Deliverable 5.1

Consolidated PHOEBE framework methodology, network-level demonstration results and tools

UK participants in Horizon Europe Project PHOEBE are supported by UKRI grant numbers 10038897 (The International Road Assessment Programme – iRAP) and 10056912 (The Floow)

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101076963

Document Control Page

Deliverable number	5.1
Deliverable title	Consolidated PHOEBE framework methodology, network-level demonstration results and tools
Deliverable version	1.0
Work Package number	WP5
Work Package Title	Systems Integration and Transferability
Due date of delivery	30/03/2026
Actual date of delivery	23/03/2026
Dissemination level	Public
Type	Report
Editor(s)	Apostolos Ziakopoulos & Stella Roussou, National Technical University of Athens
Contributor(s)	Apostolos Ziakopoulos, Stella Roussou, Maria Oikonomou, & George Yannis, National Technical University of Athens Shanna Lucchesi, EIRA Pietro Podestà, FACTUAL Consulting Sam Chapman and Mark Burke, The Floow Monica Olyslagers, iRAP Antonio Pellicer & Miguel Ángel Vazquez, Aimsun Arunava Putatunda & Santhanakrishnan Narayanan, TUM Amir Hossein Kalantari, Amna Chaudhry, Amir Pooyan Afghari, TUD
Reviewer(s)	Andréia Lopes Azevedo, Polis Network Pietro Podestà, Áron Hegedűs , FACTUAL Consulting
Project name	Predictive Approaches for Safer Urban Environments
Project Acronym	PHOEBE
Project starting date	01/11/2022
Project duration	45 months
Rights	PHOEBE consortium

PM efforts per beneficiary that contributed to the deliverable

#	Partner	PM effort in the Deliverable
1	EIRA	0.29
2	Floow	0.25
3	iRAP	0.10
4	TUD	0.75
5	NTUA	0.80
6	TUM	0.25
7	AIMSUN	0.28
Total	PHOEBE Consortium	2.72

Document History

Version	Date	Beneficiary	Description
0.1	10/12/2026		First draft of the deliverable structure
0.2	25/02/2026		Second draft of the deliverable structure
0.3	26/02/2026		Final text for the reviewers
0.4	11/03/2026		Reviewer contributions
0.5	13/03/2026		Reviewer contributions
0.6	20/03/2026		Final

Project Executive Summary

The EU-funded 'Predictive Approaches for Safer Urban Environment' (PHOEBE) project aims to develop an integrated, dynamic, human-centred predictive safety assessment framework in urban areas. This will be achieved by bringing together the interdisciplinary power of traffic simulation, road safety assessment, human behaviour, mode shift and induced demand modelling and new and emerging mobility data.

Focused on vulnerable road users' safety, the 3.5-year-long PHOEBE project will draw inspiration from real-world scenarios in the three pilot cities of Athens (GR), Valencia (ES) and the West Midlands (UK). Testing activities will be performed across the use cases to simulate and forecast the impact of changes on safety in different scenarios of disruptions or transitions across urban transport networks.

Predicting and visualising the safety and socioeconomic outcomes of new forms of transport, new technologies, or regulatory and behavioural changes from the individual (micro) level up to the network-wide (macro) level will also be a significant game-changer for urban stakeholders. The results of PHOEBE can be used as a blueprint by other European cities to develop their knowledge products, such as socioeconomic analysis models, urban road safety assessments, human behaviour and choice modelling.

PHOEBE Pilot Cities

List of participating cities:

- Athens (Greece)
- Valencia (Spain)
- West Midlands (United Kingdom)

Social Links:

https://twitter.com/Project_PHOEBE

https://www.linkedin.com/company/phoebe-project/

https://www.youtube.com/@phoebeproject

For further information please visit **WWW.PHOEBE-PROJECT.EU**

Project Partners

Organisation	Country	Abbreviation
EVROPSKI INSTITUT ZA OCENJEVANJE CEST – EURORAP	SI	EIRA
ETHNICON METSOVION POLYTECHNION	EL	NTUA
TECHNISCHE UNIVERSITEIT DELFT	NL	TUD
TECHNISCHE UNIVERSITAET MUENCHEN	DE	TUM
AIMSUN SLU	ES	AIM
POLIS AISBL	BE	POLIS
FACTUAL CONSULTING SL	ES	FC
UNIVERSITAT POLITECNICA DE VALENCIA	ES	UPV
OSEVEN SINGLE MEMBER PRIVATE COMPANY	EL	O7
THE FLOOW LIMITED	UK	FLOOW
INTERNATIONAL ROAD ASSESSMENT PROGRAMME	UK	iRAP

List of abbreviations and acronyms

Acronyms	Meaning
AADT	Annual average daily traffic
Av.	Avenue
AVs	Autonomous Vehicles
CapEX	**Capital Expenditure**
CBA	Cost-Benefit Analysis
CEA	**Cost-Effectiveness Analysis**
CMF	Crash Modification Factor
DW	**Disability Weight**
EGB	Edgbaston
EMDS	**European Mobility Data Space**
EU	Europe Union
FSIs	**Fatalities and serious injuries**
GLM	Generalised Linear Models
IHD/CHD	**Ischaemic/Coronary Heart Diseases**
IMF	Impact Modification Factor
IS	**Ischaemic Stroke**
ISO	International Organization for Standardization
ITHIM	**Integrated Transport and Health Impact Modelling**
ITU	International Telecommunication Union
KPI	**Key Performance Indicators**
KRN	Key Route Network
LTN	**Low-Traffic Neighbourhoods**
LTP	Local Transport Plan
MET	**Metabolic Equivalent of Task**
ML	Machine Learning
OpEx	**Operational Expenditure**

Acronyms	Meaning
PREDIF	State Representative Platform of People with Physical Disability
QALYs	**Quality Adjusted Life Years**
Sq.	Square
St.	**Street**
SUMP	Sustainable Urban Mobility Plan
TfWM	**Transport for the West Midlands**
T2D	Type 2 Diabetes Mellitus
UK	**United Kingdom**
UN	United Nations
UNESCO	**United Nations Educational, Scientific and Cultural Organization**
VoSL	Value Of Statistical Life
VOT	**Value Of Time**
VPF	Value Of Preventing A Fatality
VRU	**Vulnerable Road User**
WMCA	West Midlands Combined Authority region
WP	**Work Package**
WTP	Willingness To Pay
YOLO	**You Only Look Once Algorithm**

Table of Contents

List of figures

List of tables

Executive Summary

This deliverable report (D5.1) is a product of the PHOEBE project, detailing the progress and outcomes of Work Package 5 (WP5) - Systems Integration and Transferability. The document consolidates and synthesises the knowledge obtained in the previous work packages of PHOEBE and builds on the theoretical principles of the methodological framework to produce a holistic, integrated predictive safety assessment tool for urban environments.

WP5 serves as the bridge between the empirical findings of the use case trials (WP4) and the operational implementation of the PHOEBE framework, focusing on multi-scale integration, data harmonisation, socio-economic assessment, and user-oriented interface design. The activities within WP5 ensure that the PHOEBE framework can be effectively applied and transferred across diverse European urban contexts.

The main objectives of WP5 are to:

1. Develop a holistic algorithm that addresses integrated risk assessment across micro- and macro-scales, enabling the transfer of safety estimations from specific road segments to network-level analysis.
2. Quantify and predict the impacts of modelled interventions on road safety, mode shift, induced demand, and transport efficiency.
3. Integrate multi-source data, including infrastructure and telematics information, for rapid, data-driven road risk estimation across various locations and user categories, with emphasis on vulnerable road users (VRUs).
4. Establish data harmonisation and protection procedures, ensuring secure and interoperable use of PHOEBE data.
5. Design and implement a user-friendly interface that supports decision-making and enables the transferability and scalability of the PHOEBE framework beyond the project's scope.

D5.1 is structured into five interrelated tasks that collectively ensure the consolidation, integration, and transferability of the PHOEBE framework.

- T5.1 - Synthesis of Use Case Results (Lead: EIRA; Partners: iRAP, NTUA, TUM, TUD, AIM, FLOOW, FC, O7; Duration: M30-M40)

This task focuses on the synthesis and generalisation of findings obtained from the PHOEBE use cases (Athens, Valencia, West Midlands). It aims to consolidate the results into a unified framework that captures transferable insights across different contexts. The task will provide protocols for handling missing data categories, assess transferability, and develop guidelines and best practices for conducting comparable studies at the European level. The outcomes of this task will feed directly into the refinement and consolidation of the PHOEBE framework's theoretical and methodological principles.

- T5.2 - Overall Socio-Economic Impact Assessments (Lead: TUM; Partners: NTUA, iRAP, TUD, FC; Duration: M30-M40)

This task quantifies road safety impacts and evaluates the socio-economic effects of interventions using use case data. It will analyse pre- and post-intervention scenarios to estimate costs, benefits, and safety improvements using econometric and machine learning models. Results will be extrapolated to the network level, providing policy-relevant insights for authorities and feeding into PHOEBE's dynamic safety and cost-benefit framework.

- T5.3 - Dynamic Safety Prediction Tools (Lead: iRAP; Partners: NTUA, TUM, TUD, AIM, FLOOW; Duration: M25-M40)

This task develops dynamic road safety assessment models and software tools capable of processing real-time and simulated traffic information (e.g., speed, flow, user interactions). The developed functionality will be integrated into iRAP's ViDA platform and related APIs to enhance the scalability and repeatability of the PHOEBE framework. The task will enable the prediction of safety outcomes and allow the simulation of intervention effects under varying traffic and network conditions.

- T5.4 - Optimisation of Data Flow (Lead: The Floow; Partners: iRAP, AIM, POLIS, NTUA, TUM, FC, O7; Duration: M31-M40)

This task defines and optimises the way data flows throughout the PHOEBE framework. It harmonises data from multiple sources - including infrastructure, simulation, and telematics datasets - and ensures their efficient and secure integration. Emphasis will be placed on GDPR compliance, metadata consistency, and interoperability between PHOEBE tools. The resulting framework will enable transparent, replicable, and efficient data pipelines for analysis and modelling.

- T5.5 - User Interface and Implementation (Lead: POLIS; Partners: NTUA, iRAP, AIM, TUM, O7; Duration: M33-M44)

This task involves the design and implementation of the PHOEBE user interface. It will identify user requirements and functional needs through stakeholder engagement and focus groups and translate them into a practical, interactive platform. The interface will support the visualisation, interpretation, and application of PHOEBE results, improving the accessibility of the predictive framework. Refinements and user testing will ensure usability and readiness for deployment.

This deliverable is organised into six chapters:

1) Introduction: Overview of WP5 objectives and its role within the PHOEBE framework.
2) Methodology: Description of the methods used for integration, synthesis, and socio-economic assessment.
3) Task Implementation: Presentation of the progress, activities, and interrelations of Tasks 5.1–5.5.
4) Integration and Framework Consolidation: Explanation of how use case outputs and models were unified into the PHOEBE algorithm.
5) Results and Transferability Analysis: Overview of the network-level demonstration results and assessment of transferability across cities.
6) Conclusions and Next Steps: Summary of findings, remaining challenges, and preparation for D5.2 (User Interface and Implementation).

A central component of this deliverable is the detailed breakdown of the experimental design settings, interventions, and future scenarios specific to each of the three use case locations:

- Athens interventions focus on pedestrian-friendly transformations, such as the pedestrianization of Vasilisis Olgas Avenue, and traffic lane reductions in Panepistimiou St. Additionally, there is a city-wide initiative to implement a 30 km/h speed limit and establish an extensive network of bicycle routes, reflecting Athens' commitment to enhancing urban mobility and safety.
- Valencia focuses on scenario testing and simulations to gain valuable insights into the impact of proposed interventions on road user behaviour, mode selection, induced demand, and safety outcomes. With its challenges of 400-500 crashes involving cyclists annually, the city aims to test

the effectiveness of different policies and interventions in promoting road safety through the PHOEBE framework.

- West Midlands focuses on scenario testing and simulations to understand the impact of proposed transport interventions (including new cycling facilities and highway and junction remodelling) on road user behaviour, mode selection, and safety outcomes. The aim is to encourage active travel, enhance safety for cyclists and pedestrians, and improve overall transportation efficiency in the region.

The use case data and insights serve as foundational reference points for the testing of the PHOEBE Framework, ensuring that the deliverable is aligned with the project's methodological standards and provides a comprehensive view of the urban mobility and safety landscape across these three regions.

WP4, of which this deliverable is a part, is inherently linked with other work packages within the PHOEBE project. The data specifications and collection processes from WP2 play a crucial role in shaping the simulations and models in WP4. Additionally, the models and frameworks developed in WP3 are essential for the safety use case implementations in WP4. The synergies between these work packages ensure a holistic approach to addressing urban mobility challenges.

Key terms

Holistic Predictive Safety Framework: An integrated methodology combining simulation, behavioural modelling, safety assessment, and socio-economic evaluation for urban environments.

Multi-Scale Risk Assessment: Linking micro-level street interactions to macro-level network safety outcomes.

Dynamic Safety Prediction: Real-time and scenario-based forecasting of safety impacts under interventions.

Data Harmonisation: Secure integration of infrastructure, telematics, and simulation datasets.

1 Introduction

1.1 Overview

The central aim of the PHOEBE project is to develop a harmonised, prospective assessment framework for road safety integrated in the transport modelling environment. The PHOEBE Framework brings together traffic simulation, road safety assessment, human behaviour modelling, mode shift and induced demand modelling and new and emerging mobility data to form a new method for the prediction of safety and socioeconomic outcomes associated with future mobility scenarios and changes resulting from new modes of transport, novel technologies, regulatory changes, policy and behavioural changes, and changes to the physical urban road environment.

The Framework is to be supported by integrated modelling and simulation tools developed to demonstrate the application of the methodology. The Framework, both in terms of its methodological process as well as the supporting tools and models, are being applied across three use case locations: Athens, Valencia, and West Midlands.

Work Package 5 (WP5) aims to consolidate and meaningfully synthesise the knowledge obtained within the previous work packages of PHOEBE and to build on the theoretical principles of the methodological framework, creating a coherent set of knowledge products, tools, and resources for exploitation beyond the project.

The core focus of this work package is the development of a holistic algorithm capable of integrated risk assessment, bridging the gap between the microscopic and macroscopic scales. This algorithm will enable the transfer of safety and risk estimations from detailed, street-level environments to wider network-level applications, drawing upon the synthesised outcomes of the PHOEBE use case trials.

WP5 will further focus on the quantification and prediction of the effects of modelled interventions, including mode shift, induced demand, and the overall impacts on travel behaviour and safety performance. The algorithm will integrate multi-source datasets, covering network-wide conditions, infrastructure assessments, and telematics data, to deliver rapid, data-driven road risk estimations across different locations, network types, and road user categories, including vulnerable road users.

Additionally, WP5 will define data flow, harmonisation, and protection processes, ensuring compliance and interoperability across PHOEBE components. Finally, a user interface will be designed and implemented based on stakeholder input and user needs, supporting the accessibility, transferability, and long-term operational use of the PHOEBE predictive safety assessment framework.

1.2 Advancing Integration and Transferability in Predictive Safety Assessment

The PHOEBE project continues to advance the frontier of predictive urban road safety analysis by integrating microscopic and macroscopic methodologies into a unified, transferable framework. Within this context, WP5 represents a critical step in transforming PHOEBE’s research outputs into an operational, scalable, and user-oriented system for assessing and improving road safety across diverse urban environments.

Building on the empirical and modelling foundations established in previous work packages, WP5 consolidates multiple strands of knowledge to deliver a holistic algorithm that enables cross-scale transfer of safety estimations, from localised simulation environments to broader network-level applications. This advancement ensures that safety insights derived from pilot areas such as Athens, Valencia, and the West Midlands can be generalised and applied to other European contexts with comparable data structures or transport challenges.

A key innovation of WP5 lies in its capacity to integrate multi-source data, combining infrastructure information, traffic simulations, and telematics datasets into a coherent predictive framework. Through harmonised data flows and dynamic modelling, PHOEBE can estimate and forecast road safety outcomes more efficiently, accurately, and contextually than conventional static approaches.

Beyond technical innovation, WP5 also promotes knowledge transfer and user accessibility through the development of an interactive user interface. This ensures that PHOEBE’s predictive safety models can be used not only by researchers but also by city planners, policymakers, and practitioners, facilitating evidence-based decision-making in road safety management.

In summary, the advancements achieved in WP5 establish the foundation for a transferable, data-driven, and policy-relevant safety assessment system. By bridging methodological, spatial, and institutional gaps, WP5 moves PHOEBE beyond conventional state-of-the-art models, offering a powerful decision-support tool that can adapt to the evolving challenges of urban mobility and transport safety across Europe.

2 Synthesis of Use Case Results

2.1 Objectives

The synthesis of use case results in WP5 is grounded in the practical implementation of the PHOEBE framework across three different urban settings. Each pilot city explores distinct types of interventions and addresses different road safety priorities, which means that the results cannot be meaningfully interpreted through direct numerical comparisons between locations.

Instead, the objective of this synthesis is to examine how interventions perform within each use case, and to express these effects in a way that allows insights to be shared across cities. The focus is therefore placed on understanding how much key indicators change when moving from one scenario to another, rather than on their absolute values.

To support this, the synthesis introduces Impact Modification Factors (IMFs) as a common reference for summarising intervention effects. By expressing results as relative changes between scenarios, IMFs make it possible to interpret outcomes consistently across different road types, user groups, and local contexts. This approach also supports the assessment of step-by-step interventions and helps identify patterns that remain stable despite differences in data sources and modelling assumptions.

2.2 Methodological Approach

The methodological approach for synthesising use case results is built around the use of Impact Modification Factors (IMFs) as a practical means of interpreting changes observed across scenarios. Rather than treating each model output independently, IMFs are used to bring together results from different indicators under a shared comparative logic.

The concept of IMFs draws inspiration from Crash Modification Factors, which are commonly applied in road safety analysis to describe before-and-after changes following the introduction of a measure. It is derived from before-and-after crash evaluations by safety researchers:

- CMF < 1 → crash risk reduction
- CMF > 1 → crash risk increase

The Crash Modification Factor is calculated based on the equation below (FHWA, 2014):

$$Crash\ Reduction = Average\ Crashes\ (after\ CM\ implementation) - (CMF\ x\ Avg.Crashes\ (Before\ CM\ Implementation))$$

(Eq. 1)

Within PHOEBE, this idea is adapted to suit a scenario-based modelling environment and is applied to a wider range of outputs, including risk indicators, surrogate safety measures, and environmental metrics.

In practical terms, IMF values make it straightforward to understand whether a change leads to better or worse outcomes. Lower values point to an improvement, whereas higher values suggest that conditions have deteriorated. This clarity is important, as it allows complex modelling outputs to be communicated without requiring the reader to engage with technical details. The IMF calculation is based on the following equation:

$$Impact\ fluctuation = Avg.Impact\ value\ (after\ intervention) - (IMFxAvg.Impact\ value(Before\ Interventions)) \qquad (Eq.2)$$

The analysis is carried out across a sequence of scenarios, reflecting the way interventions are introduced or adjusted over time. IMFs are therefore calculated step by step, making it possible to look at the effect of a single measure on its own, as well as its interaction with other changes. Because the emphasis is placed on relative differences, the results remain comparable even when starting conditions vary substantially between use cases.

To increase confidence in the findings, the IMF-based comparisons are supported by simple consistency checks. These are used to confirm that observed changes are not driven by random fluctuations, but are repeatedly visible across different road segments and user groups. Uncertainty is communicated through confidence ranges presented alongside IMF values, helping readers judge how stable the estimated effects are and where results should be interpreted more cautiously.

Taken together, this approach provides a clear and workable way to bring together outputs from different models. By concentrating on how indicators change between scenarios, and on whether these changes are stable, the synthesis supports cross-case interpretation while remaining adaptable to other urban contexts.

2.3 Transferability of the results and Replicability

Transferability

The transferability of results concerns the extent to which the study's methodological approach and modelling framework can be applied to other contexts, rather than the direct replication of empirical outcomes. Given that road-user behaviour is highly context-dependent, transferability should be explicitly argued rather than assumed. It depends on the degree of similarity between study locations and other settings in terms of road infrastructure, traffic regulations, enforcement practices, and broader cultural norms shaping road-user behaviour. By examining non-compliance behaviours across three geographically and culturally distinct locations, the study demonstrates the applicability of its data-collection strategy and behavioural modelling approach across diverse urban contexts. While observed prevalence rates are expected to vary across settings, the underlying methods and models are transferable to environments with comparable traffic conditions and regulatory frameworks.

The transferability of results involves evaluating how well the findings can be applied to other locations with similar or different characteristics. With regard to the PHOEBE project, the questions will be whether we can transfer results from one-use-case to another or to another location within another city/region or even to a location in another country. Locations can have variations in several aspects ranging from infrastructure to traffic composition to road users' behaviours. Even if infrastructure, road environment, and traffic compositions are similar between two locations, the behaviours may possibly have variations. There may be a possibility that some aspects and part of the results are comparable, however, it is important to consider the differences in behaviours or other variations when comparing different locations and assessing the transferability of results.

Although a comprehensive transferability test would require external data from a third use-case (e.g. Budapest), which is out of the scope of this project, the transferability aspect might be looked at, to some extent, within use-cases. As an example, within the PHOEBE project, speeding behaviour was studied and analysed using same data collection method (Stated Preference Surveys) and the data analysis and modelling technique (Random Parameters Ordered Probit (RPOP) models) in the three use-cases including, West Midlands (UK), Athens (GR), and Valencia (ES). The following description of the results

(Kalantari et al., 2025) provide some insights on the extent of transferability considering the differences in road environments, traffic compositions, driver behaviours, and other factors between the three use-cases as well as limitations in the data (such as potential self reporting bias, smaller sample size in Valencia as compared to the other use-cases).

In two of the three use cases, WM and Athens, the presence of a physical median and higher traffic density were consistently linked to lower self-reported speeding, emphasizing the role of road design and traffic conditions in reducing risky driving behaviour. The presence of enforcement mechanisms, including speed cameras and higher posted speed limits, reinforced this effect. Context-specific differences were also evident: in WM, prosocial drivers reacted more to traffic and road design features; in Athens, older drivers were particularly responsive to visible enforcement; and in Valencia, small overall effects combined with large differences between drivers suggest uneven acceptance of rules. With such settings, external enforcement signals serve as clear safety indicators, and compliance may rely more on demographic characteristics such as age and sensitivity to risk than on internalised social values. In Spain, attitudes toward speeding and enforcement have been described as more mixed, with considerable differences among drivers. In Spain's lower-trust enforcement context, compliance seems to depend less on general demographic trends and more on individual orientations. Prosocial drivers were more likely to view credible speed limits and radar enforcement as legitimate, while others remained disengaged, reflecting the cultural ambivalence toward traffic rules often reported in surveys and evaluations.

The differences between the UK and Greece were also evident in socio-demographic factors: in the UK, higher education was associated with compliance, yet older men and prosocial drivers with more experience were more likely to be non-compliant. In Greece, the pattern was reversed: older age was protective, but more educated, prosocial drivers with high exposure, and prosocial drivers involved in crashes were associated with higher reported speeding. Similarly, Spain showed a related paradox, where prosocial experience reduced speeding overall but shifted toward greater risk among crash-involved drivers.

These findings suggest that certain roadway infrastructural and traffic characteristics can promote speed reduction across different locations, even though the magnitude of their effects may vary by context. In contrast, interventions related to enforcement and posted speed limits exhibit substantially lower transferability, as driver responses are strongly shaped by behavioural, demographic, cultural, and individual characteristics. These factors influence how measures such as visible enforcement, speed cameras, and posted limits are perceived and complied with, leading to heterogeneous outcomes across cities.

Replicability

Replicability refers to the extent to which the results of a study can be reproduced by applying the same models and procedures at different points in time. In the short term, replicating the analysis using identical datasets, models, and methodologies is expected to yield similar outcomes, assuming that underlying conditions remain stable. This short-term consistency supports the reliability of the analytical framework and the validity of the applied methods.

However, replicability over the long term may be more challenging. Changes in external conditions, the emergence of new patterns or modes, and shifts in underlying data distributions can influence the results when the same models are reapplied after extended periods, such as six months or one year. While the original procedures-such as data collection methods, survey instruments, data integration processes, and analytical techniques-can be repeated, the outcomes may vary due to these evolving factors rather than flaws in the methodology.

Additionally, replicability does not necessarily imply obtaining identical results. In many cases, results may be reproduced within an acceptable range rather than matching exactly. Such variability reflects inherent uncertainty in real-world data and modeling processes. Acknowledging this uncertainty is essential, as it highlights the importance of reporting confidence intervals, tolerance ranges, or error margins when interpreting replicated findings.

Replicability can also be shown in a way how we applied the framework in one use-case and learned lessons from it, adapted and applied to another use-case. Testing the PHOEBE framework across three distinct use cases demonstrated its strong replicability while highlighting the importance of contextual adaptation. The core methodological components proved to be transferable across cities; however, lessons learned from implementation confirm that the framework must be tailored to local conditions to remain effective. Adaptations were required to reflect differences in the size and typology of the road network assessed, the mobility behaviours and vulnerable user groups under analysis, and the availability, quality, and granularity of local data sources. Some examples are highlighted below:

Cities are interested in different behaviour and user groups

The PHOEBE use cases demonstrated that cities prioritise different road user groups and behavioural challenges depending on their local mobility context. For instance, Valencia focused on micromobility users, specifically addressing red-light violations associated with the rapid growth of e-scooters and shared mobility services and the resulting conflicts at intersections. Athens, by contrast, concentrated primarily on pedestrian safety in dense urban environments, analysing both red-light violations by pedestrians and unsafe crossing behaviours outside designated areas. The West Midlands use case, on the other hand, prioritised speeding behaviour among motor vehicle drivers. While the overarching methodology for identifying and assessing non-compliant behaviours was consistently applied across all use cases, its implementation required adaptation. Differences emerged in the types of data required (e.g. intersection-based versus mid-block data collection), the design and focus of survey questionnaires, and the specific behaviours monitored through AI-based systems (e.g. video-based tracking versus telematics data). The definition of the behaviours under study should be closely aligned with each city's policy objectives and the specific road safety challenges experienced by its citizens.

The same principle applies to the selection of user groups included in the simulation and impact assessment. While the PHOEBE framework is designed to account for a broad range of road users, including car drivers, public transport passengers, cyclists, motorcyclists, pedestrians, and micromobility users, not all cities have the capacity to simulate the effects of interventions for every group. This may be due to local irrelevance of certain user groups or, more commonly, limitations in data availability and modelling capabilities. Although PHOEBE promotes a holistic assessment of road safety, practical implementation may require assumptions or simplifications when comprehensive data are not available.

Data and models availability can facilitate or difficult replicability

Data and modelling availability vary widely across cities, which has direct implications for the replication of the PHOEBE framework. While the framework is designed to be applied end-to-end, some cities may require additional preparatory steps depending on their existing analytical maturity. For example, where traffic simulation models are already available and properly calibrated, these steps can be partially or fully bypassed, enabling faster deployment of the framework. In contrast, when iRAP assessments already exist, additional effort may be required to ensure network compatibility and alignment between datasets, road sections, and analytical assumptions.

The PHOEBE Framework presents high technical complexity and skills dependency

Beyond the technical lessons learned, the replication of the PHOEBE framework across the three use cases highlighted notable differences in how the framework steps, methodological logic, and resulting outputs were understood and operationalised by local teams. While the SELECT–SIMULATE–EVALUATE–EXTRAPOLATE structure proved robust and adaptable across diverse policy, data, and governance contexts, the framework's integration of multiple advanced components introduces a level of complexity that can challenge replication. Replicability may therefore be jeopardised in contexts where local authorities or practitioners lack sufficient technical capacity, modelling expertise, data readiness, or institutional support to operationalise these components without substantial external assistance. In this respect, a key role of the PHOEBE project is to explicitly identify and articulate the skills and knowledge required to build competent multidisciplinary teams capable of independently running the framework or preparing its application in new cities and regions.

2.4 Development of relevant methods to broaden the comparison

To compare variables of interest across different study scenarios, a range of analytical approaches may be employed depending on the data characteristics and study design. Within the PHOEBE project, one of the key behaviours examined is driver speeding. For this purpose, 85th percentile speeds were extracted at multiple roadway segments under various traffic simulation scenarios.

A straightforward method for comparing outcomes between scenarios is the use of a paired t-test to assess differences in mean speeds. The advantage of using this approach is that is simple to implement and interpret and it also controls within the section variability, and has high statistical power when assumptions are met. However, this method is limited to comparisons between two scenarios, sensitive to outliers, and assumes normality of differences. To address some of these limitations, Wilcoxon Signed-Rank test provides a non-parametric alternative and is robust to outliers and skewed distributions. The disadvantages include lower statistical power than paired t-test when data are normal and it tests differences in medians rather than means.

Analysis of Variance (ANOVA) method can be used to compare the means of three or more independent groups to assess whether at least one group differs significantly from the others. The method operates by dividing the total variability in the data into between-group and within-group components and computing an F-statistic to determine whether observed differences are attributable to systematic factors rather than random variation. However, depending on the data type, a suitable ANOVA approach should be used. For example, If the groups being compared are independent, standard one-way ANOVA may be appropriate. This standard approach ignores any correlation between the observations. In case the same experimental units are measured multiple times under different conditions, repeated measures ANOVA may be more suitable where the assumption is that observations within the same unit are correlated. It accounts for within-subject (or segment) variability and correlation between repeated measurements.

To demonstrate it, the following results (Table 1) are obtained by applying paired t-test on Total Fatal and Severe Injuries (FSI) estimated on Athens use-case network through iRAP models.

Table 2.1: Paired t-test results on FSIs between different Study Scenarios

Column	Scenario A	Scenario B	Mean A	Mean B	t-statistic	p-value	CI Lower (95%)	CI Upper (95%)	Significant (p<0.05)
Total - VO-FSI	**Baseline API on**	**Baseline APIoff**	0.000434	0.00043	0.681386	0.496054	-8.20E-06	1.69E-05	No
Total - MC-FSI	**Baseline API on**	**Baseline APIoff**	0.001801	0.001825	-1.59582	0.111383	-5.44E-05	5.66E-06	No
Total - Ped -FSI	**Baseline API on**	**Baseline APIoff**	0.00282	0.002867	-1.0936	0.274842	-0.000131	3.75E-05	No
Total - BC-FSI	**Baseline API on**	**Baseline APIoff**	0.001177	0.001191	-1.74532	0.081759	-3.04E-05	1.81E-06	No
Total - VO-FSI	**Baseline API on**	**With Interventions API on**	0.000435	0.000242	3.994324	7.83E-05	9.80E-05	0.000288	Yes
Total - MC-FSI	**Baseline API on**	**With Interventions API on**	0.001805	0.001474	0.834578	0.404495	-0.00045	0.001113	No
Total - Ped -FSI	**Baseline API on**	**With Interventions API on**	0.002826	0.001241	3.630844	0.000322	0.000726	0.002443	Yes
Total - BC-FSI	**Baseline API on**	**With Interventions API on**	0.001179	0.000817	3.413369	0.000713	0.000153	0.000570	Yes

The estimations are based on results from different scenarios including the study network 'without' incorporating pedestrians non-compliance behavior (Baseline API Off) and the network 'with' incorporating the behaviours (Baseline API On). Another scenario is also included where an intervention is added in the network along with the behaviours (With Interventions API On). Table 1 shows the paired t-test results comparing the differences in means of the FSIs between these scenarios.

2.5 Good Practices and European-Level Guidelines

At European level, road safety and urban mobility policies are framed by a coherent set of strategic documents and regulatory instruments. Central among these is the EU Road Safety Policy Framework 2021-2030, which establishes the long-term objective of achieving Vision Zero and the intermediate target of halving fatalities and serious injuries by 2030. This framework is explicitly based on the Safe System approach and is supported by EU-wide Key Performance Indicators to monitor progress.

The Road Safety Policy Framework is complemented by the Sustainable and Smart Mobility Strategy and the EU Urban Mobility Framework, which together emphasise the transition towards safe, sustainable, inclusive, and efficient urban transport systems. These policy instruments place a clear priority on walking, cycling, public transport, and shared mobility, while calling for better integration of safety considerations into transport planning, infrastructure design, and regulatory decision-making. Binding EU legislation, such as the General Safety Regulation (EU) 2019/2144, further reinforces this framework by strengthening vehicle safety requirements and improving protection for vulnerable road users.

Existing road safety assessment methodologies, such as crash-based network screening, Crash Prediction Models (e.g. HSM-derived approaches), infrastructure-focused risk ratings (e.g. iRAP), and hybrid frameworks developed under the Network-Wide Road Safety Assessment (NWRSA) requirements have significantly advanced the ability of road authorities to identify high-risk locations and prioritise interventions and have been supporting the implementation of EU policies. However, as highlighted by Paliotto et al. (2024), these approaches often remain constrained by reliance on historical crash data, limited transferability across national contexts, partial consideration of user behaviour, or a primary focus on infrastructure risk rather than system-wide mobility outcomes.

PHOEBE builds on and extends this state of the art by delivering a harmonised, prospective, and policy-oriented framework that directly supports the objectives of EU road safety and urban mobility policies. PHOEBE is strongly aligned with these European-level policies and directly supports their implementation. The project translates high-level policy objectives into operational, transferable assessment practices that can be applied by cities and regions across Europe.

Unlike methodologies that primarily rank risk or predict crash frequency, the integration component enables authorities to assess not only where risk exists, but how planned mobility measures affect safety, behaviour, and system performance before implementation. In doing so, PHOEBE directly operationalises the Safe System approach promoted by the EU Road Safety Policy Framework and supports compliance with Directive 2019/1936 by complementing reactive crash analysis with robust proactive assessment.By explicitly incorporating vulnerable road users and emerging mobility modes into simulation-based safety assessment, PHOEBE responds to the European Commission's call for improved tools to support Safe System implementation in urban environments.

Overall, PHOEBE functions as a bridge between European policy ambition and local implementation. Its good practices complement existing EU regulations and strategies by strengthening the analytical foundation for evidence-based mobility planning and by supporting the consistent application of the Safe System approach across different national and urban contexts.

2.5.1 National-Level Road Safety Policies in PHOEBE Pilot Countries

Across the European Union, road safety policies at national level are primarily shaped by Member States' National Road Safety Strategies and Action Plans, which operationalise EU-level objectives within national legal, institutional, and territorial contexts. While the specific instruments and timelines differ between countries, these strategies are increasingly aligned with the Safe System approach, recognising that road deaths and serious injuries are preventable and that responsibility is shared across road users, infrastructure designers, vehicle manufacturers, and policymakers.

Common elements of national road safety policies include the prioritisation of safety in urban areas, the protection of vulnerable road users (VRUs), speed management measures such as 30 km/h zones, and the systematic improvement of road infrastructure safety through audits, inspections, and targeted interventions.

Many Member States have also introduced specific regulatory and planning responses to the rapid growth of active and emerging mobility modes, including cycling, e-scooters, and shared mobility services. Progress is typically monitored through national performance indicators and crash statistics, although the level of analytical sophistication and predictive capability varies significantly across countries.

Within this policy landscape, PHOEBE also supports national safety and urban mobility regulations. It establishes good practice methodologies that strengthen the evidence base underpinning national decision-making. A key contribution of PHOEBE is the shift from predominantly reactive, crash-based safety analysis towards prospective safety assessment, allowing the safety impacts of planned mobility measures to be evaluated before implementation.

Furthermore, PHOEBE introduces a more inclusive analytical perspective by enabling the explicit assessment of impacts on different road user groups, including pedestrians, cyclists, micromobility users, and specific demographic groups such as older users or women. This supports national authorities in aligning road safety strategies with Safe System principles and broader social objectives, including equity, accessibility, and public health. In this way, PHOEBE good practices are fully compatible with existing national policies while providing advanced tools to improve their effectiveness and consistency and its alignments have been demonstrated analysing the national policies of the project pilot countries.

Greece

Road safety policy in Greece is led by the National Road Safety Strategic Plan 2021–2030, elaborated by the Hellenic Ministry of Infrastructure and Transport. The strategy explicitly takes into account the principles of Vision Zero and the Safe System approach, with a long-term objective to ultimately eradicate road fatalities by 2050 and an intermediate target to reduce road deaths and serious injuries by 50% until 2030. This plan identifies a comprehensive framework covering governance, infrastructure safety, vehicle safety, enforcement, post-crash response, and systematic performance monitoring by using defined (KPIs) (Yannis et al., 2022).

The critical approach to safety has put much more emphasis on urban road safety, reflecting the high proportion of fatalities occurring in the built-up areas and among vulnerable road users, particularly among two-wheeler riders and pedestrians. Parallel to the above, Greece has introduced a binding legal framework for Sustainable Urban Mobility Plans, making them mandatory for larger municipalities and explicitly integrating road safety objectives-such as speed management, promotion of active mobility, and protection of vulnerable users-into the local mobility planning process (Stavropoulou, E., 2022)

PHOEBE is best aligned with the Greek policy context by providing methods that enable prospective safety assessment in urban environments, fit for contributing directly to the core objectives of the national strategy. Integration of simulation-based safety analysis and behavioural modelling contributes to enabling Greek authorities to assess persistent patterns of urban risk and enhancing the safety dimension of SUMP-based decision-making.

Spain

Spain’s road safety policy is articulated through the Road Safety Strategy 2030, coordinated by the Directorate-General for Traffic (DGT) (Dirección General de Tráfico, 2022*)*. The strategy adopts a Safe System framework and aligns with EU and UN targets, aiming to reduce road deaths and serious injuries by 2030. It explicitly recognises the transformation of urban mobility systems and identifies cities as a critical focus area, given the increasing interaction between motor vehicles, pedestrians, cyclists, and micromobility users.

Key pillars of the Spanish strategy include speed management, most notably the nationwide introduction of 30 km/h limits on single-lane urban roads, safer urban infrastructure, improved enforcement, and enhanced data-driven risk management. The strategy places particular emphasis on “Safe Cities”, highlighting the responsibility of local authorities and police forces in implementing traffic calming, reallocating street space, and protecting vulnerable road users in dense urban environments.

PHOEBE’s good practices directly support the Spanish approach by enabling cities such as Valencia to test and compare urban mobility measures, assessing their impacts on safety, traffic performance, and user behaviour. The speeding behaviour tested for Valencia, for example, align closely with the National strategy first indicators: Percentage of vehicles travelling within the speed limit. As well as indicator 7, percentage of distance driven over roads with a safety rating above an agreed threshold, where the star ratings can be used as safety rating. The PHOEBE framework reinforces Spain’s shift towards risk-based, proactive safety management and provides analytical support for the implementation of urban-focused measures under the Road Safety Strategy 2030.

United Kingdom

In Great Britain, road safety policy is set out in the Road Safety Strategy published in January 2026 by the Department for Transport. This strategy represents the first comprehensive national road safety framework in over a decade and places the Safe System approach at its core, explicitly recognising that human error is inevitable and that the transport system must be designed to prevent serious injury and death when errors occur. The strategy sets national targets to reduce the number of people killed or seriously injured on roads by 2035 and prioritises evidence-led interventions, innovation, and improved data use (UK Department for Transport, 2026).

The UK's road safety strategy places strong emphasis on safe infrastructure, protection of vulnerable road users, vehicle technology, and robust enforcement. Particular attention is currently being paid to urban and metropolitan contexts, where active travel is being actively encouraged. While national policy provides the overarching framework, responsibility for delivery lies largely with regional and local authorities-such as those in the West Midlands-which are increasingly adopting regional strategies aligned to Vision Zero and embedding road safety objectives within wider transport and mobility planning frameworks (Transport for West Midlands, 2023).

The PHOEBE framework holds a very good fit within this context of the ‘Big moves’ underlining thematically regional action on transport, safety, and accessibility in the West Midlands Local Transport Plan. In particular, PHOEBE directly supports the Big Moves related to walking, wheeling, cycling and scooting, as well as the creation of a safe, accessible and reliable transport network, by enabling a deeper understanding of users’ needs, the barriers they face, and the factors influencing their travel choices.

This support is possible through PHOEBE’s integrated methodology, which can take into account traffic simulations, road safety assessments, as well as behavioural models. This will therefore provide the basis for an informed approach to planning and delivering space for cycles and walking while taking into account the safety aspects for vulnerable users. This allows for a more holistic understanding that is vital if regional ambitions are to be met regarding an increase in active travel, while regional authorities remain on track with the Safe System and Vision Zero philosophies. Additionally, it allows for short-term and long-term benefits for users, which can be vital for regional authorities seeking to understand infrastructure and policy interventions better.

Across Greece, Spain, and the United Kingdom, national road safety strategies show a clear convergence towards Safe System principles, a stronger focus on urban road safety, and closer integration between mobility planning and road safety policy. PHOEBE demonstrates how it contributes to these national frameworks by providing a harmonised and transferable methodology for prospective road safety assessment, which can be embedded within existing planning, appraisal, and decision-making processes. This demonstration supports a clearer understanding of the potential of the PHOEBE framework to inform safer urban environments and to facilitate alignment with national regulatory requirements, thereby enhancing its relevance for application in other EU Member States.

3 Framework protocols

This chapter aims to provide the blueprint for implementation of the PHOEBE Framework.

3.1 PHOEBE Framework

The PHOEBE framework is a predictive safety assessment methodology designed to support cities in understanding the future road safety and socio-economic impacts of transport interventions before they are implemented. It acts as a replicable and adaptable blueprint that integrates traffic simulation, road safety assessment, human behaviour modelling, and socio-economic evaluation within a single analytical structure. By combining established planning tools with behaviourally informed and simulation-based approaches, the framework enables cities to test alternative scenarios, capture dynamic effects such as modal shift and behavioural adaptation, and compare policy options in a consistent and evidence-based way. PHOEBE therefore supports more informed, transparent, and forward-looking decision-making in urban transport planning.

*Content already presented in D3.2, but replicated in this deliverable to support standalone guidelines

The PHOEBE framework is divided across three stages:

- Setup stage: The setup stage provides initial operational results for necessary subsequent analysis. In the baseline, the simulation in the setup stage provides skim matrices relative to travel cost, travel time, travel distance, and number of trips. All are inputs for mode shift. For the after scenario, speed and flow values are also used to generate initial risk values to identify the relative risk between modes. The relative risk is also used in the mode shift distribution. All the components run in parallel for the baseline scenario. In the after scenarios, additional steps are required to ensure the changes are well represented in each component.

- Simulation stage: At the simulation stage, the mode shift and induced demand in the case of the after scenarios – generate OD matrices segregated by gender and age. These matrices are re-imported into the simulations to run the final results. At this stage, the Behaviour components are plugged into the simulation software to account for how individuals react to baseline or intervention configurations of the network under analysis.

- Refinement stage: With the final results of the simulation, including mode shift and human behaviour component contributions, the risk models are recalculated to generate the safety KPIs. The FSIs estimations in the refinement stage are also enhanced with the behaviours' contribution.

To effectively comprehend the PHOEBE framework, it is critical to illustrate each stage within the integration flowchart clearly. The flowchart provides a visual representation that enables users to track the progression of data and processes from the initial setup through simulation and into refinement. Directional arrows are included to clarify the movement of information between stages and components, allowing for a straightforward understanding of how outputs from one phase serve as inputs for the next.

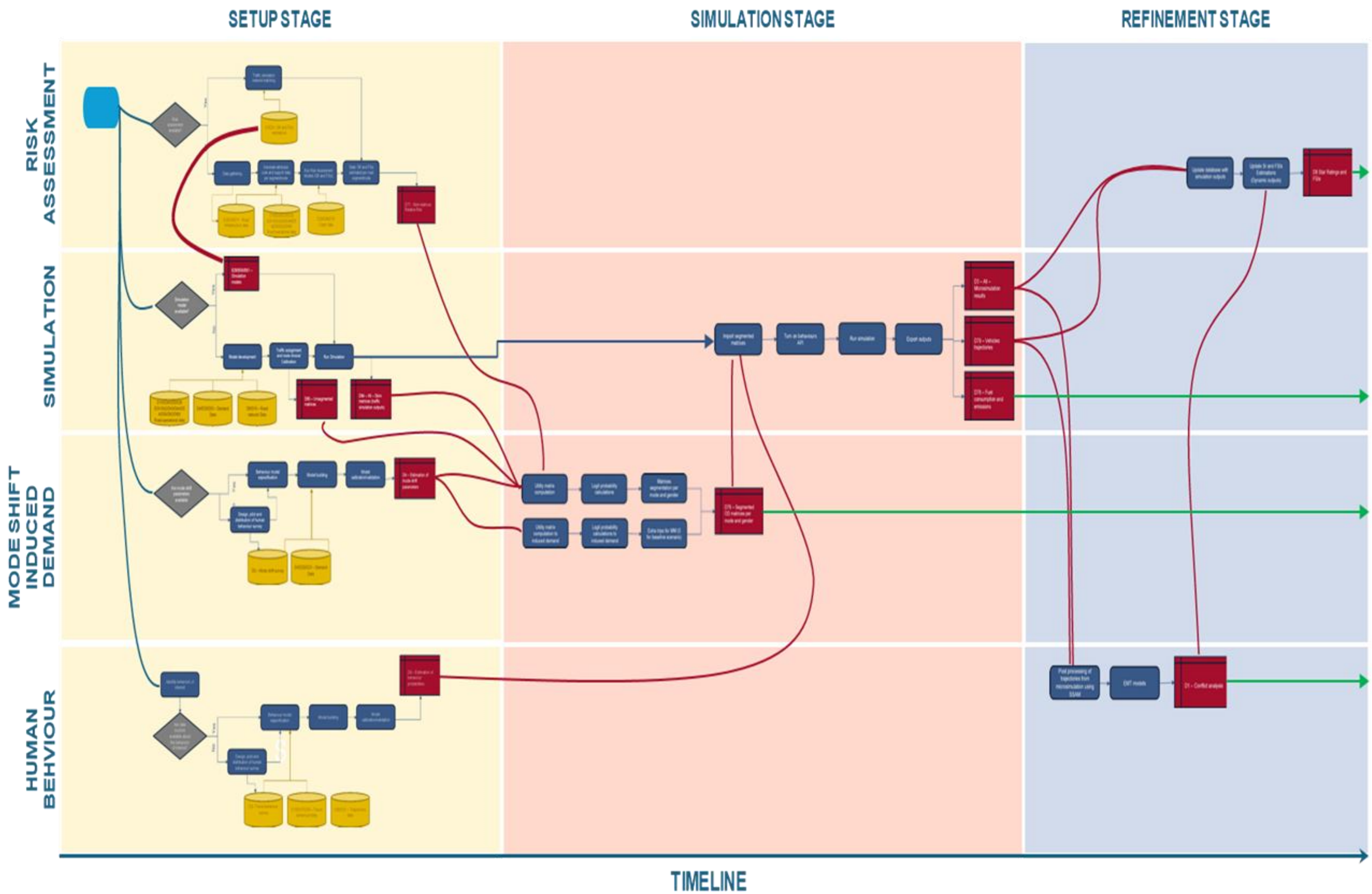


Figure 3.1: The PHOEBE Framework (*source D3.2*)

3.2 Traffic Simulation

Traffic simulation within the PHOEBE framework serves as a fundamental mechanism for using inputs from multiple modelling components (such as behavioural models, mode choice and safety assessment tools) to simulate and predict how future mobility scenarios may influence safety across urban transport systems. By running tests in a virtual environment, PHOEBE reduces real-world trial costs, avoids implementation delays, and enables full control and reproducibility of experiments.

This section outlines the full set of **requirements, processes and configuration steps** needed to run simulations effectively within the PHOEBE framework. It describes all essential components on how to activate the APIs, scripts and execute simulations. Overall, it provides a comprehensive guide to everything needed on the simulation side to run the PHOEBE framework.

3.2.1 Baseline transport model

The **baseline transport model** represents current traffic conditions against which future scenarios, interventions, or policy measures are assessed within the PHOEBE framework. It provides a calibrated and validated microsimulation environment that accurately reproduces observed traffic states, road-user interactions, and operation performance. This baseline is essential for ensuring that subsequent safety, behavioural and socioeconomic analysis are grounded in realistic traffic dynamics.

The setup of a baseline microsimulation model begins with the **definition of the transport network**. The resulting network consists of links and nodes that reflect the physical characteristics of the road infrastructure, including lane configurations, speed limits, functional road classes, and intersections layouts. Static elements such as pedestrian crossings, cycling facilities, and public transport lanes are also incorporated. It is important to emphasize that the model to be used in the PHOEBE framework needs to be developed at a **microscopic resolution**. This means that, if a macroscopic model is only available for a specific city, a micro model must be developed to forecast behavioural changes and understand its impact on safety.

Once the physical network is established, **traffic control elements** are defined. This includes the coding of traffic signals with their corresponding phase plans, timings, and coordination schemes, as well as the implementation of priority rules, yield controls and access restrictions. **Travel demand** is then introduced through the specification of origin-destination matrices, typically segmented by transport mode and time period. Demand profiles are defined to reflect temporal variations in traffic, such as morning and evening peak periods, inter-peak conditions, and weekends. Depending on data availability and modelling objectives, demand may be assigned using static or dynamic assignment procedures.

The development and execution of baselines models within PHOEBE rely on the **Aimsun Next** software package, making **license management** an important practical consideration. Appropriate Aimsun licenses must be available for model development, calibration and scenario execution, including any required add-on modules for the activation of non-compliant behaviours. Further information on how to acquire Aimsun licenses can be found in the Aimsun webpage.

Capturing this baseline accurately is the primary objective of **Step 1 of the PHOEBE framework**. Information is exchanged iteratively between the core components of the PHOEBE framework. The process begins with the road safety assessment module, which generates static estimates of crash risk and crash occurrence at the road-segment level. These estimates are provided as inputs to both the mode choice model and the behavioural modelling component. The mode choice model, including induced demand effects, produces probabilities for selecting different transport modes, while the behavioural modelling component delivers probabilities associated with specific road-user behaviour. These outputs are subsequently used as inputs to the traffic microsimulation.

After the traffic microsimulation is executed, **detailed trajectories** for individual road users are produced. These trajectories are post-processed through a surrogate safety assessment model to identify and quantify traffic conflicts, and the resulting conflict indicators are fed back into the behavioural modelling component. In parallel, the microsimulation supplies aggregated traffic flow and speed data to the road safety assessment module.

Based on the simulated interaction and behavioural outcomes, the behavioural modelling component provides calibration parameters related to fatal and serious injuries (FSI) to the road safety assessment. These parameters are used to update safety indicators and to calculate dynamic Star Ratings that account for variations in traffic speeds and flows over time. The updated Star Ratings are then returned to the traffic microsimulation for the visualization of spatially and temporally varying risk across the network.

The core outputs of Step 1 of the PHOEBE framework consist of key performance indicators generated by the traffic microsimulation and the road safety assessment components. These KPIs form the foundation for comparing the baseline scenario with future scenarios in the subsequent steps of the PHOEBE framework. Further information on the PHOEBE framework can be found in the Deliverable D1.2 Theoretical principles and methodological approach of the PHOEBE framework and selection of tools.

3.2.2 Calibration and validation

The reliability of a traffic microsimulation model depends heavily on how well it is calibrated. A key challenge is determining which behavioural model parameters and macroscopic traffic variables – such as flow, speed and travel time – most strongly influence simulation accuracy.

Calibration consists of adjusting model parameters and inputs so that the simulated network reproduces real-world traffic conditions. This requires high-quality, statistically robust observed data that represents the study area well. Essential calibration datasets include road capacities, traffic counts, travel times, speeds, delays, queues, and, at least partly, travel demand patterns.

By calibrating and validating the model using multiple data types – not just traffic counts – the microsimulation can more accurately reproduce congestion patterns, route choices, and travel times. This leads to more realistic road-user interactions and trajectories, which are crucial for safety assessment and socioeconomic analysis within the PHOEBE framework. In particular, model validity is tested in the use case regions before use by comparing simulation outputs with measured quantities.

Following the implementation of the three PHOEBE use cases, it has become evident that substantial adjustment to the simulation networks is required to achieve full compatibility with the PHOEBE framework. Across all cases, the main areas of focus for preparing a simulation model include:

- **Transitioning from macroscopic to microscopic simulation (if applicable).** When an existing model is developed at a macroscopic level, it must be converted to a microscopic simulation to support the detailed behavioural modelling required by PHOEBE. This includes increasing network granularity, adjusting traffic flow representations, and redefining vehicle interactions at a more detailed level.
- **Integrating Vulnerable Road Users and Additional transport modes**. The models must be expanded to include pedestrians, cyclists and other VRUs as well as relevant additional modes such as taxis, motorcyclists, or micro-mobility services. These user groups must be accurately represented through appropriate behavioural parameters and infrastructure elements.
- **Ensure correct traffic signal coding and integration of cycle infrastructure**. Traffic lights must be carefully validated and adjusted to ensure accurate phasing, timing and coordination. Additionally, cycle lanes and bicycle-specific features must be coded consistently and integrated across all relevant intersections so that multimodal behaviour can be simulated reliably.
- **Re-calibrating the model with updated data.** All simulation networks must be recalibrated to reflect the most recent traffic, mobility and behavioural data available. This includes updating demand matrices, route choice data, turning movements and VRU flow patterns to ensure the model reflects current real-world conditions.
- **Focusing the model on the specific area of analysis**. If only a portion of the network is relevant for the PHOEBE use case, a subnetwork may be generated to improve computational

efficiency and ensure that simulation efforts concentrate on the target study area. This involves extracting, refining, and validating the relevant portion of the main model.

More detailed information regarding the calibration procedures and model preparation steps is available on demand in Deliverable 3.2 Finalized models/simulation environment, methodology factsheet and users support materials.

3.2.3 Data requirements for simulation

The data requirements for PHOEBE are closely linked to the specific modelling components being applied. Traffic microsimulation relies primarily on detailed network and operational data that describe the physical and functional characteristics of the transport system. This includes road geometry, lane configurations, speed limits, intersection layouts, pedestrians' crossings, traffic controls, and public transport infrastructure. In addition, microsimulation requires time-dependent demand data, such as origin-destination matrices segmented by mode, as well as behavioural parameters that govern vehicle dynamics and road-user interactions.

Two broad categories of data are required for model development: **data for model building** and **data for calibration and validation**. Model-building data define the structural elements of the system, such as the transport network, demand structure, control logic and service characteristics. These datasets establish the baseline representation of the real world within the modelling environment. Calibration and validation data are used to adjust model parameters and assess model performance. This typically includes observed traffic counts, turning movements, speed, travel times, delays, queue lengths, bottleneck locations, and public transport performance indicators. For the case of the three use cases of the PHOEBE project, these are the data used for validation:

- Athens: Athens Traffic Management Centre data and video footage.
- Valencia: OD matrices, micromobility counts obtained through field observations, motorised vehicle counts provided by the Valencia City Council, and lane-specific speed records obtained from video analysis on the most representative road segments. Additionally, field measurements of traffic signal cycles at key intersections have been conducted and incorporated into the model.
- West Midlands: speeds, traffic volume estimates, turning frequencies, acceleration behaviours, excessive speeding behaviours, and OD matrices. Other data included infrastructure signal data, bus data, and mode counts from cameras, All Traffic Count (ATC) surveys to align and validate the model for the use case.

Further information on the need for validation for each of the use cases can be found in Deliverable 2.2 Conclusion of data register, feature report and accreditation.

3.2.4 Integration of PHOEBE components into traffic simulation

The integration of PHOEBE components into the traffic simulation is a central element of the PHOEBE framework, enabling richer behavioural realism and more robust safety evaluation. The following PHOEBE models are integrated into the Aimsun Next simulation software packages: non-compliant behaviours for vehicles and pedestrians; mode choice; and iRAP road assessments data.

To support the realistic reproduction of safety-critical interactions within the PHOEBE framework, several **non-compliance behaviours (NCB)** have been incorporated into the traffic simulation environment. For the PHOEBE use cases, these behaviours include speeding, red-light violation by vehicles, pedestrian crossings outside designated areas, and red-light violations by pedestrians. Because these behaviours are not available in the default Aimsun Next simulation platform, they require explicit implementation within the software. The integration was accomplished through the use of Aimsun's API capabilities, which allow external components to trigger specific behavioural events at predefined moments and locations within the simulation. The API activates the prescribed non-compliant behaviour during runtime, ensuring that the resulting interactions are embedded within the microsimulation.

Additionally, within the PHOEBE framework, **mode choice** outputs are also integrated into the simulator as these directly modify the demand used in the microsimulation. Since the simulator requires separate, disaggregated OD matrices for each transport mode, any update to the mode choice model needs regenerating all OD matrices. This process is automatised through scripting and further information will be provided in the corresponding section "Demand modelling " of this Deliverable.

The **risk assessment** module within PHOEBE is also integrated into Aimsun Next simulation software. It aligns length-based risk segments within the simulation network. Because iRAP data is organised in fixed-length records (typically 100c or 10m), the workflow converts these fixed-length iRAP segments into spatial buffers that can be matched to Aimsun sections. Intersection-level risks require node-based matching using polygonal node shapes extract from Aimsun. After validation, risks indicators are imported into Aimsun as custom attributes for use in safety-aware simulation. More detailed information is provided in Section 3.4 Risk Assessment, more specifically in Section 3.4.5 Integration.

Further information on each of these modules is available in Deliverable 3.1 New and enhanced models /simulation environments and user support, and *on demand* in Deliverable D3.2 Finalised models/simulation environments, methodology factsheets and user support materials.

3.2.4.1 How to activate the APIs for non-compliant behaviours

Enabling non-compliant behaviours within Aimsun Next requires the activation and use of the **Aimsun Next API,** which serves as the interface between the simulation engine and any external logic used to override or extend default behavioural models. The API is essential because it enables the injection of custom actions back into the simulation – precisely the mechanism required to implement behaviours that deviate from the standard Aimsun logic. The API provides controlled access to vehicle-level and network-level events through its structure lifecycle functions (.e.g AAPILoad, AAPIManage, AAPIPostManage), which act as the integration points where **custom behaviour can be activated or modified at runtime**. These callbacks allow direct manipulation of simulation states, including vehicle dynamics, detection data, and control decision – capabilities necessary for modelling non-compliant or exceptional behaviours (Aimsun Next, 2026).

Because this functionality depends on the API layer, API extension licensing must be enabled in the Aimsun Next installation.

The **activation of the APIs** is necessary to enable the **non-compliant behaviours**. The API is activated by providing the corresponding *.py file in the Aimsun Next API tab within the scenario window. This is the list of APIs currently developed for the PHOEBE project:

- Red-light violation for micromobility.
- Red-light violation for pedestrians and vehicles.
- Crossing outside designated areas.

Further information on each of these APIs and its architecture can be found *on demand* in Deliverable D3.2 Finalised models/simulation environments, methodology factsheets and user support materials.

3.2.4.2 How to activate scripts for the integration of PHOEBE modules

Scripts in Aimsun Next provide a flexible and automated way to control model processing, handle input datasets, and generate post-processing outputs efficiently. This way, **scripts can interact with Aimsun models**, read and modify objects and extract results.

Before setting up the framework, Aimsun Next (a version from April 2025 onwards) must be installed with the corresponding license, and with Python support enabled, including access to the *pandas* and *numpy* libraries.

To integrate the script into Aimsun Next, the user first opens the project and navigates to the Scripts section. A new script can be created and renamed to reflect its intended purpose. Through the script settings, the user selects the option to load the code from an external file and then browses the corresponding *.py file. Confirming these settings finalises the link between Aimsun Next and the external script. To execute the script, the user can simply right-click on the script created and run it. An alternative option is to define the new script as a post-run script, so Aimsun Next ensures that it will be automatically triggered after every simulation execution.

In PHOEBE, the following scripts have been used:

- Scripts to export and import OD matrices.
- Scripts to export skim matrices for cost, travel time, trips and distance.
- Scripts to import segmented matrices.

Additionally, to enable the execution of the PHOEBE framework, it is necessary to extract both standard and custom Aimsun outputs. Here are described the ones that have not been covered previously in this document:

- Extract the section geometries with their ID (iRAP) – Standard output.
- Extract the centroid location, areas, and ID (TUM) – Standard output.
- Extract the node shape (Spatial coverage) and ID (iRAP) – Custom output.
- Extract the whole simulation mean speed (iRAP) – Standard output.
- Extract the link 85th percentile speed for the whole simulation timeframe (iRAP) – Custom output.
- Extract the link hourly mean speed (iRAP) – Custom/Standard, depending on the statistical interval.
- Export a *.trj for analysis into the SSAM software (TUD) – Standard output.

3.2.5 How to perform simulations in Aimsun Next

The PHOEBE framework requires four simulation runs : two for the baseline scenario and two for the after-intervention scenario:

- **SIM 0.1 – Baseline setup stage**. This initial simulation runs independently and does not rely on any previous steps. Its main purpose is to generate skim matrices, which are required as inputs for the model shift model.
- **SIM 0.2/0.3 – Baseline simulation stage**. Building on SIM 0.1, this stage activates the behavioural components of the PHOEBE framework. It requires: segmented matrices, the behaviour API, and the appropriate environmental parameters to be configured. Once these conditions are met, the simulation produces two primary outputs: structured data stored in SQLite databases and SSAM trajectories for safety and behavioural analysis.
- **SIM 1.1 – After setup stage**. In this stage, interventions are incorporated into the simulation model, transitioning from baseline to modified conditions. This step depends on the correct integration of the interventions. Its main output is a new set of skim matrices reflecting the impacts of the applied interventions, which then feed into the mode shift and induced demand models.
- **SIM 1.2/1.3 – After simulation stage**. This stage builds on SIM 1.1 to enable an in-depth assessment of behavioural changes under the intervention scenario. It requires: intervention-enabled simulation models, segmented matrices, behaviour API activation, configured environmental parameters. Once these prerequisites are met, the simulation generates: SQLites files containing structured behavioural and performance data and SSAM trajectories capturing detailed behavioural patterns for further analysis

3.3 Human Behaviour Modelling

The behavioural modelling component integrates the PHOEBE framework to incorporate human factors and road user behaviour into traffic microsimulation and road safety assessment. The behavioural modelling involves developing both statistical and machine-learned models for road user behaviours based on empirical data.

3.3.1 Focus on Non-Compliance behaviours

Within the PHOEBE project non-compliance behaviours of drivers and vulnerable road users were prioritised and studied within each use-case. The specific behaviours included:

- **West Midlands (UK):** Car drivers' speeding behaviour
- **Athens (GR):** Car drivers' speeding behaviour and pedestrian red light violation and crossing outside the designated crosswalk
- **Valencia (ES):** Car drivers' speeding behaviour and micromobility users' red light violation behaviour

The main hypothesis considered and also suggested for developing a road user behavioural model is that the 'behaviour' is affected:

- At the individual level by individual characteristics (e.g. demographics), mental constructs, and personality traits
- At the road segment level, by local traffic / environmental factors

In order to estimate a behavioural model, one needs both sets of the above factors.

3.3.2 Data on Road User Behaviours

Previous research has shown that both infrastructural characteristics and individual attributes are associated with road user behaviour. Consequently, the development of robust behavioural models requires the integration of variables from both domains. Recent advances in data-collection technologies, such as telematics and radar systems, enable the detailed recording of road user behaviour (e.g. non-compliance). However, practical constraints and privacy regulations often limit the collection of personal characteristics for the same individuals. One viable approach to address this limitation is the use of stated-preference surveys, which can capture personal attributes, past behaviours, and responses to scenario-based situations on the road.

Within the PHOEBE project, this approach was implemented through the development of stated-preference surveys tailored to each study behaviour and use case. The surveys collected information on participants' demographic characteristics and self-reported past behaviours related to traffic-rule non-compliance. In addition, they included a range of scenario-based questions, supported by illustrative figures, reflecting different roadway configurations, traffic volumes, and speed limits (for drivers), as well as pedestrian crossing situations (e.g. crossing alone or in a group).

The infrastructural variables comprised roadway characteristics such as the number of lanes, median type (where present), and posted speed limit. Individual-level variables, by contrast, encompassed demographic attributes and past behavioural tendencies related to traffic-rule compliance. Examples of the latter include compliance with red traffic signals, yielding to pedestrians, use of hand signals by cyclists before turning, and overtaking from the correct side. These measures were used to capture individual safety-related behavioural dispositions and compliance tendencies.

3.3.3 Data Collection Technologies

3.3.3.1 Surveys

When direct empirical observation of individual-level behaviour is constrained, stated-preference surveys provide a practical and complementary data source. Such surveys enable the systematic collection of individual characteristics, self-reported past behaviours, and behavioural responses to controlled roadway scenarios defined by combinations of infrastructural and traffic-related variables. The resulting data can be used to develop behavioural models, while empirical sources such as telematics or video-based observations may subsequently be employed for validation purposes.

The behavioural survey is structured into two main components: (i) questions capturing demographic characteristics, past behavioural patterns, and relevant personality traits; and (ii) stated-preference questions based on scenario-based representations of roadway and traffic conditions. The total number of questions is determined by the set of applicable scenarios constructed from the key influencing factors identified in the modelling framework. The following questions provide illustrative examples of items related to self-reported past behaviours.

Please indicate on the scale provided the extent to which you agree or disagree with each statement.	**Never**	**Rarely**	**Sometimes**	**Often**	**Always**
1. I overtake a slow driver on the left.					
2. I disregard the speed limit on urban highways.					
3. I disregard the speed limit on residential roads.					
4. I enter an intersection knowing that the traffic lights have already changed against me.					
5. I disregard the red light at an intersection.					
6. I engage in hand held mobile phone conversations while driving.					
7. I engage in texting and surfing my social media platforms while driving.					

Figure 3.2: Survey example

For questions about a particular scenario with a set of roadway environment related characteristics, a picture of a study location or roadway configuration can be added followed by a set of questions to get a response on the participants' likely/unlikely behaviour under different operating conditions. As an example, a speeding behaviour related scenario is presented in the figure below while the questions include a variety of possible situations based on different speed limits and presence/non-presence of a speed camera.

Please look at the figure, consider yourself driving your car for your daily trips on a two-way highway and answer the questions in the following scenarios:

How likely would it be for you to go over the speed limit (more than 5 mph) on a highway with...	Extremely unlikely (1)	Unlikely (2)	Neutral (3)	Likely (4)	Extremely likely (5)
1. speed camera and speed limit of 30 km/h when traffic is low?	○	○	○	○	○
2. speed camera and speed limit of 30 km/h when traffic is high?	○	○	○	○	○
3. speed camera and speed limit of 50 km/h when traffic is low?	○	○	○	○	○
4. speed camera and speed limit of 50 km/h when traffic is high ?	○	○	○	○	○

How likely would it be for you to go over the speed limit (more than 5 mph) on a highway with...	Extremely unlikely (1)	Unlikely (2)	Neutral (3)	Likely (4)	Extremely likely (5)
5. **no** speed camera and speed limit of 30 km/h when traffic is low?	○	○	○	○	○
6. **no** speed camera and speed limit of 30 km/h when traffic is high?	○	○	○	○	○
7. **no** speed camera and speed limit of 50 km/h when traffic is low?	○	○	○	○	○
8. **no** speed camera and speed limit of 50 km/h when traffic is high ?	○	○	○	○	○

Figure 3.3: Snapshot of a speeding scenario based questions in the Survey

Further details on the survey questionnaire across the use-cases can be found under PHOEBE deliverable D3.2 (link).

3.3.3.2 Advanced tools

Advanced data collection tools such as **vehicle telematics, roadside radar, computer vision, and connected infrastructure sensors** have transformed the way road users' behavioural data can be observed and analysed. Telematics systems embedded in vehicles continuously record high-resolution data on speed, acceleration, braking patterns, lane changes, and headways, allowing researchers to infer behaviours such as speeding, aggressive driving, or distracted driving.

Vehicle telematics record second-by-second (or higher frequency) data on speed, acceleration, braking, GPS position, and time. Non-compliance such as speeding is identified by matching a vehicle's instantaneous or average speed to the posted speed limit for the road segment derived from digital maps. Harsh acceleration or late braking near intersections can also indicate risky or rule-breaking behaviour, such as attempts to beat a red signal. In some systems, in-vehicle sensors and event triggers explicitly flag violations (e.g., excessive speed thresholds or running a stop sign), producing labelled records of non-compliant events along with their duration and severity.

Roadside radar and LiDAR sensors complement this by capturing precise vehicle trajectories, gap acceptance, and compliance with traffic control devices, while camera-based systems can detect violations of red signals, unsafe overtaking, and pedestrian crossing behaviours. Compared with traditional surveys or manual observations, these tools provide objective, continuous, and large-scale measurements of real-world behaviour across different road types, times of day, and traffic conditions.

Video and computer-vision-based systems extend non-compliance detection to more complex behaviours involving interactions and vulnerable road users. Cameras, combined with object detection and tracking algorithms, can classify road users (vehicles, cyclists, pedestrians) and monitor their movements relative to marked crosswalks, sidewalks, and signal indications. Pedestrian non-compliance, such as crossing against a red signal, mid-block crossing, or delayed clearance, is detected by observing entry and exit times within defined crossing zones and comparing them with pedestrian signal phases. Similarly, failure to yield to pedestrians, illegal turns, or lane misuse can be identified by analysing trajectories and right-of-way relationships.

In Athens, the video analytics framework focused primarily on speeding behaviour of motorised vehicles, pedestrian red-light violations, and crossing outside designated areas. The processing pipeline combined deep-learning-based object detection models for road user identification with multi-object tracking algorithms to associate detections across consecutive video frames and reconstruct continuous movement trajectories. These trajectory-level data were subsequently used to derive surrogate safety indicators such as Time-to-Collision (TTC) and Post-Encroachment Time (PET) at a high temporal resolution, forming the empirical basis for behavioural modelling and crash-conflict relationship development.

In Valencia, the video algorithms were adapted to capture complex interactions between motor vehicles and micromobility users, particularly cyclists and e-scooter riders at signalised intersections. Convolutional neural network-based detection models were employed to recognise heterogeneous road user types, while trajectory tracking techniques enabled continuous monitoring of movement paths and inter-user spacing. Automated event detection logic identified red-light violations, unsafe overtaking manoeuvres, and close-proximity interactions indicative of elevated collision risk. The extracted trajectories supported both behavioural probability estimation and dynamic risk assessment within the PHOEBE framework, enabling quantitative evaluation of micromobility-related safety impacts under different traffic conditions.

Once collected, this rich behavioural data can be processed and transformed into meaningful indicators for road user behavioural modelling. Raw trajectories are typically cleaned, synchronised, and segmented into events (e.g., approach to an intersection, response to a signal change, or pedestrian curb-to-curb crossing). From these events, factors like speed choice distributions, compliance rates, and conflict indicators can be extracted for statistical and machine-learning models that represent how different classes of road users interact with each other under varying environmental and traffic conditions.

3.3.3.3 Challenges: Privacy and practical issues in collecting personal characteristics.

Despite their analytical power, advanced data collection techniques such as telematics, radar, and video-based sensing face significant **privacy and practical challenges**, particularly when attempting to link observed behaviours with the personal characteristics of road users. Many non-compliance behaviours such as speeding, red-light running, or unsafe pedestrian crossings are strongly influenced by factors like age, gender, driving experience, trip purpose, and risk perception. However, directly observing or inferring these characteristics from sensor data raises ethical, legal, and methodological concerns.

From a **privacy perspective**, telematics and video data are inherently sensitive because they can reveal identifiable information, including vehicle trajectories, license plates, faces, and habitual travel patterns. Regulations such as data protection laws often restrict the collection, storage, and linkage of such personally identifiable information. As a result, data are frequently anonymised, aggregated, or pseudonymized at an early stage, which limits the ability to associate behaviours with individual-level attributes. For example, while telematics can reliably detect speeding events, the driver’s age or socio-demographic background typically cannot be accessed without explicit consent, and even then may be constrained by legal and ethical review requirements.

These challenges create a **trade-off between behavioural richness and personal context**. To address this, researchers often rely on indirect approaches, such as combining anonymized sensor data with complementary sources like stated-preference surveys, travel diaries, or census-level socio-demographic data at an aggregate spatial scale.

Within the **PHOEBE** project, a stated preference survey approach has been used to collect both individual characteristics and expected behaviours and develop the behavioural models, while telematics and video data have been used to validate the survey based models' outcome.

3.3.4 Developing Analytical Models

Once the data are obtained either through the surveys or through multiple data sources including survey telematics and/or video-based observations, the next step is to translate the data into **analytical models of road users' non-compliance behaviour**. This process typically moves from data fusion and variables extraction to model estimation, validation, and interpretation, with the goal of explaining and predicting when and why non-compliance behaviour occurs.

Several analytical modelling approaches can be employed to represent road-user behaviour, including discrete choice models and, where relevant, models that explicitly account for interdependencies among traffic agents (e.g. game-theoretic frameworks). Discrete choice models describe the probability of a given behavioural outcome as a function of one or more exogenous covariates. Common examples include multinomial logit and ordered probit models. The multinomial logit model is typically applied to nominal behavioural outcomes that do not possess an inherent ordering, such as illegal crossing behaviour, and can be expressed as:

$$P(behaviour) = \frac{exp\ (\beta X)}{1 + exp\ (\beta X)} \qquad \textit{(Eq. 3)}$$

where X represents the vector of exogenous covariates and β\betaβ denotes the associated model parameters.

As an illustrative example, a speeding behaviour model was developed using data from a stated-preference survey conducted among drivers in the West Midlands (UK). Participants were randomly sampled from individuals who either resided in or frequently drove within the region and reported driving for more than one hour in a typical week. The final sample comprised 361 drivers, yielding a total of 8,664 observations. The survey collected information on demographic characteristics, past non-compliant driving behaviours, and stated speeding choices under a range of hypothetical scenarios defined by roadway infrastructure and traffic conditions. To facilitate scenario comprehension, visual representations depicting low- and high-traffic conditions were included. In addition, a dedicated section captured drivers' personality traits.

Speeding propensity was elicited using a five-level Likert scale ranging from *Extremely Unlikely* to *Extremely Likely*, yielding an ordinal response variable. Accordingly, a cumulative link mixed model (CLMM) was employed to model speeding behaviour as a function of roadway characteristics, traffic conditions, driver demographics, past behaviours, and personal attributes. CLMMs extend standard cumulative link models by incorporating random effects, allowing for the analysis of ordered responses with hierarchical or clustered structures, such as repeated observations from survey participants. The model estimates the probability of transitioning from lower to higher response categories, corresponding to an increasing likelihood of engaging in speeding under varying conditions.

For implementation within Aimsun, the ordinal response was subsequently transformed into a binary outcome, with the three lower categories (*Extremely Unlikely*, *Unlikely*, *Neutral*) grouped as 0 (not speeding) and the two higher categories (*Likely*, *Extremely Likely*) grouped as 1 (speeding).

The following equation presents the speeding behaviour model based on the above described CLMM approach.

$$\text{P (speeding)} = 1 - \frac{1}{1 + e^{(-(3.339 - U))}} \qquad (Eq.\ 4)$$

where *P(speeding)* is the probability of (proportion of traffic) going over the speed limit, and

U = – 0.460 × **Traffic** – 1.408 × **Speed Camera** – 0.184 × **Speed Limit (30mph)** –0.456× **Speed Limit (40mph)** + 0.268 × **Sex** – 0.020 × **Social Value Orientation** – 1.054 × **Past Behaviour: Speed Limit Violations**

Table 3.1 presents the description and type of each explanatory variable as well as its distribution within the above equation. The model results including the parameter estimates of the variables, p-values, and confidence intervals are presented in Table 3.2.

Table 3.1: Description, type, and distribution of the variables

Variable	Description	Type	Distribution
Traffic	Traffic (condition) at a road segment	Binary 0:Low 1:High	This variable comes from the traffic simulation. No distribution required.
Speed Camera	Presence of speed camera	Binary 0:No 1:Yes	This variable comes from the network. No distribution required.
Speed Limit (30mph)	Speed limit at the road (segment)	Binary 0:20mph 1:30mph	This variable comes from the network. No distribution required.
Speed Limit (40mph)	Speed limit at the road (segment)	Binary 0:20mph 1:40mph	This variable comes from the network. No distribution required.
Sex	The sex of a driver	Binary 0: female 1: male	$Bernoulli\ (\boldsymbol{\theta} = 0.47)$
Social Value Orientation	Personal characteristic of the driver	Continuous (degree)	$von\ Mises(\boldsymbol{\mu} = 28.13°, \boldsymbol{\kappa} = 16.66)$
Past Behaviour: Speed Limit Violations	Past speed limit violations by the driver	Continuous (Likert scale)	**Log-Normal** $(\boldsymbol{meanlog} = 0.3947, \boldsymbol{sdlog} = 0.4604)$

Table 3.2: Model results with parameter estimates, statistical significance, and 95% confidence interval

Variable*		Estimate	z value	Pr(>|z|)	95% CI	
					L	U
Intercept		3.339	1768.2	0.000	3.335	3.343
High traffic		-0.460	-207.7	0.000	-0.437	-0.435
Presence of speed camera		-1.408	-770.1	0.000	-1.389	-1.388
Speed limit	30 mph	-0.184	-100.7	0.000	-0.197	-0.196
	40 mph	-0.455	-249.1	0.000	-0.434	-0.433
Gender (Male)		0.267	146.3	0.000	0.139	0.140
Past behaviour	Disregarding speed limit (in residential areas)	1.053	575.8	0.000	2.270	2.272
Social value orientation (SVO)		-0.020	-3.9	0.000	-0.013	-0.012

Model Fit						
AIC	Loglikelihood at convergence	Participants				Observations
		Count	Variance	Std	ICC	
17558.03	-8767.02	361	5.81	2.41	0.63	8664

The models corresponding to other study behaviours within Athens and Valencia use-cases, are presented under Appendix 1 (Tables Appendix A1-A7).

3.3.5 Validation

Model validation was conducted by comparing model outputs with observed data on non-compliant behaviour collected across multiple study locations. Within the PHOEBE project, different empirical data sources were employed for validation depending on the specific use case. For example, in WM use case, validation relied on observed behavioural data derived from telematics sources provided by The Floow. In contrast, the Athens and Valencia use cases utilised data obtained from video recordings, which were subsequently processed using image-analysis algorithms to extract observed behaviours. Given the differences in data sources and processing requirements, the validation procedures are briefly outlined below.

Validation for Speeding Behaviour in WM:

Six Aimsun road segments were randomly selected from the WM network for each validation location, defined based on road geometry and characteristics. The selection criteria ensured that these segments were represented across all three datasets by iRAP (road infrastructure data), Aimsun (road and traffic related data from simulation model), and The Floow (telematics data on speeding). This approach was implemented to maximise the representation of survey scenario variability. The combined dataset was generated through the spatial integration of iRAP, Aimsun, and Floow data and subsequently linked to an additional dataset containing Floow's speeding proportions per section.

To estimate the observed speeding proportion, speeding rates from the corresponding Floow sections were aggregated for each Aimsun section. Because a single Aimsun section may encompass multiple Floow sections, the speeding proportions aligned to the Floow segments were averaged accordingly, with separate estimates computed for peak and off-peak traffic conditions. Peak hours were defined as three hours in the morning and three hours in the afternoon, while off-peak hours comprised midday and night hours each lasting three hours exactly based on the Floow's speeding proportions. These values, along with their corresponding 95% confidence intervals are observable in the figures provided under the validation results section of Appendix 1 (Figure Appendix A1, A2, A3).

For the survey proportions, the corresponding survey scenario IDs were identified for each scenario. To estimate the probability of participants selecting a given scenario, values for independent variables were integrated into the behavioural equation (explained in Equation 3, under section 3.3.4), which models decision-making based on relevant factors such as traffic conditions, speed limits, and past driving behaviour. The resulting probabilities of all participants were then aggregated, and their average across the selected scenarios was computed separately for peak and off-peak traffic hours.

Finally, calibration factors (CFs) for both peak and off-peak traffic hours were computed by dividing the actual observed proportions by the survey-derived proportions for each condition. The final CF for each location was then determined by averaging the CFs across all available segments within that location. The actual and survey based proportions and the resulting Calibration factors per location are provided under the validation results section of Appendix 1 (Figure Appendix A1, A2, A3).

Validation procedure for Athens Use-Case:

For the Athens use case, model validation involved deriving pedestrian non-compliance probabilities (specifically red-light running and crossing outside designated crosswalks) from both stated-preference survey–based models and empirical trajectory-based observations. Survey responses were first transformed into individual-level probabilities and subsequently mapped to real-world locations through scenario archetypes, resulting in spot- and traffic-state–specific behavioural parameters.

Independently, empirical observations provided realised non-compliance proportions under comparable contextual conditions, which served as an external reference for calibration. The validation procedure was applied consistently to both crossing outside designated crosswalks and red-light running behaviours, differing only in the underlying behavioural models, explanatory variables, and observational units. All available observation files were included without exclusion. Where multiple recordings, days, or sub-parts existed for the same physical location, empirical proportions were pooled across trajectory files to derive a single spot-level estimate.

Survey-based pedestrian non-compliance probabilities:

A. Observation-level probability computation

For each valid survey observation, a probability of the study non-compliance behaviour is computed using the relevant behavioural model presented under the previous section (3.3.4). Explanatory variables are grouped into:

• Scenario-defining variables, describing the crossing context (e.g. traffic state, crossing width, road direction, signalisation);

• Individual attributes, describing respondent heterogeneity (e.g. age, social value orientation);

• Past behaviour indicators, derived from Likert-scale questions and recorded to numeric form.

A linear predictor is constructed from the observation's attributes using the reported coefficients. The combined effect of the explanatory variables is converted into a probability of non-compliance using the transformation specified in the behavioural-model documentation. This yields a continuous probability of non-compliance for each respondent under each evaluated condition.

Specifically, probabilities are obtained from the latent utility formulation of the underlying CLMMs, implemented in practice via a logistic transformation of the thresholded linear predictor. For microsimulation compatibility, the ordinal response space is collapsed into a binary non-compliance outcome following the original model specification, yielding a continuous probability of non-compliance for each respondent under each evaluated condition.

B. Implementation

For each combination of real-world spot and traffic state (peak/off-peak), individual noncompliance probabilities associated with the corresponding scenario archetype are pooled and averaged to obtain a spot-level behavioural parameter. This average represents the expected likelihood of non-compliance for a randomly encountered pedestrian under the given contextual conditions. Measures of variability and statistical uncertainty are computed from the same pool of probabilities to support transparent interpretation and sensitivity analysis. Identical averages across multiple spots indicate shared behavioural archetypes defined by geometry and traffic conditions, rather than identical pedestrian behaviour.

To maintain conceptual consistency between survey-based inputs and empirical calibration targets, survey-derived probabilities were likewise aggregated by averaging peak and off-peak scenario probabilities within each scenario archetype prior to spot-level implementation. This harmonisation ensures that both survey-based behavioural parameters and empirical reference proportions represent context-specific average behaviour rather than time-of-day–specific effects, which are outside the scope of the current microsimulation configuration for Athens.

Actual observed behaviour from trajectory and gate-crossing outputs:

Two complementary outputs were analysed: (i) a dataset containing frame-level pedestrian annotations (pedestrian identity, frame index, legality of crossing, and signal state), and (ii) a dataset containing frame-level gate interaction signals without pedestrian identifiers. Frame indices were standardised, invalid records removed, and all data temporally ordered at 29.97 fps. To ensure behavioural estimates were based on sufficiently observed trajectories, only pedestrians tracked for at least 90 frames (≈3 s) were retained for pedestrian-level analyses. Trajectories were segmented to prevent artificial aggregation across tracking interruptions, with new segments defined when frame continuity was violated or short gaps exceeded a predefined tolerance. Duplicate frame–pedestrian records were removed prior to segmentation to prevent artificial fragmentation of trajectories.

A. Red-light running: pedestrian-level prevalence

Red-light running was measured at the pedestrian level using frame-level legality labels, treating red, orange, and unknown signal states as prohibitive for crossing. A pedestrian was classified as a violator if they showed at least one sustained illegal-crossing episode comprising a minimum of 90 illegal frames, allowing for brief tracking gaps of up to three frames. Prevalence was calculated as the share of eligible pedestrians meeting this criterion among those observed for a minimum duration. For identified violators, the presence of illegal frames under unknown signal states was recorded to reflect annotation uncertainty. Bootstrap confidence intervals were obtained via pedestrian-level resampling, preserving the empirical distribution of eligible pedestrians within each file.

B. Crossing outside designated crosswalks*: gate-level incidence*

Because individual pedestrians cannot be uniquely identified in the gate-based observations (due to privacy issues), crossing outside designated crosswalks is quantified as gate-level event incidence rather than pedestrian-level prevalence. For each gate position, observed legal, illegal, and, where available, ambiguous crossing streams are analysed, and the resulting event-based incidence is used as an empirical reference for calibrating survey-derived average non-compliance probabilities under equivalent contextual conditions. Gate streams may encode either identifier-based sequences, in which changes correspond to distinct crossing events, or cumulative counters, in which increases represent event accumulation. The underlying encoding is detected automatically using heuristics based on monotonicity, reset behaviour, jump magnitude, and value diversity. Crossing outside designated crosswalks incidence is defined as the proportion of illegal events among all known crossing events, with ambiguous events excluded from the denominator. Statistical uncertainty is estimated using robust binomial confidence intervals suitable for both small counts and extreme proportions. Event rates are additionally normalised by observation duration, while frame-level exposure metrics are reported only when identifier-based encoding is detected and are suppressed otherwise to avoid artefactual interpretation.

C. Aggregation across files, days, or parts

When multiple independent estimates of the same behavioural proportion are obtained from separate files, days, or sub-parts, aggregation is performed using count-based pooling rather than simple averaging, whereby numerators and denominators are summed across files prior to proportion and confidence-interval estimation. Confidence intervals for pooled estimates are computed from the aggregated counts using the same binomial framework as in the primary analyses.

The resulting calibration factors are provided under the validation results section in Appendix 1 (Figure Appendix A4, A5).

Validation for Valencia use case:

For micromobility users in Valencia, a total of three intersections were analysed. The actual non-compliant behaviour proportions, specifically red signal violations, were computed based on the video-extracted data. These were determined by dividing the number of observed red light violations by the total number of crossings for each intersection, separately for peak and off-peak traffic hours. Additionally, 95% CIs were calculated for these observed values to quantify variability in non-compliance rates.

For the survey-derived proportions, a methodology similar to that used in the WM case was applied. Relevant survey scenarios were selected for each intersection, ensuring alignment with real-world conditions. The probability of a red light violation in each scenario was estimated using a behavioural model that incorporated road characteristics, traffic conditions, and individual-level factors. The resulting probabilities of all participants were then aggregated, and their average across the selected scenarios was computed separately for peak and off-peak traffic hours. CFs were also calculated similar to the WM (See Appendix Figure A6 under Appendix 1 for results).

3.3.6 Integration

Within the PHOEBE framework, the behavioural models are integrated into Traffic Microsimulation directly and into iRAP Road Safety Assessment models indirectly (i.e. through utilising the simulation output). The process of integrating the behavioural models into microsimulation is described earlier under section 3.2.4.

The integration of behavioural factors into iRAP models is performed by using the output of traffic microsimulation integrated with behavioural models. The approach however, adopted speeding behaviour (within West midlands use-case) involved using the updated flows and speeds from the behavioural integrated microsimulation runs and using in the iRAP models. For the pedestrians' non-compliance behaviour in Athens and micromobility red light running behaviour in Valencia use-case, calibration factors have been determined. This is done by developing a crash-conflict relationship (by an extreme value theory based model) from the actual observed data on conflicts and crashes in these use-cases. Afterwards, this relationship is applied to estimate crashes from the simulated conflicts output under the scenario of **'with' and 'without'** behavioural models integration into microsimulation. The calibration factor is calculated by taking the ratio of these estimated crashes under the two scenarios. This calibration factor is then used to incorporate into and update the road safety assessment, as illustrated in the flowchart below (Figure 3.4).

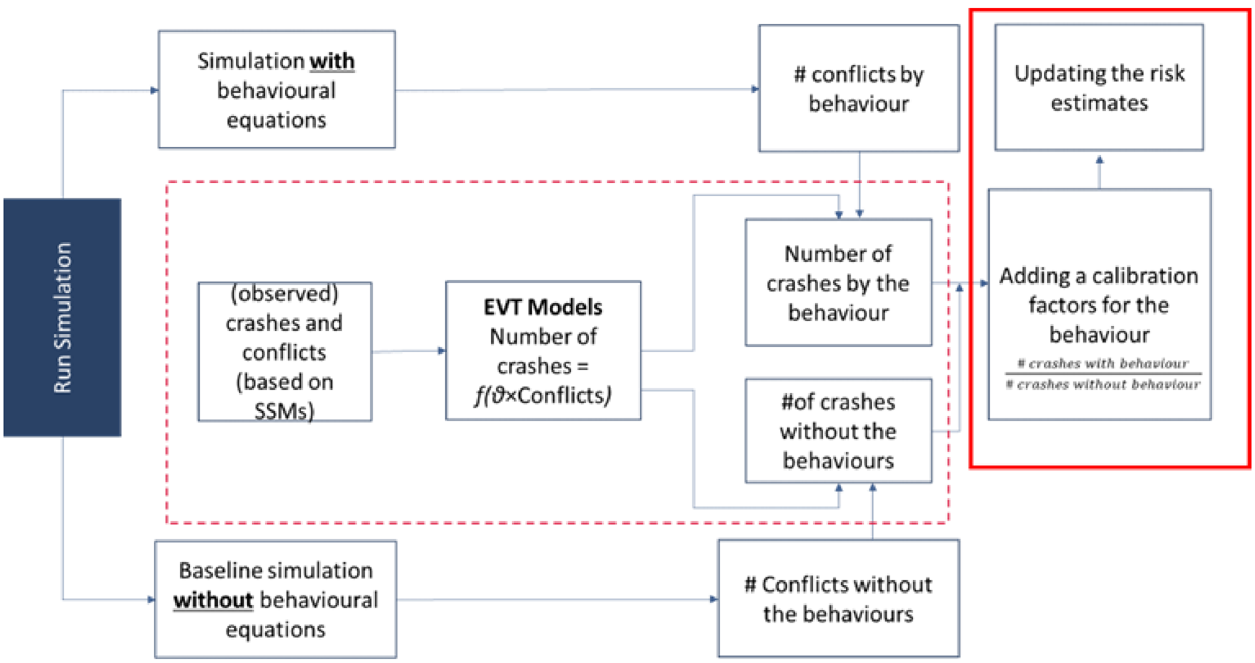


Figure 3.4: Estimation of crashes based on simulated conflicts using EVT modelling

General Methodology for Crash-Conflict Relationship:

The proposed framework for developing crash–conflict relationships is a modular and diagnostic-driven methodology that quantifies traffic safety risk using SSMs, such as TTC and PET, to represent conflict severity. The framework operates on vehicle-VRU interaction data collected across multiple driving sessions and preserves temporal integrity by explicitly identifying session boundaries (such as day transitions or discontinuities in frame indices) before any modelling is performed. SSM values are transformed to emphasise conflict imminence, such that smaller values correspond to higher surrogate severity and receive greater influence in downstream risk estimation.

Potentially safety-critical interactions are identified using an unsupervised anomaly-detection strategy that combines two complementary perspectives on abnormal traffic behaviour. One component is designed to detect rare interaction patterns in a multidimensional feature space, while the other identifies deviations from typical joint behaviour across key interaction variables through robust multivariate modelling. This dual strategy enables scalable identification of high-risk pedestrian–vehicle or micromobility user–vehicle interactions without relying on crash data, which are often sparse or unavailable. Where proxy definitions of risky behaviour are available (e.g. rule-based TTC or PET thresholds), they may be used for sensitivity analysis or model calibration; however, they do not define the extreme events used for tail-risk estimation.

To ensure compatibility with the assumptions of extreme value theory (EVT), detected interaction candidates are subjected to explicit temporal declustering. Short-term bursts of correlated events are reduced to a single representative exceedance, and a minimum separation between retained events may be enforced to further limit residual temporal dependence. EVT is applied only to the resulting set of approximately independent events. Tail risk is then modelled using complementary EVT formulations, including block-based and threshold-based approaches, with threshold selection and model adequacy evaluated using diagnostic plots, goodness-of-fit measures, and information criteria. EVT inference is treated as conditional rather than guaranteed: when data support is insufficient or diagnostics indicate unstable or implausible tail behaviour, tail extrapolation is deliberately avoided.

To maintain robustness across heterogeneous data conditions, the framework incorporates a hierarchical modelling strategy. When standard EVT models prove unreliable, uncertainty-aware or more flexible distributional alternatives are employed to characterise tail behaviour without violating statistical assumptions. Where relevant, the framework can be extended to jointly model multiple conflict indicators—such as TTC and PET—and to account explicitly for their dependence structure. All fitted models and diagnostics are retained to enable transparent comparison, uncertainty assessment, and sensitivity analysis.

The framework ultimately produces uncertainty-aware estimates of conflict risk and calibrated surrogate thresholds that can be translated into interpretable crash-frequency measures under standard stochastic assumptions. By conditioning inference on diagnostic validity and embedding principled fallback mechanisms, the proposed approach provides a statistically defensible and operationally robust basis for traffic safety risk estimation in complex urban environments.

3.4 Demand Modelling

The demand model is an umbrella term used to describe models that quantify modal shifts and induced demand caused by the introduction or implementation of road safety interventions in specific use cases - defined as particular contexts or scenarios where these interventions are assessed. The models that are used for these purposes are:

- **Mode choice models:** To estimate the modal shifts
- **Induced demand models:** To estimate the extra trips produced after any changes in a particular mode's travel cost, time, and so on.

This section provides details about these models. It starts with the setup of the demand modelling procedure, then describes data collection, model development, calibration, and integration into the main framework. This chapter also provides details of the calculation methodology for the associated risks with the modes (alternatives) directly used in the modelling procedure, which are primary input variables utilised for integrating these models with the iRAP road safety assessment models. Please refer to PHOEBE deliverable D3.1 (section: 2, subsection: Demand model) for more details.

3.4.1 Demand modelling set-ups

The demand models, i.e., the mode choice models and the induced demand models, are developed from scratch for the three use cases. For this purpose, three separate stated-preference surveys were conducted in the three use cases. The surveys comprise responses from 500 respondents for each use case, representing the age and gender composition of the particular use cases. The alternatives from the stated-preference experiments were incorporated based on the real-world modal shares of the use cases. The alternatives with small modal shares are not considered in the surveys. The considered alternatives according to the use cases are given below:

- **Athens:** Car, public transport, motorcycle, shared micro-mobility, and walking.
- **Valencia:** Car, public transport, private micro-mobility, shared micro-mobility, and walking.
- **West Midlands:** Car, public transport, private micro-mobility, shared micro-mobility, and walking.

Two trip lengths were considered: 2 km (1.2 Miles for the West Midlands) and 5 km (3.1 Miles for the West Midlands) for both mode choice and induced demand models.

The induced demand models were only considered for the shared micro-mobility mode. The changes in attributes such as travel costs, times, and associated risks were assumed to be the drivers of induced trips or trips outside the routine. The alternatives presented to the responders were +2 trips/week outside routine trips, +1 trip/week, no change, -1 trip/week, -2 trips/week and none under changing attributes. The models were developed in R using the Apollo package. The mode choice models were calibrated as per the real-world modal share of the use cases provided by the use case leaders. The induced demand models are not calibrated owing to the absence of real-world data. In project PHOEBE, the demand models address the following points:

- Effects of socio-demographics on mode choice, induced demand and modal shifts.
- Effects of the associated risks with the modes (alternatives) on mode choice, induced demand and modal shifts.
- Effects of trip distances mode choice, induced demand and modal shifts.
- Effects of quantifiable intrinsic and extrinsic characteristics associated with the modes, e.g., cost, time, etc. on mode choice, modal shift, and induced demand.

A detailed description of the demand model setup can be found in PHOEBE deliverable D3.1 (section: 2, subsection: Demand model).

3.4.2 Data on Demand models

As described in the former sections, the demand models were developed from stated-preference data. The data collection process begins with a preliminary understanding of mode usage in the region. With the help of established literature in this field, the requirements for the data of the demand models for the project PHOEBE are established. This step is followed by the step where the relevant attributes and variables, as per the use cases, are deduced. Finally, the surveys were designed to address the specific use cases. A detailed description of the surveys is provided in the sections below.

3.4.3 Data Collection Technologies

3.4.3.1 Surveys

The surveys comprise responses from 500 responders for each use case. These 500 responders represent the socio-demographic composition, in terms of age and gender, of the use cases. The surveys have four distinctive parts. These are:

- Travel behaviour (revealed inputs)
- Stated-preference experiments for mode choice
- Stated-preference experiments for induced demand
- Socio-demographics (revealed inputs)

The stated-preference experiments for the mode choice part of the demand modelling employ D-efficient design and are designed using the software NGENE, since these experiments are considerably complex. The stated-preference experiments for the induced demand part employ an orthogonal design. A total of 108 and 54 experiments are generated for mode choice and induced demand, respectively. A total of eight mode choice and four induced demand experiments are presented to the responders in a random sequence, resulting in sample sizes of 4000 and 2000 for the mode choice and induced demand models, respectively. The attributes and their levels used for the mode choice and induced demand are presented in Table 3.3, below:

Table 3.3: Attributes and their levels

	Mode choice		Induced demand	
Nr.	**Attributes**	**Levels**	**Attributes**	**Levels**
1	Travel cost	+30%, 0% (base), -30%	Change in Travel cost	+25%, 0% (base), -30%
2	Door-to-Door time	+30%, 0% (base), -30%	Change in Door-to-Door time	+25%, 0% (base), -30%
3	Access-Wait-Egress time	+30%, 0% (base), -30%	Change in Access-Wait-Egress time	+25%, 0% (base), -30%
4	Associated risk (relative)	+30%, 0% (base), -30%	Change in Associated risk (relative)	+25%, 0% (base), -30%

Multiple scenarios were also considered for the stated-preference experiments. These are listed in the Table 3.4 below.

Table 3.4: Scenarios of mode choice and induced demand experiments

	Mode choice		Induced demand	
Nr.	**Scenarios**	**Values**	**Scenarios**	**Values**
1	Trip distances	2 km (1.2 Miles), 5 km (3.1 Miles)	Trip distances	2 km (1.2 Miles), 5 km (3.1 Miles)
2	Trip purpose	Commute, shopping, and leisure	Trip purpose	Commute, shopping, and leisure
3	Condition during the trip	Daytime, Night + Light, Night + Dark (without light)	Condition during the trip	Daytime, Night + Light, Night + Dark (without light)

3.4.3.2 Challenges: Privacy and practical issues in collecting personal characteristics.

The challenges of fielding surveys with stated-preference experiments appear right after delineating the survey requirements. Survey design, including the stated-preference experiment design, is the first challenge in the process. The questionnaire must be short enough to finish, but detailed enough to capture time, cost, access, safety, and trip purpose. Induced demand surveys are more complex than standard mode choice surveys in terms of understanding. A common struggle amongst the responders is imagining taking an extra trip based on changing attributes, e.g., travel costs, time, and associated risk. To overcome these challenges, the surveys must be tested on a smaller sample (5-10% of the total planned sample) before full launch. The responses must be tested and modelled to ensure the quality of the inputs. This process indirectly sheds light on the understandability of the surveys. During survey design, specific attention must be given to the contexts of the questions and the experiments, as they vary significantly across local languages and conventions. These issues can increase the number of dropouts, along with other factors such as survey length.

Another challenge is achieving a quota that reflects the local population in terms of age, gender, and other demographics. Special care must be taken when planning surveys to target skewed population groups, e.g., senior citizens, since these groups are much harder to reach. Since a large section of the population uses mobile phones, special care must be taken in designing the survey interface, which should also enable mobile usage while answering the surveys. Bots, duplicates, and low-effort responses require attention checks and careful cleaning. Local data protection rules play a significant role in effective data collection. Special care must therefore be taken to abide by these rules and regulations. Finally, other constraints that can shift travel behaviour, e.g., weather, holidays, strikes, and local events, must also be considered before fielding the surveys.

3.4.4 Developing Analytical Models

The analytical model development phase starts after the successful completion of the data collection phase. The first step of this phase is data cleaning. The data cleaning step comprises mainly the following steps:

- Removal of duplicate and incomplete entries
- Removal of data coming from bots, fast completions, and inconsistent responses
- Tracking missing responses

After data cleaning, the collected data is coded accordingly to make it suitable as model inputs. This step mainly involves building a dataframe compatible with the software package to be used for model estimation. Some data collection interfaces also provide these facilities; i.e., they can seamlessly produce compatible dataframes that can be used directly as model estimation inputs. Specific care must be taken to break down complex variables, e.g., door-to-door time into in-vehicle time and access-wait-egress time (defined during the survey design), so they can be appropriately modelled.

Model estimation must start with a simpler model specification. After testing the signs, scale, and plausibility of the simpler estimates, more complex models can be estimated. After the estimation process, a thorough inspection of the trade-offs, such as the value of time and cost sensitivity, must be checked and validated. Finally, the estimated models must be calibrated using real-world data. In the case of the induced demand model, the models could not be calibrated and validated after these steps, since such data do not exist in the real world. The final models can then be integrated into the main framework. PHOEBE utilises Apollo package in R environment for the analytical model estimation.

3.4.5 Models calibration

Model calibration is the step in which we ensure the model reproduces a known baseline before we trust any scenario results. The process starts from a fixed utility structure. The alternative-specific constants are then tuned, keeping the behavioural coefficients fixed so that the predicted mode shares match the observed baseline shares. This process is repeated several times until the predicted modal shares converge with the observed shares. Calibration also ensures the plausibility of key outputs, e.g., average trip lengths, time distributions, and sensitivity to cost and time. The estimated models were calibrated considering the baseline year of 2024 for all the use cases. Please refer to the PHOEBE deliverable D3.1 (section: 4, subsection: Demand model) for more details.

3.4.6 Integration

Model integration of demand models (mode choice and induced demand models) in this workflow is implemented through OD demand matrices and a feedback loop between microsimulation and the risk assessment model components. The process begins with unsegmented OD matrices and the skim matrices (travel time, speed, etc.) generated by the microsimulation models for the use cases. In parallel, the iRAP component generates risk at the link level. This information is mapped in the OD matrix format for all the considered modes. These matrices are treated as the skims for the risk component. These skims are then treated as inputs to the mode choice models (utility equations) to compute choice probabilities, which are used to split unsegmented OD flows into segmented, mode-specific OD matrices.

After this process, the segmented OD matrices are fed back into the microsimulation and risk assessment models, and the interventions are deployed at this stage. This step produces the updated skim matrices, which the mode choice models then use as inputs to generate the new modal shares (new travel demand matrices). The induced demand model is introduced at this stage, with the difference between the baseline and the scenario skim metrics as inputs. The induced-demand probabilities for the shared micromobility modes are then used to inflate the segmented OD matrix to compute induced trips using those modes. These segmented travel demand matrices are then utilised in subsequent processes.

The overall process of integration can be expressed with the sequence of the following steps:

Baseline case:

a. Skim matrices (baseline) and unsegmented OD matrices (baseline) production from traffic microsimulation models.
b. Baseline relative risk per mode generation from iRAP road safety assessment models.
c. These inputs are then fed into the mode choice model to segment the baseline OD demand matrices into different modes.
d. The segmented OD demand matrices are returned to the traffic microsimulation models from the mode choice model component for intervention setup.

Intervention setup:

a. The planned interventions are set up in the traffic microsimulation models, iRAP risk assessment models and other PHOEBE components.
b. New skim matrices are produced by the traffic microsimulation models and are fed into the mode choice models.
c. New relative risks per mode are calculated by iRAP road safety assessment models and are fed to the mode choice models.

Scenario run:

a. The mode choice runs with the new skim matrices and new relative risk inputs to generate new modal shares.
b. The induced demand models (for shared micro-mobility) take the baseline and scenario skim matrices and calculate the differences in the predictors, e.g., travel times, and travel costs.
c. These differences are used by the induced demand models to generate additional demand for shared micromobility modes.
d. The extra demand is then added to the segmented OD matrix of shared micromobility produced by the mode choice models.
e. The finalised OD matrices are then returned back to the traffic micro simulations and other components of the PHOEBE framework for further analysis.

KPI calculations:

a. The segmented OD demand matrices can then be used to calculate PHOEBE KPIs, e.g., the difference between the sum of the active mode matrices of baseline and a particular scenario gives the change in active trips due to the implementation of the intervention.
b. The segmented OD demand matrices are used by the traffic microsimulation models to generate the emissions.

The sequence described above sums up the overall integration through OD demand and skim matrices.

For more details, please refer to the PHOEBE deliverable D3.1 (section: 3 subsection: Demand model).

3.5 Risk assessment

The PHOEBE Framework is designed to evaluate and predict road safety risks in urban environments. At its core, it supports strategic decision-making by providing robust safety insights that inform and is informed by demand modelling, traffic simulation and behavioural data integration.

The risk assessment module within PHOEBE is responsible for generating all key performance indicators (KPIs) used to evaluate the safety impact of interventions across all road user groups, including pedestrians, cyclists, motorcyclists, and vehicle occupants.

PHOEBE defines risk as measured risk, grounded in objective data sources such as crash statistics, exposure rates and infrastructure attributes

While the framework is flexible enough to accommodate various risk models, PHOEBE recommends the adoption of the models below:

- iRAP's Star Rating methodology
- Fatal and Serious Injury (FSI) Estimation models

These models provide a globally recognized standard for assessing road safety performance and enable comparative analysis across regions and user types.

This risk assessment component of the PHOEBE Framework guide is intended to support users in executing high-quality risk assessments within the PHOEBE framework. It provides step-by-step guidance on the following topics.

3.5.1 Data Requirements

To perform iRAP Star Ratings and Fatal and Serious Injuries (FSIs) analysis, a structured approach is essential, starting with comprehensive road attribute data gathering. The process begins by surveying the road network and coding attributes along defined chainage intervals in accordance with iRAP specifications. Traditionally, attributes are coded every 100 meters, which provides a balance between accuracy and efficiency for most environments. However, this interval can be adjusted for dense urban areas, where shorter segments (e.g., 20m or even 10m) may be necessary to capture complex features such as frequent intersections, pedestrian crossings, and variable roadside conditions.

The data required for iRAP analysis is primarily a combination of road-related attributes that describe the physical and operational characteristics of the roadway. A set of attributes can be identified through video images of the road and subsequently recording the attributes manually or automatically. Over 50 attributes are recorded and include parameters such as lane width, number of lanes, median type, shoulder width, roadside hazards, intersection control, and facilities for pedestrians and cyclists.

To know more about the requirements for the video surveys, please refer to the iRAP Survey Manual. This manual provides detailed guidelines on the collection of road survey data, which forms a key part of the process for producing Star Ratings.

For further information on coding attributes and how to perform manual coding, please consult the iRAP Coding Manual - Drive on the Left Edition or iRAP Coding Manual - Drive on the Right Edition. These manuals give detailed guidance on the coding of road attribute data, which inform the generation of Star Ratings.

To learn more about AI-driven initiatives focussed on automated detection of road attributes, visit the official AiRAP Webpage. This resource provides insights into projects and explains the procedures companies must follow to obtain accreditation for generating road attributes in accordance with the Star Ratings data specifications.

In addition to the information that can be derived from the video surveys, a number of operational parameters are also required. They are:

- Road user flows (road AADT, peak bicyclists flow and peak pedestrian flows along and across the inspected road).
- Motorcycle percentage (%)
- Heavy goods vehicle (HGV) percentage (%)
- Speed (posted/legal speed limit and operating speed)

For the setup assessment stage, the required data can be sourced from multiple inputs. Specifications for these attributes are defined in the iRAP Coding Manual (v3.10). During the refinement stage, the data is generated directly from the simulation results. When simulations are performed using AIMSUN Next, an API can be used to automatically transfer data between the two platforms (see Chapter 5). If other simulation software is used, the resulting data should be converted to comply with the required specifications.

Finally, where available, detailed crash data is required to calibrate the model using observed crash data to adjust and validate the model so that predicted risk levels align with actual crash experience. Specifically, the analysis relies on disaggregated fatality and severe injury outcomes by crash type and road user group to ensure that risk estimates and Star Ratings are grounded in real-world safety outcomes. Figure 3.5 illustrates how the data should be entered into iRAP's VIDA tool, including, where possible, the user groups and associated crash types. Although only fatality data is required at this stage, the model also requires the ratio of serious injuries to fatalities.

Percentages | Fatalities

Assigned total 0
Calibration total 0

	Vehicle occupant		Motorcyclist		Pedestrian		Bicyclist	
	Percentage (%)	Fatalities	Percentage (%)	Fatalities	Percentage (%)	Fatalities	Percentage (%)	Fatalities
User group distribution	0	0	0	0	0	0	0	0
Run-off LOC driver-side	0	0	0	0			0	0
Run-off LOC passenger-side	0	0	0	0				
Head-on LOC	0	0	0	0				
Head-on overtaking	0	0	0	0				
Intersection	0	0	0	0			0	0
Property access	0	0	0	0				
Along			0	0	0	0	0	0
Crossing intersected road					0	0		
Crossing inspected road					0	0	0	0
Other	0	0	0	0	0	0	0	0

Figure 3.5: Fatality data structure in iRAP ViDA System.

It is essential that all data undergo a robust quality assurance process. Achieving high-quality road attribute coding relies on effectively managing coder fatigue, applying structured quality control procedures, and maintaining clear oversight of the data. Quality assurance should be integrated throughout the coding workflow, including peer-to-peer reviews, daily progress monitoring, independent quality checks, and large-scale mapping verification. Peer reviews promote consistency and a shared understanding among coders, while supervisor-led progress checks enable the early detection of errors, identification of training needs, or determination of whether re-coding is required. Independent reviews, based on a representative sample of road types and conditions, provide an additional layer of validation and should be undertaken at defined milestones. Network-wide consistency is further strengthened through mapping checks and manual verification of automated outputs, acknowledging that current automation tools still require careful human oversight to achieve the required accuracy standards. The iRAP Coding Manual. (v3.10) provides detailed guidance on how to conduct quality checks on input data for the Star Rating models.

3.5.2 Generating results

3.5.2.1 Star Ratings calculations

Star Ratings are calculated within iRAP's ViDA system, the organisation's global, free-to-access digital platform for road safety analysis, reporting, and performance tracking. ViDA serves as the core platform for applying iRAP methodologies at scale, including the generation of Star Ratings, estimates of Fatal and Serious Injuries (FSI), and the development of Safer Road Investment Plans (SRIPs).

As an online data processing and analytics environment, ViDA hosts coded road attribute data, applies approved iRAP models, and generates standardised safety results. These results can then be consistently analysed, visualised, mapped, and reported across different countries and projects.

At a functional level, ViDA serves as the calculation and management platform for iRAP assessments. Users upload road attribute data that have been coded in accordance with iRAP specifications, and the ViDA system applies the embedded Star Rating models to calculate risk levels and safety outcomes for different road user groups.

To initiate Star Rating calculations, several prerequisites must first be satisfied. First, a ViDA project structure must be established, including the appropriate programme, region, and project hierarchy. Second, the coded road data must be uploaded using the required file structure and formats, ensuring that all mandatory attributes are included and expressed in the correct units. Before calculations can proceed, ViDA automatically performs validation checks on the uploaded data to confirm that the data meets structural and formatting compliance. The Star Rating model is centrally maintained within the platform; users do not install or modify it. Instead, calculations are triggered directly within ViDA once data readiness has been verified.

Responsibility for running Star Rating models in ViDA typically lies with users who have been granted the appropriate access rights and calculation permissions. Although the platform allows different user roles to perform calculations, good practice is for Star Rating analyses to be undertaken or overseen by individuals with a strong understanding of the iRAP methodology. This ensures correct interpretation of results, early identification of data issues, and the consistent application of quality assurance procedures.

Within the PHOEBE framework, star ratings are calculated using an infrastructure-based risk assessment approach, aligned with the principles of the iRAP Star Rating methodology (to know more, access iRAP

Methodology Reference Guide). The objective is to proactively quantify the inherent safety of the road environment for different road user groups based on observable characteristics, rather than retroreflectively relying on historical crash data.

The calculation is computed into ViDA and begins with the coding of the road network attributes into homogeneous segments and intersections. Each element is described using standardised attributes that influence crash likelihood and injury severity, including road layout, cross-section, roadside environment, intersection type, presence of pedestrian and cycling facilities, and the prevailing speed environment.

For each road user type (vehicle occupants, pedestrians, cyclists, and motorcyclists), risk scores are assigned to individual infrastructure attributes using established risk relationships. These relationships link specific design layout and operational features, together with operating speed, to expected crash severity outcomes. Speed plays a central role in this calculation, acting as a key modifier of injury risk across all user types.

Attribute-level risk scores are then combined to produce an overall risk score for each road segment and each user group. The resulting continuous risk score is subsequently mapped to a discrete star rating band, ranging from 1 star (highest risk) to 5 stars (lowest risk). The model detailing this is available at the iRAP Methodology Reference Guide).

The star rating methodology is expected to produce a spatially disaggregated classification of road safety performance, assigning a star rating (1 star - highest risk, to 5 stars - lowest risk) to each road segment and intersection for each road user group. The results can be visualised in an interactive map per segment in the set-up stage and/or in a link-node based on the network matching with the simulation software performed.

3.5.2.2 FSI estimation procedures

The FSI estimations follow a similar process and are also processed in ViDA. Within the PHOEBE framework, Fatal and Serious Injury (FSI) estimates are derived using the iRAP trauma estimation methodology, which builds directly on the data and outputs generating Star Ratings. FSI estimation combines infrastructure-based risk, traffic exposure, and severity relationships to estimate the expected number of fatal and serious injuries likely to occur along each road segment for each road user groups. The model per user group and crash type is available at the iRAP Methodology Reference Guide.

The procedure uses the same coded road attributes applied in the star rating assessment, including road layout, roadside environment, intersection characteristics, and speed environment. These attributes define the underlying crash risk and injury severity profile for each road user type. Traffic flow and exposure data are then incorporated into the risk scale to estimate the number of injuries, allowing the translation of relative risk into absolute FSI estimates at segment or network level. FSI estimates are calculated separately for vehicle occupants, motorcyclists, cyclists, and pedestrians, ensuring that differences in vulnerability and crash mechanisms are explicitly captured.

The FSI estimation is expected to produce quantitative estimates of fatal and serious injuries for each road segment and road user group - expressed as absolute numbers over a defined analysis period and spatial extent. The results will highlight locations and user groups contributing disproportionately to the injury burden, enabling comparison between baseline and future scenarios. When applied across scenarios, the FSI outputs will indicate the expected change in fatal and serious injuries resulting from infrastructure, speed, or policy interventions, enabling the opportunity for prioritisation of countermeasures with the greatest potential safety benefit. At network level, aggregated FSI estimates will provide an overall indication of safety performance and expected injury reduction, complementing star ratings by translating relative risk into tangible, outcome-focused safety impacts.

3.5.3 Integration

3.5.3.1 Relative Risk for Demand Models

The demand models incorporate risk indicators as one of the explanatory variables used to estimate travel decisions. Within these models, risk is treated as a relative measure, benchmarked against the level of risk experienced by vehicle occupants.

As part of the PHOEBE framework, the total SRS divided by length generated during the setup-stage road safety assessment are used to derive these relative risk values. This allows the demand models to reflect how variations in infrastructure or modal risk—normalised, for example, relative to public transport or another reference mode—can influence travellers' mode choice and induce demand responses.

How to calculate:

For each transport mode, divide the total SRS value by the length of the network segments for which SRS data is available. This step is essential because the Star Rating model only produces SRS outputs on segments where user flows exist. Normalising by segment length ensures that the resulting metric reflects the density of safety risk relative to the actual travelled network, rather than overstating risk on segments with limited or no user movement.

3.5.3.2 Network matching and connectivity with Aimsun Next

The integration of road safety assessment data into traffic simulation environments requires a robust and systematic network matching procedure. Within the PHOEBE framework, this process ensures that risk indicators derived from Star Rating are accurately represented within the Aimsun simulation network. The primary objective is to align risk data collected by a certain chainage with the link and node structure of the simulation database.

Because iRAP assessments are structured as length-based records, each representing 100-metre segments, the first task is to reconstruct network elements that correspond to these intervals. Two approaches can be taken, depending on whether (i) an iRAP assessment already exists for the simulation region, or (ii) no such assessment is available.

(i) How to calculate if an iRAP assessment exists in the simulation region

When an assessment exists, the iRAP database must be spatially linked to the Aimsun section IDs. Figure 3.6 presents the six steps proposed for network alignment.

iRAP databases provide point data corresponding to the start of each segment; therefore, continuous links must first be generated to spatially represent these 100- or 10-metre units, depending on the desired resolution (Step 1). From these links, centroids are calculated to form the basis of influence areas (Step 2). Buffer zones are then constructed around each centroid (Step 3), encapsulating the spatial footprint of each segment and enabling spatial matching with Aimsun sections (Step 4).

This influence-area approach allows risk data to be transferred from linear datasets—such as iRAP or telematics link data—to the link-node structure of the traffic simulation model. The same workflow is applied not only to iRAP data but also to telematics sources and other link-based systems, such as OpenStreetMap extracts. Once the buffer zones are defined, spatial matching assigns the corresponding data to all Aimsun sections intersecting each buffered area.

Automated validation checks are then performed to verify that key attributes—such as road name and lane count—are consistent between the matched objects (Step 5). Any inconsistencies, including mismatched lane counts or incorrectly identified road names, are flagged for manual refinement (Step 6). A secondary manual review identifies sections that either lack assigned data or contain incorrect matches, using colour-coded diagnostic maps for quality assurance. This step is particularly important in dense urban networks, where complex geometries can hinder automated alignment.

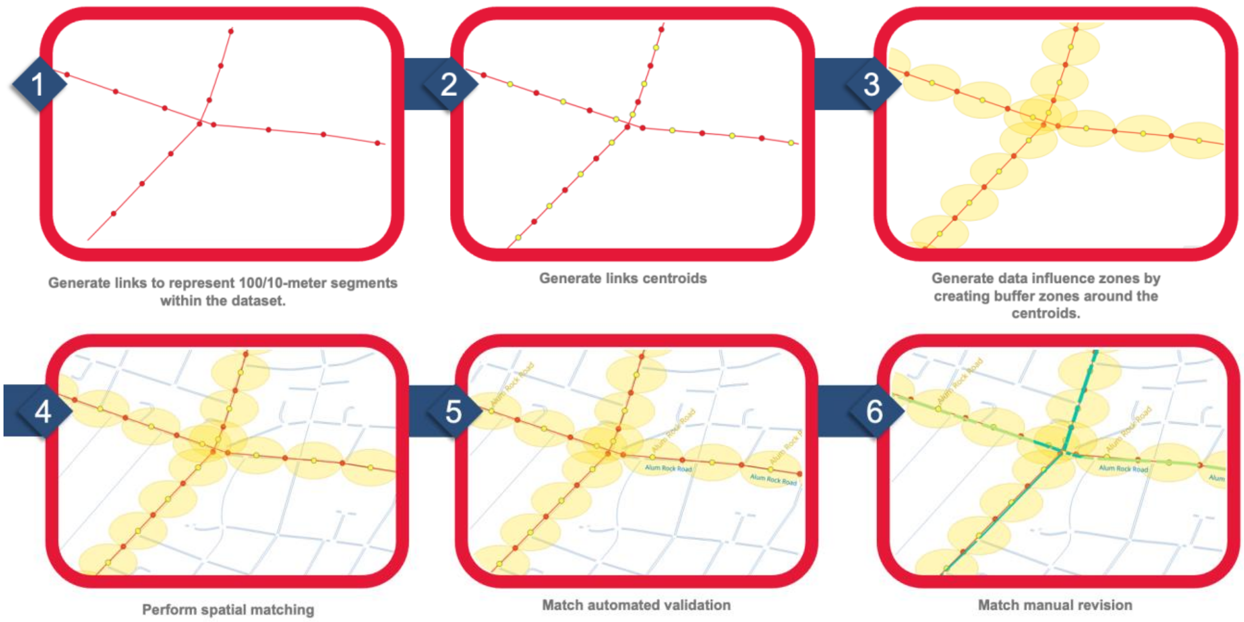


Figure 3.6: Data matching spatial procedure

(ii) How to calculate if an iRAP assessment needed to be performed.

In this case, the baseline assessment can be generated directly using the link-node configuration already established in Aimsun. The chainage in this scenario aligns with the start point of each segment in the Aimsun network, and the section identifiers in the iRAP database directly reference the corresponding Aimsun object IDs (OIDs). As a result, no spatial matching procedure is required.

3.5.3.3 Incorporation of behaviour calibration factors into risk assessment

Behavioural calibration factors support more accurate estimation of fatal and serious injuries (FSIs) by accounting for behavioural adaptations, which are particularly influential in after scenarios where interventions are in place. These calibration factors are applied as multiplicative adjustments to the FSI results.

Because the total number of fatalities in the baseline scenario must remain consistent with the observed number of fatalities across the network, the calibration factors act as redistributors rather than inflators of overall risk. They reallocate fatalities spatially, assigning greater weight to locations where non-compliant or higher-risk behaviours are more likely to occur.

Accordingly, calibration factors are constrained to values between 0 and 1, and their network-wide average is required to remain close to 1, ensuring consistency between modelled and observed fatality levels.

Depending on the type of behaviour under analysis, different approaches can be adopted to define the calibration factors. Certain non-compliant behaviours can be directly represented using operational parameters that are outputs of the traffic simulation. For example, speeding behaviour can be characterised using the standard deviation of vehicle speeds.

Other behaviours are more strongly dependent on the infrastructure characteristics of the network and cannot be directly inferred from simulation outputs alone. In these cases, additional estimations are required, based on roadway attributes and contextual factors (see Section 3.3.6).

3.6 PHOEBE performance indicators

*Content already presented in D3.2, but replicated in this deliverable to support standalone guidelines

The PHOEBE framework was designed to provide cities with a set of robust performance indicators, empowering them to measure and evaluate the effectiveness of their interventions with precision and objectivity.

A special emphasis is placed on vulnerable road user (VRU) indicators, such as monitoring speed and tracking the modal share of pedestrians and cyclists. The framework's KPIs are crafted to reflect modal split variations across different periods of the day, ensuring a nuanced understanding of mobility patterns and safety outcomes. This comprehensive approach allows urban planners and policymakers to make informed decisions, ultimately fostering safer, more sustainable, and more inclusive urban environments.

The list of performance indicators can be categorized into four groups:

- Health and Safety: 19 indicators illustrate how the intervention decreases risk and the incidence of fatal and serious injuries. These indicators also include modifications to infrastructure aimed at enhancing safety for all road users.
- Mode Shift: 24 indicators that illustrate how the interventions enhance the frequency of trips using more sustainable modes (non-motorised). These indicators take into account various age groups as well as the differences between men and women.
- Environment: 2 indicators based on the simulation calculations of pollutant emissions, calculated based on fleet vehicle types and individual vehicle speed profiles.
- Operations: 3 indicators that reflect the operational parameters of the network being studied include network travel time, travel distance, and delay.

3.6.1 Health and safety indicators

Table 6 outlines the health and safety indicators, their calculation methods, and the expected trends related to road safety.

Table 3.5: Health and Safety Indicators

ID	KPI Name	Calculation description	Desire trend
HS01	Total number of fatalities and severe injuries on the network	Sum of iRAP fatal and Serious injuries estimations per user group in a year	Smaller the better

ID	KPI Name	Calculation description	Desire trend
HS02	% of the network above road safety mark (3 stars or better) – Vehicle Occupant	Number of segments in the network with Vehicle Occupant Star Rating greater than three stars divided by the total number of segments	The greater the better
HS03	% of the network above road safety mark (3 stars or better) – Motorcyclist	Number of segments in the network with Motorcyclist Star Rating greater than three stars divided by the total number of segments	The greater the better
HS04	% of the network above road safety mark (3 stars or better) – Pedestrian	Number of segments in the network with Pedestrian Rating greater than three stars divided by the total number of segments	The greater the better
HS05	% of the network above road safety mark (3 stars or better) – Bicyclist	Number of segments in the network with Bicyclist Rating greater than three stars divided by the total number of segments	The greater the better
HS06	Risk variability within the simulated scenarios	Standard deviation of Star Rating Scores within the simulated hourly periods	Smaller the better
HS07	Average road safety risk on nodes - Vehicle Occupant	Average Vehicle Occupant Star Rating Score where the intersection is coded as present (<> None)	Smaller the better
HS08	Average road safety risk on nodes - Motorcyclist	Average Motorcyclist Star Rating Score where the intersection is coded as present (<> None)	Smaller the better
HS09	Average road safety risk on nodes - Pedestrian	Average Pedestrian Star Rating Score where the intersection is coded as present (<> None)	Smaller the better
HS10	Average road safety risk on nodes - Bicyclists	Average Bicyclist Star Rating Score where the intersection is coded as present (<> None)	Smaller the better
HS11	Average road safety risk on links/sections - Vehicle Occupant	Average Vehicle Occupant Star Rating Score where the intersection is coded as not present (= None)	Smaller the better
HS12	Average road safety risk on links/sections - Motorcyclist	Average Motorcyclist Star Rating Score, where the intersection is coded as not present (= None)	Smaller the better

ID	KPI Name	Calculation description	Desire trend
HS13	Average road safety risk on links/sections - Pedestrian	Average Pedestrian Star Rating Score where the intersection is coded as not present (= None)	Smaller the better
HS14	Average road safety risk on links/sections - Bicyclists	Average Bicyclist Star Rating Score, where the intersection is coded as not present (None)	Smaller the better
HS15	% of the road network where the operating speed (85th) is above the speed limit	Number of segments where Operating Speed (85th percentile) > Speed limit divided by the total number of segments	Smaller the better
HS16	% of road network with good quality sidewalk where operating speed (85th) is over 30km/h or 20 mph, and a pedestrian is walking along the flow	Number of segments where Operating Speed (85th percentile)> 30 km/h or 20mph, pedestrians along flows > 0, and good quality sidewalk coded (codes 1, 2,3, OR 4) divided by the number of segments where Operating Speed (85th percentile)> 30 km/h or 20mph	The greater the better
HS17	% of road network with crossing facility where operating speed (85th) is over 30km/h or 20 mph, and there is a pedestrian crossing flow	Number of segments where Operating Speed (85th percentile)> 30 km/h or 20mph, pedestrians crossing flows > 0, and pedestrian crossing is presented (<> No facility) divided by the number of segments where Operating Speed (85th percentile)> 30 km/h or 20mph	The greater the better
HS18	% of pedestrian crossings that are adequately signed and well-maintained	Number of segments where pedestrian crossing is presented (<> No facility) and crossing quality is good, divided by the number of segments where pedestrian crossing is presented	The greater the better
HS19	% of road network with bicycle facility where operating speed (85th) is over 30km/h and there is bicyclist flow	Number of segments where Operating Speed (85th percentile)> 30 km/h or 20mph, bicyclists flows > 0, and Facility for bicycle is presented (<> No facility) divided by the number of segments where Operating Speed (85th percentile)> 30 km/h or 20mph	The greater the better

3.6.2 Demand indicators

Table 3-6 outlines the demand indicators, their calculation methods, and the trends related to road safety.

Table 3.6: Demand Indicators

ID	KPI Name	Calculation description	Desire trend
MS01	Participation in non-motorised trips - Women - Age group 18-24	Total number of trips made by Women (18-24) with private micromobility, shared micromobility, and walking.	The greater the better
MS02	Participation in non-motorised trips - Women - Age group 25-34	Total number of trips made by Women (25-34) with private micromobility, shared micromobility, and walking.	The greater the better
MS03	Participation in non-motorised trips - Women - Age group 35-44	Total number of trips made by Women (35-44) with private micromobility, shared micromobility, and walking.	The greater the better
MS04	Participation in non-motorised trips - Women - Age group 45-54	Total number of trips made by Women (45-54) with private micromobility, shared micromobility, and walking.	The greater the better
MS05	Participation in non-motorised trips - Women - Age group 55-69	Total number of trips made by Women (55-69) with private micromobility, shared micromobility, and walking.	The greater the better
MS06	Participation in non-motorised trips - Women - Age group 70+	Total number of trips made by Women (70+) with private micromobility, shared micromobility, and walking.	The greater the better
MS07	Participation in motorised trips - Women - Age group 18-24	Total number of trips made by Women (18-24) with car, public transport, and motorcycle.	Smaller the better
MS08	Participation in motorised trips - Women - Age group 25-34	Total number of trips made by Women (25-34) with car, public transport, and motorcycle.	Smaller the better
MS09	Participation in motorised trips - Women - Age group 35-44	Total number of trips made by Women (35-44) with car, public transport, and motorcycle.	Smaller the better
MS10	Participation in motorised trips - Women - Age group 45-54	Total number of trips made by Women (45-54) with car, public transport, and motorcycle.	Smaller the better

ID	KPI Name	Calculation description	Desire trend
MS11	Participation in motorised trips - Women - Age group 55-69	Total number of trips made by Women (55-69) with car, public transport, and motorcycle.	Smaller the better
MS12	Participation in motorised trips - Women - Age group 70+	Total number of trips made by Women (70+) with car, public transport, and motorcycle.	Smaller the better
MS13	Participation in non-motorised trips - Men - Age group 18-24	Total number of trips made by Men (18-24) with private micromobility, shared micromobility, and walking.	The greater the better
MS14	Participation in non-motorised trips - Men - Age group 25-34	Total number of trips made by Men (25-34) with private micromobility, shared micromobility, and walking.	The greater the better
MS15	Participation in non-motorised trips - Men - Age group 35-44	Total number of trips made by Men (35-44) with private micromobility, shared micromobility, and walking.	The greater the better
MS16	Participation in non-motorised trips - Men - Age group 45-54	Total number of trips made by Men (45-54) with private micromobility, shared micromobility, and walking.	The greater the better
MS17	Participation in non-motorised trips - Men - Age group 55-69	Total number of trips made by Men (55-69) with private micromobility, shared micromobility, and walking.	The greater the better
MS18	Participation in non-motorised trips - Men - Age group 70+	Total number of trips made by Men (70+) with private micromobility, shared micromobility, and walking.	The greater the better
MS19	Participation in motorised trips - Men - Age group 18-24	Total number of trips made by Men (18-24) with a car, public transport, and motorcycle.	Smaller the better
MS20	Participation in motorised trips - Men - Age group 25-34	Total number of trips made by Men (25-34) with car, public transport, and motorcycle.	Smaller the better
MS21	Participation in motorised trips - Men - Age group 35-44	Total number of trips made by Men (35-44) with car, public transport, and motorcycle.	Smaller the better

ID	KPI Name	Calculation description	Desire trend
MS22	Participation in motorised trips - Men - Age group 45-54	Total number of trips made by Men (45-54) with a car, public transport, and motorcycle.	Smaller the better
MS23	Participation in motorised trips - Men - Age group 55-69	Total number of trips made by Men (55-69) with car, public transport, and motorcycle.	Smaller the better
MS24	Participation in motorised trips - Men - Age group 70+	Total number of trips made by Men (70+) with a car, public transport, and motorcycle.	Smaller the better

3.6.3 Operational indicators

Table 3.7 outlines the operational indicators, their calculation methods, and the trends related to road safety.

Table 3.7: Operational Indicators

ID	KPI Name	Calculation description	Desire trend
OP1	Network travel time	Addition of the time spent in the network of all vehicles and pedestrians, more info in the Aimsun manual: Calculate Traffic Statistics - Aimsun Next Users Manual (Aimsun, 2025).	Smaller the better
OP2	Network travel distance	Addition of the distance travelled through the network of all vehicles and pedestrians. More information can be found in the Aimsun manual:Calculate Traffic Statistics - Aimsun Next Users Manual (Aimsun, 2025)..	Smaller the better
OP3	Network delay	Average delay incurred by all vehicles in the network in seconds per km travelled and by vehicle.	Smaller the better* Since it does not weigh the vehicle delay per the vehicle occupants, in some instances, a larger total delay might represent a better situation, i.e, a more minor total delay per passenger.

3.6.4 Environmental indicators

Table 3.8 outlines the environmental indicators, their calculation methods, and the trends related to road safety.

Table 3.8: Environmental Indicators

ID	KPI Name	Calculation description	Desire trend
EN01	Pollutant Emissions – CO2	Based on fleet vehicle types and individual vehicle speed profiles as described in the Aimsun manual, the model used is the London Emission Model, Environmental Models - Aimsun Next Users Manual.	Smaller the better
EN02	Pollutant Emissions – NO2	Based on fleet vehicle types and individual vehicle speed profiles as described in the Aimsun manual, the model used is the London Emission Model, Environmental Models - Aimsun Next Users Manual.	Smaller the better

3.7 PHOEBE required skills and knowledge

Drawing on the experience from the three PHOEBE use cases and the framework validation activities, the following competencies are essential when building teams to replicate or prepare the application of the PHOEBE framework.

Traffic microsimulation expertise (Aimsun Next - core competency)

Table 3.9: Traffic microsimulation expertise (Aimsun Next - core competency)

Skill / Knowledge Area	Must-have	Nice-to-have
Aimsun Next microsimulation modelling	Network coding (links/nodes, lanes, speed limits), intersections, signal plans, PT lanes, VRUs	Experience with large-scale urban models and sub-network extraction
Traffic control & operations	Signal phasing/timing/coordination, priority rules, access restrictions	Advanced adaptive signal strategies

Demand handling in Aimsun Next	Import/use OD matrices by time period; static vs dynamic assignment	Advanced dynamic traffic assignment tuning
Scenario management	Baseline vs after scenarios, replication management	Automation of scenario batches
API module configuration	Activate or deactivate API modules through scenario configuration	Experience in developing Aimsun APIs
Calibration & validation	Calibration to flows, speeds, travel times, queues	Statistical goodness-of-fit metrics beyond standard practice
Multimodal modelling	Pedestrians, cyclists, micromobility representation	Complex mixed-traffic interaction tuning
Licensing & environment setup	Correct Aimsun Next licenses/modules available	License optimisation across teams

Behavioural Modelling (Human Behaviour)

Table 3.10: Behavioural Modelling (Human Behaviour)

Skill / Knowledge Area	Must-have	Nice-to-have
Behavioural theory	Understanding of compliance/non-compliance drivers	Advanced behavioural economics
Survey design (SP)	Design of stated-preference surveys, scenario framing	Adaptive or longitudinal survey methods

Statistical modelling	Discrete choice / ordinal models (e.g. CLMMs), interpretation	Game-theoretic or hybrid ML models
Data fusion	Combine survey outputs with telematics/video observations	Automated data fusion pipelines
Validation & calibration	Derive calibration factors from observed behaviour	Bayesian validation frameworks
Simulation translation	Convert behavioural probabilities into simulation-usable parameters	Dynamic behaviour adaptation
Simulation and Output and Trajectory handling	✓ Understanding SQLite data and *.trj outputs (trajectories)	High-performance trajectory processing
Surrogate safety concepts	✓ TTC, PET, conflict identification	Advanced interaction metrics
Post-processing	✓ Aggregate conflicts by location/time	Automated dashboards
Estimating Crashes from simulated Conflicts	✓ Crash-Conflict Relationship based on empirical data	Extreme Value Theory modelling
Statistical interpretation	✓ Confidence intervals, robustness checks	Bayesian EVT / advanced risk tails

Demand Modelling (Mode Choice & Induced Demand)

Table 3.11: Demand Modelling (Mode Choice & Induced Demand)

Skill / Knowledge Area	Must-have	Nice-to-have

Mode choice modelling	Detailed knowledge of discrete choice modelling, model specification for at least Multinomial Logit Model	Advanced model specifications, e.g., mixed logit formulations.
Induced demand concepts	Detailed understanding of behavioural response under changing predictors, e.g., cost, time, and risk , detailed knowledge of discrete choice modelling.	Advanced model specification extended to more modes of transport (as considered in mode choice model)
R & Apollo package OR Python & Biogeme	Model estimation, diagnostics, calibration	Advanced scripting & reproducibility
Survey-based data handling	SP experiment design awareness, survey preparation skills	D-efficient design of SP surveys and optimisation
OD matrix handling	Computing mode wise segmented OD demand matrices from the developed mode choice models, carrying out the computation under changing values of predictors , good coding skills with Python	Fully automated OD segmentation
Integration loop	Feedback between microsimulation, risk and demand models	Sensitivity & uncertainty analysis

Road Safety & Risk Assessment (Star ratings)

Table 3.12: Road Safety & Risk Assessment (Star ratings)

Skill / Knowledge Area	Must-have	Nice-to-have
iRAP methodology	Star Ratings, FSI estimation, user-group differentiation	SRIP development
Road attribute data	Road Attribute Coding standards, chainage logic, QA processes	Automated / AI-assisted coding

ViDA platform	Upload, Supporting data specifications, validation, execution of models	Advanced visualisation & reporting
Simulation–risk integration	Mapping simulation outputs (speed/flow) into risk models	Knowledge on simulation outputs
Behavioural calibration factors	Apply CFs without inflating total fatalities	Advanced redistribution logic
Interpretation for policy	Translate outputs into decision-relevant KPIs	Scenario prioritisation frameworks

Integration, Automation and data flows

Table 3.13: Road Safety & Risk Assessment (Star ratings)

Skill / Knowledge Area	Must-have	Nice-to-have
Aimsun Next API	Understanding of API lifecycle and runtime hooks for behaviour injection	Performance optimisation of API callbacks
Python scripting	pandas/numpy, file I/O, matrix transformations, automation	Modular, reusable script architecture
Non-compliant behaviour integration	Implement/activate APIs for speeding, red-light violations, illegal crossings	Extending APIs to new behaviour types
OD & skim matrix workflows	Export/import segmented OD matrices and skims; shape consistency	Advanced matrix QA and diagnostics
Post-processing outputs	Extract SQLite outputs, SSAM trajectories	Automated KPI pipelines

Data debugging	Diagnose matrix mismatch, missing IDs, failed runs	Robust logging and error-handling frameworks

Beyond the technical skills, someone in the team needs to create a PHOEBE framework understanding (system-level literacy). This requires a clear grasp of both the purpose of the framework and how its components interact, including an end-to-end understanding of the PHOEBE workflow that spans baseline and after scenarios, the setup–simulation–refinement cycle, and the iterative exchange of information between traffic simulation, behavioural modelling, demand modelling, and road safety assessment. This also involves knowing how component interfaces and data dependencies are structured—specifically which inputs and outputs are required by each model. Finally, it requires familiarity with basic principles of experiment design and reproducibility, such as consistent scenario definition, use of replications, and controlled comparisons between baseline and intervention runs, in line with PHOEBE's iterative simulation and evaluation approach.

4 Overall Socio-Economic Impact Assessments

4.1 Objectives

The overall socio-economic impact assessment of the safety interventions implemented in the project PHOEBE answers the following questions: whether these interventions make streets safer, who benefits from them, who bears the costs, and whether the benefits of implementing the interventions are significant enough to justify the effort. Project PHOEBE does not treat the interventions under study as isolated engineering choices; instead, they are seen as changes to a living transport system, where people shift modes, reroute, change trip timing, or even decide to travel more or less in a given timeframe. Therefore, the aim is to capture the behavioural responses of road users, including their mode choices, induced demand, and behavioural shifts, rather than assuming a fixed travel demand when implementing the interventions. To do this, PHOEBE employs stated preferences from use-case-specific surveys to estimate mode choice, induced demand, and behavioural shifts in response to changing attributes, e.g., travel costs, time, associated risks, and availability. It also uses revealed preferences and historical data, e.g., crash statistics, to assess the impacts of these interventions on road safety outcomes through iRAP methodology and micro-simulation models. In this way, PHOEBE links behaviour and safety outcomes consistently and comparably across the use cases.

However, implementing and operating such interventions is often costly. They require an extended construction and adoption timeline to break even on investment costs. In PHOEBE, therefore, a detailed Cost-Benefit Analysis (CBA) and Cost-Effectiveness Analysis (CEA) are carried out to answer these questions. The factors examined under the CBA and CEA frameworks are:

- **Public health benefits:** Aggregation and generation of benefits for public health from the implementation of interventions. As roads become safer, road users increasingly use active modes, e.g., biking and walking. This has a direct impact on the public health of the use case. In PHOEBE, the effects of these interventions are analysed for three lifestyle diseases in terms of cases and deaths saved. These diseases are Type 2 diabetes (T2D), Ischaemic/Coronary heart disease (IHD/CHD), and Ischaemic Stroke (IS). In the real world, the adoption of active modes may affect many more lifestyle diseases; however, only these three diseases are analysed. Such analysis is already present, e.g., the Integrated Transport and Health Impact Model (ITHIM). However, ITHIM does not explicitly include its own transport, behavioural and risk assessment models. Such studies require predefined travel scenarios as inputs. The framework provided by PHOEBE includes transport, behavioural, and risk assessment models, thereby providing more realistic changes in travel behaviour upon the adoption of the interventions, which, in turn, pushes the boundary beyond the state of the art.
- **FSI savings benefits:** One of the project PHOEBE's core objectives is to compute the changes in FSI following the implementation of the interventions. The outputs of this step are directly used to generate the benefits, which are then used in the socio-economic analysis.
- **Emission benefits:** The implemented interventions increase active mode use, reducing emissions. These reductions are considered as benefits produced by the interventions' implementation and are directly used to generate the emission benefits for the socio-economic analysis.
- **Operational benefits:** The interventions also yield savings in travel time and delays across the modelled network. These savings are also taken into account in the socio-economic analysis.

Apart from these benefits, the cost side of the intervention implementation is also modelled. Within the socio-economic analysis framework, a sub-framework is developed to scale the benefits across the whole network, which is not part of the modelled area.

4.2 Methodological framework of the overall socio-economic analysis

The overall methodological framework of the socio-economic analysis is given below in Figure 4.1.

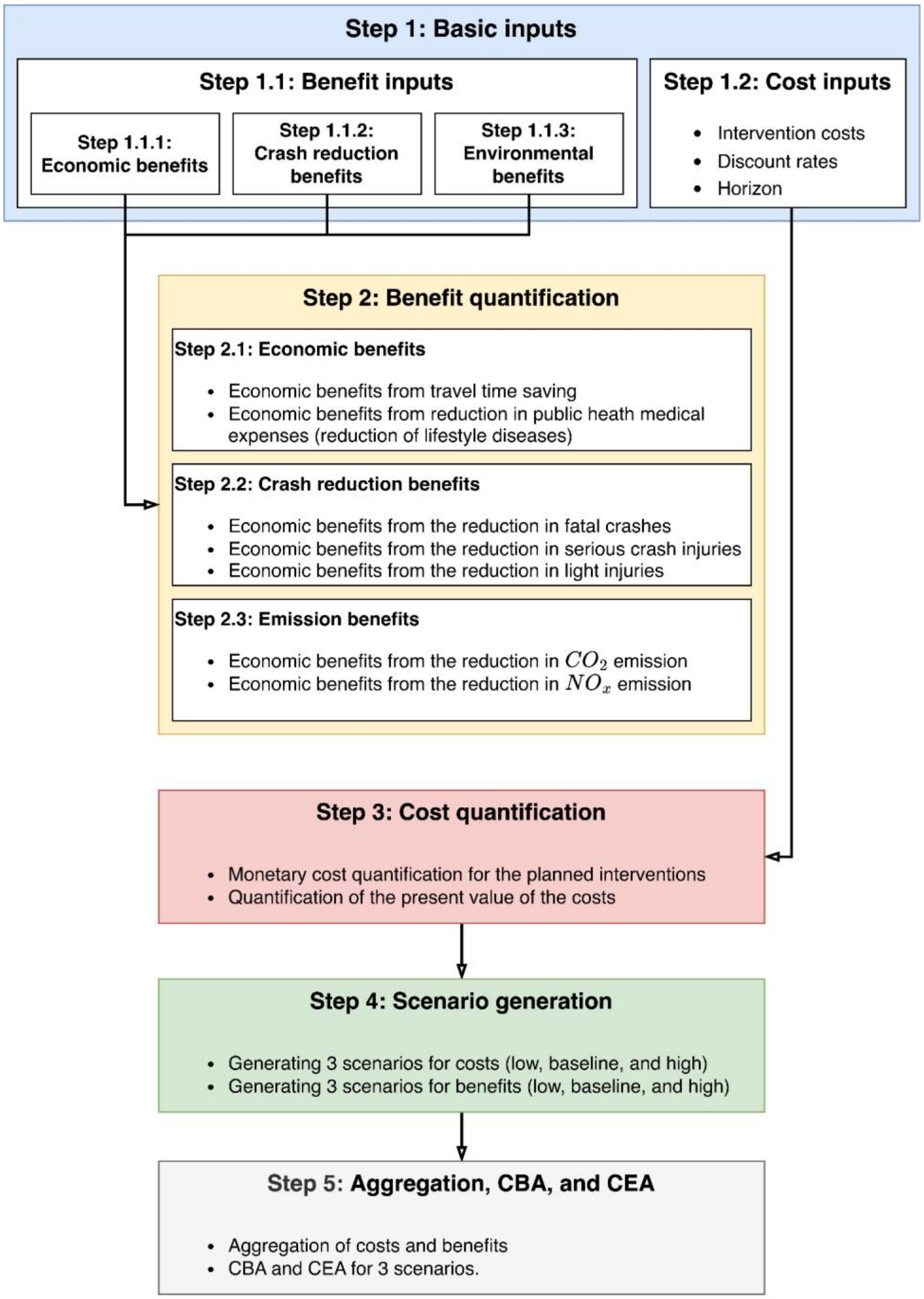


Figure 4.1: Methodological framework of the socio-economic analysis

Step 1: Basic inputs

This step gathers the minimum information needed to run a consistent appraisal across all the use cases. The Step 1.1.1, 1.1.2, and 1.1.3 draw information from the main PHOEBE framework in the form of KPIs. The benefit side inputs are explained in Section 4.1 in bullet format. For more information the reader is referred to Figure 2-1 The PHOEBE Framework flowchart (Deliverable D1.2). In Step 1.2, the basic cost side data and information are gathered, e.g., the accounting frame, the time horizon, and the baseline year.

Step 2: Benefit quantification

This step turns the output KPIs from the PHOEBE framework into yearly economic values. For each scenario, benefits are computed as scenario minus baseline for desirable outputs, with consistent sign conventions. In the section following, the detailed mathematical descriptions of the benefit quantification is presented.

Economic benefits:

The economic benefits comprise of the following:

- **Time saving:** Economic benefits from the operational KPIs. These inputs predominantly consist of the network travel time savings (with proper sign conventions) after the intervention is introduced. A monetary value of these savings can be generated by multiplying the saving values with the Value of Time (VoT) of the use cases. In case the VoT is older, an inflation adjusted VoT is used.

$$B_{Time} = \sum_{i=1}^{n} \Delta TT_{Mode(i),OD(i)} . VoT_{InflationAdjusted} \quad \text{(Eq.5)}$$

  Where, B_{Time} is total benefits (monetary) generated from the difference of $TT_{Baseline}$ and $TT_{Scenario}$, and $VoT_{InflationAdjusted}$ is the inflation adjusted value of time.

- **Public health savings from uptake of active modes:** The uptake of active modes result in more physical mobility among the road users. This enhanced active mobility tends to make the road users fit by keeping them active, thereby cutting on public medical expenses on the lifestyle and other diseases. In this section, three diseases were quantified, which are Type 2 Diabetes Mellitus (T2D), Ischaemic/Coronary Heart Diseases (IHD/CHD) and Ischaemic Stroke (IS). The calculations take into account the extra physical effort in terms of Metabolic Equivalent of Task (MET) before and after the implementation of the interventions. With the difference of the MET values and ITHIM standards, the reduction in deaths and cases of these diseases are calculated. The MET values are procured from Compendium of Physical Activities (CPA). These prevented deaths are then monetised with the help of Value of Life. The prevented cases are then used to calculate the Quality Adjusted Life Years (QALYs) saved with the help of Disability Weights (DWs) and the local Willingness To Pay (WTP) per QALYs. The DWs can be procured from IHME global disability weights database.

$$\Delta MET = \sum_{i=1}^{n} \Delta Trips.Min_{Mode(i),OD(i)} . MET_{CPA,mode(i)} \quad \text{(Eq.6)}$$

  Where, $\Delta Trips.Min_{Mode(i),OD(i)}$ is the difference in Trip-Minutes over the whole network for all considered modes (in this case the active modes), and the $MET_{CPA,mode(i)}$ is the standard MET values per mode.

  The difference in MET values are then passed through DOSE equations to generate the cases and deaths saved.

Crash reduction benefits:

The FSI reduction KPIs are procured from the Health and Safety indicators. These are the outputs of the iRAP road safety assessment models. These values are multiplied with VSL values (Van Essen et al., 2019) in order to generate the monetary benefits. Further, with the help of Disability Weights for Fatal, Serious and Light injuries, (Van Ditshuizen et al., 2021) QALYs are calculated. The QALYs are then multiplied with the WTP/QALYs to generate the monetised benefits of the prevented crashes.

$$B_{FSI} = \sum_{i=1}^{n} \Delta I_{Injurytype(i)} . MonetaryValue_{Injurytype(i)} \quad \text{(Eq.7)}$$

Where, $\Delta I_{Injurytype(i)}$ is the change in injuries according to their types (FSI).

Emission benefits:

The network wide change in emissions are procured from the Environmental Indicators. These KPIs provide CO_2 and NO_X emissions before and after implementation of the interventions. In order to monetise these benefits, the weights of the emissions are multiplied with the cost of CO_2/ton and NO_X/ton equivalent.

$$B_{ENV} = \Delta Emission[Ton]. Cost_{SocialCost}[€/ton] \quad \text{(Eq.8)}$$

After quantification of all the benefits, the aggregated monetary values of the benefits are spread over the analysis horizon (2022 to 2037). This is followed by the present value calculation of the benefits against the base year of 2020.

Step 3: Cost quantification

The approximate real costs of implementing the interventions are estimated with respect to a base year (2020). The costs are categorised under Capital Expenditure (CapEx), operational expenditure (OpEx), Revenues and Grants (one-time). An example from the Athens use case is given below (Figure 4.2).

Interventi	Category Level1 (CAPEX/OPEX)	Category Level2 (CAPEX/OPEX)	Economic life years	Cost real (Base year: 2020) in EUR
Intervention 1	CAPEX	Design and surveys	25	2,800,000.00
Intervention 1	CAPEX	Civil works	15	350,000.00
Intervention 1	CAPEX	Drainage and stormwater	20	220,000.00
Intervention 1	CAPEX	Lighting fixtures	15	480,000.00
Intervention 1	CAPEX	Utility diversions	30	650,000.00
Intervention 1	CAPEX	Traffic management & signage	10	320,000.00
Intervention 1	CAPEX	Routine maintenance	15	100,000.00
Intervention 1	CAPEX	Urban furniture & landscaping	15	180,000.00
Intervention 1	CAPEX	Traffic signals (works)	15	420,000.00
Intervention 1	CAPEX	Traffic management & control systems (initial setup)	10	280,000.00
Intervention 1	CAPEX	Monitoring cost – initial deployment	10	120,000.00
Intervention 1	CAPEX	Commissioning, training & handover	8	180,000.00
Intervention 1	OPEX/annum	Routine maintenance		55,000.00
Intervention 1	OPEX/annum	Winter services & cleaning		40,000.00
Intervention 1	OPEX/annum	Staffing & operations		25,000.00
Intervention 1	OPEX/annum	Monitoring and evaluations		10,000.00
Intervention 1	OPEX/annum	Software		6,000.00
Intervention 1	OPEX/annum	Energy		85,000.00
Intervention 1	Offsets/Revenues	Grant/External funding		1,200,000.00
Intervention 1	Offsets/Revenues	Avoided road maintenance	15	120,000.00

Figure 4.2: Cost quantification of the interventions (Athens).

The costs are then aggregated yearwise within the analysis horizon (2022 to 2037). This is followed by the calculation of the residual values according to the economic life years of the CapEx. Finally, using the discount rate, present values of these costs are quantified against the base year.

Step 4: Scenario generation

In this step, 3 scenarios were generated in order to counter the future uncertainties. The scenario analysis is carried out keeping the following points in focus:

- Handling future uncertainties
- To test the robustness of the conclusions
- To represent alternative plausible futures
- To represent the results in ranges instead of single point estimates
- To support decision-making under risk

The example of the scenario parameters are shown in the figure below.

Parameter	Scenario (Low)	Scenario (Base)	Scenario (High)	Notes
CAPEX multiplier	0.900	1.000	1.200	Multiplier for CAPEX (for scenarios); "+" is increased and "-" is decreased
OPEX multiplier (Excl. Energy)	0.900	1.000	1.150	Multiplier for OPEX (for scenarios); "+" is increased and "-" is decreased
Energy Tariff (EUR/kWh)	0.18	0.200	0.22	Energy tariff in the use case of Athens
Energy Tarif multiplier	0.900	1.000	1.100	
Utility diversion multiplier	1.000	1.200	1.500	Multiplier for moving, protecting, or modifying existing underground/overhead services/infrastructure (utilities)
Schedule slip months	0.000	6.000	12.000	Offset months the project becomes operational from the actual deadline
Discount rate	0.04	0.04	0.040	Rate for discounting the future expenditure (the base scenario is taken from assumption, fill in the rest)
Contingency rate (C)	0.100	0.150	0.200	A buffer added to the cost estimate
Optimism bias rate (OB)	0.100	0.150	0.200	Extra uplift to counter optimistic forecast of costs and timelines

Figure 4.3: Scenario parameters.

Step 5: Aggregation, CBA and CEA

After the preparation of the monetary costs and benefits, the cost benefit analysis (CBA) and cost effectiveness analysis (CEA) is carried out.

- **CBA**

$$NPV = \sum_{t=0}^{T} \frac{B_t - C_t}{(1+r)^t} \qquad \text{(Eq.9)}$$

$$BCR = \frac{\sum_t \frac{B_t}{(1+r)^t}}{\sum_t \frac{C_t}{(1+r)^t}} \qquad \text{(Eq.10)}$$

$$IRR;\ \sum_{t=0}^{T} \frac{B_t - C_t}{(1+IRR)^t} \qquad \text{(Eq.11)}$$

Where, NPV is the net present, BCR is benefit to cost ratio and IIR is the internal rate of return.

- **CEA**

In order to calculate the CEA, the QALYs are taken into account. The present value of the QALYs from Public health benefits and crash reduction benefits are calculated by multiplying them with the WTP/QALYs. Following this the CEA is calculated according to the methodology listed in Putatunda et al. (2025). The methodology is listed below.

$$ICER = \frac{ADC}{ADE} \quad \text{(Eq.12)}$$

Where, ADC is the Average Difference of Cost, ADE is the Average Difference of the Unit of Effectiveness, and $ICER$ is the Incremental Cost Effectiveness Ratio. Here, it is calculated as:

$$ICER = \frac{Costs}{PV(QALY_{CrashSavings} + QALY_{PublicHealthBenefits})} \quad \text{(Eq.13)}$$

Finally, the ADC and ADE are interpreted according to the following matrix.

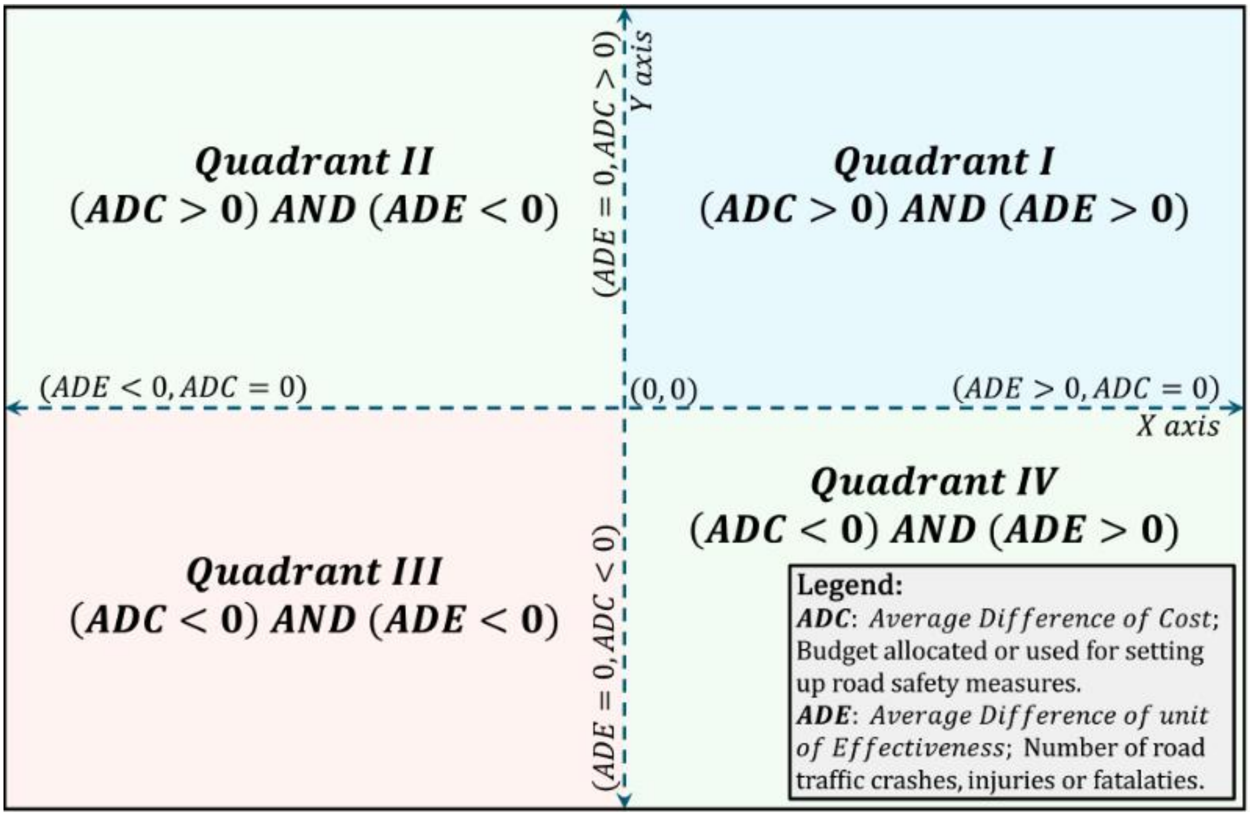


Figure 4.4: CEA ADC, ADE quadrant.

Where,

- **Quadrant I:** The planned *Road Safety Measure* is not effective, since the cost is high and the *Road Traffic Crashes*, *Road Traffic Injuries*, and fatalities remain high.
- **Quadrant II:** The planned *Road Safety Measure* is effective however the costs can still be lowered.
- **Quadrant III:** The expected goal.
- **Quadrant IV:** Outcome not expected.

4.3 Cost-Benefit Analysis (CBA)

In this section the detailed results of the CBA are presented. The section starts with the presentation of the benefits categorised under types, i.e., benefits from time savings, benefits from savings in public health, benefits from saving FSIs, and benefits from saving on emissions, all due to the implementation of the interventions. The benefits presented in this section are in monetary terms generated per year. The positive sign shows that the benefits generated are beneficial, while the negative sign shows the opposite. The generated benefits are within the modelled area of PHOEBE.

Benefits from time savings:

These benefits come from traffic microsimulation and road safety assessment models. These benefits are then directly used in socio-economic analysis.

Table 4.1: Benefits from travel time savings

Item	Scenario and use case	Benefits
Benefit from travel time saving	S1: Athens	€ 2,489,451
Benefit from travel time saving	S2: Athens	€ 154,264
Benefit from travel time saving	S1: Valencia	-€ 381,330
Benefit from travel time saving	S2: Valencia	€ 88,173.92
Benefit from travel time saving	S3: Valencia	€ 58,487.22
Benefit from travel time saving	S4: Valencia	-€ 828,292.02
Benefit from travel time saving	S1: West Midlands (Edgbaston)	£ 364,219
Benefit from travel time saving	S2: West Midlands (A45)	-£ 479,310
Benefit from travel time saving	S3: West Midlands (A45)	Not available[1]
Benefit from travel time saving	S4: West Midlands (A45)	Not Available[2]

Benefits from savings in public health:

Here, the total monetary benefits of preventing deaths and extra cases of T2D, IS and IHD/CHD are presented.

Table 4.2: Benefits from preventing T2D, IS, and IHD/CHD deaths and cases

Item	Scenario and use case	Benefits
Benefits	S1: Athens	€ 222,947
Benefits	S2: Athens	€ 450,376

[1] The travel time saving is not available from the other components of the PHOEBE framework.
[2] The travel time saving is not available from the other components of the PHOEBE framework.

Benefits	S1: Valencia	-€ 16,098
Benefits	S2: Valencia	-€ 17,076
Benefits	S3: Valencia	€ 405,782
Benefits	S4: Valencia	-€ 2,093
Benefits	S1: West Midlands (Edgbaston)	£ 66,768
Benefits	S2: West Midlands (A45)	£ 5,947,934
Benefits	S3: West Midlands (A45)	£ 9,860,720
Benefits	S4: West Midlands (A45)	-£ 5,269,134

Benefits from saving FSIs:

Here, the total monetary benefits of preventing FSIs are presented. These benefits come from road safety assessment models. These benefits are then directly used in socio-economic analysis.

Table 4.3: Benefits from FSI prevention.

Item	Scenario and use case	Benefits
Benefits	S1: Athens	€ 3,922,881
Benefits	S2: Athens	€ 12,993,887
Benefits	S1: Valencia	€ 24,829,649
Benefits	S2: Valencia	-€ 105,265,170
Benefits	S3: Valencia	-€ 7,726,249
Benefits	S4: Valencia	€ 32,159,137
Benefits	S1: West Midlands (Edgbaston)	£ 252,623
Benefits	S2: West Midlands (A45)	£ 44,695,908
Benefits	S3: West Midlands (A45)	£ 44,695,908
Benefits	S4: West Midlands (A45)	£ 44,695,908

Benefits from saving on emissions:

These benefits come from traffic microsimulation and road safety assessment models. These benefits are then directly used in socio-economic analysis.

Table 4.4: Benefits from prevented emissions

Item	Scenario and use case	Benefits
CO_2 **emission benefits/year**	S1: Athens	€ 195,178
NO_X **emission benefits/year**	S1: Athens	€ 5,888
CO_2 **emission benefits/year**	S2: Athens	€ 49,147
NO_X **emission benefits/year**	S2: Athens	-€ 6,146
CO_2 **emission benefits/year**	S1: Valencia	-€ 2,754
NO_X **emission benefits/year**	S1: Valencia	-€ 113
CO_2 **emission benefits/year**	S2: Valencia	-€ 1,502
NO_X **emission benefits/year**	S2: Valencia	-€ 66
CO_2 **emission benefits/year**	S3: Valencia	€ 598
NO_X **emission benefits/year**	S3: Valencia	€ 25
CO_2 **emission benefits/year**	S4: Valencia	-€ 7,269
NO_X **emission benefits/year**	S4: Valencia	-€ 282
CO_2 **emission benefits/year**	S1: West Midlands (Edgbaston)	£ 9,552
NO_X **emission benefits/year**	S1: West Midlands (Edgbaston)	£ 428
CO_2 **emission benefits/year**	S2: West Midlands (A45)	-£ 31,081
NO_X **emission benefits/year**	S2: West Midlands (A45)	-£ 285
CO_2 **emission benefits/year**	S3: West Midlands (A45)	£ 807,948[3]
NO_X **emission benefits/year**	S3: West Midlands (A45)	£ 140,240[4]
CO_2 **emission benefits/year**	S4: West Midlands (A45)	£ 242,307[5]
NO_X **emission benefits/year**	S4: West Midlands (A45)	-£ 2,268[6]

These benefits are yearly benefits in monetary terms, which are used to calculate the present value of the benefits against the base year.

The basic assumptions of the CBA are presented below:

[3] Calculated from modal shift outputs since the inputs from traffic microsimulation were not available.
[4] Calculated from modal shift outputs since the inputs from traffic microsimulation were not available.
[5] Calculated from modal shift outputs since the inputs from traffic microsimulation were not available.
[6] Calculated from modal shift outputs since the inputs from traffic microsimulation were not available.

Athens

Field	Value	Description
Use case	Athens	The use case under consideration
Base price year	2020	Reference year for real euros
Currency	EUR	Currency in the use case
Analysis horizon start	2022	First appraisal year
Analysis horizon end	2037	Last appraisal year
Discount rate	0.04	Rate for discounting the future expenditure
Contingency rate (c)	0.150	A buffer added to the cost estimate
Optimism bias rate (ob)	0.150	Extra uplift to counter optimistic forecast of costs and timelines
Comissioning year	2021	First service year where OPEX begin
CPI/HICP: 2020	0.9932786	
Construction Index	1	
Energy Index	0.040	
Wage Index	1.209	
Intervention 1	Increase of sidewalks in streets with high pedestrian traffic	Panepistimiou Street (from Vas. Sofias Av to Omonoia Sq.): Increase of sidewalks and decrease of traffic lanes to 4 (from 5 lanes)
Intervention 2	Increase of sidewalks in streets with high pedestrian traffic	Syntagma Square (from Karageorgi Servias to Mitropoleos): From 28/6/20: Increase of sidewalks with 4 traffic lanes and 1 traffic lane for PT (from Ermou St. to Mitropoleos St.)
Intervention 3	Exclusive lanes for pedestrians and cyclists	Vasilissis Olgas Avenue: Streets free of private vehicles. The entire 700-meter length of Vasilissis Olgas Avenue, between Vasilissis Konstantinou Avenue and Vasilissis Amalias Avenue, has been pedestrianized, and motorised traffic has been prohibite
Intervention 4	Exclusive bus lanes	Panepistimiou Street: Implementation of new exclusive bus lane (removing existing contra-flow bus lane to a new parallel-flow bus lane). Between Vasilissis Sofias Avenue and Omonoia Square, a new exclusive bus lane has been implemented along the parallel flow direction, spanning 1050 meters.
Intervention 5	Motorcycle parking management	Ermou Street (from Monastiraki to Asomaton Sq.):Parking arrangements. An adequate number of parking spaces for motorized two-wheelers has been established along Ermou Street

Figure 4.5: Assumption for CBA and CEA for Athens

Valencia:

Figure 4.6: Assumption for CBA and CEA for Valencia

Field	Value	Description
Use case	Valencia	The use case under consideration
Base price year	2020	Reference year for real euros
Currency	EUR	Currency in the use case
Analysis horizon start	2022	First appraisal year
Analysis horizon end	2037	Last appraisal year
Discount rate	0.035	Rate for discounting the future expenditure
Contingency rate (c)	0.150	A buffer added to the cost estimate
Optimism bias rate (ob)	0.150	Extra uplift to counter optimistic forecast of costs and timelines
Comissioning year	2021	First service year where OPEX begin
CPI/HICP: 2020	105.2767675	
Construction Index	1	
Energy Index	0.040	
Wage Index	1.209	
Intervention 1	Bike lane	Avenida Hermanos Machado + Avenida de los Naranjos
Intervention 2	Speed limit	Speed limit reduced to 30 km/h on motor lanes adjacent to bike lanes on: - Calle del Doctor Manuel Candela (green corridor) - Avenida/Primado Reig (blue corridor)
Intervention 3	Public e-scooter system	Extra push in shared e-scooters and related infrastructure
Intervention 4	New parking	Avenida dels Tarongers (segment Avenida de la Malva-rosa)

West Midlands

Figure 4.7: Assumption for CBA and CEA for the West Midlands

Field	Value	Description
Use case	West Midlands	The use case under consideration
Base price year	2020	Reference year for real euros
Currency	GBP	Currency in the use case
Analysis horizon start	2022	First appraisal year
Analysis horizon end	2037	Last appraisal year
Discount rate	0.035	Rate for discounting the future expenditure
Contingency rate (c)	0.150	A buffer added to the cost estimate
Optimism bias rate (ob)	0.150	Extra uplift to counter optimistic forecast of costs and timelines
Comissioning year	2021	First service year where OPEX begin
CPI/HICP: 2020	108.3000000	Base year 2015 (100)
Construction Index		
Energy Index		
Wage Index		
Intervention 1	A38 cycle lane	
Intervention 2	30mph speed limit in A45 network	
Intervention 3	Chnges in public bus services in A45 network	
Intervention 4	Implementing Automated Vehicles (A45 network)	

Based on these assumptions, the CBA was carried out. The CBA is carried under 3 scenarios for each use case (Figure 4.5 in Step 4: Scenario generation). The results of the CBA are presented below:

Athens

Table 4.5: CBA results for Athens

CBA		SCN Low		SCN Base		SCN High	
		S1	S2	S1	S2	S1	S2
Benefits	PV (Societal)	€ 74,851,819	€ 149,362,494	€ 80,999,463	€ 161,629,764	€ 84,254,809	€ 168,125,619
Costs	PV (Societal)	€ 15,987,674		€ 20,684,805		€ 28,539,975	
BCR		4.68	9.34	3.92	7.81	2.95	5.89
NPV		€ 58,864,144	€ 133,374,819	€ 60,314,659	€ 140,944,960	€ 63,570,004	€ 139,585,643

IRR		0.41	0.85	0.37	0.78	0.33	0.50

Valencia

Table 4.6: CBA results (1) for Valencia

CBA		**SCN Low**		**SCN Base**		**SCN High**	
		S1	**S2**	**S1**	**S2**	**S1**	**S2**
Benefits	PV (Societal)	€ 264,206,775	-€ 1,153,588,712	€ 285,906,305	-€ 1,248,333,948	€ 297,396,798	-€ 1,298,504,135
Costs	PV (Societal)	€ 61,707,659	€ 11,516,515	€ 73,983,215	€ 13,631,161	€ 91,766,605	€ 16,688,359
BCR		4.28	-100[7]	3.86	- 92[8]	3.24	- 78[9]
NPV		€ 211,948,216	-€ 1,211,392,348	€ 222,199,203	-€ 1,313,265,099	€ 234,775,287	-€ 1,368,175,232
IRR		0.60	Not defined	0.54	Not defined	0.44	Not defined

Table 4.7: CBA results (2) for Valencia

CBA		**SCN Low**		**SCN Base**		**SCN High**	
		S3	**S4**	**S3**	**S4**	**S3**	**S4**
Benefits	PV (Societal)	-€ 78,794,330	€ 334,318,722	-€ 85,265,777	€ 361,776,608	-€ 88,692,584	€ 376,316,307

[7] The high negative value of BCR is due to very high FSI number wrt. Baseline scenario coming from road safety assessment model.
[8] The high negative value of BCR is due to very high FSI number wrt. Baseline scenario coming from road safety assessment model.
[9] The high negative value of BCR is due to very high FSI number wrt. Baseline scenario coming from road safety assessment model.

Costs	PV (Societal)	€ 24,545,879	€ 122,646,193	€ 30,678,879	€ 145,362,722	€ 38,820,161	€ 175,131,775
BCR		-3.21	2.73	-2.78	2.49	-2.28	2.15
NPV		-€ 107,077,906	€ 221,218,745	-€ 120,227,811	€ 226,550,517	-€ 123,978,374	€ 242,463,888
IRR		Not defined	0.72	Not defined	0.63	Not defined	0.48

West Midlands

Table 4.8: CBA results (1) for the West Midlands

CBA		**SCN Low**		**SCN Base**		**SCN High**	
		S1	**S2**	**S1**	**S2**	**S1**	**S2**
Benefits	PV (Societal)	£ 7,594,177	£ 548,913,142	£ 8,217,894	£ 593,995,852	£ 8,548,168	£ 617,868,377
Costs	PV (Societal)	£ 21,832,240	£ 1,348,003	£ 27,129,992	£ 1,597,492	£ 36,275,140	£ 1,965,942
BCR		0.35	407.20[10]	0.30	371.83[11]	0.24	314.29[12]
NPV		-£ 14,207,646	£ 569,452,374	-£ 18,949,417	£ 616,637,088	-£ 18,587,938	£ 642,765,025
IRR		-0.21	63.10	Not defined	57.85	-0.21	89.9

Table 4.9: CBA results (2) for the West Midlands

[10] The high value of BCR is due to the nature of intervention (speed limit) where the implementation costs are low, and the high benefits from FSIs coming from road safety assessment models.
[11] The high value of BCR is due to the nature of intervention (speed limit) where the implementation costs are low, and the high benefits from FSIs coming from road safety assessment models.
[12] The high value of BCR is due to the nature of intervention (speed limit) where the implementation costs are low, and the high benefits from FSIs coming from road safety assessment models.

CBA		SCN Low		SCN Base		SCN High	
		S3	S4	S3	S4	S3	S4
Benefits	PV (Societal)	£ 607,727,907	£ 434,315,987	£ 657,641,124	£ 469,986,734	£ 684,071,534[13]	£ 488,875,367[14]
Costs	PV (Societal)	£ 89,614,312	£ 134,434,312	£ 106,149,232	£ 159,414,482	£ 128,860,164	£ 197,239,395
BCR		6.78	3.23	6.20	2.95	5.31	2.48
NPV		£ 540,092,412	£ 314,937,914.32	£ 575,564,349	£ 326,894,533	£ 604,491,834	£ 347,567,715
IRR		1.39	0.40	1.26	0.36	0.84	0.32

4.4 Cost-Effectiveness Analysis (CEA)

A detailed CEA is also carried out for all the use cases. For the effectiveness metrics, the QALYs loss saved due to the implementation is utilised (ADE), whereas the costs of implementation of the interventions are used for the cost metrics. In case, QALYs losses are saved, the ADE takes on negative sign, whereas it takes positive sign in case of the opposite situation. The results are presented below:

Athens:

Table 4.10: CEA results for Athens

CEA Scenario		PV (QALYs) (ADE)	PV Cost (ADC)	ICER
SCN Low	**S1**	-€ 56,376,892	€ 15,987,674	-0.28
	S2	-€ 185,923,888		-0.09
SCN Base	**S1**	-€ 61,007,175	€ 20,684,804	-0.34

[13] The benefit calculations are done without the time savings.
[14] The benefit calculations are done without the time savings.

	S2	-€ 201,193,977		-0.10
SCN High	**S1**	-€ 63,459,036	€ 28,539,975	-0.45
	S2	-€ 209,279,906		-0.14

Here, since, ADC>0 and ADE<0, therefore the outcome falls in Quadrant II: “The planned *Road Safety Measure* is effective however the costs can still be lowered.” The efficacy of the intervention can be determined by an ICER converging to zero.

Valencia:

Table 4.11: CEA results for Valencia

CEA Scenario		**PV (QALYs) (ADE)**	**PV Cost (ADC)**	**ICER**
SCN Low	**S1**	-€ 657,264,655	€ 61,707,659	-0.09
	S2	€ 1,096,200,271	€ 11,516,515	0.01
	S3	€ 203,259,042	€ 24,545,879	0.12
	S4	-€ 851,347,947	€ 122,646,193	-0.14
SCN Base	**S1**	-€ 711,246,368	€ 73,983,215	-0.10
	S2	€ 1,186,232,145	€ 13,631,161	0.01
	S3	€ 219,952,883	€ 30,678,879	0.14
	S4	-€ 921,269,888	€ 145,362,722	-0.16
SCN High	**S1**	-€ 739,831,158	€ 91,766,605	-0.12
	S2	€ 1,233,906,478	€ 16,688,359	0.01

	S3	€ 228,792,726	€ 38,820,161	0.17
	S4	-€ 958,295,463	€ 175,131,775	-0.18

Here, since, ADC>0 and ADE<0 for S1 and S4, therefore the the interventions of S1 and S4 fall in Quadrant II: The planned *Road Safety Measure* is effective however the costs can still be lowered, whereas, S2 and S3 fall in Quadrant I: "The planned *Road Safety Measure* is not effective, since the cost is high and the *Road Traffic Crashes*, *Road Traffic Injuries*, and fatalities remain high."

The efficacy of the intervention can be determined by an ICER converging to zero.

West Midlands:

Table 4.12: CEA results for West Midlands

CEA Scenario		**PV (QALYs) (ADE)**	**PV Cost (ADC)**	**ICER**
SCN Low	**S1**	-£ 5,447,506.70	£ 21,832,239.63	-0.25
	S2	-£ 888,965,939	£ 1,348,003	-0.002
	S3	-£ 917,090,391	£ 89,614,312	-0.098
	S4	-£ 806,681,539	£ 134,434,312	-0.167
SCN Base	**S1**	-£ 5,894,915	£ 27,129,992	-0.22
	S2	-£ 961,977,479	£ 1,597,492	-0.002
	S3	-£ 992,411,816	£ 106,149,232	-0.107
	S4	-£ 872,934,990	£ 159,414,482	-0.183
SCN High	**S1**	-£ 6,131,830	£ 36,275,140	-0.17
	S2	-£ 1,000,639,080	£ 1,965,942	-0.002

	S3	-£ 1,032,296,565	£ 128,860,164	-0.125
	S4	-£ 908,018,000	£ 197,239,395	-0.217

4.5 Result synthesis to refine the PHOEBE framework

The synthesis step helps identify how upstream model outputs and input availability shape the interpretation of appraisal results. For example, some benefit components that depend on safety and network assessment inputs reflect the scope and structure of the available input data. In the West Midlands A45 case, travel time savings benefits for S3 and S4 are not available, which limits direct comparison across all scenarios. Similarly, the reported FSI prevention benefits for S2, S3, and S4 are identical and comparatively large. While the scenarios differ, this consistency reflects the use of common underlying safety input assumptions and highlights the importance of interpreting these results in the context of those inputs. In Valencia, the FSI prevention benefits for S2 and S3 are large and negative, and therefore play a significant role in the overall CBA results. These observations underline the value of tracing results back to their underlying assumptions to ensure transparency and support robust interpretation.

4.6 Scaling the outputs at city wide level

The PHOEBE framework typically is implemented on a part of a big network, since implementing it on a full scale city network is computationally expensive. However, the framework produces results (here FSIs) that can be extrapolated (or scaled up) for the entire network. The scalability section of the socio-economic component is responsible for the extrapolation of the FSI to the city level. The details of the scaling component is provided in this section along with the results.

Scaling methodology:

The scaling is done in two tiers. In the first tier, the model outputs (FSIs) are converted into comparable rates, such as fatalities and serious injuries per 100,000 population. These rates are then scaled up to the full city by multiplying them with the city wide population and dividing by 100,000. This step gives city level totals that are consistent across use cases and suitable for reporting and appraisal. In the second tier, the city level safety totals of FSIs are distributed to specific network elements outside the modelled area. For each link, the required road and traffic attributes are prepared, and the fitted ML (regression) model is applied to estimate the link-wise expected fatality rate under the given conditions. The predicted rate is then converted into annual counts using an exposure measure based on AADT and link length.

The city wide scaling can be done for all major outputs of socio-economic analysis, however, the network level (outside the PHOEBE modelled area) extrapolation is only possible for the FSIs.

Scaling: Tier 1

The FSI components can be scaled up using the population extrapolation method. The values of the FSIs for the three use cases per 100k population is presented below.

FSIs crashes per scenario (per 100k population):

Table 4.13: FSIs prevented per 100k population in Athens

Athens	Fatalities	Serious injuries	Light injuries
Base	6.20	170.56	431.04
S1	5.46	150.17	379.52
S2	3.75	103.03	260.38

Table 4.14: FSIs prevented per 100k population in Valencia

Valencia	Fatalities	Serious injuries	Light injuries
Base	4.06	111.52	1532.80
S1	3.58	98.34	1351.63
S2	17.91	129.62	1781.51
S3	4.20	115.63	1589.18
S4	3.43	94.45	1298.15

Table 4.15: FSIs prevented per 100k population in the West Midlands

West Midlands	Fatalities	Serious injuries	Light injuries
Base: Edgbaston	2.55	28.05	151.72
Edgbaston: S1	2.45	26.92	145.61
Base: A45	3.97	45.65	236.16
A45: S1	2.19	25.21	130.41
A45: S2	2.19	25.21	130.41
A45: S3	2.19	25.21	130.41

These metrics are computed per 100k population. The similar metrics for the whole city (outside modelling scope) can be computed by multiplying these factors with the number of 100k population in that city. E.g., if the population of Athens is 5,3 million, the factor with which the metrics from Athens must be multiplied is 53.

Scaling: Tier 2

An extrapolation of FSIs at the link-level outside the modelled domain within the city limits is carried out in this section. A direct "copy and paste" of observed fatality rates at the link-level to the rest of the city (outside PHOEBE's modelling domain) is not justifiable because exposure and road environment differ outside the modelled boundary considerably. The purpose of the forecast is therefore transferability, i.e., to learn a systematic relationship between link attributes (e.g., operating speed, number of lanes, carriageway type, road condition) and annual fatalities, and then apply that relationship to links outside the simulation area that share similar attributes. This provides a consistent, data-driven way to extrapolate baseline safety outcomes and scenario effects beyond the pilot network, while remaining transparent enough to support cost–benefit analysis and stakeholder scrutiny.

Fatalities are rare and highly skewed at link level. Classical linear regression on raw counts tends to produce negative predictions and unstable fits. The chosen approach treats fatalities as non-negative outcomes tied to exposure, which aligns with how crash risk is typically modelled: as a rate per unit exposure rather than an unconstrained continuous variable.

The linkwise fatalities, serious injuries, and light injuries are fed to the socio-economic module by the iRAP road safety assessment model. This data consists of linkwise fatalities, serious injuries, and light injuries counts along with the attributes of the links. These attributes are used as predictors, whereas the FSI counts (linkwise) are the responses.

Model selection:

Across the three use cases (Athens, Valencia, West Midlands [EGB and A45]), a Tweedie generalized linear model (GLM) with a log link was selected for the core forecasting step. The choice is driven by three features of the data: (i) Fatalities (even expected values) are ≥ 0 and highly skewed. A log-link model ensures positive predictions, (ii) Safety outcomes behave like rare events with heavy tails. Tweedie (with power 1–2) is designed for this kind of distribution (compound Poisson–Gamma style), iii) coefficients translate into multiplicative effects (percent changes), and (iv) forecasting must be robust and interpretable for policy communication. The Tweedie family is a pragmatic compromise as it behaves like a Poisson–Gamma compound distribution for certain power parameters and is commonly used for severity problems where outcomes are non-negative and skewed

Interpretability is retained through the log link: each coefficient can be translated into a multiplicative change in the predicted fatality. For example, in Athens the coefficient for operating speed (mean) is 0.169, which corresponds to an approximately +18.4% change in the predicted fatality per one unit increase in that variable (all else equal). Similar "percent effect per unit" statements can be produced for all predictors, supporting defensible explanations in deliverables.

Modelling methodology:

The link level models use a generalized regression form that is standard in safety analysis. For each link $i$the observed annual fatalities y_i.

$$log(E[F]) = \beta_0 + \sum_{k=1}^{p} \beta_k . x_{ik} \quad \text{(Eq.14)}$$

This is a form of supervised learning because the model is trained on labeled examples, where inputs $X_i$are link attributes and the label is the observed fatality outcome.

Results:

Athens

Model type: Tweedie GLM (p = 1.99, log link)

Table 4.16: Regression outputs

Feature	Coefficient	p-Value	% change per 1 original unit	Scale used in model divisor
Intercept	-9.854	0.000		
NUM: Avg. operating speed	0.069	0.000	7.2%	1
NUM: Average Vehicle flow (AADT)	7.441e-05	0.000	Per +1,000 AADT, 7.7%	1
NUM: Speed limit	0.015	0.007	Per +10 units: 16%	1
NUM: Length of link (KM)	6.505	0.000	670 times higher fatalities per extra km	1
NUM: Average number of lanes	-0.0997	0.072	9.5% lower expected fatalities per additional lane	1

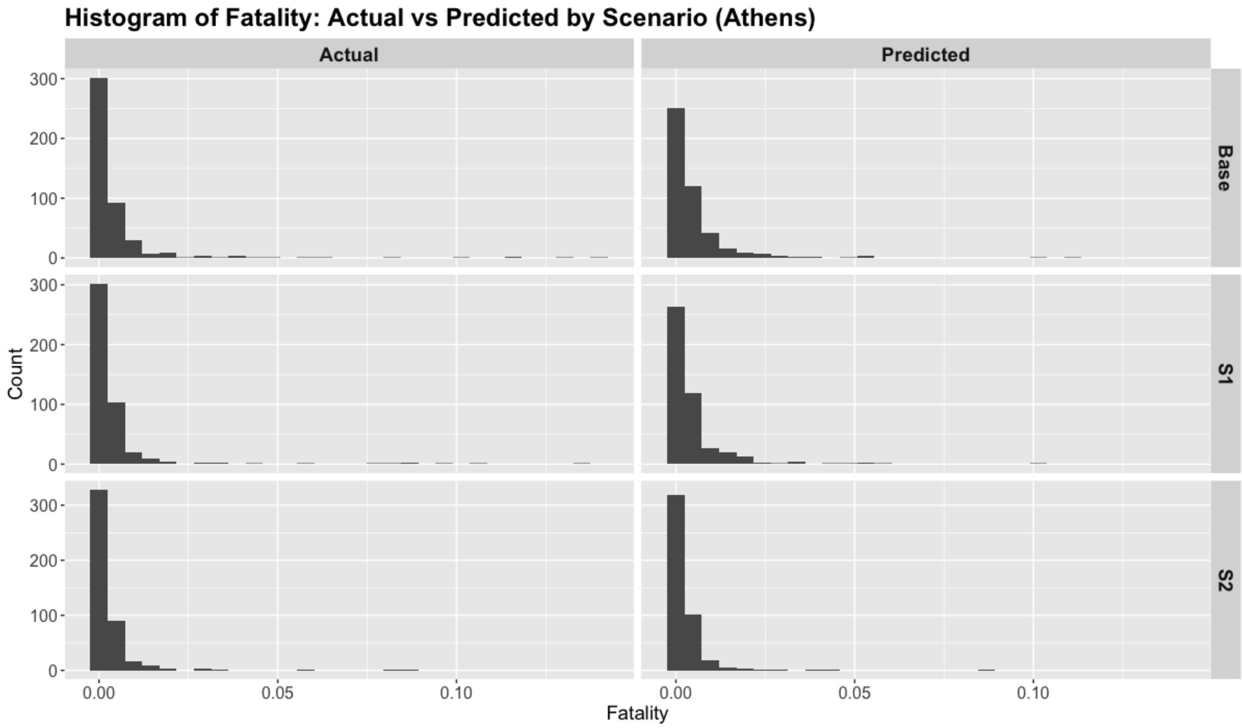
Deviance explained:

Null deviance = 3552.6; Residual deviance = 1434.5; ~59.6% deviance explained.

Table 4.17: Sum of link-wise predicted vs actual fatalities in Athens

Athens: Scenario	Actual	Predicted
Base	2.31	2.42
S1	2.04	2.13
S2	1.40	1.29

Figure 4.8: Athens, actual vs predicted fatalities scenario wise.



The serious and light injuries are 27.5 and 69.5 times that of the number of fatalities. Therefore, these metrics can be predicted by multiplying the predicted fatality with these factors.

Valencia

Model type: Tweedie GLM (p = 1.99, log link)

Table 4.18: Regression outputs

Feature	Coefficient	p-Value	% change per 1 original unit	Scale used in model divisor
Intercept	--7.899	0.000		
NUM: Avg. operating speed	0.0996	0.000	10.5% higher expected fatality per +1 speed unit	1
NUM: Average Vehicle flow AADT	2.573e-05	0.000	2.6% higher expected fatality Per +1,000 AADT	1

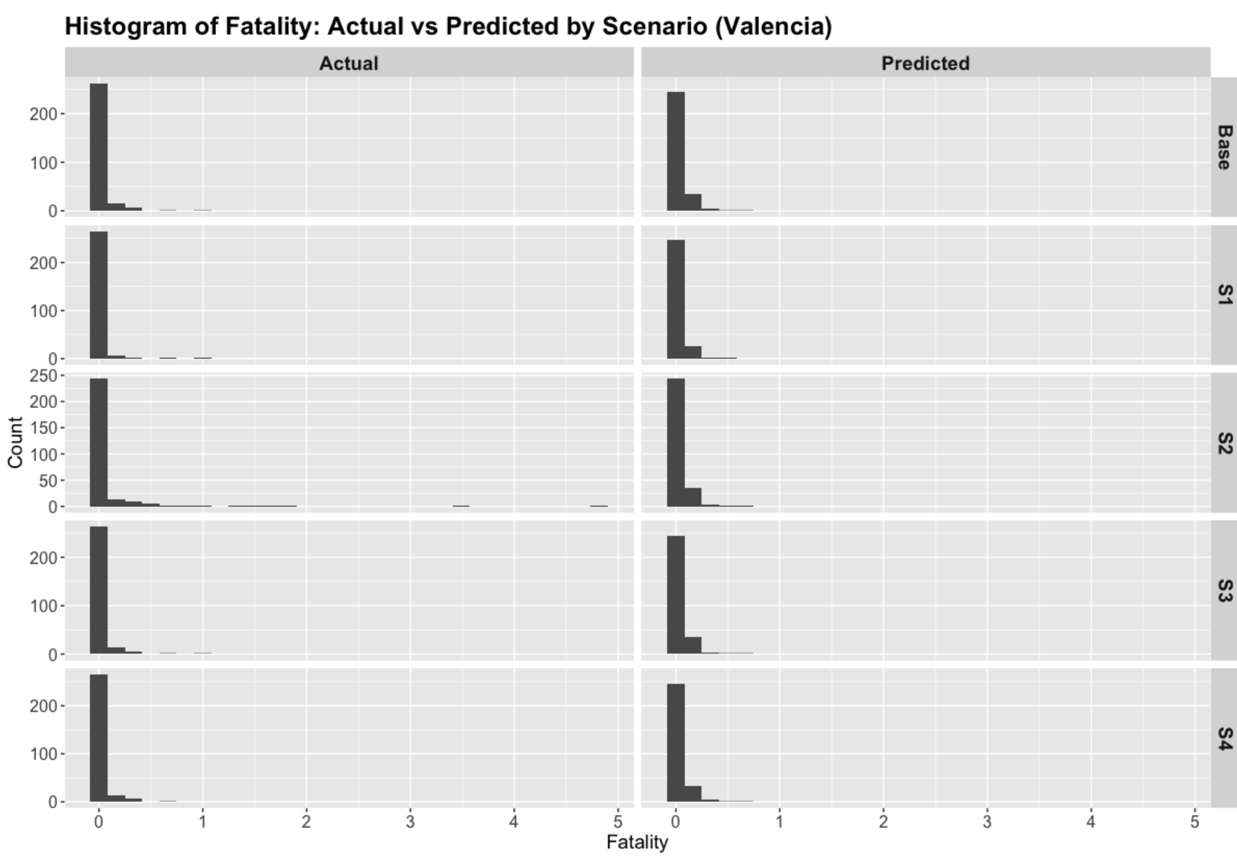
Deviance explained:

Null deviance = 6428.1; Residual deviance = 4451.5; ~30.8% deviance explained.

Table 4.19: Sum of link-wise predicted vs actual fatalities in Valencia

Valencia: Scenario	Actual	Predicted
Base	7.24	12.60
S1	6.39	9.83
S2	32.00	12.62
S3	7.51	12.51
S4	6.14	12.40

Figure 4.9: Valencia, actual vs predicted fatalities scenario wise.



The serious and light injuries are 27.5 and 377.96 times that of the number of fatalities. Therefore, these metrics can be predicted by multiplying the predicted fatality with these factors.

West Midlands

Model type: Tweedie GLM (p = 1.965, log link)

Table 4.20: Regression outputs

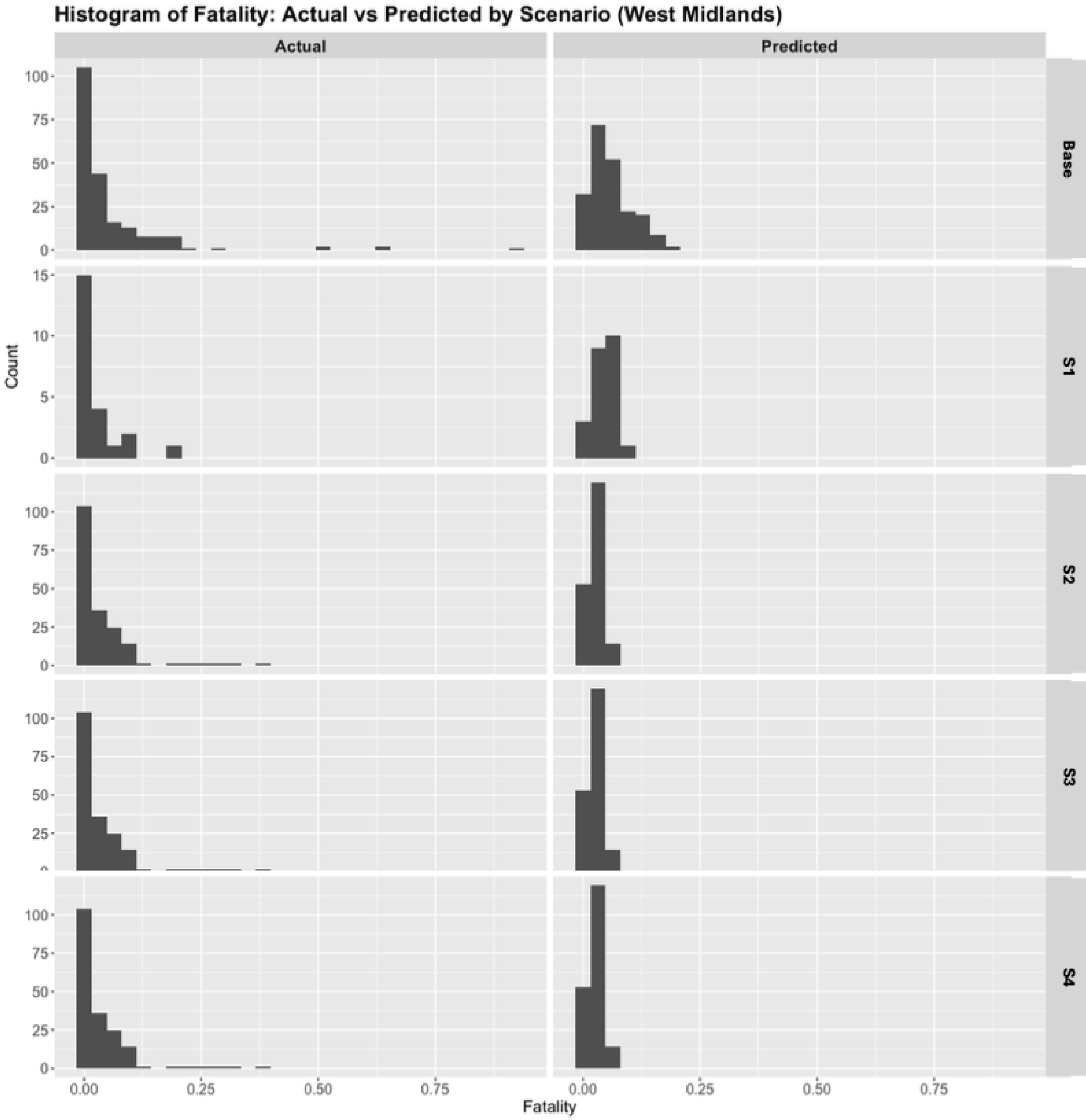
Feature	Coefficient	p-Value	% change per 1 original unit	Scale used in model divisor
Intercept	-6.184	0.000		
NUM: Average operating speed	0.060	0.000	6.1% higher expected Fatality per +1 unit increase in operating speed	1
NUM: Speed limit	0.013	0.300	1.3% higher expected fatality per +1 unit	1

Null deviance = 1387.4; Residual deviance = 1150.7; ~16.4% deviance explained.

Table 4.21: Sum of link-wise predicted vs actual fatalities in the West Midlands (Edgbaston)

West Midlands: Scenario	Actual	Predicted
Base (Edgbaston)	0.69	0.96
S1 (Edgbaston)	0.666	0.95
Base (A45)	10.80	11.00
S2 (A45)	5.96	5.35
S3 (A45)	5.96	5.35
S3 (A45)	5.96	5.35

Figure 4.10: Histogram of scenario wise fatalities on the links the West Midlands



The serious and light injuries are 11.0 (Edgbaston), 11.5 (A45) and 59.48 times that of the number of fatalities. Therefore, these metrics can be predicted by multiplying the predicted fatality with these factors.

Forecasting for links outside the modelled area:

To forecast beyond the simulated network:

1. Assemble an “outside links” dataset with the same predictor fields used in the fitted model. Ensure units match the training data.

2. Apply domain checks before predicting by verifying each numeric attribute lies within (or close to) the training range and categorical levels are known. If not, either cap values to the training domain, map categories to the closest known level, or flag those links as out-of-sample.
3. Predict fatalities using the regression outputs listed in the tables above.
4. Aggregate for reporting. Link-level forecasts are noisy; policy decisions usually rely on corridor, district, or citywide totals. Summarize by area, road class, or intervention influence zone.
5. Quantify uncertainty. Use cross-validation spread (fold-wise performance) as an empirical indication of uncertainty, and where performance is poor, prefer higher-level aggregation or alternative models (e.g., hierarchical models, corridor fixed effects, or richer features).

Overall, Athens and Valencia models are usable for transfer; Edgbaston and A45 require either more data or improved predictors before extrapolation can be carried out at a fine spatial resolution.

5 Dynamic Safety Prediction Tools

5.1 Objectives

European cities are working to make their infrastructure and transportation networks more inclusive, sustainable, and safe, despite the challenges posed by complex interactions among users, infrastructure, and transport modes. Creative yet cautious strategies are necessary. To meet these goals, cities must combine simulation with rigorous assessments. Simulations help anticipate problems and adapt systems to diverse needs, while advanced safety and accessibility evaluations address the limitations of traditional, static risk assessments. Since traffic risk levels change throughout the day with shifting conditions, adaptive assessment methods are needed to reflect real-world dynamics (Theofilatos; Yannis, 2014).

A major challenge in advancing adaptive methodologies is collecting reliable, granular data on speed and flow. Although technologies such as telematics, roadside sensors, and emerging data sources exist, acquiring these data at scale remains difficult due to technological, logistical, and privacy barriers (Ghaffarpasand et al., 2022). Moreover, these data sources often cannot represent future scenarios, which are crucial for testing new solutions. As a result, this study utilizes simulation-based operational outcomes to explore scenarios and predict safety impacts under different conditions. For example, simulation platforms like Aimsun are used to model interventions such as introducing new bike lanes or pedestrian crossings. These platforms allow urban planners to analyze how such changes affect traffic patterns under varying traffic volumes and speeds.

Simulation environments are useful for testing transportation interventions, but results may not always match real-world outcomes without careful evaluation. Safety indicators can be microscopic, mesoscopic, or macroscopic. Microscopic indicators like Time to Collision offer detail but require significant data. Mesoscopic models provide network-level trends with less data but lack conflict detail. Macroscopic measures give an overall view but miss behavioural nuances and need expert interpretation. Integrating these levels is challenging due to data harmonization and translating detailed findings into broader metrics (Kumar: Mudgal, 2025; Mustapha et al., 2024).

The primary objective of this activity is to establish a robust connecting framework between traffic simulation and risk assessments that effectively assesses the safety impacts of transportation interventions. This framework is designed to integrate simulation outputs from AIMSUN Next Software (Aimsun, 2025) with risk assessment databases from the iRAP Star Ratings (iRAP, 2024), creating a reciprocal flow of insights between the two approaches. A critical component of the framework is ensuring data compatibility and transfer. The study has tested data transfer through a manual pilot and provides guidelines for full automation of the process.

By integrating dynamic operational parameters into risk assessment, this framework delivers direct, evidence-based results that are easy for decision-makers to understand (to know more access PHOEBE **Deliverable 1.2**). It clarifies safety levels, helping stakeholders choose optimal interventions for transportation solutions. Developed under the PHEOBE Project, all methods and outputs support the project's aim of improving transport safety and efficiency. The research and implementation described were financed by PHEOBE.

5.2 Methodology

The connectivity is showcased between the AIMSUN Next Simulation Software and the iRAP Star Rating model operationalised in the ViDA System.

Aimsun Next provides a flexible simulation platform covering microscopic, mesoscopic, macroscopic and hybrid modelling approaches (Aimsun, 2025). It supports the dynamic representation of traffic operations, allowing users to replicate both individual driver's behaviour and aggregated network conditions within a unified framework. The use of the software scripting and API interfaces enables the connection with other tools. In the scope of the PHOEBE Project this has been used to integrate various modules, and it is used to integrate the connection between Aimsun software and iRAP's ViDA platform to assess dynamic changes on the road safety based on variables, like speed average, deviation, flows.

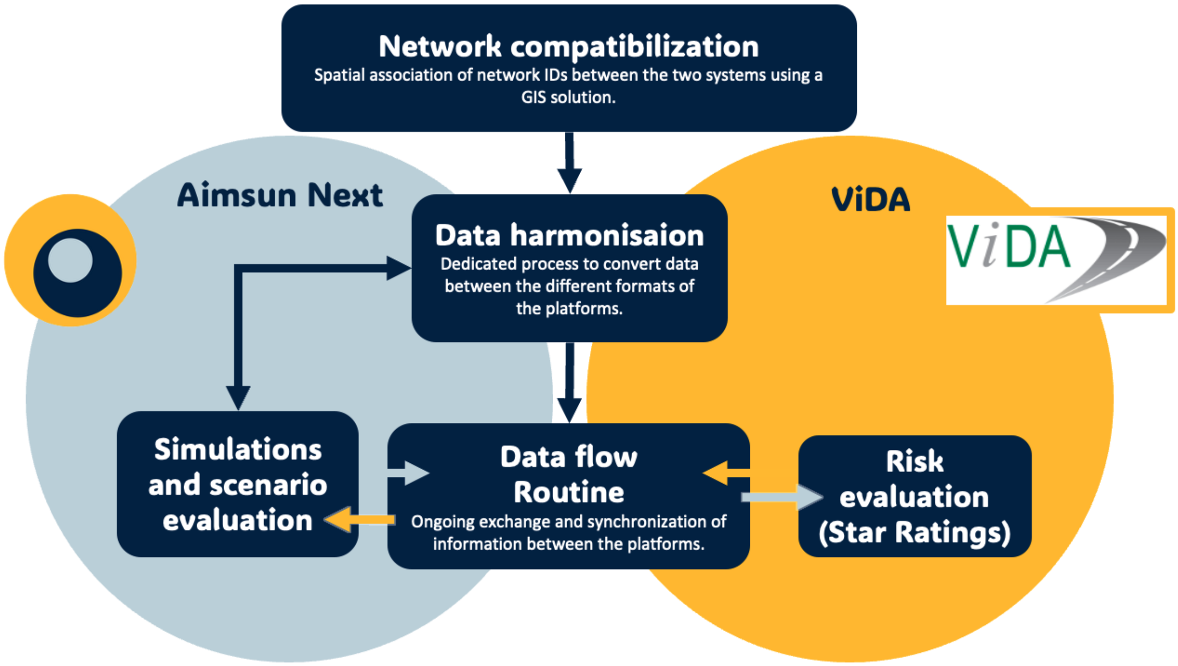
The ViDA platform and Star Rating Systems are fundamental tools in global road safety assessment frameworks. ViDA operates as a centralized analytical environment that manages road safety data and enables systematic evaluation of infrastructure risk using the Star Rating Model. The Model provides standardized metrics for rating road segments based on crash likelihood and severity, while ViDA offers capabilities such as geospatial visualization of these results (iRAP, 2021). The Star Rating Model computes star ratings for different road user categories—pedestrians, bicyclists, motorcyclists, and vehicle occupants—using algorithms that consider infrastructure attributes and traffic conditions. Key parameters include traffic volume, speed environment, intersection density, roadside hazards, and protective infrastructure such as cycle lanes and pedestrian crossings.

The connectivity between the two systems involved the three steps, described below:

1. Network Compatibilization: Differences in network setups require associating network IDs to align the systems. This ensures data exchanged between platforms refers to the same road locations. Before connecting simulation and risk assessment platforms, a GIS solution must spatially link network elements from both Aimsun Next and ViDA. This synchronizes datasets based on geographic positions. If a risk assessment exists, map current elements to it; for new assessments, use the established Aimsun Next configuration.
2. Data Harmonization: Each platform involved in the connectivity framework operates with distinct data structures and formatting requirements. To facilitate seamless interoperability, a dedicated data transformation process is essential. This process systematically converts outputs produced by one system into the specific input formats required by the other. This process is designed to be automated and happen every time data is shared between the two platforms. To further ensure the reliability and accuracy of data exchange, the transformation process is tightly integrated with validation routines that check for completeness, consistency, and alignment with both platforms' schemas before information is ingested or exported.
3. Data flow routine: After network compatibility and data transformation, the operational data flow manages ongoing information exchange between simulation and risk assessment systems. The process starts with importing the Star Rating database into Aimsun Next, where users can adjust road attributes for each simulated scenario. Post-simulation, traffic performance parameters are extracted from Aimsun Next outputs, reformatted to fit ViDA's schema, and transferred to the Star Rating database. The data is validated for ViDA processing, which recalculates star ratings and infrastructure risk scores. These updated results are then sent back to Aimsun Next to inform further simulations or analysis.

Figure 5.1 illustrates the operational workflow between the Aimsun Next simulation platform and the ViDA Star Rating model.

Figure 5.1: Connectivity framework flowchart



5.3 Model Development

5.3.1 Manual connectivity

To test the effectiveness of the proposed connectivity framework prior to automation, the project team developed and implemented a manual schedule, where the data transformation and transferability was performed using an Excel testbed. The Edgbaston region in the West Midlands was selected as a representative use case for this testing phase. This area serves as one of the pilot locations for the PHOEBE Project, specifically focusing on the simulation of a new cycle path along the A38 corridor. The manual process enabled the team to rigorously test connectivity by generating safety performance indicators that reflect the impact of the infrastructure change across various times of the day.

In this use case, a risk assessment had already been conducted, necessitating the alignment of the Aimsun Next simulation network with the existing risk assessment results. To achieve this, network compatibility was performed to establish a direct connection between the current risk assessment data and the specific network elements represented in Aimsun Next. This process involves mapping the network identifiers used in the risk assessment to the corresponding network IDs in Aimsun Next. The Object Identifiers (OIDs) were selected for this mapping process because they serve as the internal unique identifiers assigned by Aimsun Next to each object within its simulation database. Utilizing OIDs ensures that each network element in the simulation can be unambiguously linked to its counterpart in the risk assessment, enabling accurate data exchange and integration between the two systems.

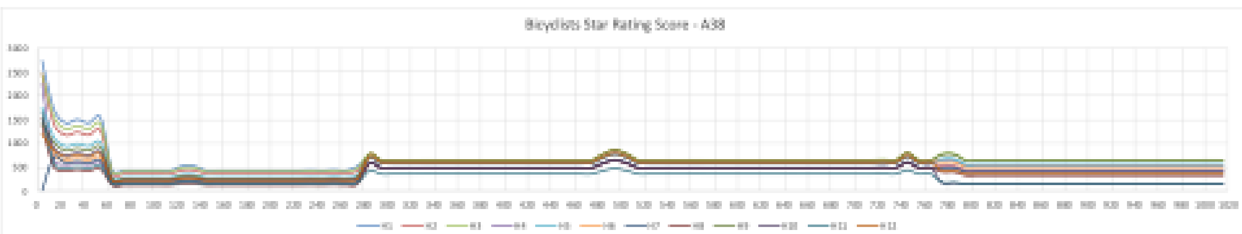
Several key challenges arose during network compatibility. First, Aimsun Next and iRAP represent road networks differently-Aimsun uses a link-and-node approach, while iRAP relies on continuous chainage. Aligning these systems requires precise mapping to ensure data corresponds correctly. Second, differences in coordinate reference systems complicate the spatial matching of network elements, making it difficult to accurately link corresponding locations. Thirdly, the platforms vary in precision, which can hinder the integration of results. To maintain accuracy, manual checks are often needed, though this process is much simpler if the risk assessment adopts the simulation network as its reference.

Following the completion of network matching between the simulation and risk assessment datasets, the next phase involved transforming the simulation outputs to align with the Star Rating specifications. Specifically, traffic flow values generated by the simulation in 15-minute intervals were converted to Annual Average Daily Traffic (AADT) figures. In addition to traffic flows, speed metrics were processed to fulfill the requirements of the Star Rating methodology. The mean operating speed, a standard output from Aimsun Next, was calculated directly from the simulation results. To estimate the 85th percentile operating speed, the method employed involved using two standard deviations above the mean operating speed as a proxy.

With the Object Identifiers (OIDs) already incorporated into the Star Rating database, linking the operational parameters from the simulation to their corresponding network elements was straightforward. This direct connection enabled the recalculation of risk results by using the updated simulation outputs as inputs, allowing for a seamless integration of simulation data into the risk assessment workflow.

Figure 5.2 draws upon the simulation outputs and risk assessment data gathered during the earlier manual integration stage. It depicts how bicyclist risk levels vary across different hours of the day and along various road segments in scenarios where a cycle path has not been installed. Each line on the graph shows the fluctuation of risk indicators—specifically Star Rating Scores—responding to shifts in traffic conditions and exposure of road users across the network and throughout the day. This visualization highlights the depth of insights made possible by seamlessly connecting risk assessment data with traffic simulation tools.

Figure 5.2: Bicyclists Star Rating Scores



5.3.2 Automatic connectivity

The goal of transferring data between ViDA system and AIMSUN Next is to support interoperability for in-depth safety and operational analysis. This requires aligning spatial references, attribute definitions, and automating Star Rating parameter integration into simulations, with feedback loops to improve risk predictions. The integration framework includes network compatibility, data harmonisation, and establishing a data flow routine. Manual tests confirmed that integration is achievable, with reproducibility ensured through standardised scripts and version-controlled datasets.

The next step refers to the integration of automated data exchange between AIMSUN Next and the iRAP Star Rating model, representing a significant advancement in the field of road safety assessment and operational decision-making. The manual pilot conducted in the Edgbaston region enabled the project team to identify and address any operational challenges, refine procedures, and establish baseline practices for subsequent automation efforts.

Automating this process is anticipated to deliver several key benefits including improving operational efficiency, enhanced Decision-Making and Safety Outcomes Prioritisation, scalability and improving collaboration. Automation streamlines the transfer and transformation of data, reducing manual intervention and the potential for human error. This enables faster, more reliable analysis and supports continuous updates for dynamic networks, making safety evaluations more responsive and robust. Additionally, by providing real-time, evidence-based safety indicators, the automated framework empowers decision-makers to identify and prioritise interventions that yield the greatest safety benefits. This leads to more informed, proactive responses to emerging risks and infrastructure needs. The project is keen to provide automated workflows, supported by API pipelines, facilitating the expansion of the methodology to larger and more complex networks.

In summary, the transition from manual to automated processes not only enhances the technical capabilities of safety assessment frameworks but also supports the broader goals of sustainable, inclusive, and safe urban mobility.

The automation is developed in two phases:

- Phase 1 - Data exchange from Aimsun Next to ViDA
- Phase 2 - Data exchange from ViDA to Aimsun Next

5.3.2.1 Phase 1 - Data exchange from Aimsun Next to ViDA

The objective of Phase 1 is to upload all operational parameters and any simulated modifications to road infrastructure into the ViDA database, ensuring strict compliance with system requirements and maintaining data integrity throughout the process. This step is critical as it enables the risk analysis results to be updated based on the latest simulation outputs, ensuring that assessments reflect current conditions and planned changes.

In order for this process to run smoothly, the following pre-setups are required:

1. **Network Matching Completed**
   Each row in the iRAP database must have a corresponding AIMSUN Next Object Identifier (OID).
2. **Project Created in ViDA**
   Ensure a project is set up in the ViDA system and linked to the correct database ID.
3. **Infrastructure Data Validation**
   All infrastructure data must pass ViDA validation checks without errors.
4. **AIMSUN OID Entry**
   The AIMSUN OID must be correctly entered in the **'Section'** column for each relevant record.

To ensure compatibility with iRAP standards, all simulation outputs must undergo a transformation process that aligns them with the required specifications for some particular attributes. This step is essential because iRAP's methodology relies on standardized data formats and calculation rules to produce accurate risk assessments. To streamline this conversion and eliminate manual intervention, an API has been developed. This API automates the necessary computations and securely transmits the information to the ViDA platform for validation. While this process has been tested with Aimsun Next, similar steps can be performed by other simulation tools.

The attributes that need processing are:

- Road AADT: Hourly counts of all motorized vehicles multiplied by a conversion factor (user to define) per 'oid' – Number input
- Operating speed (mean): Aimsun Next 'Speed' per 'oid'– Number input
- Operating speed (85th percentile): 'Spedh' + 'spdh_D' per 'oid'– Number input

- % of motorcycles: count of motorcycles / count of all motorized vehicles and transformed into iRAP specification code (see Table 5.1) per 'oid'

Table 5.1: Percentage of motorcyclists categories codes

Categories	Categories Code
Not recorded	1
0%	2
1% - 5%	3
6% - 10%	4
11% - 20%	5
21% - 40%	6
41% - 60%	7
61% - 80%	8
81% - 99%	9
100%	10

- % heavy vehicles: count of heavy vehicles / count of all motorized vehicles and transformed into iRAP specification code (see Table 5.2) per 'oid'

Table 5.2: Percentage of heavy vehicles categories codes

Categories	Categories Code
Not recorded	1
0	2
1-<5%	3
6-<10%	4
11-<15%	5
16-<20%	6
21-<30%	7
31-<40%	8
>40%	9

- Bicycle peak hourly flow: count of bicyclists in an object(oid) and transformed into iRAP specification code (see Table 5.3) per 'oid'

Table 5.3: Bicycle peak hourly flow categories codes

Categories	Categories Code
0	1
1 to 5	2
6 to 25	3
26 to 50	4
51 to 100	5
101 to 200	6
201 to 300	7
301 to 400	8
401 to 500	9

501 to 900	10
900+	11

Examples of calculations are available at https://seafile.irap.org/f/19d61aaed45d4e39882f/. They are based in the Intervention 1 of West Midlands PHOEBE Use Case. To know more about the scenario tested and the Aimsun and Star Rating available models for the scenario, refer to deliverables **D3.2** and **D4.2**. An example of expected output available for the same interventions can be found at https://seafile.irap.org/f/4bc68a43c32b4eafaedd/

For more information on the iRAP specification, please access https://irap.org/specifications/

Enhanced version

For baseline scenarios, the existing operational parameters should be sufficient, as the iRAP dataset reflects the current road configuration. However, when testing interventions, certain infrastructure parameters may also require updates. For instance, if a simulated scenario involves converting a traffic lane into a cycling lane, attributes such as the number of lanes and bicycle facilities in the iRAP database must be adjusted to accurately represent these changes.

A practical approach to enable this customization for Aimsun Next users is to load the entire iRAP dataset into Aimsun Next as editable variables. This would allow users to modify values directly within the simulation environment. The expected outcome of this process is an updated upload file that incorporates both the revised infrastructure parameters and the operational parameters after the simulation concludes. This file then is sent back to ViDA for processing, ensuring that risk analysis reflects the latest intervention scenarios. This work will be performed outside the PHOEBE scope in ongoing collaboration between the two organisations.

An example of expected output available for the same interventions can be found at https://seafile.irap.org/f/4bc68a43c32b4eafaedd/

5.3.2.2 Phase 2 - Data exchange from ViDA to Aimsun Next

The goal of Phase 2 is to enable Aimsun Next users to evaluate road safety risk results when simulating various scenarios, thereby integrating safety considerations into the decision-making process for selecting the most effective urban interventions.

In order for this process to run smoothly, the following pre-setups are required:

1. **Phase 1 Completed:**

   Initial integration and baseline configuration finalized.

2. **Results Processed in ViDA:**

   Risk analysis outputs successfully generated and exported as a CSV file for further use in simulation workflows.

Once the dataset has been processed in ViDA, it should be imported back into Aimsun Next. This step allows users to visualize and analyze road safety results alongside other performance indicators within the simulation environment. By importing the results columns, users can enable visual mapping directly in the platform, facilitating a comprehensive assessment of safety impacts in relation to operational and mobility metrics.

The data transfer and database accessibility is done through an API. The iRAP API documentation is available at https://openapi.irap.org/.

5.4 Integration between iRAP ViDA and Aimsun Next

This section describes the technical implementation of the integration between Aimsun Next and the iRAP ViDA platform. The purpose is to provide a clear and compact description of the end-to-end integration workflow, focusing on: (1) the sequence of technical steps; (2) the scripts and components involved at each stage; (3) the interaction with the ViDA REST API.

Rather than presenting a theoretical background on road safety assessment methodologies, this document is intentionally implementation oriented. It is structured around a step-by-step technical pipeline, allowing technical readers to understand how simulation outputs are transformed, uploaded, processed, retrieved, and visualized.

5.4.1 Integration overview - End-to-End pipeline

The integration between Aimsun Next and ViDA follows a single end-to-end technical pipeline, composed of a limited number of clearly defined steps. Each step produces a specific output and feeds directly into the next one.

Phase 1 (Aimsun to ViDA): Operational traffic data extracted from Aimsun simulation results is transformed to iRAP specifications, merged with existing ViDA infrastructure survey data, and uploaded to the ViDA platform for star rating computation.

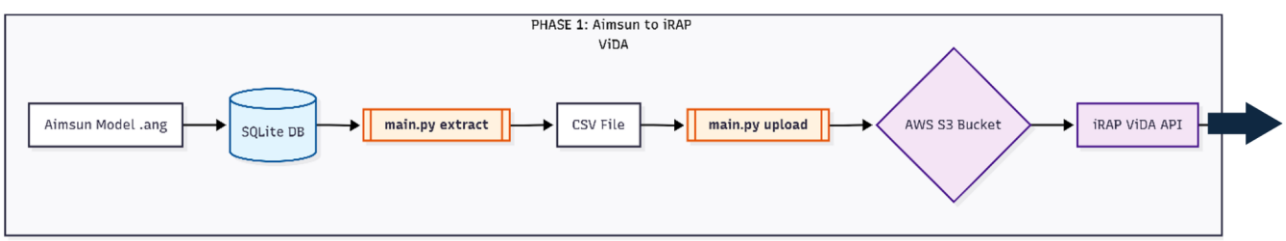
Phase 2 (ViDA to Aimsun): Computed star ratings are downloaded from ViDA, aggregated by Aimsun section, imported into the Aimsun Next model as external attributes, and visualised through automatically configured View Modes.

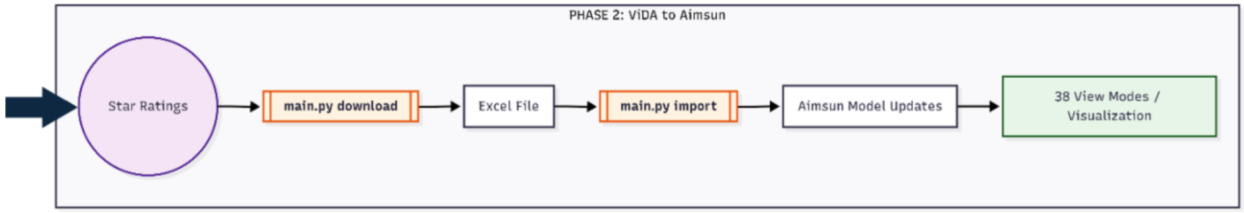
Figure 5.3 provides an overview of the complete workflow, highlighting:

- where data is processed locally.
- where interaction with the ViDA REST API occurs.
- how results are reintegrated into the Aimsun environment.

This figure is intended to serve as a reference diagram for the remainder of the document.

Figure 5.3: Aimsun Next - iRAP ViDA (and viceversa) workflow





5.4.2 End-to-End technical integration pipeline

The integration between Aimsun Next and iRAP ViDA platform is implemented as a linear, step-based technical pipeline. Each step performs a single, well-defined function and produces a clear output that feeds directly into the next.

The full pipeline can be executed automatically through a single-entry point (`main.py`) or step by step for debugging and validation purposes. A dedicated merge script (`scripts/merge_and_upload.py`) is also available for the production upload workflow.

5.4.2.1 Prerequisites and Network Matching

Before executing the pipeline, the following conditions must be met:

- An Aimsun Next model with simulation results exported in SQLite format.
- An Irap ViDA project and dataset configured with validated infrastructure attributes.
- One-to-one mapping between Aimsun section OIDs and irapn ViDA location_id (using the section field in ViDA).
- Valid application and user credentials for the ViDA API configured in config.json.
- AWS SSO credentials configured for S3 upload access.

5.4.2.2 Extract Traffic Simulation Outputs

This step handles the local processing of simulation results to prepare the data according to iRAP ViDA's technical specifications.

Table 5.4: Script to extract traffic simulation outputs

Script	main.py extract
Module	src/aimsun_database.py

Functionality	● The Aimsun SQLite database is queried at section level. ● Traffic volumes are aggregated by vehicle type (mapping directed by parameter sid) ● AADT values are computed. ● 85th percentile speed is computed using functions over the simulation results. ● Vehicle composition metrics derived.
Output	CSV file with section-level operational parameters. No API interaction occurs in this step.

5.4.2.3 Transform Data to iRAP ViDA Specifications

Data transformation ensures that continuous simulation variables, such as speeds and flows, are correctly categorized into the discrete codes required by the Irap ViDA platform.

Table 5.5: Script to transform data to iRAP ViDA Specifications

Script	main.py extract
Module	src/aimsun_database.py
Functionality	● Categorical Binning: Continuous traffic data (e.g., AADT) is processed through a discretization engine that assigns the corresponding iRAP category code (1–11). ● Vehicle Type Mapping: precise percentages for each road user group are derived using validated Aimsun identifiers, specifically mapping. ● Speed Normalization: The 85th percentile operating speed is validated and encoded into integer-based iRAP categories to ensure compatibility with ViDA's risk calculation formulas. ● Data Validation: The system performs a range check to ensure all transformed codes fall within the official iRAP specification limits before finalizing the CSV.
Output	A structured CSV file containing the processed operational parameters, where each row is indexed by the Aimsun OID (mapped to ViDA's location_id), ready for immediate API upload.

5.4.2.4 Upload Operational Parameters to iRAP ViDA platform

This stage manages the secure transmission of the processed data to the ViDA cloud environment, utilizing the platform's REST API for automated ingestion.

Table 5.6: Script to upload operational parameters to iRAP ViDA platform

Script	main.py upload
Module	src/vida_api.py
API Interaction	POST /api/v1/datasets/{id}/upload_url

Functionality	● Authentication: The script performs a handshake using Application and User keys to retrieve a session token. ● S3 Upload Workflow: Instead of a direct POST, the script requests a signed S3 URL. This allows the CSV to be streamed directly to an AWS S3 bucket via the boto3 library, ensuring efficiency for large-scale networks. ● Job Triggering: Once the file transfer is complete, the API triggers an internal ingestion job. ● Validation Monitoring: The script polls the job_id to confirm that the server-side validation (format and range checks) has been successful.
Output	A validated dataset within the ViDA portal, queued and ready for Star Rating computation.

5.4.2.5 Star Rating Computation (ViDA Internal)

Once the operational data has been successfully ingested, the ViDA platform automatically initiates the risk calculation engine.

As this is an internal, asynchronous process where ViDA's infrastructure combines the newly uploaded traffic parameters with the pre-existing infrastructure attributes in the project database. The computation generates specific safety scores for four road user groups: Vehicles, Motorcycles, Pedestrians, and Bicyclists. The integration pipeline monitors this progress through the assigned **job_id,** ensuring the calculation is finalized before proceeding to the retrieval phase

5.4.2.6 Download Star Rating results

After the computation is finished, the pipeline retrieves the final ratings to move the data from the cloud back into the local Aimsun environment.

Table 5.7: Script to download Star Rating results

Script	main.py download
Module	src/vida_api.py
API Interaction	GET /api/v1/datasets/{id}/star_ratings
Functionality	● Data Retrieval: The script authenticates and requests the full set of Star Rating results for the specified dataset ID. ● Format Handling: The JSON response from the API is parsed and converted into a local reference format (CSV/Excel).
Output	A file containing the Star Ratings, indexed by location_id (which corresponds to the Aimsun OID), ready for the import phase.

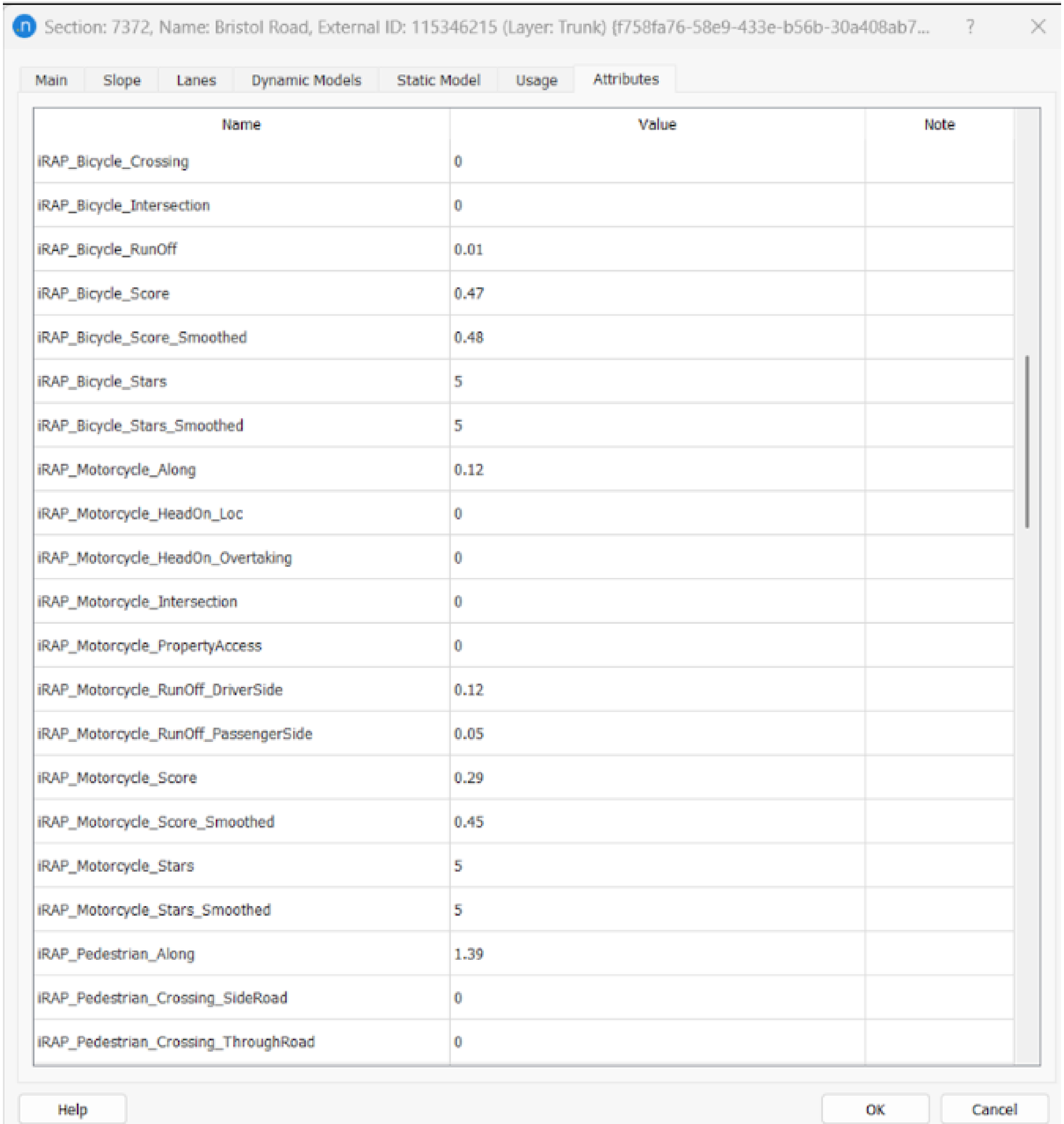
5.4.2.7 Import Star Ratings into Aimsun Next model

This step automates the enrichment of the Aimsun Next model by injecting the downloaded safety metrics back into the network's digital structure.

The import process is handled by a specialized Python script designed to interface with the Aimsun Next API. To ensure operational efficiency, the script is optimized for **aconsole.exe** execution. This approach allows the pipeline to perform structural modifications, such as creating new attributes and populating them with external data. By utilizing console.save() command, the integration ensures that all imported safety data permanently persisted in the .ang file without requiring manual intervention via the GUI.

Table 5.8: Script to import Star Ratings into Aimsun Next

Script	main.py import-ratings
Module	scripts/aimsun/import_star_ratings.py
Functionality	- The script reads the downloaded ratings and matches each record to the corresponding Aimsun section by pairing the ViDA location_id with the Aimsun OID. - Attribute Creation: It programmatically generates new external columns (e.g., iRAP_Star_Rating_Car) using the addExternalColumn() method. - Data Injection: The Star Rating values are assigned to each section via setDataValue(). - At the end of the process, the script executes a console.save() command to commit all changes to the model, ensuring the safety metrics are available for subsequent simulation runs or analysis.
Output	An updated Aimsun Next model (.ang file) containing new built-in attributes for each road section, populated with the latest Star Rating values from ViDA.



Name	Value	Note
iRAP_Bicycle_Crossing	0	
iRAP_Bicycle_Intersection	0	
iRAP_Bicycle_RunOff	0.01	
iRAP_Bicycle_Score	0.47	
iRAP_Bicycle_Score_Smoothed	0.48	
iRAP_Bicycle_Stars	5	
iRAP_Bicycle_Stars_Smoothed	5	
iRAP_Motorcycle_Along	0.12	
iRAP_Motorcycle_HeadOn_Loc	0	
iRAP_Motorcycle_HeadOn_Overtaking	0	
iRAP_Motorcycle_Intersection	0	
iRAP_Motorcycle_PropertyAccess	0	
iRAP_Motorcycle_RunOff_DriverSide	0.12	
iRAP_Motorcycle_RunOff_PassengerSide	0.05	
iRAP_Motorcycle_Score	0.29	
iRAP_Motorcycle_Score_Smoothed	0.45	
iRAP_Motorcycle_Stars	5	
iRAP_Motorcycle_Stars_Smoothed	5	
iRAP_Pedestrian_Along	1.39	
iRAP_Pedestrian_Crossing_SideRoad	0	
iRAP_Pedestrian_Crossing_ThroughRoad	0	

Figure 5.4: Irap ViDA Attributes with their correspondent values in the section 7372 after importation

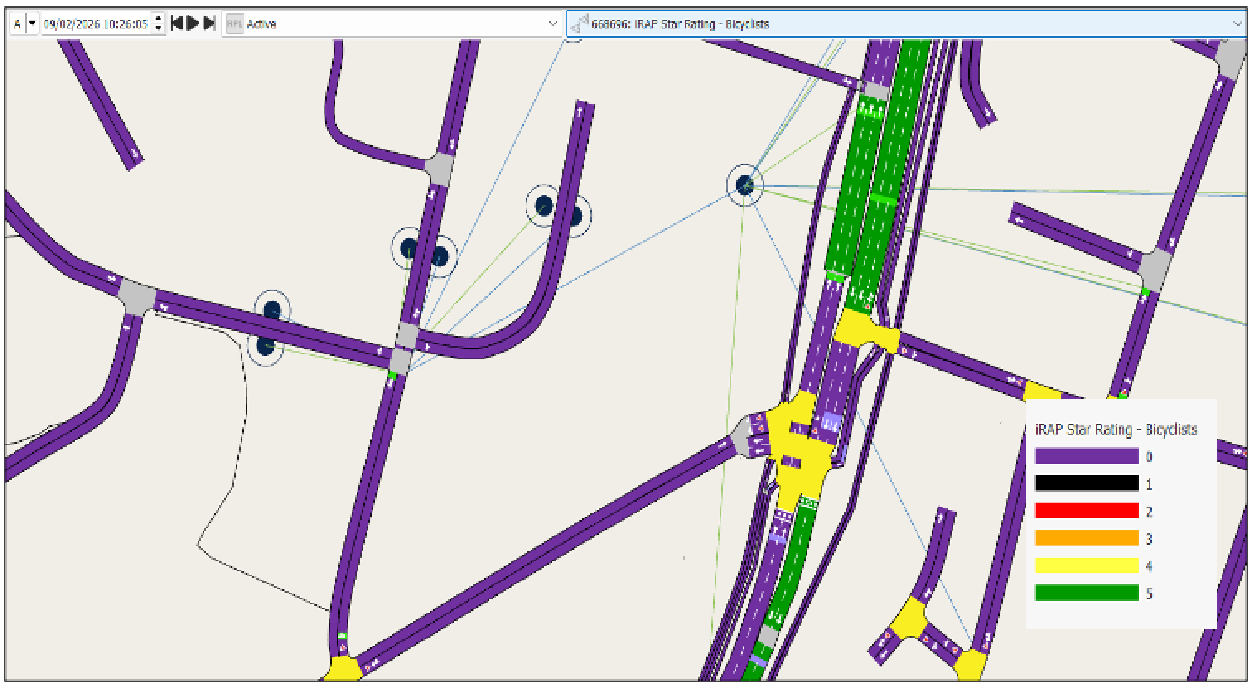
5.4.2.8 Configure Visualization in Aimsun Next model

The final stage of the pipeline focuses on translating the imported numerical star ratings into a visual safety audit, allowing for immediate identification of high-risk sections within the Aimsun Next environment.

Table 5.9: Script to configure view modes in Aimsun Next

Script	main.py view-modes
Module	scripts/aimsun/configure_view_modes.py
Functionality	- View Mode Instantiation: The script programmatically creates a new GKViewMode object in Aimsun Next specifically for iRAP safety metrics. - Standardized Symbology: The visualization follows the official iRAP color scheme confirmed by the partner, mapping risk levels to the standard 5-star palette: 5 Stars (Green), 4 Stars (Yellow), 3 Stars (Orange), 2 Stars (Red), and 1 Star (Black). - The View Mode containers, attribute linking, and color ramps (GKViewModeStyle) have been fully configured in accordance with iRAP specifications, including exact hex-code alignment with the official palette.
Output	A set of View Modes available in the Aimsun Next project, enabling users to toggle between different road user safety perspectives directly on the network geometry.

Figure 5.5: Example of a view mode called iRAP ViDA Star Ratings – Bicycles



5.4.3 Pipeline Summary

The following table provides a high-level overview of the entire technical workflow, mapping each step to its corresponding execution tool and API involvement.

Table 5.10: Pipeline summary of the integration between iRAP ViDA and Aimsun Next

Step	Action	iRAP ViDA API Interaction required	Execution
1-2	Extract and transform simulation data	No	main.py extract
3	Upload simulation results data to iRAP ViDA platform	Yes (POST)	main.py upload
4	iRAP computation	No	iRAP ViDA internal
5	Download iRAP ViDA rating scores	Yes (GET)	main.py download
6	Import iRAP ViDA rating scores to Aimsun Next	No	main.py import-ratings
7	Visualization of iRAP ViDA Rating Scores	No	main.py view-modes

5.4.4 Configuration

The pipeline is configured through a ‘config.json’ file (template provided in ‘config.example.json’). Configuration includes:

- **Database paths**: Location of the Aimsun SQLite export file.

- **ViDA API credentials**: Application and user credentials (auth_id, api_key, private_key).

- **ViDA SDK path**: Location of the ViDA SDK modules directory.

- **Aimsun paths**: Path to ‘aconsole.exe’ and the Aimsun model (‘.ang’) file.

- **S3 settings**: AWS profile, bucket, and region for upload.

- **Dataset ID**: The ViDA dataset identifier (33798 for the current project).

Configuration priority: environment variables > config.json > defaults.

5.5 Outputs

Some examples of the iRAP-ViDA - Aimsun Next integration are included below:

Figure 5.6: Example of outputs iRAP ViDA Star Ratings - Motorcyclists within Aimsun Next

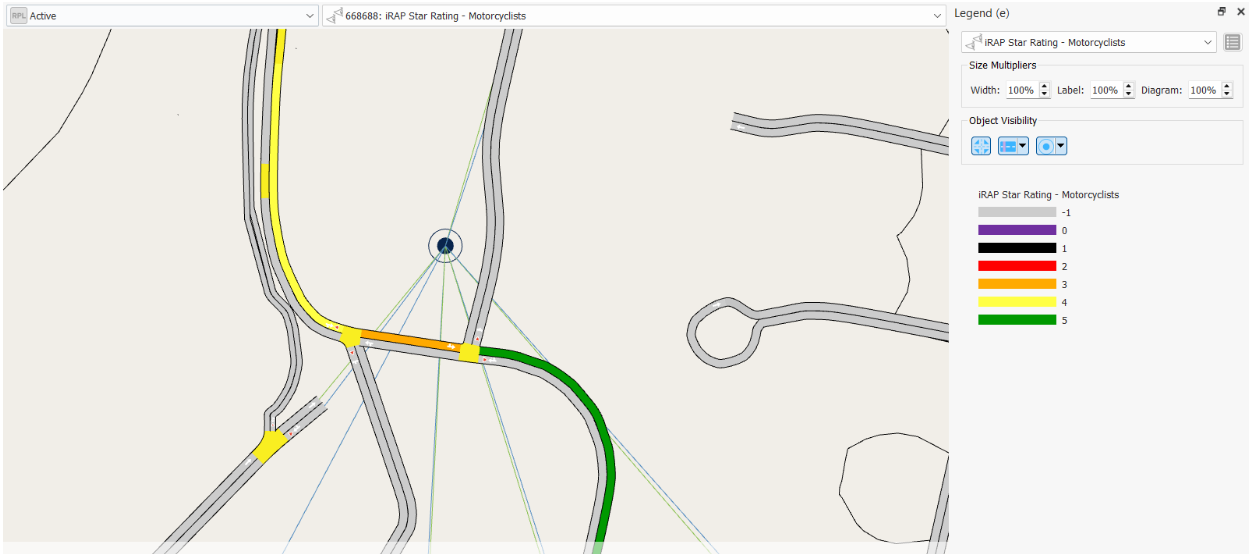


Figure 5.7: Example of outputs iRAP ViDA Star Ratings - Bicyclists within Aimsun Next

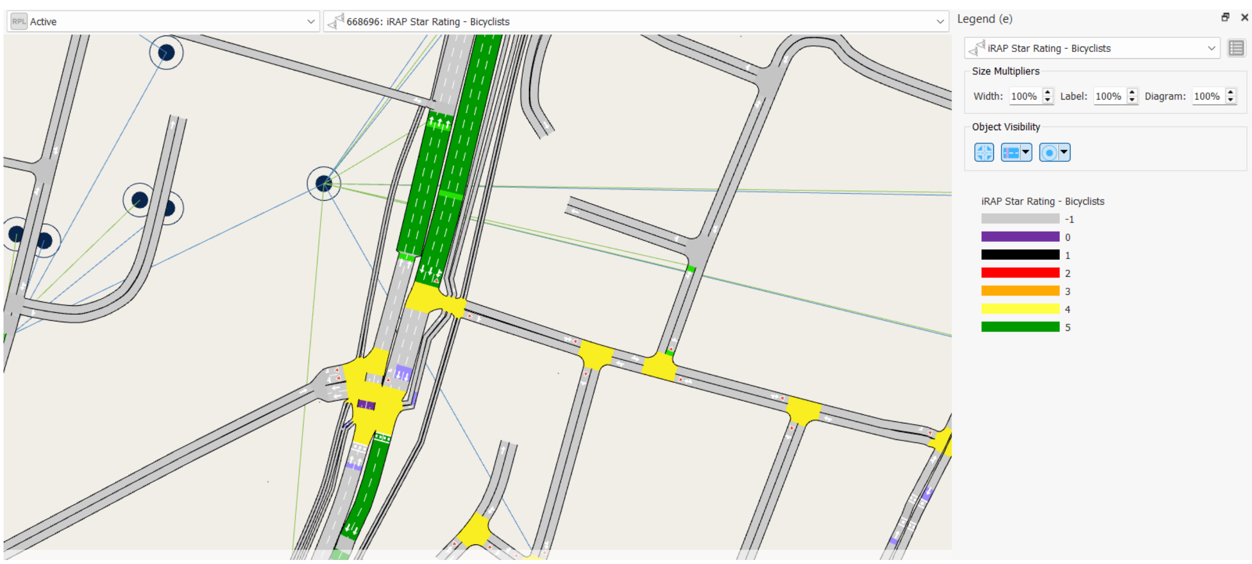
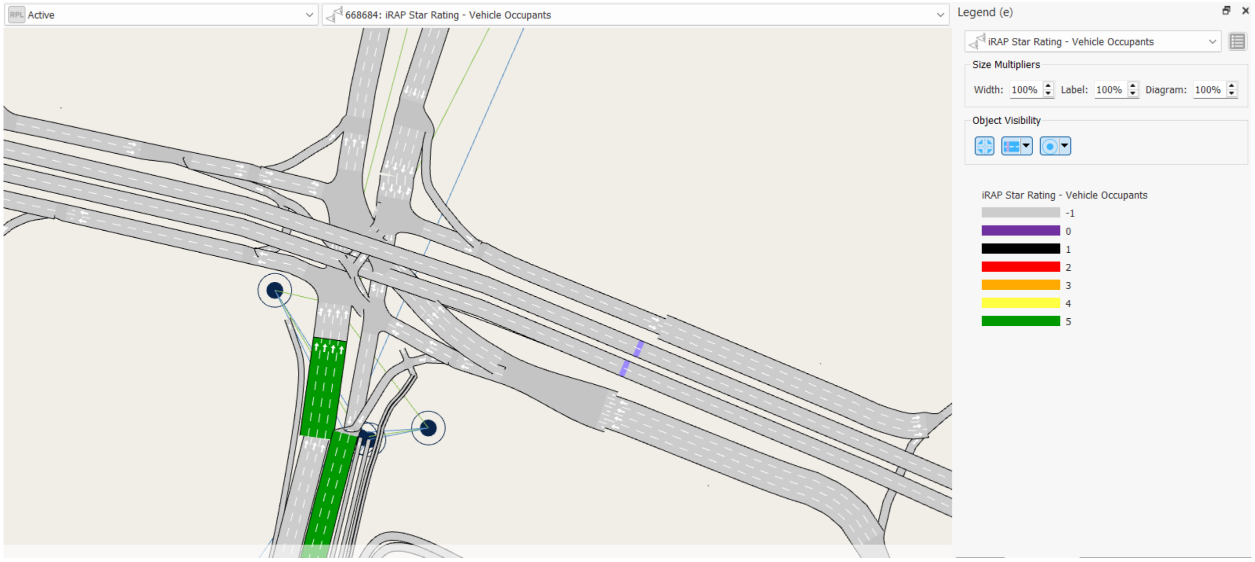


Figure 5.8: Example of outputs iRAP ViDA Star Ratings - Vehicle Occupants within Aimsun Next


Active
668684: iRAP Star Rating - Vehicle Occupants
Legend (e)
iRAP Star Rating - Vehicle Occupants
Size Multipliers
Width: 100% Label: 100% Diagram: 100%
Object Visibility
iRAP Star Rating - Vehicle Occupants
-1
0
1
2
3
4
5

6 Optimisation of Data Flow

6.1 Objectives

This section will establish how to assess the data available in terms of the needs of the project, protocols for handling missing data, detail the considerations that make sure the data adhere to FAIR principles, and outline legal considerations in terms of adherence to GDPR and the EU data act.

6.2 Methodology

PHOEBE identified several data-related challenges that had to be addressed during the project. To aid discussion, we make the distinction between *information*, the property of the system that we seek to understand, and the *data*, which is the raw material from which information can be derived. The fundamental question faced by any project is whether the data that can be obtained is sufficient to provide the requisite information to carry out the project successfully.

In many cases there are multiple data sources, following distinct methodologies and technologies, which produce data that can be used to derive a particular piece of information. However, each source will be different in terms of how well it fits the information requirements. For example, many methodologies exist for measuring the average traffic flow at a specific location. These methods include surveys that use pneumatic tubes that are laid across the carriageway, automated traffic cameras, radar, and telematics. Each of these sources is different in terms of the richness (transport mode, speed, counts, occupancy) and the quality (the coverage and resolution of the geospatial and temporal properties) of the resulting dataset. For example, a pneumatic tube survey will count only the passage of vehicles, and they are typically deployed on timescales of a day to a few weeks, which provides a very accurate record of all traversals by motor vehicles over a short timescale at a discrete location. The average traffic flow could be computed from these data; however, the results would only apply to that location and period, during which the observed behaviour could be atypical (e.g. during extreme weather events or road maintenance). Conversely, telematics datasets will have a large spatial coverage, potentially every road in a locality, but the data will be drawn from a subset of the population and will be aggregated over an extended period to allow for the sample to achieve statistical significance. Camera surveys are also carried out at discrete locations but may include pedestrian and active travel journeys; information that is increasingly in demand but for which there is a distinct dearth of credible data sources. Crucially, the consideration of different data sources is not always so trivial as to be reduced to a decision between two sources that can and cannot serve a piece of information. Indeed, PHOEBE has shown that there are many occasions where multiple datasets can be used in combination such as to mitigate for their individual limitations. This subsection aims to provide simple guidance to empower framework users to make informed decisions about their use of data, to show how these guidelines can be applied to the PHOEBE use cases, and to highlight some novel examples of how data were used in the project.

A qualitative schema was developed to formalise the process by which data are evaluated and discussed in terms of their ability to provide information. This schema is analogous to the iRAP Star Rating system, with the ability to assign a value from 1 (poor fit) to 5 (excellent fit) to map how well the source matches the needs. However, we urge users to consider the schema as a guide that will aid their evaluation of different data sources, rather than as a rulebook that should be applied meticulously.

We find that the data considerations can be grouped around six distinct types of information, which we label as:

- Road traffic collisions (RTC): Information that relates directly to the properties of collision events.

- Road properties: The geospatial layout and infrastructure of the road network.
- Traffic properties: Information that pertains to the flow of vehicles or people through the network.
- Land use: Information about the local context and environment of the road network.
- Transport properties: Information about the operation of the public transport system.
- Survey: Any information that is also tied to demographics.

Classification of data by information type

As a first step, users should classify each dataset according to both its information type (RTC, Road properties, etc.) and its data protection status:

- Non-personal data (e.g. static road geometry, aggregated land-use classifications)
- Pseudonymised personal data (e.g. vehicle-level telematics with indirect identifiers)
- Personal data (e.g. survey responses linked to personal identifiers)

This classification should be documented as part of the dataset metadata and revisited whenever datasets are combined or repurposed. To meet the needs for easy sharing and updating of models PHOEBE recommends the use of non-personal data that adheres to FAIR principles whenever possible. Pseudonymised data or data containing PII require preprocessing means such as to ensure anonymisation in full for the PHOEBE framework inputs.

Data Protection and Access Considerations

The PHOEBE framework can make use of raw data that may relate, directly or indirectly, to individuals or groups, particularly within the RTC, Traffic properties, telemetry data and Survey information types. Optimisation of data therefore requires careful consideration of data protection, privacy and emerging data access obligations, alongside the need to preserve analytical utility. As a basic principle the PHOEBE framework works with data from which any PII has been removed.

Given the large number of data types, PHOEBE provides guidance to help users determine appropriate safeguards when adapting the framework to new use cases and jurisdictions, rather than prescribing a single compliance model.

Minimisation and aggregation

Across all information types, PHOEBE encourages the use of the least granular data necessary to derive the required information. For example:

- RTC analyses may prioritise severity, location and timing without retaining individual-level identifiers
- Traffic properties can be derived from aggregated flow, speed or other measures without the need for individualised data.
- Survey data should be aggregated to demographic groupings sufficient for statistical robustness without the inclusion of PII in final used data supporting model configurations.

Where higher-resolution data are required for calibration or validation, access should be restricted and justified, with aggregated outputs used for downstream modelling wherever possible.

Anonymisation, access control and reuse

Users should assess whether datasets can be anonymised to a standard that removes all risk of identification, particularly when publishing or sharing data for reuse. Where anonymisation would significantly degrade the information content required by PHOEBE, alternative controls should be applied, such as:

- Pseudonymisation with separation of keys
- Secure access environments
- Data sharing agreements that restrict secondary or inappropriate reuse

These decisions should be made with explicit reference to the original consent and declared purpose of the original data and its intended reuse of the data, including the potential application of the PHOEBE framework to new policy questions or geographic regions. These considerations must follow data privacy and protection regulation. When processing has been performed this activity should be recorded in full as described by article 30 of GDPR.

6.2.1 Data Rating Schema

6.2.1.1 Road Traffic collisions

RTCs are relatively infrequent events; most European incident hotspots will perhaps have with maybe a handful of events per year. It is preferable for multiple years of data to be used in analysis, because of these low incidence rates. The completeness of the dataset is another major concern, i.e. the extent to which all events are captured. The severity of the incidents should also be considered, as policy makers are usually focused on tackling the number of FSI rather than all incidents of any severity. Methods exist for extrapolating the number of severe incidents at a location based on the number of observed fatalities, so the completeness at a given severity should also be considered.

5 Dataset was collected over three years or more of all serious and slight RTC along the route.

4 Dataset was collected over three years or more of all fatal RTC but has an incomplete record of serious injuries along the route.

3 Dataset is a partial (>50%) record of serious incidents along the route or is a complete record of fatalities over a period of 1-2 years.

2 Dataset is an incomplete record of fatalities along the route

1 Dataset is a fixed-point study, was collected for less than a year or on an ad hoc basis or followed an inconsistent or unusual methodology.

6.2.1.2 Road Properties

Spatial resolution and timeliness are the crucial concerns with respect to the properties of the road network. There is a strong preference for the data to relate to the operation of the road prior to the introduction of a policy change or mitigation.

5 Data were collected at a time when the road was in normal operation, over the full route, at a spatial precision of <1 m, within the previous 12 months.

4 Data were collected at a time when the road was in normal operation, over the full route, at a spatial precision of <10 m, within the previous 5 years.

3 Data were collected at a time when the road was in normal operation, or at a spatial precision of <50 m, or within the last 10 years.

2 Data were collected at a time when the road was undergoing significant changes or prior to the implementation of current infrastructure, or have a precision of > 100m, or are only viable for a fraction of the route.
1 Data are largely missing or demonstrably inaccurate.

6.2.1.3 Traffic Properties

There are many factors that are important for the evaluation of data that can be used to derive traffic properties; these include the extent to which the data are representative of the mode of interest, the time resolution of the final dataset, the applicability to the full route or sub-network, and the extent to which the traffic can be considered to be in normal operation.

5 Data were collected along the full route within the past year, to sub-hour time resolution, on days when the traffic patterns can be considered typical. Data are representative of the mode in question.
4 Data were collected along the partial route within the past two years, perhaps to sub-day time resolution, on days when the traffic patterns can be considered typical. Data are representative of the mode in question or have been reliably corrected for any biases that may be present.
3 Data were collected at fixed locations, or within the past five years, perhaps to sub-day time resolution, on days when the traffic patterns can be considered typical. Data are representative of the mode in question or have been reliably corrected for any biases that may be present.
2 Data were collected at locations outside of the route, or more than 5 years ago, perhaps time resolution greater than a day, on days when the traffic patterns can be considered typical. Data are not representative of the mode in question and have not been reliably corrected for any biases that may be present.
1 Data were collected only during times when the traffic patterns can be considered atypical.

6.2.1.4 Land use

The completeness and spatial resolution of the data are of primary concern when evaluating the extent to which a dataset can provide information on land usage. Users should also consider whether such data reflects any official view of land use in the area, such as might be maintained by national or local governments and affiliates.

5 Data are complete for the route and were obtained within the previous year, with positions that are accurate to sub-metre resolution. The resource is maintained by, or in collaboration with, local or national governments.
4 Data are complete for the route and were obtained within the previous year, with positions that are accurate to sub-metre resolution. The resource is maintained by a third party but is considered trustworthy.
3 Data are partially complete for the route or were obtained within the previous 5 years, positions are accurate to 10m resolution. The resource is maintained by a third party but is considered trustworthy.
2 Data coverage is sporadic or the dataset is older than 5 years, no precise positional information, perhaps assigned to road names. The resource is maintained by a third party or obtained as part of a bespoke survey.
1 Data are very irregular and are older than 10 years, in aggregate with no specific location information, collected in ad hoc methodology.

6.2.1.5 Transport System

Timeliness and completeness are of paramount concern when evaluating transport system data, but users should also consider the origin of the dataset. One indication that the data have been produced reliably is that they are available in General Transit Feed Specification (GTFS) format. Ideally, the data should be able to give an understanding of the number of people travelling between locations as well as the number and frequency of trips by individual vehicles.

5 Route information is maintained by official authority, collected within the previous 12 months, comprehensive account over all days, and maintains accurate records of performance for a given mode. Contains record of passenger numbers. All data are published in GTFS format.

4 Route information is maintained by official authority but could be out of date, presented in summary, and maintains accurate records of performance for a given mode. No record of passenger numbers. All data are published in GTFS format.

3 Route information comes from a third-party source and could be out of date, presented in summary, partial records of performance for a given mode. No record of passenger numbers. All data are published in accessible formats with suitable documentation.

2 Route information comes from a third-party source, is more than 5 years old, covers a part of the study region only. Data in an undocumented and unstandardised format.

1 Route information is anecdotal or contains information that is demonstrably archaic (for example, decommissioned bus stops). Data not collected in a consistent format.

6.2.1.6 Survey

The most important concerns for data that is used to understand the properties and opinions of the population is that these data are drawn from the specific subset of the population that has relevance to the project, that this dataset is both representative and has sufficient statistical power to yield meaningful results, and that the data are timely.

5 Data mapping the behaviours of interest to specific demographics. The Data collected exclusively from the study region, within the previous two years, with a sample size of >5000.

4 Data mapping the behaviours of interest to specific demographics or data containing the same acquired by combining different datasets. The Data collected exclusively from the study region, within the previous two years, with a sample size of >5000.

3 Data collected from the area containing the study region, or is more than 2 years old, with a sample size of 500 to 2500 or demographically biased groups (e.g. survey set out to contact a particular age range only).

2 Data collected from other regions, more than 5 years old, sample size < 500, unable to determine usual demographics.

1 Data do not contain the quantities of interest.

6.2.2 Data Hierarchy and Redundancy

This system should empower users to reliably appraise each data source in terms of its suitability to provide the required information. Indeed, where multiple data sources exist, the user should be able to quickly establish these sources into a hierarchy of preferred sources. Such hierarchical thinking lends itself to a system of redundancy for occasions where preferred sources are not available.

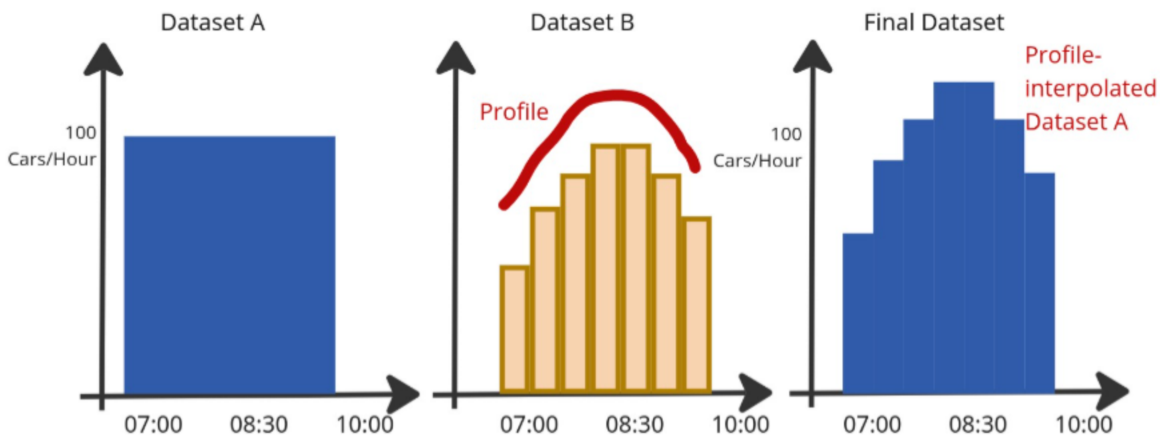
PHOEBE use cases applied the schema to their data records to establish the extent to which their needs could be met by the different sources.

6.2.2.1 Interpolation

Interpolation of data is so standard a practice in transport planning that it is rarely acknowledged explicitly. There is a great dearth of available data in the current epoch compared to what one might expect in the future, when potentially every road to be monitored in a multitude of ways such as to obtain a highly accurate and timely ‘digital twin’ of the whole network. Many datasets are obtained at discrete locations or during a brief period. Thus, the data may relate directly to a tiny proportion of the operation of the network. Often, this interpolation takes the form of an assumption of homogeneity at small scales; that a measurement at one location, or at one time, is probably applicable elsewhere in the vicinity. For example, measurements that can be used to determine the fleet mix or traffic flow might be obtained at fixed locations, but we might choose to assume their values hold for the full length of a particular road, or for a whole neighbourhood, in the absence of evidence to the contrary.

One novel example of interpolation was carried out as a part of the West Midlands use case. The microsimulation modellers needed to understand the traffic flow over the model area at a time resolution of 30 minutes. Several sources of data were available from which traffic flow could be obtained. These data included an aggregated telematics dataset (traffic properties rating 4), camera surveys (rating 3), and the Annual Average Daily Traffic Flow (AADF) dataset (rating 3). The telematics dataset presented flow estimates for every road in the sample, but those data applied only to private cars and where the traffic flow was measured over 3 hour intervals only. The AADF and camera survey data were obtained from only a handful of sites within the sample, however, they contain a record of the fleet mix measured at intervals of 15 minutes. To obtain the required information, the modellers extracted the profile of the diurnal variation in traffic flow and applied that to the aggregated telematics data to interpolate the total flow over the whole network at 30-minute resolution. Therefore, it was possible to overcome the limitations of individual data sources by combining them in a complementary manner.

Figure 6.1: Combination of individual data



A similar exercise was carried out to calibrate the car flow telematics data to those reported in the AADF at discrete locations. This analysis provided a measure of the relative flow on those streets for which AADF were not available. This analysis was submitted as part of the AiRAP accreditation process. The Floow is continuing to advance this work through experimentation with machine learning to determine whether other properties of the traffic, such as fleet mix and speeds, are recoverable from the telematics and contextual data. At the same time, the telematics analysis pipeline is being streamlined and placed into the cloud such as to improve the accuracy and readiness of the aggregate datasets.

6.3 FAIR Assessment

The PHOEBE data management approach aims to support transport management and enable scientific exploitation beyond the project by ensuring responsible data reuse and to maximise the use of the project data. This target follows the open science principles of EU legislation and good practice guidance related to data management in line with the EU Data Governance and the EU Data Acts. In particular, PHOEBE assesses the degree to which the available data sources align with so-called 'FAIR' principles, by which data can be described as being findable, accessible, interoperable, and reusable. The proposed formation of a European mobility data space (EMDS, https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52023DC0751) will lead to the creation of a FAIR-compliant framework for interlinking and federating different transport data ecosystems.

PHOEBE has embedded an assessment of all framework input data into the publicly-available data registers (Deliverable 2.2, Phoebe data registers, https://doi.org/10.5281/zenodo.16737146). This assessment operates as a complementary layer to help govern how datasets, metadata, and processing outputs are documented, stored, and shared. This register and its template format is designed for reuse by future users of the PHOEBE framework.

This register framework to support FAIR assessment is applied after a dataset has been identified as potentially suitable as a given information type from the PHOEBE data taxonomy (e.g. RTC, Road properties, Traffic properties, Land use, Transport system or Survey). For each data source an assessment is required to determine whether that source can be found, understood, combined, and used for the immediate application usage. The register also captures data supporting reuse beyond the initial framework usage to encourage reuse and wider ecosystem value.

Metadata and documentation

Figure 6.2: Register fields example

Data Resource Fields	Decsription of the field and its contents	Example singular data record
Data ID	Unique identifier for data used in the project.	D43
Data title	Textual Description of the data.	Valencia - speed limits
Appears in register	Prior linkage to D2.1 internal registers (project usage only).	PRODUCER Data records Information-Data-Producers
Data type	Type of Data in relation to the project.	Stakeholder/Regional data resources
Data Owner	Data ownership in relation to the project.	EXTERNAL STAKEHOLDERS
Data provider to the project	Project partner responsible for the data resource used within the project.	UPV
Use Cases covered	Use cases where data is applicable to allow filtering.	Valencia
Known data VS expected data	If data is 'known' and available or 'expected' data that may not yet be available	Known already
Part of project dataset (T2.1.2)	Link to managed data areas as declared in D7.3 Data Management Plan and its related records.	DMPDS4 - Collated Data from the use case regions and stakeholders
Relation to workpackage(s)	Relationship to specific developemntal workpackages in the PHOEBE project to allow filtering by parts of the project - or its usage after the project.	WP2
Ethical review identifier	Ethical review identifiers that pertain to the data usage within the project to ensure meeting the requirements of ISO 26000 for ethical handling. (Each identifier relates to an ethical check documented in D7.4 Ethical Management Plan and its related registers).	E8 E9
DPIA ID	Data privacy Impact Assessment (DPIA) records related to this data and its usage in the project data processing. Each record relates to records held in the project register 'GDPR and AI processing register' related to D7.3. Thjis register incorporates all data privacy impact assessments for project data and processing.	DPIA 5 DPIA 6 DPIA 7
Brief description of data	A textual description of the data resource to enable understanding of its content and potential usage.	Georeferenced database of speed limit for each street
Location of Data	Location or means of access to data used in the project to enable reuse of prior and new data resources used in project work. This may be a URL or can be a contact point for requested access if given a need to request access to the resource for onward usage. PLEASE NOTE not all data is available for onward reuse to protect propriatry data that the project may of utilised OR for partners to protect potential exploitation of data involved in the project work.	https://valencia.opendatasoft.com/explore/dataset/velocitat-carrers-velocidad-calles/export/ AND https://valencia.opendatasoft.com/explore/dataset/velocitat-carrers-velocidad-calles/information/?location=16,39.50133,-0.37493&basemap=e4bf90
FAIR Rating	An indicator reviewing the FAIR staus of the data resource and its overall potential for sharing and reuse.	Already FAIR
FAIR Reasoning	This details any reasons behind and OPEN or restricted designation for FAIR sharing of data.	Used in project model configuration and enhancements
Data Size Estimation	This is an upper estimate of data size bands to help understand the size of the resource should a user wish to obtain the data for direct or indirect reuse.	Under 100 Meg
Temporal extent	Information regarding the temporal extent of the data resource to understand the range of real world measurement, representation or capture that the data may contain.	current only
Geographical extent	The geographical coverage extent detailing the coverage of the available resource.	Valencia
Current Accessibility	Accessibility filter by means of accessibility of the data resource.	Accessible (open URL)
Current Interoperability	Interoperability filter to select by means of interoperability readyness from the data sources.	In common format allowing reuse
License for reuse	License that applies for potential reuse of the data beyond the project.	Creative Commons License
Keywords for reusability - to aid findability	Keywords about the data resource to allow finding and fast understanding of its content.	Valencia Tabular data Driver Behaviour Car Occupants supporting data

To encourage the reuse of data, each dataset used within PHOEBE, the existing registers (detailed in Deliverable D2.2 https://doi.org/10.5281/zenodo.17206476) capture information such as the data source and its FAIR assessment. The register fields (part of https://doi.org/10.5281/zenodo.16737146) are summarised in image Figure 6.1 with an example entry. These records provide a minimal but consistent metadata record that allows datasets to be recorded for use both in and beyond the PHOEBE framework.

Interoperability across data types

PHOEBE use cases frequently required the integration of datasets that differ significantly in origin, scale and structure. For example, linking traffic properties with RTC or demographic records required additional processing to relate the one source to another. To create interoperable data, users should:

- Use standardised, open, machine-readable formats whenever possible.
- Using a clearly defined geospatial referencing system (preferably WGS 84) and date & time conventions (preferably ISO 8601 date formatting, YYYY-MM-DDThh:mm:ss),
- Create mapping tables or schemas where variables in different sources represent the same quantity.

- Record and share any assumptions or alignment procedures that are followed to create interoperable data.

These steps are particularly important where data is likely to be reused in new areas or policy interventions, where implicit assumptions from the original use case may no longer hold.

Reproducibility and archiving

To support reproducibility, all procedures that are followed to derive information from raw data should be preserved. These procedures will include scripts, configuration files, any adjustable parameter settings, and a record of software versions. Users are encouraged to archive curated datasets, metadata, and processing artefacts in a persistent research repository (e.g. Zenodo, see PHOEBE archive at https://zenodo.org/communities/phoebe/). The use of repositories enables the assignment of a DOI and long-term discoverability.

When the sharing of data is restricted due to sensitivity or licensing constraints, metadata and descriptions of the processing pipeline should still be published with clear details on FAIR restrictions and how access to the restricted data may be requested.

Emerging data access obligations

Two significant EU acts pertaining to the access and use of data became applicable during the course of the PHOEBE project; the Data Governance Act (DGA) and the Data Act. Both contain additional data access and sharing obligations that are particularly pertinent for discussion here in the context of connected vehicles or the provision of digital services. Where relevant, these considerations should be documented alongside GDPR assessments so that future users of the PHOEBE framework understand both the legal constraints and the conditions under which data may be reused.

European Data Governance Act

The Data Governance Act (DGA, EU 2022/868) became applicable in September 2023 and aims to regulate the sharing and reuse of protected data held by the public sector. These regulations include limits on the use of exclusive reuse agreements between public bodies and third parties and stipulations that public bodies should assist potential users whenever possible, such as by removing PII from the data or by providing access from secure locations only. The commission created a searchable register for protected data held by the public sector, the ERPD, which acts as a single tool for accessing information compiled by the information contact points for each nation.
The DGA has positive implications for the future use of the PHOEBE framework in that public sector data is increasingly more accessible and that there are clear routes and guidelines to follow when crucial data is seemingly beyond reach. These provisions also allow for the easier sharing of data between sectors, which could open the door to much more novel and innovative data reuse in the future. For example, there has been some recent research activity that has sought to understand the completeness of RTC data using health records, and to tie incidents to health outcomes[15]. The combined datasets offer very promising avenues for further investigation that could greatly improve our understanding of when road risk poses a significant threat to quality of life.

[15] For summary, Linking police and health data on road collisions 2025, UK DfT

European Data Act

The Data Act (EU 2023/2854) became applicable in September 2025 and is designed to increase the availability of data held by the private sector. At an individual level, the Data Act empowers users to access the raw data that they generate while using connected products or services. These data can then be passed on to other businesses for the supply of additional services.

From the perspective of the PHOEBE framework, the most significant policy changes concern the interaction between the public sector and private (industrial) data handlers. Under chapter 5 of the act private sector bodies will now be obliged to provide access to data on request from a public body when that data is deemed to be essential, within certain conditions, to help to address an exceptional need. 'Exceptional need' refers to both time-sensitive emergencies, such as pandemics, as well as ongoing threats to public safety, such as road risk. For example, Birmingham City Council declared there to be a road safety emergency in the city, requiring emergency action, in late 2024. Had the UK been subject to the EU Data Act at that point then the city council or other public stakeholders could have obliged private companies to provide any raw data that has the potential to provide important insights into the road safety situation. In this scenario the onus is on the public body to demonstrate that the same or equivalent data cannot be obtained by another means and the scope to use the data would be limited to tackling the designated exceptional need. Furthermore, these data would not become public sector information and would not be eligible for re-use under the European Open Data Directive.

There are significant stipulations regarding data interoperability. A party wishing to participate in the EMDS must adhere to strict requirements that ensure the properties, structure, and restrictions on use of the data are well documented and machine readable. These data should also have adjacent technical means and documentation such as to facilitate their automatic access and continuous transfer. This approach bodes well for the future implementation of the PHOEBE framework as users will be able to quickly obtain and utilise multiple data sources by expending considerably less effort.

7 Conclusions

7.1 Recapitulation

This deliverable has brought together the main outcomes of this section of the project and clarified how the PHOEBE framework can be applied beyond individual use cases. Focus has been given on consolidating methods, results, and workflows developed earlier in the project and aligning them into a coherent, transferable process for predictive safety assessment at the urban network level.

Across the three pilot cities-Athens, Valencia, and the West Midlands-the analysis showed that direct comparison of absolute values is not appropriate due to differences in network structure, interventions, and data availability. To address this, the framework adopts Impact Modification Factors (IMFs) to express changes between scenarios in relative terms. This approach allows safety, mobility, and behavioural impacts to be interpreted consistently while preserving local context.

It has also been demonstrated how different modelling components can be combined in practice. Traffic simulation, risk assessment, behavioural modelling, demand modelling, and socio-economic analysis are integrated in a way that supports the transfer of insights from detailed, local analyses to broader network-level assessments. This multi-scale perspective is essential for evaluating urban interventions that affect multiple road user groups and objectives simultaneously.

A further contribution of this work is the testing of dynamic links between simulation outputs and road safety assessment tools. The manual and semi-automated connectivity between Aimsun Next and the iRAP Star Rating methodology has shown that operational traffic conditions, such as speed and flow, can be systematically translated into updated safety indicators. This creates the basis for assessing how safety performance evolves under different traffic and policy scenarios.

Finally, key issues have been addressed related to data flow, interoperability, and governance. The protocols described in this deliverable aim to support reproducibility, transparency, and compliance with data protection requirements, while enabling the framework to be scaled to other cities and applications.

7.2 Practical Recommendations

Based on the experience gained, several practical recommendations can be drawn for future applications of the PHOEBE framework and similar predictive safety assessment approaches.

First, scenario-based analysis should prioritise relative changes rather than absolute indicators. Using metrics such as IMFs helps decision-makers focus on the direction and magnitude of change introduced by an intervention, which is more robust across different urban contexts. Second, early attention should be given to network compatibility and data alignment. Establishing stable identifiers and consistent spatial references between simulation models and risk assessment databases significantly reduces effort later in the workflow and improves reliability of results. Third, automation of data exchange should be pursued wherever possible. While manual pilots are valuable for validation, automated workflows are essential for scaling the framework to larger networks, testing multiple scenarios, and supporting iterative planning processes.

Additionally, behavioural and demand responses should be treated as integral components of safety assessment rather than as secondary effects. Interventions may alter compliance, mode choice, and exposure, particularly for vulnerable road users, and these changes can strongly influence overall safety outcomes. Finally, transparency in assumptions, data limitations, and uncertainty remains critical. Results should be communicated with clear explanations of their scope and constraints, especially when transferred to new contexts or used to support policy decisions.

These recommendations provide guidance for the continued development and application of the PHOEBE framework and support its transition toward a fully operational, user-oriented implementation, which is further addressed in the subsequent stages of the project.

8 References

Aimsun (2025). Aimsun Next version 24.0.2 Users Manual [Online documentation]. Recuperado em 14 de agosto de 2025, de https://docs.aimsun.com/next/24.0.2/

Aimsun Next (2026). Aimsun Next User Manual, https://docs.aimsun.com/next/26.0.0/UsersManual/ApiArchitecture.html

Dirección General de Tráfico. (2022). Road Safety Strategy 2030 – Summary. Ministry of the Interior, Government of Spain. Available at: https://seguridadvial2030.dgt.es/export/sites/sv2030/.galleries/descargas/Road_Safety_Strategy_2030_Summary_EN.pdf

FHWA-1 Federal Highway Administration, "Crash Modification Factors (CMF) Clearinghouse, 2014.

iRAP. (2024). iRAP Star Rating and Investment Plan Manual (Version 1.5, English). iRAP. https://www.irap.org/

Ghaffarpasand, O., Burke, M., Osei, L. K., Ursell, H., Chapman, S., & Pope, F. D. (2022). Vehicle telematics for safer, cleaner and more sustainable urban transport: A review. Sustainability, 14(24), 16386. https://doi.org/10.3390/su142416386

Kumar, A., & Mudgal, A. (2025). Surrogate safety assessment in heterogeneous traffic environment prevailing in developing countries: a systematic literature review. International Journal of Injury Control and Safety Promotion, 32(2), 182–200. https://doi.org/10.1080/17457300.2025.2494209

Mustapha, A., Abdul-Rani, A. M., Saad, N., & Mustapha, M. (2024). Advancements in traffic simulation for enhanced road safety: A review. Simulation Modelling Practice and Theory, 137, 103017. https://doi.org/10.1016/j.simpat.2024.103017

Paliotto, A., Meocci, M., Terrosi, A., & La Torre, F. (2024). Systematic review, evaluation and comparison of different approaches for the implementation of road network safety analysis. Heliyon, 10, e28391. Elsevier. https://doi.org/10.1016/j.heliyon.2024.e28391

Putatunda, A., Al Haddad, C., & Antoniou, C. (2025). A comprehensive review of the socio-economic appraisal methodologies of the road safety measures. Accident Analysis & Prevention, 217, 108021.

Theofilatos, A., & Yannis, G. (2014). A review of the effect of traffic and weather characteristics on road safety. Accident Analysis and Prevention, 72, 244–256. https://doi.org/10.1016/j.aap.2014.06.017

Transport for West Midlands. (2023). Refreshed Regional Road Safety Strategy 2023–2030. West Midlands Combined Authority. Available at: https://www.tfwm.org.uk/media/rmyd1vzf/refreshed-regional-road-safety-strategy-2023-2030.pdf

UK Department for Transport. (2026). Road safety strategy (large print version). Department for Transport.Available at: https://assets.publishing.service.gov.uk/media/69677c093eea543d3e7034ba/road-safety-strategy-large-print.pdf

Van Ditshuizen, J. C., Sewalt, C. A., Palmer, C. S., Van Lieshout, E. M., Verhofstad, M. H., Den Hartog, D., & Dutch Trauma Registry Southwest Soesman NMR 5 Jakma TSC 6 Waleboer M. 7 Staarink M. 8 Bruijninckx MMM 9 Cardon AYMVP 10 den Hoed PT 11 Roukema GR 12 van der Vlies CH 13 Schep NWL 14 van de Schoot L. 15. (2021). The definition of major trauma using different revisions of the abbreviated injury scale. Scandinavian journal of trauma, resuscitation and emergency medicine, 29(1), 71.

Van Essen, H., Van Wijngaarden, L., Schroten, A., Sutter, D., Bieler, C., Maffii, S., ... & El Beyrouty, K. (2019). Handbook on the external costs of transport, version 2019 (No. 18.4 K83. 131).

Yannis, G., Folla, K., Nikolaou, D. (2022). The Greek National Road Safety Plan 2021–2030: Linking and monitoring road safety targets, KPIs and measures. Paris: OECD Publishing.Available at: https://www.itf-oecd.org/sites/default/files/docs/greece-road-safety-strategy.pdf

Kalantari, A., Chaudhry, A., Burke, M., Zuriaga, A. M. P., Ziakopoulos, A., Papadimitriou, E., … Afghari, A. P. (2025, October 2). Social values behind the wheel: How does social value orientation shape speeding behaviour across drivers in Europe?. https://doi.org/10.31234/osf.io/348cf_v1

9 Appendix 1

Behavioural Models for Athens Use Case

Speeding Behaviour:

$$P\text{ (speeding)} = 1 - \frac{1}{1 + e^{(-(2.762 - U))}} \qquad (Eq.15)$$

Where:

P(speeding) is the probability of (proportion of traffic) going over the speed limit, and

U = – 0. 437× Traffic+0.894×Lanes-0.2120×Presence of physical median-0.676 ×Speed limit (50kmph)-0.932 ×Presence of speed camera –0.023 ×Age- 0.024 ×Social Value Orientation+1.012 × Past Behaviour:Disregarding the speed limit on urban highways

Table Appendix A1: Speeding Behaviour results for Athens Use Case

<table>
<tr><th colspan="2" rowspan="2">Variable*</th><th rowspan="2">Estimate</th><th rowspan="2">z value</th><th rowspan="2">Pr(>|z|)</th><th colspan="2">95% CI</th></tr>
<tr><th>L</th><th>U</th></tr>
<tr><td colspan="2">Intercept</td><td>2.762</td><td>4.849</td><td>0.000</td><td>1.645</td><td>3.879</td></tr>
<tr><td colspan="2">High traffic</td><td>-0.437</td><td>-13.033</td><td>0.000</td><td>-0.503</td><td>-0.371</td></tr>
<tr><td colspan="2">Multi-lane</td><td>0.894</td><td>20.935</td><td>0.000</td><td>0.810</td><td>0.978</td></tr>
<tr><td colspan="2">Presence of physical median</td><td>-0.212</td><td>-5.245</td><td>0.000</td><td>-0.291</td><td>-0.132</td></tr>
<tr><td>Speed limit</td><td>50 km/h</td><td>-0.676</td><td>-19.961</td><td>0.000</td><td>-0.742</td><td>-0.609</td></tr>
<tr><td colspan="2">Presence of speed Camera</td><td>-0.932</td><td>-27.213</td><td>0.000</td><td>-0.999</td><td>-0.865</td></tr>
<tr><td colspan="2">Age</td><td>-0.023</td><td>-2.359</td><td>0.018</td><td>-0.043</td><td>-0.004</td></tr>
<tr><td colspan="2">Social value orientation (SVO)</td><td>-0.024</td><td>-2.930</td><td>0.003</td><td>-0.041</td><td>-0.004</td></tr>
<tr><td>Past behaviour</td><td>Disregarding the speed limit on urban highways</td><td>1.012</td><td>7.961</td><td>0.000</td><td>0.763</td><td>1.262</td></tr>
</table>

<table>
<tr><th colspan="7">Model Fit</th></tr>
<tr><th>AIC</th><th>Loglikelihood at convergence</th><th colspan="4">Participant</th><th>Observations</th></tr>
<tr><td rowspan="2">31419.44</td><td rowspan="2">-15696.72</td><td>Count</td><td>Variance</td><td>Std</td><td>ICC</td><td rowspan="2">14976</td></tr>
<tr><td>468</td><td>5.79</td><td>2.40</td><td>0.637</td></tr>
</table>

Red Light Running Behaviour:

$$P\text{ (red light running by pedestrians)} = 1 - \frac{1}{1 + e^{(-(2.833 - U))}} \qquad \textit{(Eq.6)}$$

Where:

P(red light running by pedestrians) is the probability of (proportion of traffic) going over the speed limit, and

U = –1.098× Traffic -0.682×Crossing width+0.497×Beingin a rush-0. 024×Age+ 0.777 ×Past Behaviour:Red light violation+0.483 × Past Behaviour:Not looking when in hurry

Table Appendix A2: Red Light Running behavior results for Athens Use Case

<table>
<tr><th colspan="2" rowspan="2">Variable*</th><th rowspan="2">Estimate</th><th rowspan="2">z value</th><th rowspan="2">Pr(>|z|)</th><th colspan="2">95% CI</th></tr>
<tr><th>L</th><th>U</th></tr>
<tr><td colspan="2">Intercept</td><td>2.833</td><td>5.914</td><td>0.000</td><td>1.894</td><td>3.772</td></tr>
<tr><td colspan="2">High traffic</td><td>-1.098</td><td>-22.642</td><td>0.000</td><td>-1.193</td><td>-1.003</td></tr>
<tr><td colspan="2">Wide crossing width</td><td>-0.682</td><td>-14.383</td><td>0.000</td><td>-0.774</td><td>-0.589</td></tr>
<tr><td colspan="2">Being in a rush</td><td>0.497</td><td>10.534</td><td>0.000</td><td>0.404</td><td>0.589</td></tr>
<tr><td colspan="2">Age</td><td>-0.024</td><td>-2.950</td><td>0.003</td><td>-0.040</td><td>-0.008</td></tr>
<tr><td rowspan="2">Past behaviour</td><td>Red light violation</td><td>0.777</td><td>6.378</td><td>0.000</td><td>0.538</td><td>1.015</td></tr>
<tr><td>Not looking when in hurry</td><td>0.483</td><td>2.923</td><td>0.003</td><td>0.159</td><td>0.808</td></tr>
<tr><th colspan="7">Model Fit</th></tr>
<tr><th>AIC</th><th colspan="2">Loglikelihood at convergence</th><th colspan="3">Participant</th><th>Observations</th></tr>
</table>

<table>
<tr><td rowspan="2">31419.44</td><td rowspan="2">-15696.72</td><td>Count</td><td>Variance</td><td>Std</td><td>ICC</td><td rowspan="2">7488</td></tr>
<tr><td>468</td><td>4.32</td><td>2.07</td><td>0.567</td></tr>
</table>

*Age (years) is continuous.

Past behaviour is based on a 5-point Likert scale starting from ‘Never = 1’ to ‘Always = 5’.

The dependent variable, ‘Tendency to go over the speed limit’ was based on a 5-point Likert scale starting from ‘Extremely Unlikely = 1” to “Extremely Likely = 5” but for Aimsun implementation, the response variable is simplified to a binary outcome, grouping the three lower categories (Extremely Unlikely, Unlikely, Neutral) as 0, and the two higher categories (Likely, Extremely Likely) as 1.

Jaywalking behaviour of pedestrians, Athens

$$\text{P (jaywalking)} = 1 - \frac{1}{1 + e^{(-(0.569 - U))}} \qquad (Eq.16)$$

Where,

P(jaywalking) is the probability of (proportion of traffic) going over the speed limit, and

U= –0.560 × Traffic - 1.003×Crossing width+0.239×Two-way road+0.132 ×Distracted while crossing+0.409 × Past Behaviour: crossing despite the oncoming vehicle(s) approaching-0.026×Age-0.011×Social value orientation (SVO)

Table Appendix A3: Jaywalking behavior results for Athens Use Case

Variable*	Estimate	z value	Pr(>|z|)	95% CI	
				L	U
Intercept	0.569	1.589	0.112	-0.132	1.271
High traffic	-0.560	-17.562	0.000	-0.622	-0.497
Wide crossing width	-1.003	-30.876	0.000	-1.067	-0.940
Two-way road	0.239	7.507	0.000	0.177	0.302

<table>
<tr><td colspan="2">Distracted while crossing</td><td>0.132</td><td>4.177</td><td colspan="2">0.000</td><td colspan="2">0.0701</td><td>0.194</td></tr>
<tr><td>Past behaviour</td><td>Crossing despite the oncoming vehicle(s) approaching</td><td>0.409</td><td>4.733</td><td colspan="2">0.000</td><td colspan="2">0.239</td><td>0.578</td></tr>
<tr><td colspan="2">Age</td><td>-0.026</td><td>-4.257</td><td colspan="2">0.000</td><td colspan="2">-0.038</td><td>-0.014</td></tr>
<tr><td colspan="2">Social value orientation (SVO)</td><td>-0.011</td><td>-2.011</td><td colspan="2">0.044</td><td colspan="2">-0.022</td><td>-0.0002</td></tr>
<tr><td colspan="9">Model Fit</td></tr>
<tr><td>AIC</td><td colspan="2">Loglikelihood at convergence</td><td colspan="5">Participant</td><td>Observations</td></tr>
<tr><td rowspan="2">37822.72</td><td rowspan="2" colspan="2">-18899.36</td><td>Count</td><td>Variance</td><td colspan="2">Std</td><td>ICC</td><td rowspan="2">14976</td></tr>
<tr><td>468</td><td>2.43</td><td colspan="2">1.56</td><td>0.425</td></tr>
</table>

*Age (years) and SVO (degree) are continuous.

Past behaviour is based on a 5-point Likert scale starting from ‘Never = 1’ to ‘Always = 5’.

The dependent variable, ‘Tendency to go over the speed limit’ was based on a 5-point Likert scale starting from ‘Extremely Unlikely = 1” to “Extremely Likely = 5” but for Aimsun implementation, the response variable is simplified to a binary outcome, grouping the three lower categories (Extremely Unlikely, Unlikely, Neutral) as 0, and the two higher categories (Likely, Extremely Likely) as 1.

Valencia Use-Case:

Valencia Speeding Behaviour:

Speeding Behaviour:

$$P\ (\text{speeding}) = 1 - \frac{1}{1 + e^{(-(0.410 - U))}} \qquad \textit{(Eq.17)}$$

Where:

P(speeding) is the probability of (proportion of traffic) going over the speed limit, and

U = – 0.190 × Traffic+0.489×Multi-lane road+0.308×Presence of physical median-0.768 ×Speed limit (30kmph)-1.341 ×Speed limit (40kmph)-1.767× Speed limit (50kmph)–1.076 ×Presenceof speed radar – 0.050 ×Age+ 0.065 ×Driving exposure+0.578 × Past Behaviour: Disregarding minimum headway

Variable Distributions

Table Appendix A4: Variable Distributions for Valencia Use Case

Variable	Description	Type	Distribution
*Crossing Width	Width of the crossing facility	Binary 0: Narrow 1: Wide	This variable comes from the network. No distribution required.
**Traffic	Traffic (condition) at the intersection approach (with crossing)	Binary 0: Low 1: High	This variable comes from the traffic simulation. No distribution required.
Distance to Approaching Vehicle(s)	Distance of an approaching vehicle to a micromobility user at the crossing	Binary 0: Close 1: Far	This variable comes from the traffic simulation. No distribution required.
Social Value Orientation	Personal characteristic of a micromobility user	Continuous (degree)	
Age	The age of a micromobility user	Continuous	
Past Behaviour: Red Light Running on Crossings	Past red light running events by a micromobility user at marked crossings (scale)	Continuous (Likert scale)	Log-Normal
Past Behaviour: Riding out of Bicycle Lane	Past events by a micromobility user involving riding out of bicycle lane (scale)	Continuous (Likert scale)	Log-Normal

Model Results

Table Appendix A5: Speeding Behaviour results for Valencia Use Case

Variable*		Estimate	z value	Pr(>|z|)	95% CI	
					L	U
Intercept		-0.410	-0.421	0.673	-2.321	1.500
High traffic		-0.190	-2.613	0.008	-0.333	-0.047
Multi-lane road		0.489	4.543	0.000	0.278	0.700
Presence of physical median		0.308	2.874	0.004	0.098	0.518
Speed limit	30 km/h	-0.768	-3.507	0.000	-1.197	-0.338
	40 km/h	-1.341	-4.679	0.000	-1.902	-0.779
	50 km/h	-1.767	-7.382	0.000	-2.236	-1.298
Presence of speed radar		-1.076	-14.241	0.000	-1.224	-0.928
Age		-0.050	-3.260	0.001	-0.080	-0.020
Driving exposure (hr/week)		0.065	2.957	0.021	0.020	0.108
Past behaviour	Disregarding minimum headway	0.578	2.717	0.006	0.161	0.995
Model Fit						

AIC	Loglikelihood at convergence	Participant				Observations
		Count	Variance	Std	ICC	
7447.72	-3708.86	124	4.19	2.06	0.560	3224

*Age (years) is continuous.

Past behaviour is based on a 5-point Likert scale starting from ‘Never = 1’ to ‘Always = 5’.

The dependent variable, ‘Tendency to go over the speed limit’ was based on a 5-point Likert scale starting from ‘Extremely Unlikely = 1” to “Extremely Likely = 5” but for Aimsun implementation, the response variable is simplified to a binary outcome, grouping the three lower categories (Extremely Unlikely, Unlikely, Neutral) as 0, and the two higher categories (Likely, Extremely Likely) as 1.

Red light running behaviour of micromobility users, Valencia

Speeding Behaviour:

$$P\text{ (red light running by micromobility users)} = 1 - \frac{1}{1 + e^{(-(2.867 - U))}} \qquad (Eq.18)$$

Where

P(redlight running by micromobility users) is the probability of a micromobility (cyclist/e-scooter) user to run the red light at an intersection with marked crossing

and

U = – 0.256 × Crossing Width – 0.541× Traffic + 1.087 × Distance To Approaching Vehicle(s)– 0.036 × Social Value Orientation – 0.032 × Age + 0.587 × Past Behaviour: Red light Running On Crossings + 0.802 × Past Behaviour: Riding Out Of Bicycle Lane

Model Results:

Table Appendix A7: Red Light Running by micromobility users behavior results for Valencia Use Case

Variable*	Estimate	z value	Pr(>|z|)	95% CI	
				L	U
Intercept	2.867	3.479	0.000	1.251	4.483

Wide crossing width		-0.256	-2.614	0.008	-0.448	-0.064
High traffic		-0.541	-5.493	0.000	-0.734	-0.348
Far distance to approaching vehicle(s)		1.087	0.100	0.000	0.889	1.285
Social value orientation (SVO)		-0.035	0.163	0.015	-0.064	-0.006
Age		-0.032	0.164	0.018	-0.059	-0.005
Past behaviour	Red light violation	0.587	0.014	0.000	0.265	0.908
	Riding out of bicycle lane	0.801	0.013	0.000	0.478	1.124

Model Fit						
AIC	Loglikelihood at convergence	Participant				Observations
		Count	Variance	Std	ICC	
4115.25	-2045.63	114	2.7	1.6	0.450	1824

*Age (years) and SVO (degree) are continuous.

Past behaviour is based on a 5-point Likert scale starting from ‘Never = 1’ to ‘Always = 5’.

The dependent variable, ‘Tendency to go over the speed limit’ was based on a 5-point Likert scale starting from ‘Extremely Unlikely = 1” to “Extremely Likely = 5” but for Aimsun implementation, the response variable is simplified to a binary outcome, grouping the three lower categories (Extremely Unlikely, Unlikely, Neutral) as 0, and the two higher categories (Likely, Extremely Likely) as 1.

Validation Results:

WM Use-case:

Figure Appendix A1: Two-way, Two-lane, Undivided Road for WM

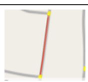
Name of Road	Configuration	Characteristics	Actual proportions_Peak	Actual proportions_Peak_95% CI	Actual proportions_off_peak	Actual proportions_off_peak_95% CI	Survey proportions_peak	Survey proportions_off_peak	Survey proportions__peak (Speeding Camera Present)	Survey proportions_low_off_peak (Speeding Camera Present)	CF_peak	CF_off_peak
Wash Lane		Two-way, two lane roadway (1 lane in each direction) Lane width (≥ 3.25m) Speed limit category=20mph No speed camera present Section Length=98.75m	0,1329	[0.1224-0.1433]	0,1537	[0.1474-0.1600]	0,6655	0,7497	0,3778	0,4701	0,1997	0,2050
Clay Lane		Two-way, two lane roadway (1 lane in each direction) Lane width (≥ 3.25m) Speed limit category=30mph No speed camera present Section length=66.77m	0,1262	[0.1089-0.1434]	0,1907	[0.1781-0.2033]	0,6290	0,7174	0,3431	0,4324	0,2006	0,2658
Tennant Street		Two-way, two lane roadway (1 lane in each direction) Medium (≥ 2.75m to < 3.25m) Speed limit category=20mph No speed camera present Section Length=86.6	0,0000	[nan-nan]	0,0000	[nan-nan]	0,6655	0,7497	0,3778	0,4701	—	—
Wash Ln		Two-way, two lane roadway (1 lane in each direction) Wide (≥ 3.25m) Speed limit category=30mph No speed camera present Section Length=157.263	0,1387	[0.1046-0.1728]	0,1361	[0.1165-0.1556]	0,6290	0,7174	0,3431	0,4324	0,2205	0,1896
Church Road: Sheaf Lan		Two-way, two lane roadway (1 lane in each direction) Wide (≥ 3.25m) Speed limit category=30mph No speed camera present Section Length=75.43	NaN	NaN	NaN	NaN	0,6290	0,7174	0,3431	0,4324	—	—
Holder Road		Two-way, two lane roadway (1 lane in each direction) Wide (≥ 3.25m) Speed limit category=30mph No speed camera present Section Length=154.46	0,0725	[0.0691-0.0759]	0,1150	[0.1096-0.1205]	0,6290	0,7174	0,3431	0,4324	0,1152	0,1603

Figure Appendix A2: Two-way Multilane Divided Road for WM

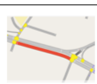
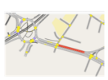
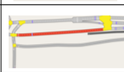
Name of Road	Speed limit (km/h)	Speed limit	Speed camera present	Configuration	No. of lanes (in each	Lane width (m)	Section length (m)	Actual proportions_Peak	Actual proportions_Peak_95% CI	Actual proportions_off_peak	Actual proportions_off_peak_95% CI	Survey proportions_peak	Survey proportions_off_peak	CF_peak	CF_off_peak
Islington Row Middleway	64.4 km/h in Aimsun (40km/h in iRAP data)	40 mph	Yes		3	Wide (≥ 3.25m)	163.35	0,0206	[0.0185-0.0227]	0,0260	[0.0247-0.0272]	0,2949	0,3785	0,0700	0,0686
Islington Row Middleway	64.4 km/h in Aimsun (40km/h in iRAP data)	40 mph	Yes		3	Wide (≥ 3.25m)	131.13	0,0090	[0.0084-0.0095]	0,0164	[0.0153-0.0176]	0,2949	0,3785	0,0305	0,0433
Lee Bank Middleway	64.4 km/h in Aimsun (40km/h in iRAP data)	40 mph	Yes		3	Wide (≥ 3.25m)	375.4	0,0222	[0.0210-0.0235]	0,0262	[0.0255-0.0269]	0,2949	0,3785	0,0754	0,0692
Lee Bank Middleway	64.4 km/h in Aimsun (40km/h in iRAP data)	40 mph	No		3	Wide (≥ 3.25m)	137.82	0,0220	[0.0206-0.0233]	0,0409	[0.0389-0.0430]	0,5730	0,6663	0,0383	0,0614
Lee Bank Middleway	64.4 km/h in Aimsun (40km/h in iRAP data)	40 mph	No		2	Wide (≥ 3.25m)	248.47	0,06047584	[0.0549-0.0660]	0,0965	[0.0896-0.1033]	0,5730	0,6663	0,1055	0,1448
Coventry Road	64.4 km/h in Aimsun (40km/h in iRAP data)	40 mph	No		2	Wide (≥ 3.25m)	180.57	0,25905192	[0.2448-0.2733]	0,2839	[0.2723-0.2956]	0,5730	0,6663	0,4521	0,4261

Figure Appendix A3: Two-way ,Multilane Undivided Road for WM

Name of Road	Speed limit	Speed limit Cat	Speed camera	Configuration	no. of lanes	Lane width	Section length	Actual proportions_	Actual proportions_Pe	Actual proportions	Actual proportions_	Survey proportions_p	Survey proportions_off_	CF_peak	CF_off_peak
Bristol Rd (A38 bristol Rd approaching Speedwell Rd)	<30km/h	20mph	Yes	Exact section NA (Lnad use residential)	2	Medium (≥ 2.75m to < 3.25m)	224.14m	0,0365	[0.0340-0.0389]	0,0625	[0.0584-0.0665]	0,3778	0,4701	0,0965	0,1329
Bristol Rd (A38 bristol rd approaching Speedwell Rd)	<30km/h	20mph	Yes	Exact section NA (Land use: residential)	2	Medium (≥ 2.75m to < 3.25m)	224.4	0,0365	[0.0340-0.0389]	0,0625	[0.0584-0.0665]	0,3778	0,4701	0,0965	0,1329
Bristol Rd (A38 bristol rd approaching Speedwell Rd)	<30km/h	20mph	Yes	Exact section NA (Land use residential)	2	Medium (≥ 2.75m to < 3.25m)	257.13m	0,2617	[0.2484-0.2750]	0,3075	[0.3055-0.3094]	0,3778	0,4701	0,6927	0,6540
Bristol Rd	<30km/h	20mph	No	Exact section NA (Land use residential)	2	Medium (≥ 2.75m to < 3.25m)	256.84	0,2617	[0.2484-0.2750]	0,3075	[0.3055-0.3094]	0,6655	0,7497	0,3932	0,4102
Bristol Rd	<30km/h	20mph	No	Exact section NA (Land use residential)	2	Medium (≥ 2.75m to < 3.25m)	68.6m	Nan	Nan	Nan	Nan	0,6655	0,7497	—	—
Bristol Rd	<30km/h	20mph	No	Exact section NA (	2	Medium (≥ 2.75m to < 3.25m)	76.7m	0,2285	[0.2130-0.2440]	0,2754	[0.2675-0.2834]	0,6655	0,7497	0,3434	0,3674

Athens Use-Case:

Figure Appendix A4: Red Light Running for Athens

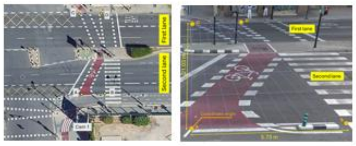
Actual RLR proportions_Peak_95% CI	Survey RLR proportions_peak	Survey RLR proportions_off_peak	Survey RLR proportions_avg	CF
[0, upper]	0.101733	0.230614	0.166	0.000
[0.058, 0.267]	0.056866	0.141425	0.099	1.412
[0.058, 0.267]	0.056866	0.141425	0.099	1.412
[1.46e-05, 0.003]	0.101733	0.230614	0.166	0.003
[1.46e-05, 0.003]	0.056866	0.141425	0.099	0.006
[1.46e-05, 0.003]	0.056866	0.141425	0.099	0.006
[0.013, 0.043]	0.056866	0.141425	0.099	0.259
[0.106,0.121]	0.056866	0.141425	0.099	1.145

Figure Appendix A5: Crossing outside the designated crosswalk for Athens

Spot	Name of Street/Intersection	Geometry	Actual JW proportions_Peak	Actual JW proportions_off_peak_95% CI	Survey JW proportions_peak	Survey JW proportions_off_peak	Survey JW proportions_avg	CF
1	Leof. Vasileos Konstantinou	Narrow (≈4 lanes total)	0.241	[0.220, 0.262]	0.165009	0.257226	0.211	1.461
2	Leof. Vasilissis Amalias	Wide (6–8 lanes)	0.093	[0.0856, 0.101]	0.071809	0.117503	0.095	1.298
3	Leof. Vasilissis Amalias (Syntagma)	Wide (6–8 lanes)	0.093	[0.0856, 0.101]	0.071809	0.117503	0.095	1.298
4	Leof. Vasilissis Sofias	Narrow (≈4 lanes total)	0.218	[0.206, 0.229]	0.165009	0.257226	0.211	1.322
5	Leof. Vasilissis Sofias-Panepistimiou	Wide (major node)	0.218	[0.206, 0.229]	0.071809	0.117503	0.095	3.038
6	Panepistimiou-Korai	Wide (arterial)	0.218	[0.206, 0.229]	0.071809	0.117503	0.095	3.038
7	Panepistimiou-Santaroza	Wide (arterial)	0.015	[0.014, 0.015]	0.071809	0.117503	0.095	0.208
8	Panepistimiou-Plateia Omonoias	Wide (plaza-scale)	0.070	[0.068, 0.071]	0.071809	0.117503	0.095	0.972

Valencia Use-Case:

Figure Appendix A6: Red Light Running of Micromobility Users for Valencia

Configuration	Actual proportions_Peak	Actual proportions_Peak_95% CI	Actual proportions_off_peak	Actual proportions_off_peak_95% CI	Survey proportions_peak	Survey proportions_off_peak	CF_peak	CF_off_peak
	0.1383	(0.1228, 0.1538)	0.1510	(0.1321, 0.1698)	0.0940	0.1379	1.4713	1.0951
	0.1851	(0.1696, 0.2006)	0.2127	(0.1972, 0.2282)	0.1131	0.1634	1.6366	1.3014
	0.1617	(0.1303, 0.1929)	0.2525	(0.2243, 0.2806)	0.1131	0.1634	1.4293	1.5450