DAY-OF-WEEK, MONTH, AND SEASONAL DEMAND VARIATIONS: COMPARING FLOW ESTIMATES ACROSS NEW TRAVEL DATA SOURCES

Kentaro Mori
Graduate Research Assistant
Department of Civil, Architectural and Environmental Engineering
The University of Texas at Austin
mori@utexas.edu

Kara M. Kockelman, Ph.D., P.E. (corresponding author)
Dewitt Greer Centennial Professor of Transportation Engineering
Department of Civil, Architectural, and Environmental Engineering
The University of Texas at Austin
kkockelm@mail.utexas.edu
Tel: 512-471-0210

Submitted for presentation at BTR#6, April 2024

ABSTRACT
Transportation planners and engineers are increasingly interested in incorporating demand variations into travel models. Regression models are used here to predict and compare variations in permanent traffic recorder (PTR) counts along Texas highways to vehicle-kilometers traveled (VKT) inferred from INRIX's probe-vehicle data across days of the year. Results suggest INRIX data do not illuminate month-of-year variations in network use, but significant day-of-week differences are clear in both. Interestingly, INRIX appears to capture much more travel away from PTRs on Saturdays.

Keywords: Travel demand prediction, demand variations, seasonal and day of week effects, big data, traffic counts, VKT and VMT estimation

QUESTIONS
While travel demand models are calibrated to offer traffic forecasts for a single, "typical" weekday - and sometimes a typical weekend, actual travel varies substantially throughout the year. For example, Americans make more person-trips during spring and summer than in winter (Huang et al. 2024). Similarly, 2001-2017 NHTS data show Americans making 15-16% of person-trips on Fridays versus just 12% on Sundays, as shown in Table 1 (FHWA 2023). The latest (2022) NHTS shows Saturdays carrying the most person-trips (15.6%) and, together with Friday, the most VKT (16.0%), suggesting a post-pandemic shift in American's busiest day of travel.

Table 1. Percentages of Annual Person-Trips (PT) and VKT Occurring on Each Day-of-Week in NHTS

	2022		**2016/2017**		**2008/2009**		**2000/2001**	
Day-of-week	PTrips	HH VKT	PTrips	HH VKT	PTrips	HH VKT	PTrips	HH VKT
Sunday	11.7%	12.3%	12.4%	12.2%	12.2%	11.3%	12.8%	11.7%

Monday	14.4%	15.8%	14.1%	13.9%	14.2%	15.0%	13.9%	14.1%
Tuesday	14.4%	13.4%	14.5%	13.6%	14.8%	14.1%	14.0%	14.1%
Wednesday	14.7%	12.4%	14.8%	14.4%	14.8%	15.2%	14.7%	15.3%
Thursday	14.9%	13.8%	14.8%	17.0%	14.4%	15.5%	14.5%	14.9%
Friday	14.3%	16.4%	15.3%	15.4%	15.4%	15.7%	15.6%	16.5%
Saturday	15.6%	16.0%	14.1%	13.7%	14.3%	13.3%	14.6%	13.5%
Sample size	16,997 persons	7,893 households	264,234 persons	129,696 households	308,901 persons	150,147 households	160,758 persons	69,817 households

Note: All percentages are based on population-weighted values from each NHTS.

Annual average daily traffic (AADT) estimates across local and national roadway networks are used to inform infrastructure investments, validate travel demand models, evaluate road safety, and more (Gadda et al. 2007; Selby and Kockelman 2013). AADT values come from PTR stations' counts (in roughly 100 sites per US state) and short-period traffic counts (SPTC) from portable counters (roughly every 5 to 20 miles of state-managed highway). SPTC values are adjusted using day-of-week and month-of-year factors from relevant PTRs. Several U.S. states, including Vermont, Pennsylvania, and Washington, have published recent average seasonal factors (VTrans 2023; PennDOT 2022; WSDOT 2020).

Travel modelers are beginning to assess the impacts of demand variations across wide networks. Huang et al. (2024) simulated shared autonomous vehicle fleets in Austin, Texas across four seasons for both weekdays and weekends using demand adjustment factors calculated from the 2017 NHTS. They found network and fleet performance (including VKT, operator profits, and traveler wait times) differed significantly across seasons and days of the week. However, local demand patterns often deviate substantially from national averages. Moreover, due to the ever-shifting dynamics of travel, "big data", such as probe vehicle data derived from measurements pervasively obtained from instrumented vehicles, may be better suited for this task than travelr surveys (which are administered to less than 1 percent of the population every few or more years). This paper explores two datasets for capturing regional demand variations to address the associated questions:

1. What are the advantages and disadvantages of each data source in capturing travel demand variations across the year?
2. What are notable variation patterns?
3. Are the apparent demand variations consistent across data sources?

METHODS

Two datasets reveal how travel demand changes throughout the year in Texas, the USA's second most-populated and second physically largest state. Nearly 400 PTRs track traffic counts using loop detectors embedded in Texas highway pavements. This conventional dataset differs greatly from the trip tables (OD matrices) obtained from the Regional Integrated Transportation Information System (RITIS) Nextgen Trip Analytics interface (CATT Lab 2024). RITIS relies on INRIX's probes: "connected vehicles" (with in-vehicle GPS systems) and "location-based services" data (from smartphone applications).

From 2013 through 2022, the Texas Department of Transportation (TxDOT) maintained 398 distinct PTR stations (Figure 1). 214 of these stations classify vehicles into 13 categories. The number of stations fluctuates day to day, as local receivers or loops trip off unexpectedly. Of the 214, just 51 of these stations offer traffic counts for more than 90% of days in 2019 through 2022, and Figure 2 shows their total light-duty vehicle (LDV) counts, after imputing for missing values and outliers. Missing values and values over 5 standard deviations away from the mean in 2019, 2021, and 2022 were imputed as the average for the day of the week for the same month over the 3 years. However, counts from February 13-17, 2021, and February 3, 2022, were kept because the cause can be identified as winter storms Uri and Landon descending on Texas. For 2020, missing values were imputed using the same method but averaged using data from just 2020. 3,047 values (4.1%) were imputed out of 74,210 total data points used to derive Figure 2. The 4 years of demand patterns align well (with peaks on Fridays), excepting extraordinary events – like the COVID-19 pandemic and winter storms in 2021 and 2022.

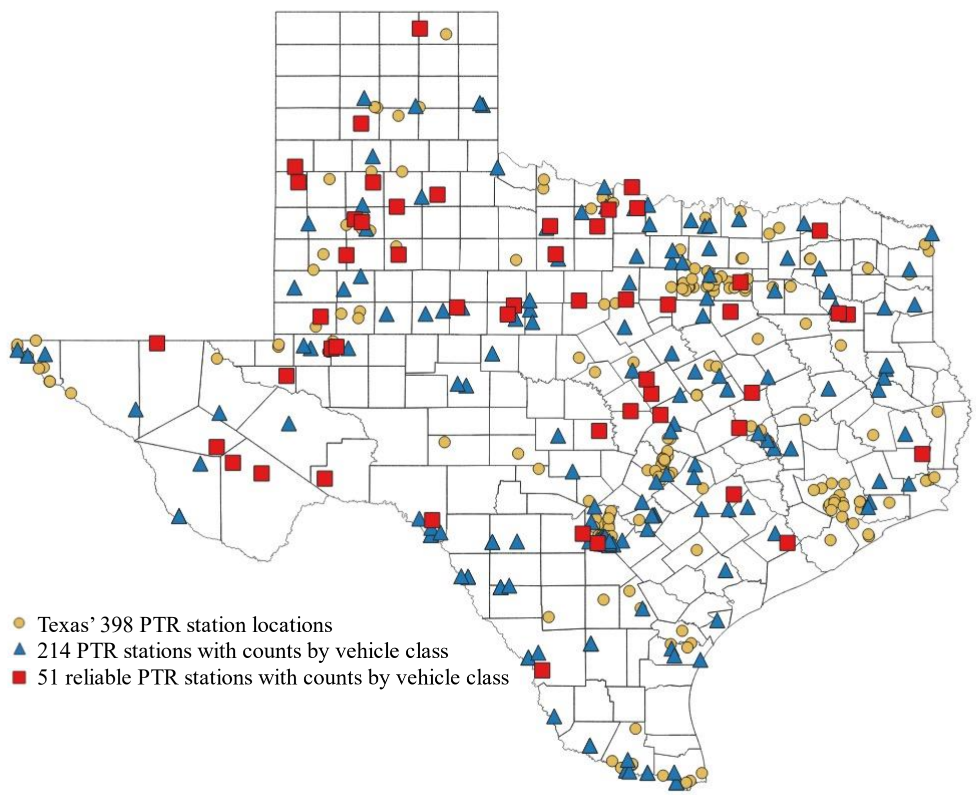

Figure 1. Texas' 398 PTR Station Locations (plotted on map of Texas' 254 counties) - with blue triangles used for 214 vehicle-class counting stations and red squares for the 51 reliable plotted stations

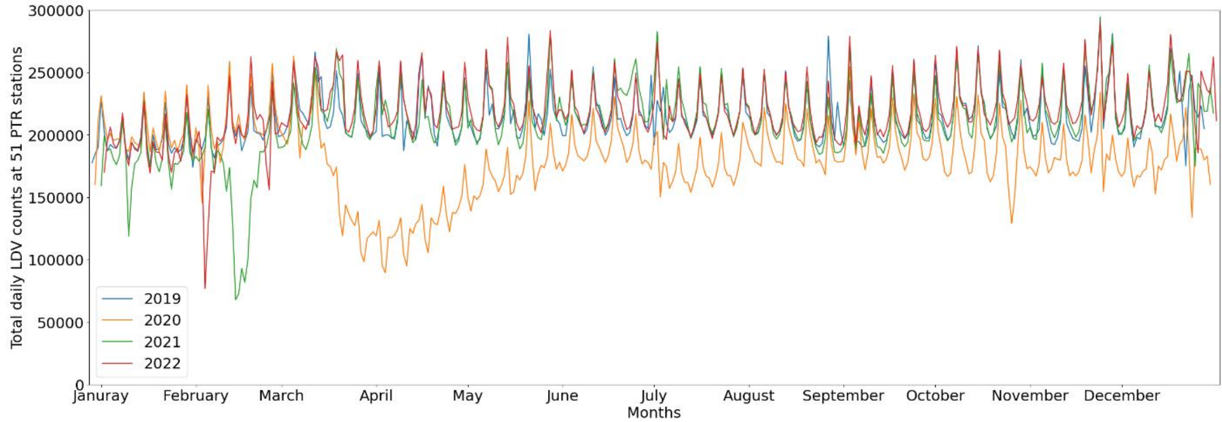

Figure 2. Total Daily LDV Traffic Counts across 178 PTR Stations from 2019 through 2022 (data shifted to correspond to the 2021 days of week)

RITIS's processed probe vehicle data were purchased by TxDOT (and then made available to the research team) for the spring and fall months of 2021 (February through April and September through November) and 2022 (February through April and August through October). Figure 3 shows the daily LDV VKT approximated using shortest-paths on the Statewide Analysis Model network. Unlike the PTR data, the INRIX/RITIS 2021 and 2022 demands are inconsistent and their oscillations are relatively irregular (though Friday versus Sunday peaks and valleys are largely visible).

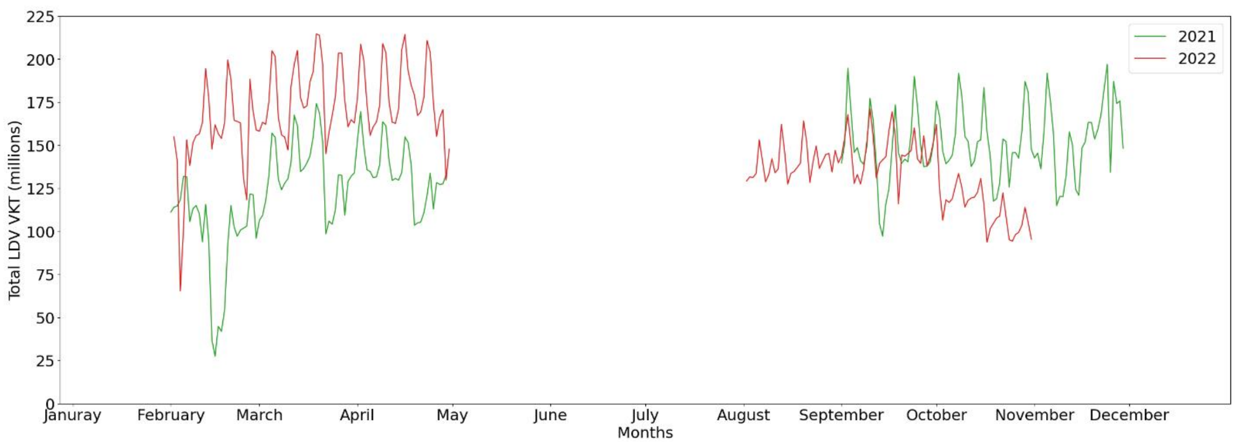

Figure 3. Total Daily LDV VKT in INRIX/RITIS 2021 and 2022 (data shifted to correspond to the 2021 days of week)

Log-linear ordinary least squares regression was used to evaluate and compare how demand variations can be explained in each dataset. The marginal effects of variables on Y are expressed as ratio changes, allowing for comparisons of coefficient estimates between the two models. The PTR regression model for daily LDV traffic counts by station (using data from 214 stations with counts by vehicle classes across 10 years) uses fixed effects to reflect PTR-specific levels of demand and unobserved heterogeneity (Hedges 1994). Explanatory variables include days of the week, month, year, and select holidays (and days before/after). The model for state-level VKT controls only for days of week and months of year, due to the limited number of observations (n = 357 INRIX/RITIS 24-hr trip tables, versus n = 386,132 daily counts across stations and years).

FINDINGS

Table 2 shows the PTR daily LDV count model results. On average (across sites), LDV counts vary by 27.5% across the week: falling 4.1% on Tuesdays (relative to the base day: Mondays) and rising 23.4% on Fridays (relative to Mondays). January is the least busy month, and June and July are the busiest (with counts roughly 13% above January's). Counts rose year after year, excepting the notable drop in 2020 (and recovery by 2022).

Table 2. PTR Daily LDV Traffic Count Fixed-Effect Log-Linear Regression Result

Variable	**Estimate**	**t-statistic**	**Percentage effect**
Constant	7.780	3844.4	-
Monday (base day)	0	-	-
Tuesday	-0.042	-29.5	-4.1
Wednesday	-0.032	-22.5	-3.1
Thursday	0.035	24.3	+3.6
Friday	0.210	144.5	+23.4
Saturday	0.084	58.2	+8.8
Sunday	0.018	12.7	+1.8
January (base month)	0	-	-
February	0.010	5.2	+1.0
March	0.120	64.9	+12.7
April	0.059	31.8	+6.1
May	0.103	55.3	+10.8
June	0.128	68.8	+13.7
July	0.132	70.5	+14.1
August	0.096	52.2	+10.1
September	0.066	35.5	+6.8
October	0.100	54.7	+10.5
November	0.095	48.3	+10.0
December	0.118	62.3	+12.5
Year 2013 (base year)	0	-	-
Year 2014	0.051	29.2	+5.2
Year 2015	0.112	63.6	+11.9
Year 2016	0.137	75.4	+14.7
Year 2017	0.155	85.9	+16.8
Year 2018	0.181	100.2	+19.8
Year 2019	0.211	118.4	+23.5
Year 2020	0.047	26.4	+4.8
Year 2021	0.187	103.2	+20.6
Year 2022	0.217	119.5	+24.2
Friday before Memorial Day	0.107	14.5	+11.3
Memorial Day	0.062	8.3	+6.4

July 3	0.111	13.3	+11.7
Fourth of July	-0.088	-11.8	-8.4
Friday before Labor Day	0.137	18.8	+14.7
Labor Day	0.137	18.6	+14.7
Tuesday before Thanksgiving	0.266	35.8	+30.5
Day before Thanksgiving	0.449	60.1	+56.7
Thanksgiving Day*	0.023	3.1	+2.3
Day after Thanksgiving	-0.049	-6.5	-4.8
Saturday after Thanksgiving	0.120	16.1	+12.7
Sunday after Thanksgiving	0.263	35.4	+30.1
December 23	0.202	27.0	+22.4
Christmas Eve (Dec. 24)	0.061	8.0	+6.3
Christmas (Dec. 25)	-0.184	-25.2	-16.8
December 26	0.131	16.3	+14.0
R-squared	0.200		
#Observations	n = 386,132 site-days		
#Entities	214 PTR sites		

** Thanksgiving is the third or fourth Thursday of November.*
Final column contains percentage effect, which is exp(β)-1 or approximately β if β is close to 0.

Table 3 provides regression results for the INRIX/RITIS-based estimates of daily LDV VKT. Although fewer parameters are statistically significant in the INRIX/RITIS-based VKT estimates (and no month-of-year effects are evident), the day-of-week variations are mostly consistent between the two datasets (Table 3 vs Table 2 results), with daily LDV use appearing steady from Sunday through Thursday and rising roughly 20% on Fridays. However, the INRIX/RITIS data suggest a 17.4% rise in LDV flows on Saturdays, rather than the 8.8% rise indicated by the PTR data. Adding a separate Saturday indicator for the two INRIX/RITIS data years (2021 and 2022) to the PTR model (Table 2) led to an even lesser increase of 7.3%, suggesting there may be substantial Saturday LDV driving away from PTR sites.

Table 3. INRIX/RITIS Daily LDV VKT (millions) Log-Linear Regression Result

Variable	Estimate	t-statistic	Percentage effect
Constant	4.463	151.5	-
Monday through Thursday + Sunday (base days)	0	-	-
Friday	0.186	7.9	+20.4
Saturday	0.160	6.7	+17.4
February 2021 (base month)	0	-	-
March 2021	0.378	9.5	+45.9
April 2021	0.379	9.3	+46.1
September 2021	0.473	11.7	+60.5
October 2021	0.499	12.5	+64.7
November 2021	0.512	12.6	+66.9
February 2022	0.511	12.4	+66.7

March 2022	0.672	16.8	+95.8
April 2022	0.658	16.2	+93.1
August 2022	0.432	10.8	+54.0
September 2022	0.469	11.6	+59.8
October 2022	0.199	4.9	+22.0
R-squared	0.618		
Adj R-squared	0.603		
#Observations	n = 357 days' trip tables		

Although log-linear regression was used for both PTR and INRIX/RITIS datasets to compare day of week, month of year, and holiday effects between the two differing measures of travel demand, unique models for each dataset were also explored. Table 4 shows the regression results of the daily LDV and total traffic counts at each PTR station standardized to z-scores using the mean and standard deviation at the station over the 10 years. The LDV count z-score model has a slightly higher R-squared value of 0.217 compared to 0.200 in the log-linear model presented in Table 2, and there are no large differences in the statistically significant variables. The model of total traffic counts at all 398 stations reveals some differences in traffic count variations (by day of week, month of year, and holidays) when predicting total traffic counts (at 398 stations) vs just LDV counts (at 214 stations). For example, although LDV counts on Saturdays and Sundays are higher than that of Mondays, the total traffic decreases on those days. Similarly, many days around holidays see increased LDV traffic but lower total traffic. This is because heavier commercial vehicles have demand patterns that are distinct from LDVs. Additionally, the R-squared value of the total traffic count model is substantially higher than the LDV count models, likely due to the elimination of the uncertainties surrounding vehicle type inference.

Table 4. PTR Daily Total Traffic Count Z-Score Regression Result

	Y = LDV count		**Y = Total count**	
Variable	**Estimate**	**t-statistic**	**Estimate**	**t-statistic**
Constant	-0.969	-142	-0.977	-250
Monday/Tuesday (base day)	0	-	0	-
Wednesday	-0.008	-1.70	0.086	33.1
Thursday	0.235	50.2	0.285	109
Friday	0.976	206	0.869	329
Saturday	0.362	77.4	-0.191	-73.2
Sunday	0.019	4.07	-0.766	-292
January (base month)	0	-	0	-
February	0.113	16.0	0.165	41.5
March	0.489	70.1	0.511	131
April	0.289	41.5	0.347	89.3
May	0.411	58.4	0.474	120
June	0.480	68.4	0.585	150
July	0.475	67.3	0.552	141
August	0.360	51.9	0.481	124

September	0.265	37.7	0.417	106
October	0.392	57.2	0.514	133
November	0.381	51.7	0.502	122
December	0.481	67.3	0.450	112
Year 2013 (base year)	0	-	0	-
Year 2014	0.133	21.2	0.155	43.0
Year 2015	0.334	53.5	0.361	99.6
Year 2016	0.429	67.0	0.518	145
Year 2017	0.464	73.3	0.629	176
Year 2018	0.540	85.9	0.729	205
Year 2019	0.651	105	0.870	244
Year 2020	0.032	5.16	0.225	62.4
Year 2021	0.548	87.9	0.823	230
Year 2022	0.688	110	0.915	256
Federal holiday*	-0.004	-0.311	-0.306	-49.0
Friday before Memorial Day	0.533	19.1	0.440	28.0
Memorial Day	0.157	5.21	-0.300	-17.7
July 3	0.499	16.0	0.303	17.6
Fourth of July	-0.391	-13.0	-0.748	-44.5
Friday before Labor Day	0.661	24.0	0.545	35.2
Labor Day	0.487	16.5	-0.071	-4.26
Tuesday before Thanksgiving	0.947	34.1	0.695	44.4
Day before Thanksgiving	2.004	71.2	1.320	83.5
Thanksgiving Day*	-0.065	-2.18	-1.077	-63.7
Day after Thanksgiving	-0.331	-11.9	-1.186	-74.9
Saturday after Thanksgiving	0.481	17.2	-0.081	-5.08
Sunday after Thanksgiving	1.313	47.0	0.905	57.4
December 23	0.848	30.1	0.372	23.4
Christmas Eve (Dec. 24)	0.131	4.56	-0.691	-42.0
Christmas (Dec. 25)	-0.883	-29.8	-1.630	-97.5
December 26	0.421	13.9	-0.200	-11.5
R-squared	0.217		0.358	
Adj R-squared	0.217		0.357	
#Observations	n = 386,172 (214 PTR sites)		n = 1,017,371 (398 PTR sites)	

**Federal holidays include New Year's Day, MLK Day, Presidents' Day, Memorial Day, Independence Day, Labor Day, Veterans Day, Thanksgiving, and Christmas. Thanksgiving is the third or fourth Thursday of November.*

Table 5 provides the best linear regression results for the INRIX/RITIS-based estimates of daily LDV VKT. Compared to the log-linear model presented in Table 3, the day-of-week variable for Thursday emerged as statistically significant, and the goodness-of-fit is slightly better (adjusted R-squared of 0.697 vs 0.603).

Table 5. INRIX/RITIS Daily LDV VKT (millions) Linear Regression Result

Variable	**Estimate**	**t-statistic**
Constant	89.23	28.0
Monday through Wednesday + Sunday (base days)	0	-
Thursday	5.14	2.0
Friday	28.23	11.0
Saturday	23.43	9.0
February 2021 (base month)	0	-
March 2021	37.39	8.7
April 2021	36.74	8.5
September 2021	50.50	11.7
October 2021	53.98	12.6
November 2021	56.46	13.0
February 2022	57.56	13.1
March 2022	81.93	19.2
April 2022	80.13	18.4
August 2022	43.44	10.2
September 2022	48.72	11.3
October 2022	14.48	3.36
R-squared	0.709	
Adj R-squared	0.697	
#Observations	n = 357 days' trip tables	

The R-squared values of the log-linear regressions are not substantially lower than those of the models tailored to the individual datasets. Therefore, this paper focuses on the comparison of the log-linear regression results.

While INRIX's big-data assembly should be able to out-perform traditional data sources in sensing day-to-day demand variations, these results suggest it is not yet the case. Apparently, INRIX's sampling rates can and do vary dramatically: day to day and month to month (Figure 3).

Although directly capturing statewide demand variations from trip tables was not possible, INRIX and RITIS offer tools for more local analyses, such as speeds and flows along specific corridors and to/from specific zones, essentially acting as virtual traffic recorders across vast regions. PTR sites are limited, often offline/not working, and cannot exceed capacity, while vehicles head to less congested links during peak times of the year or capacity-lowering events (like a crash upstream or downstream of the PTR). To make more sense of INRIX/RITIS values, INRIX's evolving daily vehicle sampling rates are needed.

ACKNOWLEDGEMENTS
The authors thank the Texas Department of Transportation for data access, the University of Maryland's Center for Advanced Transportation Technology Laboratory staff for answering

questions, undergraduate Connor Moening for data downloads, and Aditi Bhaskar for her editing (and administrative) support.

REFERENCES

Center For Advanced Transportation Technology (CATT Lab). (2024). *Regional Integrated Transportation Information System (RITIS).* https://ritis.org/

Federal Highway Administration (FHWA). (2023). *National Household Travel Survey*. Washington D.C.: U.S. Department of Transportation. https://nhts.ornl.gov/

Gadda, S., Kockelman, K. M., & Magoon, A. (2007). Estimates of AADT: Quantifying the uncertainty. Paper presented at the *World Conference on Transportation Research.*

Hedges, L. V. (1994). Fixed effects models. *The Handbook of Research Synthesis*, *285*, 299.

Huang, Y., Kockelman, K. M., & Gurumurthy, K. M. (2024). Agent-based simulations of shared automated vehicle operations: reflecting travel-party size, season and day-of-week demand variations. *Transportation*, https://doi.org/10.1007/s11116-023-10454-5

INRIX. (2024) *INRIX*. https://inrix.com/

Pennsylvania Department of Transportation (PennDOT). (2022). *2022 Pennsylvania Traffic Data.* https://gis.penndot.pa.gov/BPR_PDF_FILES/Documents/Traffic/Traffic_Information/Annual_Report/2022/2022_Traffic_Information_Report.pdf

Selby, B., & Kockelman, K. M. (2013). Spatial prediction of traffic levels in unmeasured locations: applications of universal kriging and geographically weighted regression. *Journal of Transport Geography*, *29*, 24-32. https://doi.org/10.1016/j.jtrangeo.2012.12.009

Vermont Agency of Transportation (VTrans). (2023). *Continuous Traffic Counter Report.* https://vtrans.vermont.gov/sites/aot/files/documents/Redbook%202022.pdf

Washington State Department of Transportation (WSDOT). (2020). Short Count Factoring Guide. https://wsdot.wa.gov/sites/default/files/2021-07/Short%20Count%20Factoring%20Guide%202020.pdf