

# Evaluation of the Accuracy of Global Positioning System Coordinates for Collision Locations in California

John Bigham, Garrett Strang, and Sang Hyouk Oum

Safe Transportation Research and Education Center (SafeTREC), UC Berkeley

Transportation Research Board  
94th Annual Meeting  
January 2015 • Washington, D.C.

## Objective

To evaluate collision data in California:

- Determine the overall accuracy of GPS coordinates
- Categorize the types of errors or discrepancies
- Investigate error trends
- Provide recommendations for using GPS coordinates

## Methodology

### Data

The California collision database (SWITRS) was updated in 2006 to allow inclusion of GPS coordinates.

- Over 90% of the GPS coordinates from CHP officers.
- 30% of all collisions include a GPS coordinate.

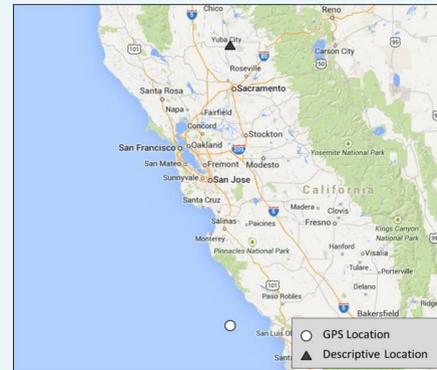
Year	2009	2010	2011
Collision Count	166,329	163,614	161,743
Include GPS Coordinates	40,341 (24.3%)	48,620 (29.7%)	52,381 (32.4%)

### Methods

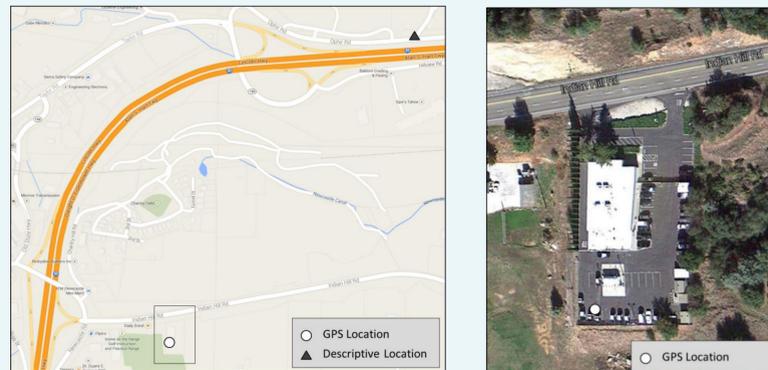
- A random sample of 383 collisions for year each from 2009 to 2011 was extracted (1,149 total).  
- Sample size was determined based on 5% confidence interval and 0.5 of an estimator of the accuracy.
- The random sample was imported into Google My Maps using the GPS coordinate and adjusted based on descriptive location associated with the collision.
- The distance from the original GPS coordinate to adjusted point was calculated for each pair.
- Error trends were reviewed to determine a classification system.

## Classification System

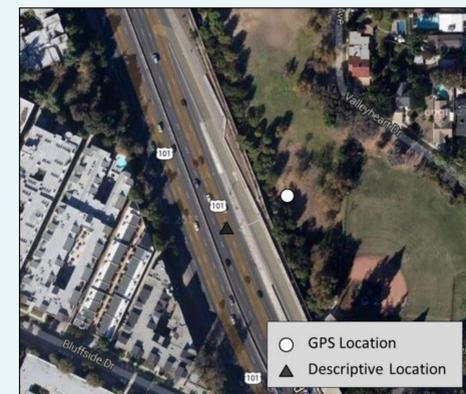
Systematic GPS Error



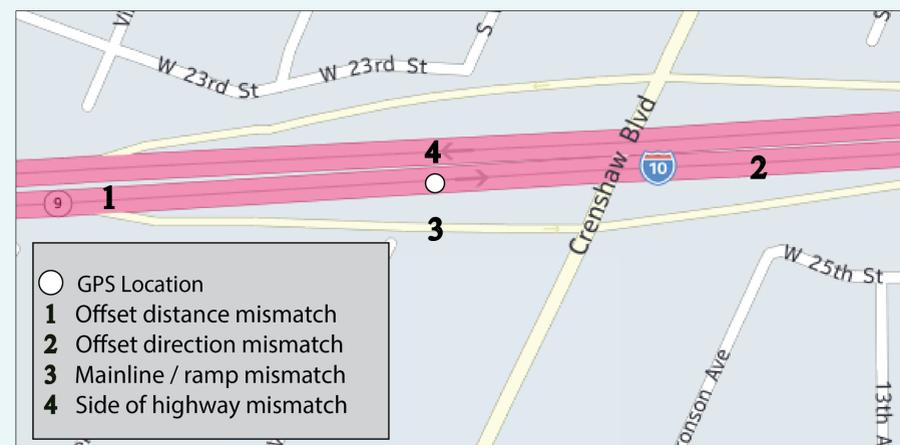
GPS Coordinate in Parking Lot



GPS Coordinate Off Roadway



Intersection Mismatch, Offset Distance Mismatch, Side of Highway Mismatch, Mainline/Ramp Mismatch, and Offset Direction Mismatch



**Multiple Mismatch Issues** A GPS coordinate exhibited two or more of the mismatch types identified.

**Unknown** The descriptive location references an unidentifiable road or landmark.

**Correct** A GPS coordinate was deemed correct if the location was within 500 feet (154.2 meters) of the descriptive location and could not be classified into the other categories.

## Results

Categorization Counts by Year

Category	2009	2010	2011	2009-2011
Systematic GPS error	31 (8.1%)	24 (6.3%)	27 (7%)	82 (7.1%)
GPS coordinate in parking lot	12 (3.1%)	11 (2.9%)	13 (3.4%)	36 (3.1%)
GPS coordinate off roadway	99 (25.8%)	94 (24.5%)	33 (8.6%)	226 (19.7%)
Intersection mismatch	0 (0%)	1 (0.3%)	4 (1.3%)	5 (0.4%)
Offset distance mismatch	37 (9.7%)	38 (9.9%)	71 (18.5%)	146 (12.7%)
Side of highway mismatch	23 (6%)	13 (3.4%)	5 (1.3%)	41 (3.6%)
Ramp/mainline mismatch	9 (2.3%)	17 (4.4%)	6 (1.6%)	32 (2.8%)
Offset direction mismatch	10 (2.6%)	8 (2.1%)	7 (1.8%)	25 (2.2%)
Multiple mismatch issues	9 (2.3%)	18 (4.7%)	6 (1.6%)	33 (2.9%)
Unknown	12 (3.1%)	12 (3.1%)	5 (1.3%)	29 (2.5%)
Correct	141 (36.8%)	147 (38.4%)	206 (53.8%)	494 (43%)
Total	383	383	383	1149

- Overall, 43% of the GPS coordinates were categorized as correct, 2.5% were unknown, and the other 54.5% exhibited some type of discrepancy.

- GPS coordinates located off the roadway were the most frequent error type, comprising nearly 20% of the sample, followed by offset distance mismatch in 13% of the collisions, and systematic GPS errors in 7% of the collisions.

## Discussion & Conclusion

- Location information in crash data is inherently imprecise, especially when officers are required to estimate long distances from the nearest intersection.
- Using GPS can offer greater precision, however GPS coordinates for collision data still suffer from numerous types of inaccuracies reflected in the various error categorizations.
- A major outcome of this research was the inability to verify the accuracy of mismatch categories given the lack of true known collision locations.
- GPS coordinate errors are typically attributed to operator or data entry error, which could be minimized through quality control procedures. However, for California collision data through 2011, it is recommended to thoroughly review GPS coordinates or apply traditional geocoding processes in GIS software to obtain revised coordinates before conducting spatial analyses.

