

**Exploring Naturalistic Driving Study (NDS) Data and Roadway  
Information Database (RID) for Emerging Applications in Traffic Safety**

by  
Sherif Ishak  
Osama A. Osman  
Grace Ashly  
Matthew Theriot  
Mengqiu Ye  
Peter R. Bakhit

Department of Civil and Environmental Engineering  
Louisiana State University  
Baton Rouge, LA, 70803

NCITEC Project No. 2016-11

conducted for

NCITEC

August 2016

## 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 Department of Transportation University Transportation Centers Program, in the interest of information exchange. The U.S. Government assumes no liability for the contents or use thereof.*

## ABSTRACT

Recently, several transportation-related databases were introduced and archived in order to analyze the behavior of the different components of the transportation system. The Naturalistic Driving Study (NDS), which provides data on drivers' behavior during their typical commutes, and the Roadway Information Database (RID), which includes information about the geometric characteristics of the roads where the NDS participants drove, are two of the main data sources that have recently grabbed researchers' attention. These two databases provide ample amount of data that have the potential for transportation research in the area of traffic safety and human factors. While many researchers studied drivers' behavior using the NDS data, few studies were performed on the link between drivers' behavior and the characteristics of roadways. Thus, in this study, the research team thoroughly explored the RID dataset and identified its ability to link to NDS. Specifically, analysis was conducted to map RID and NDS data using ArcGIS. In addition, researchers provided several potential research topics integrating the combined RID and NDS database. A short list of potential applications includes quantifying the effect of warning and regulatory signs on driver compliance, measuring the effectiveness of safety campaigns with driver compliance, examining statistically significant clusters of crash and near crash events taking into account roadway conditions, and others.



## **ACKNOWLEDGMENTS**

The research work reported was sponsored by the National Center for Intermodal Transportation for Economic Competitiveness (NCITEC).



# TABLE OF CONTENTS

ABSTRACT.....	III
ACKNOWLEDGMENTS .....	V
TABLE OF CONTENTS.....	VII
LIST OF FIGURES .....	IX
INTRODUCTION .....	1
OBJECTIVE .....	3
SCOPE .....	5
LITERATURE REVIEW .....	7
100-Car NDS Studies.....	7
SHRP2 NDS and RID Studies .....	8
METHODOLOGY .....	11
Data Description .....	11
NDS Dataset.....	11
RID Dataset.....	13
Data Extraction .....	14
The Process of Linking RID and NDS .....	14
DISCUSSION OF RESULTS.....	16
Exploration of RID Database.....	16
Mapping RID in ArcGIS.....	21
Applications of Street View and Bird’s Eye View .....	22
NDS and RID Road Alignment .....	24
Potential Topics .....	26
CONCLUSION.....	29
RECOMMENDATIONS .....	30
ACRONYMS, ABBREVIATIONS, AND SYMBOLS .....	31
REFERENCES .....	33





## LIST OF FIGURES

Figure 1 Crash Data and Near Crash Selection vs Speed .....	7
Figure 2 Driver Glance Duration's Impact on Crashes .....	9
Figure 3 DAS Channels (2).....	12
Figure 4 RID Highway Coverage .....	14
Figure 5 Part of Lighting Information of RID Data.....	16
Figure 6 RID Sample Data in ArcGIS Program .....	17
Figure 7 FUGRO Software Screenshot.....	19
Figure 8 Screenshot of Data Input in ArcMap Program.....	22
Figure 9 Roadway data for Tampa, Florida.....	22
Figure 10 Bird's Eye View in ArcMap.....	23
Figure 11 Street View in ArcMap.....	24
Figure 12 Newly and Previously Mapped Points .....	25
Figure 13 Newly Mapped Points .....	26



## INTRODUCTION

With the advancements of technology and data storage capability, huge amounts of transportation data are now being collected and stored with higher frequencies, which introduced the concept of Big Data to the transportation field. Big Data, in general, is a terminology that points to structured and unstructured data that are collected at high frequency. By definition, structured data are those having a fixed field within a file. This includes common data stored in relational databases and spreadsheets. Unstructured data are those that cannot be contained into databases and spreadsheets such as, photos, videos, and text messages. Recently, several transportation-related databases were introduced including the Naturalistic Driving Study (NDS) and the Roadway Information Database (RID). The two databases mainly focus on human factors and safety, which provide ample amount of data that have the potential for transportation research in the areas of human factors and traffic safety.

The purpose of the Strategic Highway Research Program 2 (SHRP 2) NDS data is to provide researchers with realistic information on driving behavior, which allows them to make more informed decisions on how to improve drivers' safety. The NDS data affirms the importance of the study by showing the fact that 80 percent of crashes recorded were caused by drivers' inattention (1). This establishes a demand for more naturalistic data studies to obtain more crash data and determine what factors could lead to distracted driving. There are three phases to the Naturalistic Driving Study, including conducting test planning activities (Phase 1), conducting a field test (Phase 2), and preparing for a larger scale field data collection effort (Phase 3). Phase 1 is comprised of 15 tasks that are needed to plan the NDS study. These include establishing intelligent vehicle initiative data needs, developing test requirements, selecting candidates test areas and evaluating crash frequency data among others (1). Phase 2 can be further broken down into 10 goals. The ten goals to be accomplished by the field test include characterization of crashes, near crashes, and incidents, quantification of near-crash events, and characterization of driver inattention among others. Phase 3 includes evaluating the performance of the hardware, sensors, and data collection system, along with the performance of the data reduction plan, triggering methods, and data analysis (1).

The Roadway Information Database (RID) is a collection of roadway data compiled by the Center for Transportation Research and Education of Iowa State University (CTRE). The roadway information is gathered from state highway departments and private sources. The private sources use Furgo Roadware equipped vans to measure roadway characteristics along the highways used in the SHRP 2 study (2). The RID information is considered an additional critical resource for analyzing the NDS data. By linking the RID data to the NDS data, researchers will be able to study the impacts that roadway design has on drivers' behavior. The

design of the database was carried out in three phases. First, a standard road network was designed using Esri StreetMap Premium for each of the six locations. This network used link-node design and GPS coordinates to portray roadway segments and intersections. Second, a framework route system was created, which stored each route as an individual linear feature. This allowed researchers to determine the distance along roadway segments in which events occurred. The final step involved locating these events using a route/ measure linear referencing method (LRM). This model for the RID database was based off the Unified Network-Transportation network data model (UNETRANS) proposed by Butler in 2008 (3).

The exploration of the Naturalistic Driving Study (NDS) with respect to the Roadway Information Database (RID) requires researchers' attention. While many researchers studied drivers' behavior using the NDS data, not many studies were performed on the link between drivers' behavior and the geometric characteristics of roads, which is now possible with the introduction of the RID database. Thus, in this study, these databases will be explored to identify the potential research opportunities they offer.

## **OBJECTIVE**

Transportation databases are growing over time, especially with the introduction of the Naturalistic Driving Study (NDS) data and the Roadway Information Database (RID). The NDS and the RID databases are part of the second Strategic Highway Research Program (SHRP2) funded by the USDOT. The NDS data provides data on the normal drivers' behavior in their own vehicles while driving their typical commutes. Whereas, the RID data includes information about the geometric characteristics of the roads where the naturalistic driving study participants drove. This includes roadway location, lane widths, guardrails, grade, median presence, curvature, intersection characteristics, and environmental data such as time of day and weather conditions.

These databases provide great opportunity for researchers to investigate the safety aspects of transportation networks. While there has been significant research performed using the NDS database, very limited research was conducted with the RID. In essence, the primary goal of this study is to explore the RID database and investigate the elements in NDS database that can be linked to the RID data. To accomplish this goal, the following specific objectives will be achieved: (1) explore the RID database, (2) connect the RID and NDS databases for potential safety studies, and (4) identify the possible research questions these data can answer.



## **SCOPE**

The study is limited to the use of NDS and RID databases. Only a small amount of NDS dataset could be accessed from website. Most of the original data is stored at Virginia Tech Transportation Institute (VTTI). Interested research institutions need to purchase the necessary data from VTTI after completing a Data Use License and Contract.





## LITERATURE REVIEW

This section lists the most recent research studies related to RID and NDS. The first part of this section focuses on VTTI's 100-car NDS study, while the second part includes research studies related to RID and SHRP 2.

### 100-Car NDS Studies

Jade Montgomery et al's report (4) analyzed the impact that a driver's age and gender has on their ability to brake in normal driving situations. A total of 527,861 braking instances from 11,503 trips were collected. The results determined that males Time to Collision (TTC) at braking was 1.3 seconds lower on average than women's TTC. This showed a significant difference in TTC ( $P < 0.01$ ) for gender. The results for age showed participants aged over 30 had a TTC at braking of 1.7 seconds higher than participants aged under 30 years. This was also a significant result ( $P < 0.01$ ). With such a significant difference in TTC for both age and gender it was determined Forward Collision Warnings (FCWs) should be designed based on the demographic of their particular vehicle to maximize the effectiveness of this feature.

A study designed to validate near crash events was performed by Jenny Jonasson (5) at Chalmers and Gothenburg University. The study used extreme value statistics to develop two methods to determine bias in the near crash selection. The goal of the study was to find a better method for distinguishing near crash events from baseline events for further analysis of the SHRP2 study. The results of their study showed a discrepancy between the distribution of maximum speeds for crashes and maximum speeds for near crashes. This confirmed that there was considerable bias in the selection of near crashes as shown in Figure 1.

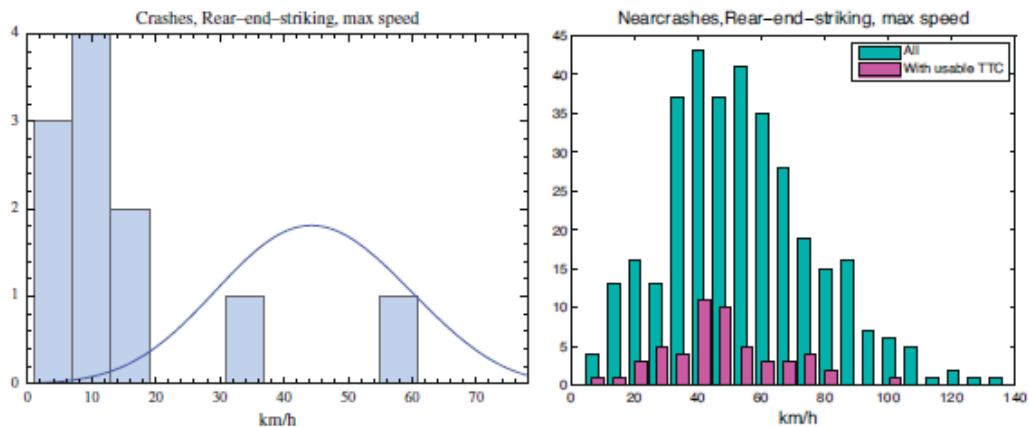


Figure 1 Crash Data and Near Crash Selection vs Speed

Klauer et al's report (6) on the NDS-100 car summary analyzed the impact of driver

inattention on near crash and crash risk. Since the NDS data occurred in a naturalistic setting, drivers experience realistic distractions that can lead to crash and near crash events. To determine the impact, driving behavior was compared from baseline data to situations where crashes and near crashes were occurring. For this report, 20,000 six-second segments were randomly selected from vehicles that experienced near crash events. For 5,000 of the baseline data samples, eye glances were reported. The test found that drivers age and personality played a part in the frequency of inattention related crashes/near crashes. Younger, more inexperienced drivers had a more significant likelihood to be involved in crashes and near crashes. These drivers also reported more traffic violations in a survey taken before the study. The results also found that drivers who were more frequently involved in inattention related crashes were often more drowsy and scored lower on personality tests. The data showed that drivers' involvement in near crash/crash events was often not a fluke. The correlation between drivers who often participated in driving inattention activities and those who were involved in inattention related crashes was 0.72, which is a significant result.

Chen et al's study (7) analyzed the behavior of drivers during lane change maneuvers. The goal of the study is to gain a better understanding of the drivers' behavior when overtaking another vehicle. Therefore, researchers could more precisely design forward crash warnings so that they do not erroneously display during these scenarios. The research team validated the lead vehicle detection algorithm with the video to determine the algorithm correctly located the lead vehicle in 84% of the lane changes as shown below. The algorithm had a sensitivity of 0.87 and a specificity of 0.983. The algorithm also determined that vehicles were more likely to change lanes with travel speeds ranging from 30-60 mph. The TTC values were found to increase when vehicles were traveling at a faster speed, due to the drivers' tendency to allow larger gaps at higher travel speeds.

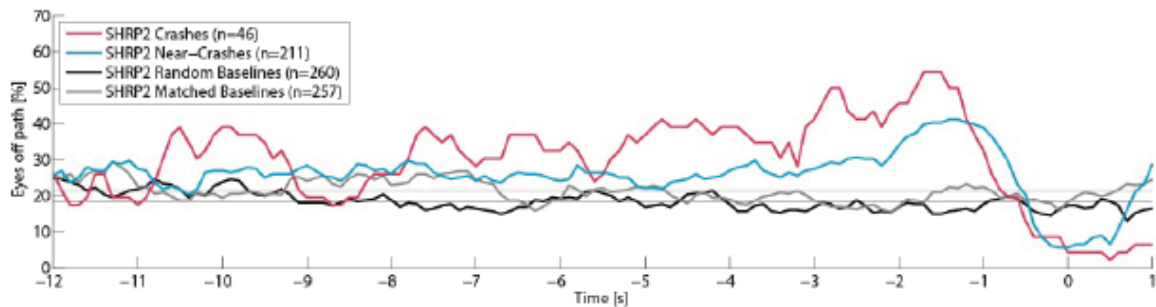
Sangster et al's study (8) used the data from the 100car-NDS to test and calibrate car-following models. The data was used to calibrate four different models including the Gipps, intelligent driver, Gaxis-Herman-Rothery, and the Rakha-Pasumarthy-Adjerid model. Up until the NDS study, car following models have been validated using test tracks, simulators, and loop detectors. The results of the calibrations found that the RPA model is the best suited calibration model to the naturalistic data set. Although the RPA model is more complex, it created far less error than the second best, Gipps model.

### **SHRP2 NDS and RID Studies**

The following section summarizes the past studies related to NDS and RID. The first four studies awarded to assess the SHRP 2 data in association with the RID data include Iowa State University Center for Transportation Research and Education (CTRE): Lane departures on rural two-lane curves; MRIGlobal: Offset left-turn lanes; University of Minnesota Center

for Transportation Studies (CTS): Rear-end crashes on congested freeways; and SAFER Vehicle and Traffic Safety Centre at Chalmers University, Sweden: Driver inattention and crash risk (9). These projects were only able to use limited data from the SHRP2 study, because they began before the data collection process was complete. Of the four projects, three were selected for Phase II and completed. These include CTRE's study on lane departures on rural two lane curves, Chalmers' study driver inattention, and MRIGlobal's study on offset left-turn lanes (10).

Researchers at Chalmers University of Technology in Sweden performed the first study (10) incorporating the SHRP 2 and RID data. Their study analyzed the effects of driver distractions using the SHRP 2 data. The primary goal of the study was to develop inattention-risk relationships that determine the relationship between driver inattention and crash risk in lead-vehicle precrash scenarios. These relationships helped in determine which glances are most dangerous for drivers. Many of these distractions did not occur frequently enough to have statistical significance. The analysis confirmed some of the findings from previous studies. It confirmed that distracting activities occurred more frequently in near crash events, visually demanding tasks involved more risk, and texting had the highest odds ratio, meaning it leads to a significant risk. The danger of glances was quantified using a three metric model including inopportune glance, mean glance duration, and the driver's uncertainty of the driving scenario. **Error! Reference source not found.** Figure 2 shows an increase in danger, the longer a driver's eyes are off path. The results also found that lead vehicle crashes are caused by a combination of glance duration and closure rate. The researchers noted that their results suggest the need for FCW, autonomous cruise control, and autonomous emergency braking.



**Figure 2 Driver Glance Duration's Impact on Crashes**

The second project assigned to the SHRP 2 data was performed by researchers at MRIGlobal. The study (11) showed how the NDS and RID data can be used to provide guidance for safety countermeasures to offset left turn lanes. Gap acceptance behavior was a contributing analysis factor to this study. The main goal of their research was to evaluate left turning gap acceptance by an extensive sample of drivers at different intersections that

incorporate left turn lane offsets. The analysis used a logistic regression to predict the critical gap from left turning vehicles in each offset category. The results determined that as the offset became more negative, the critical gap length increased. Critical gaps were also 2 seconds longer when the sight was restricted from an oncoming left turning vehicle, but this result is not considered a statistically significant amount. It was also determined that intersections designed to allow vehicles' view to be blocked from oncoming left turn vehicles decreased the operation efficiency of the intersection. Since there was no crash data from these intersections, data was too limited to determine crash related safety risks at these intersections.

University of Minnesota's study was not completed, so only preliminary analysis is available. The primary goal of the study was to determine how drivers behave when encountering a freeway stopping wave. This information can be used to reduce congestion on urban roadways. The NDS data included 250 freeway trips containing break-to-stop events. From the NDS data, researchers can obtain braking deceleration data, along with following vehicle reaction time and following distance. With this information, it is possible to gain more insight on drivers' behavior on congested freeways (11).

Iowa State University's research team performed a study to analyze roadway departures on rural two lane curves. The purpose of the study (13) was to use the NDS and RID data to determine how driver behavior, roadway factors, and environmental factors relate to these departures. Only paved roadways over one mile out of the urban area with speeds posted 40-60 mph were included in the study. Time series models were also used to lane position and speed of the vehicles. The results showed that drivers tended to maintain their upstream position during the curve and that drivers' distractions caused them to shift in the lane. If they were on the inside and encountered a distraction, they tended to shift 0.14 m towards the right at the next point in the curve. This showed the need for rumble strips or paved shoulders as a counter safety measure. Younger drivers were found to speed into curves more than older drivers by 0.5 mph per every 10 years. Four multivariate logistic regression models were used to evaluate how environmental factors affect roadway departure. The results showed that right side lane departure is 6.8 times more likely on the inside of a curve. The presence of a guardrail decreased inside departures by 66 percent. Males were found to have outside lane departures four times as often as females.

## **METHODOLOGY**

The objectives of this study are achieved through three main steps: data description, data extraction, and linkage of RID and NDS. In the first step, researchers describe in detail the two datasets that were used in the study: NDS and RID. The second step briefly explains the process of downloading and purchasing RID and NDS. The last step demonstrates the process of linking RID and NDS using ArcGIS program.

### **Data Description**

#### **NDS Dataset**

The NDS data evaluated in this study is provided by the Second Strategic Highway Research Program (SHRP 2), performed by Virginia Tech Transportation Institute (VTTI), which is the largest scale NDS study performed to date. The study utilized 2,800 drivers, of ages 16-80, to collect 200,000 highway miles of data in six locations. Having finished the study in 2014, research incorporating the 200,000 highway miles of data has only just begun. The six locations where the data was collected include two counties surrounding Tampa, Florida; ten counties in central Indiana, Erie County in New York, four counties in North Carolina, ten counties in central Pennsylvania, and four counties in Washington (12). This includes 150 vehicles from Indiana, 150 from Pennsylvania, 441 from Florida, 441 from New York, 300 from North Carolina, and 409 from Washington (2). Each participant was assessed in a screening process that tested his/her qualifications for the study.

The Data Acquisition System (DAS) for the SHRP 2 study is developed in house by VTTI. The DAS includes a forward radar, four video cameras, accelerometers, vehicle network information, Geographic Positioning System, computer based lane tracking, built in computer vision algorithms, and data storage capabilities. The data is recorded continuously, creating an extensive dataset of driving behavior, which is critical for safety analysis studies. Figure 3 from Campbell's paper, "The SHRP 2 Naturalistic Driving Study Addressing Driver Performance and Behavior in Traffic Safety" provides all the DAS Channels.

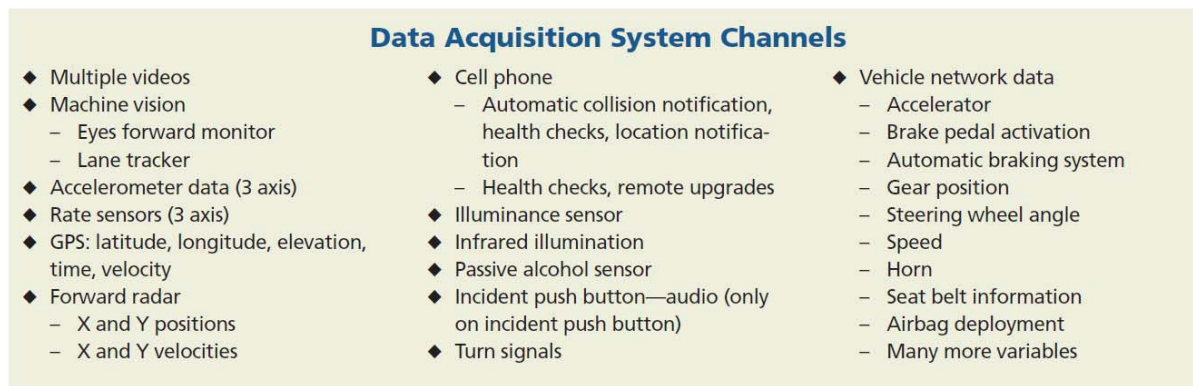


Figure 3 DAS Channels (2)

The SHRP 2 Data includes 50 million vehicle miles, 5 million trips, over 3900 vehicle years, and over 1 million hours of video recording. The crash data varies for existing studies due to the timing of when the research was performed. Since many of the projects were contracted while data collection was ongoing, their data collection was incomplete (10).

The overall purpose of the SHRP2 study is to extend the database collected from the 100-Car NDS study to include a much larger sample size. This helps to account for the limitations of the 100-car study, such as the lack of crash and near crash data. The SHRP2 study accounts for regions spread across the country, preventing the bias from the original study, which was only in northern Virginia. The data also has far more subjects, which allows for more crash data. This information is useful for accomplishing the 10 goals, previously mentioned, from the original study, because it provides more insight into the causes of crashes and near crashes.

This extensive database of drivers' behavior is essential to further research on drivers' safety. To analyze driver behavior, researchers examined the drivers' pre-event maneuvers, precipitating factors, contributing factors, associative factors, avoidance maneuvers, and event types. The inattention data collected includes secondary task distractions, driving related inattention, drowsiness, and non-driving related eye glances. Secondary tasks include wireless devices, passenger related distractions, internal distractions, dining, vehicle related secondary tasks, daydreaming, talking or singing, external distractions, personal hygiene, smoking, and other distractions.

The events captured through the NDS footage can be categorized into five separate groups. These groups include crashes, near crashes, crash relevant conflicts, proximity conflicts, and non-conflict events. Crashes include any form of contact the vehicle makes with another object, whether that object is another vehicle, a pedestrian, or a stationary object. Near crashes include any instances that require a rapid evasive maneuver by the

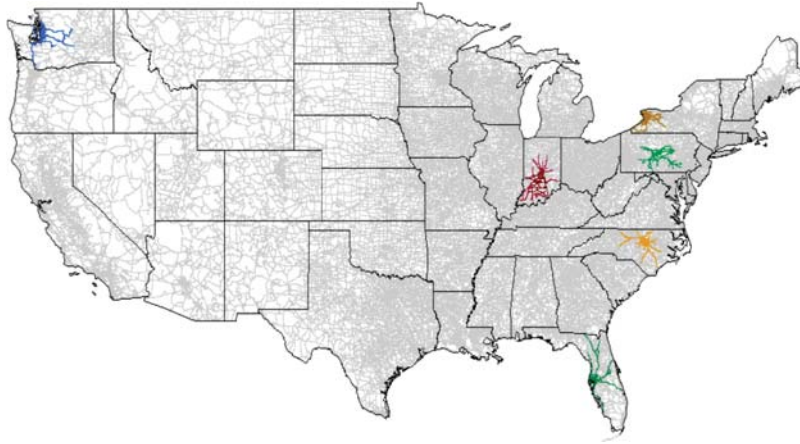


driver, but does not ultimately result in a crash (1). The determination of what classifies an event as a near crash is a debatable topic that is covered in Jonasson's study (5). Crash relevant conflicts are all instances that require the driver to utilize any form of crash avoidance response. This means that the driver's behavior is not "normal" and does not fall under the 99 percent confidence limit for baseline data. Proximity conflict is any event that causes the driver's vehicle to be within close proximity of another object due to unawareness by the driver. This does not include situations where the driver is intentionally within close proximity of another object, i.e. when a driver is at a red light. Non-conflict events include all situations where the drivers' risk is increased, but no conflicts, near crashes, or crashes occur (1). Any situation where no abnormal events are occurring is considered baseline data. This information is also important, because it allows researchers to compare drivers' behavior during events to those when events are not occurring.

### **RID Dataset**

The RID data set is compiled from government, public, and private databases. The data collected from the mobile data collection project consists of roadway data including road class, roadway characteristics, roadway features, geometric features, roadside features, pavement features, and other additional data. Roadway characteristics include number of lanes, on street parking, pavement markings, medians, rumble strips/stripes, speed limits, and shoulders. Roadway features are comprised of intersections, driveways, bridges, approaches, rail crossings, and ramps. The geometric features in RID are grades, cross slopes, horizontal curvature, and vertical curvature. Roadside features include barrier systems, obstacles, sidewalks, signs, and street lighting. Pavement features are comprised of edge drop off, pavement condition, and friction factors. Additional RID data includes crashes, roadway safety projects, traffic, work zones, weather, and enforcement/education projects.

RID data also includes supplementary data from government agencies such as the information provided in Smadi's SHRP report (3). Figure 4 shows the highway networks covered by the RID database.



**Figure 4 RID Highway Coverage**

### **Data Extraction**

As a first step, the research team has contacted Smadi, a Research Scientist at CTRE, through email to obtain the “Term of Use” document, which listed detail steps in order to acquire the dataset. At the end, Smadi provided a hard-drive with the comprehensive RID dataset and login access through the Fugro system. According to the CTRE website, Fugro Roadware, which provides Infrastructure Asset Management Technology and Data Collection service, was the selected vendor for the mobile data collection project. Later on, the RID data was investigated thoroughly to identify the common links with the NDS data.

The RID contains extensively detailed data on 25,000 centerline miles of roadways in six sites, less detailed data on 200,000 centerline miles of roadways in the six states in which the sites were located, and supplemental data on topics such as crash histories, travel volumes, construction, and weather in the six states. Due to the large size of the RID, a representative sample from Tampa, FL was taken for further analysis.

### **The Process of Linking RID and NDS**

The main objective of this step was to create a pseudo dataset of points with coordinates, which could be mapped onto the same location as the RID data from Tampa, Florida. This was done in order to work with a dataset similar to the NDS data since buying the NDS data is not part of this project’s scope. For the purposes of this project, the excel dataset will be called the PseudoNDS Data.

The excel data could have been mapped either through catalog, with the “Add data” button or added directly by using the “Add XY data” option in the “File” menu. The “Add XY data” option was used to successfully map the data into ArcMap because it was the most



direct method. When adding the data, the coordinate system used was GCS\_WGS\_1984. This coordinate system was specified because it was the system in which the RID data was originally projected.

XY data added from excel files directly into ArcMap can be viewed only for the session during which the program is open. The mapped data is automatically deleted when the program is closed. In order to view the points on the map for every session the program was opened, the PseudoNDS data was exported into a shapefile and added as a map layer.

The NDS data can be mapped in ArcMap using the same processes described above. The NDS data is often incorrectly aligned with the road alignment so the points used in the PseudoNDS dataset were realigned with positioning of the RID dataset in order to have a correct representation of the data. The output of such alignment will be presented in the result section.

# DISCUSSION OF RESULTS

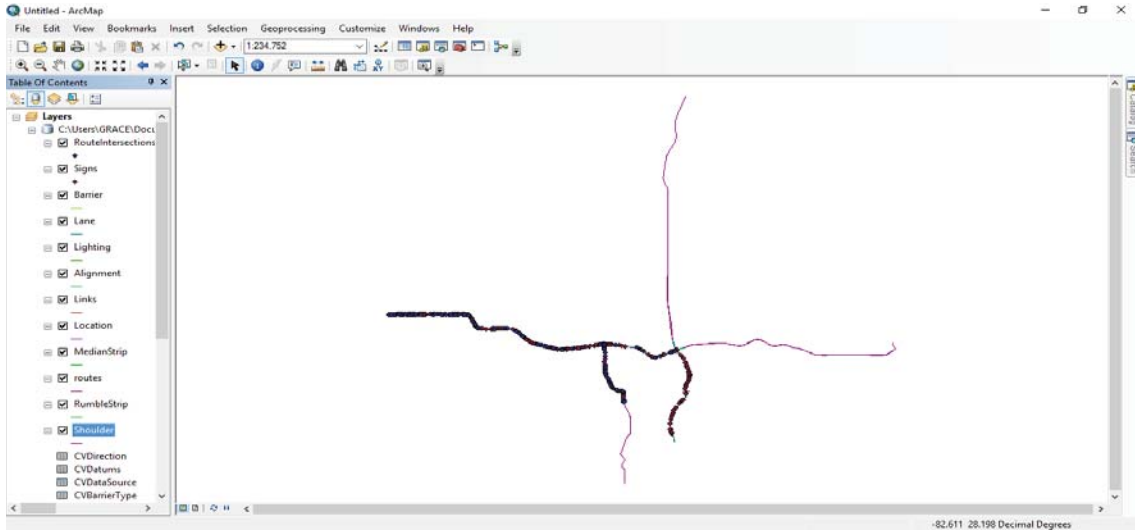
## Exploration of RID Database

As explained in the methodology section, the research team has contacted the CTRE to purchase the RID. Then, a representative sample was taken for further analysis. The sample dataset included two types of data, namely: a spreadsheet and an ArcGIS data files. The spreadsheet listed each roadway geometry item’s information in more than 12 categories including lighting, lane, median strip, shoulder and others. The ArcGIS file provided all items in accordance to the spreadsheet in Tampa, FL. Figure 5 is a screenshot of a part of the spreadsheet under the “lighting” tag. As shown in Figure 5, the spreadsheet provided field name, field type, description and other information. This information could be used to connect to NDS dataset.

Field Name	Units	Format	Field Type	Example Value	Description
AspectID		{xxxxxxxx-xxxx-xxxx-xxxx-xxxxxxxxxxxx}	GUID	{6AA84D21-AEA3-4D81-B874-5E343F8A6CC7}	UniqueID
AspectTypeID		xxxxxxxx	Short		Lighting
Comment		xxxxxxxx	String		Any Comments deemed necessary at the time
SourceID		xxxxxxxx	String		Unique ID retained from data source
DateSourceID		xxxxxxxx	Short		Unknown

**Figure 5 Part of Lighting Information of RID Data**

Figure 6 shows the selected sample dataset from Tampa, FL within the ArcGIS program. The roadway geometry items shown in ArcGIS were identical to the spreadsheet from the RID, including signs, barriers and other categories. Each category has around 10 components including Source ID, Aspect ID, Side of Road and others. The roads in darker color are the sample data points. The tool bar on left side lists the categories that are included in this package.



**Figure 6 RID Sample Data in ArcGIS Program**

Furthermore, Fugro system, which is associated with the RID dataset, is considered an essential resource for RID exploration. Users can easily access different routes, analyze the map location, and replay the video used in data collection. Figure 7 displays a screenshot of the Fugro system. In this figure, “ID segment” 964891 was selected. The map on the left of the figure shows the section’s geometrical alignment using a thick blue color line. The street view on the right of the figure shows the starting point of this segment, which is marked using the yellow arrow in the map. Several tool icons at the top of the figure are available for users to help analyze the segment.



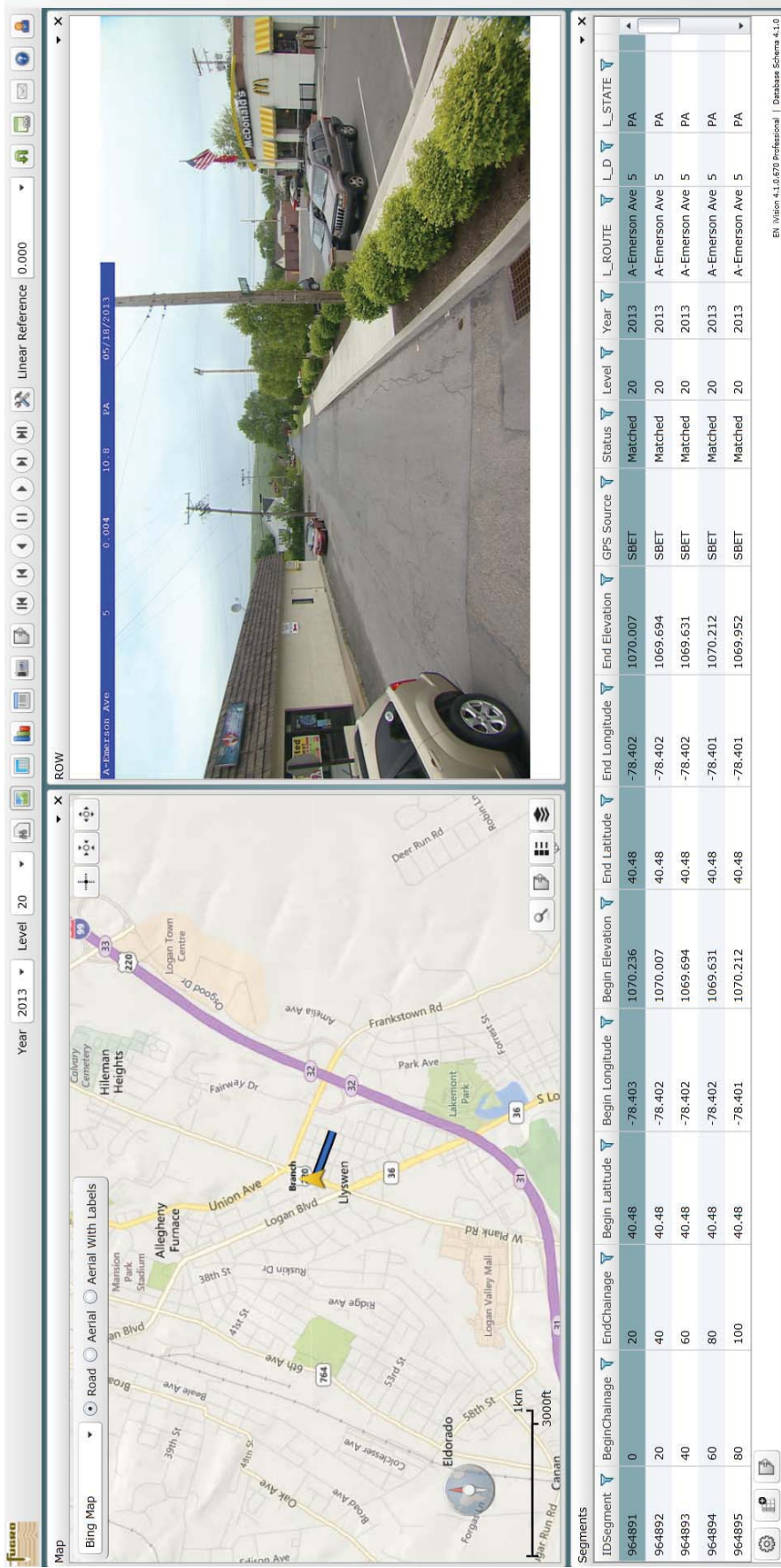


Figure 7 FUGRO Software Screenshot



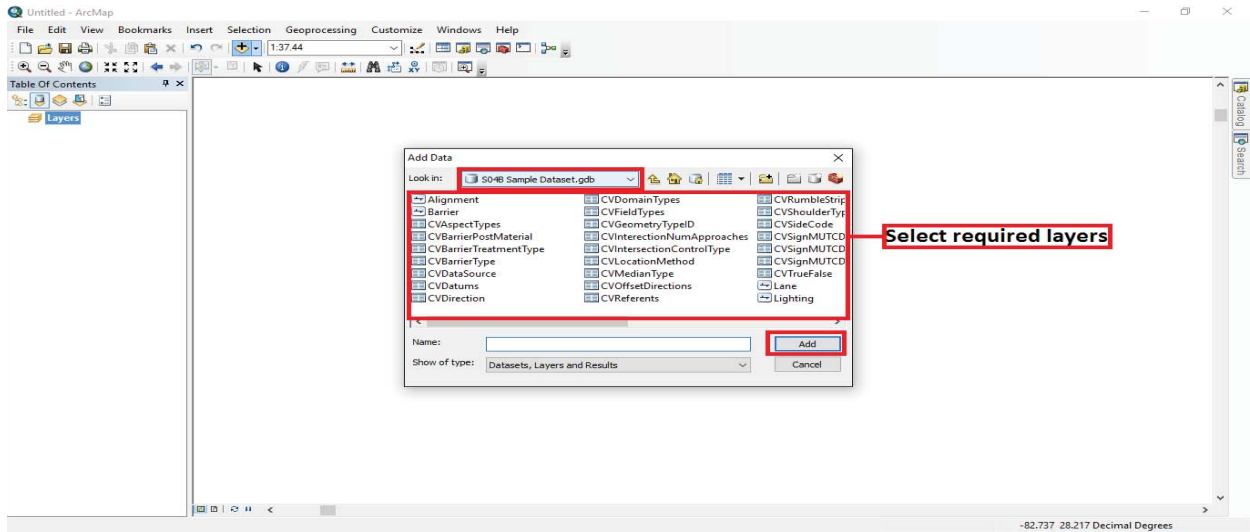
## Mapping RID in ArcGIS

ArcMap software was used to map the RID data. ArcMap is the main component of the ArcGIS suite of geospatial processing programs. It is used primarily to view, edit, create, and analyze geospatial data.

The RID data includes a geodatabase and a metadata file containing descriptions of all data. The selected RID data includes information for 20 miles of roadway in Florida includes details of the following.

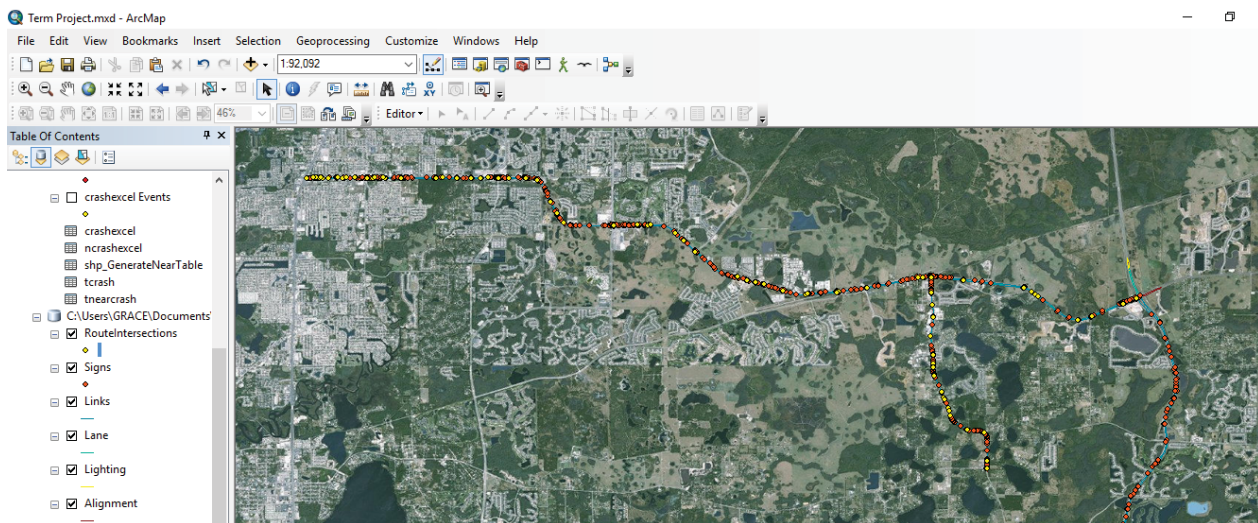
- Horizontal curvature:
  - Radius
  - Length
  - Point of curvature (PC)
  - Point of tangency (PT)
  - Direction of curve (left or right based on driving direction)
- Grade
- Cross-slope/Superelevation.
- Lanes: number, width, and type (turn, passing, acceleration, car pool, etc.).
- Shoulder type/curb (and paved width, if it exists).
- *All Manual on Uniform Traffic Control Devices (MUTCD) signs.*
- Guardrails/Barriers.
- Intersection: location, number of approaches, and control (uncontrolled, all-way stop, two-way stop, yield, signalized, roundabout). Ramp termini were considered intersections.
- Median presence: type (depressed, raised, flush, barrier).
- Rumble strip presence: location (centerline, edgeline, shoulder).
- Lighting presence.

The RID geodatabase file was opened via ArcMap and the road features listed above were added as layers. The screenshot of this process is expressed as shown in Figure 8.



**Figure 8 Screenshot of Data Input in ArcMap Program**

After inputting the dataset into the ArcMap program, the data was then viewed geospatially in ArcMap after adding a base map. Figure 9 displays an example for roadway data within ArcMap. Each colored dot attached to the roadway segment represents a different category of the roadway characteristics. In this figure, the red dots indicate signage within the selected roadway links.



**Figure 9 Roadway data for Tampa, Florida**

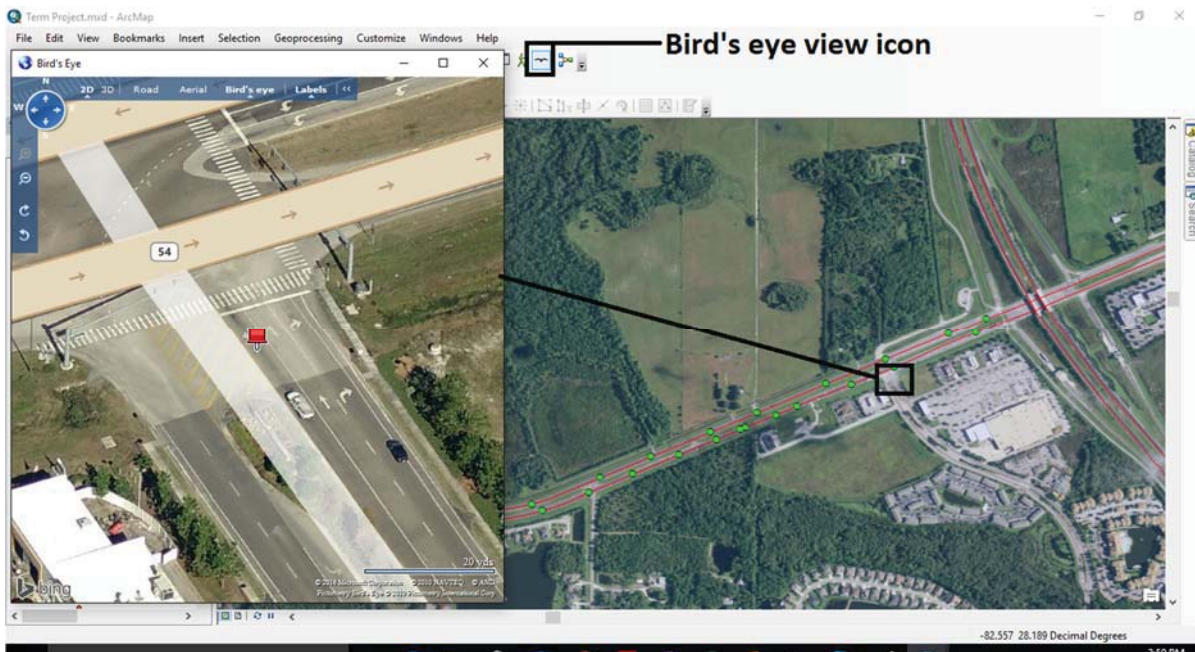
### Applications of Street View and Bird's Eye View

The ArcGIS Desktop Street View AddIn is a useful tool when the street view of specific



event locations is required. The recent version (Version 1.2) has an added tool for Bird's Eye view.

In order to explore the street views of the mapped data, the ArcGIS Desktop Street View Add-In was downloaded from ESRI's ArcGIS website. The add-in provides you with an access to include the Google Street View and Microsoft Bird's eye together. To view a location in the bird's eye view or the street view, the icon corresponding to the required view must be selected. The cursor can then be used to select the location to be viewed. Figure 10 and Figure 11 show the bird's eye view and the street view of a particular location on the map in ArcMap.



**Figure 10 Bird's Eye View in ArcMap**



**Figure 11 Street View in ArcMap**

### **NDS and RID Road Alignment**

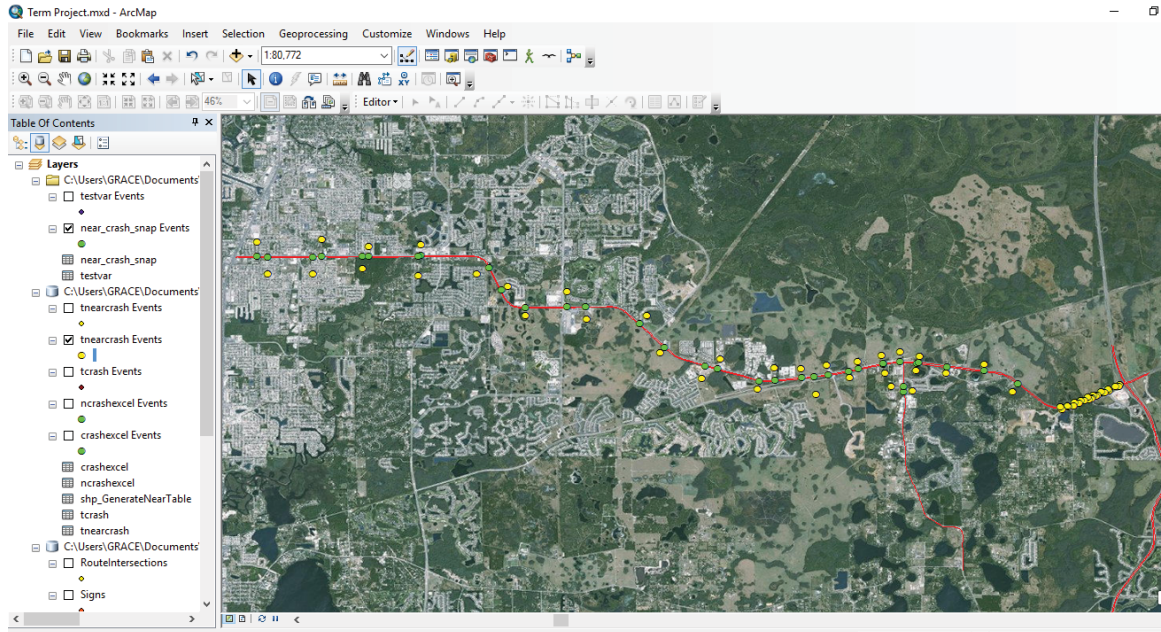
In order to obtain appropriate data visualization for further analysis, it is imperative to adjust the data. First, the data was loaded on the map to check whether the events are located correctly or not. Second, the “Near” ArcMap toolbox was used to align the incorrectly aligned data with the road geometry. This tool calculates the misalignment distance and gives some proximity information between the PseudoNDS data and road alignment. Hence, the PseudoNDS was used as the input feature in the dialogue box and the roadway was selected as a “near” feature. Finally, the data locations were updated based on the correct alignment and the “Location” option was used to extract this new information.

When the command was executed, four new attributes were added to your PseudoNDS data attribute table. These attributes include “NEAR\_FID”, “NEAR\_DIST”, “NEAR\_X” and “NEAR\_Y”. The “NEAR\_FID” describes the object ID of the road segment, while “The NEAR\_DIST” provides information on the distance between each event and the alignment. The “NEAR\_X” and “NEAR\_Y” fields display the updated X and Y coordinates where the data was placed correctly on the roadway.

Knowing the updated data locations, the attribute table that contains the four new

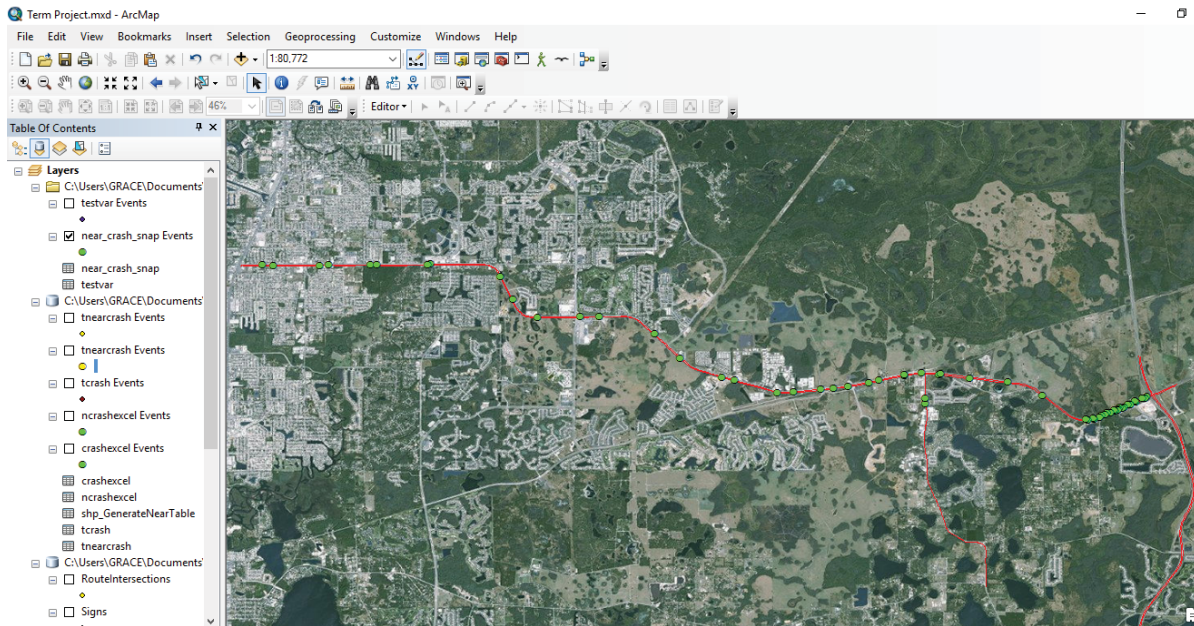
attributes was exported as a database file. This file was then added to the map where the data were placed correctly using the new XY fields. Finally, the new coordinate fields (“NEAR\_X”, “NEAR\_Y”) were saved permanently as a shape file with two different layers.

Figure 12 shows an example of the updated data points (green points) and the raw data points (yellow points). The Figure displays how the green points have been moved onto the road alignment compared to their old positions. Figure 13 shows the locations of only the newly mapped data points.



**Figure 12 Newly and Previously Mapped Points**





**Figure 13 Newly Mapped Points**

### **Potential Topics**

Since the SHRP 2 NDS data and RID data are still relatively new, there are countless ways that their vast information can be utilized to improve drivers' safety and transportation efficiency. There has yet to be a defined algorithm that is most efficient for extracting the necessary information from the surplus of data. The NDS data itself can also be expanded to incorporate more locations across the country and the RID data can be improved through more input from private and public sources. In this section, researchers will post several ideas of potential research that combines RID and NDS.

One potential use involves extracting the speed limit value from RID and incorporating it into distracted driving studies. In NDS dataset, trips involving distracted driving provide the vehicle's GPS\_speed. However, it is difficult to compare speed values when the vehicles were driving on different routes. In this case, GPS\_speed from NDS dataset could be transferred to the percentage of speed limit after obtaining the speed limit from RID.

In addition, roadway geometric elements could be used in safety research. NDS provides details of crash and near crash events, including the video and vehicle performances. However, NDS does not include the roadway elements, which are core elements in evaluating the causes of crash and near crash events. A short list of potential safety issues that can be addressed using RID data includes but is not limited to: the effect of

roadway grade on potential crashes, the effect of horizontal curve on crash and near crash events, and the effect of guardrails, median, and rumble strips on crash and near crash events.

Other potential topics could focus on the relationship between driving behavior and road characteristics. A few suggested studies include: measuring the effectiveness of safety campaigns with driver compliance, understanding passenger influence on driver behavior and comparing it to survey-based studies, quantifying the effect of warning and regulatory signs on driver compliance, validating simulation studies on effects of text messaging on driver compliance, and examining statistically significant clusters of crash and near crash events taking into account roadway conditions.



## CONCLUSION

The Naturalistic Driving Study (NDS) and the Roadway Information Database (RID) are two valuable tools for improving roadway safety and efficiency. The RID data includes information about physical characteristics of the roadways whereas the naturalistic driving study provides information of vehicle trips performed on these roadways. Since the SHRP 2 NDS study was only completed in 2014, research using the NDS data has merely scratched the surface of its potential for improving drivers' safety. When coupled with the RID, the NDS data is able to explain drivers' behavior within the context of a roadway environment. Linking the NDS data to the RID allows researchers to correlate driving behavior to specific roadway conditions.

In this research project, objectives of exploring and mapping the RID, linking the NDS data, and exploring the data using the Street View and Bird's eye view were discussed. Using ArcGIS software, the NDS and RID data were linked, providing a map with overlaid information from both NDS and RID databases. This new shapefile, containing both sets of data, can be utilized for several potential safety studies.

Many of these potential research topics incorporating the linked RID and NDS databases are included. A short list of potential applications include quantifying the effect of warning and regulatory signs on driver compliance, measuring the effectiveness of safety campaigns with driver compliance, examining statistically significant clusters of crash and near crash events taking into account roadway conditions, and others.

## RECOMMENDATIONS

The Naturalistic Driving Study (NDS) and the Roadway Information Database (RID) are two valuable tools for improving roadway safety and efficiency. The RID data includes information on roadway features that becomes useful for safety research when linked with the naturalistic data from the NDS database. The following section lists recommended research work that can utilize information from both databases.

A potential study can be performed to incorporate speed limit data from the RID into distracted driving studies performed previously using NDS data. This allows the researcher to distinguish between speed changes due to distraction and those caused by changes in roadway features. There are also several opportunities for research involving the geometric data and roadway features from the RID. By combining the information from the RID and the NDS databases, researchers can analyze effects of roadway grade, horizontal and vertical curves, guardrails, medians, rumble strips, and other various roadway features effects on crash and near crash events. Other potential studies utilizing information on roadway characteristics include but are not limited to measuring the effectiveness of safety campaigns with driver compliance, understanding passenger influence on driver behavior and comparing it to survey-based studies, quantifying the effect of warning and regulatory signs on driver compliance, validating simulation studies on effects of text messaging on driver compliance, and examining statistically significant clusters of crash and near crash events taking into account roadway condition.



## ACRONYMS, ABBREVIATIONS, AND SYMBOLS

NDS	Naturalistic Driving Study
RID	Roadway Information Database
SHRP	Strategic Highway Research Program
CTRE	Center for Transportation Research and Education
UNETRANS	Unified Network-Transportation network data model
VTTI	Virginia Tech Transportation Institute
TTC	Time to Collision
FCW	Forward Collision Warnings
DAS	Data Acquisition System
GIS	Geographic Information Systems
GPS	Global Positioning System



## REFERENCES

1. Dingus, T.A., Klauer, S.G., Neale, V.L., Petersen, A., Lee, S.E., Sudweeks, J.D., Perez, M.A., Hankey, J., Ramsey, D.J., Gupta, S. and Bucher, C., 2006. The 100-car naturalistic driving study, Phase II-results of the 100-car field experiment (No. HS-810 593).
2. Campbell, K.L., 2012. The SHRP 2 naturalistic driving study: Addressing driver performance and behavior in traffic safety. TR News, (282).
3. Smadi, O., Hawkins, N., Hans, Z., Bektas, B.A., Knickerbocker, S., Nlenanya, I., Souleyrette, R. and Hallmark, S., 2015. Naturalistic Driving Study: Development of the Roadway Information Database (No. SHRP 2 Report S2-S04A-RW-1).
4. Montgomery, Jade. "Age and Gender Differences in Time to Collision at Braking From the 100-Car Naturalistic Driving Study." (n.d.): n. pag. Web.
5. Jonasson, J., & Rootzen, H. (2014). Internal validation of near-crashes in naturalistic driving studies: a continuous and multivariate approach. *Accident Analysis & Prevention* Vol. 62, 102-109.
6. Klauer, S.G., Dingus, T.A., Neale, V.L., Sudweeks, J.D. and Ramsey, D.J., 2006. The impact of driver inattention on near-crash/crash risk: An analysis using the 100-car naturalistic driving study data.
7. Chen, Rong. "Driver Behavior during Lane Change from the 100-Car Naturalistic Driving Study." Diss. Virginia Tech, n.d. Abstract. (n.d.): n. pag. Print.
8. Sangster, J., Rakha, H., & Du, J. (2013). Application of naturalistic driving data to modeling of driver car-following behavior. *Transportation Research Record: Journal of the Transportation Academies*, 20-33.
9. Hallmark, S., Hsu, Y.Y., Boyle, L., Carriquiry, A., Tian, Y. and Mudgal, A., 2011. Evaluation of data needs, crash surrogates, and analysis methods to address lane departure research questions using naturalistic driving study data (No. SHRP 2 Report S2-S01E-RW-1).
10. Victor, T., Dozza, M., Bärghman, J., Boda, C.N., Engström, J., Flannagan, C., Lee, J.D. and Markkula, G., 2015. Analysis of naturalistic driving study data: Safer glances, driver inattention, and crash risk (No. SHRP 2 Report S2-S08A-RW-1).
11. Hutton, J.M., Bauer, K.M., Fees, C.A. and Smiley, A., 2015. Analysis of Naturalistic Driving Study Data: Offset Left-Turn Lanes (No. SHRP 2 Report S2-S08B-RW-1).

12. Hallmark, S., McGehee, D., Bauer, K.M., Hutton, J.M., Davis, G.A., Hourdos, J., Chatterjee, I., Victor, T., Bargman, J., Dozza, M. and Rootzen, H., 2013. Initial analyses from the SHRP 2 naturalistic driving study: addressing driver performance and behavior in traffic safety (No. SHRP 2 Safety Project S08).
13. Hallmark, S.L., Oneyear, N., Tyner, S., Wang, B., Carney, C. and McGehee, D., 2015. Analysis of Naturalistic Driving Study Data: Roadway Departures on Rural Two-Lane Curves (No. SHRP 2 Report S2-S08D-RW-1).