El Paso County
ROAD SAFETY PLAN

EL PASO COUNTY
COLORADO

February 2021

MEMORANDUM

TO: El Paso County

FROM: Cambridge Systematics

DATE: 2/26/2021

RE: El Paso County Road Safety Plan Extension: Geolocating

This memo presents the deliverable for the geolocating task added to the El Paso County Road Safety Plan project, including a summary of the data sources and challenges, geolocating process, and resulting crash totals. Crash geolocating is the process of identifying the geographical location and collecting the GIS coordinates of crashes using other location information provided by officers within the corresponding crash report. Crash locations are a necessity for completing the safety analysis to identify overrepresented contributing crash factors and developing an accurate understanding of a region's crash trends to support the El Paso County Road Safety Plan project.

The GIS location of a crash event allows the analysis to consider roadway characteristics, roadway classes, jurisdictions, roadway ownership, and additional fields only available within disparate datasets. This includes network analysis and screening the roadway network for local roadway crashes to map and prioritize locations with the highest potential for safety improvement. Specific crash locations are also needed to accurately analyze a specific location and select appropriate countermeasures to address the risk characteristics.

Crash Data Sources

Several sources were contacted and utilized in the data collection effort:

- Colorado Department of Transportation

- Pikes Peak Area Council of Governments
- Colorado Springs Police Department
- El Paso County

CDOT provided crash data from 2010-2019 within the county. The Colorado Department of Transportation (CDOT) is responsible for geolocating crashes in Colorado on state-maintained routes. However, CDOT does not provide latitude/longitude for crashes on the local roads in the state, and coordinates provided by officers on local roads are not validated by CDOT quality assurance staff

The Pikes Peak Area Council of Governments (PPACG) is a Metropolitan Planning Organization (MPO) encompassing 16 counties, including El Paso County. Crash locations on local roads within their planning area were previously developed by PPACG staff. However, the local road crashes were only provided from 2010 through 2015. After 2015 crash geolocating services were no longer provided, making local roadway crash data beyond 2015 unavailable.

The Colorado Springs Police Department also located crashes for all roadways with Colorado Springs, including from 2016-2019. The combination of the three sources provided a high percentage of the total crashes in county. However, the local roadway crashes outside of Colorado Springs from 2016-2019 (5,114 crashes) still remained unlocated for the county. For those crashes without coordinates from CDOT, PPACG or the Colorado Springs Police Department the project team developed a process for pin-pointing the specific location of the event through available fields in the crash data provided.

Geolocating Local Road Crashes

The data provided by CDOT for the county included several location fields for the local roadway crashes that did not have GIS coordinates. These fields include the route name, intersecting route name or address, and offset from the intersection or mile marker. A two-step process was developed to utilize the route address or intersecting routes to pin-point the initial GIS coordinate and adjust the location based on the offset provided. The following details each of these steps.

Step 1 – Geolocating Addresses and Intersections

The initial step was to prepare the data format for the google application programming interfaces (APIs). The 5,114 unlocated crashes from 2016-2019 were cleaned by removing unknown locations and removing extra notes typed in route name fields. The address field was created using the existing location fields.

All crash records were geolocated to the provided intersection or address, using Python and Google API to return the GIS coordinates. Next, all coordinates were checked to be within the El Paso County boundaries. Crashes initially located outside the boundaries were separated and manually checked. Following the initial API geocoding, 2,986 of the crashes were successfully geocoded to the point of reference and required no further offsetting or adjustment, while 2,070 of the crashes progressed to Step 2 to account for offset distances provided in the crash data.

The remaining 58 crash locations referenced mile markers and could not be located via the API. These crashes were located by identifying mile marker 0.00 in the GIS roadway network file for the route listed in the crash data

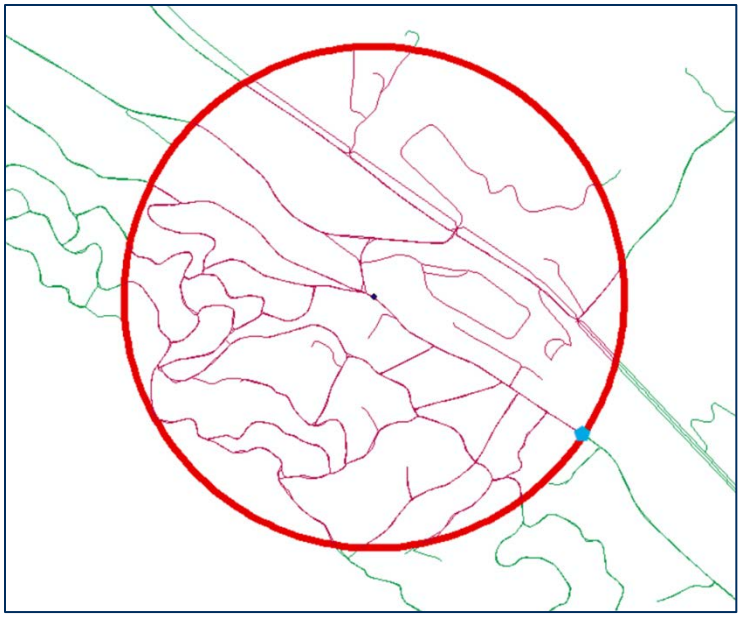
and tracing the roadway network file to the mile marker distance listed and extracting the GIS coordinates. Crashes with mile markers outside the county were labeled as “unknown” and removed.

A manual geolocating process was used to check the accuracy of the coordinates provided through the API by confirming the automated result matched the location that was manually determined by searching the street addresses or intersections by hand. For validation purposes the top-100 highest crash locations, which produce 911 of the crash records, were manually geolocated by hand and compared with Google API results. The validation test showed near 90 percent accuracy. The 10 percent of records that did not match the API results were manually located and updated. There were 67 crashes unable to be located due to an “unknown” location field. This validation, as well as the aforementioned process of manually locating crashes outside the county boundary, comprised the quality assurance check.

Step 2 – Correcting Locations for Offsets

Crashes with an offset from the intersection (2,070 records) progressed to the second step for offsetting the crash point from the intersection using the distance provided by the “link” field in the crash data set. There were 146 crashes that contained an offset measurement in the “link” field that was either blank or contained “99999” as distance. These crashes were pinned to the intersection location identified in Step 1. There were also nine crashes to be offset with an “unknown” route reference, where the initial intersection could not be located. Crashes with an actual offset distance, 1,915 crash records, were processed in GIS to adjust coordinates from the previous step. Six of these records contained an offset distance greater than seven miles and were flagged to be manually geolocated.

Figure 1: Crash Offsetting Example



Offsets are provided in crash records for a given direction from the intersection (found in Step 1) and a distance measured in feet or miles. The offset distance is used to create a buffer around the intersection point equal at that distance. Next, the adjacent route names intersecting the buffer are checked to ensure they match one of the two intersecting routes listed in the crash record in the approximate direction provided. *Figure 1: Crash Offsetting Example* shows an example where the offset was listed to east and the matching route name farthest east within the buffer was successfully identified. The new crash coordinates are then placed at the route intersection with the offset buffer (larger blue dot).

In this automated step, out of 1,915 crashes, 1,768 crash points were successfully attributed to routes matching the location field in the crash record. The remaining 147 crashes retained the coordinates from their intersection geolocated in Step 1.

Geolocating Results

The following summarizes the crash geolocating results in each step. *Appendix A* further break downs the fields and details used to geolocate the crashes:

- 5,114 crashes from 2016-2019 were processed for geolocating;
- 2,986 crashes were processed through API geocoding in Step 1 alone;
- 58 crashes were located using the GIS roadway file to identify the referenced mile marker; and
- 2,070 crashes were processed for offsetting in Step 2, following use of the Google API.

Throughout each of aforementioned steps combined, 67 of the 5,114 crashes (1.3 percent) provided “unknown” locations. A complete geodatabase with the located crashes from 2011-2019 has been compiled for use in the regional trend analysis and the network screening tasks. The database will be delivered to the county and safety partners for future use.

APPENDIX

Appendix A

Table 1: Located Crash Summary

Step	Category	Description	Total Number of Records Processed	Unknown Records*
Google API	Geolocating without offset	• Crashes that did not have offsets or use mile markers	2986	51
Google API	Geolocating with offset	• Crashes that have offset and *progress to Step 2*	2070	9
ArcGIS	Mile Marker	• Use roadway file to determine Mile Marker location	58	7
Step 1	**Final Location Subtotal**	• Those that do not progress to Step 2	**3,044**	**58**
Google API + Offset	Adjust Coordinates	• Identified offset distance and direction along matched route name	1,915	0
Google API + Offset	Unknown Offset	• Crashes with an offset that is zero, blank, or unknown retained coordinates from Step 1	155	9
Step 2	**Final Location Subtotal**		**2,070**	**9**
Total			**5,114**	**67**

*Unknown records (67 of 5,114) in the Google API step are tagged based on “unknown” location fields or locations specified outside the county boundary, and unknowns in the offsetting step are tagged based on mile markers that are not within the county boundary, and/or the location fields are “unknown”.