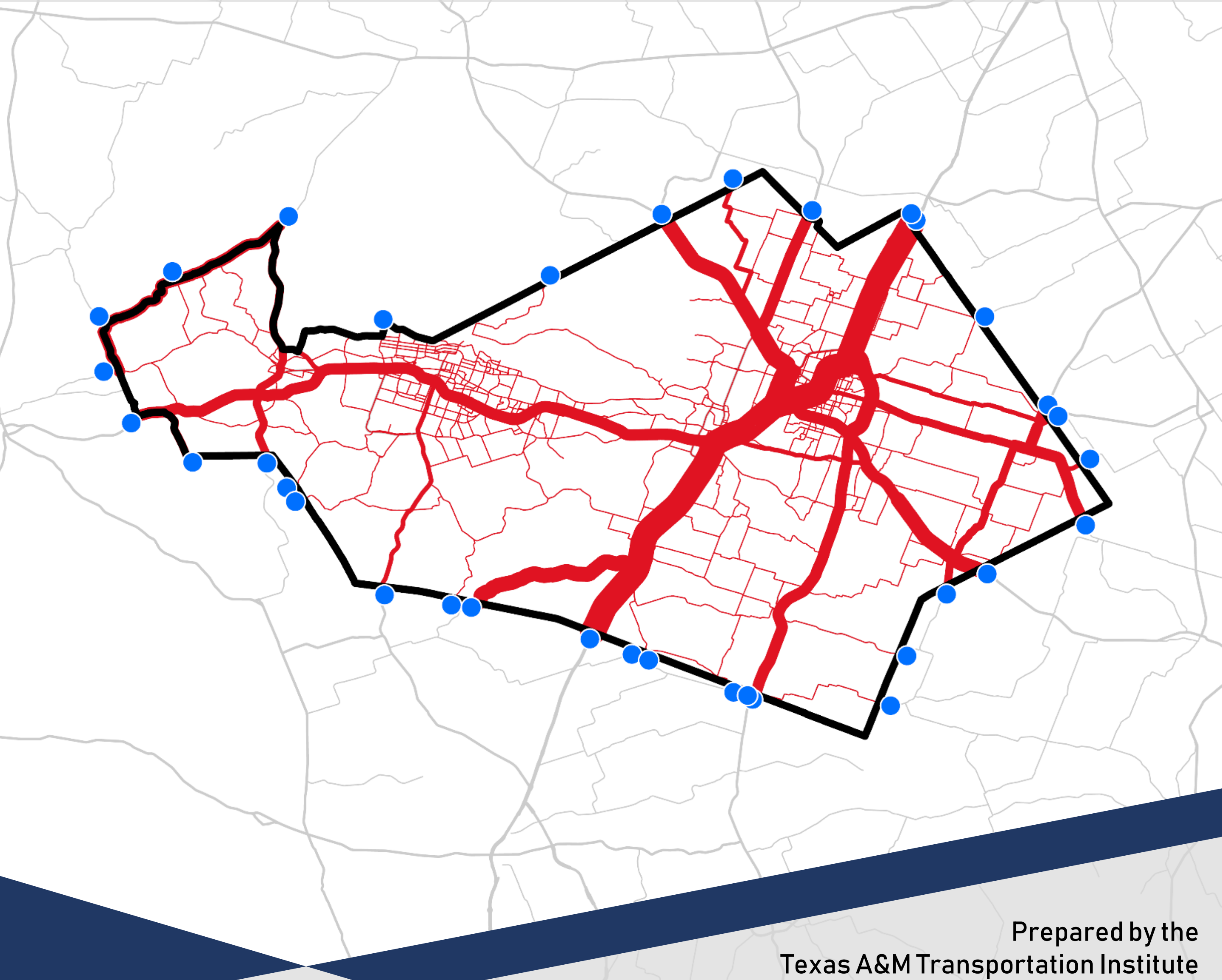
2015 KILLEEN-TEMPLE MPO EXTERNAL STUDY

Technical Summary

August 2018



Prepared by the
Texas A&M Transportation Institute

Texas A&M Transportation Institute

Texas Department of Transportation

2015 Killeen-Temple MPO External Study Using Cellular and GPS Data

Technical Summary

Texas Department of Transportation Travel Survey Program

By

Byron Chigoy, Associate Research Scientist
Michael Martin, Assistant Research Scientist
and
Edwin N. Hard, Research Scientist

Performed in cooperation with the
Texas Department of Transportation
and the
Killeen-Temple Metropolitan Planning Organization
and the
Federal Highway Administration

August 2018

TEXAS A&M TRANSPORTATION INSTITUTE
The Texas A&M University System
College Station, Texas 77843-3135

DISCLAIMER

The contents of this report reflect the views of the authors who are responsible for the data, findings, and conclusions presented herein. The content does not necessarily reflect the official views or policies of the Federal Highway Administration or the Texas Department of Transportation (TxDOT). This report does not constitute a standard, specification, or regulation.

ACKNOWLEDGEMENTS

The authors would like to thank Christeen Pusch, TxDOT Travel Survey Program Manager, and the Department for its continuing program to collect and analyze urban travel data to support travel demand modeling.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

STUDY BACKGROUND AND OBJECTIVE

This technical memorandum provides a summary of an external travel study performed for the Killeen-Temple Metropolitan Planning Organization (KTMPO) in the spring of 2018 for origin-destination (O-D) data acquired for the spring of 2015. The Texas Department of Transportation (TxDOT) acquired the O-D data for this period to coincide with traffic count data collected by TxDOT in the spring of 2015. The primary outcome of the study includes external-to-external (E-E), external-to-internal (E-I), and internal-to-external (I-E) travel data for passenger vehicles and trucks. TxDOT purchased cellular and Global Positioning System (GPS)-based O-D data for the study to provide estimates of external travel needed as input for the KTMPO travel demand model. The GPS data were acquired from INRIX and the cellular data from Airsage.

The use of passive data for external studies has become state-of-the-practice in recent years and provides significant benefits over prior methods such as roadside interviews and license plate matching techniques that garner safety and privacy concerns. The results of this study affect many uses of the travel model such as evaluation of freeway improvements, new location corridors and tolling studies, as well as air quality/conformity assessments. The study's primary objective was the development of detailed traffic matrices of external O-D interactions for the regional travel demand model. Since the use of passive technology to study travel is still evolving, this study also included a comparison of GPS and cellular data results. Researchers used data management techniques, algorithms, and analytical computations to address the technical and data processing challenges inherent in using large amounts of passive data.

Table 9 provides the expanded external trips for the Killeen-Temple MPO area. They represent summary results of the study and are explained with supporting information in later sections of this report.

KTMPO Internal Study Area

The KTMPO area includes all of Bell County and portions of Coryell and Lampasas counties. Core elements of the study area and design include Traffic Analysis Zones (TAZs) and 'external' locations where major roadways and thoroughfares cross the study area boundary. Figure 1 shows the study area, the study area boundary, the external locations, and the TAZs. The research team aggregated the KTMPO's TAZs into larger zones to compensate for cellular data's lower positional accuracy. These larger 'cell capture' TAZs consider census tract boundaries because they often align with the model zones and serve as the basis for cellular data expansion. In addition, the use of census tract geography coincides with American Community Survey (ACS) data products. The KTMPO's model TAZ structure contains 731 zones and 34 external stations. The research team aggregated the TAZs into 275 larger TAZs to better capture cellular data.

Identifying External Trips with GPS and Cellular Data

Identifying the external station (or roadway) used by trips that cross the study area boundary is an essential part of an external travel study. For GPS data, the external station is determined using a process known as map matching, where a trip's GPS points are electronically associated with the roadway network. This process reveals the route used between the trip's origin and destination that allows the external station to be identified.

Cellular data is polygon-based rather than point-based (as with GPS data), which means that the analysis assigns trip O-D pairs to TAZs and the route used between the trip's origin and destination often must be estimated. Typical analysis methods for cellular O-D data assign the route or roadway by "association" and assume that trips occurring within a geographic travel shed drawn around a road to be on the road. This process works relatively well for isolated roadways but is challenging in developed and/or urban areas where other parallel or nearby roadways exist. Where this occurs, roadways in close proximity to each other would in fact serve distinct O-Ds that are prohibitive to distinguish empirically. To identify the distinct O-Ds served by all the region's external roadways, the study area for cellular data, as shown in Figure 1, includes 88 "halo" zones around the periphery of the study area in counties bordering the area. The development of halo zones involved a careful review of regional populations, roadway locations, and travel patterns. The halo zones provide a basis for a select link analysis (see the Map Matching section for details) that routes observed O-D flows through the appropriate external roadway.

Halo zones serve several additional purposes. The first is to distinguish between resident and non-resident travel, with non-residents being those that live in the halo zones. Second, they capture cellular signals before they enter and after they exit the study area. Third, depending on forecasting needs, modelers can forecast each external zone independently, which allows for modeling different assumptions of growth in external traffic.

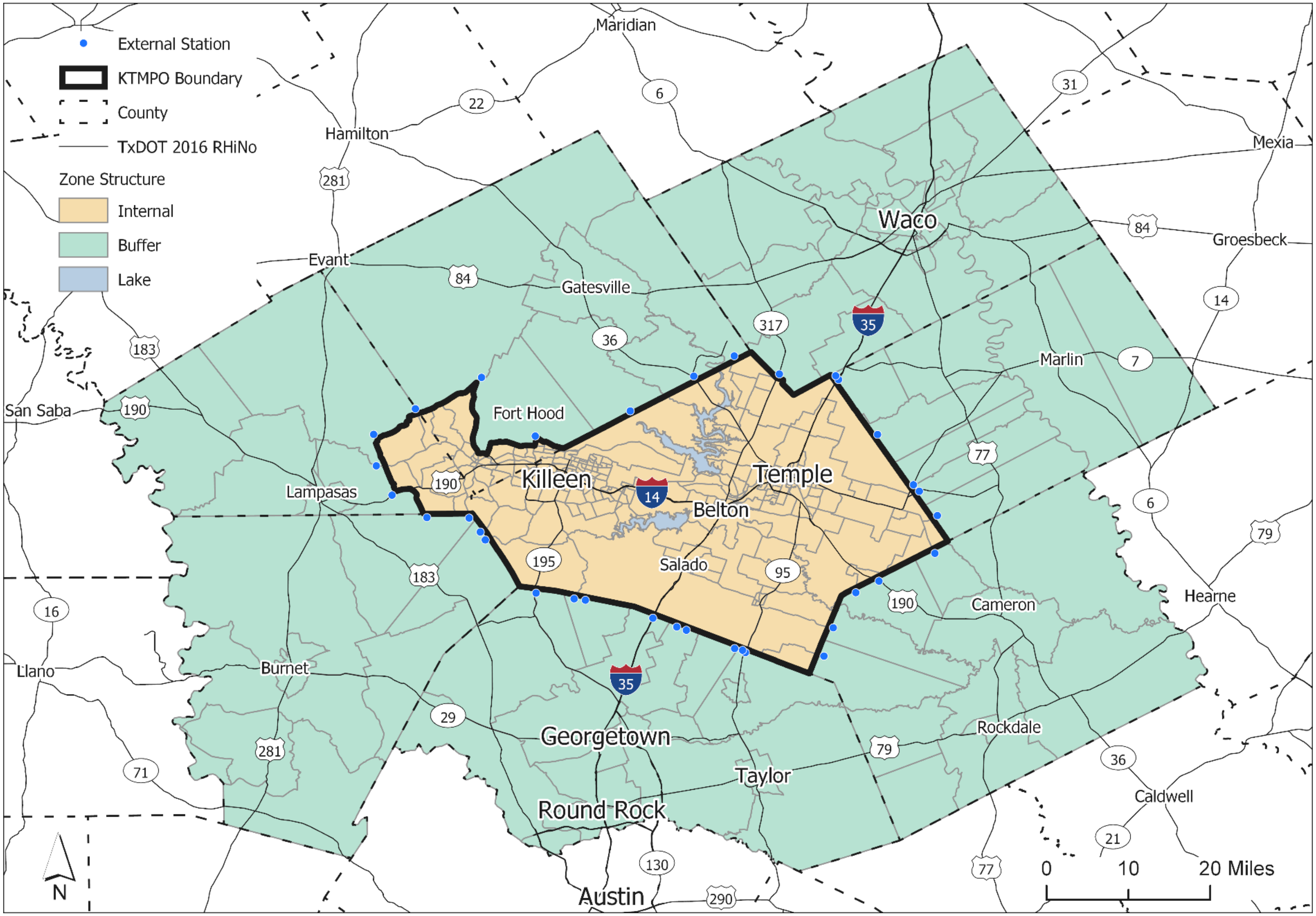


Figure 1. KTMPO External Survey Geography and Zone Structure.

DATA COLLECTION

Vendors of cellular and GPS data maintain large archives of data over many years and have extraction tools that efficiently pull data from their archives based on a range of dates and geographies provided for each study. The traffic counts are data collected in the "field" that require direct interaction with the roadway under study but provide relatively accurate ground truth data to baseline and expand the cellular and GPS data.

Traffic Counts

As with previous data collection methodologies, cellular and GPS data sources represent a sample of traveler information for a study area and therefore these data must be expanded to provide an estimate for the entire study area. The most common method to expand the data is the use of expansion factors derived from traffic counts. For external surveys, this equates to traffic counts at all external stations for a study area. Furthermore, to provide modelers the data needed to prepare a model for a particular study area, the survey results need to be disaggregated by non-commercial/passenger and commercial vehicle types. To meet this need, the traffic data required for developing the expansion factors must be classification counts.

The passive O-D data purchased for the Killeen-Temple area are from the spring 2015. This is the same year when the TxDOT performed saturation counts in the Waco and Brownwood districts where the KTMPO is located. Additionally, TxDOT collects annual counts of on-system roads throughout the state every year. The combination of saturation and annual counts in the KTMPO area is useful for model calibration due to the abundance of counts for links within the model network. However, both saturation and annual counts are volume counts and do not provide the classification data needed for external station data expansion.

Due to the lack of current count data, the research team compiled historical classification count data from the 2006 KTMPO External Survey to develop 2015 classification count estimates for the study area. These data included 24-hour daily averages for non-commercial and commercial vehicles at each of the external locations for the 2006 survey. Additionally, the team extracted 2015 saturation and/or annual count data from the STARS II database for each 2015 external location. They also queried the STARS II database to identify permanent count locations that aligned or were near external locations and had classification count data. Upon review, it was determined that there were three 2015 external sites that had permanent count stations nearby and those stations collected classification count data. For those permanent stations, classification count data for the same time period as the passive data were extracted, if possible. Those stations and classification count dates are:

1. External 732 – MS197 (December 2015 data);
2. External 746 – S215 (April 2015), W513 (April 2016), LW513 (Nov. 2015); and
3. External 761 – S374 (April 2015).

For some locations, no prior classification data was available. For these locations, the research team estimated classification data using a density analysis of the GPS data, nearby classifications, and/or data from a similar facility (e.g., freeway).

Cellular Data

TxDOT coordinated with TTI to purchase one month of cellular data from AirSage for the period between April 15, 2015 and May 15, 2015 for an average weekday. This time period when data were acquired corresponds with dates of TxDOT's traffic saturation counts for the region. Analysis of the data show that there are an estimated 2,054,702 (expanded) average weekday trips in the KTMPO study area, including 1,791,147 internal trips and 263,555 external trips. TxDOT purchased the data for average weekday trips, as only weekday travel is modeled for the KTMPO area. Additional options that were purchased included trips by five time periods, by trip purpose, and trip classification based on whether the trip was made by a resident of the study area or a visitor (non-resident) from outside the study area. The time periods included early AM (12 a.m.-6 a.m.), the AM Peak (6 a.m.-10 a.m.), Mid-Day (10 a.m.-3 p.m.), the PM Peak (3 p.m.-7 p.m.), late PM (7 p.m.-12 a.m.), and a full 24-hour day.

A notable drawback of cellular data is the lack of clear data for commercial vehicle travel. Presumably, commercial vehicle travel is present in the data. However, cellular data primarily tracks cell phone usage ultimately tied back to a subscriber's home, making it impossible to determine whether a subscriber left home and started driving a truck when they got to work. Additionally, for cellular studies, regional trips made by non-residents, and transients (i.e., briefly observed devices), are detected. For these trips, it is difficult to determine the residence, or primary activity place of the device, which precludes methods to determine if that place is associated with commercial vehicle activity. These trips are often external; however, their mode is unknown. For example, in the Tyler, Texas study, TTI observed a significant number of interstate I-E and E-I trips stopped at a large gas station/truck stop. This study did not focus on making similar observations as in Tyler, but researchers would expect similar findings in Killeen-Temple.. This indicates that trucks are present, but again the quantity of trucks cannot be determined via the data.

The E-E, I-E, and E-I trip-type designations use the trip's direction of travel relative to the study area and whether its origin or destination is inside or outside of the internal model region. Accordingly, a trip with an external zone origin and an internal zone destination is an E-I trip and vice versa for an I-E trip. E-E trips are those with both ends of the trip being external. It is important to note that for a trip to be E-E it must pass through the modeling region. Table 1 shows the percentage breakdown of each of these types of trips. Based on the cellular data, approximately 11 percent of all external traffic entering and exiting the KTMPO area are E-E trips.

Table 1. Cellular Area-Wide Weekday Averages.

	E-I	I-E	E-E	I-I	All Types
Total Trips	116,547	117,346	29,662	1,791,147	2,054,702
Percent Total Trips*	5.67%	5.71%	1.44%	87.17%	100.00%
Percent Total External Trips*	44.22%	44.52%	11.25%	NA	100.00%

* Total may not sum to 100 percent due to rounding.

The sample size calculation for cellular data is the penetration rate, which is the amount of cellular subscribers captured by AirSage data in the study area versus the total population of the study area. Based on prior studies using AirSage data, researchers estimate that the penetration rate for this study was about 17 percent of the study area population, or about 63,500 persons based on the 2015 model population. This rate is largely dependent on the market penetration of cellular carriers (or carrier) with whom AirSage has an agreement. During the data collection period for this study, AirSage indicated an agreement with one major U.S. carrier in the region.

GPS Data

TxDOT and TTI acquired three months of pre-processed GPS trip data from INRIX for the KTMPO area. These data are from INRIX's Insights™ Trips product and include individual trip records with waypoints. The trip records contain time-stamped trip start and trip end locations. Waypoint data is included for each trip record and provides points between the trip ends to identify the trip's route. INRIX provided a combined dataset of GPS data sources from non-commercial passenger cars/vehicles, commercial/freight vehicles (including trucks), and mobile applications.

The collection period for GPS data extended from February 2015 to May 2015. As with the cellular data, the data purchase dates correspond to TxDOT's traffic saturation counts for the region. The initial dataset includes 988,717 trip records. The study team processed each trip to determine if the trip was a weekday trip, whether or not it entered the study area, and if entered the study area by way of a roadway with an external station. This process revealed that 183,777 trips from the dataset did not meet these criteria (180,454 were weekend trips) making the final trip sample for analysis 804,940 trips. Each record contained a unique device identification (ID) and the trip ID with time-stamped trip end locations. Trip ends were pre-processed by INRIX using a 10-minute dwell time (e.g., could be thought of as a stop time) threshold and one-mile dwell distance. The actual trip end locations were anonymized by truncating the origin and destination coordinates to three digits of precision. The time between each GPS waypoint ranged between .017 and 10.0 minutes with an average of .83 minutes and median of 1.0 minutes. The waypoints for all trips totaled approximately 105 million records with an average of 130 waypoints per trip.

There are several ways to consider the sample size of GPS data relative to daily traffic. One way is to consider it as the average sample per day, or the total sample divided by the number of days in the collection period. Another way to calculate the sample size is to calculate it as a pooled sample, which is the total of all trips for observation days. This analysis reports sample size as the average sample per day. Accordingly, sampled passenger vehicles, including I-I trips, accounted for 27.7 percent of the average sample per day and 11.8 percent of the total sample for all external stations. Sampled trucks, including I-I trips, accounted for 72.3 percent of the average sample per day and 88.2 percent of the total sample for all external stations.

The sources of GPS data included in this sample are passenger vehicles, field service/local delivery fleets, and for hire/private trucking fleets. Within the sample the distribution of these sources is 26.1 percent passenger vehicles, 37.4 percent field service/local delivery fleets, and 36.5 percent for hire/private trucking fleets.

Unlike cellular data, the GPS data require significant post-processing by the data user. This processing requires large computing resources and data management to extract and impute additional details of the data. These details include TAZ attribution, map matching to roadways, and study area determinations of trip ends for E-E, E-I, and I-E attribution. Additional metrics that need to be calculated are speed and time between GPS waypoints. Despite the additional labor and computations involved in the analysis of GPS data, they offer a rich source of routing data and flexibility as compared to cellular data. Table 2 through Table 4 provide unweighted distributions of total, passenger vehicles and truck trips.

Table 2. GPS Weekday Total Sample.

	E-I	I-E	E-E	I-I	All Types
Total Trips*	108,594	109,052	161,879	425,415	804,940
PercentTotal Trips	13.5%	13.5%	20.1%	52.9%	100.0%
Percent Total External Trips	28.6%	28.7%	42.7%	NA	100.0%

*Unweighted results.

Table 3. GPS Weekday Total Sample Cars/Light Trucks.

	E-I	I-E	E-E	I-I	All Types
Total Trips*	15,900	16,385	12,471	165,418	210,174
Percent Total Trips	7.6%	7.8%	5.9%	78.7%	100.0%
Percent Total External Trips	35.5%	36.6%	27.9%	NA	100.0%

*Unweighted results.

Table 4. GPS Weekday Total Sample Medium and Heavy Trucks.

	E-I	I-E	E-E	I-I	All Types
Total Trips*	92,694	92,667	149,408	259,997	594,766
Percent Total Trips	15.6%	15.6%	25.1%	43.7%	100.0%
PercentTotal External Trips	27.7%	27.7%	44.6%	NA	100.0%

*Unweighted results.

MAP MATCHING

Map matching is the process of matching coordinate data to a digital roadway network. Most mobile device users see this process when using routing directions on mapping apps. It is most commonly associated with GPS data, but in this study, it also includes the process of matching cellular data to a roadway network. As with GPS data, there is no intrinsic association of the data to a given roadway.

Cellular Data

The map matching process for cellular data uses a combination of the KTMPO model network and the Texas Statewide Analysis Model (SAM) network. Stitching the two networks together allows for assigning trips outside of the study area in the external halo zones to the internal network based on the shortest path. The shortest path is typically a measure of distance or time between O-D pairs and the route to take, which minimizes one or the other. The analysis for this study uses the shortest time path, which is calculated using the speed limits of the network.

TransCAD is software used for the development of travel demand models. One of its core abilities is to assign an O-D matrix of traffic to a network. Built into TransCAD is an analytical component, termed select link, to obtain the O-Ds of traffic that cross a particular link during assignment. Accordingly, the analysis methodology uses a select-link analysis of roadway segments located at each of the 34 external stations. The select-link analysis process assigned E-E and I-E/E-I trips to the roadway network, while tracking the O-D of the trips that passed through each external link (see external stations in Figure 1).

Figure 2 shows that the E-E traffic appears logically assigned to the interstates and major roadways that bisect the study area. Figure 3 alternatively shows that I-E/E-I traffic additionally appears logically assigned to the major roadway network since more of the internal network is used to distribute the internal trips. Overall, the percentage breakdown of cellular derived external traffic per station is logical and consistent.

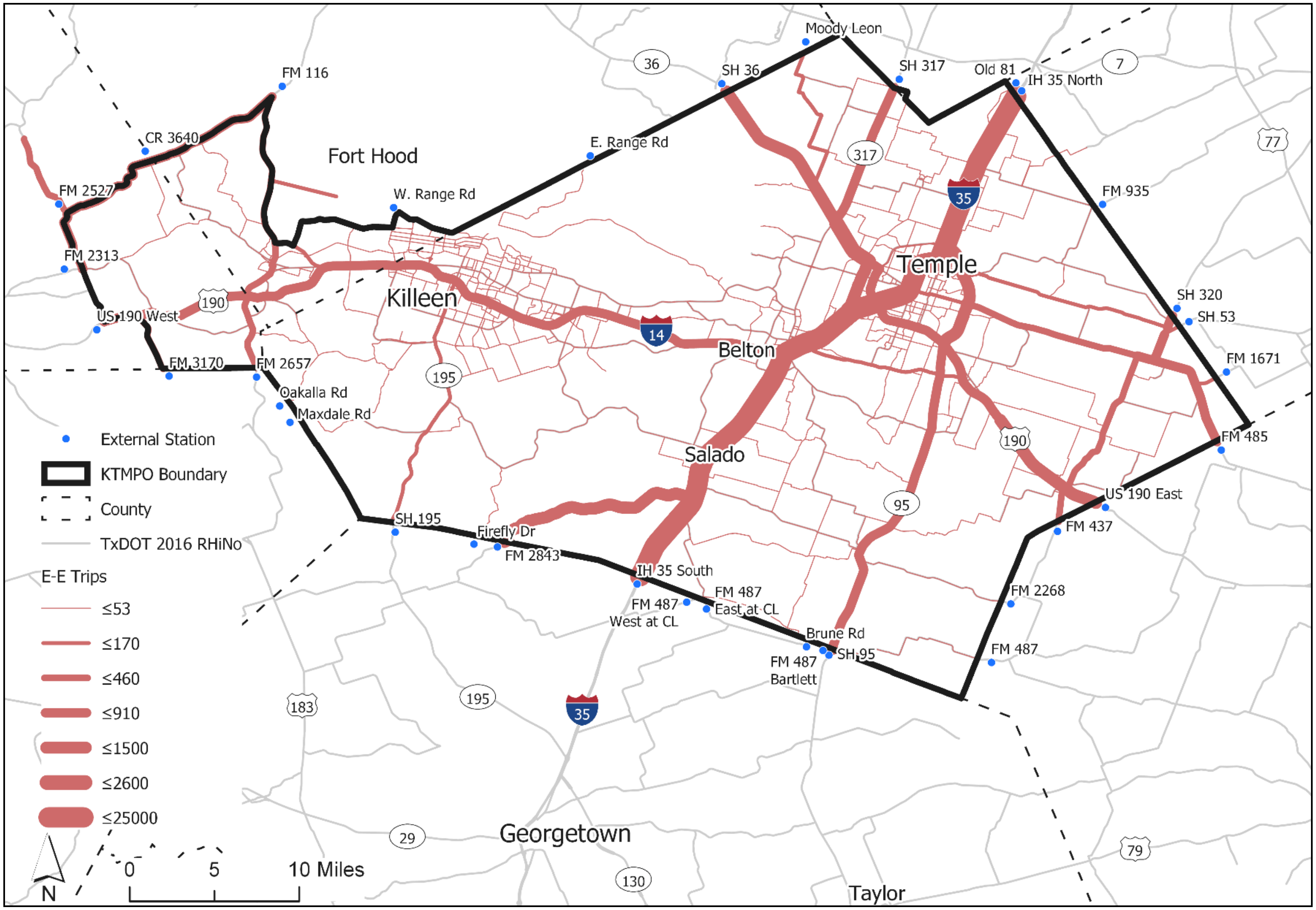


Figure 2. Cellular External Trip (E-E) Assignment Flow Bands.

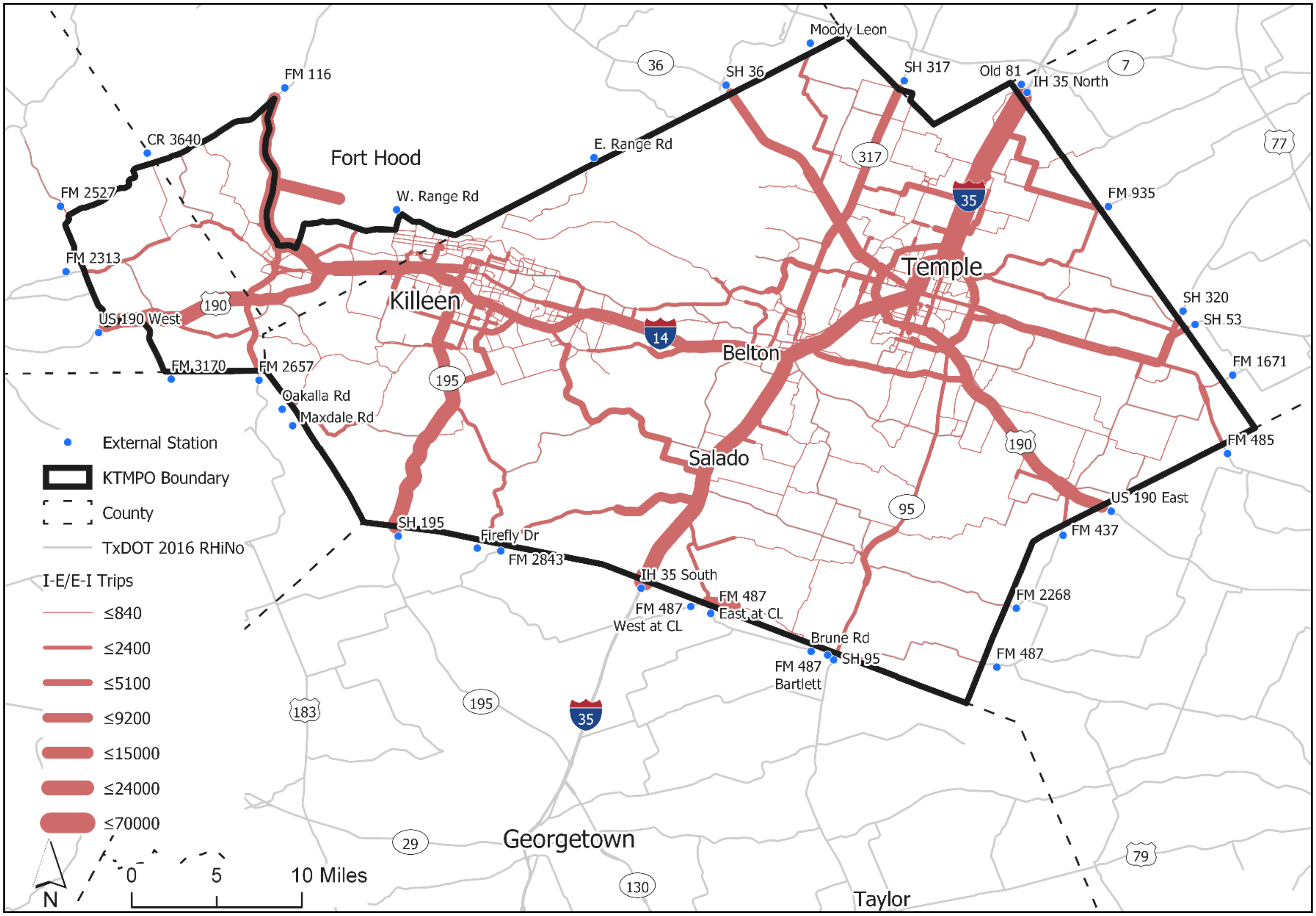


Figure 3. Cellular External-Local (I-E/E-I) Assignment Flow Bands.

GPS Data

Assigning GPS data to distinct road sections and determining ingress and egress external stations has significantly more challenges than those present with cellular data. The positional accuracy of GPS data is very good along with the time scale of the data (i.e., waypoints at one-minute intervals or less). On the other hand, these data are large and difficult to work with using traditional GIS applications and methods. Data management strategies are key to working with large amounts of GPS data. These strategies include data chunking, efficient sorting of data in memory or on a disk, and establishing efficient indexing routines. Figure 4 demonstrates a typical map-matching problem using GPS data with waypoints superimposed on a highway system.

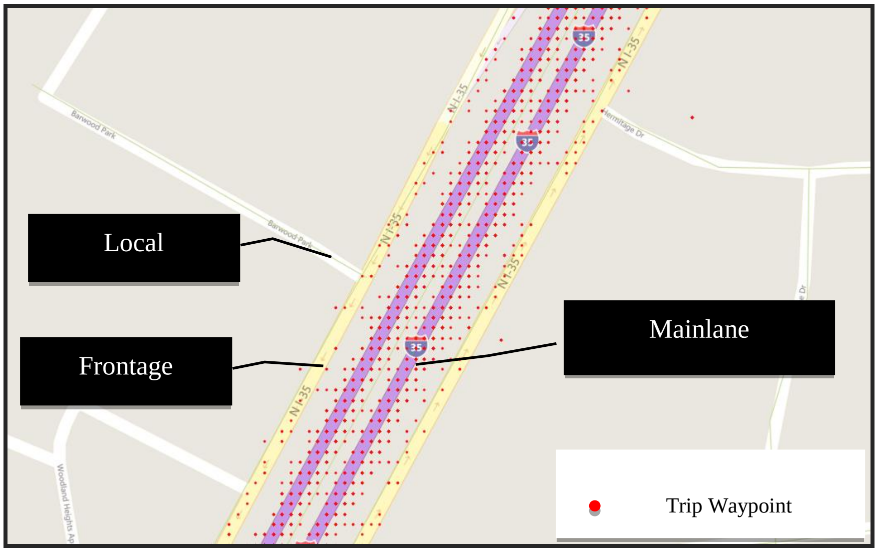


Figure 4. Matching GPS Data to Road Segments.

As illustrated, depending on the network and the precision of the coordinates, each GPS point matches to several different roadway sections. One way of solving this problem is to do a K-Nearest Neighbor (KNN) search for each point. KNN is an algorithm that attempts to locate a limited or unlimited set of results that are closest in space to a given point or object. However, this too has its drawbacks, as the return of the search can quickly become very large. To limit this problem, the algorithm imposes a limit of the three nearest roadway segments within a 100-meter radius.

The results of the KNN search determine the external roadway used when each trip crossed the study area boundary. Figure 5 illustrates an example of this process. The flag "insa" designates whether the point was in or out of the study area, and the flag "gaptest" indicates the first point that was not in the study area. In this illustration, each flagged waypoint cross-references to a table of external stations based on the highway names from the KNN search and highway names in the external look-up table. The algorithm selects the highway name that matches the KNN search and is closest to the point.

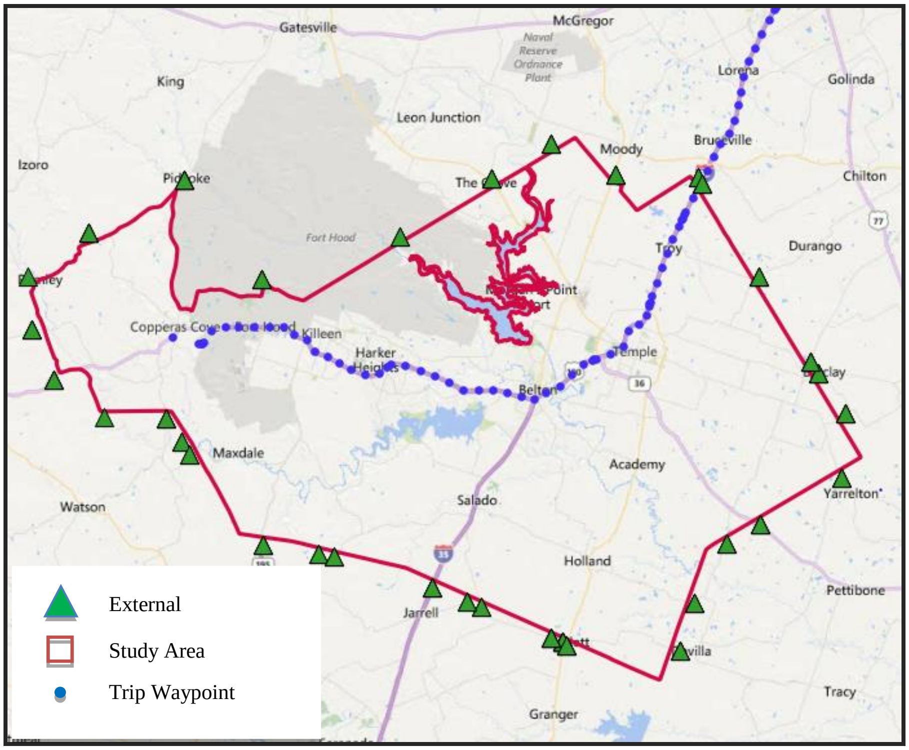


Figure 5. External Station Matching Example.

The results of the external station-matching algorithm indicate the following:

- Most matches are associated with the first KNN result.
- With no initial match, the algorithm searches the subsequent KNN results.
- When this process fails to find a match, the search broadens to include the immediate additional 10 points (+/-5 waypoint sequence) for each flagged waypoint and the results of the KNN for these points.

However, in some cases not all trips pass through one of the external stations in the study area, as illustrated in Figure 6. The reason for this is that travel demand model development does not account for all points of ingress/egress and focuses on major roads and local roads to major attractors. Although these roads are often the most logical routes, drivers often take "shortcuts."

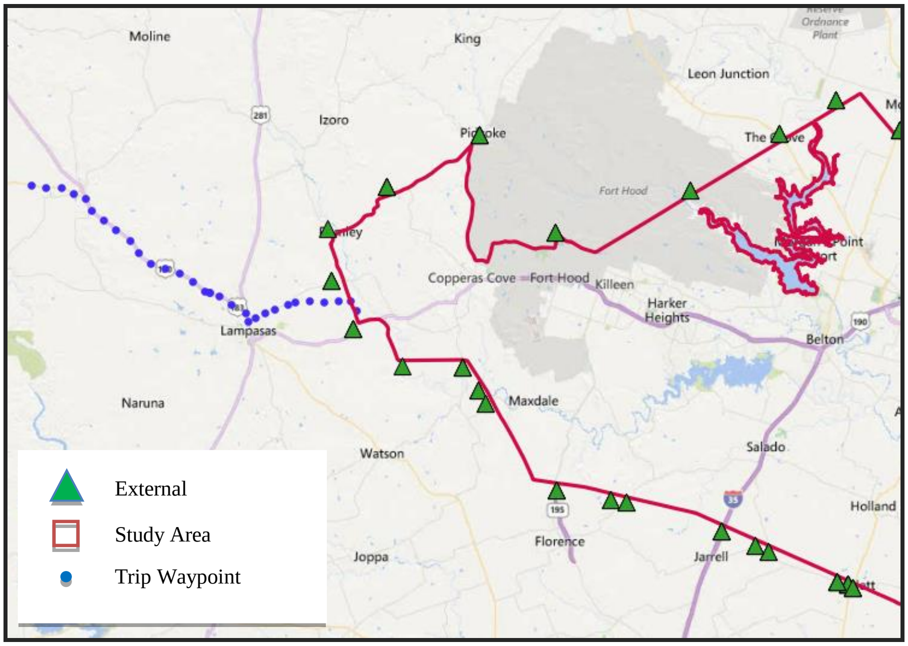


Figure 6. Unmatched External Trips.

DATA EXPANSION

Data expansion is the process of growing sample data to match a known population. Part of data expansion is calculating different weights representing each category of the population relative to the sample, and vice-versa. This process controls for biases in the sample data that may be over-represented. For external analysis, the basis for expansion is traffic counts, which represent the known population. There are many techniques for weighting data. This study uses a process known as Iterative Proportional Fitting (IPF). IPF is a procedure to fix values in a matrix to a known total. The process is iterative because it estimates initial weights. From the initial estimate, the resultant error is calculated, and the process then recalculates the weight using the error until it minimizes a specified minimum error. The IPF procedure seeks to minimize the weighted mean absolute percentage error (MAPE). This objective best controls for errors in higher volume roadways rather than large errors in smaller volume roadways.

Cellular Data

Cellular data from AirSage represents expanded results. AirSage expands cellular data by aggregating households in the cellular data, also known as household clusters, to census tracts. The ratio of household clusters to the total households in the census tract is the penetration rate of the cellular data. AirSage applies this ratio to each trip made by each household cluster. AirSage provides only expanded results, which the study team post-processes as part of the external analysis.

Figure 7 shows the locations that produce and attract the highest number of I-E/E-I trips per square mile within the KTMPO area.

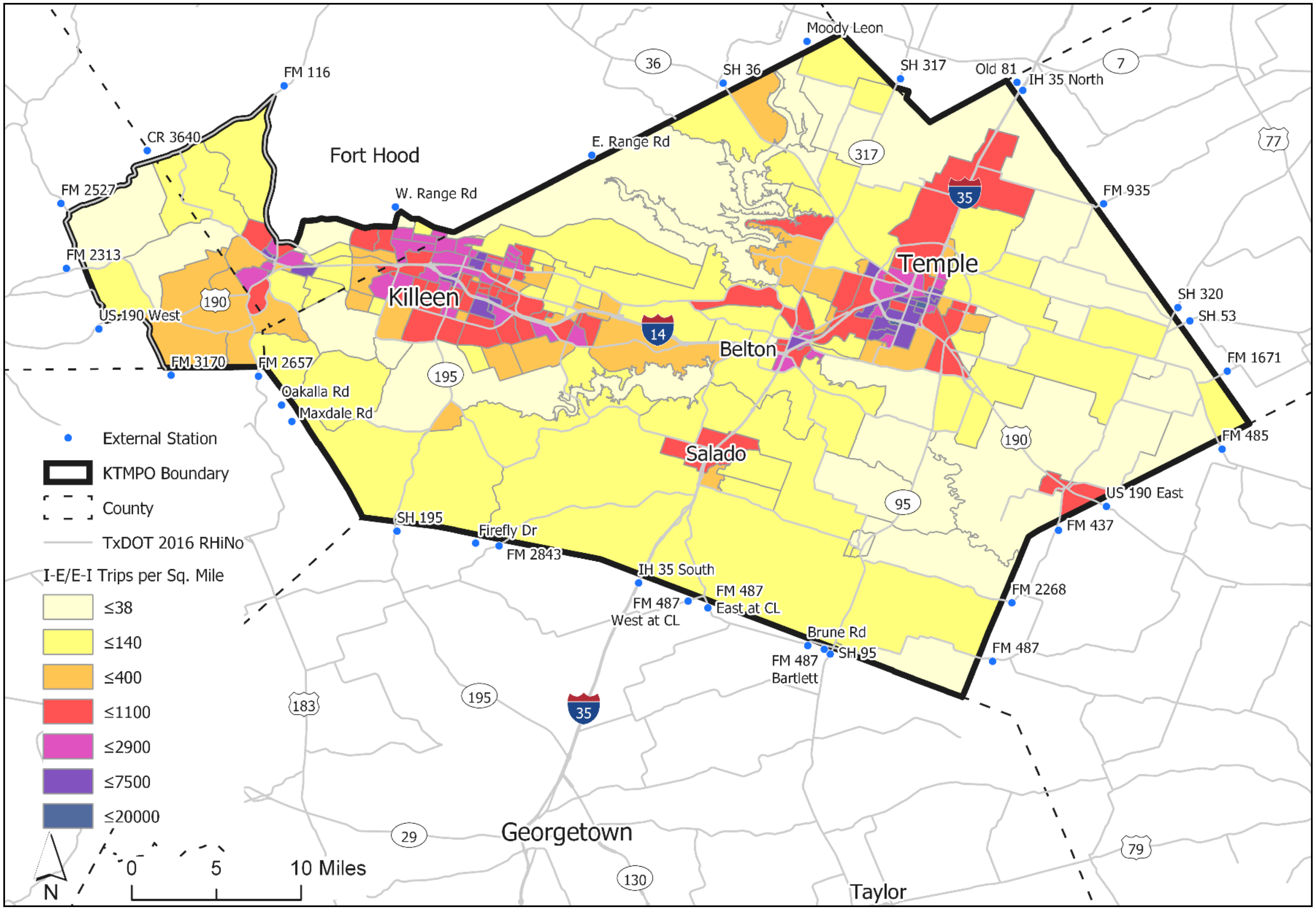


Figure 7. Cellular External-Local Trips per Square Mile by Zone.

Table 5 provides sample size distributions for each of the external roadways for cellular and GPS data, sorted by roadway total volume. In reviewing Table 5, it is important to keep in mind that cellular data represents an expanded sample, which is subject to error. Ultimately, what the cellular data provides is an approximate distribution of trips that would utilize the roadways. In later analysis, this distribution and overages are corrected for by fitting the data to the traffic counts. Presuming that cellular data is mostly passenger vehicles, it is approximately 213 percent of the traffic counts. Alternatively, presuming that cellular data is representative of trucks and passenger vehicles, it is approximately 156 percent of the traffic counts. AirSage and TTI researchers believe that cellular data has a significant passenger vehicle bias and the proportion of passenger to commercial vehicles in the data is unknown.

The sample size overages contained in the cellular data reflect that the cellular data is device trips and not vehicular trips. In light of this, the study team calculates correction factors that negatively expanded the cellular results relative to traffic counts of cars. These cellular correction factors apply only to the external trips in the cellular data.

In Table 5, the average GPS data is based on the total workdays during the study dates, which are 65 of a total of 88 days during the study period from February 9, 2015 to May 9, 2015

Table 5. Passive Data Sample Size Distributions.

Highway	Count of Cars	Count of Trucks	Total Count	Cellular	GPS Cars	GPS Trucks	% Cellular Sample to Cars	% Cellular Sample to Counts	% GPS Car to Counts (Avg)*	% GPS Trucks to Counts (Avg)*
IH 35 North	42,588	23,466	66,054	95,373	22,576	231,571	224%	144%	0.8%	15%
IH 35 South	37,091	20,315	57,406	58,572	21,239	206,481	158%	102%	0.9%	16%
US 190 West	9,477	802	10,279	20,424	1,359	6,184	216%	199%	0.2%	12%
SH 195	8,039	756	8,795	22,246	3,791	5,790	277%	253%	0.7%	12%
SH 317	5,760	394	6,154	8,357	1,333	3,669	145%	136%	0.4%	14%
US 190 East	5,003	853	5,856	16,135	1,542	9,262	323%	276%	0.5%	17%
SH 95	4,993	646	5,639	4,842	779	5,119	97%	86%	0.2%	12%
SH 36	4,688	923	5,611	11,488	1,211	6,642	245%	205%	0.4%	11%
FM 116	4,134	346	4,480	13,181	456	1,088	319%	294%	0.2%	5%
FM 2657	1,828	204	2,032	5,233	492	743	286%	258%	0.4%	6%
SH 320	1,519	246	1,765	3,541	135	719	233%	201%	0.1%	4%
FM 2313	1,374	270	1,644	2,386	185	311	174%	145%	0.2%	2%
FM 485	1,301	175	1,476	2,785	683	1,428	214%	189%	0.8%	13%
FM 487 Bartlett	1,291	124	1,415	0	81	436	0%	0%	0.1%	5%
FM 437	1,030	147	1,177	2,665	40	133	259%	226%	0.1%	1%
SH 53	933	128	1,061	4,606	164	569	494%	434%	0.3%	7%
Brune Rd	687	109	796	0	32	128	0%	0%	0.1%	2%
FM 487 East	616	159	775	4,203	70	244	682%	542%	0.2%	2%
FM 2843	533	98	631	2,193	108	224	411%	348%	0.3%	4%
FM 935	522	94	616	2,403	42	203	460%	390%	0.1%	3%
FM 487 West at CL	458	119	577	2,788	51	569	609%	483%	0.2%	7%
E. Range Rd	436	64	500	0	40	77	0%	0%	0.1%	2%
CR 3640	440	60	500	1,024	64	34	233%	205%	0.2%	1%
W. Range Rd	443	57	500	0	124	367	0%	0%	0.4%	10%
FM 2268	429	59	488	897	23	36	209%	184%	0.1%	1%
Firefly Dr	411	56	467	1,773	233	308	431%	380%	0.9%	8%
FM 487	271	49	320	260	38	162	96%	81%	0.2%	5%
Oakalla Rd	237	31	268	1,156	17	23	488%	431%	0.1%	1%
FM 1671	163	21	184	428	4	9	263%	233%	0.0%	1%
Moody Leon	153	24	177	1,581	24	107	1,033%	893%	0.2%	7%
FM 2527	147	20	167	1,033	17	13	703%	619%	0.2%	1%
Maxdale Rd	155	11	166	988	10	11	637%	595%	0.1%	2%
Old 81	143	20	163	0	245	1,445	0%	0%	2.6%	111%**
FM 3170	108	25	133	653	19	72	605%	491%	0.3%	4%
Total	**137,401**	**50,871**	**188,272**	**293,214**	**57,227**	**484,177**	**213%**	**156%**	**0.6%**	**15%**

*The percentage of GPS trips being greater than 100 percent is possibly a result of potential inaccuracies in the traffic count.

GPS Data

Table 5 shows that commercial vehicles dominate the GPS data. This makes it ideal for developing external truck O-D matrices. Overall, the average sample for trucks was 15 percent, which generally increases as the roadway volume increases. This matches expectations as commercial vehicles make up a greater portion of freeway traffic than non-freeway traffic. The results from cars was not as promising, with a very low average sample size of 0.6 percent. As with trucks, the average sample size increases with volume.

Expansion of the GPS data results use the total of all weekday trips observed in the data. This total compensates for empty cells within a trip O-D matrix. This phenomenon occurs due to the lack of observations of trips between an O-D pair on any given day. Using the total trips, rather than a daily average, allows for the capture of trips within these empty cells over time, rather than using a snapshot. As with cellular data, the expansion process uses IPF to calculate correction factors. The procedure expands trucks negatively, and cars positively.

O-D External Trip Matrices

The travel demand model directly applies the final O-D trip matrices. These matrices are calculated by developing a ratio of the sample trips to the expanded trips for each external roadway. This ratio is then applied to all O-D pairs associated with that external roadway. For trucks, the O-D matrix only uses the GPS truck results. For passenger vehicles, the resulting trip table is an average of the cellular data. However, the study team's goal is to refine this process using revealed biases and other results.

RESULTS

Table 6 and Table 7 show the expanded results for cellular and GPS data. Overall, the expansion process did not significantly alter the distribution of external trips, with less than 0.5 percent difference in E-E trips between the expanded and unexpanded result, for trucks and cars. However, the E-E as a percent of all trips did increase, which is a reflection of the unexpanded data's bias toward trucks.

Table 6. Cellular Area-Wide Weekday Expanded Trips.

	E-I	I-E	E-E	Total	Unexpanded E-E
Total Trips	54,140	54,140	13,058	121,338	29,660
Percent Total Trips	44.6%	44.6%	10.8%	100.0%	11.3

Table 7. GPS Area-Wide Weekday Expanded Trips.

	E-I	I-E	E-E	Total	Unexpanded E-E
GPS Car Trips	42,959	42,959	25,708	111,626	12,471
GPS Truck	10,298	10,298	15,366	35,962	149,408
Total Trips	53,257	53,257	41,074	147,588	161,879
Percent Car Trips*	38.5%	38.5%	23.0%	100.0%	27.9%
Percent Truck Trips*	28.6%	28.6%	42.7%	100.0%	44.6%
Percent Total Trips*	36.1%	36.1%	27.8%	100.0%	42.7%

*Total may not sum to 100 percent due to rounding.

Table 8 shows a summary comparison of the total and percentage of E-E trips between the cellular and GPS data. The percentage of E-E trips reflects the percent of trips in relation to I-E/E-I trips for the entire study area.

Table 8. Comparison of E-E Trips and Total External Traffic Percentages between GPS and Cellular Data.

Vehicle Category	GPS-Expanded		Cellular Expanded	
	E-E Trips	% E-E	E-E Trips	% E-E
Car	25,708	23.0%	13,058	10.8%
Truck	15,366	42.7%	NA	NA
Total	41,074	27.8%	13,058	10.8%

The percent of car E-E trips for the study area is 23 percent from GPS data and 10.8 percent from cellular data. One potential reason for this may be the sparse sample provided by the GPS data.

Expanded External Trips and Recommended Estimations

Table 9 provides expanded, aggregate external estimations as a percent of total trips by station. The recommended estimations are those of the greatest confidence. The external estimations for cars are developed from the cellular results and, where the cellular results were inconclusive, the GPS results were used. The external estimations for trucks use the GPS data, as this data is predominantly truck data.

Table 9 KTMPO Expanded External Trips.

Highway	Count of Cars*	Count of Trucks*	Total Count*	% I-E+E-I Cars (Recommended)	% E-E Cars (Recommended)	% I-E+E-I GPS Cars (For Reference)	% E-E GPS Cars (For Reference)	% I-E+E-I GPS Trucks (Recommended)	% E-E GPS Trucks (Recommended)
IH 35 North	42,588	23,466	66,054	73.3%	26.7%	50.6%	49.4%	37.8%	62.2%
IH 35 South	37,091	20,315	57,406	71.8%	28.2%	47.2%	52.8%	30.6%	69.4%
US 190 West	9,477	802	10,279	95.7%	4.3%	92.3%	7.7%	80.2%	19.8%
SH 195	8,039	756	8,795	100.0%	0.0%	98.8%	1.2%	94.7%	5.3%
SH 317	5,760	394	6,154	91.8%	8.2%	89.4%	10.6%	84.7%	15.3%
US 190 East	5,003	853	5,856	94.1%	5.9%	60.9%	39.1%	71.6%	28.4%
SH 95	4,993	646	5,639	81.1%	18.9%	65.3%	34.7%	52.8%	47.2%
SH 36	4,688	923	5,611	89.6%	10.4%	60.9%	39.1%	70.7%	29.4%
FM 116	4,134	346	4,480	94.1%	5.9%	82.0%	18.0%	71.5%	28.5%
FM 2657	1,828	204	2,032	95.7%	4.3%	96.0%	4.1%	88.6%	11.4%
SH 320	1,519	246	1,765	85.2%	14.8%	82.1%	17.9%	84.6%	15.5%
FM 2313	1,374	270	1,644	84.7%	15.3%	60.1%	39.9%	69.9%	30.1%
FM 485	1,301	175	1,476	79.6%	20.4%	86.3%	13.7%	76.1%	23.9%
FM 487 Bartlett**	1,291	124	1,415	81.4%	18.6%	81.4%	18.6%	43.5%	56.5%
FM 437	1,030	147	1,177	84.5%	15.5%	55.0%	45.1%	73.0%	27.0%
SH 53	933	128	1,061	97.2%	2.8%	86.5%	13.5%	90.5%	9.5%
Brune Rd**	687	109	796	81.3%	18.7%	81.3%	18.7%	58.9%	41.1%
FM 487 East at CL	616	159	775	100.0%	0.0%	84.1%	15.9%	76.0%	24.0%
FM 2843	533	98	631	51.5%	48.5%	59.4%	40.6%	51.0%	49.0%
FM 935	522	94	616	98.1%	1.9%	83.5%	16.5%	84.8%	15.2%
FM 487 West	458	119	577	100.0%	0.0%	82.5%	17.5%	81.0%	19.1%
W. Range Rd**	443	57	500	88.7%	11.3%	88.7%	11.3%	75.0%	25.0%
CR 3640	440	60	500	97.3%	2.7%	87.7%	12.3%	73.3%	26.7%
E. Range Rd**	436	64	500	97.7%	2.3%	97.7%	2.3%	93.8%	6.3%
FM 2268	429	59	488	98.6%	1.4%	61.2%	38.8%	37.1%	62.9%
Firefly Dr	411	56	467	100.0%	0.0%	97.6%	2.4%	87.1%	12.9%
FM 487	271	49	320	79.4%	20.6%	55.1%	44.9%	42.3%	57.7%
Oakalla Rd	237	31	268	100.0%	0.0%	100.0%	0.0%	87.5%	12.5%
FM 1671	163	21	184	84.0%	16.1%	24.7%	75.3%	63.6%	36.4%
Moody Leon	153	24	177	90.8%	9.2%	96.1%	4.0%	91.7%	8.3%
FM 2527	147	20	167	63.5%	36.5%	87.8%	12.2%	80.0%	20.0%
Maxdale Rd	155	11	166	98.7%	1.3%	89.7%	10.3%	83.3%	16.7%
Old 81**	143	20	163	68.1%	31.9%	68.1%	31.9%	55.6%	44.4%
FM 3170	108	25	133	50.0%	50.0%	83.6%	16.4%	69.2%	30.8%
Total Expanded Unique Trips**	**124,105**	**35,633**	**159,738**	**89.20%**	**10.8%**	**77.0%**	**23.0%**	**57.3%**	**42.7%**

**Unique trips are differentiated from total counts (e.g., the total count of cars is not the sum of the count of cars) because each E-E trip crosses two external count stations and is thus double counted. Accordingly, to develop the total unique vehicle count, the E-E trips at each station are divided by a factor of two.

**External estimation uses GPS results.