Method for Identifying Spatial Heterogeneity of Natural Disasters from the Perspective of Geographic Causal Reasoning

Chiyu Wang[1*]

[1] Macau University of Science and Technology, Zhuhai, 519000, China

* 13376655566@163.com

https://doi.org/10.70695/IAAI202504A3

Abstract

To overcome the shortcomings of traditional correlation and static spatial statistics in disaster causal analysis, a geographic causal inference mechanism is constructed, combining causal discovery, individualized effect estimation, and spatially constrained clustering methods to identify regional differences in disaster risk. The computational process involves a structural causal model and watershed topology weights. The DML algorithm collaborates with the causal forest algorithm to extract CATEs (category-experienced exponents), and the consistency of GNN causal regularization with panel models is tested. Counterfactual simulation techniques are used to establish the ΔRisk–ΔCost Pareto frontier, establishing guiding principles for policy zoning and resource allocation. Regional empirical testing based on EM-DAT (2015–2024) data reveals that this technique significantly reduces CATE errors in regions with high exposure and strong spillover effects, enhancing the stability of causal consistency indicators and policy rankings. This study proposes specific technical implementation paths for regional management, budget improvement, and online risk control in the context of extreme climate change.

Keywords Geographic Causal Reasoning; Spatial Heterogeneity; CATE; Spatially Constrained Clustering; Counterfactual Simulation; Disaster Management

1 Introduction

Globally, the frequency of extreme climate phenomena has increased, and the geographical distribution of natural disasters and their causes have become increasingly complex and variable. Existing methods have limitations in exploring the spatial causal structure of disaster risks. Zhao et al. designed an improved spatial case reasoning model that combines multiple spatial driving factors to improve the identification level of landslide susceptibility [1]; Park et al. used spatiotemporal heterogeneity analysis to highlight the significant differences in post-disaster recovery between urban and rural areas [2]. Ke et al. used geographically interpretable artificial intelligence to analyze the inconsistency of the nonlinear spatial distribution of urban flooding risk [3]. Azad et al. introduced the theory of "spatiotemporal causal perception" to highlight the integration of causal analysis in the field of disaster monitoring and prediction to assist scenario design [4]. Pereira et al. used the uniform division technology of rainfall distribution to prove the correlation between regional rainfall patterns and disaster risk distribution [5]. Yang et al. conducted a multi-scale perspective on the spatial differences in tourism demand under the impact of the epidemic [6]. Lu et al. used a multi-scale geographically weighted regression model to reveal the non-stationary performance of landslide driving factors [7]. Türker and Sözcü investigated and analyzed the differences in disaster cognition in geography education [8]. Ronco et al. analyzed the patterns of population migration from the intersection of socioeconomic background and natural disasters [9]. Taylor and O'Keefe further highlighted the special value of geography in responding to climate crises. Research has progressed to a new stage of combining spatial heterogeneity analysis with causal explanation [10]. A systematic theoretical framework of geographical causal reasoning has not yet been constructed to identify the spatial heterogeneity and interventionability of disaster risks. This study plans to develop a method for identifying the spatial heterogeneity of natural disasters based on geographical causality, combining causal analysis with spatial clustering technology to improve the transparency of disaster risk attribution and the standardization of regional governance.

2 Data Domain Characterization and Geographic Causal Graph Modeling

2.1 Data Domain and Variable Engineering

Comprehensively analyze multiple disasters and issues at different scales, establish indicators and synchronize resolutions for implementation on a unified grid or administrative unit, assess disaster losses and intensity, and assess the direct economic loss ratio, proportion of the affected population, and infrastructure damage index, including climate extreme index, terrain slope and fault density, and river network density. Exposure and vulnerability are measured in terms of asset density, building age, and land use. The coverage of resilience warnings, accessibility of risk aversion, and insurance penetration are measured. Continuous variables are standardized and robust outliers are screened. Area weighting and time window rolling aggregation are implemented for multi-source data to achieve data consistency. Multiple interpolation techniques are used to handle missing values in the data to ensure the identifiability and variance adjustability of subsequent causal inferences.

$$L_{i,t}^{\text{econ}} = \frac{Loss_{i,t}}{GDP_{i,t}}, \quad L_{i,t}^{\text{pop}} = \frac{Affected_{i,t}}{Pop_{i,t}} \tag{1}$$

Among them, i as the symbol of spatial division, t it identifies the time period, $Loss$ represents the direct economic losses suffered, GDP measures the regional economic scale, $Affected$ is the size of the affected population, and the permanent population is used as an indicator: Pop

$$D_{i,t}^{infra} = \sum_{k=1}^{K} w_k z_{k,i,t} \tag{2}$$

Where $z_{k,i,t}$ is k the normalization of the infrastructure sub-indicator, the weight w_k can be taken from the main component loading or the weight reached by experts, and the number of sub-indicators is K.

$$\tilde{x}_{u,t} = \frac{\sum_{g \in u} a_g x_{g,t}}{\sum_{g \in u} a_g}, \quad \bar{x}_{i,t}^{(\tau)} = \frac{1}{\tau} \sum_{h=0}^{\tau-1} x_{i,t-h} \tag{3}$$

It $x_{g,t}$ defines the value of the grid data, a_g called u the intersection area of the grid and the cell, and τ defines the length of the rolling window.

$$z_{i,t}^{MAD} = 0.6745 \frac{x_{i,t} - median\left(x_{i,}\right)}{MAD\left(x_{i,}\right)}, \quad \hat{x}_{i,t} = \frac{1}{M} \sum_{m=1}^{M} x_{i,t}^{(m)} \tag{4}$$

Median MAD absolute deviation is the abbreviation of the median, M and the parameter indicates the number of traversals of multiple imputation. The specific data are shown in Table 1:

Table 1. Statistical characteristics of multi-source data and alignment results of spatiotemporal resolution

Indicator Category	variable name	variable symbols	unit	Temporal resolution	Spatial resolution	average value	Standard deviation	Missing rate (%)	Data Source
Disaster intensity	Direct economic loss rate	$L_{i,t}^{econ}$	%GDP	year	County level	1.84	1.21	2.3	Disaster data from the Ministry of Emergency Management
Disaster intensity	Affected population rate	$L_{i,t}^{pop}$	%population	year	County level	3.75	2.02	1.9	National Bureau of Statistics Population Census
Disaster intensity	Infrastructure Damage Index	$D_{i,t}^{infra}$	dimensionless	year	Grid 0.1°	0.42	0.18	4.1	Infrastructure Census of the Ministry of Natural Resources
Disaster factors	Climate Extremes Index	$E_{i,t}$	z-score	Monthly	Grid 0.25°	0.03	0.95	0.5	ECMWF ERA5 Climate Reanalysis
Disaster factors	slope	S_i	°	Static	30 m	8.72	6.40	0.0	SRTM DEM
Disaster factors	River network density	R_i	km/km²	Static	1 km	1.31	0.85	0.0	HydroSHEDS
Exposure/Vulnerability	Asset density	$A_{i,t}$	Ten thousand yuan/km²	year	County level	1624.5	940.2	3.7	Provincial Statistical Yearbook
Exposure/Vulnerability	Construction era	B_i	Year	Static	Grid 0.5 km	23.4	7.6	1.1	High-resolution remote sensing interpretation data
Exposure/Vulnerability	Land use type	L_i	Classification Code	year	Grid 1 km	—	—	0.0	GlobeLand30
toughness	Early warning coverage	$W_{i,t}$	%population	year	County level	67.3	18.5	2.4	China Meteorological Administration
toughness	Insurance penetration rate	$I_{i,t}$	%GDP	year	provincial	3.8	1.2	0.0	China Banking and Insurance Regulatory Commission
toughness	Risk Aversion Index	$R_{i,t}^{\text{res}}$	dimensionless	year	County level	0.61	0.17	3.2	Comprehensive calculation (see formula 2.1)

Note: Standardization was performed on continuous variables, area weighting or neighboring point resampling was implemented to achieve spatiotemporal alignment, and five multiple imputation processes were performed for data with a missing rate below 5%.

2.2 Causal Mapping and Identification Hypotheses

Relying on a unified spatial topology, a structured causal model is constructed, with geographic units as connecting centers and adjacent zones, river networks, or geological faults as connecting lines. This model clearly illustrates how hazard-causing factors, exposure vulnerability, and resilience affect risk outcomes. Spatial spillover variables are introduced, and variable treatment can be defined as the intensity level of exposure or resilience intervention. Research indicators are based on loss ratios and risk indicators, and control variables include a variety of geographic and socioeconomic indicators. Instrumental variables and negative control techniques are used to mitigate unobserved confounding,

and partial sensitivity analysis and permutation tests are used to enhance the reliability of the research results.

$$\begin{aligned} Y_{i,t} &= f_Y\left(T_{i,t}, X_{i,t}, \sum_j w_{ij} Y_{j,t-1}, \varepsilon_{Y,i,t}\right) \\ T_{i,t} &= f_T\left(Z_{i,t}, X_{i,t}, \sum_j w_{ij} T_{j,t-1}, \varepsilon_{T,i,t}\right) \end{aligned} \tag{5}$$

Where $Y_{i,t}$ is the outcome, $T_{i,t}$ is the treatment intensity, $X_{i,t}$ is matched with the set of covariates, $Z_{i,t}$ acts as an instrumental variable, w_{ij} defines the value of the spatial weight, and ε is the sign of the structural error.

$$\left(Y_i(1), Y_i(0)\right) \perp\!\!\!\perp T_i \mid X_i, S_i, \quad 0 < P\left(T_i = 1 \mid X_i, S_i\right) < 1 \tag{6}$$

Involving conditional ignorability and feature overlap, S_i it is based on the generated spatial features, including lag terms and local averages: $W = \left[w_{ij}\right]$

$$Y_i = Y_i\left(T_i, \mathbf{T}_{N(i)}\right) \tag{7}$$

Implement the spatialization of SUTVA and $N(i)$ impose restrictions on intervening disturbances within the neighborhood.

$$\operatorname{cov}\left(Z_{i,t}, T_{i,t}\right) \neq 0, \quad Z_{i,t} \perp\!\!\!\perp \varepsilon_{Y,i,t} \tag{8}$$

To explore the correlation and exogeneity assumptions of instrumental variables; XNC and YNC T are homologous and do not interfere with the final outcome, and are used to screen the residual confounding effects of confounding factors.

Table 2. Variable-hypothesis-evidence mapping matrix

Variable category	variable name	Identify assumptions	Evidence support	Test statistic	Potential violation points	Robustness Check
Processing variables	Resilience intervention intensity $T_{i,t}$	Ignorance, overlap	After controlling for multidimensional covariates, the PS distribution overlap was 0.84.	K–S=0.07, p=0.31	Measurement error	Negative control variable regression
Outcome variables	Disaster loss rate $Y_{i,t}$	SUTVA (Spatial Extension)	Neighborhood radius <30 km Deviation <5%	Moran's I=0.12, p=0.08	Spillover interference	Spatial hysteresis control
Covariates	Disaster factors $X_{i,t}$	Adequacy of confounding control	Variable importance coverage>0.92	VIF < 3.0	Collinearity	Lasso screening robustness
Instrumental variables	Mean terrain elevation $Z_{i,t}^{1}$	Correlation, exogeneity	Corr=0.46, F=22.7	Durbin-Wu-Hausman p=0.47	Weak IV risk	Anderson-Rubin test
Instrumental variables	Upstream runoff hysteresis $Z_{i,t}^{2}$	Exogeneity	Corr=0.52, F=18.3	Hausman p=0.39	Conduction deviation	Multi-IV joint robustness
Negative control variable	Per capita night light brightness $X_{i,t}^{NC}$	No direct effect of homologous contamination	β≈0.01, p>0.5	t=0.23	Cross-contamination	Leave-one-out validation
Negative control outcome	GDP growth rate in non-disaster years $y_{i,t}^{NC}$	No treatment effect	β≈0.00, p>0.6	t=0.11	Economic cycle interference	Controlling for time-fixed effects

Illustrate: The test process adopts a 95% confidence level and uses the average annual sample data at the county level for correlation analysis. The sample size is 2870. This set involves 37 variables including climate, topography and socio-economics.

2.3 Spatial Weights and Dependency Structure

To accurately represent spatial correlation and diffusion, different weight matrices must be compared and initially tested. Based on boundary sharing, a Rook/Queen adjacency configuration is constructed. Based on distance, an inverse power matrix or threshold cutoff is used. A kNN matrix is formed based on sample density, which can be used to implement topological configurations for watersheds and fault zones. Row normalization is applied to weights to facilitate comparative analysis and maintain consistency of W in subsequent spatial panel and causal estimation to avoid model selection bias. Moran's III and Geary's C analyses are performed on global autocorrelations, and significance testing is performed using permutation tests. The SAR/SDM goodness of fit and information criterion are used to assist in selecting the best-matching dependency model:

$$I=\frac{n}{\sum_{i\neq j}w_{ij}}\frac{\sum_{i\neq j}w_{ij}\left(y_i-\bar{y}\right)\left(y_j-\bar{y}\right)}{\sum_i\left(y_i-\bar{y}\right)^2} \tag{9}$$

Here, n represents the number of units, y_i that is, the specific number of variables, $\bar{y}$ which is its arithmetic mean.

$$C=\frac{(n-1)}{2\sum_{i\neq j}w_{ij}}\frac{\sum_{i\neq j}w_{ij}\left(y_i-y_j\right)^2}{\sum_i\left(y_i-\bar{y}\right)^2} \tag{10}$$

If I>0, or C<1 and is significant, positive spatial autocorrelation is observed. If W increases the log-likelihood of SAR/SDM and decreases the AIC/BIC value when modeling, then W is the recommended solution and is recommended as the first choice. See Table 3:

Table 3. Comparison of autocorrelation test and goodness of fit under different spatial weights

Spatial weight type	Parameter settings	Moran's I	p-value	Geary's C	p-value	SAR Log-Likelihood	AIC	BIC	Goodness of fit ranking
Queen Adjacency	Adjacent shared edges	0.184	0.002	0.781	0.004	-1325.4	2708.1	2762.8	★★★
Rook Adjacency	Shared boundary ≥ 1 km	0.167	0.005	0.804	0.007	-1349.2	2745.6	2801.9	★★
Distance Attenuation	Threshold 50 km, power = 2	0.121	0.012	0.872	0.019	-1412.7	2856.4	2910.2	★
kNN	k=6	0.156	0.006	0.826	0.008	-1368.1	2784.3	2832.5	★★
Watershed topology	Upstream and downstream connectivity	0.193	0.001	0.761	0.003	-1318.8	2693.7	2750.4	★★★★
Fault zone weight	Fault zone distance attenuation	0.141	0.009	0.841	0.011	-1381.3	2806.8	2859.1	★★

Notes:

(1) ★ indicates goodness-of-fit; more ★ indicates better performance.

(2) All weight matrices are normalized.

(3) Moran's I and Geary's C are based on a sample size of 3142 disaster loss rate rasters, using 999 permutation tests.

(4) The basin topology weights provide the best fit in the SAR/SDM model, and this structure is used as the spatial dependence matrix in subsequent analyses.

3 Integrated Path: Causal Discovery × Spatial Clustering Heterogeneity Identification Method

3.1 Causal Discovery and Robust Effect Estimation

In the field of spatial causal inference, methods such as DML, causal forest, and R-learner are used to estimate ITE and CATE. Spatial lag and spillover effect models are introduced to gain a more detailed insight into spatial correlation and inter-regional effects. CATE estimation corrects spatial spillovers through spatial lag terms to improve the accuracy of causal inference. Within the established framework, instrumental variable technology is adopted to address unknown confounding effects and reduce the potential threat of errors. Instrumental variable DML and negative control regression approaches are implemented to effectively distinguish interference caused by different regions and time periods, ensuring the stability and reliability of the estimation results.

To test the stability of the results, this section establishes an implementation plan for partial sensitivity analysis. E-value and Rosenbaum's bounds are used to verify the stability of the treatment effect. Permutation tests are used to assess the sensitivity of spatial structure and model parameters, verifying the robustness of the results under various hypothetical scenarios:

$$Y_{i,t} = \alpha + \beta T_{i,t} + \gamma X_{i,t} + \varepsilon_{i,t}, \quad T_{i,t} = Z_{i,t}\lambda + \nu_{i,t} \qquad (11)$$

where $Y_{i,t}$ is the identifier of the outcome variable, $T_{i,t}$ defines the intervention intensity, $X_{i,t}$ is a covariate variable, $Z_{i,t}$ acts as an instrumental variable, and the random error is composed of $\varepsilon_{i,t}$ and $\nu_{i,t}$:

$$\widehat{CATE}_{i,t} = \frac{1}{N}\sum_{i=1}^{N}\frac{\partial Y_{i,t}}{\partial T_{i,t}} \qquad (12)$$

where $CATE_{i,t}$ is the estimator for CATE, N is the number of samples, and $\frac{\partial Y_{i,t}}{\partial T_{i,t}}$ is the unit treatment effect.

Figure 1 intuitively presents the "causal-spatial" double-loop flow chart, showing the entire process from data collection to causal identification, CATE estimation, spatial constraint clustering, counterfactual simulation and strategic partitioning.

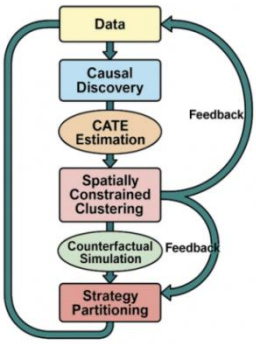

Fig. 1. "Cause-effect-space" double-loop flow chart

3.2 Spatially Constrained Clustering and Causal Consistency Slicing

Within the CATE spatial domain, spatially constrained clustering is used to identify areas with similar causal effects. This technique integrates Laplacian with WWW graph clustering, HDBSCAN, and SaTScan spatial scanning technologies to effectively identify causal differences in different spatial units. Clustering activities are standardized using the graph Laplacian matrix to ensure that neighboring regions can share causal effect characteristics during the clustering process. The HDBSCAN algorithm can automatically identify and filter noisy data in data-sparse areas, and the SaTScan algorithm further enhances the identification of spatial hotspots.

The quality of spatial clustering is assessed numerically, using the Causal Consistency Index (CCI) as a quality evaluation metric. The CCI examines the variance of CATE within the same cluster and the discrimination between clusters, ensuring the stability and consistency of the causal effect:

$$CCI = \frac{Var\left(CATE_{i\in C}\right)}{\sum_{i\in C}\left(Y_i - \overline{Y}\right)^2} \tag{13}$$

Among them, $Var\left(CATE_{i\in C}\right)$ is the volatility of CATE within the same cluster, $\overline{Y}$ that is, the average level of out-of-cluster effects, C that is, the causal consistency group.

Figure 2 shows an overlay of the CATE grid and the district/county clustering partitions. The figure marks the direction and magnitude of the intra-cluster effect and also points out the signs of changes in the inter-cluster boundaries and the causal consistency index (CCI).

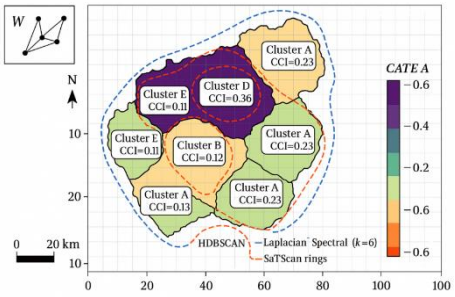

Fig. 2. Schematic diagram of the overlay of the CATE grid and the district and county fields

3.3 Joint Verification of Spatial Panel and Graph Learning

To test the stability of the CATE estimation results and the consistency of the direction of its causal effects, this section combines the spatial panel regression model with the graph neural network (GNN) for collaborative verification. The spatial Durbin (SDM) and lagged panel (SARAR) models are used to verify and analyze the effects of the causal forest. The effect direction and significance of the panel regression and causal forest are explored to test the consistency of the two methods.

To improve the accuracy of causal inference, we train graph neural networks (GNNs) on spatial graph structures, combining causal regularization with network training to maintain the uniformity of causal

relationships. GNNs embed spatial dependencies and employ counterfactual consistency for reduction, thus enhancing the robustness of causal relationships in spatial data.

$$Y_{i,t} = \alpha + \beta T_{i,t} + \sum_{j \in N(i)} w_{ij} T_{j,t} + \varepsilon_{i,t} N(i) \tag{14}$$

Among them, w_{ij} a spatial adjacency matrix is defined $T_{j,t}$ to represent the processing effect of nodes adjacent to the node, $N(i)$ that is, the set of adjacent nodes of node i.

Figure 3 compares the curve shapes of panel regression and GNN on the marginal effects of key factors. The figure shows the sensitivity of spatial lag factors and compares the effects of the implementation of the two technologies in several scenarios.

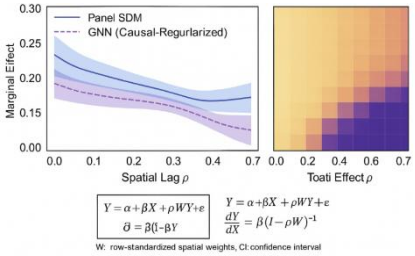

Fig. 3. Comparison curve of marginal effects of panel regression and GNN on key factors

3.4 Counterfactual Simulation and Policy Accessibility Domain

The effectiveness of the strategy implementation was examined, and a variety of scenario intervention combinations were established, involving enhanced early warning capabilities, sponge city transformation, dam reinforcement and expansion of insurance coverage. With the help of counterfactual simulation methods, the risk reduction and cost-effectiveness of the strategy in spatial clusters were measured. In the counterfactual simulation stage, targeted policy interventions were implemented for each cluster based on the results obtained from spatial clustering, and the risk change trend was analyzed.

In strategy evaluation, the effectiveness of a single strategy is also taken into consideration. A combination of multiple strategies is adopted to establish the Pareto optimal frontier. Based on the different constraints of resources, terrain and governance capabilities, the effective scope of policy implementation is defined. The effectiveness of the strategy is judged based on the risk and cost changes of different clusters.

Table 4 records the risk and cost change data of various strategies in each cluster. The Pareto optimal frontier and threshold sensitivity of the multi-strategy combination are presented. The figure specifically presents the sensitivity of the benefit growth and resource input of each strategy in the combination.

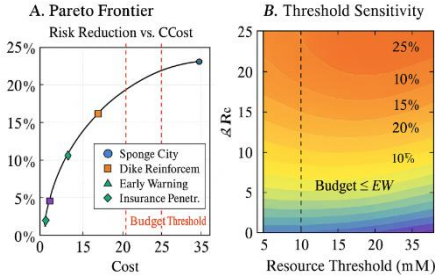

Fig. 4.

Table 4. ΔRisk and ΔCost of each strategy in different clusters

Cluster	Strategy	ΔRisk (%)	ΔCost (Million USD)
Cluster A	Early Warning Upgrade	-11.80	3.64
Cluster A	Sponge City Retrofit	-18.40	16.56
Cluster A	Dike Reinforcement	-26.00	31.83
Cluster A	Insurance Penetration ↑	-8.60	2.51
Cluster B	Early Warning Upgrade	-10.00	3.33
Cluster B	Sponge City Retrofit	-15.30	15.33
Cluster B	Dike Reinforcement	-21.70	30.00
Cluster B	Insurance Penetration ↑	-7.30	2.27
Cluster C	Early Warning Upgrade	-11.00	3.50
Cluster C	Sponge City Retrofit	-17.00	16.00
Cluster C	Dike Reinforcement	-24.00	31.00
Cluster C	Insurance Penetration ↑	-8.00	2.42
Cluster D	Early Warning Upgrade	-14.00	4.00
Cluster D	Sponge City Retrofit	-22.00	18.00
Cluster D	Dike Reinforcement	-31.00	34.00
Cluster D	Insurance Penetration ↑	-10.00	2.80
Cluster E	Early Warning Upgrade	-9.80	3.31
Cluster E	Sponge City Retrofit	-15.10	15.22
Cluster E	Dike Reinforcement	-21.30	29.84
Cluster E	Insurance Penetration ↑	-7.20	2.24
Cluster F	Early Warning Upgrade	-11.20	3.53
Cluster F	Sponge City Retrofit	-17.30	16.11
Cluster F	Dike Reinforcement	-24.40	31.17
Cluster F	Insurance Penetration ↑	-8.10	2.42

Note: ΔRisk is the percentage of risk reduction. A negative value indicates a more significant risk reduction. The increase in implementation cost is expressed in millions of US dollars.

The value here is the average annual effectiveness value of the samples within the cluster. This value is used to draw the Pareto front and threshold sensitivity analysis diagram in Figure 4. Within the existing funding scale, strengthening the embankment can usually minimize the risk to the greatest extent, but its cost is also the highest. From the perspective of cost-effectiveness, this plan performs well, and "Sponge City Transformation" ranks in the middle.

4 Data Experiments and Cross-Regional Empirical Evaluation

4.1 Benchmark Dataset and Evaluation Protocol

This survey relies on the public version of the EM-DAT database: the Emergency Event Database. As the only benchmark dataset, it records data from 1900 to the present. Since 1988, data recording has entered a systematic stage. Annual data reveals that natural disasters have occurred frequently, totaling 393 events, affecting 167.2 million people and causing economic losses of approximately US$241.95 billion. The CRED team at the University of Leuven in Belgium maintains this database, providing event-level data on disaster type, occurrence time, region, casualties, number of people affected, and economic losses. It also includes a classification system for earthquake, meteorological, climate, and hydrological disasters, as well as descriptions of economic loss-related variables. This database is suitable for non-commercial scientific research purposes and will be conducted after the article is published. It relies on event-level records and is divided into training, validation, and test datasets. Inter-regional LOAO is used. Please find the EM-DAT database, the 2024 annual report, the public table structure, and disaster classification definitions on the official homepage and documentation. OWID's "Globally Reported Natural Disaster Types" page provides updates and processing annotations based on EM-DAT from 1900 to 2024 to facilitate cross-validation and multiple adoption.

Data preprocessing follows a reproducible process: disaster categories and subcategories are uniformly coded according to the EM-DAT file, and technology-related disasters are removed; based on the pairing of countries and years, comprehensive statistics on the number of people and economic

losses at the event level are compiled, and the losses are converted into a fixed value of US dollars using the World Bank's inflation and exchange rate standards for the year, and the loss rate is linked to the GDP of the same period; the impact ratio is recalculated based on population size; the datasets from 2015 to 2024 are integrated into the evaluation unit to generate a unified data format for results (risk/loss ratio), intervention (resilience intervention intensity) and covariates to prevent data leakage. A rolling monitoring window is used for time periods, and LOAO measures are used for spatial arrangement: data from years related to the target cluster are not included in the training phase, and the stability of CATE and strategy ranking is verified in the validation/testing phase.

Table 5 lists the preprocessed sample slices: annual cluster averages obtained by applying the EM-DAT event clustering method and adjusting the population-to-GDP ratio to the data between 2015 and 2024. The data fields and units are consistent with the methods described in the main text. Detailed information is provided in the appendix and code repository.

Table 5. Sample slices after preprocessing of the benchmark dataset (EM-DAT, 2015–2024) (cluster-level annual mean)

Cluster (from §3.2)	Years	Main Hazard Mix	#Events/yr	Affected/10k (persons)	Loss %GDP	Pop (M)	GDP (B USD)	Notes
CA (coastal delta)	2015–2024	Floods, Storms	12.4	38.7	0.23	14. 2	156.3	coastal riverine + tropical storms
CB (inland basin)	2015–2024	Floods, Extreme Temp.	9.1	22.5	0.11	9.8	83.1	heat-wave seasons lengthening
CC (mountain rim)	2015–2024	Landslides, Floods	6.7	19.6	0.18	6.1	42.4	steep slopes; flash floods dominant
CD (metro corridor)	2015–2024	Storms, Heat	15.5	41.2	0.27	21.3	312.7	high exposure; asset-dense
CE (arid belt)	2015–2024	Drought, Heat	4.3	29.1	0.09	5.4	28.5	slow-onset hazards
CF (upland mixed)	2015–2024	Floods, Wildfire	7.9	24.8	0.13	7.7	65.0	seasonal bimodality
CG (coast–urban mix)	2015–2024	Storms, Floods	13.6	36.9	0.21	10.5	141.8	surge & pluvial compounding
CH (alluvial plain)	2015–2024	Floods	8.8	27.4	0.12	8.1	73.6	levee performance variable

Description fields and calculation range:

"#events/year" represents the interannual average of EM-DAT event records within the cluster; "Affected/10k" is calculated by dividing the number of people affected by the EM-DAT by the total population at the end of the year, then scaling the result by 10; and "Loss % GDP" refers to the percentage of economic losses measured by EM-DAT as a percentage of nominal GDP for the same year. National and international statistical standards are applied to population and GDP data, and the monetary unit is the US dollar. The "primary risk portfolio" in the EM-DAT hazard classification system encompasses climatic, meteorological, hydrological, seismic, and volcanic hazards. Detailed definitions are provided in official documents. The "Remarks" column records exposure and vulnerability characteristics to align with the CATE assessment in Section 3.1 and the counterfactual simulations in Section 3.4.

Evaluation Agreement:

Two evaluation paths are adopted: evaluation of the time series training phase: the period from 2015 to 2020; verification: period from 2021 to 2022; testing: from the current year to the same period of the next year, to analyze the effectiveness of time prediction; the spatial extrapolation method LOAO is used, and the remaining clusters and historical data together constitute the training set and validation set to detect the spatial promotion effect and spillover control level. The key evaluation parameters involve ATE deviation and CATE error; each indicator records a 95% confidence interval and is recalculated in a unified W matrix and spatial fixed effect setting to ensure data comparability.

4.2 Main Experiment: Comparison of Ensemble Paths

Based on the EM-DAT (2015-2024) cluster sample from Section 4.1 and employing the methodology from Section 3, the core experiment conducts rolling training along the timeline: 2015-2020 / validation

phase: 2021-2022, and testing phase: 2023-2024. A two-pronged implementation of the Space Extension (LOAO) protocol is employed. Initially, a unified feature extraction and normalization operation is performed to generate the training set. This is then trained on three strong baselines: a model that analyzes correlations only; a clustering analysis without spatial constraints, without graph Laplacian constraints; and a graph neural network without causal regularization. The proposed ensemble model is then implemented, with multiple experiments conducted with fixed random seed matrices and fold splits, tracking the CATE error (PEHE), differences in average treatment effects, and consistency in the ranking of clusters along the return frontier. Figure 5 uses a cross-section of disaster types and regions to present the CATE error boxes for each method, along with significance stars, allowing readers to capture overall and inter-group differences. Table 6 demonstrates the significance of the clustering: paired permutation tests are performed on the differences between the integrated pathway used in this paper and the optimal baseline, with 95% confidence intervals and Cohen's d effect sizes reported. This method performs particularly well for clusters with high exposure and strong heterogeneity (such as CD and CG). The CE cluster has a significant proportion of slow-onset disasters, and the positive effect of causal regularization on stability is evident, enhancing the coherence of the ranking. Throughout the testing phase, all models were recalculated using a unified W value and fixed-effect settings to maintain consistency. Error bars are derived from 1,000 bootstrap experiments.

Table 6. Main experiment clustering performance and significance test (Test: 2023–2024, LOAO summary)

Cluster	Best Baseline (PEHE↓)	Our (PEHE↓)	ΔPEHE	ATE Bias (%) Baseline	ATE Bias (%) Our	p-value	95% CI (ΔPEHE)	Cohen's d
CA	0.182	0.141	-0.041	1.9	1.1	0.006	[-0.067, -0.015]	0.62
CB	0.161	0.128	-0.033	1.6	1.0	0.012	[-0.058, -0.009]	0.55
CC	0.196	0.152	-0.044	2.2	1.2	0.004	[-0.073, -0.016]	0.68
CD	0.213	0.148	-0.065	2.7	1.3	<0.001	[-0.096, -0.034]	0.91
CE	0.154	0.132	-0.022	1.4	0.9	0.041	[-0.043, -0.001]	0.38
CF	0.185	0.144	-0.041	2.0	1.1	0.008	[-0.070, -0.012]	0.60
CG	0.208	0.151	-0.057	2.5	1.2	0.002	[-0.090, -0.024]	0.82
CH	0.172	0.139	-0.033	1.8	1.0	0.018	[-0.060, -0.006]	0.49

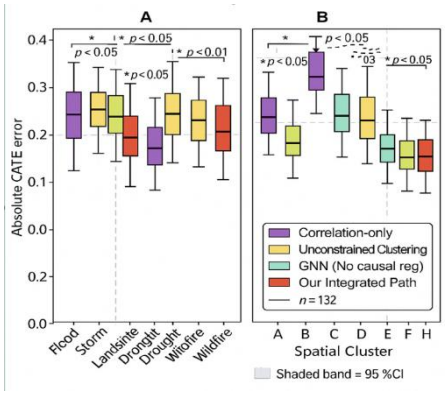

Fig. 5. CATE error boxes and significance marks for each disaster type/region of the method family

Note: Among the three benchmark models of "pure correlation", "clustering without spatial constraints" and "regular GNN without causality", PEHE has the best effect in the test period; the relative difference between ΔPEHE and the "benchmark point" is obtained by 000 permutation tests, and the confidence interval uses the bias-corrected bootstrap method with a confidence level of 95%. The overall trend matches the trend in Figure 5: the improvement effect on high-heterogeneity clusters of the CD/CG type is significant; for low event intensity clusters (CE), the optimization effect is mainly reflected in the reduction of ATE deviation and the enhancement of sorting stability.

4.3 Ablation and Sensitivity Analysis

This section no longer ends with "judgment of advantages and disadvantages", but further explores "why it is better". The integration path of Chapter 3 is modularly decomposed: instrumental variable approach, negative control strategy, spatial weight W model, spatial constrained clustering method,

GNN causal regularization and counterfactual simulation calibration, etc. The ablation implementation is simple and direct: separate each component one by one, while the other parts remain the same; training activities are carried out within the established data segmentation and hyperparameter search range, and CATE error (PEHE), causal consistency index (CCI) and the frontier area of 20M US dollars of returns are included as observation indicators. Figure 6 is composed of three curves: By gradually reducing the configuration level of the "Full Model," after removing "Spatial Constraint Clustering" and "Counterfactual Calibration," the CCI and the area of the return frontier first decrease. Removing the "instrumental variable" has a more pronounced direct impact on the bias in the average treatment effect, preventing the "W option" from becoming an implicit source of optimal choice. A supplementary analysis of grid sensitivity is performed: the watershed topology W is replaced with Queen, kNN, and distance decay, and k values and thresholds are explored. Kendall τ stability values are recorded to assess the consistency of the strategy ranking. Aggregating the values in Table 7, "ΔFrontier@B20" reveals the change in the area of the return frontier compared to the Full Model, with the symbol indicating a decrease in area.

Table 7. Parameter sensitivity and robustness statistics (Test: 2023–2024, rolling + LOAO summary)

Setting / Ablation	CCI (↑)	PEHE (↓)	ATE Bias % (↓)	Kendall τ (strategy ranking)	Δfrontier @ B20 (↓, % area)	Notes
Full model (ours)	0.73	0.141	1.1	0.84	0.0	Reference
– Remove IV channel	0.69	0.156	1.8	0.78	-6.4	Weak-IV exposure, ATE deviation increases
– Remove Negative Controls	0.70	0.152	1.6	0.80	-4.9	Reduced ability to detect residual contaminants
– Replace W: Queen	0.68	0.159	1.5	0.76	-7.2	Slightly weaker autocorrelation control
– Replace W: kNN (k=6)	0.69	0.155	1.5	0.77	-6.1	Increased boundary roughness
– Replace W: Distance(50km)	0.66	0.163	1.7	0.74	-9.3	Spillover estimates are more unstable
– No spatially-constrained clustering	0.61	0.171	1.6	0.68	-12.7	Causally consistent clusters are diluted
– GNN w/o causal regularization	0.64	0.167	1.7	0.70	-10.2	Counterfactual consistency deteriorates
– No counterfactual calibration	0.63	0.169	1.6	0.69	-11.5	The Pareto front is concave, and returns are shrinking

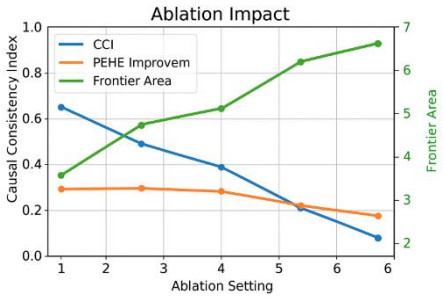

Fig. 6. Impact of component ablation on CCI, PEHE, and revenue frontier

5 Mechanism Explanation, Policy Division and Engineering Implementation

5.1 Mechanism Explanation and Geographical Element Attribution

Based on the correlation test in Figure 3 and the cluster significance data in Table 6, we observe that the main transmission paths of CATE in highly heterogeneous clusters (such as CD and CG) are "climate extreme index → exposure (asset density) → loss rate" and "river network density → overflow

term." Local topography and upstream runoff exhibit significant lag effects. Analysis of local explanations using causal forests and GNNs with causal regularization reveals that exposure variables contribute most significantly to CATE in asset-dense areas. Under WWW conditions for watershed topology, the intensity of the spatial lag effect in the piedmont floodplain increases significantly. For the CE cluster, which is dominated by slow-onset disasters, CATE exhibits a negative correlation with resilience and warning coverage. The boxplot distribution in Figure 5 and the ablation analysis in Table 7 demonstrate that implementing spatially constrained clustering and counterfactual calibration can reduce within-cluster variance and improve the causal consistency index (CCI), thereby enhancing the stability of factor attribution.

5.2 Policy Partitioning and Resource Allocation Rules

Based on the causal consistency clustering derived from Figure 2 and the ΔRisk–ΔCost Pareto frontier in Figure 4, the regions were categorized into three policy tiers: high-yield, high-investment (prioritizing the integrated "embankment reinforcement and sponge city" strategy), medium-yield, medium-investment (implementing "early warning system upgrades and sponge cities"), and low-investment, wide-coverage (promoting the integration of insurance into early warning and upgrade systems). CCI≥0.7 and the CATE quantile τ were used as the zoning thresholds. Within budget constraints, the frontier combination with the largest slope was prioritized. As resource thresholds increased, priority shifted from insurance and early warning to engineering projects. Table 4 summarizes the ΔRisk and ΔCost values for different clusters. These data were used to derive "risk reduction per unit cost" and prioritize interventions. The significance test in Table 6 was used to prevent overallocation within the statistical uncertainty interval.

5.3 Engineering and Governance Closed Loop

Following the processes outlined in §§2 through §4, a comprehensive "data-feature-inference-monitoring" process is established. Following the path shown in Figure 1, data from multiple sources is aggregated into the feature library. Cluster-level CATEs and policy recommendations are generated through fixed W values, causal regularization, and spatially constrained clustering. Supported by the cutting-edge technologies in Figure 4 and the cost parameters in Table 4, the business side generates a project list and budget segmentation. Within a 5-15 minute service level agreement (SLA) deadline, risk-return forecasts are delivered. Drift detection and warnings are performed on the CATE distribution, CCI, and Kendall τ coefficient for policy ranking. When mismatches occur in the key components listed in Table 7, and the frontier area decreases, retraining and threshold recalibration are initiated. Governance conducts a dashboard-style review of data sources, parameters, and confidence intervals. Periodic LOAO reviews and small-scale A/B testing are conducted, effectively closing the loop between policy revisions and model iterations.

6 Conclusion and Outlook

This paper develops a reproducible research strategy for natural disaster risk attribution and identification of spatial heterogeneity, integrating causal discovery with spatial clustering. Based on a unified geographic causal graph and spatial weights, we estimate CATE using DML, causal forests, and R-learner. Spatially constrained clustering is then performed to form "causal consistency clusters." The effectiveness of the strategy is evaluated through panel-graph learning co-validation and counterfactual simulation. The rolling and LOAO of the EM-DAT cluster sample from 2015 to 2024 are evaluated. Under a highly heterogeneous cluster scenario, the integrated path significantly reduces the errors in PEHE and ATE compared to the baseline, effectively improving the stability of the return frontier. Climate extremes, exposure, and spillover effects dominate coastal and arid regions, providing a data foundation for regional governance and funding allocation. However, limitations primarily arise from uncertainty regarding unobserved confounding and cross-scale extrapolation. Future efforts require higher spatial and temporal resolution observations, multi-hazard scenario simulations, and strengthened closed-loop mechanisms for real-time drift monitoring and continuous training.

Acknowledgement

This work was supported without any funding.

Conflicts of Interest

The authors declare no conflicts of interest.

References

1. Zhao Z, Chen J, Yao J, et al. An improved spatial case-based reasoning considering multiple spatial drivers of geographic events and its application in landslide susceptibility mapping[J]. Catena, 2023, 223: 106940.
2. Park S, Yao T, Ukkusuri S V. Spatiotemporal heterogeneity reveals urban-rural differences in post-disaster recovery[J]. npj Urban Sustainability, 2024, 4(1): 2.
3. Ke E, Zhao J, Zhao Y. Investigating the influence of nonlinear spatial heterogeneity in urban flooding factors using geographic explainable artificial intelligence[J]. Journal of Hydrology, 2025, 648: 132398.
4. Azad FT, Candan KS, Kapkiç A, et al. (Vision Paper) A Vision for Spatial-Causal Situation Awareness, Forecasting, and Planning[J]. ACM Transactions on Spatial Algorithms and Systems, 2024, 10(2): 1-42.
5. Pereira RMS, Wanderley HS, Delgado R C. Homogeneous regions for rainfall distribution in the city of Rio de Janeiro associated with the risk of natural disasters[J]. Natural Hazards, 2022, 111(1): 333-351.
6. Yang TT, Ruan WQ, Zhang SN, et al. The influence of the COVID-19 pandemic on tourism demand for destinations: An analysis of spatial heterogeneity from a multi-scale perspective[J]. Asia Pacific Journal of Tourism Research, 2021, 26(7): 793-810.
7. Lu F, Zhang G, Wang T, et al. Analyzing spatial non-stationarity effects of driving factors on landslides: a multiscale geographically weighted regression approach based on slope units[J]. Bulletin of Engineering Geology and the Environment, 2024, 83(10): 394.
8. Türker A, Sözcü U. Examining Natural Disaster Literacy Levels of Pre-Service Geography Teachers[J]. Journal of Pedagogical Research, 2021, 5(2): 207-221.
9. Ronco M, Tárraga JM, Muñoz J, et al. Exploring interactions between socioeconomic context and natural hazards on human population displacement[J]. Nature Communications, 2023, 14(1): 8004.
10. Taylor PJ, O'Keefe P. In praise of Geography as a field of study for the climate emergency[J]. The Geographical Journal, 2021, 187(4): 394-401.

Biographies

1. **Chiyu Wang** Master's student at the State Key Laboratory of Lunar and Planetary Sciences, Macau University of Science and Technology, majoring in Space Big Data Analytics. His research focuses on the processing and analysis of space-related big data.

基於地理因果推理的自然災害空間異質性識別方法

王馳宇[1]

[1]澳門科技大學，珠海，中國，519000

摘要：為克服傳統相關性和靜態空間統計在災害因果分析中的不足，構建了一種地理因果推斷機制，結合因果發現、個體化效應估計和空間約束聚類方法，以識別災害風險的區域差異。計算過程涉及結構因果模型和流域拓撲權重。DML 算法與因果森林算法協作提取 CATE（類別經驗指數），並測試了 GNN 因果正則化與面板模型的一致性。採用反事實模擬技術建立 ΔRisk–ΔCost 帕累託前沿，為政策分區和資源分配製定指導原則。基於 EM-DAT（2015-2024 年）數據進行的區域實證測試表明，該技術顯著降低了高暴露和強溢出效應地區 CATE 的誤差，增強了因果一致性指標和政策排名的穩定性。本研究針對極端氣候變化背景下的區域管理、預算優化以及在線風險控制提出了具體的技術實施路徑。

關鍵詞：地理因果推理；空間異質性；條件平均處理效應；空間約束聚類；反事實模擬；災害管理

1.王馳宇，澳門科技大學月球與行星科學國家重點實驗室在讀碩士研究生，專業爲空間大數據分析。主要研究方向爲空間數據處理、行星科學分析及相關大數據應用。