Evaluating the Influence of Vehicle Lateral Deviation  
on Extra Widening Requirements for Two‑Lane Undivided Roads

Internship Report  
Abhay Singh  
Supervisor: Dr. Saurabh Kumar, IIT Guwahati

To Professor C.Mallikarjuna  
Date: 5 July 2025

✦ **1. Introduction and Background**

The geometric design of roads plays a critical role in ensuring both the operational efficiency and safety of vehicular traffic. Among the geometric elements, **extra widening on horizontal curves** is of particular importance, as it directly affects a vehicle’s ability to negotiate a curve without encroaching on adjacent lanes or departing the carriageway. Traditional standards, such as **IRC 73:1980** in India and guidelines provided by AASHTO and Austroads internationally, have historically relied on simplified formulas derived from swept path analysis of idealized vehicles. These provisions, although valuable, often fail to account for **dynamic real-world behaviors**, such as steering corrections, inconsistent lane positioning, and overtaking maneuvers, all of which are common in mixed-traffic scenarios.

This project was initiated to address these limitations by integrating empirical data analysis with a structured literature review and software tool development. Specifically, the work aimed to quantify how lateral deviation varies across different vehicle types, curve geometries, and driver behavior patterns, and to benchmark these findings against codal provisions. The study sought to develop recommendations that would enable road designers and policymakers to better align design standards with actual operational demands.

✦ **2. Objectives of the Study**

The research was structured around four core objectives:

1. **Developing a Modular Software Toolkit:**  
   A set of Python scripts capable of automating data cleaning, trajectory processing, behavior classification, encroachment detection, and visualization, thereby ensuring consistent, reproducible analysis.
2. **Conducting a Systematic Literature and Codal Review:**  
   A rigorous synthesis of international research and guidelines to contextualize findings within a broader framework and identify gaps between theory and practice.
3. **Empirically Measuring Lateral Deviation:**  
   Using high-resolution trajectory data, the project quantified the actual curve widening required across vehicle categories, curve types (transition versus circular), and maneuver scenarios (overtaking versus lane-following).
4. **Developing Evidence-Based Recommendations:**  
   Proposing refined widening allowances and design considerations that more accurately reflect observed behavior, thereby improving safety margins and operational reliability.

✦ **3. Data Acquisition and Field Measurement**

The cornerstone of the analysis was a set of **high-precision trajectory datasets** collected through Total Station surveys and drone-based video capture. These methods allowed the accurate recording of vehicle positions at fine temporal resolutions, capturing not just the geometric path but also **dynamic variables** such as speed, heading angle, and yaw rate.

The data covered a diverse range of vehicles—cars, bikes, LCVs, and trucks—across multiple road sections. Particular attention was given to **transition curves**, where drivers tend to make more significant steering adjustments compared to constant-radius curves. This comprehensive dataset enabled a detailed analysis of vehicle positioning and variability.

✦ **4. Systematic Literature Review and Codal Benchmarking**

To anchor the findings in a broader context, a **PRISMA-compliant literature review** was conducted. The process included:

* **Database Searches:** Using keywords related to curve widening, lateral deviation, driver behavior, and geometric design, 844 records were identified across Scopus, IEEE Xplore, Google Scholar, ScienceDirect, and PubMed.
* **Screening and Filtering:** After duplicate removal and title/abstract screening, 172 articles underwent full-text assessment.
* **Inclusion Criteria:** Papers were selected if they provided empirical measurements of vehicle lateral behavior, modeling approaches for curve widening, or codal specifications.
* **Final Dataset:** 95 studies and design documents were included, covering a wide spectrum of methodologies and standards.

This review highlighted the persistent reliance on simplified geometric models and the lack of integration of empirical driver behavior into widening guidelines.

✦ **5. Development of the Analytical Toolkit**

Recognizing the need for **scalable, consistent, and reproducible analysis**, this project involved developing **12 modular Python scripts**, each targeting a specific aspect of data processing, behavior classification, and visualization.  
Below is a structured description of each script as presented:

•**ULTIMATE CODE**  
The foundational all-in-one pipeline designed to process, project, and visualize vehicle trajectories in a single workflow. This script takes raw smoothed CSV files and homography files as input and performs several critical tasks: projecting each vehicle trajectory onto the road reference geometry (centreline and edges), calculating signed and absolute deviations, estimating dynamic curve widening, and computing other parameters such as longitudinal speed, lateral speed, yaw rate, and heading angles. The output includes a detailed CSV with all computed columns, along with trajectory alignment plots and high-resolution deviation graphs. This module serves as the starting point for most downstream analysis.

•**DRIVER BEHAVIOR CLASSIFICATION CODE**  
This script categorizes driver behavior by computing **third-order steering entropy** for each vehicle. It uses a **3-component Gaussian Mixture Model (GMM)** to derive data-driven thresholds, classifying driving patterns into Normal, Distracted, or Erratic categories. The script also applies **K-Means clustering** to provide a reference split of the entropy values. Outputs include two CSV files: one listing each vehicle’s entropy metric and behavior class, and another summarizing the GMM model thresholds and diagnostic statistics. These outputs support consistent behavioral segmentation of all vehicle passes.

•**ASA/SDSA/SAR/SRR STEERING ENTROPY CODE**  
A specialized module for analyzing steering variability using complementary metrics. This script calculates four key indicators—**Average Steering Angle (ASA), Standard Deviation of Steering Angle (SDSA), Steering Angle Range (SAR), and Steering Reversal Rate (SRR)**—either from raw heading angle time series or steering entropy outputs. Thresholds can be applied to categorize the steering behavior, and the script generates a time-stamped CSV with all metrics and behavior labels. This level of detail provides a nuanced view of steering stability and control.

•**CSV MERGING CODE**  
This utility script consolidates multiple processed trajectory files into a single dataset. It automatically appends file index suffixes (e.g., \_F1, \_F2) to Vehicle IDs to maintain unique identifiers and concatenates all rows in a consistent format. The merged CSV output becomes the master dataset for further analysis and ensures that all files are combined without duplication or loss of metadata.

•**VEHICLE SUMMARY CODE**  
A reporting module that creates a **comprehensive summary of all numeric variables** in the processed or merged dataset. For each metric, it calculates descriptive statistics such as mean, minimum, maximum, standard deviation, and coefficient of variation (CV), both by vehicle type and by individual Vehicle ID. The script generates two output CSVs—one summarizing results by vehicle category (Car, Bike, Truck, LCV) and another detailing results per vehicle. This information helps identify patterns of variability, operational differences between vehicle classes, and any outlier behaviors.

•**CSV WITHOUT OVERTAKING CASES CODE**  
A specialized script to **remove overtaking episodes** from the processed trajectory data. Using the output of the overtaking detection script, this module identifies and eliminates all intervals during which a vehicle crosses into the opposing lane or initiates overtaking maneuvers. The resulting dataset focuses solely on standard lane-following behavior, ensuring that curve widening and deviation estimates are not artificially inflated by passing events.

•**CURVE CLASSIFICATION CODE**  
This script labels each point in a vehicle trajectory file by its location along the road alignment: **tangent, transition spiral 1, circular curve, or transition spiral 2**. Each (X, Y) coordinate is projected onto the surveyed centreline, and a corresponding "Region" label is assigned. The output includes a CSV with the new classification column and a high-resolution visualization plot showing the centreline color-coded by segment. This classification enables analysis of how driver behavior varies by curve element.

•**GRAPHS GENERATOR CODE**  
An automation script for creating a **full set of visualization outputs**. For each numeric variable, the script produces boxplots, time-series plots, and correlation matrices, along with variance inflation factor (VIF) tables to assess multicollinearity. All figures are saved as high-resolution PNG images, providing ready-to-use visuals for reports, presentations, and publications.

•**BOXPLOTS VISUALIZATION CODE**  
This script generates twelve standardized box-and-whisker plots capturing the **dispersion, median, and outlier behavior** for key kinematic and geometric metrics. Outputs include plots for variables such as signed deviation, lateral and longitudinal speeds, accelerations, yaw rate, and estimated curve widening. The plots help quickly assess data quality, identify extreme cases, and compare behavior across vehicles.

•**OVERTAKING DETECTION CODE**  
A robust module that automatically detects overtaking maneuvers within processed trajectory files. It records entry and exit coordinates, calculates the overtaking distance and duration, and flags whether oncoming traffic was present during each maneuver. The output CSV provides a structured summary of overtaking events, serving as the foundation for further filtering and classification.

•**OVERTAKING CLASSIFICATION CODE**  
Building on the detection results, this script assigns **categorical labels** to every vehicle in the dataset: Overtaking with Oncoming Vehicles, Overtaking without Oncoming Vehicles, In-Lane with Oncoming Vehicles, or Free-Flowing. This labeling enables targeted analysis of risk exposure and behavior under different traffic interactions. The output is a clean CSV with overtaking categories directly attached to each Vehicle ID.

•**ENCROACHMENT CLASSIFICATION CODE**  
A specialized script for detecting **lane encroachment and off-tracking events**. It scans each trajectory and identifies when a vehicle crosses the centerline or drifts outside the pavement edge. For each event, it computes encroachment area, duration, and distance, then labels the behavior accordingly. The output includes a detailed table quantifying where, when, and how drivers violated lane boundaries.

Collectively, this toolkit enabled the efficient, reproducible, and detailed analysis of vehicle trajectories, behavioral patterns, and geometric design compliance, forming the technical foundation for all insights and recommendations in this study.

✦ **6. Behavior Classification and Entropy Analysis**

One of the unique contributions of this project was the application of **steering entropy metrics** to categorize driver behavior. Steering entropy was computed by measuring the unpredictability in the time series of heading angles. Specifically:

* **Average Steering Angle (ASA):** Mean value of the steering angle.
* **Standard Deviation of Steering Angle (SDSA):** Variability around the mean.
* **Steering Angle Range (SAR):** Difference between maximum and minimum steering angle.
* **Steering Reversal Rate (SRR):** Frequency of steering direction changes.

By applying a **3-component Gaussian Mixture Model (GMM)** to the distribution of third-order steering entropy values, thresholds were derived for classifying driver behavior into **Normal, Distracted, and Erratic categories**. In this approach, each vehicle’s trajectory was assigned a probability of belonging to each of the three clusters based on its entropy score, and the cluster with the highest probability was taken as the final behavior class.

After model fitting, the resulting cluster proportions indicated that a significant fraction of drivers exhibited non-normal steering behavior. Specifically, approximately **40–45%** of all classified vehicle passes were categorized as either Distracted or Erratic. This figure was calculated by summing the proportions of these two higher-entropy clusters, as shown in the model summary output:

|  |  |
| --- | --- |
| **Cluster** | **Proportion of Vehicles** |
| Normal | 58% |
| Distracted | 25% |
| Erratic | 17% |

Adding the proportions of Distracted and Erratic categories gives:

25% + 17% = 42\%

This means that **about 4 in every 10 drivers deviated from the smooth, consistent steering patterns assumed in traditional design standards**. The high prevalence of such non-normal behavior was especially pronounced on **transition curves**, where drivers tended to make more frequent and abrupt steering corrections to align with the changing roadway geometry.

This insight underscores the importance of including **driver variability and behavioral inconsistency in geometric design considerations**. Conventional codal provisions generally rely on idealized swept path assumptions and assume that most drivers maintain a stable lateral position. In reality, the data clearly demonstrated that a substantial proportion of drivers operate with elevated entropy and unpredictability, directly contributing to larger and more variable curve widening demands.

In summary, the GMM-based classification not only provided an objective, data-driven framework to segment steering behavior but also quantified the scale of variability in the real-world driving population. This evidence forms a strong justification for **recalibrating extra widening allowances and safety margins** to reflect the actual distribution of driver behaviors observed under mixed-traffic conditions.

✦ **7. Quantitative Findings on Curve Widening**

**Estimated Curve Widening:**

* Mean widening ranged from **0.58 m for cars** to **0.89 m for LCVs**.
* Maximum widening reached **2.53 m for trucks**, more than double the upper limit of IRC guidelines.
* Transition curves exhibited a higher median widening (~0.6 m) and greater dispersion compared to circular curves (~0.55 m median).

**Variability and Safety Margins:**

* Standard deviations in widening exceeded **0.4 m** across vehicle types.
* Coefficients of variation above **0.5** demonstrated significant inconsistency in lateral positioning.
* Encroachment events were frequently recorded, further underscoring the mismatch between standard assumptions and real behavior.

These findings suggest that **current design provisions substantially underestimate lateral clearance needs**, particularly under mixed-traffic conditions.

✦ **8. Comparative Codal Analysis**

A multi-standard comparison was conducted, evaluating IRC, AASHTO, Austroads, DMRB, and JTG guidelines. Observations include:

* All standards relied heavily on **theoretical swept path models**, ignoring behavioral factors such as steering correction and psychological margins.
* Even the more conservative AASHTO guidelines did not account for the observed maximum deviations in transition curves.
* No codal document examined included allowances for entropy-based behavior classification or overtaking-driven variability.

These insights demonstrate the necessity of updating geometric design criteria to reflect empirical evidence.

✦ **9. Recommendations for Practice**

Drawing on these findings, several recommendations are proposed:

* **Integration of Empirical Data:** Design standards should incorporate trajectory-based measurements, particularly for transition and circular curves.
* **Adaptive Widening Provisions:** Widening allowances should be dynamic, adjusting for vehicle category, behavior class, and observed lateral deviation.
* **Restricted Overtaking Policies:** Given the distortion of lane usage during overtaking, specific restrictions should be considered on critical curve sections.
* **Continuous Data Collection Programs:** Establishing permanent monitoring stations will enable regular calibration of design standards.
* **Behavioral Risk Weighting:** Road design should account for the proportion of Distracted and Erratic drivers in the traffic mix.

By implementing these recommendations, roadway designs can better accommodate real driver behavior, enhancing safety and efficiency.

✦ **10. Implications for Future Research**

This project establishes a replicable framework for trajectory analysis and design evaluation. Future studies could:

* Apply the methodology to **urban and mountainous road networks**.
* Extend analysis to **wet or low-friction conditions**, where deviations are likely to increase.
* Evaluate the impact of **driver assistance systems (ADAS)** on lane-keeping variability.
* Develop **predictive models of lateral deviation** under different traffic densities and environmental conditions.

✦ **11. Conclusion**

In summary, this research demonstrates that **real-world curve widening requirements far exceed those prescribed by existing standards**, primarily due to dynamic steering corrections, behavioral variability, and overtaking behavior. By combining high-resolution trajectory data, advanced classification techniques, and systematic benchmarking, the study provides a compelling case for revising geometric design guidelines.

The toolkit developed in this project can serve as a robust foundation for future analyses, offering an accessible, data-driven approach to improving roadway safety and performance.

✦ **12. Acknowledgment**

I am deeply grateful to **Dr. Saurabh Kumar**, whose guidance, expertise, and support were instrumental in every phase of this project. His mentorship has not only enriched my technical understanding but also strengthened my commitment to evidence-based research and innovation in the field of transportation engineering.