Report No. UT-19.25

HERE Data Validation

Prepared For:

Utah Department of Transportation
Research & Innovation Division

Submitted By:

RSG, Inc.

Authored By:

Robert Chamberlin, P.E., PTOE
Nima Haghighi, Ph.D.
Kiavash Fayyaz, Ph.D.

Final Report

December 2019

RESEARCH

Keeping Utah Moving

Utah Department of Transportation - Research Division
4501 South 2700 West - P.O. Box 148410 - SLC,UT 84114-8410

DISCLAIMER

The authors alone are responsible for the preparation and accuracy of the information, data, analysis, discussions, recommendations, and conclusions presented herein. The contents do not necessarily reflect the views, opinions, endorsements, or policies of the Utah Department of Transportation or the U.S. Department of Transportation. The Utah Department of Transportation makes no representation or warranty of any kind and assumes no liability, therefore.

ACKNOWLEDGMENTS

The authors would like to acknowledge the Utah Department of Transportation (UDOT) for funding this research, and the following individuals from UDOT serving on the Technical Advisory Committee (TAC) for helping to guide the research:

- Kelly Njord (UDOT)
- Jamie Mackey (UDOT)
- Grant Farnsworth (UDOT)

The authors would also like to acknowledge Shana Lindsey from UDOT Research who served as Project Manager.

TECHNICAL REPORT ABSTRACT

<table>
<tr><td>1. Report No.
UT- 19.25</td><td>2. Government Accession No.
N/A</td><td colspan="2">3. Recipient's Catalog No.
N/A</td></tr>
<tr><td colspan="2" rowspan="2">4. Title and Subtitle
HERE Data Validation</td><td colspan="2">5. Report Date December 2019</td></tr>
<tr><td colspan="2">6. Performing Organization Code</td></tr>
<tr><td colspan="2">7. Author(s)
Robert Chamberlin, Nima Haghighi and Kiavash Fayyaz</td><td colspan="2">8. Performing Organization Report No.</td></tr>
<tr><td colspan="2" rowspan="2">9. Performing Organization Name and Address
Resource System Group
307 W 200 S, Suite 2004
Salt Lake City, Utah 84101</td><td colspan="2">10. Work Unit No.
5H08266H</td></tr>
<tr><td colspan="2">11. Contract or Grant No.
19-8240</td></tr>
<tr><td colspan="2" rowspan="2">12. Sponsoring Agency Name and Address
Utah Department of Transportation
4501 South 2700 West
P.O. Box 148410
Salt Lake City, UT 84114-8410</td><td colspan="2">13. Type of Report & Period
Final
Feb 2013 to Feb 2015</td></tr>
<tr><td colspan="2">14. Sponsoring Agency Code
PIC No. UT18.703</td></tr>
<tr><td colspan="4">15. Supplementary Notes
Prepared in cooperation with the Utah Department of Transportation and the U.S. Department of Transportation, Federal Highway Administration</td></tr>
<tr><td colspan="4">16. Abstract
UDOT has purchased a contract from HERE for providing travel-time information for Utah highways. HERE travel-time estimates are primarily derived from GPS data obtained from in-vehicle navigation devices.
In this research, UDOT is interested in learning how the travel-time estimates from HERE compare to another source of travel-time data, obtained from the Google Directions Application Program Interface (API). The source of this travel-time information is from users of Google Maps directions who, by virtue of using the direction service, agree to provide Google with their geographic position. Google aggregates this data to produce travel-time estimates for highway segments.
Both sources of travel-time information utilize an algorithm for estimating travel times that is partially based on real-time data and partially based on other factors that are not publicized by either service.
The research analyzed 59 highway segments and compared travel-time estimates from January 21 to February 25, 2019. The research finds that, under normal operations, travel times from HERE and Google are very comparable. The research focused on specific highway segments and times when a crash was reported on I-15. A total of 113 crash incidents were analyzed to determine how travel times from the two services differed under the potentially congested conditions occurring during the clearance or resolution of a crash incident. In 71% of these cases, travel times from Google were significantly lower than those reported from HERE.</td></tr>
<tr><td colspan="2">17. Key Words
HERE, Google Directions API, passive data, travel time estimation</td><td colspan="1">18. Distribution Statement
Not restricted. Available through:
UDOT Research Division
4501 South 2700 West
P.O. Box 148410
Salt Lake City, UT 84114-8410
www.udot.utah.gov/go/research</td><td rowspan="2">23. Registrant's Seal

N/A</td></tr>
<tr><td>19. Security Classification (of this report)

Unclassified</td><td>20. Security Classification (of this page)

Unclassified</td><td>21. No. of Pages
53
22. Price
N/A</td></tr>
</table>

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ACRONYMS

NPMRDS National Performance Management Research Data Set

UDOT Utah Department of Transportation

EXECUTIVE SUMMARY

UDOT has purchased a contract from HERE for providing travel-time information for Utah highways. HERE travel-time estimates are primarily derived from GPS data obtained from in-vehicle navigation devices.

In this research, UDOT is interested in learning how the travel-time estimates from HERE compare to another source of travel-time data, obtained from the Google Directions Application Program Interface (API). The source of this travel-time information is from users of Google Maps directions who, by virtue of using the direction service, agree to provide Google with their geographic position. Google aggregates this data to produce travel-time estimates for highway segments.

Both sources of travel-time information utilize an algorithm for estimating travel times that is partially based on real-time data and partially based on other factors that are not publicized by either service.

The research analyzed 59 highway segments and compared travel-time estimates for a 36-day period, from January 21 to February 25, 2019. The research finds that, under normal operations, travel times from HERE and Google are very comparable. An analysis of travel-time estimates for work zones also confirms that the estimates from HERE and Google are very comparable, though it is noted that the work zones were not field checked for the extent of activity.

The research also investigated specific highway segments and times when a crash was reported on I-15. A total of 113 crash incidents were analyzed to determine how travel times from the two services differed under the potentially congested conditions occurring during the clearance or resolution of a crash incident. In 71% of these cases, travel times from Google were lower than those reported from HERE.

1. INTRODUCTION

1.1 Problem Statement

UDOT receives data products from the HERE data aggregator, which include travel-time estimates and other data products. Based on presentations by UDOT staff, the HERE data are based on GPS traces obtained from in-vehicle navigation devices representing approximately 2% of the Utah vehicle fleet.

The HERE passive data set presents an opportunity for accurate travel-time data at unprecedented levels of geographic resolution. Although passive data provides a very large sample, including thousands of trips, it is still only a sample. Because it is not a controlled random sample, it is not representative of all travelers or trips.

Commercially available datasets such as HERE include only travelers with certain devices, carriers, and/or apps installed. Moreover, short-distance trips or short-duration activities are often under-represented in the data because they require more frequent observations of position which are not always available due to several factors including battery management and device usage, including the usage of in-vehicle navigation services. Other sources of error include internal algorithms that assert rules for defining “trips” which may result in inaccurate travel-time estimates, particularly in work zones or other cases of non-recurring congestion such as crash incidents.

How representative is the HERE data that UDOT has obtained? This research seeks to validate the HERE dataset by utilizing another “big data” product from Google obtained through Google’s Directions API. This API obtains real-time travel-speed information from users of Google’s Directions functionality. There is no publicly available information on the market penetration or user base of Google’s Directions function, but, for the purposes of this research, it is considered a comparable data set to the HERE data set for travel-time comparisons. The Google Directions functionality is from smartphone users while the HERE travel-time dataset is primarily estimated from in-vehicle navigation devices, but there are no reliable estimates of market penetration for either data set.

1.2 Objectives

This research will expand our understanding of when and/or where travel-time inaccuracies appear in the HERE data when compared to a more prolific crowd-sourced data mine such as Google Traffic. The research will shed light on the strengths and weaknesses of the HERE data regarding geographic locations throughout Utah under normal operating conditions and under conditions where unusual congestion might occur due to crash incidents.

1.3 Scope

The research included six tasks, as follows:

1. Review of recent Big Data travel-time comparisons
2. Selection of Utah corridors for conducting the travel-time comparison
3. Preparation of HERE and Google datasets for identical corridors and time periods
4. Analysis of travel-time datasets.
5. Analysis of travel times associated with crash incidents
6. Findings and final report

1.4 Outline of Report

This report documents the findings of the research and proceeds with the following sections:

- Literature Review
- Research Methods
- Data Collection
- Data Analysis
- Conclusion

2. LITERATURE REVIEW

2.1 Overview

Three studies were reviewed to understand the types of statistical analysis that have been utilized in conducting similar travel-time comparisons. These studies are:

- I-95 Corridor Coalition Vehicle Probe Project: Validation of INRIX Data, July-September 2008. *-determine statistical measures to use*
- Validating the NPMRDS Data Set for South Dakota (SD2013-08-F)
- Considerations for Calculating Arterial System Performance Measures in Virginia (Final Report VTRC 17-R2)

The I-95 Corridor Coalition analysis evaluated the real time-travel time/speed data for 1500 miles of freeways and 1000 miles of arterials in New Jersey, Pennsylvania, Delaware, Maryland, Virginia and North Carolina. The purpose of the analysis was to validate vehicle probe data provided by INRIX using two other speed data sources obtained from the traditional floating car method and from a system of Bluetooth detectors. The two statistical comparison methods used were the Average Absolute Speed Error (AASE) and the Speed Error Bias (SEB).

The South Dakota validation of the NPMRDS was conducted to determine whether the NPMRDS was a valid data source to use for performance reporting. The NPMRDS was compared to 31 sites where SDDOT maintains permanent traffic monitoring equipment that measures traffic counts and travel speeds. The 31 sites were situated on 14 Interstate Highways and 17 principal arterial, non-Interstate locations.

Two statistical measures were used to determine statistical fit:

Absolute Speed Error, where a 10-mph threshold was selected as a maximum allowable error limit; and,

Speed Error Bias, where a 5mph threshold was selected as a maximum allowable error limit.

The analysis found that the NPMRDS generally underestimated speeds across all speed classes (0-30 mph, 30-45mph, 45-60mph, >60mph) for both site types, Interstates and principal arterials, though errors decreased as facility speed and volume increased. NPMRDS free-flow speeds are biased low in all cases. NPMRDS free-flow speeds were found to be 7 to 9 miles per hour (mph) slower than SDDOT-reported free-flow speeds.

The largest error biases were found on lower volume, lower speed roadways.

The Virginia DOT study focused on arterial travel speeds as an arterial system performance measure, as arterials had not been studied extensively. The VDOT study used arterial travel speeds obtained from Bluetooth detectors as their benchmark and compared travel speeds from INRIX and from the NPMRDS against this benchmark.

3. RESEARCH METHODS

3.1 Overview

The research consisted of accumulating a travel-time estimate data set from both HERE and Google Directions for 59 specified highway segments in Utah over an approximate 30-day period. RSG also obtained travel-time estimates from Bing for the same highway segments over the same period. This research focuses on the HERE-Google comparison, though some graphics include Bing data for comparison purposes.

3.2 Google Directions API

To obtain a comparable travel-time dataset from Google, RSG utilized the Google Directions API. This is a Google service that calculates directions between locations using an HTTP address. The Directions API returns a “duration_in_traffic” estimate, which is the time vehicles spend traversing a specific route defined by origin and destination latitudes and longitudes.

The origins and destinations were paired with the 59 routes set up in iPems within the “here_it_is” tag.

As an additional analysis, RSG also obtained travel-speed information from the Bing Maps service for the same 59 selected highway segments. Generally, the Bing data showed lower speeds for all highway segments and are not considered reliable. Some of the Bing results are included in the report graphics, but the focus of the research is on the Google Directions-HERE comparison.

3.3 Selection of Trial Highway Segments and Comparison of Average Travel Speeds

RSG coordinated with UDOT's Traffic Management Division staff to select 59 highway segments for conducting the travel-time comparison. There was interest in achieving geographic and functional class diversity in the selections and in assessing some specific highway segments where UDOT had specific interest, including work zones.

Figure 1- Figure 31 show these segments as they are defined within the iPems online portal. Except for two highway segments -- I-15 to I-80 Eastbound and 5750 S to Redwood Road, Taylorsville – all highway segments were analyzed for travel speed separately for both directions of travel (only one direction of travel is shown in the figure maps).

The figures also provide a comparison of travel speeds from HERE and Google. The charts show average travel speeds across all 15-minute periods for the trial period, January 21-February 25, 2019.

Figure 1: Comparison Highway Segment – I-15 4800S – 900 S (NB/SB)

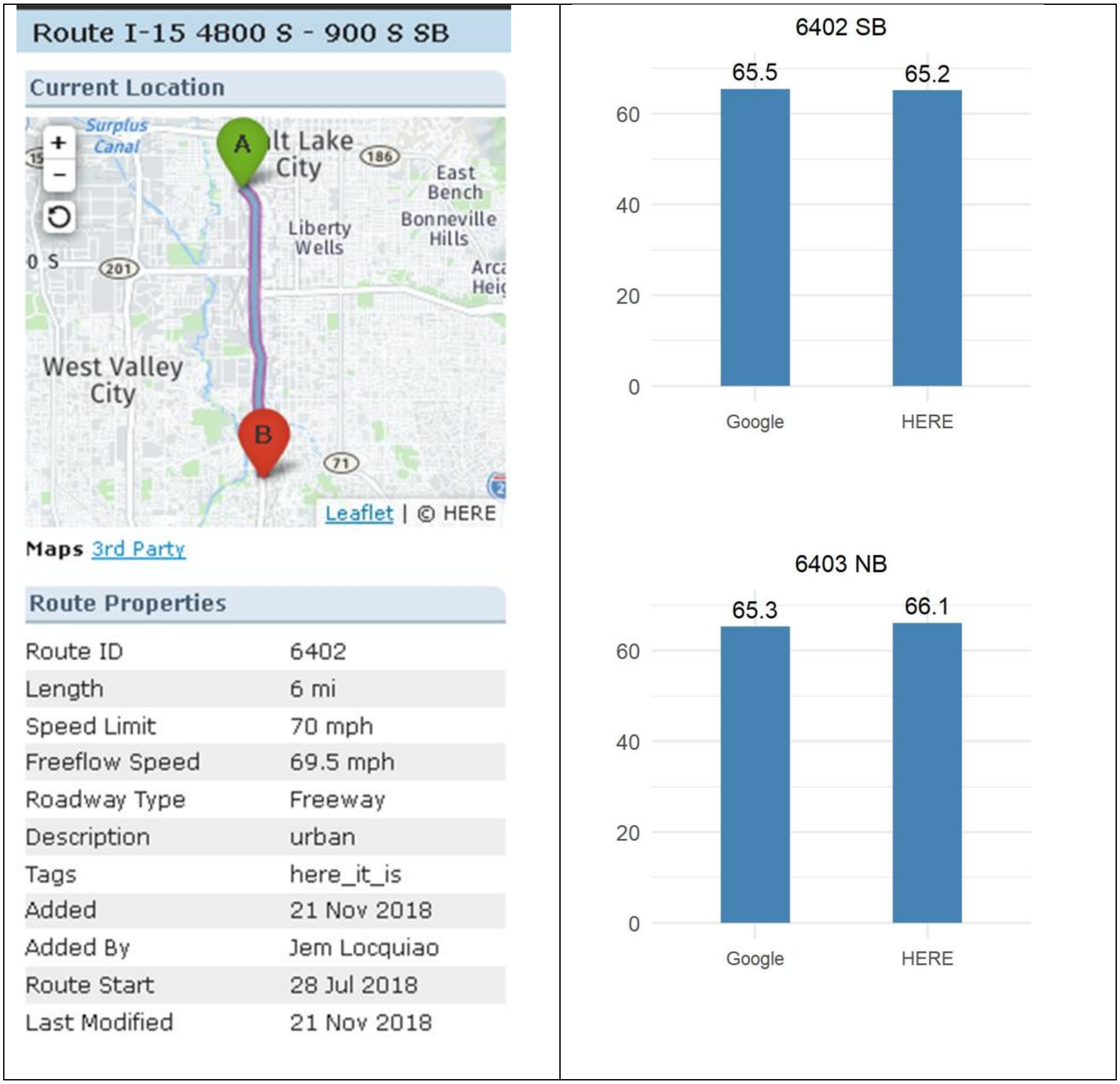


Figure 2: Comparison Highway Segment – Foothill Drive (NB/SB)

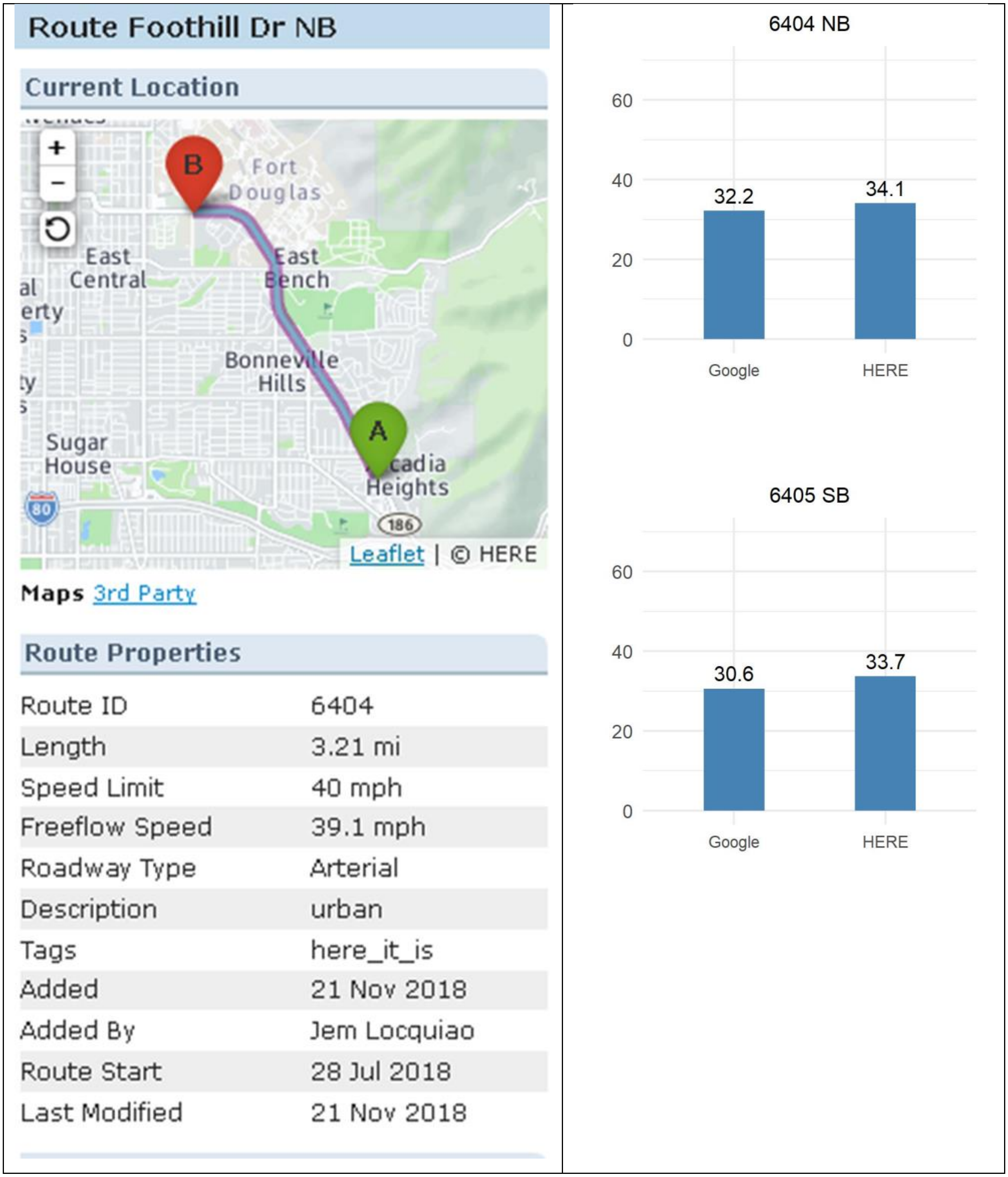


Figure 3: Comparison Highway Segment – I-80 Mountain Dell (EB/WB)

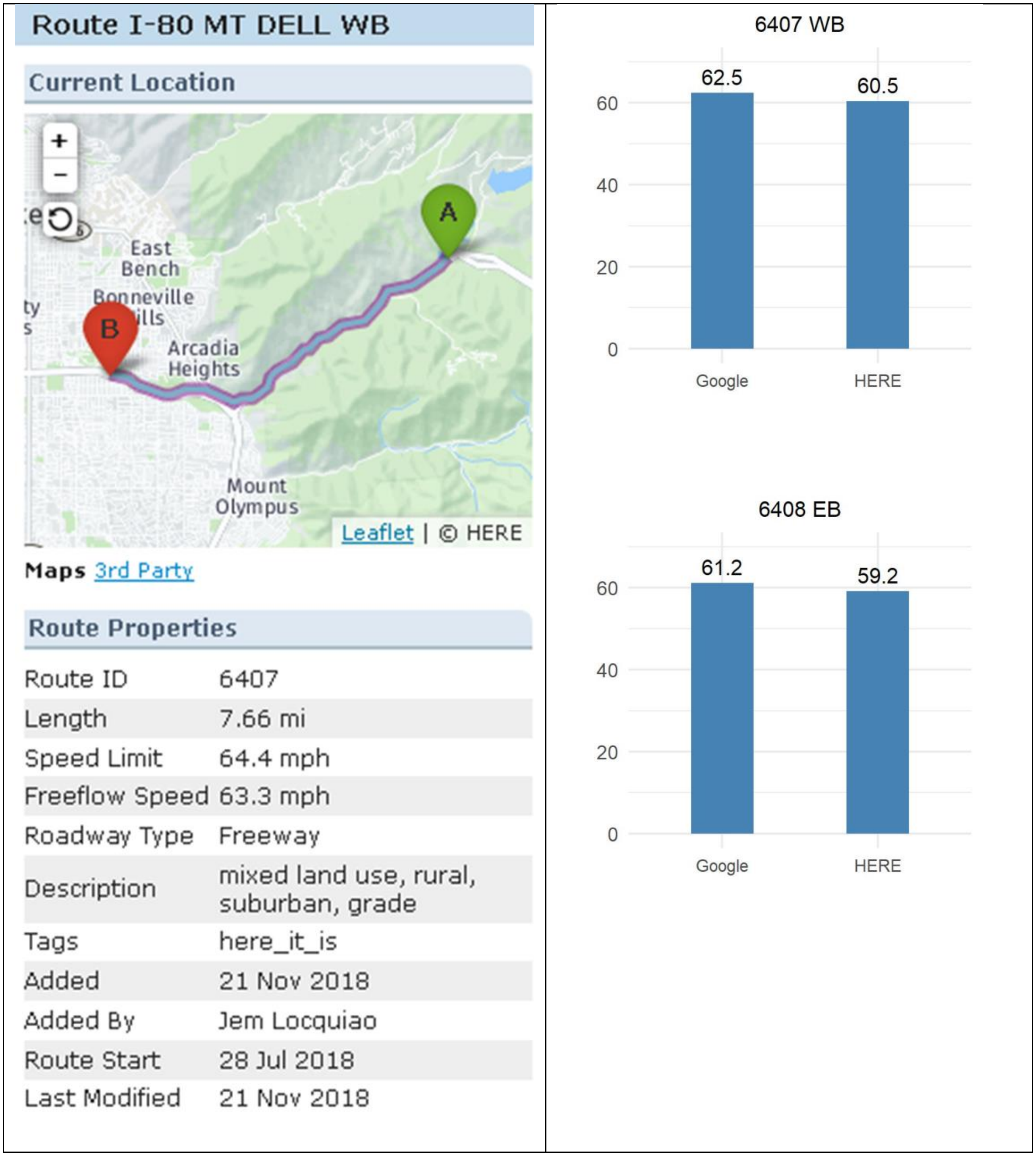


Figure 4: Comparison Highway Segment – Redwood Road 2120 N – 735 N (NB/SB)

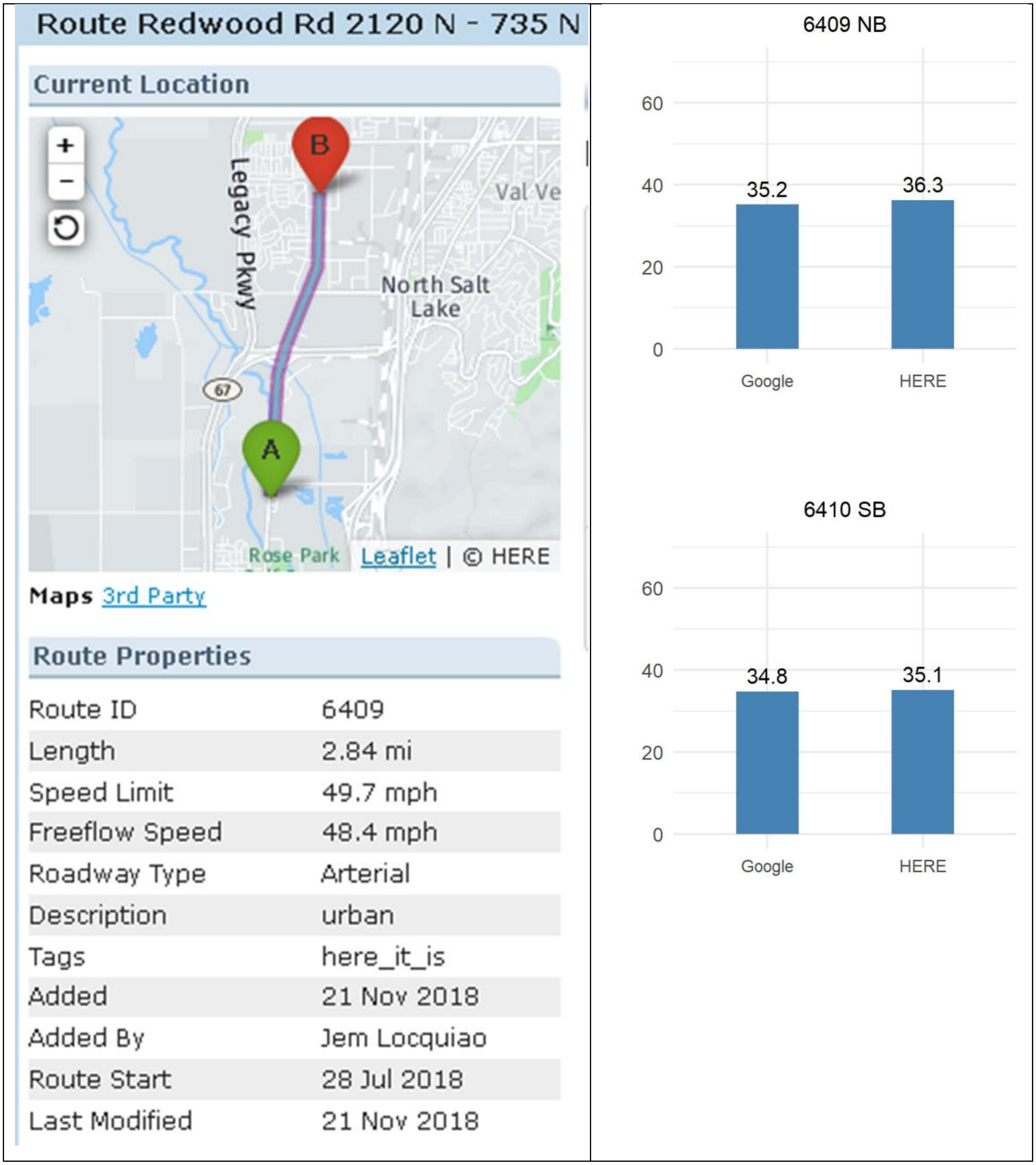


Figure 5: Comparison Highway Segment – I-80 Great Salt Lake (WB/EB)

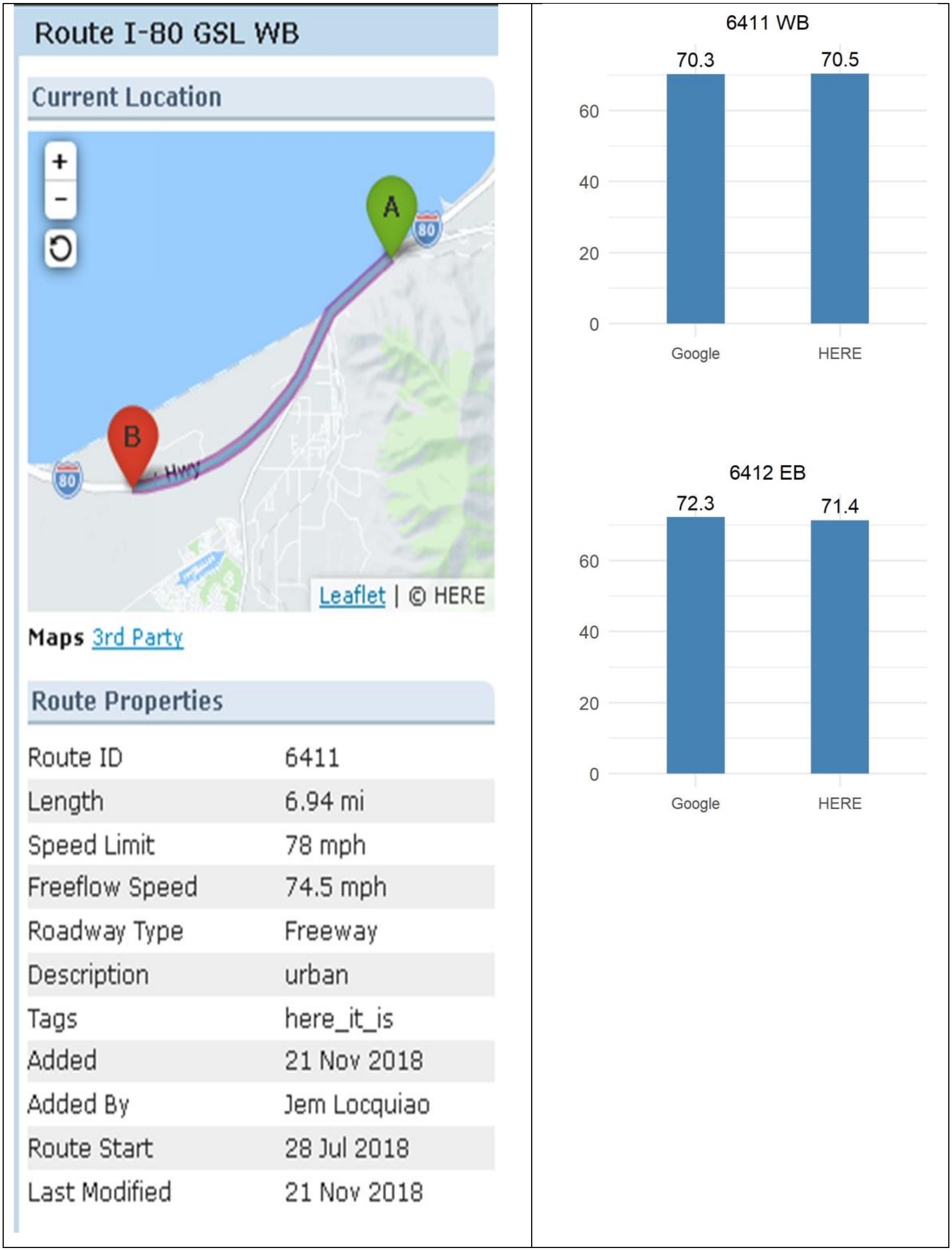


Figure 6: Comparison Highway Segment – Route 9000S/9400S (WB/EB)

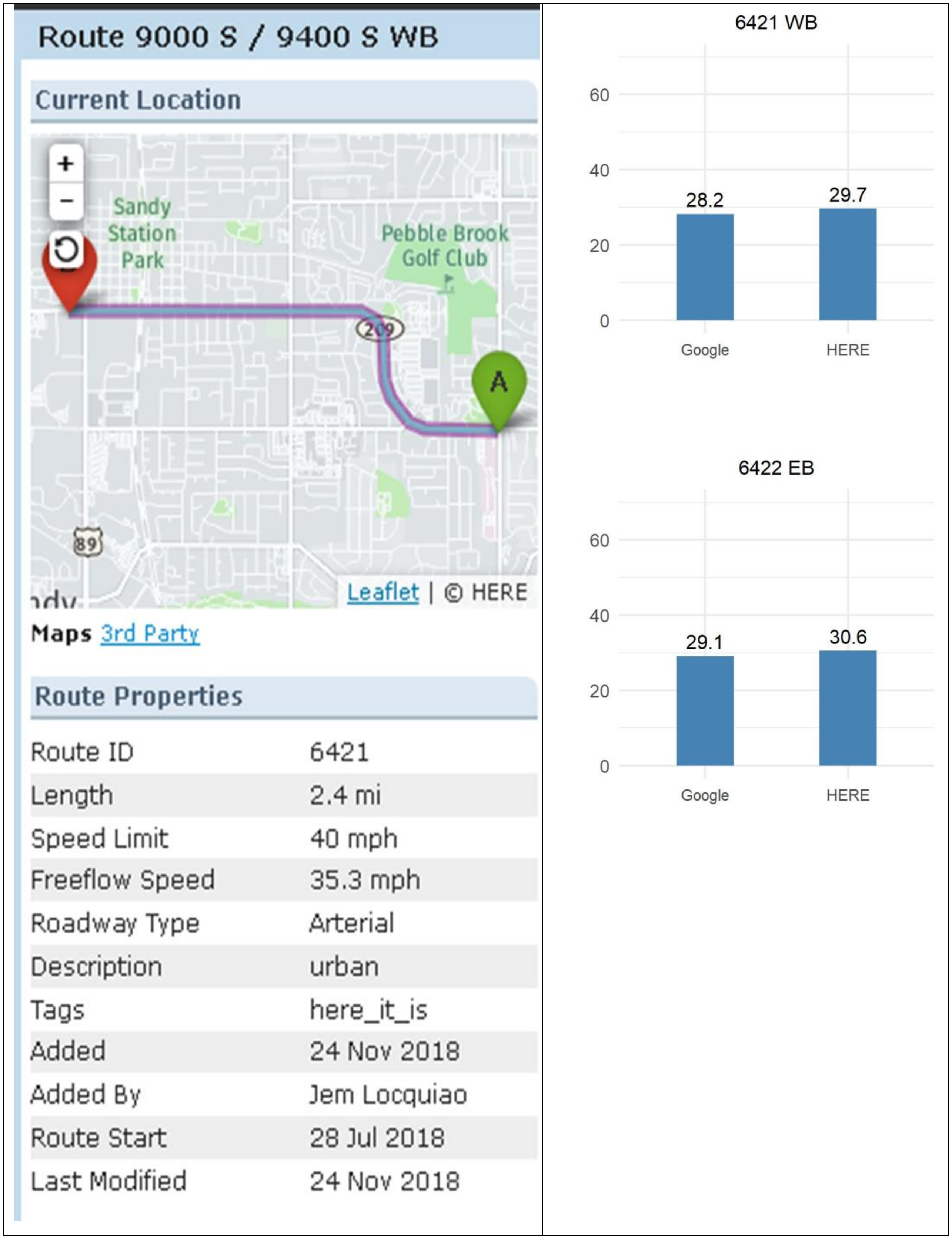


Figure 7: Comparison Highway Segment – Moab Main Street (NB/SB)

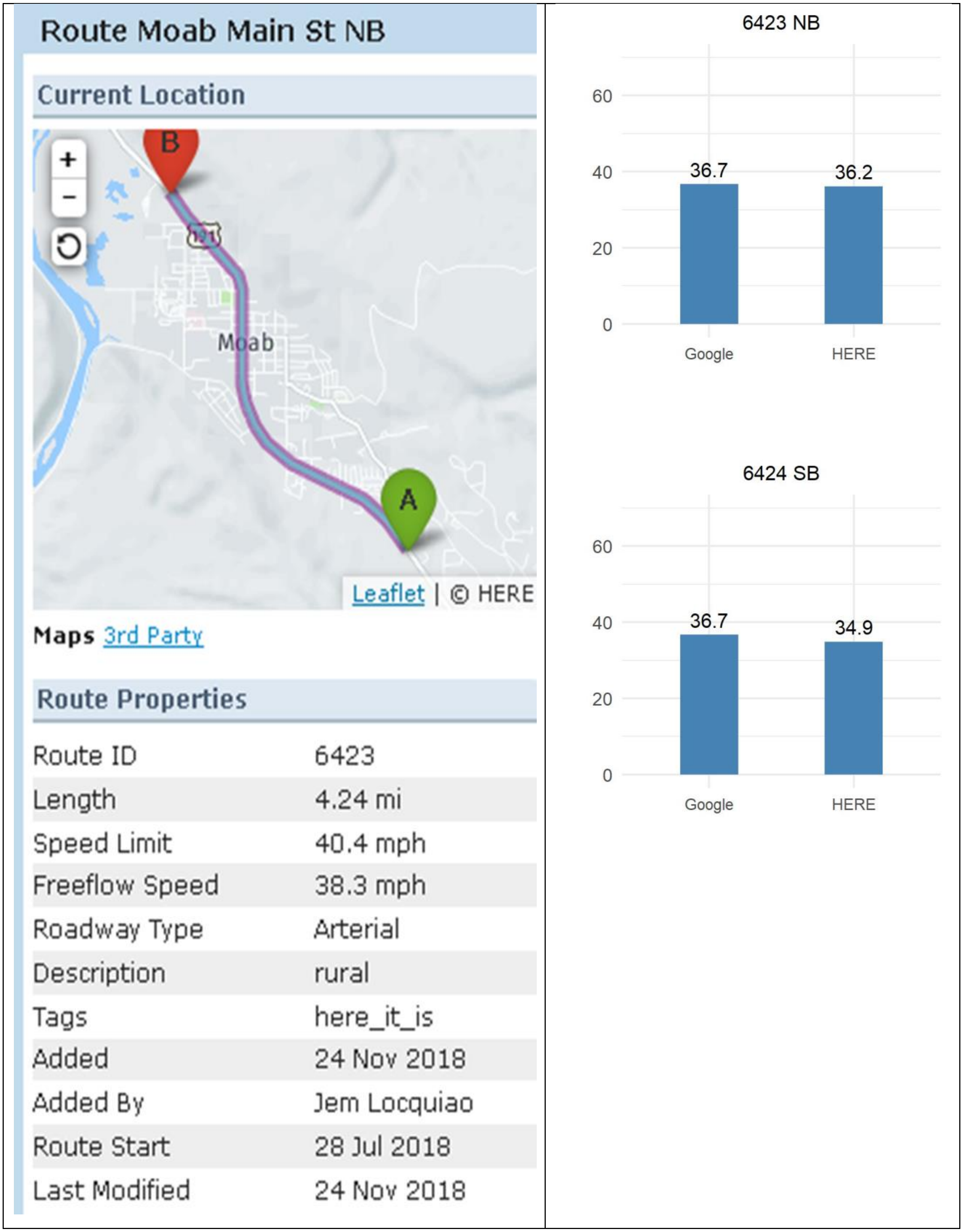


Figure 8: Comparison Highway Segment – SR 46 La Sal (NB/SB)

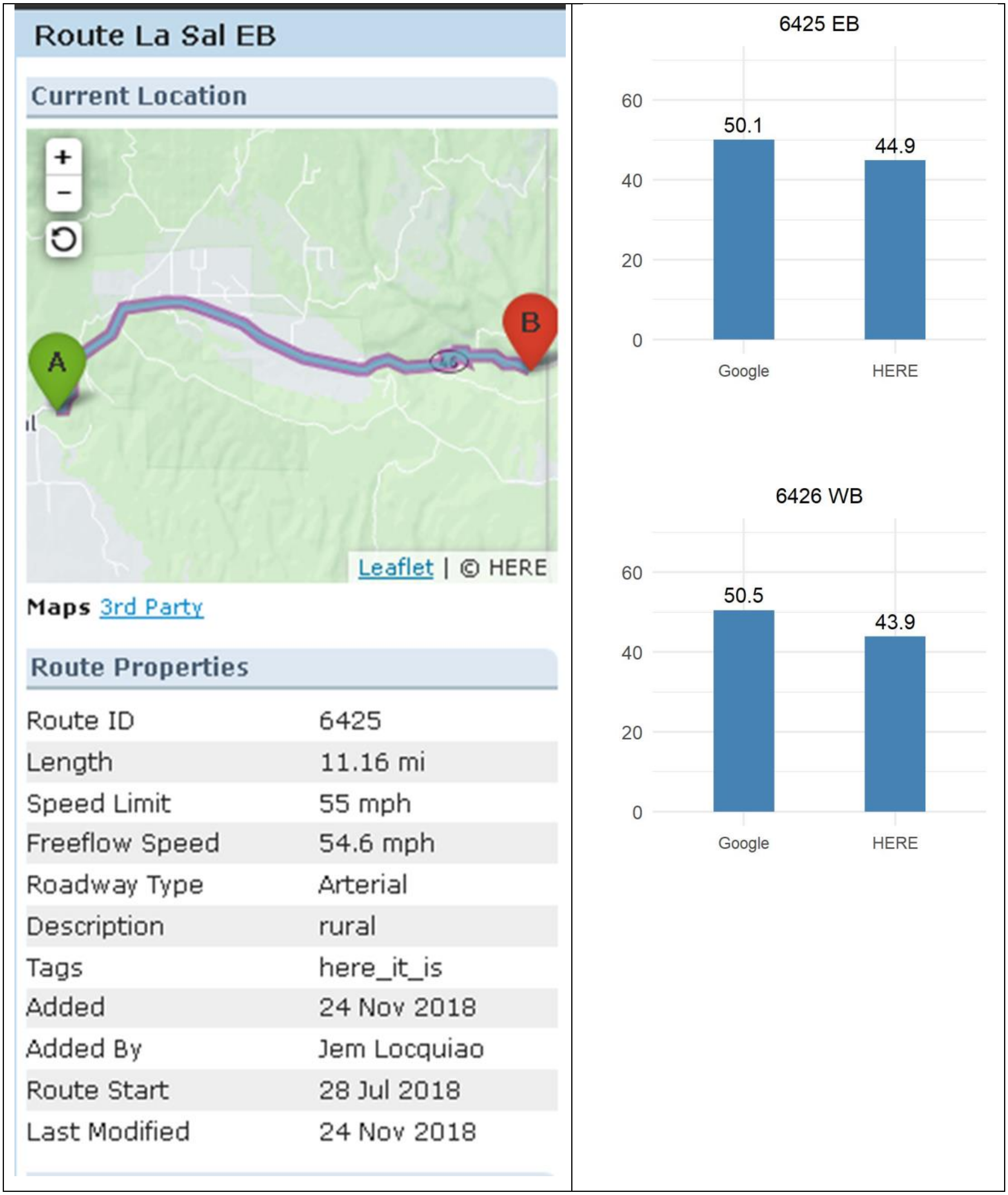


Figure 9: Comparison Highway Segment – Bluff Street, St. George (NB/SB)

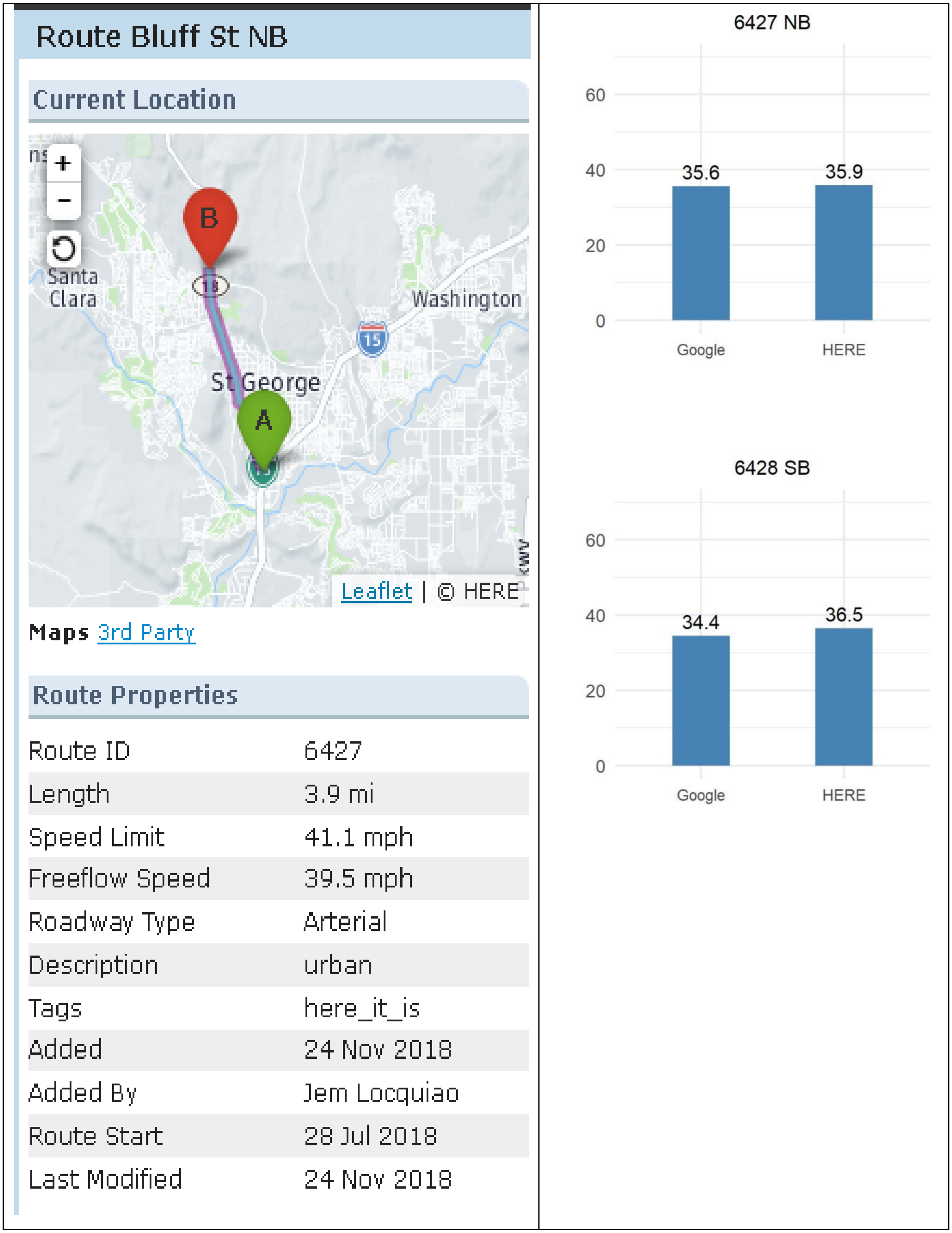


Figure 10: Comparison Highway Segment – I-15, St. George (NB/SB)

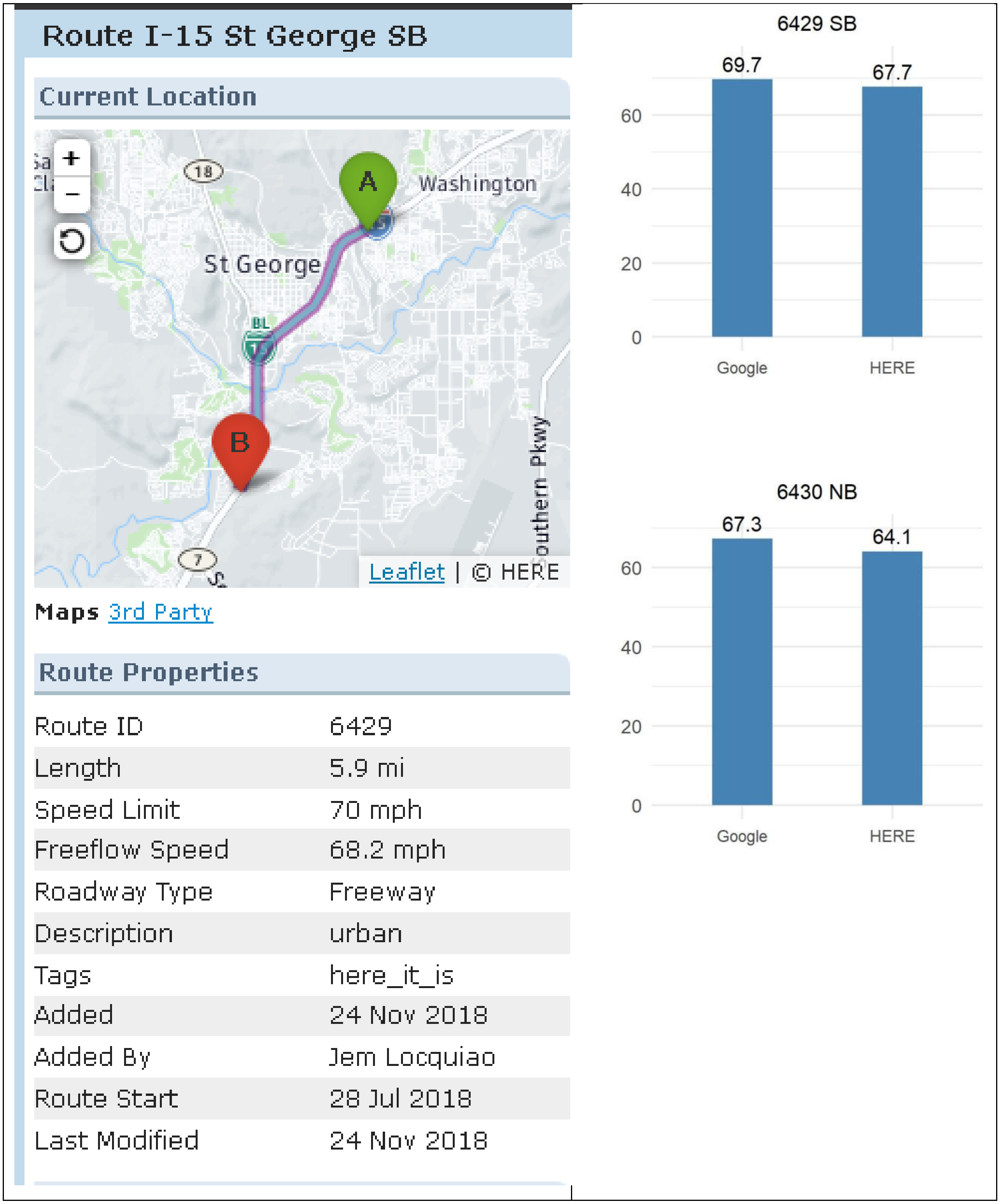
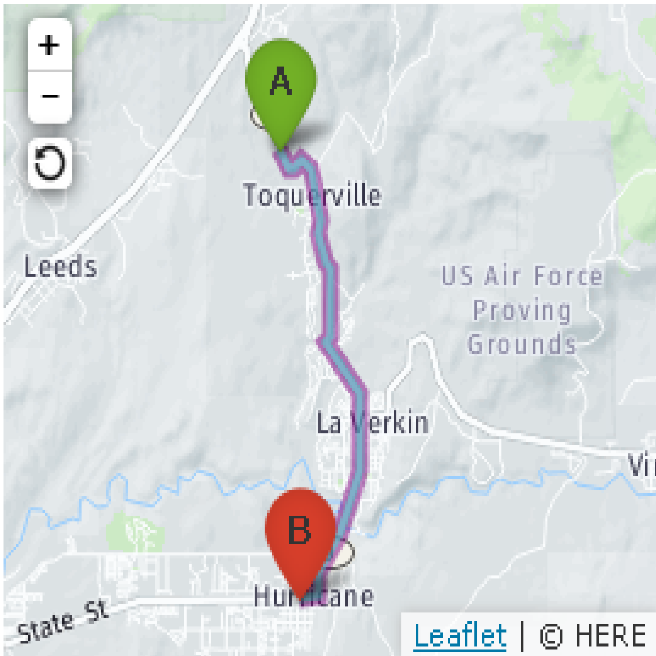


Figure 11: Comparison Highway Segment – SR 17, Toquerville-La Verkin-Hurricane (NB/SB)

Route Toquerville - La Verkin - Hurr

Current Location



Maps 3rd Party

Route Properties

Route ID	6431
Length	7.13 mi
Speed Limit	43 mph
Freeflow Speed	42.6 mph
Roadway Type	Freeway
Description	mixed land use
Tags	here_it_is
Added	24 Nov 2018
Added By	Jem Locquiao
Route Start	28 Jul 2018
Last Modified	24 Nov 2018

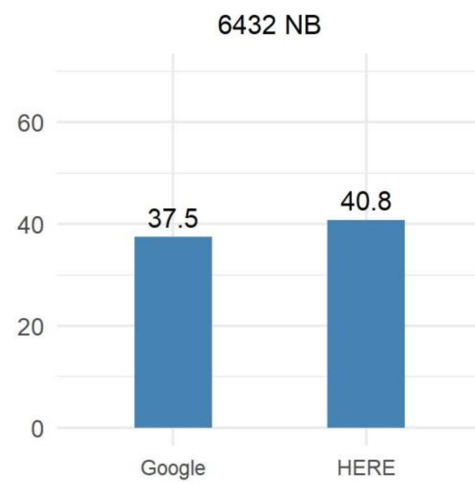
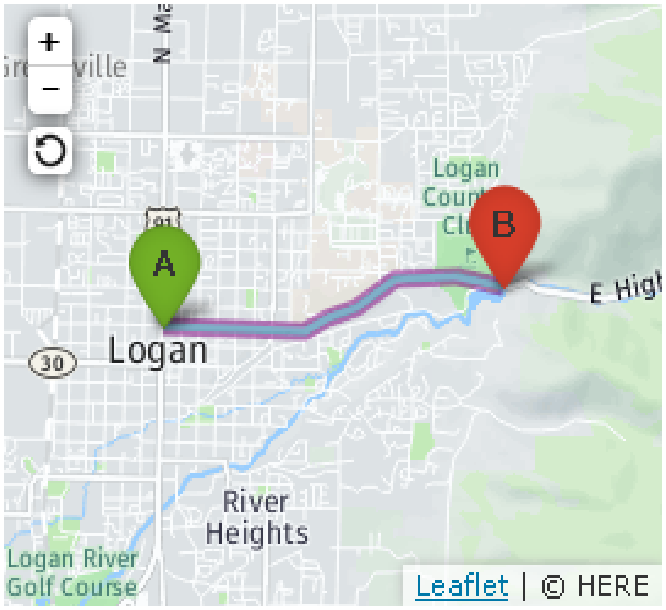


Figure 12: Comparison Highway Segment – US89, Logan (EB/WB)

Route US-89 Logan EB

Current Location



Maps 3rd Party

Route Properties

Route ID	6433
Length	2.58 mi
Speed Limit	42.6 mph
Freeflow Speed	38.8 mph
Roadway Type	Arterial
Description	urban
Tags	here_it_is
Added	24 Nov 2018
Added By	Jem Locquiao
Route Start	28 Jul 2018
Last Modified	24 Nov 2018

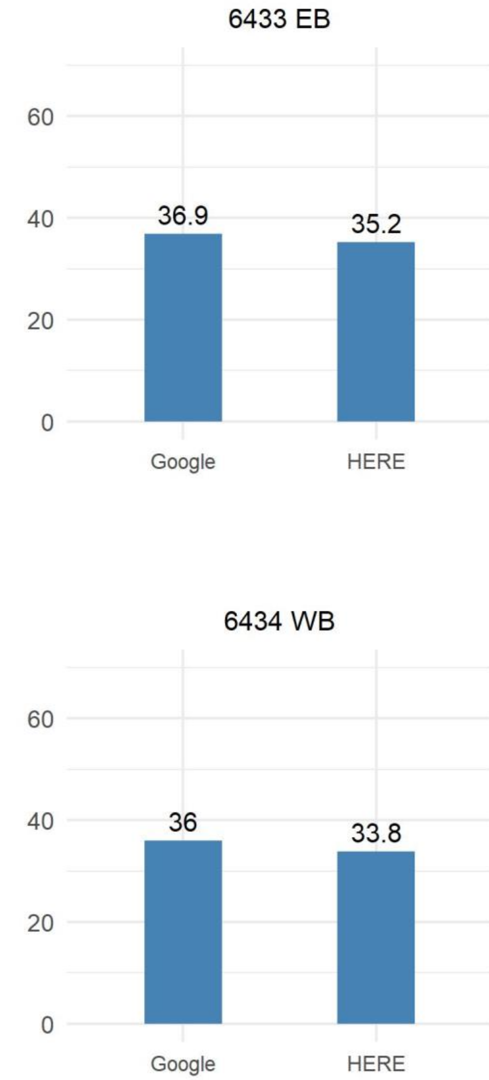
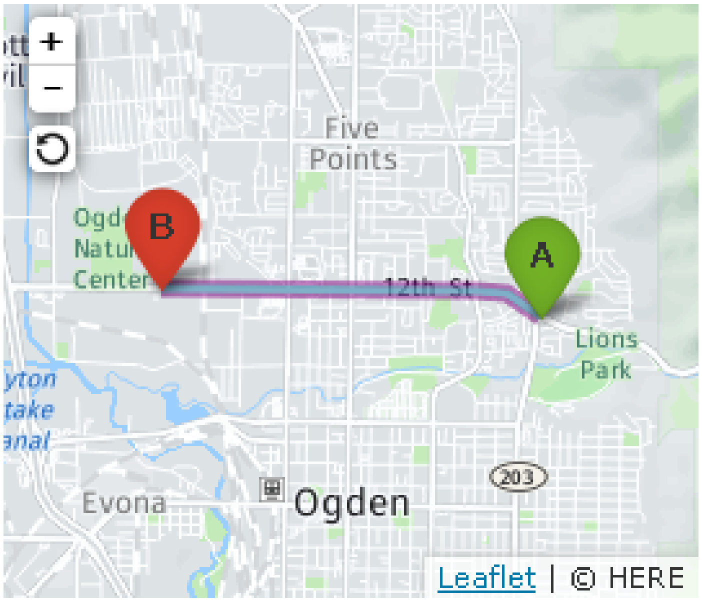


Figure 13: Comparison Highway Segment – SR39, Ogden (EB/WB)

Route Ogden SR-39 WB

Current Location



Maps 3rd Party

Route Properties

Route ID	6435
Length	2.67 mi
Speed Limit	44.7 mph
Freeflow Speed	40.1 mph
Roadway Type	Arterial
Description	urban
Tags	here_it_is
Added	24 Nov 2018
Added By	Jem Locquiao
Route Start	28 Jul 2018
Last Modified	24 Nov 2018

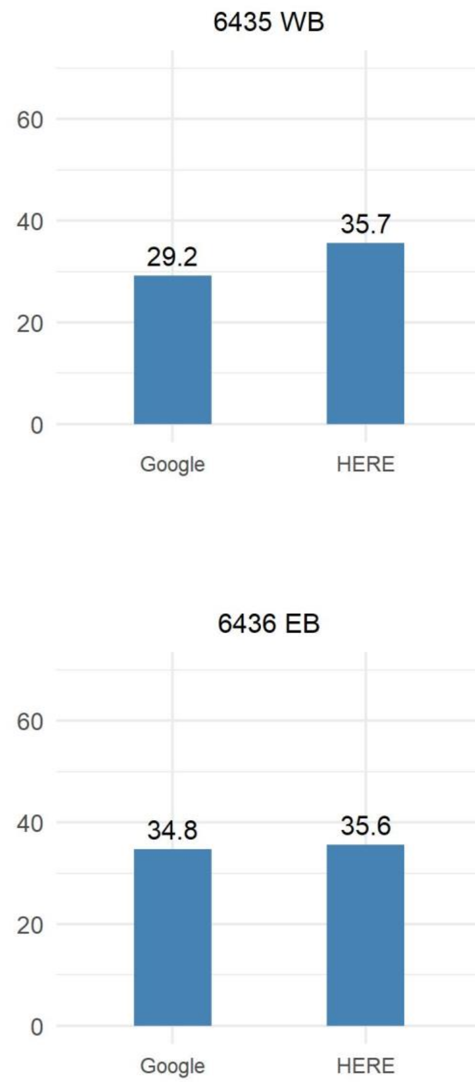
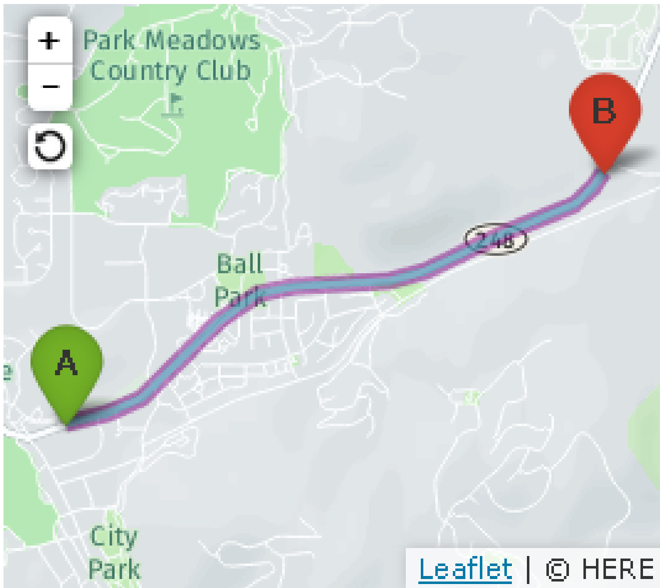


Figure 14: Comparison Highway Segment – SR248, Park City (EB/WB)

Route Kearns SR-248 EB

Current Location



Maps 3rd Party

Route Properties

Route ID	6437
Length	2.23 mi
Speed Limit	40.6 mph
Freeflow Speed	38.3 mph
Roadway Type	Arterial
Description	mixed land use
Tags	here_it_is
Added	24 Nov 2018
Added By	Jem Locquiao
Route Start	28 Jul 2018
Last Modified	24 Nov 2018

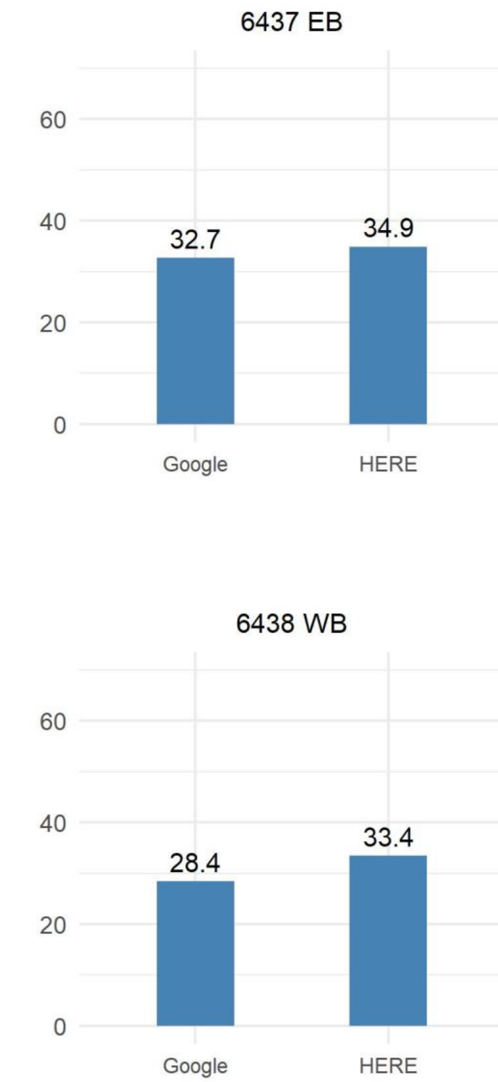


Figure 15: Comparison Highway Segment – Bangerter Highway (NB/SB)

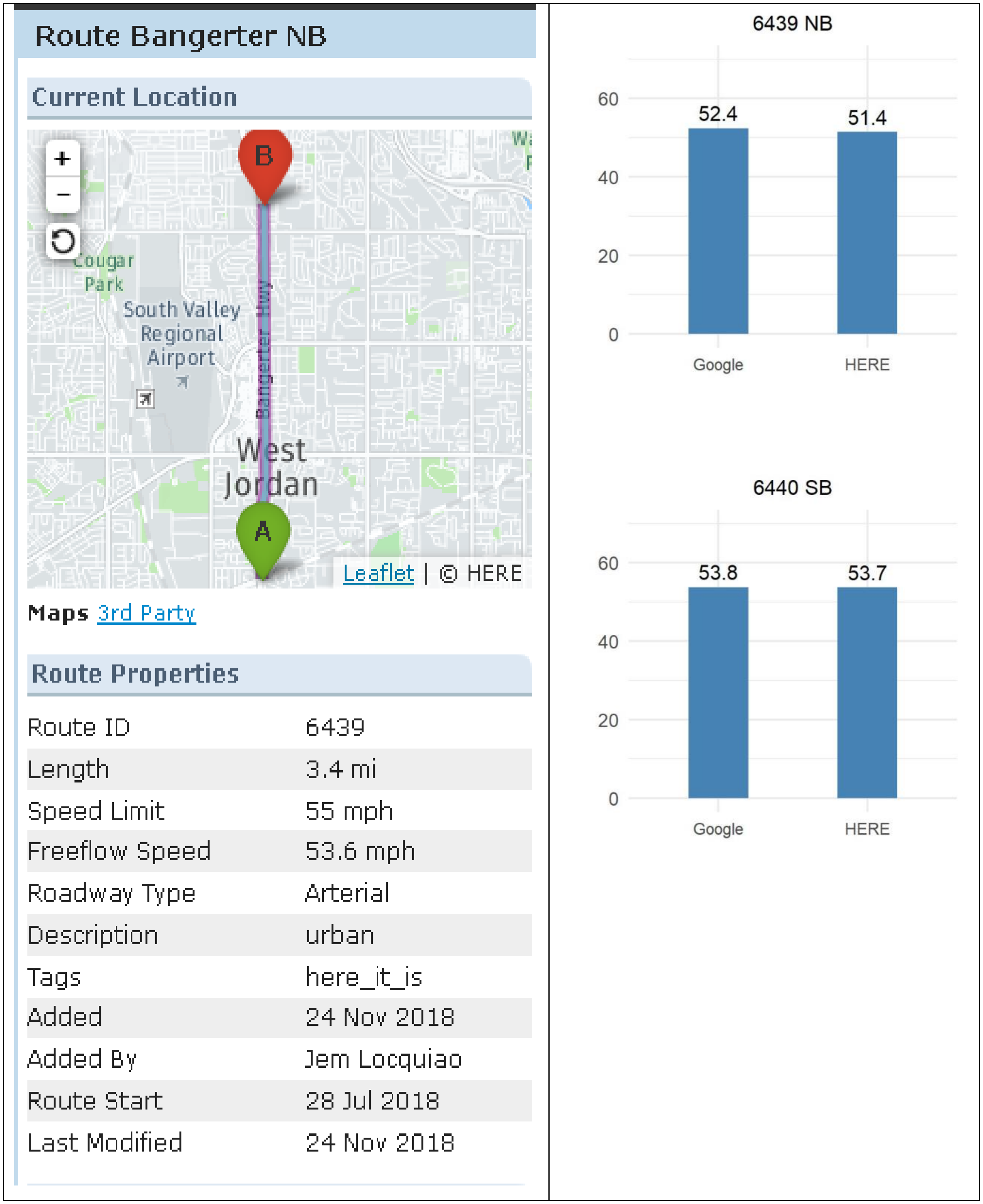
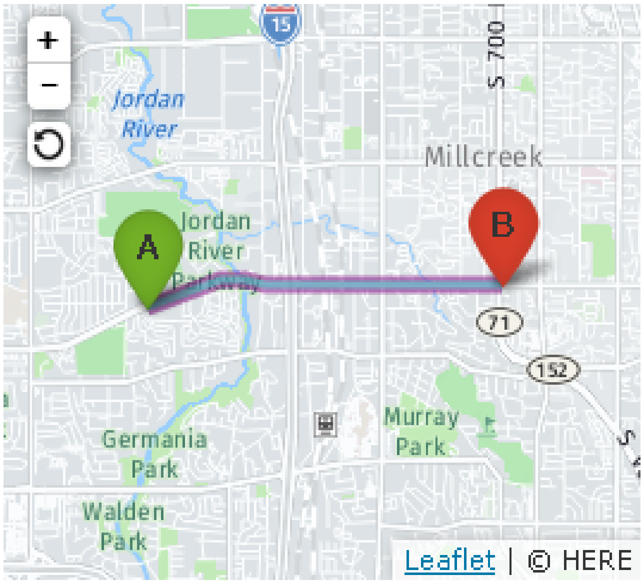


Figure 16: Comparison Highway Segment – SR 266 (EB/WB)

Route SR-266 EB

Current Location



Maps 3rd Party

Route Properties

Route ID	6441
Length	2.71 mi
Speed Limit	42.6 mph
Freeflow Speed	38 mph
Roadway Type	Arterial
Description	urban
Tags	here_it_is
Added	24 Nov 2018
Added By	Jem Locquiao
Route Start	28 Jul 2018
Last Modified	24 Nov 2018

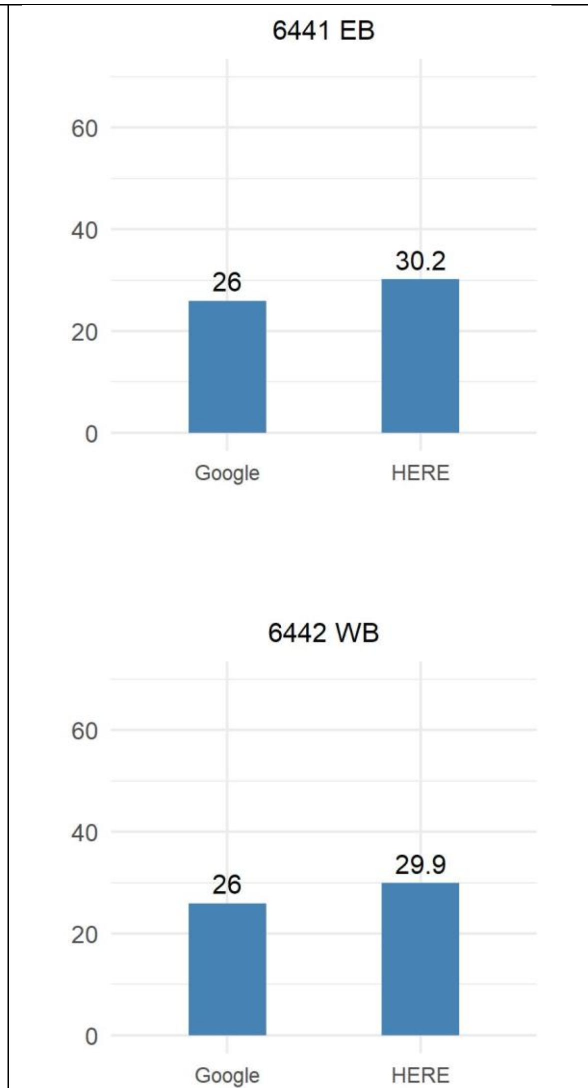
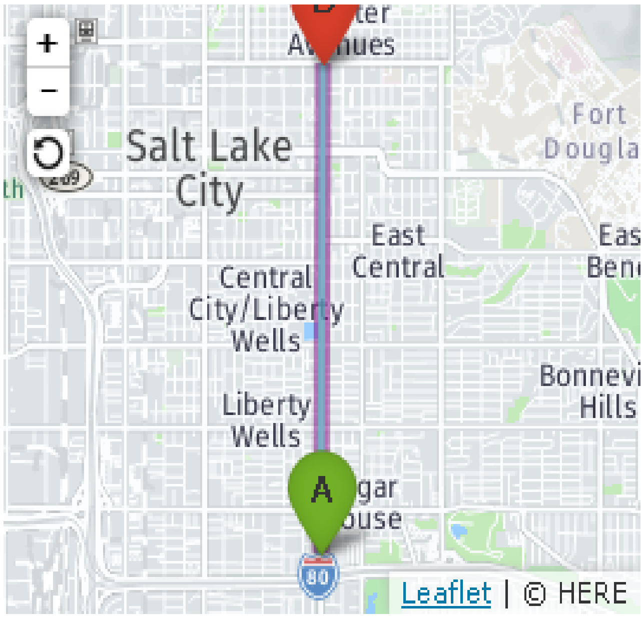


Figure 17: Comparison Highway Segment – 700E, Salt Lake City (NB/SB)

Route 700 E NB (extended)

Current Location



Maps 3rd Party

Route Properties

Route ID	6443
Length	3.32 mi
Speed Limit	39.1 mph
Freeflow Speed	34.5 mph
Roadway Type	Arterial
Description	urban
Tags	here_it_is
Added	24 Nov 2018
Added By	Jem Locquiao
Route Start	28 Jul 2018
Last Modified	24 Nov 2018

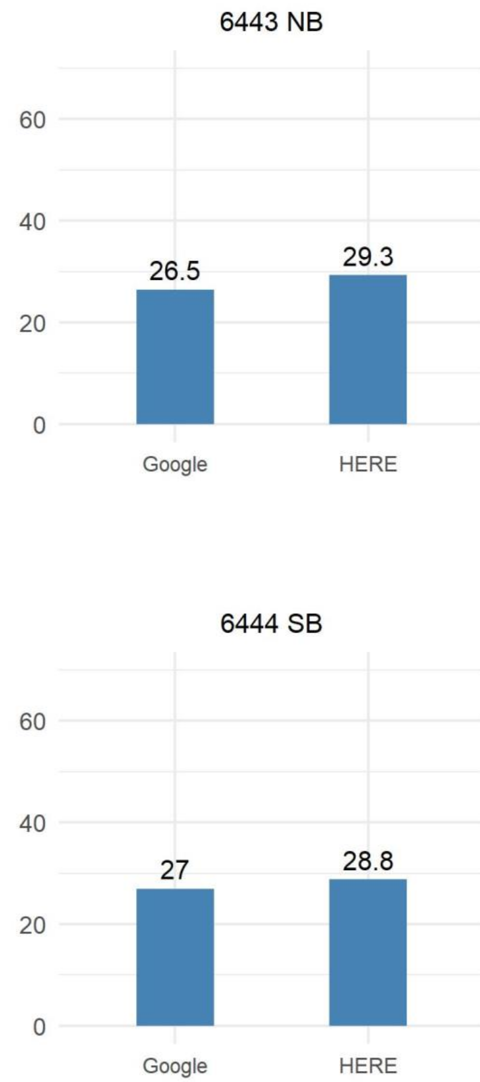
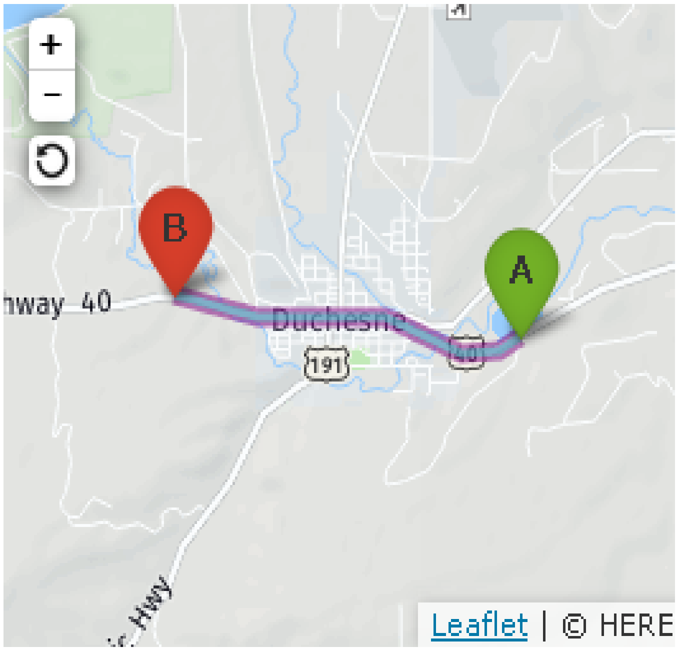


Figure 18: Comparison Highway Segment – Highway 40, Duchesne (EB/WB)

Route HWY 40 Duchesne WB

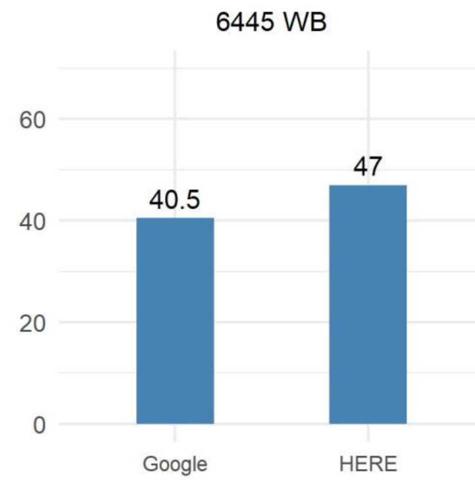
Current Location



Maps 3rd Party

Route Properties

Route ID	6445
Length	2.56 mi
Speed Limit	45.4 mph
Freeflow Speed	48.7 mph
Roadway Type	Arterial
Description	rural
Tags	here_it_is
Added	24 Nov 2018
Added By	Jem Locquiao
Route Start	28 Jul 2018
Last Modified	24 Nov 2018



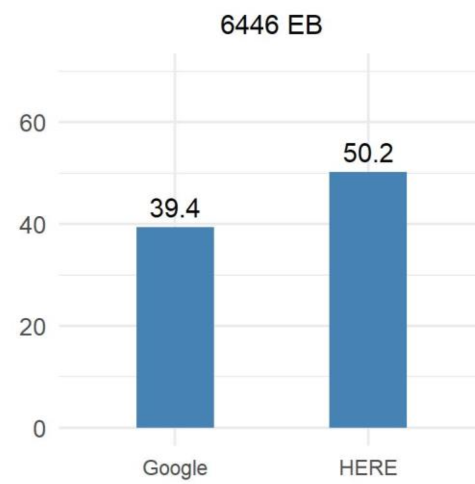


Figure 19: Comparison Highway Segment – University Avenue, Provo (NB/SB)

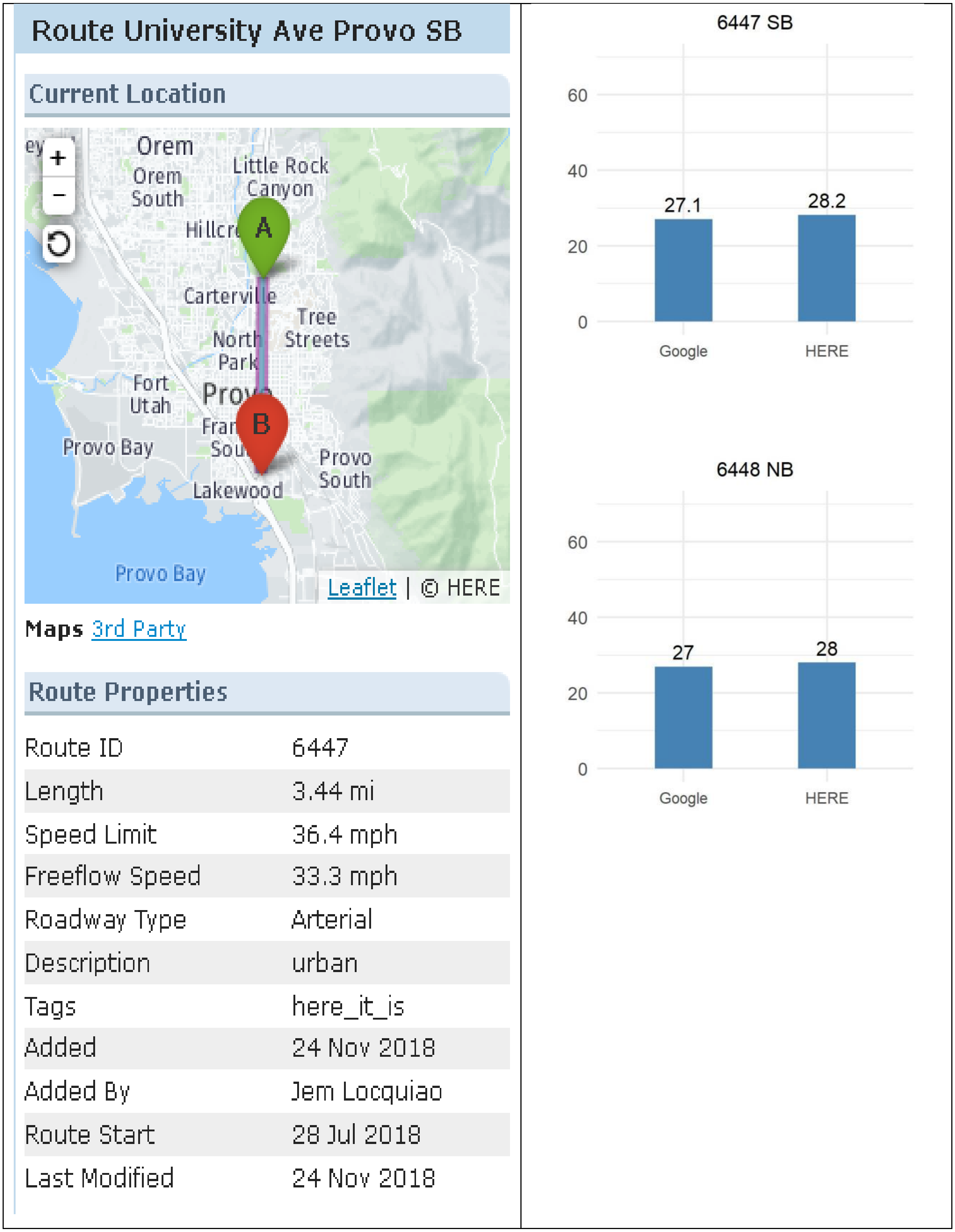


Figure 20: Comparison Highway Segment – I-80-Echo Reservoir (NB/SB)

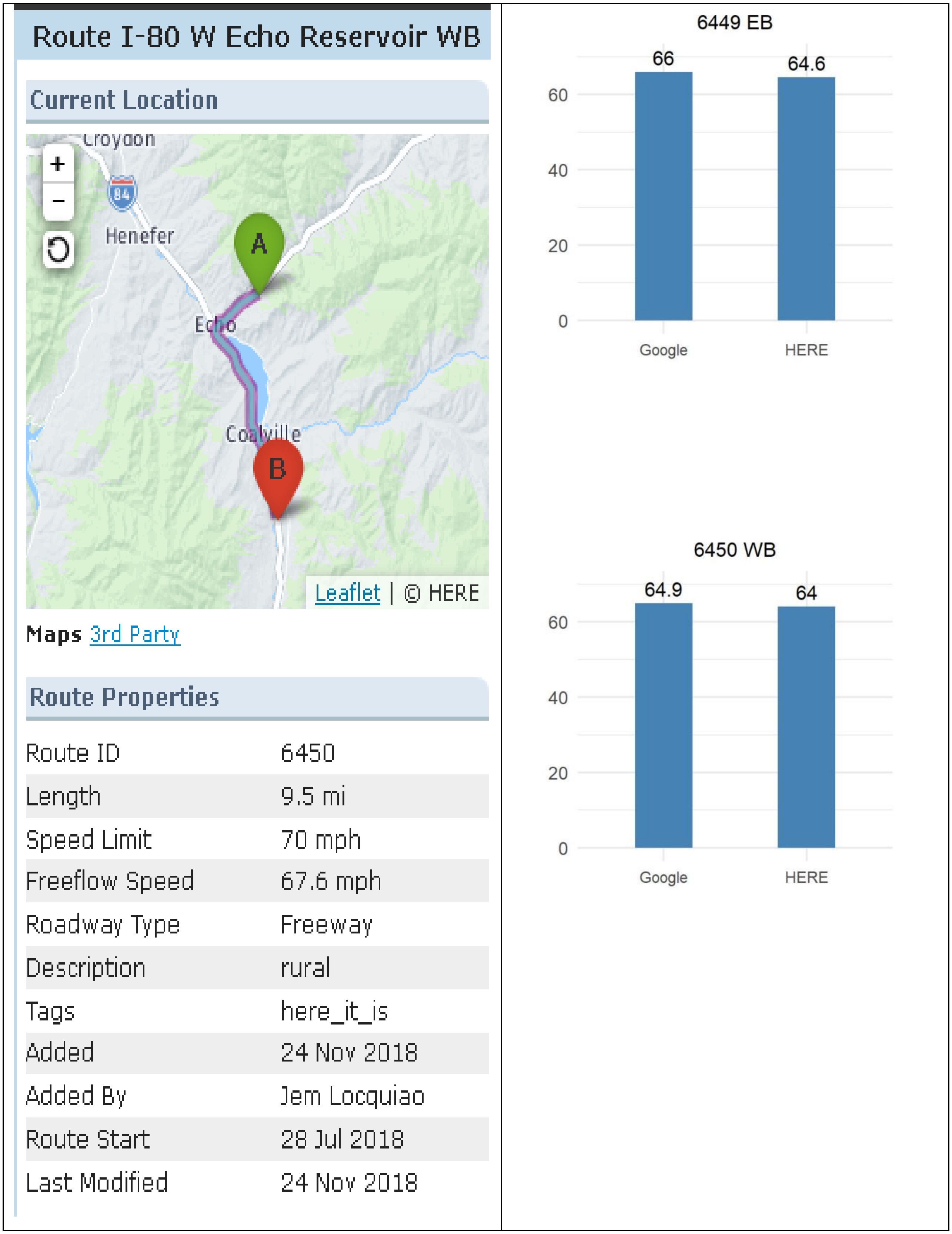
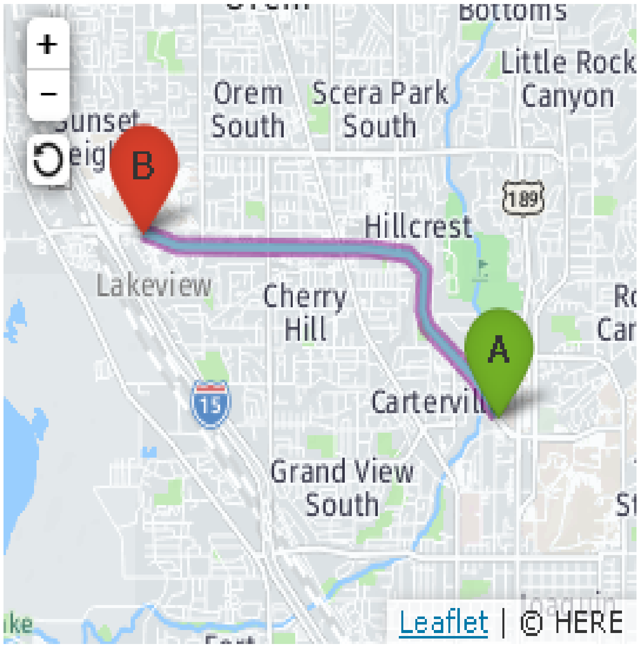


Figure 21: Comparison Highway Segment – University Parkway, Orem (EB/WB)

Route University Parkway Orem W

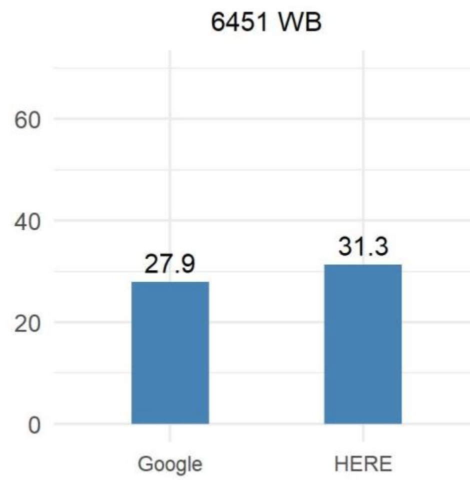
Current Location



Maps 3rd Party

Route Properties

Route ID	6451
Length	3.39 mi
Speed Limit	46.4 mph
Freeflow Speed	38.2 mph
Roadway Type	Arterial
Description	urban
Tags	here_it_is
Added	24 Nov 2018
Added By	Jem Locquiao
Route Start	28 Jul 2018
Last Modified	24 Nov 2018



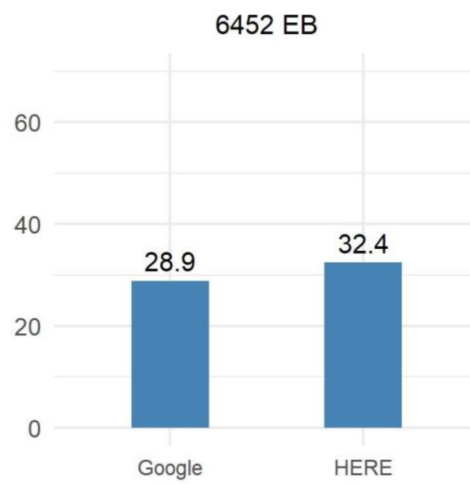


Figure 22: Comparison Highway Segment – SR 92 Timpanogos (EB/WB)

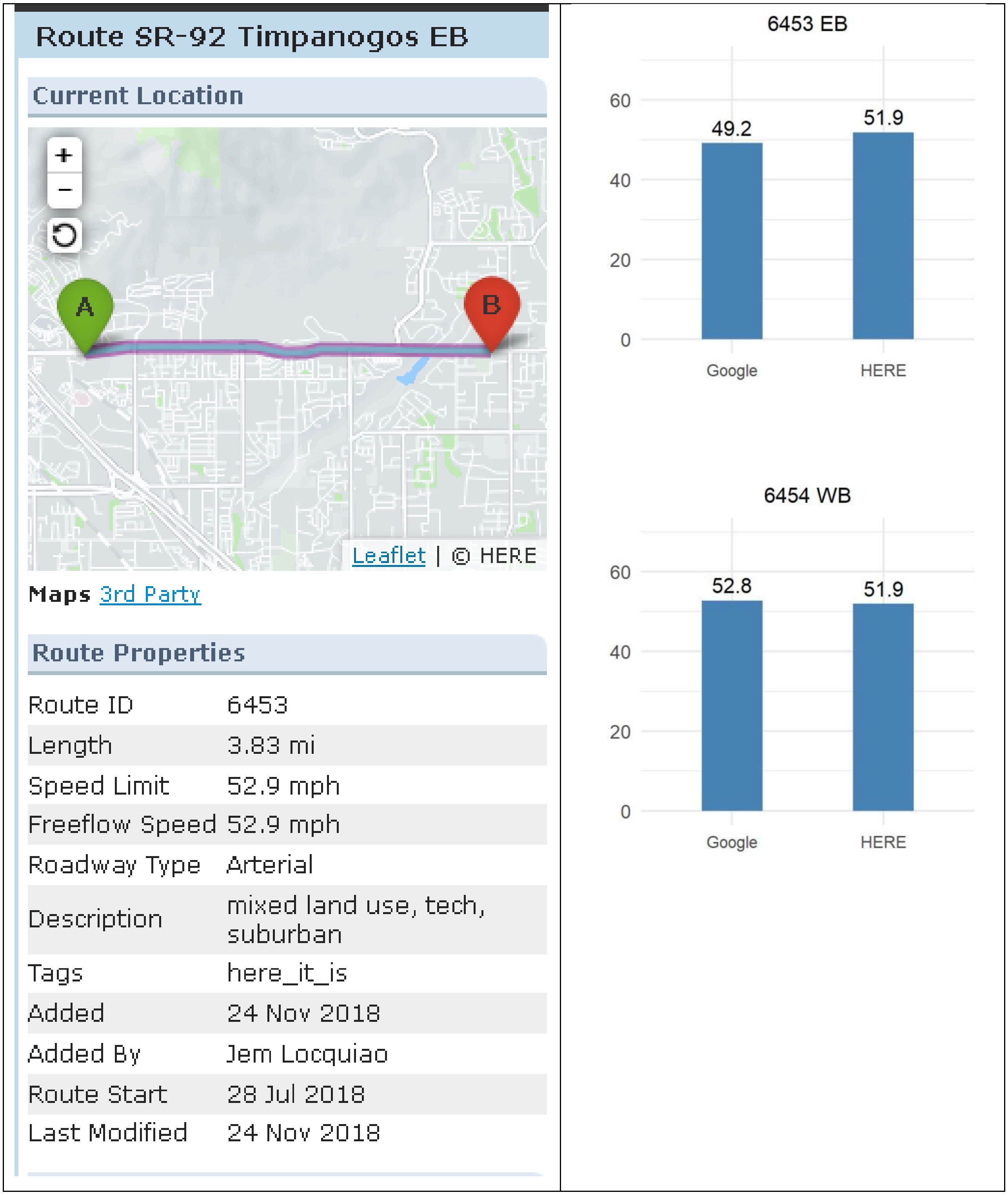
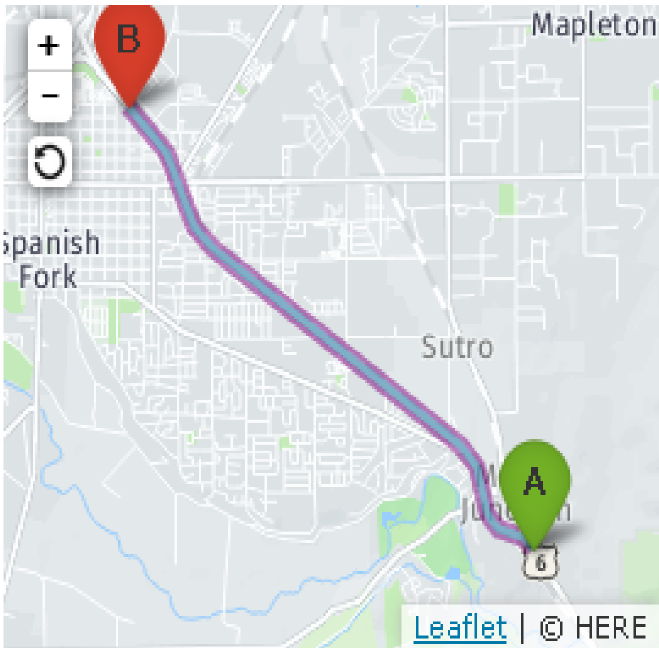


Figure 23: Comparison Highway Segment – US 6 (NB/SB)

Route US-6 NB

Current Location



Maps 3rd Party

Route Properties

Route ID	6457
Length	4.28 mi
Speed Limit	57 mph
Freeflow Speed	53.2 mph
Roadway Type	Arterial
Description	mixed land use, suburban
Tags	here_it_is
Added	24 Nov 2018
Added By	Jem Locquiao
Route Start	28 Jul 2018
Last Modified	24 Nov 2018

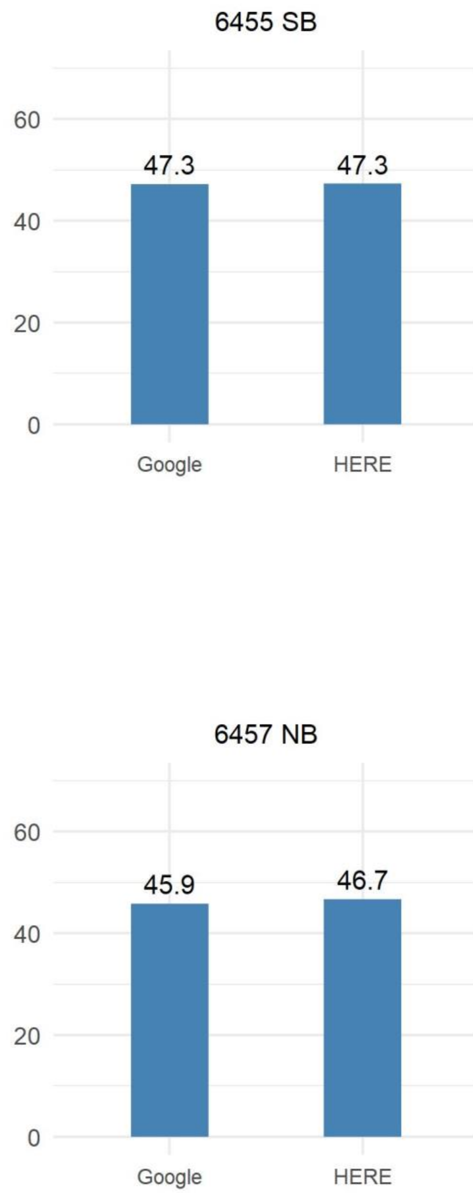
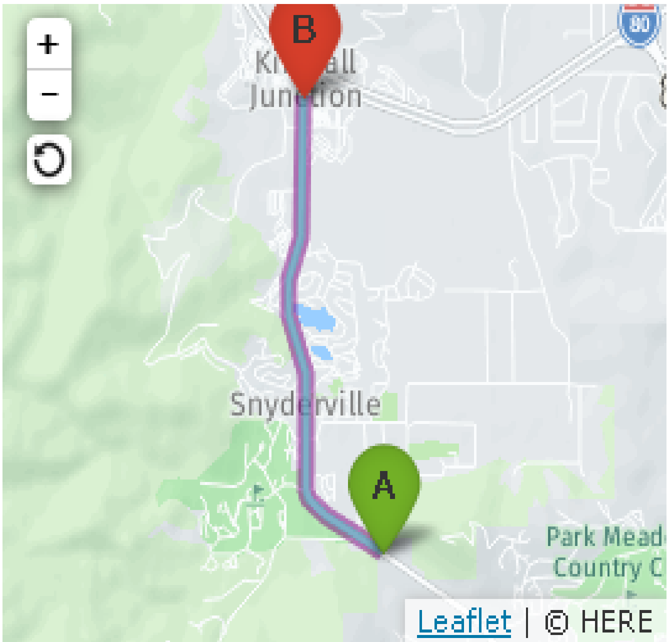


Figure 24: Comparison Highway Segment – SR 224, Kimball Jct/Snyderville (NB/SB)

Route SR-224 Snyderville NB

Current Location



Maps 3rd Party

Route Properties

Route ID	6458
Length	3.27 mi
Speed Limit	50.4 mph
Freeflow Speed	45 mph
Roadway Type	Arterial
Description	mixed land use, suburban
Tags	here_it_is
Added	24 Nov 2018
Added By	Jem Locquiao
Route Start	28 Jul 2018
Last Modified	24 Nov 2018

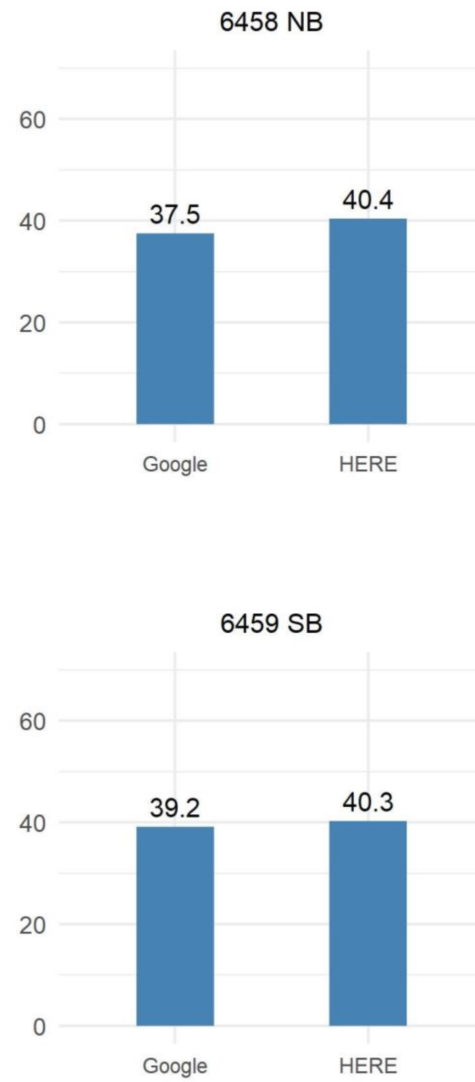
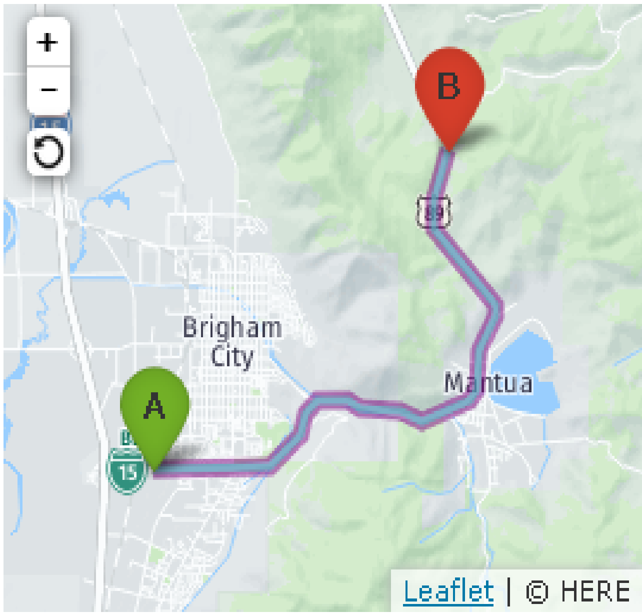


Figure 25: Comparison Highway Segment – US 89 Brigham City

Route US-89 Brigham City EB

Current Location



Maps 3rd Party

Route Properties

Route ID	6460
Length	9.14 mi
Speed Limit	59 mph
Freeflow Speed	57.3 mph
Roadway Type	Arterial
Description	mixed land use, suburban, rural
Tags	here_it_is
Added	24 Nov 2018
Added By	Jem Locquiao
Route Start	28 Jul 2018
Last Modified	24 Nov 2018

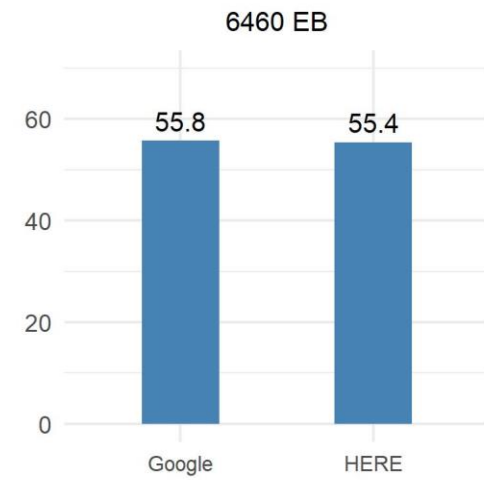


Figure 26: Comparison Highway Segment – SR 9, Rockville/Springdale

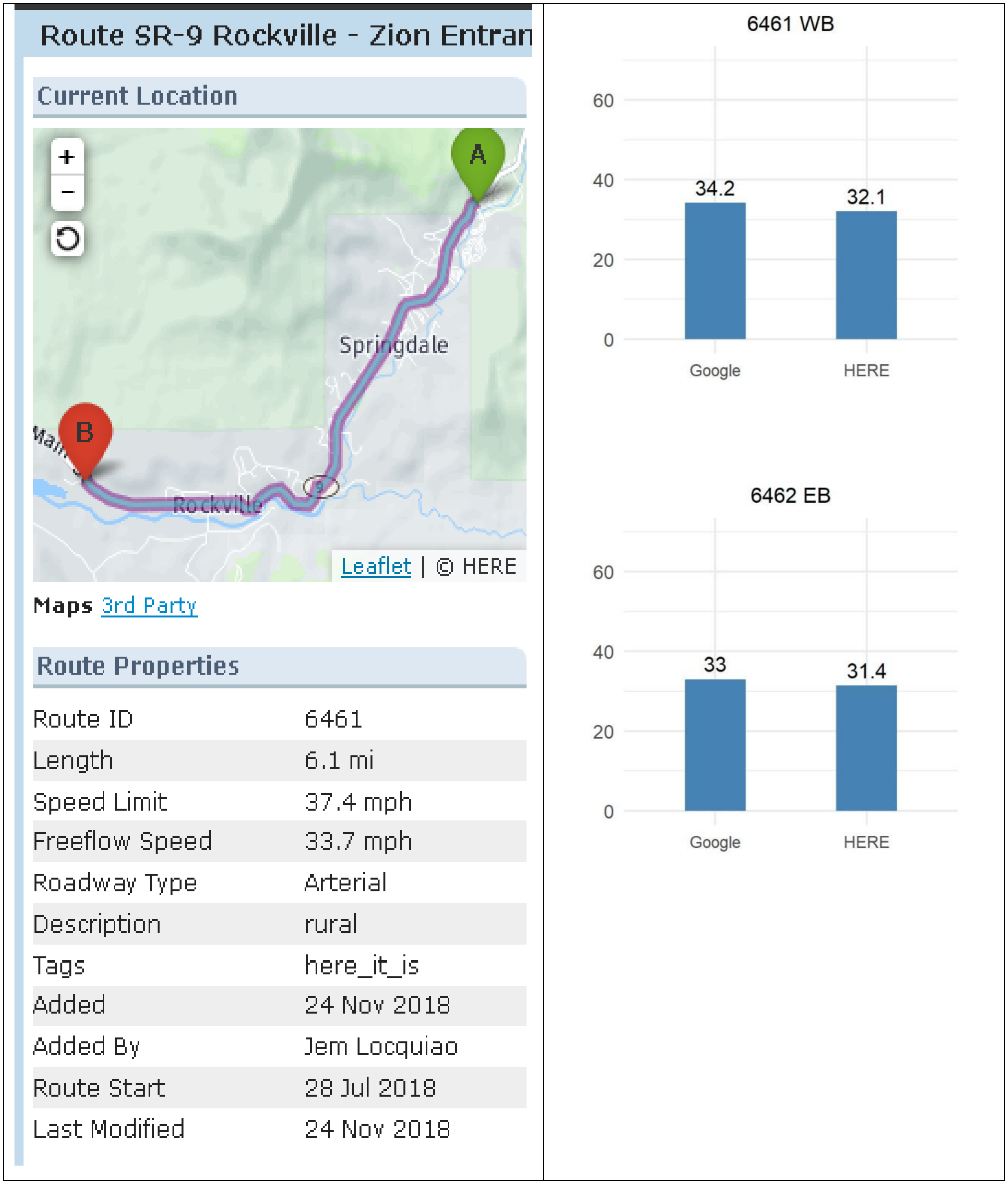
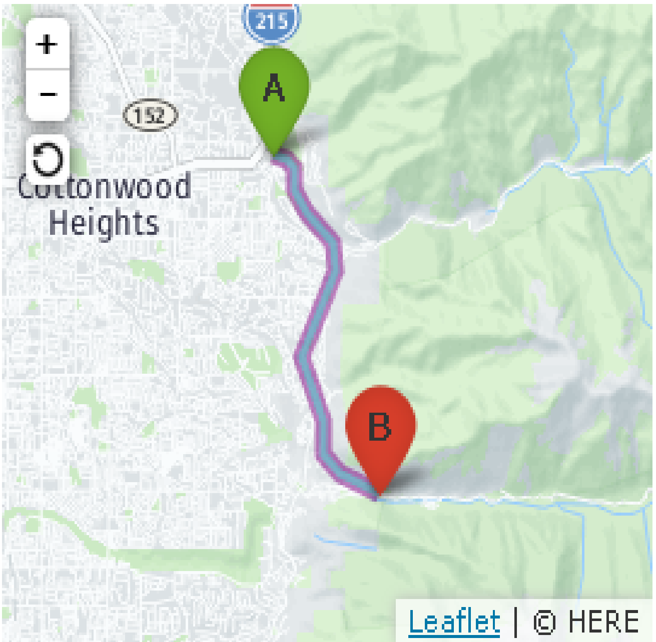


Figure 27: Comparison Highway Segment – 6200S/Wasatch Blvd, Cottonwood Heights

Route 6200 S - Little Cottonwood

Current Location



Maps 3rd Party

Route Properties

Route ID	6463
Length	5.5 mi
Speed Limit	48.6 mph
Freeflow Speed	45.4 mph
Roadway Type	Arterial
Description	urban, suburban
Tags	here_it_is
Added	24 Nov 2018
Added By	Jem Locquiao
Route Start	28 Jul 2018
Last Modified	24 Nov 2018

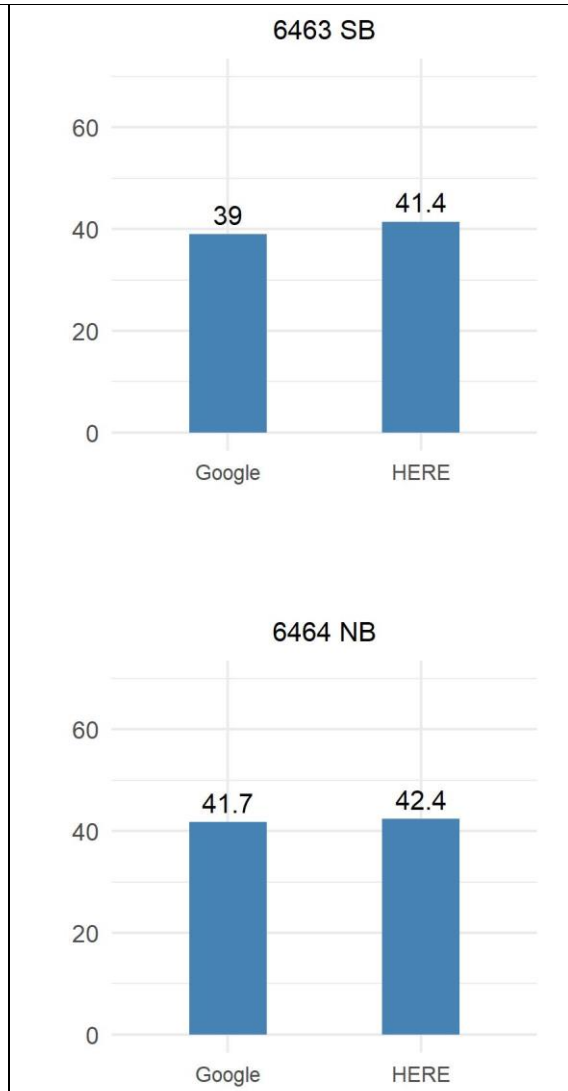
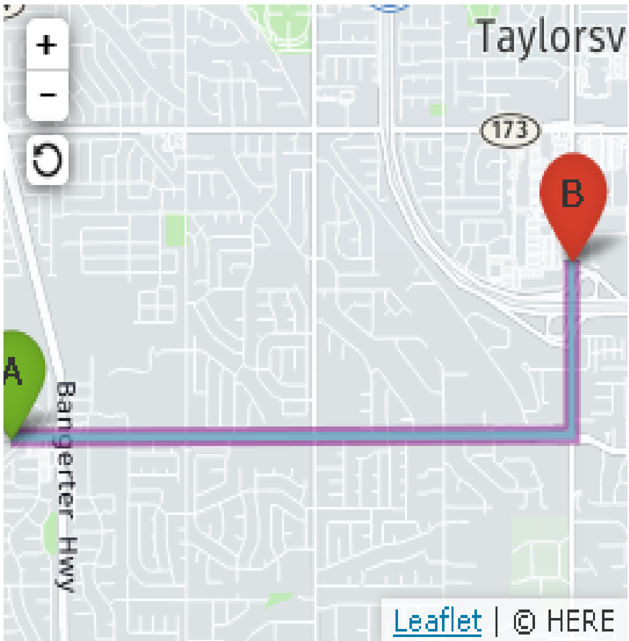


Figure 28: Comparison Highway Segment – 5750 S to Redwood Road, Taylorsville (EB to NB)

Route Bennion WBL 5750 S Redwo

Current Location



Maps 3rd Party

Route Properties

Route ID	6466
Length	2.75 mi
Speed Limit	36.4 mph
Freeflow Speed	34.3 mph
Roadway Type	Arterial
Description	suburban
Tags	here_it_is
Added	24 Nov 2018
Added By	Jem Locquiao
Route Start	28 Jul 2018
Last Modified	24 Nov 2018

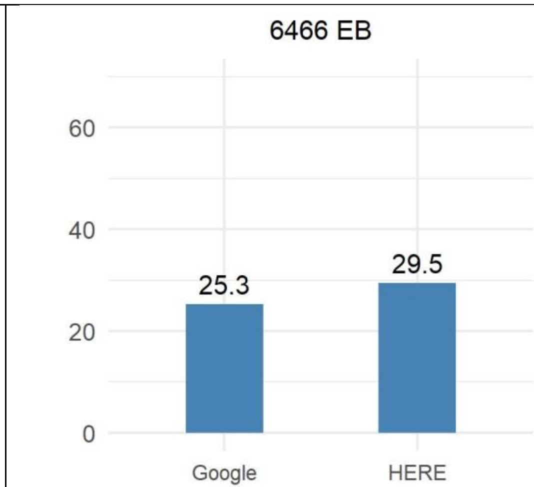
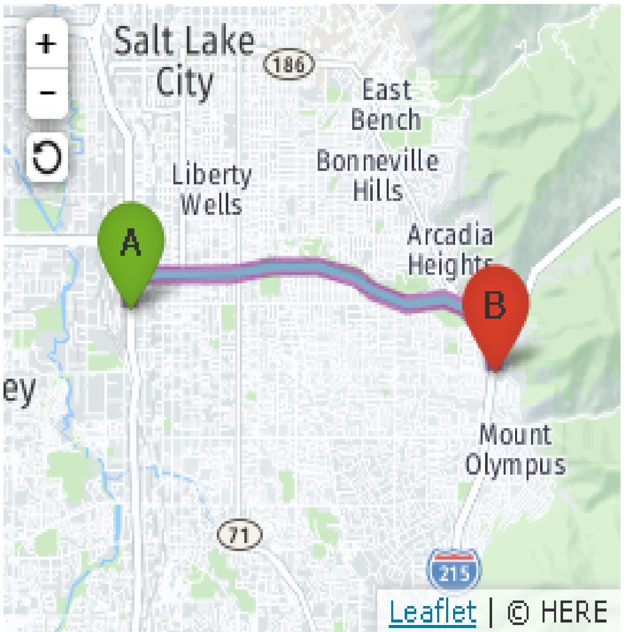


Figure 29: Comparison Highway Segment – I-15 to I-80 (EB)

Route I-15 NB to I-80 EB to I-215

Current Location



Maps 3rd Party

Route Properties

Route ID	6467
Length	6.77 mi
Speed Limit	68.6 mph
Freeflow Speed	63.1 mph
Roadway Type	Freeway
Description	mixed land use
Tags	here_it_is
Added	24 Nov 2018
Added By	Jem Locquiao
Route Start	28 Jul 2018
Last Modified	24 Nov 2018

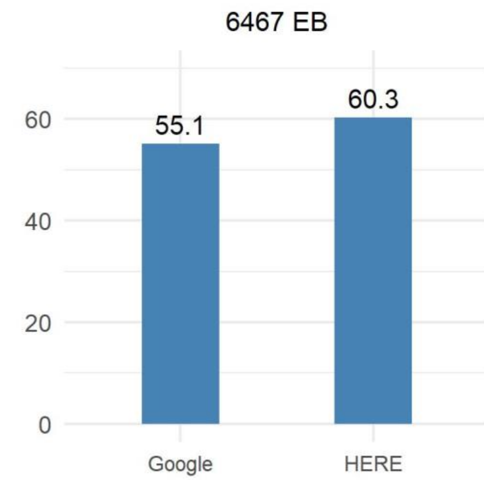
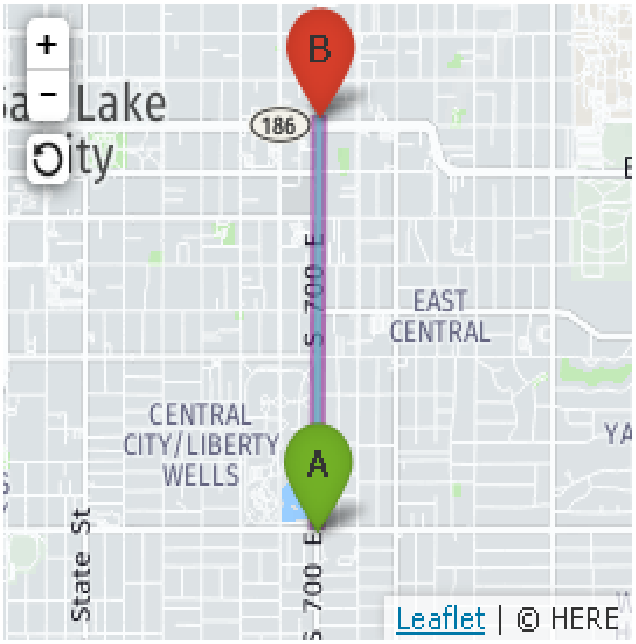


Figure 30: Comparison Highway Segment – 700E, Salt Lake City (NB/SB)

Route 700 E NB (shorter sub-sectio

Current Location



Maps 3rd Party

Route Properties

Route ID	6481
Length	1.35 mi
Speed Limit	39.8 mph
Freeflow Speed	33.7 mph
Roadway Type	Arterial
Description	urban
Tags	here_it_is
Added	26 Nov 2018
Added By	Jem Locquiao
Route Start	28 Jul 2018
Last Modified	26 Nov 2018

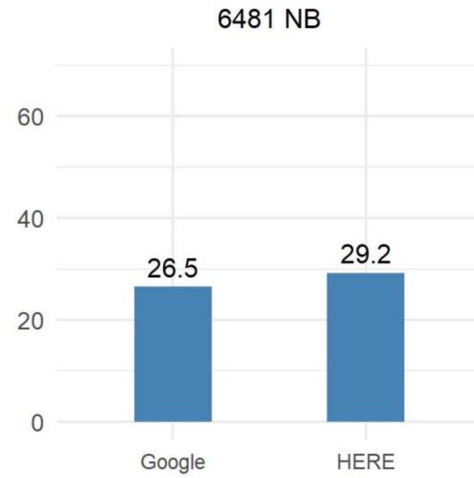
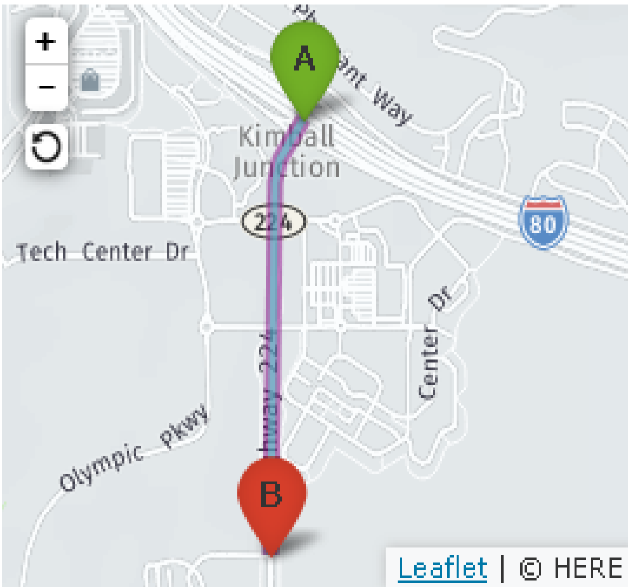


Figure 31: Comparison Highway Segment – SR 224, Park City (NB/SB)

Route SR-224 Snyderville SB (short

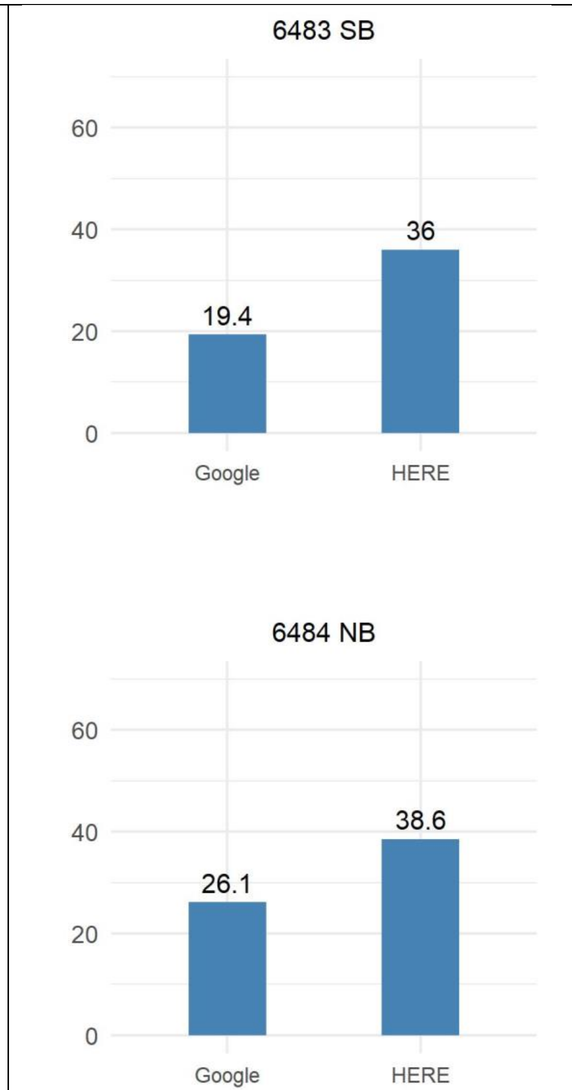
Current Location



Maps 3rd Party

Route Properties

Route ID	6483
Length	0.79 mi
Speed Limit	47.4 mph
Freeflow Speed	43.1 mph
Roadway Type	Arterial
Description	suburban
Tags	here_it_is
Added	26 Nov 2018
Added By	Jem Locquiao
Route Start	28 Jul 2018
Last Modified	26 Nov 2018



3.4 Data Collection

After establishing the 59 highway segments for comparison, the research team set up a data collection system on UDOT's iPems site, labeled here_it_is. A simultaneous API download of the Google Directions travel speeds for corresponding highway segments was established which allowed querying of real-time-travel time through the Google Directions API. The Google data were stored in a database for the trial period, January 21-Febrary 25, 2019. This process resulted in two datasets matched for highway segment and time, enabling a statistical comparison.

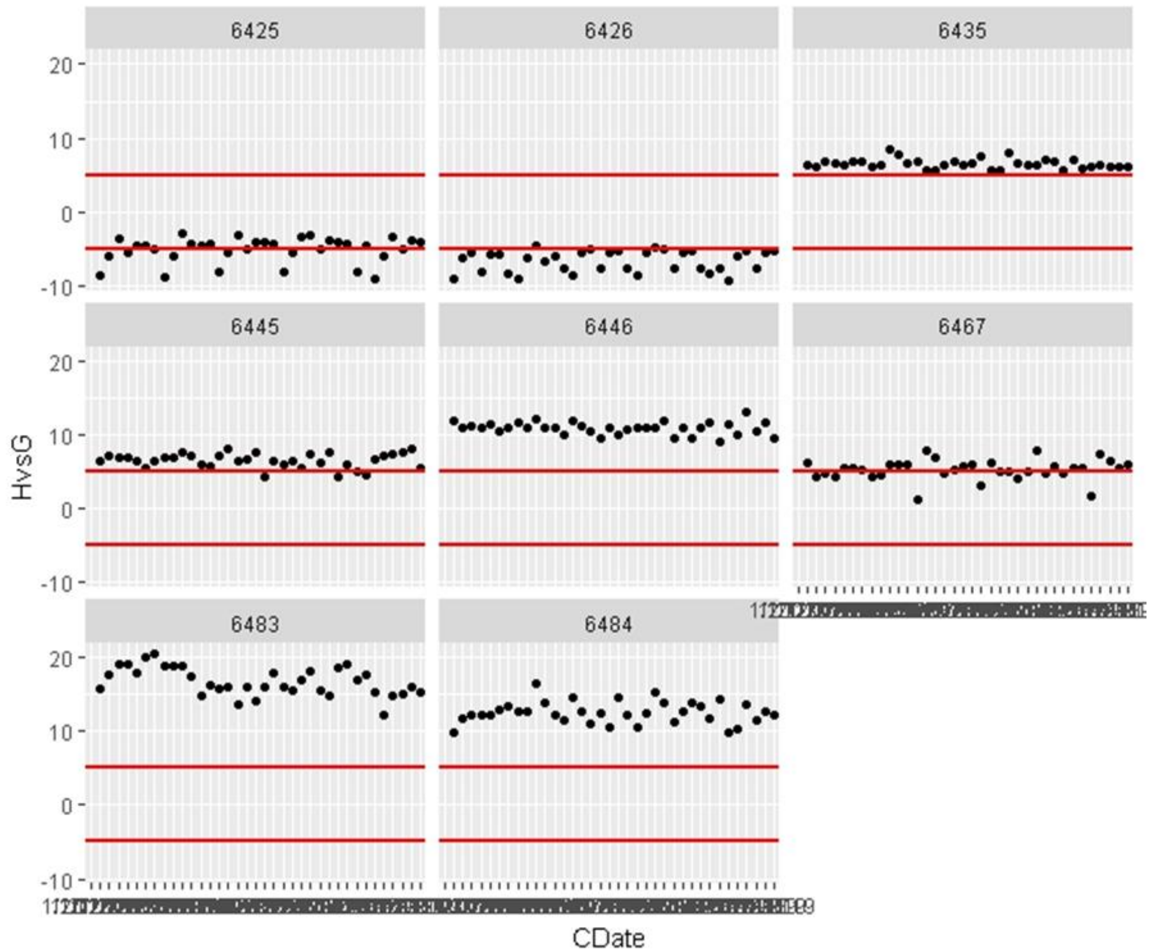
The data were cleaned to remove records with no speed or time data (<1%). The resulting dataset was reviewed for distance, time and speed data between the data sets in order to identify "problematic" routes, where there were systematic differences between the average travel speeds between HERE and Google. A total of eight problematic routes were identified:

- 6425 Lasal EB
- 6426 Lasal WB
- 6435 Ogden SR 35 WB
- 6445 Highway 40 Duchesne EB
- 6446 Highway 40 Duchesne WB
- 6483 SR 224 Kimball Jct. SB
- 6484 SR 224 Kimball Jct. NB
- 6467 I-15 NB to I-80 EB

The systematic nature of the speed comparisons can be seen in

Figure 32 which shows the difference (HERE speed minus Google speed), with a +/- 5mph bracket. In six of these cases, the HERE travel speeds are >5mph higher than the Google travel speeds. In two cases, both of which are for Highway 40 in Duchesne, the Google speeds are >5 mph higher than those obtained from HERE.

Figure 32: Average Travel Speed Comparison for Eight Routes Determined to Have Systematic Speed Errors



These routes are highlighted as they are outliers in the analysis. Under normal travel conditions, the HERE and Google travel times tracked very closely, with minor variability. In these cases, systematic data bias has occurred. In analyzing the underlying causes of the systematic bias for two routes, 6483 and 6484 (SR 224 in Kimball Jct, Park City), it appears that the Google data were canvassing travel times from vehicles that were traveling on SR 224 and also vehicles that were on adjacent roads, some of which were within shopping centers. This would account for the systematically lower speeds for this highway segment. Further analysis of the underlying causes of the systematic bias was not conducted.

In addition, the research team received information on crash incidents on several trial segments on I-15. The purpose of this data was to highlight the time slices in which incidents occurred such that a more specific analysis of travel speeds could be conducted for cases where incident-related delays might have occurred.

3.5 Data Analysis

The two data sets were initially compared for overall average speeds by route. The summary speed data (all days, all times) are shown in Figure 33 and Figure 34, with the problematic routes highlighted in a box.

Figure 33: Average Overall Speeds by Highway Segment, HERE, Google, Bing (1)

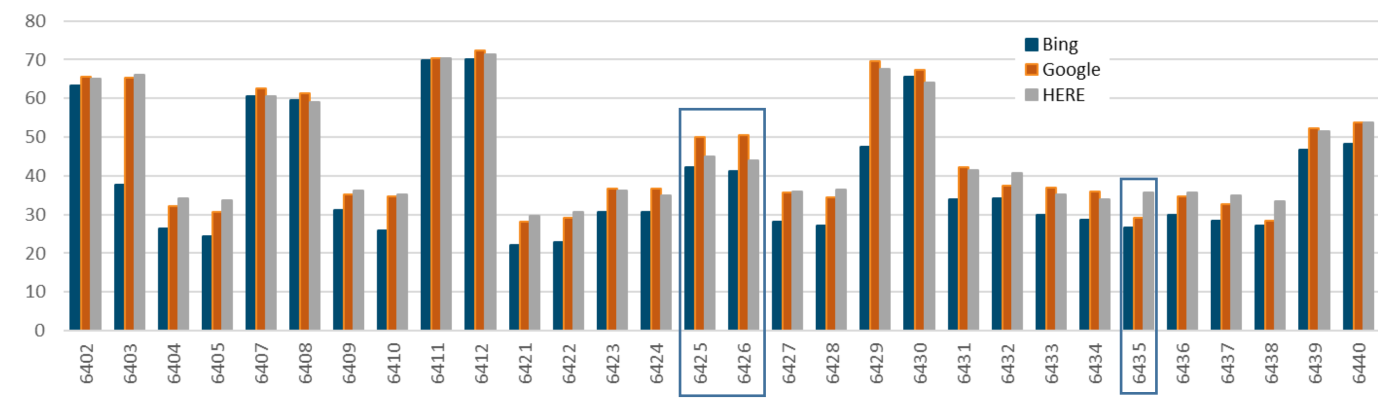


Figure 34: Average Overall Speeds by Highway Segment, HERE, Google, Bing (2)

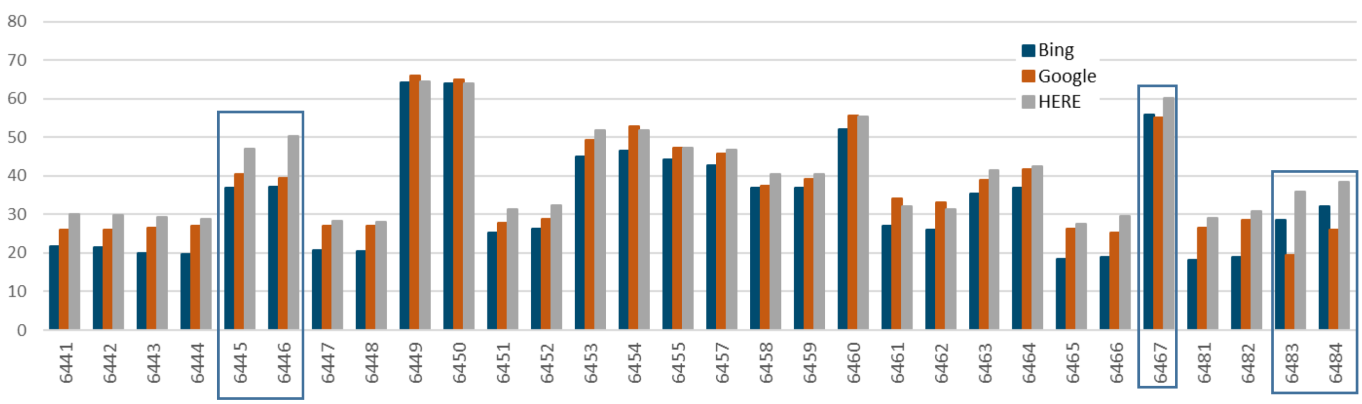


Figure 35 provides another look at the data, showing a speed equivalency line and the comparison of travel speeds for each highway segment, averaged across all 15-minute time slices. The plot shows good agreement of travel speeds. However, lower speed roadways show a tendency for the Google speed estimate to be lower than that provided by HERE (lower left of the plot).

Figure 35: Google vs HERE Average Travel-time estimates, All Routes

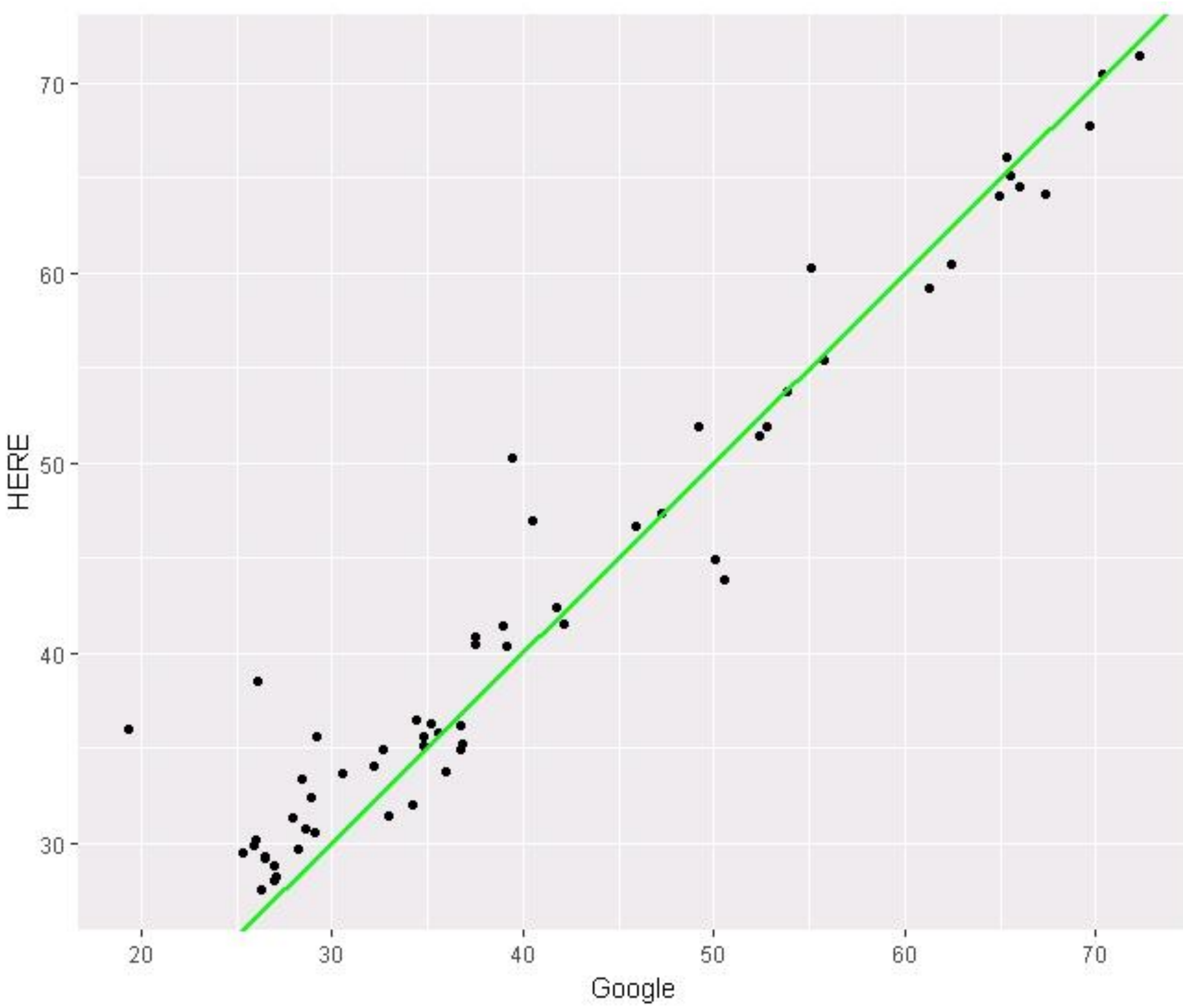
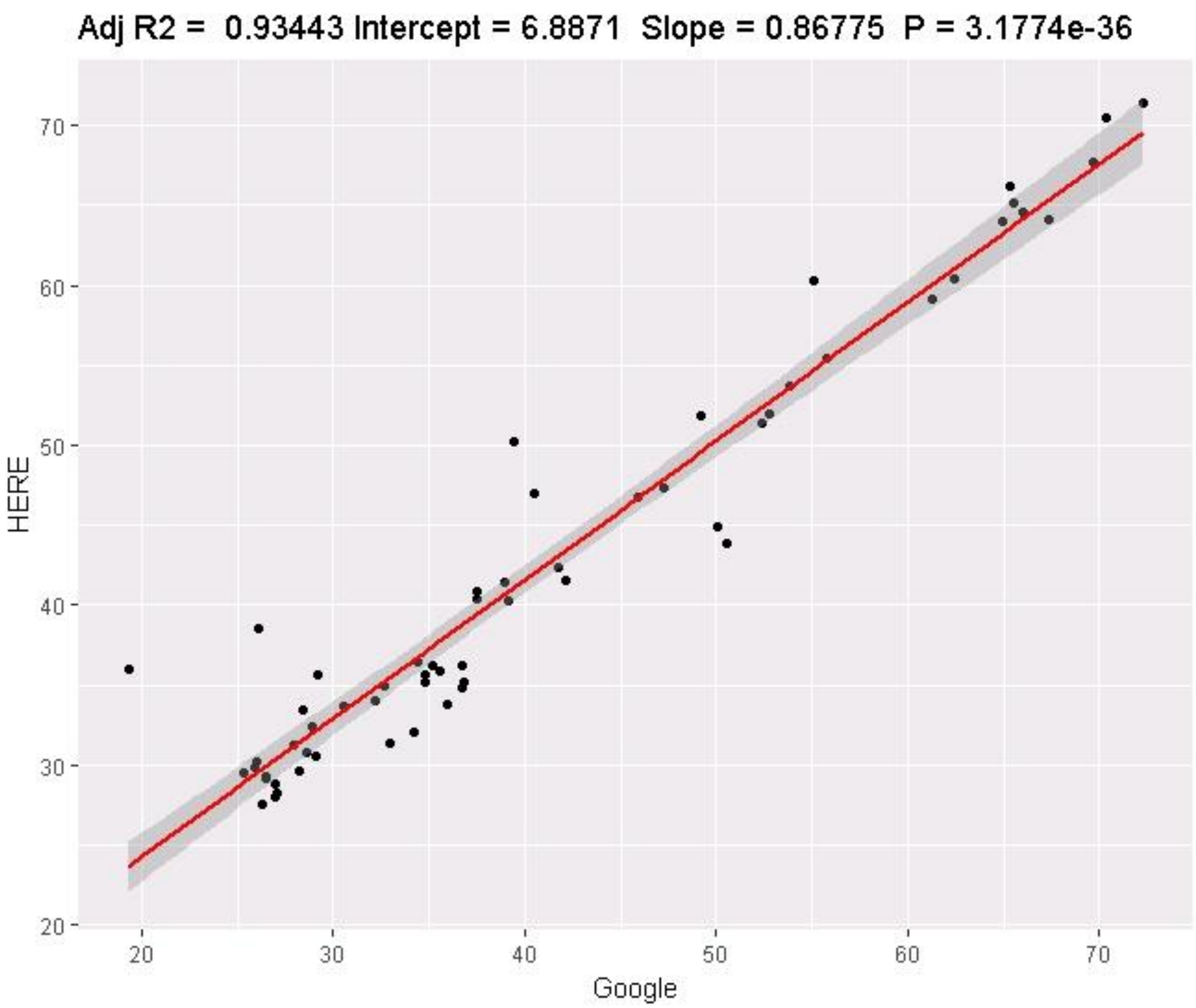


Figure 36 shows a linear model to fit the Google data to the HERE data for all routes, days and times. The fit has an r-squared of 0.93 with a positive y-intercept indicating overall lower speeds for Google than for HERE. The model is significant with a p-value of close to zero.

Figure 36: Statistical Fit of the Google Travel Speed Data to the HERE Travel Speed Data



There are 36 days of data for 59 routes (total of 2124 pairs of route-day). Route-Day pairs with significant average speed differences were compared. Using a threshold of 20% difference in average speed, it was determined that approximately 95% of route-day pairs were not significantly different when comparing HERE and Google.

Routes 6446 (Highway 40 Duchesne), 6483 (SR224 SB), and 6484 (SR224 NB), are the ones with the most average days different from Google and account for more than 95% of the problematic route-day pairs.

Figure 37 compares the three travel-time estimates from Bing, Google and HERE through a box-whisker plot displaying the minimum and maximum speeds, the speed distribution encompassing 50% of the samples (box) and the mean speed. From a visual inspection, the Google data match the HERE data, though tend toward lower speeds.

Figure 37: Average Speeds, All Routes/Days/Times

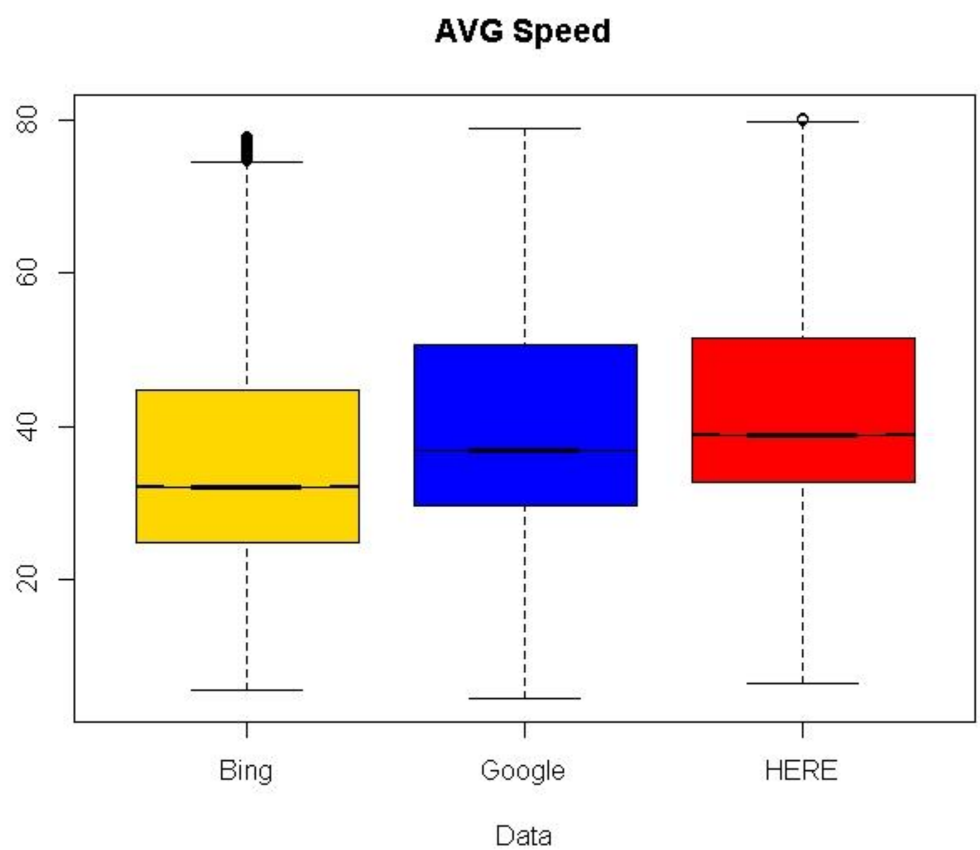
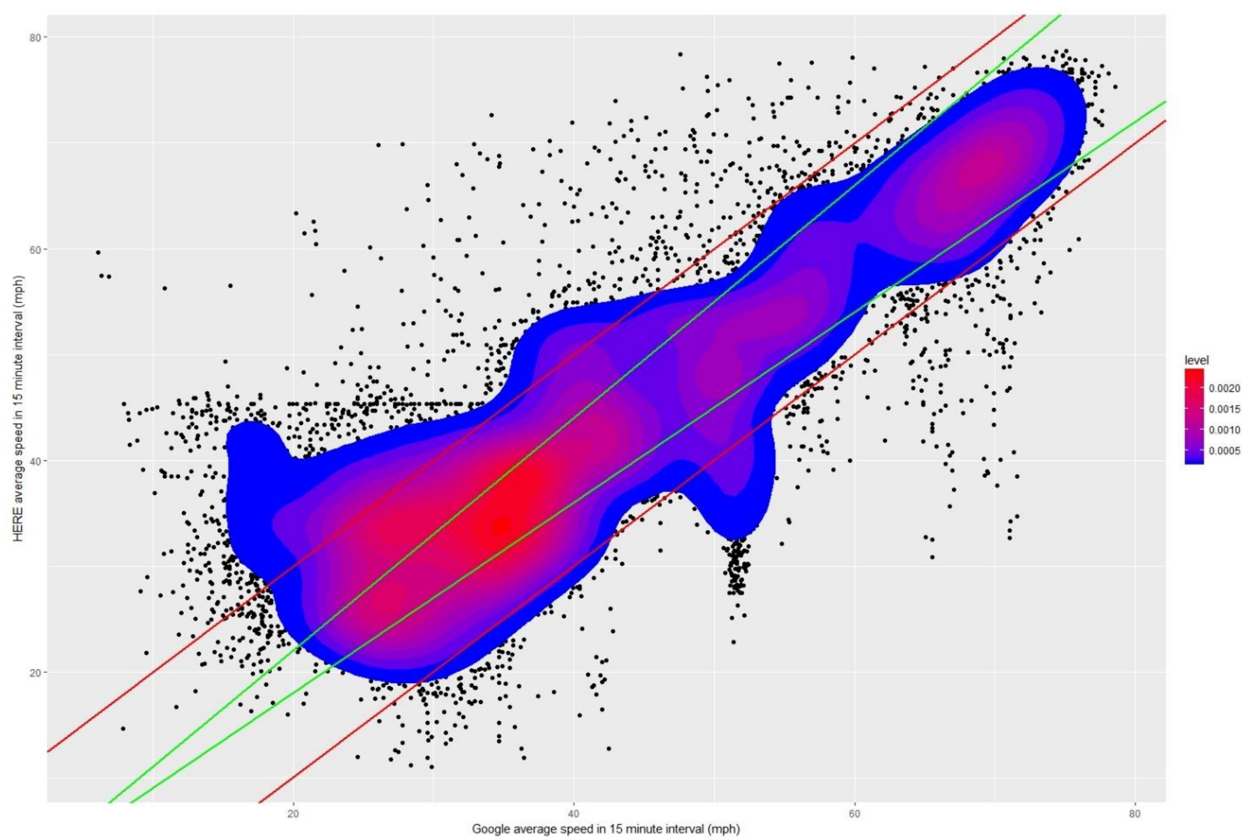


Figure 38 visualizes HERE and Google average speeds (in 15-min intervals) in a 2D density plot. The green lines are lines with intercept = 0 and slope = 0.9 or 1.0, showing the limits of the ratio of HERE speeds to Google speeds. Red lines are lines with intercept = -10 or 10 and slope = 1, showing the boundary for cases where the difference between Google and HERE speeds are less than or greater than 10 mph. The density colors are levels of density. The redder the color is, the more data points (average 15-min interval speeds for Google and HERE). For example, in this graphic, there is a high density of speeds between 30-40 mph and between green lines (slopes 0.9 and 1), meaning that the average speeds are very similar.

Figure 38: Google-HERE 15-minute Average Speed Density Plot



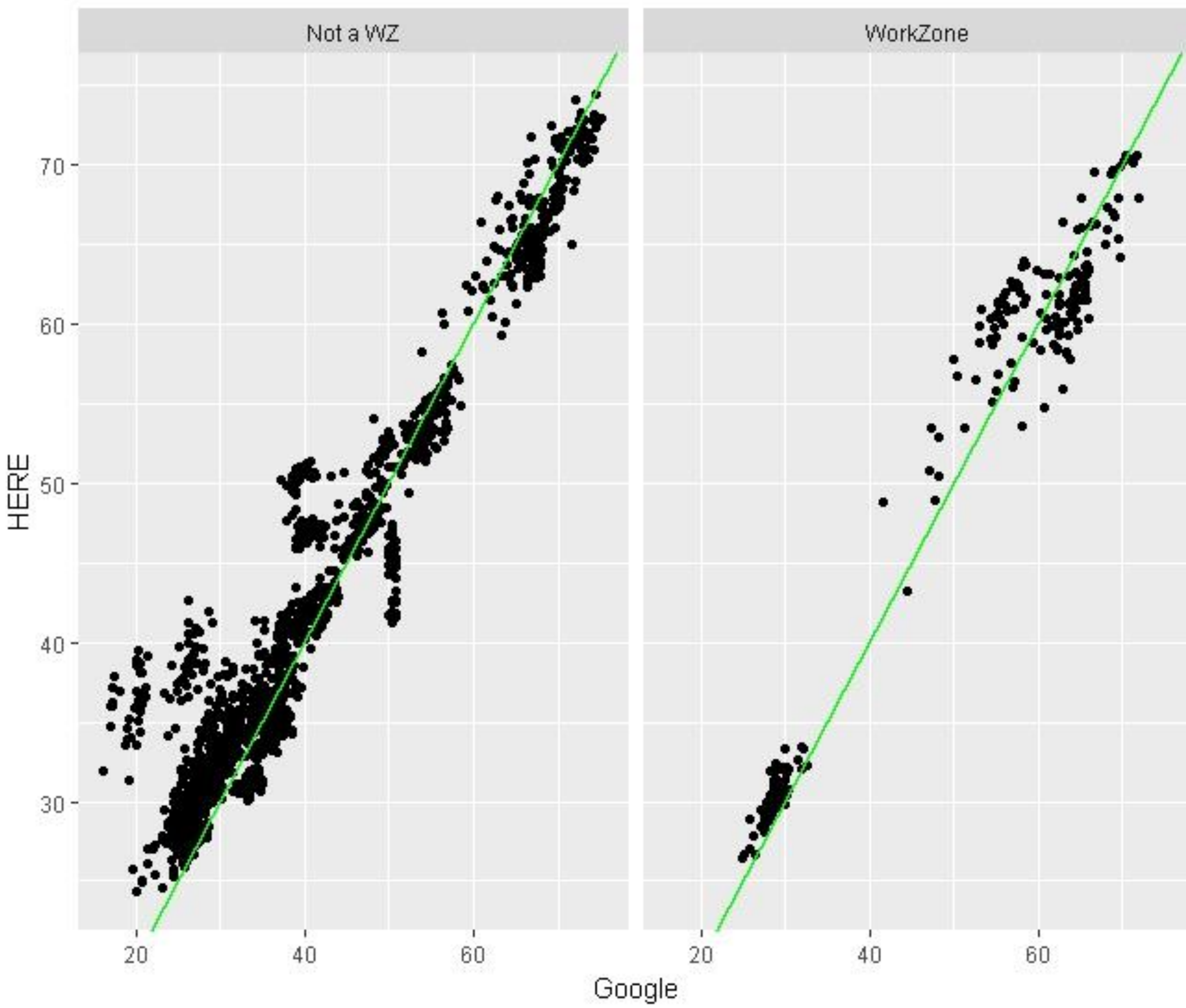
3.6 Comparison of Travel Speed Estimates for Work Zones

An original impetus for this research was to better understand the validity of the HERE data travel-time estimates for work zones. Of the 59 study highway segments (115 unique segment directions), the following were flagged as work zones:

- 6402 – I-15 SB, 900S to 4800S
- 6407 – I-80 WB, Mountain Dell
- 6408 – I-89 EB, Mountain Dell
- 6421 – 9000S/9400S WB
- 6422 -- 9000S/9400S EB
- 6467 – I-15 NB to I-80 EB to I-215 (EB)

Figure 39 shows the HERE-Google average daily travel-time comparisons for work zones and non-work zones across all day-Route pairs. The data indicate that both datasets are in general agreement about work-zone-travel-speed estimates. It should be noted that the work zones were not field checked to determine the extent of the work zone or their impact on traffic operations.

Figure 39: Comparison of Travel Speed Estimates for Work Zones vs Non-Work Zones



3.7 Comparison of Travel-Speed Estimates for Identified Crash Locations and Periods

The crash dataset contains crash event ID, logging time, clearance time, street, mile post, and severity for 647 crashes that occurred from January 2019 to February 2019 in four counties: Weber, Davis, Salt Lake, Utah, and Juab within Utah. Crash-logging time and clearance time represent the time when the crash was logged into the crash-record database and the time when the crash was cleared from the roadway, respectively. Of the larger dataset (647 crashes) a total of 113 crashes were identified as occurring within the here_it_is study segments 6402 and 6407 on I-15 Northbound or Southbound from mileposts 301 to 307, as shown in

Figure 40 and listed in Table 1.

Figure 40: I-15 Reported Locations of I-15 Crashes Analyzed for HERE-Google Travel Times

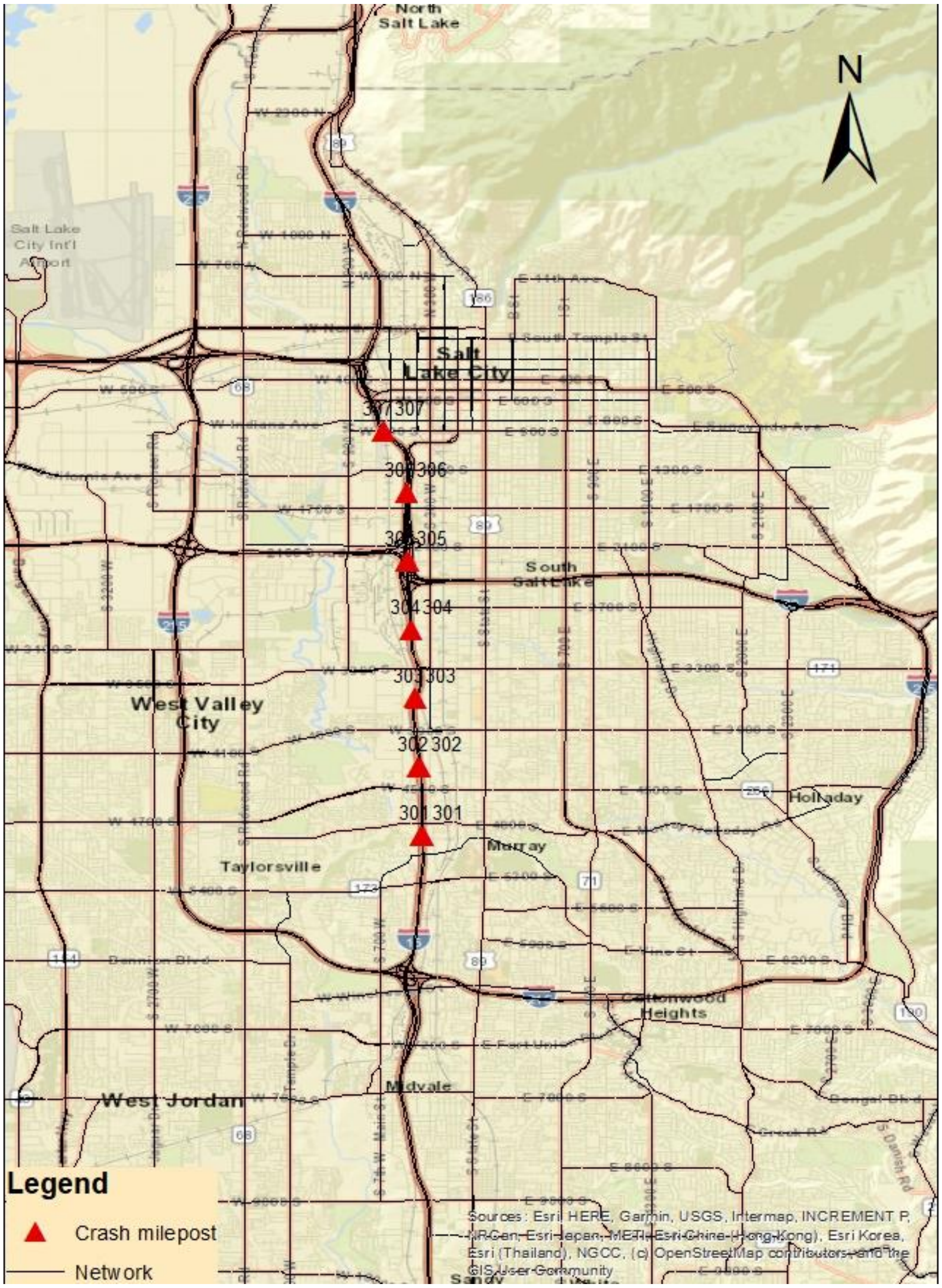


Table 1: Milepost Location and Direction of I-15 Crashes Analyzed for Travel-time comparisons

Main Street	Milepost	Direction	Number of Crashes
I-15	301	NB	8
I-15	301	SB	2
I-15	302	NB	9
I-15	302	SB	4
I-15	303	NB	13
I-15	303	SB	4
I-15	304	NB	14
I-15	304	SB	11
I-15	305	NB	14
I-15	305	SB	10
I-15	306	NB	7
I-15	306	SB	6
I-15	307	NB	8
I-15	307	SB	3

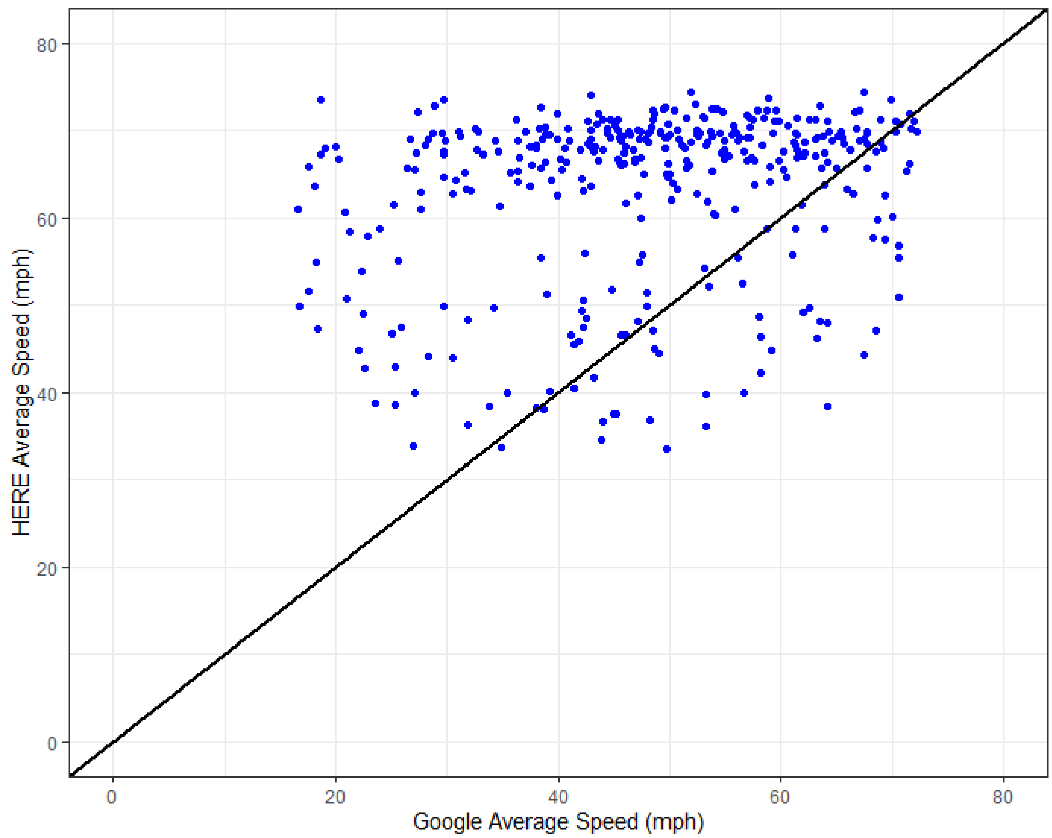
Since the actual crash-occurrence time can be different from crash-logging time, an algorithm was developed to estimate crash-occurrence time. The algorithm searches for significant reduction in average speed during a 2-hour period prior to the crash-logging time. Each 15-minute average speed is compared to the previous interval's average speed to find the first interval with a 10% or greater reduction in average speed, which is then considered the "Crash-Occurrence Time".

Google and HERE average speeds for 15-min time intervals with crash events are compared to investigate the accuracy of two datasets in average speed estimation. Figure 41 illustrates comparison of Google and HERE average speed estimations for intervals with crash events. As shown in Figure 41, the HERE algorithm tends to overestimate average speed compared to the Google algorithm. There are, however, several observations that Google algorithm estimated faster average speeds than HERE algorithm.

To further analyze the accuracy of average speed estimation, we focused on the two cases. Case 1 includes 15-minute intervals during crash events where HERE estimated faster average speeds for most of 15-min intervals associated with the crash event and Case 2 contains 15-minute intervals during crash events where Google tended to estimate faster average speeds for most of time intervals.

A total of 337 15-minute intervals are represented over the 113 crash incidents. Of these, there are 240 instances of 15-minute intervals with significantly higher HERE average speeds (71%), 27 cases with significantly higher Google average speeds (8%), and 70 cases (21%) with approximately the same average speed estimates (HERE travel times within +/-10% of Google-estimated travel times).

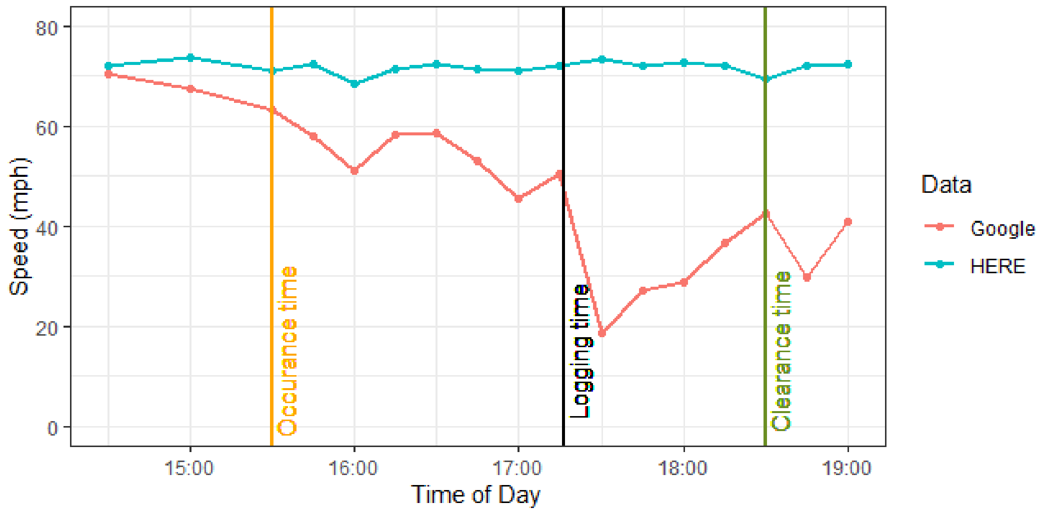
Figure 41:Comparison of HERE and Google Average Speeds for 15-Minute Intervals During Crash Incidents



Case 1

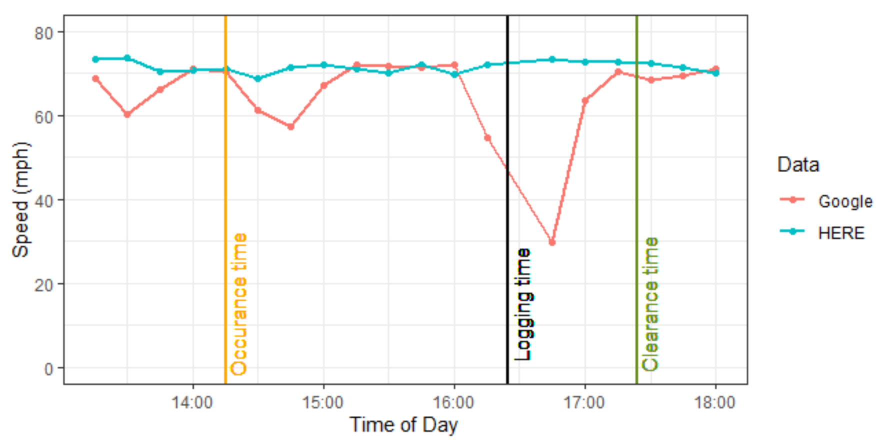
Case 1 consists of crash events that HERE algorithm estimated faster average speed for most of 15-min time intervals. We focused on 5 different crashes as examples of case 1 and compared Google and HERE estimates before and after crash occurrence. Figure 42 shows the average speed profiles associated with these crashes. As illustrated in the figure, while Google dataset reports significant speed reduction after crash occurrence, HERE estimates are not impacted significantly, and average speed trends remain stable even after crash occurrence. Consequently, it seems that Google dataset can better reveal the impact of crash events on average speed and travel time. Smaller sample size of HERE dataset might explain its lower accuracy in average speed estimation compared to the Google dataset.

Figure 42: Speed Profiles Associated with I-15 Crash Incident 87950, Case 1



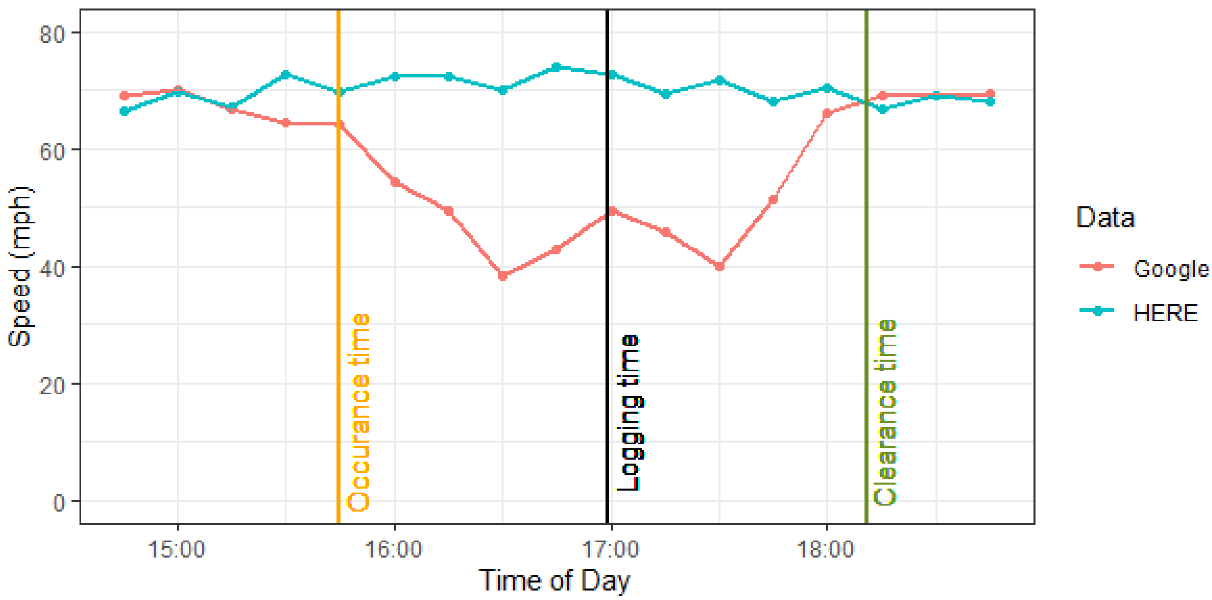
Crash ID: 87950, February 1, 2019 (I-15 NB at MP302)

Figure 43: Speed Profiles Associated with I-15 Crash Incident 88554, Case 1



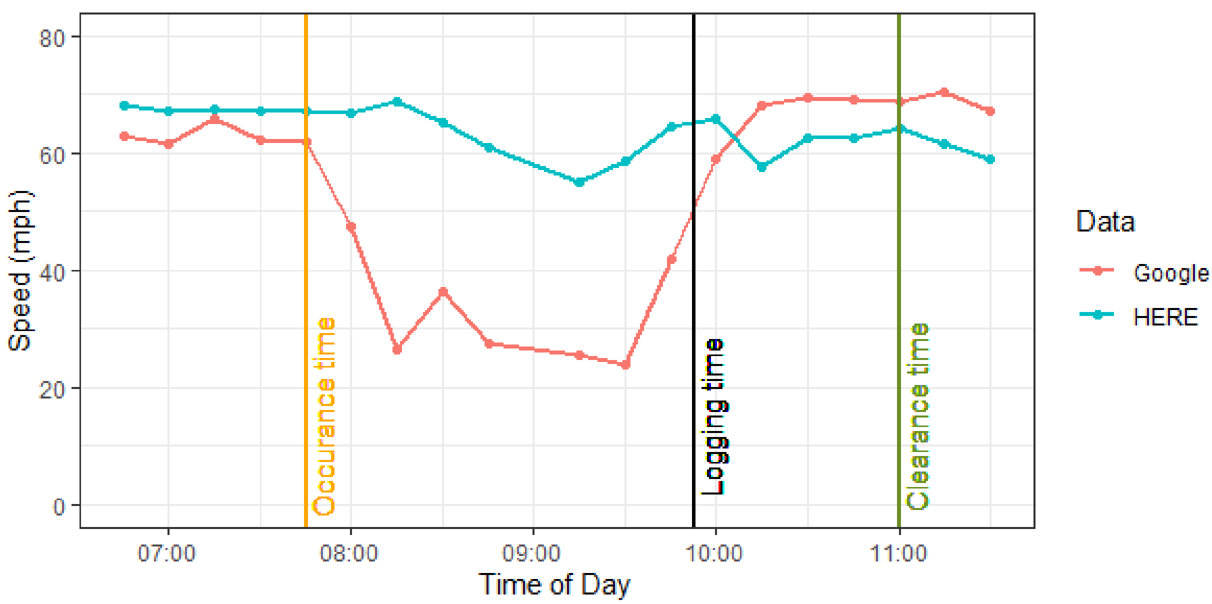
Crash ID: 88554, February 9, 2019 (I-15 NB at MP307)

Figure 44: Speed Profiles Associated with I-15 Crash Incident 87670, Case 1



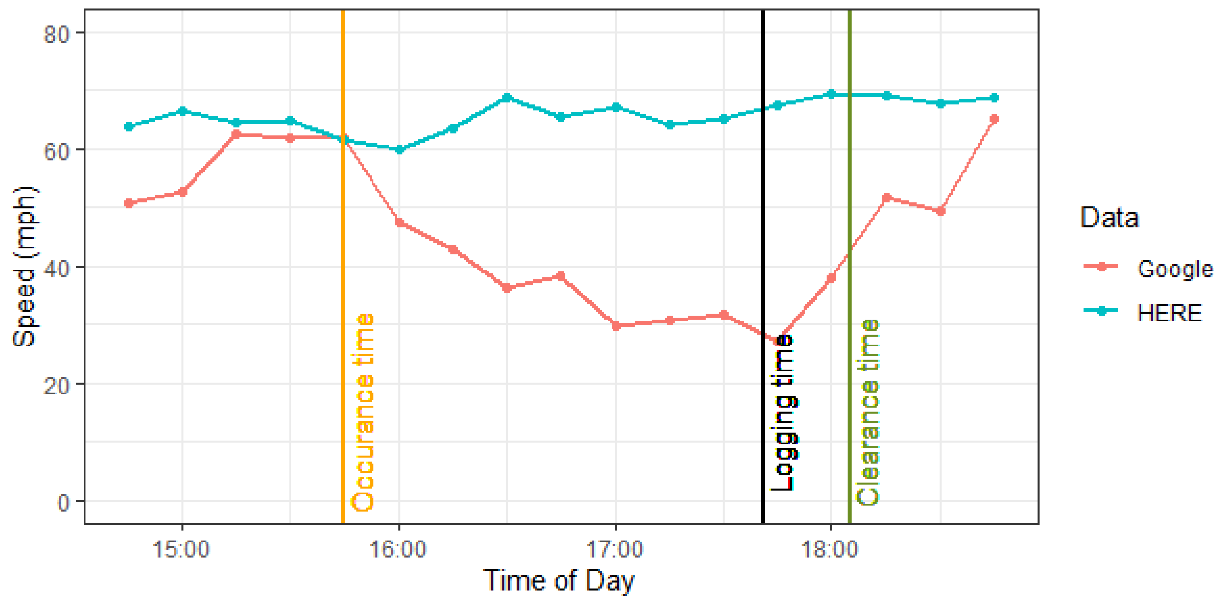
Crash ID: 87670, January 25, 2019 (I-15 SB at MP307)

Figure 45: Speed Profiles Associated with I-15 Crash Incident 89454, Case 1



Crash ID: 89454, February 22, 2019 (I-15NB at MP304)

Figure 46: Speed Profiles Associated with I-15 Crash Incident 88845, Case 1



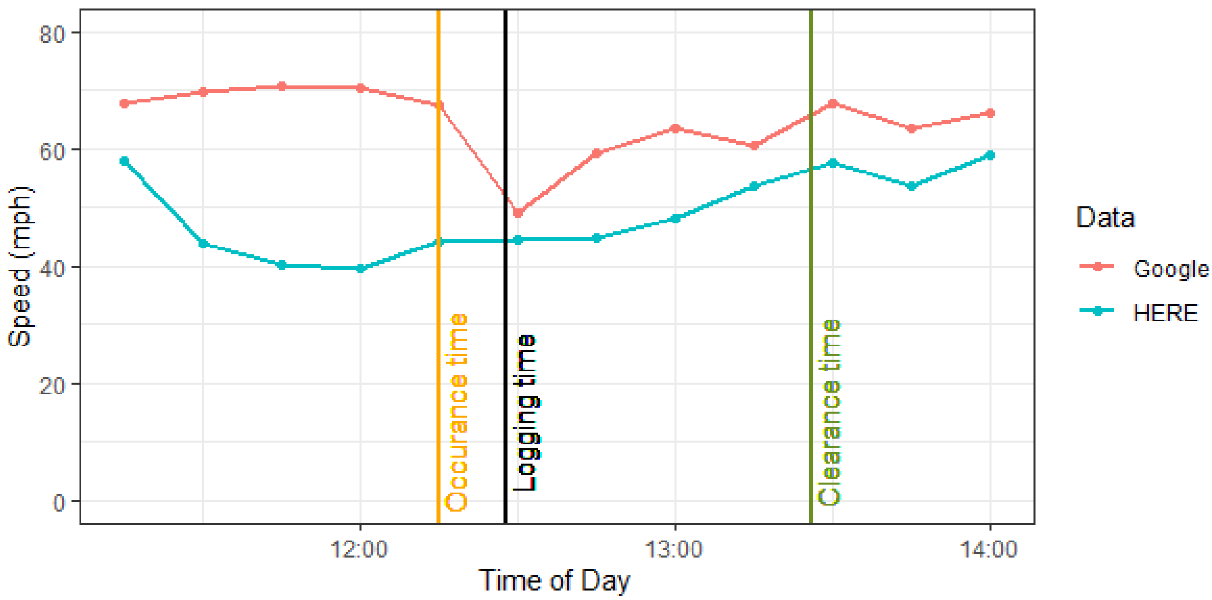
Crash ID: 88845, February 13, 2019 (I-15SB at MP305)

Case 2

Case 2 consists of crash events where Google estimated faster average speeds for most of 15-min time intervals. We focused on 2 different crashes as examples of Case 2 and compared Google and HERE speed estimates before and after crash occurrence.

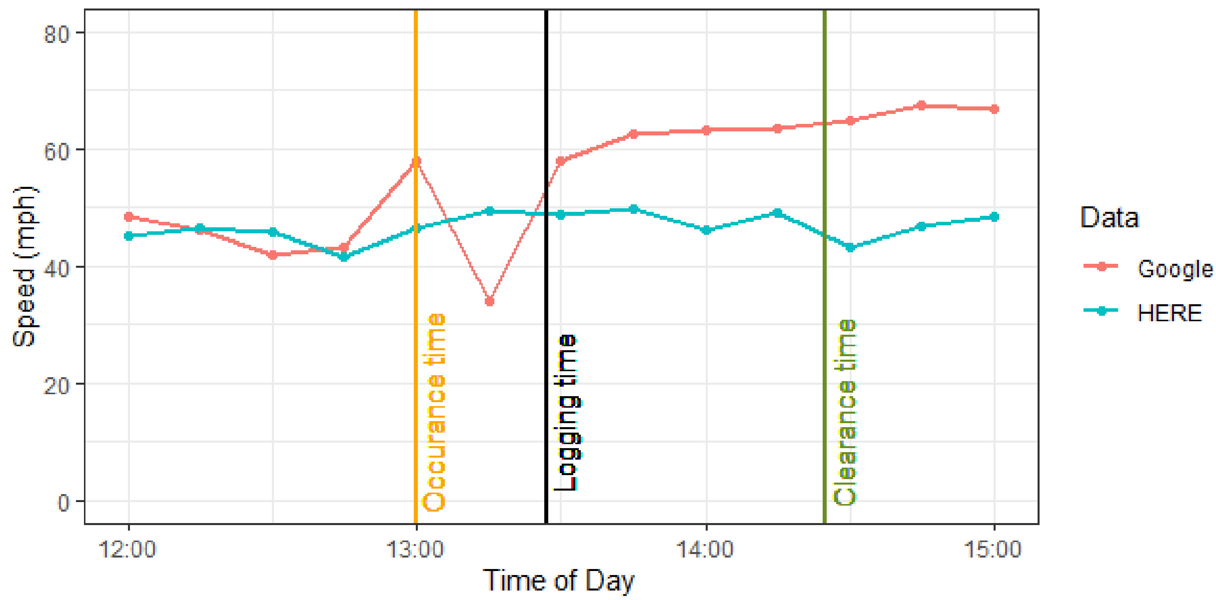
Figure 47 shows the average speed profiles associated with these crashes. As illustrated in the figure, although the HERE dataset estimated lower speeds for most of time intervals associated with crash events, speed estimates are not significantly altered after crash occurrence and the average speed trend remains stable before and after crash occurrence. The Google dataset on the other hand, estimated faster average speeds but average speed is significantly reduced after the crash occurrence.

Figure 47: Speed Profiles Associated with I-15 Crash Incident 88128, Case 2



Crash ID: 88128, February 5, 2019 (I-15SB at MP306)

Figure 48: Speed Profiles Associated with I-15 Crash Incident 88334, Case 2



Crash ID: 88334, February 6, 2019 (I-15NB at MP301)

3.8 Summary

The comparison of HERE and Google travel-time data shows that, under normal operating conditions without congestion such as that which would be associated with a crash event, travel-time estimates from both sources are very comparable. The crash analysis tends to show greater sensitivity to spot congestion related to responding to and clearing the crash incident. For 113 crashes on I-15 during the study period, 71% of the 15-minute time intervals logged between the estimated crash occurrence and clearing showed Google travel times lower than those reported by HERE.

Without information on the market penetration of devices reporting to HERE vs. those reporting to Google, it is impossible to state with certainty that the Google data are more accurate. However, it is fair to assume that more devices are reporting to Google through the Google Directions service than to HERE, which data comes primarily from in-vehicle navigation devices. Neither the HERE nor the Google service has provided information on the fraction of the traffic stream their service is sampling.

4. CONCLUSION

The research confirms the validity of the HERE travel-time estimates across different functional class highways in Utah for periods under normal operation. An analysis of speed estimates within work zones suggests that travel-time estimates from both sources track very well. This may be due to learning algorithms that adjust travel-speed estimates based on consistently recurring patterns (i.e. from recent days and weeks).

Under periods of non-recurring congestion, however, such as during crash incidents, the HERE data tend to over-estimate travel speeds when compared to the data obtained from the Google Directions API. Caution should be exercised when using the HERE data to estimate travel speeds for non-recurring congestion events.

REFERENCES

I-95 Corridor Coalition Vehicle Probe Project: Validation of INRIS Data, July-September 2008. I-95 Corridor Coalition. January 2009.

Validating the National Performance Management Research Data Set (NPMRDS) for South Dakota. Study SD2013-08. March 2018.

Considerations for Calculating Arterial System Peformance Measures in Virgina. Final Report VTRC 17-R2. February 2017