applied sciences

MDPI

Article

Route Risk Index for Autonomous Trucks

Ryan Jones [1], Raj Bridgelall [1,*] and Denver Tolliver [2]

[1] Transportation, Logistics, & Finance, College of Business, North Dakota State University, P.O. Box 6050, Fargo, ND 58108-6050, USA; ryan.jones.2@ndsu.edu
[2] Upper Great Plains Transportation Institute, North Dakota State University, P.O. Box 6050, Fargo, ND 58108-6050, USA; denver.tolliver@ndsu.edu
* Correspondence: raj@bridgelall.com

Abstract: The proliferation of autonomous trucking demands a sophisticated understanding of the risks associated with the diverse U.S. interstate system. Traditional risk assessment models, while beneficial, do not adequately address the state and regional variations in factors that significantly impact the safety and efficiency of autonomous freight transport. This study addresses the problem by developing a composite risk index that evaluates the safety of U.S. interstate routes for autonomous trucking, considering both state and regional differences in traffic volumes, road conditions, safety records, and weather patterns. The potential for autonomous trucking to transform the freight industry necessitates a risk assessment model that is as dynamic and multifaceted as the system it aims to navigate. This work contributes a regionally sensitive risk index using GIS methodologies, integrating data from national databases, and applying statistical analysis to normalize risk factors. The findings reveal significant state and regional disparities in risk factors, such as the predominance of precipitation-related risks in the Southeast and traffic in the Far West. This work provides a targeted approach to risk assessment for policymakers and infrastructure planners and offers a strategic tool for logistics companies in optimizing autonomous trucking routes. The long-term benefit is a scalable model that can adapt to evolving data inputs and contribute to the broader application of risk assessment strategies in various domains.

Keywords: continuous utilization; geographic information systems (GISs); labor cost reduction; national regulations; risk assessment methodologies; supply chain agility; transportation efficiency

check for updates

Citation: Jones, R.; Bridgelall, R.; Tolliver, D. Route Risk Index for Autonomous Trucks. *Appl. Sci.* **2024**, *14*, 2892. https://doi.org/10.3390/app14072892

Academic Editors: Reynold C.K. Cheng, Xiaolin Han, Chenhao Ma and Matthias Renz

Received: 5 March 2024
Revised: 26 March 2024
Accepted: 27 March 2024
Published: 29 March 2024

1. Introduction

Autonomous trucking has the potential to revolutionize freight transport, promising enhanced safety, reduced emissions, and improved efficiency [1]. However, the safe integration of autonomous trucking into the national transportation infrastructure presents a complex challenge. While the literature widely acknowledges the potential safety and efficiency benefits of autonomous vehicles (AVs), the focus on autonomous trucks (ATs) warrants specific attention. ATs will have a disproportionate impact on the freight industry compared with passenger cars due to their larger size, weight, and operational complexity, which introduce distinct challenges in traffic management, infrastructure wear, and accident severity. Additionally, the economic implications of AT deployment, given their role in the supply chain, justify a dedicated analysis of their unique risk factors.

The safety and efficiency of autonomous trucking are contingent upon myriad factors, including traffic volumes, road quality, accident rates, and weather conditions that span a diverse geography [2]. The heterogeneity of risk factors across the U.S. interstate system presents a considerable obstacle. For example, while the I-95 corridor in the Southeast might be prone to high traffic volumes and severe weather patterns, the I-5 in the Pacific Northwest may contend with road roughness due to heavy vehicle traffic and varied topography and climate conditions. Such variability demands a complex yet refined approach to risk assessment that current methodologies do not sufficiently address. Current models for risk assessment in transportation are either overly broad, neglecting regional variations, or are too fragmented, lacking a holistic view of interstate networks. The literature lacks a granular yet comprehensive risk analysis to facilitate strategic planning and policymaking for AT deployments across different regions of the United States.

The goal of this paper is to develop and validate a composite risk index that quantitatively evaluates the safety of U.S. interstate routes for the emerging autonomous trucking industry. The aim is to determine the relative contributions of traffic volumes, road quality, accident rates, and weather conditions to the overall risk profile of interstate segments, and how these contributions vary by state and region across the United States. To achieve this goal, this work employs geographic information systems (GISs) for spatial analysis of traffic and environmental data, integrates data from national databases such as the Highway Performance Monitoring System (HPMS) and the Fatality Analysis Reporting System (FARS), and utilizes statistical methods to normalize and analyze risk factors.

The contributions of this work are multifold. First, it provides a robust, data-driven methodology for risk assessment specific to autonomous trucking. Second, this paper identifies regional patterns in risk factors, enabling targeted risk mitigation strategies. Finally, it supports the assertion that a uniform national strategy for the deployment of autonomous trucking is less effective than a region-specific approach. This research thus offers a significant value proposition for stakeholders in transportation, logistics, and policy development, aiming to navigate the complexities of autonomous freight transportation safely and efficiently.

The organization of the rest of this paper is as follows: Section 2 reviews the literature on autonomous trucking but with a focus on deployment strategies and risk evaluation. Section 3 describes the workflow developed to source and process data for the effective visualization of risk patterns. Section 4 discusses the analytical results and implications for stakeholders seeking to benefit from AT deployments. Section 5 concludes the research and suggests future work.

2. Literature Review

This section reviews existing literature, focusing on the opportunities and challenges of ATs, deployment risk assessments, and identifying gaps in the literature.

2.1. Opportunities and Challenges

ATs offer significant benefits, including safety improvements, operational efficiency, reduced labor and fuel costs, and enhanced driving ease [3]. Significant economic advantages come from replacing the driver [4], achieving almost continuous vehicle utilization [5], avoiding traffic by prioritizing nighttime operations, and reducing road damage by operating at night when lower temperatures reduce the impact of load stress [6]. AT advancement could profoundly affect the supply chain and society at large, potentially boosting productivity, GDP, and employment [7]. ATs can also address rising e-commerce demands sustainably through increased fuel efficiency, utilization, and reduced accidents [8].

ATs promise to mitigate supply chain bottlenecks and the bullwhip effect, enhancing supply chain agility and efficiency [9]. This strategic advantage allows for operational redesign by leveraging the reliability and flexibility of ATs. Moreover, there is consensus in the literature that autonomous vehicles in general will significantly contribute to accident reduction by eliminating causes due to human error [10]. However, the transition to ATs faces barriers, including societal acceptance, privacy, security, and legal challenges [11]. Addressing these concerns requires developing standards for data use, enhancing cybersecurity, and clarifying legal liabilities. Moreover, the absence of unified national regulations presents another hurdle, necessitating a balanced approach by states to encourage testing while ensuring public safety [12].

2.2. Deployment Risk Assessments

The literature on autonomous trucking deployment risk assessment spans various methodologies aimed at safe and efficient integration into transportation networks. Road roughness poses a risk to AV operation because they must accurately assess roughness levels to adjust actuations such as speed and acceleration to maintain operational stability [13]. Climate change has also emerged as a catalyst for road deterioration, particularly rutting damage caused by extreme weather changes [14]. Precipitation poses challenges because AVs must automatically assess the levels of skidding and hydroplaning on wet surfaces to adjust for safe operation [6]. Rain that leads to flooding can also challenge the performance of autonomous navigation systems. Advanced sensor fusion and more advanced machine learning methods become necessary to enhance operational reliability in adverse weather conditions [15].

Studies have also shown that truck operation is associated with higher crash risks relative to cars [16]. High traffic variability also challenges the safe operation of AVs because their sensors and actuators must react quickly [17]. For example, in congested situations, erratic human driving behavior, like sudden lane changes or aggressive maneuvers, can be unpredictable for AVs to handle, potentially causing crashes. Studies have also shown that crash risk is a dynamic quantity that has considerable spatial and temporal variations [18]. These studies highlight the importance of tailored risk assessment frameworks to address the unique challenges of ATs. Frameworks leveraging driving simulators [19], trajectory planning [20], liability considerations [21], and cybersecurity [22] illustrate the diverse application and critical nature of comprehensive risk assessments. Collectively, this body of work highlights the need for interdisciplinary approaches to ensure the safe integration of ATs into the transportation ecosystem.

2.3. Literature Gaps

Despite consensus on the transformative potential of ATs, their deployment presents unique challenges in risk assessment, particularly due to the dynamic nature of interstate transportation. Current research often focuses on risk factors in maintenance [23] and trajectory planning [24], steering away from the complex interplay of risk factors in route selection. There is a noted gap in utilizing GISs for comprehensive autonomous trucking route risk assessments and in integrating national databases into risk models [12]. The variability in regional risk factors and the lack of comprehensive risk indices tailored to autonomous trucking expose the need for sophisticated risk assessment frameworks. This study aims to fill these gaps by proposing a composite risk index that leverages GISs and national transportation databases, offering a novel contribution to the transportation risk assessment field.

3. Methodology

This workflow integrates GISs for spatial analysis, data management techniques for processing and normalizing data, and statistical methods for risk assessment, showcasing an interdisciplinary approach to develop a comprehensive risk index to inform AT deployment planning.

3.1. Data

The workflow utilized data from three large publicly available datasets from various United States government agencies as follows:

HPMS: The Highway Performance Monitoring System, updated in 2020 and maintained by the U.S. Bureau of Transportation Statistics [25]. The data contain layers of individual GIS shapefiles, one per U.S. state. Each shapefile layer contains many rows of linear objects, each representing a road segment. Attributes for each linear segment include state and county information, route number, geometric characteristics, the international roughness index (IRI), and the average annual daily traffic (AADT), a measure of the average number of vehicles that travel on the segment per day.

FARS: The Fatality Analysis Reporting System, updated in 2021 and maintained by the U.S. Bureau of Transportation Statistics [26]. The shapefile contains a single layer with geospatially encoded points representing a single fatal accident. Attributes include characteristics of the crash and environmental conditions at the time.

NCEI: The National Centers for Environmental Information (NCEI) dataset contains the mean annual precipitation from 1901 to 2000 for each U.S. county tracked [27].

3.2. Analytical Workflow

Figure 1 illustrates the data synthesis workflow consisting of individual procedures that are either coded in software or implemented using a GIS tool.

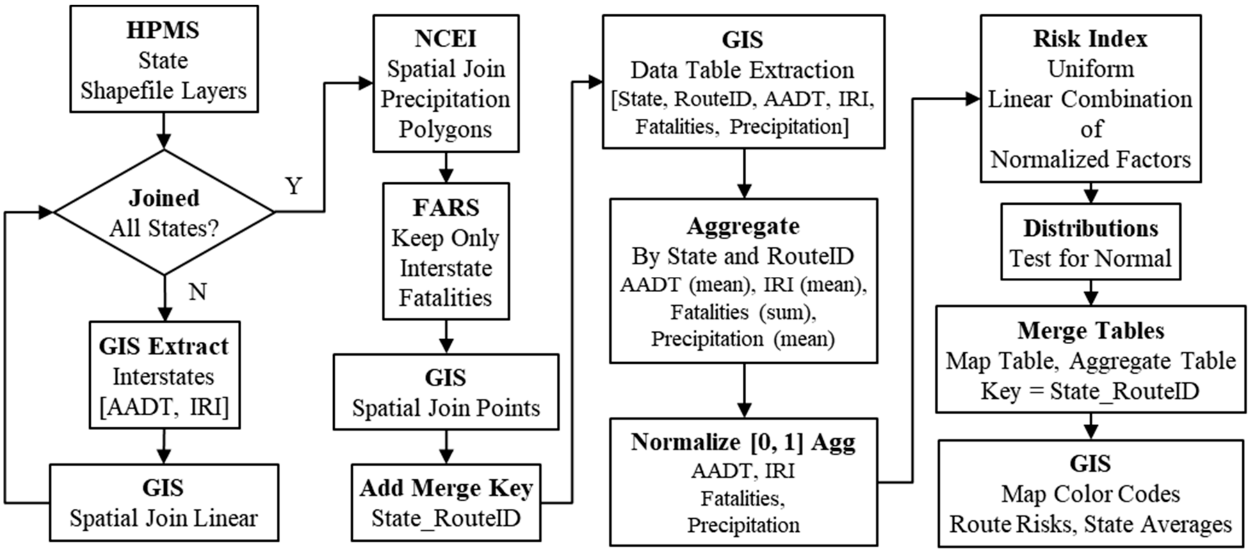

Figure 1. The data synthesis workflow.

The "GIS Extract" procedure of the workflow kept only those linear objects representing interstate routes, along with their AADT and IRI features, by setting a filter for F_SYSTEM = 1. For example, Figure 2 shows the result of extracting the linear GIS objects representing the two interstate routes (red lines) in North Dakota. The blue lines represent other roadway types. The workflow then iteratively joined the extracted interstate shapefiles until it processed the shapefile layers for all states of the contiguous United States (CONUS).

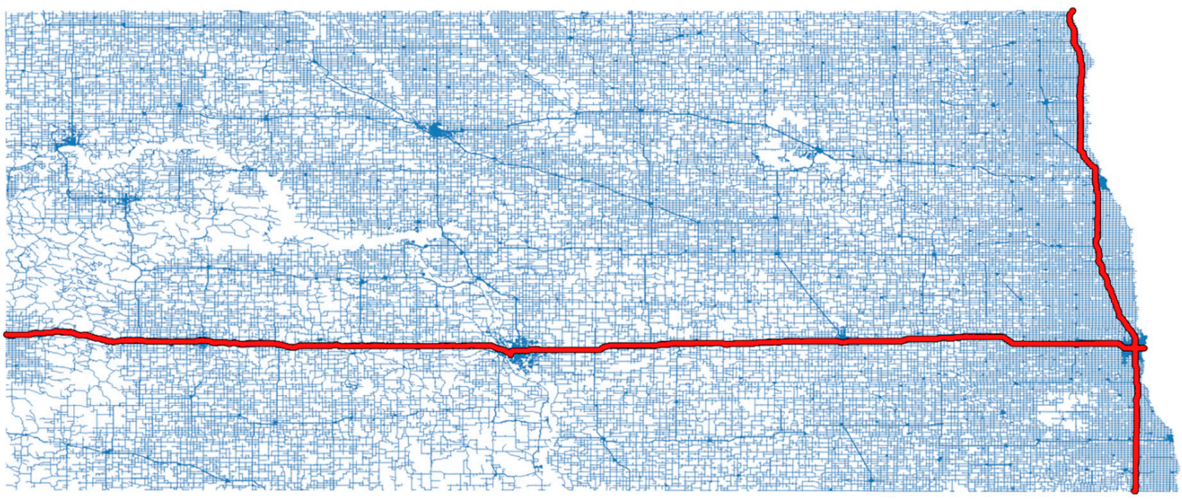

Figure 2. Extraction of the linear GIS objects representing the two interstates in North Dakota.

The spatial joining of the precipitation polygons of U.S. counties averaged the mean annual precipitation across intersecting road segments. Figure 3 depicts the mean annual precipitation for CONUS counties from 1901 to 2000. This illustration helps to interpret the results of where risk factors dominate.

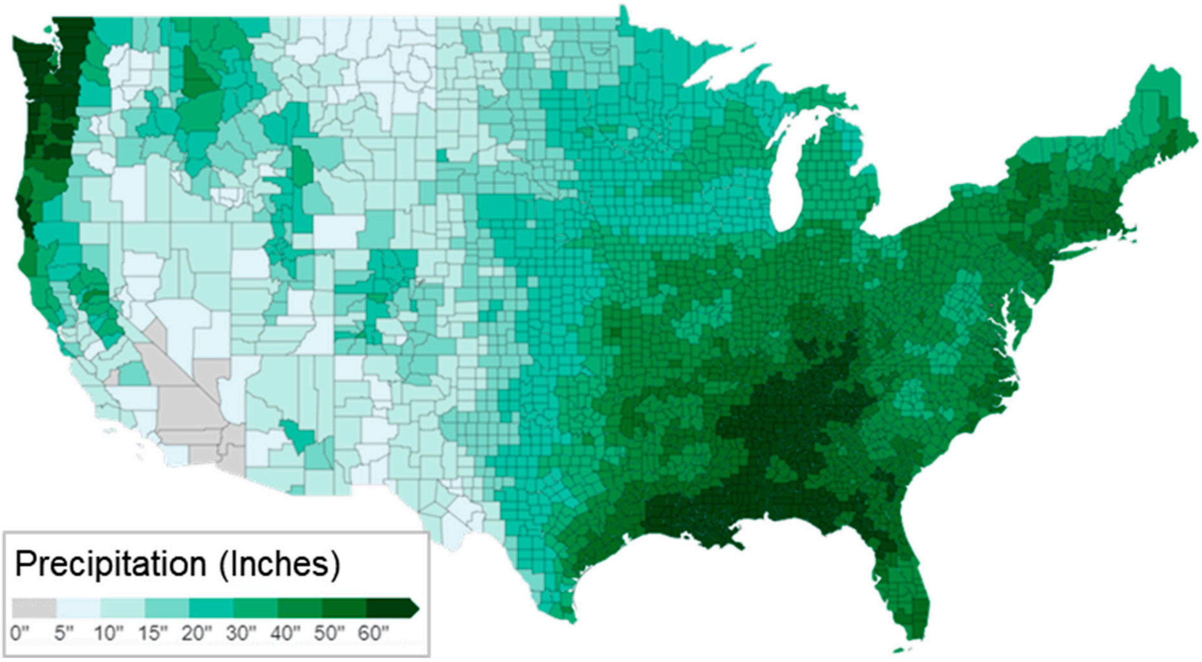

Figure 3. Mean annual precipitation for CONUS counties from 1901 to 2000.

In processing the FARS data, the workflow kept only the data for crashes on the interstate routes. The workflow then spatially joined those to the previously joined shapefiles containing the AADT, IRI, and precipitation values for each interstate segment.

Due to the fragmented regulatory constraints among states, autonomous trucking companies often make deployment decisions at the state level [28]. Therefore, the workflow aggregated the risk index for an entire interstate section that falls within individual states. In preparation for merging aggregated values at the state and road levels, the GIS procedure created a merge key containing a string of the state abbreviation and the interstate route number. The data aggregation that followed computed the mean values for AADT, IRI, and precipitation and the sum of fatalities along the entire interstate within a state. Performing the aggregation on an extracted data table was more convenient due to the enhanced programming functions available outside of the GIS environment. Specifically, the authors used Python along with programming libraries such as pandas, seaborn, and scipy to plot descriptive statistics of each factor such as bar charts and box plots to visualize the risk factors across states and regions.

The next strategy was to normalize the four factors within the [0, 1] range so that they became more comparable in a unitless dimension. The normalization is as follows:

$$N(x) = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \tag{1}$$

This study defines the composite risk index (*CRI*) as a uniformly weighted linear combination of the four normalized factors as follows:

$$CRI = \frac{w_1 \cdot N(TV) + w_2 \cdot N(RC) + w_3 \cdot N(SR) + w_4 \cdot N(WP)}{\sum_{i=1}^{4} w_i} \tag{2}$$

where the following are true:

- *TV* represents the traffic volume.

- RC denotes the road condition in terms of roughness.
- SR stands for safety records such as fatalities.
- WP indicates weather patterns, with mean annual precipitation in this context.
- $N()$ is the normalizing function in Equation (1).
- w_1, w_2, w_3, and w_4 are weights assigned to each risk factor, reflecting their relative importance. The summation in the denominator ensures that the CRI remains normalized and comparable.

This baseline research sets all the weights to unity because there is currently no other information about their relative importance in the overall risk, which future work will explore. This definition ensures a uniform approach to risk assessment across all states and routes. The results section further justifies this definition by showing that, statistically, the factors lack correlation.

The selected measures of correlation between x and y variables were the Spearman's rank and the Kendall's tau correlation coefficients because these methods do not assume that both variables are normally distributed [29]. The Spearman's rank correlation coefficient ρ is a nonparametric measure of the monotonicity of the relationship between two variables. Spearman's correlation converts the variables into rank data and then assesses how well a monotonic function can describe the relationship between these ranks. The Spearman's rank correlation coefficient S is

$$S = 1 - \frac{6\sum d_i^2}{n(n^2 - 1)} \tag{3}$$

where the following are true:

- d_i is the difference between the ranks of corresponding variables x_i and y_i.
- n is the number of observations.

A Spearman correlation of +1 or -1 occurs when there is a perfect monotonic relationship between the two variables, while a correlation of 0 indicates no monotonic relationship.

Kendall's correlation coefficient K assesses the strength and direction of the association between the two variables. It evaluates their ordinal association by considering the number of concordant and discordant pairs of data points. The formula for Kendall's tau is

$$K = \frac{C - D}{\sqrt{(n_0 - n_1)(n_0 - n_2)}} \tag{4}$$

where the following are true:

- C and D are the numbers of concordant and discordant pairs, respectively.
- $n_0 - n(n-1)/2$.
- $n_1 = \sum t_i(t_i - 1)/2$ for each group of tied ranks in the first quantity.
- $n_2 = \sum u_j(u_j - 1)/2$ for each group of tied ranks in the second quantity.
- n is the number of observations.
- Concordant pairs are pairs of observations where the ranks for both elements agree (i.e., both ranks are higher or both are lower in each pair), while discordant pairs are those where the ranks disagree (one rank is higher in the first element of the pair, and the other rank is higher in the second element).

The last three procedures of the workflow evaluated the composite risk index in terms of the distribution of its components and composite index, including testing for a normal distribution and interpreting its spatial distribution on a map. This entailed merging the aggregated factors by state and route back into the GIS platform. The analysis also provided a regional view of the risk index components to assess the spatial distribution of its dominant components.

4. Results

The following subsections discuss the distribution of risk factors, the relationship among risk factors, the distribution of risk by route, the distribution of risk by state, and the distribution of risk by region.

4.1. Risk Factor Distribution

Table 1 summarizes the overall statistics of the risk factors for both the original and normalized values. The mean AADT was approximately 52,727 vehicles, with a standard deviation (STD) of 39,141, indicating a wide range of traffic levels across the road segments. The average IRI value was around 92.57, with a standard deviation of 31.74, suggesting variability in road surface conditions. The mean precipitation level was 38.43, with a 14.24 standard deviation, reflecting differing weather conditions across locations. On average, there were approximately 11 fatalities per road segment per state, but with a high standard deviation of 19, indicating significant variability in safety outcomes.

Table 1. Risk factor statistics.

	Original		Normalized	
Risk Factor	**Mean**	**STD**	**Mean**	**STD**
AADT	52,727	39,141	0.212	0.167
IRI	92.57	31.74	0.266	0.159
Fatalities	11	19	0.058	0.106
Precipitation	38.43	14.24	0.486	0.199

Figure 4 shows box plots of the distribution of the four risk components by region. For reference, Figure 5 shows the U.S. regional definitions according to the U.S. Bureau of Economic Analysis, which mirror current AT testing. The box plots provide a visualization of the central tendency, dispersion, and skewness of the data distribution, highlighting outliers as diamond shaped points. The edges of the colored boxes represent the interquartile range (IQR) from Q1 to Q3, representing the middle 50% of the data. The line within the box represents the median value. The whiskers above and below the boxes represent the maximum and minimum values, respectively, excluding outliers. The individual dots above the whiskers represent outlier values, defined as 1.5 times the IQR.

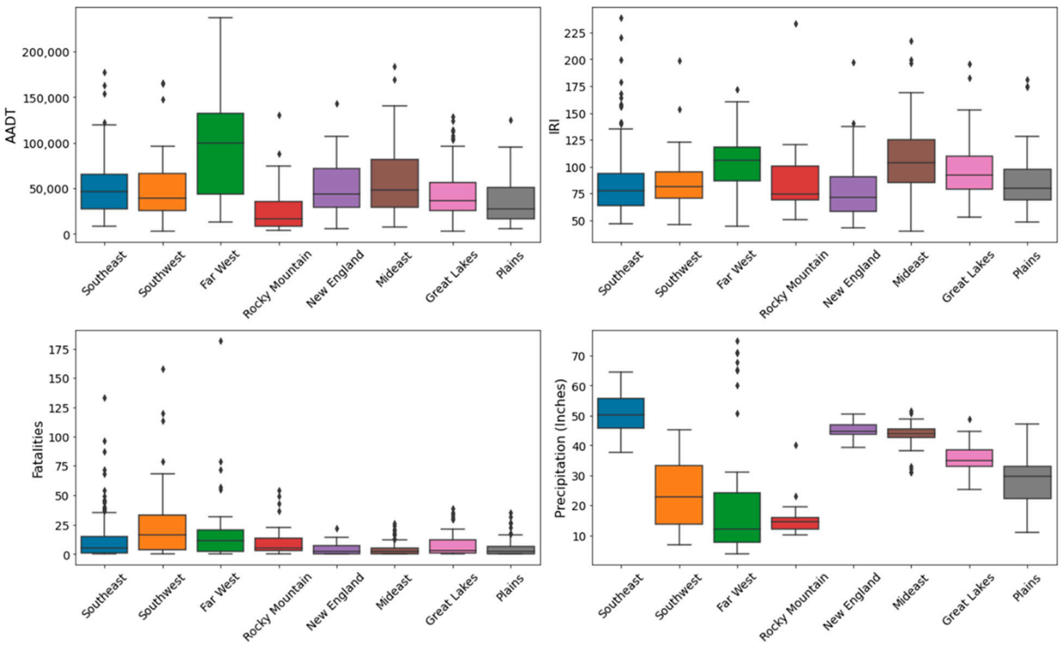

Figure 4. The regional distribution of the risk components.

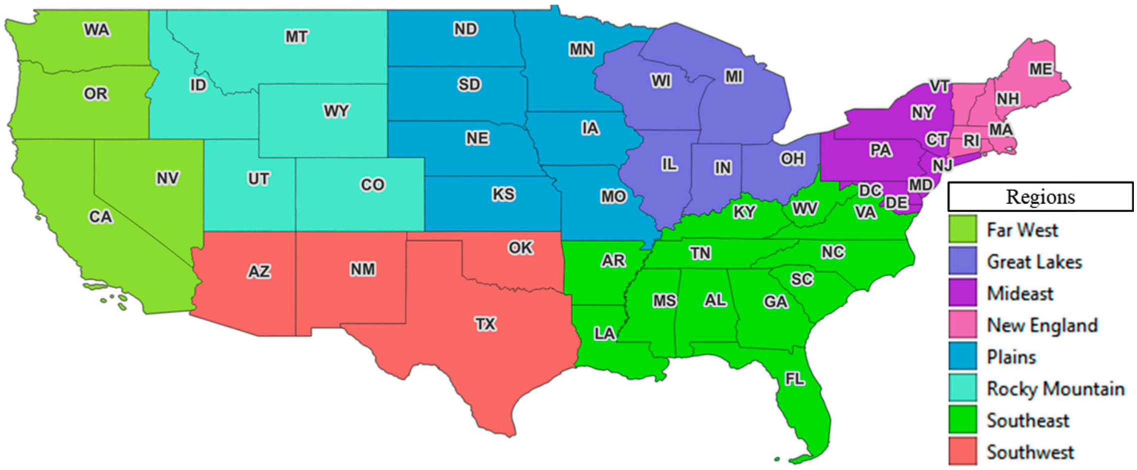

Figure 5. The definition of U.S. regions.

AADT: The Far West region exhibits the widest IQR, suggesting a high variability in AADT. The Rocky Mountain region shows a comparatively low median and small IQR, which could indicate more uniform traffic patterns with fewer extreme values.

IRI: The median IRI values across the regions do not vary significantly; however, the Southeast, Far West, and Mideast regions have a larger spread of IRI scores, indicating a wide range of road quality. Outliers are present in several regions, most notably in the Southeast, which could be indicative of particularly rough patches of interstate.

Fatalities: The Southwest stands out with the highest median value for fatalities and several outliers, which might be a point of concern for transportation safety in that region. The New England, Mideast, Great Lakes, and Plains regions have notably lower median values and fewer outliers, suggesting fewer fatalities on their interstates.

Precipitation: The Southeast has the highest median precipitation, which is consistent with its climatic patterns, as shown in Figure 3. The Far West has a significant number of outliers, showing that some areas receive much higher rainfall. The Rocky Mountain region shows lower precipitation levels, with a particularly tight IQR, indicating consistent weather patterns in terms of precipitation. These box plots provide valuable insights for AT deployment planning. For example, high variability in AADT could require dynamic planning for freight movement. High IRI values may necessitate more frequent maintenance schedules, and high fatality rates could lead to increased insurance costs and the need for enhanced safety measures. Finally, precipitation data could be crucial when planning routes to avoid weather-related delays. These results provide a quantitative foundation for policy recommendations and strategic planning.

Figure 6 presents a pair plot with histograms and scatter plots, illustrating the relationships between the four normalized variables. The histograms on the diagonal show the distribution of each variable, while the scatter plots show the relationships between pairs of variables. The scatter plots in the upper right triangle of the grid annotate the Spearman correlation coefficient and its corresponding *p*-value. Similarly, the lower left triangle of the grid annotates the Kendall correlation coefficient and its corresponding *p*-value. These values can lead to the following interpretations:

AADT vs. IRI: A positive correlation (S = 0.28, K = 0.19), suggesting that as traffic volume increases, so does the roughness of the road, potentially due to more wear and tear.

AADT vs. Fatalities: A positive correlation (S = 0.26, K = 0.18), implying that higher traffic volumes might be associated with an increase in fatalities, due to a greater risk of accidents.

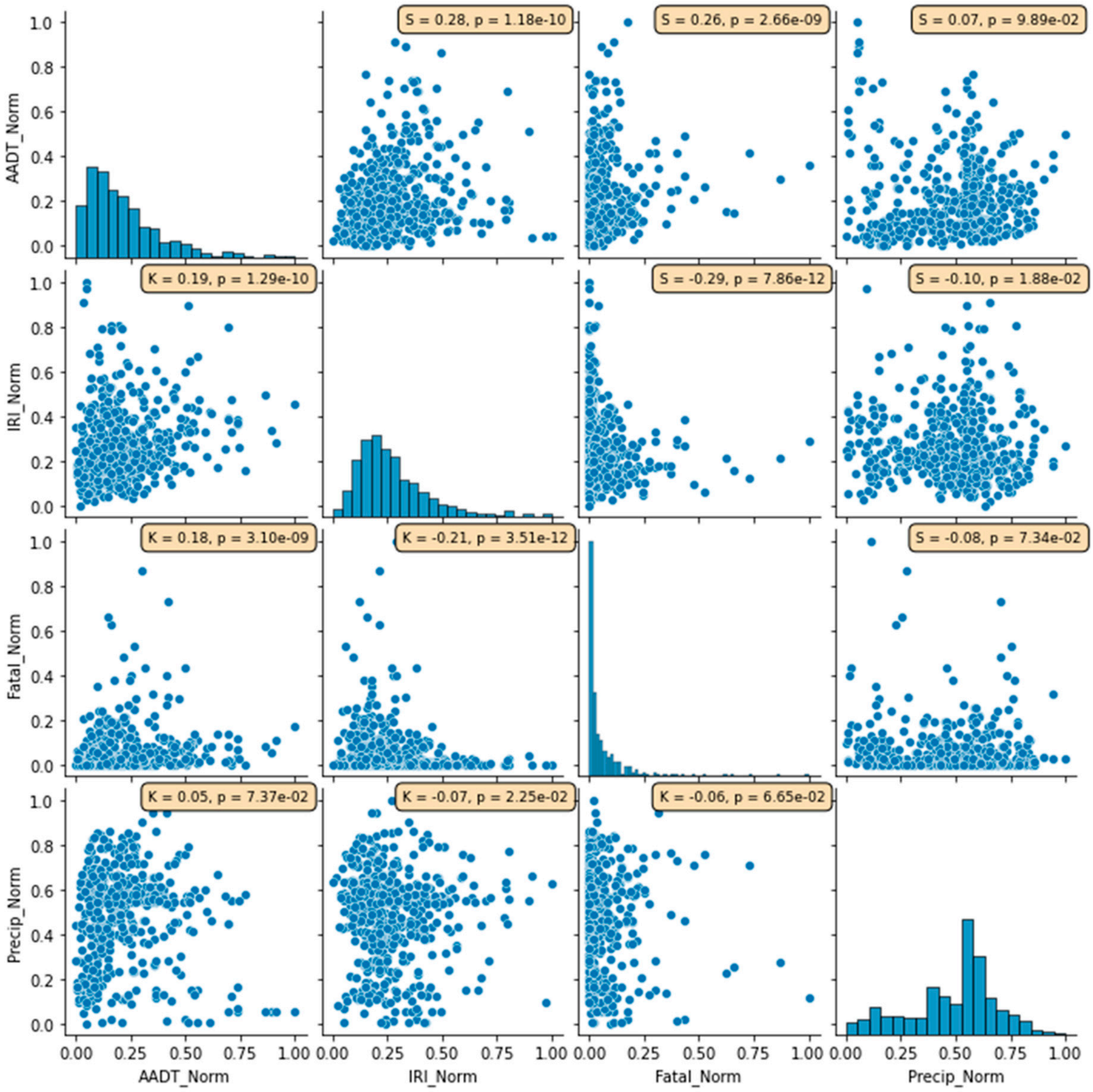

Figure 6. Pairwise plot of the normalized risk components.

AADT vs. Precipitation: A positive correlation (S = 0.07, K = 0.05), which might indicate that regions with higher traffic also tend to have more precipitation, though the correlation is weak and might not be significant.

IRI vs. Fatalities: A negative correlation (S = −0.29, K = −0.21), suggesting that better road quality (lower roughness) could be associated with more fatalities, potentially due to higher speeds on smoother roads.

IRI vs. Precipitation: A negative correlation (S = −0.10, K = −0.07), which is also weak, indicating a negligible relationship between road roughness and precipitation.

Fatalities vs. Precipitation: A negative correlation (S = −0.08, K = −0.06), which could suggest that higher precipitation might correlate with fewer fatalities, though this is counterintuitive; the weak correlation suggests it may not be a significant predictor.

The p-values provide a measure of confidence in the existence of a monotonic relationship. All the p-values are extremely low, indicating that all the results are statistically significant. This indicates that there is sufficient evidence to reject the null hypothesis, which states there is no association (i.e., the correlation is zero) between the two variables. This suggests that there is a monotonic relationship between the two variables.

4.2. Risk Factor Relationships

The histograms along the diagonal plots of Figure 6 show the frequency distributions of each variable, normalized to fit the same scale for comparison. The AADT distribution is right-skewed, indicating a higher frequency of lower traffic volumes. The IRI distribution is also right-skewed, indicating a higher frequency of lower smoother interstate segments. The distribution for fatalities is heavily right-skewed, with most data points toward the lower end, indicating fewer fatalities are more common. The precipitation data show a peak near the center of the distribution, indicating that the precipitation of most interstate segments is near the average, but a minority have either low or high precipitation levels. One caution in interpreting these results is that they do not account for possible confounding variables. Furthermore, correlation does not imply causation, and the relationships observed here would need further analysis to establish any causal links.

Figure 7 is a histogram overlaid with a Gaussian (normal) probability density function. The histogram represents the observed distribution of the composite risk index (Risk Index), while the red curve indicates the estimated normal distribution based on the data. The inset provides key statistics including the mean (1.023) and the standard deviation (STD), which is 0.322. The histogram shows the frequency of the risk index in terms of density, with the x-axis representing the risk index value and the y-axis representing the density of these values. The distribution is symmetric around the mean, which aligns with the assumption of normality. The formal test for normality evaluated the Kolmogorov–Smirnov (KS) statistic (0.023), which is a measure of the maximum distance between the observed cumulative distribution and the expected cumulative distribution. The KS test is preferable when assessing the similarity between distributions without making assumptions about their underlying shapes or parameters [30]. The null hypothesis is that the data are normally distributed. The *p*-value (0.941) of the test indicates the probability of observing the KS statistic (or one more extreme) under the null hypothesis. The high *p*-value suggests that the test cannot reject the null hypothesis, implying that the distribution of the composite risk index does not significantly differ from a normal distribution. Hence, the interpretation is that risks associated with interstate routes are normally distributed. This facilitates the categorization of risks at the standard deviation boundaries. For example, a risk index of less than or greater than one standard deviation, respectively, can represent low and high risk, respectively, whereas risk indices within one standard deviation can represent medium risk.

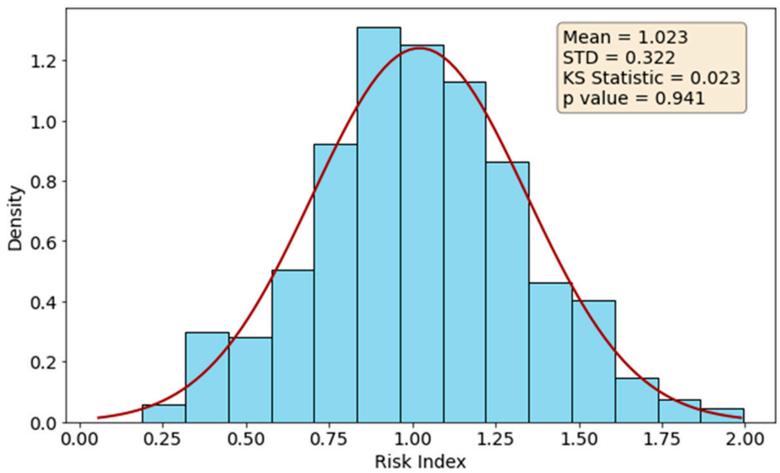

Figure 7. Distribution of the risk index and results of the test for normality.

4.3. Route Risk Distribution

Figure 8 is a map displaying the U.S. Interstate Highway System, with routes color-coded according to the deviation from the mean of the risk index. This map helps with visualizing geographic variations in risk across the interstate network. The routes in blue have an average risk index that is less than one standard deviation below the mean, indicating lower-than-average risk. Conversely, the routes in red have an average risk index that is more than one standard deviation above the mean, indicating higher-than-average risk. The other two colors represent routes that have an average risk index within one standard deviation of the mean, representing slightly lower (green) and slightly higher (orange) than average risk, respectively. The color gradation provides a quick reference to identify areas that might require more in-depth analysis or increased safety measures for AT deployments.

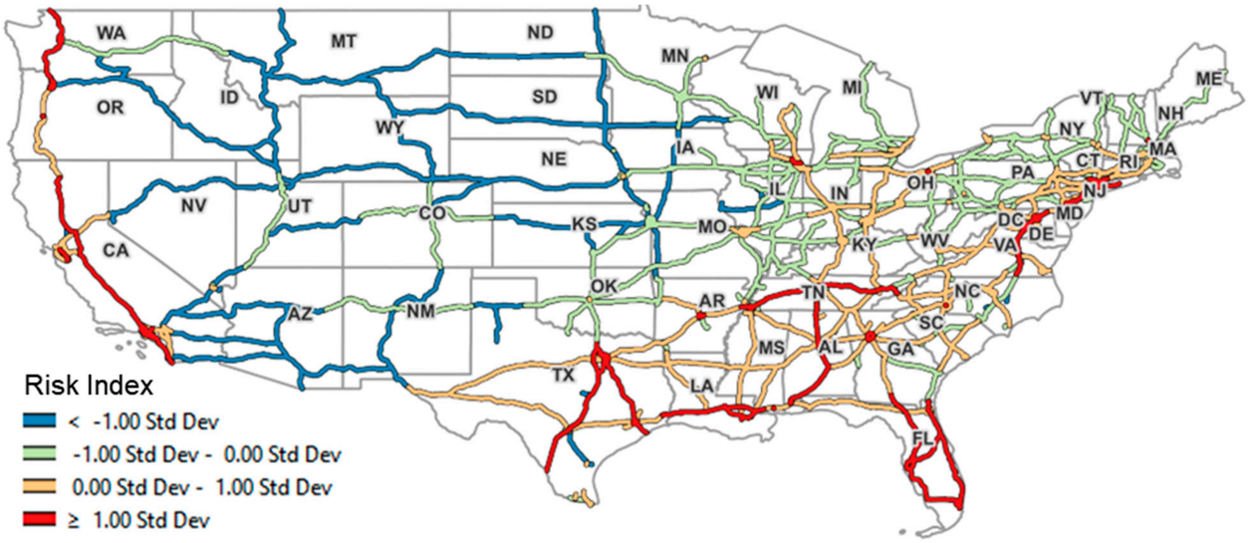

Figure 8. Risk index for interstate routes within each state.

The map integrates the risk index with spatial analytics to inform decision-making processes for transportation planning and safety management. It will help stakeholders to visualize and prioritize resource allocation for infrastructure improvements, enforcement of safety regulations, and implementation of advanced technologies aimed at risk reduction. The concentration of higher-risk routes (indicated in red) in certain areas could be a result of multiple factors, such as high traffic volumes, poor road conditions, higher accident rates, or severe weather conditions. Conversely, routes with lower risk scores are associated with better infrastructure, less traffic, safer driving conditions, or milder weather.

Figure 9 is a stacked bar chart that illustrates the composition of risk indices for various U.S. interstate segments, categorized by regions. The risk components include mean daily traffic, road roughness, fatalities, and precipitation. Each bar represents an interstate segment identified by its number and associated region (e.g., MD_495 in the Mideast region), with the total length of the bar corresponding to the composite risk index for that segment. The chart displays the interstate segments ordered by risk index, with the highest at the bottom. The grayscale color coding within each bar denotes the contribution of each risk component to the total risk index as noted in the "Risk Components" legend. This visualization allows for the comparison of risk profiles across different interstate segments. For example, one can observe that precipitation is a predominant risk factor for many of the interstate segments, especially those with the highest overall risk index. The regional labels provide additional context by showing the geographical distribution of these interstate segments. For instance, segments in the Southeast (e.g., FL_95, FL_395, LA_110) rank near the top risk indices, which could be related to various factors such as higher traffic volumes, road conditions, or regional driving behaviors.

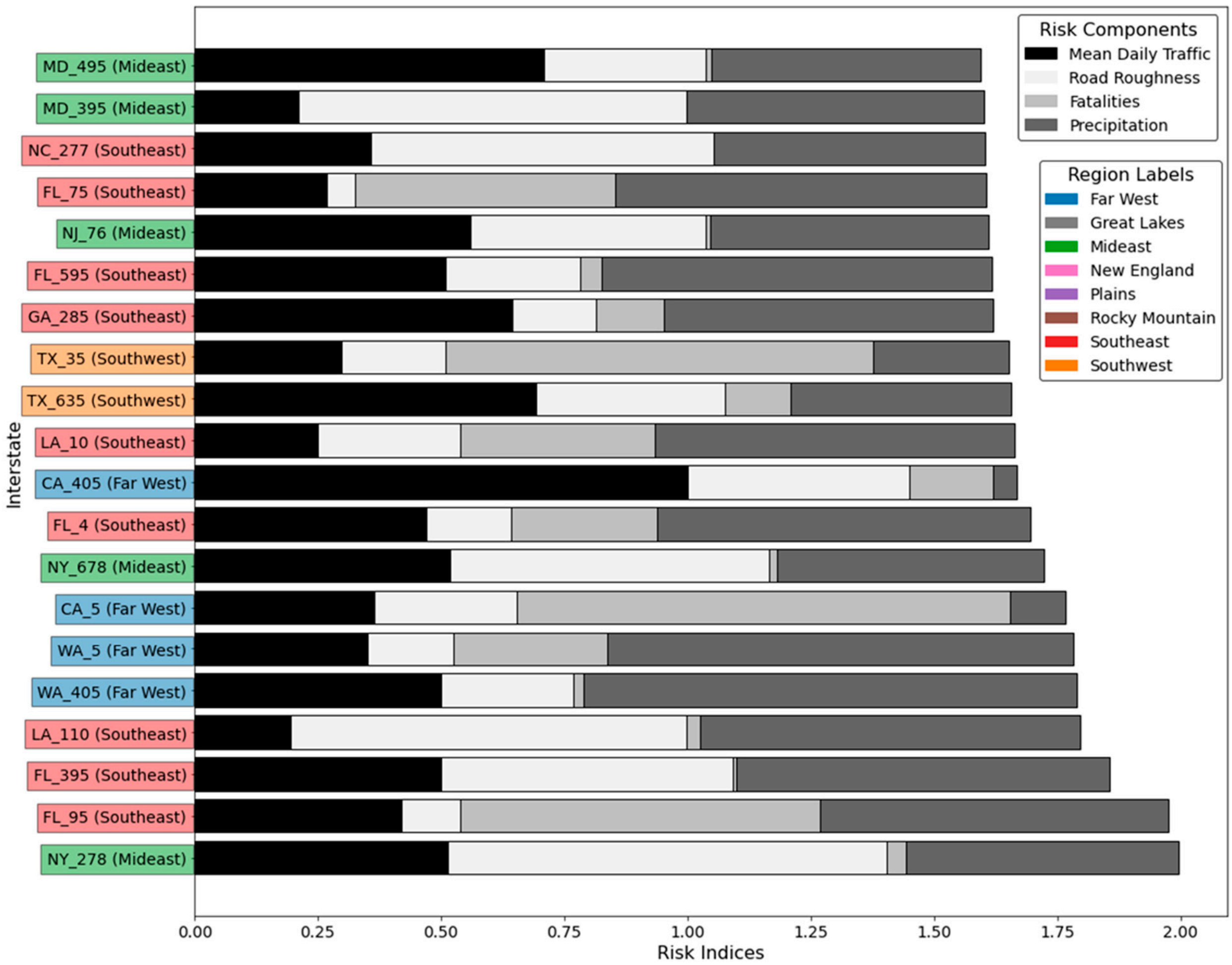

Figure 9. Ranked risk indices by route and region.

Figure 9 is essential for understanding the relative importance of each risk component within each interstate segment and can guide targeted interventions. For transportation planning, this figure emphasizes the need to address specific risk factors on certain routes, which could be beneficial for prioritizing infrastructure investments, policymaking, and risk mitigation strategies, especially in the context of enhancing safety for autonomous trucking operations. Stakeholders can reference this figure when evaluating the application of the risk index in identifying high-priority areas for AT deployment and the need for tailored interventions based on the specific risk profiles of interstate segments considered.

4.4. State Risk Distribution

Figure 10 is a grouped bar chart displaying the average components of the risk index for each state. The states are along the x-axis, ordered by the average risk index from highest to lowest. The y-axis represents the average value of each risk component. Notably, the variation across states is less than the variation across regions, as shown previously in Figure 9. The chart's color-coding identifies four different risk components, as indicated in the legend. The segmentation of each state's bar into these four colors shows the contribution of each risk component to that state's average risk index. This visualization allows for cross-state comparisons of how different risk factors contribute to the overall risk index. For instance, the chart shows that in the state with the highest average risk index, precipitation is the largest contributor, followed by IRI, AADT, and then fatalities. The pattern appears consistent in several states with the highest risk indices, indicating that precipitation and road conditions are significant contributors to the risk index across various states.

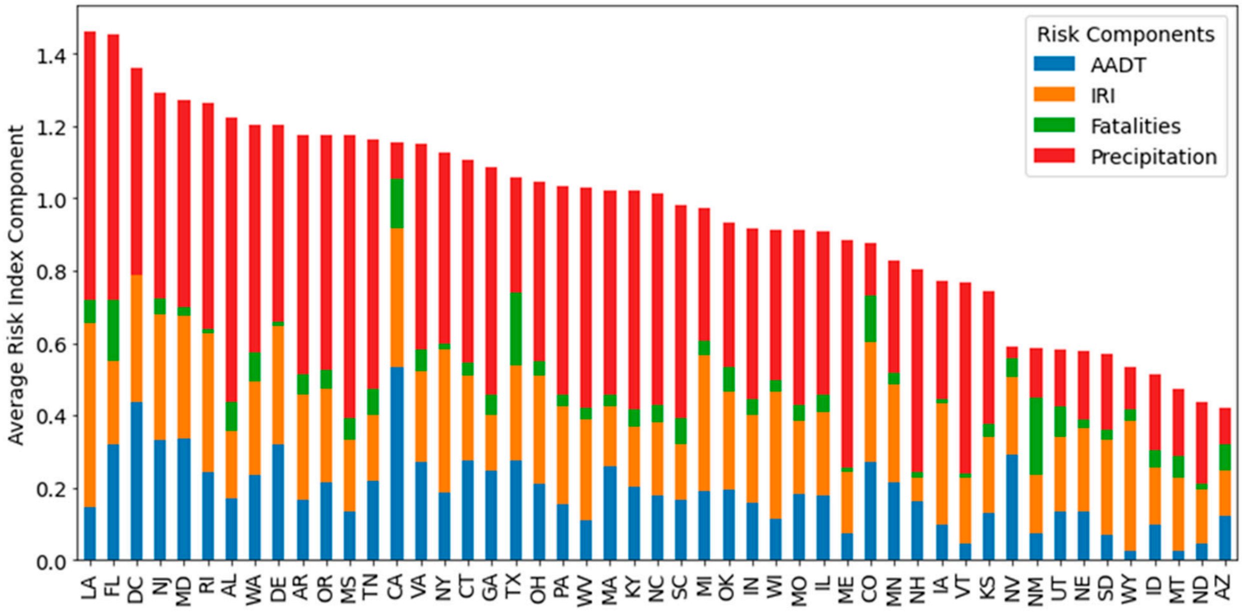

Figure 10. Ranked risk indices by state.

Figure 11 is a choropleth map of the United States, where the color code for each state is the deviation of its average risk index from the mean, measured in standard deviations. This type of map visually conveys the geographic distribution of derived risk indices across states. States colored in blue and red have risk indices that are more than one standard deviation below and above the mean, respectively. The other colors show states with intermediate risk indices, as indicated in the legend. States with higher risk levels are mostly in the Southeast and Mideast. In contrast, the Rocky Mountain and Plains parts of the country exhibit lower-than-average risk levels.

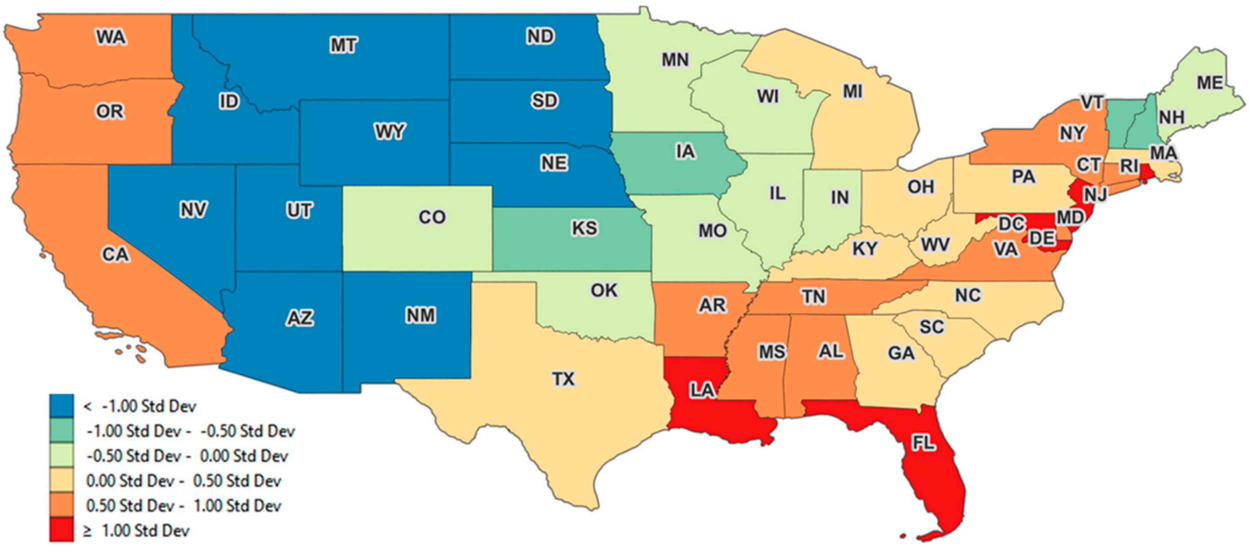

Figure 11. Risk index by state.

The bar chart and accompanying map serve as tools for stakeholders assessing the variation in risk profiles across states, informing federal and state-level decision making on resource allocation for road safety improvements crucial to AT deployment. They can highlight where investment in infrastructure is most needed, particularly in areas where precipitation and road quality are predominant risk factors, thereby enhancing the safety and efficiency of autonomous trucking operations. Moreover, these visualizations aid in understanding regional disparities in traffic volume, road conditions, weather patterns, and safety statistics, which collectively shape the risk landscape. The insights derived from these resources highlight the need for tailored interventions or technology in autonomous trucking routes to address specific regional risk profiles, ensuring safe and efficient transport across the national network.

4.5. Regional Risk Distribution

Figure 12 displays a choropleth map of the United States categorized by a region risk index, which quantifies the risk level across different regions. The color-coded risk index value for each region is according to the legend. The map illustrates significant regional variations in risk, with the Southeastern states exhibiting the highest risk indices, while the Rocky Mountain states have lower risk indices. The central states display a moderate level of risk.

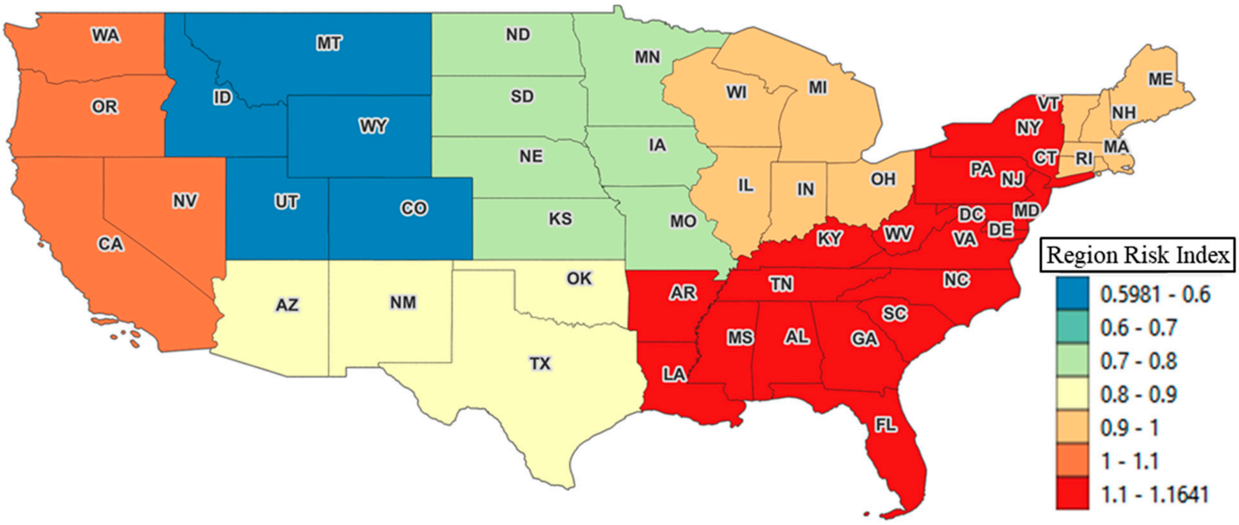

Figure 12. Risk index by region.

Figure 13 is a stacked bar chart that breaks down the composition of the average risk indices for regions in the United States. Each bar corresponds to a region, and the segments within each bar represent the contribution of the four risk components as indicated in the legend. The length of each colored segment within the bars indicates the proportion of each risk component contributing to the region's overall average risk index. The sum of these segments equates to the total average risk index for the region. This visual representation facilitates a comparative analysis of how different risk factors contribute to the overall risk across regions. One observation is that precipitation, IRI, and AADT are significant contributors to risk in all regions. However, the proportion of risk attributed to these factors varies by region, reflecting regional differences in traffic patterns, infrastructure quality, weather conditions, and safety outcomes.

Figure 14 provides further visualization of the relative strength of risk components across regions. The bar chart depicts the mean proportional contribution of various risk factors across different U.S. regions. The y-axis represents the mean proportional contribution, calculated as the average contribution of each risk factor to the total risk index for that region. For instance, in the Far West, AADT makes up the most significant portion of the risk index, while in the Southeast and New England regions, precipitation is the dominant risk factor. The chart provides a clear visual comparison of how the significance of different risk factors varies regionally.

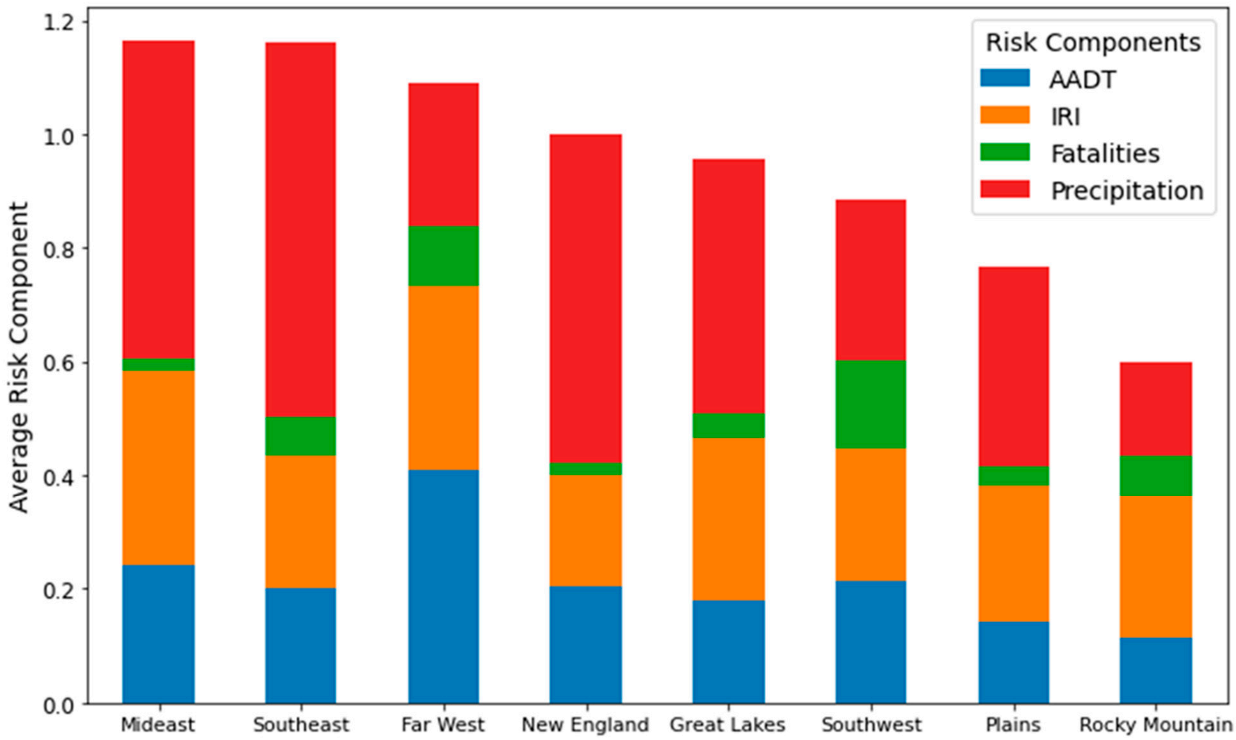

Figure 13. Risk index makeup by region.

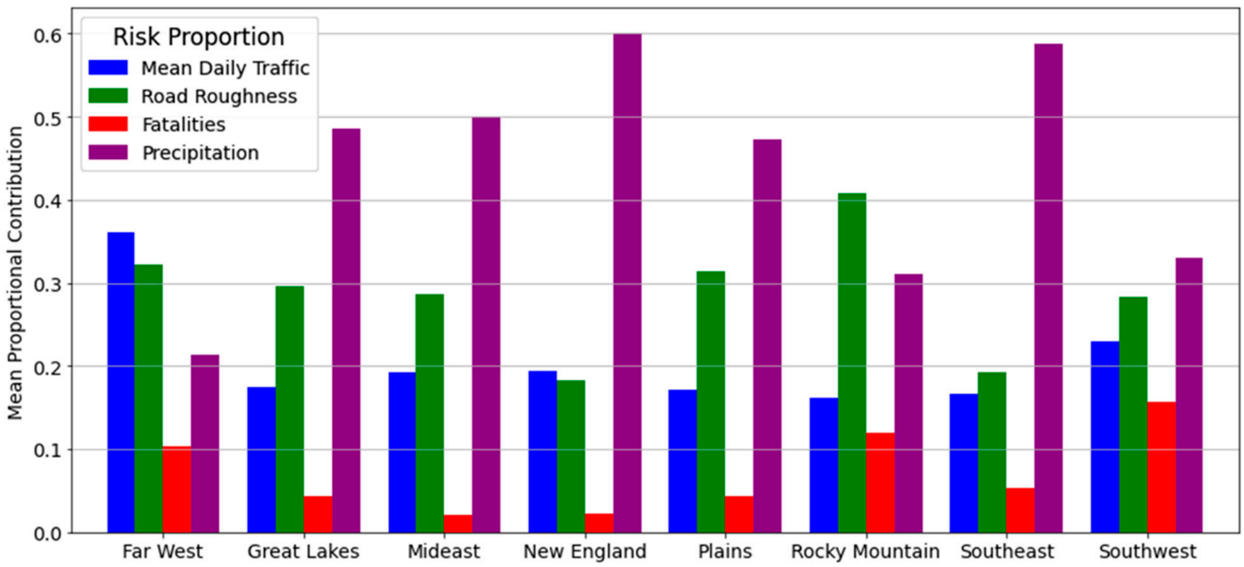

Figure 14. Risk component proportions by region.

Figure 15 is a choropleth map of the United States with each state color-coded to represent the dominant risk factor from the developed risk index. The map provides a clear visual representation of which risk factor is most significant in each state, based on the analysis. For instance, states colored in purple have precipitation as their dominant risk factor, while those in orange are most affected by IRI.

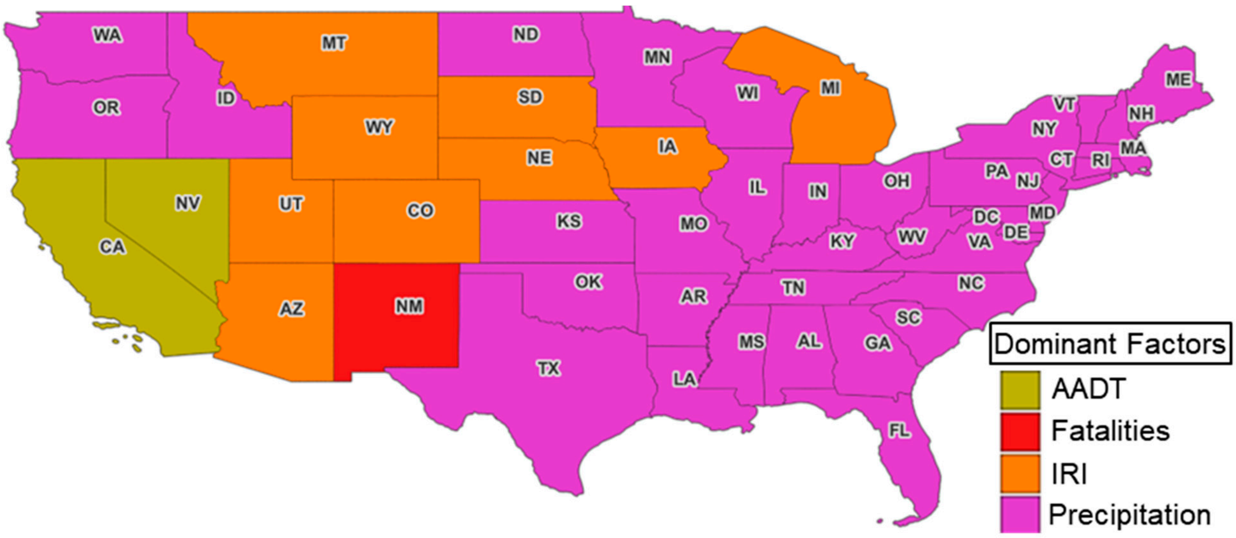

Figure 15. Dominant risk factors by state.

The above maps and their complementary bar charts are useful tools for stakeholders in transportation and logistics to dissect the geographic distribution of risks and strategize accordingly. They facilitate an integrated analysis by illuminating how these regional factors coalesce to form a multifaceted risk landscape. The spatial visualization supports the case for regional prioritization in transportation planning and the development of infrastructure, especially within high-risk areas such as the Southeastern states. Targeting investments to mitigate specific risks could bolster the safety of autonomous trucking routes in those regions. Moreover, the map highlights a critical need for region-specific policy frameworks that adeptly address the varied nature of transportation risks. It highlights the intricate nexus of factors behind regional transportation risks to help stakeholders craft focused, data-driven strategies in transportation management.

The regional disparities in risk contributions suggest a targeted allocation of resources. For instance, regions plagued by road roughness may require heftier infrastructural investments, while those where precipitation poses significant risk could benefit from the adoption of weather-adaptive logistics planning. Together, these visual tools emphasize the imperative for a multifaceted approach to infrastructure enhancement, technological advancement, and nuanced policymaking, all of which are fundamental to amplifying safety and efficiency within the national transportation framework.

5. Discussion

The development of the composite risk index presented in this paper supports the need for a multifaceted assessment of risks associated with the deployment of autonomous trucking across the U.S. interstate system. While previous studies have examined transportation risks in isolation, this work synthesizes these risks into a singular, actionable index. This work also highlighted the critical importance of regional variations in risk factors, a dimension often overlooked in existing models. The analysis of the risk index, as visualized through the series of figures in the results section, reveals a complex tapestry of regional risk factors that influence the safety and efficiency of autonomous freight transportation. By integrating GIS methodologies with national traffic and safety databases, the constructed model captures the multifaceted nature of transportation risks. Using a quantitative approach offers a replicable and scalable model for risk assessment, capable of processing complex datasets to yield precise, actionable insights. The quantitative analysis enables a rigorous, data-driven exploration of risks, distinguishing it from more subjective approaches and emphasizing its value in strategic decision making. Identifying key drivers of risk in different regions contributes a more targeted approach for informed decisions in route planning, infrastructure development, and policy formulation.

General observations are that precipitation emerges as a significant risk factor across most regions, aligning with the pre-existing understanding that weather conditions are a primary contributor to transportation risks. To reduce regional bias, the analysis did not make a distinction between different types of precipitation such as rain versus snow and ice. This is based on the notion that any type of weather condition will affect the accuracy of sensors such as cameras, which can impede operational performance. The findings show how factors such as road roughness, traffic levels, precipitation, and fatality rates contribute differentially across regions. For example, in the Southeast and New England regions, precipitation emerged as a dominant risk factor, whereas, in the Far West, traffic volume and road roughness played more substantial roles. This distinction is crucial, as it highlights the regional specificity of risk profiles and challenges the notion of a one-size-fits-all approach to transportation risk management. Furthermore, the distinction between risk factors across regions emphasizes the need for customized risk management strategies, moving beyond a generic approach to address the specific challenges and requirements of each region.

The implications of these findings are vast. For policymakers, this research provides a data-driven foundation for regional infrastructure planning and AV legislation. For the logistics industry, it offers a strategic guide for long-term route planning and risk mitigation that can enhance the safety and reliability of autonomous trucking operations. Furthermore, this work advances the body of knowledge by providing a methodological framework that can be adapted and applied to other areas of transportation risk assessment and infrastructure management.

Despite its strengths, this research has some limitations. The risk index developed relies on the availability and accuracy of data from national databases, which may not capture real-time changes or the variability in local road conditions and traffic patterns. Furthermore, the normalization of risk factors to a uniform scale, while useful for comparison, may oversimplify the complexity of individual factors. The uncorrelatedness of factors may not capture non-linear effects among them, requiring a more complex evaluation such as employing empirical orthogonal function (EOF) analysis.

The methodology outlined in this study serves as a blueprint for future investigations to adopt advanced statistical models and real-time data integration, further refining the precision and applicability of the risk index. Acknowledging the dynamic nature of transportation risks, future endeavors will explore the potential of machine learning techniques to predict emerging risk patterns, enhancing the adaptability and resilience of autonomous trucking networks to evolving environmental and traffic conditions. For instance, the inclusion of instantaneous precipitation intensity data could enrich the granularity and immediacy of the risk assessment, particularly for short-term operational decisions under specific weather conditions and specific routes. Future research should consider integrating qualitative assessments to enrich the quantitative analysis, offering a more comprehensive understanding of risk factors and their impact on autonomous trucking operations.

The contributions of this work are instrumental for various stakeholders, including transportation planners, policymakers, and the autonomous trucking industry. By providing a comprehensive risk assessment tool, it paves the way for more targeted and effective strategies in transportation safety and logistics management. The methodology and findings of this study set a precedent for future research, which should continue to refine risk assessments and adapt to the evolving landscape of autonomous transportation.

6. Conclusions

The strategic implementation of autonomous trucking hinges on an intricate understanding of the risks involved in navigating the diverse U.S. interstate landscape. This study addressed the critical need for a regionally sensitive risk assessment tool, complementing broad-brush approaches that do not consider the localized intricacies of risk factors influencing transportation safety and efficiency. This comprehensive analysis informed the creation of a composite risk index, meticulously constructed using GISs and corroborated by data from established national databases. The findings illuminate the pronounced regional variations in risk factors, with AADT, IRI, fatalities, and precipitation contributing unevenly across different states. This study substantiated these findings by employing robust methods that synergize spatial analysis with statistical normalization to distill complex datasets into a coherent risk profile. The benefits of this contribution are manifold, particularly for entities involved in the integration of autonomous trucking networks, such as transportation agencies, policymakers, and logistics companies. The detailed maps of risk factors provided can help to inform decision making about where to focus enhanced road safety protocols, prioritize infrastructural investments, and select routes to reduce risk.

This research provides a blueprint for adapting risk assessment methodologies to other domains within transportation and beyond, where state and regional factors significantly influence operational risks. The practical applications of this work range from informing national transportation safety guidelines to aiding in the design of AV navigation systems that can dynamically respond to the risk profiles of the routes they traverse. The ability to adapt and generalize this method to address different scales and scopes of risk assessment is central to its long-term value. This study's approach echoed broader themes such as the interplay between technology and infrastructure, and the importance of data-driven policy.

Future work will focus on addressing the limitations discussed earlier, particularly the need for real-time data integration and the refinement of risk factor weighting within the index. Further research may also explore the application of machine learning techniques to predict and adapt to emerging risk patterns, thereby enhancing the resilience and reliability of autonomous trucking routes. The workflow presented ensures that future explorations into transportation risk assessment are both grounded in empirical evidence and attuned to regional variability.

Author Contributions: Conceptualization, R.B. and R.J.; methodology, R.B. and R.J.; software, R.B. and R.J.; validation, R.B., R.J. and D.T.; formal analysis, R.B. and R.J.; investigation, R.B. and R.J.; resources, D.T.; data curation, R.B. and R.J.; writing—original draft preparation, R.B. and R.J.; writing—review and editing, R.B., R.J. and D.T.; visualization, R.B. and R.J.; supervision, R.B. and D.T.; project administration, D.T.; funding acquisition, D.T. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the United States Department of Transportation, Federal Motor Carrier Safety Administration, grant number FM-MHP-23-001.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are openly available as cited.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Fritschy, C.; Spinler, S. The impact of autonomous trucks on business models in the automotive and logistics industry—A Delphi-based scenario study. *Technol. Forecast. Soc. Chang.* **2019**, *148*, 119736. [CrossRef]
2. Xiaoxiang, M.; Zhimin, T.; Feng, C. A reliability-based approach to evaluate the lateral safety of truck platoon under extreme weather conditions. *Accid. Anal. Prev.* **2022**, *174*, 106775. [CrossRef] [PubMed]
3. Simpson, J.R.; Sharma, I.; Mishra, S. Modeling trucking industry perspective on the adoption of connected and autonomous trucks. *Res. Transp. Bus. Manag.* **2022**, *45*, 100883. [CrossRef]
4. Sindi, S.; Woodman, R. Implementing Commercial Autonomous Road Haulage in Freight Operations: An Industry Perspective. *Transp. Res. Part A Policy Pract.* **2021**, *152*, 235–253.
5. Lee, S.; Cho, K.; Park, H.; Cho, D. Cost-Effectiveness of Introducing Autonomous Trucks: From the Perspective of the Total Cost of Operation in Logistics. *Appl. Sci.* **2023**, *13*, 10467. [CrossRef]
6. Rana, M.M.; Hossain, K. Impact of autonomous truck implementation: Rutting and highway safety perspectives. *Road Mater. Pavement Des.* **2022**, *23*, 2205–2226. [CrossRef]
7. Zhao, J.; Lee, J.Y. Effect of Connected and Autonomous Vehicles on Supply Chain Performance. *Transp. Res. Rec.* **2023**, *2677*, 402–424. [CrossRef]

8. Gružauskas, V.; Baskutis, S.; Navickas, V. Minimizing the trade-off between sustainability and cost effective performance by using autonomous vehicles. *J. Clean. Prod.* **2018**, *184*, 709–717. [CrossRef]
9. Chen, C.-M.; Lu, Y. Shipment sizing for autonomous trucks of road freight. *Int. J. Logist. Manag.* **2021**, *32*, 413–433. [CrossRef]
10. Rana, M.M.; Hossain, K. Connected and autonomous vehicles and infrastructures: A literature review. *Int. J. Pavement Res. Technol.* **2023**, *16*, 264–284. [CrossRef]
11. Mishler, S.; Chen, J. Effect of Automation Failure Type on Trust Development in Driving Automation Systems. *Appl. Ergon.* **2023**, *106*, 103913. [CrossRef] [PubMed]
12. Bridgelall, R.; Jones, R.; Tolliver, D. Ranking Opportunities for Autonomous Trucks Using Data Mining and GIS. *Geographies* **2023**, *3*, 806–823. [CrossRef]
13. Jiang, J.; Seaid, M.; Mohamed, M.S.; Li, H. Inverse algorithm for real-time road roughness estimation for autonomous vehicles. *Arch. Appl. Mech.* **2020**, *90*, 1333–1348. [CrossRef]
14. Matulla, C.; Hollósi, B.; Andre, K.; Gringinger, J.; Chimani, B.; Namyslo, J.; Fuchs, T.; Auerbach, M.; Herrmann, C.; Sladek, B.; et al. Climate Change driven evolution of hazards to Europe's transport infrastructure throughout the twenty-first century. *Theor. Appl. Climatol.* **2018**, *133*, 227–242. [CrossRef]
15. Zhang, Y.; Carballo, A.; Yang, H.; Takeda, K. Perception and sensing for autonomous vehicles under adverse weather conditions: A survey. *ISPRS J. Photogramm. Remote Sens.* **2023**, *196*, 146–177. [CrossRef]
16. Kim, E.; Kim, Y.; Park, J. The necessity of introducing autonomous trucks in logistics 4.0. *Sustainability* **2022**, *14*, 3978. [CrossRef]
17. Al-Turki, M.; Ratrout, N.T.; Rahman, S.M.; Reza, I. Impacts of autonomous vehicles on traffic flow characteristics under mixed traffic environment: Future perspectives. *Sustainability* **2021**, *13*, 11052. [CrossRef]
18. Ghoul, T.; Sayed, T.; Fu, C. Real-time safest route identification: Examining the trade-off between safest and fastest routes. *Anal. Methods Accid. Res.* **2023**, *39*, 100277. [CrossRef]
19. Elgharbawy, M. A big testing framework for automated truck driving. *Urban Transp. Constr.* **2019**, *4*, e27.
20. Shi, Y.; Huang, Y.; Chen, Y. Trajectory planning of autonomous trucks for collision avoidance with rollover prevention. *IEEE Trans. Intell. Transp. Syst.* **2021**, *23*, 8930–8939. [CrossRef]
21. Alawadhi, M.; Almazrouie, J.; Kamil, M.; Khalil, K.A. Review and Analysis of the Importance of Autonomous Vehicles Liability: A Systematic Literature Review. *Int. J. Syst. Assur. Eng. Manag.* **2020**, *11*, 1227–1249. [CrossRef]
22. Girdhar, M.; Hong, J.; Moore, J. Cybersecurity of Autonomous Vehicles: A Systematic Literature Review of Adversarial Attacks and Defense Models. *IEEE Open J. Veh. Technol.* **2023**, *4*, 417–437. [CrossRef]
23. Tao, X.; Mårtensson, J.; Warnquist, H.; Pernestål, A. Short-term maintenance planning of autonomous trucks for minimizing economic risk. *Reliab. Eng. Syst. Saf.* **2022**, *220*, 108251. [CrossRef]
24. Zhang, F.; Xia, R.; Chen, X. An optimal trajectory planning algorithm for autonomous trucks: Architecture, algorithm, and experiment. *Int. J. Adv. Robot. Syst.* **2020**, *17*, 2. [CrossRef]
25. BTS. Highway Performance Monitoring System (HPMS) 2020. 6 June 2023. Available online: https://geodata.bts.gov/datasets/highway-performance-monitoring-system-hpms-2020/about (accessed on 25 February 2024).
26. BTS. 2021 Fatal Motor Vehicle Accidents (FARS). 16 November 2023. Available online: https://geodata.bts.gov/datasets/usdot::2021-fatal-motor-vehicle-accidents-fars/about (accessed on 25 February 2024).
27. NCEI. Climate at a Glance County Mapping. National Oceanic and Atmospheric Administration (NOAA). 2024. Available online: https://www.ncei.noaa.gov/access/monitoring/climate-at-a-glance/county/mapping/110/pcp/202201/12/mean (accessed on 26 February 2024).
28. Bridgelall, R.; Tolliver, D. Deciphering autonomous vehicle regulations with machine learning. *Appl. Sci.* **2024**, *14*, 1396. [CrossRef]
29. Siegel, S.; Castellan, N.J., Jr. *Nonparametric Statistics for The Behavioral Sciences*, 2nd ed.; McGraw-Hill: New York, NY, USA, 1988.
30. Boslaugh, S. *Statistics in a Nutshell: A Desktop Quick Reference*; O'Reilly Media, Inc.: Sebastopol, CA, USA, 2013.