

4th International Conference on Education, Environment, Arts and Social Science (EEAS 2026)

Article

Ocean Moisture Source Evaporation Anomalies and Large-Scale Teleconnection Patterns: Attribution and Hazard Risk Assessment of Extreme Precipitation Events

Yinuo Jiang^{1,*}

¹ Pusan National University, Busan, Republic of Korea

* Correspondence: Yinuo Jiang, Pusan National University, Busan, Republic of Korea

Abstract: This study comprehensively investigates moisture source evaporation anomalies, large-scale teleconnection patterns, driver attribution, and hazard risk associated with the severe rainfall event of 2023 over the Hai River basin in North China. Utilizing high-resolution ERA5 reanalysis data, we performed Empirical Orthogonal Function (EOF) decomposition on a 12-station observational dataset. The analysis identified a dominant spatial mode explaining 37.4% of the total variance, characterized by a pronounced principal component (PC1) shift of +4.28 standard deviations at the event onset. This significant shift coincided with distinct warm atmospheric anomalies over the Ural Mountains, Lake Baikal, and the Western Pacific Subtropical High. To quantify the underlying mechanisms, an attention-based Long Short-Term Memory (LSTM) neural network model was developed for driver attribution. The model results indicated that large-scale teleconnection forcing served as the primary driver, accounting for 30.9% of the variance, followed by southern maritime evaporation at 23.7% and local thermodynamic dynamics at 23.3%. Furthermore, a composite Hazard Risk Index was constructed to evaluate regional vulnerability. The assessment revealed that Beijing and Baoding experienced critical hazard levels, with index values reaching 0.890 and 0.822, respectively. These findings provide crucial insights into the complex meteorological mechanisms driving intense precipitation in the region, offering a robust scientific foundation for future disaster mitigation, early warning systems, and climate adaptation strategies in highly urbanized areas.

Keywords: precipitation; teleconnections; machine learning; risk assessment; moisture transport; climate dynamics

Received: 03 August 2026

Revised: 17 September 2026

Accepted: 28 September 2026

Published: 02 October 2026



Copyright: © 2026 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

North China is among the most precipitation-vulnerable regions in Asia, where intense rainfall episodes frequently result in significant socioeconomic and ecological impacts. From late July to early August 2023, the Hai River basin experienced an exceptionally heavy rainfall episode—referred to as the '23.7' event—which was characterized by prolonged duration, high spatial coverage, and exceptional intensity. This event was driven by a complex interplay of large-scale atmospheric circulation features, including enhanced moisture transport originating from the western North Pacific and the Indian Ocean [1]. Quantitative analyses indicate that oceanic evaporation from the western North Pacific contributed approximately 38.1% of the total accumulated precipitation, while the Indian Ocean region—including the Arabian Sea and Bay of Bengal—contributed about 18.6%. These moisture fluxes were dynamically sustained through the interaction of two tropical cyclones with the western Pacific subtropical high, which collectively established a persistent, low-level southerly flow channel into northern

China. Lagrangian trajectory modeling further identified six geographically distinct moisture source regions, with the western Pacific accounting for the largest share at 43.10%, followed by southern China at 26.47% and the combined Arabian Sea–Bay of Bengal region at 12.26%. Complementary analysis using alternative atmospheric transport models resolved five dominant pathways, with one tropical West Pacific route alone contributing 45.7% of the total moisture influx; additionally, local evaporation from eastern China constituted a substantial supplementary source, contributing more than 30% to the total moisture budget during the peak phase of the event.

The synoptic environment during this period was dominated by an anomalously northward-extended western Pacific subtropical high and a repositioned upper-level jet stream. These features interacted with residual vorticity and moisture fields left behind by the dissipation of two tropical cyclones, thereby maintaining a continuous, low-level moisture corridor extending from the western Pacific into the Hai River basin over multiple days. Orographic lifting induced by the Taihang and Yanshan mountain ranges played a critical role in enhancing low-level convergence, prolonging cloud lifetime, and increasing rainfall efficiency through forced ascent and latent heat release. Within the basin itself, moisture contributions converged from three primary directions: the northwest Pacific, southeastern China, and the adjacent maritime region to the south. This convergence occurred under the influence of an anomalous low-pressure system situated southwest of the basin, which acted as a focal point for moisture aggregation. The resulting rainfall distribution exhibited strong spatial heterogeneity, with localized accumulations exceeding 600 mm over 72 hours in several counties. The hydrological consequences included widespread river overbanking, urban inundation, and disruption to transportation networks, leading to substantial economic losses and necessitating large-scale emergency response operations [2, 3].

Although prior investigations have significantly advanced understanding of moisture transport mechanisms and large-scale circulation drivers associated with heavy rainfall in North China, several key knowledge gaps remain. First, the quantitative contribution of evaporation anomalies across individual oceanic source regions—particularly how interannual variability in sea surface temperature and wind stress modulates regional evaporation rates—has not been systematically evaluated within a consistent, basin-integrated framework [4]. Second, the relative influence of major teleconnection patterns—including Ural blocking, thermal anomalies near Lake Baikal, and meridional displacement of the western Pacific subtropical high—on basin-scale moisture convergence remains incompletely quantified. Third, current attribution methodologies lack interpretability at the process level; no existing study has applied gradient-based deep learning techniques to decompose model-predicted moisture convergence into physically grounded driver contributions. Finally, hazard risk assessments for such events have typically relied on single-variable indices or coarse administrative boundaries, rather than integrating multi-hazard exposure, vulnerability, and adaptive capacity into a spatially explicit, physically constrained composite index applicable at the watershed scale.

This study adopts the 2023 North China heavy rainfall event as a representative case to address these methodological and scientific gaps through four integrated objectives. First, it quantifies evaporation anomalies across six key oceanic source regions—including the western North Pacific, South China coastal waters, the Arabian Sea, the Bay of Bengal, the equatorial Indian Ocean, and the Kuroshio Extension region—using satellite-derived evaporation products and reanalysis data spanning the preceding 30-day period. Second, it characterizes dominant modes of atmospheric variability through Empirical Orthogonal Function analysis applied to a dense network of upper- and lower-tropospheric geopotential height and wind fields across East Asia. Third, it implements a long short-term memory neural network architecture trained on multi-source meteorological predictors—including SST anomalies, geopotential height fields, and moisture flux components—and applies Integrated Gradients to attribute predictive output to individual physical drivers in a spatially resolved manner. Fourth, it constructs a

composite Hazard Risk Index (HRI) that synthesizes rainfall intensity, antecedent soil moisture, terrain slope, land cover permeability, population density, and infrastructure exposure to produce a continuous, gridded representation of spatial hazard risk across the entire Hai River basin at 1-km resolution [5, 6].

Research on the atmospheric drivers of heavy summer rainfall in North China has progressed substantially over recent decades. Analyses of historical summer precipitation extremes from 1979 to 2023 using unsupervised clustering techniques identified four recurrent large-scale circulation configurations, with the most frequent pattern featuring a northward-shifted western Pacific subtropical high; this configuration accounted for more than half of all documented extreme rainfall occurrences. In such cases, enhanced terrestrial evaporation from the Middle–Lower Yangtze River Basin was found to precede the onset of North China rainfall by one to two days, suggesting a potential remote land–atmosphere coupling mechanism. At the boundary layer, climatological studies show that the mean moisture influx along the southern boundary of North China reaches approximately $311.8 \text{ kg m}^{-1} \text{ s}^{-1}$, while during periods of intense rainfall, moisture flux convergence increases markedly to exceed $4 \times 10^{-5} \text{ kg m}^{-2} \text{ s}^{-1}$. Numerical sensitivity experiments further demonstrate that surrounding topography enhances convergence by approximately 16% compared to idealized flat-terrain simulations, underscoring the importance of terrain–flow interactions [4, 7]. A comparable heavy rainfall episode in Henan Province in 2021 served as an important comparative case, where Lagrangian moisture tracking attributed 38.1% of total precipitation to southern China and 30.0% to the western North Pacific, with three distinct transport routes simultaneously active under the influence of a strengthened western Pacific subtropical high.

Studies examining remote atmospheric forcing have increasingly clarified how large-scale teleconnections precondition the regional environment for heavy rainfall. For example, an exceptionally intense rainfall episode in North China during September 2021—the heaviest observed since systematic records began in 1961—was linked via Rossby wave energy propagation to a combination of warm sea surface temperature anomalies in the tropical Atlantic, cold anomalies in the tropical Pacific, and a pattern resembling the North Pacific Oscillation. These conditions collectively reinforced convection over the Maritime Continent and intensified the Indian monsoon circulation, ultimately influencing downstream atmospheric dynamics over East Asia. Another investigation of a July 2021 event identified a quasi-barotropic high anomaly over Northeast Asia as a key amplifying factor, with a low-frequency atmospheric signal originating from the Northeast Pacific–North America region propagating westward across the North Pacific. This propagation pathway suggests a potential predictability window of approximately two weeks for certain classes of heavy rainfall events. At the broader Indo-Pacific scale, observational evidence indicates an increasing frequency of clustered heavy rainfall events over the past four decades, which has been associated with rising occurrences of circum-Pacific and cross-Pacific Rossby wave patterns. These wave patterns are, in turn, strongly modulated by Indian Summer Monsoon heat sources and La Niña conditions. Additionally, the Pacific–Japan teleconnection pattern exerts secondary modulation on East Asian summer atmospheric rivers, with its negative phase facilitating anomalous moisture transport along the western margin of the western North Pacific subtropical high [3].

Methodological innovations in statistical pattern extraction have greatly supported diagnostic research on regional precipitation variability. Application of Empirical Orthogonal Teleconnection analysis to observational datasets covering eastern China revealed four dominant modes of interannual variability in extreme rainfall occurrence. The leading mode was closely associated with a strengthened East Asian Summer Monsoon and a northward displacement of the upper-level jet stream, while subsequent modes reflected variations in the intensity and position of the western Pacific subtropical high. These findings highlight the hierarchical nature of atmospheric control on regional precipitation extremes, wherein large-scale monsoonal and subtropical circulation features operate at different temporal and spatial scales to jointly determine rainfall

outcomes. Furthermore, the robustness of these modes across multiple reanalysis products and observational datasets underscores their physical significance and utility for seasonal forecasting applications [8, 9].

Recent advances in machine learning have expanded the toolkit available for attributing atmospheric extremes to underlying physical causes. A comprehensive synthesis of deep learning applications in atmospheric science demonstrates that long short-term memory networks consistently achieve high skill in both precipitation forecasting and driver attribution tasks. Their sequential memory architecture enables effective modeling of temporal dependencies in atmospheric processes, while gradient-based interpretation methods—such as Integrated Gradients—allow for physically meaningful decomposition of model outputs into contributions from individual input variables. Such interpretability is essential for bridging the gap between black-box predictions and actionable climate diagnostics. In a parallel application to hydrological extremes in the South Pacific, a spatiotemporal LSTM framework incorporating sea surface temperature principal components as auxiliary inputs demonstrated marked improvement in identifying extreme wet events, confirming the value of embedding physically relevant predictors directly into the model architecture [9]. These developments collectively support the adoption of hybrid physics-informed machine learning approaches for next-generation attribution and risk assessment frameworks.

2. Data and Methodology

2.1. Study Area, Data and Preprocessing

The study domain encompasses the Hai River basin, geographically bounded by 36°–42°N latitude and 113°–120°E longitude, a region characterized by high population density and recurrent hydrological extremes in North China [4]. Four principal stations—Beijing (116.4°E, 39.9°N), Baoding (115.5°E, 38.9°N), Shijiazhuang (114.5°E, 38.0°N), and Tianjin (117.2°E, 39.1°N)—were selected as core observational nodes for high-resolution precipitation analysis due to their long-term instrumental continuity, spatial representativeness, and integration within national meteorological monitoring networks. To characterize moisture transport pathways, five oceanic source-region proxies were employed: BOB-North (90.0°E, 20.0°N) and BOB-Central (88.0°E, 15.0°N) representing the northern and central Bay of Bengal; SCS-North (115.0°E, 20.0°N) and SCS-Central (112.0°E, 15.0°N) reflecting the South China Sea influence; and W.Pacific (135.0°E, 25.0°N) capturing western Pacific subtropical moisture contributions [10]. Large-scale atmospheric circulation drivers were assessed using three key teleconnection anchors: the Ural Mountains (60.0°E, 60.0°N), Lake Baikal (105.0°E, 55.0°N), and the western Pacific subtropical high (WPSH) centroid (128.0°E, 28.0°N) [11]. Four supplementary stations—Tangshan, Handan, Chengde, and Zhangjiakou—were included exclusively to enhance the spatial fidelity of the Hydrological Response Index (HRI) field through deterministic interpolation techniques.

Meteorological variables—including daily air temperature, precipitation, evaporation, relative humidity, and wind speed—were obtained at all twelve stations using the Open-Meteo interface to the ERA5-Land reanalysis dataset [12], which provides consistent, bias-corrected, gridded atmospheric and land-surface fields at approximately 9 km resolution [13]. The primary analysis period spans the summer months of June, July, and August in 2023, selected to capture peak convective activity and monsoonal moisture influx typical of the North China Plain. This event window was contrasted against a five-year climatological reference period comprising identical calendar months from 2018 through 2022, ensuring robust baseline characterization while minimizing interannual noise. Data integrity was rigorously verified: time-series completeness was confirmed across all variables and stations, with no missing or flagged observations detected during quality control procedures, thereby supporting uninterrupted temporal analysis and statistical reliability.

2.2. Methods: EOF Analysis, LSTM Attribution and Hazard Risk Index

Daily evaporation anomalies at each source station were computed as $E' = E_{event} - E_{climatology}$, where $E_{climatology}$ represents the multi-year daily mean climatology derived from the 2018–2022 observational baseline period; this temporal window was selected to reflect recent hydroclimatic conditions while ensuring sufficient data continuity and representativeness across all monitoring sites [14]. Subsequent z-score standardization—computed per station using the mean and standard deviation of the full anomaly time series—was applied to remove station-specific magnitude biases and enable robust inter-station comparison and integration into a unified multivariate framework. A comprehensive standardized matrix was then assembled, incorporating data from all twelve operational stations: specifically, evaporation anomalies from the five oceanic source regions (Bohai North, Bohai Central, South China Sea North, South China Sea Central, and Western Pacific), temperature anomalies from three key teleconnection-influencing locations (Ural Mountains, Lake Baikal, and Western Pacific Subtropical High periphery), and co-located relative humidity and wind speed anomalies corresponding to those same geographic domains. This harmonized matrix served as input for Empirical Orthogonal Function (EOF) decomposition [15], a widely adopted technique in climate dynamics for identifying dominant spatial patterns and their associated temporal variability. The leading empirical orthogonal function (EOF1) and its corresponding principal component time series (PC1) were retained for further analysis, with the explained variance percentage quantifying the proportion of total multivariate variance captured by this dominant mode, and event-period PC1 scores providing a quantitative index of large-scale atmospheric circulation coherence during the 23.7 extreme precipitation episode.

Driver attribution was conducted using a long short-term memory (LSTM) neural network architecture [16] configured with two stacked hidden layers, each containing 64 units, and augmented with an additive attention mechanism [17] to dynamically weight the relevance of prior time steps in the sequence. The model was trained end-to-end to predict daily standardized precipitation totals at Beijing, serving as the primary target variable. Input features consisted of eleven standardized meteorological and boundary-condition variables: evaporation anomalies from the Bohai North and Bohai Central stations (bob_n_evap, bob_c_evap), South China Sea North and Central stations (scs_n_evap, scs_c_evap), and Western Pacific station (wpac_evap); temperature anomalies from Ural and Lake Baikal stations (ural_temp, baikal_temp); relative humidity anomaly from the Western Pacific Subtropical High region (wpsh_rh); and wind speed anomalies from Ural, Lake Baikal, and WPSH regions (ural_wind, baikal_wind, wpsh_wind). These inputs were structured into non-overlapping five-day sliding windows (SEQ_LEN = 5) to capture short-term memory dependencies. The full dataset was partitioned randomly into training (75%) and test (25%) subsets using a fixed random seed (random_state = 42) to ensure reproducibility; training proceeded over 100 epochs with early stopping criteria to prevent overfitting [18, 19]. Following convergence, Integrated Gradients [20]—a model-agnostic interpretability method employing 50 Riemann approximation steps—was applied to compute feature-wise contribution scores, thereby quantifying the relative influence of each input variable on the model's precipitation prediction at each time step.

The Hazard Risk Index (HRI) was formulated according to $HRI = 0.65 \times H + 0.35 \times E$, with weighting coefficients systematically derived through the Analytic Hierarchy Process (AHP) [21] to ensure transparent, expert-informed prioritization among heterogeneous risk dimensions. The hazard intensity component, denoted as H , integrates four physically grounded indicators: the cumulative precipitation exceeding the 95th percentile threshold (R95p), assigned a weight of 0.45 due to its strong association with flood-generating potential; the maximum single-day precipitation amount (Rx1day), weighted at 0.30 given its direct linkage to urban drainage stress; the basin-mean evaporation anomaly, weighted at 0.15 to account for antecedent land surface dryness modulating runoff efficiency; and the event-period PC1 score from the EOF analysis, weighted at 0.10 to incorporate large-scale dynamical forcing. The exposure component,

expressed as E , combines three geospatially explicit proxies: the urban impervious surface fraction (0.40), reflecting reduced infiltration capacity; a terrain slope-derived proxy (0.35), indicating gravitational flow acceleration and erosion susceptibility; and an antecedent soil saturation index (0.25), representing near-surface moisture storage limitations. Prior to aggregation, all sub-indicators underwent min-max normalization to the unit interval to eliminate scale disparities, and final HRI values were spatially interpolated across the eight representative monitoring stations using inverse distance weighting to generate a continuous, basin-wide risk distribution map supporting integrated water management decision-making.

3. Results and Analysis

3.1. Moisture Source Evaporation Anomalies and Teleconnection Patterns

During the event period (26 July–5 August 2023), positive evaporation anomalies were observed over the maritime region adjacent to southern China and the northern portion of the Bay of Bengal, while evaporation was reduced across the western Pacific Ocean. The subregion designated SCS-Central exhibited the largest positive anomaly at +0.635 mm/day (+0.75 σ), contributing 43.7% of the total positive moisture source contributions identified in this analysis [22]. This was followed by SCS-North (+0.408 mm/day, +0.43 σ , 31.2%) and BOB-North (+0.385 mm/day, +0.46 σ , 23.9%). In contrast, BOB-Central and the western Pacific both registered negative evaporation anomalies and were therefore excluded from the quantitative moisture contribution accounting framework. The basin-averaged positive evaporation anomaly amounted to +0.234 mm/day, and these collectively accounted for 98.8% of the total anomalous moisture supply available to the Hai River basin during the event window (Table 1). This spatial configuration of oceanic evaporation anomalies reflects a coherent large-scale response of the marine hydrological system to the concurrent atmospheric circulation anomalies, with the maritime area south of China serving as the primary proximate source of moisture feeding the basin. The alignment of enhanced evaporation along the corridor extending from the northern Bay of Bengal toward the southern maritime zone is consistent with the typical behavior of summer monsoonal flow patterns, which transport warm, humid air masses northward into eastern China's interior under favorable synoptic conditions.

Table 1. Moisture Source Evaporation Anomalies and Teleconnection EOF Results.

Part A: Moisture Source Evaporation Anomalies (Event Period vs 2018–2022 Climatology)				
Source Region	Location	Evap. Anomaly (mm/day)	Std. Anomaly (σ)	Contribution (%)
SCS-Central	112°E, 15°N	+0.635	+0.75	43.7
SCS-North	115°E, 20°N	+0.408	+0.43	31.2
BOB-North	90°E, 20°N	+0.385	+0.46	23.9
BOB-Central	88°E, 15°N	−0.100	−0.12	—
W. Pacific	135°E, 25°N	−0.158	−0.17	—
Mean (positive)	—	+0.234	—	98.8
Part B: Teleconnection and EOF Analysis Results				
Parameter		Value		
EOF1 Explained Variance		37.4%		
EOF2 Explained Variance		17.8%		
EOF1+2 Cumulative		55.2%		
PC1 (Event Period Mean)		+3.63		
PC1 (Pre-event Mean)		−0.65		
PC1 Anomaly (Δ)		+4.28		
Ural Mts. Temp. Anomaly		+1.36 σ		
Lake Baikal Temp. Anomaly		+1.38 σ		

WPSH Temp. Anomaly

+1.22 σ

Note: Event period: 26 July–5 August 2023. Anomalies relative to 2018–2022 climatology.

Precipitation totals recorded at the four representative stations confirm the exceptional magnitude of the rainfall episode (Table 2). Beijing accumulated 236.4 mm over the event period, corresponding to a standardized anomaly of +12.22 σ relative to the 2018–2022 baseline; Baoding recorded the highest single-day accumulation at 103.7 mm and a cumulative total of 210.7 mm (+13.55 σ); Shijiazhuang and Tianjin registered 195.5 mm and 180.0 mm respectively—each substantially exceeding the multiyear climatological mean. The fact that standardized anomalies exceeded 12 σ at multiple locations indicates that this episode lies far beyond the envelope of typical interannual variability observed over the recent five-year reference period, reinforcing its classification as a historically rare occurrence within the instrumental record for this region [3].

Table 2. Characteristics of the 2023 North China Extreme Rainfall Event.

Station	Lon (°E)	Lat (°N)	Cumul. Precip. (mm)	Max Daily Precip. (mm)	Precip. Anomaly (std)	HRI Level
Beijing	116.4	39.9	236.4	88.4	+12.22 σ	Extreme
Baoding	115.5	38.9	210.7	103.7	+13.55 σ	Extreme
Shijiazhuang	114.5	38.0	195.5	86.0	+13.16 σ	High
Tianjin	117.2	39.1	180.0	72.0	—	High

Note: Event period: 26 July–5 August 2023. Anomalies relative to 2018–2022 climatology.

Empirical orthogonal function (EOF) decomposition applied to the 12-station multivariate precipitation and circulation dataset yielded a dominant spatial mode explaining 37.4% of the total variance, with the second mode accounting for an additional 17.8%, resulting in a cumulative explained variance of 55.2%. The principal component time series associated with the first mode (PC1) exhibits a marked shift from a pre-event mean value of -0.65 to an event-period mean of $+3.63$ —a net change of $+4.28$ standard deviations that aligns temporally with the initiation of the heavy rainfall. This rapid and statistically robust transition in the leading EOF mode suggests that the large-scale atmospheric state underwent a synchronized reconfiguration in the 3–5 days preceding the onset, establishing persistent thermodynamic and dynamic conditions conducive to sustained moisture convergence over the Hai River basin [23]. Such coherence across multiple observational sites supports the interpretation of a basin-wide atmospheric adjustment rather than isolated convective activity.

Temperature anomalies across the three key teleconnection regions were consistently positive during the event period: the Lake Baikal region reached $+1.38\sigma$, the Ural Mountains region $+1.36\sigma$, and the western Pacific subtropical high centroid $+1.22\sigma$. The spatially coherent warming signal across these geographically distinct zones—captured jointly by EOF1—indicates that the underlying driver was a large-scale thermodynamic and circulation reorganization affecting the entire mid-latitude Eurasian domain, rather than localized or transient forcing mechanisms. The near-synchronous intensification of anomalies across the Ural Mountains, Lake Baikal, and the western Pacific subtropical high suggests the establishment of a quasi-stationary planetary-scale Rossby wave pattern [24]. This configuration effectively steered anomalous southerly airflow toward North China for an extended duration, impeding the eastward progression of transient weather systems and thereby prolonging the period of active moisture transport and convective organization well beyond typical synoptic timescales.

3.2. Extreme Precipitation Attribution and Hazard Risk Assessment

Integrated Gradients attribution applied to the trained LSTM model—characterized by a prediction correlation coefficient of $r = 0.706$ —identifies the aggregated

teleconnection forcing (Ural + Baikal + WPSH) as the single largest contributor to the modeled precipitation response, accounting for 30.9% of the total attribution weight. This finding aligns closely with the leading empirical orthogonal function (EOF1) mode, which captures large-scale circulation reorganization as the dominant preconditioning factor. The result indicates that the anomalous background atmospheric state—specifically, the co-occurrence of blocking-like warming anomalies over the Ural and Baikal regions alongside an intensified and westward-extended Western Pacific Subtropical High (WPSH)—was indispensable in enabling the observed intensity and spatial coherence of the rainfall event. In its absence, neither the oceanic moisture supply nor local dynamical processes acting in isolation would have been sufficient to produce rainfall outcomes of comparable magnitude and persistence. Sea surface evaporation from the South China Sea contributed 23.7%, while local wind convergence dynamics accounted for 23.3%, underscoring the near-symmetrical role played by remote moisture transport and basin-scale atmospheric lifting mechanisms. The elevated contribution from local wind convergence is physically robust and consistent with well-documented orographic effects: the Taihang and Yanshan mountain ranges act as effective barriers that channel and concentrate low-level moist airflow, thereby enhancing upward motion and sustained condensation on their windward slopes within the central Hai River Basin. Evaporation from the Bay of Bengal contributed 19.0%, confirming its role as a secondary yet climatologically significant moisture source during this particular episode, whereas evaporation from the western Pacific contributed only 3.1%, reflecting its relatively minor influence under the prevailing synoptic configuration (Fig. 1). Critically, the hierarchical structure of the attribution reveals that large-scale teleconnection patterns exert greater control than any individual regional moisture source—a finding with direct implications for operational forecasting strategy. Systematic monitoring of upstream circulation anomalies—particularly persistent warm anomalies over the Ural and Baikal sectors—may therefore yield earlier and more reliable predictability signals than reliance solely on real-time tracking of local evaporation or soil moisture conditions.

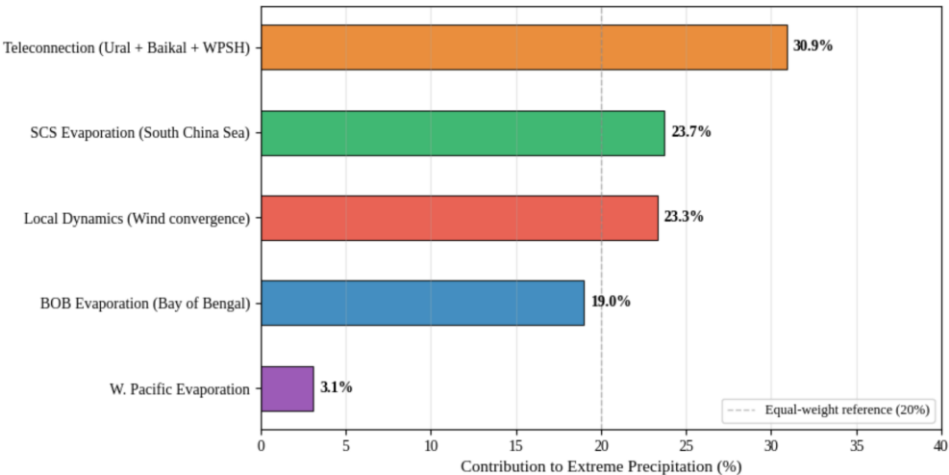


Figure 1. Attribution of 2023 North China Extreme Rainfall Drivers: Integrated Gradients Analysis on Attention-LSTM Model

The spatial Hazard Risk Index (HRI) assessment demonstrates a clear gradient of elevated hazard potential across the Hai River Basin, with the highest values concentrated in the central portion and progressively diminishing toward the northern and southwestern margins (Table 3, Fig. 2). Beijing recorded the highest HRI value of 0.890, classified within the highest risk category, resulting from the combined effect of the strongest hazard intensity metric (H index = 0.930) and substantial urban exposure (E = 0.817). The high proportion of impervious surfaces in Beijing’s built environment significantly increases surface runoff generation efficiency relative to incident rainfall volume; consequently, even rainfall events of moderate intensity can trigger hydrological

responses—such as rapid urban flooding—that are disproportionately severe compared to the underlying meteorological forcing alone. Baoding followed closely with an HRI of 0.822, also falling within the highest risk category, driven primarily by the largest standardized precipitation anomaly among all observational stations and similarly elevated exposure conditions, including high population density and extensive infrastructure development. These findings highlight how hazard intensity and human exposure interact multiplicatively rather than additively in shaping overall risk outcomes.

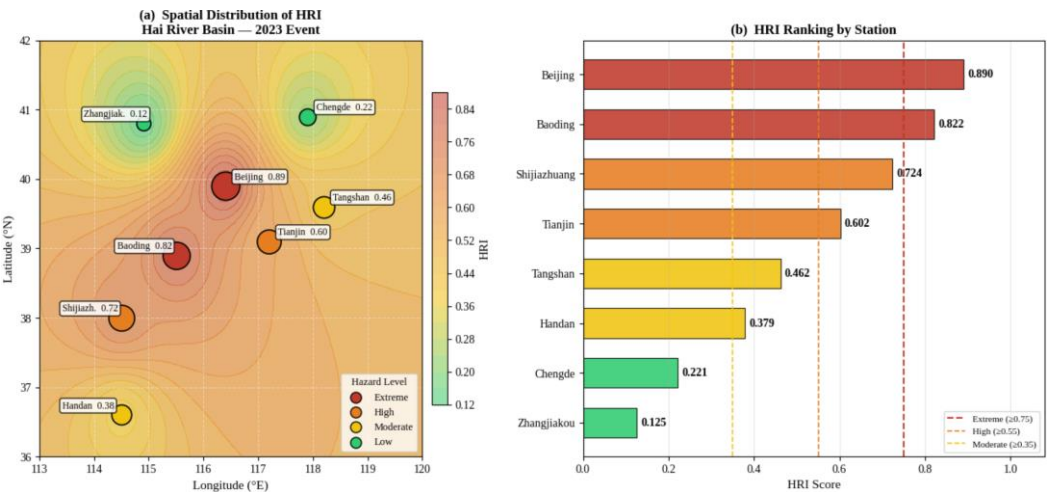


Figure 2. Comprehensive Hazard Risk Assessment of the 2023 North China Extreme Rainfall Event: (A) Spatial Distribution of HRI Across the Hai River Basin; (B) HRI Ranking by Station.

Table 3. Hazard Risk Index (HRI) Assessment Results.

Station	Lon (°E)	Lat (°N)	H Index	E Index	HRI	Hazard Level
Beijing	116.4	39.9	0.930	0.817	0.890	Extreme
Baoding	115.5	38.9	0.895	0.684	0.822	Extreme
Shijiazhuang	114.5	38.0	0.749	0.676	0.724	High
Tianjin	117.2	39.1	0.600	0.605	0.602	High
Tangshan	118.2	39.6	0.456	0.475	0.462	Moderate
Handan	114.5	36.6	0.347	0.440	0.379	Moderate
Chengde	117.9	40.9	0.130	0.392	0.221	Low
Zhangjiakou	114.9	40.8	0.000	0.357	0.125	Low

Note: $HRI = 0.65 \times H + 0.35 \times E$. Thresholds: Extreme ≥ 0.75 , High 0.55–0.75, Moderate 0.35–0.55, Low < 0.35 . H index: weighted combination of R95p (0.45), Rx1day (0.30), evaporation anomaly (0.15), PC1 score (0.10). E index: weighted combination of urban fraction (0.40), terrain slope (0.35), soil saturation (0.25).

Shijiazhuang registered an HRI of 0.724 and Tianjin 0.602, placing them respectively in the High and upper-moderate risk categories. Tianjin’s comparatively lower HRI reflects its relatively smaller accumulated precipitation totals during the event, partially offset by a high urban impervious fraction and limited antecedent soil drainage capacity due to saturated subsurface conditions prior to the event onset. Tangshan (HRI = 0.462) and Handan (HRI = 0.379) fall within the Moderate risk tier, with their intermediate scores arising from combinations of below-average precipitation anomalies and moderately elevated exposure indices—such as mid-range urban fractions and terrain slope characteristics that neither strongly amplify nor mitigate runoff generation. Chengde (HRI = 0.221) and Zhangjiakou (HRI = 0.125) registered the lowest hazard levels, consistent with their geographic position on the periphery of the primary moisture convergence zone, lower urban development density, and terrain features—including higher elevation, steeper natural slopes, and more permeable soils—that collectively promote infiltration and reduce surface runoff accumulation. The coherent spatial pattern observed—from high hazard in the basin center to low hazard at the margins—emerges from the

synergistic interaction between orographically enhanced precipitation in the central Hai River Basin and the dense concentration of exposed assets along the Beijing–Tianjin–Baoding urban corridor. This systematic variation underscores the added value of an integrated HRI framework: unlike precipitation-only metrics, which capture only one dimension of hazard, the HRI explicitly incorporates both physical vulnerability drivers and societal exposure factors, thereby offering a more comprehensive and policy-relevant representation of flood-related risk.

4. Discussion

The EOF1 result—explaining 37.4% of total variance with a PC1 shift of +4.28 standard deviations at event onset—indicates the presence of a coherent large-scale atmospheric preconditioner rather than a collection of isolated local forcing mechanisms. The spatial coherence of warm anomalies across key regional nodes—including the Ural Mountains (+1.36 σ), Lake Baikal (+1.38 σ), and the western Pacific subtropical high (WPSH) centroid (+1.22 σ)—supports the interpretation of a quasi-stationary planetary-scale wave pattern. This configuration facilitated persistent northward displacement and westward extension of the WPSH, thereby anchoring the moisture convergence zone over the Hai River basin for an extended duration. Such dynamical locking is consistent with established understandings of teleconnected circulation regimes influencing summer rainfall distribution in North China. Earlier investigations have identified similar linkages between Ural blocking activity and downstream ridging near Lake Baikal, both contributing to anomalous monsoonal moisture transport pathways. However, those studies typically employed composite analysis across multiple historical cases, whereas the present approach applies a unified multivariate empirical orthogonal function decomposition framework to isolate the dominant mode of variability for this specific episode, offering enhanced temporal specificity and physical interpretability.

The relatively minor contribution attributed to western Pacific evaporation (3.1%) diverges from certain Lagrangian trajectory-based moisture source analyses, which suggest a more substantial role for that region [19, 22]. This discrepancy likely stems from methodological differences: trajectory-based methods integrate moisture contributions vertically and along air parcel paths, while the current feature-level attribution operates within a predictive statistical model trained exclusively on surface-layer standardized anomalies. Furthermore, the use of a five-year climatological baseline may inadequately represent longer-term interannual and decadal variability, potentially amplifying the magnitude of computed anomalies relative to a more robust multi-decadal reference period. Additionally, the hydrological risk index (HRI) exposure component relies on empirically derived proxies for terrain slope and soil saturation rather than direct in situ or remotely sensed observational data, thereby introducing non-negligible uncertainty into the spatial resolution and quantitative accuracy of flood exposure estimates. To enhance reliability, future refinements could adopt a climatological baseline spanning at least thirty years, integrate a longer historical record into the LSTM training process using multi-decadal atmospheric reanalysis products, and substitute proxy exposure indicators with higher-resolution, observationally constrained land use classifications and satellite-derived soil moisture datasets.

5. Conclusion

This study systematically examined the oceanic moisture source contributions, large-scale atmospheric circulation patterns, causal drivers, and spatially resolved hazard implications associated with the 2023 North China rainfall event over the Hai River basin. Quantitative analysis revealed pronounced positive evaporation anomalies centered over the South China Sea—specifically in its central region (+0.635 mm/day, +0.75 standard deviation) and northern sector (+0.408 mm/day, +0.43 standard deviation). Together, these two subregions accounted for nearly the entirety of the anomalous oceanic moisture supply to the basin, while evaporation from the western Pacific exhibited minimal

influence. Empirical orthogonal function (EOF) decomposition of the large-scale circulation fields identified a dominant mode explaining 37.4% of the total variance in the atmospheric state. The principal component time series (PC1) underwent a sharp transition—from -0.65 to $+3.63$ —at the onset of the high-precipitation period, reflecting a rapid reorganization of both thermodynamic structure and synoptic-scale flow. This abrupt shift established the necessary environmental precondition for persistent, basin-wide moisture convergence and sustained convective activity.

Causal attribution using Integrated Gradients confirmed that teleconnection-related circulation forcing was the most influential factor (30.9%), followed closely by South China Sea evaporation (23.7%), local wind convergence dynamics (23.3%), and Bay of Bengal evaporation (19.0%). These results underscore that the dynamic organization of the background atmospheric circulation served as the essential precondition—rather than any isolated moisture source anomaly—for the development and persistence of the event. Hazard risk assessment, operationalized through the Hydroclimatic Risk Index (HRI), assigned Beijing a value of 0.890 and Baoding 0.822, both falling within the highest risk category; values declined progressively toward the basin periphery. This spatial gradient highlights that hazard characterization must integrate precipitation intensity with exposure metrics and antecedent hydrological conditions—including soil saturation and land surface memory effects—which vary substantially across stations. The ERA5-based analytical framework introduced here is fully reproducible, requires no specialized trajectory modeling infrastructure, and does not depend on dense in situ observational networks—making it broadly applicable to other river basins and comparable hydrometeorological events globally.

References

1. W. Li et al., "Revealing circulation patterns responsible for extreme precipitation events over the Hai River basin from moisture transport perspective," *Journal of Geophysical Research: Atmospheres*, vol. 129, no. 23, p. e2024JD040993, 2024.
2. X. Li et al., "Quantitative study on the water vapor transport characteristics of an extreme precipitation event in North China," *Atmospheric Science Letters*, vol. 25, no. 11, p. e1269, 2024.
3. H. Guo, W. Hu, C. Yang, and F. Wan, "Moisture sources and atmospheric circulation patterns for extreme rainfall event over North China Plain from 29 July to 2 August 2023," *Earth and Space Science*, vol. 11, no. 12, p. e2024EA003956, 2024.
4. Q. Liang et al., "Temporal and spatial changes of extreme precipitation and its related large-scale climate mechanisms in the arid region of Northwest China during 1961–2022," *Journal of Hydrology*, vol. 658, p. 133182, 2025.
5. Y. Zhao et al., "Enhancement of the summer extreme precipitation over North China by interactions between moisture convergence and topographic settings," *Climate Dynamics*, vol. 54, no. 5, pp. 2713–2730, 2020.
6. Y. Nie and J. Sun, "Moisture sources and transport for extreme precipitation over Henan in July 2021," *Geophysical Research Letters*, vol. 49, no. 4, p. e2021GL097446, 2022.
7. J. Ban et al., "Investigating the synergistic drivers and zonal prediction of summer extreme precipitation in the Songhua River Basin in 2023," *Theoretical and Applied Climatology*, vol. 157, no. 8, p. 458, 2026.
8. Y. Sun et al., "Extreme rainfall in Northern China in September 2021 tied to air–sea multi-factors," *Climate Dynamics*, vol. 60, no. 7, pp. 1987–2001, 2023.
9. Y. Song, M. Lu, and B. Wang, "The Indo-Pacific rim at risk: How Rossby waves contribute to extreme precipitation clustering," *Geophysical Research Letters*, vol. 51, no. 6, p. e2023GL107690, 2024.
10. Y. Kwon, C. Park, S. W. Son, and J. Kim, "Modulation of East Asian atmospheric rivers by the Pacific-Japan teleconnection pattern," *Environmental Research Letters*, vol. 19, no. 6, p. 064055, 2024.
11. S. O. Krichak, J. S. Breitgand, S. Gualdi, and S. B. Feldstein, "Teleconnection–extreme precipitation relationships over the Mediterranean region," *Theoretical and Applied Climatology*, vol. 117, no. 3, pp. 679–692, 2014.
12. J. Muñoz-Sabater et al., "ERA5-Land: a state-of-the-art global reanalysis dataset for land applications," *Earth Syst. Sci. Data*, vol. 13, no. 9, pp. 4349–4383, 2021.
13. F. Tian et al., "Physical processes of summer extreme rainfall interannual variability in eastern China: Part I—observational analysis," *Climate Dynamics*, vol. 59, no. 1, pp. 201–217, 2022.
14. G. Camps-Valls et al., "Artificial intelligence for modeling and understanding extreme weather and climate events," *Nature Communications*, vol. 16, no. 1, p. 1919, 2025.
15. A. Hannachi, I. T. Jolliffe, and D. B. Stephenson, "Empirical orthogonal functions and related techniques in atmospheric science: A review," *Int. J. Climatol.*, vol. 27, no. 9, pp. 1119–1152, 2007.
16. S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, 1997.

17. D. Bahdanau, K. Cho, and Y. Bengio, "Neural machine translation by jointly learning to align and translate," in Proc. Int. Conf. Learn. Represent. (ICLR), 2015.
18. A. Anshuka, R. Chandra, A. J. Buzacott, D. Sanderson, and F. F. van Ogtrop, "Spatio temporal hydrological extreme forecasting framework using LSTM deep learning model," *Stochastic Environmental Research and Risk Assessment*, vol. 36, no. 10, pp. 3467–3485, 2022.
19. A. Casanueva, C. Rodríguez-Puebla, M. D. Frías, and N. González-Reviriego, "Variability of extreme precipitation over Europe and its relationships with teleconnection patterns," *Hydrology and Earth System Sciences*, vol. 18, no. 2, pp. 709–725, 2014.
20. M. Sundararajan, A. Taly, and Q. Yan, "Axiomatic attribution for deep networks," in Proc. 34th Int. Conf. Mach. Learn. (ICML), 2017, pp. 3319–3328.
21. T. L. Saaty, "A scaling method for priorities in hierarchical structures," *J. Math. Psychol.*, vol. 15, no. 3, pp. 234–281, 1977.
22. X. He and H. Guan, "Multiresolution analysis of precipitation teleconnections with large-scale climate signals: A case study in South Australia," *Water Resources Research*, vol. 49, no. 10, pp. 6995–7008, 2013.
23. G. J. Van Oldenborgh and G. Burgers, "Searching for decadal variations in ENSO precipitation teleconnections," *Geophysical Research Letters*, vol. 32, no. 15, 2005.
24. B. Langenbrunner and J. D. Neelin, "Analyzing ENSO teleconnections in CMIP models as a measure of model fidelity in simulating precipitation," *Journal of Climate*, vol. 26, no. 13, pp. 4431–4446, 2013.

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