



HOKKAIDO UNIVERSITY

Title	Risk Assessment of Bridge Damage Due to Heavy Rainfall Considering Landslide Risk and Driftwood Generation Potential Using Convolutional Neural Networks and Conventional Machine Learning
Author(s)	Ren, Fudong; Isobe, Koichi; Ando, Miku
Citation	Water, 16(23), 3471 https://doi.org/10.3390/w16233471
Issue Date	2024-12-02
Doc URL	https://hdl.handle.net/2115/100111
Rights(URL)	https://creativecommons.org/licenses/by/4.0/
Type	journal article
File Information	water_16-23_3471.pdf



Article

Risk Assessment of Bridge Damage Due to Heavy Rainfall Considering Landslide Risk and Driftwood Generation Potential Using Convolutional Neural Networks and Conventional Machine Learning

Fudong Ren ^{1,*}, Koichi Isobe ²  and Miku Ando ¹

¹ Graduate School of Engineering, Hokkaido University, Sapporo 060-8628, Hokkaido, Japan; ando.miku.i4@elms.hokudai.ac.jp

² Faculty of Engineering, Hokkaido University, Sapporo 060-8628, Hokkaido, Japan; kisobe@eng.hokudai.ac.jp

* Correspondence: fudong.ren.v3@elms.hokudai.ac.jp; Tel.: +81-011-706-6203

Abstract: This study addresses the assessment of bridge damage risks associated with heavy rainfall, focusing on landslide susceptibility and driftwood generation potential. By integrating convolutional neural networks (CNNs) with traditional machine learning methods, the research develops an advanced predictive framework for estimating driftwood accumulation at river bridges—a recognized challenge in disaster management. Concentrating on the Tokachi River basin in Hokkaido, Japan, the research utilizes diverse environmental and geographical data from authoritative sources. The findings demonstrate that the innovative approach not only enhances the accuracy of driftwood volume predictions but also distinguishes the effectiveness of CNNs compared to conventional methods. Crucially, areas prone to landslides are identified as significant contributors to driftwood generation, impacting bridge safety. The study underscores the potential of machine learning models in improving disaster risk assessment, while suggesting further exploration into real-time data integration and model refinement to adapt to changing climate conditions and ensure long-term infrastructure safety.



Citation: Ren, F.; Isobe, K.; Ando, M. Risk Assessment of Bridge Damage Due to Heavy Rainfall Considering Landslide Risk and Driftwood Generation Potential Using Convolutional Neural Networks and Conventional Machine Learning. *Water* **2024**, *16*, 3471. <https://doi.org/10.3390/w16233471>

Academic Editor: Giuseppe Oliveto

Received: 15 November 2024

Revised: 28 November 2024

Accepted: 30 November 2024

Published: 2 December 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: bridge disaster risk; driftwood generation potential; landslide risk; CNN; machine learning

1. Introduction

In recent years, rainfall patterns have become increasingly concentrated and intense, a trend that is likely to continue. Such heavy rainfall not only increases river flow but also leads to a significant rise in the sediment and driftwood that rivers carry. Notably, driftwood often becomes trapped by bridges, blocking river channels and exacerbating flooding and erosion upstream. In this context, it is essential to identify bridges that are highly susceptible to these hazards and prioritize them for the installation of effective disaster prevention measures and necessary repairs to ensure a resilient transportation network. Research into the mechanisms of driftwood generation, its movement downstream, and its deposition in waterways is ongoing. Figure 1 shows the causes of driftwood and forms of bridge damage due to driftwood. In August 2023, the Erosion and Sediment Control Department of the Ministry of Land, Infrastructure, Transport, and Tourism underscored the need for driftwood management [1]. Establishing a rational and efficient method for predicting the volume of driftwood that may accumulate at river bridges is recognized as a critical challenge.

Evaluating the risks associated with driftwood disasters at river bridges is a critical aspect of ensuring infrastructure resilience, particularly in regions prone to severe weather events and flooding. Minami and his team examined driftwood conditions generated during the August 1998 disaster at the Yosasa River in Tochigi Prefecture, identifying the

sources, distribution, and impact of driftwood following the flood, contributing to a better understanding of driftwood behavior and its implications for disaster management [2]. Yamada with his partner investigated the damage to riparian forests and the generation and accumulation of driftwood in the Atsubetsu River basin during the Typhoon No. 10 disaster in 2003, providing insights into the impacts of severe weather on forested riverbanks and the subsequent movement of driftwood, aiding in developing strategies to manage these natural hazards [3]. Yano and his team introduced a concept for assessing potential woody debris risks to bridges from nearby forests, using GIS data and forest classification. It successfully quantified debris generation potential and correlated it with historical flood risks at bridges along Japan's Kageetsu River [4]. Dozono developed a risk assessment method for predicting driftwood disasters at bridges by analyzing the potential for driftwood generation in river basins. It provides a systematic approach to evaluate the risks and impacts of driftwood on bridge structures, enhancing the preparedness and mitigation strategies against such natural hazards [5]. Bangnira and his team explored the susceptibility of UK bridges to instream wood accumulation, analyzing the environmental and structural factors contributing to wood buildup at bridge sites and emphasizing the need for improved design and river management practices [6]. Yano analyzed the causes of driftwood generation during the 2017 Northern Kyushu heavy rain, identifying unstable slopes, poor forest management, and extreme rainfall as key factors [7]. Kakiyama constructed a model to estimate driftwood risk based on topographical factors like slope gradient and land use, aiming to identify high-risk areas [8]. Tanaka assessed woody debris risk by examining terrain characteristics linked to slope failure, highlighting the relationship between slope steepness, instability, and woody debris production [9]. Collectively, these studies emphasize the essential need to integrate driftwood risk assessments into bridge design and river management, ultimately contributing to the enhanced safety and sustainability of infrastructure planning.

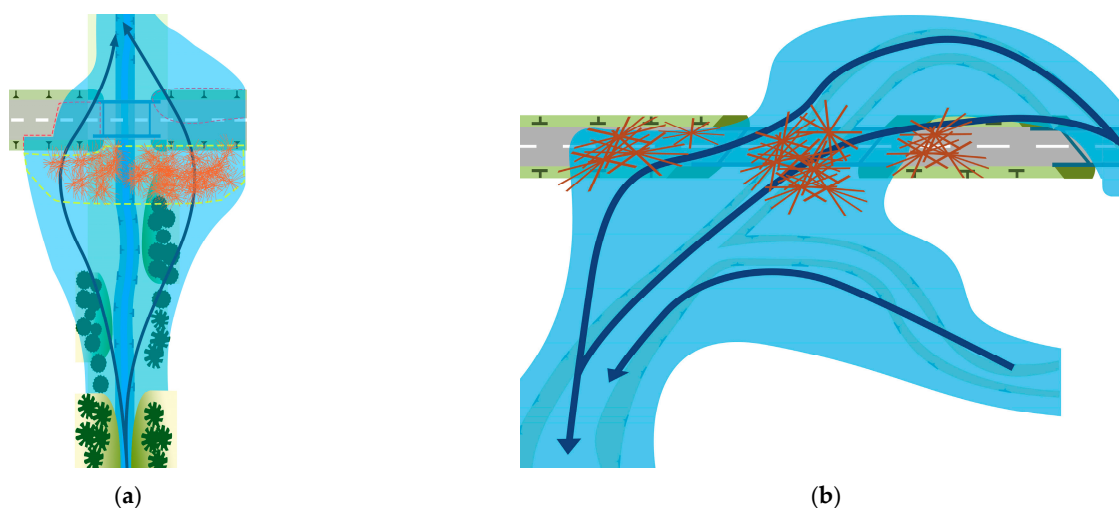


Figure 1. Bridge damage caused by driftwood, as reported in survey findings from the Civil Engineering Bureau, Road Division, Hokkaido Prefectural Government. (a) River channel blockage due to driftwood and embankment erosion behind the abutment. (b) Subsidence of bridge piers due to riverbed scouring and channel blockage from driftwood.

The innovation of this study lies in its refined approach to identifying sources of driftwood generation compared to previous research. The primary goal of this study is to develop a more accurate and robust method for predicting and managing the risks associated with driftwood in river environments, thereby improving safety in bridge infrastructure. While earlier studies [4,5] mainly focused on areas with slopes exceeding 30° as sources, this study advances the methodology by incorporating landslide risk maps generated from CNNs and traditional machine learning predictions. This integration aims

to more precisely delineate the impact of slope collapses on driftwood generation potential. Moreover, the calculated driftwood generation potential and driftwood capture amount are utilized as key features in assessing bridge risk. These features, combined with existing environmental and bridge characteristics, enhance the predictive framework for bridge risk. Additionally, this approach enables a comparison between CNN-based and traditional machine learning techniques in bridge disaster risk assessment.

2. Study Area and Data Use

2.1. Study Area

The focus of this study is the Tokachi River basin, as shown in Figure 2a,b, located in the central and eastern parts of Hokkaido, near Obihiro City in Japan. In 2016, Typhoon No. 10 inflicted severe damage on the area, leading to the discharge of substantial quantities of driftwood. Over 20,000 m³ of driftwood accumulated near the mouth of the Tokachi River, causing significant disruption. Additionally, several bridges in the area sustained significant damage, exhibiting various types of failure patterns [10].

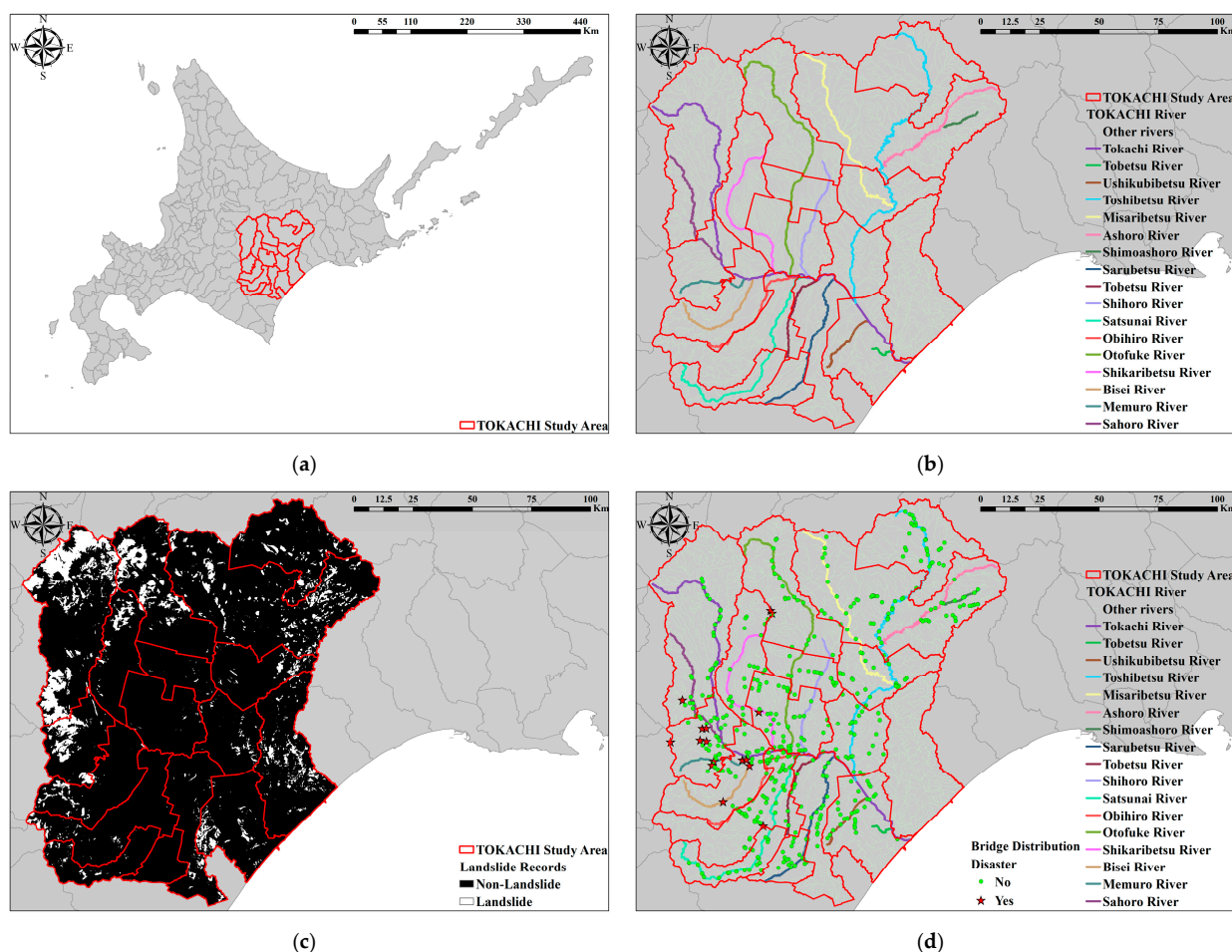


Figure 2. (a) Location of study area, (b) Distribution of major rivers, (c) Distributions of historical landslides, (d) Distribution of bridges in the study area.

2.2. Data Used in This Study

2.2.1. Data Used for Landslide Risk Assessment

In this study, we divided the entire target basin into 10 m mesh grids to evaluate slope disaster risk and estimate driftwood potential. For landslide risk prediction, we utilized the landslide topography distribution map database provided by the Disaster Prevention Science Institute [11]. The data for the landslide factors were generated using

national numerical information from the Geospatial Information Authority of Japan and the Ministry of Land, Infrastructure, Transport, and Tourism [12]. Terrain information, including elevation, slope angle, slope direction, and average annual rainfall, was derived from 10 m DEM data provided by the Geospatial Information Authority of Japan's base map information download service, using ArcGIS tools. Qualitative factors such as terrain, soil, and surface conditions were also extracted every 10 m from the same service [13]. Additionally, a vegetation classification map was created using vegetation survey data from the Ministry of the Environment [14]. These are detailed in Table 1, Figures 3 and A1.

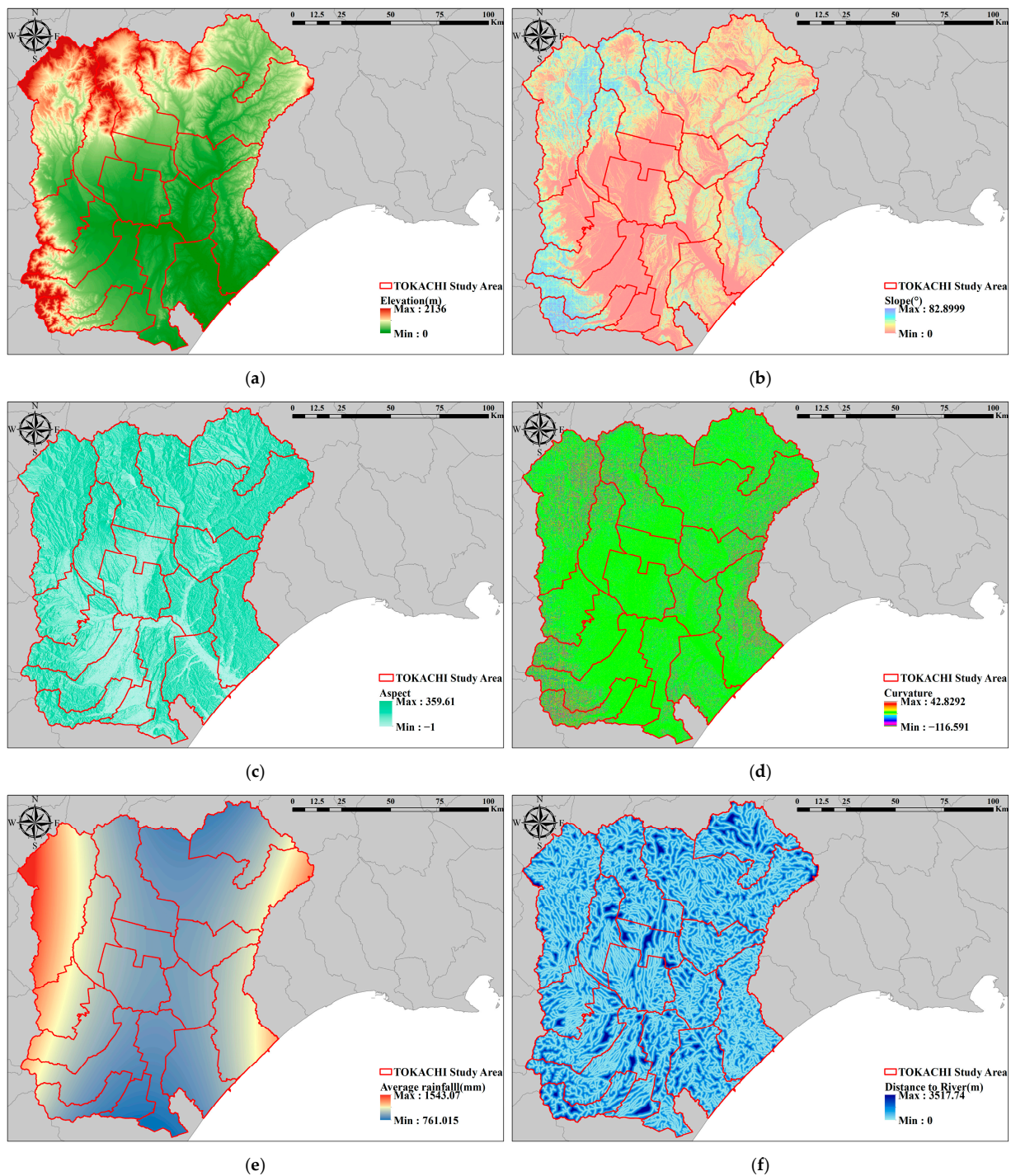


Figure 3. Cont.

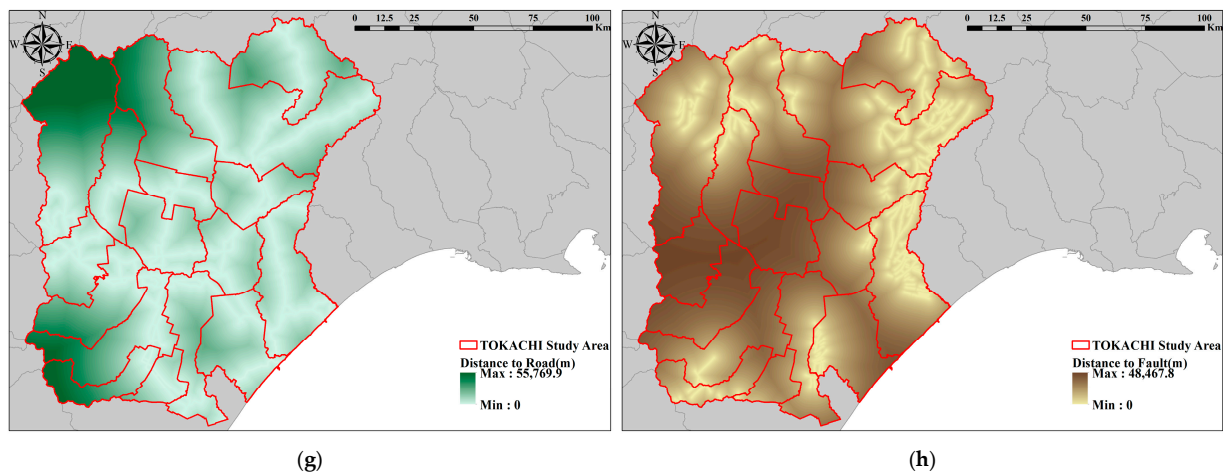


Figure 3. Maps of landslide features in the study area. (a) Elevation (m), (b) Slope (degrees), (c) Aspect, (d) Curvature, (e) Average rainfall (mm), (f) Distance to rivers (m), (g) Distance to roads (m), (h) Distance to faults (m).

Table 1. Features and data range in the study area for landslide risk.

Feature	Resource	TOKACHI
HLS	NIED ¹	0, 1
DEM (m)	GIAJ ²	0 to 2136
Slope (°)	DEM ³	0 to 82.8999
Aspect	DEM	−1 to 359.61
Curvature (dimensionless)	DEM	−116.591 to 42.9292
Average Rainfall (mm)	JMA ⁴	761.015 to 1543.07
Distance to River (m)	MLIT ⁵	0 to 3517.74
Distance to Road (m)	MLIT	0 to 55,769.9
Distance to Fault (m)	GSJ ⁶	0 to 48,467.8
Land Use	MLIT	12 types
Soil	MLIT	46 types
Lithology	GSJ	66 types
Vegetation	BiodiC-J ⁷	79 types
Catchment Area (km ²)	DEM	1 to 12,762
Openness (°)	DEM	0 to 90

Notes: ¹ NIED: National Research Institute for Earth Science and Disaster Resilience. ² GIAJ: Geospatial Information Authority of Japan. ³ DEM: Digital Elevation Map with 10 m × 10 m resolution published by Geospatial Information Authority of Japan. ⁴ JMA: Japan Meteorological Agency. ⁵ MLIT: Ministry of Land, Infrastructure, Transport, and Tourism. ⁶ GSJ: Geological Survey of Japan, National Institute of Advanced Industrial Science and Technology. ⁷ BiodiC-J: Biodiversity Center of Japan.

2.2.2. Data Used to Assess the Risk of River Crossing Bridges During Heavy Rain

The dataset for disaster risk prediction focuses on road bridges along Tokachi prefectural roads and national highways. It excludes structures such as overpasses, rail bridges, and the bridges for which data were not available from the general bridge diagrams [15]. In total, the dataset encompasses 515 bridges, 16 of which sustained damage. The history of bridge disasters was compiled by reviewing local government websites and disaster reports from the Japanese Geotechnical Society [16–18], specifically for the heavy rainfall event in August 2016. Bridge registers were provided by the Ministry of Land, Infrastructure, Transport, and Tourism Hokkaido Development Bureau, and general bridge diagrams were supplied by the Hokkaido Department of Construction, as shown in Table 2. Figure 2d illustrates the distribution of bridges in the study area.

Table 2. Features and data range in the study area used for bridge risk assessment.

Feature	Resource	TOKACHI
Bridge Disaster History	MLIT-HRDB ¹ HPG-CD ²	0, 1
River Width (m)	GIAJ ³	0.3 to 99.920436
Curvature (dimensionless)	GIAJ	−116.591 to 42.9292
Catchment Area (km ²)	DEM ⁴	1 to 12,762
Riverbed Slope (%)	DEM	0 to 4.22
Surface Geology	MLIT ⁵	6 types
Intersection Angle (°)	GIAJ	33.3273 to 145.3979
Span Number		1 to 23
Completion Year	MLIT-HRDB	6 types
Bridge length (m)	HPG-CD	3.0 to 984.2
River Flow Obstruction Rate		0 to 0.0932
Driftwood Capture Amount (m ³)		0 to 319.406113 by CNN

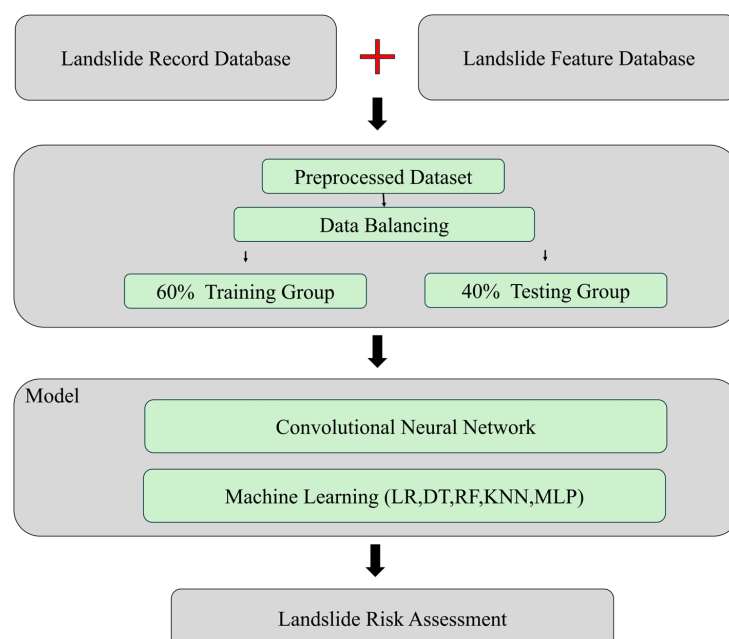
Notes: ¹ MLIT-HRDB: Ministry of Land, Infrastructure, Transport, and Tourism, Hokkaido Regional Development Bureau, Construction Department, Road Construction Division. ² HPG-CD: Hokkaido Prefectural Government, Construction Department, Civil Engineering Bureau Road Division. ³ GIAJ: Geospatial Information Authority of Japan. ⁴ DEM: Digital Elevation Map with 10 m × 10 m resolution published by Geospatial Information Authority of Japan. ⁵ MLIT: Ministry of Land, Infrastructure, Transport, and Tourism.

3. Method

This section may be divided by subheadings. It should provide a concise and precise description of the experimental results, their interpretation, as well as the experimental conclusions that can be drawn.

3.1. Landslide Risk Assessment

The detailed workflow of landslide risk is shown in Figure 4, and can be broadly categorized into four stages. Initially, the database is established, and all the feature data are integrated and unified using ArcGIS. Subsequently, the data are preprocessed, and the significance of the features is assessed. Following that, a training model is constructed, and its performance is evaluated. Finally, the preprocessed data are reintroduced into the training model for testing and prediction in the fourth step.

**Figure 4.** Flowchart of landslide risk assessment.

3.2. Driftwood Generation Potential and Bridge Risk

The origins of driftwood include the outflow of standing trees, fallen trees, logged timber, and materials used for construction. In our study, we identify landslides as the primary cause of driftwood generation, particularly in mountainous regions where landslides involve coniferous or broadleaf tree species. We assume that tree collapses due to landslides are the main natural source of driftwood. This study focuses on the sliding of standing trees associated with slope failures to understand the mechanism of tree movement and its environmental impact. By integrating these natural sources with bridge feature data, we use a machine learning model to predict and quantify the risk to bridges, effectively assessing bridge risk within our framework based on the potential for driftwood generation and capture, as shown in Figure 5.

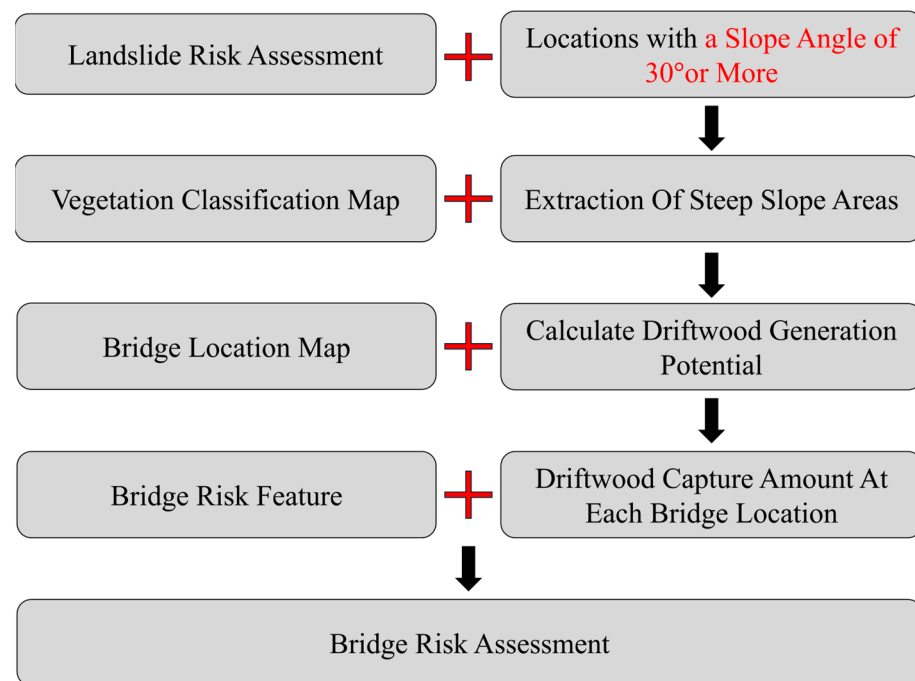


Figure 5. Flowchart for assessing the risk of river-crossing bridges during heavy rain.

As previously discussed, the estimation of driftwood generation potential in this study is based on previously proposed methodology [4,5]. The potential volume of driftwood is calculated by assuming that all the areas identified as having a high risk of landslides actually experience collapses. This approach allows for the estimation of the maximum potential driftwood volume that could be generated by landslides.

$$V_i = \sum_k \beta_k A_{ki} \quad (1)$$

Here, V_i (km^3) represents the driftwood generation potential at any given bridge, β_k (m^3/km^2) denotes the volume of driftwood runoff per unit area of forest for each specific tree type k , and A_{ki} (km^2) represents the forest area within the driftwood supply source at location i in the watershed. The value of β_k varies based on forest vegetation types [19], with coniferous trees assigned a value of 1000, broadleaf trees 100, and barren lands 0, as illustrated in Figure 6a.

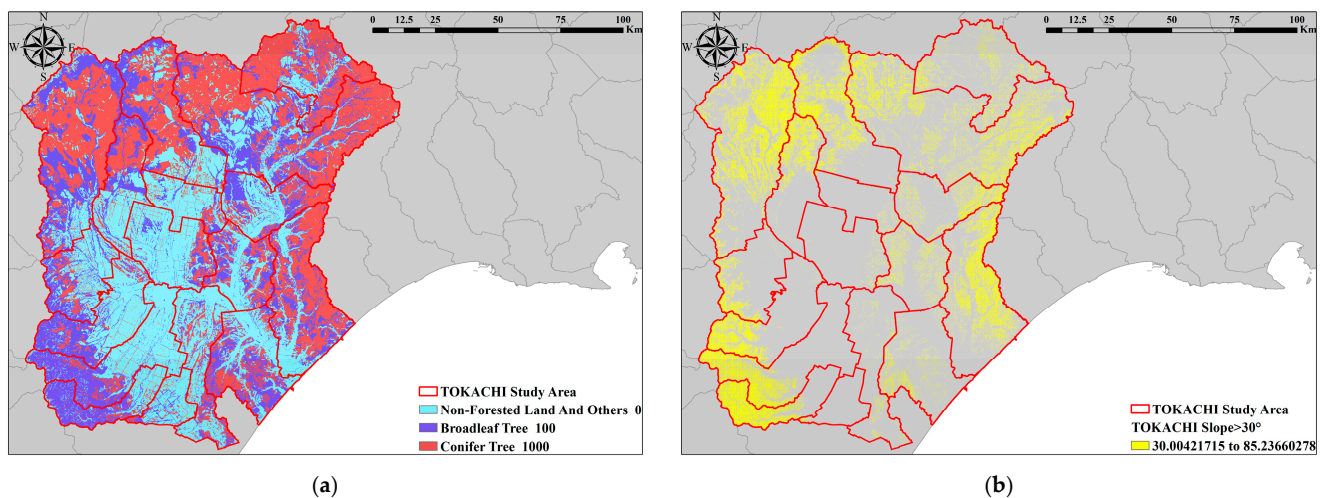


Figure 6. Distribution map of (a) vegetation reclassification and (b) landslide susceptibility area with slopes exceeding 30 degrees.

The driftwood capture rate estimates the amount of driftwood intercepted by a specific bridge. Driftwood originating upstream is carried downstream and often accumulates at river crossings, including bridges and dams. The likelihood of this driftwood being captured is influenced by factors such as the river's width and the bridge's span. The formula that describes this relationship is provided below.

$$\alpha_i = \frac{L_w}{W_i} \times \frac{1}{k} \quad (2)$$

Here, L_w represents the assumed length of driftwood determined from field surveys in forests within the watershed. However, in this study, we assumed $L_w = (\text{DSM} - \text{DEM})/2$, which will be explained in detail later. $W_i(m)$ represents the minimum channel width at high water levels at bridge i , and k is a coefficient representing the difficulty of driftwood accumulation at the bridge. Due to insufficient data on bridges and driftwood, it is impossible to determine the constant k , but regardless of the numerical value, it is possible to represent the amount of driftwood captured relatively [5]. In this study, we set $k = 1$ so that the driftwood capture rate does not exceed 1, calculate the relative amount of driftwood captured by each bridge, and compare the driftwood capture trends between different bridges.

To determine the amount of driftwood each bridge captures based on the calculated driftwood generation potential, as shown in Figure 7, the following recurrence formula is used:

$$\tilde{V}_i = \left(V_i - \tilde{V}_{i-1} - \tilde{V}_{i-2} \dots \tilde{V}_1 \right) \alpha_i = \alpha \left(V_i - \sum_{k=1}^{i-1} \tilde{V}_i \right) \quad (3)$$

Here, $\tilde{V}_i$ represents the amount of driftwood captured at bridge i , and α_i represents the driftwood capture rate for bridge i .

In this research, areas with slopes exceeding 30 degrees, as indicated on the landslide susceptibility map (see Section 3.1 and Figure 6b), are identified as potential driftwood sources. By concentrating on high-risk landslide zones with steep slopes, the study targets primary sources of driftwood, which are largely derived from trees that fall due to landslides. The maximum distance L_{max} that driftwood can travel to reach river channels is determined using an empirical formula that estimates soil displacement from slope collapses, taking into account the slope's vertical profile [11]. Trees located where the distance from the river to the slope exceeds the estimated maximum distance are excluded

from the driftwood production estimate, while trees within this range are included, as shown in Figure 8 [4,20].

$$\frac{H}{L_{max}} = 0.73 \tan \theta - 0.21 \quad (4)$$

Here, H' , L' , and θ ($\tan \theta = H'/L'$) represents the height, length, and angle of inclination of the slope. H and L_r represent the elevation and distance of the slope from the nearest river channel, respectively.

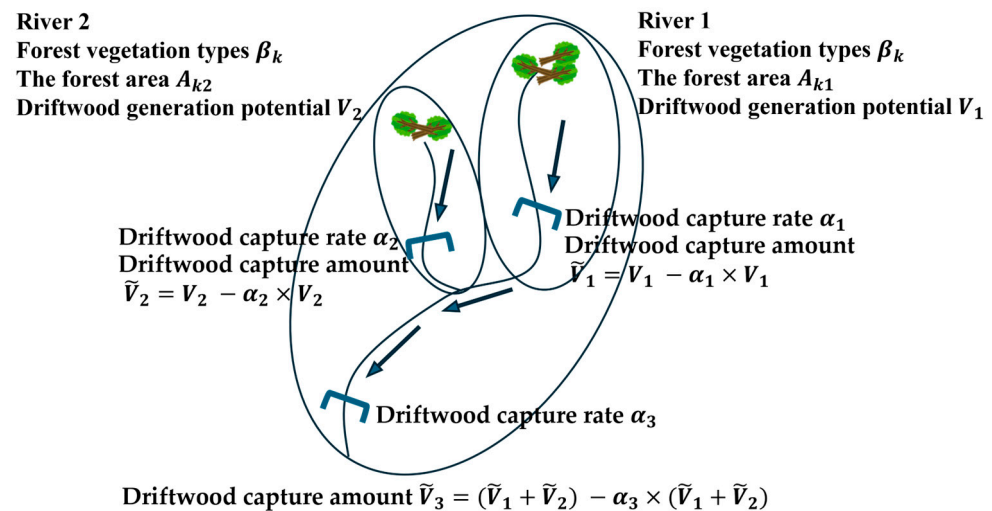


Figure 7. Method for calculating the driftwood capture amount.

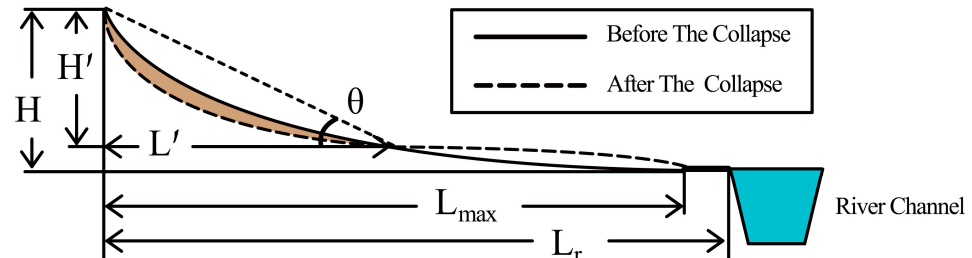


Figure 8. Method for calculating the travel distance of sediment from landslide collapses [4,20].

3.3. Models

3.3.1. CNNs

CNNs, short for convolutional neural networks, are a type of artificial neural network known for their ability to effectively extract feature representations from images. This capability allows CNNs to recognize visual patterns in images without the need for complex, human-designed rules [21].

The structure of a CNN can be broken down into three main layers: the convolutional layer, the pooling layer, and the fully connected layer, as shown in Figure 9. In the convolutional layer, filters are applied to the input image to detect various features. To manage this large number of data, the output is downscaled in the pooling layer, effectively reducing the dimensionality while preserving important features. The fully connected layer is responsible for classification, where the features extracted from the previous layers are used to classify objects. The key operations performed in any CNN can be succinctly described by the following equation [22]:

$$O^l = P(\sigma(O^{l-1} \times W^l + b^l)) \quad (5)$$

where O^{l-1} is the output feature data from the previous layer of the l th layer, W^l and b^l indicate the weights of the layer, respectively, that convolve O^{l-1} by linear convolution, and $\sigma(\cdot)$ denotes the nonlinearity function outside the convolutional layer. These steps are often followed by a pooling operation, which is represented by P in Equation (5) [22].

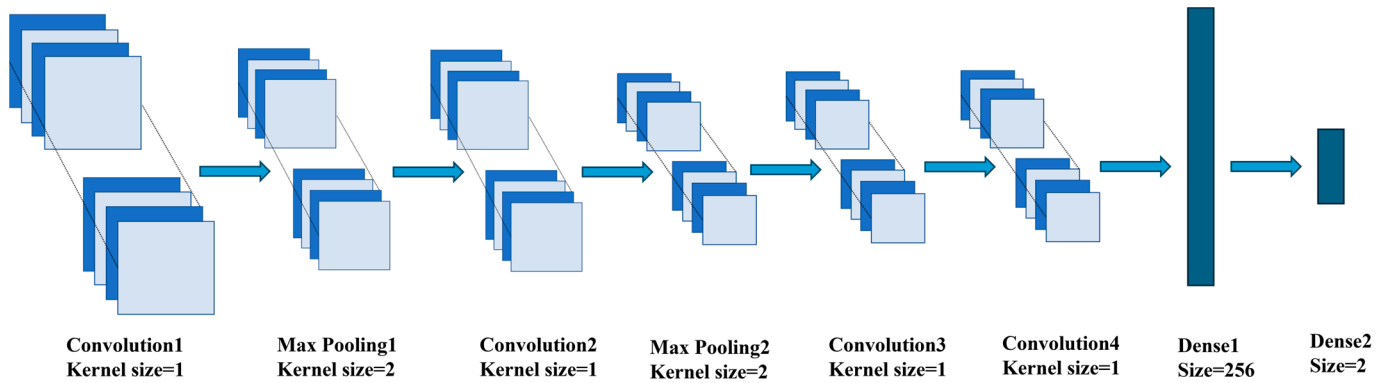


Figure 9. Diagram of CNN model structure.

In this study, 60% of the data are used for training and 40% for testing. We employ the Adam optimizer with a 0.001 learning rate. The network consists of a 512-channel 1×1 convolutional layer, followed by Dropout and Max Pooling layers, and then additional 1×1 convolutional layers reducing from 256 to 32 channels, each with Dropout. The architecture concludes with a Flatten layer, a 256-node Dense layer, and a final 2-node Dense layer with SoftMax for binary classification [23].

3.3.2. Machine Learning

Logistic Regression

Logistic regression is an extremely easy to understand model that is equivalent to $y = f(x)$, showing the relationship between the independent variable x and the dependent variable y . The logistic regression (LR) prediction equation is shown below:

$$P(Y = 1|x) = \frac{1}{1 + e^{-(w^T x + b)}} \quad (6)$$

In other words, the logarithmic probability of output $Y = 1$ is a model represented by a linear function of input x , which is a logistic regression model. When the value of $(w^T x + b)$ is closer to positive infinity, the probability value of $P(Y = 1|x)$ is closer to 1. Therefore, the idea of logistic regression is to first fit the decision boundary (not limited to linear, but also polynomial), and then establish the relationship between the boundary and the probability of classification, to obtain the probability in the case of two classifications.

One of the main advantages of logistic regression is its interpretability, allowing us to clearly understand the contribution of each independent variable to the outcome. The logistic regression model also has a low computational cost, making it ideal for preliminary analysis of large-scale datasets and especially for classification problems that require quick and easily interpretable results.

Decision Tree (DT) and Random Forest (RF)

The decision tree model is a tree structure consisting of nodes that represent features and corresponding decision rules. It involves classifying or regressing instances by recursively dividing data into sub-regions based on features, continuing until a condition is met and forming leaf nodes. Decision trees are particularly effective in modeling nonlinear relationships by creating complex decision boundaries through recursive partitioning, allowing them to adapt to nonlinear data distributions without making strong assumptions about the underlying relationships.

Random forest is an ensemble learning algorithm that integrates multiple decision trees. Each decision tree acts as a classifier, and for a given input, multiple trees produce classification results. The random forest aggregates these votes and assigns the most common category as the final output. By combining multiple decision trees, random forests enhance the stability and accuracy of the model, making them well-suited to high-dimensional data and complex classification tasks. These models are selected for their flexibility, high performance, and ability to effectively capture nonlinear relationships.

KNN

KNN (K-Nearest Neighbor) is one of the simplest machine learning algorithms, which can be used for classification and regression, and is a supervised learning algorithm. It is based on the idea that if most of the most similar K (i.e., most closely neighboring) samples in the feature space belong to a particular class, then that sample also belongs to that class. In other words, this method only determines the category of the sample to be classified according to the category of the nearest one or several samples in the classification decision. The mathematical description of KNN is as follows:

$$X = (x_1, x_2, x_3, \dots, x_n), Y = (y_1, y_2, y_3, \dots, y_n) \\ \left(\sum_{i=1}^n |x_i - y_i|^p \right)^{\frac{1}{p}} \quad (7)$$

The n means feature numbers, and p normally equals 1 (Manhattan distance), 2 (Euclidean distance), and ∞ (Chebyshev distance).

A major advantage of KNN is that it does not require any assumptions and directly classifies samples based on the category of the nearest samples. This makes it highly adaptable in scenarios involving small sample sizes and multi-class problems. The reason for choosing KNN is its straightforward approach, which is well-suited to preliminary data exploration and handling unstructured data that are difficult to model.

MLP

Artificial neural networks are divided into multiple layers and single layers. Each layer contains several neurons. Each neuron is connected by a direct arc with variable weights. The network is trained through repeated learning and training of known information, and through gradual adjustments. The method of changing the weight of the neuron connection can achieve the purpose of processing information and simulating the relationship between input and output.

It does not need to know the exact relationship between input and output and does not need many parameters. It only needs to know the non-constant factors that cause output changes, that is, non-constant parameters. Therefore, compared with traditional data processing methods, neural network technology has obvious advantages in processing fuzzy data, random data, and nonlinear data. It is especially suitable for systems with large scales, complex structures, and unclear information. An ANN with multiple hidden layers is called a multilayer perceptron (MLP). The basic mathematical expression of the neuron is:

$$y = f \left(\sum_{i=1}^n w_n x_n + b \right) \quad (8)$$

Suppose $W = [w_1, w_2, w_3, \dots, w_n]$ is the weight vector, and $X = [x_1, x_2, x_3, \dots, x_n]$ is the input vector.

$$y = f(W^T X + b) \quad (9)$$

The network can learn and train on known information repeatedly, making it highly effective in processing dynamic and complex data relationships. The reason for choosing MLP is its powerful modeling capability, which allows it to perform exceptionally well in complex nonlinear problems.

Machine Learning Strategy

When training with datasets, it is common to split the data into two parts: a training set and a test set. The training set is used to train model parameters, while the test set evaluates the model's performance on unseen data. This separation helps ensure the model's effectiveness.

However, models often perform well using training data but less so using test data. To address this, cross-validation is used. This method involves dividing the initial dataset into K sub-samples, using one for validation and the rest for training. This process is repeated K-times, rotating the validation subset each time. The results are averaged to provide a robust estimate of model performance. K-fold cross-validation, particularly 5-fold and 10-fold, is widely used due to its efficiency in utilizing data for both training and validation, as shown in Figure 10.

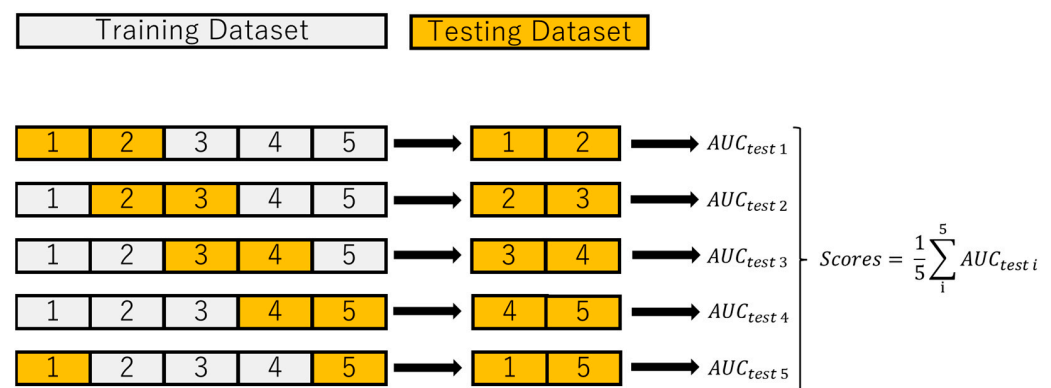


Figure 10. Diagram of cross-validation.

3.3.3. AUC and ROC

A confusion matrix is a table that is often used to evaluate the performance of a classification algorithm. It provides a comprehensive way to assess the performance of a model by breaking down the predicted and actual classes into four categories. The confusion matrix is typically arranged as shown in Table 3.

Table 3. The confusion matrix.

True category		Predicted Category	
		Positive (1)	Negative (0)
		True positive (TP)	False negative (FN)
	Positive (1)		
	Negative (0)	False positive (FP)	True negative (TN)

AUC stands for Area Under the ROC Curve, and it is a metric used to evaluate the performance of a binary classification model. The ROC (Receiver Operating Characteristic) curve is a graphical representation that illustrates the trade-off between a true positive rate and false positive rate at various classification thresholds. A higher AUC indicates better discriminative ability of the model. It measures the model's ability to distinguish between positive and negative instances.

4. Result and Discussion

4.1. Feature Importance

In this study, we employed the information gain ratio (IGR) method to analyze the training dataset. A higher IGR value indicates a greater importance of a causative factor for landslides or bridges within the models. Factors with an IGR of zero are deemed non-contributive and are recommended for exclusion from the models [24,25]. As illustrated in

Figures 11 and 12, the IGR values for all 15 factors associated with landslides were above zero, confirming their relevance in modeling within this study. And 12 factors related to bridges when span = 1, span number and river flow obstruction rate are 0, so they will be removed in the subsequent study. Notably, as shown in Figure 12, the driftwood capture amount plays a crucial role in predicting the hazard levels of bridges, providing a robust basis for this research. Consequently, all the identified causative factors for landslide risk and bridge damage risk were included in the modeling process.

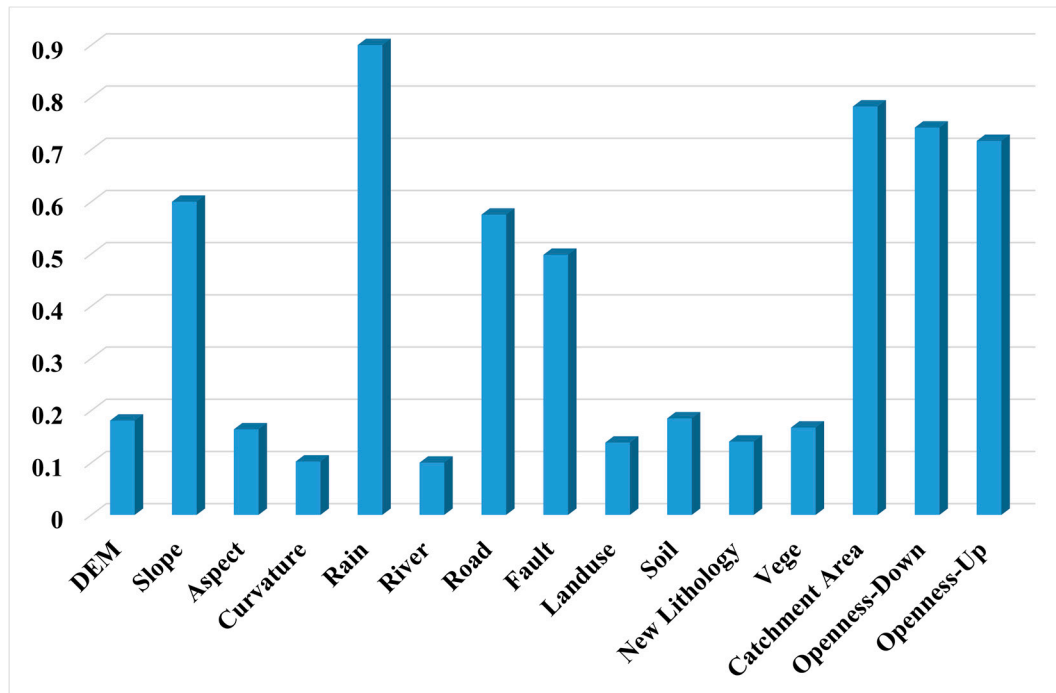


Figure 11. Information gain ratio of the features used for the landslide risk assessment.

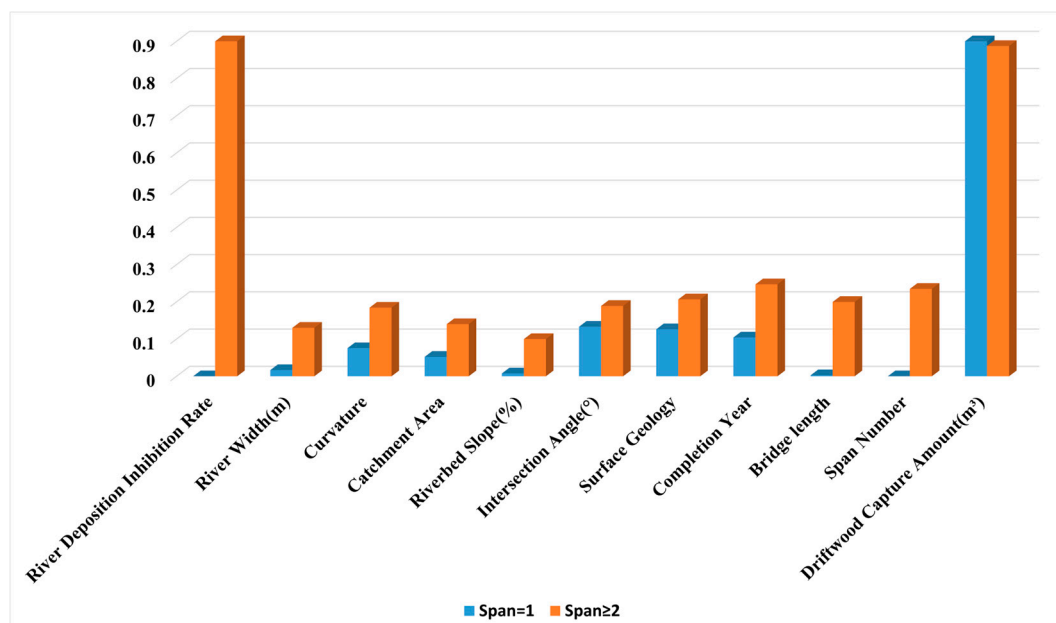


Figure 12. Information gain ratio of the features used for the bridge disaster risk assessment.

4.2. Landslide Risk Assessment Result

Table 4 and Figure 13 depict a comparative analysis of different machine learning models applied for landslide susceptibility mapping (LSM) in the Tokachi region. Each model's performance in classifying the risk levels of landslide occurrence is assessed using Receiver Operating Characteristic (ROC) curves and Area Under the Curve (AUC) metrics, and their predictive power visualized through landslide susceptibility maps.

Table 4. Landslide risk prediction AUC for different models.

Model	DT	KNN	LR	RF	MLP	CNN
AUC	0.832	0.946	0.681	0.889	0.982	0.993

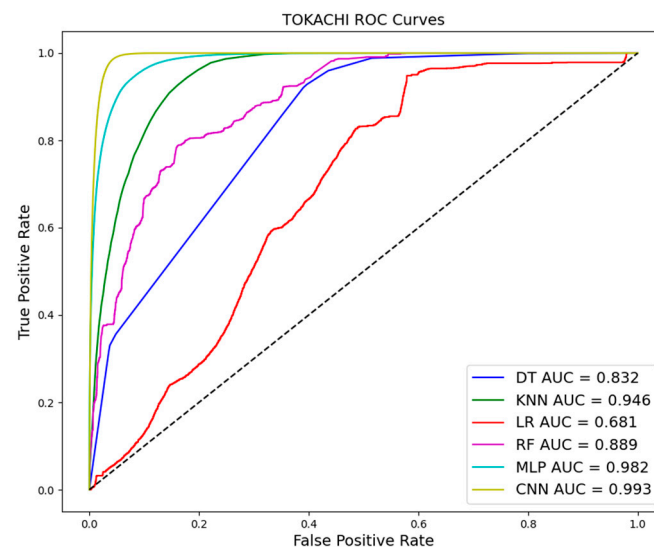


Figure 13. Landslide risk prediction AUC for different models.

Decision trees (DTs) exhibit an AUC of 0.832, demonstrating good performance. K-Nearest Neighbor (KNN) reaches an impressive AUC of 0.946, indicating excellent prediction accuracy. Logistic regression (LR) shows a lower AUC of 0.681, suggesting less effectiveness in this application. Random Forests (RFs) have a high AUC of 0.889, reflecting strong predictive capabilities. Multilayer perceptron (MLP) and convolutional neural networks (CNNs) lead with AUCs of 0.982 and 0.993, respectively, suggesting outstanding performance in predicting landslide susceptibility.

The susceptibility maps generated by the CNN and MLP models are detailed, showing varying risk levels across the region. In the CNN-based map, a broader area in the “very high-” and “high” risk categories is highlighted compared to the MLP model. The CNN seems to identify more potential high-risk zones, which could indicate a higher sensitivity to factors contributing to landslides. The MLP-based map displays a more conservative distribution of “high” and “very high” risk areas. This might suggest that the MLP model prioritizes minimizing false positives or is less sensitive to the variables that influence landslide occurrence.

The CNN model's superior performance in ROC and AUC assessments and its detailed visualizations on the susceptibility map suggest that CNNs are particularly effective for applications involving spatial data. CNNs excel at identifying patterns and spatial relationships, which are essential for analyzing landslide susceptibility. They are especially skilled at capturing contextual information from surrounding areas, providing significant advantages over other machine learning models that rely on pointwise evaluations for tasks like image and pattern recognition. The CNN model effectively manages complex, nonlinear interactions among the variables that contribute to landslides. Meanwhile, the MLP model also performs well, with a high AUC of 0.982, but it distributes risk differently. This

feature could be advantageous in scenarios where minimizing the overprediction of risk areas is crucial, thereby avoiding unnecessary costs from excessive mitigation measures. This nuanced understanding of each model's strengths enables more informed decision-making when selecting the appropriate tool for specific landslide risk assessment situations.

The difference in performance and risk mapping between CNNs and other models like KNN and RF also highlights the importance of choosing the right algorithm based on the specific characteristics of the dataset and the objectives of the analysis. While KNN and RF show robustness and high accuracy, their simplistic or ensemble nature may not capture as detailed spatial patterns as CNNs. The CNN and MLP results are shown in Figure 14. Based on these results, by selecting the CNN's result, the driftwood generation potential is calculated.

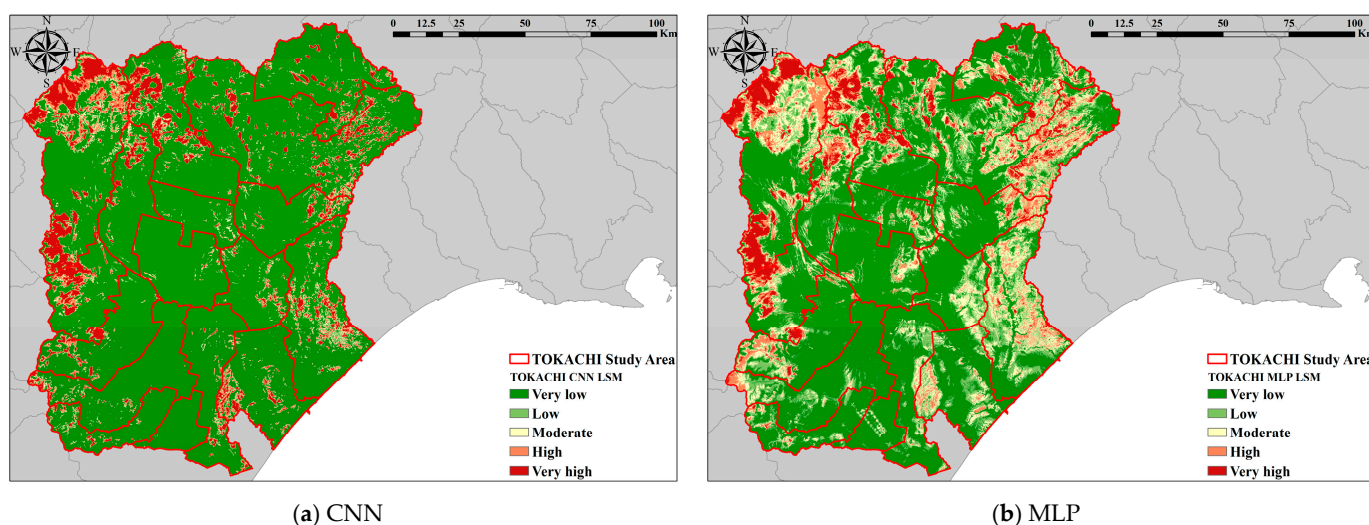


Figure 14. Landslide risk assessment using CNN and MLP.

4.3. Driftwood Generation Potential (DGP) and Driftwood Capture Amount (DCA)

In this study, we estimated the potential driftwood length at the landslide site by calculating the difference between the Digital Surface Model (DSM) and the Digital Elevation Model (DEM). While the DEM provides terrain data excluding artificial structures, the DSM includes the heights of all surface objects, such as vegetation. Therefore, tree heights can be determined by subtracting the DEM from the DSM. For this analysis, DSM data [26] were obtained from the Japan Aerospace Exploration Agency (JAXA). We calculated the driftwood length as half of the observed height difference during descent [4], resulting in an average estimate of 5.24 m in the landslide-prone area, obtained using convolutional neural networks (CNNs).

The regions with very high susceptibility are concentrated in the northern and mountainous areas, characterized by steep slopes and loose soil, which make them highly prone to landslides during heavy rainfall or seismic events (Figure 14). Figure 15a shows that approximately 17% of the areas predicted to be at high risk of landslides are identified as the area having driftwood source. In steep upstream regions, landslides dislodge trees and vegetation, significantly increasing driftwood generation, especially following intense rainfall events. In contrast, areas with low susceptibility experience minimal driftwood production due to the stable terrain and a lower risk of landslides.

High driftwood capture zones are commonly found along major river channels and at confluence points, particularly in downstream areas (Figure 15b). These zones function as natural driftwood traps due to reduced flow velocities or shifts in channel width. The overlap between driftwood generation and capture areas indicates that driftwood from landslide-prone upstream regions is transported downstream, where it accumulates at confluences or areas with slower flow. This buildup can present significant risks to infrastructure and exacerbate flooding. Areas with high landslide susceptibility are also

driftwood sedimentation hotspots, as landslides uproot vegetation, introducing it into river systems as driftwood. Major river channels, confluences, and bends, where flow slows, align with high driftwood capture zones, facilitating accumulation. Effective management in the Tokachi area should focus on stabilizing slopes in landslide-prone areas and monitoring driftwood buildup in rivers. Preventive actions, such as reforestation, slope reinforcement, and installing driftwood barriers, can help mitigate the impacts of landslides and driftwood. Driftwood generated from landslides poses risks to both the environment and human infrastructure.

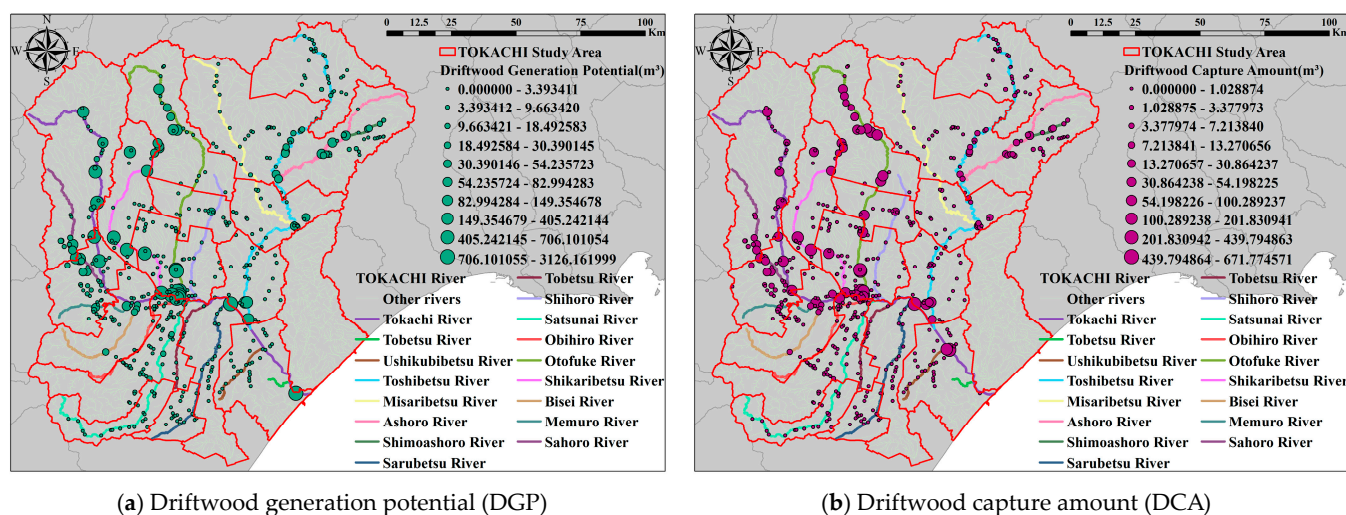


Figure 15. Calculation results of driftwood generation potential (DGP) and driftwood capture amount (DCA) based on CNN.

4.4. Bridge Disaster Risk Assessment

In the field of predictive modeling for bridge disaster prediction, a thorough analysis encompassing both single-span (Span = 1), as shown in Figure 16a and Table 5a, and multi-span bridges (Span ≥ 2), as shown in Figure 16b and Table 5b, reveals distinct outcomes and feature influences (Figure 12) for different bridge structures. The decision tree and random forest models emerge as the top performers in both categories. For Span = 1 bridges, these models achieved high accuracy and strong AUC scores, demonstrating their effectiveness in simpler structural contexts. The decision tree offered a well-balanced trade-off between precision and recall. However, the random forest model's near-perfect scores across both span categories, particularly for multi-span bridges, suggest a potential overfitting issue due to its complexity and the smaller dataset size for Span ≥ 2 , raising concerns about its generalizability.

Additionally, other models such as logistic regression, K-Nearest Neighbor (KNN), artificial neural network (ANN), and multilayer perceptron (MLP) exhibited varying results. Logistic regression struggled with precision in both datasets, achieving moderate accuracy for Span = 1 but a significant drop for Span ≥ 2 , despite a high recall, indicating challenges with class discrimination. KNN consistently encountered issues with precision, while ANN and MLP performed reasonably well for Span = 1 but saw declines in Span ≥ 2 . These results highlight the need for model adjustments or improvements when working with more complex datasets, where simpler models or those with built-in regularization may provide more consistent performance.

The significance of various predictive features, as shown in Figure 12, further highlights the varying impacts of environmental and structural variables on bridge safety. For Span = 1 bridges, "Driftwood Capture Amount (m³)" stands out as a crucial factor, far surpassing other features, emphasizing the importance of driftwood deposition in influencing the stability of single-span bridges. In contrast, for Span ≥ 2 models, both the "River Deposition Inhibition Rate" and "Driftwood Capture Amount (m³)" play a significant

role, indicating that debris accumulation is a major risk for larger, more complex bridge structures. This variation in feature importance suggests that while factors such as “Curvature”, “Intersection Angle”, and “Surface Geology” maintain moderate relevance across both bridge types, the specific dynamics of bridge span and structure require customized modeling approaches.

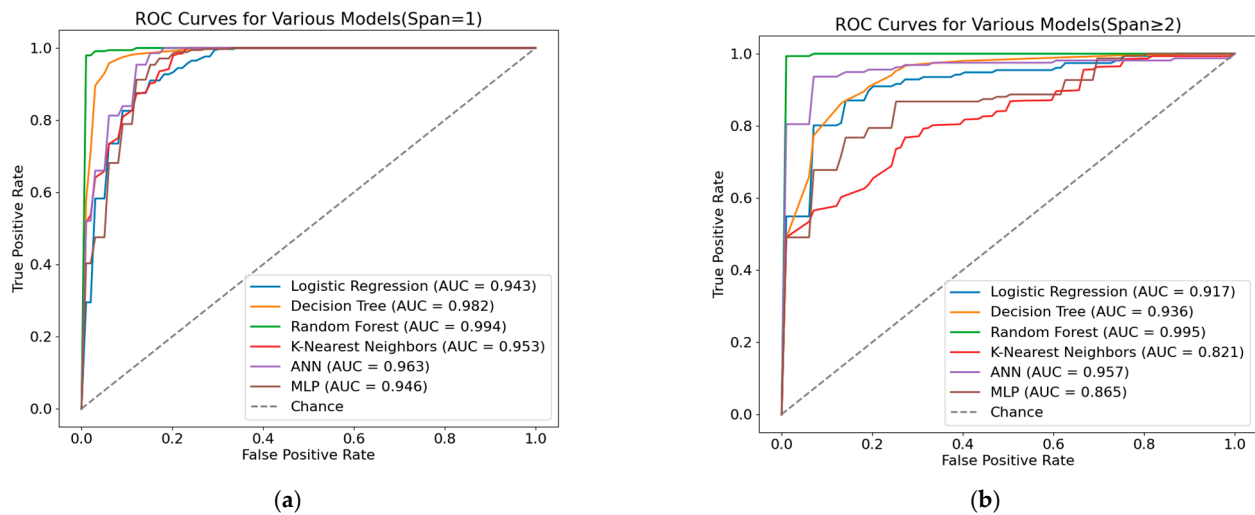


Figure 16. Comparison of AUC for bridge disaster risk prediction in training groups using different models based on various bridge structures. (a) Span = 1, (b) Span ≥ 2.

Table 5. (a) Bridge disaster risk assessment results for a Span = 1 bridge, obtained from each model. (b) Bridge disaster risk assessment results for a Span ≥ 2 bridge, obtained from each model.

(a)										
Span = 1	Model	TP	Confusion Matrix		FN	ACC	P-Score	R-Score	F-Score	AUC
			FP	TN						
	Logistic Regression	6	41	303	1	0.880	0.128	0.857	0.222	0.940
	Decision Tree	5	2	342	2	0.989	0.714	0.714	0.714	0.990
	Random Forest	5	11	333	2	0.963	0.313	0.714	0.435	0.986
	K-Nearest Neighbor	7	47	297	0	0.866	0.130	1.000	0.230	0.952
	ANN	6	20	324	1	0.940	0.231	0.857	0.364	0.968
	MLP	5	24	320	2	0.926	0.172	0.714	0.278	0.934
(b)										
Span ≥ 2	Model	TP	Confusion Matrix		FN	ACC	P-Score	R-Score	F-Score	AUC
			FP	TN						
	Logistic Regression	8	32	123	1	0.799	0.200	0.889	0.327	0.891
	Decision Tree	5	6	149	4	0.939	0.455	0.556	0.500	0.849
	Random Forest	7	1	154	2	0.982	0.875	0.778	0.824	0.994
	K-Nearest Neighbor	8	66	89	1	0.591	0.108	0.889	0.193	0.746
	ANN	8	6	149	1	0.957	0.571	0.889	0.696	0.984
	MLP	8	32	123	1	0.799	0.200	0.889	0.327	0.885

This thorough analysis highlights the importance of considering bridge type, as well as relevant structural and environmental features, when developing predictive models for bridge disasters. Tailoring model strategies to the specific conditions of each bridge type and prioritizing the most influential features can help reduce risks such as overfitting in more complex models, ultimately enhancing the accuracy and reliability of predictions related to bridge safety.

Figure 17 illustrates bridge risk and disaster predictions, categorizing bridges into five risk levels, ranging from very low to very high. The highest-risk bridges are predominantly

located in the southern and southwestern regions of the study area. A clear correlation is observed between the driftwood capture amounts shown in Figure 15b and the bridge risk predictions in Figure 17. Specifically, areas with significant driftwood accumulation align with regions where bridges are assessed as having a higher risk of disaster. This connection is likely due to the impact of driftwood on bridge structures during flood events. Large quantities of driftwood can obstruct water flow, increase hydraulic pressure on bridge piers, and cause direct damage to bridge components, all of which heighten the risk of failure or severe damage, particularly in areas prone to heavy rainfall and flooding.

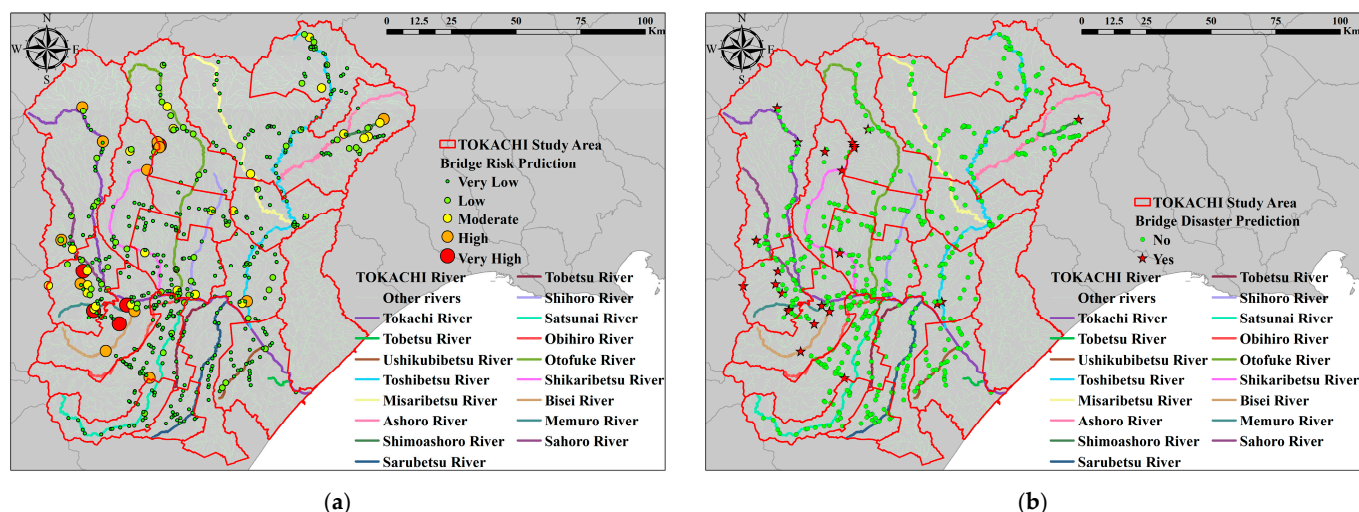


Figure 17. Prediction results for bridge disaster risk. (a) Bridge disaster prediction results, (b) Distribution map of high disaster risk bridges.

The overlap between driftwood-heavy zones and high-risk bridges underscores the need for targeted interventions to mitigate risk and prevent disasters. Strategies such as reinforcing bridge structures, ensuring unobstructed river flow, and establishing proactive inspection programs are essential to reducing the chances of bridge damage and failure. Additionally, installing driftwood capture devices upstream of bridges could alleviate pressure on vulnerable structures by decreasing the volume of driftwood reaching critical areas.

In conclusion, the visual data in these maps highlight the relationship between natural driftwood dynamics and bridge vulnerability. The findings suggest that managing driftwood accumulation and addressing structural risks can significantly enhance regional resilience against flood-related disasters. By understanding the links between driftwood capture, bridge risk levels, and disaster predictions, local authorities can develop targeted, prioritized plans to improve bridge safety and mitigate potential hazards in the Tokachi area.

5. Conclusions

In this study, we have demonstrated the critical application of advanced machine learning techniques, particularly convolutional neural networks (CNNs), in assessing and predicting the risks associated with bridge infrastructure in the Tokachi area of Hokkaido, Japan. Given the increasing prevalence of extreme weather events and their resultant impacts such as landslides and driftwood accumulation, our research highlights the imperative need for sophisticated risk assessment models that can accurately forecast infrastructure vulnerabilities and guide preemptive measures.

Our research applied a comprehensive approach that integrated environmental and structural data with convolutional neural networks (CNNs) and conventional machine learning techniques to predict and evaluate disaster risks for bridges. The study differentiated between single-span and multi-span bridges, finding that while “Driftwood Capture Amount (m^3)” was a key predictor for single-span bridge risk, both “River Deposi-

tion Inhibition Rate" and "Driftwood Capture Amount (m^3)" were critical for multi-span bridges. This distinction is essential for accurate risk assessment and supports more informed decision-making. Additionally, the study quantified the impact of factors like slope inclination, vegetation types, and river dynamics on landslides and driftwood accumulation, enabling detailed risk profiles for bridges that guide targeted intervention and maintenance strategies.

The findings of our study underscore the superior performance of CNNs in spatial data analysis and pattern recognition, which are crucial for detecting areas with high susceptibility to landslides and driftwood blockages. These capabilities allow for a more nuanced understanding of risk distributions, facilitating targeted interventions that are both cost-effective and timely. Moreover, the application of these models has proven to be instrumental in identifying potential hotspots for disaster, thereby enhancing the strategic planning of infrastructure resilience measures.

Furthermore, the integration of machine learning into risk assessment processes not only improves the accuracy of predictions but also provides a dynamic tool for infrastructure management under changing climatic conditions. This adaptability is particularly relevant in the context of global climate change, which is expected to intensify the frequency and severity of extreme weather events.

In conclusion, our research contributes to the ongoing efforts in disaster risk reduction by providing a methodological framework that technology to enhance the bridge against environmental threats. The insights gained from this study advocate for the broader adoption of machine learning techniques in civil engineering practices, particularly for regions prone to natural disasters. Future studies could expand upon this work by incorporating real-time data analytics and exploring the potential of hybrid models that combine the strengths of multiple machine learning approaches to further refine risk assessment and management strategies.

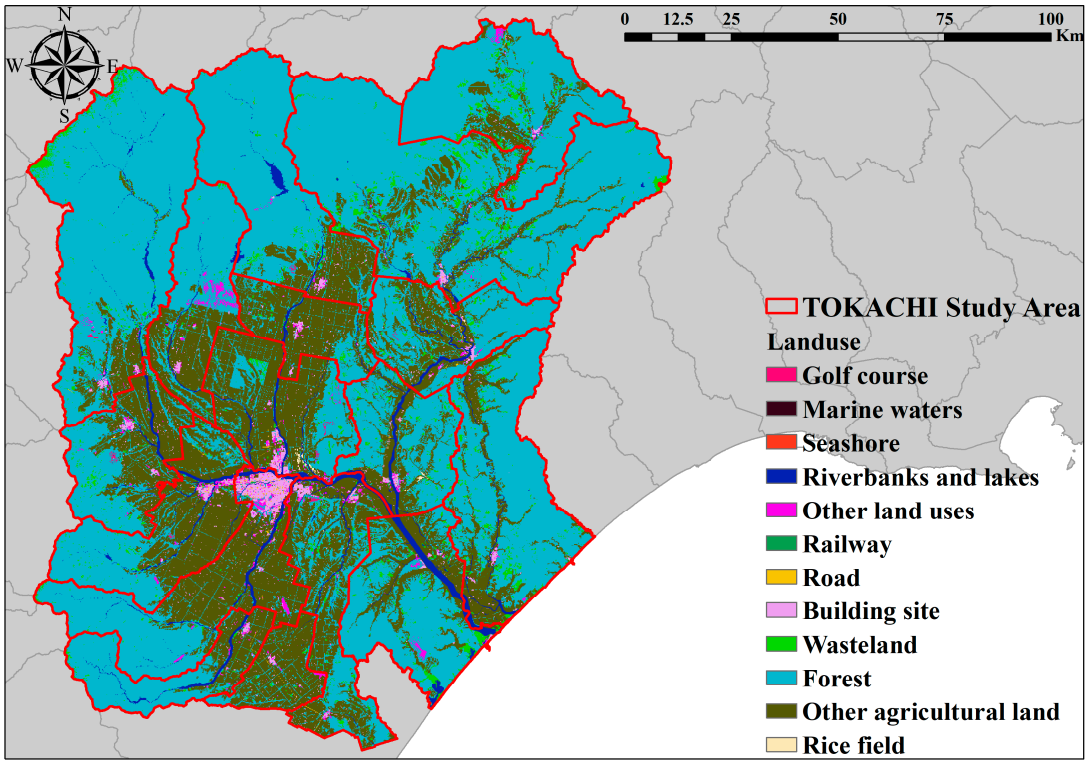
Author Contributions: Conceptualization, F.R., M.A. and K.I.; methodology, F.R. and M.A.; software, F.R., M.A. and K.I.; validation, F.R. and M.A.; formal analysis, F.R. and M.A.; investigation, F.R. and M.A.; resources, F.R., M.A. and K.I.; data curation, F.R.; writing—original draft preparation, F.R.; writing—review and editing, K.I.; visualization, F.R.; supervision, K.I. All authors have read and agreed to the published version of the manuscript.

Funding: This study was partially funded by JST SPRING Grant Number JPMJSP2119, KAKENHI Grant Numbers 23K04019 and 24K00985, as well as Sato Kogyo Co., Ltd. And it was made possible through the provision of bridge inspection reports, general maps, and bridge databases for national and prefectural highway bridges from the Hokkaido Regional Development Bureau of the Ministry of Land, Infrastructure, Transport, and Tourism, as well as Hokkaido Prefecture. Additionally, it incorporates the contributions of several graduates, including Wu Zhenyu, Ota Ryohei, Watanabe Sakurako, Taniguchi Naoki, Orito Kazuki, Kusumoto Yumi, Ohira Natsuo, Tateyama Tomoya, and Li Diyuan. We also received valuable guidance from the Foundation Structure Subcommittee of the Geotechnical Engineering Research Committee in the Civil Engineering Technology Association of Hokkaido. We would like to express our sincere gratitude for their support.

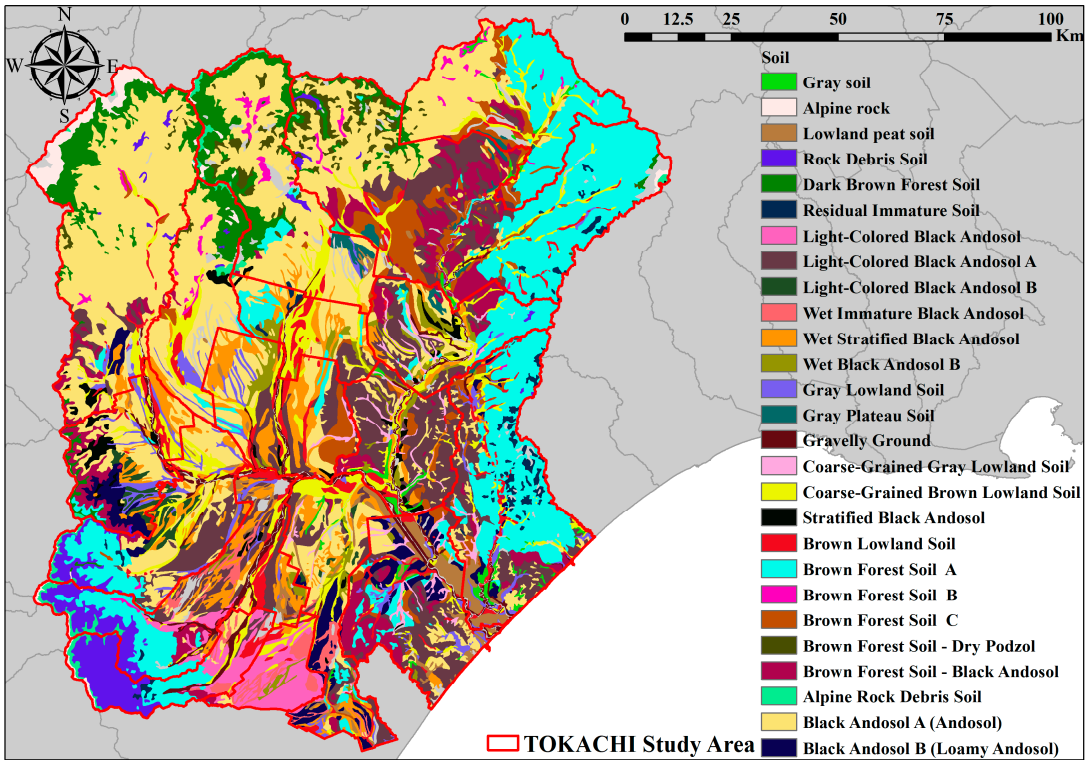
Data Availability Statement: The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding author.

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

Appendix A

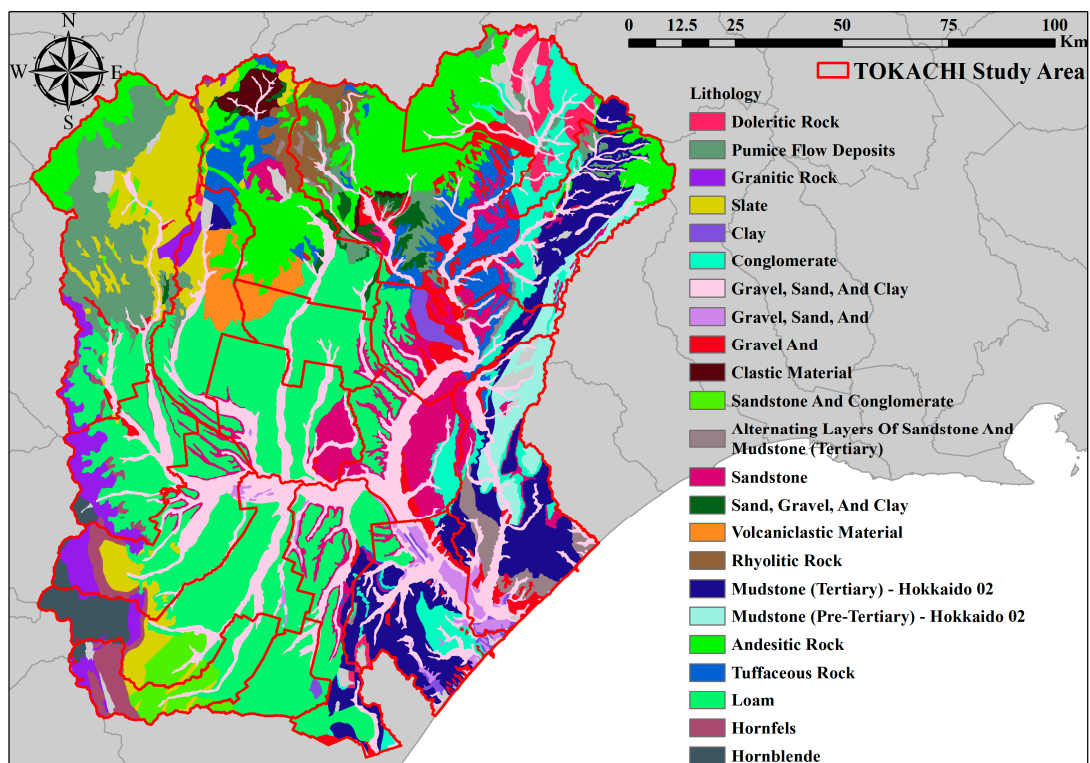


(a)

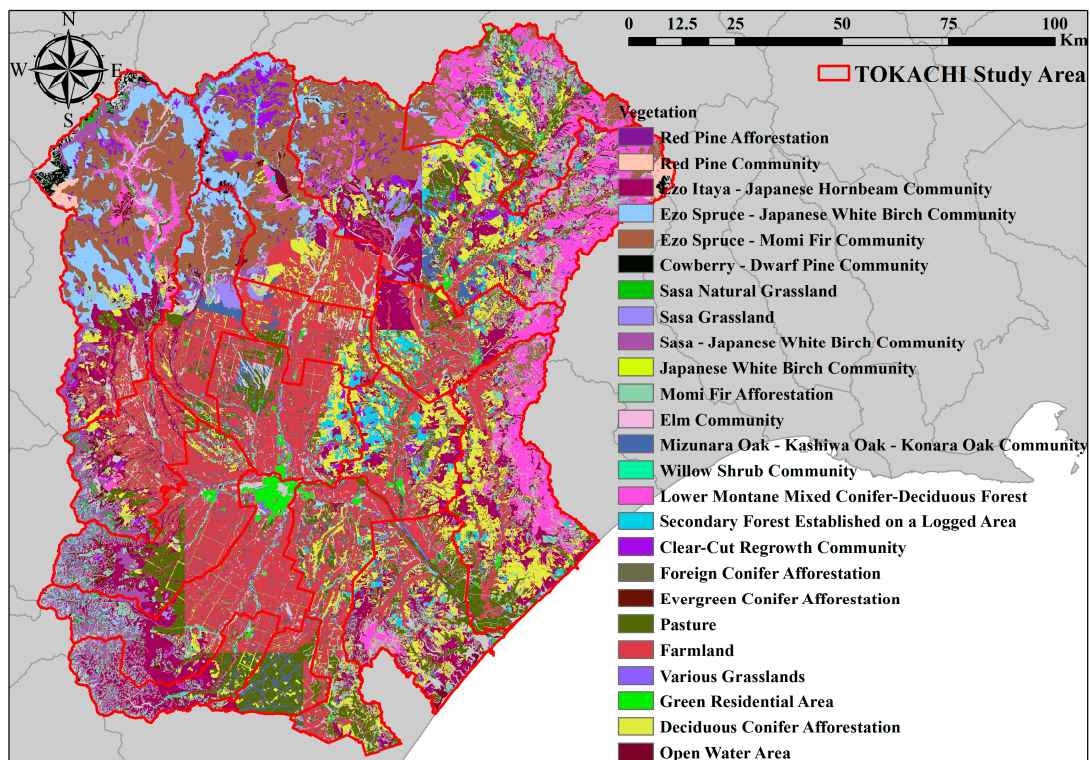


(b)

Figure A1. Cont.



(c)



(d)

Figure A1. Maps showing parts of landslide features: (a) Land use, (b) Soil, (c) Lithology, and (d) Vegetation.

References

1. *Basic Considerations for the Driftwood Management Plan During Sediment and Flood Disasters (Trial Version) [Draft]*; Ministry of Land, Infrastructure, Transport and Tourism: Chiyoda-ku, Japan, 2023.
2. Minami, N.; Doi, Y.; Osanai, N.; Takesaki, S.; Nakayama, Y.; Nonaka, K.; Mashiko, S. Report of woody debris disaster at Yosasa River on August 1998 in Tochigi prefecture, Japan. *Sabo Gakkaishi* **2000**, *53*, 44–51.
3. Yamada, K.; Nagasaka, Y.; Sato, H.; Tsushima, T.; Abe, T. Quantitative study of the area of damage and the amount of lost or trapped timber in a riparian forest in the Appetsu River basin caused by Typhoon 0310 (Etau). *Sabo Gakkaishi* **2006**, *59*, 13–20.
4. Yano, S.; Tsuchihashi, S.; Dozono, S.; Kazama, K.; Kita, T. Assessment of woody debris disaster risk at all bridges along the Kagetsu River according to its generation potential concept. *J. Jpn. Soc. Civ. Eng. Ser. B1* **2016**, *72*, I_289–I_294.
5. Dozono, S. *Development of a Risk Assessment Method for Driftwood Disaster at Bridges Based on the Potential for Driftwood Generation in River Basins*; Kyushu University: Fukuoka, Japan, 2017.
6. Bangnira, F.; Marti-Cardona, B.; Imam, B.; Ruiz-Villanueva, V. Identifying bridges prone to instream wood accumulation: Insights from bridges across the UK. *Nat. Hazards* **2024**, *120*, 25–40. [\[CrossRef\]](#)
7. Yano, S.; Okubo, R.; Tsusue, A.; Takemura, D.; Tomita, K.; Kasama, K.; Nihei, Y. Analysis on cause of driftwoods generation by 2017 northern Kyushu heavy rain. *J. Jpn. Soc. Civ. Eng. Ser. B1* **2018**, *74*, I_1063–I_1068. [\[CrossRef\]](#)
8. Kakihara, N.; Minowa, Y.; Nagashima, K. Model construction for estimating risk of driftwood by topographical factors. *Jpn. For. Soc. Congr.* **2022**, *133*, 525.
9. Tanaka, Y.; Nagashima, K. Woody debris risk assessment based on terrain analysis of slope failure. *Jpn. For. Soc. Congr.* **2021**, *132*, 548.
10. *Tokachi River Basin Sediment Disaster Countermeasure Examination: Considering the Characteristics of the Tokachi River Basin*; Ministry of Land, Infrastructure, Transport and Tourism: Chiyoda-ku, Japan, 2019.
11. *Landslide Topography GIS Data*; National Research Institute for Earth Science and Disaster Resilience: Tsukuba-Shi, Japan, 2016.
12. Ministry of Land, Infrastructure, Transport and Tourism. Download Site for Digital Land Information. Available online: <https://nlftp.mlit.go.jp/ksj/> (accessed on 18 November 2023).
13. Geospatial Information Authority of Japan. 10m Digital Elevation Model. Available online: <https://fgd.gsi.go.jp/download/mapGis.php?tab=dem> (accessed on 24 October 2023).
14. Ministry of the Environment. Natural Environment Survey Web-GIS. Available online: <http://gis.biodic.go.jp/webgis/index.html> (accessed on 29 November 2024).
15. Hokkaido Regional Development Bureau. Hokkaido Bridge and Road Large Structure Data Storage and Management System. Available online: <https://www.hkd.mlit.go.jp/ky/category/u23dsn00000001h3.html> (accessed on 20 April 2024).
16. Cabinet Office, Japan. Damage Caused by Typhoon No. 10 of 2016. Available online: https://www.bousai.go.jp/updates/h28typhoon10/pdf/h28typhoon10_22.pdf (accessed on 15 May 2024).
17. *Report on the Investigation into Geological Disasters Caused by the Heavy Rains in Hokkaido in August 2016*; Japanese Geotechnical Society: Bunkyo-ku, Japan, 2017.
18. Hokkaido Shimizu Town. Report on the Response to the Heavy Rain Disaster Caused by Typhoon No. 10 in 2016. Available online: <https://www.town.shimizu.hokkaido.jp/disaster/files/taihuu10gouoamesaigaihoukokusyo.pdf> (accessed on 12 April 2024).
19. *Guidelines for Preventing Debris Flows and Driftwood*; Forestry Agency: Chiyoda-ku, Japan, 2012.
20. Moriwaki, H. A Prediction of the Runout Distance of a Debris. *Landslides* **1987**, *24*, 10–16_1. [\[CrossRef\]](#) [\[PubMed\]](#)
21. Ding, A.; Zhang, Q.; Zhou, X.; Dai, B. Automatic recognition of landslide based on CNN and texture change detection. In Proceedings of the 2016 31st Youth Academic Annual Conference of Chinese Association of Automation (YAC), Wuhan, China, 11–13 November 2016; IEEE: Piscataway, NJ, USA, 2016.
22. Zhang, C.; Sargent, I.; Pan, X.; Li, H.; Gardiner, A.; Hare, J.; Atkinson, P.M. An object-based convolutional neural network (OCNN) for urban land use classification. *Remote Sens. Environ.* **2018**, *216*, 57–70. [\[CrossRef\]](#)
23. Ren, F.; Isobe, K. Versatility Evaluation of Landslide Risk with Window Sizes and Sampling Techniques Based on Deep Learning. *Appl. Sci.* **2024**, *14*, 10571. [\[CrossRef\]](#)
24. Shirzadi, A.; Solaimani, K.; Roshan, M.H.; Kavian, A.; Chapi, K.; Shahabi, H.; Keesstra, S.; Ahmad, B.B.; Bui, D.T. Uncertainties of prediction accuracy in shallow landslide modeling: Sample size and raster resolution. *Catena* **2019**, *178*, 172–188. [\[CrossRef\]](#)
25. Wang, Y.; Fang, Z.; Hong, H. Comparison of convolutional neural networks for landslide susceptibility mapping in Yanshan County, China. *Sci. Total Environ.* **2019**, *666*, 975–993. [\[CrossRef\]](#) [\[PubMed\]](#)
26. Japan Aerospace Exploration Agency. ALOS Global Digital Surface Model (DSM) “ALOS World 3D–30m (AW3D30)”. Available online: https://www.eorc.jaxa.jp/ALOS/jp/dataset/aw3d30/aw3d30_j.htm (accessed on 10 April 2024).

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.