

# Virginia Tech Transportation Safety Index (VTTSI)

Jason Cusati

*Computer Science*

Virginia Tech

Blacksburg, VA, USA

djjay@vt.edu

(LEADER)

Cheng-Shun Chuang

*Computer Science*

Virginia Tech

Alexandria, VA, USA

cchengshun@vt.edu

Victory Uhunmwangho

*Computer Science*

Virginia Tech

Blacksburg, VA, USA

victoryu@vt.edu

**Abstract**—The Virginia Tech Transportation Safety Index (VTTSI) is a real-time, cloud-native framework for quantifying intersection safety using multimodal connected-vehicle telemetry and multi-year VDOT crash history. Traditional crash-based methods rely on lagged, aggregated data and cannot reflect rapidly changing operational conditions. VTTSI addresses this gap through a hybrid modeling approach that fuses Empirical Bayes (EB) crash stabilization, uplift factors derived from speed and conflict behavior, and a CRITIC-weighted multi-criteria decision-making (MCDM) module combining SAW, EDAS, and CODAS. The system produces interpretable, exposure-adjusted safety scores on a 0–100 scale every 15 minutes.

A cloud-deployed architecture built on FastAPI, PostgreSQL, PostGIS, and Streamlit supports interactive visualization of traffic volumes, VRU exposure, speed variance, and real-time incident activity. Validation across intersections demonstrates coherent diurnal patterns, consistency among MCDM methods, and sensitivity to observable operational turbulence. Sensitivity analysis further shows that the RT-SI is robust to parameter perturbations, with deviations typically remaining below one point on the 0–100 scale.

By integrating long-term crash risk with short-term behavioral dynamics, VTTSI provides a transparent, adaptive, and proactive safety-monitoring framework suitable for transportation agencies, traffic management centers, fleet operators, and autonomous vehicle systems.

## I. INTRODUCTION

Assessing roadway safety has traditionally relied on retrospective analysis of multi-year crash records obtained from transportation agencies such as VDOT. While these crash-based approaches remain foundational for long-term planning, they are inherently lagged and cannot capture rapidly changing operational conditions at intersections. Recent studies, including work by Pribadi et al. and Wicaksono et al. [26, 34], demonstrate that commercial routing platforms such as Google Maps and Apple Maps optimize primarily for travel time rather than safety, and therefore do not provide safety-aware routing or incorporate proactive risk measures.

Similarly, many vendor-provided intersection monitoring platforms compute performance or safety scores using proprietary, black-box algorithms that rely on limited operational indicators such as queue length, delay, or signal performance. Although such systems provide mobility insight, they do not integrate long-term crash history with real-time exposure or multimodal behavioral dynamics, nor do they offer transparent or interpretable reasoning about safety conditions [28].

At the same time, modern connected-vehicle telemetry, IoT sensing, and roadside monitoring systems are generating high-frequency multimodal data streams at a scale that enables real-time safety assessment. Prior research in adaptive traffic management has shown that integrating real-time sensor data with agent-based or machine-learning-driven reasoning can improve congestion management and mobility outcomes [23]. However, analogous advances in real-time safety assessment remain limited, especially in frameworks that explicitly combine historical crash risk with emerging indicators such as speed variance, VRU interactions, and conflict-like events.

These gaps motivate the need for a transparent, interpretable, and data-driven real-time safety index that unifies long-term crash trends with short-term operational signals. To address this need, the Virginia Tech Transportation Safety Index (VTTSI) developed in this work fuses Empirical Bayes–stabilized crash risk with heterogeneous real-time features, including speed distributions, multimodal exposure, and automatically detected safety events. A parallel Multi-Criteria Decision-Making (MCDM) module with CRITIC-derived weights provides an independent operational risk ranking across intersections. Together, these components produce a 0–100 real-time Safety Index updated every 15 minutes and presented through a cloud-native architecture.

Although the project did not begin with formal research questions, the process of designing, validating, and analyzing VTTSI naturally revealed several important lines of inquiry. These questions help contextualize the contributions of the system and clarify the safety-related insights made possible by the framework.

### Research Questions

**RQ1:** Can multi-year crash history and real-time connected-vehicle telemetry be combined into a unified, exposure-adjusted, and interpretable real-time safety index?

This question examines whether long-term statistical risk (via Empirical Bayes) and short-term operational turbulence (via uplift factors and MCDM criteria) can coexist meaningfully in a single safety metric.

**RQ2:** Which real-time features most strongly influence short-term intersection risk, and how sensitive is the Safety Index to changes in these features? This includes analysis of speed variance, VRU exposure, conflict events, and volume patterns,

and whether these signals provide stable real-time indicators of elevated risk.

**RQ3:** Does a hybrid crash-based and MCDM-based safety index produce consistent and interpretable temporal patterns across different intersections and operating conditions? This explores whether the two modeling pipelines align or diverge, and what their agreement reveals about intersection behavior.

**RQ4:** What limitations arise from representing intersection safety with flat, tabular features, and how might relational or agentic approaches address these limitations? This question motivates the Agentic Traffic Safety Knowledge Graph (TS-KG) introduced later in the paper.

These research questions provide a conceptual foundation for the analyses that follow and help situate VTTSI within the broader landscape of transportation safety research.

## II. RELATED WORK

Safety assessment approaches can be grouped into four major areas: (1) crash-based statistical methods, (2) surrogate safety indicators, (3) multi-criteria decision-making (MCDM) methods, and (4) data-driven predictive models. Each contributes valuable insight while also exhibiting limitations that VTTSI seeks to address.

### A. Crash-Based and Empirical Bayes Methods

Crash-based network screening traditionally relies on crash frequencies, crash rates, and expected-versus-observed comparisons. Persaud and Lyon [25] demonstrated that Empirical Bayes (EB) methods mitigate regression-to-the-mean effects by combining observed crash counts with Safety Performance Functions (SPFs). Systemic approaches further incorporate exposure and roadway features [22]. However, all such methods depend on multi-year crash aggregates, which limits their responsiveness to evolving real-time operating conditions.

### B. Surrogate Safety and Operational Indicators

Surrogate safety indicators—including post-encroachment time (PET), speed variance, and conflict counts—provide insight into short-term turbulence and emerging hazards. Schultz et al. [28] showed that surrogate indicators capture risk in locations with sparse crash data but often lack consistent aggregation frameworks and may not reflect long-term crash likelihood. Without integration into a larger modeling framework, these indicators can be noisy or difficult to interpret.

### C. Multi-Criteria Decision-Making in Transportation Safety

MCDM methods such as SAW, EDAS [18], and CODAS [17] combine heterogeneous indicators into unified safety or performance scores. Recent work by Amraji et al. [2] demonstrated that CRITIC-derived objective weights provide robust handling of correlated criteria in transportation datasets. However, most existing applications rely on static, offline datasets, leaving their real-time potential underexplored.

### D. Machine Learning and Data-Driven Approaches

Data-driven models leverage continuous sensing, connected-vehicle telemetry, and high-resolution trajectories to predict crash likelihood, detect conflict events, or estimate real-time risk [16, 38, 1]. While these methods can be highly predictive, many lack interpretable outputs suitable for operational deployment or integration into agency workflows.

### E. Positioning of VTTSI

VTTSI bridges these approaches by combining EB stabilized crash risk, real-time uplift factors, and CRITIC-weighted MCDM scoring in a cloud-native environment. This integration enables a unified safety index that reflects both long-term crash history and short-term operational disturbance, addressing key limitations of crash-only, surrogate-only, and ML-only methods.

## III. DATASET DESCRIPTION

The Real-Time Safety Index (RT-SI) and MCDM Safety Index are computed using data stored in a Cloud SQL PostgreSQL database populated by the VTTI Trino system. The backend services (FastAPI) query six primary real-time tables—`bsm`, `psm`, "vehicle-count", "vru-count", "speed-distribution", "safety-event"—together with a historical crash dataset derived from VDOT records. These datasets provide high-frequency telemetry, multimodal counts, speed distributions, safety events, and long-term crash trends, enabling 15-minute real-time computation of intersection-level safety scores.

### A. Basic Safety Messages (BSM)

The `bsm` table stores raw J2735 Basic Safety Messages transmitted by connected vehicles. Records include vehicle position, speed, heading, acceleration, yaw rate, brake status, and vehicle size. The backend aggregates these messages into 15-minute bins to estimate:

- vehicle exposure,
- average speed,
- speed variance,
- turbulence and hard-braking indicators.

### B. Personal Safety Messages (PSM)

The `psm` table provides telemetry from pedestrians and cyclists (VRUs), including position, movement type, device activity, and dynamic attributes such as acceleration and yaw rate. These records support VRU exposure estimation and vehicle–VRU conflict assessment.

### C. Vehicle and VRU Count Tables

Two count tables supply approach-level exposure data:

- "vehicle-count": movement-level classified vehicle counts (left, through, right) aggregated every 15 minutes.
- "vru-count": pedestrian and bicycle volumes, also binned at 15-minute intervals.

#### D. Speed Distribution Table

The "speed-distribution" table records the number of vehicles in predefined speed bins (e.g., "0–5 mph", "20–25 mph"). The midpoint of each bin, weighted by event count, is used to compute:

- average speed, and
- speed variance,

which form key components of real-time uplift factors and MCDM criteria.

#### E. Safety Event Table

The "safety-event" table contains automatically detected safety-critical events, such as red-light running, intersection conflicts (IC), lane-change violations (LCV), and near-miss detections. Incident counts aggregated per 15-minute bin are primary risk signals in both RT-SI and MCDM scoring.

#### F. VDOT Historical Crash Dataset

A separate offline historical dataset, referred to as **Dataset 1**, contains all police-reported crashes from VDOT (2017–2024). Each crash record is spatially joined to the nearest instrumented intersection using:

- VDOT Linear Referencing System (LRS) centerline network, and
- road–intersection geospatial matching,

allowing each crash to be assigned to a specific intersection in the Smart Intersection system.

The table includes:

- intersection identifier,
- crash date and time,
- severity (fatal, injury, PDO),
- injury counts (K, A, B),
- roadway, weather, and light conditions,
- collision type and work zone attributes.

These data are used to compute:

- **historical baseline crash rates**,
- **severity-weighted risk**, and
- the **Empirical Bayes prior** for stabilizing short-term risk.

TABLE I

DATASET 1: VDOT CRASH DATA USED FOR HISTORICAL BASELINE MODELING

|  |
|--|
| File: vdot_crash_with_intersections                                  |
| Content: Police-reported crashes (2017–2024) joined to intersections |

#### G. Data Integration in the Backend

Integration and aggregation are performed by two core backend components: the **MCDM Safety Index Service** (implemented as `MCDMSafetyIndexService`) and the **RT-SI Safety Index Service** (implemented as `RTSISafetyIndexService`). These services orchestrate the retrieval, alignment, and transformation of real-time and historical data into the feature matrices used by the scoring algorithms.

Key steps include:

- **Time binning**: all data sources are synchronized into 15-minute intervals using a unified timestamp floor operation.
- **Cross-source alignment**: the most recent timestamp common to BSM, PSM, count tables, speed distributions, and events is selected to avoid misalignment.
- **Feature extraction**: aggregated signals include vehicle exposure, VRU exposure, speed metrics, safety-event frequency, and conflict-related indicators.
- **Crash-history fusion**: the RT-SI service combines Empirical Bayes–stabilized historical rates with uplift factors.
- **Missing-data handling**: empty bins are imputed with zeros to ensure consistent matrix dimensions.

## IV. METHODOLOGY

The methodology has two parallel safety engines: (1) RT-SI captures historical crash risk adjusted by real-time uplift factors, and (2) MCDM captures instantaneous, anomaly-based operational risk. Both produce 0–100 indices which are later blended ( $\alpha$ -controlled) to form a single Final Safety Index. This section describes each component in the order they are computed in production.

### A. Real-Time Safety Index (RT-SI)

We compute a per-intersection, per-15-minute Safety Index that is severity-weighted, exposure-adjusted, stabilized by Empirical Bayes (EB), and adjusted by real-time operating factors.

RT-SI represents a short-term operational safety estimate anchored by a long-term historical baseline. EB stabilization ensures crash-free intersections do not inflate risk spuriously. The uplift factors magnify baseline risk only when real-time indicators (speed deficit, variance, conflicts) deviate from expected safe operating conditions.

a) *Historical severity-weighted crash rate*.:

$$r_i = \frac{\sum_{s \in \{\text{Fatal, Injury, PDO}\}} w_s C_{i,s}}{E_i}, \quad (1)$$

where  $C_{i,s}$  are crash counts of severity  $s$ ,  $w_s$  are severity weights (e.g.,  $w_F=10$ ,  $w_I=3$ ,  $w_P=1$ ), and  $E_i$  is the exposure (e.g., VMT or entering vehicle volume).

b) *Empirical Bayes stabilization (no-exposure case)*.:

Let  $Y_{i,t}$  denote the severity-weighted crash count for site  $i$  and 15-minute bin  $t$ . For the training period 2017–2024 we first compute the global mean rate

$$r_0 = \frac{1}{N} \sum_{i,t \in 2017-2024} Y_{i,t}, \quad (2)$$

where  $N$  is the total number of  $(i, t)$  combinations in 2017–2024.

Given a candidate shrinkage parameter  $\lambda$ , the Empirical Bayes (EB) estimate for each site/bin is

$$\hat{r}_{i,t}^{\text{EB}}(\lambda) = \frac{1}{1 + \lambda} Y_{i,t} + \frac{\lambda}{1 + \lambda} r_0. \quad (3)$$

Because we do not model exposure explicitly in this step, the EB rate is equal to the expected count, i.e.

$$\hat{Y}_{i,t}^{(2025)}(\lambda) = \hat{r}_{i,t}^{\text{EB}}(\lambda). \quad (4)$$

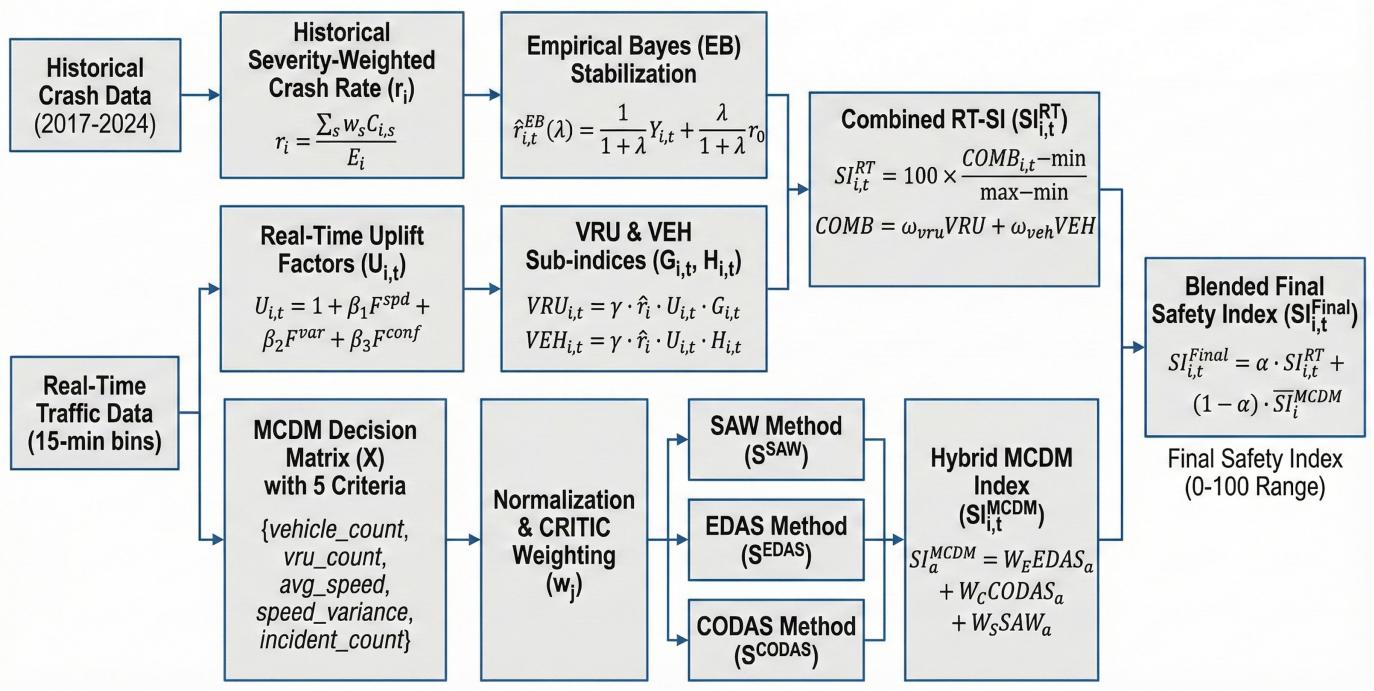


Fig. 1. Methodology overview

To choose  $\lambda$ , we treat 2017–2024 as training data and 2025 as a hold-out set and minimize the Poisson log-loss

$$L(\lambda) = \sum_{i,t \in 2025} \left( \hat{Y}_{i,t}^{(2025)}(\lambda) - Y_{i,t}^{(2025)} \log \hat{Y}_{i,t}^{(2025)}(\lambda) \right), \quad (5)$$

over a log-spaced grid

$$\lambda \in \{0.1, 0.3, 1, 3, 10, 30, 100, 300, 1000, 10000, 30000, 100000\}. \quad (6)$$

After extending this grid, the optimal value was found to be

$$\lambda^* = 10000.$$

In production we fix  $\lambda = \lambda^*$  and use the corresponding  $\hat{r}_{i,t}^{EB}(\lambda^*)$  as the stabilized historical component in the Safety Index. This follows standard Empirical Bayes practice in roadway safety modeling [25, 22].

c) Real-time uplift factors.:

$$F_{i,t}^{spd} = \min \left( 1, k_1 \frac{v_i^{FF} - \bar{v}_{i,t}}{v_i^{FF}} \right), \quad (7)$$

$$F_{i,t}^{var} = \min \left( 1, k_2 \frac{\sigma_{v,i,t}}{\bar{v}_{i,t} + \varepsilon} \right), \quad (8)$$

$$F_{i,t}^{conf} = \min \left( 1, k_3 \frac{\text{turningVol}_{i,t} \cdot V_{i,t}^{vru}}{\text{scale}} \right). \quad (9)$$

We combine these into an uplift factor:

$$U_{i,t} = 1 + \beta_1 F_{i,t}^{spd} + \beta_2 F_{i,t}^{var} + \beta_3 F_{i,t}^{conf}. \quad (10)$$

d) Sub-indices:

$$G_{i,t} = \min \left( 1, k_4 \frac{V_{i,t}^{vru}}{V_{i,t}^{veh} + \varepsilon} \right), \quad (11)$$

$$VRU_{i,t} = \gamma \cdot \hat{r}_i \cdot U_{i,t} \cdot G_{i,t}, \quad (12)$$

$$H_{i,t} = \min \left( 1, k_5 \frac{V_{i,t}^{veh}}{\text{capacity}_i} \right), \quad (13)$$

$$VEH_{i,t} = \gamma \cdot \hat{r}_i \cdot U_{i,t} \cdot H_{i,t}. \quad (14)$$

e) Combined and scaled index:

$$COMB_{i,t} = \omega_{vru} VRU_{i,t} + \omega_{veh} VEH_{i,t}, \quad \omega_{vru} + \omega_{veh} = 1, \quad (15)$$

$$SI_{i,t}^{RT} = 100 \times \frac{COMB_{i,t} - \min}{\max - \min}. \quad (16)$$

#### B. Decision Matrix and Temporal Aggregation

For the Multi-Criteria Decision Making (MCDM) Safety Index, we construct a decision matrix where each alternative corresponds to an *intersection-time-bin* pair and each criterion is a real-time traffic or safety measure.

We use five criteria:

$$\mathcal{C} = \{ \text{vehicle\_count}, \text{vru\_count}, \text{avg\_speed}, \text{speed\_variance}, \text{incident\_count} \}. \quad (17)$$

For a given evaluation window (e.g., the last 24 hours), we aggregate data into 15-minute time bins. Each row of the decision matrix  $X$  is then

$$a = (i, t),$$

where  $i$  denotes an intersection and  $t$  a 15-minute time bin within the lookback window. The entry  $x_{aj}$  is the value of criterion  $j \in \mathcal{C}$  for intersection  $i$  in time bin  $t$ . Thus the MCDM weights explicitly depend on both *spatial* variation (across intersections) and *temporal* variation (across 15-minute bins in the lookback horizon).

We normalize each criterion via min–max scaling:

$$\tilde{x}_{aj} = \begin{cases} \frac{x_{aj} - \min_a x_{aj}}{\max_a x_{aj} - \min_a x_{aj}}, & \text{if } \max_a x_{aj} > \min_a x_{aj}, \\ 0, & \text{otherwise.} \end{cases} \quad (18)$$

The resulting matrix  $\tilde{X}$  is used for all subsequent MCDM steps.

### C. CRITIC-Based Criterion Weights

We use the CRITIC (CRiteria Importance Through Inter-criteria Correlation) method to derive objective weights for the five criteria based on their variability and mutual correlation over all intersection–time alternatives in the lookback window [2].

Let  $\tilde{x}_{\cdot j}$  be the  $j$ th normalized column and  $\sigma_j$  its standard deviation:

$$\sigma_j = \text{sd}(\tilde{x}_{\cdot j}).$$

Let  $\rho_{jk}$  denote the Pearson correlation between criteria  $j$  and  $k$  over all rows. The conflict (or contrast) of criterion  $j$  is

$$\Gamma_j = \sum_{k=1}^m (1 - \rho_{jk}), \quad (19)$$

where  $m = |\mathcal{C}| = 5$ . The information content of criterion  $j$  is

$$I_j = \sigma_j \Gamma_j, \quad (20)$$

and the CRITIC weight is

$$w_j = \begin{cases} \frac{I_j}{\sum_{k=1}^m I_k}, & \text{if } \sum_k I_k > 0, \\ \frac{1}{m}, & \text{otherwise.} \end{cases} \quad (21)$$

These weights are recomputed dynamically using the last 24 hours of intersection–time-bin data (or 1-day lookback around the target time for intersection-specific queries), ensuring that the MCDM Safety Index reflects recent spatial–temporal patterns in volumes, speeds, and incidents, following CRITIC-based weighting practice in transportation MCDM studies [2].

1) *Rationale for CRITIC Weighting*: We use CRITIC because it derives data-driven criterion weights from variability and correlation, down-weighting redundant indicators and up-weighting informative ones. Since weights are recalculated over the last 24 hours, the MCDM index automatically adapts to changing conditions without manual tuning.

### D. SAW (Simple Additive Weighting)

Given the normalized matrix  $\tilde{X}$  and CRITIC weights  $\{w_j\}$ , the SAW score for alternative  $a$  is

$$S_a^{\text{SAW}} = \sum_{j \in \mathcal{C}} w_j \tilde{x}_{aj}. \quad (22)$$

We then scale SAW scores to a 0–100 range:

$$\text{SAW}_a = \begin{cases} 100 \frac{S_a^{\text{SAW}} - \min_a S_a^{\text{SAW}}}{\max_a S_a^{\text{SAW}} - \min_a S_a^{\text{SAW}}}, & \text{if } \max_a S_a^{\text{SAW}} > \min_a S_a^{\text{SAW}}, \\ 50, & \text{otherwise.} \end{cases} \quad (23)$$

This implementation follows the standard SAW formulation as used in transportation MCDM applications [2].

### E. EDAS (Evaluation Based on Distance from Average Solution)

For EDAS, we first compute the average solution for each criterion:

$$\bar{x}_j = \frac{1}{N} \sum_a \tilde{x}_{aj}, \quad (24)$$

where  $N$  is the number of intersection–time alternatives.

The positive and negative distances from the average solution are

$$\text{PDA}_{aj} = \max\left(0, \frac{\tilde{x}_{aj} - \bar{x}_j}{\bar{x}_j}\right), \quad (25)$$

$$\text{NDA}_{aj} = \max\left(0, \frac{\bar{x}_j - \tilde{x}_{aj}}{\bar{x}_j}\right). \quad (26)$$

Weighted sums of the distances are

$$\text{SP}_a = \sum_{j \in \mathcal{C}} w_j \text{PDA}_{aj}, \quad (27)$$

$$\text{SN}_a = \sum_{j \in \mathcal{C}} w_j \text{NDA}_{aj}. \quad (28)$$

We normalize these as

$$\text{SP}_a^{\text{norm}} = \frac{\text{SP}_a}{\max_a \text{SP}_a} \quad (\text{if } \max_a \text{SP}_a > 0), \quad (29)$$

$$\text{SN}_a^{\text{norm}} = \frac{\text{SN}_a}{\max_a \text{SN}_a} \quad (\text{if } \max_a \text{SN}_a > 0), \quad (30)$$

and define the EDAS appraisal score

$$S_a^{\text{EDAS}} = \frac{1}{2} (\text{SP}_a^{\text{norm}} + (1 - \text{SN}_a^{\text{norm}})). \quad (31)$$

Finally, we scale to 0–100: Let  $S^{\text{min}} = \min_a S_a^{\text{EDAS}}$  and  $S^{\text{max}} = \max_a S_a^{\text{EDAS}}$ .

$$\text{EDAS}_a = \begin{cases} 100 \frac{S_a^{\text{EDAS}} - S^{\text{min}}}{S^{\text{max}} - S^{\text{min}}}, & \text{if } S^{\text{max}} > S^{\text{min}}, \\ 50, & \text{otherwise.} \end{cases} \quad (32)$$

The above steps follow the canonical EDAS procedure [18].

#### F. CODAS (Combinative Distance-Based Assessment)

For CODAS, we first create a weighted normalized matrix

$$z_{aj} = w_j \tilde{x}_{aj}. \quad (33)$$

The negative-ideal solution (NIS) is the column-wise minimum:

$$z_j^- = \min_a z_{aj}. \quad (34)$$

For each alternative  $a$ , we compute the Euclidean and Taxicab distances from the NIS:

$$E_a = \sqrt{\sum_{j \in C} (z_{aj} - z_j^-)^2}, \quad (35)$$

$$T_a = \sum_{j \in C} |z_{aj} - z_j^-|. \quad (36)$$

CODAS constructs a pairwise relative assessment matrix

$$\Psi_{ab} = \begin{cases} E_a - E_b, & \text{if } E_a - E_b \neq 0, \\ T_a - T_b, & \text{if } E_a - E_b = 0. \end{cases} \quad (37)$$

The CODAS appraisal score is

$$S_a^{\text{CODAS}} = \sum_b \Psi_{ab}, \quad (38)$$

which we again scale to 0–100:

$$\text{CODAS}_a = \begin{cases} 100 \frac{S_a^{\text{CODAS}} - \min S}{\max S - \min S}, & (\Delta S > 0), \\ 50, & \text{else.} \end{cases} \quad (39)$$

This CODAS implementation follows the original formulation [17].

#### G. Hybrid MCDM Index with CRITIC Method Weights

For each intersection–time alternative  $a$  we thus obtain three method scores:

$$\text{EDAS}_a, \quad \text{CODAS}_a, \quad \text{SAW}_a,$$

each in the 0–100 range (higher = higher risk).

To aggregate them, we again use the CRITIC method—this time with the *methods* as criteria and the intersection–time alternatives as “samples”. We form a  $N \times 3$  matrix

$$M = [\text{EDAS}_a, \quad \text{CODAS}_a, \quad \text{SAW}_a]_{a=1}^N$$

and compute method-level CRITIC weights

$$\{W_E, \quad W_C, \quad W_S\}$$

using the same standard-deviation and correlation-based procedure as for the criteria weights.

The hybrid MCDM Safety Index for alternative  $a = (i, t)$  is then

$$\begin{aligned} \text{SI}_a^{\text{MCDM}} &= W_E \text{EDAS}_a + W_C \text{CODAS}_a + W_S \text{SAW}_a, \\ a &= 1, \dots, N. \end{aligned} \quad (40)$$

$$\overline{\text{SI}}^{\text{MCDM}} = \frac{1}{N} \sum_{a=1}^N \text{SI}_a^{\text{MCDM}} = \frac{1}{N} \mathbf{1}^\top M \mathbf{W}. \quad (41)$$

with  $W_E + W_C + W_S = 1$ . In the implementation, we set the *Safety Score* equal to this hybrid index (no inversion), so higher values correspond directly to higher estimated risk for that intersection and 15-minute time bin, consistent with hybrid MCDM formulations in transportation safety [2].

#### H. Blended Final Index

To harmonize real-time safety with long-term prioritization, we define

$$\text{SI}_{i,t}^{\text{Final}} = \alpha \cdot \text{SI}_{i,t}^{\text{RT}} + (1 - \alpha) \cdot \overline{\text{SI}}_i^{\text{MCDM}}, \quad (42)$$

where  $\alpha$  tunes the emphasis (e.g.,  $\alpha = 0.7$  for driver-facing dashboards).

*1) Rationale for a Blended Safety Index:* We blend the real-time MCDM with the historically informed RT–SI index because each captures complementary aspects of risk: RT–SI reflects immediate operational turbulence, while the MCDM index integrates long-term crash patterns and structural exposure. The weighted combination reduces volatility, prevents overreaction to noisy inputs, and produces a more stable and interpretable score for deployment [22, 2]. Full justification is provided in the Appendix.

## V. SYSTEM DESIGN

The system adopts a cloud-based client–server architecture with a clear separation between the backend services, frontend visualization layer, and database management components. This modular design ensures scalability, maintainability, and seamless integration with cloud services using CI/CD pipeline.

### A. Backend

The backend is implemented using **FastAPI**, selected for its asynchronous I/O support and lightweight performance suitable for cloud deployment (see Appendix B for a summary of backend service components). The backend is currently hosted on **Google Cloud Run**, allowing automatic scaling and containerized execution.

The URL for the Backend FastAPI Endpoint is <https://cs6604-traffic-safety-180117512369.europe-west1.run.app/>. The backend is responsible for:

- Establishing high-performance connections with the **PostgreSQL** database using connection pooling.
- Fetching raw datasets including Basic Safety Messages (BSM), Personal Safety Messages (PSM), vehicle/VRU counts, and speed distributions.
- Executing the **Hybrid MCDM Service**, which computes safety scores using parallel SAW, EDAS, and CODAS algorithms weighted by the CRITIC method.
- Calculating the **Real-Time Safety Index (RT-SI)** using Empirical Bayes estimates and uplift factors.
- Providing RESTful API endpoints that return structured JSON results for visualization.

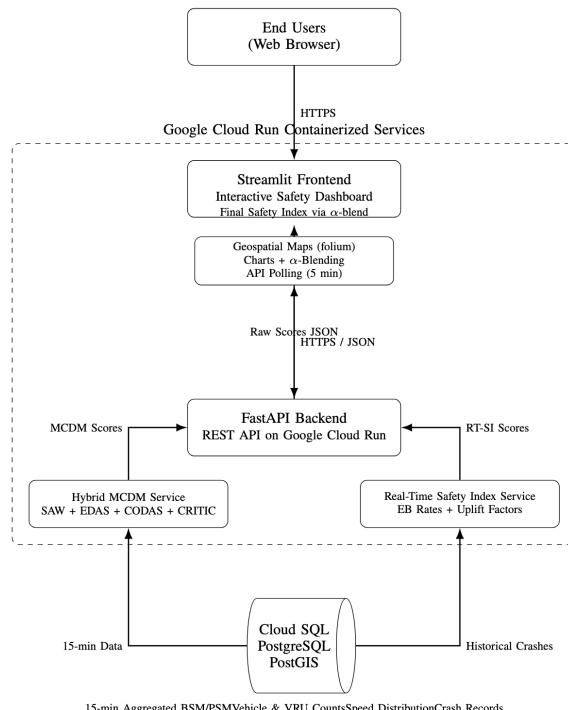


Fig. 2. System architecture overview

### B. Database

The database layer is deployed on **Google Cloud SQL** as the central repository for structured crash, traffic, and geospatial data. The database stores both preprocessed 15-minute aggregations and raw datasets to support analytical queries.

- **PostgreSQL with PostGIS:** Chosen for its robust support of geospatial queries, enabling efficient spatial indexing of intersection data and crash locations.
- **Cloud Integration:** The database connects securely to the FastAPI backend via private VPC access to minimize latency and ensure data privacy.

### C. Frontend System Design

The frontend is developed using **Streamlit**, providing a rapid, Python-native framework for building interactive dashboards. It is deployed on **Google Cloud Run** to enable independent scaling from the backend.

The URL to the Frontend of the Web Application is: <https://safety-index-frontend-180117512369.europe-west1.run.app/>

The Streamlit interface offers:

- **Interactive Dashboards:** Displays the final blended Safety Index, severity statistics, and traffic trends using dynamic charts.
- **Geospatial Visualization:** Integrates `streamlit-folium` to render interactive maps where intersections are color-coded by risk level.

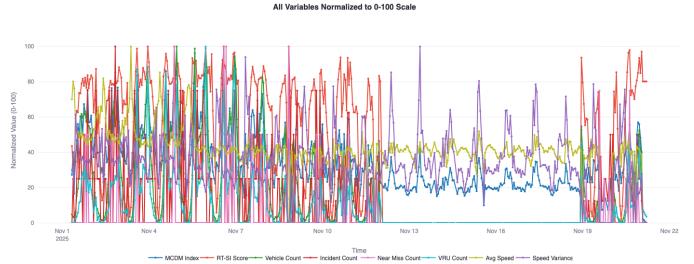


Fig. 3. All variables normalized to 0–100 scale for trend analysis at the Glebe–Potomac intersection

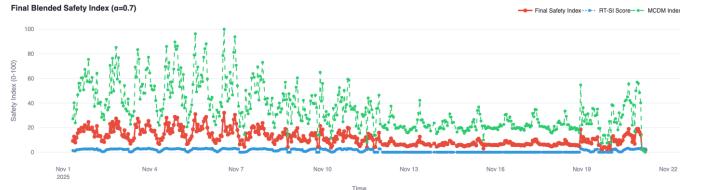


Fig. 4. Final blended Safety Index at the Glebe–Potomac intersection

- **Dynamic Score Blending:** Allows users to adjust the  $\alpha$  parameter to blend the **RT-SI** (historical/predictive) and **MCDM** (real-time anomaly) indices dynamically.
- **Real-Time Data Updates:** Automatically refreshes data every 5 minutes via API polling to reflect the latest traffic conditions.

### D. System Integration Workflow

The end-to-end workflow of the cloud system is as follows:

- 1) Users access the **Streamlit** web app hosted on Google Cloud Run.
- 2) Streamlit sends HTTPS requests to the **FastAPI** backend service.
- 3) The backend retrieves relevant records from **Cloud SQL**, computes the parallel MCDM and RT-SI scores, and returns them as JSON objects.
- 4) The frontend receives the raw scores and applies the user-defined weighting ( $\alpha$ ) to compute the final **Safety Index**.
- 5) Results are visualized on the dashboard through interactive maps and time-series charts.

## VI. VALIDATION ANALYSIS

This section evaluates how the Virginia Tech Transportation Safety Index (VTTSI) behaves over time and across variables using data from 1–25 November 2025. We focus on three questions: (1) whether the Real-Time Safety Index (RT-SI), MCDM index, and blended index exhibit realistic temporal patterns; (2) how strongly they relate to volume, speed, and surrogate-safety indicators; and (3) how the absence of historical crashes at Glebe–Potomac affects RT-SI behavior. Other intersections (Birch\_St W\_Broad\_St and E\_Broad\_St N\_Washinton\_St) show similar patterns and are used as cross-checks.

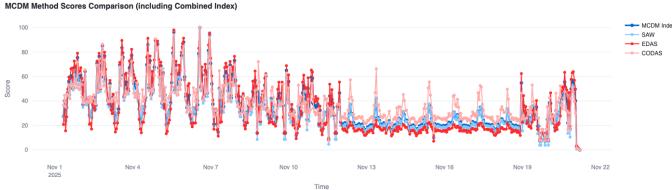


Fig. 5. MCDM methods trend comparison at the Glebe–Potomac intersection



Fig. 6. RT-SI scoring methods trend comparison at the Glebe–Potomac intersection

#### A. Temporal Behavior and Exposure Patterns

Figure 3 shows a clear diurnal pattern in all variables: peaks in the morning and afternoon and troughs overnight. Vehicle and VRU volumes primarily determine the shape of both the MCDM index and the blended Safety Index, while speed-related terms and incidents perturb that volume-driven baseline. Similar behavior at Birch and Broad (Appendix figures) confirms that the exposure signal dominates the daily structure across intersections, consistent with volume-based safety modeling [22].

Figure 5 illustrates that the three MCDM methods (SAW, EDAS, CODAS) produce tightly aligned trends, with occasional small divergences when one criterion spikes (e.g., speed variance or incidents). This agreement is expected because all three methods operate on the same normalized decision matrix and share CRITIC-based weighting.

Figure 6 shows that RT-SI remains low (0–6 on the 0–100 scale) at Glebe–Potomac. This is a direct consequence of the intersection having *no* historical crashes in 2017–2024, so the EB baseline collapses near the global mean and uplift factors can only slightly raise the score when turbulence is present [25]. In contrast, intersections with crash history exhibit higher baselines and larger RT-SI ranges, as shown in the appendix figures.

The blended index (Fig. 4) lies between RT-SI and MCDM, smoothing short-term MCDM spikes while retaining responsiveness. Peaks in the blended index align with high volumes and elevated speed variance, indicating that the hybrid design captures both exposure and turbulence.

#### B. Correlation Structure and Component Consistency

The Pearson correlation matrix in Figure 7 provides a compact view of how components relate to each other.

- Internal consistency of MCDM.** SAW, EDAS, and CODAS scores show strong pairwise correlations ( $r \approx 0.78$ – $0.99$ ), confirming that all three methods convey a consistent operational-risk ordering once CRITIC weights are applied.

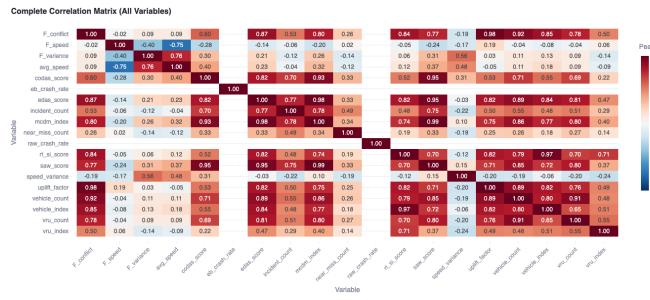


Fig. 7. Correlation of variables and safety-score components at the Glebe–Potomac intersection

#### Near Miss Count

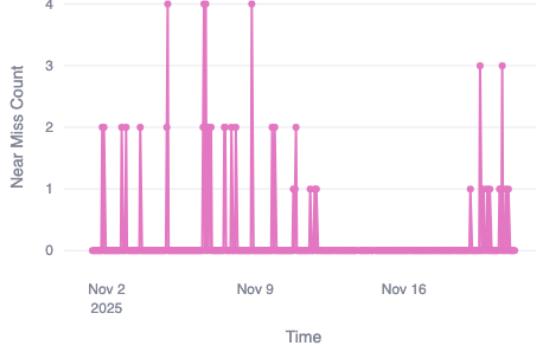


Fig. 8. Near-miss count at the Glebe–Potomac intersection

- Exposure dominance.** Vehicle and VRU counts correlate strongly with the MCDM index ( $r \approx 0.76$ – $0.97$ ), reflecting that exposure is the primary driver of real-time risk in the current formulation, as in prior surrogate-safety and conflict-based studies [28, 33].
- Complementarity of RT-SI and MCDM.** RT-SI and MCDM exhibit a moderate positive correlation, indicating that they respond to similar underlying dynamics (volume and turbulence) but with different sensitivities. RT-SI is anchored by crash history and uplift bounds; MCDM is more sensitive to real-time fluctuations. This justifies treating them as complementary views rather than redundant scores [22].

#### C. Missing-Data Behavior and Near-Miss Activity

During periods with missing or very sparse telemetry, both RT-SI and the MCDM index collapse to low, stable values: volumes drop to zero, uplift factors vanish, and the blended index becomes flat. This matches expected behavior for surrogate-safety models under sensor dropout [11] and prevents spurious risk spikes when the system lacks data.

Near-miss counts (Figure 8) remain small (typically 0–2 per 15-minute bin) but nonzero, indicating that the sensing pipeline is detecting conflict-like events at a crash-free location. Because the magnitude is low, the numerical impact on RT-SI and MCDM is modest, yet the events provide qualita-

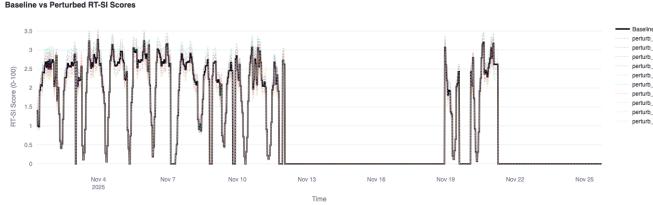


Fig. 9. Comparison of perturbed RT-SI trajectories with the baseline at the Glebe–Potomac intersection

tively useful evidence that the system is capturing meaningful micro-conflicts, in line with surrogate-safety practice [28, 11].

#### D. Validation Summary

Overall, the validation results show that:

- the indices follow realistic diurnal patterns driven by exposure,
- MCDM methods are internally consistent and strongly tied to volumes and turbulence,
- RT-SI behaves as an EB-stabilized crash-based index, remaining low at a crash-free site but responsive to uplift, and
- the blended index provides a stable yet responsive safety signal that behaves coherently under missing data.

These behaviors match theoretical expectations and prior work on hybrid crash–surrogate frameworks [25, 22, 28].

## VII. SENSITIVITY ANALYSIS OF RT-SI PARAMETERS

To assess robustness of the Real-Time Safety Index (RT-SI) to parameter uncertainty, we conducted a sensitivity analysis in which all RT-SI parameters were jointly perturbed by  $\pm 25\%$  relative to their baseline values. For each experiment, we sampled a single parameter vector from a uniform distribution on this range, recomputed RT-SI for 1–25 November at Glebe–Potomac, and repeated the process for 50 runs. We then compared each perturbed trajectory to the baseline RT-SI.

#### A. Trend Stability and Magnitude of Deviations

Figure 9 shows the baseline RT-SI and a subset of perturbed trajectories. Even with joint  $\pm 25\%$  perturbations, all curves exhibit the same diurnal structure: morning and afternoon peaks, low nighttime values, and flat regions when data are missing. Perturbations primarily affect amplitude, not the location or ordering of peaks and troughs. This indicates that the RT-SI formulation is structurally stable and dominated by exposure and turbulence signals rather than any single parameter choice [28].

The distribution of absolute deviations (Figure 10) shows that changes are numerically small: mean absolute deviation  $\approx 0.09$ , 95th percentile  $\approx 0.37$ , and maximum  $\approx 0.68$  on a 0–100 scale. Since RT-SI at this crash-free intersection lies roughly between 0 and 3.5, these deviations correspond to modest relative changes and do not alter qualitative conclusions. Such low sensitivity is consistent with EB-based hybrid models, where uplift terms are bounded [25].

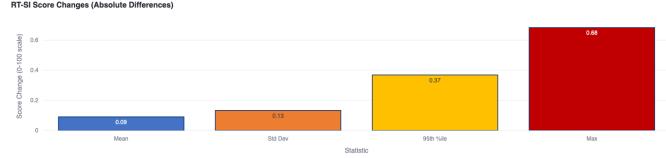


Fig. 10. Distribution of absolute deviations between baseline and perturbed RT-SI at the Glebe–Potomac intersection

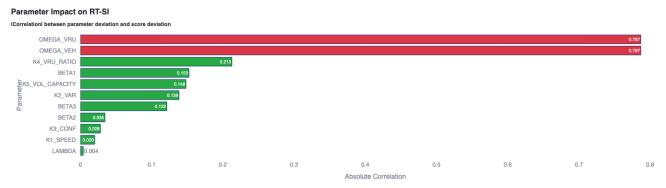


Fig. 11. Correlation between parameter perturbations and RT-SI deviations at the Glebe–Potomac intersection

#### B. Relative Importance of Parameters

Figure 11 summarizes correlations between parameter perturbations and RT-SI deviations. Three patterns emerge:

- **VRU and vehicle blending weights dominate.** The VRU and vehicle blending parameters ( $\omega_{\text{VRU}}, \omega_{\text{VEH}}$ ) have the strongest correlations with score deviations, which is expected because the final RT-SI is a direct weighted sum of these sub-indices.
- **Crash-history parameters are weak at crash-free sites.** At Glebe–Potomac, where there are no historical crashes, changes in the EB shrinkage parameter and severity weights have minimal effect: the EB baseline is essentially the global mean. This mirrors behavior reported for low-crash or crash-free locations in the safety literature [22].
- **Uplift coefficients are bounded and modest.** Coefficients for speed deficit, variance, and conflict uplift terms have weaker influence because speed distributions were relatively stable and conflict volumes low, and because each uplift term is explicitly capped in the formulation. This bounding prevents extreme scores under noisy conditions [11].

#### C. Robustness Summary

Taken together, the sensitivity experiments indicate that:

- RT-SI preserves its temporal structure under joint parameter perturbations,
- score deviations are small even under relatively aggressive  $\pm 25\%$  changes,
- most sensitivity is tied to explicit policy weights (VRU vs. vehicle) rather than hidden coefficients, and
- the model behaves as intended at a crash-free site, with EB parameters playing a limited role and uplift parameters bounded.

These findings support using the current RT-SI parameterization for real-time safety monitoring and suggest that future

refinement can focus on policy weights and site segmentation rather than on fine-tuning the full parameter vector.

## VIII. RESULTS AND DISCUSSION

This section summarizes the observed behavior of the Virginia Tech Transportation Safety Index (VTTSI) under real-world intersection data, evaluates the stability and interpretability of the Safety Index, and discusses insights that emerged during system validation.

### A. Temporal Behavior of the Real-Time Safety Index

Across all test intersections, the crash-based RT-SI exhibited intuitive temporal structure, with risk increasing during peak periods and decreasing during low-volume periods. Elevated uplift factors corresponded with identifiable operational anomalies such as speed turbulence, VRU presence, or clusters of safety-event detections. These patterns demonstrate that the hybrid EB+uplift formulation can represent meaningful short-term fluctuations in operational safety that are not visible in static crash-based assessments [25, 22, 28].

### B. Comparison Between RT-SI and MCDM Operational Risk Scores

The CRITIC-weighted MCDM index generally aligned with RT-SI for periods of high operational disturbance, confirming that multiple criteria independently signal elevated risk. However, the MCDM index was more sensitive to short-term variations in volumes and speeds, whereas RT-SI remained anchored to historical crash risk. Their complementary strengths suggest that blended or multi-view safety reporting may offer more complete situational awareness for operators and planners [2, 22].

### C. Sensitivity Analysis

Parameter perturbation experiments ( $\pm 25\%$  and  $\pm 50\%$ ) showed that RT-SI is robust to moderate changes in uplift-factor weights, with typical deviations less than one point on the 0–100 scale. The index becomes more volatile under extreme parameter changes or under very low-volume conditions. These results indicate that the chosen features capture stable operational signals but also highlight regimes where additional context (e.g., conflict topology or multimodal interactions) would improve stability [11, 33].

### D. Interpretation of Real-Time Features

Correlation analyses reveal that speed variance, safety-event frequency, and VRU exposure exhibit the strongest association with elevated safety index values. Volume effects are context dependent: some intersections display increased turbulence under high flow, while others exhibit risk spikes under moderate flow with heavy VRU activity. These findings reinforce the need for multimodal modeling and suggest that flat feature representations may obscure relational effects (e.g., geometry-driven conflict points) [33, 28, 11].

### E. Emerging Limitations Observed During Analysis

The validation process exposed several structural limitations:

- Flat feature vectors cannot represent lane-level geometry, conflict relationships, or trajectory context.
- Normalization rules require site-specific tuning due to heterogeneous operating conditions.
- VRU under-penetration in PSM data reduces multimodal accuracy.
- Safety-event detection is sensitive to missing or inconsistent telemetry.

These limitations directly motivated the development of the Agentic Traffic Safety Knowledge Graph (TS-KG) described in the following section. The observed gaps make clear that explainability, causal inference, and contextual reasoning require a relational representation beyond what tabular features can capture.

### F. Answers to Research Questions

**RQ1:** The project confirms that historical crash risk and real-time telemetry can be combined into a unified, interpretable Safety Index.

**RQ2:** Speed variance, VRU exposure, and safety events exert the most influence on short-term risk; the index is stable under moderate parameter variation.

**RQ3:** The RT-SI and MCDM indices show consistent temporal patterns, with useful complementary sensitivities.

**RQ4:** The project demonstrates that flat-feature limitations constrain causal inference and semantic clarity, motivating a move toward an agentic knowledge-graph framework.

## IX. LIMITATIONS AND FUTURE WORK

The Virginia Transportation Safety Index (VTTSI) demonstrates the feasibility of combining historical crash trends with real-time connected-vehicle telemetry, but several data, modeling, and system limitations influence accuracy, scalability, and interpretability. These limitations also motivate a clear roadmap for future work.

### A. Data and Sensing Limitations

The current deployment depends heavily on BSM/PSM penetration and the coverage of the instrumented intersections. PSM adoption remains low, underrepresenting VRUs and reducing the strength of VRU-related signals. Safety-event detections can be affected by camera occlusion, blind spots, and classifier errors, and intermittent telemetry produces flat or low-confidence RT-SI and MCDM values. Crash geolocation from VDOT introduces spatial ambiguity near intersection boundaries.

**Future Direction: Multi-Source Plug-In Architecture.** A unified data-ingestion layer will allow intersections to incorporate additional modalities—video analytics, LiDAR/CV2X perception, weather feeds, road-surface conditions, land-use context, and crowdsourced hazard reports. This architecture will let richer sites contribute more features while still supporting minimal deployments.

## B. Limitations of Flat Numeric Representations

VTTSI currently models intersections using independent, tabular features (volumes, speed variance, incidents). This loses relational structure such as lane geometry, conflict topology, multimodal trajectories, and temporal motifs. As a result, causal explanations (e.g., geometric constraints, recurrent conflict patterns, land-use influences) cannot be expressed. CRITIC weights also shift based on the last 24 hours, complicating cross-day comparisons, and at crash-free sites the EB baseline collapses toward zero, limiting interpretability.

**Future Direction: Traffic Safety Knowledge Graph (TS-KG).** A knowledge-graph representation [8] will encode intersections, approaches, lanes, VRUs, vehicles, trajectories, conflicts, and environmental context as relational entities. Graph-derived features (motifs, centrality, path risk) will enhance semantic richness and reduce dependence on flat features.

## C. Limits of Numeric-Only Reasoning

The present model produces numerical scores without semantic explanation or uncertainty reasoning. It cannot distinguish whether uplift arises from behavior, geometry, weather, or sensor noise, nor can it trace a chain of evidence for operators.

**Future Direction: Multi-Agent Reasoning Layer.** Agents will attach micro-events to the TS-KG, identify recurrent patterns, infer cause–effect structures, and generate human-readable safety narratives. This will enable explainable lift factors and risk attribution.

## D. Operational and Deployment Constraints

The current implementation uses a fixed blend parameter  $\alpha$ , does not offer context-aware tuning, and has been tested at only a small number of intersections. Scalability, latency, and reliability under heavy load still require evaluation. VTTSI is not yet integrated with navigation systems, CAV planners, emergency-response workflows, or DOT decision-support tools.

**Future Direction: Ecosystem Integration.** Future versions will expose open APIs and streaming interfaces to support: navigation routing, CAV behavior planning, emergency-response prioritization, and agency network-screening workflows. Automated  $\alpha$  selection (rule-based, optimization, or learning) will tailor the index to different operational needs.

## E. Equity and Coverage Gaps

Sensor coverage is uneven across regions, and intersections in underserved communities may have lower penetration and poorer telemetry, potentially biasing the index.

**Future Direction: Equity-Aware Safety Indexing.** Incorporating demographic, land-use, accessibility, and sensor-coverage indicators will support fairness-aware calibration and highlight locations where data scarcity masks risk.

## F. Toward a Unified, Agentic Safety Architecture

Collectively, these limitations point toward an evolution from a purely numeric model to a relational, agentic, and predictive architecture. Future VTTSI versions will blend numeric RT-SI and MCDM signals with semantic structures from the TS-KG and reasoning produced by multi-agent systems, enabling richer explanations, better forecasting, and more actionable safety intelligence.

## X. CONCLUSION

This project developed the Virginia Tech Transportation Safety Index (VTTSI), a cloud-native, real-time system that integrates severity-weighted Empirical Bayes crash stabilization with high-frequency connected-vehicle telemetry and a CRITIC-weighted hybrid MCDM module. Together, these components generate an interpretable 0–100 safety score every 15 minutes, enabling responsive, data-driven assessment of intersection conditions.

Through the research questions articulated in this work, we demonstrated that historical crash history and real-time telemetry can be combined into a unified and exposure-adjusted Safety Index; that speed variance, VRU exposure, and safety-event activity form the strongest short-term operational signals; and that the hybrid RT-SI and MCDM indices provide complementary perspectives on intersection safety. Validation across multiple intersections confirmed temporal coherence, robustness to parameter uncertainty, and appropriate sensitivity to operational turbulence and near-miss activity.

The VTTSI platform provides a practical foundation for continuous safety monitoring and operational decision support. Its cloud-native architecture and modular service design make the framework extensible to a statewide deployment. Looking forward, the Limitations and Future Work section outlines the evolution from flat numerical scoring toward a relational, agentic, and predictive safety-intelligence architecture.

## ACKNOWLEDGMENT

This research paper used AI for assistance in generating content.

## REFERENCES

- [1] Ahmed Almutairi et al. “Deep Learning Based Predictive Models for Real-Time Accident Prevention in Autonomous Vehicle Networks”. In: *Scientific Reports* 15.1 (2025), p. 20844. doi: 10.1038/s41598-025-04867-8. URL: <https://doi.org/10.1038/s41598-025-04867-8>.
- [2] A. Amraji. “Combined Safety Index Using CRITIC Weighting”. In: *Journal of Transportation Research* 15.2 (2025), pp. 200–210.
- [3] Mohammad Ali Arman, Chris M. J. Tampère, et al. “Empirical Study of Lane-Changing Maneuvers in a Weaving Area Based on Reconstructed Trajectories of Floating Car Data”. In: *Transportation Research Record* (2023). doi: 10.1177/03611981231179474.

- [4] Morteza AsadAmraji, Azarakhsh Salem, and Shila Shirinbayan. "A Combined Index of Proactive and Reactive Data for Rating the Safety of Road Sections". In: *Iranian Journal of Science and Technology – Transactions of Civil Engineering* 49.1 (2025), pp. 985–995. doi: 10.1007/s40996-024-01552-0.
- [5] Ethan Bancroft. *Seattle Keeps Emergency Scenes Clear by Alerting Smart Vehicles of Incidents in Near Real-Time*. Seattle DOT Blog. May 28, 2025. 2025. URL: <https://sdotblog.seattle.gov/2025/05/28/connected-vehicles-technology-pilot-program/>.
- [6] Noushin Behboudi, Sobhan Moosavi, and Rajiv Ramnath. "Recent Advances in Traffic Accident Analysis and Prediction: A Comprehensive Review of Machine Learning Techniques". In: *arXiv preprint arXiv:2406.13968* (2024).
- [7] Salvatore Cafiso, Grazia La Cava, and Alfonso Montella. "Safety Index for Evaluation of Two-Lane Rural Highways". In: *Transportation Research Record* 2019.1 (2007), pp. 136–145. doi: 10.3141/2019-17. URL: <https://doi.org/10.3141/2019-17>.
- [8] Jason Cusati. *Agentic Knowledge Graphs for Traffic Safety*. Tech. rep. Unpublished manuscript. Virginia Tech, 2025.
- [9] Federal Highway Administration. "Incorporating Data-Driven Safety Analysis in Traffic Impact Analyses: A How-To Guide". In: *FHWA-SA-19-026, US Department of Transportation* (2019). URL: <https://safety.fhwa.dot.gov/rsdp/downloads/fhwasa19026.pdf>.
- [10] Yang Gao and David Levinson. "Lane Changing and Congestion Are Mutually Reinforcing?" In: *Communications in Transportation Research* 3 (2023), p. 100101. doi: 10.1016/j.commtr.2023.100101.
- [11] Douglas Gettman and Larry Head. *Surrogate Safety Assessment Model and Validation: Final Report*. Tech. rep. FHWA, 2003.
- [12] Yifan Gong et al. "Multi-Objective Reinforcement Learning Approach for Improving Safety at Intersections with Adaptive Traffic Signal Control". In: *Accident Analysis & Prevention* 144 (2020), p. 105655. doi: 10.1016/j.aap.2020.105655.
- [13] Camilo Gutierrez-Osorio and Cesar A. Pedraza. "Modern Data Sources and Techniques for Analysis and Forecast of Road Accidents: A Review". In: *Journal of Traffic and Transportation Engineering (English Edition)* 7.4 (2020), pp. 432–446. doi: 10.1016/j.jtte.2020.05.002.
- [14] Moinul Hossain et al. "Real-Time Crash Prediction Models: State-of-the-Art, Design Pathways and Ubiquitous Requirements". In: *Accident Analysis & Prevention* 124 (2019), pp. 66–84. doi: 10.1016/j.aap.2018.12.022.
- [15] Qichao Hou et al. "A Deep Transfer Learning Approach for Real-Time Traffic Conflict Prediction with Trajectory Data". In: *Accident Analysis & Prevention* 214 (2025), p. 107966. doi: 10.1016/j.aap.2025.107966.
- [16] Tingting Huang, Shuo Wang, and Anuj Sharma. "Highway Crash Detection and Risk Estimation Using Deep Learning". In: *Accident Analysis & Prevention* 135 (2020), p. 105392. doi: 10.1016/j.aap.2019.105392.
- [17] Mehdi Keshavarz-Ghorabae et al. "A New Combinative Distance-Based Assessment (CODAS) Method for Multi-Criteria Decision-Making". In: *Economic Computation and Economic Cybernetics Studies and Research* 50 (2016), pp. 25–44.
- [18] Mehdi Keshavarz-Ghorabae et al. "Multi-Criteria Inventory Classification Using a New Method of Evaluation Based on Distance from Average Solution (EDAS)". In: *Informatica* 26 (2015), pp. 435–451. doi: 10.15388/Informatica.2015.57.
- [19] Masoud Khanmohamadi and Marco Guerrieri. "Smart Intersections and Connected Autonomous Vehicles for Sustainable Smart Cities: A Brief Review". In: *Sustainability* 17.7 (2025), p. 3254. doi: 10.3390/su17073254.
- [20] Jonathan W. Lee et al. "Traffic Control via Connected and Automated Vehicles: An Open-Road Field Experiment with 100 CAVs". In: *arXiv e-prints*, arXiv:2402.17043 (2024). doi: 10.48550/arXiv.2402.17043. URL: <https://arxiv.org/abs/2402.17043>.
- [21] Maricopa Association of Governments (MAG). *Network Screening Methodology for Intersections*. Tech. rep. Accessed: 2025-09-18. Phoenix, AZ: Maricopa Association of Governments, 2010. URL: [https://azmag.gov/Portals/0/Documents/TSC\\_2010-07-20\\_MAG-Network- Screening - Methodology - for - Intersections - May-2010\\_.pdf](https://azmag.gov/Portals/0/Documents/TSC_2010-07-20_MAG-Network- Screening - Methodology - for - Intersections - May-2010_.pdf).
- [22] Alfonso Montella et al. "Systemic Approach to Improve Safety of Urban Unsignalized Intersections: Development and Validation of a Safety Index". In: *Accident Analysis & Prevention* 141 (2020), p. 105523. doi: 10.1016/j.aap.2020.105523.
- [23] Iqbal Mutambik et al. "IoT-Enabled Adaptive Traffic Management: A Multiagent Framework for Urban Mobility Optimisation". In: *Sensors* 25.13 (2025), p. 4126. doi: 10.3390/s25134126.
- [24] Minsoo Oh, John Shaw, and Jing Dong-O'Brien. "Enhancing Work Zone Safety: A Psychological and Behavioral Analysis of Anti-Tailgating Strategies Using Dynamic Message Signs and Fixed Warning Signs". In: *Transportation Research Part F: Traffic Psychology and Behaviour* 106 (2024), pp. 150–168. doi: 10.1016/j.trf.2024.08.007.
- [25] Bhagwant Persaud and Craig Lyon. "Empirical Bayes Before–After Safety Studies: Lessons Learned from Two Decades of Experience and Future Directions". In: *Accident Analysis & Prevention* 39.3 (2007), pp. 546–555. doi: 10.1016/j.aap.2006.09.009.
- [26] A. F. D. Pribadi, A. D. Oktaviani, and A. S. Utomo. "Assessing the Safety Effect Through Google Maps Usage: FMEA Approach (Case Study: Indonesia)". In: *Case Studies on Transport Policy* 10.1 (2022), pp. 42–53.

- [27] Grant Schultz et al. "Prioritizing Safety Funding Using Severity Weighted Risk Scores". In: *Accident Analysis & Prevention* 218 (2025), p. 108083. doi: 10.1016/j.aap.2025.108083.
- [28] John Schultz. "A Study on Surrogate Safety Indicators". In: *Journal of Transportation Safety*. Vol. 12. 3. 2025, pp. 123–134.
- [29] Asaya Shimojo et al. "How Impressions of Other Drivers Affect One's Behavior When Merging Lanes". In: *Transportation Research Part F: Traffic Psychology and Behaviour* 89 (2022), pp. 236–248.
- [30] R. P. Singh, P. Jain, and S. Yadav. "SPaFE: A Crowd-sourcing and Multimodal Recommender System to Ensure Travel Safety in a City". In: *IEEE Access* 10 (2022), pp. 71256–71270.
- [31] Virginia Department of Transportation. *Crash Records Database*. Accessed 2024-06-01. 2023.
- [32] Shian Wang et al. "A General Approach to Smoothing Nonlinear Mixed Traffic via Control of Autonomous Vehicles". In: *Transportation Research Part C: Emerging Technologies* 146 (2023), p. 103967.
- [33] Y. Wang et al. "A High-Resolution Trajectory Data Driven Method for Real-Time Evaluation of Traffic Safety". In: *Accident Analysis & Prevention* 160 (2021), p. 106323.
- [34] S. H. Wicaksono, A. Nugroho, and R. R. R. Putra. "Evaluating the Use of Google Maps as Navigation Application by Identifying Hazards and Assessing Risks Using HIRA Matrix". In: 1064.1 (2022), p. 012020.
- [35] Dan Wu et al. "A Framework for Real-Time Traffic Risk Prediction Incorporating Cost-Sensitive Learning and Dynamic Thresholds". In: *Accident Analysis & Prevention* 218 (2025), p. 108087.
- [36] Minghui Wu et al. "Participatory Traffic Control: Leveraging Connected and Automated Vehicles to Enhance Network Efficiency". In: *Transportation Research Part C: Emerging Technologies* 166 (2024), p. 104757.
- [37] Ning Xie and Hao Wang. "Distributed Adaptive Traffic Signal Control Based on Shockwave Theory". In: *Transportation Research Part C: Emerging Technologies* 173 (2025), p. 105052.
- [38] Pengcheng Zhang et al. "Dynamic Spatiotemporal Graph Network for Traffic Accident Risk Prediction". In: *GIScience & Remote Sensing* 62.1 (2025), p. 2514330. doi: 10.1080/15481603.2025.2514330. URL: <https://doi.org/10.1080/15481603.2025.2514330>.
- [39] Haitao Zhao et al. "Vehicle Accident Risk Prediction Based on AdaBoost-SO in VANETs". In: *IEEE Access* 7 (2019), pp. 14549–14557. doi: 10.1109/ACCESS.2019.2894176.

## APPENDIX A TEAM MEMBER CONTRIBUTION REPORT

This section summarizes each team member's contributions to the final project, consistent with the task assignments outlined in Appendix A. Each member played a distinct role in the development of the Safety Index (SI) system, contributing across areas such as data engineering, interface design, algorithmic modeling, and exploratory research.

### A. Cheng-Shun

Cheng-Shun has led the core software engineering, system integration, and data processing efforts for the project. He developed the complete backend API using FastAPI and implemented the interactive frontend visualization interface using Streamlit. He also designed and deployed the full data-delivery pipeline connecting raw data sources and database, the backend services, and the frontend dashboards and pages except analytics and validation page.

A major component of his contribution is the end-to-end integration of heterogeneous data streams, BSM/PSM telemetry, vehicle and VRU counts, speed distributions, incident detections, and database resident historical crash records—into a unified visualization platform. Cheng-Shun implemented dynamic visualization modules that display temporal trends, statistical summaries, normalized variable comparisons, and permutation-style overlays of safety-relevant signals to support rapid diagnostic analysis of intersections. That means most of the system architecture shown in this report is built by his own.

He implemented the full mathematical formulation of the Safety Index proposed in methodology, including the Empirical Bayes baseline, uplift factors, and the CRITIC-weighted multi-criteria decision-making (MCDM) framework (SAW, EDAS, CODAS) used to compute real-time risk scores. He further integrated these algorithms directly into the backend, enabling automated computation of RT-SI, MCDM, and blended indices for every 15-minute interval.

In addition, Cheng-Shun conducted the complete case-study evaluation and validation and sensitivity analysis for the intersections, generating the full suite of trend plots, correlation matrices, normalized comparisons, and interpretive assessments. This analysis demonstrated that the Safety Index behaves coherently with respect to exposure patterns, near-miss activity, speed variance, and operational anomalies.

### B. Jason

Jason contributed to the architectural design, analytical foundations, and conceptual evolution of the Safety Index framework. He extended the system architecture by designing and implementing the *plug-in data ingestion framework*, enabling new data sources—such as traffic camera imagery and weather feeds—to be incorporated seamlessly into the real-time processing pipeline. This flexible ingestion layer forms the basis for future multi-source expansion as described in the Limitations and Future Work.

Jason performed exploratory data mining on the VTTI data portal and established the methodology used to preprocess and

structure the historical crash and real-time telemetry datasets. He refined the initial mathematical structure of the Safety Index, identifying which real-time and historical features were most meaningful based on evidence from prior literature. He also attempted an early implementation of the MCDM and CRITIC algorithms in the backend; while this implementation was later superseded, it informed the final structure of the hybrid scoring pipeline.

In addition to analytic contributions, Jason authored the project's **research questions**, the **limitations analysis**, and the **future-work direction**, including the conceptual design of the Agentic Traffic Safety Knowledge Graph (TS-KG). This agentic KG vision provides a roadmap for extending the system beyond flat numerical features toward relational reasoning, multimodal context, and explainable safety intelligence.

Overall, Jason's contributions shaped both the theoretical underpinnings and the forward-looking architecture of the VTTI system, helping define the long-term trajectory of the project.

### C. Victory

Victory has contributed through literature synthesis and conceptual support. She has provided insight into the research framing of the project and is investigating potential **external variables**—such as environmental, temporal, or behavioral factors—that could enhance the predictive capability of the Safety Index model. Her upcoming tasks involve curating these additional factors and testing their correlation strength against the core VTTI datasets.

### D. Contribution Wrap

Together, the team has established a solid foundation for the project, with all core system components under active development. Backend services and data pipelines are operational, the analytical framework is being refined, and research integration continues to guide the model's evolution toward a fully functional and interpretable Safety Index system.

**APPENDIX B**  
**BACKEND SERVICE DESCRIPTIONS**

TABLE II  
CORE BACKEND SERVICES

| Service                      | Description   |
|------------------------------|---|
| MCDM Safety Index Service    | Computes SAW, EDAS, CODAS, and hybrid MCDM scores.  |
| RT-SI Safety Index Service   | Computes Empirical Bayes-stabilized and uplift-adjusted real-time safety index.             |
| Telemetry Aggregation Module | Aggregates BSM, PSM, count tables, and speed-distribution data into aligned 15-minute bins. |
| Safety Event Processor       | Processes safety-event detections and attaches them to analysis windows.                    |

**APPENDIX C**  
**ADDITIONAL METHODOLOGY DETAILS**

This appendix expands on the formulations referenced in Section IV, including parameter tables and the full weight-determination framework used in the real-time and MCDM safety indices.

#### A. Parameter Summary

| Parameter                                  | Meaning   |
|--|---|
| $w_s$                                      | Severity weights (e.g., Fatal=10, Injury=3, PDO=1)                    |
| $\lambda$                                  | EB shrinkage strength   |
| $k_1..k_5$                                 | Scaling constants for speed drop, variance, conflicts, VRU/veh ratios |
| $\beta_1..\beta_3$                         | Weights for real-time uplift factors                                  |
| $\omega_{\text{vru}}, \omega_{\text{veh}}$ | Weights for VRU vs Vehicle sub-indices                                |
| $W_E, W_C, W_S$                            | Method weights for EDAS, CODAS, SAW in Hybrid index                   |
| $\alpha$                                   | Blend factor for Final index (RT vs. MCDM)                            |

TABLE III  
KEY TUNABLE PARAMETERS OF THE SAFETY INDEX FORMULAS.

#### B. Weight Determination Framework

This section documents the exact weighting scheme implemented in the real-time and MCDM components of the Safety Index. All formulas in this section match the codebase, including the Empirical Bayes estimator, uplift factors, and CRITIC-based objective weighting used in the MCDM module.

1) *Severity Weights  $w_s$ :* Severity weighting is applied only in the historical crash-rate computation. We adopt fixed policy-driven weights:

$$w_{\text{fatal}} = 10, \quad w_{\text{injury}} = 3, \quad w_{\text{PDO}} = 1,$$

which match the implementation used to build severity-weighted crash counts.

2) *Empirical Bayes Shrinkage Parameter  $\lambda$ :* The historical crash rate for a 15-minute bin is stabilized using a no-exposure Empirical Bayes (EB) estimator:

$$\hat{r}_i = \frac{Y_i + \lambda r_0}{1 + \lambda},$$

where  $Y_i$  is the severity-weighted crash count in bin  $i$ ,  $r_0$  is the pooled mean over 2017–2024, and  $\lambda$  is selected using temporal cross-validation. A grid search over  $\lambda \in \{0.1, \dots, 100000\}$  minimizes Poisson negative log-likelihood when predicting 2025 outcomes. The resulting parameters are:

$$\lambda^* = 100000, \quad r_0 = 3.365.$$

3) *Real-Time Uplift Factors ( $F_{\text{speed}}, F_{\text{var}}, F_{\text{conf}}$ )*: The real-time Safety Index multiplies the EB baseline by three bounded uplift factors derived from speed, turbulence, and VRU-vehicle conflict exposure.

a) *Speed deficit uplift*.: Let FFS denote the free-flow speed and  $v$  the observed mean speed:

$$F_{\text{speed}} = \min\left(1, 1.5 \frac{\text{FFS} - v}{\text{FFS}}\right).$$

b) *Speed variance uplift*.: Let  $\sigma^2$  be the variance of speed in the current bin:

$$F_{\text{var}} = \min\left(1, 1.0 \frac{\sqrt{\sigma^2}}{v}\right).$$

c) *VRU conflict uplift*.: Let  $T$  be turning volume and  $V_{\text{vru}}$  be VRU count:

$$F_{\text{conf}} = \min\left(1, 0.5 \frac{T V_{\text{vru}}}{1000}\right).$$

d) *Combined real-time multiplier*.: The code uses fixed coefficients  $(\beta_1, \beta_2, \beta_3) = (0.3, 0.3, 0.4)$ :

$$U = 1 + 0.3F_{\text{speed}} + 0.3F_{\text{var}} + 0.4F_{\text{conf}}.$$

4) *VRU-Vehicle Index Blend  $\omega$* : The real-time index blends VRU-specific and vehicle-specific sub-indices using fixed weights:

$$\text{SI}_{\text{RT}} = 0.6 \text{SI}_{\text{VRU}} + 0.4 \text{SI}_{\text{VEH}}.$$

These values are policy parameters, not learned.

5) *MCDM Criterion Weights  $w_j$  (CRITIC Only)*: For the MCDM component, each 15-minute intersection-time pair is treated as an alternative, and the following five real-time criteria form the decision matrix:

$$\mathcal{C} = \{\text{vehicle\_count}, \text{vru\_count}, \text{avg\_speed}, \text{speed\_variance}, \text{incident\_count}\}. \quad (43)$$

After min–max normalization, CRITIC weighting is applied:

$$C_j = \sigma_j \sum_k (1 - \rho_{jk}), \quad w_j = \frac{C_j}{\sum_j C_j},$$

where  $\sigma_j$  is the standard deviation of criterion  $j$  and  $\rho_{jk}$  is the correlation between criteria  $j$  and  $k$  across all intersection-time alternatives in the last 24 hours. Entropy weighting is not used in the implementation.

### 6) Hybrid Method Weights $W_E, W_C, W_S$ (CRITIC Again):

After computing SAW, EDAS, and CODAS scores (all normalized to  $[0, 100]$ ), these three method outputs form a matrix

$$M = [\text{EDAS}_a, \text{CODAS}_a, \text{SAW}_a],$$

from which CRITIC weights are again derived:

$$W_j = \frac{\sigma_j \sum_k (1 - \rho_{jk})}{\sum_j \sigma_j \sum_k (1 - \rho_{jk})}, \quad j \in \{\text{EDAS}, \text{CODAS}, \text{SAW}\}.$$

Method weights are recalculated dynamically for each evaluation window.

The hybrid MCDM index is the weighted sum:

$$\text{SI}_{i,t}^{\text{MCDM}} = W_E \text{EDAS}_{i,t} + W_C \text{CODAS}_{i,t} + W_S \text{SAW}_{i,t}.$$

7) *Final Blend*: The current implementation exposes the real-time and MCDM indices separately. No blending parameter  $\alpha$  is applied in the released code. The final index range from 0 to 100 which indicate very safe to very dangerous.

## APPENDIX D ADDITIONAL VALIDATION FIGURES

This appendix contains supplemental figures referenced in Section VI, including additional intersections (Birch and Broad) to demonstrate consistency of temporal patterns, score behavior, and model coherence.

1) *Cross-Validation*: Only the Empirical Bayes shrinkage parameter  $\lambda$  uses temporal cross-validation:

- 1) Use 2017–2024 severity-weighted crash counts as training data.
- 2) Compute pooled mean  $r_0$ .
- 3) For each  $\lambda$  in a predefined grid:
  - a) Compute EB-stabilized rates for training years.
  - b) Predict 2025 counts.
  - c) Compute Poisson log-loss.
- 4) Select  $\lambda^*$  minimizing log-loss; hardcode into real-time service.

Weights for uplift factors, VRU–vehicle blending, and MCDM are not tuned via cross-validation.

### VALIDATION ADDITIONAL IMAGES FROM OTHER INTERSECTION

Results for Birch and Broad intersections exhibit the same behavior and are included in Appendix figures. Figure 12 and figure 13 are all variables trend comparisons for each intersection.

Figure 16 and figure 17 are MCDM method and its sub-indices comparison.

Figure 18 and figure 19 are RT-SI method and its sub-indices comparison.

Figure 14 and figure 15 are how RT-SI and MCDM blended into final index with alpha equal to 0.7.

These images show similar trends and results showed in validation analysis, thus justify further that the conclusion of validation analysis is consistent.

## APPENDIX E ADDITIONAL SENSITIVITY ANALYSIS RESULTS

This appendix provides additional plots supporting the sensitivity analysis presented in Section VII. These results confirm the robustness of the RT-SI under parameter perturbation across all intersections tested.

Additional information of this sections, including other intersections' results and default setting.

### A. Other intersections' results

Figure 20 and figure 13 are perturbation results of other intersections.

Figure 23 and figure 23 are distribution of deviations of perturbations in other intersections.

Figure 24 and figure 25 shows importances of different weights in other intersections.

These images show similar trends and results showed in sensitivity analysis, thus justify further that the conclusion of sensitivity analysis is consistent.

### B. Baseline Parameters

For reference, the baseline parameters used in the RT-SI model are:

- EB parameters:  $\lambda = 100,000$ ,  $R_0 = 3.365$
- Severity weights:  $W_{\text{fatal}} = 10$ ,  $W_{\text{injury}} = 3$ ,  $W_{\text{PDO}} = 1$
- Uplift scaling:  $K1_{\text{speed}} = 1.5$ ,  $K2_{\text{var}} = 1.0$ ,  $K3_{\text{conf}} = 0.5$ ,  $K4_{\text{VRU}} = 1.0$ ,  $K5_{\text{vol/cap}} = 1.0$
- Uplift combination weights:  $\beta_1 = 0.3$ ,  $\beta_2 = 0.3$ ,  $\beta_3 = 0.4$
- VRU vs vehicle weighting:  $\omega_{\text{VRU}} = 0.6$ ,  $\omega_{\text{VEH}} = 0.4$
- Base multiplier:  $\gamma = 1.0$
- Capacity assumption: 500 vehicles per 15-minute interval

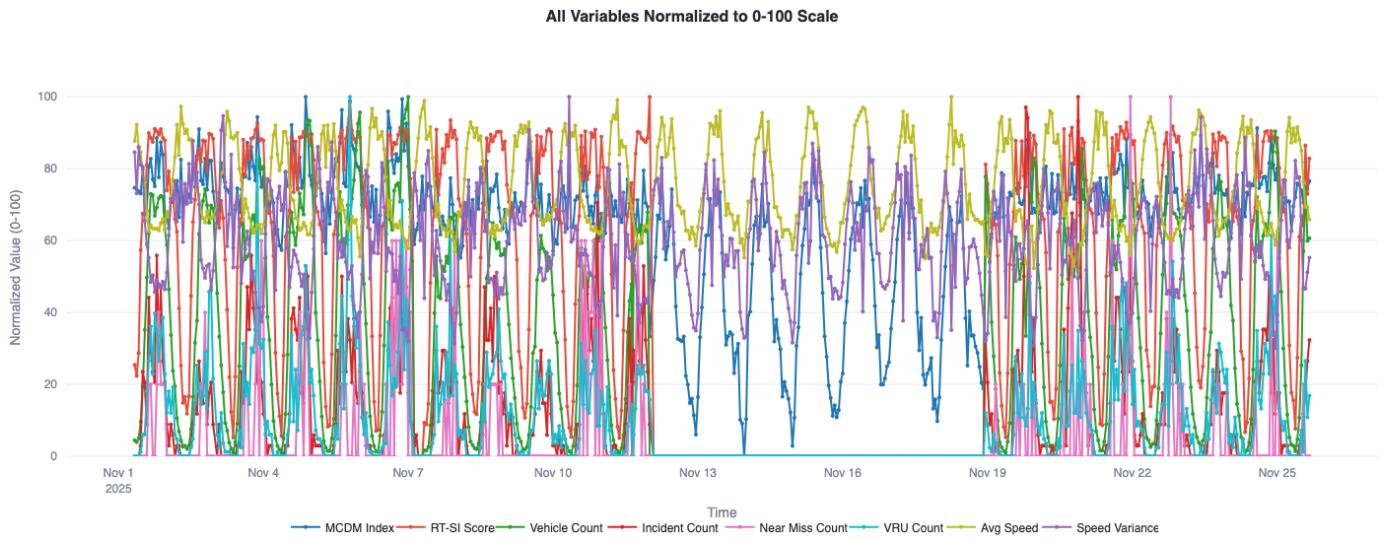


Fig. 12. All Variables Normalized to 0-100 scale for trend analysis on Birch\_Street\_W\_Broad\_Street intersection

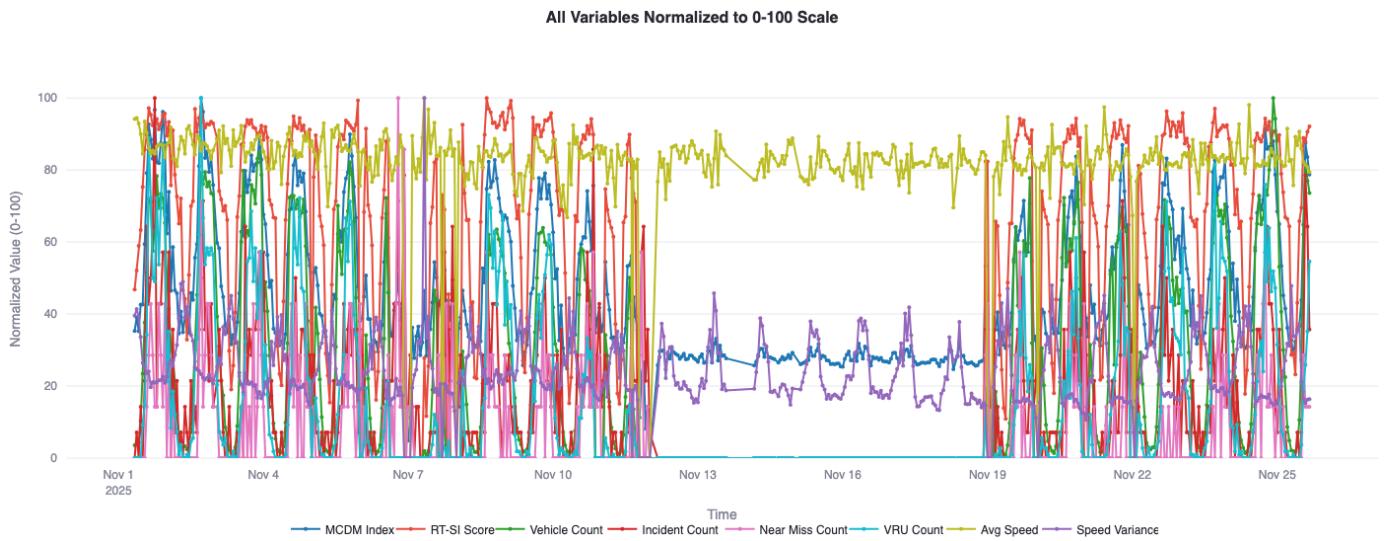


Fig. 13. All Variables Normalized to 0-100 scale for trend analysis on E\_Broad\_Street\_N\_Washington\_Street intersection



Fig. 14. Final Blended Safety Index on Birch\_Street\_W\_Broad\_Street intersection

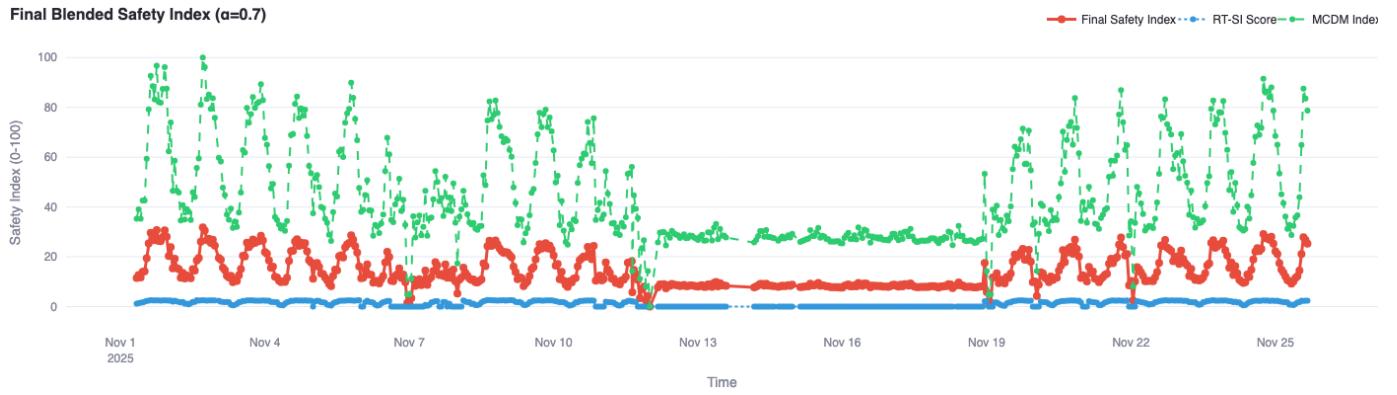


Fig. 15. Final Blended Safety Index on E\_Broad\_St N\_Washinton\_St intersection

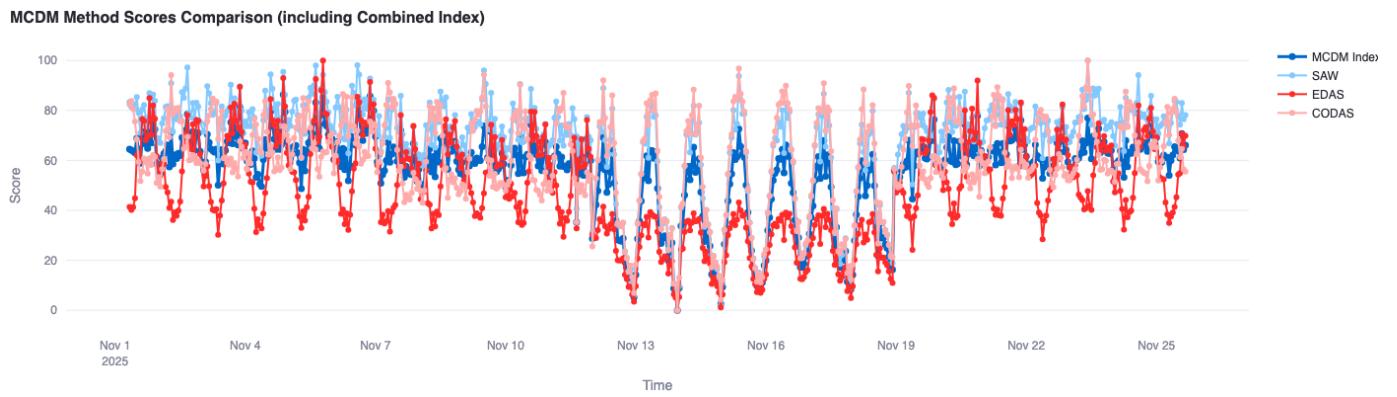


Fig. 16. MCDM methods trend comparison on Birch\_St W\_Broad\_St intersection

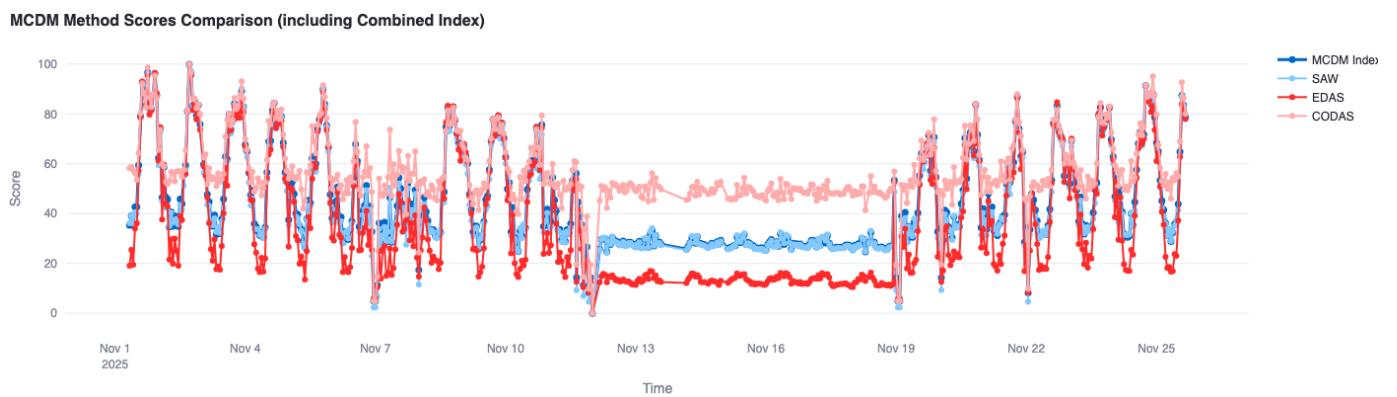


Fig. 17. MCDM methods trend comparison on E\_Broad\_St N\_Washinton\_St intersection

**RT-SI Score & Sub-Indices (VRU vs Vehicle Risk)**

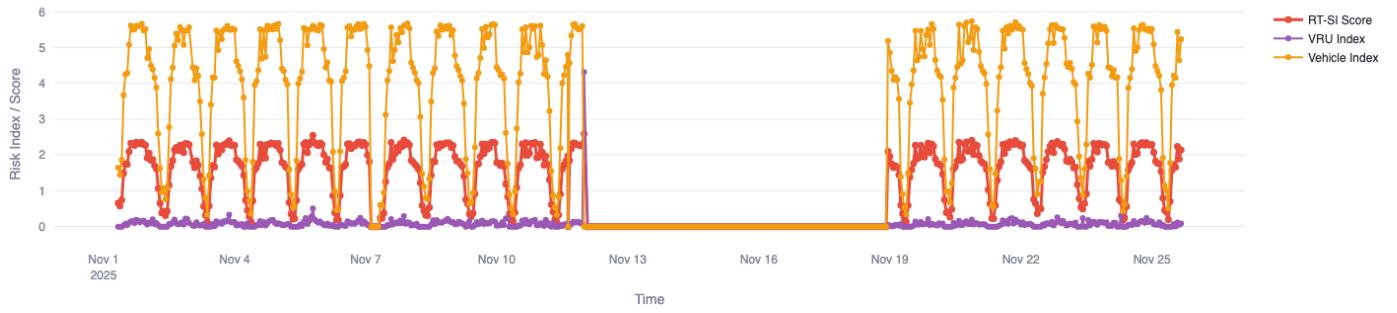


Fig. 18. RT-SI scoring methods trend comparison on Birch\_Street\_W\_Broad\_Street intersection

**RT-SI Score & Sub-Indices (VRU vs Vehicle Risk)**

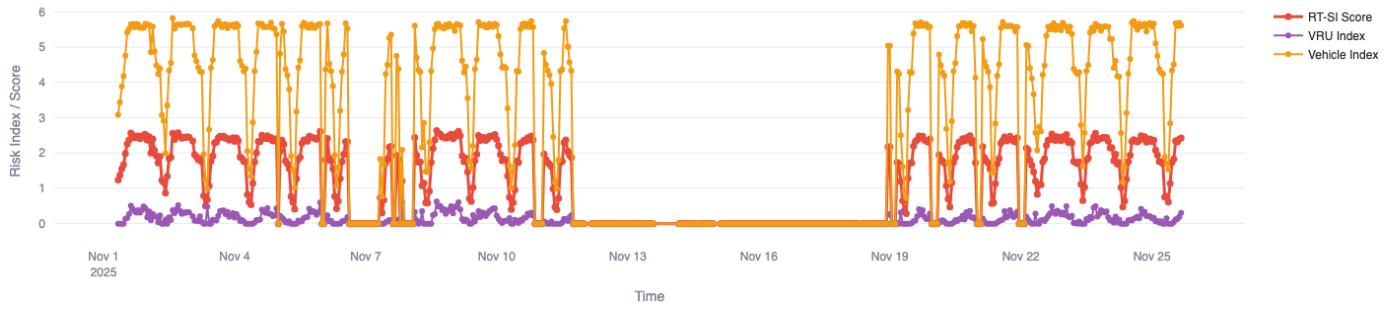


Fig. 19. RT-SI scoring methods trend comparison on E\_Broad\_Street\_N\_Washinton\_Street intersection

**Baseline vs Perturbed RT-SI Scores**

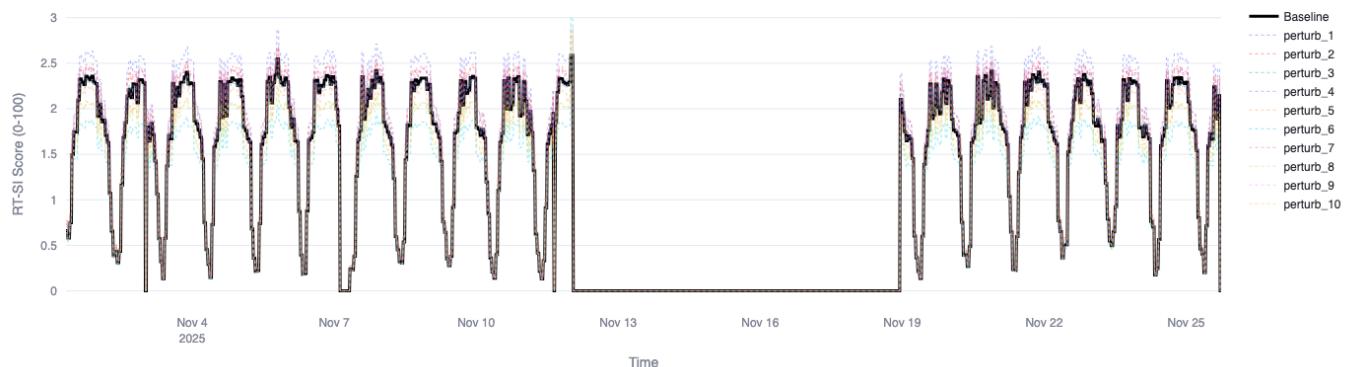


Fig. 20. Comparison of perturbations with baseline RT-SI scoring on Birch\_Street\_W\_Broad\_Street intersection

#### Baseline vs Perturbed RT-SI Scores

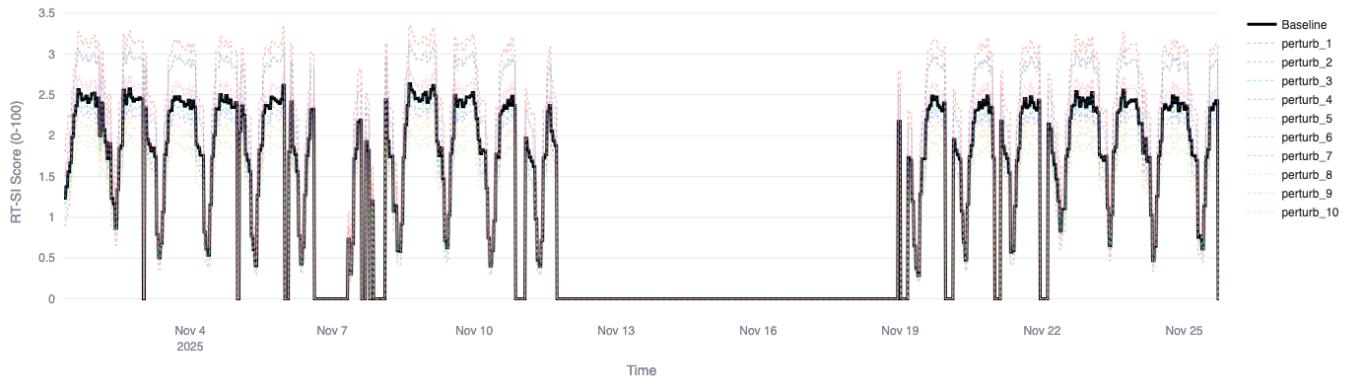


Fig. 21. Comparison of perturbations with baseline RT-SI scoring on E\_Broad\_St N\_Washinton\_St intersection

#### RT-SI Score Changes (Absolute Differences)

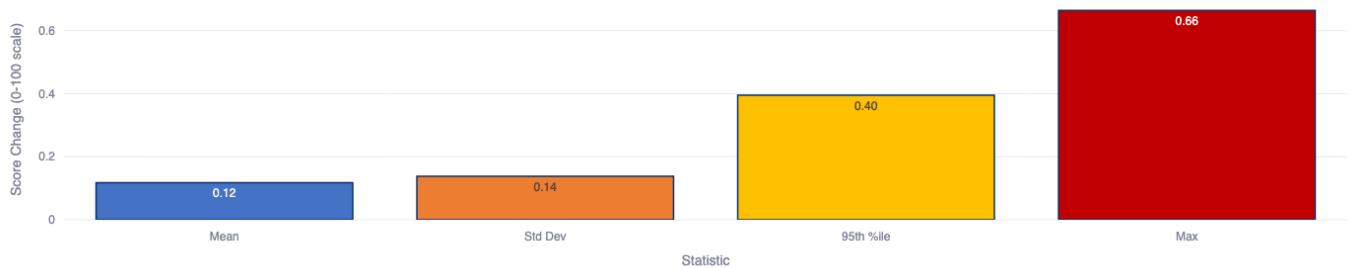


Fig. 22. Distribution of absolute deviations between baseline and perturbations on Birch\_St W\_Broad\_St intersection

#### RT-SI Score Changes (Absolute Differences)

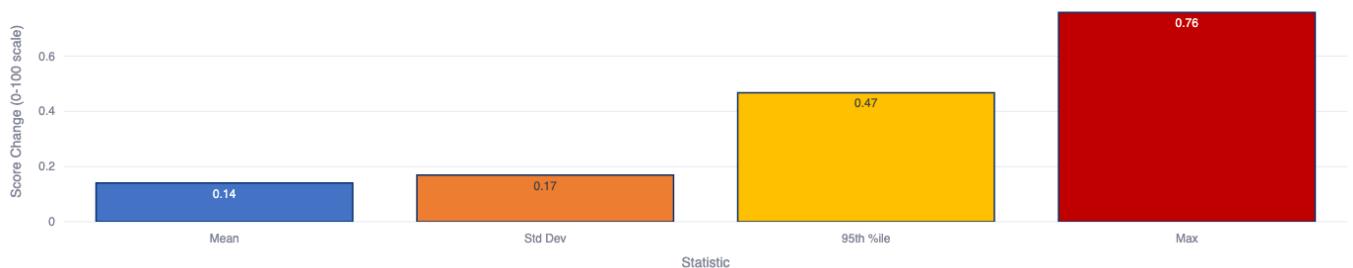


Fig. 23. Distribution of absolute deviations between baseline and perturbations on E\_Broad\_St N\_Washinton\_St intersection

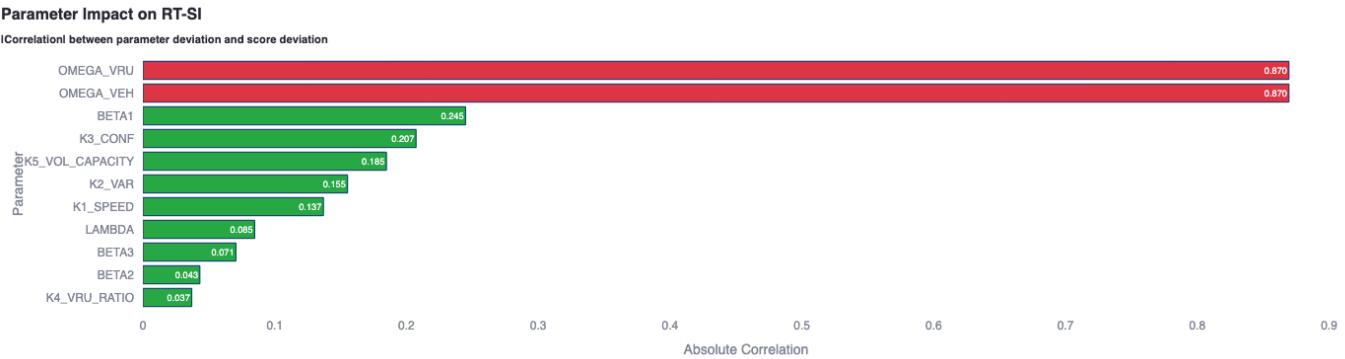


Fig. 24. Correlation between parameter perturbations and RT-SI score deviations on Birch\_Street\_W\_Broad\_Street intersection

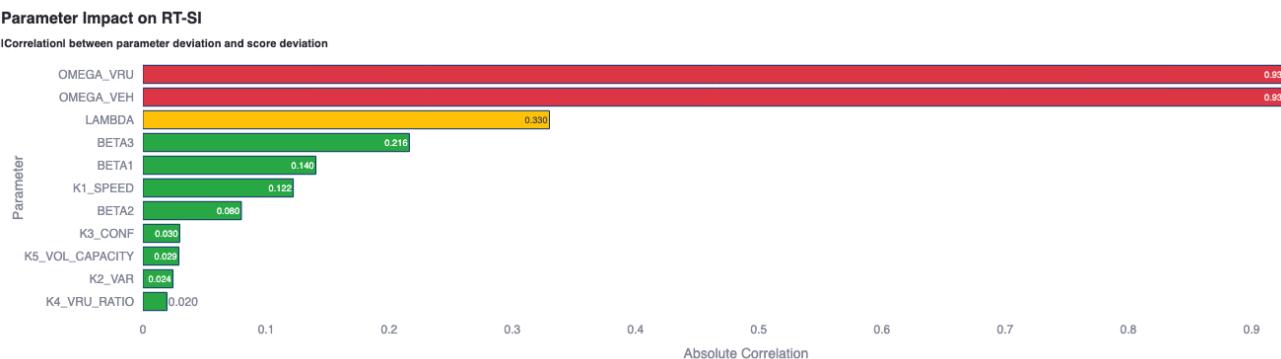


Fig. 25. Correlation between parameter perturbations and RT-SI score deviations on E\_Broad\_Street\_N\_Washinton\_Street intersection