LOCATING FAST CHARGING STATIONS FOR SAFE AND RELIABLE INTERCITY ELECTRIC VEHICLE TRAVEL IN WASHINGTON

(FINAL PROJECT REPORT)

by

Parastoo Jabbari
Moein Khaloei
Don MacKenzie
University of Washington, Department of Civil & Environmental Engineering

Project Sponsors

Pacific Northwest Transportation Consortium (PacTrans)
Washington State Department of Transportation (WSDOT)

for

Pacific Northwest Transportation Consortium (PacTrans)
USDOT University Transportation Center for Federal Region 10
University of Washington
More Hall 112, Box 352700
Seattle, WA 98195-2700

In cooperation with US Department of Transportation-Research and Innovative Technology Administration (RITA)

Disclaimer

The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the information presented herein. This document is disseminated under the sponsorship of the U.S. Department of Transportation's University Transportation Centers Program, in the interest of information exchange. The Pacific Northwest Transportation Consortium, the U.S. Government and matching sponsor assume no liability for the contents or use thereof.

Technical Report Documentation Page

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.	
4. Title and Subtitle Locating fast charging stations for safe and reliable intercity electric vehicle travel in Washington		**5. Report Date** **6. Performing Organization Code**	
7. Author(s) Parastoo Jabbari Moein Khaloei Don MacKenzie		**8. Performing Organization Report No.**	
9. Performing Organization Name and Address PacTrans Pacific Northwest Transportation Consortium University Transportation Center for Region 10 University of Washington More Hall 112 Seattle, WA 98195-2700		**10. Work Unit No. (TRAIS)** **11. Contract or Grant No.** DTRT-G-UTC40	
12. Sponsoring Organization Name and Address United States of America Department of Transportation Research and Innovative Technology Administration		**13. Type of Report and Period Covered** Research **14. Sponsoring Agency Code**	
15. Supplementary Notes Report uploaded at www.pacTrans.org			

16. Abstract

It is generally believed that direct current fast charging (DCFC) infrastructure has a significant effect on consumer adoption of electric vehicles (Evs), especially for long trips. To examine this relationship in Washington state, the expected adoption rate of EVs in each zip code was estimated by using a zero-inflated negative binomial model and then compared with the current adoption rate of EVs. The difference between expected and actual number of EVs in each zip code was then examined to determine whether it corresponded to the density of current DC fast charging stations in that zip code or adjacent zip codes. Results showed that in the northern and eastern parts of the state there are few DCFC charging stations and also lower than expected EV adoption rates. In contrast, the western part of the state showed higher EV adoption than predicted based on sociodemographic characteristics, along with a much greater concentration of DCFC stations. These results could be used for making policies related to future investments in this field.

17. Key Words		18. Distribution Statement No restrictions.	
19. Security Classification (of this report) Unclassified.	**20. Security Classification (of this page)** Unclassified.	**21. No. of Pages** 24	**22. Price** NA

Table of Contents

List of Figures

List of Tables

Executive Summary

Charging infrastructure has a significant effect on consumer adoption of electric vehicles (EVs). It is generally accepted that the relative attractiveness of EVs and other alternative fuel vehicles depends on several factors. These include up-front cost; operating costs, including fuel (electricity) and maintenance; range; refueling/recharging time; the availability of refueling infrastructure; environmental impacts; and government incentives. In the case of EVs, many of these factors are determined by the charging infrastructure: the number, type, locations, and pricing of charging stations. Previous research has generally indicated that to make plug-in electric vehicles (PEVs) more attractive to consumers, we should make charging opportunities ubiquitous, fast, and inexpensive.

The goal of this work was to establish a foundation for the data-driven prioritization of investments in direct current fast charging (DCFC) infrastructure along Washington's highways. The central question to be addressed in this work was: Where should Washington's DCFC system be expanded to ensure that current and future EV owners can travel safely over long distances without running out of charge? This project provided some initial answers and lays the groundwork for deeper investigations that will require the integration of several lines of research.

The number of plug-in hybrid electric vehicles (PHEVs) and battery electric vehicles (BEVs) registered in Washington increased by 70 percent between mid-2014 and the end of 2015. If Washington is to continue its EV leadership and achieve Governor Inslee's goal of 50,000 plug-in vehicles on the road by 2020, it will need to further increase EV adoption. Therefore, it is important to know which factors influence the adoption of EVs among residents of Washington. In this study, we sought to estimate the expected rates of EV adoption based on residents' demographics in each ZIP code and examined qualitatively whether there is a

relationship between EV adoption rates in zip codes in Washington state and the number of nearby DCFC charging infrastructures. First, we needed the number of EVs registered by each zip code, which was provided by the Washington State Department of Licensing and Department of Transportation. To set up a model based on residents' demographics, we used U.S. Census data.

To examine this relationship in Washington state, the expected adoption rate of EVs in each zip code was estimated by using a zero-inflated negative binomial model and then compared with the current adoption rate of EVs. The difference between expected and actual number of EVs in each zip code was then examined to determine whether it corresponded to the density of current DC fast charging stations in that zip code or adjacent zip codes.

Results showed that in the northern and eastern parts of the state there are few DCFC charging stations and also lower than expected EV adoption rates. In contrast, the western part of the state showed higher EV adoption than predicted based on sociodemographic characteristics, along with a much greater concentration of DCFC stations.

The final outcome of this project is a foundation for subsequent work on designing recharging networks and prioritizing public investments in DCFC in Washington.

Chapter 1 Introduction and Literature Review

With advances in technology, data collection methods for transportation-related studies have expanded. Data sets collected from smartphone navigation apps and GPS devices have opened a new window into travel behavior. Because of the extensive coverage of cellular networks, using smartphones to collect data about traffic and travel behavior is becoming a more cost-effective method of data collection than using loop detectors and surveys *(1)*. Previously, GPS data have been used to measure traffic flow speed *(1)*, perform transit analysis *(2)*, and estimate travel patterns (origin-destination flows) *(3, 4)*.

One of the areas in which this type of data can be useful is predicting the optimal locations of charging infrastructure for plug-in electric vehicles (PEVs). Because of the specific characteristics of PEVs, including limited and uncertain range, the high costs of charging infrastructure, and time-consuming charging events, the locations of charging stations influence PEV adoption and reliability.

Beyond location factors alone, the time required to "refuel" a PEV also influences user behavior. Direct current fast charging (DCFC) enables users to charge their vehicle in less than an hour *(5)*, a significantly shorter time than with level 1 and level 2 chargers. Hidrue et al. *(6)* showed that people value less time-consuming options for charging their PEVs. Therefore, increased DCFC infrastructure investments are essential to make PEVs attractive and PEV travel reliable, and to enable drivers to satisfy their desire for long-distance trips. However, high capital and fixed costs and low utilization of DCFC stations make such stations an unattractive investment. Therefore, it is crucial to find the appropriate locations for such investments, to serve as many people as possible while maintaining equitable and reliable access to public charging for PEV users.

Nicholas et al. *(7)* noted that “The demand for charging is derived from the demand for plug-in vehicles (PEVs) and from the travel patterns expected for these vehicles.” Studying the travel behavior of current and future PEV users can identify appropriate locations for DC fast charging infrastructures. Using GPS data, we can understand the types of long-distance trips that travelers (whatever their vehicle‘s powertrain) make. Then, on the basis of that, we can estimate the highest value locations for DC fast charging stations to enable PEV users to meet their travel needs without having to make major changes in them.

Even though GPS data provide us with solid information about traveler behavior, they need to go through a cleaning and calibration process before they can be used for further analysis and decision making. Some of the issues with raw GPS data are as follows:

- Not everyone uses GPS devices or navigation application while travelling.
- People may use GPS devices only for a portion of their trips.
- In areas with a weak signal, some trips may be missing or interrupted.

In order to use these data correctly, cleaning and calibrating the data accurately becomes crucial. Cleaning the data includes devising rules to address missing data and interruptions when they occur. Calibrating the data means adjusting the sample data to represent real travel patterns and control for differences in representativeness. However, a lack of supplementary information and ground-truth data is one of the major barriers to accurately calibrating GPS data.

In this report, we discuss the progress we made in calibrating and using GPS data for identifying high value locations for DC fast charging stations and modeling electric vehicle (EV) adoption rates at the ZIP code level. First, INRIX *(8)* provided us with one month of data on trips that started or ended in Washington state, including time, date, and coordinates of origin, destination, and waypoints along the route. This report presents the methodology and results for data cleaning, trip identification, and adjusting trip counts using a statewide gravity model to fill in gaps in demand for origins and destinations for which no trips were observed by INRIX.

Chapter 2 Methodology and Results

2.1 Data Cleaning and Trip Identification

One of the challenges of using GPS data, particularly for analyzing long-distance travel patterns, is trip identification. The presence of short stops within the dataset is one of the issues that must be addressed. To do this, we must first decide whether two consecutive trips of a user are actually two different trips or whether they are part of one longer trip *(8)*. We found in the literature that using rule-based algorithms is the most common way of identifying trips and making decisions about the stops *(8)*. For example, Wolf *(9)* used the dwell time of the stop as a rule to decide whether a trip actually ended or was part of a longer trip.

In addition to short stops, signal loss is another issue that must be addressed in the identification stage *(8)*. Tsui and Shalabi *(10)* discussed different types of signal loss based on the length and type of trip. On the basis of the data sets, time and distance thresholds must be assumed to determine whether or not signal loss between two consecutive trips happened.

To identify trips, two rules are necessary to decide whether two consecutive trips should be merged: 1) time threshold and 2) distance threshold. If both time and distance gaps between two consecutive trips of a user are small, there is a high likelihood that there is an interruption in signal. If so, we can assume that the two trips are actually part of one longer trip and neglect the gap. However, if these gaps are large for two consecutive trips, it means that the GPS device was off or interrupted for a long time. Since we do not have any information about stops and routes taken in between two trips, we must treat them as two different trips. If we have full GPS information but a time gap between two trips, it means that there was a stop between the trips and, depending on the length of the stop, we should decide whether to treat them as two different trips or as one long trip.

To determine the rules, we looked at the distribution of time and distance gaps between consecutive trips of a single user within our data. Given that we were ultimately interested in long distance trips of PEVs, we decided to be more flexible in choosing the time thresholds. The majority of the dwell times at stops were less than 15 minutes before the user started the second part of the trip. Assuming that if the driver's trip exceeded the range of the vehicle for a PEV, s/he would probably need to charge the vehicle, and since 15 minutes is not enough time to charge a PEV, we decided to use a threshold of 15 minutes to merge the trips with the stops between them. For the distance threshold, we decided that two consecutive trips needed to be less than 10 km apart to consider them part of a longer trip. For distance, ideally, we did not expect to see any distance gap between the ending and starting location of two consecutive trips. If we observed a distance gap, it meant that some portion or a whole trip was not recorded. If this gap exceeded 10 km, the uncertainty about the routes taken would be higher, and the trips would be treated as different trips.

In the one month of data from INRIX, 1,298,017 trips were connected to a device ID, which allowed for sequential trips to be identified. Out of those, 85,495 consecutive trips had time and distance gaps below the threshold of 15 minutes and 10 km and were therefore consolidated into longer trips. After these trips were joined and other trips were removed that had either their origin or destination outside of Washington, 846,632 trips remained in our data set.

2.2 Adjusting Trip Counts Using a Statewide Gravity Model

Washington state has 598 standard zip codes and therefore 357,604 OD pairs. As expected, for many of the ODs we had zero observations, which means that no trips were recorded between that origin and destination in our data. Even with a full month of trips from INRIX, there were still 330,258 zip code pairs for which no trip was recorded. However, observing zero trips does not mean that there is no demand for travel between a pair of zip codes; it simply means the demand is low and no trips were observed in the INRIX data set. Therefore, we decided to model the expected trip counts for OD pairs by using a gravity model based on the attractiveness of the destinations, productiveness of the origins, and the travel time between the origin and destination, along with a representativeness factor based on origin and destination counties. The logic behind incorporating a representativeness factor was that GPS data are not unbiased. The data collection was skewed toward regions where the adoption rate of navigation apps was higher. Also, no data were available on how representative the GPS samples were of the actual population. With the county representativeness factor, we aimed to capture some of these biases.

The gravity model is a simple method for calculating the number of trips, $T_{ij,}$ happening between two zones based on Newton's gravitational law. The simplest form of the gravity model is as follows *(11)*:

$$T_{ij} = CO_iD_jf(c_{ij}) \tag{2.1}$$

In this function, C is a balancing factor, Oi is the measure of attractiveness of origin zone i, and Dj is the measure of attractiveness for destination zone j. f(cij) is a cost function that is usually the function of distance or travel time. For this function, it was assumed that total trips attracted to a zone could be a measure of the attractiveness of that zone and total trips produced at a zone could be measure of the productiveness of that zone.

To determine the expected number of trips between ODs based on the number of trips presented in the data, we used a count model. Because the data set did not provide information on the home location of each trip, we could not know whether a given trip was a departure or return trip. Therefore, among trips happening between OD pairs, we had both departure and return trips. To capture that, we modified the above gravity model to the following:

$$T_{ijmn} = \frac{Pop_{im}^{\alpha} Pop_{jn}^{\beta}}{TT_{ijmn}^{\gamma}} \cdot \rho_m + \frac{Pop_{im}^{\beta} Pop_{jn}^{\alpha}}{TT_{ijmn}^{\gamma}} \cdot \rho_n \qquad i \neq j \tag{2.2}$$

where

T_{ijnm}: Number of trips starting from zip code i in county m and ending in zip code j in county n

Pop_{im} : Population of zip code i in county m

Pop_{jn} : Population of zip code j in county n

TT_{ijmn}: Travel time between zip code i in county m and zip code j in county m

α: Origin population coefficient

β: Destination population coefficient

γ: Travel time coefficient

ρ: Representativeness factor, which is specific for each home county

The first term of equation (2.2) represents trips departing from zip code i and traveling to zip code j, and the second term represents trips returning to zip code j after previously traveling from j to i. The representativeness factor is based on origin county (home county). In this model, we used travel time between ODs as our cost measure.To measure travel time between OD pairs we used the Google Distance Matrix API. For some of the zip codes, there was no driving option (e.g., zip codes in the middle of national parks or forests, etc); these observations were dropped from the modeling process.

We assumed that the trip counts between zip codes could be modeled with a negative binomial regression model. However, because our model was nonlinear (equation 2.2), we could not use many estimation software packages. We estimated the parameters by using an optimization algorithm to maximize the log-likelihood. The log-likelihood for nonlinear negative binomial models is as follows *(16)*:

$$L(\theta) = \sum_{i=1}^{n}[\log\{\Gamma(\phi + y_i)/\Gamma(y_i + 1)\Gamma(\phi)\} + y_i\log\{\mu_i/(\phi + \mu_i)\} + \phi\log\{\phi/(\phi + \mu_i)\}]. \tag{2.3}$$

In the above equation, $\Gamma(.)$ is gamma function and ϕ is dispersion parameter for the gamma distribution that will be estimated. y_i is the dependent variable for observation i, μ_i is the expected value of y_i (16). We used a genetic algorithm to find the unknown parameters of the model and ϕ.

The results of parameter estimates were as shown in table 2.1.

Table 2.1 Estimated values for parameters

Parameters	**Estimated value**
α(origin parameter)	0.97
β(destination parameter)	0.35
γ(travel time parameter)	2.801

The results for county representativeness factors are shown in table 2.2.

Table 2.2 Estimated county representativeness factors

County	***ρ***	**County**	***ρ***	**County**	***ρ***	**County**	***ρ***
Adams	0.4069	Franklin	0.1432	Lewis	0.1312	Snohomish	0.1124
Asotin	0.2615	Garfield	0.0297	Lincoln	0.2326	Spokane	0.2553
Benton	0.1314	Grant	0.2601	Mason	0.1910	Stevens	0.3101
Chelan	0.3781	Grays	0.3859	Okanogan	0.3892	Thurston	0.2429
Clallam	0.0496	Island	0.4350	Pacific	0.4871	Wahkiakum	0.3001
Clark	0.0635	Jefferson	0.2064	Pend Oreille	0.6041	Walla Walla	0.2931
Columbia	0.0222	King	0.1901	Pierce	0.1073	Whatcom	0.2617
Cowlitz	0.1375	Kitsap	0.4868	San Juan	0.3312	Whitman	0.3510
Douglas	0.3058	Kittitas	0.1867	Skagit	0.2690	Yakima	0.4911
Ferry	0.0262	Klickitat	0.3715	Skamania	0.6370		

Figure 2.1 shows the relationships of observed counts and fitted counts.

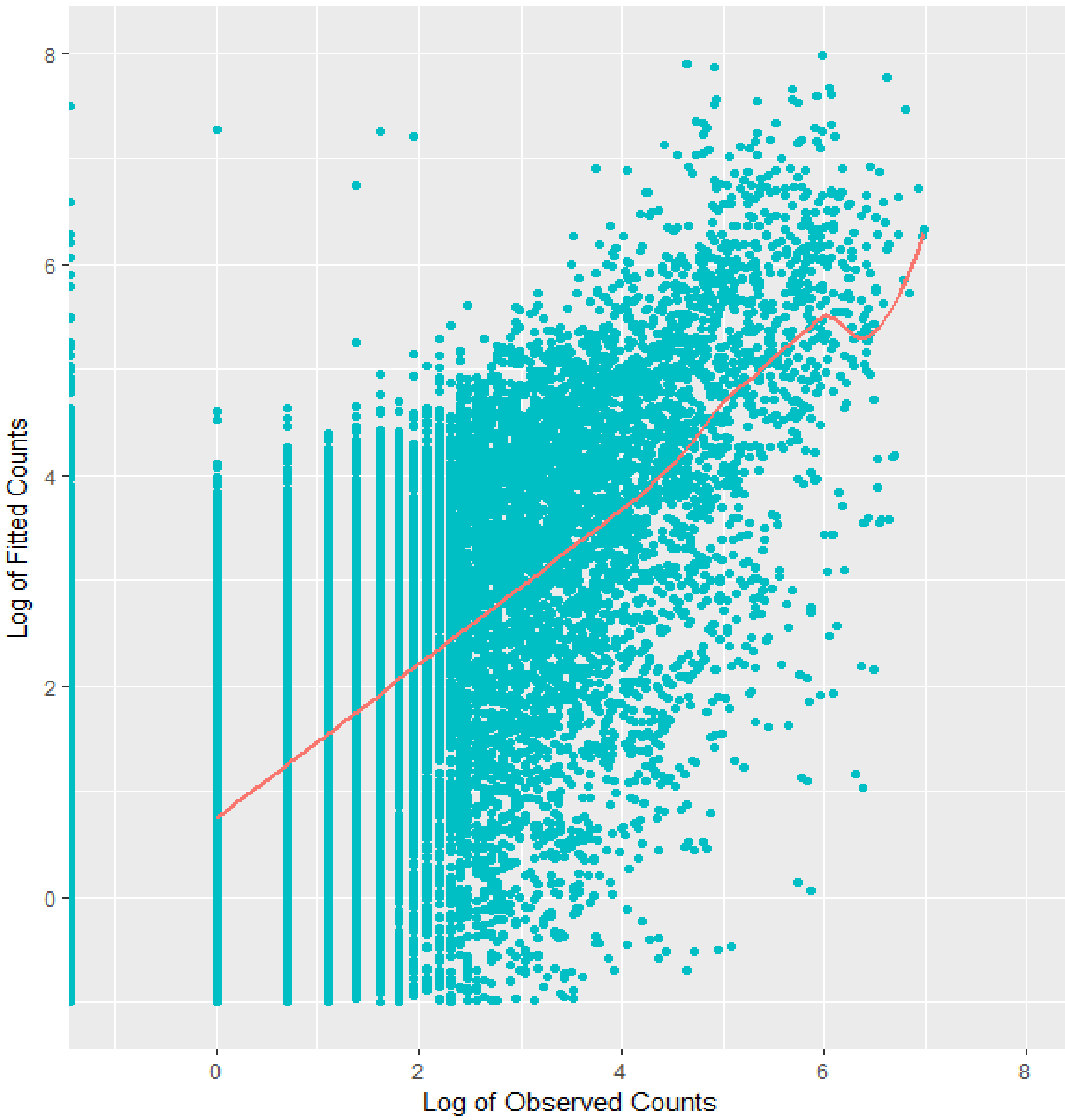

Figure 2.1 Log of observed counts from data vs. log of estimated values from nonlinear negative binomial count model.

In the estimated values, we no longer had zero counts, which helped with calibration (because if the counts were zero, the calibrated counts would be zero as well). Also, because we estimated representativeness factors, we captured some of the biases due to underrepresented areas and differing rates of adoption of mobile devices in different regions of the state.

2.3 Calibrating Data

The method we used to calibrate the data was called iterative screenline fitting (ISF) or matrix partitioning. This method uses traffic counts to derive an expansion factor to address systematic biases in the data *(17)*. This method does not require network assignment, which was advantageous because using another model (for network assignment) would have required another set of assumptions, introducing additional errors *(17)*.

To implement this method we needed to identify a screenline or screenlines. We started by dividing Washington state into two sub-regions with one screenline, as can be seen in figure 2.2.

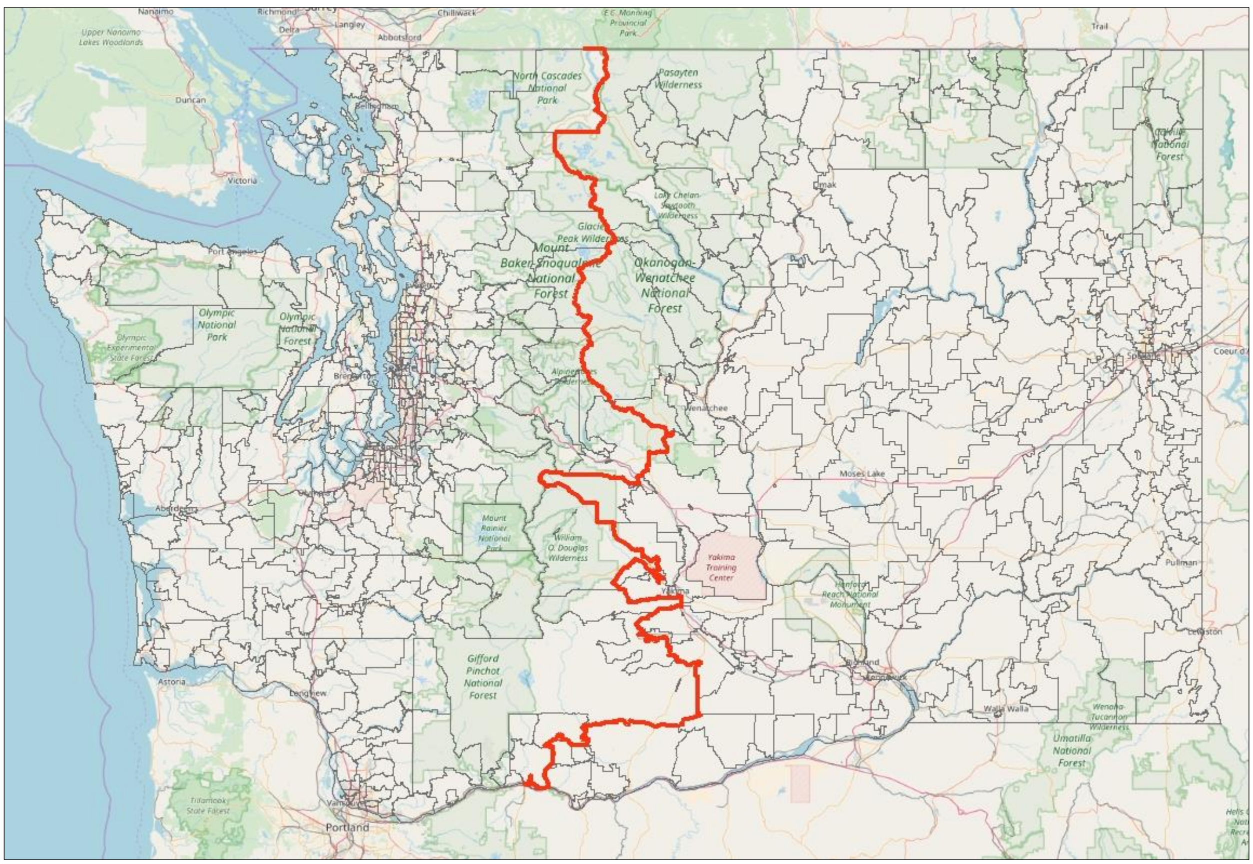

Figure 2.2 Snapshot of Washington state zip codes and a screenline for calibration

It was important that the screenline followed the geographical border of the studies' zones. Therefore, we followed zip code borders. We used AADT traffic counts where road networks crossed the screenline. Then we compared the sum of counts along the screenline with the number of trips in the OD matrix. By dividing the sum of the screenline counts by the total observed trips from one sub-region to the other, we were able to calculate the expansion factor. If we had had multiple screenlines, different expansion factors from each screen line would have disagreed with each other. To address this issue, we iterated on expansion factors for individual OD pairs to get values stabilized to the values that minimized errors versus screenline counts *(17)*. Here, we used one screenline mainly because we had an OD matrix from INRIX data that showed travel behavior, but in a smaller order of magnitude, and we needed to expand the data. Adding more screenlines would have complicated the calculations and would not have guaranteed smaller errors. Calibrated counts calculated as follows:

$$T'_{ijmn} = \vartheta \; T_{ijmn} \tag{2.4}$$

In the above equation, ϑ is the calibration factor, T_{ijmn} is trip counts from the gravity model, and T'_{ijmn} is the calibrated trip count from zip code i in county m to zip code j to county n.

2.4 EV OD Matrix

After estimating and calibrating the OD matrix for trips happening in Washington state, we expanded the analysis to estimate the potential demand for EV trips if EVs were unconstrained by charging. We calculated the EV OD matrix as follows:

$$t_{ijmn} = \frac{T'_{ijmn}}{2} \cdot \left(\frac{EV_i}{Cars_i}\right) + \frac{T'_{ijmn}}{2} \cdot \left(\frac{EV_j}{Cars_j}\right) \quad (2.5)$$

where t_{ijnm} is the number of EV trips from zip code i in county m to zip code j in county n. T'_{ijmn} is the number of calibrated car trips from zip code i in county m to zip code j in county n. EV_i is the number of EVs, and $Cars_i$ is the total number of cars registered in ZIP code i.

2.5 Visualization Tool

To better understand travel patterns and high value location for charging infrastructures, using Tableau® software, we developed an interactive map. Two snapshots of the map can be seen in figure 2.3. The purpose of this map is to show where the traffic was coming from and going to for each segment of the road. For example, the top right panel in figure 2.3 shows where the traffic passing a selected point was coming from or going to, as a subset of all the trips observed in the top left panel. The lower panel displays the traffic count for the segment. A working draft of the tool can be found here:

https://public.tableau.com/profile/parasto.jabbari#!/vizhome/EVODs/MonthlyEVTripsVisualization

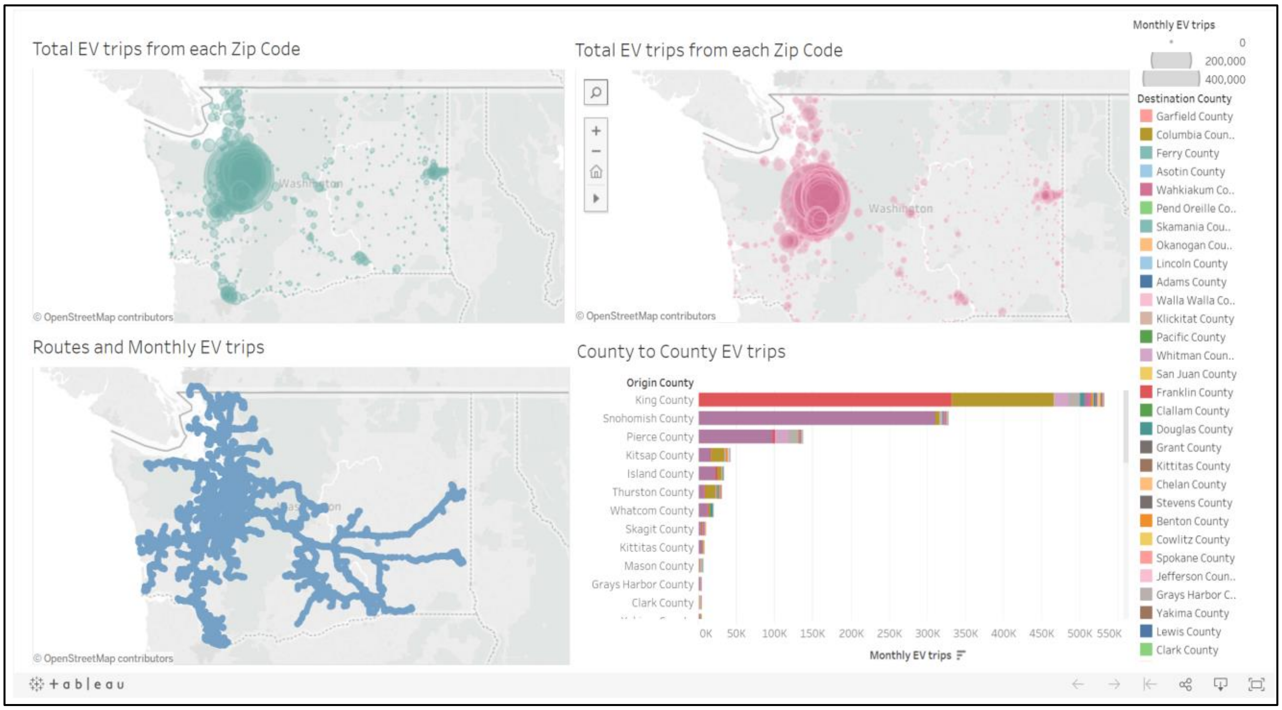

Figure 2.3 Snapshot of statewide visualization tool

<u>2.6 Adoption Model</u>

The number of plug-in hybrid electric vehicles (PHEVs) and battery electric vehicles (BEVs) registered in Washington increased by 70 percent between mid-2014 and the end of 2015. If Washington is to continue its EV leadership and achieve Governor Inslee's goal of 50,000 plug-in vehicles on the road by 2020, it will need to further increase EV adoption. Therefore, it is important to know which factors influence the adoption of EVs among residents of Washington.

This section describes how we estimated the expected rates of EV adoption based on residents' demographics in each ZIP code and examined qualitatively whether there is a relationship between EV adoption rates of zip codes in Washington state and the number of nearby DCFC charging infrastructures. First, we needed the number of EVs registered by each zip code, which was provided by the Washington State Department of Licensing and Department of Transportation. To set up a model based on residents' demographics, we used U.S. Census data *(18)*. The data included socio-economic characteristics of households (e.g., employment per household) or individuals (e.g., percentage of residents with a bachelor degree or higher) in each zip code, as well as the geographical attributes (e.g., area) of each zip code. For further study, the geographic coordinates of DCFC stations were obtained from the U.S. Department of Energy's Alternative Fuels Data Center *(19)*.

To estimate the expected rate of EV adoption in each zip code, we used a zero-inflated count model, allowing for spatial correlation among contiguous zip codes. Thus, we estimated a logit model for the probability of EV adoption being zero/not zero in each zip code, and a count model for the EV adoption rate in zip codes with non-zero adoption.

In spatial models, as shown in equation 2.6, the error term was assumed to be made up of a spatially weighted errors vector ($\lambda \mathbf{W}^* \boldsymbol{\varepsilon}$), as well as a vector of independent and identically distributed (IID) errors, **u**.

$$Y = \beta X + \boldsymbol{\varepsilon}, \quad \textit{where} \quad \boldsymbol{\varepsilon} = \underbrace{\lambda \mathbf{W}^* \boldsymbol{\varepsilon}}_{\text{Spatial error component}} + \underbrace{\mathbf{u}}_{\text{IID error component}} \tag{2.6}$$

To model spatial correlation, neighborhoods had to be defined among the study areas. Using a binary neighbors criterion, in which regions were either listed as neighbors or not, the function created an n-by-n weights matrix with values given by the coding scheme style chosen. Here, we used a 0/1 coding scheme to define neighbors as 1 and otherwise as 0. Two areas are often considered neighbors if they share a common boundary. Here, a queen adjacency criterion was used, i.e., two areas were considered neighbors if they shared at least a point along their borders. Figure 2.4 shows the queen adjacency criterion in comparison with other types.

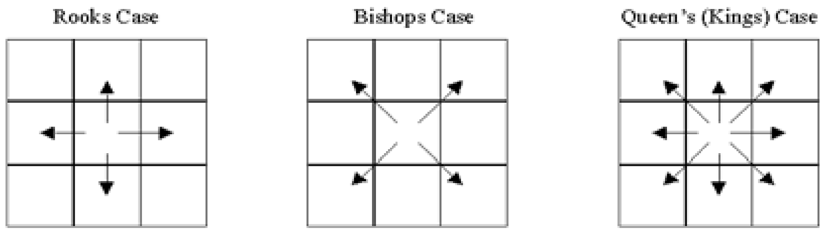

Figure 2.4 Adjacency criteria

The final model set up by R-INLA is presented in table 2.3. The dependent variable is the number of EVs registered in the zip code.

As indicated in table 2.3, all the predictors had a statistically significant relationship to the response variable. In general, we found that the EV adoption rate increases as the number of well-educated residents increases in each zip code. The population density of a zip code has a negative effect on the rate of EV adoption. The EV adoption rate in a zip code rises as the percentage of high income households increases in that zip code. Employment per household positively affects the rate of EV adoption in each zip code. The percentage of children (under 18 years old) in each zip code has a negative effect on the EV adoption rate. Interestingly, a higher average income makes it less likely that any EVs will be adopted (per the zero-inflated model), but a larger fraction of high-income households increases the numbers of EVs in adopting areas (per the negative binomial model). We rephrased the output from table 2.3 to show the count model and logit model in tables 2.4 and 2.5, respectively.

Table 2.3 Spatial zero-inflated negative binomial model of EV adoption rates in Washington state

Count model coefficients (negative binomial with log link)

Variable	**estimate**	**Std. Error**	**z value**	**Pr(>|z|)**
(Intercept)	-5.65	3.64e-02	-155.04	< 2e-16 ***
Log (population)	1.00 ˟			
% residents with bachelor degree or Higher	2.26e-02	2.80e-03	4.67	3.06e-06 ***
Population density	-1.00e-04	<0.0001	-2.16	3.1e-02 *
% households with annual income > $100k	2.14e-02	3.40e-03	5.51	3.59e-08 ***
Average # of workers per household	8.07e-01	1.34e-01	6.54	6.02e-11 ***
% children in zip code	-5.21e-02	5.80e-03	-5.67	1.42e-08 ***
Log(theta)	0.58	0.08	7.16	7.96e-13***
Variance (σ^2) of spatial error component	0.25	0.11		
Variance (σ^2) of IID error component	0.04	0.05		

Zero-inflation model coefficients (binomial with logit link)

Variable	**estimate**	**Std. Error**	**z value**	**Pr(>|z|)**
Intercept	1.71	0.62	2.78	5.51e-03 **
Mean income in zip code	-0.02	0.01	-3.15	1.63e-03**
% children in zip code	-0.10	0.02	-4.53	5.85e-06 ***
LL d.f.	-2673 10			

*: significance at α=0.10. **: significance at α=0.05. ***: significance at α=0.01.
˟ coefficient on log(population) was fixed to 1.00

Table 2.4 Specification of final model (count part)

	Count	intercept	X_1: population	X_2: % residents with bachelor degree or Higher	X_3: Population density (people per sq km)	X_4: % households with annual income > $100k	X_5: Average # of workers per household	X_6: % children in zip code
model	**Y=**	**-5.65**	**$+\log(X_1)$**	**$+2.26e\text{-}02*X_2$**	**$-1.00e\text{-}04*X_3$**	**$+2.14e\text{-}02*X_4$**	**$+8.07e\text{-}01*X_5$**	**$-5.21e\text{-}02*X_6$**
Mean	156.00		11952.00	27.20	528.0	21.10	1.10	19.30
Median	19.00		4861.00	22.90	40.60	18.30	1.10	19.60

Table 2.5 Specification of final model (logit part)

	Utility (Zero inflation)	intercept	Z_1: Mean income/1000	Z_2: % children in zip code
model	**U=**	**+1.71**	**$-0.02*Z_1$**	**$-0.10*Z_2$**
Mean			71.0	19.30
Median			65.5	19.60

Next, we used the fitted model to estimate the expected rate of EV adoption in each zip code and then compared these expected rates with the actual adoption numbers. In this way, we identified the zip codes in which actual adoption rate was higher or lower than the estimated adoption rate. Figure 2.5 shows areas, shaded in green, where actual EV adoption exceeded the predicted option. Likewise, grey areas are those where actual EV adoption fell short of the levels expected based on sociodemographic characteristics. Furthermore, the positions of existing DCFC charging stations are shown by red points to explore whether such differences were due to an inadequate number of fast charging stations in the surrounding region. It is worth noting that our data may not support a precise investigation of the effects of the availability of DCFC charging stations on EV adoption rates in Washington state, and such an investigation needs a more detailed and precise survey.

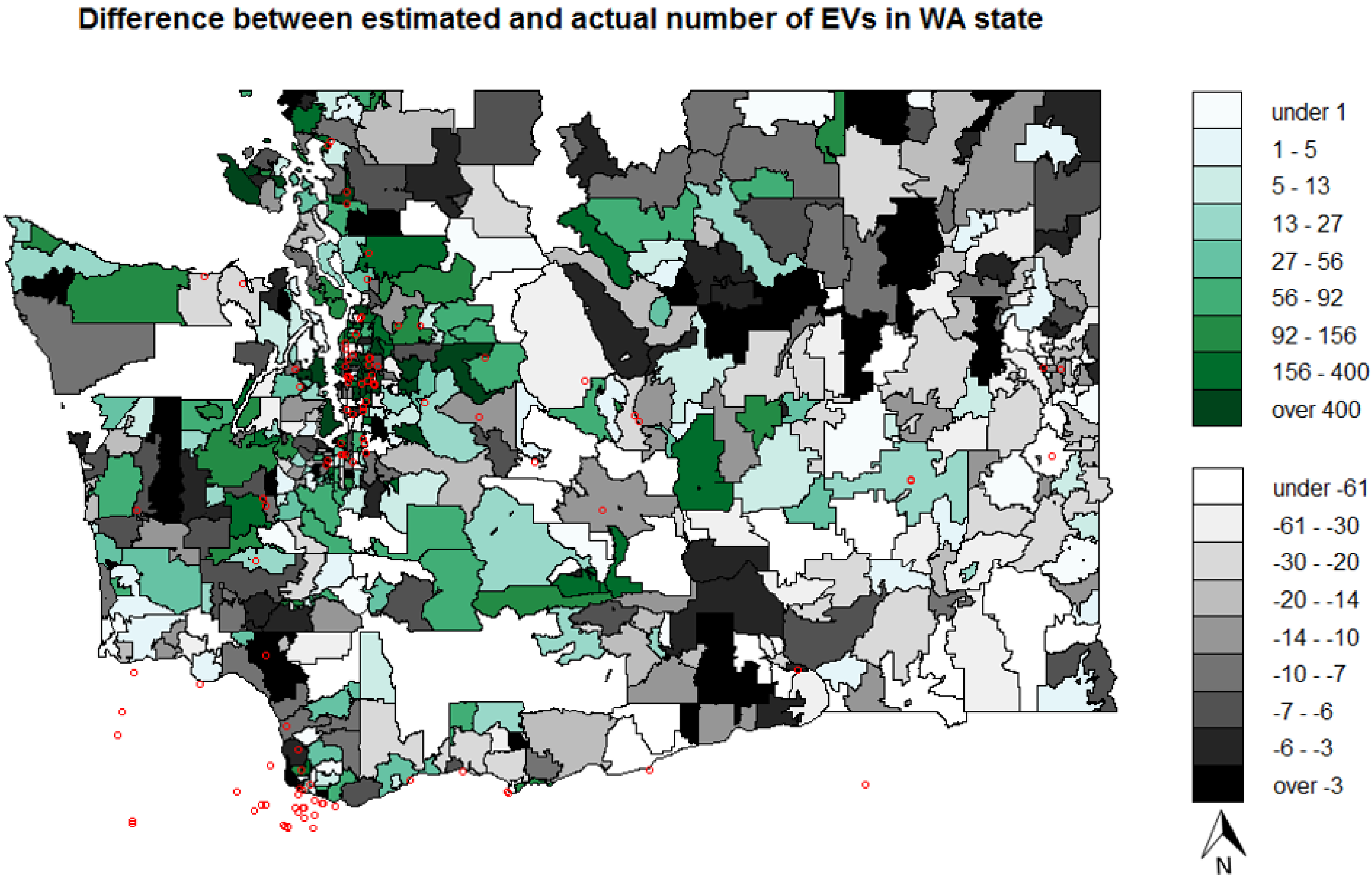

Figure 2.5 Zip codes of Washington state with actual EV adoption rates higher and lower than estimated

A comparison of figure 2.6 with figure 2.7 shows that in the northern and eastern parts of the state there are few DCFC charging stations and also lower than expected EV adoption rates. In contrast, the western part of the state shows higher EV adoption rates than predicted based on sociodemographic characteristics, along with a much greater concentration of DCFC stations.

On the local scale, if we compare figure 2.6 with figure 2.7, we can observe that in the central areas where DCFC stations are highly concentrated, there are more green zip codes with higher than expected EV adoption rates (figure 2.8). This increase is rather obvious, especially along the north-south I-5 corridor. While there are a few grey zip codes with fewer than expected EV adoption rates, they represent small differences between estimated and actual EV adoption rates.

We also looked specifically at the Puget Sound region. In figure 2.9 and 2.10, the difference between estimated and actual EV adoption rate is presented for Puget Sound area.

Altogether, it seems that the zip codes contiguous to DCFC charging stations are more likely to have higher rates of EV adoption. From the figures we can conclude that the EV adoption rate has the potential to be improved in the north, east and southeast parts of the state in which actual EV adoption rates, although lower, are very close to the expected rates. According to the predictors of the EV adoption model, we should consider the factors affecting the EV adoption in policies related to investments in this field. Education, income, population density, employment per household, and percentage of children in each zip code are the factors related to the rate of EV adoption in Washington state.

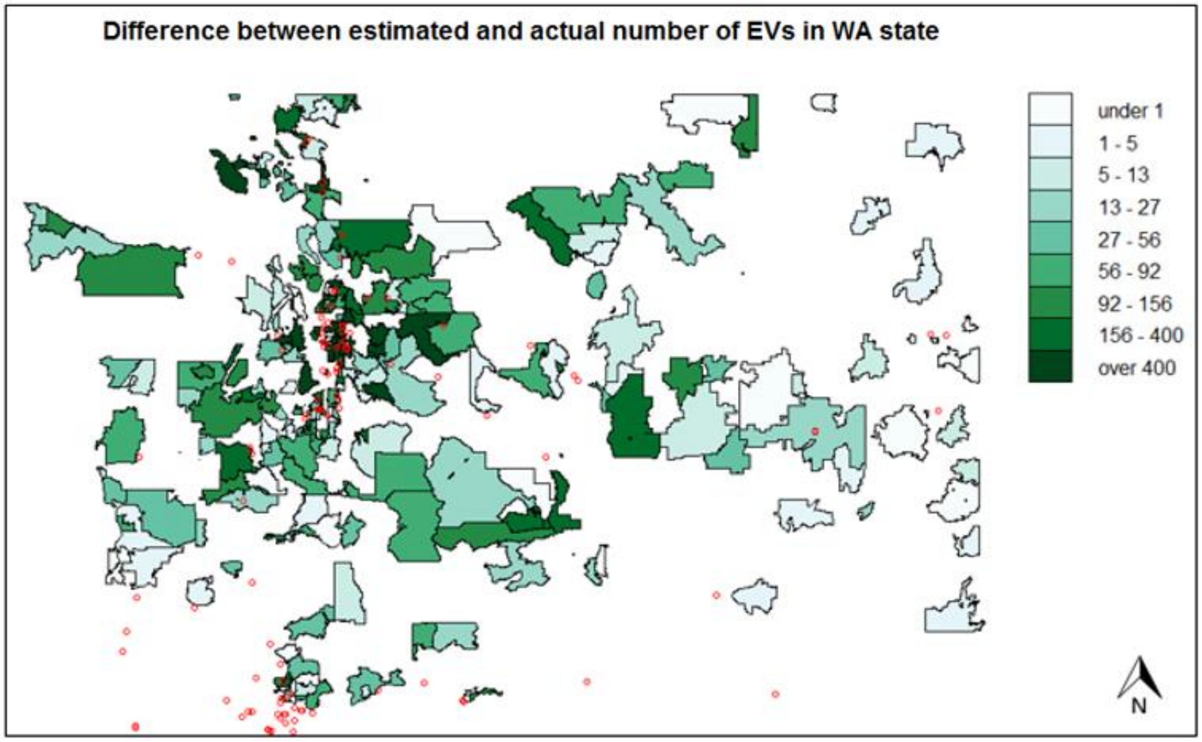

Figure 2.6 Zip codes of Washington state with actual EV adoption rates higher than estimated

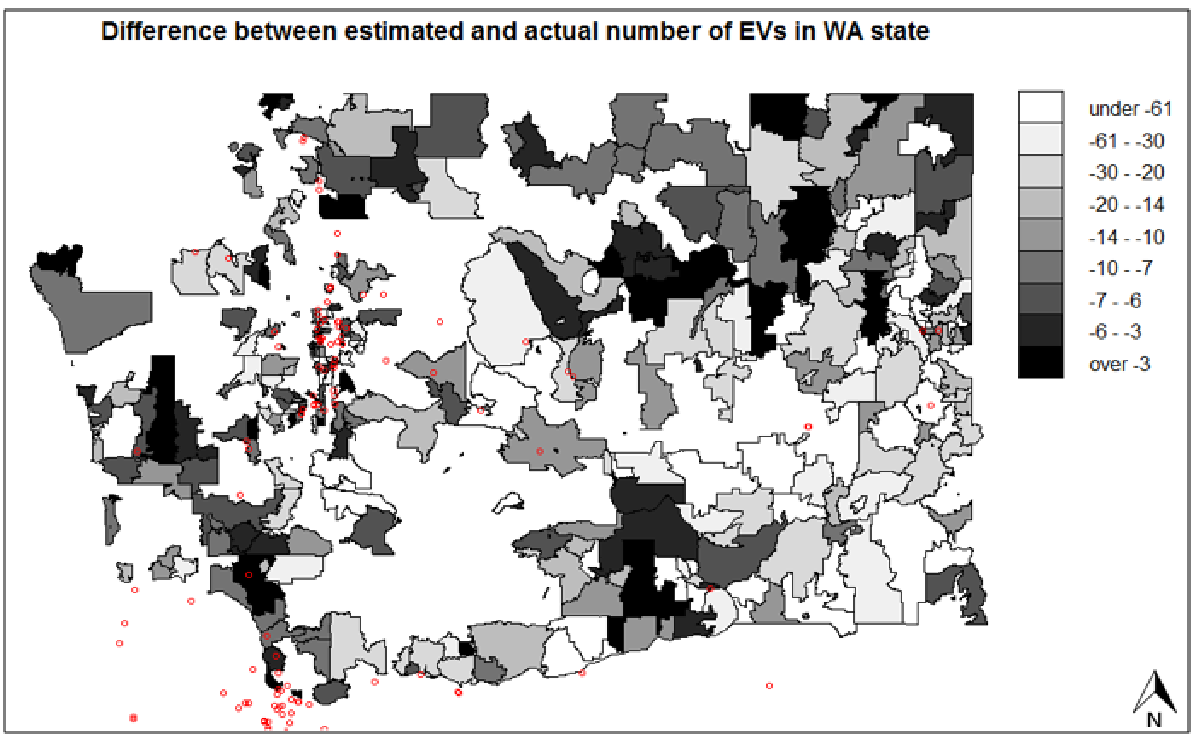

Figure 2.7 Zip codes of Washington state with actual EV adoption rates lower than estimated

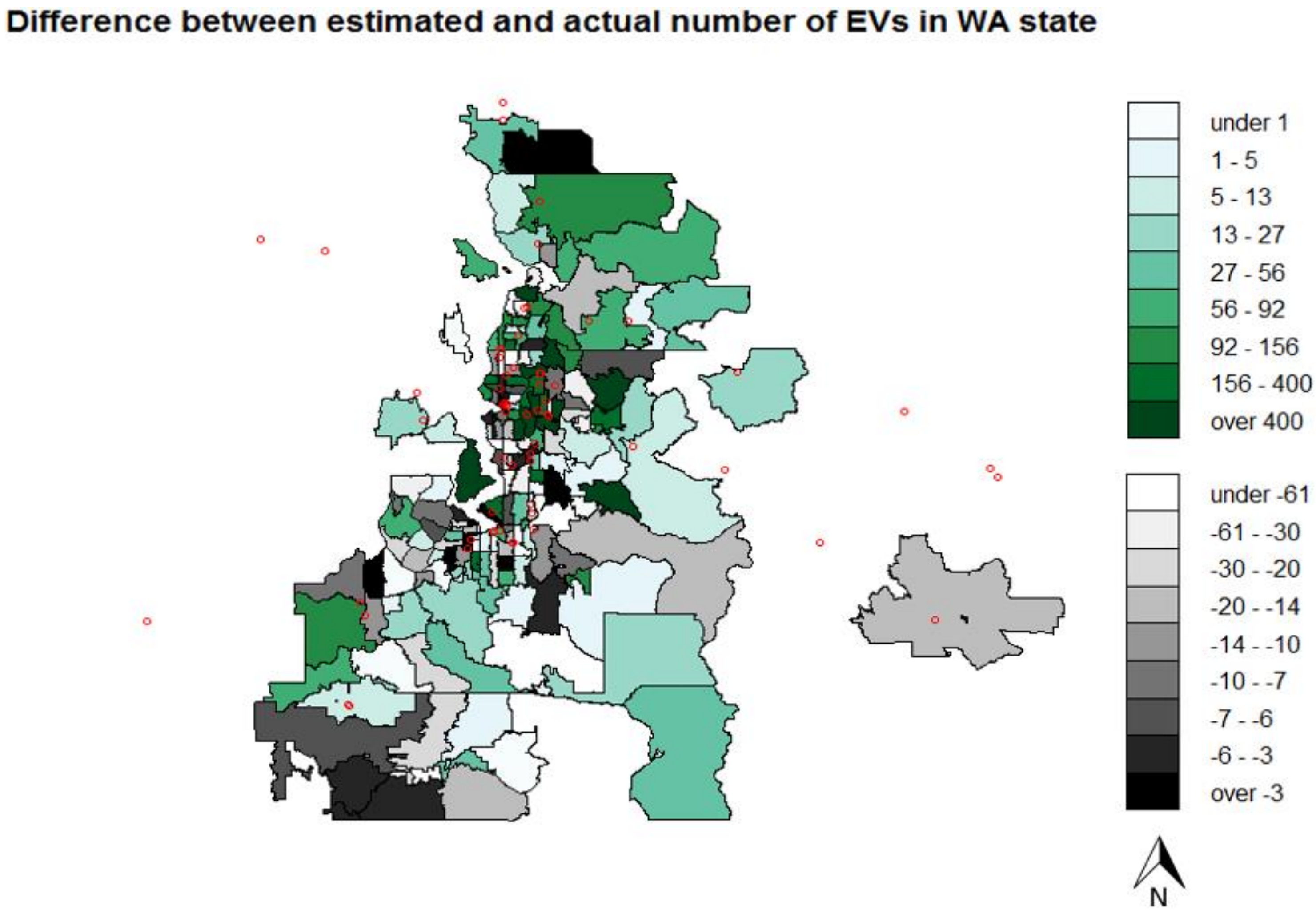

Figure 2.8. Zip codes of Puget Sound with actual EV adoption rates higher and lower than estimated

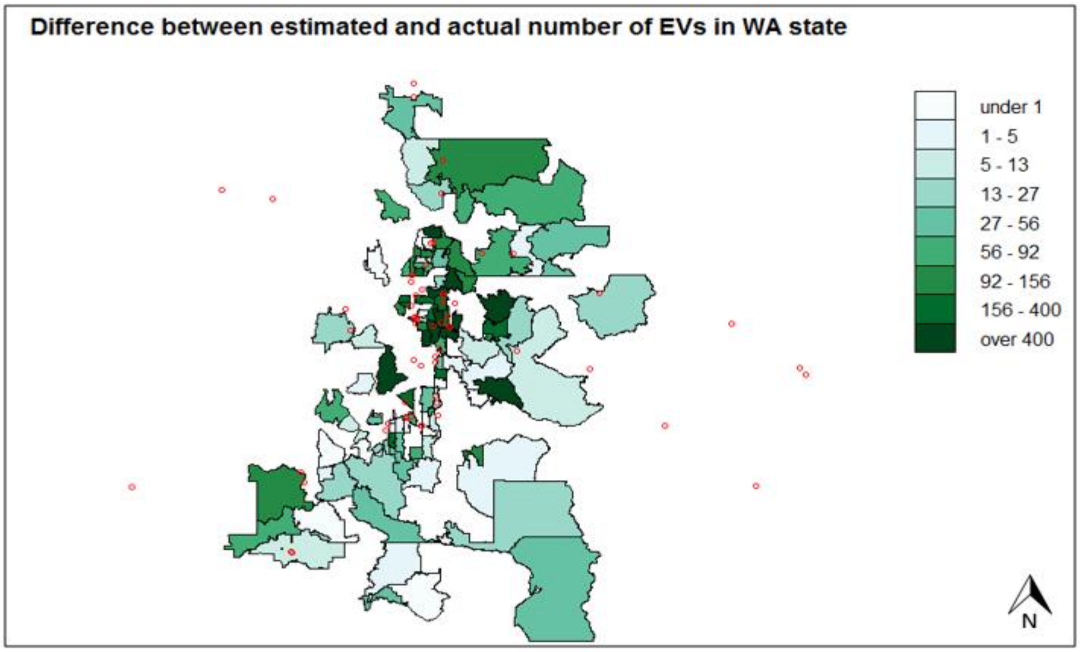

Figure 2.9 Zip codes of Puget Sound with actual EV adoption rates higher than estimated

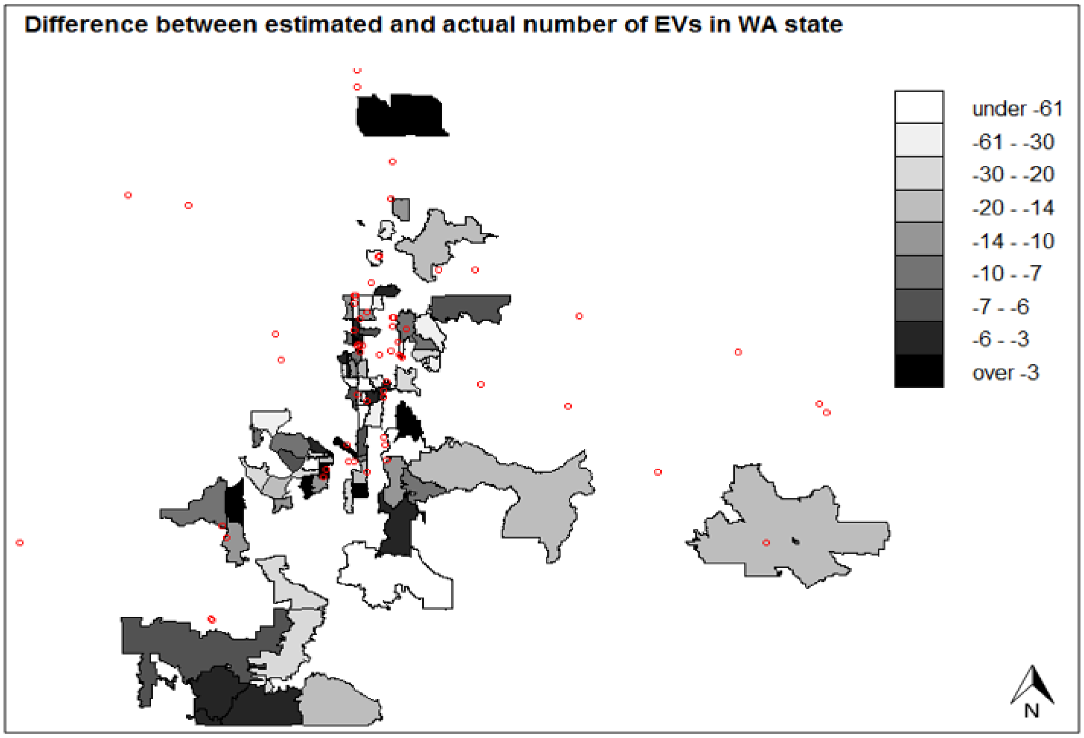

Figure 2.10 Zip codes of Puget Sound with actual EV adoption rates lower than estimated

References

1. Herrera, Juan C., et al. Evaluation of traffic data obtained via GPS-enabled mobile phones: The Mobile Century field experiment. *Transportation Research Part C: Emerging* Technologies 18, No 4, 2010, pp. 568-583.

2. Berlingerio, Michele, et al. AllAboard: a system for exploring urban mobility and optimizing public transport using cellphone data. Joint European Conference on Machine Learning and Knowledge Discovery in Databases. Springer Berlin Heidelberg, 2013.

3. Calabrese, Francesco, et al. Estimating Origin-Destination flows using opportunistically collected mobile phone location data from one million users in Boston Metropolitan Area. IEEE Pervasive Computing 99, 2011.

4. Wang, Pu, et al. Understanding road usage patterns in urban areas. arXiv preprint arXiv:1212.5327, 2012.

5. The EV Project: Characterize the demand and energy characteristics of direct current fast chargers, http://www.avt.inl.gov/, June 2015.

6. Hidrue, Michael K., et al. "Willingness to pay for electric vehicles and their attributes." Resource and energy economics 33.3, 2011, pp. 686-705.

7. Nicholas, Michael A., et al. DC Fast as the Only Public Charging Option? Scenario Testing from GPS-Tracked Vehicles. Transportation Research Board 91st Annual Meeting. No. 12-2997, 2012.

8. INRIX, http://inrix.com/

9. Shen, Li, and Peter R. Stopher. Review of GPS travel survey and GPS data-processing methods. Transport Reviews 34.3, 2014, pp. 316-334.

10. Wolf, Jean Louise. Using GPS data loggers to replace travel diaries in the collection of travel data. Diss. School of Civil and Environmental Engineering, Georgia Institute of Technology, 2000.

11. Tsui, Sheung, and Amer Shalaby. Enhanced system for link and mode identification for personal travel surveys based on global positioning systems. Transportation Research Record: Journal of the Transportation Research Board 1972, 2006, pp. 38-45. https://doi.org/10.3141/1972-07.

12. Mathew, Tom V., and KV Krishna Rao. Trip distribution. Introduction to transportation engineering, 2006.

13. Alexander, Lauren, et al. Origin–destination trips by purpose and time of day inferred from mobile phone data. Transportation Research Part C: Emerging Technologies 58, 2015, pp. 240-250.

14. StreetLight Insight calibration feature: Use Local Data to Estimate Vehicle Counts, Laura Schewel, 2016, http://blog.streetlightdata.com/estimate-vehicle-counts.

15. Soetaert, Karline, Karel Van den Meersche, and Dick Van Oevelen. Package limSolve, solving linear inverse models in R.

16. Svetliza, Carolina F., and Gilberto A. Paula. "Diagnostics in nonlinear negative binomial models." Communications in Statistics-Theory and Methods 32.6 (2003): 1227-1250.

17. Bernardin, Vincent L, Sadrsadat, Review of methods for data validation and expansion of passively collected origin-destination data, Transportation Research Board 2018.

18. https://factfinder.census.gov/faces/nav/jsf/pages/index.xhtml

19. https://www.afdc.energy.gov/data_download