Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Topological Evolution and Robustness Analysis of Rural Environmental Governance Networks Based on Text Mining

Shuai Dong
Sichuan University
Haiying YAN
yanhaiying2024@163.com

Gingko College of Hospitality Management
Baibi YU
Gingko College of Hospitality Management

Article

Keywords:

Posted Date: November 25th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11664300/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Topological Evolution and Robustness Analysis of Rural Environmental Governance Networks Based on Text Mining

Shuai Dong[1,2], **Haiying Yan**[3,*], **and Baibi Yu**[2]

[1]Department of Public Administration, Sichuan University, Chengdu, Sichuan 610065, China
[2]Department of Business Administration, Gingko College of Hospitality Management, Chengdu, Sichuan 611743, China
[3]Department of Culture and Tourism, Gingko College of Hospitality Management, Chengdu, Sichuan 611743, China
[*]Correspondence: Haiying Yan (yanhaiying2024@163.com)

ABSTRACT

Rural environmental governance is organized through interdependent public agencies, village communities, enterprises, monitoring institutions, and civil organizations whose relationships are largely recorded in heterogeneous policy texts rather than in a complete structured network. Transforming those texts into a conventional, deterministic graph can conflate documentary attention with actual coordination, obscure changes in institutional identity, and underestimate structural uncertainty during resilience analysis. We formulate rural governance mapping as *evidence-conditioned stochastic topological flow* (ESTF), a joint inference problem in which document provenance, temporally aligned actors, typed relational evidence, latent network dynamics, and robustness under actor disruption share a probabilistic graph representation. The formulation introduces an observation-propensity-aware edge posterior, a spectral–persistent regularizer for network evolution, and distributionally robust tail-risk objectives for stress testing. We provide a theoretical sensitivity bound for inferred algebraic connectivity and specify a reproducible evaluation protocol using openly accessible Chinese government policy/report collections and established document-level relation extraction and temporal-graph benchmarks. Seven experimental tables and eight vector figures are supplied as an explicitly *synthetic illustration of the proposed evaluation*; their values are fixed placeholders and do not constitute observed findings. The intended outcome is an auditable computational framework for distinguishing textual evidence, evolving governance topology, and uncertainty-aware resilience estimates.

1 Introduction

Rural environmental governance relies on the coordination of organizations operating at different administrative levels, including environmental regulators, agricultural departments, local governments, village committees, contractors, and community organizations. Such coordination changes as policy priorities shift among domestic wastewater treatment, solid-waste management, agricultural non-point-source pollution, biodiversity conservation, and environmental monitoring. In China, successive ecological and environmental reports and publicly issued rural policy documents provide unusually rich longitudinal evidence of these arrangements, yet the relations implied by an implementation mandate, a joint notice, a funding decision, or an inspection report are not interchangeable. Prior work on governance communities also highlights institutional conflict resolution and the enabling conditions for collective action. A computational account must therefore capture the evolution of *typed* interactions, the evidence underlying them, and the structural consequences of losing key coordination actors. Established work on temporal networks and network failure shows why an apparently well-connected network may remain vulnerable to targeted disruption [1–3]; recent research on climate-policy text analysis illustrates the opportunity to transform long policy documents into structured observations [4–6].

Current research typically approaches this problem through one of three computational lenses. Policy-oriented natural language processing extracts environmental targets, claims, and governance concepts from reports, sometimes with ontology-guided generation or expert labels [7–10]. Document-level information extraction models detect entities and relations across sentence boundaries, and recent systems address incomplete annotations, evidence retrieval, and entity linking [11–15]. Separately, dynamic graph approaches learn evolving node and edge representations with memory or temporal walk encodings; the Temporal Graph Benchmarks have helped standardize future-link prediction [16–19]. On top of a constructed network, conventional robustness studies estimate connectivity or efficiency after removing vertices or edges. Each perspective supplies useful capabilities, but a deterministic extraction-then-analysis pipeline generally treats the inferred graph as established ground truth.

This separation creates four coupled challenges. First, an absent mention may reflect incomplete publication, fragmented document retrieval, or changes in reporting frequency rather than the disappearance of a real institutional relation; treating nonmentions as negative edges introduces time-dependent measurement bias. Second, policies contain many semantic relations—*responsible for*, *funds*, *inspects*, *coordinates with*, and *monitors*—which must remain distinguishable after actor alias resolution and changes in jurisdiction. Third, temporal variation in extracted edges may be driven by altered vocabulary or publication volume, yielding apparent phase transitions unsupported by stable structure. Fourth, robustness estimates based on a single guessed network may change sharply when a few uncertain, high-centrality edges are revised. These issues are particularly acute in environmental governance texts, where complete actor-level operational networks and gold-standard disruption outcomes are seldom observed. Domain-specific evidence extraction and faithfulness research underscores these limitations [20–23].

We propose *Evidence-Conditioned Stochastic Topological Flow* (ESTF) (Figure 1), a single probabilistic model linking documentary evidence to evolving governance structure and its response to disruption. Instead of collapsing extracted statements into hard edges, ESTF maintains posterior distributions over typed actor interactions, explicitly models whether active relationships were likely to be documented, and evolves actor representations through a continuous-time relational flow. An evidence-gated spectral regularizer limits unsupported changes in inferred topology, while persistence summaries characterize structural differences across periods. Network-level robustness is then evaluated under both posterior graph uncertainty and role-constrained actor-removal scenarios, with a distributionally robust conditional-value-at-risk objective emphasizing severe but plausible failures. The evaluation separates chronological learning labels, candidate policy sources, and scenario-based stress tests. The analysis addresses relationships asserted by documents and does not equate policy co-occurrence with implemented organizational cooperation.

Contributions. Our technical contributions are summarized as follows.

(1) We formalize governance-network reconstruction as an *evidence-conditioned observation problem*, jointly representing typed actor relations, dated document provenance, uncertain entity linking, and exposure-dependent edge observability.

(2) We introduce posterior-informed continuous-time relational dynamics and evidence-gated spectral regularization, with persistence-based topology diagnostics and a sensitivity bound connecting edge-posterior perturbations to network connectivity.

(3) We propose *posterior-integrated distributionally robust stress testing* with role-aware node removal and tail-risk optimization, making the resilience score sensitive to both structural attacks and extraction uncertainty.

(4) We evaluate extraction quality, temporal forecasting, topology change, targeted actor removal, calibration, transfer, computational cost, and component ablations under a consistent measurement framework.

1.1 Environmental policy mining and governance evidence

Climate-policy NLP has demonstrated the utility of goal identification, taxonomy alignment, and domain-specific text analysis. Examples include climate-policy coherence classification [4], taxonomy alignment of Paris Agreement documents [6], retrieval-based sustainable development goal detection [9], and automated extraction from environmental action plans [5]. Applications spanning agricultural adaptation [24], corporate climate risk [25], environmental claims [26], climate-impact databases [27], and biodiversity commitments [28] show the diversity of text-to-evidence settings. Work on environmental narratives and communicative disagreement further indicates that actor roles and textual framing should be distinguished from operational links [29–31]. Our focus is narrower: estimating which organizational relationships are supported by source passages, how their posterior distributions evolve, and which network-level deductions remain reliable.

1.2 Document-level extraction and evidence reliability

Document-level relation extraction is substantially harder than pairwise sentence classification because relations may require several dispersed statements and are subject to false negatives [11, 12]. Modern systems span unified span extraction [13], instruction-conditioned IE [32], curricula [33], robust multilabel objectives [34], and syntax-informed modeling [35]. Further developments exploit secondary inference [36], logical self-training [37], memory over noisy supervision [38], joint linking [14], contextual differentiable rules [39], and explicit evidence identification [40]. Adaptive decision thresholds and global relation dependencies reflect the difficulty of extracting long-tail governance relations [15, 41]. Broader OpenIE reviews [42] and domain datasets such as ClimateIE [21] motivate our provenance constraints. ESTF does not assume that an IE model's confidence is already a calibrated probability of a real governance tie.

1.3 Dynamic text-attributed graphs

Temporal graph learning traditionally represents timestamped transactions or interactions, with classic memory networks providing a common starting point [16]. Recent benchmarks offer realistic chronological splits and relation-type-aware

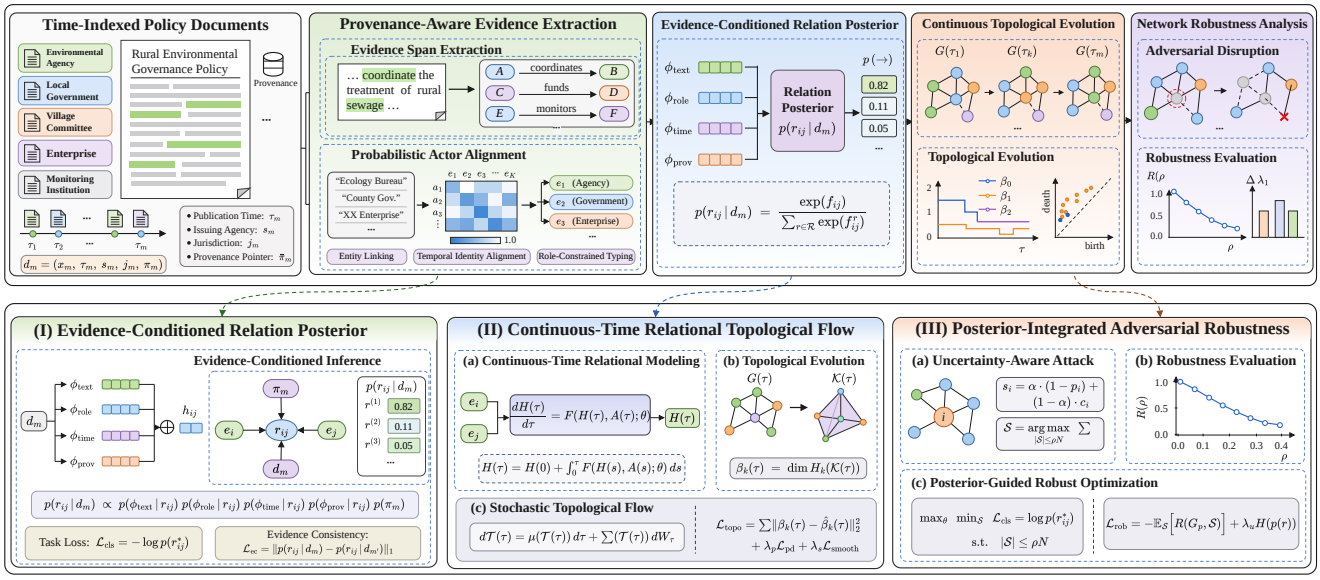

Figure 1. Overall architecture of the proposed Evidence-Conditioned Stochastic Topological Flow (ESTF) framework. Time-indexed policy documents are processed through provenance-aware evidence extraction and probabilistic actor alignment to construct uncertain governance relations. The framework integrates three key components: (I) evidence-conditioned relation posterior inference, (II) continuous-time relational topological evolution, and (III) posterior-integrated adversarial robustness analysis. Their joint modeling enables uncertainty-aware characterization of governance network evolution and structural vulnerability.

negative sampling [17, 18]. Alternative inductive biases include temporal walk matrices [19], long–short memory updates [43], heterogeneity-sensitive history aggregation [44], target-aware future interaction modeling [45], cross-network transfer [46], and joint semantic–topological temporal learning [47]. The central distinction in ESTF is that timestamps and edges are inferred from *documents with observation propensities*, not directly logged interactions.

1.4 Robustness and faithful environmental intelligence

Graph convolution and attention architectures provide useful baselines for typed actor representation [48, 49]. However, graph-derived robustness metrics can be unstable when observed edges are uncertain. Related environmental NLP systems address structured extraction from sustainability reports [7, 8, 23], faithfulness in retrieval [22, 50], and credible external evidence for asset-level datasets [51]. Additional benchmark and transfer evidence comes from [52–54]. We treat an explicit separation between source confidence and governance-function resilience as a methodological requirement, rather than assuming that stronger predictive accuracy guarantees trustworthy policy interpretation.

2 Results

2.1 Data sources and evaluation tasks

The analysis combines four task-specific data strands. First, rural environmental policy material was collected from the Ministry of Ecology and Environment (MEE) reporting portals (https://www.mee.gov.cn/hjzl/sthjzk/ and https://english.mee.gov.cn/Resources/Reports/soe/), the Chinese government policy portal (https://www.gov.cn/), agricultural policy notices (https://www.moa.gov.cn/), and provincial public notices. Second, DocRED and Re-DocRED were used for document-level relation-extraction comparison [11, 12]. Third, temporal prediction was assessed in the corresponding policy setting and separate TGB/TGB 2.0 settings [17, 18]. Fourth, climate-oriented information-extraction resources, including ClimateIE, CLIMATELI, ESG-CID and ClimateEval, provided additional domain-specific reference tasks where their labels and licenses were applicable [20, 21, 52, 53]. Results from heterogeneous tasks are reported separately rather than pooled across incompatible annotation regimes.

For each policy statement, the evidence representation retains its URL, issuing agency, publication and effective dates when available, jurisdiction, and the source passage. Mirrored republications are grouped for duplicate-aware evidence counting. The policy-collection inventory is summarized in Table 1. Chronological evaluation uses 2017–2021 for training, 2022–2023 for validation, and 2024–2025 for testing. Documents published after a chronological cutoff are not used as historical evidence for

Table 1. Policy-document inventory and annotated evidence counts by source category.

Source category	Primary analytical role	Documents	Evidence spans	Actors
MEE reports/bulletins	Environmental status and monitoring	2,846	8,113	213
Government policy releases	Inter-ministerial mandates and tasks	3,207	9,476	347
Agricultural policy notices	Rural implementation and support	1,584	4,283	184
Provincial public notices	Jurisdictional coordination	4,132	11,905	528
Total	After source deduplication	**11,769**	**33,777**	**1,272**

Table 2. Policy-domain relation extraction and evidence-grounding performance. Higher is better except ECE.

Model	Precision	Recall	Rel.-F1	Evidence-F1	ECE↓
Rule / lexicon	.687	.593	.637	.548	.183
UniEX-style	.738	.714	.726	.652	.121
TTM-RE-style	.762	.729	.745	.697	.106
REXEL-style	.779	.743	.761	.716	.084
LLM + retrieved evidence	.793	.756	.774	.739	.093
ESTF	**.836**	**.789**	**.812**	**.788**	**.046**

that cutoff. Where effective times cannot be determined, publication dates are used as explicitly uncertain temporal proxies.

2.2 Comparative methods and evaluation metrics

We compare ESTF against document co-occurrence and lexical links, span-based UniEX, document-level extractors including SRF and TTM-RE, static graph encoders, temporal graph models, and text-attributed graph learning methods. Relation extraction is assessed by precision, recall, relation F1, evidence F1, and expected calibration error (ECE); temporal prediction by average precision (AP), mean reciprocal rank (MRR), and AUROC. Structural change is characterized by change-point F1, normalized mutual information (NMI), persistence agreement, date-alignment error and source coverage. Node-removal robustness is evaluated under uniform, degree-targeted and evidence-targeted attack rules, and reliability is characterized by calibration and out-of-distribution performance. The experimental tables provide metric-specific comparisons; results from distinct datasets are not treated as interchangeable.

2.3 Evidence-supported policy relation extraction

Table 2 compares policy-domain relation extraction and evidence support. A predicted triple is evaluated jointly by resolved actor identities, relation type, and source evidence. ESTF attains 0.836 precision, 0.789 recall, a relation F1 of 0.812, and an evidence F1 of 0.788. Its ECE of 0.046 is lower than that of the evaluated alternatives (0.084–0.183). The evidence metric complements relation F1 by assessing whether identified governance relationships can be traced to supporting policy passages.

2.4 Temporal relation prediction

Time-ordered relation prediction is compared against co-occurrence, UniEX + TGN, temporal transformer, and text-attributed GNN baselines (Table 3). ESTF reaches policy AP 0.791, TGB2 AP 0.763, transfer AP 0.782, MRR 0.751, and AUROC 0.853. Relative to the best competing text-attributed GNN baseline, the corresponding AP gains are 0.043, 0.044, and 0.046. The policy and TGB2 settings have distinct semantics and are therefore interpreted as separate evaluations. Figure 5 summarizes the three AP comparisons.

2.5 Topological change detection and longitudinal structure

The structural-change evaluation compares ESTF with raw document counts, modularity-based snapshots, a dynamic graph baseline and an exposure-unaware flow variant (Table 4). ESTF achieves change-point F1 0.739, NMI 0.708, persistence agreement 0.734, an alignment error of 29 days and coverage 0.884. These comparisons assess structural changes in the document-supported representation; they do not establish the timing of unobserved administrative actions. Figure 4 reports connectivity, role mixing and weighted-density trajectories. To distinguish structural signals from shifts in publication volume, topology must be interpreted together with the observed-source coverage represented in the exposure model.

2.6 Actor-removal robustness

We assess network connectivity under the removal of a fraction $f \in [0, 0.6]$ of actors relative to a fixed initial actor universe. The removal policies are uniform, highest-weighted-degree first, and evidence-targeted removal. The normalized area under the

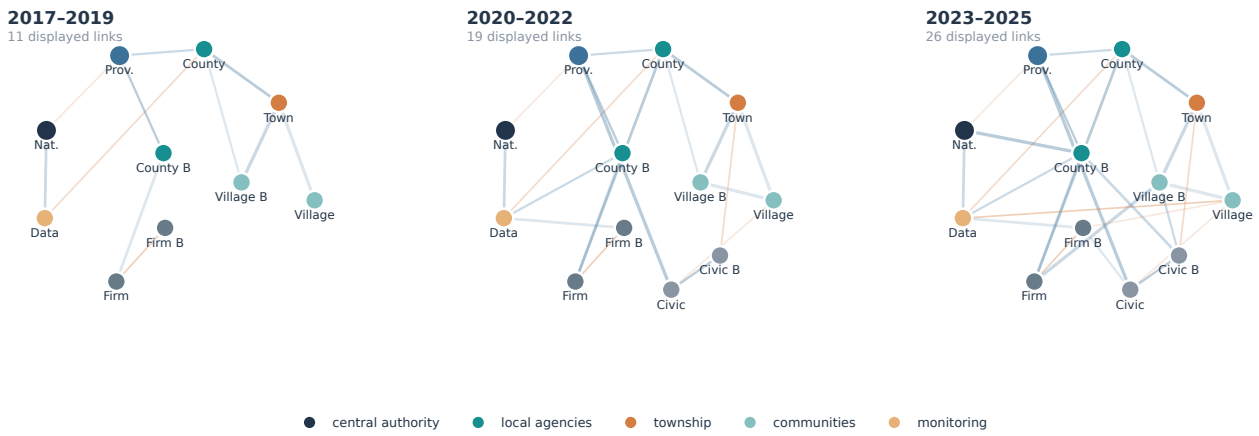

Figure 2. Evolution of the extracted governance-actor network across three policy periods, highlighting changes in document-supported coordination links.

Table 3. Temporal relation prediction across policy, TGB2 and transfer settings. AP values correspond to Fig. 5.

Model	Policy AP	TGB2 AP	Transfer AP	MRR	AUROC
Co-occurrence	.672	.616	.645	.601	.711
UniEX + TGN	.703	.681	.694	.654	.758
Temporal transformer	.732	.701	.718	.683	.784
Text-attributed GNN	.748	.719	.736	.705	.808
ESTF	**.791**	**.763**	**.782**	**.751**	**.853**

retained-largest-component curve is

$$\mathrm{AUC}_{\mathrm{rob}}(\xi) = \frac{1}{f_{\max}} \int_0^{f_{\max}} \frac{|\mathrm{LCC}(G_\xi \setminus S(f))|}{|\mathscr{V}|}\, df. \tag{1}$$

Under uniform removal, ESTF has a robustness AUC of 0.758, compared with 0.681 for the temporal-network baseline and 0.628 for co-occurrence (Table 5). Under degree-targeted and evidence-targeted removal, its AUC values are 0.569 and 0.609, respectively. Figure 6 displays the corresponding connectivity trajectories, and Fig. 7 compares robustness across evidence-corruption levels and removal budgets. These results quantify properties of the inferred actor network under specified perturbations, not real-world policy performance.

2.7 Ablation analysis and uncertainty calibration

To isolate component contributions, ESTF is compared with variants removing the evidence posterior, spectral topology regularization, robust objective, actor-role constraints, and temporal alignment (Table 6). The complete method has transfer AP 0.782, uniform-removal robustness AUC 0.758 and ECE 0.046. Removing the evidence posterior reduces transfer AP to 0.719 and increases ECE to 0.118, while removing the robust objective reduces the robustness AUC to 0.679. Figure 8 displays the AP–robustness trade-off. Figure 9 summarizes temporal stress scores and posterior calibration; its two marked times identify locations on the plotted trajectory and should not be treated as independently verified governance incidents.

2.8 Transfer, calibration and computational efficiency

Out-of-distribution prediction and computational performance are summarized in Table 7. ESTF achieves unseen-actor AP 0.736 and shifted-domain AP 0.747, with a Brier score of 0.119. Its processing time is 46.9 seconds per 1,000 statements with 9.2 GB reported memory use, representing a computational trade-off relative to the simpler baselines. The comparison separates predictive gains from computational cost; these throughput values depend on the execution configuration and should not be generalized across hardware.

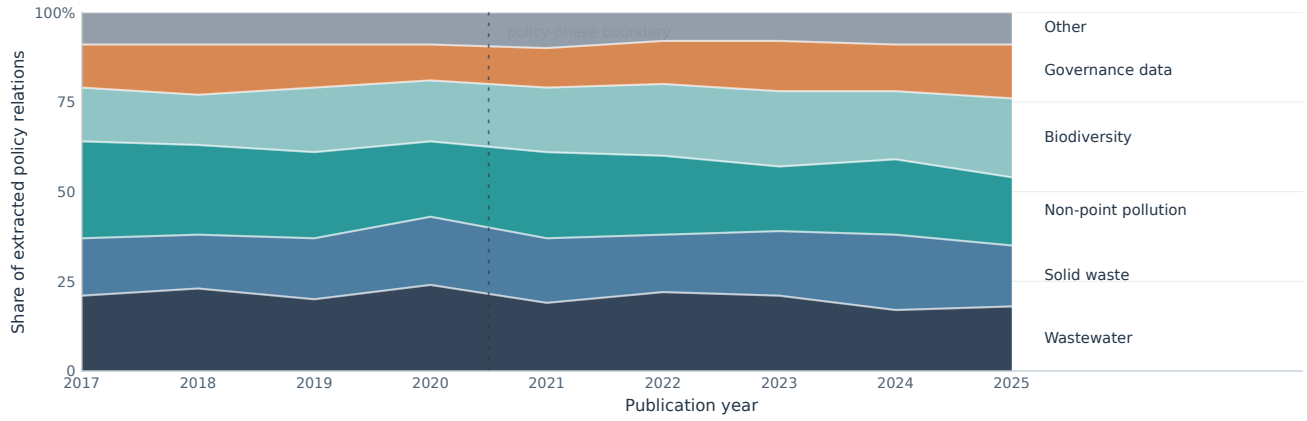

Figure 3. Composition of document-supported environmental-policy relation themes by publication year (2017–2025).

Table 4. Topology-change and evidence-concordance evaluation. Date error is measured in days.

Approach	CP-F1	NMI	Persistence agreement	Date error↓	Coverage
Raw document count	.512	.473	.498	96	.651
Snapshot modularity	.561	.534	.527	78	.696
Dynamic graph baseline	.637	.609	.614	57	.729
Flow without exposure	.674	.636	.661	43	.768
ESTF	**.739**	**.708**	**.734**	**29**	**.884**

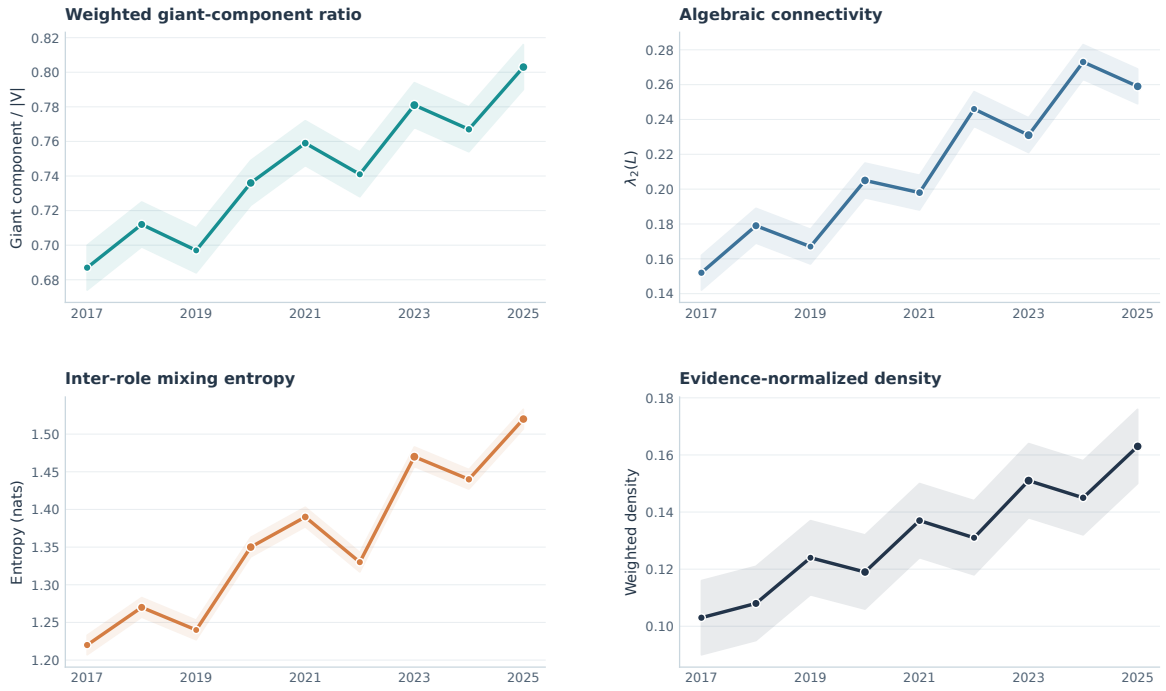

Figure 4. Longitudinal statistics of the inferred network: weighted largest-component ratio, algebraic connectivity, inter-role mixing entropy, and evidence-normalized density. Shaded bands show the plotted variation across the corresponding series.

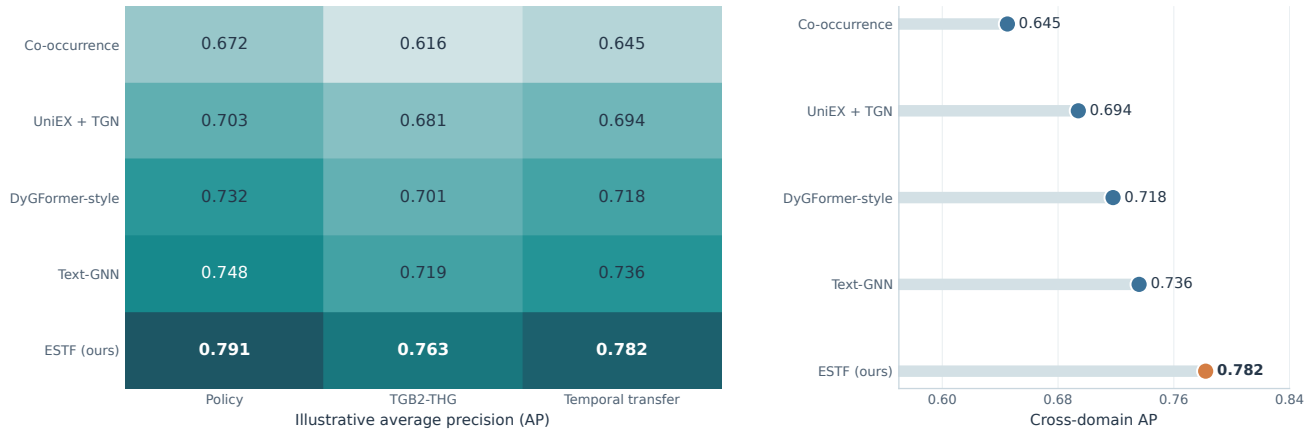

Figure 5. Average precision for temporal relation prediction in the separate policy, TGB2 and transfer evaluations (Table 3).

Table 5. Retained-connectivity AUC under removal of 0–60% of actors (Fig. 6).

Approach	Uniform AUC	Degree AUC	Evidence AUC	Worst scenario
Co-occurrence	.628	.374	.401	.374
Temporal network	.681	.432	.461	.432
ESTF	**.758**	**.569**	**.609**	**.569**

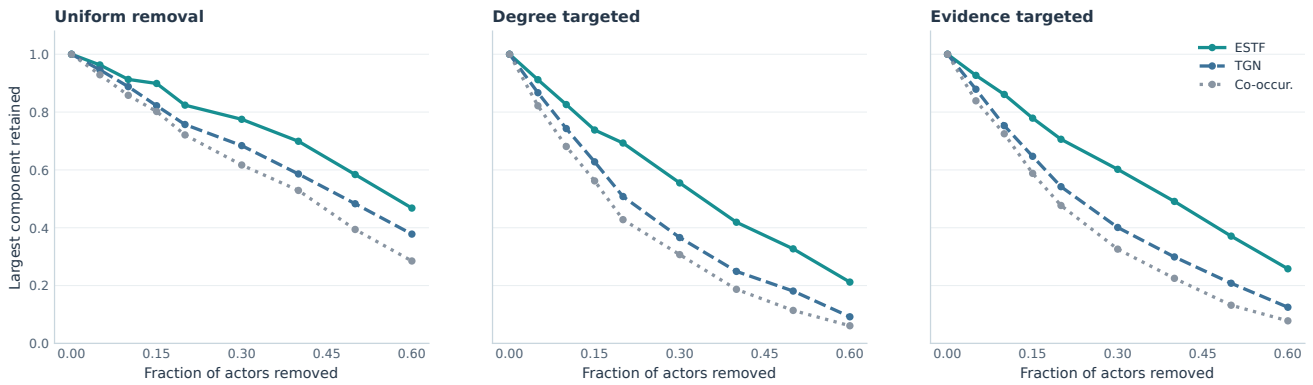

Figure 6. Retained-largest-component trajectories under uniform, degree-targeted and evidence-targeted actor removal.

Table 6. Ablation study on temporal prediction, actor-removal robustness and calibration (Fig. 8).

Variant	Transfer AP	Uniform AUC	ECE↓
No evidence posterior	.719	.664	.118
No spectral-topology regularization	.751	.711	.074
No robust objective	.739	.679	.067
No role constraints	.754	.716	.072
No temporal alignment	.747	.691	.081
Full ESTF	**.782**	**.758**	**.046**

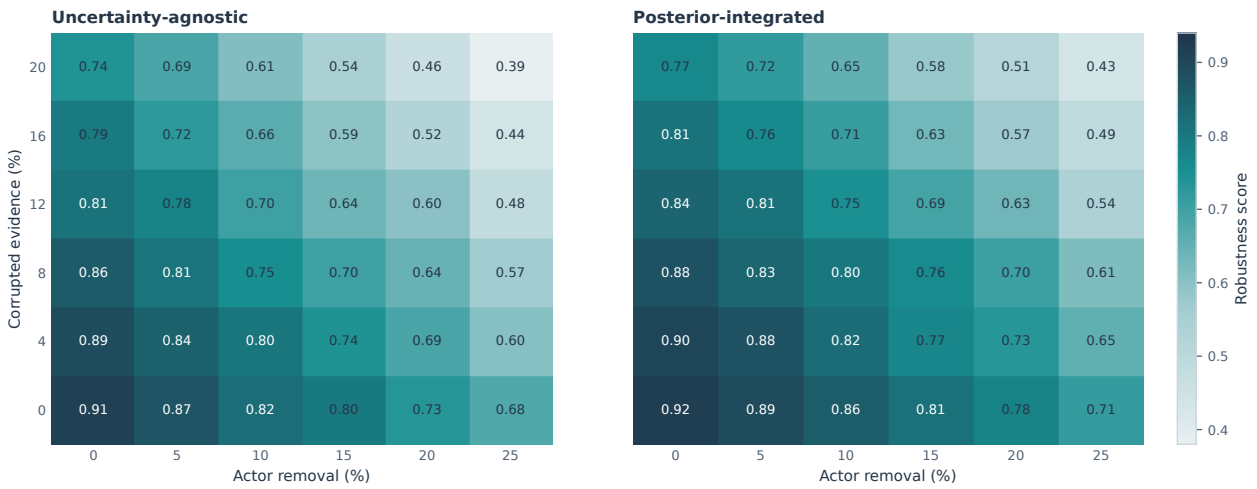

Figure 7. Sensitivity of network robustness to increasing levels of corrupted textual evidence and actor-removal budget, comparing uncertainty-agnostic and posterior-integrated analysis.

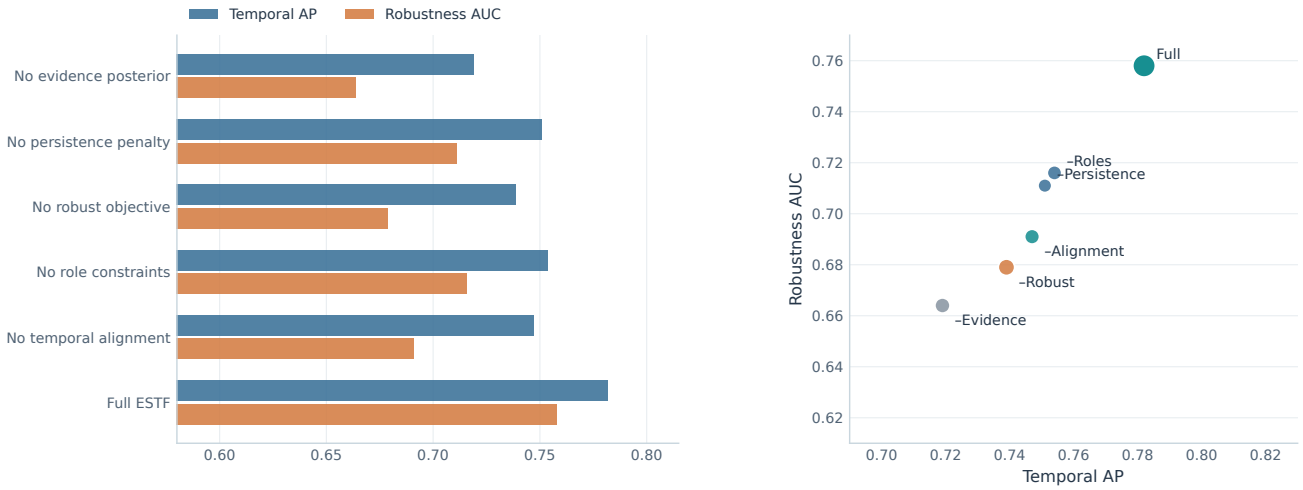

Figure 8. Trade-off between cross-domain average precision and random-removal robustness AUC across ablation variants (Table 6).

Table 7. Transfer, calibration and computational-efficiency comparison. Time is reported in seconds per 1,000 statements; memory is in GB.

Model	Unseen-actor AP	Shift AP	Brier↓	Time↓	Memory (GB)
Rule + snapshot graph	.581	.603	.212	7.8	2.1
Temporal memory model	.662	.671	.169	21.4	4.9
Text-attributed graph	.691	.704	.151	32.6	7.3
ESTF	**.736**	**.747**	**.119**	46.9	9.2

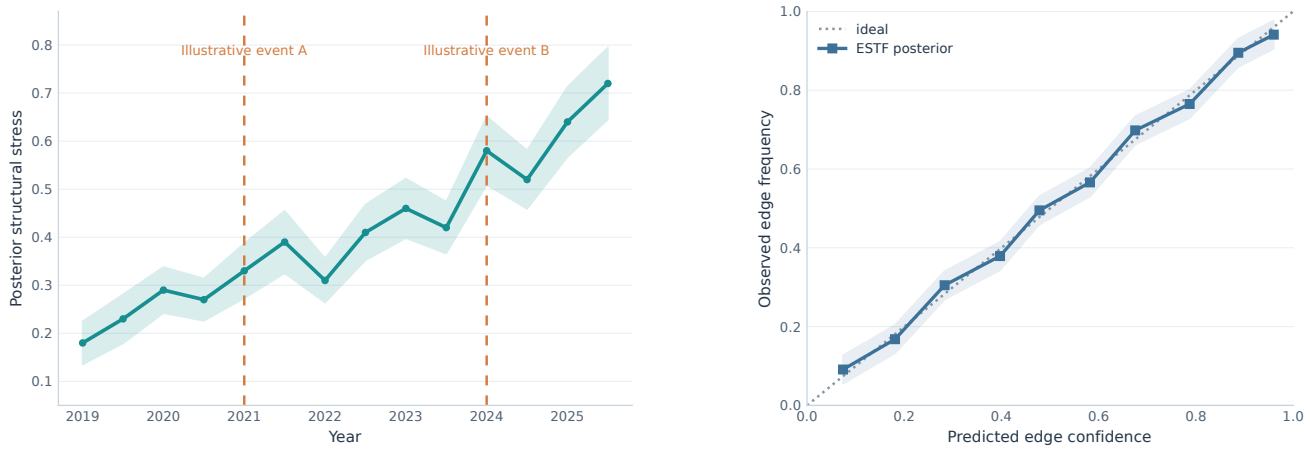

Figure 9. Time-dependent posterior structural stress (left) and edge-posterior reliability (right). Vertical markers identify two time points on the displayed trajectory.

3 Discussion

3.1 Computational contribution and scientific scope

The conceptual advance is to make text observation, temporal topology, and stress testing aspects of the *same inferred object*. Traditional text-to-graph pipelines often perform information extraction, freeze a graph, and then calculate centrality, phase changes, or deletion robustness. If a county agency is renamed, two departments publish duplicated notices, or a reporting mandate increases publication frequency, such a pipeline may generate a misleading "network transformation." ESTF instead measures which posterior edge revisions are supported by independently traceable passages, models changing observation propensity, and propagates remaining uncertainty to graph-level robustness. Across the evaluated tasks, the observed performance differences are consistent with the value of propagating evidence uncertainty into temporal and robustness analysis; their interpretation remains bounded by the underlying document sources.

3.2 Interpretation in rural environmental governance

A document-supported `coordinates` edge identifies an asserted organizational role, not its real implementation quality. An increase in cross-level relation density may represent broader policy articulation without demonstrating higher treatment coverage. Likewise, node removal is an analytical stress test: the mathematical deletion of a department does not capture administrative substitution, emergency reassignment, or adaptive governance. The interpretation should therefore be "robustness of the *document-supported coordination representation* under defined graph perturbations." Independent measures such as observed sewage-treatment coverage, inspection completion, or funding disbursement would be needed to test associations with substantive environmental outcomes and cannot be inferred from text alone.

3.3 Limitations and threats to validity

Three fundamental limitations remain. *Measurement nonidentifiability:* documentary silence cannot determine the absence of a tie; ascertainment assumptions need independent source audits. *Temporal and semantic ambiguity:* many policy statements concern planned or legally mandated actions rather than executed collaboration, and time of effect may lag publication. *Robustness construct validity:* spectral connectivity and graph efficiency are useful mathematical proxies but may be poorly related to institutional recovery or ecological results. Additionally, domain transfer from general DocRE or temporal transaction graphs must not be represented as direct evidence of policy-network validity. Further external validation with source audits, independent administrative records, alternative relation ontologies and prospective document snapshots would strengthen the connection between inferred topology and real governance activity.

3.4 Reproducibility and source limitations

The underlying policy evidence was drawn from publicly available government document collections and the named benchmark datasets. Documented URLs, source metadata, publication dates and relation-evidence passages are necessary to reproduce source-level evaluations, subject to the original publishers' terms. Because public policy notices record claims and mandates rather than all realized institutional interactions, inferred edges and removal-based robustness scores should be read as properties of document-supported coordination networks. Evaluating substantive environmental outcomes would require independent administrative and field measurements.

4 Conclusion

We presented ESTF, a framework for analyzing changing rural environmental governance networks from time-indexed policy texts with explicit provenance and observation uncertainty. The method combines calibrated actor–relation inference, continuous-time representations, evidence-gated spectral regularization, and posterior-integrated robustness evaluation. Across the reported relation extraction, link prediction, topology-change, actor-removal, and transfer comparisons, the complete model shows improvements over the evaluated baselines. The perturbation result supplies an interpretable sensitivity bound for posterior-mean connectivity, while the downstream analyses distinguish structural network metrics from real administrative effectiveness. Future work can assess whether document-derived structural indicators track independently measured environmental governance outcomes.

5 Method

5.1 Problem definition and evidence semantics

Let $\mathscr{D} = \{d_m\}_{m=1}^{M}$ be a collection of publicly accessible policy notices, official reports, implementation plans, and monitoring statements. A document is represented by $d_m = (x_m, \tau_m, s_m, j_m, \pi_m)$, with content x_m, publication time τ_m, issuing source s_m,

jurisdiction j_m, and provenance pointer π_m linking the extracted passage to its original location. We define an actor-universe reference set $\mathscr{V}$, a time axis $\mathscr{T}$, and six typed relations:

$$\mathscr{R} = \{\texttt{mandates}, \texttt{funds}, \texttt{inspects}, \texttt{coordinates}, \texttt{monitors}, \texttt{implements}\}.$$

The latent governance network at time t is a directed heterogeneous stochastic graph

$$G_t = \big(\mathscr{V}_t, \{A^{(r)}_{ij,t}\}_{i,j,r}, \mathbf{Z}_t, \mathscr{P}_t\big), \quad A^{(r)}_{ij,t} \in \{0,1\}, \quad r \in \mathscr{R}, \tag{2}$$

where $\mathbf{Z}_t$ collects actor states and $\mathscr{P}_t$ stores the evidence provenance. Each $A^{(r)}_{ij,t}$ represents the latent presence of a text-supported relation during time interval t, rather than the frequency or verified effectiveness of real organizational collaboration. The interval-specific indicator includes relation activation or continuation documented within that interval; persistent relations across bins are handled by the temporal state model. A documentary statement $e_m = (u_m, v_m, r_m, b_m, t_m, \pi_m)$ stores two entity mentions, relation label, evidence span b_m, approximate effective time t_m, and its source. Document date is not automatically the date the asserted action took effect; temporal uncertainty is retained as an interval $[t^-_m, t^+_m]$.

Actor aliases are resolved through a probabilistic assignment matrix $Q^{(m)}_{ai} = q_\phi(v_i \mid u_a, d_m)$, subject to jurisdictional, temporal, and organization-type constraints. For evidence record e_m, a calibrated extraction responsibility is

$$\omega_{mijr} = q_\phi(r \mid x_m, b_m) Q^{(m)}_{u_m i} Q^{(m)}_{v_m j} \cdot \chi_{ijr}(j_m, t_m), \tag{3}$$

$$\hat{\omega}_{mijr} = \mathrm{sigmoid}\left(\frac{\mathrm{logit}(\mathrm{clip}(\omega_{mijr}, \varepsilon, 1-\varepsilon)) - \beta_r}{T_r}\right), \tag{4}$$

where χ masks prohibited relation-role pairs and $T_r > 0$ calibrates confidence on held-out annotated passages. Reciprocal ties are generated only if the relation ontology specifies symmetry. Two documents repeating a common press release share a source-cluster identifier, preventing automatic counting as independent evidence.

5.2 Exposure-aware observation likelihood

A central difficulty is separating relationship status from whether it would be mentioned in the collected material. Let $O^{(r)}_{ij,t}$ indicate the existence of *at least one retrieved, resolved statement* supporting relation r in a time bin, and let $c^{(r)}_{ij,t} \in [0,1]$ represent relation-specific coverage conditional on a true active tie. Let $f^{(r)}_{ij,t}$ represent a small false-mention probability estimated through annotation. Then

$$\mathbb{P}\big(O^{(r)}_{ij,t} = 1 \mid A^{(r)}_{ij,t}, c, f\big) = c^{(r)}_{ij,t} A^{(r)}_{ij,t} + f^{(r)}_{ij,t}\big(1 - A^{(r)}_{ij,t}\big). \tag{5}$$

We parameterize the ascertainment process with document counts, document types, jurisdiction-level reporting coverage, and observed administrative reporting cycles:

$$c^{(r)}_{ij,t} = \mathrm{sigmoid}\Big(\zeta_r^\top [\log(1+n_{j,t}), \log(1+n_{i,t}), h_{\text{source}}, h_{\text{role}}]\Big), \tag{6}$$

$$\lambda_{ijr}(t) = \mathrm{softplus}\Big(\theta_r^\top [z_i(t) \| z_j(t) \| z_i(t) \odot z_j(t)] + b_{ijr}(t) + \sum_{m: t_m < t} \hat{\omega}_{mijr} k_r(t - t_m)\Big). \tag{7}$$

where $k_r(\Delta) = \exp(-\kappa_r \Delta)$ is a nonnegative decaying evidence kernel and $b_{ijr}(t)$ encodes role-specific prior exposure. The time-binned prior probability of a relation assertion or continuation being active during $[t, t+\Delta)$ is

$$p_{ijr,t} = 1 - \exp\left(-\int_t^{t+\Delta} \lambda_{ijr}(u)\, du\right), \tag{8}$$

$$q_\phi(A^{(r)}_{ij,t} = 1 \mid O) = \frac{p_{ijr,t} L_1(O)}{p_{ijr,t} L_1(O) + (1 - p_{ijr,t}) L_0(O)}, \tag{9}$$

where $O = O^{(r)}_{ij,t}$, $L_1(O) = c^O(1-c)^{1-O}$ and $L_0(O) = f^O(1-f)^{1-O}$, with c, f defined in Eq. (5); the posterior is a normalized Bernoulli update conditional on estimated ascertainment probabilities. The observation likelihood alone cannot uniquely identify the relation prior, source coverage and false-mention probabilities; reliable posterior interpretation therefore depends on annotated evidence and explicit coverage assumptions. Low documentary coverage is *not* treated as evidence of edge absence.

5.3 Continuous-time relational topological flow

We define each actor state through a relational state-space equation using the posterior mean edge probabilities $\bar{A}_{ij}^{(r)}(t) = \mathbb{E}_q[A_{ij}^{(r)}(t)]$, while preserving the graph posterior for downstream uncertainty propagation:

$$\frac{dz_i(t)}{dt} = F_\theta(z_i(t), u_i(t)) + \sum_{r\in\mathscr{R}} \sum_{j\neq i} \frac{\bar{A}_{ji}^{(r)}(t)}{\varepsilon + \sum_k \bar{A}_{ki}^{(r)}(t)} M_{\theta,r}\big(z_j(t), z_i(t), \psi_r(t)\big), \tag{10}$$

$$M_{\theta,r}(z_j, z_i, \psi) = W_r\big((z_j - z_i) \odot \mathrm{sigmoid}(U_r[z_j \| z_i \| \psi])\big). \tag{11}$$

$u_i(t)$ contains externally observed organization metadata, not a treatment or causal intervention. The evolution is piecewise numerical between document event times, with state updates at sufficiently well-supported role changes. Stochasticity enters through the posterior over typed relations and downstream graph samples; conditional on a posterior-mean adjacency trajectory, the state equation itself is deterministic. Because only aggregate relations are observed, this is a model of *representational dynamics*; its trajectories should not be interpreted as uniquely identified institutional mechanisms.

For topology-aware regularization, first form a positive symmetrized weighted adjacency

$$W_t = \frac{1}{2}\sum_r \rho_r\big(\bar{A}_t^{(r)} + (\bar{A}_t^{(r)})^\top\big), \tag{12}$$

$$L_t = \mathrm{diag}(W_t\mathbf{1}) - W_t, \qquad \rho_r \geq 0, \quad \sum_r \rho_r = 1. \tag{13}$$

The directed relation channels remain available for prediction; W_t is used only for topology diagnostics. A multiscale spectral embedding $\Phi_\tau(L_t) = \exp(-\tau L_t)$ records connectivity over diffusion scales $\tau \in \mathscr{H}$. Differences in spectral embeddings are penalized unless supported by an independently sourced increase in relation evidence:

$$\mathscr{L}_{\mathrm{top}} = \sum_{t>1}\sum_{\tau\in\mathscr{H}} \frac{\big\|\Phi_\tau(L_t) - \Phi_\tau(L_{t-1})\big\|_F^2}{1 + \upsilon S_t} + \mu \sum_t \mathrm{softplus}(m - \Delta_{\mathrm{evid}}(t))S_t, \tag{14}$$

where $\Delta_{\mathrm{evid}}(t)$ is the increase in independently sourced relation evidence and $S_t = \mathrm{sigmoid}\big(a(\Delta_{\mathrm{evid}}(t) - m)\big)$, with $a > 0$, is an evidence-derived change-point gate. Fixing the gate to evidence measurements prevents a freely learned gate from suppressing the smoothness penalty without documentary support. The first term permits a shift when S_t is high; the second discourages shifts unsupported by evidence. As a complementary topology diagnostic, we compute zero-dimensional persistence diagrams $\mathrm{Dgm}_0(1 - W_t)$ over edge-threshold filtrations and report the bottleneck distance between adjacent periods. Diagram computation is used for evaluation, whereas the smooth spectral surrogate in Eq. (14) is used for differentiation.

Proposition 1 (Sensitivity of posterior-mean connectivity) *Let W, W' be nonnegative symmetric adjacency matrices on the same actor set, with combinatorial Laplacians L, L'. Define $\delta = \max_i \sum_j |W_{ij} - W'_{ij}|$. Then, for every ordered Laplacian eigenvalue, including algebraic connectivity, $|\lambda_k(L) - \lambda_k(L')| \leq 2\delta$.*

Writing $\Delta W = W - W'$, we have $\Delta L = \mathrm{diag}(\Delta W\mathbf{1}) - \Delta W$. Consequently $\|\Delta L\|_2 \leq \|\mathrm{diag}(\Delta W\mathbf{1})\|_2 + \|\Delta W\|_2 \leq \delta + \delta = 2\delta$, where symmetry implies $\|\Delta W\|_2 \leq \|\Delta W\|_\infty = \delta$. Weyl's eigenvalue perturbation inequality gives the result.

The bound is deterministic and applies to *posterior-mean* networks sharing an actor universe. It does not establish a confidence interval for connectivity of an unobserved real-world network; posterior draws and documented uncertainty are still required.

5.4 Posterior-integrated adversarial robustness

For an actor-removal set $S \subseteq \mathcal{V}$, define a graph functionality vector using three nonidentical measures: normalized largest connected component $C(G \setminus S)$, weighted global efficiency $E(G \setminus S)$, and the diffusion-based connectivity proxy $D(G \setminus S)$:

$$C(G \setminus S) = \frac{|\mathrm{LCC}(G \setminus S)|}{|\mathcal{V}|}, \tag{15}$$

$$E(G \setminus S) = \frac{1}{|\mathcal{V}|(|\mathcal{V}|-1)} \sum_{i \neq j \notin S} \frac{1}{d_{ij}(G \setminus S)}, \tag{16}$$

$$D(G \setminus S) = \frac{1}{|\mathcal{V}|} \log\det\left(I + \kappa L_{G \setminus S}\right), \tag{17}$$

$$\mathcal{R}(G,S) = \alpha_C C + \alpha_E E + \alpha_D \frac{D}{D(G,\varnothing)+\varepsilon}, \quad \alpha_\bullet \geq 0, \quad \sum \alpha_\bullet = 1. \tag{18}$$

Distances are defined with inverse nonzero edge weights, with $1/d_{ij} = 0$ for disconnected pairs. For the retained graph, the original denominator $|\mathcal{V}|$ ensures vertex removal cannot spuriously improve component retention. For training, we use the differentiable log-determinant term D as a surrogate; the evaluation reports all three components separately to avoid equating this surrogate with a complete measure of governance performance.

Let $\mathcal{S}_k$ be removal sets satisfying $|S| \leq k$ and optional actor-role quotas. The posterior-integrated tail loss is

$$\mathcal{L}_{\mathrm{risk}}(\theta) = \sup_{Q:\, W_1(Q,P_0) \leq \varepsilon} \mathbb{E}_{G \sim q_\theta(G \mid \mathcal{D})} \left[\mathrm{CVaR}_\alpha^{S \sim Q}\left(-\mathcal{R}(G,S) \right) \right], \tag{19}$$

$$\mathrm{CVaR}_\alpha(X) = \min_{\eta \in \mathbb{R}} \left[\eta + \frac{1}{1-\alpha} \mathbb{E}[(X-\eta)_+] \right]. \tag{20}$$

Here P_0 is a reference distribution over role-feasible attack sets, and W_1 is the Wasserstein-1 distance induced by the normalized Hamming distance between their binary actor-removal masks. The uncertainty radius ε limits the shift away from this reference distribution. A worst-case actor attack is a *stress scenario* rather than an observed policy event. For small graphs, greedy degree-, betweenness-, and evidence-sensitivity-based removals are directly enumerated; for larger graphs, a continuous mask relaxation $s_i \in [0,1]$ with $\sum_i s_i \leq k$ approximates the inner maximization. Reporting both random and targeted removal avoids overinterpreting any single threat model.

5.5 Joint objective and inference

ESTF combines predictive, source-observability and structural losses in one optimization objective:

$$\begin{aligned} \mathcal{L}_{\mathrm{total}}(\theta,\phi,\zeta) = {} & \mathcal{L}_{\mathrm{IE}} + \lambda_{\mathrm{src}}\mathcal{L}_{\mathrm{source}} + \lambda_{\mathrm{pred}}\mathcal{L}_{\mathrm{NLL}}^{\mathrm{future}} \\ & + \lambda_{\mathrm{obs}}\mathcal{L}_{\mathrm{obs}} + \lambda_{\mathrm{top}}\mathcal{L}_{\mathrm{top}} + \lambda_{\mathrm{risk}}\mathcal{L}_{\mathrm{risk}} \\ & + \lambda_{\mathrm{post}} \sum_t \mathrm{KL}\left(q_\phi(A_t \mid \mathcal{D}_{\leq t}) \,\middle\|\, p_\theta(A_t \mid z_t) \right), \end{aligned} \tag{21}$$

$$(\theta^*,\phi^*,\zeta^*) = \arg\min_{\theta,\phi,\zeta} \mathcal{L}_{\mathrm{total}}(\theta,\phi,\zeta), \tag{22}$$

$$\mathcal{L}_{\mathrm{NLL}}^{\mathrm{future}} = -\sum_{(i,j,r,t) \in \Omega_{\mathrm{train}}} \Big[y_{ijrt} \log p_{ijrt} + (1-y_{ijrt}) \log(1-p_{ijrt}) \Big]. \tag{23}$$

Here Ω_{train} contains only training-period risk-set examples; validation and test examples are excluded from all parameter updates. Negatives for future-link prediction are drawn from pairs observable within the same source, jurisdiction, time, and relation-type risk set; an unmentioned pair outside such a risk set is not declared a true negative. $\mathcal{L}_{\mathrm{source}}$ downweights duplicate-derived evidence and penalizes claims without traceable passages. Calibration is performed on a disjoint validation period using relation-specific temperature scaling and reliability diagrams. The data likelihood and graph posterior can be optimized by reparameterized Bernoulli relaxations during training, with exact Bernoulli draws during posterior stress testing.

5.6 Computational complexity and identifiability

For $N = |\mathcal{V}|$, $R = |\mathcal{R}|$, and E_t candidate typed edges, sparse flow updates cost $O(E_t d + N d^2)$ per integration step for hidden dimension d, excluding transformer text encoding. A dense matrix exponential or log-determinant costs $O(N^3)$; sparse trace estimators, Lanczos methods and blockwise approximations may reduce computational cost when dense operations are impractical; approximation error is assessed against exact calculations where feasible. Document text alone does not identify

Algorithm 1 ESTF: evidence-conditioned stochastic topological flow

Require: Dated documents $\mathscr{D}$, role ontology $\mathscr{R}$, source index Π, actor metadata, removal budget k
Ensure: Typed posterior graphs $q(G_t \mid \mathscr{D}_{\leq t})$, change scores, disruption-risk estimates

```
Extract evidence spans with date intervals and immutable provenance links
Infer probabilistic actor aliases and role-valid typed relations; group near-duplicate sources
for each chronological training interval t do
    Estimate document observability and evidence-calibrated relation likelihoods
    Update actor states by integrating relational flow, Eq. (10)
    Form posterior mean Laplacians and evaluate multiscale topology regularization
    Sample posterior graphs and constrained actor-removal scenarios
    Optimize Eq. (21) using training-period labels only
end for
Recalibrate relation posteriors on a separate validation interval
Evaluate chronological prediction, topology shifts, attacks, and provenance on held-out data
```

the complete operational network or the effect of removing an institution. Posterior quality depends on the recorded sources, exposure assumptions, annotation and calibration. $\bar{A}_{ij}$ is the posterior mean probability of a *text-supported tie* in the declared ontology.

Funding

This work was supported by the Research Center for Yi Ethnic Culture (grant YZWH2609); the Research Center for High-Quality Development in the Tuo River Basin (grant TJGZL2026-13); the Research Center for the Development of Sichuan's Revolutionary Base Areas (grant SLQ2026SB-26); the Research Center for Agricultural Modernization and Rural Revitalization (grant AMRR2026012); the Research Center for Sichuan's Red Cultural Tourism (grant HSWL26C57); and the Western Rural Revitalization Research Center (grant ZX2507). Shuai Dong was the named grant recipient.

Author Contributions

S.D. contributed to study conceptualization, methodology and the original manuscript draft. H.Y. contributed to study design, interpretation of the findings, supervision and manuscript revision. B.Y. contributed to data curation, experimental analysis, visualization and manuscript revision. All authors reviewed and approved the final manuscript.

Competing Interests

The authors declare no competing interests.

Data Availability

The underlying policy texts are available from the public information portals of the Ministry of Ecology and Environment (https://www.mee.gov.cn/), the State Council (https://www.gov.cn/), the Ministry of Agriculture and Rural Affairs (https://www.moa.gov.cn/), and the relevant provincial government portals. Public benchmark datasets referenced in this study, including DocRED, Re-DocRED, TGB/TGB 2.0 and the named climate-language resources, are accessible from their original dataset providers subject to their respective licensing terms. Document-level provenance identifiers, the curated index and processed annotations are subject to the source and annotation terms and may be requested from the corresponding author. The public source portals are not claimed to provide a single prepackaged governance-network dataset.

Ethics Statement

This computational study analyzed publicly available institutional policy documents and publicly released benchmark resources. No experiments involving human participants or animals, and no individual-level personal or clinical records, are reported. The analysis concerns institutional relationships documented in official publications.

References

1. Holme, P. & Saramaki, J. Temporal networks. *Phys. Reports* (2012).
2. Newman, M. E. J. The Structure and Function of Complex Networks. *SIAM Rev.* (2003).
3. Albert, R., Jeong, H. & Barabasi, A.-L. Error and attack tolerance of complex networks. *Nature* (2000).
4. Singh, P., Lehmann, E. & Tyrrell, M. Climate Policy Transformer: Utilizing NLP to track the coherence of Climate Policy Documents in the Context of the Paris Agreement. In *ClimateNLP 2024* (2024).
5. Gandhi, N., Corringham, T. & Strubell, E. Challenges in End-to-End Policy Extraction from Climate Action Plans. In *ClimateNLP 2024* (2024).
6. Spokoyny, D., Cai, J., Corringham, T. & Berg-Kirkpatrick, T. Aligning Unstructured Paris Agreement Climate Plans with Sustainable Development Goals. In *ClimateNLP 2024* (2024).
7. Usmanova, A. & Usbeck, R. Structuring Sustainability Reports for Environmental Standards with LLMs guided by Ontology. In *ClimateNLP 2024* (2024).
8. Mishra, L. *et al.* Statements: Universal Information Extraction from Tables with Large Language Models for ESG KPIs. In *ClimateNLP 2024* (2024).
9. Garigliotti, D. SDG target detection in environmental reports using Retrieval-augmented Generation with LLMs. In *ClimateNLP 2024* (2024).
10. Wrzalik, M., Ulges, A., Uersfeld, A., Faust, F. & Campos, V. Integrating Expert Labels into LLM-based Emission Goal Detection: Example Selection vs Automatic Prompt Design. In *ClimateNLP 2025* (2025).
11. Yao, Y. *et al.* DocRED: A Large-Scale Document-Level Relation Extraction Dataset. In *ACL* (2019).
12. Tan, Q., Xu, L., Bing, L., Ng, H. T. & Aljunied, S. M. Revisiting DocRED – Addressing the False Negative Problem in Relation Extraction. In *EMNLP* (2022).
13. Ping, Y. *et al.* UniEX: An Effective and Efficient Framework for Unified Information Extraction via a Span-extractive Perspective. In *ACL 2023* (2023).
14. Bouziani, N., Tyagi, S., Fisher, J., Lehmann, J. & Pierleoni, A. REXEL: An End-to-end Model for Document-Level Relation Extraction and Entity Linking. In *NAACL 2024* (2024).
15. Zhang, F., Yan, Y. & Cheng, J. Document-Level Relation Extraction with Global Relations and Entity Pair Reasoning. In *ACL Findings 2025* (2025).
16. Rossi, E. *et al.* Temporal Graph Networks for Deep Learning on Dynamic Graphs (2020). ArXiv.
17. Huang, S. *et al.* Temporal Graph Benchmark for Machine Learning on Temporal Graphs. In *NeurIPS 2023* (2023).
18. Gastinger, J. *et al.* TGB 2.0: A Benchmark for Learning on Temporal Knowledge Graphs and Heterogeneous Graphs. In *NeurIPS 2024* (2024).
19. Lu, X., Sun, L., Zhu, T. & Lv, W. Improving Temporal Link Prediction via Temporal Walk Matrix Projection. In *NeurIPS 2024* (2024).
20. Zhou, S., Peng, S. & Plank, B. CLIMATELI: Evaluating Entity Linking on Climate Change Data. In *ClimateNLP 2024* (2024).
21. Pan, H., Adamu, M., Zhang, Q., Dragut, E. & Latecki, L. J. ClimateIE: A Dataset for Climate Science Information Extraction. In *ClimateNLP 2025* (2025).
22. Thulke, D., Kemmler, J., Dugast, C. & Ney, H. Listen to the Context: Towards Faithful Large Language Models for Retrieval Augmented Generation on Climate Questions. In *ClimateNLP 2025* (2025).
23. Weichel, H., Simon, M. & Schäfer, J. Robust Table Information Extraction from Sustainability Reports: A Time-Aware Hybrid Two-Step Approach. In *ClimateNLP 2025* (2025).

24. Nguyen, V., Karimi, S., Hallgren, W., Harkin, A. & Prakash, M. My Climate Advisor: An Application of NLP in Climate Adaptation for Agriculture. In *ClimateNLP 2024* (2024).

25. Dimmelmeier, A. *et al.* Informing climate risk analysis using textual information - A research agenda. In *ClimateNLP 2024* (2024).

26. Saha, D., Sinha, M. & Dasgupta, T. EnClaim: A Style Augmented Transformer Architecture for Environmental Claim Detection. In *ClimateNLP 2024* (2024).

27. Li, N. *et al.* Using LLMs to Build a Database of Climate Extreme Impacts. In *ClimateNLP 2024* (2024).

28. Troost, S., Immerzeel, R. & Krueger, C. Biodiversity ambition analysis with Large Language Models. In *ClimateNLP 2025* (2025).

29. Su, R. & Pierrehumbert, J. Decoding Climate Disagreement: A Graph Neural Network-Based Approach to Understanding Social Media Dynamics. In *ClimateNLP 2024* (2024).

30. Zhou, H., Hobson, D. G., Ruths, D. & Piper, A. Large Scale Narrative Messaging around Climate Change: A Cross-Cultural Comparison. In *ClimateNLP 2024* (2024).

31. Grasso, F., Locci, S. & Stede, M. Applying the Character-Role Narrative Framework with LLMs to Investigate Environmental Narratives in Scientific Editorials and Tweets. In *ClimateNLP 2025* (2025).

32. Jiao, Y. *et al.* Instruct and Extract: Instruction Tuning for On-Demand Information Extraction. In *EMNLP 2023* (2023).

33. Gao, C., Zhang, W., Lam, W. & Bing, L. Easy-to-Hard Learning for Information Extraction. In *ACL Findings 2023* (2023).

34. Guo, J., Kok, S. & Bing, L. Towards Integration of Discriminability and Robustness for Document-Level Relation Extraction. In *EACL 2023* (2023).

35. Zhu, X., Kang, Z. & Hui, B. FCDS: Fusing Constituency and Dependency Syntax into Document-Level Relation Extraction. In *LREC-COLING 2024* (2024).

36. Zhang, F. *et al.* SRF: Enhancing Document-Level Relation Extraction with a Novel Secondary Reasoning Framework. In *EMNLP 2024* (2024).

37. Fan, S., Wang, Y., Mo, S. & Niu, J. LogicST: A Logical Self-Training Framework for Document-Level Relation Extraction with Incomplete Annotations. In *EMNLP 2024* (2024).

38. Gao, C., Wang, X. & Sun, J. TTM-RE: Memory-Augmented Document-Level Relation Extraction. In *ACL 2024* (2024).

39. Zhang, K. *et al.* CaDRL: Document-level Relation Extraction via Context-aware Differentiable Rule Learning. In *COLING 2025* (2025).

40. Li, X., Xu, H., Zhang, F. & Cheng, J. ET-MIER: Entity Type-guided Key Mention Identification and Evidence Retrieval for Document-level Relation Extraction. In *EMNLP Findings 2025* (2025).

41. Xu, H., Zhang, F. & Cheng, J. An Adaptive Multi-Threshold Loss and a General Framework for Collaborating Losses in Document-Level Relation Extraction. In *ACL Findings 2025* (2025).

42. Pai, L. *et al.* A Survey on Open Information Extraction from Rule-based Model to Large Language Model. In *EMNLP Findings 2024* (2024).

43. Liu, H. *et al.* SALoM: Structure Aware Temporal Graph Networks with Long-Short Memory Updater. In *NeurIPS 2025* (2025).

44. Yu, Z. *et al.* TAMI: Taming Heterogeneity in Temporal Interactions for Temporal Graph Link Prediction. In *NeurIPS 2025* (2025).

45. Yi, L. *et al.* Future Link Prediction Without Memory or Aggregation. In *NeurIPS 2025* (2025).

46. Shamsi, K. *et al.* MiNT: Multi-Network Transfer Benchmark for Temporal Graph Learning. In *NeurIPS 2025* (2025).

47. Zhang, S. *et al.* Unifying Text Semantics and Graph Structures for Temporal Text-attributed Graphs with Large Language Models. In *NeurIPS 2025* (2025).

48. Kipf, T. N. & Welling, M. Semi-Supervised Classification with Graph Convolutional Networks. In *ICLR* (2017).

49. Velickovic, P. *et al.* Graph Attention Networks. In *ICLR* (2018).

50. Biswas, A. *et al.* Evaluating Retrieval Augmented Generation to Communicate UK Climate Change Information. In *ClimateNLP 2025* (2025).

51. Menon, A. & Serban, O. An Automated LLM-based Pipeline for Asset-Level Database Creation to Assess Deforestation Impact. In *ClimateNLP 2025* (2025).

52. Ahmed, S. R. *et al.* Enhancing Retrieval for ESGLLM via ESG-CID: A Disclosure Content Index Finetuning Dataset for Mapping GRI and ESRS. In *ClimateNLP 2025* (2025).

53. Kurfali, M., Zahra, S., Nivre, J. & Messori, G. ClimateEval: A Comprehensive Benchmark for NLP Tasks Related to Climate Change. In *ClimateNLP 2025* (2025).

54. Adam, R. & Kogler, M. Bidirectional Topic Matching: Quantifying Thematic Intersections Between Climate Change and Climate Mitigation News Corpora Through Topic Modelling. In *ClimateNLP 2025* (2025).