Supplementary Material for

Impact of environmental and operational stress on defect formation in China’s high-speed rail subgrades

This PDF file includes:

Supplementary Methods

Supplementary Results and Discussion

Supplementary Figure 1 to 7

Supplementary Table 1 to 9

Supplementary References

Supplementary Methods

Data preprocessing:

To prevent errors in the SEM analysis caused by extreme values and outliers, we examined the data distribution of each variable and identified values outside the Interquartile Range (IQR). During this process, no significant outliers were found that would substantially impact the analysis. Additionally, we conducted a comprehensive check for missing values and found none. Therefore, no alterations or removals of data points were necessary.

High-speed rail mileage calculation:

Given that longer mileage increases the likelihood of defects, we calculated the HSR mileage within each district and county. First, we validated the shapefiles (.shp) of the HSR lines against the official Chinese HSR map to ensure data accuracy. Then, in ArcGIS, we overlaid the map of China, divided by district and county boundaries, with the HSR line map. Using the split intersecting lines function, we segmented the HSR lines within each district and county. Afterward, we applied ArcGIS's zonal statistics function to calculate the HSR line lengths within each district and county. After calculating the mileage, we compared the results with pre-processed data to verify the accuracy of the segmentation. Additionally, we cross-referenced the HSR mileage data for each district or county with official records and conducted manual measurements for randomly selected counties to further validate data accuracy and consistency.

Methodology for multi-defect model construction:

We based our selection on Disaster System Theory (DST) hypotheses (H1-H16). The process of variable selection followed a systematic approach:

(1) Initial variable selection: All initial variable subclasses and variables were imported into AMOS to construct a preliminary measurement model. At this stage, we screened the variables based on their statistical significance to the target defects. Specifically, variables with a P-value greater than 0.05 were excluded, as they were not statistically significant at the 95% confidence level. This step was essential to ensure that only statistically significant factors were retained, thereby improving the model's accuracy and explanatory power. The variables excluded during this process included: silty clay, clay loam, loam, sand, effective water storage capacity -50 mm/m, effective water storage capacity-0 mm/m, soil depth-10cm, drainage class -well drained, drainage class -somewhat poorly drained, drainage class -poorly drained, drainage class -very poorly drained, distance to fault, drylands, forested lands, shrublands, vegetable woodlands, other forested lands, canals, lakes, reservoirs, mudflats, beach land, urban land, rural residential, other unused land (alpine desert, tundra), marine, and frost damage. This step was essential to ensure that only statistically significant factors were retained, thereby improving the model's accuracy and explanatory power.

(2) Path fitting and further screening: After the initial screening, the remaining variables were connected according to the hypothesized theoretical paths to construct a preliminary structural model. We excluded paths with a P-value greater than 0.05 to ensure statistical significance. Variables with a significance level greater than 0.1 and lacking significant causal relationships with other variables were also removed. During this phase, we excluded the entire category of geological variables due to their minimal impact on defects and weak correlations with other variables.

(3) Model optimization: After the initial path fitting, we further optimized the model by evaluating its fit indices (e.g., SRMR, RMSEA, GFI, CFI, TLI, NFI, IFI, RFI). Variables that were difficult to integrate into the model during the fitting process, such as wind speed and consecutive 5-day rainfall, were excluded. The final set of variables that were statistically significant and contributed to the model fit are listed in Supplementary Table 1.

This systematic approach ensured that the final model was both statistically robust and highly explanatory, thereby enhancing the validity of the findings.

Methodology for single-defect model construction:

(1) Detection and addressing of multicollinearity: During the construction of the independent defect models, we first used variance inflation factors (VIF) to detect and address multicollinearity, excluding factors with VIF greater than 10 to improve model stability and accuracy.

(2) Initial model building and variable screening: We then built preliminary models by associating each defect with relevant variable subclasses. During this phase, we used AMOS software for an initial screening, excluding variable subclasses with a P-value greater than 0.05.

(3) Categorization and path fitting: Based on Disaster System Theory, candidate variables were categorized into two groups: those that could not establish causal relationships with variables having P-values less than 0.05 and those that could. Variables unable to establish such relationships were excluded. Remaining candidate variables were fitted into path models based on causal relationships, and paths with P-values greater than 0.05 were removed.

(4) Model optimization: Finally, we optimized the model using fit indices such as SRMR, RMSEA, GFI, CFI, TLI, NFI, IFI, and RFI. Variables that were difficult to integrate into the model during this process were excluded. This structured approach ensured the final models were both statistically robust and explanatory.

Structural Equation Modelling fit indicators:

(1) Standardized Root Mean Square Residual (SRMR): SRMR is a standardized measure of the difference between observed data and model predictions[32]. It reflects the gap between the model's predicted mean square residuals and the standardized residuals' root mean square. The smaller the SRMR value, the better the model fits the observed data. Generally, SRMR < 0.1 indicates an acceptable model fit, while SRMR < 0.05 suggests an excellent fit.

(2) Root Mean Square Error of Approximation (RMSEA): RMSEA measures the degree of discrepancy between the model and an ideal model, considering model complexity and adjusting for sample size[33]. Smaller RMSEA values indicate a closer fit to the ideal model. Typically, RMSEA < 0.1 is considered an acceptable fit, and RMSEA < 0.05 indicates a model that fits the data very well.

(3) Goodness-of-Fit Index (GFI): GFI measures the proportion of variance in the observed data that is explained by the model. Higher GFI values indicate better model performance in explaining data variance. GFI > 0.9 is generally considered a good fit, while GFI > 0.95 indicates an excellent model fit.

(4) Adjusted Goodness-of-Fit Index (AGFI): AGFI is an adjusted version of GFI that takes model degrees of freedom into account, adjusting for model complexity. Higher AGFI values suggest better model performance in explaining the data. AGFI > 0.8 is usually seen as acceptable, while AGFI > 0.9 is considered a good fit.
(5) Comparative Fit Index (CFI): CFI is used to compare the fit of the proposed model to an independent baseline model[34]. A CFI value closer to 1 indicates a better fit. Typically, CFI > 0.9 indicates a good fit, and CFI > 0.95 suggests an excellent model fit.
(6) Tucker-Lewis Index (TLI): TLI compares the fit of the proposed model to an independent baseline model while considering model complexity. Higher TLI values indicate better model fit. TLI > 0.9 is considered good, and TLI > 0.95 indicates an excellent model fit.
(7) Normed Fit Index (NFI): NFI compares the fit of the proposed model to a null model. Higher NFI values indicate greater improvement over the null model. Generally, NFI > 0.9 is considered good, while NFI > 0.95 suggests an excellent model fit.
(8) Incremental Fit Index (IFI): IFI measures the improvement in model fit relative to an independent baseline model, accounting for model complexity. Higher IFI values suggest better model fit. IFI > 0.9 is considered good, and IFI > 0.95 indicates an excellent model fit.
(9) Relative Fit Index (RFI): RFI compares the fit of the proposed model to an independent baseline model. Higher RFI values indicate better model fit. Typically, RFI > 0.9 is considered good, and RFI > 0.95 indicates an excellent model fit.

Mediation effect testing steps:
(1) Model setup: Begin by drawing the final multi-defect and single-defect models in AMOS.
(2) Effect output: In the "Analysis Properties" menu, select "Output" and check the "Indirect, direct & total effects" option to display these effects.
(3) Bootstrap settings: Under the "Bootstrap" tab, enable "Perform bootstrap" and set the number of bootstrap samples to 4000, balancing computational time with precision. Check "Bias-corrected confidence intervals" and set the confidence level to 95%.
(4) Custom formulas: Input custom formulas for indirect effects, such as "StdIndA = StdH10×StdH11×StdH6" represents the product of paths H10, H11, and H6. Here, StdH10 represents the path from geomorphological variables to economic variables, StdH11 indicates the path from economic variables to HSR operational characteristics, and StdH6 refers to the path from HSR operational characteristics to subgrade defects.
(5) Mediation analysis: AMOS calculates bias-corrected bootstrap confidence intervals for mediation effects, providing lower (BCLI) and upper (BCUI) bounds. If the interval includes zero, the mediation effect is not significant. If the interval excludes zero, the mediation effect exists.

Bias-corrected and percentile methods:
(1) bias-corrected method: The bias-corrected method in Bootstrap analysis is used to adjust for the asymmetry of confidence intervals caused by skewed sample distributions. This adjustment results in more accurate confidence intervals, especially important in mediation analysis where effect distributions may not be symmetrical[37]. AMOS software automatically applies this method, providing bias-corrected 95% confidence intervals.
(2) percentile method: The percentile method, one of the simplest Bootstrap approaches, calculates confidence intervals based on the distribution of bootstrap samples[38]. It directly determines the upper and lower bounds of effect estimates, making it particularly useful for cases with unclear distributions or small sample sizes. AMOS includes this method, offering Percentile 95% confidence intervals.
Combining the bias-corrected and percentile methods ensures the robustness of mediation effect estimates. The bias-corrected method adjusts for distributional biases, while the percentile method provides a straightforward confidence interval. Together, they offer a comprehensive understanding of the mediation effect's range and stability, ensuring reliable results across different distribution assumptions.

Validation of resampling iterations for mediation effect:
To confirm the validity of using 4000 resamples, we tested 3000, 5000, and 6000 resamples to observe the changes in the lower and upper bounds of the confidence intervals in both the Bias-Corrected and Percentile methods. Based on Supplementary Table 5, resampling with 3000, 4000, 5000, and 6000 iterations showed slight variation in the confidence interval results, verifying the reasonableness of the chosen number of iterations.

Spatial autocorrelation calculation:
HSR subgrade defects often exhibit spatial clustering, meaning that defects in one location are not independent of those in nearby locations. This spatial dependence, known as spatial autocorrelation, can lead to biased estimates in traditional statistical models, including SEM. To address this issue and accurately capture the spatial characteristics of subgrade defects, we performed a spatial autocorrelation analysis using the Spatial Autocorrelation (Moran's I) tool in ArcGIS. The conceptualization of spatial relationships was defined as Inverse Distance, with the distance method set to Euclidean Distance. This configuration ensures that nearby features have a stronger influence on each other, aligning with the spatial nature of defect data. Finally, the analysis generated Moran's I indices, along with corresponding Z-scores and P-values, for each type of defect: settlement, frost damage, uplift deformation, and mud pumping. These indices serve as the basis for determining whether spatial lag analysis should be performed.

Spatial lag value calculation:
To capture spatial dependence in the subgrade defects, we constructed a spatial lag model. The process involved the following steps:
(1) Coordinate system transformation: To ensure the accuracy of boundary distance calculations, the coordinate system of the geospatial data was reprojected into the Albers Equal-Area Projection (suitable for the China region, with units in meters). This projection minimizes distortion in distance measurements across the study area.
(2) Spatial weight matrix construction: To quantify spatial relationships, a spatial weight matrix was constructed. A threshold of 100 meters was set to define spatial neighbors, meaning that two polygons were considered neighbors if the minimum distance between them was less than or equal to this threshold. For neighboring polygons, a binary weight of 1 was assigned, while non-neighboring polygons were assigned a weight of 0. To address variations in the number of neighbors across polygons and ensure comparability among observations, the spatial weight matrix was row-standardized. This approach ensures that the spatial relationships captured in the matrix are both meaningful and

consistent.

(3) Spatial lag model estimation: Each defect type (settlement, frost damage, uplift deformation, and mud pumping) was incorporated into the spatial lag model. Using the Maximum Likelihood Estimation (MLE) method, the spatial lag values were calculated, accounting for the influence of neighboring regions.

(4) Model outputs: The model produced key outputs for each defect type, including the spatial lag coefficient (rho), along with its Z-score and P-value. These outputs were used to evaluate the strength and statistical significance of the spatial lag effect.

Random forest model construction:

(1) Data preparation: To ensure the effectiveness of model training and validation, the dataset was randomly split into a training set and a validation set in a 7:3 ratio. The training set was used to build the model, while the validation set was employed to evaluate the model's generalization performance. The dependent variable of the multi-defect model was obtained by aggregating single defects, including subgrade settlement, uplift deformation, and mud pumping. In addition, our study adopted a class imbalance control sampling design, ensuring a 1:3 ratio between presence and pseudo-absence samples. This process was repeated 30 times using guided resampling techniques, ensuring robust and unbiased results. To eliminate the influence of feature scaling, all input features were standardized using StandardScaler to have a mean of 0 and a standard deviation of 1. This ensured comparability among features during model training.

(2) Hyperparameter optimization: To improve the performance of the RF models, the Optuna framework was employed for hyperparameter optimization. Optuna is a Bayesian optimization-based automated tuning tool that efficiently searches the hyperparameter space and identifies the optimal parameter combinations. The optimized hyperparameters included the number of trees (n_estimators), tree depth (max_depth), node split and leaf sample requirements (min_samples_split and min_samples_leaf), the number of features considered for splits (max_features), and whether bootstrap sampling was used (bootstrap). The optimization objective was to maximize the mean AUC on the training set using 5-fold cross-validation, ensuring the selection of robust parameter configurations. A total of 300 trials were conducted to identify the best-performing hyperparameters, further improving the model's predictive performance.

(3) Model training and evaluation: To evaluate the model's performance and reduce the risk of overfitting, two separate 5-fold cross-validations were conducted to ensure robust optimization and evaluation. The first 5-fold cross-validation, as described in step (2) Hyperparameter optimization, was used to identify the best-performing parameter configurations, with the mean AUC score as the evaluation metric. The second 5-fold cross-validation was employed to assess the generalization ability of the optimized models, where the training set was divided into five subsets, with four subsets used for training and the remaining subset for validation in each iteration. This process was repeated five times to ensure that each subset served as a validation set once. For each fold, the AUC (Area Under the Curve) was calculated, and ROC curves (Receiver Operating Characteristic curves) were plotted to visually assess the model's classification performance. Finally, the Random Forest model was retrained on the entire training set using the optimal hyperparameters to fully utilize all available data. This approach not only optimized the model's configuration but also ensured a comprehensive evaluation of its predictive performance and generalizability.

(4) Feature importance analysis: After training the model, the importance of each feature was evaluated using the feature_importances_ attribute of the RF algorithm, which calculates the Gini importance for each feature. To ensure robustness, the feature importance values were averaged across 30 independently constructed models, resulting in the final ranking of the features. To visualize the results, a lollipop chart was used to display the importance of each feature.

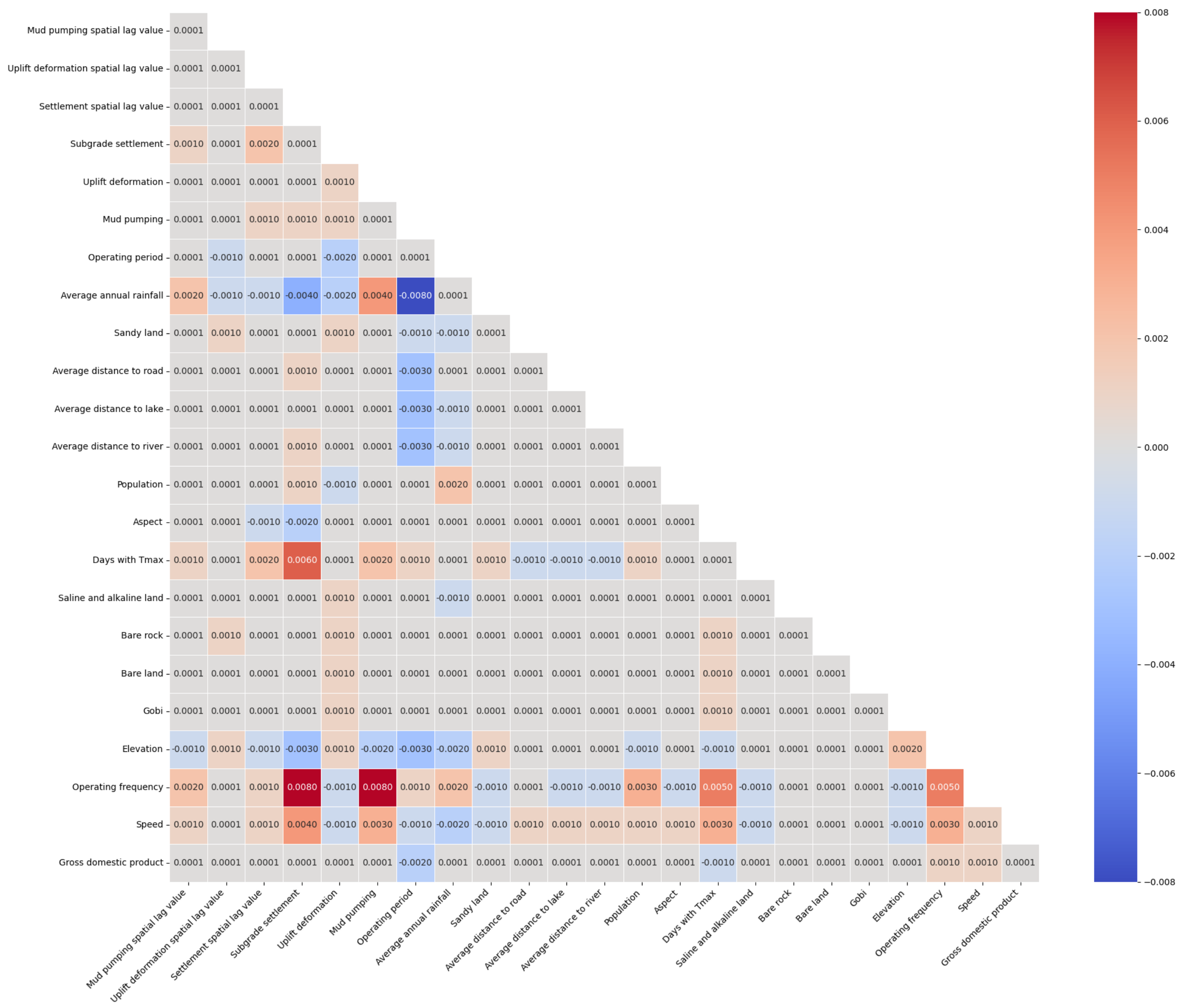
Supplementary Results and Discussion

The complementary role of SEM and RF in defect modeling:
Since RF can capture the complex nonlinearities and interaction effects that SEM fails to model, we introduced RF as a complement to SEM. In the multi-defect model, the importance analysis results of RF (see Supplementary Figure 4) indicate that spatial lag and HSR operational characteristics are the most critical variables driving the co-occurrence of subgrade defects. These results are highly consistent with the findings of SEM, further validating the significant role of spatial lag and operational characteristics in the co-occurrence of defects. However, RF also revealed that climate variables play a more prominent role. In SEM, the effects of climate variables are mainly transmitted to subgrade defects through indirect pathways (e.g., via land use composition), which weakens their direct effects in the total effect calculation. In contrast, RF directly evaluates the overall contribution of variables to the target outcome, highlighting the importance of climate variables. This result indicates that RF effectively supplements SEM in capturing direct effects.

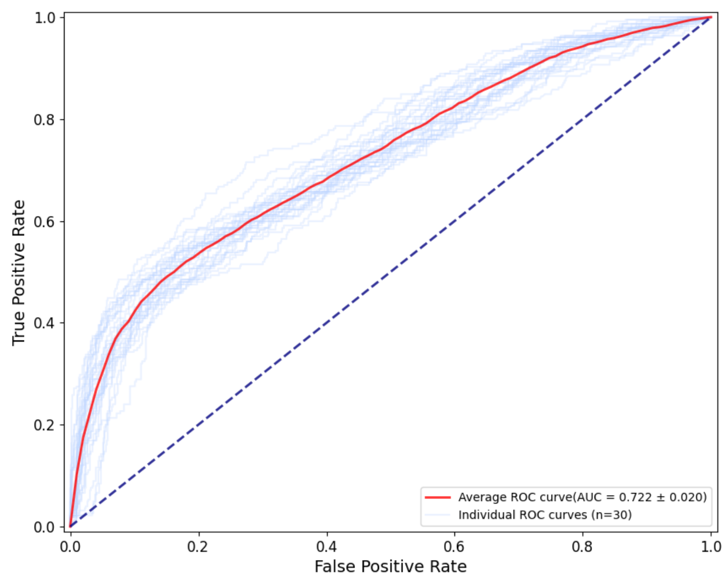
In the single-defect model, the analysis results of RF indicate that spatial lag is the primary driving factor for all types of defects, which is highly consistent with the conclusions of SEM. This consistency suggests that spatial lag has a strong direct effect that can be consistently identified by both methods. Furthermore, RF and SEM exhibit a high degree of agreement in the importance rankings of individual factors. For instance, in the case of settlement defects, both SEM and RF identified construction length as a significant factor. Similarly, for frost damage defects, SEM identified annual freezing days as a key factor, and RF ranked it as the second most important factor after spatial lag. However, for uplift deformation defects, RF assigned a higher importance to Elevate (0.177), while SEM assigned it a lower importance. This discrepancy arises because SEM, through its path decomposition, found that the direct effect of Elevate was weak, while its indirect effects, mediated by operating frequency, exerted a negative impact, thus reducing its total effect value. RF, on the other hand, captures the interactions between elevate and other variables (e.g., spatial lag or climate variables), thereby increasing its importance. Additionally, for mud pumping, RF also highlighted the significant role of operating frequency (0.235), further reflecting the strength of RF in modeling complex variable relationships.

In summary, the results of SEM and RF exhibit a high degree of consistency. Simultaneously, RF addresses SEM's limitations in modeling direct effects by capturing nonlinearities and interaction effects, whereas SEM provides a clear explanation of causal mechanisms through path decomposition. This complementary relationship not only enhances the predictive performance of the models but also deepens the theoretical understanding of defect mechanisms, offering a more comprehensive perspective for the integrated modeling and practical application of subgrade defect analysis.

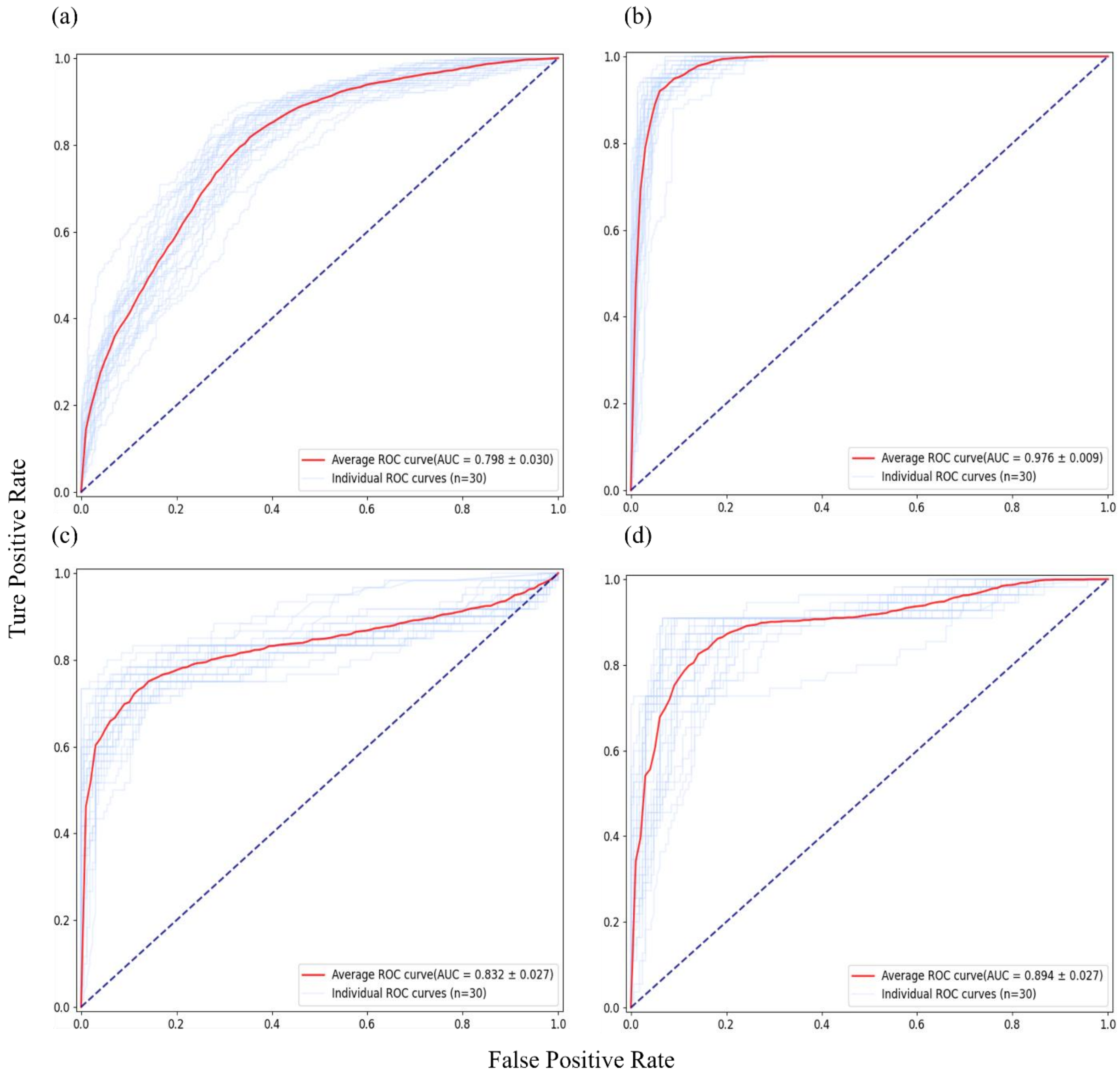
Supplementary Figures



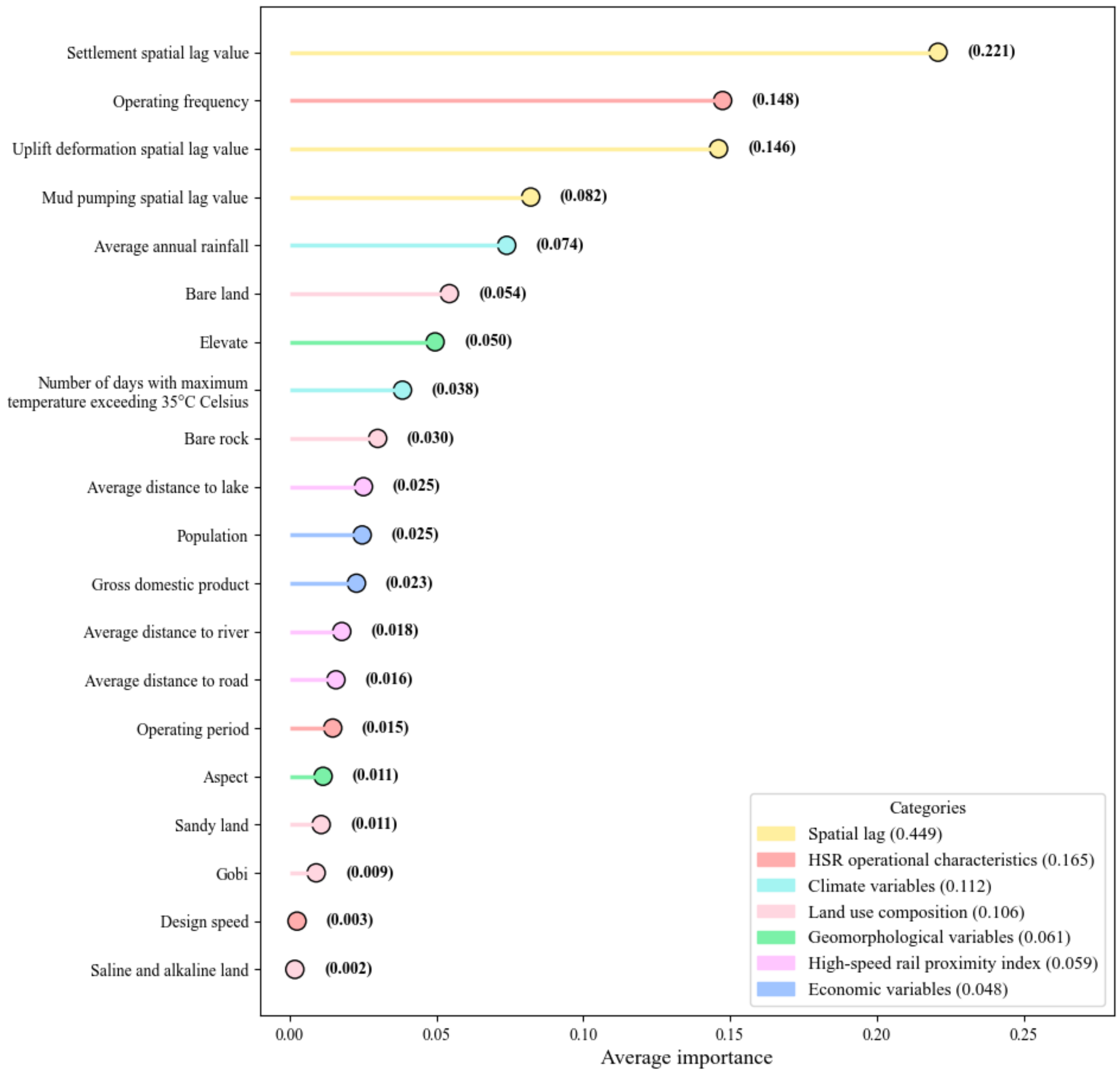
Supplementary Figure 1. Residual covariances plot (multi-defect model)



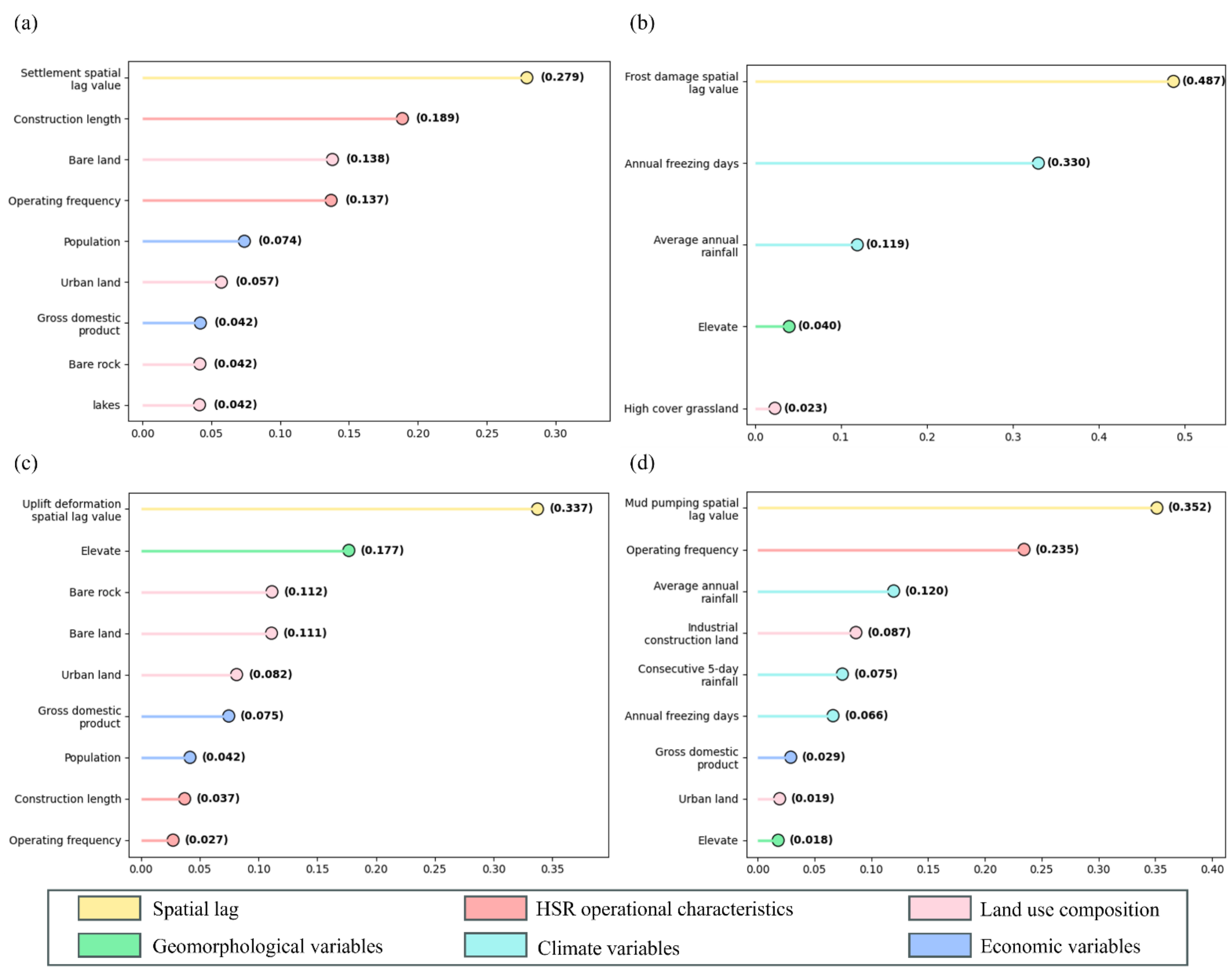
Supplementary Figure 2. Evaluation of the multi-defect model using Random Forest with AUC values and ROC curves. The confidence interval in the legend represents the standard deviation of all AUC values.



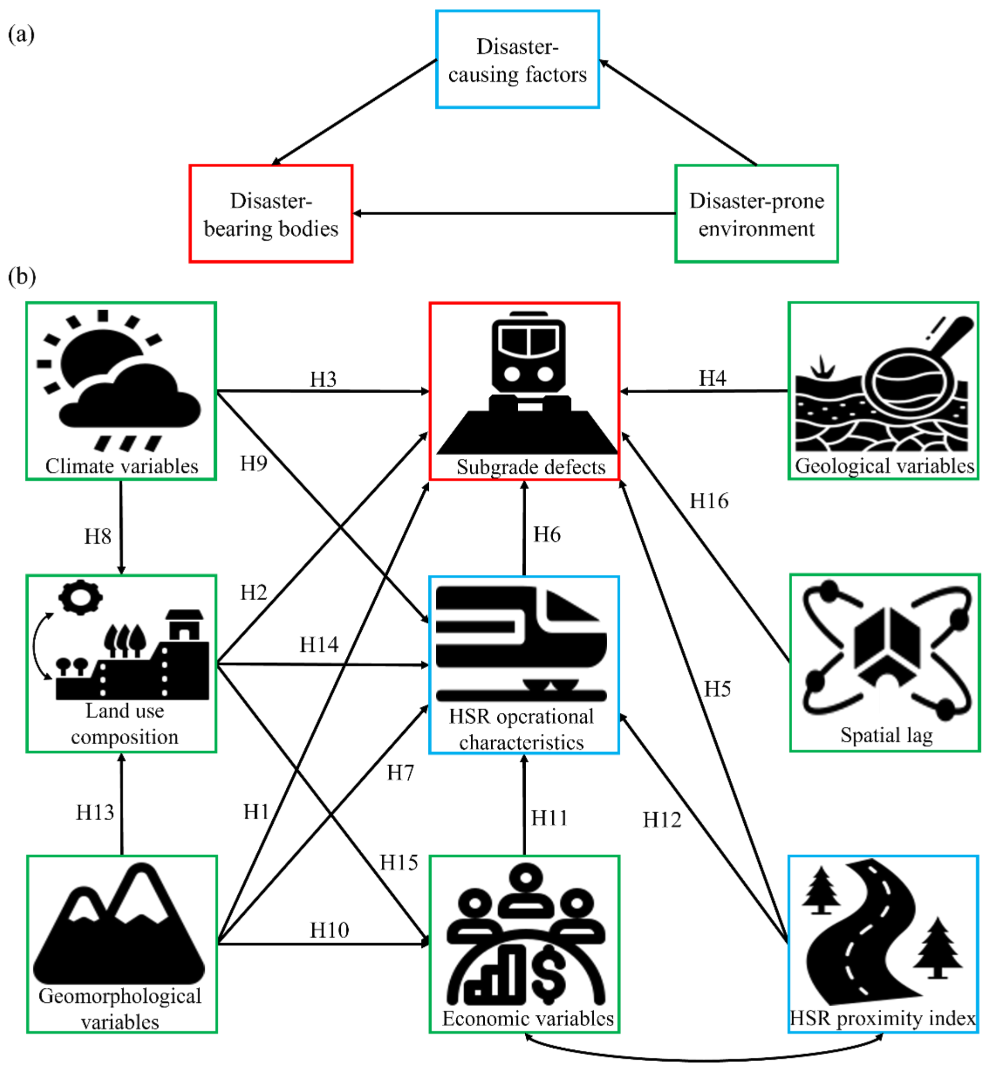
Supplementary Figure 3. Evaluation of the Single-defect model using Random Forest with AUC values and ROC curves: (a) Subgrade settlement, (b) Frost damage, (c) Uplift deformation, and (d) Mud pumping. The confidence intervals in the legends represent the standard deviations of all AUC values for each subfigure. Each "Individual ROC Curves" in the subfigures represents the aggregation of the ROC curves from the 5-fold cross-validation.



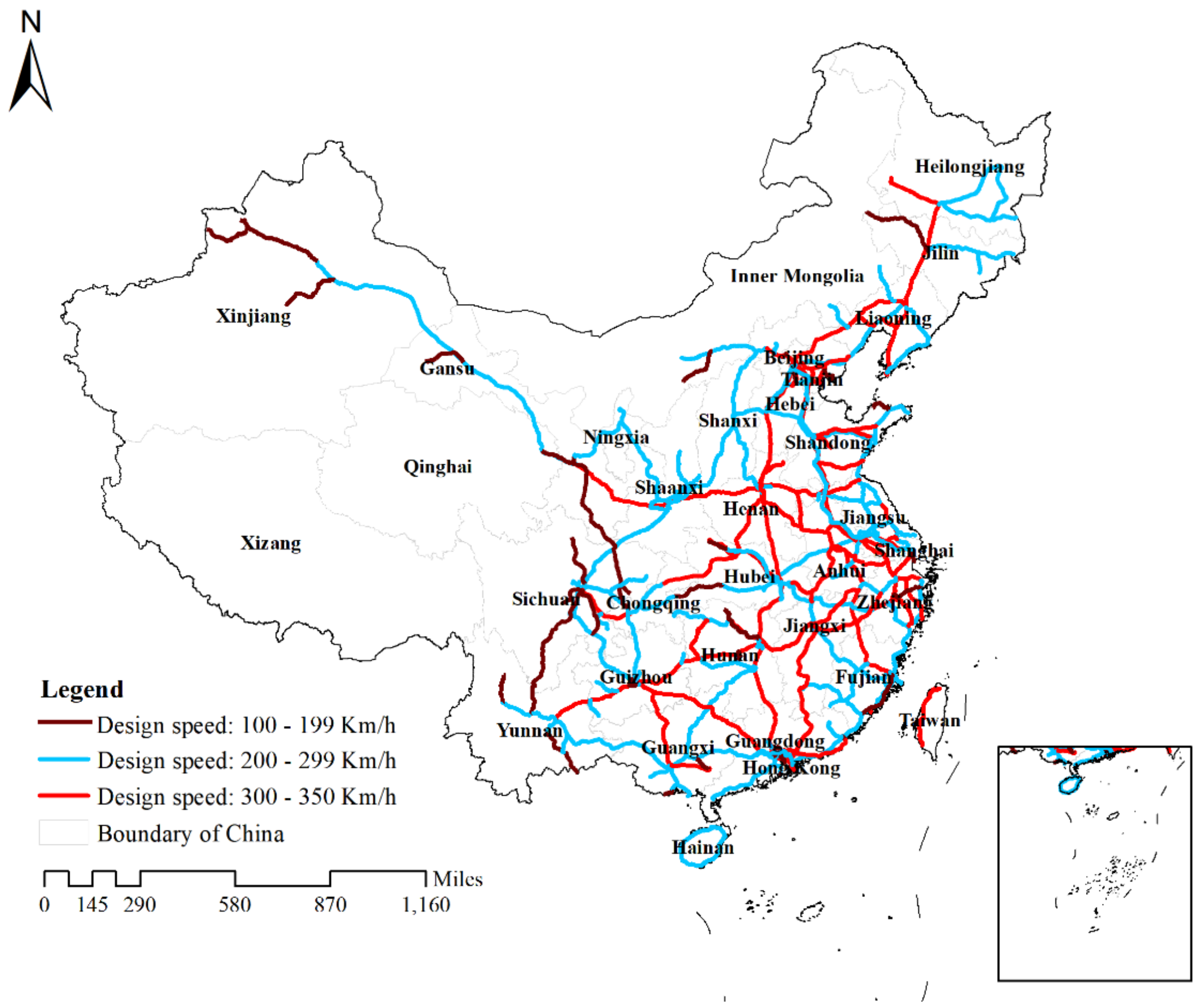
Supplementary Figure 4. Random forest variable importance plot (multi-defect model). The values in parentheses in the "Categories" section in the bottom-right corner represent the total importance of each category.



Supplementary Figure 5. Random forest variable importance plot (Single-defect model): (a) Subgrade settlement, (b) Frost damage, (c) Uplift deformation, and (d) Mud pumping.



Supplementary Figure 6. Hypothesized structural equation model for high-speed rail (HSR) subgrade defects. (a) Disaster System Theory framework. (b) Structural Equation Model based on Disaster System Theory. This model illustrates the proposed relationships between subgrade defects and various predictive factors categorized. The climate variables, land use composition, spatial lag, geomorphological variables, economic variables, and geological variables fall under the Disaster-prone environment, while the HSR proximity index and High-speed rail operational characteristics are classified as Disaster-causing factors. The Subgrade defects represent the outcomes occurring as Disaster-bearing bodies. The hypothesized pathways are labeled H1 through H16, indicating the assumed direction and influence of one variable over another. The unidirectional arrows represent hypothesized causal effects, originating from the variable at the base of the arrow. Note that the presented full (a priori) model in this supplementary figure includes pathways that may not be significant; these are assessed and refined in the main article, where non-significant pathways have been excluded (refer to Figure 1 and the Results section for the final model). Credit: icons, Flaticon.com.



Supplementary Figure 7. Distribution of design speeds for China's HSR (Map boundaries follow China's "Regulations on the Content Representation of Public Maps[44]")

Supplementary Tables

Supplementary Table 1 Path coefficients for the measurement model (multi-defect model)

Path relationship			Path coefficients[a]	95%CI (lower)[b]	95%CI (upper)	Standard error[c]	Critical ratio[d]	P-value[e]
Climate variables	→	Average annual rainfall	-3.624	-5.103	-2.144	0.755	-4.798	***
Climate variables	→	Number of days with maximum temperature exceeding 35 degrees Celsius	1					
Geomorphological variables	→	Aspect	0.129	0.101	0.156	0.014	9.202	***
Geomorphological variables	→	Elevation	1					
HSR operational characteristics	→	Design speed	0.774	0.670	0.878	0.053	14.707	***
HSR operational characteristics	→	Operating frequency	1					
HSR operational characteristics	→	Operating period	0.288	0.219	0.356	0.035	8.322	***
Economic variables	→	Gross domestic product	5.121	3.070	7.171	1.046	4.897	***
Economic variables	→	Population	1					
HSR proximity index	→	Average distance to lake	0.985	0.979	0.991	0.003	285.661	***
HSR proximity index	→	Average distance to road	0.975	0.967	0.982	0.004	236.827	***
HSR proximity index	→	Average distance to river	1					
Land use	→	Gobi	1.499	1.357	1.640	0.072	20.697	***
Land use	→	Bare land	1.660	1.493	1.826	0.085	19.568	***
Land use	→	Bare rock	1.833	1.656	2.009	0.090	20.305	***
Land use	→	Saline and alkaline land	0.431	0.346	0.515	0.043	9.968	***
Land use	→	Sandy land	1					
Spatial lag	→	Settlement spatial lag value	1					
Spatial lag	→	Uplift deformation spatial lag value	2.524	1.977	3.071	0.279	9.059	***
Spatial lag	→	Mud pumping spatial lag value	0.333	0.155	0.511	0.091	3.683	***
Subgrade defects	→	Mud pumping	0.154	0.059	0.248	0.048	3.216	***
Subgrade defects	→	Uplift deformation	1					
Subgrade defects	→	Subgrade settlement	0.724	0.541	0.906	0.093	7.759	***

[a] Path coefficients: These represent the estimated strength and direction of the relationship between two variables in a structural equation model. A positive coefficient indicates a positive relationship, while a negative coefficient indicates a negative relationship. For example, if the path coefficient between train frequency and subgrade defects is positive, it suggests that increasing train frequency directly contributes to a higher occurrence of subgrade defects.

[b] 95% Confidence Interval: The range within which the true path coefficient is expected to lie 95% of the time. The lower and upper bounds define the limits of this interval. A narrower confidence interval indicates the robustness of the model.

[c] Standard Error: Measures the uncertainty of the estimate[35]. Smaller standard errors indicate more reliable estimates.

[d] Critical Ratio: The ratio of the estimate to its standard error, equivalent to the t-value[36], used for significance testing. Generally, a C.R. value greater than 1.96 indicates significance at the 95% confidence level.

[e] P-value: Indicates the significance level of the path coefficient, with $P < 0.05$ generally considered significant. Significant P-values help to determine the main pathways of defect occurrence (*p < .05, **p < .01, ***p < .001).

Supplementary Table 2 Detailed explanation of pathways (multi-defect model)

Pathways	Cause	Effect	Hypothetical relationship	Brief introduction	References	P[a]
H1	Geomorphological variables	Subgrade defects	+	Geomorphological variables, such as aspect and slope, directly impact the stability of the railway subgrade. For example, different slopes can cause varying temperatures within and outside the subgrade, potentially leading to uneven settlement; slopes may also affect moisture flow, thereby disrupting the drainage system of the subgrade.	2、7、8	0.003
H2	Land use composition	Subgrade defects	+	Land use types, such as urban construction land, affect subgrade stability. For instance, groundwater extraction in urban areas can lead to subgrade defects.	14、15	***
H3	Climate variables	Subgrade defects	+	Climate variables, such as rainfall and temperature changes, directly influence the structural performance of the railway subgrade. For example, frequent rainfall can saturate and soften the subgrade soil, triggering settlement or	1、2、3、5、10	0.625

				mud pumping issues.		
H4	Geological variables	Subgrade defects	+	Geological conditions, such as soil composition and rock hardness, directly determine the load-bearing capacity of the subgrade. For example, weak soils are prone to settlement under train loads, while hard rock layers are likely to remain more stable, reducing the occurrence of subgrade defects.	9、10、12	0.863
H5	HSR proximity index	Subgrade defects	+	The proximity of high-speed rail (HSR) to natural or man-made obstacles, such as rivers and roads, influences subgrade stability. For example, the presence of rivers increases the amount of groundwater in the surrounding geological environment, thus affecting subgrade performance.	4、14、15	0.980
H6	HSR operational characteristics	Subgrade defects	+	HSR operational characteristics, such as frequency and speed of operations, directly impact subgrade load and stability. For instance, frequent HSR operations increase subgrade fatigue and wear, leading to settlement and deformation.	2、3、6、12	0.006
H7	Geomorphological variables	HSR operational characteristics	-	Geomorphological variables affect HSR operational characteristics, such as route design and speed limits. For example, complex terrain may require more winding routes, thereby reducing train speeds and impacting operational efficiency.	7、8	0.003
H8	Climate variables	Land use composition	+	Climate conditions influence land use types. For example, regions with abundant rainfall are more suitable for agriculture, while arid areas may be more appropriate for pastureland or barren land.	11、17	0.067
H9	Climate variables	HSR operational characteristics	-	Climate change negatively impacts HSR operational characteristics. For instance, extreme weather events like heavy rain or high temperatures may cause train speeds to reduce or schedules to be altered to ensure safe operation.	6、18、21、28	0.068
H10	Geomorphological variables	Economic variables	-	Geomorphological conditions may limit economic activities. For example, the complexity of mountainous or hilly terrain may increase transportation and infrastructure costs, thereby restricting economic development.	20、27、30	***
H11	Economic variables	HSR operational characteristics	+	The intensity of economic activity influences HSR operational demand. For instance, economically developed areas with high population density typically have greater demand for HSR services, leading to increased operational frequency and speed.	24、25	0.002
H12	HSR proximity index	HSR operational characteristics	-	The proximity of HSR to rivers or other obstacles impacts its operational characteristics. For example, sections of rail near rivers may require higher design standards and maintenance to ensure safe HSR operation.	13、14、15	0.669
H13	Geomorphological variables	Land use composition	+	Geomorphic conditions can limit the way land is used. For example, in areas with higher altitudes, the amount of land that is difficult to utilize will increase.	19、22、29	***
H14	Land use composition	HSR operational characteristics	+	Land use types influence HSR operational demand. For example, urbanized areas with dense populations and frequent economic activity require more HSR services, affecting operational frequency and speed.	14、16	0.445
H15	Land use composition	Economic variables	+	Land use types directly influence regional economic development. For instance, larger urban areas typically have higher populations and GDP.	23、26、31	0.853
H16	Spatial lag	Subgrade defects	+	The spatial lag effect directly influences the occurrence of subgrade defects. For instance, defects occurring at a specific location may gradually spread to adjacent areas through stress transfer, leading to an expansion of the affected area.	39、40、41、42、43	***

[a]The significance of the structural model was evaluated after fitting the paths according to Supplementary Figure 6. Some paths showed low significance (H3 at 0.625, H4 at 0.863, H5 at 0.980, H12 at 0.669, H15 at 0.853, and H9 at 0.068) and were sequentially removed. After removing these paths, the significance of H14 decreased from 0.030 to 0.445, leading to its exclusion as well. The remaining paths all met the significance criterion (H8 changed from 0.067 to 0.021 and was retained). Removing these paths significantly improved the model's fit indices. For example, the NFI increased from 0.820 to 0.843, the RFI from 0.788 to 0.819, the IFI from 0.826 to 0.850, the TLI

from 0.795 to 0.827, and the RMSEA decreased from 0.118 to 0.109. These improvements indicate that the overall model fit was notably enhanced after path optimization. The final retained paths were H1, H2, H6, H7, H8, H10, H11, H13, and H16, all of which met the significance standard (P < 0.05). Please note, this P-value does not represent the final significance of the paths (for the final path significance, please refer to Supplementary Table 3).

Supplementary Table 3 Path coefficients for the structural model (multi-defect model)

Path relationship	Path coefficients	95%CI (lower)	95%CI (upper)	Standard error	Critical ratio	P-value
H10	-0.006	-0.009	-0.002	0.002	-2.640	0.008(**)
H8	0.115	0.054	0.175	0.031	3.735	***
H11	1.209	0.521	1.896	0.351	3.443	***
H7	-0.802	-0.923	-0.680	0.062	-13.024	***
H13	0.015	0.007	0.023	0.004	3.918	***
H1	0.121	0.037	0.205	0.043	2.819	0.005 (**)
H2	0.785	0.401	1.169	0.196	4.004	***
H6	0.116	0.049	0.182	0.034	3.441	***
H16	6.117	4.833	7.401	0.655	9.341	***

Supplementary Table 4 Standardized bootstrap mediation effects test for composite subgrade defects over 4000 iterations (multi-defect model)

Parameter	Effect size	Standard error	Bias-corrected 95%CI			Percentile 95%CI		
			Lower	Upper	P	Lower	Upper	P
H10-H11-H6	-0.001	0.001	-0.004	-0.001	0.001	-0.003	-0.001	0.003
H13-H2	0.017	0.006	0.007	0.031	0	0.006	0.029	0.001
H8-H2	0.031	0.010	0.015	0.056	0	0.013	0.053	0.001
H7-H6	-0.135	0.048	-0.241	-0.046	0.002	-0.241	-0.045	0.002

Supplementary Table 5 Distribution of 95% CI values under different bootstrap sampling times. (multi-defect model)

Parameter	3000/Bias-corrected		3000/ Percentile		5000/Bias-corrected		5000/ Percentile		6000/Bias-corrected		6000/ Percentile	
	Lower	Upper	Lower	Upper	Lower	Upper	Lower	Upper	Lower	Upper	Lower	Upper
H10-H11-H6	-0.004	-0.001	-0.003	-0.001	-0.004	-0.001	-0.004	-0.001	-0.004	-0.001	-0.004	-0.001
H13-H2	0.007	0.030	0.006	0.029	0.007	0.030	0.006	0.029	0.007	0.031	0.006	0.029
H8-H2	0.015	0.056	0.013	0.053	0.015	0.056	0.013	0.053	0.015	0.056	0.013	0.053
H7-H6	-0.240	-0.045	-0.241	-0.046	-0.241	-0.046	-0.241	-0.046	-0.243	-0.046	-0.242	-0.046

Supplementary Table 6 Goodness-of-fit measures (GOFs) for the structural equation models

GOFs *	Adequate level	Recommended level	Multi-defect	Subgrade settlement	Frost damage	Uplift deformation	Mud pumping
Standardised root mean sq. residual (SRMR)	< 0.1	< 0.05	0.082	0.059	0.047	0.035	0.046
Root mean sq. error of approx. (RMSEA)	< 0.1	< 0.05	0.074	0.074	0.090	0.057	0.076
Goodness-of-fit index (GFI)	> 0.9	> 0.95	0.903	0.967	0.989	0.984	0.973
Adj. goodness-of-fit index (AGFI)	> 0.8	> 0.9	0.866	0.941	0.940	0.957	0.935
Comparative fit index (CFI)	> 0.9	> 0.95	0.936	0.942	0.984	0.980	0.957
Tucker–Lewis index (TLI)	> 0.9	> 0.95	0.919	0.916	0.939	0.957	0.916
Normal fit index (NFI)	> 0.9	> 0.95	0.929	0.936	0.982	0.976	0.953
Incremental fit index (IFI)	> 0.9	> 0.95	0.936	0.943	0.984	0.980	0.957
Relative fit index (RFI)	> 0.9	> 0.95	0.910	0.907	0.934	0.948	0.907

*χ^2 test statistics (χ^2 probability level and χ^2/degrees of freedom) are also frequently used GOFs. However, they are recommended not to use when the sample size is large. *(Jöreskog, Karl G., and Dag Sörbom. LISREL 8: Structural equation modeling with the SIMPLIS command language. Scientific Software International, 1993.)*

Supplementary Table 7 Variance inflation factor (VIF) values for variable (Single-defect model)

Type	Variable subclasses	Adequate level	Variance inflation factors
Subgrade settlement	Gross domestic product	< 10	1.17
	Population	< 10	1.94
	Lakes	< 10	1.02
	Urban land	< 10	1.83
	Bare rock	< 10	5.52
	Bare land	< 10	5.39
	Construction length	< 10	1.33
	Operating frequency	< 10	1.16
	Settlement spatial lag value	< 10	1.10
Frost damage	Average annual rainfall	< 10	1.80
	Annual freezing days	< 10	2.82
	Elevation	< 10	1.21

	High cover grassland	< 10	1.14
	Frost damage spatial lag value	< 10	1.86
	Gross domestic product	< 10	1.17
	Population	< 10	1.94
	Elevation	< 10	1.29
Uplift deformation	Bare rock	< 10	5.52
	Urban land	< 10	1.77
	Bare land	< 10	5.51
	Construction length	< 10	1.34
	Operating frequency	< 10	1.26
	Uplift deformation spatial lag value	< 10	1.24
	Gross domestic product	< 10	1.19
	Annual freezing days	< 10	1.68
	Average annual rainfall	< 10	5.73
	Consecutive 5-day rainfall	< 10	5.51
Mud pumping	Elevation	< 10	1.39
	Industrial construction land	< 10	1.06
	Urban land	< 10	1.18
	Operating frequency	< 10	1.42
	Mud pumping spatial lag value	< 10	1.17

Supplementary Table 8 Standardized bootstrap mediation effects test for composite subgrade defects over 4000 iterations (single-defect model)

Type	Parameter	Effect size	Standard error	Bias-corrected 95%CI			Percentile 95%CI		
				Lower	Upper	P	Lower	Upper	P
Subgrade settlement	Population -Urban land – Construction length -Subgrade settlement	0.019	0.004	0.012	0.028	0	0.012	0.027	0.001
	Population-Urban land - Operating frequency -Subgrade settlement	0.022	0.005	0.013	0.033	0	0.013	0.032	0.001
	Gross domestic product - Operating frequency -Subgrade settlement	0.018	0.005	0.010	0.028	0	0.009	0.028	0.001
	Bare rock - Construction length - Subgrade settlement	0.064	0.011	0.044	0.087	0	0.043	0.086	0.001
Frost damage	Elevation - Average annual rainfall - Annual freezing days - Frost damage	0.039	0.007	0.026	0.054	0	0.026	0.054	0.001
	Elevation - Average annual rainfall-Frost damage	-0.017	0.008	-0.033	-0.002	0.024	-0.033	-0.002	0.023
Uplift deformation	Bare rock - Construction length -Uplift deformation	0.037	0.010	0.018	0.057	0.001	0.019	0.057	0.001
	Urban land - Population - Operating frequency -Uplift deformation	0.005	0.002	0.001	0.010	0.044	0.001	0.010	0.047
	Urban land - Population - Construction length -Uplift deformation	0.016	0.005	0.008	0.026	0.001	0.008	0.026	0.001
	Urban land - Gross domestic product - Operating frequency -Uplift deformation	0.001	0.001	0.001	0.002	0.024	0.001	0.002	0.024
	Urban land - Gross domestic product - Construction length -Uplift deformation	-0.002	0.001	-0.004	-0.001	0	-0.004	-0.001	0.001
	Elevation- Operating frequency- Uplift deformation	-0.017	0.009	-0.035	-0.001	0.047	-0.035	-0.001	0.047
Mud pumping	Urban land - Operating frequency- Mud pumping	0.015	0.005	0.006	0.025	0.001	0.005	0.025	0.001
	Urban land - Gross domestic product - Operating frequency- Mud pumping	0.001	0.001	0.001	0.002	0.001	0.001	0.003	0.002
	Elevation- Operating frequency- Mud pumping	-0.026	0.008	-0.042	-0.010	0.001	-0.042	-0.010	0.001
	Elevation - Average annual rainfall - Mud pumping	-0.025	0.012	-0.049	-0.003	0.021	-0.049	-0.003	0.024
	Elevation - Consecutive 5-day rainfall - Mud pumping	0.028	0.015	0.001	0.058	0.047	0.001	0.059	0.048
	Annual freezing days- Average annual rainfall -Mud pumping	-0.063	0.029	-0.121	-0.006	0.027	-0.123	-0.008	0.024
	Annual freezing days- Consecutive 5-day rainfall-Mud pumping	0.043	0.023	0.001	0.091	0.048	0.001	0.092	0.047

Supplementary Table 9 Spatial autocorrelation (Moran's I) and spatial lag analysis results of subgrade defects

Type	Spatial autocorrelation (Moran's I) analysis			Spatial lag analysis		
	Moran's I index	Z-score	P-value	Spatial lag coefficient (ρ)	Z-score	P-value
Subgrade settlement	0.103	12.019	0	0.271	8.897	0
Frost damage	0.229	26.598	0	0.625	30.069	0
Uplift deformation	0.061	7.247	0	0.444	16.778	0
Mud pumping	0.171	20.039	0	0.421	15.588	0

Supplementary References

1.Zhao, J., Liu, K. & Wang, M. Exposure analysis of Chinese railways to multihazards based on datasets from 2000 to 2016. *Geomat. Nat. Haz. Risk*. **11**, 272-287 (2020).
2. Zhang, S., Sheng, D., Zhao, G., Niu, F. & He, Z. Analysis of frost heave mechanisms in a high-speed railway embankment. *Can. Geotech. J*. **53**, 520-529 (2016).
3. Wan, Z., Bian, X. & Chen, Y. Mud pumping in high-speed railway: in-situ soil core test and full-scale model testing. *Railway Eng. Sci.* **30**, 289-303 (2022).
4. Guo, Y., Sun, Q. & Sun, Y. Dynamic evaluation of vehicle-slab track system under differential subgrade settlement in China's high-speed railway. *Soil Dyn. Earthq. Eng*. **164**, (2023).
5. Wang, J., Zhou, Y., Wu, T. & Wu, X. Performance of cement asphalt mortar in ballastless slab track over high-speed railway under extreme climate conditions. *Int. J. Geomech*. **19**, (2019).
6. Wang, Z., Ling, X., Wang, L., Zhao, Y. & Cunha, A. Attenuation law of train-induced vibration response of subgrade in Beijing–Harbin railway. *Shock Vib*. **2021**, 1-8 (2021).
7. Zhang, Y., Sun, B., Wen, A. & Cheng, B. Transverse thermal difference of high-speed railway roadbed in seasonally frozen regions. *Proc. Inst. Civ. Eng-Gr*. **172**, 264-273 (2019).
8. Liu, H., Niu, F., Niu, Y., Xu, J. & Wang, T. Effect of structures and sunny–shady slopes on thermal characteristics of subgrade along the Harbin–Dalian Passenger Dedicated Line in Northeast China. *Cold Reg. Sci. Technol*. **123**, 14-21 (2016).
9. Li, Z., Liu, S., Feng, Y., Liu, K. & Zhang, C. Numerical study on the effect of frost heave prevention with different canal lining structures in seasonally frozen ground regions. *Cold Reg. Sci. Technol*. **85**, 242-249 (2013).
10. Kong, L.-W., Zeng, Z.-X., Bai, W. & Wang, M. Engineering geological properties of weathered swelling mudstones and their effects on the landslides occurrence in the Yanji section of the Jilin-Hunchun high-speed railway. *B. Eng. Geol. Environ*. **77**, 1491-1503 (2017).
11. Wang, S. et al. Application of Bayesian hyperparameter optimized random forest and XGBoost model for landslide susceptibility mapping. *Front. Earth Sc-Switz*. **9**, (2021).
12. Qiu, M.-m., Yang, G.-l., Shen, Q., Yang, X., Wang, G., & Lin, Y.-l. (2017). Dynamic behavior of new cutting subgrade structure of expensive soil under train loads coupling with service environment. *J. Cent. South Univ*. **24**(4), 875-890. doi:10.1007/s11771-017-3490-0
13.Hong, L., Ouyang, M., Peeta, S., He, X. & Yan, Y. Vulnerability assessment and mitigation for the Chinese railway system under floods. *Reliab. Eng. Syst. Safe*. **137**, 58-68 (2015).
14.Zhang, T., Shen, W.-B., Wu, W., Zhang, B. & Pan, Y. Recent surface deformation in the Tianjin area revealed by sentinel-1A data. *Remote Sens*. **11**, (2019).
15.Sundell, J., Haaf, E., Tornborg, J. & Rosén, L. Comprehensive risk assessment of groundwater drawdown induced subsidence. *Stoch. Env. Res. Risk*. A. **33**, 427-449 (2019).
16.Wang, W.-D., Li, J. & Han, Z. Comprehensive assessment of geological hazard safety along railway engineering using a novel method: a case study of the Sichuan-Tibet railway, China. *Geomat. Nat. Haz. Risk*. **11**, 1-21 (2019).
17.Bununu, Y. A., Bello, A., & Ahmed, A. Land cover, land use, climate change and food security. Sustainable Earth Reviews, **6**(1), (2023).
18.Castro, G., Pires, J., Motta, R., Bernucci, L., Marinho, F., & Merheb, A. Unsaturated numerical analysis of a railroad track substructure considering climate data. *Transp. Geotech*. **31**, (2021).
19.Esen, F. Determining the effects of geomorphological factors on the distribution of land use and plant cover by different statistical methods. *Environ. Monit. Assess*. **194**(7), 465, (2022).
20.Gao, D., & Li, S. Spatiotemporal impact of railway network in the Qinghai-Tibet Plateau on accessibility and economic linkages during 1984–2030. *J. Transp. Geogr*. **100**, (2022).
21.Garmabaki, A. H. S., Thaduri, A., Famurewa, S., & Kumar, U. Adapting railway maintenance to climate change. *Sustainability*, **13**(24), (2021).
22.Hieu, N., Huong, H. T. T., Hens, L., Hieu, D. T., Phuong, D. T., & Canh, P. X. Sustainable livelihoods development by utilization of geomorphological resources in the Bai Tu Long Bay, Quang Ninh Province, *Vietnam. Environ. Dev. Sustain*. **20**(6), 2463-2485, (2017).
23.Keola, S., Andersson, M., & Hall, O. Monitoring economic development from space: using nighttime light and land cover data to measure economic growth. *World Dev*. **66**, 322-334, (2015).
24.Li, X., Zhang, M., & Wang, J. The spatio-temporal relationship between land use and population distribution around new intercity railway stations: A case study on the pearl river delta region, China. *J. Transp. Geogr*. **98**, (2022).
25.Niu, F., & Wang, F. Economic spatial structure in China: evidence from railway transport network. *Land*, **11**(1), (2022).
26.Qun, W., Yongle, L., & Siqi, Y. The incentives of China's urban land finance. *Land Use Policy*, **42**, 432-442, (2015).
27.Rito, C., Boretto, G., Bazzano, G., & Cioccale, M. Assessment of geomorphodiversity and the impacts of urban growth in Puerto Madryn, Patagonia, Argentina. *J. S. Am. Earth Sci*. **119**, (2022).
28.Soleimani-Chamkhorami, K. et al. Life cycle cost assessment of railways infrastructure asset under climate change impacts. Transport. *Res. D-Tr. E*. **127**, (2024).
29.Trappe, J., Büdel, C., Meister, J., & Baumhauer, R. Combining geophysical and geomorphological data to reconstruct the development of relief of a medieval castle site in the Spessart low mountain range, Germany. *Earth Surf. Proc. Land*. **47**(1), 228-241, (2021).
30.Wang, H., Gao, J., & Hou, W. Quantitative attribution analysis of soil erosion in different geomorphological types in karst areas: based on the geodetector method. *J. Geogr. Sci*, **29**(2), 271-286, (2019).
31.Yu, J., Zhou, K., & Yang, S. Land use efficiency and influencing factors of urban agglomerations in China. *Land Use Policy*, **88**, (2019).
32. McDonald, R. P., & Ho, M.-H. R. Principles and practice in reporting structural equation analyses. *Psychol. Methods*, **7**(1), 64–82 (2002).
33. Raykov, T., Tomer, A. & Nesselroade, J. R. Reporting structural equation modeling results in psychology and aging: some proposed guidelines. *Psychol. Aging* **6**, 499-503 (1991).
34. Schreiber J. B., Nora, A., Stage, F. K., Barlow, E.A. & King, J. Reporting structural equation modeling and confirmatory factor analysis results: A review. *J. Educ. Res*. **99**, 323-337 (2006).

35. Hair, J.F., Hult, G.T.M., Ringle, C.M., Sarstedt, M., Danks, N.P., Ray, S. An introduction to structural equation modeling. In: partial least squares structural equation modeling (PLS-SEM) using R. Classroom companion: business. Springer, Cham. https://doi.org/10.1007/978-3-030-80519-7_1(2021).
36. Bollen, K. A., & Davies, W. R. Causal indicator models: Identification, estimation, and testing. *Struct. Equ. Modeling*. **16**(3), 498–522 (2009).
37. Streukens, S., Leroi-Werelds, S. Bootstrapping and PLS-SEM: A step-by-step guide to get more out of your bootstrap results. *Eur. Manag. J*. **34**, 618-632 (2016).
38. Cheung, M. W.-L. Comparison of approaches to constructing confidence intervals for mediating effects using structural equation models. *Struct. Equ. Modeling*. **14**, 227-246 (2007).
39. Ge, F.-J., Chen, W.-X., Zeng, Y.-Y., & Li, J.-F. The nexus between urbanization and traffic accessibility in the middle reaches of the Yangtze River urban agglomerations, China. Int. J. Env. Res. Pub. He. **18**, 3828 (2021).
40. Ouni, F., & Belloumi, M. Pattern of road traffic crash hot zones versus probable hot zones in Tunisia: A geospatial analysis. Accident. Anal. Prev. 128, 185-196 (2019).
41. Mou, Z.-H., Jin, C.-C., Wang, H.-B., Chen, Y.-Q., Li, M. Spatial influence of engineering construction on traffic accidents: A case study of Jinan. Accident. Anal. Prev. **177**, 106825 (2022).
42. Liang, L., Zhang, F., Wu, F., Chen, Y.-X., & Qin, K.-Y. Coupling coordination degree spatial analysis and driving factors between socio-economic and eco-environment in northern China. Ecol. Indic. 135, 108555 (2022).
43. Wang, C., Lim, M. K., Zhang, X.-Y., Zhao, L.-F., & Lee, P. T.-W. Railway and road infrastructure in the Belt and Road Initiative countries: Estimating the impact of transport infrastructure on economic growth. Transport Res. A-Pol. **134**, 288-307 (2020).
44. Ministry of Natural Resources of the People's Republic of China. Notice of the Ministry of Natural Resources on Issuing the "Specifications for the Content Representation of Public Maps". The Central People's Government of the People's Republic of China. https://www.gov.cn/gongbao/content/2023/content_5752310.htm (2023).