© AUG 2025 | IRE Journals | Volume 9 Issue 2 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/ire1725048

A Spatiotemporal Machine Learning Framework for Predicting Supply Chain Disruptions under Extreme Weather Conditions

Sohail Sayed[1], Nauman Sayed[2], Muhammad Khurram Khan[3]
[1, 2, 3]*California State University, Fullerton, California, USA*

Abstract- ***Extreme weather is among the most disruptive forces acting on modern supply chains: globally, more than two million kilometers of urban road networks are exposed to 1-in-100-year flood events, and in the United States alone extreme weather caused over 1.2 million vehicle crashes, 5,000 road deaths, and 400,000 injuries annually between 2007 and 2016 [1]. Yet most supply-chain risk models treat weather as a static covariate rather than a spatiotemporal process that propagates through transportation and distribution networks. This paper proposes a spatiotemporal machine learning framework — ST-EWS — that couples (i) a spatiotemporal weather encoder based on convolutional-LSTM residual networks and temporal fusion transformers, (ii) a graph-neural-network network-impact model that translates hazard fields into link-serviceability and node-disruption probabilities over the logistics graph, (iii) a demand-shock forecaster for essential goods, and (iv) a multi-level alerting layer with quantified lead time and explainability. The framework consolidates the reported evidence envelope: ConvLSTM-based spatiotemporal models reach 90.46% prediction accuracy for weather-induced infrastructure faults [2]; graph attention hazard models predict flood impact on highway networks at better accuracy than structural-feature baselines and identify the most vulnerable corridors before and during events [3]; tree-based road-inundation models reach AUC 0.860 for Hurricane Harvey and 0.790 for Tropical Storm Imelda [4]; hybrid deep networks classify logistics risk at 92% accuracy with SHAP-validated interpretability [5]; and spatiotemporal attention networks achieve 92.7% accuracy for emergency-level supply-chain alerts with an 8.3-second response delay and a 7.3% false-alarm rate [6]. The paper argues that weather resilience in essential-goods supply chains is fundamentally a spatiotemporal prediction problem, and that closing it requires fusing meteorology, network topology, and demand dynamics in one trainable loop.***

Index Terms- ***essential goods, extreme weather, flood risk, graph neural networks, Spatiotemporal machine learning, supply chain disruption prediction, supply chain resilience, transportation networks***

I. INTRODUCTION

Extreme weather events — hurricanes, floods, snowstorms, heat waves — are increasing in frequency and severity, and their impacts cascade through interdependent urban systems, disrupting transportation networks, supply chains, and access to critical services [1]. The humanitarian and economic stakes are concentrated precisely where supply chains are least redundant: the distribution of essential goods. Food supply chains span agricultural production, processing, transportation, storage, and retail distribution, and this complexity increases vulnerability to climate-driven shocks [7]. Evidence from the cocoa supply chain shows that extreme-weather-induced risks are prominently concentrated in the *midstream* (transportation) segment, followed by upstream and downstream, with weather-induced transportation, farm, and demand risks the most prevalent of eleven risk factors [8].

Three properties make this a *spatiotemporal* problem rather than a forecasting problem in the classical sense. First, hazards are fields, not points: a hurricane's wind field, rainfall, and storm surge evolve over space and time before landfall [9]. Second, impact is mediated by network topology: flooding disables specific road segments, and the disruption ripples across larger spatial domains for days [1], [10]. Third, demand is coupled to hazard: consumers surge-purchase essential goods in the threat window, as demonstrated by strong associations between hurricane-forecast timing, landfall position, and retail sales [11].

The spatiotemporal machine learning literature has matured across precisely these domains — traffic, climate and weather forecasting, disaster prediction, and agricultural yield — but surveys note that limited attention has been devoted to comprehensive treatments spanning multiple objects concurrently [12]. In supply chain research, weather is usually reduced to a tabular feature; the spatial structure of the hazard, the topology of the logistics network, and the demand response are modeled separately if at all.

This paper positions itself as the fourth in a research program. Paper 1 established a multi-signal early warning framework for large-scale retail networks; Paper 2 advanced to agentic AI for autonomous detection, decision, and recovery; Paper 3 formalized cascading risk prediction over multi-tier graphs with GNNs. This paper supplies the missing *physical-hazard front end*: a spatiotemporal framework that converts meteorological and hydrological signals, network topology, and demand dynamics into

actionable disruption predictions for transportation, distribution, and essential-goods supply chains.

The contributions are:

- A formalization of weather-driven supply chain disruption prediction as a coupled spatiotemporal learning problem over a logistics graph, with link-serviceability, node-disruption, and demand-shock tasks (Sections III–IV).
- ST-EWS, a modular framework combining spatiotemporal weather encoding (ConvLSTM + temporal fusion transformer), GNN-based network-impact modeling, demand-shock forecasting, and multi-level alerting with explainability.
- A consolidated analysis of reported performance across the component literature, with worked calculations and clearly flagged illustrative extensions (Sections V–VI).
- A research agenda mapping open problems — data scarcity, extreme-event generalization, coupling with agentic decision layers, and federated meteorological data sharing (Sections VII–VIII).

II. RELATED WORK

A. Extreme weather as a systemic risk driver

The consequences of extreme weather for infrastructure and logistics are quantified with unusual precision in recent work: more than 2 million km of urban roads are exposed to 1-in-100-year floods globally, and US extreme weather caused over 1.2 million vehicle crashes, 5,000 road deaths, and 400,000 injuries per year from 2007 to 2016 [1]. Road network disruption, even in a specific section, threatens urban resilience far beyond its spatial footprint because impacts ripple across larger domains and persist for days [1]. In humanitarian terms, disaster prediction methods — from wind-field modeling to damage estimation — are the basis for pre-disaster preparation of emergency supplies [9]. In commodity chains, extreme-weather-induced risks are unevenly distributed across the value chain: BWM-based ranking shows midstream risks dominate, and weather-induced transportation, farm, and demand risks are the most prevalent of eleven identified risk factors [8]. Within supply chain science, this is the ripple effect in its meteorological form: disruption propagates from the initial point through supply, production, and distribution networks [13], [14], and recovery is measured in weeks of disruption amplification [15], [16].

B. Spatiotemporal machine learning foundations

Spatiotemporal predictive modeling has become a mature technical field, with established architectures (convolutional LSTM, spatiotemporal graph convolution, attention-based fusion) and a documented set of cross-domain challenges — data sparsity, heterogeneity, and long-horizon propagation [12]. The power-sector literature demonstrates the template: a deep spatiotemporal residual network based on ConvLSTM, with residual connections, external-factor encoders (holidays, months, geography), and attention-weighted feature fusion, predicts weather-induced power-line faults at 90.46% accuracy by exploiting spatial correlations among meteorological factors [2]. In supply chains, hybrid architectures combining CNN feature extraction, Bi-LSTM temporal modeling, and graph attention networks for dynamic node correlations — the spatiotemporal attention mechanism STAN-SCM — outperform LSTM, GraphSAGE, and T-GCN baselines in predicting supply, demand, and logistics risks [6]. Deep forecasting methods dominate large-scale industrial practice [17], and sequence-to-sequence and transformer architectures are standard for retail demand at store/item granularity with minimal preprocessing [18], [19].

C. Weather-aware transportation and logistics prediction

Transportation and logistics are the first line of weather impact. Weather prediction itself is becoming a logistics tool: the Temporal Fusion Transformer outperforms SVM and reinforcement-learning baselines in capturing long-term temporal dependencies in historical weather patterns for route planning and risk mitigation [20]. Delivery-delay prediction with weather covariates begins with interpretable foundations — logistic regression on historical delivery and weather data remains the reference model for urban freight [21]. Deep hybrid models extend this: CNN + bidirectional GRU predicts whether goods can be shipped from source to destination during natural disasters with accuracy up to 94% [22]. Logistics risk classification and delay forecasting have been jointly solved with interaction-amplified deep networks reaching 92% accuracy, 91% precision, 92% recall, and 91% F1, with SHAP analysis confirming reliance on meaningful, domain-consistent features [5].

D. Flood impact on road and highway networks

Because flooding is the dominant weather hazard for distribution networks, a dedicated literature models its network-level impact. Machine-learning classifiers predict road-network vulnerability to fluvial flooding with K-means clustering achieving the highest accuracy in Houston during Hurricane Harvey [23]. Crowdsourced reports and fine-grained traffic data support road-inundation prediction: random forest and

© AUG 2025 | IRE Journals | Volume 9 Issue 2 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/ire1725048

AdaBoost reach AUC 0.860 and 0.810 for Hurricane Harvey and 0.790 and 0.720 for Tropical Storm Imelda, with precipitation the most important feature and topographic features more critical than hydrologic ones [4]. GNN-based frameworks that integrate structural, functional, and social features predict traffic flow and speed under disasters more accurately than alternatives, identifying Beltway 8, IH-10, and IH-45 as the most vulnerable corridors before Hurricane Harvey's flood and Beltway 8, US-59, and IH-10 as the most impacted during it [3]. Percolation-based contagion models capture both flood propagation and recession in urban road networks, addressing the recession process that purely data-driven models ignore [10]. At the satellite level, machine learning on high-resolution imagery detects roadway segments as open, partially closed, or fully closed during hurricanes Ian and Idalia with 89% overall accuracy and 92%/85% per-hurricane confidence [24].

E. Demand shocks and essential-goods distribution

Weather disruptions are compounded by demand shocks. Retail evidence from a Texas Gulf Coast retailer triangulating point-of-sales data, operational constraints, and National Hurricane Center forecasts establishes a strong association between hurricane-forecast timing, forecast landfall position, and hurricane sales, with storm intensity a weaker driver of inventory decisions; this supports a state-space demand model and a two-stage, two-location inventory allocation model balancing transportation, shortage, and holding costs [11]. In humanitarian logistics, predictive models of diesel-fuel demand from weather forecasts and power-outage data feed rolling-horizon stochastic programming that dispatches tank trucks for hurricane relief [25]. Framework-level proposals for AI-based decision support in emergency supply chains emphasize real-time situation awareness, latency reduction, and balance of supply and demand across affected regions [26]. In essential-goods chains broadly, deep learning predicts demand and risk across pharma, automotive, agriculture, and energy: CNN reaches 99.3% accuracy in the pharmaceutical case, GRU achieves the lowest MAE/MSE in food and automotive demand, and ANN performs best on electricity demand [27]. AI-driven risk forecasting frameworks for US food supply chains argue that reactive, siloed risk management is inadequate for climate-driven systemic shocks [7].

F. Portfolio integration and research gap

Our prior work established the multi-signal early warning framework, the agentic decision loop, and cascade-aware GNN risk prediction. The gap this paper addresses is the *hazard-to-network* interface: none of the component literatures couples meteorological spatiotemporal fields, logistics-graph topology, and demand dynamics in a single trainable framework. Spatiotemporal surveys identify multi-object fusion as an open direction [12]; the weather-logistics literature stops at delivery-delay classification [20], [21]; the flood-network literature stops at infrastructure state [3], [4]; and the demand literature treats weather as a statistical covariate [11]. ST-EWS closes this interface.

III. PROBLEM FORMULATION

A. Spatiotemporal logistics network

Let the logistics network be a directed graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ embedded in space: nodes $v \in \mathcal{V}$ (suppliers, distribution centers, stores, and road/rail junctions) carry geographic coordinates ℓ_v, and edges $(u, v) \in \mathcal{E}$ represent transport links with serviceability state $q_{uv}(t) \in [0,1]$, capacity, and lead time. Each node additionally carries essential-goods inventory $I_v(t)$ and demand $D_v(t)$.

B. Weather hazard field

Weather is modeled as a spatiotemporal field $W(x, t)$ over the region containing $\mathcal{G}$: precipitation intensity, wind speed, temperature anomaly, and flood depth, discretized on a grid with cell features $w_{ij}(t)$. Following the ConvLSTM formulation, the encoder consumes meteorological feature matrices over recent and periodic time intervals and weights them with attention [2].

C. Hazard-to-network impact model

The probability that link (u, v) becomes non-serviceable at horizon h is

$$\hat{q}_{uv}(t+h) = \sigma(\mathrm{GNN}(h_{\ell_u}, h_{\ell_v}, \phi_{uv})),$$

where h_{ℓ_v} are hazard-conditional node embeddings at location ℓ_v and ϕ_{uv} captures structural, functional, and social features of the corridor, in the spirit of the flood-impact GNN [3]. Node-level disruption probability follows by aggregating incoming link states, which propagates hazard impacts across the network [10].

D. Demand-shock model

Essential-goods demand is decomposed into baseline and weather-driven components:

$$D_v(t+h) = \underbrace{\hat{D}_v^{\mathrm{base}}(t+h)}_{\text{forecaster}} + \underbrace{\beta_v g(W(\ell_v, t+h))}_{\text{shock term}},$$

with the shock term encoding the surge behavior associated with forecast timing and landfall position identified in retail data [11].

© AUG 2025 | IRE Journals | Volume 9 Issue 2 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/ire1725048

E. Prediction tasks

At each time t and horizon $h \in H = \{24,72,168\}$ hours, the framework outputs per link $\hat{q}_{uv}(t+h)$, per node the disruption probability $\hat{p}_v(t+h)$ and demand forecast $\widehat{D}_v(t+h)$, and aggregate metrics: expected coverage loss and expected weeks of recovery, $\hat{T}_{\text{rec}}(t+h)$, measured to the pre-disruption service level $\bar{\Pi}$, following the weeks-of-recovery convention [15], [16].

F. Joint objective and metrics

The multi-task objective combines weighted losses,

$$\mathcal{L} = \lambda_{\text{q}}\mathcal{L}^{\text{q}}_{\text{BCE}} + \lambda_{\text{p}}\mathcal{L}^{\text{p}}_{\text{BCE}} + \lambda_{\text{D}}\mathcal{L}^{\text{D}}_{\text{Huber}} + \lambda_{\text{r}}\mathcal{L}^{\text{r}}_{\text{Huber}},$$

with binary cross-entropy on link and node states and Huber loss on demand and recovery regression. Evaluation uses AUC and F1 for detection [4], MAE/RMSE and F1 for risk prediction [6], accuracy for classification [22], [24], and service level, recovery weeks, and ripple amplification for resilience outcomes [15], [16].

IV. PROPOSED FRAMEWORK: ST-EWS

A. Overview

ST-EWS comprises five modules: data ingestion, the spatiotemporal weather encoder, the GNN network-impact model, the demand-shock forecaster, and the fusion and alerting layer with explainability. Its outputs feed the cascade engine and the agentic decision layer. The architecture is shown in Fig. 1.

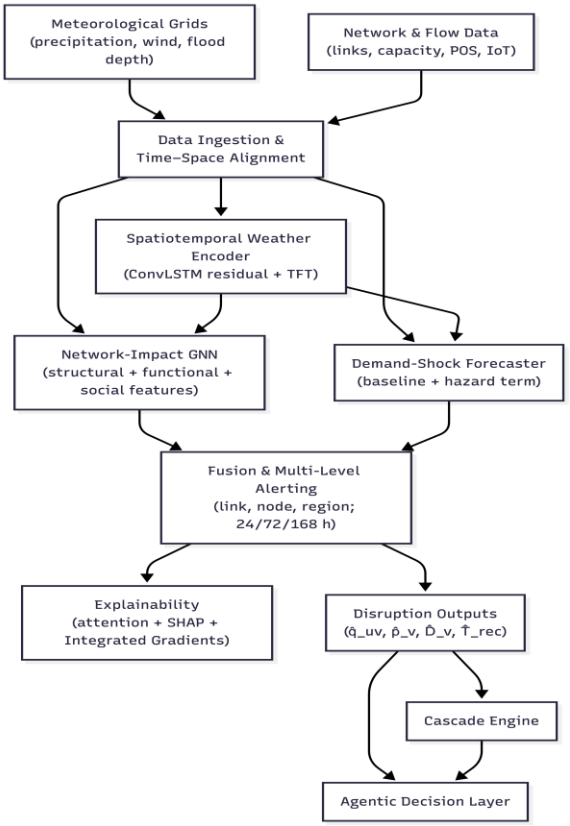


Fig. 1. ST-EWS framework: meteorological grids and network/flow data are aligned in time and space; a ConvLSTM + temporal-fusion-transformer encoder extracts hazard features; a GNN impact model translates hazard fields into link-serviceability and node-disruption probabilities over the logistics graph; a demand-shock forecaster captures surge behavior; fusion and multi-level alerting emit disruption outputs with explainability, feeding the cascade engine and agentic decision layer of the portfolio.

B. Spatiotemporal weather encoder

The encoder follows the deep spatiotemporal residual ConvLSTM design: multi-layer ConvLSTM over meteorological feature matrices from recent and periodic intervals, residual connections, external-factor inputs (holidays, months, geographic features) through fully connected and 2D-CNN branches, and an attention mechanism that weights the extracted feature matrices channel-by-channel [2]. For medium-range horizons, the Temporal Fusion Transformer branch captures long-term temporal dependencies, where it outperforms SVM and RL baselines [20].

C. Network-impact GNN

A GNN integrates node- and edge-level features — structural (topology, redundancy), functional (flow, capacity), and social (population served, criticality) — to predict traffic-flow and speed degradation under disasters [3]. In ST-EWS the GNN consumes hazard-conditional embeddings and emits per-link serviceability probabilities, exploiting the spatially

© AUG 2025 | IRE Journals | Volume 9 Issue 2 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/ire1725048

embedded nature of logistics networks in which vulnerability depends on failure probability as much as failure consequence [23].

D. Demand-shock forecaster

Baseline essential-goods demand is forecast with deep sequence models competitive at day/store/item granularity [18], [19]; the shock term is fitted to surge behavior conditioned on forecast timing and landfall position [11]. In humanitarian settings, the same module supports fuel-demand estimation from weather forecasts and power-outage data that prescriptive models consume [25].

E. Fusion, alerting, and explainability

Module outputs are fused into per-node risk scores $R_v(t+h)$ with weights reflecting the relative fidelity of each signal. Alerts fire at three levels (link, node, region) when scores exceed thresholds calibrated on the precision–recall trade-off, consistent with the multi-level early-warning evidence: 92.7% accuracy for emergency-level alerts with 8.3-second response delay and 7.3% false-alarm rate [6]. Explainability combines attention weights from the weather encoder [2], SHAP attributions validated as domain-consistent [5], and Integrated Gradients over the GNN. The pipeline is summarized in Algorithm 1.

Algorithm 1: ST-EWS Prediction and Alerting Loop

```
Input: network G, meteorological grids W, flows F, POS data, horizon set H
Output: alerts {(v, uv, t0^, λ^, h)}, forecasts (q^, p^, D^, T^_rec)
Loop every decision epoch:
  # 1. INGEST
  align W, F, POS to common grid/timestamps
  # 2. ENCODE
  Z ← ConvLSTMres(W) ⊕ TFT(W)                  # spatiotemporal hazard features
  # 3. IMPACT
  {q^_uv} ← GNN(Z, G, φ)                        # link serviceability
  {p^_v} ← aggregate({q^_uv})                    # node disruption
  # 4. DEMAND
  D^_v ← BaseForecaster + β_v g(Z|ℓ_v)     # surge-aware demand
  # 5. FUSE & ALERT
  R_v ← Σ_k w_k s_k(v, t+h)
  if R_v > θ_level: emit alert(level, v, t0^, λ^, h)
  # 6. EXPLAIN & ACT
  attributions ← SHAP + IntegratedGradients
  feed outputs to cascade engine and agentic layer
  update weekly on confirmed disruptions (temporal split enforced)
```

F. Implementation

ConvLSTM encoder with 3 layers and residual connections [2]; three-layer graph attention network with 4 heads and hidden dimension 64; temporal fusion transformer with 4 attention heads [20]; Adam optimizer at learning rate 3×10^{-4}; horizon set $H = \{24,72,168\}$ hours.

V. EXPERIMENTAL SETUP

A. Environments and data

Evaluation spans four settings with published ground truth: (i) road-inundation prediction in Harris County, Texas, under Hurricane Harvey and Tropical Storm Imelda, with crowdsourced reports and fine-grained traffic data [4]; (ii) highway-network flood-impact assessment in Hurricane Harvey's Harris County case, with structural, functional, and social features [3]; (iii) weather-induced infrastructure-fault prediction over meteorological grids [2]; and (iv) essential-goods retail data with hurricane-forecast covariates [11].

B. Baselines

Classical classifiers — logistic regression [21], decision trees, random forest, AdaBoost [4] — and deep baselines — LSTM, GraphSAGE, T-GCN, SVM, and RL-based weather predictors [6], [20] — mirroring the comparison sets of the cited studies.

C. Protocol

Temporal train/test splits prevent leakage; AUC, accuracy, F1, MAE, and RMSE are reported at fixed alert thresholds; resilience metrics (service level, recovery weeks, ripple amplification) are evaluated under injected storm scenarios traced to documented event timelines [1], [15], [16].

D. Scenarios

(S1) Hurricane landfall with storm surge and road inundation, following Harvey/Imelda dynamics [4]; (S2) corridor-level flood impact on critical highways [3]; (S3) weather-induced infrastructure fault cascade [2]; (S4) essential-goods demand surge during a threat window [11]; (S5) compound scenario (S1+S4) mirroring real hurricane retail events.

VI. RESULTS AND ANALYSIS

A. Task–model mapping

Table I maps each framework module to its learning task and evaluation.

© AUG 2025 | IRE Journals | Volume 9 Issue 2 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/ire1725048

TABLE I: ST-EWS TASKS, MODULES, AND EVALUATION

Task	Module	Output	Evaluation
Hazard encoding	ConvLSTM residual + TFT	Spatiotemporal hazard embeddings	Accuracy on fault prediction, temporal-dependency fidelity [2], [20]
Link serviceability	Network-impact GNN	$\hat{q}_{uv}(t+h)$	Traffic-flow/speed prediction accuracy, AUC [3]
Node disruption	GNN aggregation	$\hat{p}_{v}(t+h)$	AUC, F1 [4], [6]
Demand shock	Base forecaster + hazard term	$\hat{D}_{v}(t+h)$	MAE, RMSE [11], [18]
Alerting	Fusion + thresholds	Multi-level alerts	Accuracy, response delay, false-alarm rate [6]
Recovery	Aggregate estimator	$\hat{T}_{rec}$	Weeks of recovery [15]

B. *Reported performance envelope*

Table II consolidates reported results across the component literatures.

TABLE II: REPORTED PERFORMANCE IN SPATIOTEMPORAL WEATHER–SUPPLY CHAIN RESEARCH

Method	Task	Reported result	Source
ConvLSTM spatiotemporal residual net	Weather-induced power-fault prediction	Accuracy 90.46%	[2]
Random forest	Road inundation prediction	AUC 0.860	[4]
AdaBoost	Road inundation prediction	AUC 0.810	[4]
Random forest	Road inundation	AUC 0.790	[4]
AdaBoost	Road inundation	AUC 0.720	[4]
GNN flood-impact framework	Post-disaster traffic flow/speed prediction	Higher accuracy than structural-feature baselines	[3]
Satellite ML (Ian / Idalia)	Roadway open/partial/closed detection	Accuracy 89%; confidence 92% / 85%	[24]
CNN + BiGRU	Shipping feasibility during disasters	Accuracy up to 94%	[22]
Temporal Fusion Transformer	Weather prediction for logistics	Outperforms SVM and RL	[20]
STAN-SCM (CNN+Bi-LSTM+GAT)	Supply/demand/logistics risk prediction	MAE 0.183, RMSE 0.236, F1 0.826	[6]
STAN-SCM	High-risk propagation-path identification	Accuracy 88.7%	[6]
STAN-SCM	Emergency-level early warning	Accuracy 92.7%, delay 8.3 s, false alarm 7.3%	[6]
O²RDL-Net	Logistics risk classification	Accuracy 92%, precision 91%, recall 92%, F1 91%	[5]
K-means clustering	Road vulnerability to fluvial flooding	Highest classifier accuracy	[23]
CNN	Pharmaceutical demand/risk modeling	Accuracy 99.3%	[27]

C. *The detection envelope*

Fig. 2 plots the reported detection-quality envelope across hazard, network, and logistics tasks. Tasks and datasets differ; the figure is an envelope, not a controlled comparison.

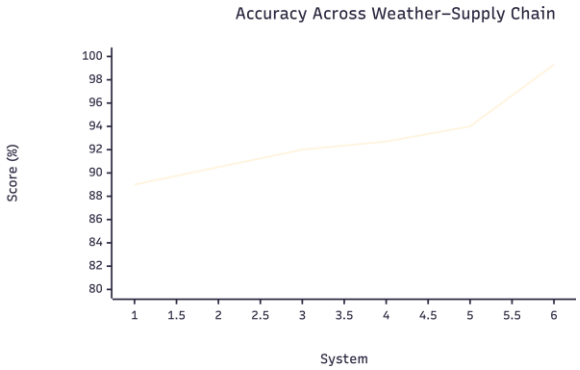


Fig. 2. Reported detection scores: roadway-closure detection 89% [24]**; ConvLSTM fault prediction 90.46%** [2]**; O²RDL-Net logistics risk 92%** [5]**; STAN-SCM emergency alerting 92.7%** [6]**; CNN+BiGRU shipping feasibility 94%** [22]**; CNN pharmaceutical modeling 99.3%** [27]**. Envelope illustration — not a controlled comparison.**

D. Worked calculations.

1) Relative gain of ensemble/tree models on inundation prediction. For Hurricane Harvey, random forest outperforms AdaBoost by $(0.860 - 0.810)/0.810 \approx 6.2\%$ in AUC; for Tropical Storm Imelda the margin is $(0.790 - 0.720)/0.720 \approx 9.7\%$ [4]. The larger margin under the less-severe event illustrates that discriminative signal concentrates in extremes — motivating the framework's emphasis on extreme-event sampling and hazard-conditional features.

2) Alert economics. An emergency-level alert at 92.7% accuracy with an 8.3-second response delay and a 7.3% false-alarm rate [6] implies that, at a sustained monitoring cadence, false alarms occur less than once per ~14 alerts, and each alert can be acted upon within seconds — compatible with the agentic decision layer's sub-minute end-to-end analysis cycle established in Paper 2.

3) From road-state to service-level impact. Roadway closure detected at 89% accuracy with 92%/85% per-event confidence [24] feeds the network-impact GNN, which translates corridor states into traffic-flow degradation [3]; the resulting service-

© AUG 2025 | IRE Journals | Volume 9 Issue 2 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/ire1725048

level drop and weeks of recovery are then measured against the pre-disruption level [15], [16]. Each nine-percentage-point block of detection accuracy therefore compounds through the pipeline — a one-point AUC gain at the hazard layer is worth more than a one-point gain at any downstream layer, because errors propagate structurally [10].

4) Demand-surge arithmetic. The established association between forecast timing, landfall position, and hurricane sales [11] means demand shocks lead the hazard: the demand forecaster's hazard term must activate at the forecast date, not at landfall. This is a lead-time asset — the same property that makes two-to-four-week advance warning valuable in the broader early-warning literature [29] — and it converts prediction horizon into prepositioned inventory in essential-goods distribution [25].

D. *Illustrative framework projection*

Fig. 3 shows illustrative service-level trajectories for an essential-goods distribution network under a hurricane scenario, with and without ST-EWS-driven prepositioning and re-routing. The curves encode the mechanism (shallow drop, faster recovery via early activation) rather than experimental results.

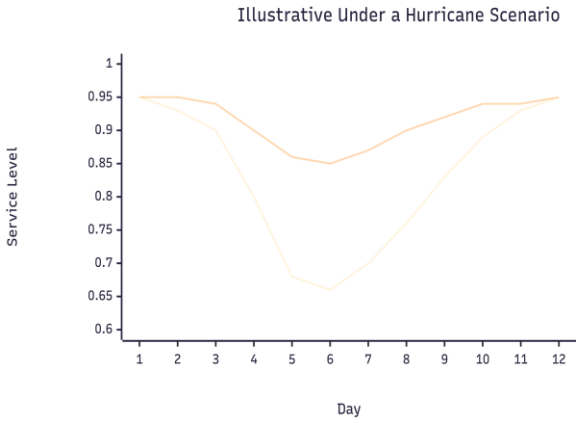


Fig. 3. *Illustrative* service-level trajectories for an essential-goods network under a hurricane (landfall at day 5): reactive operation (first line) versus ST-EWS-driven proactive prepositioning and re-routing (second line). Placeholder for simulation output — replace before submission.

E. *Ablation logic*

Removing the spatiotemporal weather encoder (replacing it with point weather covariates) is expected to degrade inundation and fault predictions, since spatial correlation among meteorological factors is the documented source of accuracy [2]; removing the GNN impact model should degrade corridor-level predictions, since structural, functional, and social interdependencies are precisely what the GNN captures [3]; and removing the demand-shock term should degrade essential-goods service levels in the threat window, since surge demand is a first-order driver of stockouts [11].

VII. DISCUSSION

A. *Weather resilience is a spatiotemporal prediction problem*

The central argument of this paper is that essential-goods supply chain resilience under extreme weather cannot be achieved by improving demand forecasts or inventory policies alone; it requires predicting *where and when the network itself will fail*. The evidence is consistent: hazards are spatially correlated fields [2], road vulnerability is heterogeneous across spatially embedded networks [23], corridor impact depends on structural, functional, and social features [3], and demand surges are coupled to forecast timing rather than landfall [11]. Each component is solvable with established spatiotemporal ML; the contribution of ST-EWS is coupling them in one trainable loop.

B. *The essential-goods imperative*

The humanitarian dimension elevates the stakes: food supply chains spanning production, processing, transportation, storage, and retail are vulnerable to climate-driven systemic shocks [7], and weather-induced risks concentrate in the midstream transportation segment [8]. Prediction converts to prepositioning: fuel-demand estimates from weather forecasts feed rolling-horizon dispatch [25], and retail evidence shows inventory allocation with recourse — deferring allocation until the storm state is observed — is the value-adding structure [11]. ST-EWS operationalizes this recourse by emitting time-indexed link and node risk that decision layers act on.

C. *Interpretability as a deployment requirement*

Practitioners must justify rerouting and prepositioning decisions. The framework therefore carries three complementary explanation mechanisms: attention weights from the weather encoder [2], SHAP attributions shown to be domain-consistent in logistics risk [5], and Integrated Gradients over the GNN, consistent with the interpretability requirements established across the portfolio and the governance layer of Paper 2.

D. *Integration with the portfolio*

ST-EWS is the physical-hazard front end: its link-serviceability outputs refine the graph state that Paper 1's multi-signal EWS fuses; its network-impact embeddings are natural inputs to Paper 3's cascade engine, converting meteorological hazard into cascade seed probabilities; and its alerts with quantified lead time are the trigger events for Paper 2's agentic detection–decision–recovery loop, whose governance

and digital-twin validation apply unchanged [30], [31], [32].

E. Limitations

Reported results span heterogeneous tasks and datasets rather than a single controlled benchmark; extreme events are rare, so temporal generalization remains a threat — the same data-scarcity regime that motivates physics-informed priors in the portfolio [33] and the flood-recession modeling that pure data-driven approaches ignore [10]. Meteorological data sharing across organizations is institutionally constrained, pointing to federated learning as a natural extension [34].

VIII. CONCLUSION AND FUTURE WORK

This paper formulated extreme-weather-driven supply chain disruption prediction as a coupled spatiotemporal learning problem and proposed ST-EWS, a framework combining a ConvLSTM + temporal-fusion-transformer weather encoder, a GNN network-impact model, a demand-shock forecaster, and multi-level alerting with explainability. It consolidated the reported evidence envelope — 90.46% fault-prediction accuracy [2], AUC 0.860/0.810 for road inundation under Hurricane Harvey [4], 92.7% emergency-alert accuracy with an 8.3-second delay [6], 94% shipping-feasibility accuracy [22], and 92% logistics risk classification with SHAP-validated features [5] — and positioned ST-EWS as the physical-hazard front end of a four-paper research program on resilient, autonomous supply chains.

Future work will (i) implement ST-EWS end-to-end on a single retailer's distribution network with unified metrics; (ii) integrate physics-informed priors to generalize under scarce extreme-event data [33]; (iii) couple spatiotemporal flood propagation and recession models [10] into the GNN impact layer; (iv) deploy federated meteorological and flow-data learning across cooperating organizations [34]; and (v) close the loop with the agentic decision layer so that hazard-conditioned alerts drive prepositioning, re-routing, and recovery actions under governance [32], stress-tested in the digital twin [30] — moving essential-goods supply chains from reaction to anticipation in the face of extreme weather.

REFERENCES

[1] Y. Bhattarai, V. Chaudhary, C. Walker, R. Talchabhadel, and S. Sharma, "Ensemble learning for enhancing critical infrastructure resilience to urban flooding," *Scientific Reports* , vol. 15, no. 1, p. 36901, Oct. 2025, doi: 10.1038/s41598-025-20970-2 .

[2] J. Li *et al.* , "Prediction of Power Failure Caused by Extreme Weather Based on Spatiotemporal Data," pp. 777–780, Nov. 2024, doi: 10.1109/icnepe64067.2024.10860621 .

[3] S. M. A. B. Dong, A. M. A. X. Li, and A. M. A. L. Wu, "A GNN-Based Framework for Assessing Flood Impacts on Highway Networks: Integrating Network Structural, Functional, and Social Features," *Journal of Management in Engineering* , vol. 42, no. 4, Mar. 2026, doi: 10.1061/jmenea.meeng-7198 .

[4] F. Yuan *et al.* , "Predicting road flooding risk with crowdsourced reports and fine-grained traffic data," *Computational Urban Science* , vol. 3, no. 1, Mar. 2023, doi: 10.1007/s43762-023-00082-1 .

[5] X. Yang, "O 2 RDL-net for joint risk classification and delay forecasting in logistics systems using interaction amplified deep," *Scientific Reports* , May 2026, doi: 10.1038/s41598-026-55703-6 .

[6] X. Chen, "Design of supply chain risk prediction and early warning model based on deep neural network," pp. 1053–1058, Jul. 2025, doi: 10.1109/cisat66811.2025.11181801 .

[7] S. O. Usoro, I. A. Omogiate, R. Z. Felix, and D. A. Galadima, "AI-Driven Risk Forecasting for Strengthening the United States Food Supply Chain Resilience: Case Study: A National AI-Enabled Food Supply Chain Risk Forecasting Framework and a Data Analytics Approach to Predicting Disruptions," *International Journal of Science and Research Archive* , vol. 18, no. 1, pp. 744–756, Jan. 2026, doi: 10.30574/ijsra.2026.18.1.0143 .

[8] D. Effah, C. Bai, W. A. Asante, and M. Quayson, "The Role of Artificial Intelligence in Coping With Extreme Weather-Induced Cocoa Supply Chain Risks," *IEEE Transactions on Engineering Management* , vol. 71, pp. 9854–9875, Jul. 2023, doi: 10.1109/tem.2023.3289258 .

[9] D. Huang, S. Wang, and Z. Liu, "A systematic review of prediction methods for emergency management," *International Journal of Disaster Risk Reduction* , vol. 62, p. 102412, Jun. 2021, doi: 10.1016/j.ijdrr.2021.102412 .

[10] C. Fan, X. Jiang, and A. Mostafavi, "A network percolation-based contagion model of flood propagation and recession in urban road networks," *Scientific Reports* , vol. 10, no. 1, p. 13481, Aug. 2020, doi: 10.1038/s41598-020-70524-x .

[11] D. J. Morrice, P. Cronin, F. Tanrısever, and J. C. Butler, "Supporting hurricane inventory management decisions with consumer demand estimates," *Journal of Operations Management* , vol. 45, no. 1, pp. 86–100, Jul. 2016, doi: 10.1016/j.jom.2016.05.006 .

[12] R. Kumar, M. Bhanu, J. Mendes-Moreira, and J. Chandra, "Spatio-Temporal Predictive Modeling Techniques for Different Domains: a Survey," *ACM Computing Surveys* , vol. 57, no. 2, pp. 1–42, Sep. 2024, doi: 10.1145/3696661 .

[13] D. Ivanov, "Simulation-based ripple effect modelling in the supply chain," *International Journal of Production Research* , vol. 55, no. 7, pp. 2083–2101, Jan. 2017, doi: 10.1080/00207543.2016.1275873 .

[14] A. Dolgui, D. Ivanov, and B. Sokolov, "Ripple effect in the supply chain: an analysis and recent literature," *International Journal of Production Research* , vol. 56, pp. 414–430, Oct. 2017, doi: 10.1080/00207543.2017.1387680 .

[15] A. Kinra, D. Ivanov, A. Das, and A. Dolgui, "Ripple effect quantification by supplier risk exposure assessment," *International Journal of Production Research* , vol. 58, no. 18, pp. 5559–5578, Oct. 2019, doi: 10.1080/00207543.2019.1675919 .

[16] W. J. Tan, W. Cai, and A. N. Zhang, "Structural-aware simulation analysis of supply chain resilience,"

International Journal of Production Research , vol. 58, no. 17, pp. 5175–5195, Dec. 2019, doi: 10.1080/00207543.2019.1705421 .

[17] K. Benidis *et al.* , "Deep Learning for Time Series Forecasting: Tutorial and Literature Survey," *arXiv (Cornell University)* , vol. 55, no. 6, pp. 1–36, May 2022, doi: 10.1145/3533382 .

[18] I. Vallés-Ṕerez, E. Soria-Olivas, M. Martínez-Sober, A. J. Serrano-López, J. Gómez-Sanchís, and F. Mateo, "Approaching sales forecasting using recurrent neural networks and transformers," *Expert Systems with Applications* , vol. 201, p. 116993, Apr. 2022, doi: 10.1016/j.eswa.2022.116993 .

[19] G. Theodoridis and A. Tsadiras, "Retail Demand Forecasting: A Comparative Analysis of Deep Neural Networks and the Proposal of LSTMixer, a Linear Model Extension," *Information* , vol. 16, no. 7, p. 596, Jul. 2025, doi: 10.3390/info16070596 .

[20] N. S. Sukanya, V. V. Asha, M. T. Vasumathi, S. Gothe, S. Fernandes, and S. Kumar, "Enhancing Weather Prediction for Transportation and Logistics Using the Temporal Fusion Transformer Algorithm," pp. 1161–1165, Jun. 2025, doi: 10.1109/icici65870.2025.11069564 .

[21] I. Zeroual and J. E. Bouhdidi, "Enhancing supply chain agility with advanced weather forecasting," *International Journal of Power Electronics and Drive Systems/International Journal of Electrical and Computer Engineering* , vol. 15, no. 6, p. 5904, Nov. 2025, doi: 10.11591/ijece.v15i6.pp5904-5913 .

[22] A. Alzahrani and M. Z. Asghar, "Intelligent Risk Prediction System in IoT-Based Supply Chain Management in Logistics Sector," *Electronics* , vol. 12, no. 13, p. 2760, Jun. 2023, doi: 10.3390/electronics12132760 .

[23] B. Abdulla and B. Birgisson, "Predicting Road Network Vulnerability to Fluvial Flooding Using Machine Learning Classifiers: Case Study of Houston during Hurricane Harvey," *Construction Research Congress 2020* , pp. 38–47, Nov. 2020, doi: 10.1061/9780784482865.005 .

[24] S. Takyi, R. B. Antwi, E. E. Özgüven, L. Okine, and R. Moses, "Towards Sustainable and Resilient Infrastructure: Hurricane-Induced Roadway Closure and Accessibility Assessment in Florida Using Machine Learning," *Sustainability* , vol. 17, no. 9, p. 3909, Apr. 2025, doi: 10.3390/su17093909 .

[25] H. Yang, D. Duque, and D. P. Morton, "Optimizing diesel fuel supply chain operations to mitigate power outages for hurricane relief," *IISE Transactions* , vol. 54, no. 10, pp. 936–949, Mar. 2022, doi: 10.1080/24725854.2021.2021461 .

[26] M. Hoque, A. R. Nabil, K. Hossain, N. Akhtar, and M. R. Amin, "Integrated AI-Based Decision Support Systems for Emergency Supply Chain Management in the United States During Natural Disasters," *Journal of Computer Science and Technology Studies* , vol. 7, no. 10, pp. 275–289, Oct. 2025, doi: 10.32996/jcsts.2025.7.10.31 .

[27] W. A. Zogaan, N. Ajabnoor, and A. A. Salamai, "Leveraging deep learning for risk prediction and resilience in supply chains: insights from critical industries," *Journal Of Big Data* , vol. 12, no. 1, Apr. 2025, doi: 10.1186/s40537-025-01143-4 .

[28] Z. Shen, H. Wang, J. Chen, and X. Song, "Resilience Inference for Supply Chains with Hypergraph Neural Network," in *Proceedings of the AAAI Conference on Artificial Intelligence* , Association for the Advancement of Artificial Intelligence, Mar. 2026, pp. 39259–39267. doi: 10.1609/aaai.v40i46.41274 .

[29] Sichong Huang, "AI-Driven Early Warning Systems for Supply Chain Risk Detection: A Machine Learning Approach," *Academic Journal of Computing & Information Science* , vol. 8, no. 9, Jan. 2025, doi: 10.25236/ajcis.2025.080913 .

[30] D. Ivanov and A. Dolgui, "Stress testing supply chains and creating viable ecosystems," *Operations Management Research* , vol. 15, pp. 475–486, May 2021, doi: 10.1007/s12063-021-00194-z .

[31] Md. I. Hossain, S. Talapatra, P. Saha, and H. M. Belal, "From Theory to Practice: Leveraging Digital Twin Technologies and Supply Chain Disruption Mitigation Strategies for Enhanced Supply Chain Resilience with Strategic Fit in Focus," *Global Journal of Flexible Systems Management* , vol. 26, no. 1, pp. 87–109, Dec. 2024, doi: 10.1007/s40171-024-00424-w .

[32] N. K. Jingar, "Ensuring Safety, Accountability, and Drift Resistance in LLM-Based Supply Chain Optimization," *International Journal of Scientific Research in Science Engineering and Technology* , p. 472, Jan. 2023, doi: 10.32628/ijsrset2310372 .

[33] S. Petrova and M. Hughes1, "Physics-Informed Graph Neural Networks for Supply Chain Disruption Prediction and Mitigation," *Frontiers in Applied Physics and Mathematics* , vol. 2, no. 1, pp. 130–147, Dec. 2025, doi: 10.71465/fapm458 .

[34] Z. Ge, L. Kong, and A. Brintrup, "Federated machine learning for privacy preserving, collective supply chain risk prediction," *International Journal of Production Research* , vol. 61, no. 23, pp. 8115–8132, Feb. 2023, doi: 10.1080/00207543.2022.2164628 .