A top-down approach for a multi-scale identification of risk areas in infrastructures: particularization in a case study on road safety

Erik Rúa*, Lino Comesaña-Cebal, Pedro Arias and Joaquín Martínez-Sánchez

Abstract

Introduction: Transport infrastructures have an important function in society and the development of a country. In Spain, the most used modes of traveler transport are road and rail, far ahead of other means of transport such as air or maritime transport. Both rail and road infrastructures can be affected by numerous hazards, endangering their performance and the safety of users. This study proposes a methodology with a multiscale top-down approach to identify the areas affected by fire, landslide, and safety in road and rail infrastructures in Galicia (Northwest Spain).

Methodology: The methodology is developed in three steps, coinciding with the three scales considered in this work: network-, system-, and object-level. In the first step, risk areas are identified and prioritized, resulting in the most critical safety risk in a motorway section. This area defines a study scenario composed of a location (A-55 motorway) and the associated risk (road safety). In the second step, the road safety factors within this scenario are selected, hierarchized, and weighted using a combination of Multi-Criteria Decision-Making methods including the Analytical Hierarchy Process and the Best–Worst Method. Finally, a risk map is generated based on the weighting of infrastructure-related safety factors and compared to real historical accident data for validation. The methodology is based on road and risk assessment standards and only information in the public domain is used.

Results: Results show that only 3 segments out of 153 were classified incorrectly, which supports a probability higher than 95% of agreement with real data (at 5% significance level). In a conclusion, the overall methodology exhibits a high potential for hazard prevention and road-safety enhancement.

Keywords: Multiscale assessment, Transport infrastructure, Risk identification, Road safety, Risk map, Multicriteria decision analysis

1 Introduction

Transport infrastructures are important components in modern societies and in developed countries to ensure communication and mobility of people and goods, thus favoring social and economic growth and development.

Focusing only on rail and road transport infrastructure, Europe has 79,142 km of motorways (year 2018) with Spain the country with the greatest length with 15,585 km. In terms of railway, Europe has 231,284 km (year 2019) with Spain the fourth European country with 21,988 km of railways [1].

Both rail and road infrastructures play a very important role, with both road and rail being the most used means of transport by travelers, well ahead of maritime and air transport [2]. For this reason, the development of reliable and resilient infrastructures is essential, being one of the Sustainable Development Goals (SDG) for UN's 2030 Agenda [3]. Within the current societal context, including SDGs in a challenging environmental, economic, and political situation, a new concept for resilience-oriented...
maintenance that efficiently maximizes transport infrastructure performance is mandatory.

The assessment of risks, considered as the combination of the probability of occurrence of an event and the magnitude of its consequences, is crucial to improving infrastructure resilience [4]. Risk assessment is a general process for establishing the context of the infrastructure, the identification of hazards, and risk appreciation [4], which are addressed more in detail in the following paragraphs.

Regarding the contextualization of infrastructures, this article will focus on road and rail infrastructures in Galicia (located in northwest Spain). These infrastructures will be considered from a global-to-specific point of view in a multiscale approach. For this purpose, the terms network, system, object, component, and element are defined as follows:

- **Network** an aggregate of interconnected systems of objects that collectively fulfill a function [5, 6].
- **System** a delimited group of interrelated, independent, or interacting objects [4, 6].
- **Object** individually identifiable part of a system with a specific function in the system [5].
- **Component** individually identifiable part of an object with a specific function in the object [5, 6].
- **Element** the smallest unit of a system for which the internal structure and relationships are no longer considered [5].

Hazards are defined as a potential source of undesirable consequences [4]. In terms of safety, in the last decade (2010–2019), 18,419 people have died in road traffic accidents in Spain, 1503 of them in Galicia [7]. As for railways, the number of fatalities in Spain was 334 for the same period [1]. Several natural hazards are considered in Europe, such as among others: heat waves, heavy precipitation, river floods, windstorms, landslides, droughts, forest fires, and avalanches. However, most of these hazards are influenced by other hazards. Forest fires are influenced by heat waves and droughts, and landslides are influenced by heat waves, heavy precipitation, river floods, windstorms, and droughts [8]. Both fires and landslides are hazards that severely affect safety in road infrastructures and railways [9–12]. Accordingly, we focus on fire, landslides, and safety.

Finally, risk appreciation aims to provide evidence-based data and information to make decisions on how to deal with risk. Risk appreciation is divided into three phases: identification, analysis, and evaluation of the risk. Risk identification includes how risks are discovered, recognized, and recorded. Risk analysis is the foundation for understanding the risk, providing inputs for risk appreciation and decision-making. Lastly, risk evaluation consists of the comparison of risk levels under the defined criteria [13].

The purpose of this work is to determine which areas of road and rail infrastructures are subject to the aforementioned hazards. After these risk areas are identified, their prioritization is carried out to select the riskiest one to be studied in detail. For that purpose, we define and validate a multiscale methodology developed in three steps, accounting for the considered scales with a top-down approach. The core of the work consists of identifying risky areas in road infrastructures based on proven and standardized procedures using only information in the public domain. The main novelty of the proposed methodology consists of the multiscale procedure that ranges from a global risk identification at network-level to a site-specific object-level risk assessment, with a particular focus in this case study on road safety, in which a risk map has been defined by quantifying the safety factors related to the infrastructure. These factors have been weighted through expert opinion, obtaining a map that allows updating the risk zones according to changes in the reality observed through access to public databases, which is relevant in situations of climate change. This methodology allows an individualized analysis of the risk factors to estimate which are more relevant and to be able to act in a specific and preventive way. Obtaining the risk map can be done with real-time information, being the basis of a simulation tool for a digital twin focused on the analysis of road behavior in terms of road safety.

## 2 Related work

In this section, we describe the literature related to the objectives of our work. This review is carried out in a top-down fashion starting with that works focused on multiscale approaches, describing their limitations, and remarking on the differences with our proposal. Then, we progress with the literature related to the risk assessment in our scenario, regarding road safety issues, multi-criteria methods for decision-making and, to conclude, the illustration of the results as risk maps.

The top-down multiscale approach is one of the novelties of the methodology proposed in this paper. Some works related to this multiscale approach are mentioned as follows. Berres et al. [14] presented techniques for the exploration of interconnected traffic dynamics at intersections and highways. These techniques are based on sensors at different scales, which were named microscale, mesoscale, and macroscale. Considering the multiscale approach and also in a road safety context is the work of Thorisson and Lambert [15], in which they integrated road safety metrics of road segments in re-scaleable straight-line diagrams. They identified the road segment...
under stress by searching for one or more metrics that are outliers concerning the contextual data. In the work of Achillopoulou et al. [16] provided the link between the components of multiple hazard resilience assessment in transport infrastructures based on a variety of Structural Health Monitoring, considering three scales: components, assets, and networks. The work carried out by Dragićević et al. [17] develops a multiscale analysis and fuzzy sets combined with GIS-based multicriteria evaluation to determine landslide susceptibility for regional, municipal and local scales at resolutions of 50, 10 and 1 m respectively. Similar to the previous work, Bernardo et al. [18] also treated the problem of landslides. They created a map of susceptibility with which identified the areas of the road most exposed to landslide and they focused on local monitoring of those parts identified.

Many of the previous works mentioned [14, 15, 17, 18] only consider one hazard (such as road safety or landslide). The only one that considers several risks [16] does not have the same approach as this work because it does not treat the hazards in a geolocated way to identify risk areas. However, it is a good example of the different scale that they consider in an infrastructure. In addition of consider several hazards, in the present work, such hazards are studied from a general scale to a site-specific scale, identifying and prioritizing both risk areas (general scale) and factors affecting road safety (site scale). None of the previous works has studied the hazards of this method. Another aspect that differentiates this work from previous ones is the identification and prioritization that has been done in a Geographic Information System (GIS), which allows the creation of risk maps in the last step.

In our specific case study, the risk analyzed in detail is that of road safety. Multi-Criteria Decision-Making (MCDM) models are used to weigh the factors affecting this risk. MCDM are models that analyze many conflicting criteria in decision-making, thus aiding and improving the decision-making process. These models were used in numerous works for the selection and weighting of the most important road safety factors, and the most used method is the Analytical Hierarchy Process (AHP). The most outstanding works that use AHP in road safety aspects are the following.

Seven causes of accidents were selected by Nanda & Singh [19] and created a table indicating the number of accidents associated with each factor. Then, using AHP, they obtained a weighting for each factor. With the weighting and the number of accidents, they generated a ranking of the states with the highest accident rate. In the case of the work carried out by Keymanesh et al. [20], they identified nine factors that contribute to accidents on an entire road. They then divided a road into 8 sections and selected potential black spots in each one. For each section, the most important factors and the most dangerous potential black spots were weighted by five experts using AHP. Both the identification and the prioritization of black spots were compared with the data collected from accidents by the police.

The following articles should also be highlighted, in which a very similar methodology was followed in all of them. Cheng et al. [21] classified road factors by taking into account the factors of driver, vehicle, road, and environment. They also defined their subfactors and obtained the weighting with the AHP. Farooq et al. [22] selected 20 driver behavior factors that have a critical impact on road safety and weights them in a three-level hierarchical structure using AHP. Sordyl [23] performed a hierarchy of the road traffic safety factors in a general way under the levels: driver, vehicle, and environment. It used the AHP to obtain the weighting of these factors.

The AHP not only stands out for being the most used individually but also when combined with other methods, such as an integrated multi-criteria decision-making model combining AHP and Best Worst Method (BWM) to weight driver behavior factors according to Moslem et al. [24].

In all the above-mentioned works, many different road safety factors have been considered and prioritized through different MCDM [19, 21–24]. However, these works neither validate the results with real data nor generate risk maps.

One of the previous works that compare its MCDM with real data is the work carried out by Keymanesh et al. [20]. However, they first identified the hot spots and then prioritized them considering the factors that affect by AHP. The last step was comparing their results with the data from the police, but they do not generate a risk map. In contrast to this work, we select the factors that influence road accidents, weighting them and then we obtain the segments with higher risk.

3 Materials and methodology
3.1 Location
This study will focus on analyzing transport infrastructures located in the southern part of Galicia (Northwest Spain). This area involves two of the four provinces of Galicia, in this case, the provinces of Ourense and Pontevedra. The location of the region and the transport infrastructures are shown in Fig. 1. The transport infrastructure considered are 10 Motorways, 12 State Roads, 2 High-Speed Railways, and 3 Conventional Rail Lines.

3.2 Materials
As stated above, the hazards considered are landslides, fires, and infrastructure safety. These selection criteria
apply equally to both road and rail infrastructure so that the location of risky areas does not depend on the type of infrastructure.

To locate risk areas in step 1 of the methodology, both geographic and non-geographic information has been used. In addition to the fact that the information used is common for rail and road, this information is in the public domain and is explained in the following Table 1.

The information used for the scenario definition in steps 2 and 3 of the methodology is also publicly available and is defined in Table 2.

3.3 Methodology

The general workflow in this work is shown in Fig. 2. The methodology has been divided into three different steps, which coincide with the network, system, and object level, each of the levels belonging to different scales. In this way, the public information used is better adapted to each of the levels according to its scale.

Regarding the methods, the first step (network level) consists of using the RAMS criteria approach (Reliability, Availability, Maintainability, and Safety) [39–41] to identify and prioritize the areas where the considered hazards affect rail and road. The second step (system level) is devoted to studying the area with the highest risk to establish the risk scenario, defined as the inter-relationship between hazards and a certain location in the infrastructure [4]. In this case study, the scenario results to be a road safety risk in a motorway location and, accordingly, road safety factors and their weights are defined using Multi-Criteria Decision-Making methods. The last step (object level) of the methodology consists of creating a road safety risk map with weighted road safety factors that is validated with historical accident data.

The main novelty of the methodology presented lies in the multi-scale approach, adapting to the different scales of the public information used and coinciding with the network, system, and object level.

This methodology is implemented and processed in QGIS, an open-source software framework, and the three steps are described in detail in the following sections.

3.3.1 Step 1: Network level

The publicly available information explained in Table 1 is used to identify the risk areas. In the case of the identification of fire zones, the same weighted factors have been
used as in the work by Novo et al. [10]. For the identification of landslide zones, the landslide susceptibility map has been used. In the case of the safety risk areas, the four inputs have been used as explained in Table 1. Therefore, three different types of hot spots are obtained: fire, landslide, and safety hot spots.

Following, hot spots that coincide in the same location are identified to determine which one presents the greatest risk. For this purpose, they are weighted under the RAMS approach obtaining a prioritization of hot spots. Considering the RAMS, fire and landslide are aggregated as Reliability and Availability indicators in contrast to Safety-related indicators such as accidents, injuries, and fatalities. The weighting of these indicators is based on the work carried out by Li et al. [42], which has been adapted by associating system
preservation with reliability and mobility with availability. Subsequently, the weights are averaged and normalized by summation [43] to obtain the first level hierarchy described in Table 3.

The resulting hierarchy to prioritize hot spots from highest to lowest risk is created and shown in Fig. 3, where No. of infrastructures indicate the number of rail and/or road infrastructures in each hot spot, and accidents, injuries, and fatalities are indicators to quantify safety in each hot spot.

### 3.3.2 Step 2: System level

Once the most critical hot spot has been selected, the scenario must be defined as the location, the hazards involved and the infrastructure to be studied. In this case, such critical hot spot results to be the conjunction of safety-related risks affecting road infrastructure.

It is necessary to examine the factors that affect the road safety study to determine the assets to be analyzed. According to Cheng et al. [21], Sordyl [23], and Alonso et al. [44], such factors can be divided into three main
groups, including human factor, vehicle factor, and traffic factor. Considering that the traffic factor encompasses the infrastructure and its environment, we focus on the traffic factor as the target of the study. These factors have been included, hierarchized, and adapted from Toledo et al. [45], and are shown in Fig. 4.

All the items in the traffic factor must be weighted to determine the most critical through a multi-criteria decision-making model based on Analytic Hierarchy Process (AHP) and the Best–Worst Method (BWM). Following the work of Moslem et al. [24], a series of surveys were carried out and the weighting of each of the factors was obtained.

The AHP allows for the weighting of different factors to create a hierarchy based on the pairwise comparison, allowing the consistency of the process to be checked through the consistency ratio [46]. If the pairwise comparison matrices do not fulfill this requirement, a weighted goal programming model is applied [43, 47].

The BWM is a method for obtaining the weights of the criteria by comparing the best and worst criteria with the rest of the criteria [47].

The combination of these two methods is useful to deal with weighting factors including many criteria, as in this case, “Endowments and conservation”, which includes six criteria. In this case, the BWM is used, while the weighting of the rest of the criteria in the hierarchy is achieved using the AHP method [24]. As a result, the number of comparisons to be made is lowered.

3.3.3 Step 3: Object level
We consider a subset of the weighted factors to create a risk map of the most dangerous points of the motorway. The selection of the optimal subset is performed in three steps.

The first step is sieving of the factors, where the least weighted factors with non-available public information are pruned. The second step is to simplify constant elements within the scale such as meteorological factors measured at a lower scale. The last step is the aggregation of traffic data from industrial parks, shopping centers, and hospitals that are jointly considered under the traffic flow and composition factor.

To model each factor in the subset, a georeferenced layer is created describing the level of risk with a numerical label ranging from 1 to 5, being 1 the lowest and 5 the highest. The layers for the risk mapping are enumerated as follows, whereas their weights and risk map are shown in the results section:
Traffic flow and composition: this layer is elaborated considering that the higher the Average Daily Index, the higher the risk of that segment [45].

In-plant layout: the radii of the motorway curves are calculated and compared to the minimum radii established by the standard [48], finding that some of these curves do not meet these requirements. Therefore, the smaller the radius of the curve, the greater the risk of the segment.

Cross-section: the width of both verges and lanes is calculated, resulting in a fairly constant value complying with the standard [48]. Accordingly, the cross-section was discarded from the risk map.

Speed: the speed limits for this road depend on each section, ranging among 50, 60, 80, 100 and 120 km/h. This layer was derived by associating each section with the speed limit, with the higher speed, the higher risk [49].

Interchange: dimensions for these rails were compared to the standard [48], finding that the minimum dimensions are not met in many cases; therefore, the smaller the dimensions, the greater the risk in that section.

Elevation: the slope of the road is compared with the maximum slope indicated in the standard [48], with the higher the slope, the greater the risk.

4 Results
This section shows the results obtained by applying the proposed methodology. As the methodology has been divided and explained in three steps, the results are also shown in this way.

4.1 Step 1: Network level
The first result obtained after cross-referencing the layers is the hot spots. Figure 5 shows the three types of generated hot spots, appearing either individually (fire, landslide, or safety hot spots) or jointly. In total there were 30 hot spots.

With the location of the 30 hot spots, it is necessary to select the most critical one. According to the hierarchy shown in Fig. 3, the weights for the first consisted of 64% for Reliability and Availability (32.2 and 31.8%, respectively) and a weight of 36% for Safety, as was shown in Table 3.
At the second level of the hierarchy, fires and landslides have been weighted and normalized according to the number of affected infrastructure assets \[11, 50, 51\], being 10 affected by fires and 12 by landslides, which are shown below:

- **Fires** (1) Pavement, (2) Track, (3) Retaining walls, (4) Embankments, (5) Bridges, (6) Vegetation, (7) Lighting columns, (8) Road gantries, (9) Vehicle restraint systems, (10) Tunnel.
- **Landslide** (1) Tunnel, (2) Viaduct, (3) Cut slope and embankment, (4) Drains and culverts, (5) Fencing and protection elements, (6) Vegetation, (7) Track, (8) Ballast, (9) Pavement, (10) Lighting columns, (11) Road gantries, (12) Vehicle restraint systems.

As a result, this gave a weight of 45% to fires and 55% to landslides.

Within the Safety criterion, accidents and injuries plus fatalities are classified with an equal weighting, according to \[52\].

Table 4 shows a summary of all the weights previously indicated.

With the above, all weights obtained are used to generate the equation for the prioritization of hot spots shown in Eq. 1.

Prioritization hot spots
\[
= 0.64 \times (R & A) \times [(0, 45F_1) + (0, 55L_1)] + 0, 365 \times [(0, 5A) + (0, 5x(I & F))]
\]

where R&A is Reliability and Availability; F is fire; I is the number of infrastructures in each hot spot; L is landslides; S is Safety; A is accidents; I&F is injuries and fatalities.

We obtained the most critical hot spots applying Eq. 2 to all the hot spot candidates, being the 6 most critical ones shown in Table 5.

### 4.2 Step 2: System level

For the definition of the scenario, the most critical location was selected. The location, in the municipality named Mos, includes three main road infrastructures: A-55, A-52, and AP-9. Since it contains several accident blackspots, the main infrastructure to be studied is the A-55, being the others assessed in terms of their interlinks (Fig. 6). The A-55 links the towns of Vigo and Tui and continues to the Portuguese border where it connects with the A3. It also joins the A-52 with the AP-9 and is especially important because it connects the city of Vigo with the nearby industrial area, resulting in large traffic of both light and heavy vehicles. It has a length of 31.24 km located entirely in the province of Pontevedra. Since its construction, it has presented numerous problems in terms of road safety, as the section linking Vigo and Porriño was built on the old national road N-120.

For the weighting of safety factors defined in Fig. 4, following the AHP and BWM method previously described, two groups were distinguished in the pairwise comparisons of the safety factors to obtain their weighting: one group of experts and the other regular road users.

In total, 31 people belonging to the expert group and 36 people belonging to the regular road user group carried out the weighting of the matrices, consisting of 7 matrices for the AHP (two 2 × 2 matrices, four 3 × 3 matrices, and one 4 × 4 matrix) and a single 6 × 6 matrix for the BWM method.

The regular user group performed the pairwise comparison of levels 1 and 2 of the hierarchy since the factors at these levels are considered more generic and easily understood by all users. The expert user group performed that of level 3, because of the level of detail and difficulty of these factors. Being non-homogeneous

| Location | Infrastructures involved | Risk |
|----------|--------------------------|------|
| Mos      | 3 Roads                  | Safety |
| A Canda  | 2 Roads                  | Landslide |
| A Cañiza | 2 Roads                  | Landslide |
| Ribadavia| 2 Roads, 2 Railway lines  | Landslide |
| Cesantes | 1 Road                   | Safety |
| Vilavella| 2 Roads, 2 Railway lines  | Fire  |
groups comparing different levels, it was not necessary to combine the weights obtained from the comparisons for each group at the same level, only the aggregation of preferences of the whole hierarchy was performed by the geometric mean. With all this, the weights of the factors obtained from both multi-criteria decision methods are shown in Table 6.

Table 7 shows the overall weights of level 3, where the weights of levels 1 and 2 were taken into account. In other words, the weight of each level 3 factor has been multiplied by the weight of level 2 and the weight of level 1 to which it belongs.

### 4.3 Step 3: Object level

Risk map creation was done considering a subset of the weighted factors in Table 7. As was explained in the corresponding methodology section, the optimal subset was performed in three steps and the factors with the weight of such subset are depicted in Table 8.

The Traffic Flow and composition factor was divided into two sub-factors: traffic flow and composition of heavy vehicles. Since the presence of industrial parks increases the percentage of heavy vehicles, the weight of the Industrial Park factor was the weight of the composition of heavy vehicles, accounting for 19% (0.054/0.278) of the 0.45 and traffic flow accounts for 81% (0.224/0.278) of the 0.45.

Based on the above, the risk maps created in QGIS software for both increasing and decreasing directions are shown in Fig. 7. Table 9 shows the numerical values of the extreme and high-risk sections for the kilometer points in increasing and decreasing directions, and Table 10 shows the summary of all data for both directions. A value according to risk level ranging from 1 to 5 was assigned to each layer. The results obtained also varied in the same range and were divided into 5 levels.
Discussion
This section analyses and discusses the results obtained. First, the risk map obtained is compared with historical accident data [53] and validated by performing an exact Binomial test. In this way, the results obtained in steps 2 and 3, which refer to the system and object level respectively, are discussed. The results obtained in step 1 at the network level and the proposed methodology, in general, are discussed at the end of this section concerning previous works.

The risk map obtained has been validated with real accident data providing the location of the accident with an accuracy of 100 m, between the years 2016 and 2019. It should be noted that these real accident data come from another source and have a different scale than those used in network level 1 and shown in Table 2. This comparison supports the validation of the risk map derived through the multi-criteria decision methodology based on AHP and BWM. The hazard index (Haz) [45] is calculated using Eq. 2.

---

### Table 6
Weights of the factor of traffic safety obtained from AHP and BMW

| Level | Weight | Level 2 | Weight | Level 3 | Weight |
|-------|--------|---------|--------|---------|--------|
| 1. Infrastructure | 0.71 | 1.1. Section characteristics | 0.53 | 1.1.1. Cross-section | 0.30 |
| | | | | 1.1.2. Speed | 0.25 |
| | | | | 1.1.3. Traffic flow and composition | 0.45 |
| 1.2. Layout characteristics | 0.31 | 1.2.1. In-plant layout | 0.57 | 1.2.2. Slope | 0.15 |
| | | | | 1.2.3. Interchange | 0.28 |
| 1.3. Endowments and conservation | 0.09 | 1.3.1. Signalling | 0.11 | 1.3.2. Lighting | 0.14 |
| | | | | 1.3.3. Containment systems | 0.12 |
| | | | | 1.3.4. Beaconing | 0.12 |
| | | | | 1.3.5. Road surface | 0.35 |
| | | | | 1.3.6. Drainage | 0.16 |
| 1.4. Works | 0.07 | 1.4.1. Works | 1 |
| 2. Environment | 0.29 | 2.1. Meteorology | 0.67 | 2.1.1. Fog | 0.37 |
| | | | | 2.1.2. Rain | 0.32 |
| | | | | 2.1.3. Ice | 0.31 |
| 2.2. Other factors | 0.33 | 2.2.1. Industrial Park | 0.54 | 2.2.2. Shopping centres | 0.19 |
| | | | | 2.2.3. Hospitals | 0.27 |

### Table 7
Overall weighted weights of the factors of level 3

| Factor | Weight |
|--------|--------|
| Traffic flow and composition | 0.169 |
| In-plant layout | 0.125 |
| Cross-section | 0.113 |
| Speed | 0.094 |
| Fog | 0.072 |
| Rain | 0.062 |
| Interchange | 0.062 |
| Ice | 0.060 |
| Industrial Park | 0.052 |
| Works | 0.050 |
| Slope | 0.033 |
| Hospitals | 0.026 |
| Road surface | 0.022 |
| Shopping centres | 0.018 |
| Drainage | 0.010 |
| Lighting | 0.009 |
| Containment systems | 0.008 |
| Beaconing | 0.008 |
| Signaling | 0.007 |

### Table 8
Weights of the factors considered in the subset for the risk map

| Factor | Weight | Normalized weight |
|--------|--------|-------------------|
| Traffic flow and composition | 0.169 + 0.052 + 0.026 + 0.018 = 0.265 | 0.46 |
| In-plant layout | 0.125 | 0.22 |
| Speed | 0.094 | 0.16 |
| Interchange | 0.062 | 0.11 |
| Slope | 0.033 | 0.06 |
Though Haz is usually calculated for longer segments of roads, in our case it is derived for 100-m segments to be directly compared to the risk map. With the hazard indices obtained for both the increasing and decreasing directions, 5 equal intervals are defined for visualization, where the higher the index, the greater the hazard. The three highest levels of the hazard index are shown in Figs. 8 and 9 with the risk map previously drawn up for both directions. A zoom-in of the most critical points is also shown.

There were not any sections with values 1 (corresponding to “no risk”) or 5 (corresponding to “extreme risk”).

**Table 9** Extreme risk calculated values for increasing and decreasing direction

| Increasing          |        |       | Decreasing          |        |       |
|---------------------|--------|-------|---------------------|--------|-------|
| Calculated risk     | Kilometre Point |       | Calculated risk     | Kilometre Point |       |
| 4.2645              | 12.0   |       | 4.3654              | 12.0   |       |
| 4.2145              | 9.9    |       | 4.3145              | 12.1   |       |
| 4.2045              | 12.1   |       | 4.3122              | 12.2   |       |
| 4.2045              | 12.2   |       | 4.1454              | 9.2    |       |
| 4.1595              | 9.2    |       | 4.1448              | 9.5    |       |
|                     |        |       | 4.1187              | 9.6    |       |
|                     |        |       | 4.0945              | 9.3    |       |

\[ Haz = \frac{\text{No. accidents with casualties}}{10^8 \text{ vehicles km}} \]  

(2)
As the values obtained were classified into these levels with linear interpolation, there are segments with intermediate colors, and so their risk was between these two values.

As can be seen in Figs. 8 and 9, maps show a good correlation, as points with a higher danger index coincided with high-risk areas. This good correlation can be seen in zoom 3 in the increasing direction (Fig. 8), in one point of zoom 1, and in one point of zoom 4 in the decreasing direction (Fig. 9). In these three examples, the hazard index points had a moderate level and corresponded to medium risk on the map. The same happens in zoom 3 of the decreasing direction (Fig. 9), but in this case, the hazard index point was very high, and the risk was extreme.

Cases with a lower hazard index (or simply no accidents) than the risk index cannot be considered incorrect, as this indicates that these segments present some potential risk factors. Although no accidents have occurred yet, they should be considered because, in the case of future changes of different nature (traffic flow increase, new types of vehicles, climate change, aging of the infrastructure, etc.), these sections may reach high Hazard Indexes. Therefore, predictive maintenance policies and early decision-making should be followed, with special attention to these segments, before accidents

|                  | Mean | Median | Std Dev | Minimum | Maximum |
|------------------|------|--------|---------|---------|---------|
| Increasing risk  | 3.090| 3.070  | 0.419   | 2.325   | 4.264   |
| Decreasing risk  | 3.082| 3.032  | 0.431   | 2.325   | 4.365   |

Fig. 8 Validation of the risk map for increasing direction
occur. An example of this was two points in zoom 1 and one point in zoom 2 in the increasing direction (Fig. 8), and all points in zoom 2 in the decreasing direction (Fig. 9). However, in the case of a high hazard index and a low-risk index, this is considered an incorrect result, as all points with a high hazard index should be perfectly identified within the risk map. This occurred only for 3 points: one point in zoom 1 in the increasing direction (Fig. 8), one point of zoom 1, and one point of zoom 4 in the decreasing direction (Fig. 9). In the example of the increasing direction, the calculated risk was high, and the hazard index was very high. In both examples of decreasing direction, the calculated risk was moderate and hazard indexes were high.

Figure 10 shows the risk map for the increasing and decreasing direction with all the points of the hazard index with which the validation is done.

Criteria followed for the validation were the following to be on the side of security:

- Segment CORRECT identified: Calculated risk equal to or higher than the hazard index
- Segment INCORRECT identified: Calculated risk lower than hazard index

The road was divided into segments of 100 m, which resulted in 170 segments for the decreasing and the same for the increasing direction. Therefore, the risk was calculated for 340 segments. However, real data was only provided for 153 of these segments, because there were not any accidents in the rest of the segments, and the hazard index was very low. With this, the analysis of the results was derived considering only the segments with real data.

To know whether the segments were correctly or incorrectly classified, the assessment takes into account that these results follow a binomial distribution where the outcomes are Boolean-valued: success or failure. According to the results, we tested if this methodology had a probability of correctly classifying the segments with a
probability greater than 95%, with a hypothesis contrast where the statements are (Eq. 3):

\[ H_0 : p \leq 0.95 \]
\[ H_1 : p > 0.95 \]

The hypothesis established in Eq. 3 was calculated with an exact Binomial test in the software R with a 5% significance level. The results show that the p-value was less than 0.05 (5% significance level), resulting in the rejection of the null hypothesis, and the acceptance of the alternative hypothesis. This concluded that the methodology correctness was greater than 95%.

The advantages of the proposed methodology over alternative methods reviewed in the related work are explained in detail below.

If we focus on discussing our results compared to results obtained in previous works, we can say the top-down multi-scale methodology shows results at the three scales: location and prioritization of the hot spots at the network level, weighing of the safety factor at the system level and risk map creation at the object level. None of the previous works [14–18] related to this multiscale approach, showed results for the different scales and none considered together the hazards that are taken into account in the present work.

Table 11 Results for the exact Binomial test to assess the probability of classifying with 95% success (5% significance level)

| Exact Binomial test |       |
|---------------------|-------|
| Total segments      | 153   |
| Failures            | 3     |
| Success             | 150   |
| p value             | 0.04944 |

![Fig. 10 All hazard index points for increasing and decreasing direction](image-url)
Regarding the MCDM and validation of the result that is developed in parts 2 and 3 of our methodology, the works revised [19, 21–24, 54] used some MCDM but they do not validate their results with real data. In the work carried out by Keymanesh et al. [20] despite using the AHP method and comparing it with real data, they do not generate a risk map. The work carried out by Fuller et al. [55] does generate a risk map and compares it with real data, but they only take into account four road safety factors and use a multi-criteria evaluation method that does not take into account the weighting of either experts or road users. The work of Driss et al. [56] validates the generated map, but they did not obtain it using MCDM. In the work carried out by Hu et al. [57], in addition to not using any MCDM, they only compared the generated map with field-based interviews and not with real accident data.

6 Conclusions

In this work, a multiscale methodology with 3 steps has been proposed and carried out. The first step consists of identifying the hot spots at the network level, the second step consists of identifying and prioritizing the factors that affect the road safety of the system, and the last step is the creation and validation of a risk map.

For the identification of risky areas, landslide, fire, and safety hazards have been considered. Once obtained the hot spots, they were prioritized to obtain the area with the highest risk, where the second step was applied. The risk factor in this area was safety in road infrastructure, so the factors affecting road safety were analyzed. Only factors related to traffic were considered, leaving other factors which could affect road safety, such as human or vehicle factors, unanalyzed. These factors were classified into different levels and weighted using a Multi-Criteria Decision Method. In this case, the Analytical Hierarchy Process (AHP) and the Best Worst Method (BWM) were applied, and the opinions of 67 participants were obtained through surveys.

Once the results of the weighting of factors were obtained, in the third step two risk maps were created with these factors: one for the increasing and one for the decreasing direction. The resulting risk maps were compared with the hazard index of real accidents on the road to validate the risk map and the methodology applied. This validation has concluded that the methodology has a probability of success in classification greater than 95%. In this way, we obtain a risk map based on the safety factors related to the infrastructure. Thanks to the weighting of factors obtained with the methodology, we have quantified each 100-m segment according to the risk presented by these factors. Validation with real data shows that these infrastructure-related safety factors do affect the occurrence of accidents. However, if we only obtained the risk map with these real data, we would not have quantified the influencing factors. Therefore, preventive action can be taken on those factors that have the greatest influence on the risk.

In general conclusion, this methodology shows a good overall result as the area identified as the riskiest has a probability of successful classification higher than 95%. All this methodology is developed with a multiscale top-down approach and in as standardized a way as possible since numerous road and risk assessment standards are followed. The fact that the information is in the public domain means that the methodology developed here is applicable to any environment (at least in Spain), and its application is easily automated so that it is quick to apply as the public data used hardly needs any processing. The advantage over previous works is that this methodology presents results for the three scales that have been considered in this multi-scale approach and it is suitable for analyzing a transport infrastructure, from a network completeness point of view to the quantification of safety factors in 100-m sections to obtain a risk map. This methodology allows an individualized analysis of the risk factors to estimate which are most relevant, and to act in a specific and preventive manner. Obtaining the risk map from these factors could be done in real-time. For this reason, the proposed methodology can be the basis for a simulation tool of a Digital Twin focused on the analysis of road behavior in terms of road safety. A weakness of this methodology is given by its application in the case study of road safety on the A-55 motorway. In this case, the hierarchy of factors can be applied to other segments of the motorway, but not the weightings obtained since these weights are particular to this road due to its special characteristics.

Acknowledgments

This work has been supported by Gobierno de España: Ministerio de Ciencia, Innovación y Universidades through the Grant PID2019-108816RB-I00 funded by MCIN/AEI/10.13039/501100011033 and the Grant for Predoctoral Contracts with reference PRE2020-096222. This work has also received funding from the European Union’s Horizon 2020 research and innovation program under Grant Agreement No 958171.

Author contributions

Conceptualization, E.R.; formal analysis, E.R.; methodology, E.R.; software, L.C.-C.; validation, E.R.; and L.C.-C.; writing—original draft preparation, E.R., and J.M.-S.; writing—review and editing, E.R. and J.M.-S.; visualization E.R.; supervision P.A. and J.M.-S.; project administration, P.A. and J.M.-S.; funding acquisition, P.A. All authors have read and agreed to the published version of the manuscript.

Funding

This research received no external funding.

Availability of data and materials

Not applicable.
42. Li, Z., & Sinha, K. C. (2004). Methodology for multicriteria decision making in highway asset management. Transportation Research Record, 1885(1), 79–87. https://doi.org/10.3141/1885-12

43. Aznar Bellver, J. and Guijarro Martínez, F. (2012). Nuevos métodos de valoración: modelos multicriterio. Universidad Politécnica de Valencia.

44. Alonso, F., Esteban, C., Casquero, E., Velarte, J. L., Suárez, J. and Pastor, J. C. (2019). Enseñanzas para la gestión de la accidentalidad en emplazamientos peligrosos en las carreteras. Paltaforma Tecnológica la Carretera. 4. p 20–22.

45. Toledo-Castillo, F., Peiró-Torres, M., Hidalgo-Fuentes, S., Sospedra-Baeza, M. and Toledo-Castillo, S.-B. M., Peiró-Torres, M. P., Hidalgo-Fuentes, S. (2014). Variables relevantes en la seguridad de la infraestructura. Paltaforma Tecnológica la Carretera. 6.

46. Saaty, T. and Vargas, L. (2012). Models, methods, concepts & applications of the analytic hierarchy process. Springer.

47. Rezaei, J. (2015). Best-worst multi-criteria decision-making method. Omega (United Kingdom), 53, 49–57. https://doi.org/10.1016/j.omega.2014.11.009

48. Dirección General de Carreteras. (2016). Trazado. Instrucción de Carreteras. Norma 3.1-IC, Boletín Oficial del Estado. Gobierno de España.

49. Van den Bergh, W. European Road Safety Observatory: Road Safety Thematic Report - Speed. European Commission. https://road-safety.transport.ec.europa.eu/system/files/2021-07/road_safety_thematic_report_speeding.pdf.

50. UVIGO. (2020). Emergency Management Plan. SAFEWAY Deliverable D8.1.

51. Ntzeremes, P., & Kriytopoulos, K. (2019). Evaluating the role of risk assessment for road tunnel fire safety: A comparative review within the EU. Journal of Traffic and Transportation Engineering (English Ed.). https://doi.org/10.1016/J.JTTE.2018.10.008

52. TECNALIA. (2019). Validation Report No 1. RAGTIME Deliverable D6.2. [Online]. Available: http://ragtime-asset.eu/.

53. Gobierno de España: Ministerio del Interior. Dirección General de Tráfico (DGT): Ficheros microdatos de accidentalidad con víctimas. https://www.dgt.es/es/seguridad-vial/estadisticas-e-indicadores/ficheros-microdatos-accidentalidad/ (accessed Jun. 21, 2021).

54. Kanuganti, S., Agarwala, R., Dutta, B., Bhanegaonkar, P. N., Singh, A. P., & Sarkar, A. K. (2017). Road safety analysis using multi criteria approach: A case study in India. Transportation Research Procedia, 25, 4649–4661. https://doi.org/10.1016/j.trpro.2017.05.299

55. Fuller, D. O., Williamson, R., Jeffe, M., & James, D. (2003). Multi-criteria evaluation of safety and risks along transportation corridors on the Hopi Reservation. Applied Geography, 23(2–3), 177–188. https://doi.org/10.1016/j.apgeog.2003.08.010

56. Driss, M., Benabdellah, K., Saint-Gerand, T., & Hamadouche, M. A. (2014). Traffic safety prediction model for identifying spatial degrees of exposure to the risk of road accidents based on fuzzy logic approach. Geocarto International, 30(3), 243–257. https://doi.org/10.1080/10106049.2014.883554

57. Hu, J., Zhong, G., Cheng, Z., and Wang, D. (2012). GIS-based road safety evaluation model for cyclist in campus of Higher Education Mega Center. IEEE Conf. Intell. Transp. Syst. Proceedings, ITSC, pp. 1127–1131. https://doi.org/10.1109/ITSC.2012.6338868.

Publisher’s Note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.