%
I-80/I-90 W. SR 9

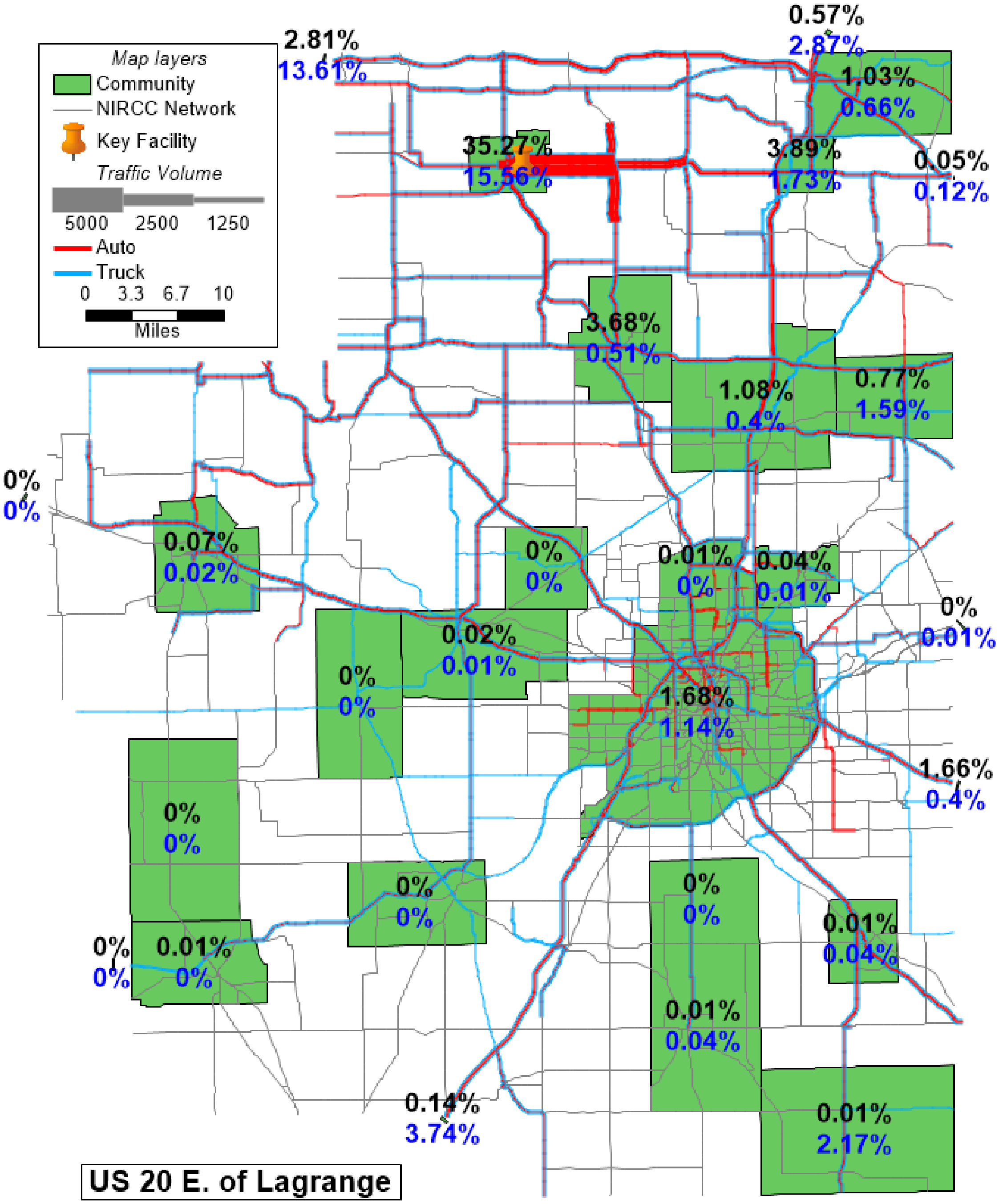
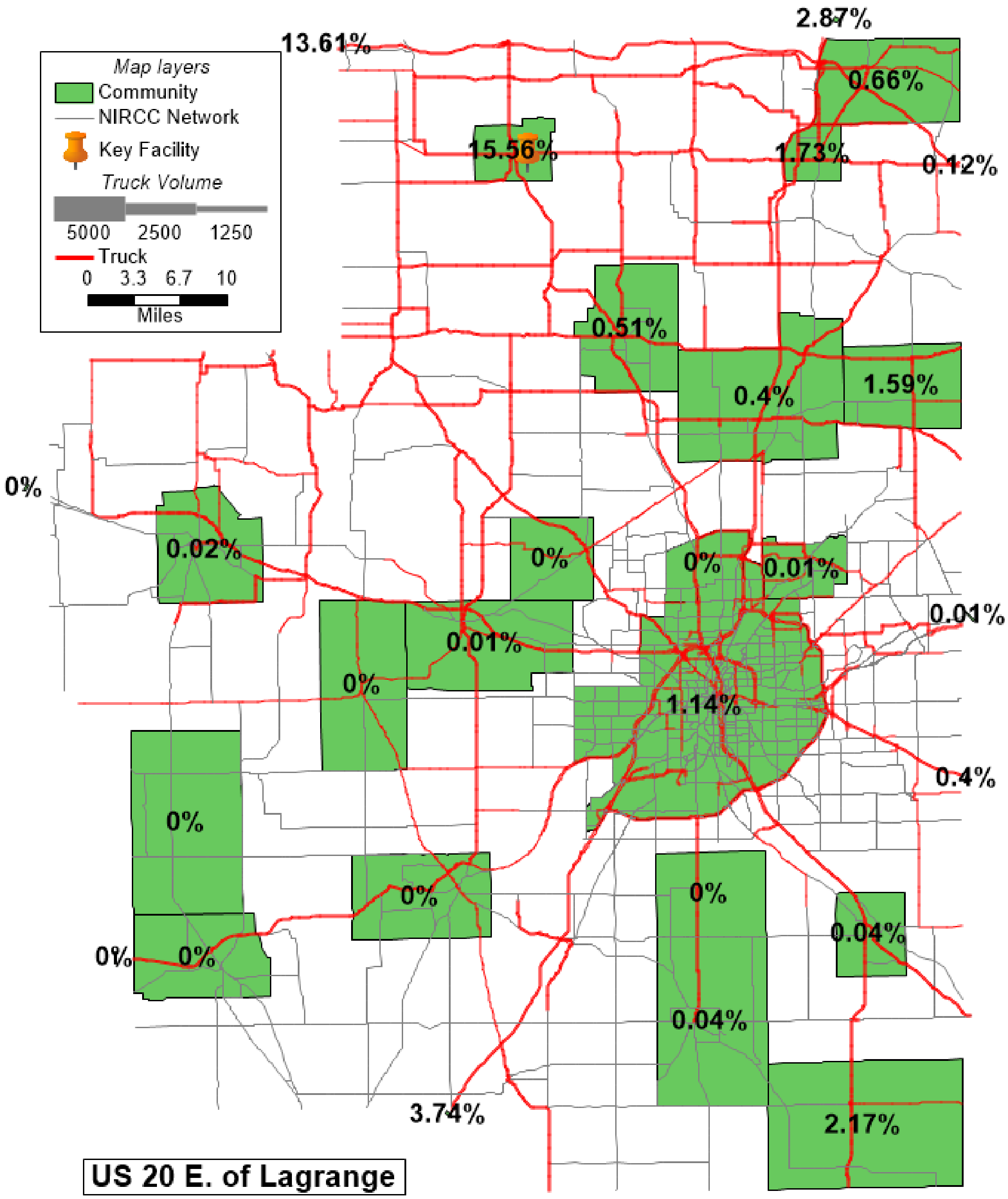


FIGURE 92. DEMAND AT US 20 E. OF LAGRANGE


Map layers
Community
NIRCC Network
Key Facility
Truck Volume
5000
2500
1250
Truck
0
3.3
6.7
10
Miles
13.61%
2.87%
0.66%
15.56%
1.73%
0.12%
0.51%
0.4%
1.59%
0%
0.02%
0%
0%
0.01%
0.01%
0.01%
0%
1.14%
0.4%
0%
0%
0%
0%
0.04%
0%
0%
0.04%
3.74%
2.17%
US 20 E. of Lagrange

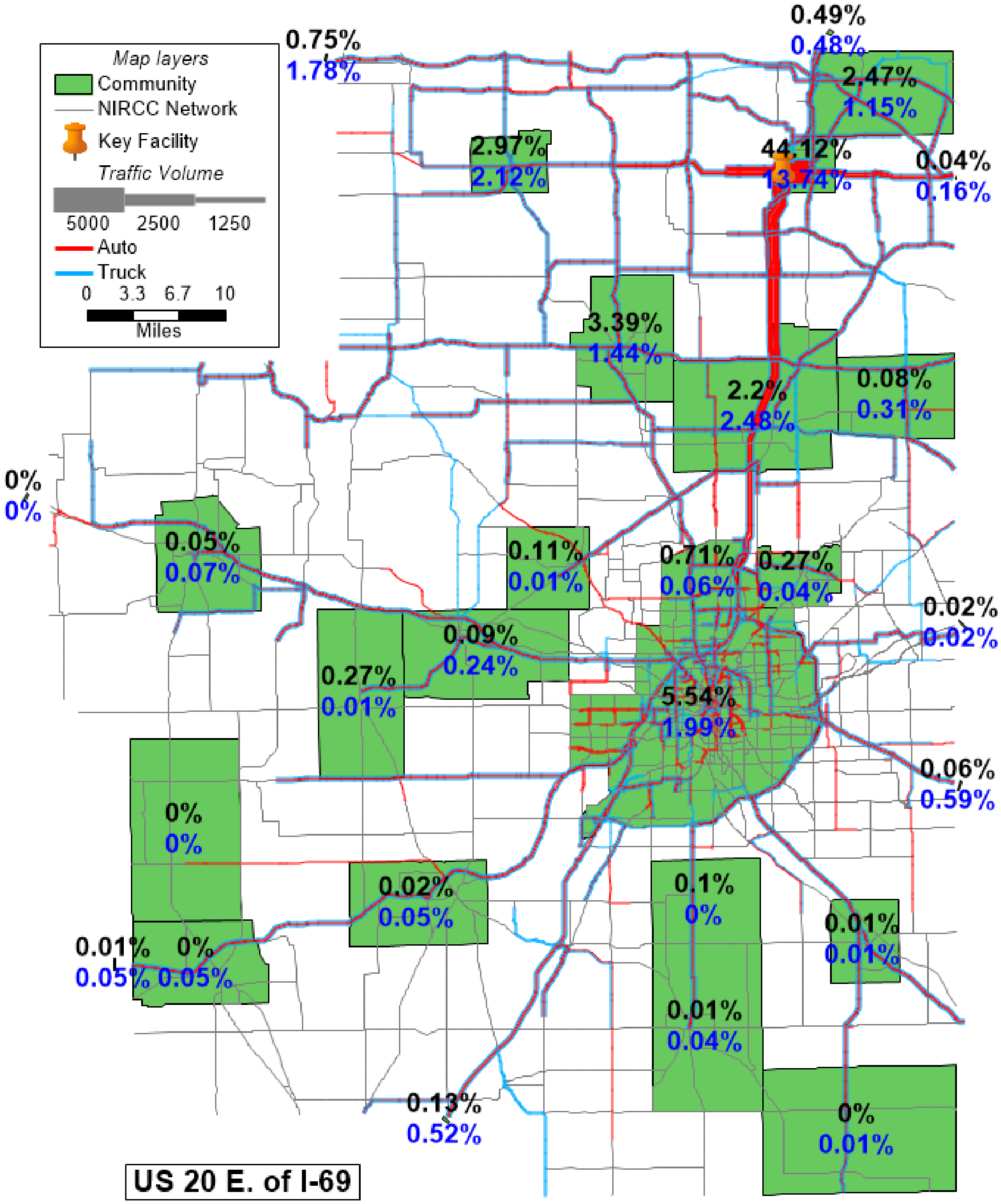
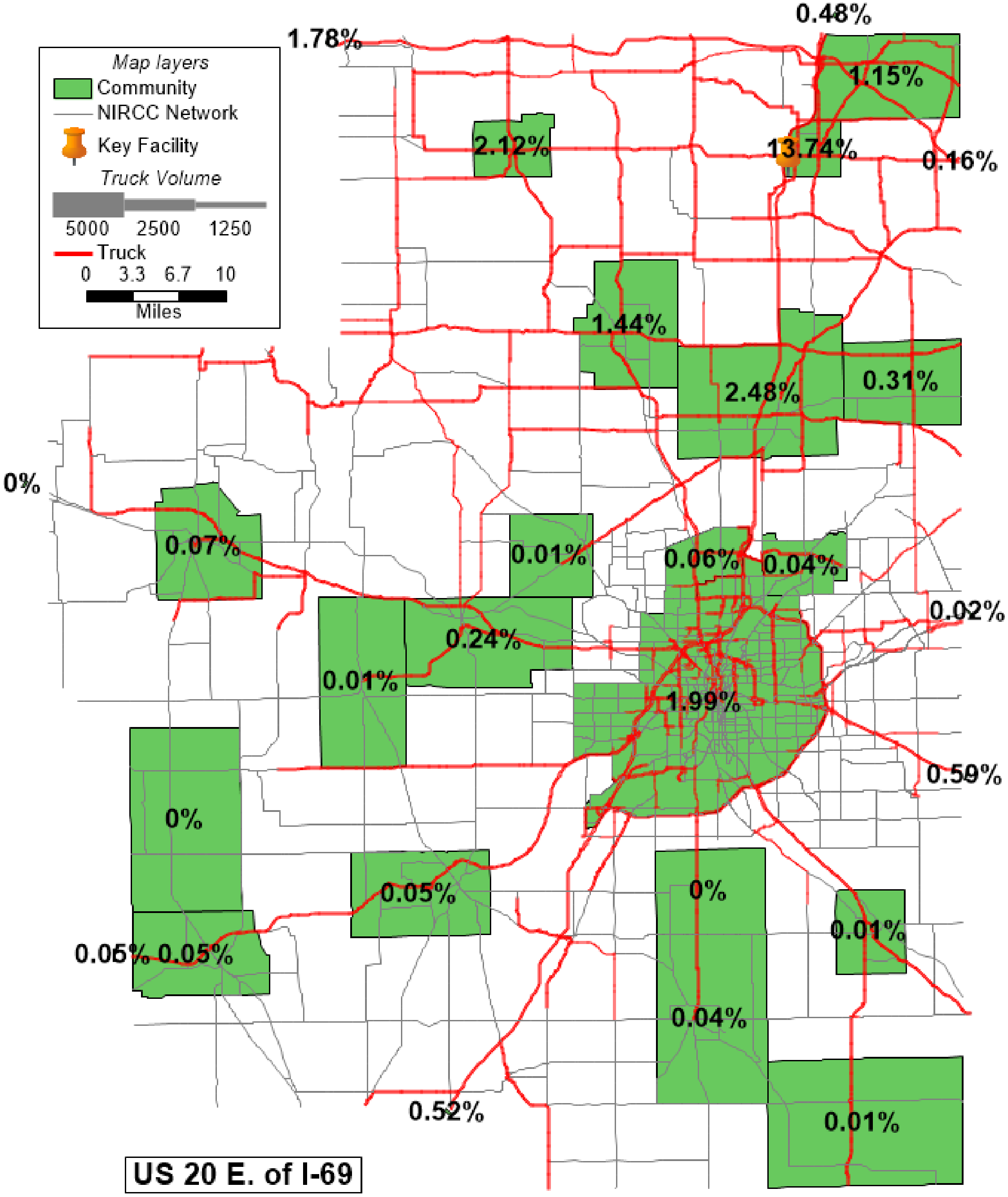


FIGURE 93. DEMAND AT US 20 E. OF I-69



Map layers
Community
NIRCC Network
Key Facility
Truck Volume
5000
2500
1250
Truck
0
3.3
6.7
10
Miles
1.78%
0.48%
1.15%
2.12%
13.74%
0.16%
1.44%
2.48%
0.31%
0%
0.07%
0.01%
0.06%
0.04%
0.02%
0.24%
0.01%
1.99%
0.59%
0%
0.05%
0%
0.01%
0.05%
0.05%
0.04%
0.52%
0.01%
US 20 E. of I-69

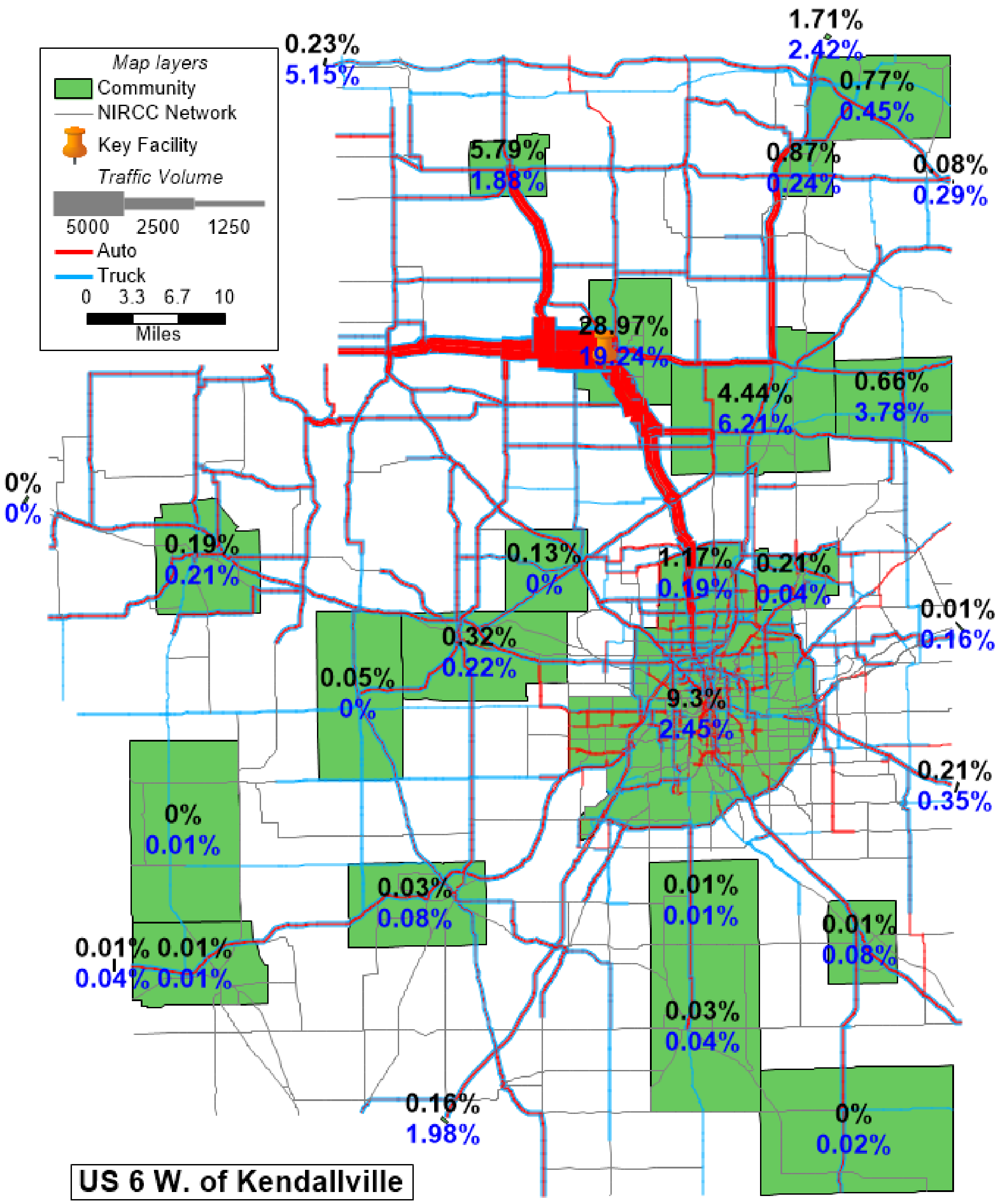
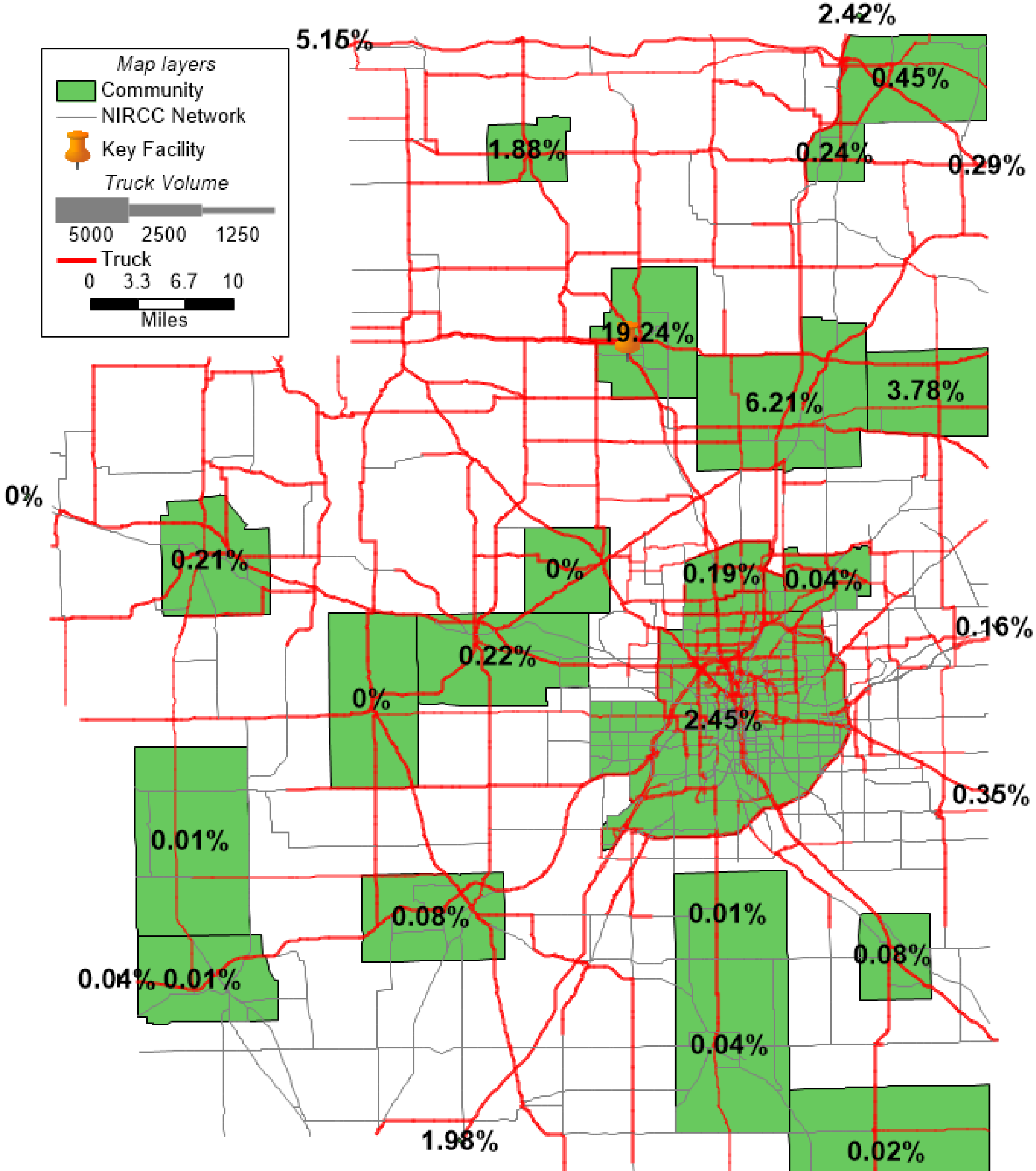


FIGURE 94. US 6 W. OF KENALLVILLE



US 6 W. of Kendallville

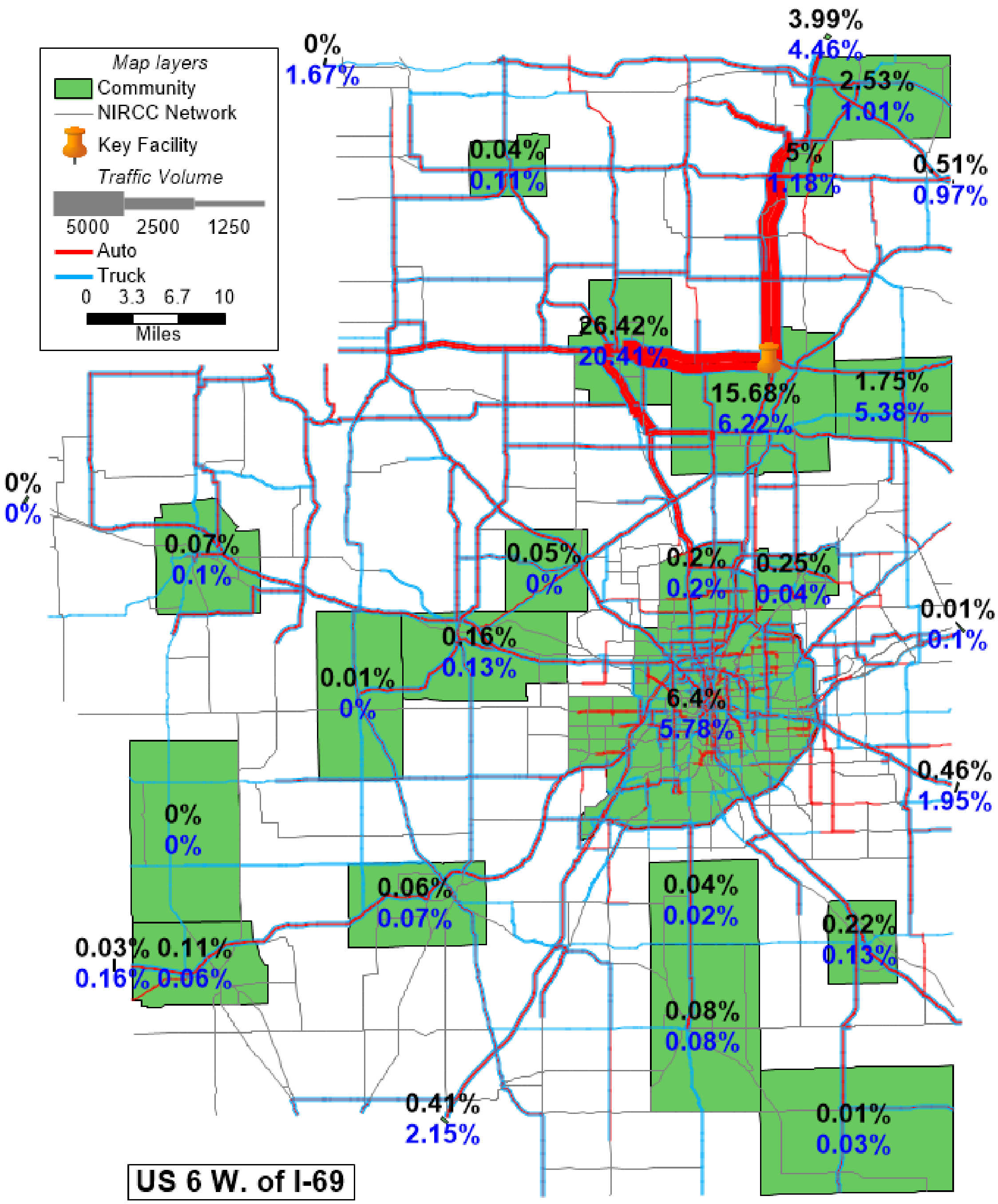
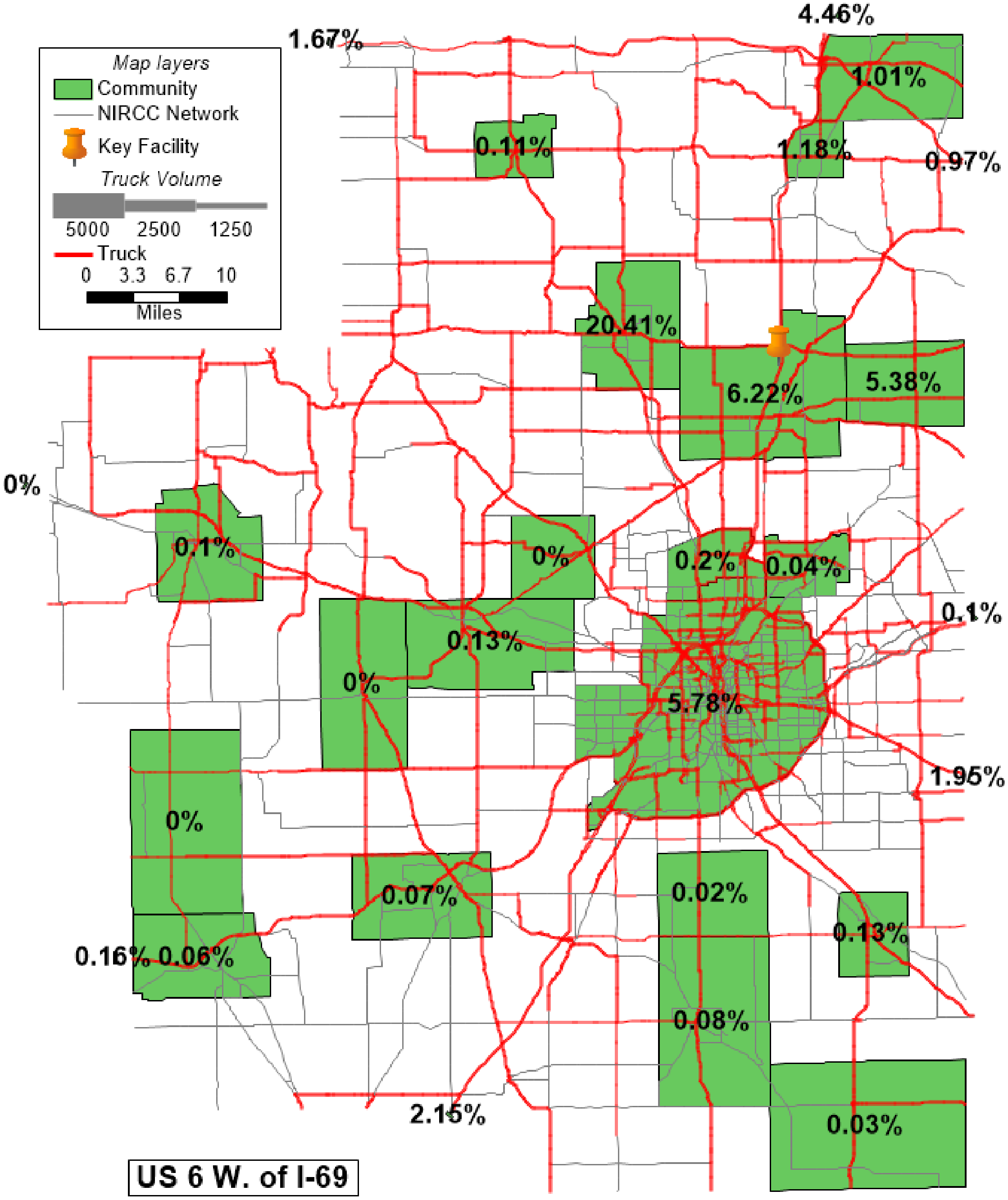


FIGURE 95. DEMAND AT US 6 W. OF I-69



Map layers
Community
NIRCC Network
Key Facility
Truck Volume
5000
2500
1250
Truck
0
3.3
6.7
10
Miles
1.67%
4.46%
1.01%
0.11%
1.18%
0.97%
20.41%
6.22%
5.38%
0%
0.1%
0%
0.2%
0.04%
0.1%
0.13%
0%
5.78%
1.95%
0%
0.07%
0.02%
0.13%
0.16%
0.06%
0.08%
2.15%
0.03%

US 6 W. of I-69

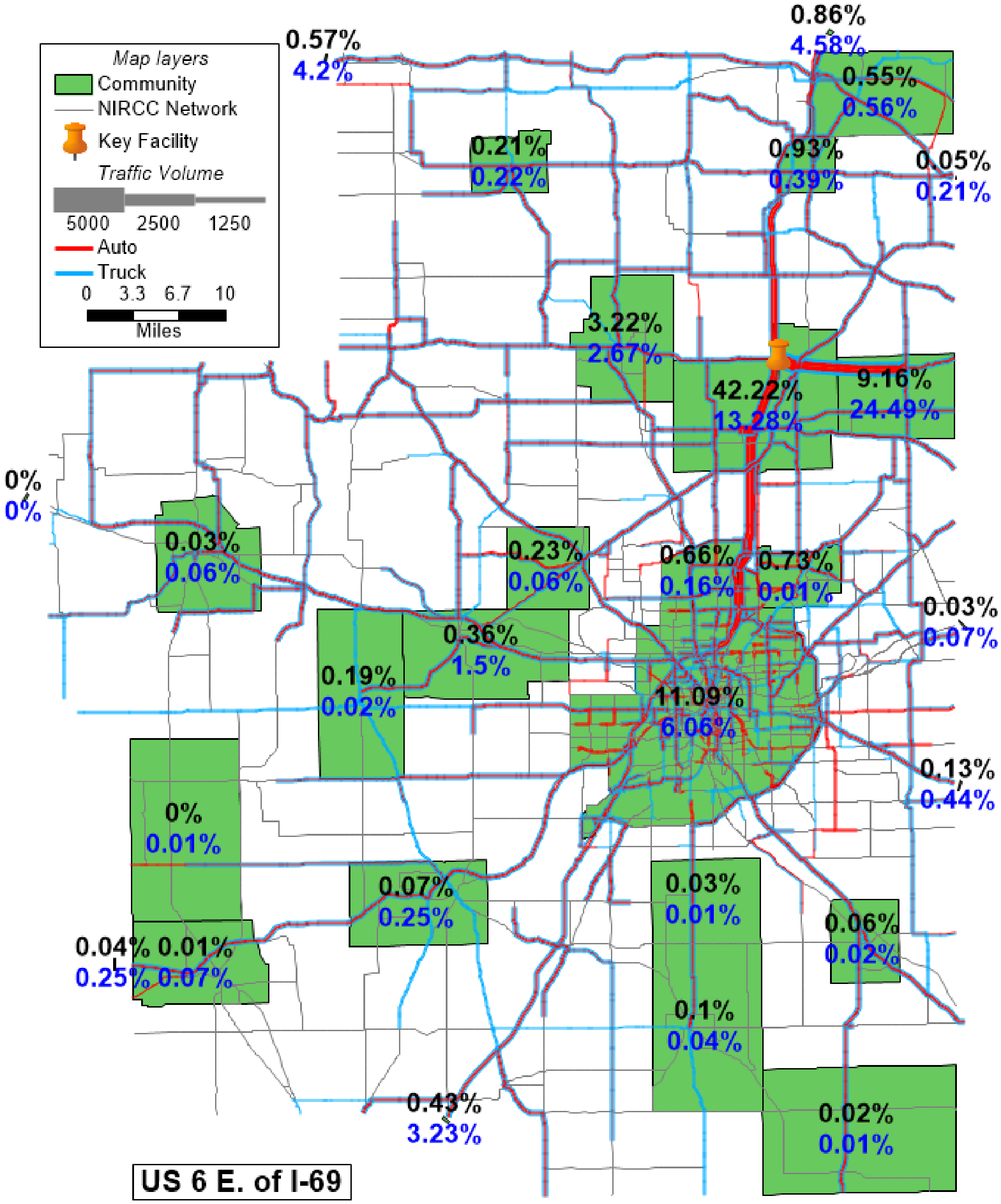
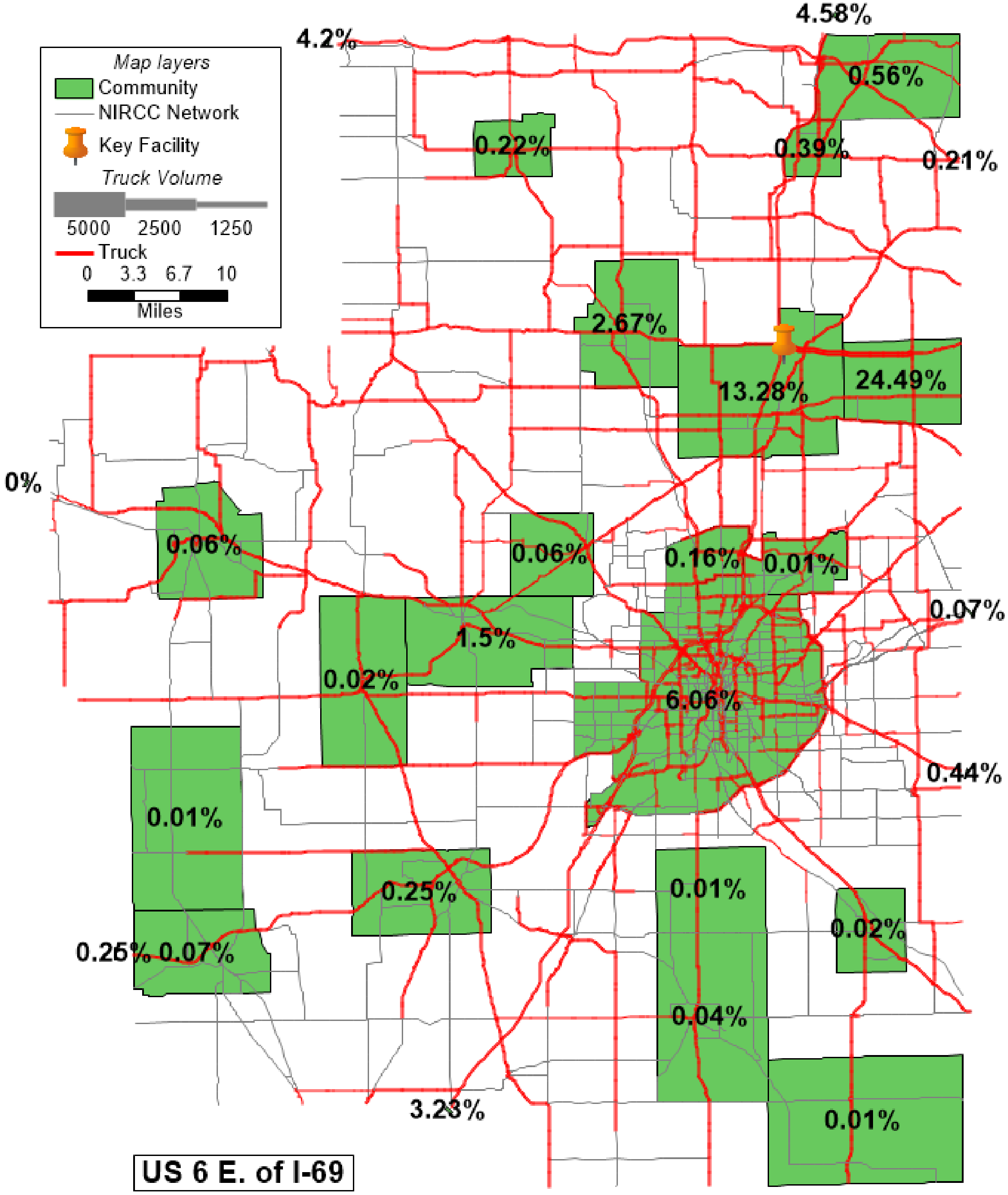


FIGURE 96. DEMAND AT US 6 E. OF I-69



Map layers
Community
NIRCC Network
Key Facility
Truck Volume
5000
2500
1250
Truck
0
3.3
6.7
10
Miles
4.2%
4.58%
0.56%
0.22%
0.39%
0.21%
2.67%
13.28%
24.49%
0%
0.06%
0.06%
0.16%
0.01%
0.07%
1.5%
0.02%
6.06%
0.44%
0.01%
0.25%
0.01%
0.02%
0.26%
0.07%
0.04%
3.23%
0.01%
US 6 E. of I-69

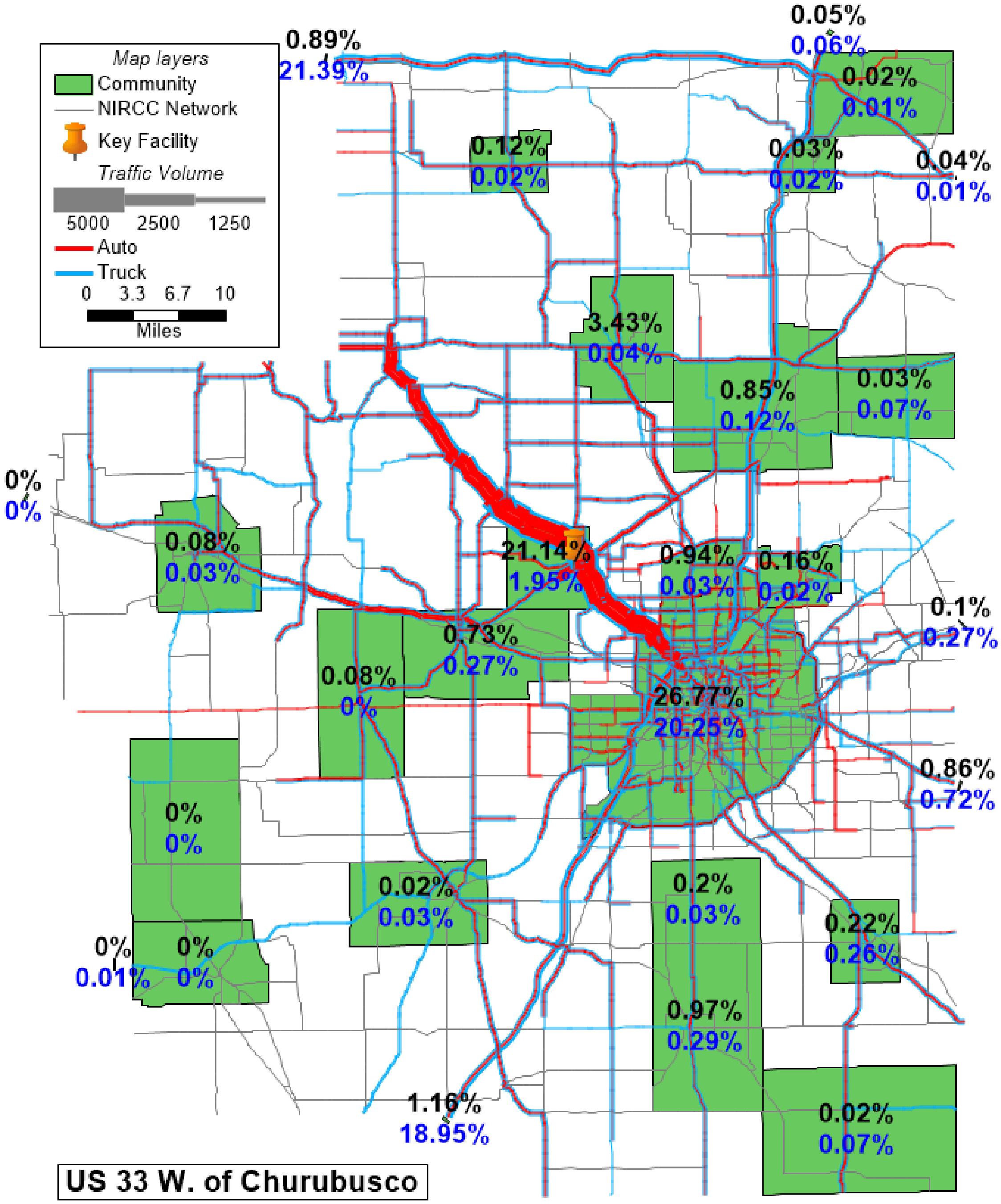
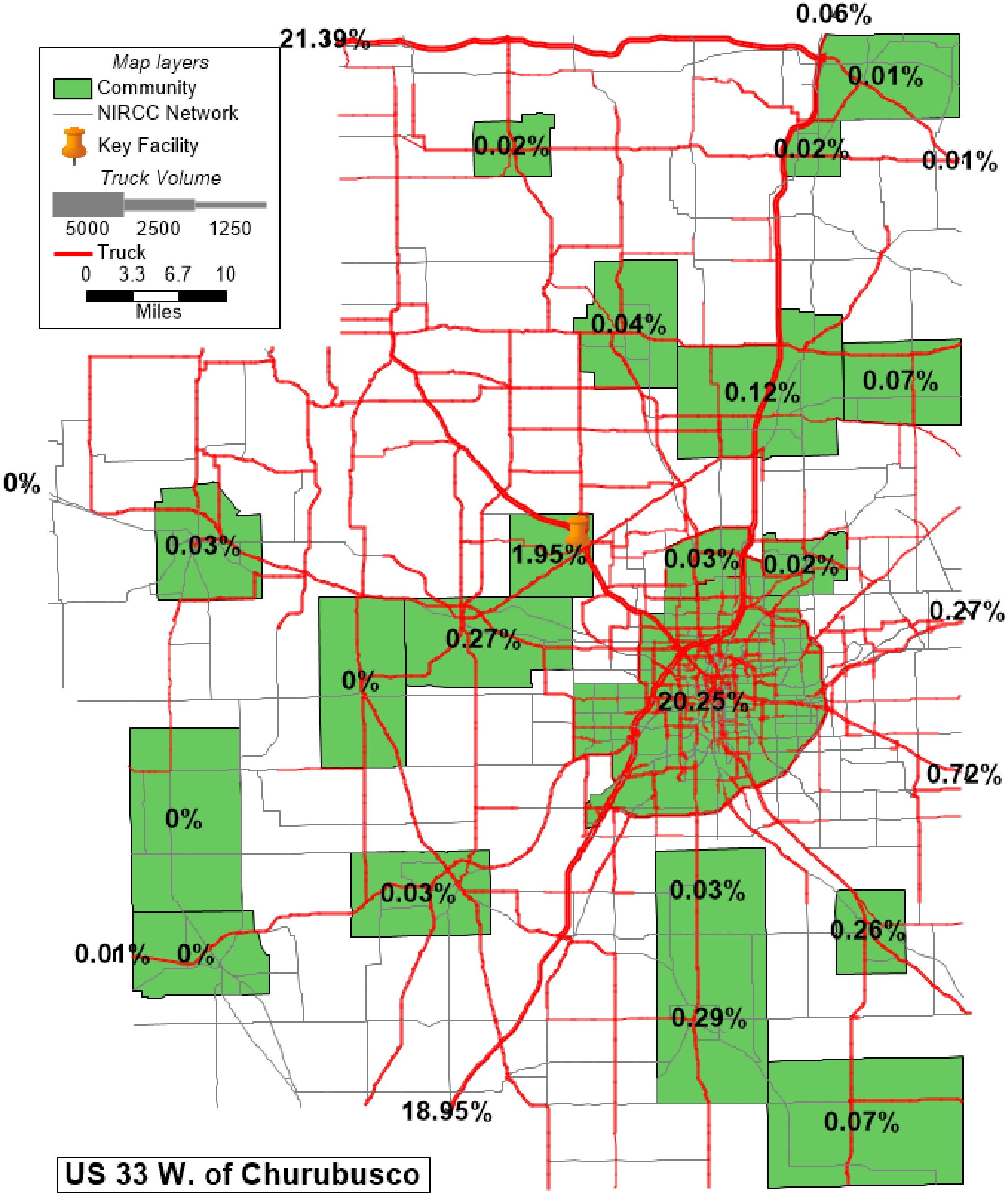


FIGURE 97. DEMAND AT US 33 W. OF CHURRUBUSCO



Map layers
Community
NIRCC Network
Key Facility
Truck Volume
5000
2500
1250
Truck
0
3.3
6.7
10
Miles
21.39%
0.06%
0.01%
0.02%
0.02%
0.01%
0.04%
0.12%
0.07%
0%
0.03%
1.95%
0.03%
0.02%
0.27%
0.27%
0%
20.25%
0.72%
0%
0.03%
0.03%
0.26%
0.01%
0%
0.29%
18.95%
0.07%
US 33 W. of Churubusco

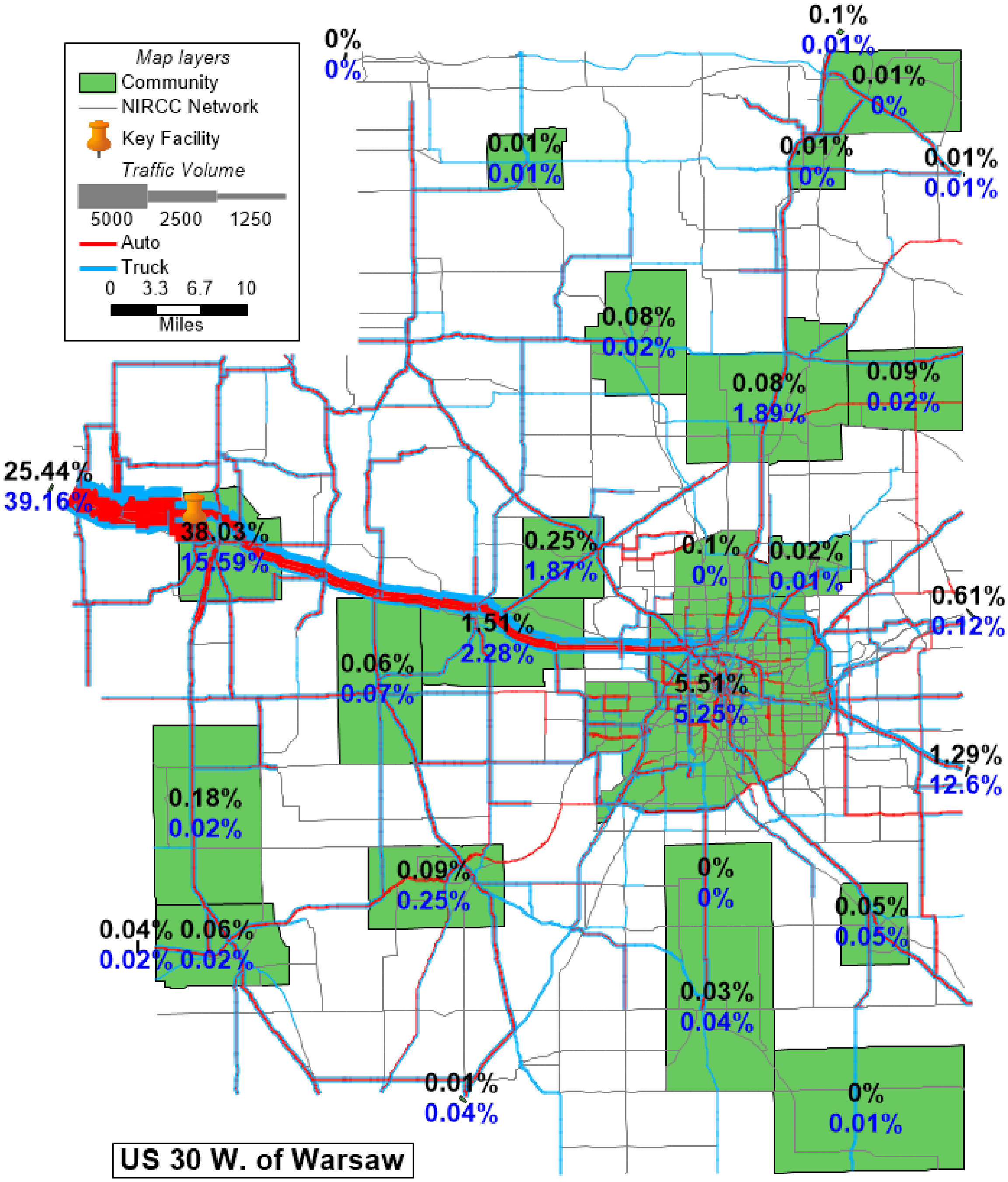
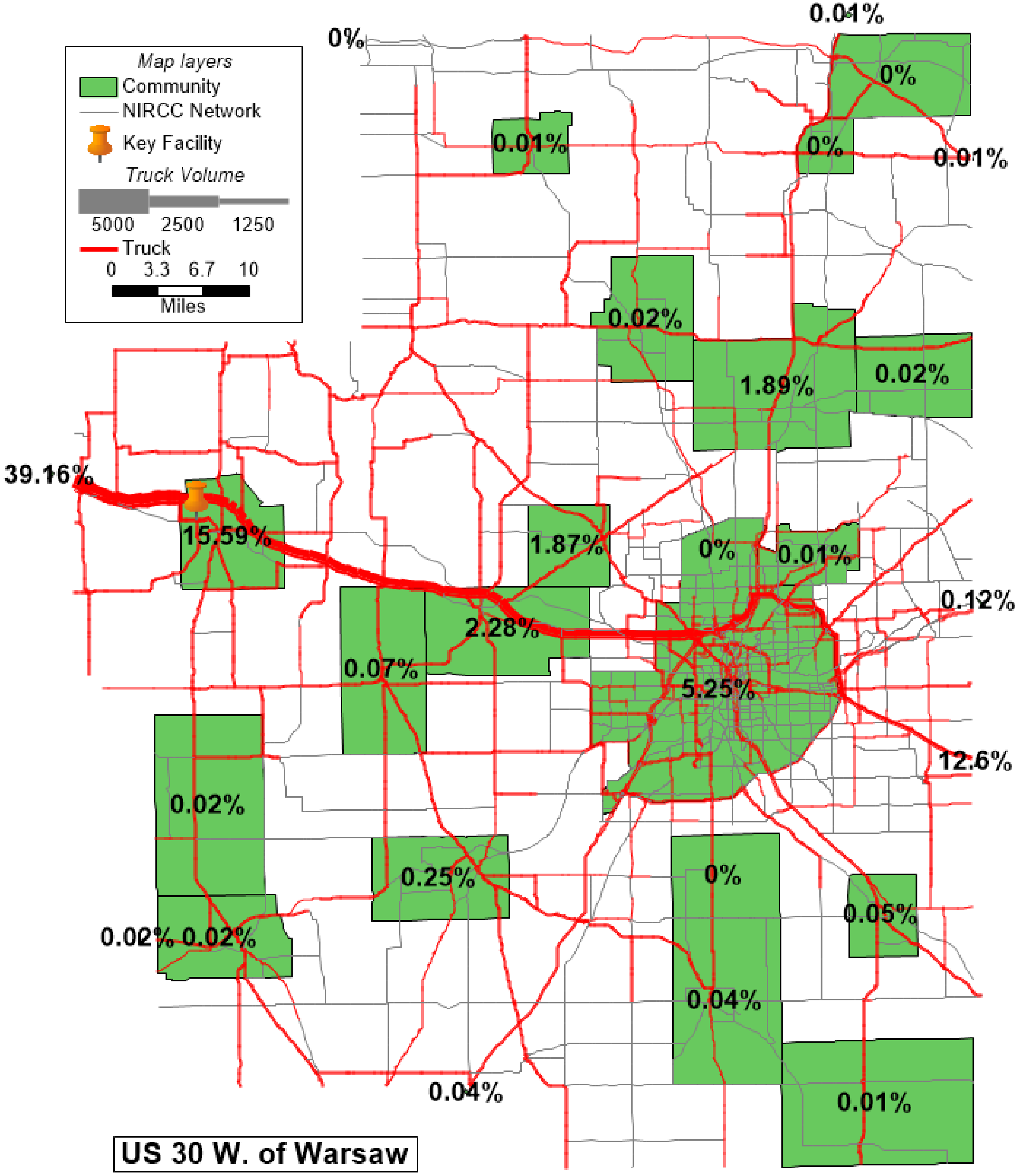


FIGURE 98. DEMAND AT US 30 W. OF WARSAW



US 30 W. of Warsaw

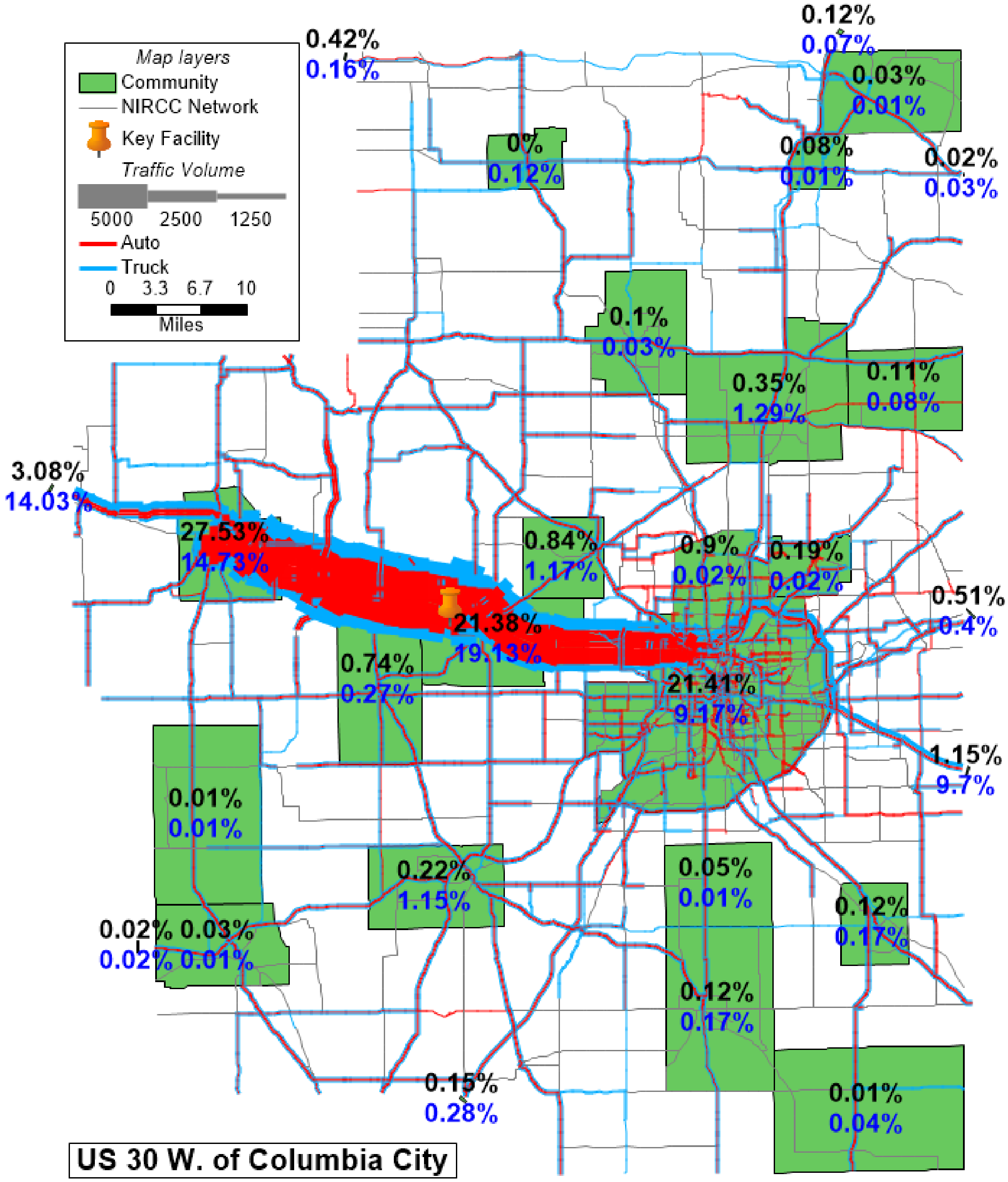
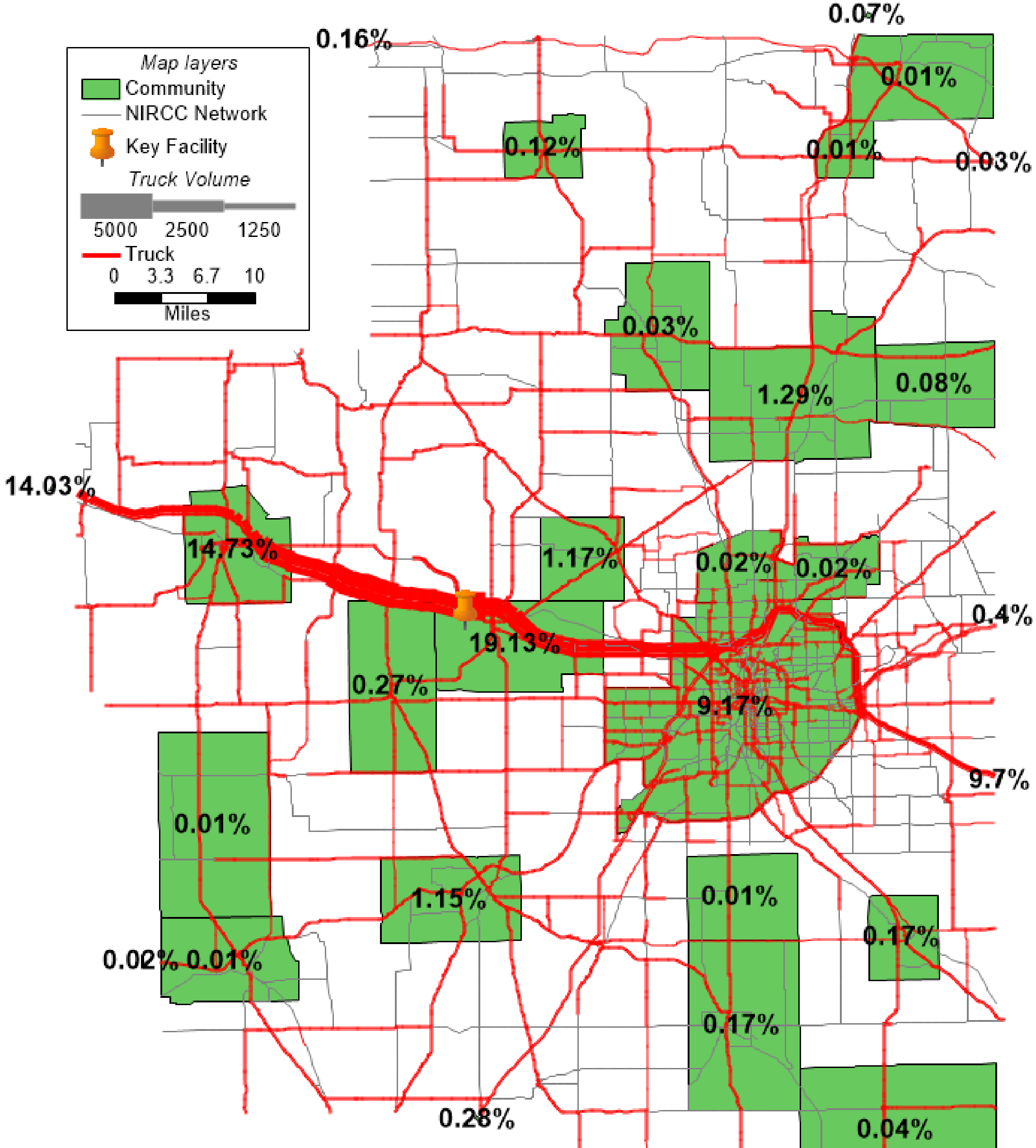


FIGURE 99. DEMAND AT US 30 W. OF COLUMBIA CITY



US 30 W. of Columbia City

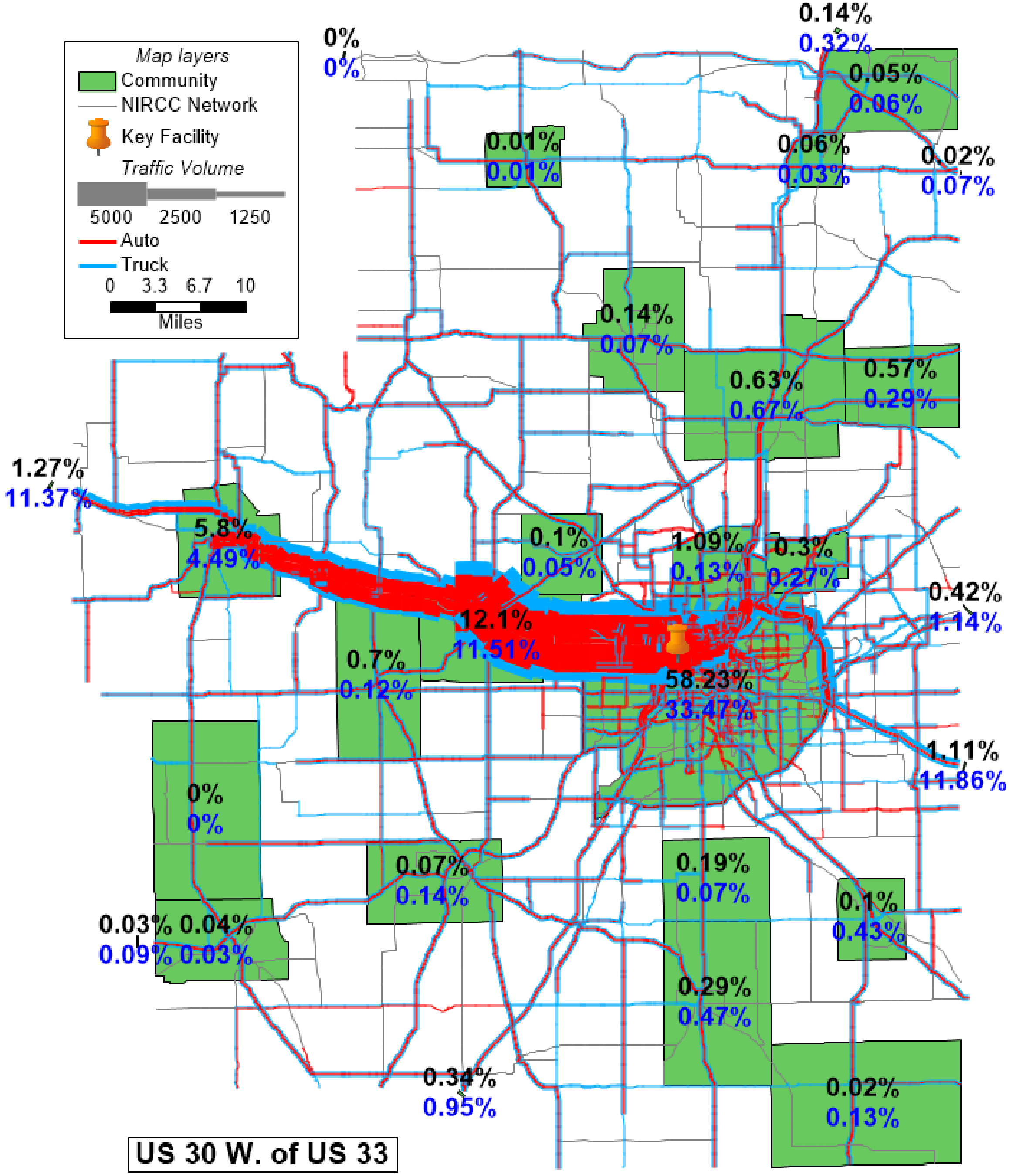
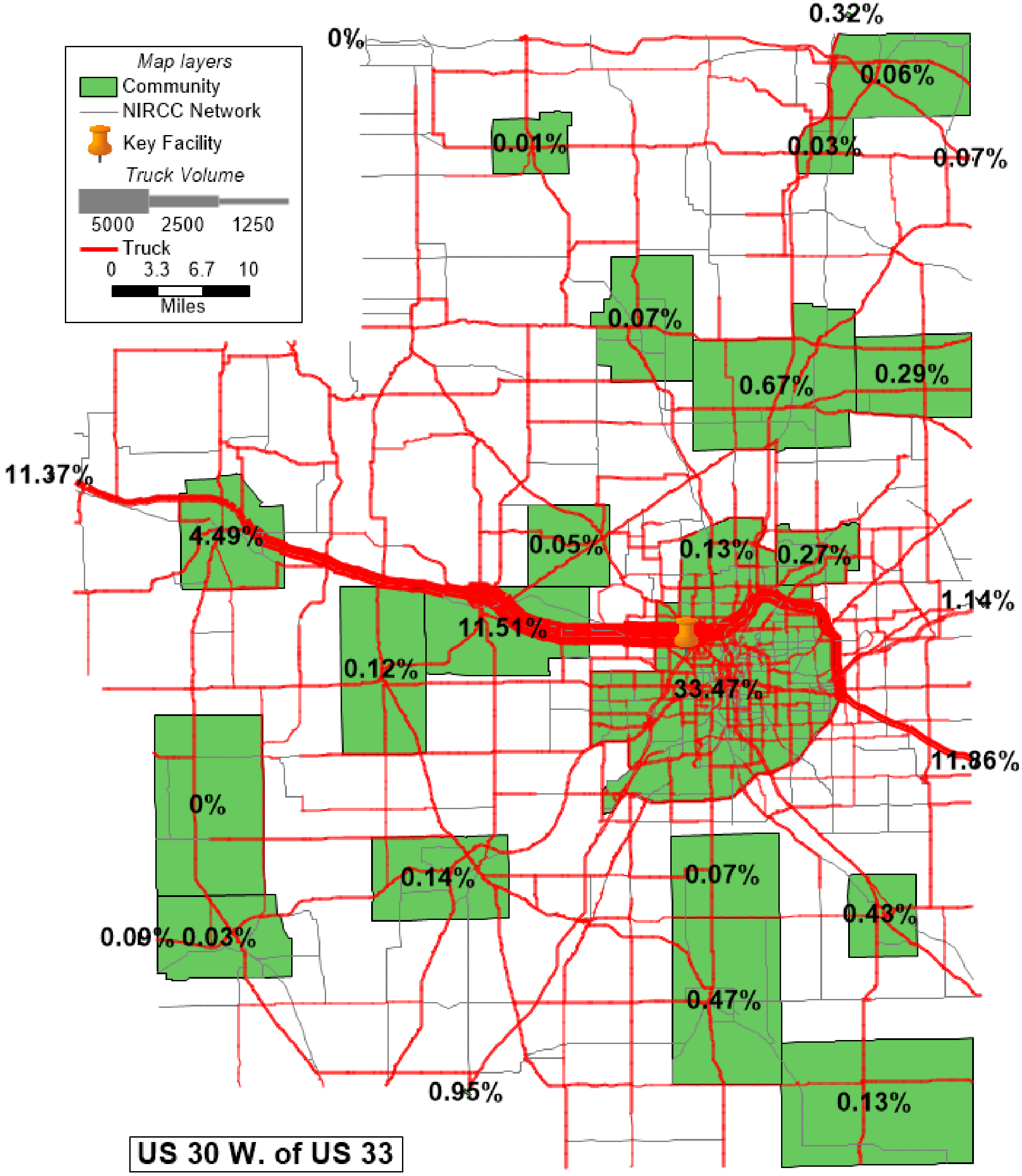


FIGURE 100. DEMAND AT US 30 W. OF US 33


Map layers
Community
NIRCC Network
Key Facility
Truck Volume
5000
2500
1250
Truck
0
3.3
6.7
10
Miles
0%
0.32%
0.06%
0.01%
0.03%
0.07%
0.07%
0.67%
0.29%
11.37%
4.49%
0.05%
0.13%
0.27%
1.14%
11.51%
0.12%
33.47%
11.86%
0%
0.14%
0.07%
0.43%
0.09%
0.03%
0.47%
0.95%
0.13%
US 30 W. of US 33

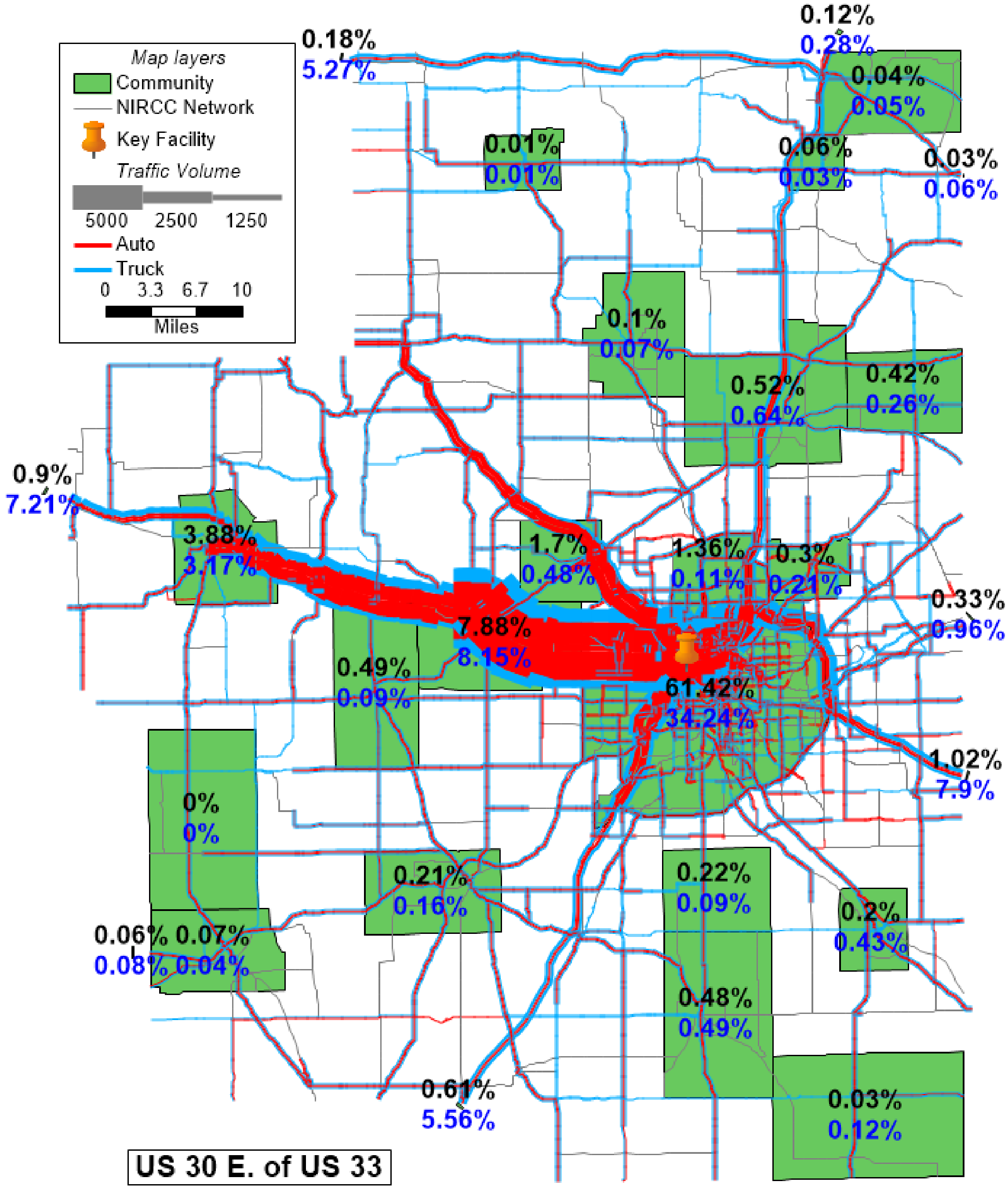
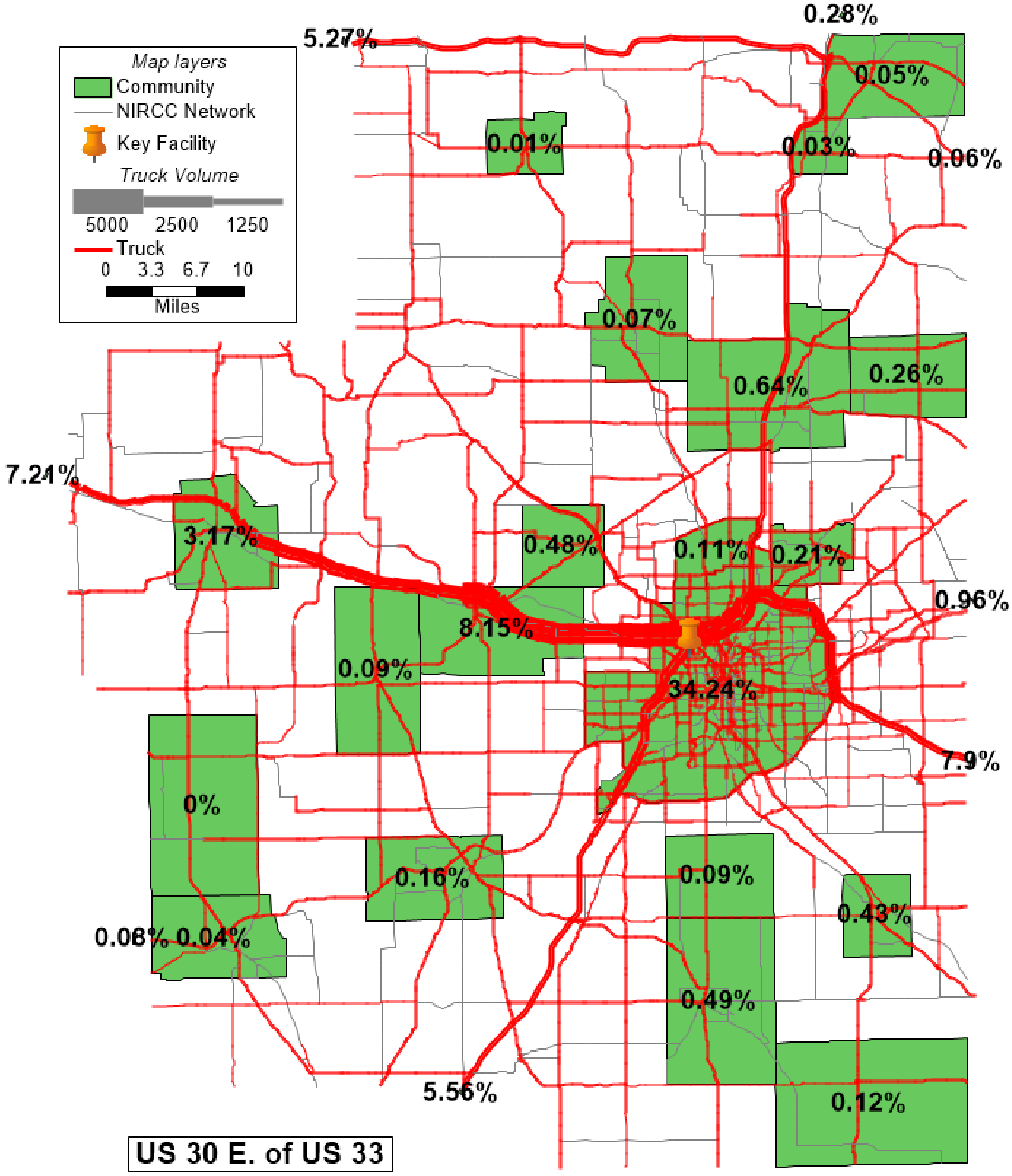


FIGURE 101. DEMAND AT US 30 E. OF US 33



Map layers
Community
NIRCC Network
Key Facility
Truck Volume
5000 2500 1250
Truck
0 3.3 6.7 10
Miles
5.27%
0.28%
0.05%
0.01%
0.03%
0.06%
0.07%
0.64%
0.26%
7.21%
3.17%
0.48%
0.11%
0.21%
0.96%
8.15%
0.09%
34.24%
7.9%
0%
0.16%
0.09%
0.43%
0.08%
0.04%
0.49%
5.56%
0.12%
US 30 E. of US 33

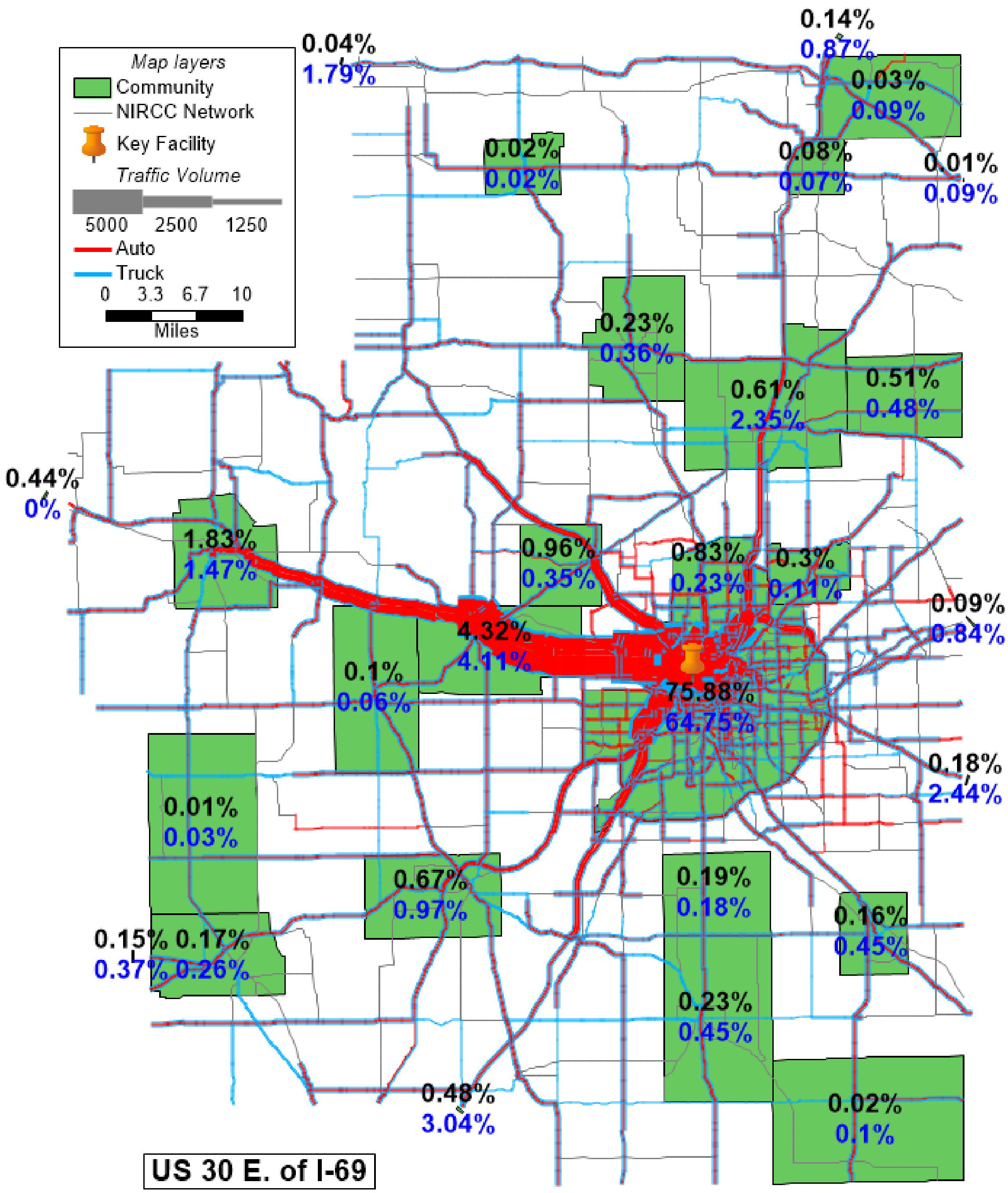
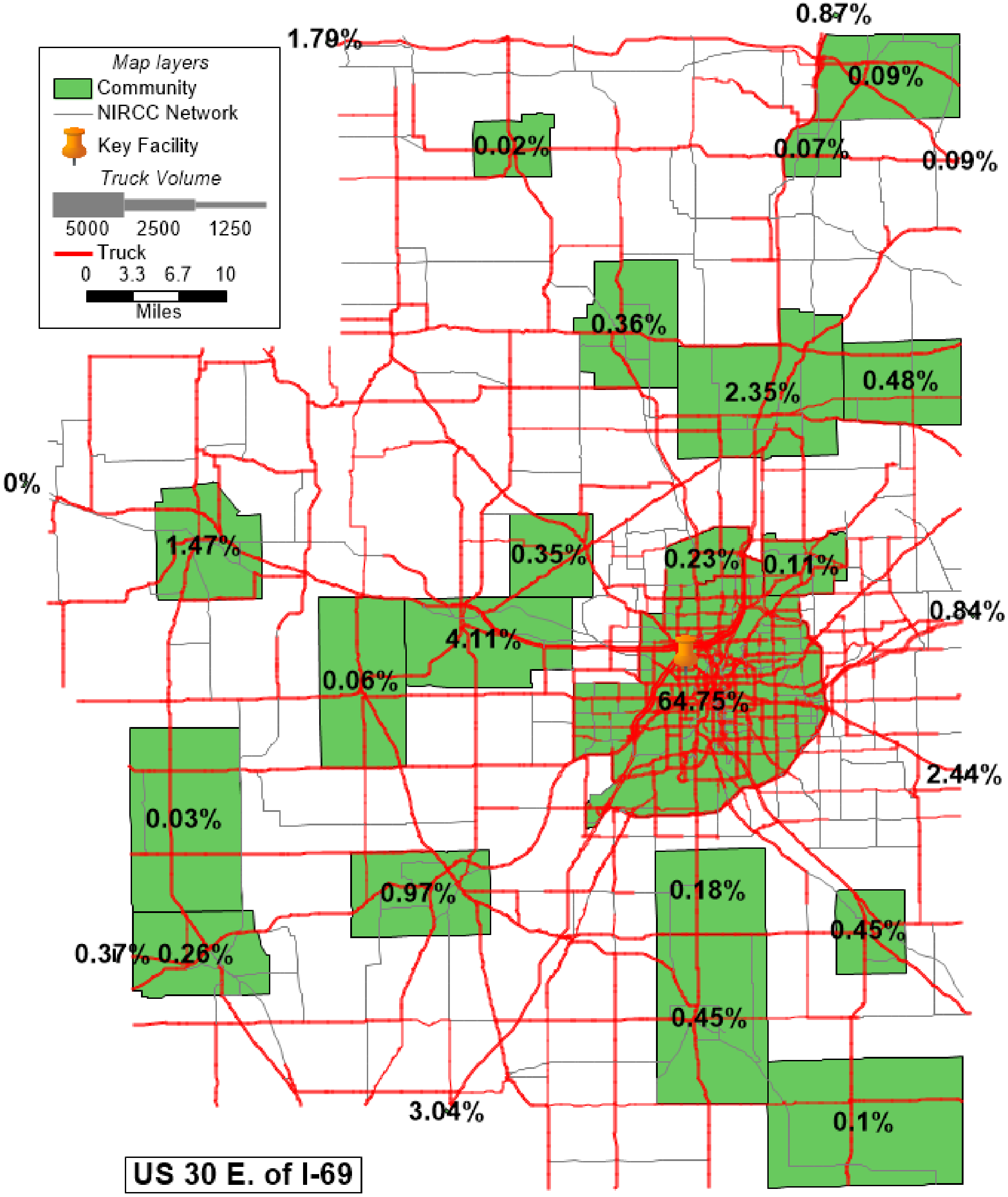


FIGURE 102. DEMAND AT US 30 E. OF I-69


Map layers
Community
NIRCC Network
Key Facility
Truck Volume
5000
2500
1250
Truck
0
3.3
6.7
10
Miles
1.79%
0.87%
0.09%
0.02%
0.07%
0.09%
0.36%
2.35%
0.48%
0%
1.47%
0.35%
0.23%
0.11%
0.84%
4.11%
0.06%
64.75%
2.44%
0.03%
0.97%
0.18%
0.45%
0.37%
0.26%
0.45%
3.04%
0.1%
US 30 E. of I-69

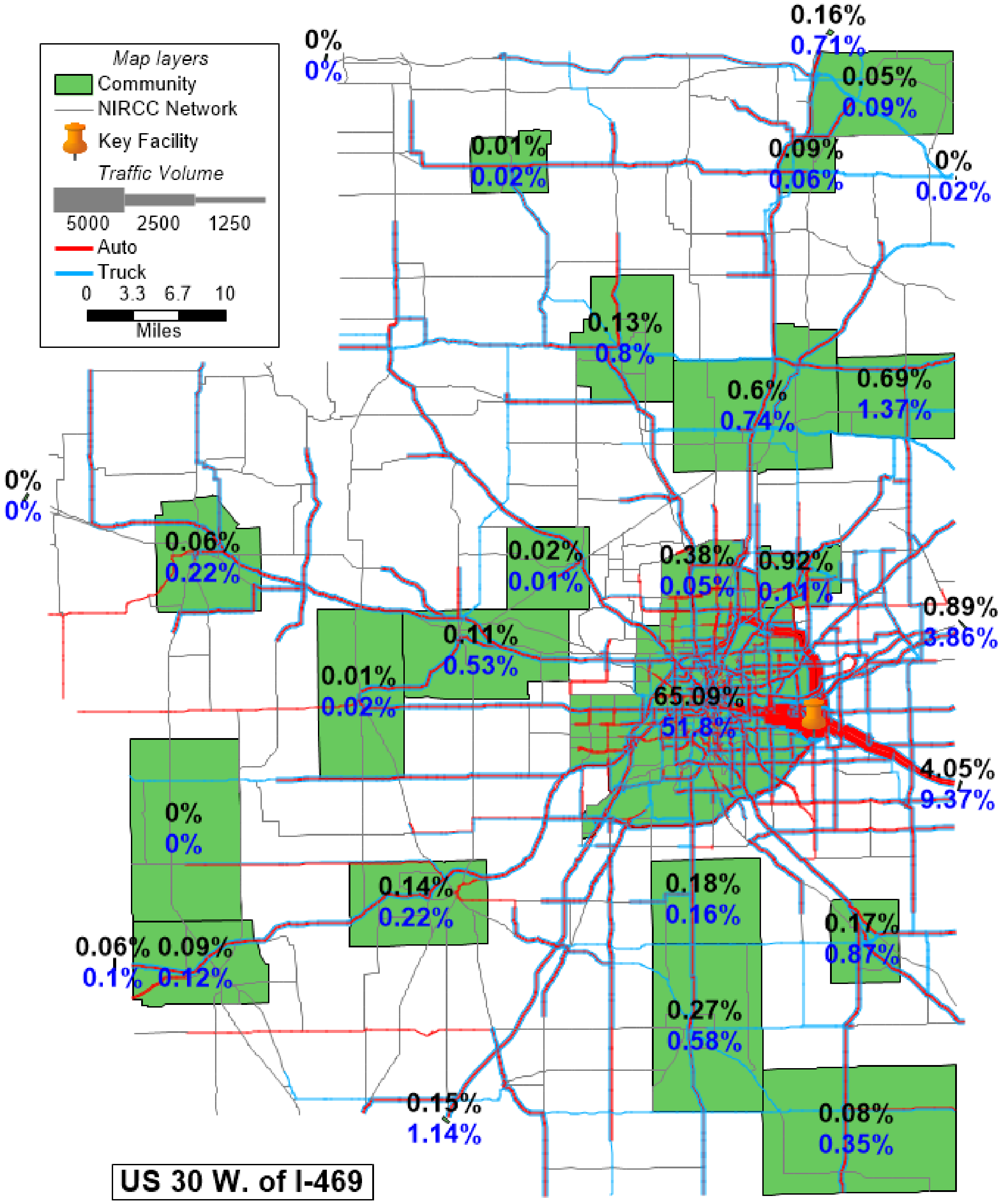
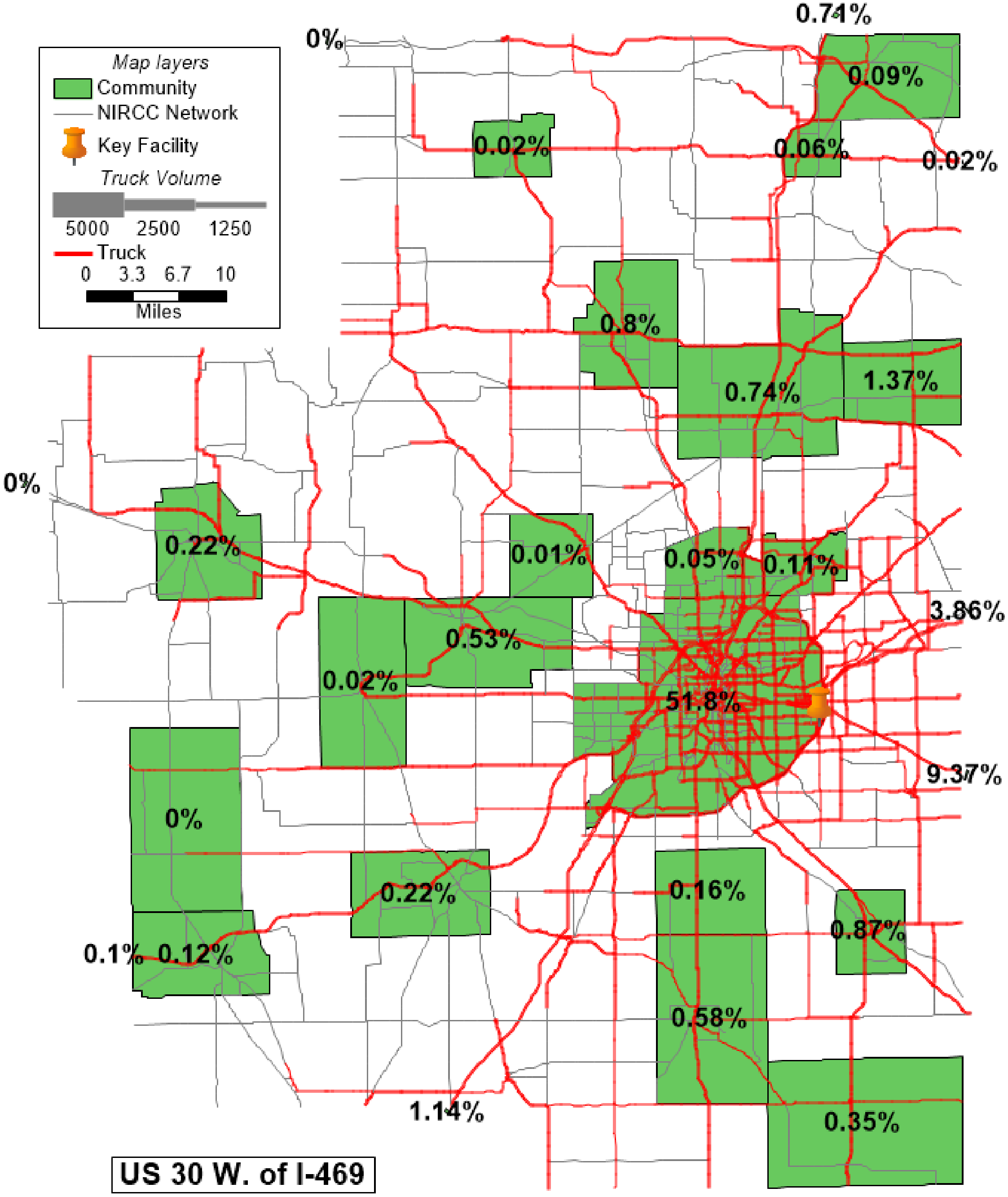


FIGURE 103. DEMAND AT US 30 W. OF I-469



Map layers
Community
NIRCC Network
Key Facility
Truck Volume
5000 2500 1250
Truck
0 3.3 6.7 10
Miles
0%
0.74%
0.09%
0.02%
0.06%
0.02%
0.8%
0.74%
1.37%
0%
0.22%
0.01%
0.05%
0.11%
3.86%
0.53%
0.02%
51.8%
9.37%
0%
0.22%
0.16%
0.87%
0.1%
0.12%
0.58%
1.14%
0.35%
US 30 W. of I-469

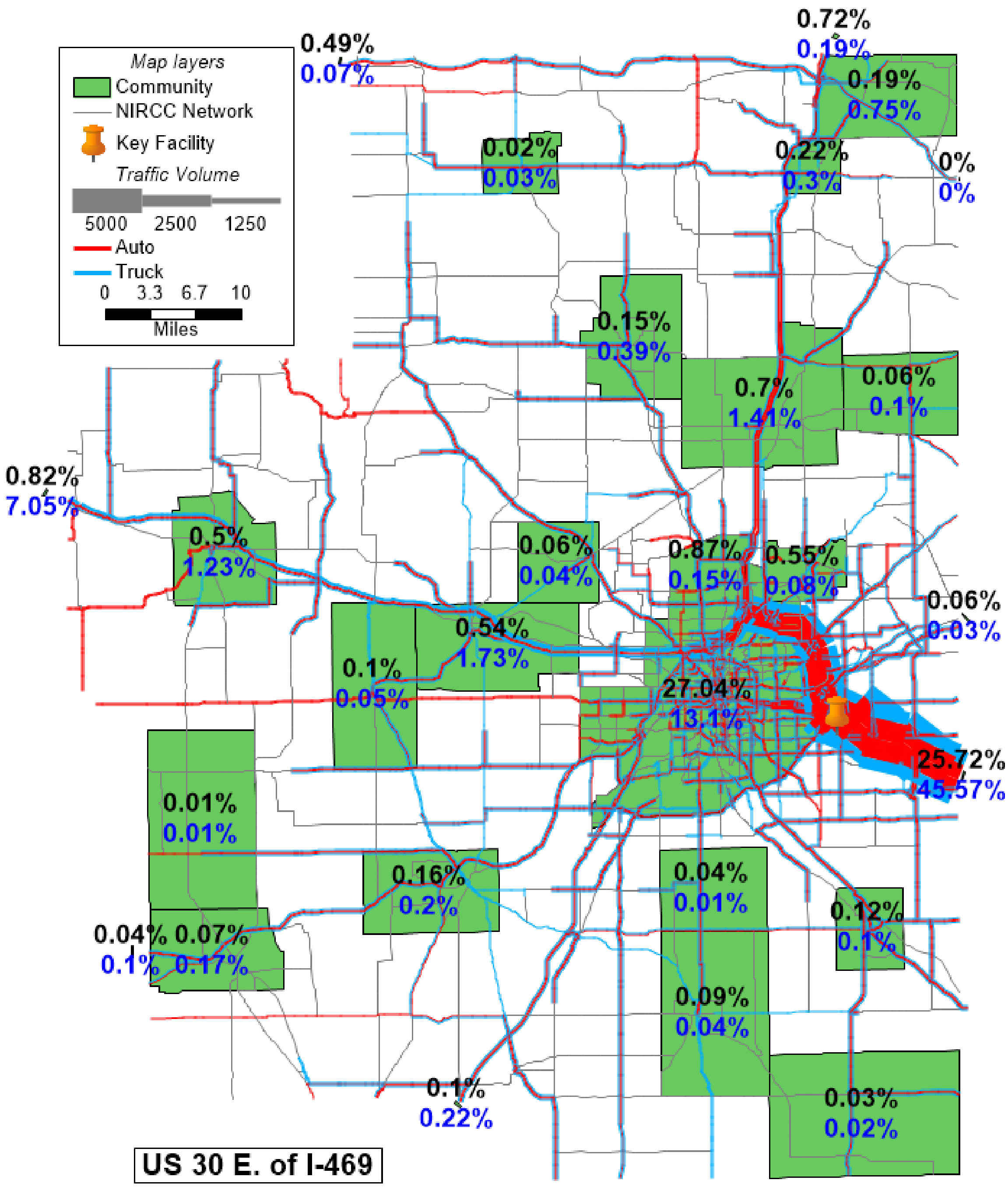
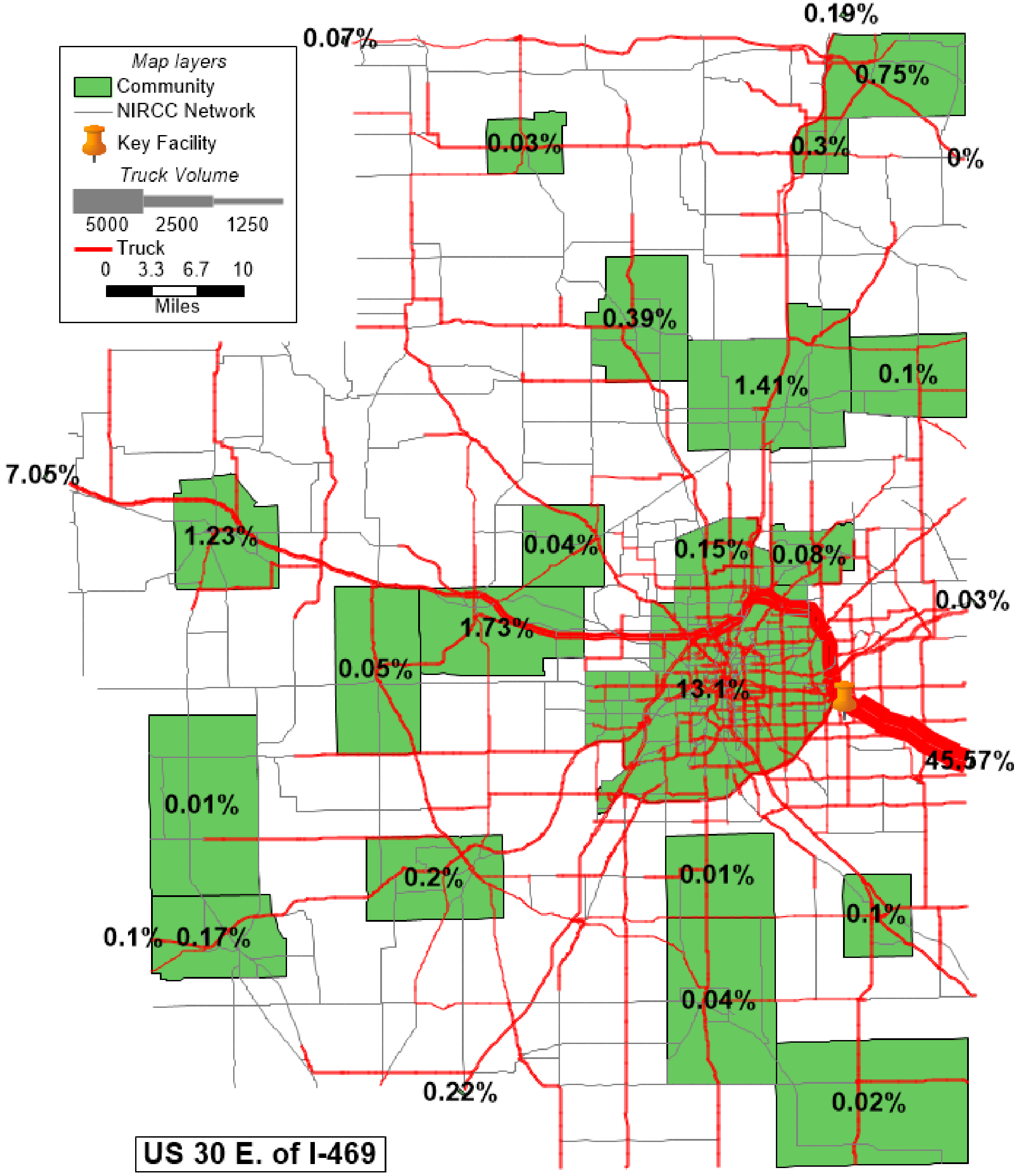


FIGURE 104. DEMAND AT US 30 E. OF I-469


Map layers
Community
NIRCC Network
Key Facility
Truck Volume
5000 2500 1250
Truck
0 3.3 6.7 10
Miles
0.07%
0.19%
0.75%
0.03%
0.3%
0%
0.39%
1.41%
0.1%
7.05%
1.23%
0.04%
0.15%
0.08%
0.03%
1.73%
0.05%
13.1%
45.57%
0.01%
0.2%
0.01%
0.1%
0.1%
0.17%
0.04%
0.22%
0.02%
US 30 E. of I-469

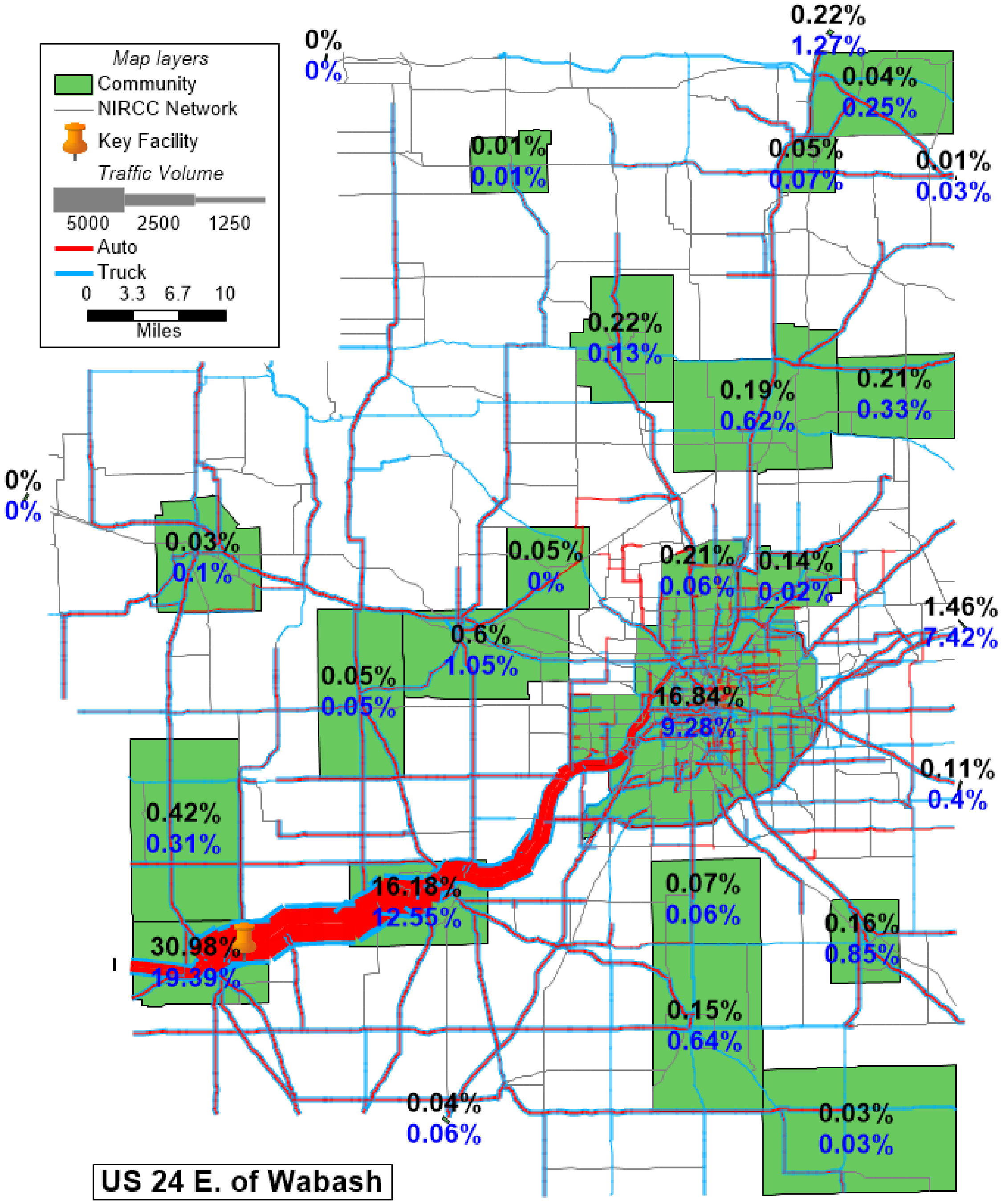
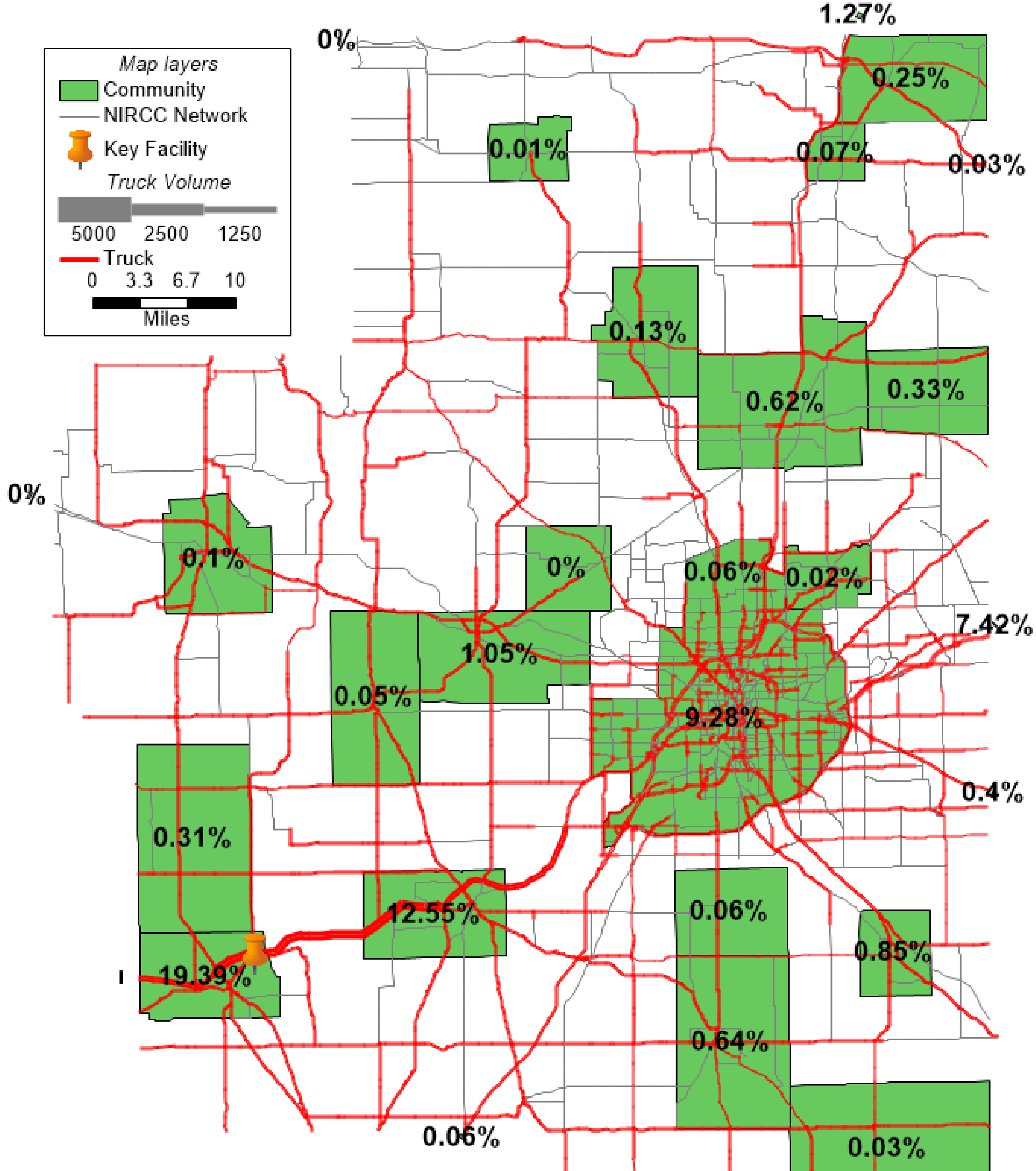


FIGURE 105. DEMAND AT US 24 E. OF WABASH



US 24 E. of Wabash

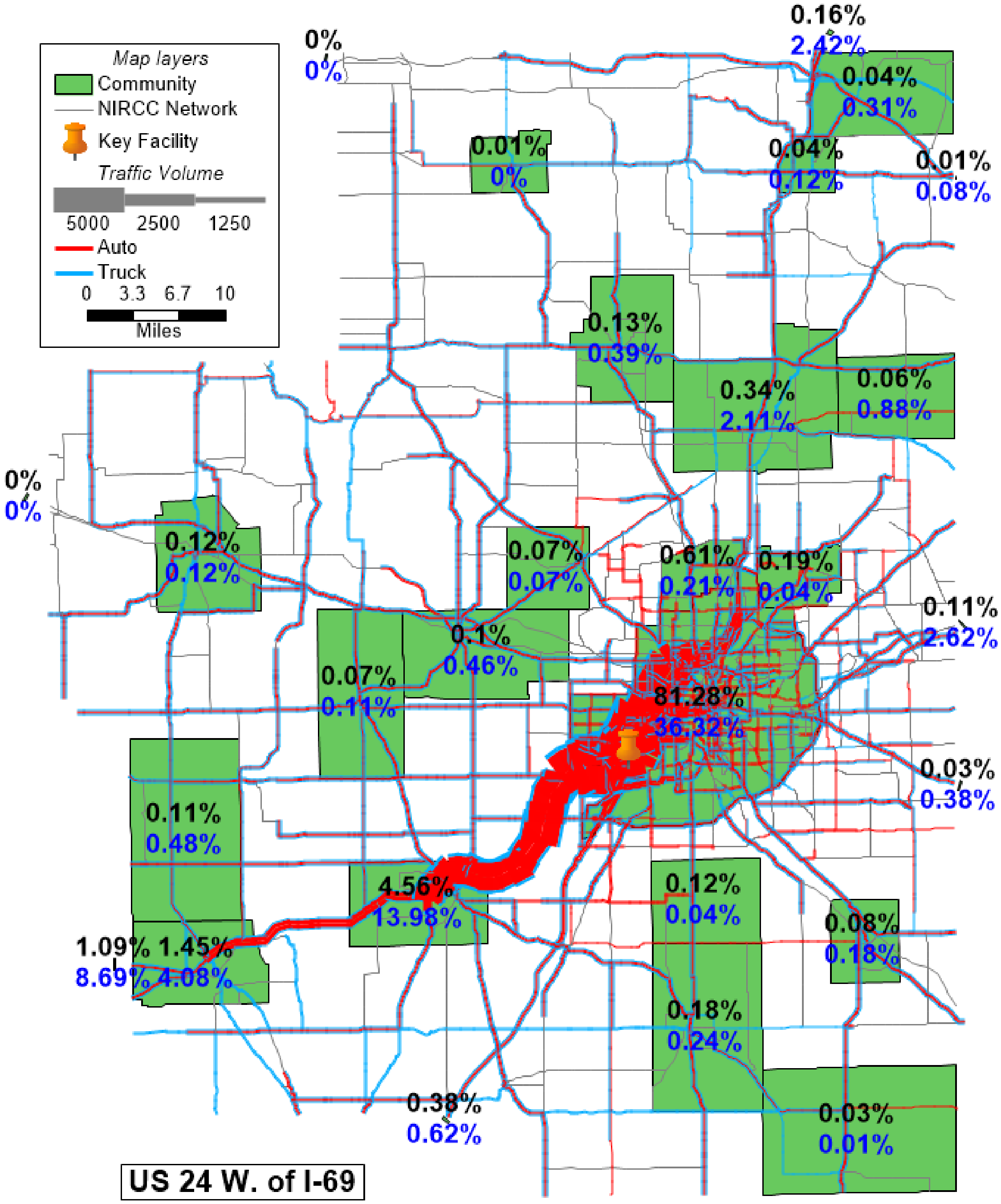
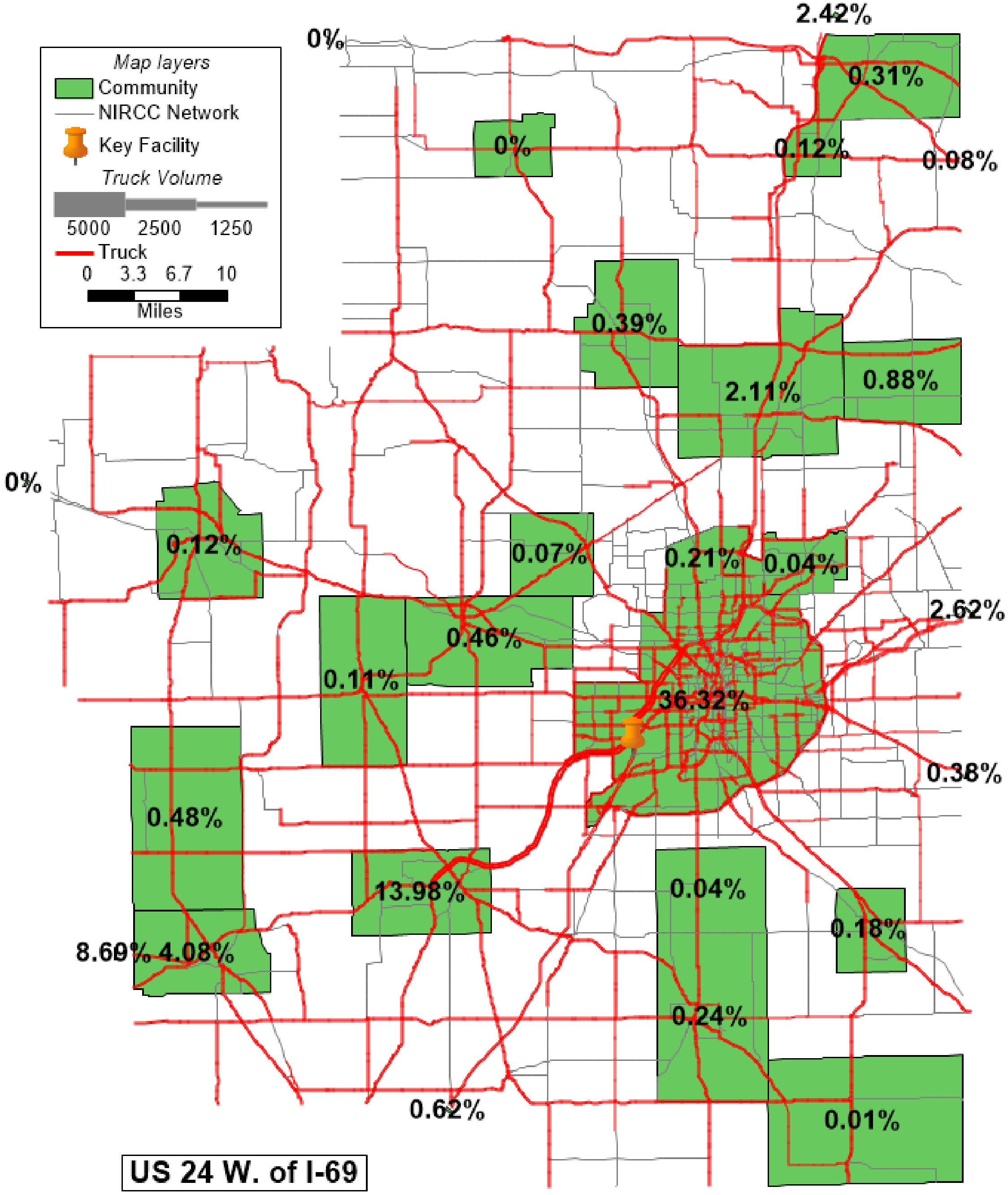


FIGURE 106. DEMAND AT US 24 W. OF I-69



Map layers
Community
NIRCC Network
Key Facility
Truck Volume
5000 2500 1250
Truck
0 3.3 6.7 10
Miles
0%
2.42%
0.31%
0%
0.12%
0.08%
0.39%
2.11%
0.88%
0%
0.12%
0.07%
0.21%
0.04%
2.62%
0.46%
0.11%
36.32%
0.38%
0.48%
13.98%
0.04%
0.18%
8.69%
4.08%
0.24%
0.62%
0.01%
US 24 W. of I-69

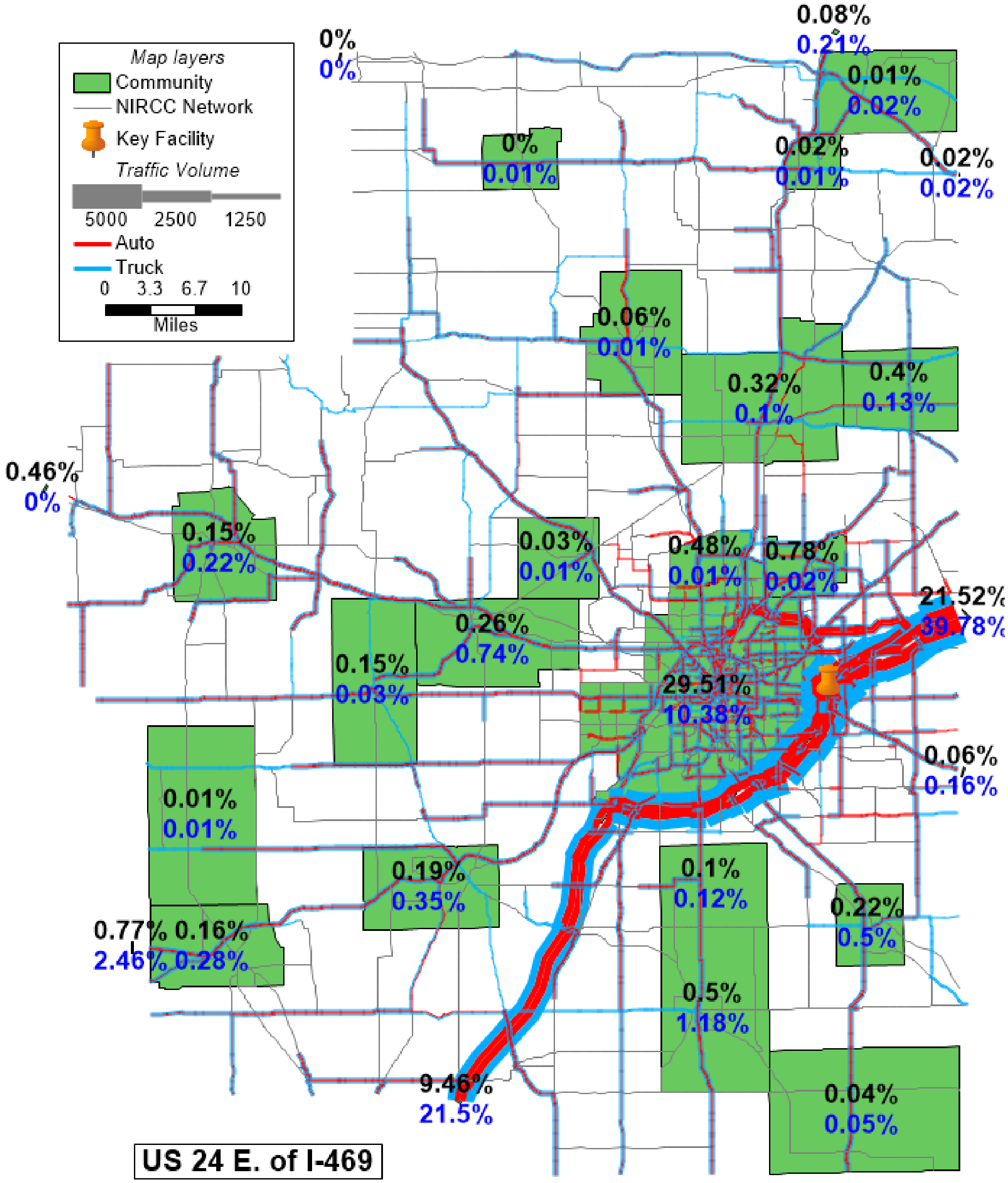
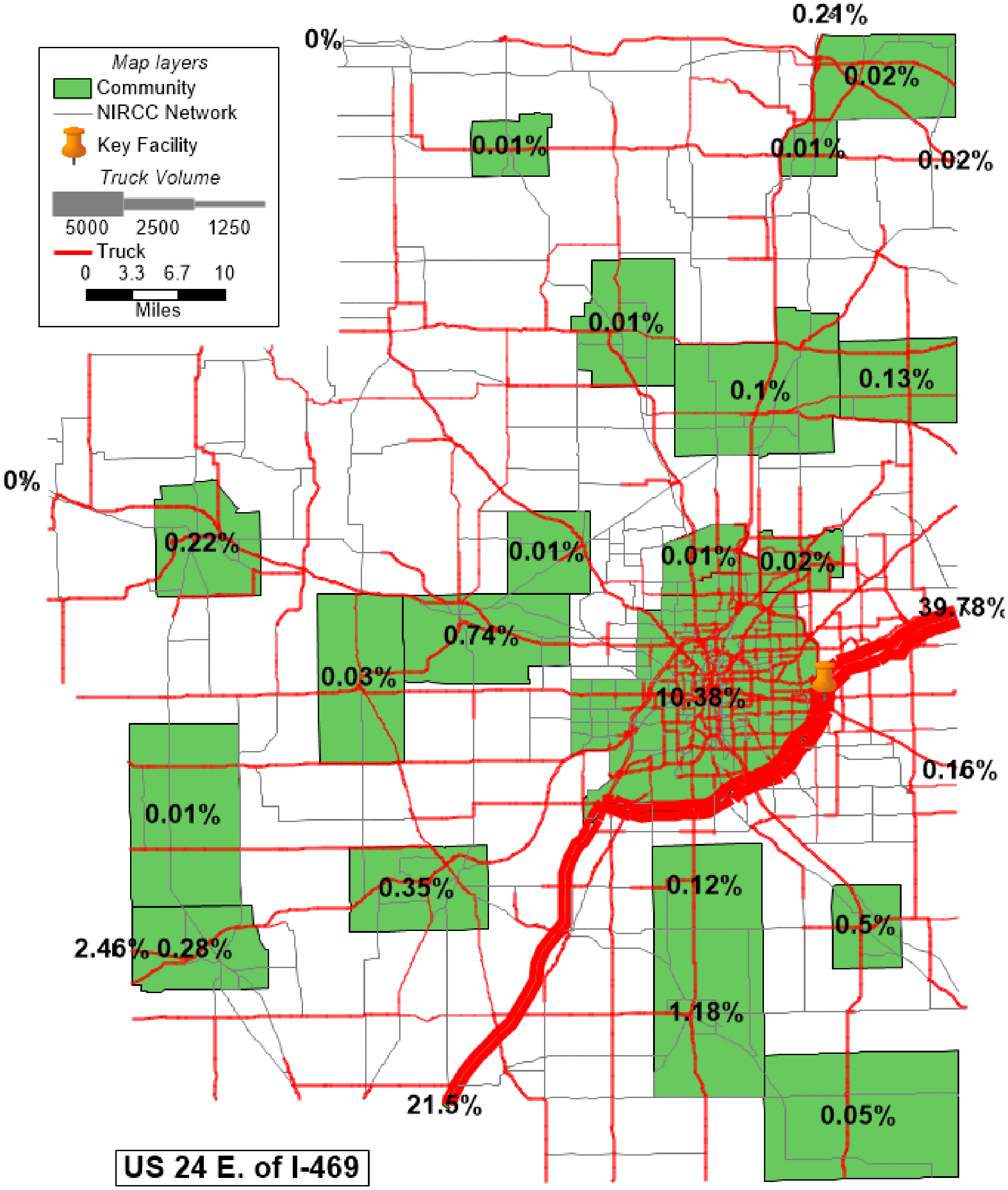


FIGURE 107. DEMAND AT US 24 E. OF I-469



US 24 E. of I-469

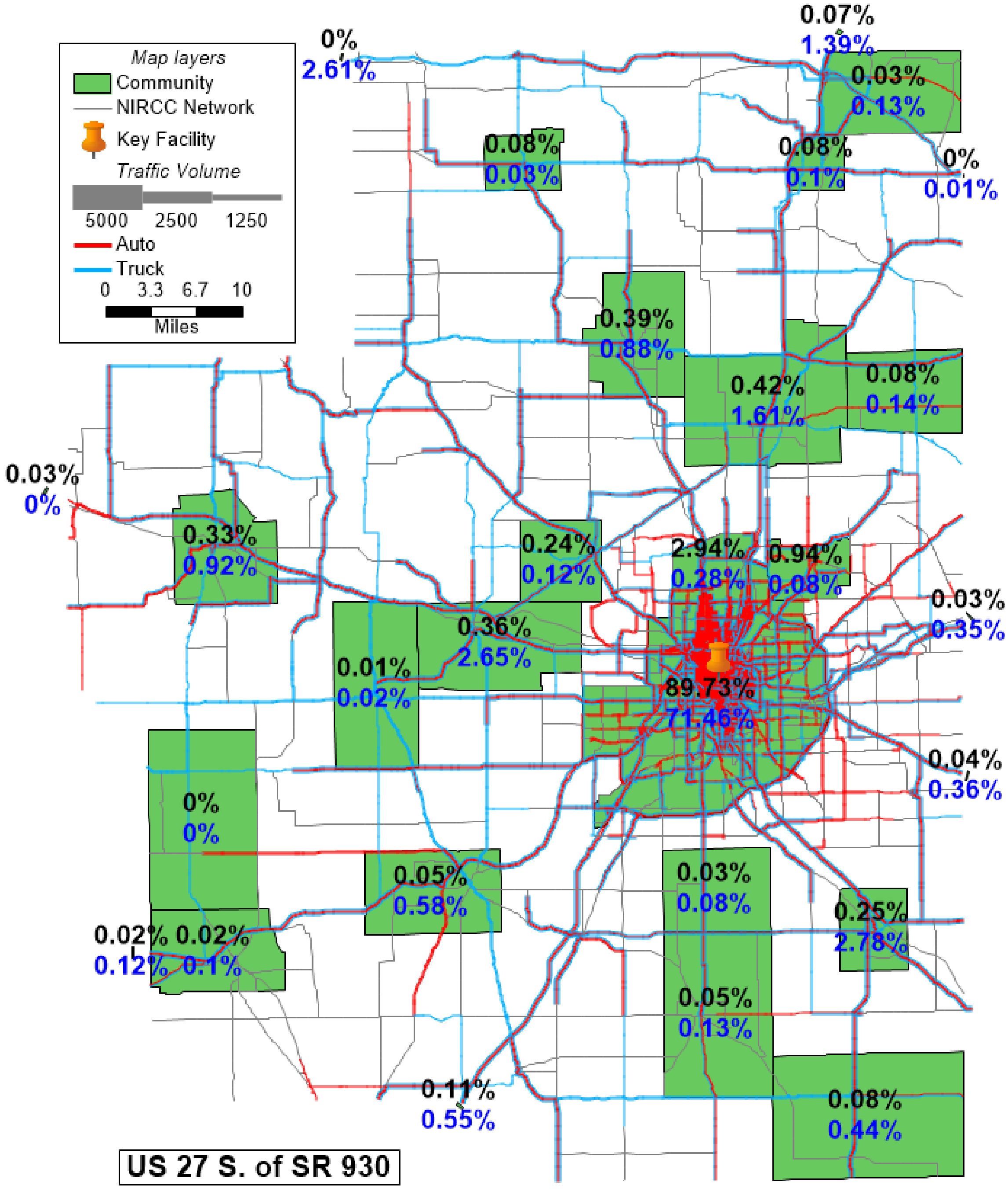
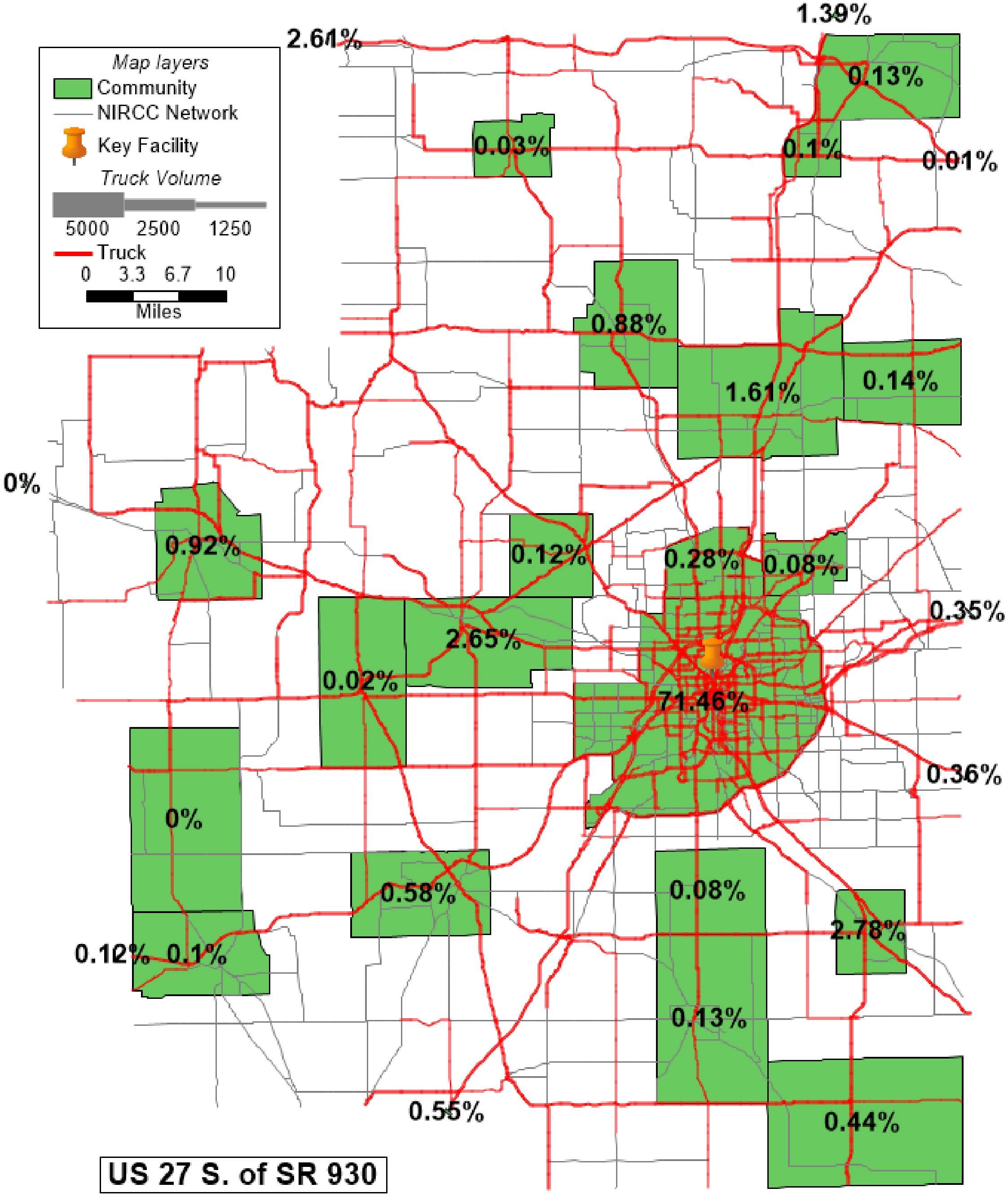


FIGURE 108. DEMAND AT US 27 S. OF SR 930


Map layers
Community
NIRCC Network
Key Facility
Truck Volume
5000
2500
1250
Truck
0
3.3
6.7
10
Miles
2.61%
1.39%
0.13%
0.03%
0.1%
0.01%
0.88%
1.61%
0.14%
0%
0.92%
0.12%
0.28%
0.08%
0.35%
2.65%
0.02%
71.46%
0.36%
0%
0.58%
0.08%
2.78%
0.12%
0.1%
0.13%
0.55%
0.44%
US 27 S. of SR 930

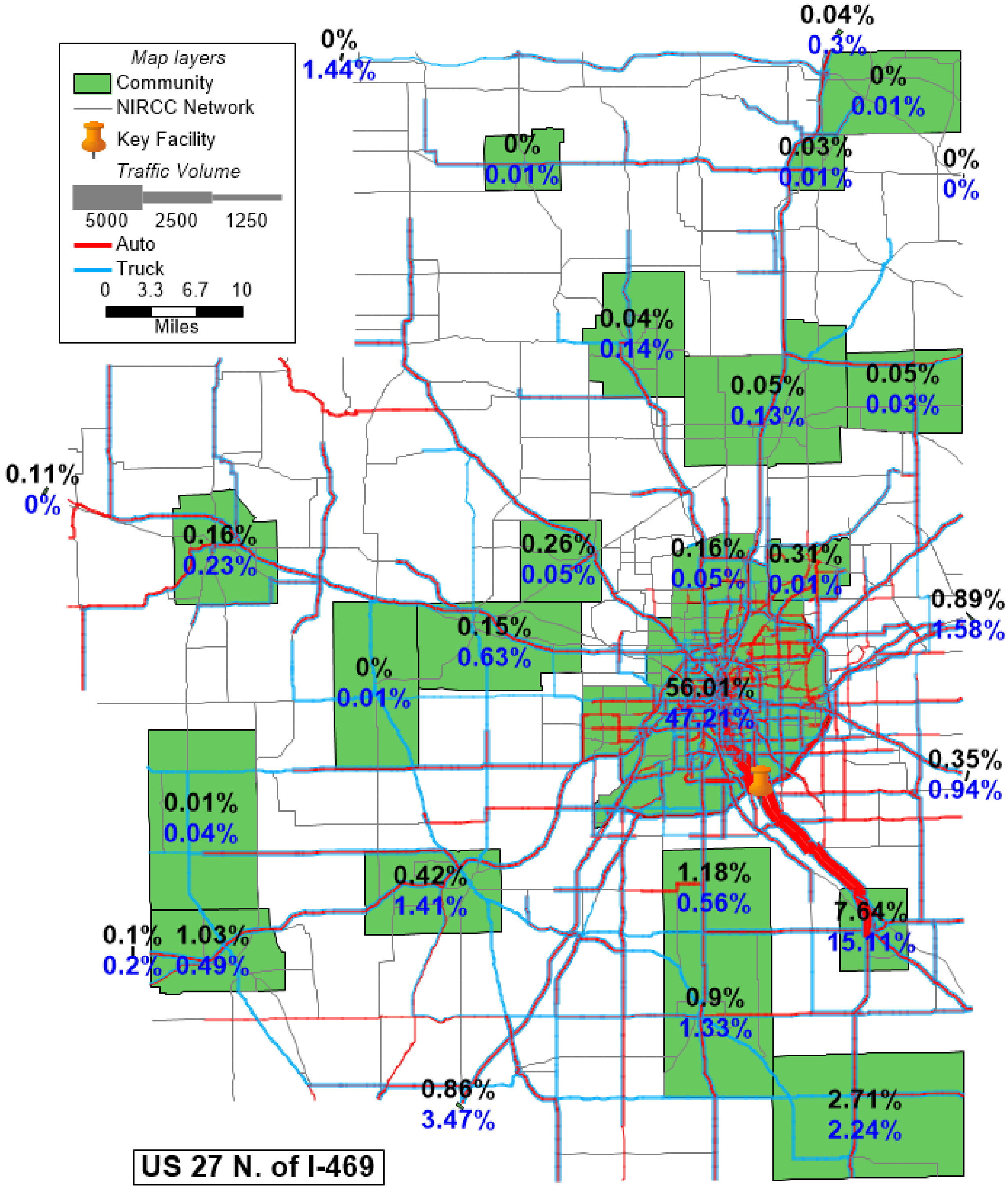
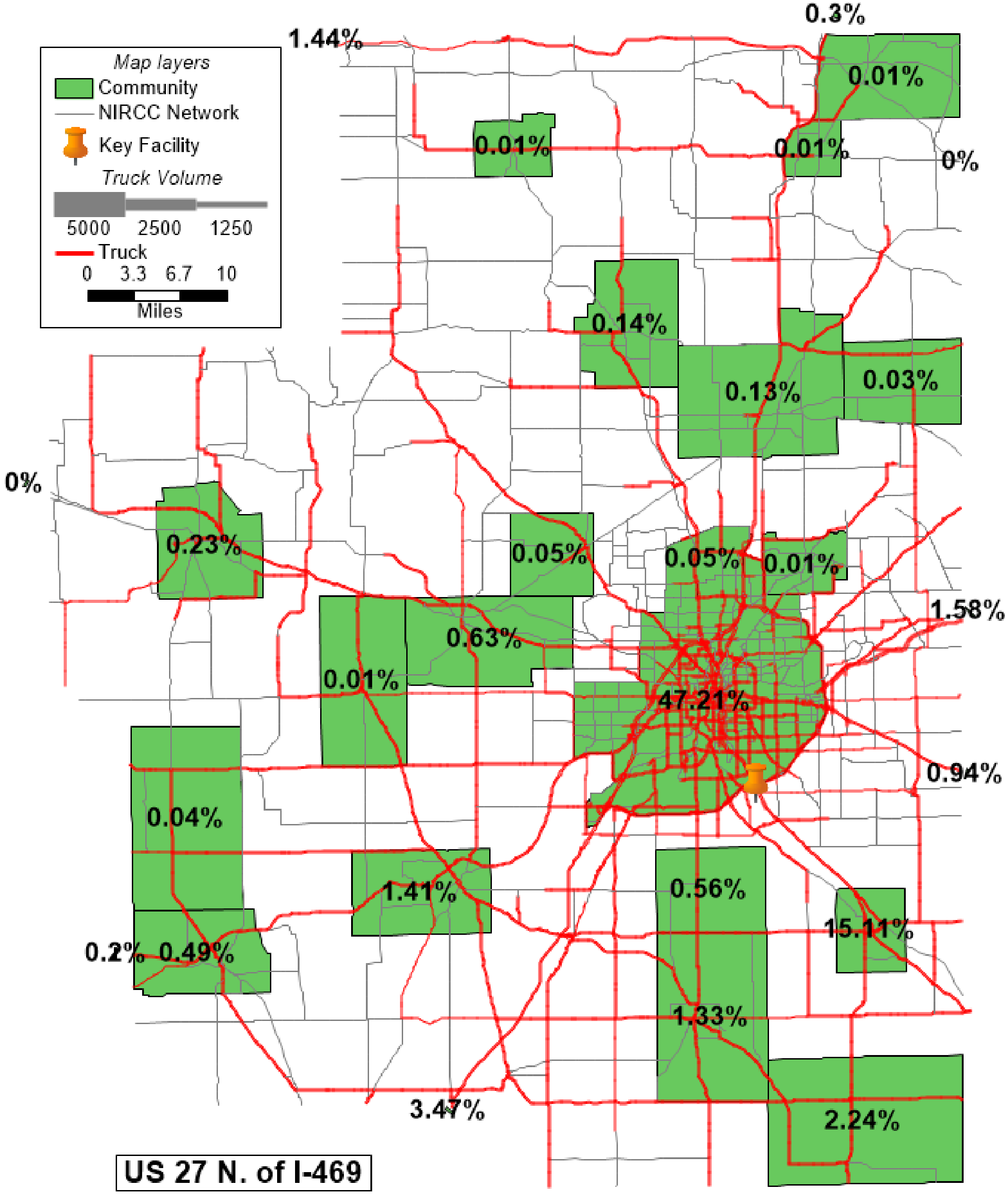


FIGURE 109. DEMAND AT US 27 N. OF I-469



Map layers
Community
NIRCC Network
Key Facility
Truck Volume
5000 2500 1250
Truck
0 3.3 6.7 10
Miles
1.44%
0.3%
0.01%
0.01%
0.01%
0%
0.14%
0.13%
0.03%
0%
0.23%
0.05%
0.05%
0.01%
1.58%
0.63%
0.01%
47.21%
0.94%
0.04%
1.41%
0.56%
15.11%
0.2%
0.49%
1.33%
3.47%
2.24%
US 27 N. of I-469

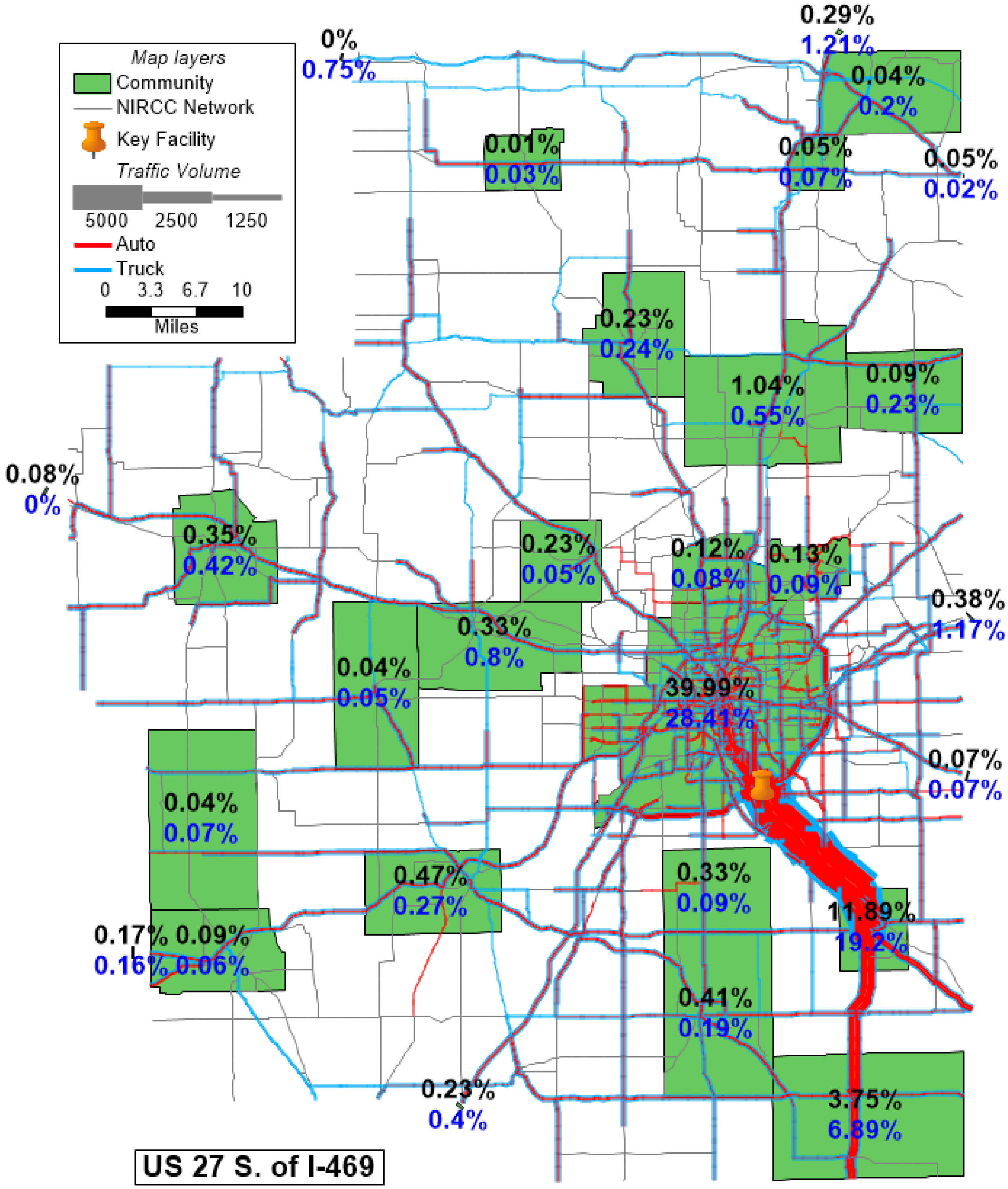
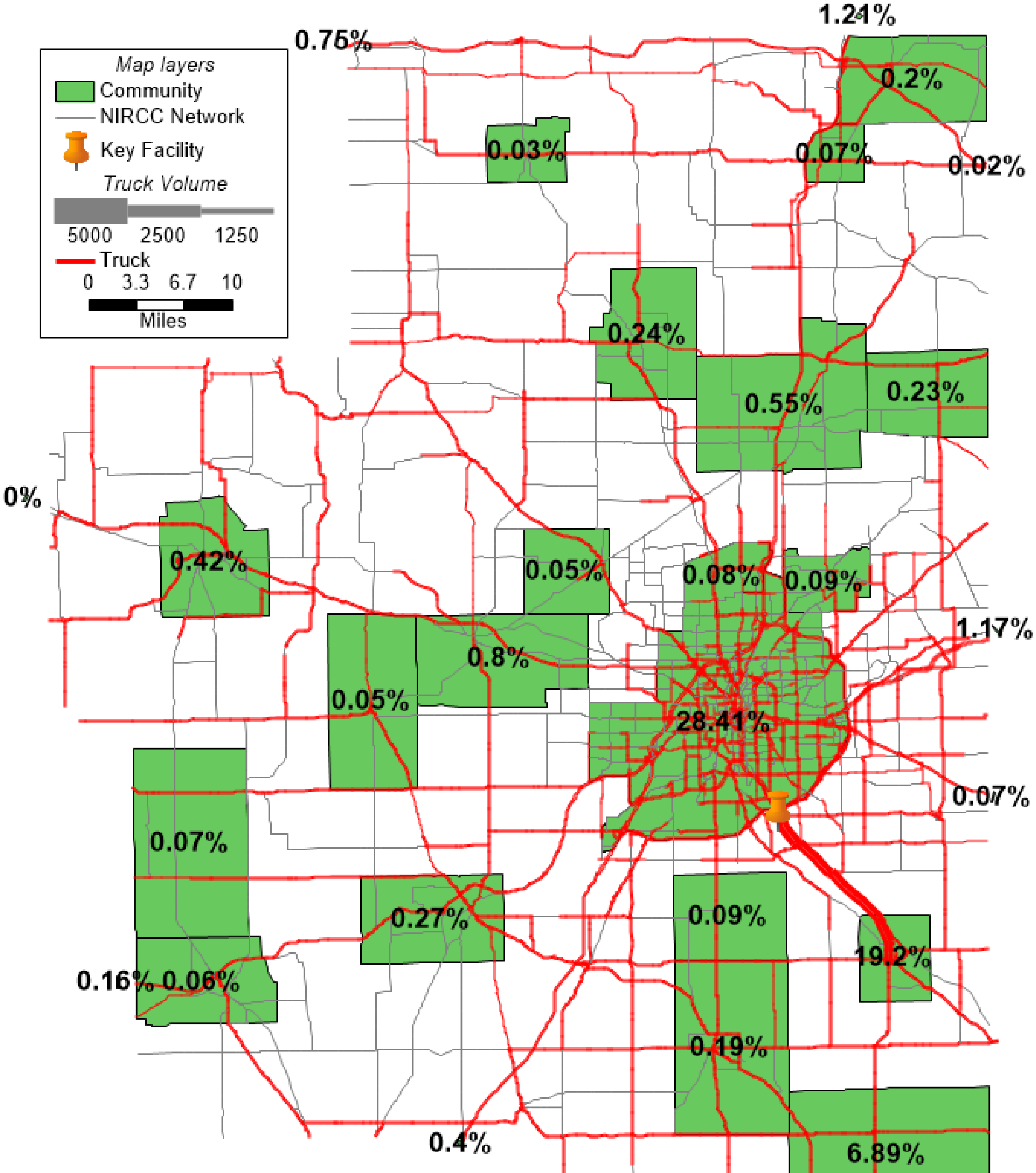


FIGURE 110. DEMAND AT US 27 S. OF I-469



US 27 S. of I-469

APPENDIX C. DATA VALIDATION DETAILS

This appendix provides further details on and results of the data validation summarized in Section 3.0. As in that section, results are presented from two validation efforts, the screening for poor coverage and scaling and test comparison with traffic count data.

SCREENING FOR POOR COVERAGE

Below are presented additional details of the screening for poor coverage.

Figure 111 and Figure 112 present the selected zones with the new threshold for auto, commercial (medium and heavy trucks together), medium, and heavy truck trips for the INRIX data. As it can be seen, there are couple zones in Allen County with lower coverage than others (mostly for heavy truck trips). Figure 113 and Figure 114Figure 114**Error! Reference source not found.** illustrate the zone coverage inside Allen County.

FIGURE 111: INRIX AUTO AND COMMERCIAL TRIP COVERAGE

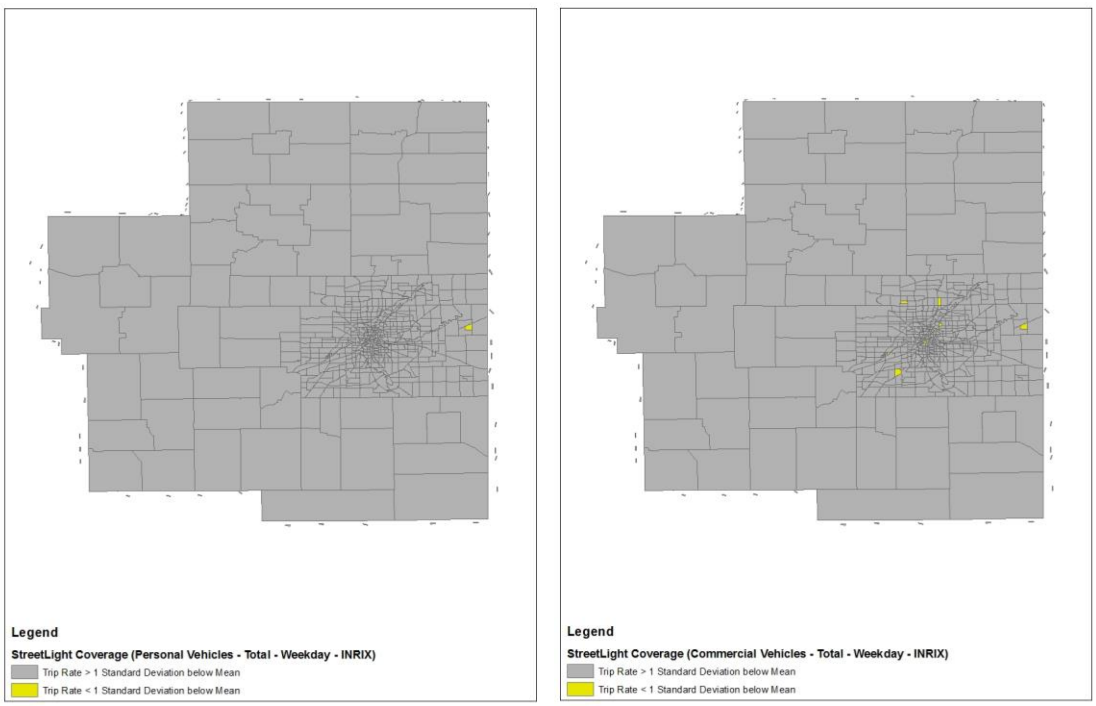


FIGURE 112: INRIX MEDIUM AND HEAVY TRUCK TRIP COVERAGE

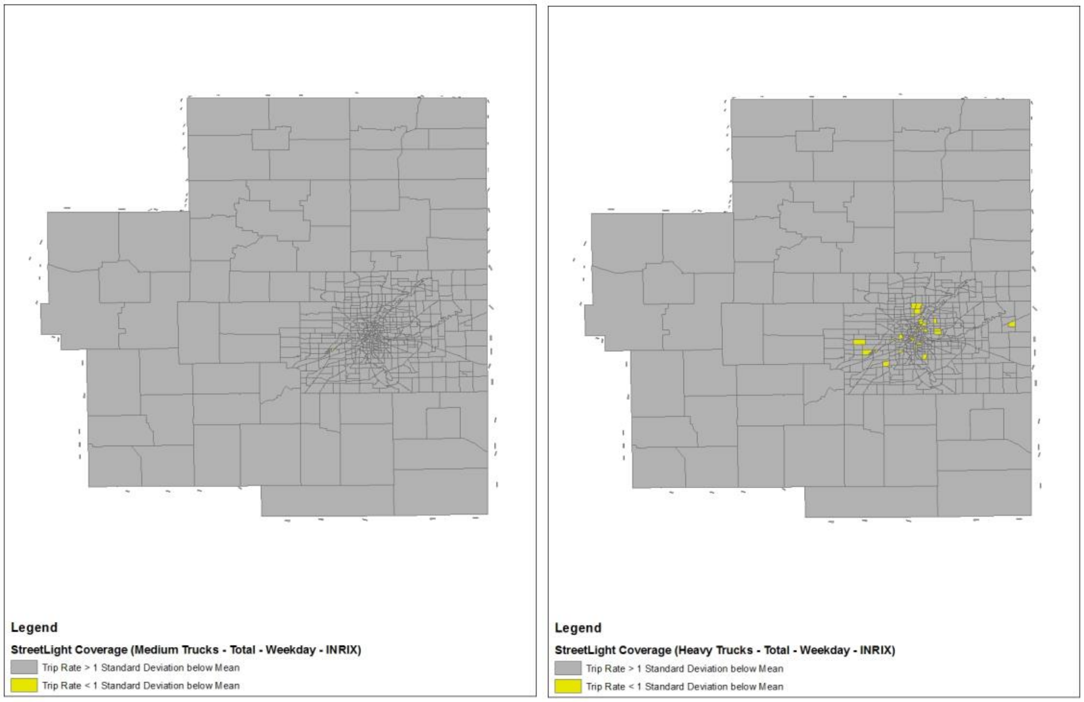


FIGURE 113: INRIX AUTO AND PERSONAL TRIP COVERAGE IN ALLEN COUNTY

FIGURE 114: INRIX MEDIUM AND HEAVY TRUCK TRIP COVERAGE IN ALLEN COUNTY

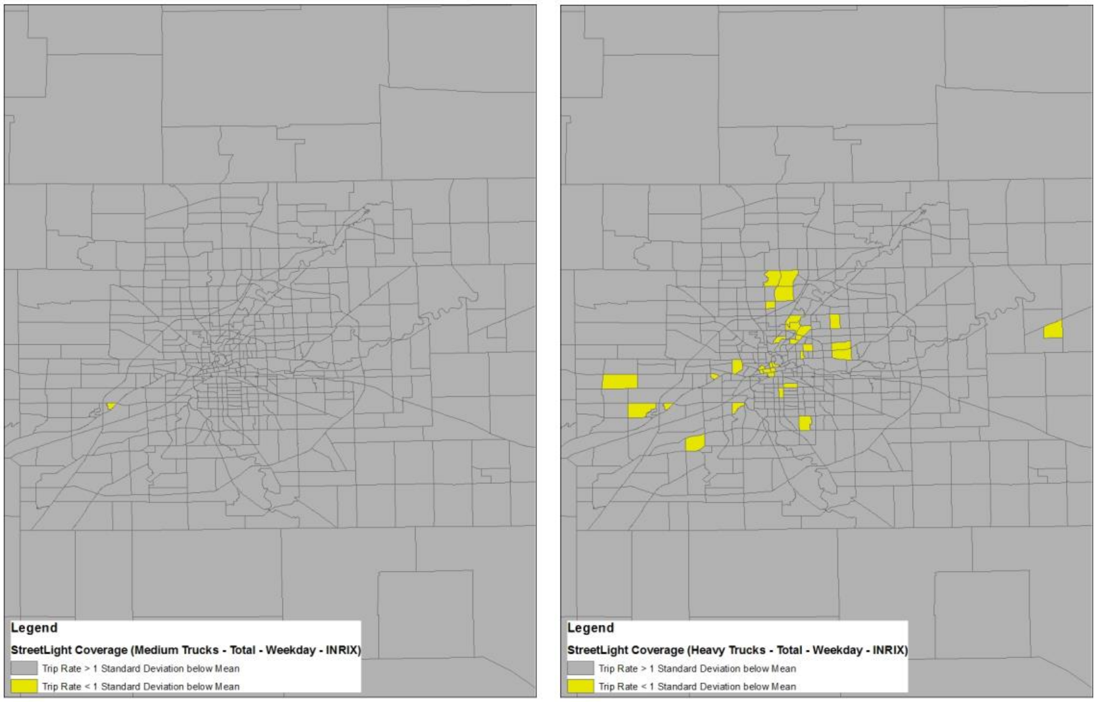
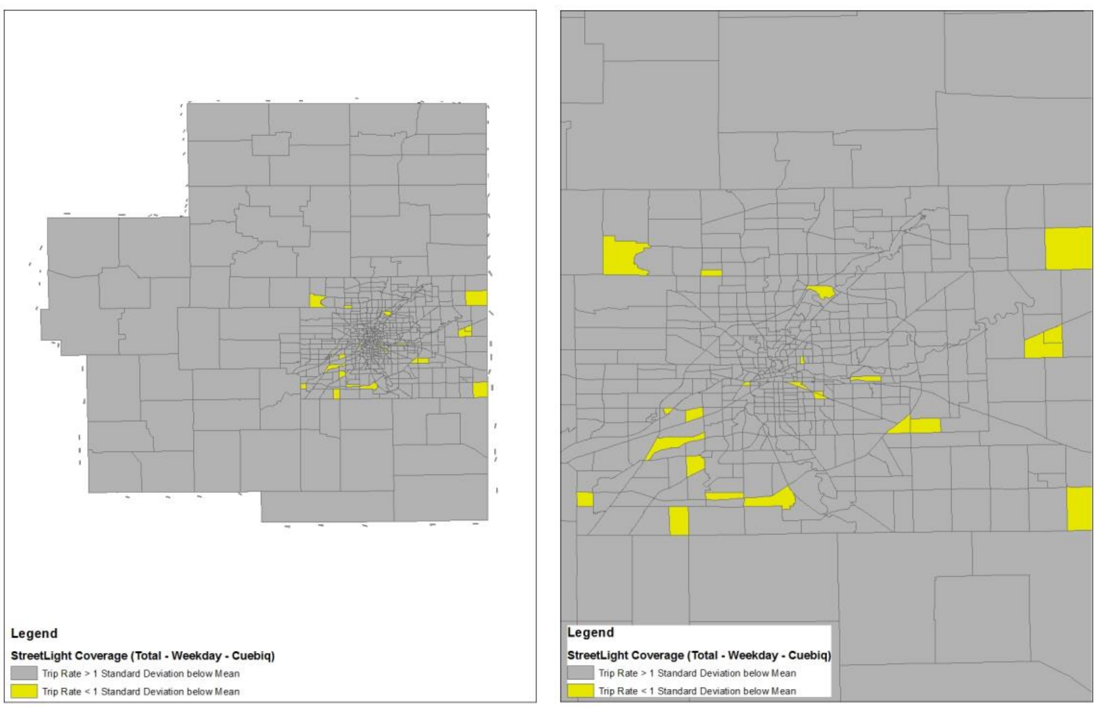


FIGURE 115: CUEBIQ TRIP COVERAGE IN THE STUDY AREA AND ALLEN COUNTY



According to Figure 111 to Figure 114, for INRIX and Figure 115 for Cuebiq, the data provides reasonable coverage in the study area. Although there are couple zones with lower coverage, they are still showing enough coverage that they can be improved in the data expansion.

Additional results of the further comparison of INRIX and Cuebiq trip rates with quick response (QR) rates are compared below.

FIGURE 116: TRIP RATES BY INRIX AND QR METHOD

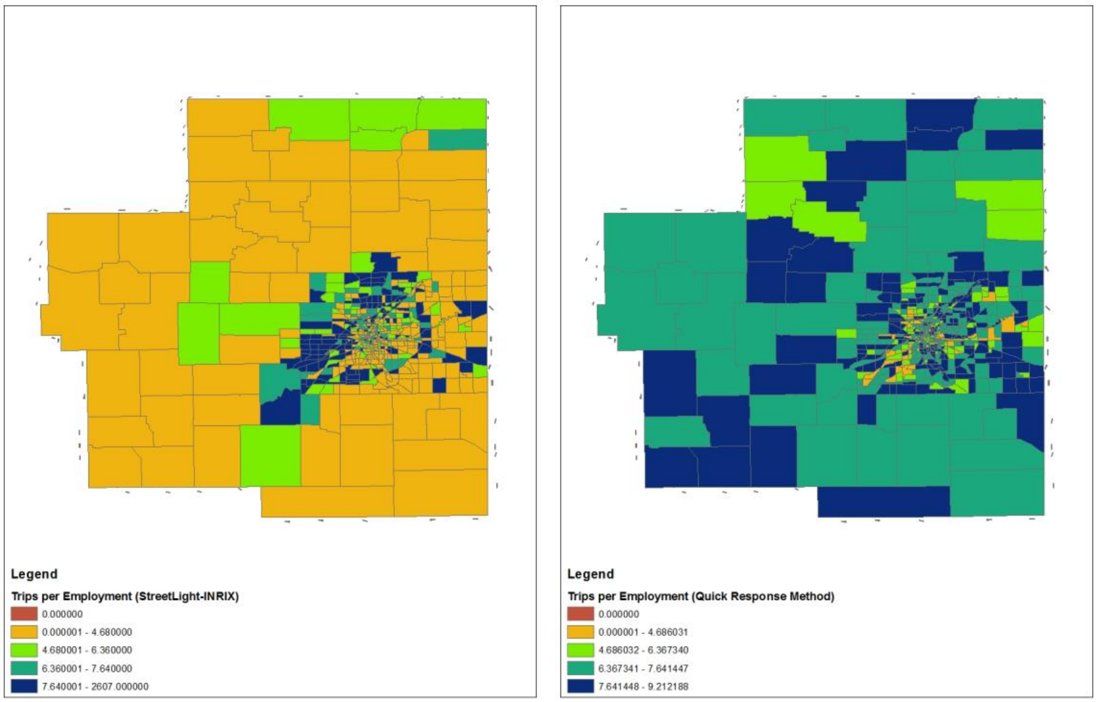


FIGURE 117: TRIP RATES BY INRIX AND QR METHOD IN ALLEN COUNTY

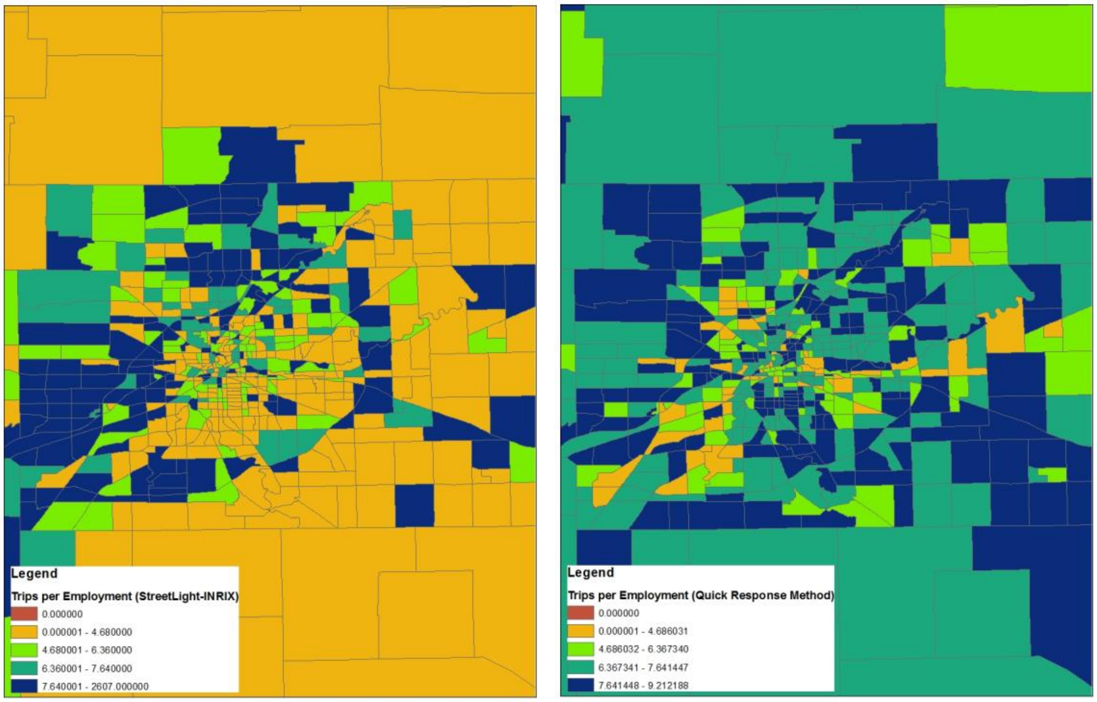
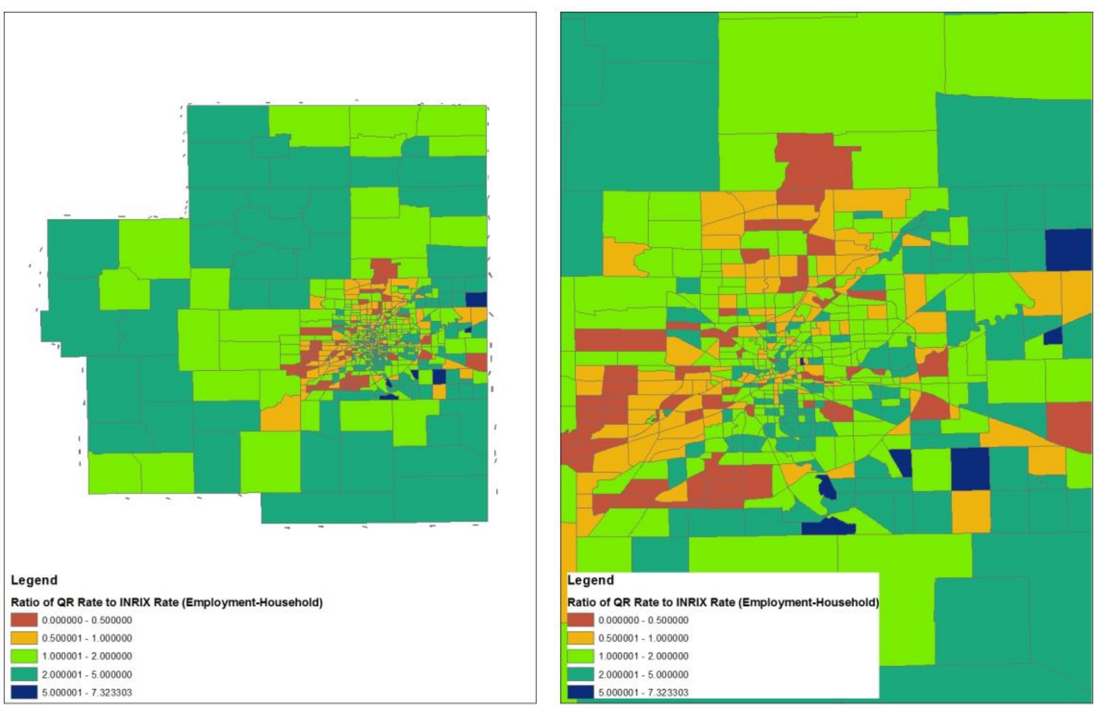
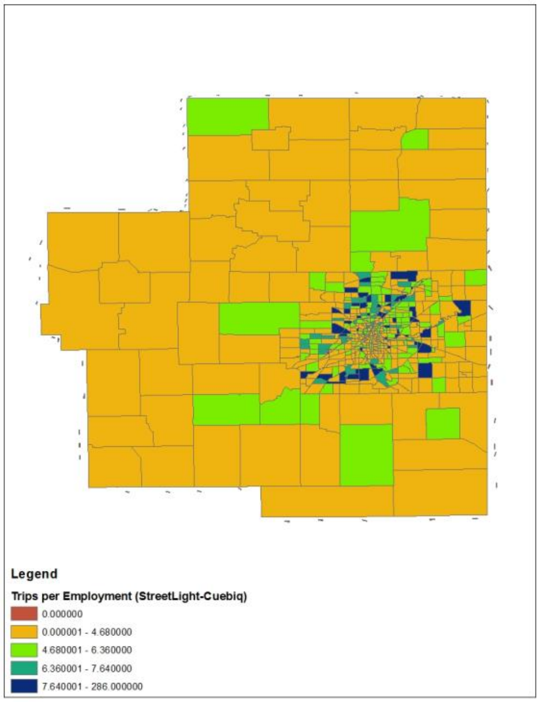


FIGURE 118: RATIO OF QR TO INRIX TRIP RATE IN THE STUDY AREA AND ALLEN COUNTY



INRIX trip rates are roughly comparable with trip rates estimated by QR method although there are some differences. Most importantly, there are no "holes" in terms of trip coverage in the data provided by INRIX. The trip rates implied by the INRIX data are a little bit low in rural areas relative to the QR rates. This could be due to poorer coverage in rural areas but is equally likely due to the fact that the QR rates were developed for urban areas and rural areas are known to have lower trip rates. If there is deficient coverage in the rural areas, it is clearly of a magnitude that can be addressed in the data expansion.

TRIP RATES BY CUEBIQ AND QR METHOD



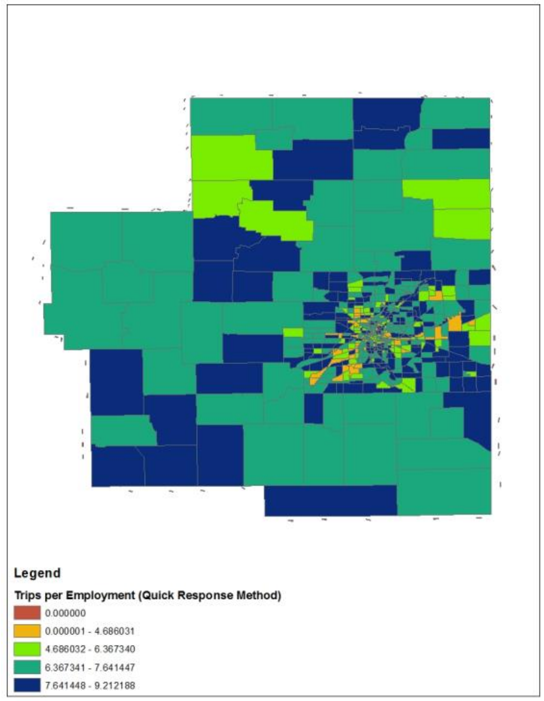


FIGURE 119: TRIP RATES BY CUEBIQ AND QR METHOD IN ALLEN COUNTY

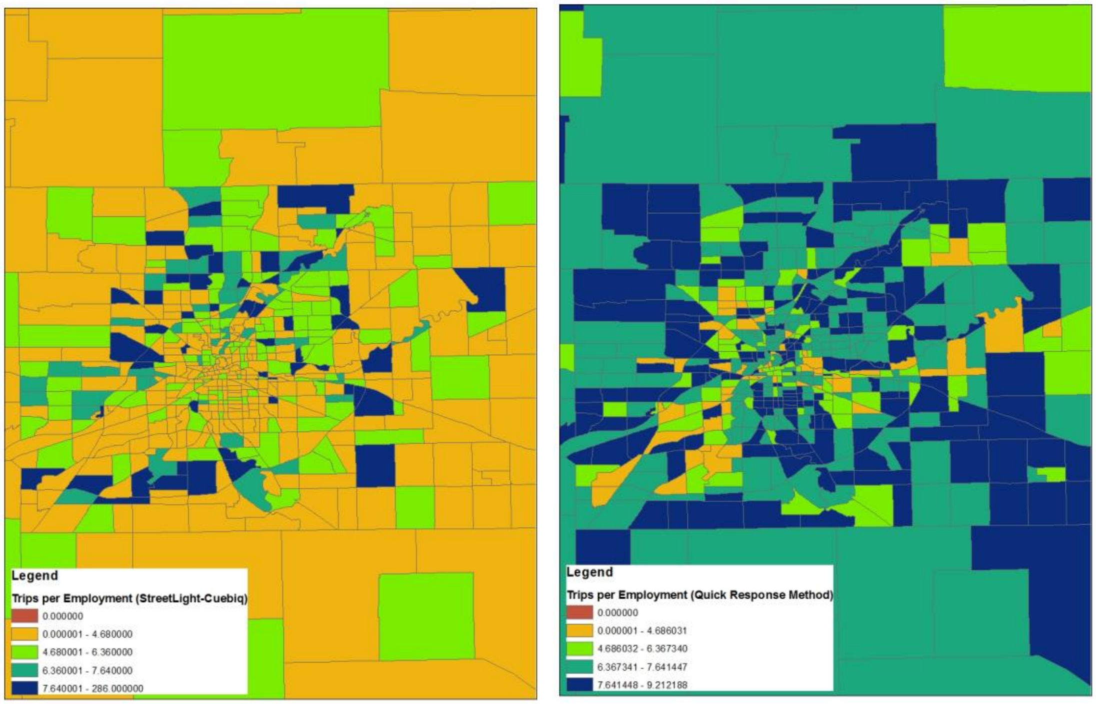
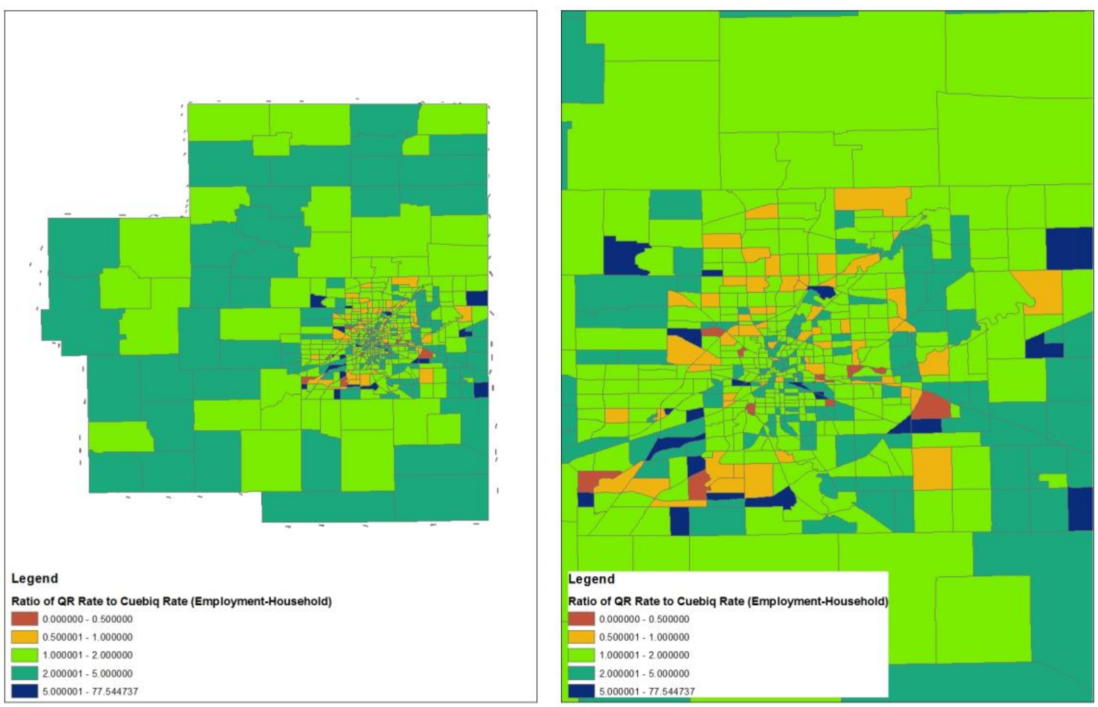
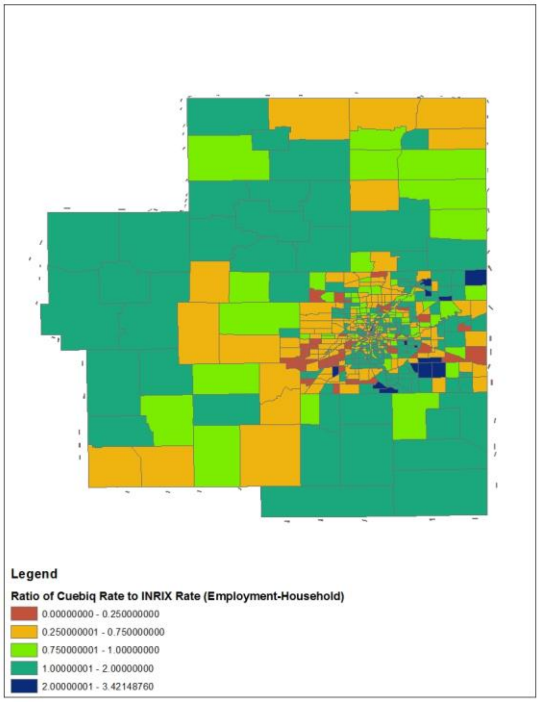


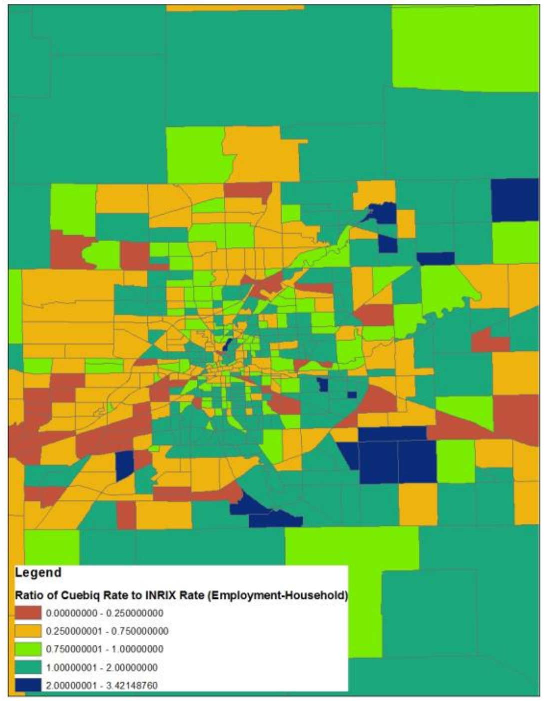
FIGURE 120: RATIO OF QR TO CUEBIQ TRIP RATE IN THE STUDY AREA AND ALLEN COUNTY



Finally, for CUEBIQ and INRIX trip rate comparison, the ratio of them was calculated which is reported in Figure 121. As seen in the figure, there is no consistent pattern over all zones. In fact, some zones have higher CUEBIQ trip rate, while others show higher INRIX trip rate. Overall, CUEBIQ trips are as reasonable as INRIX trips in terms of coverage and the minor issues can be fixed in the data expansion process.

FIGURE 121: RATIO OF CUEBIQ TO INRIX TRIP RATE IN THE STUDY AREA AND ALLEN COUNTY





SCALING AND TEST COMPARISON WITH TRAFFIC COUNT DATA

Redundant interzonal E-E trips which are not passing through the study area were removed from the trip tables. The new trip tables were then assigned to the network. The INRIX trip tables (auto, medium-size truck, and heavy-size truck) were separately assigned. The model volumes by class, therefore, can be compared to the corresponding counts (auto, SUT, and MUT) to validate the trip tables. In contrast, the CUEBIQ trip table includes all trips together. Thus, the assignment volumes are compared with the total ADT. Table 21 to Table 24 present INRIX and CUEBIQ assignment statistics by vehicle class. According to these tables, truck trips in INRIX are still very high. The overall loading error for INRIX auto trips is very low but the overall RMSE is high. CUEBIQ assignment results also show high overall loading error and RMSE for all facilities. The statistics by facility type or area type are high too, which indicate the trip tables must be properly scaled and expanded to provide more reasonable traffic volumes.

TABLE 21: ASSIGNMENT STATISTICS FOR INRIX AUTO TRIPS AFTER REMOVING E-E TRIPS

Assignment Statistic	Average Error	Loading Error (%)	RMSE (%)	MAPE
All Facilities	-56.68	-0.84	88.63	268.94
Freeways	-8,415.99	-43.97	61.84	114.03
Arterials	-2,116.96	-20.60	64.39	129.64
Collectors	2,488.98	85.77	166.60	412.99
Locals	2,454.77	194.53	300.48	948.62
Urban Links	-1,351.96	-13.97	70.04	93.67
Rural Links	1,503.80	40.99	134.66	477.28

TABLE 22: ASSIGNMENT STATISTICS FOR INRIX MEDIUM TRUCKS AFTER REMOVING E-E TRIPS

Assignment Statistic	Average Error	Loading Error (%)	RMSE (%)	MAPE
All Facilities	9,132.71	5,461.62	6,456.95	18,068.88
Freeways	15,079.00	1,461.85	1,592.55	3463.94
Arterials	10,060.56	4,542.53	5,223.68	11463.39
Collectors	8,034.63	11,135.10	13,282.00	23430.84
Locals	8,140.33	27,005.83	34,456.45	23872.19
Urban Links	9,400.79	5,110.54	5,981.06	11,470.61
Rural Links	8,794.63	5,852.66	6,958.09	25,085.41

TABLE 23: ASSIGNMENT STATISTICS FOR INRIX HEAVY TRUCKS AFTER REMOVING E-E TRIPS

Assignment Statistic	Average Error	Loading Error (%)	RMSE (%)	MAPE
All Facilities	30,282.71	9,021.01	11,694.94	157,332.54
Freeways	67,995.06	1,973.93	2,184.11	24924.95
Arterials	29,277.53	7,334.69	9,106.83	92906.49
Collectors	27,927.50	37,274.37	48,408.47	213594.44
Locals	29,560.55	272,268.26	395,457.22	354955.68
Urban Links	25,379.96	9,516.38	12,016.39	150,418.90
Rural Links	35,199.15	8,677.20	11,026.92	167,761.20

TABLE 24: ASSIGNMENT STATISTICS FOR CUEBIQ TRIPS AFTER REMOVING E-E TRIPS

Assignment Statistic	Average Error	Loading Error (%)	RMSE (%)	MAPE
All Facilities	1,545.51	22.92	86.04	215.92
Freeways	7,795.95	40.73	74.96	290.98
Arterials	1,454.80	14.16	60.06	197.94
Collectors	1,296.27	44.67	156.68	234.98
Locals	504.95	40.02	140.79	234.05
Urban Links	518.10	5.35	57.30	84.81
Rural Links	2,627.30	71.61	162.85	368.90

Figure 122 and Figure 123 illustrate the assignment loading error after removing redundant E-E trips from INRIX and CUEBIQ trips, respectively.

According to the statistics and the loading error maps, trips needs to be scaled properly as the loading error and RMSE are very high for all facility types. Part of this issue is due to external trips as it can be seen in Figure 122 and Figure 123; however, I-I trips have a major role in undesirable statistics.

FIGURE 122: LOADING ERROR OF INRIX TRIPS AFTER REMOVING REDUNDANT E-E TRIPS

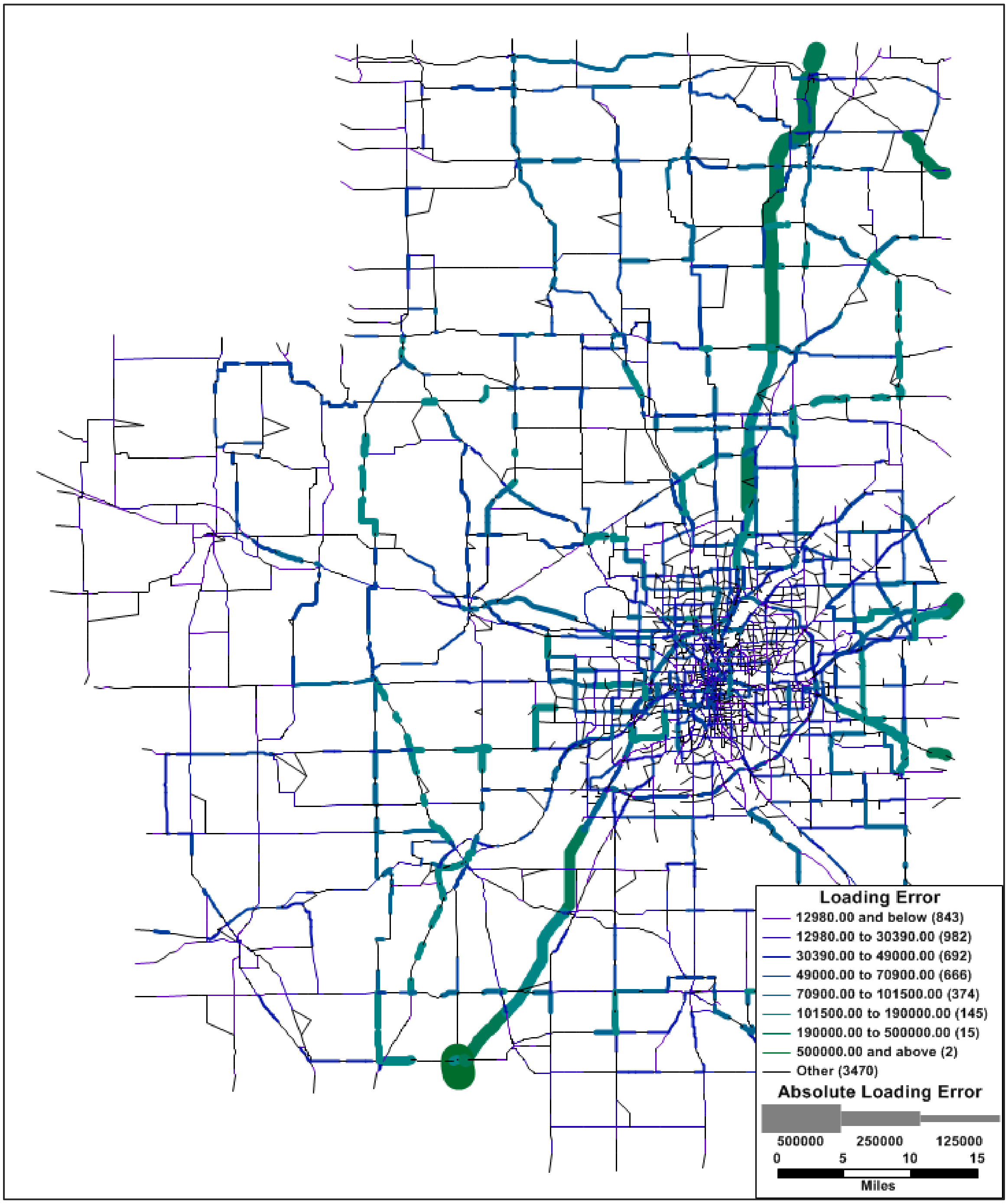
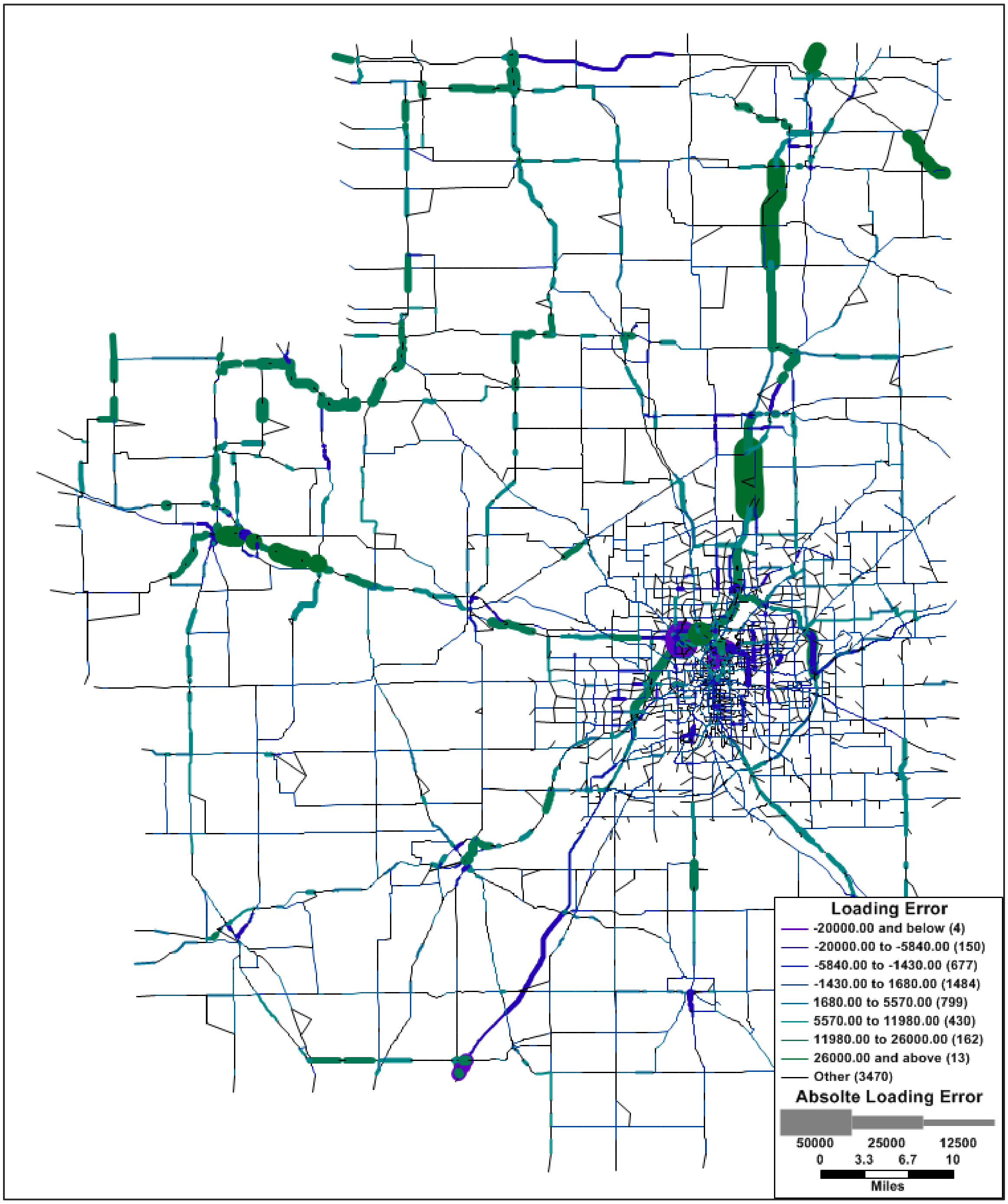


FIGURE 123: LOADING ERROR OF CUEBIQ TRIPS AFTER REMOVING REDUNDANT E-E TRIPS



As the first step to scale trips to become more aligned with the network counts, the total external trips including external-internal (E-I), internal-external (I-E), and E-E trips were scaled to match counts at the external stations. The network has 64 external stations or gates and 2015 counts are available for all of them. The external trips, therefore, were fratared to match external gates' counts. The process was separately conducted for CUEBIQ and INRIX trips as total AADT was used as ground truth for CUEBIQ trips but car AADT and truck AADT were used for INRIX passenger car and truck trip tables. Since the breakdown of SUT and MUT is not available at all external count stations, the INRIX medium and heavy truck trip tables were added up and the total truck trip table was scaled to match total truck AADTs. The resulted truck trips were then proportionally divided to SUT and MUT trips based on the medium and heavy truck trips before fratar. Table 25 to Table 28 present the assignment statistics after fratering external trips.

TABLE 25: ASSIGNMENT STATISTICS FOR INRIX AUTO TRIPS AFTER EXTERNAL FRATAR

Assignment Statistic	Average Error	Loading Error (%)	RMSE (%)	MAPE
All Facilities	-2,736.32	-40.59	85.94	124.21
Freeways	-8,624.02	-45.06	60.14	118.36
Arterials	-4,989.67	-48.56	72.11	106.55
Collectors	-312.53	-10.77	93.59	136.49
Locals	517.52	41.01	154.89	455.71
Urban Links	-4,668.59	-48.24	74.59	64.77
Rural Links	-646.08	-17.61	95.42	193.54

TABLE 26: ASSIGNMENT STATISTICS FOR INRIX MEDIUM TRUCKS AFTER EXTERNAL FRATAR

Assignment Statistic	Average Error	Loading Error (%)	RMSE (%)	MAPE
All Facilities	3,092.32	1,849.29	2,367.15	4,754.47
Freeways	5,492.91	532.52	690.43	1,304.50
Arterials	3,793.53	1,712.85	2,024.73	3,198.09
Collectors	2,462.80	3,413.17	4,461.37	6,108.87
Locals	2,453.90	8,140.92	9,532.25	9,348.23
Urban Links	3,527.45	1,917.62	2,302.37	3,883.10
Rural Links	2,644.14	1,759.63	2,433.67	5,777.65

TABLE 27: ASSIGNMENT STATISTICS FOR INRIX HEAVY TRUCKS AFTER EXTERNAL FRATAR

Assignment Statistic	Average Error	Loading Error (%)	RMSE (%)	MAPE
All Facilities	2,866.11	853.79	1,281.78	10,385.83
Freeways	12,159.84	353.01	388.00	4,857.10
Arterials	2,748.70	688.61	913.78	6,830.11
Collectors	2,350.17	3,136.73	4,587.56	13,162.29
Locals	2,364.85	21,781.54	26,771.48	46,340.71
Urban Links	2,189.72	821.05	1,283.74	7,495.01
Rural Links	3,659.91	902.23	1,269.24	13,838.62

TABLE 28: ASSIGNMENT STATISTICS FOR CUEBIQ TRIPS AFTER EXTERNAL FRATAR

Assignment Statistic	Average Error	Loading Error (%)	RMSE (%)	MAPE
All Facilities	1,038.78	15.41	77.81	181.25
Freeways	6,453.33	33.72	64.76	266.96
Arterials	849.60	8.27	54.96	169.94
Collectors	935.65	32.24	142.13	192.81
Locals	580.12	45.97	144.54	238.49
Urban Links	247.20	2.55	55.75	80.20
Rural Links	1,873.77	51.07	136.06	301.05

New assignment statistics show significantly improvement compared to the statistics before running fratar especially for CUEBIQ trips and INRIX truck trips. INRIX auto trips overall loading error increased after fratar; however, the RMSE has improved. Since fratar affects E-I and I-E trips, it is entirely possible the new internal auto trips become farther from counts after fratar especially when the trip table before fratar has good agreement with counts. This issue will be addressed in the next step of adjustment. In general, the new trip tables completely satisfy the goal of this step which is matching external gates' counts. Figure 124 and Figure 125 presenting the loading errors after frataring externals also confirm that the new trip tables perform better than the ones before external frataring for external gate links. Although fratar somewhat improved the assignment statistics, the overall performance is not desirable yet. More scaling, therefore, is needed to generate more aligned volumes with the counts.

FIGURE 124: LOADING ERROR OF INRIX TRIPS AFTER EXTERNAL FRATAR

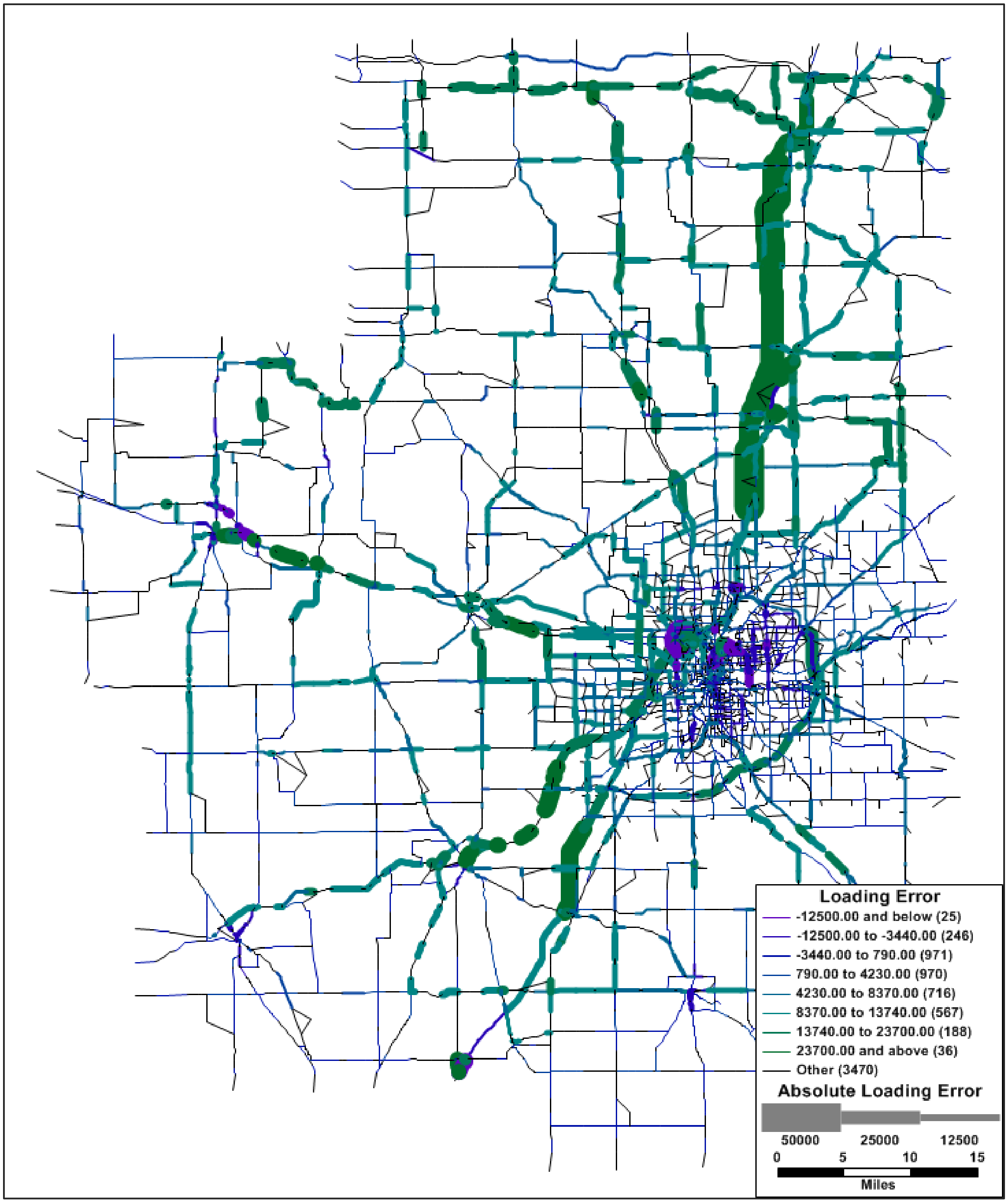
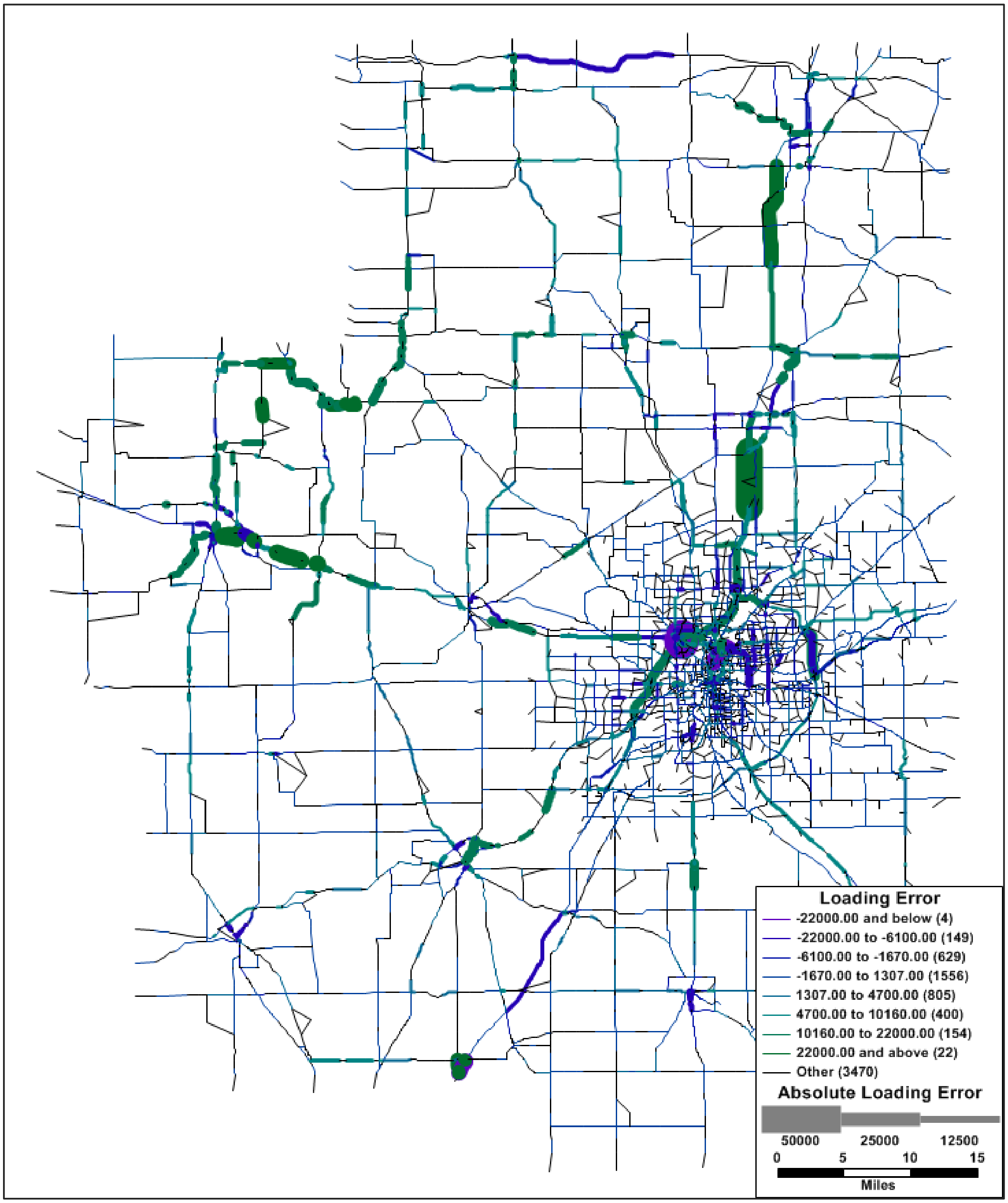


FIGURE 125: LOADING ERROR OF CUEBIQ TRIPS AFTER EXTERNAL FRATAR



The goal of the second step of trip adjustment is minimizing the overall loading error by vehicle class to have 0 percent overall loading error. This process was conducted through iterative scaling of the I-I trips because the external trips are completely matching the external gates' counts after fratating. Different scales were considered for INRIX auto, medium, and heavy truck trip tables. Table 29 summarizing the scaled trip tables indicates that total CUEBIQ trips decreased a little bit while INRIX truck trips needed to be drastically reduced. On the other hand, total INRIX auto trips needed to be significantly increased. The scaled trips, therefore, are much more reasonable that previous set of trip tables. The total CUEBIQ and INRIX scaled trips are in same order of magnitude, although a large share of the trips are intrazonal and especially when these trips are removed it is clear that there are significantly more trips in the CUEBIQ dataset than the INRIX dataset. Since both datasets are now scaled to represent roughly the same VMT, this indicates that the INRIX interzonal trips are significantly longer on average than the CUEBIQ interzonal trips.

TABLE 29: CUEBIQ AND INRIX TRIP TABLES AFTER OVERALL SCALING

Trip	Total Trips before Overall Scaling	Total Trips after Overall Scaling	Interzonal Trips
CUEBIQ	2,608,969	2,568,895	580,144
INRIX – Total	4,319,635	2,208,019	316,163
– Personal	1,066,780	2,079,812	304,455
– Medium Trucks	1,218,353	57,179	7,731
– Heavy Trucks	2,034,502	71,028	3,977

Table 30 to Table 33 report assignment statistics for different trip tables and according to them, the assignment has significantly improved especially truck trips. This scaling process fixed the issue of INRIX auto trips which did not have desirable loading error after external frataring. The RMSE of INRIX auto trips is now as low as CUEBIQ which shows the improvement applied by scaling. Although the new assignment results illustrate significant enhancement, new RMSEs are still high and above desired thresholds. In other words, the scaling process helped improve the assignment; however, expansion adjustment factors beyond simple scaling factors are needed.

TABLE 30: ASSIGNMENT STATISTICS FOR INRIX AUTO TRIPS AFTER OVERALL SCALING

Assignment Statistic	Average Error	Loading Error (%)	RMSE (%)	MAPE
All Facilities	1.83	0.03	72.21	139.95
Freeways	3,631.21	18.97	59.25	208.10
Arterials	-599.73	-5.84	52.75	132.66
Collectors	386.45	13.32	122.88	146.85
Locals	247.20	19.59	113.91	198.07
Urban Links	-636.80	-6.58	56.70	74.71
Rural Links	659.80	17.98	107.65	217.86

TABLE 31: ASSIGNMENT STATISTICS FOR INRIX MEDIUM TRUCKS AFTER OVERALL SCALING

Assignment Statistic	Average Error	Loading Error (%)	RMSE (%)	MAPE
All Facilities	-0.14	-0.08	406.63	128.09
Freeways	-43.45	-4.21	333.08	113.05
Arterials	-12.48	-5.64	94.05	103.11
Collectors	7.43	10.30	137.36	126.26
Locals	15.16	50.28	137.08	134.44
Urban Links	-18.88	-10.26	100.04	78.20
Rural Links	16.09	10.71	650.31	159.77

TABLE 32: ASSIGNMENT STATISTICS FOR INRIX HEAVY SIZE TRUCK TRIPS AFTER OVERALL SCALING

Assignment Statistic	Average Error	Loading Error (%)	RMSE (%)	MAPE
All Facilities	0.09	0.03	166.53	451.04
Freeways	283.25	8.22	62.94	1181.42
Arterials	-96.18	-24.09	129.53	528.10
Collectors	36.32	48.47	229.28	332.63
Locals	19.00	175.00	353.84	432.71
Urban Links	-47.40	-17.77	163.87	198.31
Rural Links	36.66	9.04	160.64	722.07

TABLE 33: ASSIGNMENT STATISTICS FOR CUEBIQ TRIPS AFTER OVERALL SCALING

Assignment Statistic	Average Error	Loading Error (%)	RMSE (%)	MAPE
All Facilities	6.20	0.09	71.94	161.13
Freeways	4,059.22	21.21	54.99	225.13
Arterials	-600.58	-5.85	51.43	155.37
Collectors	325.40	11.21	131.57	167.85
Locals	374.05	29.64	125.48	203.94
Urban Links	-1,192.39	-12.32	52.87	72.93
Rural Links	1,249.80	34.07	120.67	265.91

Figure 126 and Figure 127 illustrate the loading error maps of INRIX and CUEBIQ trips after overall scaling. Comparing these maps with Figure 124 and Figure 125 confirms that INRIX scaled trips have lower loading error than trips after external fratar. CUEBIQ maps does not show significant improvement like INRIX, which was expected since total CUEBIQ trips did not change a lot after scaling. The assignment statistics for CUEBIQ trips show improvement though.

FIGURE 126: LOADING ERROR OF INRIX TRIPS AFTER EXTERNAL FRATAR AND OVERALL SCALING

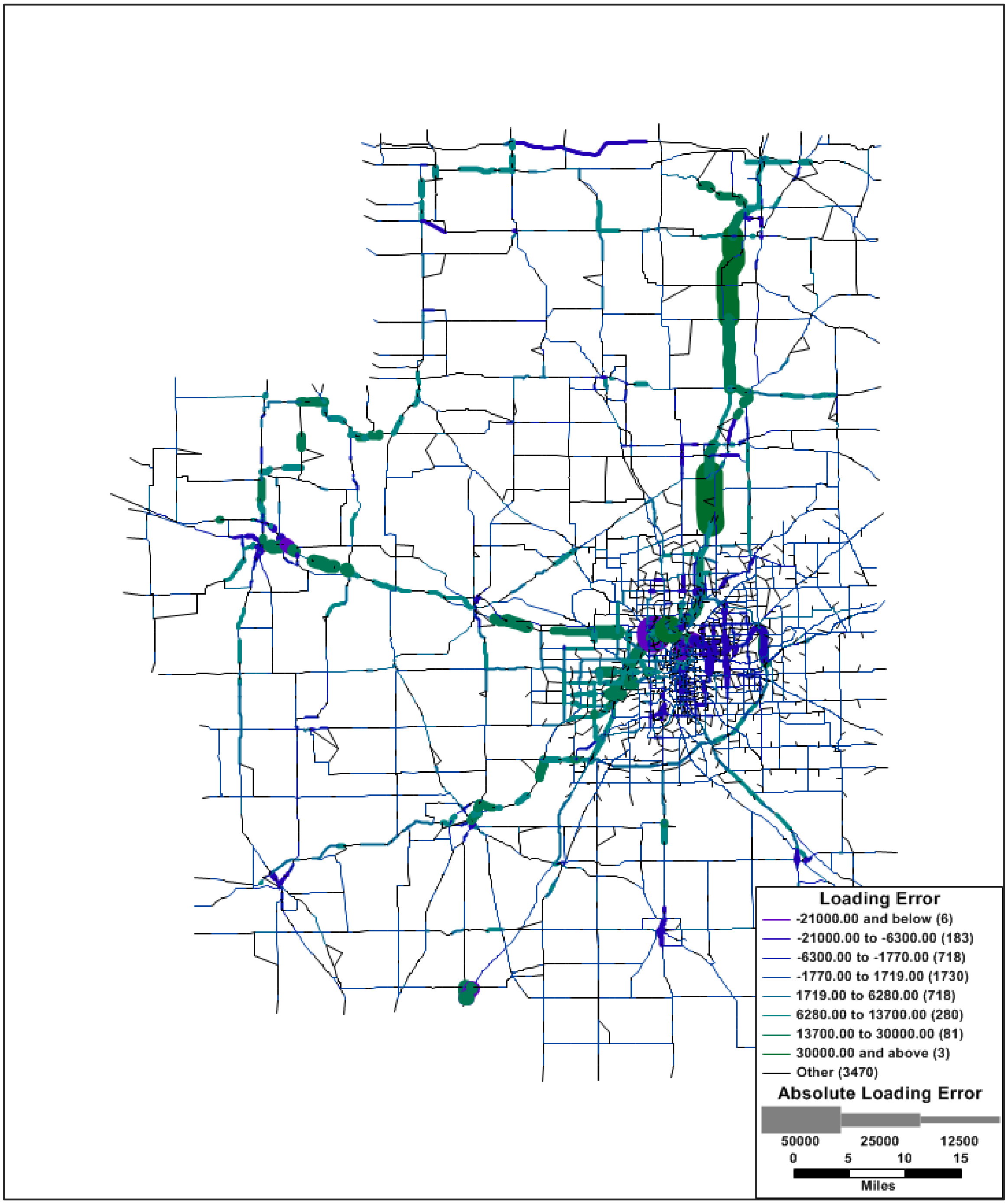
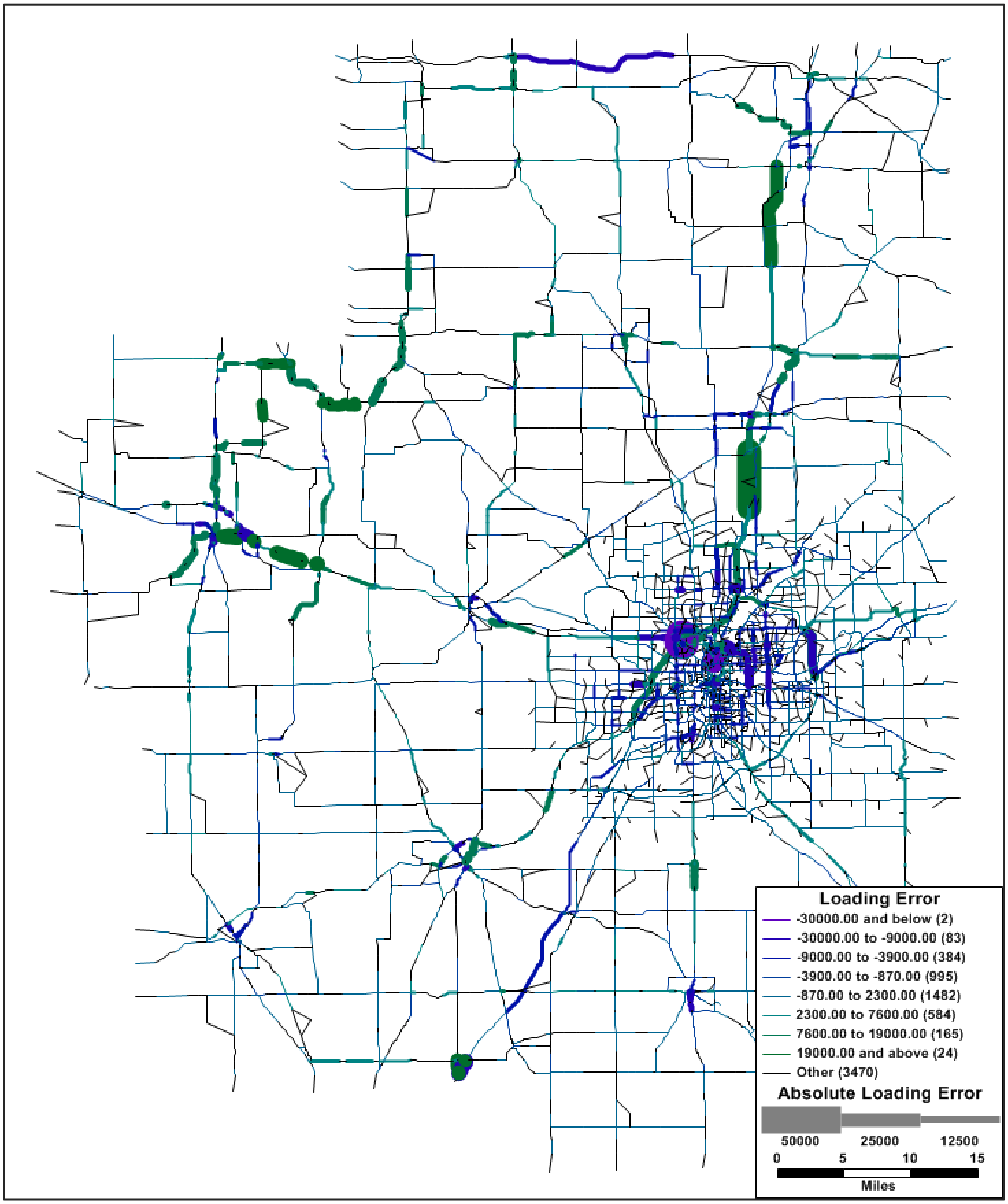


FIGURE 127: LOADING ERROR OF CUEBIQ TRIPS AFTER EXTERNAL FRATAR AND OVERALL SCALING



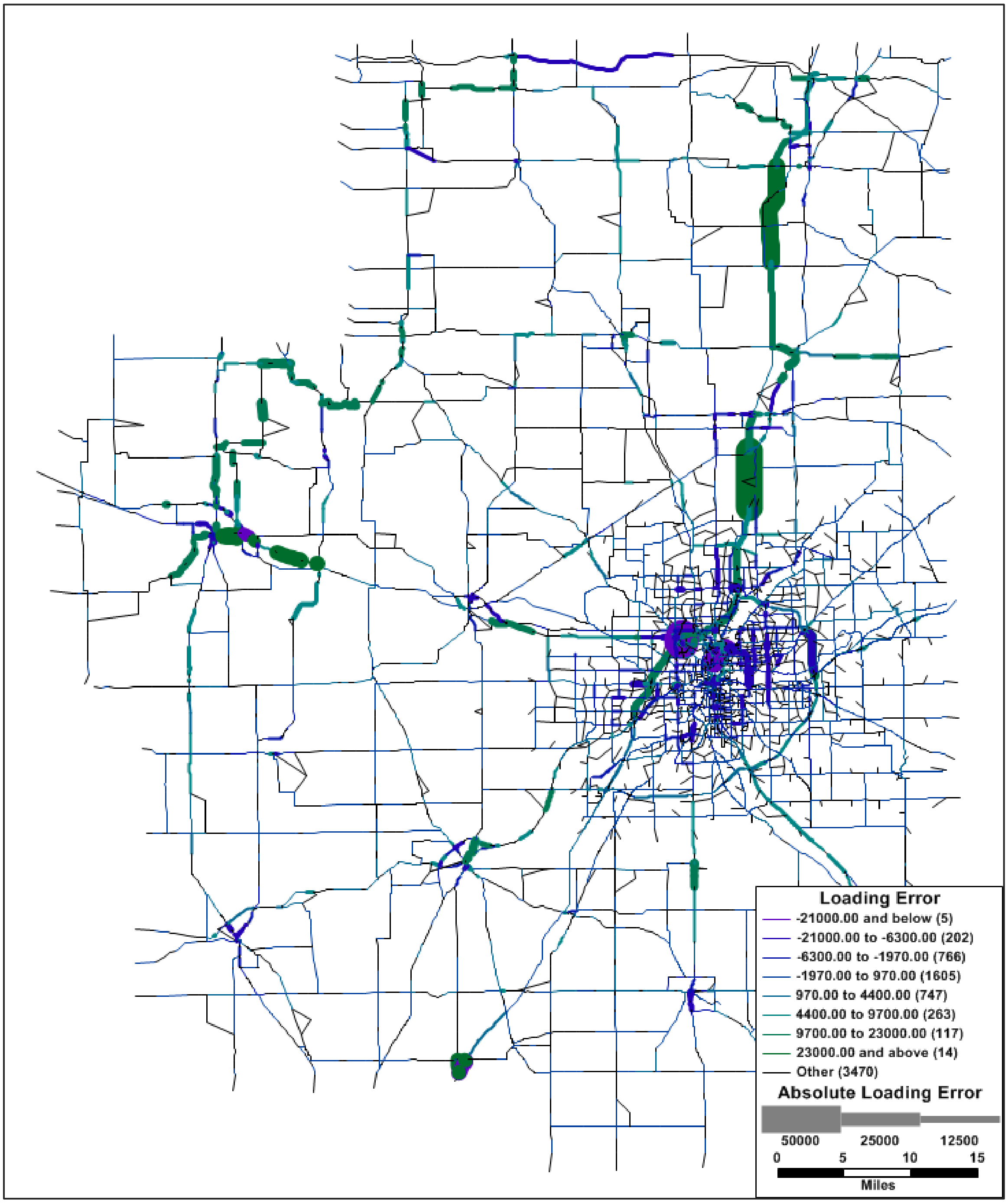
RSG mixed the two trip tables to generate the hybrid trip table as follows: all I-I trips come from CUEBIQ trip table while E-I, I-E, and E-E trips come from INRIX trip tables. The reason behind this strategy is that LBS data has very large catchment zones as externals which ends up with very high external trips. On the other hand, INRIX I-I truck trips are very high; however, CUEBIQ internal trips are in more reasonable range.

Since CUEBIQ trip table has only one vehicle class, INRIX auto, medium, and heavy truck trip tables were added together to make external segments of the hybrid trip table consistent with the I-I section. The hybrid trip table has 2,172,704 vehicle trips which 1,919,674 trips are I-I (about 88 percent). The hybrid trip table was then assigned and the I-I trips were rescaled to get 0 percent overall loading error. Table 34 presents assignment statistics after rescaling the hybrid trip table and Figure 128 illustrates the loading error.

TABLE 34: ASSIGNMENT STATISTICS FOR HYBRID TRIPS AFTER RESCALING I-I

Assignment Statistic	Average Error	Loading Error (%)	RMSE (%)	MAPE
All Facilities	1.78	0.03	71.91	155.08
Freeways	7,028.98	36.73	63.10	257.13
Arterials	-593.48	-5.78	51.25	154.04
Collectors	154.18	5.31	124.57	151.10
Locals	332.96	26.39	117.83	196.49
Urban Links	-1,135.10	-11.73	54.14	73.45
Rural Links	1,156.19	31.52	115.79	248.26

FIGURE 128: LOADING ERROR OF THE HYBRID TRIPS AFTER RESCALING I-I



APPENDIX D. DATA EXPANSION DETAILS

DETAILED CUEBIQ ISF RESULTS

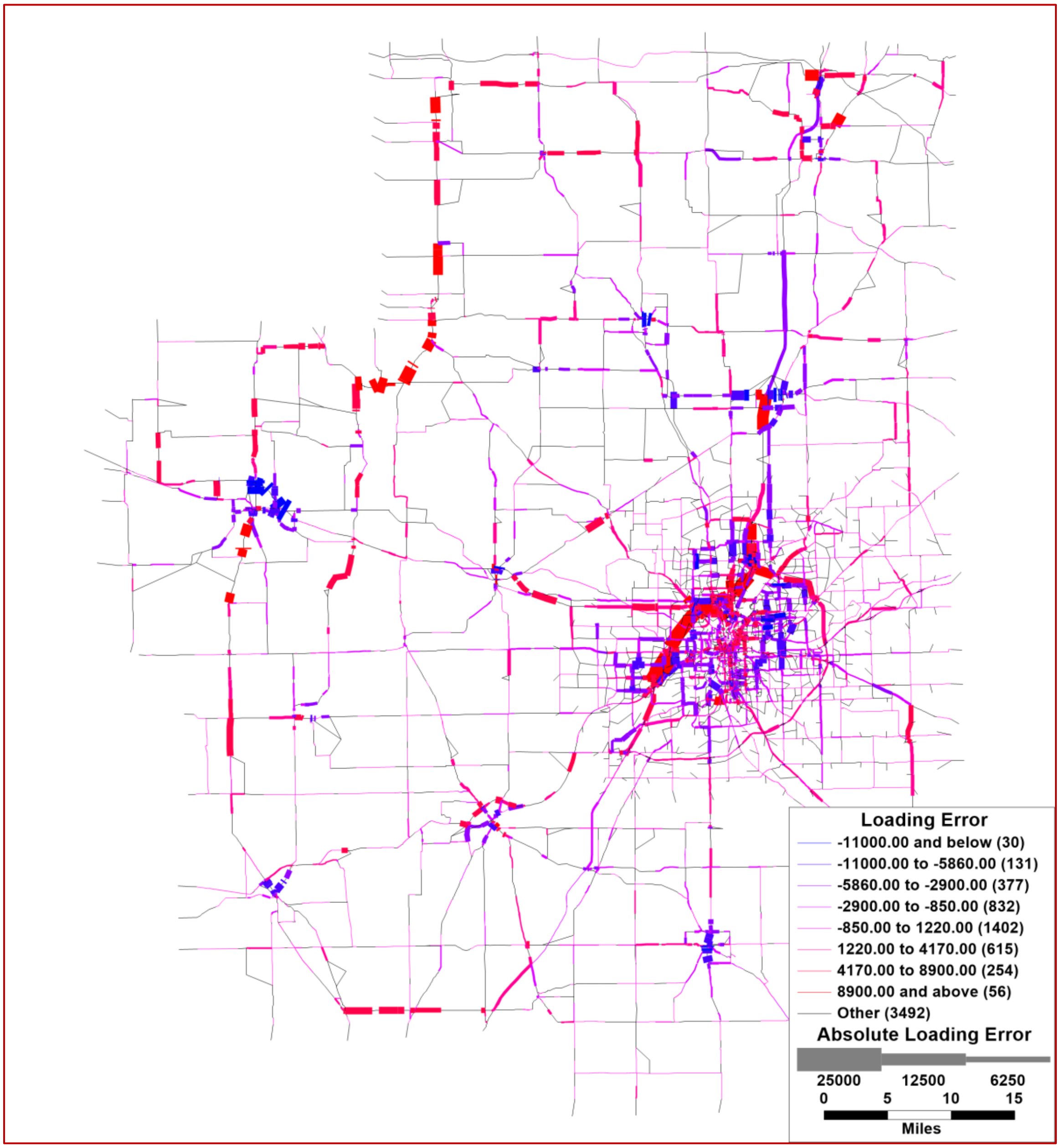
This ISF algorithm was applied to CUEBIQ and INRIX trip tables to expand passively collected OD trips. Table 35 reports assignment statistics of running expanded CUEBIQ trips and

Figure 129 illustrates the loading error.

TABLE 35: ASSIGNMENT STATISTICS FOR EXPANDED CUEBIQ TRIPS

Assignment Statistic	Average Error	Loading Error (%)	RMSE (%)	MAPE
All Facilities	-47.8	-0.7	52.9	78.7
Freeways	4007.0	20.5	37.0	21.3
Arterials	-0.7	-0.0	39.0	80.6
Collectors	-413.9	-14.4	96.0	137.0
Locals	-628.8	-49.8	118.8	220.0
Urban Links	-509.7	-5.3	43.5	49.2
Rural Links	310.6	8.4	72.9	112.9

FIGURE 129: ASSIGNMENT LOADING ERROR FOR CUEBIQ TRIPS AFTER ISF



Comparison between Table 35 and Table 33 indicates that ISF significantly improved the assignment results. The overall RMSE after applying ISF has been improved by 19 percent while freeways' RMSE has been improved by 18 percent. The same improvement can be also seen for urban and rural areas, especially for rural links, showing 48 percent improvement. The loading error map also confirms that ISF has resolved some of the initial trip table issues. The statistics in Table 35 are for any link with AADT; however, links should be prioritized by their facility type and count accuracy as model volumes on low-facility-type links are not as important as freeways or arterials. As a result, reviewing the model performance with more focus on high-ranking links and high-count accuracy is desirable. For this purpose, the statistics were regenerated only for the links with AADT greater than

1000 vehicles. This threshold is fair and reasonable enough according to the size of the model area and network resolution.

TABLE 36: ASSIGNMENT STATISTICS FOR EXPANDED CUEBIQ TRIPS FOR ANY LINK WITH AADT>1000

Assignment Statistic	Average Error	Loading Error (%)	RMSE (%)	MAPE
All Facilities	-137.1	-1.8	48.7	48.2
Freeways	4007.0	20.5	37.0	25.9
Arterials	-60.9	-0.6	38.0	35.3
Collectors	-664.9	-18.5	84.9	63.7
Locals	-1057.5	-45.4	97.4	56.4
Urban Links	-547.8	-5.6	42.7	42.4
Rural Links	243.1	5.1	61.8	54.4

As Table 36 indicates, the overall fit is better when the links with very low counts are filtered out. Most improvements are for local and rural links.

DETAILED INRIX ISF RESULTS

INRIX trips were also expanded by ISF methodology. Since INRIX trips include three different trip tables for auto, SUT, and MUT and the counts are available by vehicle classes, the expansion methodology was conducted on each vehicle class separately. The best trip table for each vehicle class was then selected and combined to create a complete trip table and rescaled to zero overall loading error for all vehicle classes.

Table 37 to Table 40 report the assignment statistics after expansion for INRIX trips by vehicle class and all classes together.

Figure 130 shows the INRIX trip assignment loading error.

Although ISF did not improve the model performance by vehicle class significantly, it improved the overall performance for all vehicle classes together by 19 percent. It also reduced RMSE for auto and medium-size trucks. Similar to CUEBIQ expansion results, rural links have been enhanced more than urban links and local links show more refinement than other facility types. High RMSE for SUT and MUT trips after ISF expansion demonstrates that truck counts are not as accurate even after data cleaning and network edits.

TABLE 37: ASSIGNMENT STATISTICS FOR EXPANDED INRIX AUTO TRIPS

Assignment Statistic	Average Error	Loading Error (%)	RMSE (%)	MAPE
All Facilities	38.3	0.8	71.4	86.9
Freeways	6238.8	42.2	57.1	65.1
Arterials	-86.5	-1.1	51.8	67.3
Collectors	-348.1	-13.5	100.8	101.8
Locals	-1100.5	-62.1	120.7	68.6
Urban Links	-756.4	-10.0	56.7	61.3
Rural Links	422.5	13.2	88.8	103.8

TABLE 38: ASSIGNMENT STATISTICS FOR EXPANDED INRIX MEDIUM SIZE TRUCK TRIPS

Assignment Statistic	Average Error	Loading Error (%)	RMSE (%)	MAPE
All Facilities	4.7	2.8	415.5	109.3
Freeways	-197.3	-19.5	333.0	88.0
Arterials	33.7	15.3	146.6	92.6
Collectors	-2.8	-3.9	178.8	123.4
Locals	8.8	29.0	93.9	94.4
Urban Links	-29.8	-16.4	97.1	71.8
Rural Links	43.0	28.7	659.6	151.6

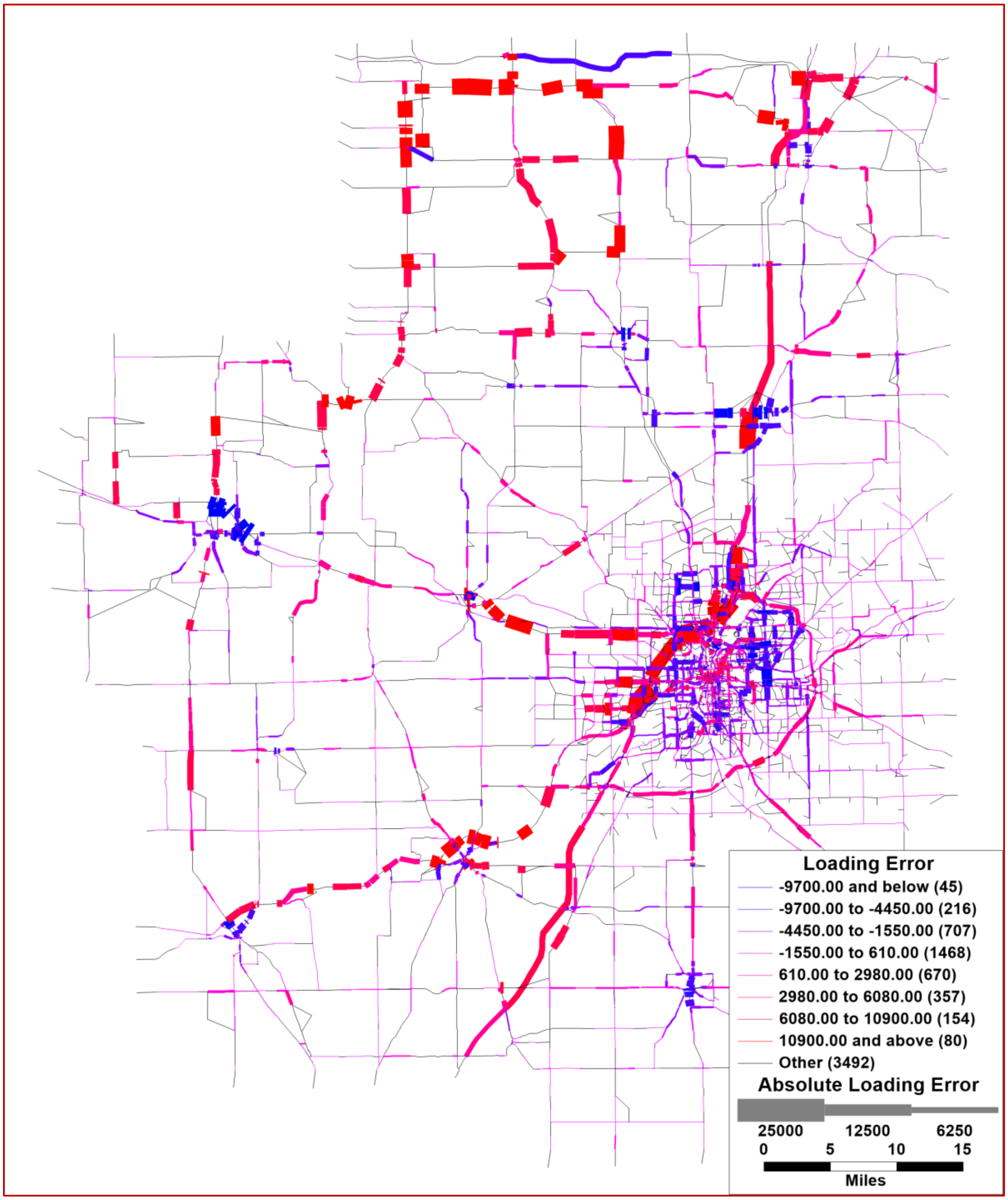
TABLE 39: ASSIGNMENT STATISTICS FOR EXPANDED INRIX HEAVY SIZE TRUCK TRIPS

Assignment Statistic	Average Error	Loading Error (%)	RMSE (%)	MAPE
All Facilities	1.7	0.5	141.5	299.6
Freeways	321.0	9.5	62.7	1288.0
Arterials	-40.4	-10.0	76.5	344.8
Collectors	12.3	16.1	211.2	204.9
Locals	-10.9	-100.0	122.0	100.0
Urban Links	-19.2	-7.3	111.8	166.2
Rural Links	27.0	6.6	149.3	453.3

TABLE 40: ASSIGNMENT STATISTICS FOR EXPANDED ALL INRIX TRIPS

Assignment Statistic	Average Error	Loading Error (%)	RMSE (%)	MAPE
All Facilities	138.2	2.1	59.3	91.4
Freeways	6545.7	33.4	43.0	38.6
Arterials	71.3	0.7	43.4	59.1
Collectors	-191.1	-6.7	110.3	126.2
Locals	-580.6	-46.0	119.4	83.4
Urban Links	-557.5	-5.8	45.5	54.0
Rural Links	779.5	21.0	96.2	133.4

FIGURE 130: ASSIGNMENT LOADING ERROR FOR INRIX TRIPS AFTER ISF



Comparison between

Figure 129 and

Figure 130 indicates that the model performance with expanded CUEBIQ trips is slightly better than INRIX expanded trips. For instance, I-90 corridor in the model area is underloaded in the INRIX loaded network; however, it is in a good agreement with the counts in the CUEBIQ loaded network. Reported assignment statistics also confirm this observation.

According to Table 37Table 38 and Table 39, performance of MUT trips is better than SUT trips. One reason might be inaccurate SUT counts which make the RMSE very high for SUT trips, another is that small absolute errors can still correspond to large relative errors when the total volume is small, as is the case for most SUT counts. The project team, therefore, developed a new approach for INRIX truck trips, using the same ISF methodology. In this new approach, SUT and MUT trips are combined and assigned to the network together and the model volumes are compared with the total truck AADT. This approach takes advantage of the fact that there are more links with total truck counts than SUT and MUT counts and that these total truck counts are believed to be more accurate than the class specific truck counts. The resulting expanded total truck trips were then split back out to SUT and MUT trips proportionally according to their original values. Table 41 to

Table 44 report the assignment statistics with this new approach after overall scaling and Figure 131 illustrates the assignment loading error.

TABLE 41: ASSIGNMENT STATISTICS FOR EXPANDED INRIX AUTO TRIPS (APPROACH 2)

Assignment Statistic	Average Error	Loading Error (%)	RMSE (%)	MAPE
All Facilities	38.3	0.8	71.4	87.3
Freeways	6183.7	41.8	56.3	64.3
Arterials	-49.4	-0.7	52.1	69.0
Collectors	-374.6	-14.5	101.1	101.6
Locals	-1101.0	-62.1	120.8	68.7
Urban Links	-741.3	-9.8	56.7	62.7
Rural Links	409.0	12.8	89.0	103.8

TABLE 42: ASSIGNMENT STATISTICS FOR EXPANDED INRIX MEDIUM SIZE TRUCK TRIPS (APPROACH 2)

Assignment Statistic	Average Error	Loading Error (%)	RMSE (%)	MAPE
All Facilities	-1.8	-1.1	405.3	89.3
Freeways	-288.6	-28.5	335.6	62.1
Arterials	36.0	16.3	95.4	79.6
Collectors	-10.6	-14.7	104.2	97.6
Locals	20.6	68.3	134.6	113.4
Urban Links	1.4	0.8	86.6	74.4
Rural Links	-6.2	-4.1	645.2	105.5

TABLE 43: ASSIGNMENT STATISTICS FOR EXPANDED INRIX HEAVY SIZE TRUCK TRIPS (APPROACH 2)

Assignment Statistic	Average Error	Loading Error (%)	RMSE (%)	MAPE
All Facilities	-18.3	-5.5	138.2	274.7
Freeways	330.8	9.8	59.4	1198.7
Arterials	-73.1	-18.1	90.0	307.7
Collectors	-0.9	-1.1	144.2	192.6
Locals	-10.9	-100.0	122.0	100.0
Urban Links	-33.9	-12.8	122.6	146.2
Rural Links	0.9	0.2	140.8	422.4

TABLE 44: ASSIGNMENT STATISTICS FOR ALL EXPANDED INRIX TRIPS (APPROACH 2)

Assignment Statistic	Average Error	Loading Error (%)	RMSE (%)	MAPE
All Facilities	-29.8	-0.5	55.7	80.9
Freeways	6403.5	32.7	42.7	38.1
Arterials	-78.7	-0.8	41.5	52.8
Collectors	-394.1	-13.7	97.1	110.4
Locals	-576.8	-45.7	118.5	83.3
Urban Links	-584.1	-6.1	44.9	49.8
Rural Links	440.2	11.9	81.8	115.4

Table 41 to Table 44 show improvement in RMSE in Approach 2 compared to Approach 1 especially for trucks and all vehicles together. The overall RMSE went down by almost 4 percent by combining trucks and expanding them together. This result would seem to confirm that truck counts by vehicle size are not as accurate as total truck AADT.

Similar to CUEBIQ trips, to reduce the impact of low-volume, less accurate counts, the statistics were regenerating by filtering out any link with AADT less than 1000 and truck AADT less than 50. The new statistics are reported in Table 45Table 41 to Table 48

Table 44. According to these tables, MUT RMSE significantly improved while the slight enhancement can be seen in other modes and overall. According to Table 35 and Table 36, performance of CUEBIQ trips is slightly better than INRIX trips.

FIGURE 131: ASSIGNMENT LOADING ERROR FOR INRIX TRIPS AFTER ISF (APPROACH 2)

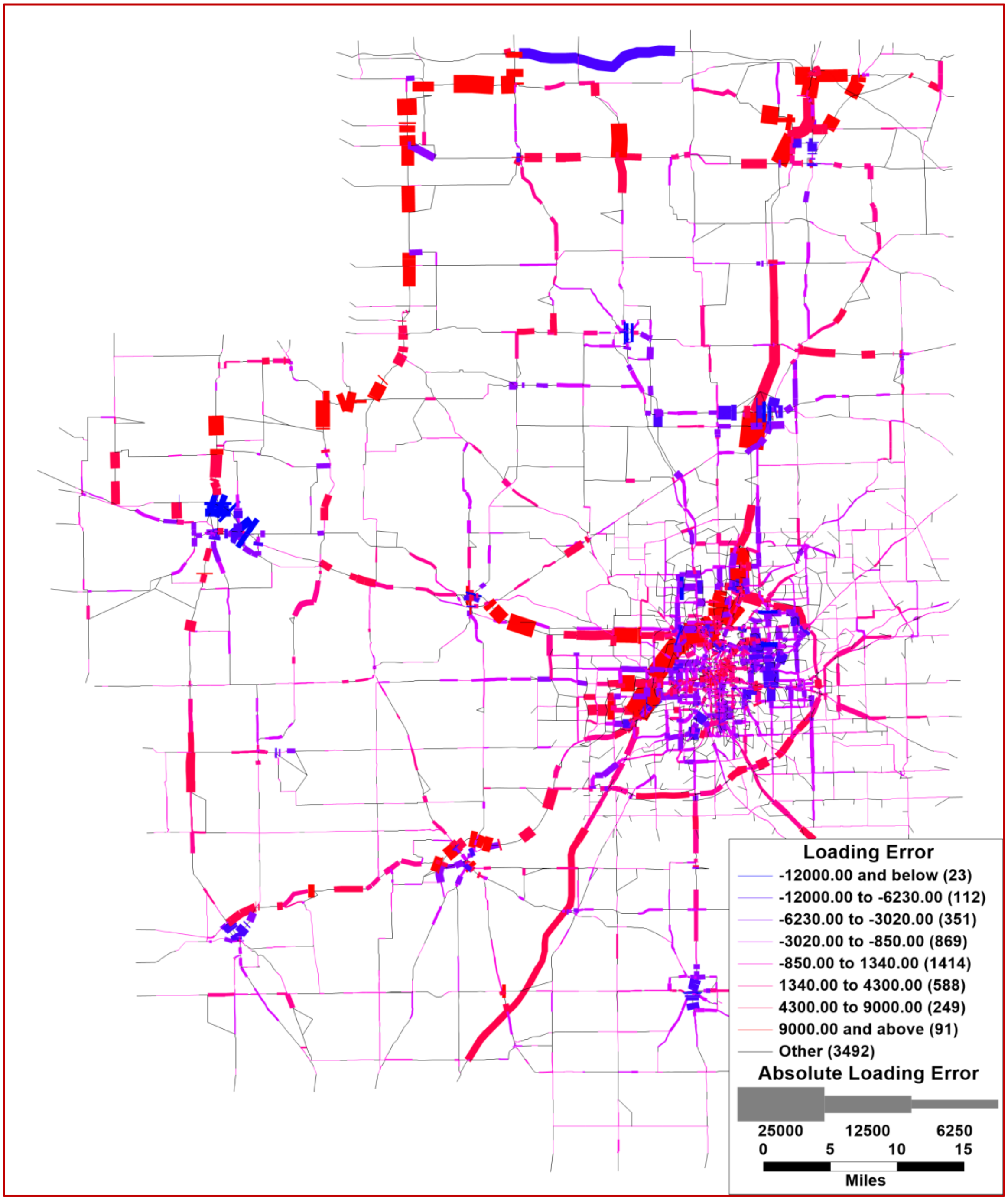


TABLE 45: ASSIGNMENT STATISTICS FOR EXPANDED INRIX AUTO TRIPS FOR ANY LINK WITH AADT > 1000 (APPROACH 2)

Assignment Statistic	Average Error	Loading Error (%)	RMSE (%)	MAPE
All Facilities	-65.6	-1.2	64.6	54.9
Freeways	6183.7	41.8	56.3	64.3
Arterials	-135.6	-1.7	50.1	44.0
Collectors	-637.6	-19.5	88.1	62.7
Locals	-1429.7	-56.0	107.7	50.4
Urban Links	-805.5	-10.2	55.2	51.4
Rural Links	373.0	9.3	76.7	56.6

TABLE 46: ASSIGNMENT STATISTICS FOR EXPANDED INRIX MEDIUM SIZE TRUCK TRIPS FOR ANY LINK WITH AADT > 1000 AND TRUCK AADT > 50 (APPROACH 2)

Assignment Statistic	Average Error	Loading Error (%)	RMSE (%)	MAPE
All Facilities	-9.7	-4.2	354.4	62.1
Freeways	-295.3	-28.7	332.6	61.0
Arterials	35.6	14.8	91.2	64.1
Collectors	-31.7	-28.2	81.4	59.4
Locals	64.8	99.6	99.6	99.6
Urban Links	-4.6	-2.0	78.0	58.7
Rural Links	-18.4	-7.8	536.5	66.9

TABLE 47: ASSIGNMENT STATISTICS FOR EXPANDED INRIX HEAVY SIZE TRUCK TRIPS FOR ANY LINK WITH AADT > 1000 AND TRUCK AADT > 50 (APPROACH 2)

Assignment Statistic	Average Error	Loading Error (%)	RMSE (%)	MAPE
All Facilities	-78.6	-11.4	74.7	68.9
Freeways	127.7	3.6	39.2	163.0
Arterials	-133.1	-22.2	72.9	50.2
Collectors	-33.6	-15.3	76.9	78.9
Locals	0.0	0.00	0.0	0.0
Urban Links	-92.9	-15.3	81.5	80.4
Rural Links	-64.1	-8.4	69.5	58.8

TABLE 48: ASSIGNMENT STATISTICS FOR ALL EXPANDED INRIX TRIPS FOR ANY LINK WITH AADT > 1000 (APPROACH 2)

Assignment Statistic	Average Error	Loading Error (%)	RMSE (%)	MAPE
All Facilities	-126.3	-1.6	51.3	48.4
Freeways	6403.6	32.7	42.7	38.1
Arterials	-143.7	-1.4	40.4	37.3
Collectors	-643.3	-17.9	85.4	61.1
Locals	-1070.5	-45.9	97.3	53.6
Urban Links	-627.4	-6.4	44.0	42.7
Rural Links	409.6	8.6	69.9	54.5

DETAILED HYBRID ISF RESULTS

As shown in Section 3.2 |, the hybrid scenario is very similar to the CUEBIQ trip table. Since the purpose of the hybrid scenario is to capture strengths and advantages of both datasets, a new hybrid scenario using CUEBIQ and INRIX trips as defined as follows:

1- The new HYBRID scenario has two vehicle classes: Auto and Trucks. Truck trips are the summation of SUT and MUT trips.

2- All external trips including external-external, internal-external, and external-internal trips are coming from INRIX

3- Internal-internal auto trips are obtained from subtracting INRIX truck trips from CUEBIQ trips.

4- Internal-internal truck trips come directly from INRIX.

It should be noted that if the resulted internal-internal auto trip for any OD pairs would become negative (meaning that there are more INRIX trucks than CUEBIQ total trips), it is changed to either 0 or 0.5 depending on INRIX auto trips for those cells. If INRIX data show auto trips for those OD pairs, 0.5 is assigned to the cells with hope that the expansion will adjust the trips properly. In contrast, if INRIX auto trips are zero for those cells, it is assumed that there are no auto trips for those OD pairs.

ISF was also run for the new hybrid trips. Assignment statistics of hyrbid trips after running ISF for auto and trucks are reported in Table 49 and

Table 50. Table 51 also shows the assignment statistics for all vehicle classes together.

Table 50 reports the statistics for the links with AADT greater than 1000 vehicles per hour.

Figure 132 also presents the assignment loading error of hybrid trips.

TABLE 49: ASSIGNMENT STATISTICS FOR EXPANDED HYBRID AUTO TRIPS

Assignment Statistic	Average Error	Loading Error (%)	RMSE (%)	MAPE
All Facilities	-4.9	-0.1	66.0	81.9
Freeways	4846.4	32.8	50.2	51.1
Arterials	-138.2	-1.8	47.6	62.5
Collectors	-331.1	-12.8	94.8	96.8
Locals	-988.2	-55.7	115.6	71.0
Urban Links	-486.9	-6.4	53.0	58.5
Rural Links	156.2	4.9	79.5	97.0

TABLE 50: ASSIGNMENT STATISTICS FOR EXPANDED HYBRID TRUCK TRIPS

Assignment Statistic	Average Error	Loading Error (%)	RMSE (%)	MAPE
All Facilities	7.0	1.2	124.5	160.1
Freeways	9.2	0.2	51.2	39.0
Arterials	-100.0	-11.8	88.4	157.8
Collectors	65.0	34.6	234.3	169.6
Locals	45.0	103.1	216.7	182.7
Urban Links	-73.5	-12.7	110.2	125.3
Rural Links	45.1	7.6	128.8	184.7

TABLE 51: ASSIGNMENT STATISTICS FOR EXPANDED ALL HYBRID TRIPS

Assignment Statistic	Average Error	Loading Error (%)	RMSE (%)	MAPE
All Facilities	113.7	1.70	55.0	75.4
Freeways	4615.4	23.6	35.9	26.8
Arterials	232.1	2.3	42.4	51.3
Collectors	-335.5	-11.7	92.0	100.0
Locals	-554.1	-43.9	111.5	84.8
Urban Links	-91.3	-1.0	45.8	49.3
Rural Links	188.0	5.1	73.8	103.1

According to Table 49 to Table 51, the HYBRID scenario falls between CUEBIQ and INRIX auto trips from CUEBIQ data improved INRIX autos and combining SUT and MUT resulted in better truck trips. The overall RMSE for all vehicles is almost the same as INRIX overall RMSE though. Although the HYBRID trips did not generate the best fit overall, the scenario indicated that its mixed trips might provide higher quality data if analysis by vehicle class is required.

FIGURE 132: ASSIGNMENT LOADING ERROR FOR HYBRID TRIPS AFTER ISF

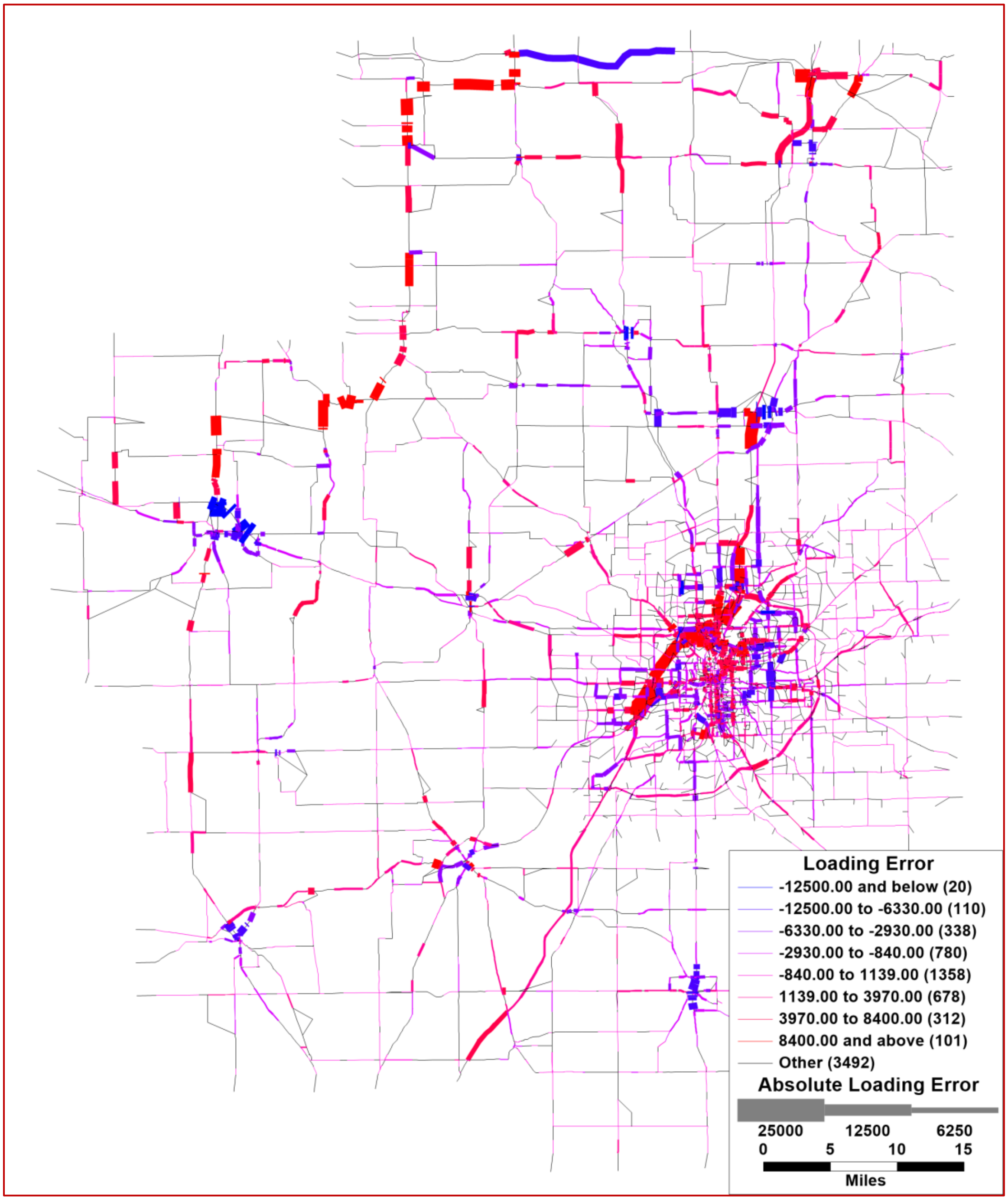


Table 52 to

Table 54 presents statistics by filtering out any link with AADT less than 1000 and truck AADT less than 50 to reduce impact of low-accurate countsTable 41

Table 44 for comparison.

TABLE 52: ASSIGNMENT STATISTICS FOR EXPANDED HYBRID AUTO TRIPS FOR ANY LINK WITH AADT > 1000

Assignment Statistic	Average Error	Loading Error (%)	RMSE (%)	MAPE
All Facilities	-109.2	-1.9	59.6	52.0
Freeways	4846.4	32.8	50.2	51.1
Arterials	-217.3	-2.8	45.8	40.9
Collectors	-571.4	-17.5	82.5	60.5
Locals	-1224.2	-47.9	102.8	52.7
Urban Links	-534.8	-6.8	51.6	48.5
Rural Links	51.8	1.3	68.0	53.5

TABLE 53: ASSIGNMENT STATISTICS FOR EXPANDED HYBRID TRUCK TRIPS FOR ANY LINK WITH AADT > 1000 AND TRUCK AADT > 50

Assignment Statistic	Average Error	Loading Error (%)	RMSE (%)	MAPE
All Facilities	-22.0	-3.0	108.1	79.7
Freeways	9.2	0.2	51.2	39.0
Arterials	-130.5	-14.3	83.3	62.3
Collectors	60.7	23.2	192.2	98.4
Locals	37.2	33.8	47.8	43.2
Urban Links	-105.9	-15.4	99.9	70.0
Rural Links	24.1	3.1	109.6	85.7

TABLE 54: ASSIGNMENT STATISTICS FOR ALL EXPANDED HYBRID TRIPS FOR ANY LINK WITH AADT > 1000

Assignment Statistic	Average Error	Loading Error (%)	RMSE (%)	MAPE
All Facilities	49.1	0.6	50.8	47.6
Freeways	4615.4	23.6	35.9	26.8
Arterials	181.8	1.7	41.4	37.5
Collectors	-551.3	-15.3	81.2	59.2
Locals	-914.3	-39.2	91.0	55.5
Urban Links	-115.5	-1.2	44.9	42.9
Rural Links	94.6	2.0	63.1	51.9

INDEPENDENT ODME RESULTS

In this section, results of running independent ODME on passively collected trips after fratar is reviewed and analyzed.

Table 55 reports the assignment statistics for CUEBIQ trips after running independent ODME. Comparison between Table 35 and

Table 55 indicates ODME improved CUEBIQ trips more than ISF as expected; however, the difference is only about 7 percent. In fact, ISF reduced overall RMSE from 65 percent to 53 percent which is more than half of what ODME could improve. Figure 133 shows the assignment loading error of CUEBIQ trips after independent ODME.

TABLE 55: ASSIGNMENT STATISTICS FOR CUEBIQ TRIPS AFTER INCEDEPNDENT ODME

Assignment Statistic	Average Error	Loading Error (%)	RMSE (%)	MAPE
All Facilities	13.52	0.20	45.82	68.91
Freeways	1602.17	8.18	28.11	22.57
Arterials	356.03	3.45	35.71	47.02
Collectors	-523.17	-18.21	75.11	92.68
Locals	-685.92	-54.36	123.28	81.85
Urban Links	-224.81	-2.35	38.30	44.26

Rural Links	140.31	3.78	60.47	96.70

Table 56 to Table 59 report the assignment statistics of INRIX trip table by vehicle class after running independent ODME. Figure 134 also presents the assignment loading error of the same run. Similar to CUEBIQ trips, ISF improved trips more than half of ODME for auto, heavy-size truck, and all vehicles. ISF did not enhance medium-size truck trips while ODME improved them by 12 percent.

FIGURE 133: ASSIGNMENT LOADING ERROR FOR CUEBIQ TRIPS AFTER INCEDEPNDENT ODME

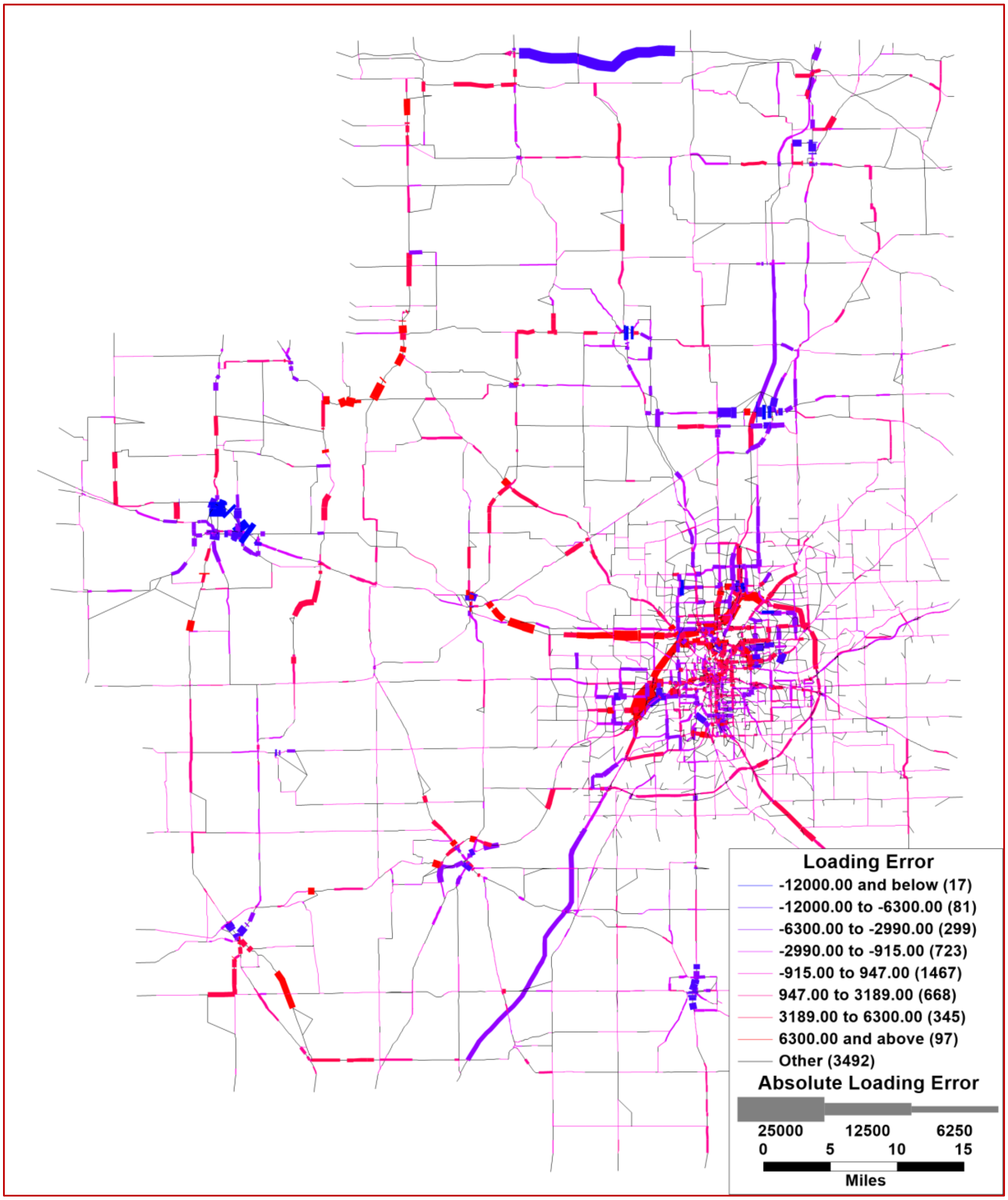


TABLE 56: ASSIGNMENT STATISTICS FOR INRIX AUTO TRIPS AFTER INCEDEPNDENT ODME

Assignment Statistic	Average Error	Loading Error (%)	RMSE (%)	MAPE
All Facilities	107.5	2.2	65.2	88.7

Freeways	4130.5	27.9	46.0	45.7
Arterials	172.7	2.3	50.0	73.0
Collectors	-283.7	-11.0	88.7	102.6
Locals	-1020.0	-57.5	124.7	73.3
Urban Links	-633.5	-8.4	51.9	62.3
Rural Links	466.4	14.6	80.5	106.7

TABLE 57: ASSIGNMENT STATISTICS FOR INRIX MEDIUM TRUCKS AFTER INCEDEPNDENT ODME

Assignment Statistic	Average Error	Loading Error (%)	RMSE (%)	MAPE
All Facilities	4.5	2.7	393.1	87.4
Freeways	-252.3	-25.0	326.7	62.8
Arterials	26.0	11.8	80.3	74.2
Collectors	6.1	8.4	119.9	98.6
Locals	20.8	68.9	141.8	134.9
Urban Links	12.7	7.0	88.0	75.3
Rural Links	-4.9	-3.2	624.9	101.4

TABLE 58: ASSIGNMENT STATISTICS FOR INRIX HEAVY TRUCKS AFTER INCEDEPNDENT ODME

Assignment Statistic	Average Error	Loading Error (%)	RMSE (%)	MAPE
All Facilities	-8.7	-2.6	117.6	261.2
Freeways	75.5	2.2	51.2	965.5

Arterials	-42.5	-10.5	70.7	297.8
Collectors	11.9	15.6	151.1	191.1
Locals	-10.9	-100.0	122.0	100.0
Urban Links	-18.1	-6.9	98.3	151.4
Rural Links	3.9	1.0	122.1	387.6

TABLE 59: ASSIGNMENT STATISTICS FOR ALL INRIX TRIPS AFTER INCEDEPNDENT ODME

Assignment Statistic	Average Error	Loading Error (%)	RMSE (%)	MAPE
All Facilities	-65.8	-1.0	51.8	79.5
Freeways	4283.0	21.9	34.6	24.8
Arterials	-134.8	-1.3	39.8	54.4
Collectors	-317.0	-11.0	87.0	106.5
Locals	-636.0	-50.4	125.5	80.6
Urban Links	-747.5	-7.8	41.5	48.8
Rural Links	551.3	14.8	76.5	113.3

FIGURE 134: ASSIGNMENT LOADING ERROR FOR INRIX TRIPS AFTER INCEDEPNDENT ODME

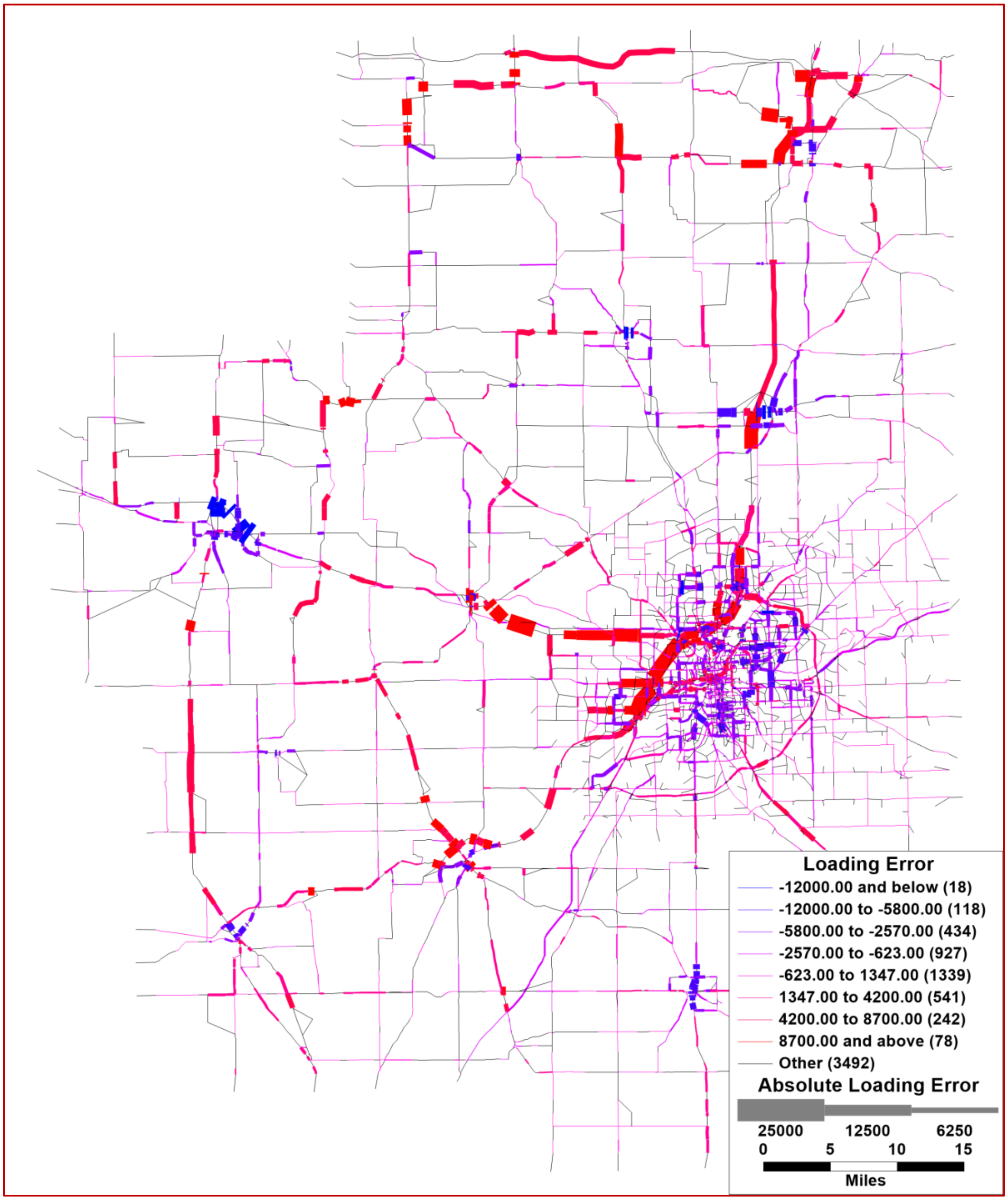


Table 60 to Table 62 report the assignment statistics of HYBRID trips after running independent ODME and

Figure 135 portrays the loading error. Although independent ODME shows a slight improvement for all vehicle classes, it did not improve trips by vehicle classes compared to ISF. This fact reconfirms the efficiency of ISF while it uses only a portion of all counts.

TABLE 60: ASSIGNMENT STATISTICS FOR HYBRID AUTO TRIPS AFTER INCEDEPNDENT ODME

Assignment Statistic	Average Error	Loading Error (%)	RMSE (%)	MAPE
All Facilities	10.8	0.2	66.5	92.9
Freeways	4611.1	31.2	46.4	52.9
Arterials	31.1	0.4	50.1	75.1
Collectors	-402.3	-15.6	93.8	108.5
Locals	-988.7	-55.8	125.4	79.4
Urban Links	-994.3	-13.1	51.6	64.5
Rural Links	525.0	16.4	85.3	112.5

TABLE 61: ASSIGNMENT STATISTICS FOR HYBRID TRUCK TRIPS AFTER INCEDEPNDENT ODME

Assignment Statistic	Average Error	Loading Error (%)	RMSE (%)	MAPE
All Facilities	-10.1	-1.7	132.0	166.8
Freeways	-487.4	-10.2	51.7	37.6
Arterials	-111.7	-13.2	96.7	166.9
Collectors	71.6	38.1	255.0	175.0
Locals	41.3	94.5	215.5	186.5
Urban Links	-91.3	-15.7	117.0	138.5
Rural Links	30.3	5.1	137.1	186.9

TABLE 62: ASSIGNMENT STATISTICS FOR ALL HYBRID TRIPS AFTER INCEDEPNDENT ODME

Assignment Statistic	Average Error	Loading Error (%)	RMSE (%)	MAPE
All Facilities	-242.4	-3.6	50.6	79.1
Freeways	3963.4	20.2	30.8	25.6
Arterials	-326.8	-3.2	38.2	53.2
Collectors	-495.3	-17.2	90.7	106.6
Locals	-703.2	-55.7	124.6	77.7
Urban Links	-1067.2	-11.1	40.3	49.0
Rural Links	515.2	13.9	75.9	112.0

FIGURE 135: ASSIGNMENT LOADING ERROR FOR HYBRID TRIPS AFTER INCEDEPNDENT ODME

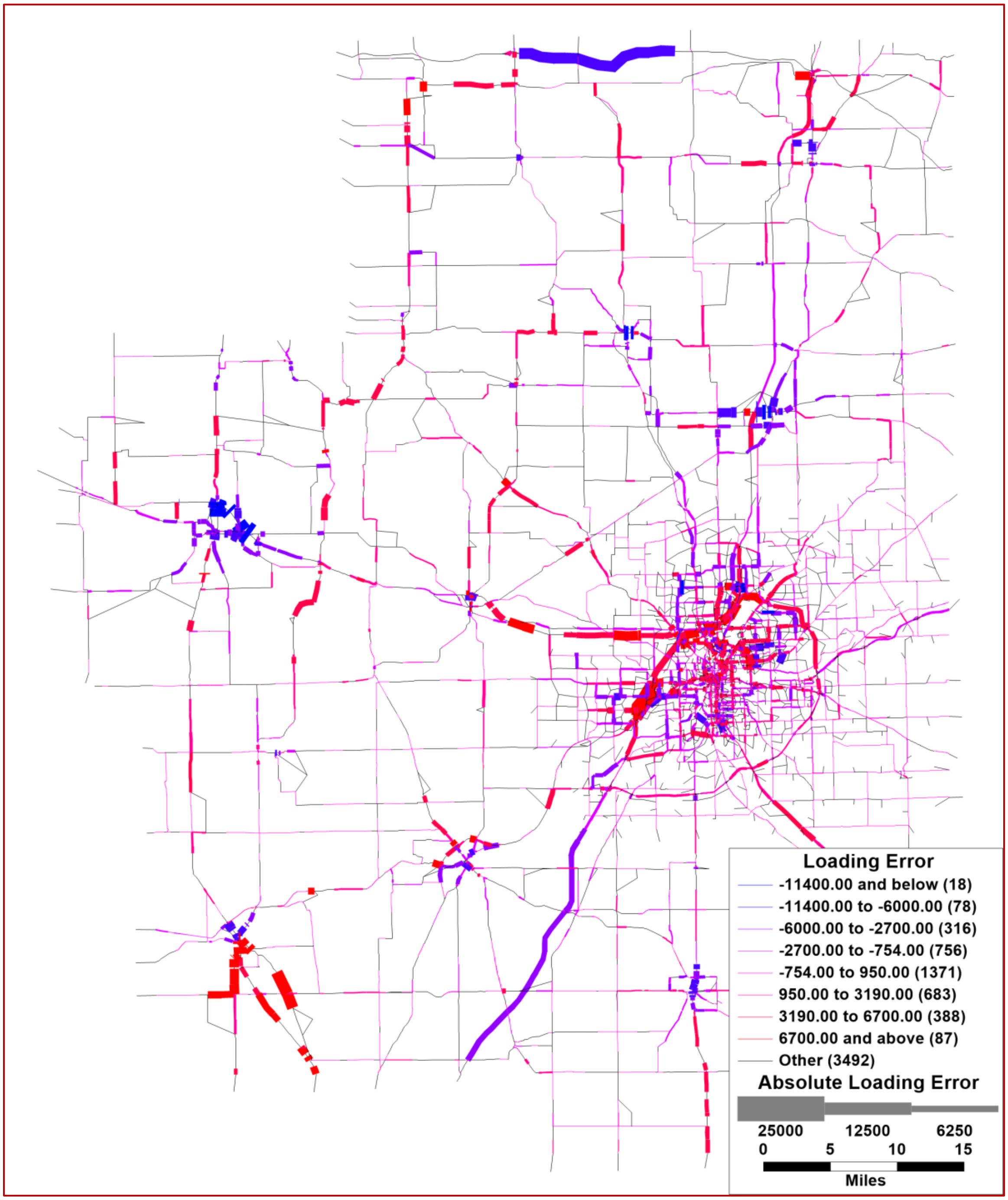


TABLE 63: MATRIX MEASURES BETWEEN INDEPENDENT ODME AND FRATAR

Statistic	CUEBIQ	INRIX	HYBRID
MAE	2.6	1.5	1.1
MAPE (%)	19.8	26.7	15.5

DETAILED SEQUENTIAL ODME RESULTS

Independent ODME results confirm that ISF significantly improved trips using only a portion of traffic counts. In fact, ISF enhanced trips more than half of the amount ODME could while using many less traffic counts. The project team, therefore, decided to run ODME on the trip table obtained by ISF to see whether or not better result than independent ODME can be found. The same ODME methodology with the same bounds and number of iterations was run on CUEBIQ, INRIX, and HYBRID trip tables, separately. Table 64 shows the statistics for CUEBIQ trips and Table 65 reports the statistics for the same trips only on the links with AADT greater than 1000 vehicle per hour.

TABLE 64: ASSIGNMENT STATISTICS FOR CUEBIQ TRIPS AFTER SEQUENTIAL ODME

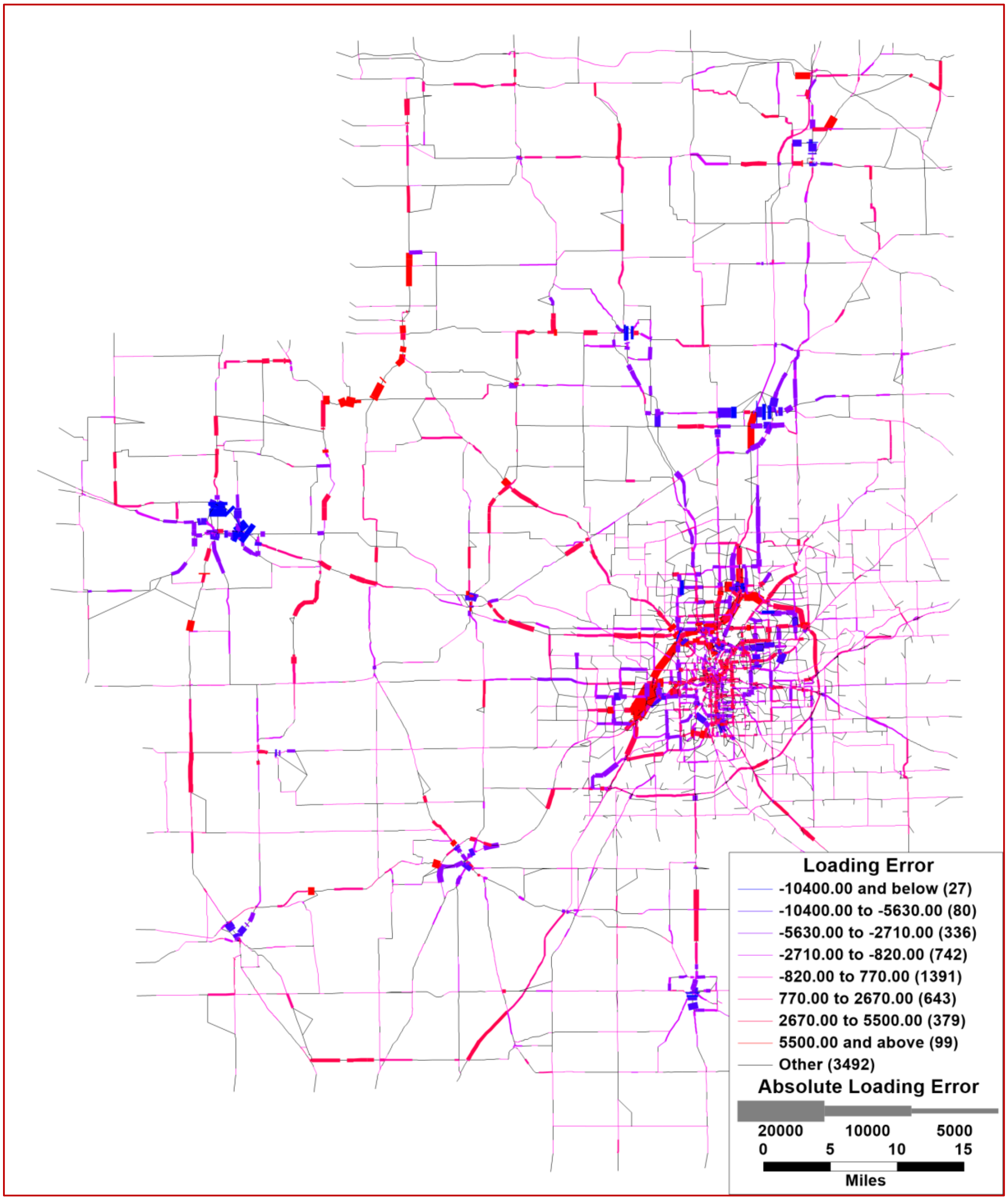
Assignment Statistic	Average Error	Loading Error (%)	RMSE (%)	MAPE
All Facilities	-95.1	-1.4	42.5	64.2
Freeways	2367.9	12.1	23.8	16.9
Arterials	165.0	1.6	33.3	43.2
Collectors	-568.0	-19.8	70.9	87.8
Locals	-654.8	-51.9	118.4	80.0
Urban Links	-341.7	-3.6	36.0	41.2
Rural Links	62.5	1.7	53.9	91.0

TABLE 65: ASSIGNMENT STATISTICS FOR CUEBIQ TRIPS AFTER SEQUENTIAL ODME FOR ANY LINK WITH AADT > 1000

Assignment Statistic	Average Error	Loading Error (%)	RMSE (%)	MAPE
All Facilities	-167.5	-2.2	39.2	38.1
Freeways	2367.9	12.1	23.8	16.9
Arterials	121.6	1.2	32.4	30.4
Collectors	-823.8	-22.9	62.5	47.4
Locals	-1127.3	-48.4	97.2	57.2
Urban Links	-367.9	-3.7	35.4	35.9
Rural Links	-36.0	-0.8	45.4	39.7

According to Table 64, the RMSE of sequential ODME is the best among all runs of expanded CUEBIQ trips. Thus, ISF helped ODME find a better solution than ODME alone. ISF and ODME can be run sequentially and the integrated approach improves the passively collected trips better than either of the approaches alone. Figure 136 shows the loading error after running sequential ODME.

FIGURE 136: ASSIGNMENT LOADING ERROR FOR CUEBIQ TRIPS AFTER SEQUENTIAL ODME



Sequential ODME was also run on INRIX trips to see if the same improvement can be seen for this dataset. The sequential ODME was run on the INRIX trips expanded by Approach 2. Table 66 to Table 69 reports the assignment statistics of running sequential ODME on expanded INRIX trips and

Figure 137 presents the loading errors.

TABLE 66: ASSIGNMENT STATISTICS FOR INRIX AUTO TRIPS AFTER SEQUENTIAL ODME

Assignment Statistic	Average Error	Loading Error (%)	RMSE (%)	MAPE
All Facilities	10.9	0.2	59.3	77.7
Freeways	2214.6	15.0	32.7	35.9
Arterials	140.2	1.8	44.6	62.1
Collectors	-307.5	-11.9	90.8	91.3
Locals	-819.5	-46.2	117.0	74.4
Urban Links	-599.6	-7.9	44.0	52.0
Rural Links	304.6	9.5	80.7	95.2

TABLE 67: ASSIGNMENT STATISTICS FOR INRIX MUDIUM TRUCKS AFTER SEQUENTIAL ODME

Assignment Statistic	Average Error	Loading Error (%)	RMSE (%)	MAPE
All Facilities	7.4	4.5	413.9	105.9
Freeways	11.9	1.2	332.9	115.5
Arterials	30.0	13.6	146.6	97.0
Collectors	-10.0	-13.9	137.7	111.3
Locals	25.0	82.8	164.4	160.4
Urban Links	-14.4	-7.9	129.9	83.3
Rural Links	31.0	20.7	647.4	131.4

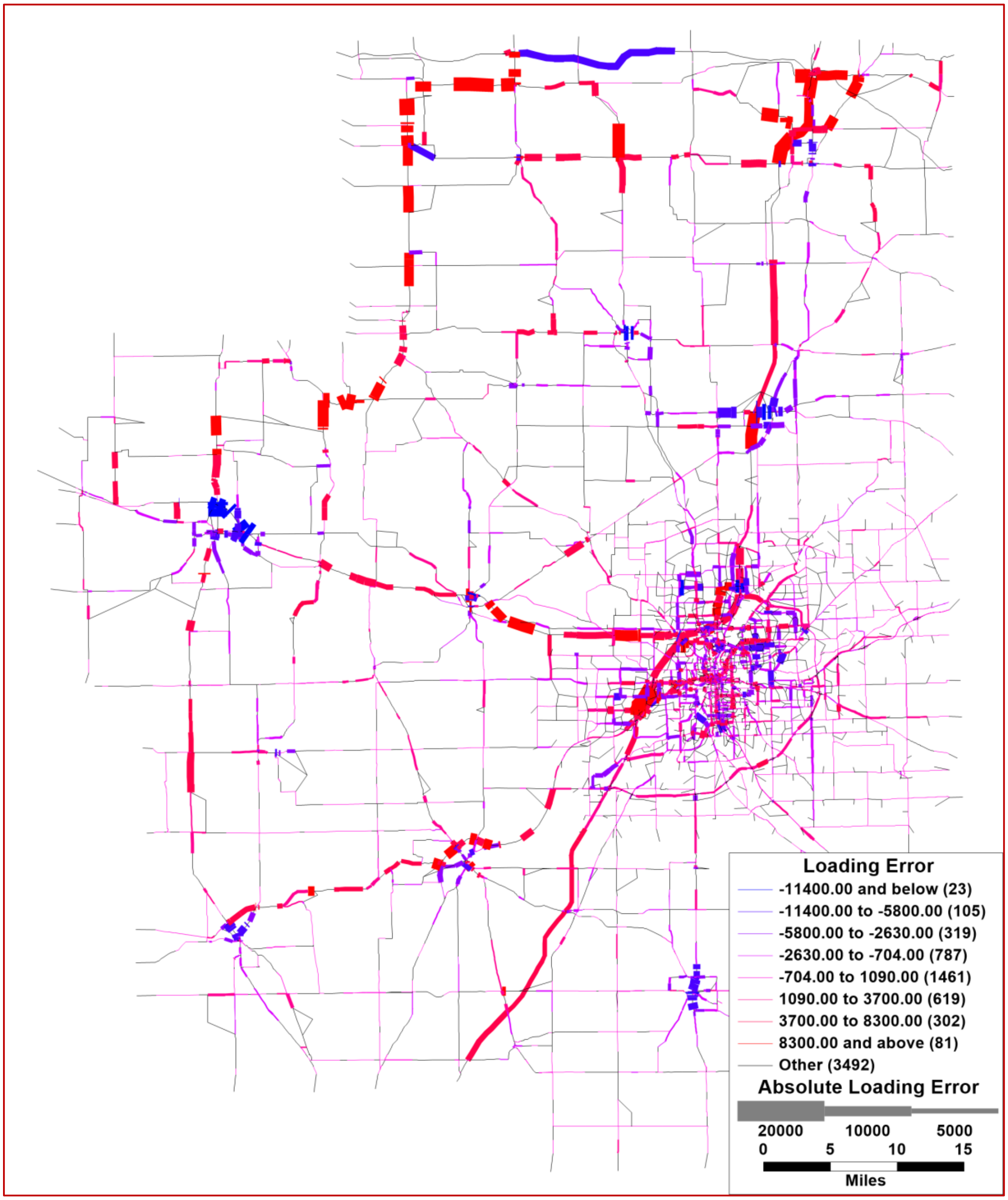
TABLE 68: ASSIGNMENT STATISTICS FOR INRIX HEAVY TRUCKS AFTER SEQUENTIAL ODME

Assignment Statistic	Average Error	Loading Error (%)	RMSE (%)	MAPE
All Facilities	13.8	4.2	150.5	295.2
Freeways	863.0	25.6	69.2	1346.6
Arterials	-49.4	-12.2	75.7	343.3
Collectors	4.0	5.2	151.7	193.9
Locals	-10.9	-100.0	122.0	100.0
Urban Links	-2.2	-0.8	130.3	168.4
Rural Links	33.9	8.3	154.6	441.3

TABLE 69: ASSIGNMENT STATISTICS FOR ALL INRIX TRIPS AFTER SEQUENTIAL ODME

Assignment Statistic	Average Error	Loading Error (%)	RMSE (%)	MAPE
All Facilities	91.1	1.4	50.1	74.5
Freeways	3357.3	17.1	28.8	22.7
Arterials	221.8	2.2	38.1	49.3
Collectors	-280.6	-9.8	90.1	101.3
Locals	-503.5	-39.9	115.0	82.2
Urban Links	-335.1	-3.5	39.4	43.7
Rural Links	455.4	12.3	77.6	108.4

FIGURE 137: ASSIGNMENT LOADING ERROR FOR INRIX TRIPS AFTER SEQUENTIAL ODME



Sequential ODME produced trips with lower overall RMSE compared to ISF and independent ODME; however, the improvement is not as good as the CUEBIQ trips' enhancement. RMSE for trucks goes up with sequential ODME but auto trips show reduction in RMSE. Even so, sequential ODME reconfirms that ISF could decrease overall RMSE by 10 percent which is about 63 percent of ultimate improvement by sequential ODME and 71 percent of enhancement by independent ODME. Table 70 to Table 73 reports the assignment statistics for the links with determined count threshold. As expected, the AADT threshold reduced the RMSE in all vehicle classes. The overall RMSE for all INRIX vehicle classes is higher than CUEBIQ trips with and without the count threshold.

TABLE 70: ASSIGNMENT STATISTICS FOR INRIX AUTO TRIPS AFTER SEQUENTIAL ODME FOR ANY LINK WITH AADT > 1000

Assignment Statistic	Average Error	Loading Error (%)	RMSE (%)	MAPE
All Facilities	-96.8	-1.7	53.0	44.8
Freeways	2214.6	15.0	32.7	35.9
Arterials	65.0	0.8	42.7	35.9
Collectors	-546.4	-16.8	78.1	52.5
Locals	-977.0	-38.3	104.5	64.3
Urban Links	-656.0	-8.3	42.7	40.4
Rural Links	237.5	5.9	68.7	47.7

TABLE 71: ASSIGNMENT STATISTICS FOR INRIX MEDIUM TRUCKS AFTER SEQUENTIAL ODME FOR ANY LINK WITH AADT > 1000 AND TRUCK AADT > 50

Assignment Statistic	Average Error	Loading Error (%)	RMSE (%)	MAPE
All Facilities	6.2	2.7	361.7	81.2
Freeways	12.5	1.2	329.9	116.9
Arterials	28.1	11.7	139.6	76.4
Collectors	-23.6	-21.1	115.1	82.6
Locals	-15.1	-23.2	23.2	23.2
Urban Links	-18.7	-8.3	118.4	74.2
Rural Links	42.0	17.9	537.9	91.4

TABLE 72: ASSIGNMENT STATISTICS FOR INRIX HEAVY TRUCKS AFTER SEQUENTIAL ODME FOR ANY LINK WITH AADT > 1000 AND TRUCK AADT > 50

Assignment Statistic	Average Error	Loading Error (%)	RMSE (%)	MAPE
All Facilities	-14.2	-2.1	80.5	76.2
Freeways	658.2	18.7	48.2	200.0
Arterials	-101.0	-16.9	60.4	54.5
Collectors	-18.3	-8.3	82.0	84.5
Locals	0.0	0.00	0.0	0.0
Urban Links	-21.1	-3.5	86.6	92.6
Rural Links	-5.6	-0.7	75.8	61.7

TABLE 73: ASSIGNMENT STATISTICS FOR ALL INRIX TRIPS AFTER SEQUENTIAL ODME FOR ANY LINK WITH AADT > 1000

Assignment Statistic	Average Error	Loading Error (%)	RMSE (%)	MAPE
All Facilities	11.6	0.2	459	41.9
Freeways	3357.3	17.1	28.8	22.7
Arterials	163.8	1.6	37.1	32.6
Collectors	-498.3	-13.9	78.6	53.0
Locals	-910.5	-39.1	94.3	56.2
Urban Links	-370.1	-3.8	38.6	36.2
Rural Links	423.8	8.9	65.8	48.4

Table 74 to

Table 76 report the assignment statistics for HYBRID trips after sequential ODME, and as expected, sequential ODME improved HYBRID trips more than independent ODME for vehicle classes. The overall RMSE in sequential ODME is a bit higher than independent ODME RMSE but higher loading error in independent ODME might be the reason of this observation. Table 77 to

Table 79 also report the assignment statistics with the count threshold. ODME confirms that ISF fixes many of the representativeness issues using less information and assumptions than ODME.

Figure 138 presents the loading error of sequential ODME assignment.

TABLE 74: ASSIGNMENT STATISTICS FOR HYBRID AUTO TRIPS AFTER SEQUENTIAL ODME

Assignment Statistic	Average Error	Loading Error (%)	RMSE (%)	MAPE
All Facilities	5.4	0.1	62.0	81.3
Freeways	3871.3	26.2	43.8	45.4
Arterials	-48.9	-0.6	46.2	65.6
Collectors	-315.2	-12.2	87.3	93.9
Locals	-895.8	-50.5	116.1	74.8
Urban Links	-487.7	-6.4	49.5	59.7
Rural Links	181.9	5.7	75.4	95.3

TABLE 75: ASSIGNMENT STATISTICS FOR HYBRID TRUCK TRIPS AFTER SEQUENTIAL ODME

Assignment Statistic	Average Error	Loading Error (%)	RMSE (%)	MAPE
All Facilities	-20.8	-3.6	126.6	167.0
Freeways	-826.3	-17.3	52.4	37.6
Arterials	-105.5	-12.4	90.2	170.5
Collectors	66.6	35.4	239.6	175.4
Locals	32.4	74.1	166.9	142.0
Urban Links	-104.8	-18.1	114.7	134.9
Rural Links	21.3	3.6	130.5	191.7

TABLE 76: ASSIGNMENT STATISTICS FOR ALL HYBRID TRIPS AFTER SEQUENTIAL ODME

Assignment Statistic	Average Error	Loading Error (%)	RMSE (%)	MAPE
All Facilities	95.7	1.4	51.9	74.1
Freeways	2853.0	14.6	27.8	19.6
Arterials	276.7	2.7	40.9	52.7
Collectors	-333.1	-11.6	85.3	96.5
Locals	-562.7	-44.6	110.7	81.3
Urban Links	-107.0	-1.1	43.1	49.1
Rural Links	177.3	4.8	69.9	100.9

FIGURE 138: ASSIGNMENT LOADING ERROR FOR HYBRID TRIPS AFTER SEQUENTIAL ODME

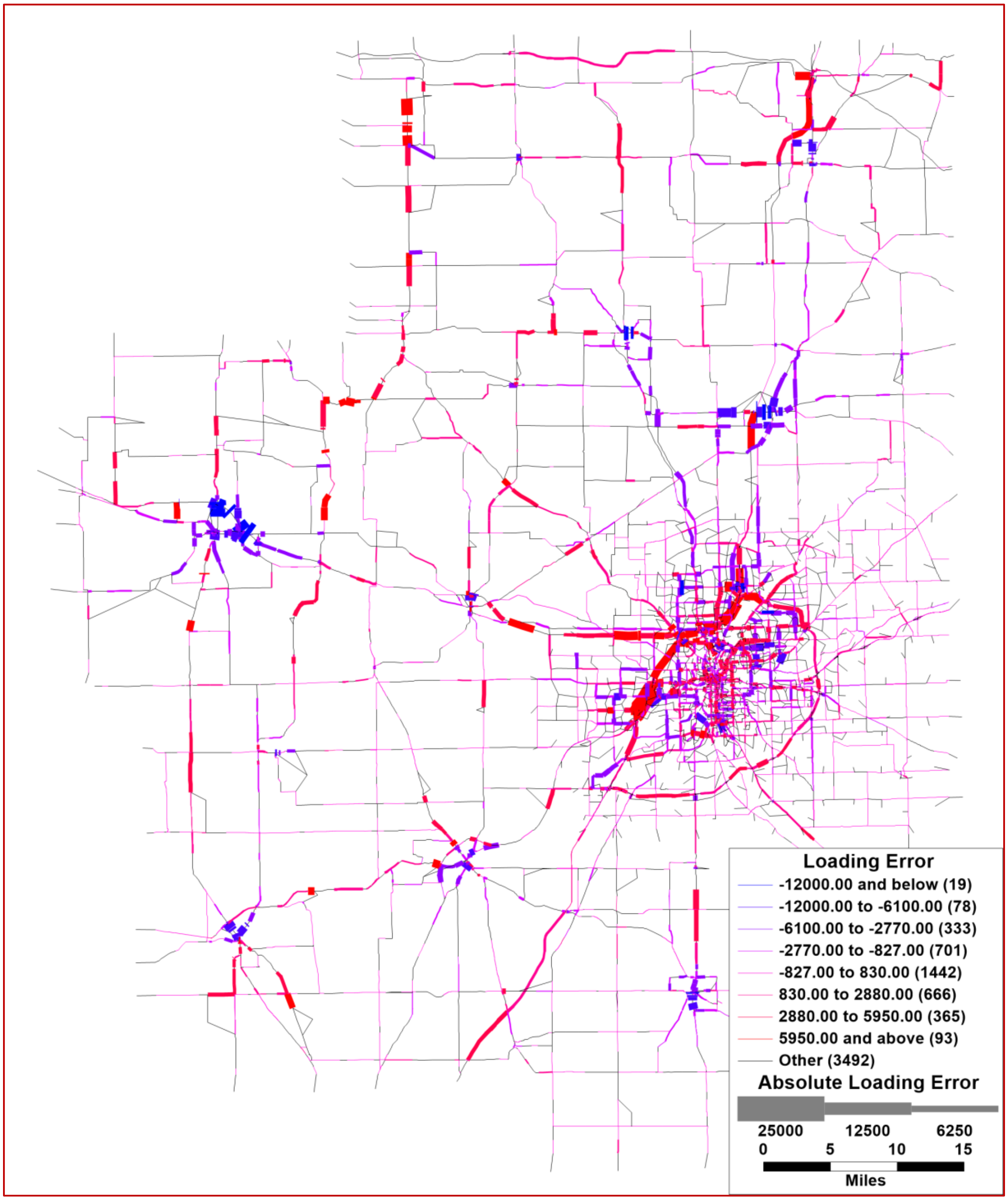


TABLE 77: ASSIGNMENT STATISTICS FOR HYBRID AUTO TRIPS AFTER SEQUENTIAL ODME FOR ANY LINK WITH AADT > 1000

Assignment Statistic	Average Error	Loading Error (%)	RMSE (%)	MAPE
All Facilities	-102.9	-1.8	55.8	48.8
Freeways	3871.3	26.2	43.8	45.4
Arterials	-134.4	-1.7	44.2	39.6
Collectors	-556.0	-17.0	75.8	55.3
Locals	-1079.0	-42.3	103.4	60.8
Urban Links	-546.5	-6.9	48.0	46.1
Rural Links	80.7	2.0	64.4	49.3

TABLE 78: ASSIGNMENT STATISTICS FOR HYBRID TRUCK TRIPS AFTER SEQUENTIAL ODME FOR ANY LINK WITH AADT > 1000 AND TRUCK AADT > 50

Assignment Statistic	Average Error	Loading Error (%)	RMSE (%)	MAPE
All Facilities	-59.0	-8.0	110.0	78.1
Freeways	-826.3	-17.3	52.4	37.6
Arterials	-139.7	-15.3	84.6	60.6
Collectors	63.5	24.3	198.4	97.0
Locals	31.8	28.9	40.8	36.9
Urban Links	-145.6	-21.1	103.8	68.1
Rural Links	-8.8	-1.1	111.0	84.4

TABLE 79: ASSIGNMENT STATISTICS FOR ALL HYBRID TRIPS AFTER SEQUENTIAL ODME FOR ANY LINK WITH AADT > 1000

Assignment Statistic	Average Error	Loading Error (%)	RMSE (%)	MAPE
All Facilities	23.5	0.3	47.7	44.9
Freeways	2853.0	14.6	27.8	19.6
Arterials	220.0	2.1	39.8	36.3
Collectors	-551.4	-15.3	75.0	54.9
Locals	-934.2	-40.1	90.4	54.2
Urban Links	-138.3	-1.4	42.2	40.6
Rural Links	76.2	1.6	59.4	48.7

In conclusion, ODME improved CUEBIQ, INRIX, and HYBRID compared to ISF and sequential ODME produced slightly better results compared to independent ODME.