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.

Spatial Information Pathways for Coastal Water-Level Forecasting Across Diverse Gauge Networks

Jia Rong

rong.j@kean.edu

Wenzhou-Kean University https://orcid.org/0009-0007-9772-7917

Guanchao Tong

tguancha@kean.edu

Wenzhou-Kean University https://orcid.org/0000-0002-3503-2745

Research Article

Keywords: coastal water level, spatial coupling, tide-gauge network, atmospheric forcing, high-water forecasting, hindcasting

Posted Date: August 12th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10666609/v1

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

Additional Declarations: The authors declare no competing interests.

Spatial Information Pathways for Coastal Water-Level Forecasting Across Diverse Gauge Networks

Jia Rong[1,2] Guanchao Tong[1,2]

[1]College of Science, Mathematics and Technology, Wenzhou–Kean University, 88 Daxue Road, Ouhai, Wenzhou, Zhejiang Province 325060, China
[2]College of Science, Mathematics and Technology, Kean University, 1000 Morris Avenue, Union, New Jersey 07083, USA
Jia Rong and Guanchao Tong contributed equally and are corresponding authors. Correspondence: rongj@kean.edu; tguancha@kean.edu.

Abstract

Coastal gauge networks contain spatial information that can strengthen water-level hindcasts beyond local temporal memory. This study compares stable station connectivity, time-ordered exchange, regional atmospheric conditioning, and event-dependent pair information above a common local forecast across seven tide-gauge networks spanning the Thames–Southern North Sea, New York Bight, Charleston, Gulf/Galveston, Southeast Australia, Tokyo–Tokai, and Los Angeles Bight. Spatial coupling strengthened trajectory reconstruction across the study systems and revealed distinct regional high-water patterns. Combined spatial information was most valuable in the Thames–Southern North Sea, regional atmospheric conditioning produced a coherent Gulf/Galveston profile, and time-ordered exchange was prominent in Southeast Australia. The comparison further showed that mean trajectory skill, high-water magnitude, event detection, and hindcast-derived warning timing illuminate complementary aspects of coastal forecast performance. The analysis links regional gauge-network structure to the engineering outcomes of coastal water-level hindcasts.

Keywords: coastal water level; spatial coupling; tide-gauge network; atmospheric forcing; high-water forecasting; hindcasting

1 Introduction

Forecasting coastal water level requires both local temporal memory and information distributed across a gauge network. At one station, the observed trajectory combines astronomical tide, atmospheric response, shelf and estuarine storage, and the persistence of earlier anomalies. Across stations, an exposed or leading gauge can provide advance information about a developing regional response. Recurrent neural networks provide compact temporal representations [1, 2], while graph recurrent, convolutional, and attention-based architectures support structured exchange in multivariate time series [3–11]. These capabilities have been adopted for tide-level and storm-surge prediction in typhoon, estuarine, and regional settings [12–20].

Cross-gauge dependence has been represented in several distinct ways. Variable-topology and causal-information graph models learn station exchange directly in multi-gauge coastal prediction [12, 13, 20]. Multipoint and attention-based networks combine several gauges or gridded predictors and vary the influence of those inputs over a forecast [16, 17]. Regional neural storm-surge models use localized atmospheric features [18, 19], whereas global studies provide data-driven storm-surge reconstructions and databases of coastal response under large-scale forcing [14, 15]. Fixed graphs encode stable distance or adjacency relationships. Dynamic and attention graphs allow the edge weights to change with the event state. Lagged graphs read a sender state from an earlier time, reflecting repeatable ordering among gauges. Meteorological forcing can enter as a station-local feature, a receiver condition on graph state,

or an ordered sender–receiver feature containing the two local forcing vectors and their signed difference. These configurations can achieve similar aggregate errors while producing different corrections to event amplitude, phase, and threshold crossing. Existing studies establish the value of multi-station and atmosphere-aware forecasting, but the ingredients are commonly combined inside one architecture and evaluated mainly through full-record error. Their separate predictive value across coastal networks and forecast objectives remains difficult to compare.

Building on a three-system analysis of lag-aligned gauge exchange and pairwise atmospheric forcing [21], this study extends the comparison to seven gauge networks and to route effects on trajectory, high-water, detection, and warning outcomes.

Root mean squared error (RMSE) and mean absolute error (MAE) summarize the full trajectory and are dominated by ordinary water-level conditions. High-water forecasting adds event magnitude, first threshold crossing, peak timing, and warning lead. Tide–surge interaction, wind stress, pressure anomalies, coastal geometry, and event timing all shape these outcomes [22–28]. Global analyses of storm surge, extreme sea level, and coastal flood frequency likewise motivate joint evaluation of aggregate and upper-tail performance [29–32]. A spatial route can improve the full trajectory and still occupy a different rank for high-water magnitude or warning time.

Coastal Graph Applicability and Tail Evaluation (CoastGATE) attaches five bounded spatial corrections to one local gated recurrent unit (Local-GRU) trajectory. Static connectivity represents stable distance-weighted association. Directed-delay propagation adds time-ordered sender access estimated from historical residuals [33–35]. Regional-forcing coupling conditions the receiving state on local wind and pressure. Dynamic pair coupling uses simultaneous station states and ordered forcing contrasts, and Lagged pair coupling combines those pair features with directed sender delays. Across seven networks with different layouts, forcing contrasts, and high-water event counts, we quantify each correction through mean trajectory error, event-weighted 95th- and 99th-percentile (q95 and q99) high-water magnitude, event detection, peak behaviour, and warning lead. This comparison tests how the value of spatial information changes with coastal setting and forecast objective.

2 Materials and methods

2.1 Study systems, observations, and topology

The seven regional datasets cover 2015–2024 and combine tide-gauge observations, deterministic tide, atmospheric fields, and bathymetry-derived connectivity. Each dataset uses a fixed station order, source metadata, and hourly time coordinate. United States stations were obtained from the National Oceanic and Atmospheric Administration Center for Operational Oceanographic Products and Services (NOAA CO-OPS) archive and metadata service [36]. The six Thames–Southern North Sea records were obtained from the UK Environment Agency (EA) and the British Oceanographic Data Centre (BODC) National Tide and Sea Level Facility (NTSLF), with Global Extreme Sea Level Analysis (GESLA) Version 4.0 used as a supplementary source; station-specific gauge datum metadata and provider quality-control fields were preserved [37, 38]. Southeast Australia combines Australian Bureau of Meteorology (BoM) Australian Baseline Sea Level Monitoring Project (ABSLMP) records with the GESLA archive [38, 39]. Tokyo–Tokai combines tide-gauge records from the Japanese Oceanographic Data Center (JODC), Japan Coast Guard (JCG), and Japan Meteorological Agency (JMA) [40].

Meteorological forcing was obtained from European Centre for Medium-Range Weather Forecasts Reanalysis version 5 (ERA5) hourly single-level fields [41, 42]. The input channels were 10 m zonal wind, 10 m meridional wind, mean-sea-level-pressure anomaly, and wind-stress magnitude. The model received the station-forcing vector available at the forecast issue time; ERA5 values after the issue time were not supplied as predictors. The geographical eligibility graph was derived from the General Bathymetric Chart of the Oceans (GEBCO) 2026 Grid at its native 15-arc-second resolution [43]. A water-path screen defined eligible station pairs, and water-path distance formed the static spatial prior. Natural Earth 1:10m land and coastline vectors provide reproducible map backgrounds in Figure 1 [44]. The resulting mask

Table 1: Regional datasets and common forecasting task. Directed physical edges count receiver–sender entries. The cycle counts report the training, validation, and test portions after the 72 h boundary embargo.

System	Water-level source	Stations / edges	Train / validation / test cycles	q95 / q99 clusters	Forecast task
Thames–Southern North Sea	UK EA/BODC NTSLF; GESLA 4.0 supplement	6 / 30	5587 / 676 / 688	13 / 7	96 h history; 72 h horizon; 6 h issue stride
New York Bight	NOAA CO-OPS	8 / 56	11080 / 1385 / 1385	11 / 20	96 h history; 72 h horizon; 6 h issue stride
Charleston	NOAA CO-OPS	7 / 42	11069 / 1383 / 1385	11 / 14	96 h history; 72 h horizon; 6 h issue stride
Gulf/Galveston	NOAA CO-OPS	8 / 56	10144 / 1268 / 1268	15 / 13	96 h history; 72 h horizon; 6 h issue stride
Southeast Australia	BoM ABSLMP; GESLA supplement	7 / 30	10220 / 1277 / 1278	12 / 17	96 h history; 72 h horizon; 6 h issue stride
Tokyo–Tokai	JODC, JCG and JMA	11 / 110	9767 / 1198 / 1211	18 / 15	96 h history; 72 h horizon; 6 h issue stride
Los Angeles Bight	NOAA CO-OPS	6 / 30	11655 / 1434 / 1447	21 / 15	96 h history; 72 h horizon; 6 h issue stride

defines the eligible station pairs used by each spatial model.

The water-path calculation used an eight-fold spatial downsampling of the native grid, giving a 120-arc-second working grid. Cells with elevations at or below mean sea level were treated as water, and adjacent water cells were connected through the eight-neighbour grid. Each gauge was snapped to the nearest water cell. Dijkstra's algorithm then calculated the shortest wet-cell path between every ordered station pair using great-circle segment lengths. A pair entered the eligibility mask when a finite path shorter than 1000 km existed within the regional grid; self-edges were excluded. The same path length supplied the continuous distance prior in Equation 12. This construction preserves navigable coastal separation around land barriers while allowing compact regional networks to remain densely connected.

Provider quality flags were applied before the records were placed on a common hourly Coordinated Universal Time (UTC) grid. NOAA and National Tide Gauge Network water levels were paired with their provider tide-prediction series. Where a provider prediction was unavailable, the deterministic tide was reconstructed by least-squares harmonic regression on the first 80% of the station record using the principal semidiurnal and diurnal periods. Water level and tide remained in the same station datum, preserving the residual $r_{i,t} = y_{i,t} - T_{i,t}$ within each station datum. Forecast issue cycles were included when at least one station provided a valid 96 h history. Missing station-history cells were assigned a finite zero value and identified by the hourly availability mask; unavailable target cells were excluded from the loss and evaluation. Tide and atmospheric forcing were complete for every included cycle. Water-level and target gaps were not interpolated.

ERA5 fields were matched to each gauge using the nearest native grid cell. Mean-sea-level pressure was centred by the station mean over the fitting record, and wind-stress magnitude was calculated with an air density of 1.225 kg m^{-3} and drag coefficient 1.3×10^{-3}. No spatial interpolation was applied to the atmospheric fields. These choices provide one station-resolved forcing vector at each issue time and preserve the forcing differences used by the ordered pair routes. The astronomical tide sequence for the next 72 h was supplied separately as the deterministic auxiliary feature.

Table 1 lists the seven systems. The station identities are given in Supplementary Table S1, and the corresponding maps in Figure 1 show the dataset coordinates and directed physical mask. The Thames–Southern North Sea system contains six stations and 30 directed edges. Tokyo–Tokai contains eleven stations and 110 directed edges. The other systems span six to eight stations and include 30–56 directed edges. The event-cluster counts for the matched q95 and q99 summaries appear in Table 1 and Supplementary Table S7.

The seven systems form a common experiment with distinct network geometries and event counts. Station counts range from six to eleven, and the regional networks differ in gauge spacing, coastal orientation, and spatial forcing contrast. Thames–Southern North Sea and Gulf/Galveston contain compact multi-station networks with clear along-coast organization. New York Bight, Charleston, and Los Angeles Bight combine embayments with adjacent open coasts. Southeast Australia and Tokyo–Tokai extend across broader coastal margins and include wider differences in station spacing and orientation. These layouts provide different opportunities for stable distance-weighted exchange, repeatable station ordering, and atmosphere-conditioned coupling.

The q95 event counts range from 11 to 21 independent clusters and the q99 counts from 7 to 20. For the Thames–Southern North Sea, the test catalogue contains 13 q95 and 7 q99 clusters. The station-resolved ERA5 fields likewise differ in their within-network contrast, so the same regional weather system can appear nearly coherent across one network and more heterogeneous across another. Several compact

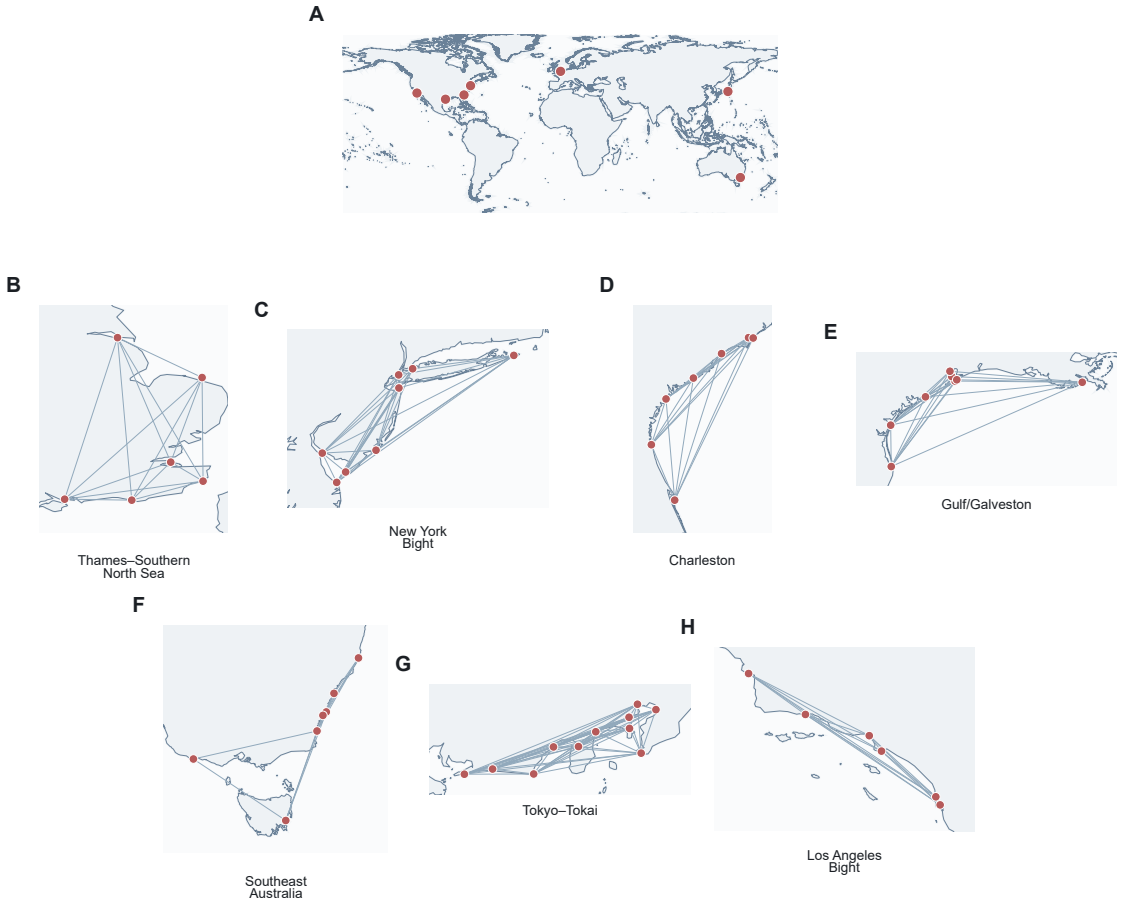


Figure 1: Global setting and physical station networks. Panel A shows the seven coastal systems using their dataset station coordinates on Natural Earth 1:10m coastlines. Panels B–H show, in order, the Thames–Southern North Sea, New York Bight, Charleston, Gulf/Galveston, Southeast Australia, Tokyo–Tokai, and Los Angeles Bight networks, including their local coastlines, station locations, and masks. Lines indicate the eligible station connections used by the graph routes.

systems remain densely connected after the water-path screen; in those systems, geography acts mainly through the continuous distance prior. Southeast Australia has a sparser eligibility mask and a broader coastal arrangement. Using the same forecast horizon, issue stride, target construction, initialisation count, and route definitions across the seven systems allows route effects to be compared alongside network scale, station arrangement, atmospheric contrast, and the amount of independent high-water evidence.

2.2 Forecast target and chronological evaluation

The model predicted the non-tidal water-level residual. Let $y_{i,t}$ denote total water level at station i and time t, and let $T_{i,t}$ denote the deterministic tide. Equation 1 defines the residual target.

$$r_{i,t} = y_{i,t} - T_{i,t}. \tag{1}$$

Every forecast used 96 h of station history and produced a 72 h residual trajectory at a six-hour issue stride. Deterministic future tide was supplied as a known astronomical feature and restored for the total-water evaluation. The hourly availability indicator distinguished valid observations from the finite zero value used for unavailable history cells. Target masks restricted optimization and scoring to available observations. The total-water trajectory was used for all reported accuracy, tail, detection, peak, and warning metrics.

Each dataset used a chronological 8:1:1 partition with a 72 h boundary embargo. Station thresholds, topology priors, lag associations, and correction scales were fitted from the training portion of the record. The four atmospheric-forcing channels were standardized separately within each region before entering a forcing-dependent route. For channel c, the standardized value was

$$\widetilde{f}_{i,t,c} = \frac{f^{\mathrm{raw}}_{i,t,c} - \mu^{\mathrm{train}}_{r,c}}{\max\left(\sigma^{\mathrm{train}}_{r,c}, 10^{-6}\right)}, \qquad c = 1, \ldots, 4, \tag{2}$$

where $\mu^{\mathrm{train}}_{r,c}$ and $\sigma^{\mathrm{train}}_{r,c}$ were calculated over all stations and issue cycles in the training portion of regional system r. The same statistics were applied to validation and test cycles. Pair forcing was reconstructed from the standardized receiver and sender vectors and their signed difference. Water-level targets, deterministic tide, thresholds, distance priors, and correction scales remained in physical units. Validation loss selected the local and route checkpoints, and the test segment was then scored without parameter updates. Three independent random initialisations were used for every principal route.

2.3 Regional experiment and local forecast

The same analysis was applied to all seven regional networks. Before fitting, the station histories, deterministic tide, atmospheric fields, high-water thresholds, and 96–72 h windows were aligned for each dataset. For each regional system and random initialisation, a Local-GRU was trained with shared recurrent parameters across stations while its output preserved station-resolved trajectories through the node dimension and future-tide input. For station i and historical hour s, the encoder input and issue-time state were

$$\begin{aligned} x_{i,s} &= \left[r_{i,s},\ \Delta r_{i,s},\ T_{i,s},\ \Delta T_{i,s},\ y_{i,s},\ \Delta y_{i,s}\right]^{\top}, \\ h_{i,t} &= \mathrm{GRU}_{\theta}\left(x_{i,t-L+1:t}\right) + E_a\left(\frac{1}{L}\sum_{s=t-L+1}^{t} a^{\mathrm{hist}}_{i,s}\right), \end{aligned} \tag{3}$$

where r, T, and y denote non-tidal residual, deterministic tide, and observed total water level; Δ is the one-hour first difference; $a^{\mathrm{hist}}_{i,s}$ is the recorded history-availability indicator; E_a is its learned embedding; t is issue time; and $L = 96$ h. The time-indexed encoder states used by the delayed routes receive the corresponding hourly availability embedding $E_a(a^{\mathrm{hist}}_{i,s})$. Future tide and issue-time atmospheric forcing are supplied separately and are not channels of $x_{i,s}$. The local head produced a residual trajectory and restored deterministic tide for lead $k = 1, \ldots, 72$:

$$r^{\mathrm{loc}}_{i,t+k} = f_{\mathrm{loc}}\left(h_{i,t}, T_{i,t+1:t+72}\right)_k, \qquad B_{i,t+k} = T_{i,t+k} + r^{\mathrm{loc}}_{i,t+k}. \tag{4}$$

The local model was optimized under the settings reported in Supplementary Table S8. Let $\widehat{r}$ denote the residual output being optimized, let $\mathcal{V}$ contain valid station–issue–lead samples, and let $Q_{95} = \{(i,t,k) \in \mathcal{V} : y_{i,t+k} \geq q^{\mathrm{tot}}_{0.95,i}\}$ contain samples whose observed total water level exceeds the station-specific training q95 threshold. The optimized loss was

$$\mathcal{L}_{\mathrm{tail}} = \frac{1}{|\mathcal{V}|}\sum_{(i,t,k)\in\mathcal{V}}\left(\widehat{r}_{i,t+k} - r_{i,t+k}\right)^2 + \lambda_{95}\frac{1}{|Q_{95}|}\sum_{(i,t,k)\in Q_{95}}\left(\widehat{r}_{i,t+k} - r_{i,t+k}\right)^2. \tag{5}$$

Both terms measure residual forecast error, and λ_{95} gives additional weight to samples occurring above the total-water q95 threshold. Validation loss selected the checkpoint under the stopping rule in Supplementary Table S8. The selected encoder, availability embedding, and local head were then fixed. Their forecast $B_{i,t+k}$ and state $h_{i,t}$ supplied the local reference for every spatial model in the corresponding system and initialisation. Supplementary Table S9 summarizes this setup for all seven systems.

2.4 Spatial models and training

$$\widehat{y}^{(m)}_{i,t+k} = B_{i,t+k} + R^{(m)}_{M,i,t+k}, \qquad k = 1, \ldots, 72. \tag{6}$$

Table 2: Spatial model configurations and the information tested. D_m activates dynamic state attention, P_m activates ordered pair forcing, and R_m activates receiver-level forcing.

Route configuration	D_m	P_m	R_m	Delay state	Information tested
Local-GRU reference	0	0	0	None	What is the added value of spatial information?
Static connectivity	0	0	0	Zero	Does fixed distance-weighted adjacency add information?
Directed-delay propagation	1	0	0	Directed	Does repeatable cross-station ordering add information?
Regional-forcing coupling	0	0	1	Directed	Does receiver-level atmospheric conditioning add a spatial increment?
Dynamic pair coupling	1	1	0	Zero	Does event-dependent pair information add value at zero lag?
Lagged pair coupling	1	1	0	Directed	Do directed delays and ordered forcing contrasts add complementary information?

For regional system r, let $\mathcal{I}_r$ denote its physical mask, water-path prior, directed lag matrix, forcing fields, and station-availability indicators. The spatial model uses these quantities to produce graph state $g^{(m)}$ and bounded correction $R_M^{(m)}$:

$$(B, h, \mathcal{I}_r) \longrightarrow g^{(m)} \longrightarrow R_M^{(m)} \longrightarrow \widehat{y}^{(m)} = B + R_M^{(m)}. \tag{7}$$

The architecture, regularization, and optimization settings for the local and spatial components are listed in Supplementary Table S8. The station-forcing control supplied receiver forcing to a bounded residual head without cross-station exchange. Table 2 states the forecasting question addressed by each configuration, and Figure 2 summarizes the comparison.

Static connectivity tests the fixed water-path and distance prior. Directed-delay propagation adds receiver–sender ordering without pairwise atmospheric features. Regional-forcing coupling retains the directed graph and lagged sender state and adds forcing at the receiving station. Dynamic pair coupling uses zero-lag state exchange and ordered receiver–sender forcing. Lagged pair coupling combines directed state access with the same ordered pair forcing. The station-forcing control has no cross-station exchange.

Each targeted comparison addresses a specific operation. No-delay evaluates simultaneous exchange in place of the estimated lag. Shuffled-delay tests whether the learned organization of the lag matrix carries information. No-direction removes sender–receiver asymmetry by setting each pair of delays to $\min(\tau_{ij}, \tau_{ji})$ and averaging the two distance-prior weights before row normalization. No-pair-forcing removes the ordered atmospheric contrast, and no-common-mode removes the regional centring operation. These comparisons assess the contribution of delay, directional asymmetry, pair forcing, and common-mode centring within the selected regional configurations.

2.5 Spatial representation and bounded correction

For model m, graph state $g_{i,t}^{(m)}$ produces bounded correction $R_{M,i,t+k}^{(m)}$ at station i, issue time t, and forecast lead k. The local reference has $R_{M,i,t+k}^{(m)} = 0$.

For routes with common-mode removal, the message-passing branch centres the hidden representation across available stations. Let $a_{i,t}$ be the availability indicator. Equation 8 defines the centred state.

$$\bar{h}_t = \frac{\sum_i a_{i,t} h_{i,t}}{\max\left(\sum_i a_{i,t}, 1\right)}, \qquad \widetilde{h}_{i,t} = h_{i,t} - \bar{h}_t. \tag{8}$$

For compact notation, Static connectivity, Directed-delay propagation, Regional-forcing coupling, Dynamic pair coupling, and Lagged pair coupling are denoted by S, D, R, P_0, and P_τ, respectively. Equation 9 specifies the route state.

$$z_{i,t}^{(m)} = \begin{cases} \widetilde{h}_{i,t}, & m \in \{\mathrm{D}, \mathrm{P}_0, \mathrm{P}_\tau\}, \\ h_{i,t}, & m \in \{\mathrm{S}, \mathrm{R}\}. \end{cases} \tag{9}$$

The corresponding lag is given by Equation 10.

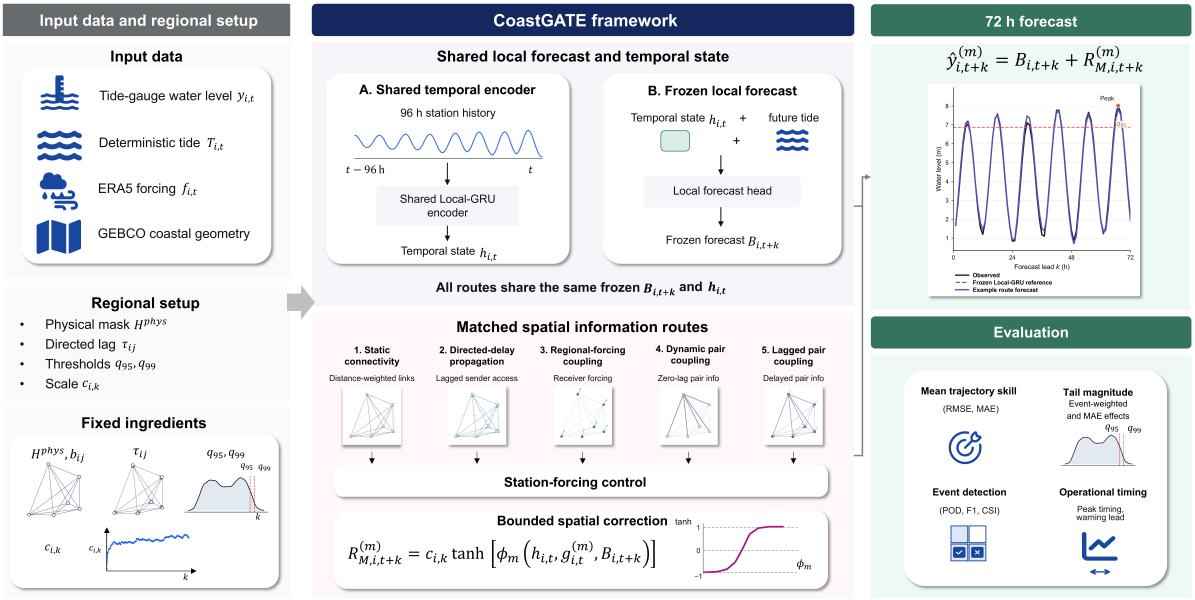


Figure 2: CoastGATE forecasting and evaluation framework. Tide-gauge observations, deterministic tide, atmospheric forcing, coastal geometry, regional graph inputs, and high-water thresholds support a shared local forecast and temporal state. Five spatial information routes generate bounded corrections for the 72 h forecast, which is evaluated through mean trajectory skill, high-water magnitude, event detection, and hindcast-derived warning timing.

$$\tau_{ij}^{(m)} = \begin{cases} \tau_{ij}, & m \in \{\mathrm{D}, \mathrm{R}, \mathrm{P}_\tau\}, \\ 0, & m \in \{\mathrm{S}, \mathrm{P}_0\}. \end{cases} \tag{10}$$

Directed lags were estimated from the issue-time non-tidal residual series in the training segment. Let $u_{i,n}$ be the residual at station i for the nth six-hour issue cycle. For receiver i and sender j, the lag statistic and selected delay were

$$\rho_{ij}(\ell) = \mathrm{corr}\left(u_{j,n}, u_{i,n+\ell/6}\right), \qquad \tau_{ij} = \underset{\ell \in \{0,6,12,18,24\}\,\mathrm{h}}{\arg\max} |\rho_{ij}(\ell)|. \tag{11}$$

The Pearson correlation was evaluated on finite paired cycles. When two candidate lags had the same absolute correlation, the smaller lag was selected. Rows index receivers and columns index senders, so $\tau_{ij} > 0$ means that receiver i accesses the state of sender j at $t - \tau_{ij}$. This correlation step determines τ_{ij}; the distance term enters separately through b_{ij}.

Water-path distance enters the graph independently through the static edge prior. For water-path distance d_{ij} in kilometres and eligibility mask H_{ij}^{phys}, the implemented prior is

$$\widetilde{w}_{ij} = H_{ij}^{\mathrm{phys}} \exp\left(-\frac{d_{ij}}{200\ \mathrm{km}}\right), \qquad w_{ij} = \frac{\widetilde{w}_{ij}}{\sum_{j'} \widetilde{w}_{ij'}}, \qquad b_{ij} = \begin{cases} \log\left(\max(w_{ij}, 10^{-6})\right), & H_{ij}^{\mathrm{phys}} = 1, \\ -10^{9}, & H_{ij}^{\mathrm{phys}} = 0. \end{cases} \tag{12}$$

The lag matrix controls which historical sender state is read; b_{ij} controls its static weight before any dynamic score is added.

The static bias b_{ij} combines the eligible edge with the water-path-distance prior. The standardized atmospheric-forcing vector at station i is $\widetilde{f}_{i,t}$, with channels for 10-m zonal wind, 10-m meridional wind, mean-sea-level-pressure anomaly, and wind-stress magnitude. Equation 13 defines the ordered pair feature used by Dynamic pair coupling and Lagged pair coupling.

$$p_{ij,t} = \left[\widetilde{f}_{i,t},\ \widetilde{f}_{j,t},\ \widetilde{f}_{i,t} - \widetilde{f}_{j,t}\right]. \tag{13}$$

The Regional-forcing coupling route adds a receiver-level forcing embedding $\rho(\widetilde{f}_{i,t})$ to its graph state. With indicators D_m, P_m, and R_m for dynamic attention, pair forcing, and receiver-level forcing, respectively, Equation 14 defines the attention score.

$$e_{ij,t}^{(m)} = b_{ij} + D_m \frac{Q\left(z_{i,t}^{(m)}\right)^{\mathsf{T}} K\left(z_{j,t-\tau_{ij}^{(m)}}^{(m)}\right)}{\sqrt{d_h}} + P_m \psi\left(p_{ij,t}\right). \tag{14}$$

Here $H_{ij}^{\mathrm{phys}} = 1$ marks an admissible directed physical edge, $a_{i,t} = 1$ marks an available station state, Q and K are learned query and key projections, V is the value projection, d_h is the query–key dimension, b_{ij} is the topology log-prior, ψ maps ordered forcing to an edge score, and ρ maps receiver forcing to the route state. The route-specific correction head is denoted by ϕ_m, and $c_{i,k}$ sets its station- and lead-specific amplitude.

Equation 15 gives the effective edge mask, normalized weights, and graph update for the route definitions above.

$$\begin{aligned} M_{ij,t}^{\mathrm{eff},(m)} &= H_{ij}^{\mathrm{phys}} a_{i,t} a_{j,t} a_{j,t-\tau_{ij}^{(m)}}, \\ \alpha_{ij,t}^{(m)} &= \frac{M_{ij,t}^{\mathrm{eff},(m)} \exp\left(e_{ij,t}^{(m)}\right)}{\sum_{j'} M_{ij',t}^{\mathrm{eff},(m)} \exp\left(e_{ij',t}^{(m)}\right)}, \\ g_{i,t}^{(m)} &= \mathrm{GRUCell}\left(\sum_j \alpha_{ij,t}^{(m)} V\left(z_{j,t-\tau_{ij}^{(m)}}^{(m)}\right),\ z_{i,t}^{(m)} + R_m \rho(\widetilde{f}_{i,t})\right). \end{aligned} \tag{15}$$

Equation 16 applies the graph correction with the station- and lead-specific scale $c_{i,k}$.

$$R_{M,i,t+k}^{(m)} = c_{i,k} \tanh\left(\phi_m\left(h_{i,t}, g_{i,t}^{(m)}, B_{i,t+k}\right)\right), \qquad \widehat{y}_{i,t+k}^{(m)} = B_{i,t+k} + R_{M,i,t+k}^{(m)}. \tag{16}$$

Equation 17 sets the scale to the 0.90 quantile of the local residual error distribution.

$$c_{i,k} = Q_{0.90}\left(\left\{\left|r_{i,t+k} - r_{i,t+k}^{\mathrm{loc}}\right| : t \in \mathcal{T}_{\mathrm{fit}}\right\}\right), \tag{17}$$

where $\mathcal{T}_{\mathrm{fit}}$ denotes the training cycles used to calibrate the local residual scale and r^{loc} is the Local-GRU prediction on those cycles. The water-path screen defines eligible station pairs and the distance term supplies a continuous geographic prior. Several compact networks remain densely connected after screening, so geographic information enters mainly through edge eligibility and weighting.

2.6 Metrics and uncertainty

Continuous metrics were RMSE, MAE, signed bias, absolute bias, and coefficient of determination [45, 46]. Station-specific q95 and q99 thresholds were fixed from the training distribution of total water level. Tail magnitude was evaluated by MAE above the appropriate threshold. Probability of detection (POD), precision, F1 score, critical success index (CSI), and false-alarm ratio (FAR) characterized event detection. Peak-amplitude error, peak-timing error, and warning-lead difference described the event trajectory and threshold-crossing timing.

The principal tail estimand was the matched event-cluster effect. High-water catalogues were constructed from the unique hourly observation timeline. Forecast errors kept their original issue-cycle and lead-time indices, so overlapping forecast windows remained separate forecast instances. Station-specific q95 and q99 thresholds were estimated from the training distribution, and consecutive exceedance periods separated by more than 48 h formed distinct clusters. The q95 and q99 catalogues were constructed independently. Every six-hour issue cycle whose 72 h horizon contained a cluster peak received the same event label, giving each event equal weight. For candidate route m and the frozen Local-GRU reference b, Equation 18 averages the q95 MAE difference across independent high-water clusters. Negative values

indicate lower candidate error. The q99 effect uses the corresponding q99 clusters. For each initialisation, the q95 and q99 intervals were computed from exactly 2,000 equal-event resamples [47–49]. Table 3 reports the mean effect and 95% t interval across the three initialisations. Supplementary Tables S2 and S3 list route-level point metrics, while Supplementary Table S4 gives the three-initialisation results for Lagged pair coupling. Event-bootstrap intervals quantify sampling variation within an initialisation; the t intervals quantify variation among the three initialisations.

The primary event definition uses total water because it corresponds to the level relevant to coastal flood warning. A secondary residual-tail view used station-specific residual q95 and q99 thresholds fitted on the training segment. The event groups and issue-cycle membership remained those of the total-water catalogue, while the pointwise exceedance mask was recomputed from the residual thresholds. This view reused the frozen residual predictions and the 2,000-resample event-bootstrap procedure to describe residual forecast behaviour during the same engineering-relevant high-water events.

$$\Delta_{q95}^{(m)} = \frac{1}{|\mathcal{E}_{q95}|} \sum_{e \in \mathcal{E}_{q95}} \left[\mathrm{MAE}_{q95,e}^{(m)} - \mathrm{MAE}_{q95,e}^{(b)} \right]. \tag{18}$$

Figure 3 presents the regional route matrix with square cells and no in-cell numerical values. The numerical effects and intervals appear in Table 3 and Supplementary Tables S2–S4. Validation loss fixed each route checkpoint, and all routes were then scored on the test segment. The regional route labels below summarize those test results.

3 Results

3.1 Overall value of spatial coupling

Lagged pair coupling reduced total-water RMSE relative to the common Local-GRU forecast in every system, with reductions of 6.98% in the Thames–Southern North Sea, 1.41% in New York Bight, 1.89% in Charleston, 2.71% in Gulf/Galveston, 3.38% in Southeast Australia, 4.89% in Tokyo–Tokai, and 1.91% in Los Angeles Bight. Spatial exchange added mean-trajectory value across the full range of station layouts. Upper-tail effects varied more strongly by region and route than the aggregate RMSE effects. The seven-system route matrix and the corresponding route-level RMSE distributions are shown in Figure 3 and Supplementary Fig. S2.

The station-forcing control provides a direct comparison for meteorological input without cross-station exchange. It lowered RMSE in all seven systems, showing that station-resolved atmospheric inputs contribute to broad trajectory reconstruction. In Gulf/Galveston, the control produced a q95 effect of −0.0083 m, compared with −0.0107 m for Regional-forcing coupling. In the Thames–Southern North Sea, the corresponding values were −0.0118 m and −0.0213 m. The control also improved q95 in Tokyo–Tokai, whereas its q95 effects were positive in New York Bight and Charleston. Station-level atmospheric input contributed to the forecast, and the graph-based forcing route produced a larger high-water effect in Gulf/Galveston and the Thames–Southern North Sea.

A persistence forecast supplied a direct reference for the 72 h horizon. It carried the last issue-time non-tidal residual through the forecast window and restored the deterministic tide. Lagged pair coupling had lower test RMSE than this persistence reference in all seven systems, with the largest separation in the Thames–Southern North Sea (Supplementary Table S10). This comparison places the learned route results alongside the information available in the latest observation.

The residual-tail analysis in Supplementary Table S11 uses the same event clusters and recomputes the exceedance mask from the residual thresholds. The paired views show route behaviour in the residual and reconstructed total-water representations of each high-water event.

Table 3 and Figure 4 show that the lowest-RMSE and q95-leading models differ in several systems. Broad trajectory reconstruction and high-water magnitude therefore favour different spatial information in some coastal networks.

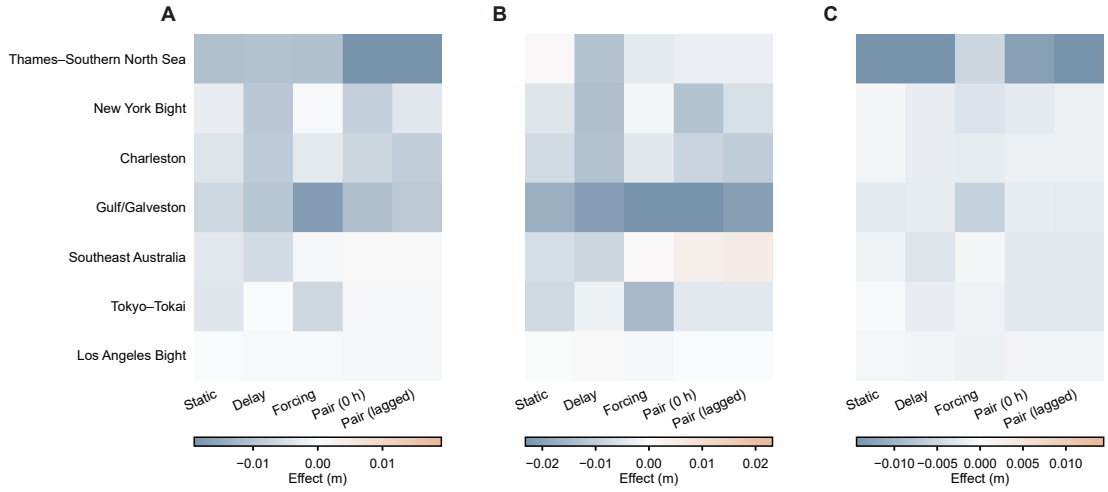


Figure 3: Route effects across the seven coastal networks. Panel A shows q95 event-weighted MAE effects, Panel B shows q99 effects, and Panel C shows RMSE effects. Every cell is candidate minus the frozen Local-GRU reference. Blue denotes lower candidate error and peach denotes higher candidate error. Numerical values and intervals are given in Table 3 and Supplementary Tables S2–S4.

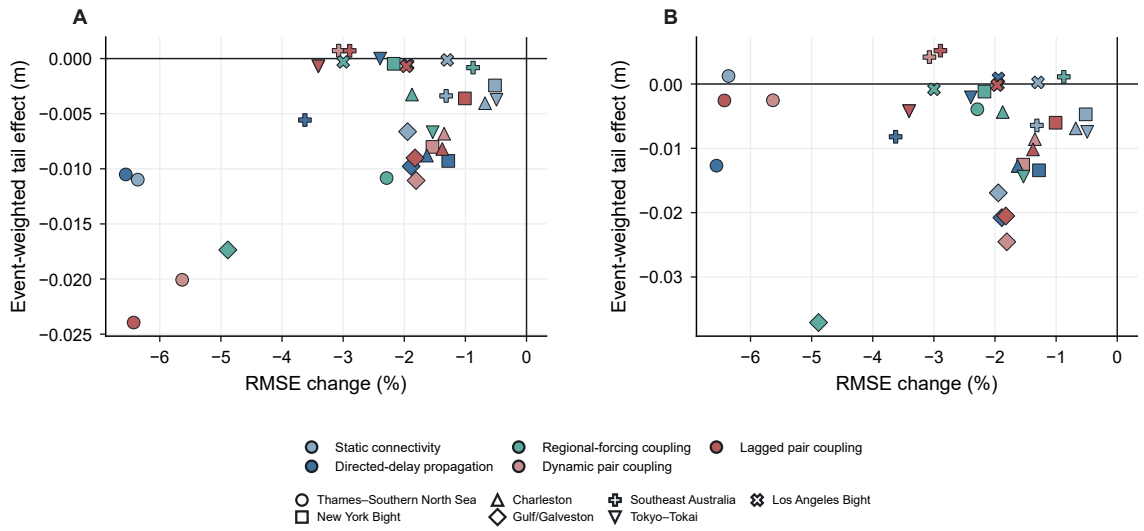


Figure 4: Relationship between mean and high-water effects. Panel A relates RMSE change to q95 event effects and Panel B gives the corresponding q99 relationship. The horizontal and vertical reference lines mark zero route-minus-local-reference effect.

Table 3: Regional spatial-model profiles. The upper block reports q95 and q99 effects for the model with the lowest test q95 effect. The lower block identifies the lowest-RMSE model. Negative effects denote lower candidate error. Intervals summarize variation across the three initialisations; warning-lead results are reported in Section 3.3 and Supplementary Fig. S3.

High-water profile

System	q95-leading route	q95 effect [95% interval]	q99 effect [95% interval]
Thames–Southern North Sea	Lagged pair coupling	-0.0213 [-0.0364, -0.0094]	-0.0107 [-0.0311, +0.0038]
New York Bight	Directed-delay propagation	-0.0020 [-0.0037, -0.0004]	-0.0035 [-0.0057, -0.0015]
Charleston	Directed-delay propagation	-0.0007 [-0.0025, +0.0009]	-0.0022 [-0.0059, +0.0010]
Gulf/Galveston	Regional-forcing coupling	-0.0107 [-0.0191, -0.0023]	-0.0349 [-0.0485, -0.0214]
Southeast Australia	Directed-delay propagation	-0.0038 [-0.0066, -0.0016]	-0.0047 [-0.0107, +0.0018]
Tokyo–Tokai	Lagged pair coupling	-0.0018 [-0.0028, -0.0007]	-0.0057 [-0.0120, +0.0006]
Los Angeles Bight	Regional-forcing coupling	-0.0006 [-0.0008, -0.0003]	-0.0004 [-0.0013, +0.0006]

Mean-error profile

System	Lowest-RMSE model	RMSE change (%)
Thames–Southern North Sea	Dynamic pair coupling	-7.43
New York Bight	Regional-forcing coupling	-4.07
Charleston	Regional-forcing coupling	-3.74
Gulf/Galveston	Regional-forcing coupling	-5.58
Southeast Australia	Directed-delay propagation	-3.63
Tokyo–Tokai	Dynamic pair / Lagged pair coupling	-4.89
Los Angeles Bight	Regional-forcing coupling	-4.08

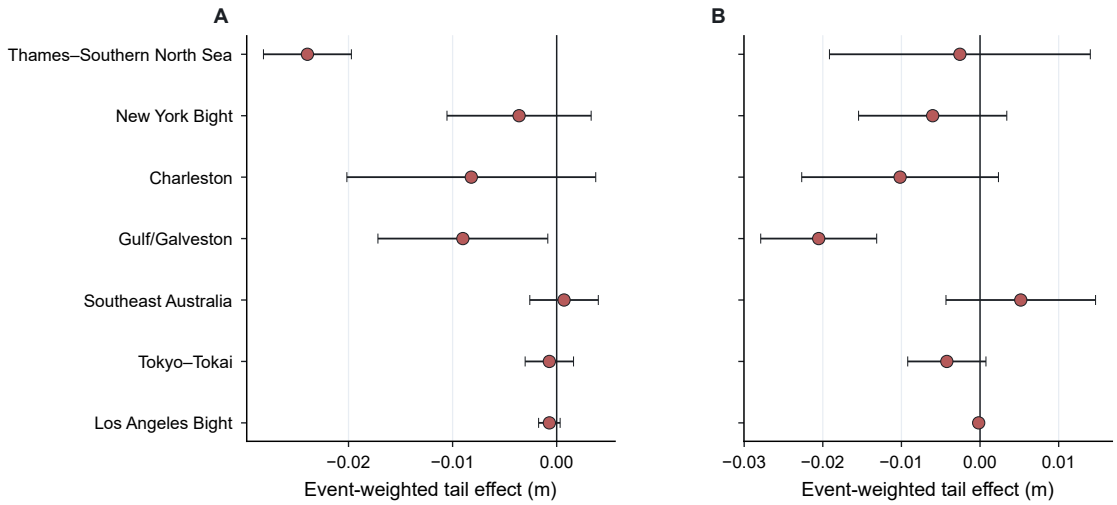


Figure 5: High-water effects of Lagged pair coupling across the seven systems. Panel A shows q95 event effects and Panel B q99 effects. Points give route-minus-local-reference effects; negative values indicate lower candidate error, and horizontal intervals summarize initialisation variation.

3.2 Regional route profiles

The Thames–Southern North Sea provides the clearest combined-coupling result. Lagged pair coupling produced the largest q95 event reduction, while Dynamic pair coupling gave the lowest RMSE at 7.43% below the local reference. The added directed sender access in Lagged pair coupling produced the strongest tail effect, separating the route that best reconstructed the full trajectory from the route that best corrected the q95 events. Figure 5 shows the Lagged pair effects across the three initialisations.

Gulf/Galveston gives the strongest regional-forcing profile. Regional-forcing coupling led q95, q99, and RMSE, with q95 and q99 effects of -0.0107 m and -0.0349 m. Their three-initialisation intervals were $[-0.0191, -0.0023]$ m and $[-0.0485, -0.0214]$ m. This route combines the directed distance-weighted graph, delayed sender states, and receiver-level forcing. Its agreement across mean and upper-tail outcomes distinguishes Gulf/Galveston from the other six systems.

Southeast Australia supplies the clearest directed-delay result. Directed-delay propagation led RMSE and q95, with a q95 interval below zero; its q99 point effect was also negative, with a wider interval. New York Bight placed Directed-delay propagation first for q95 and q99, while Charleston placed it first for q95 and q99 point estimates. Their interval patterns differ, and the rankings describe regional route profiles with varying interval support.

Tokyo–Tokai and Los Angeles Bight showed smaller upper-tail increments. In Tokyo–Tokai, Dynamic pair coupling and Lagged pair coupling gave the largest RMSE reduction, while Lagged pair coupling had the most negative q95 and q99 point effects. Los Angeles Bight showed small negative q95 effects across the principal routes, with q99 effects close to zero. Figure 6 shows the q95 and q99 route profiles for all seven systems.

Taken together, the systems form four regional profiles. The Thames–Southern North Sea favours combined delayed and pair information for q95 magnitude, Gulf/Galveston favours receiver-level forcing across mean and upper-tail measures, and Southeast Australia shows the clearest directed-delay profile. New York Bight and Charleston place Directed-delay propagation first in the point comparison, while Tokyo–Tokai and Los Angeles Bight show smaller, metric-dependent tail increments. RMSE and q95 rankings agree in Gulf/Galveston and Southeast Australia, with different route orderings elsewhere. The comparison shows which spatial information pathway contributes most strongly within each regional network and how its value changes with the forecast objective.

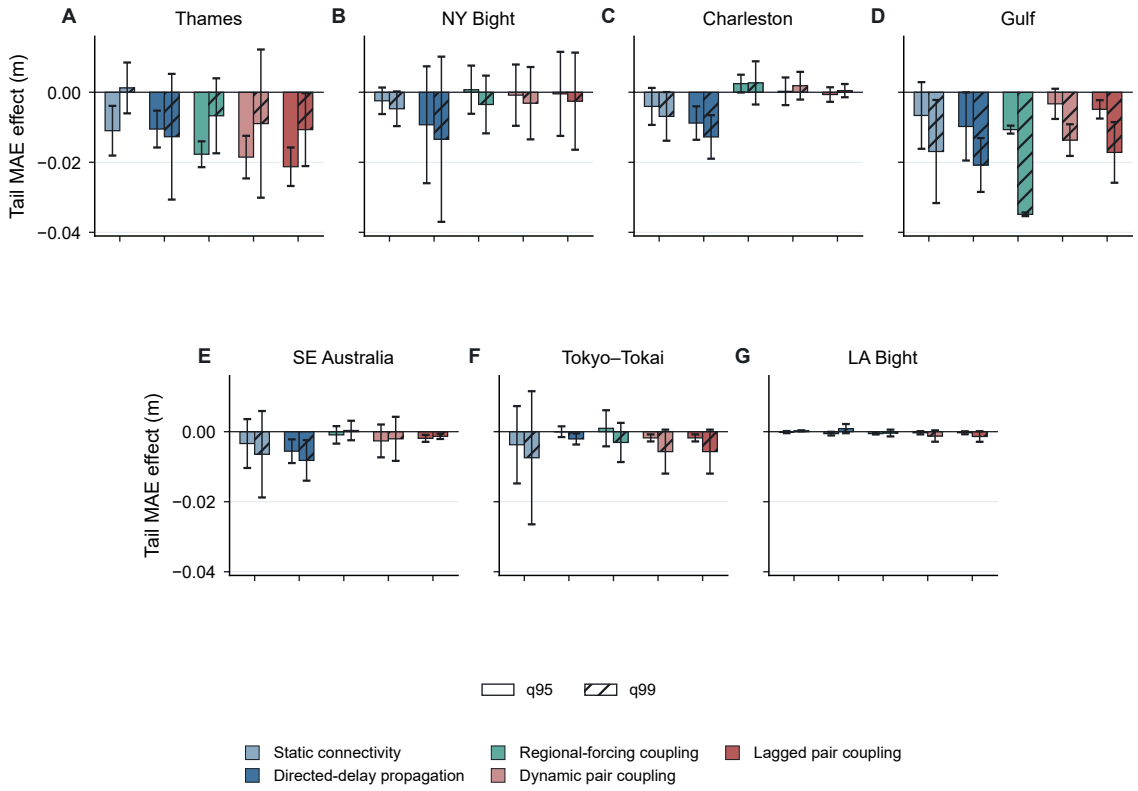


Figure 6: Regional q95 and q99 profiles. Panel titles identify the seven coastal systems. Each panel compares the five spatial routes with the frozen Local-GRU reference; solid bars show q95 effects, hatched bars show q99 effects, and whiskers show initialisation intervals. Route colours follow Table 2.

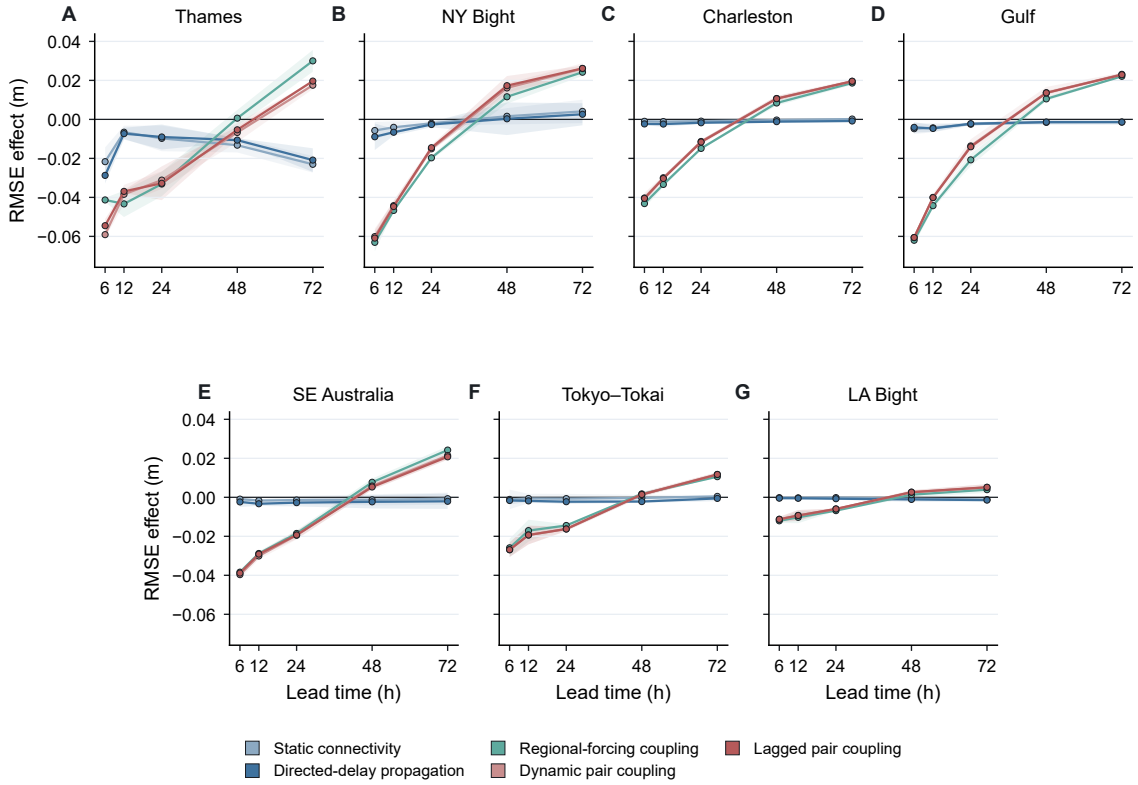


Figure 7: Horizon-wise RMSE effects. Panel titles identify the seven coastal systems. Curves give route-minus-local-reference effects at 6, 12, 24, 48, and 72 h; shaded bands show initialisation intervals.

3.3 Horizon, detection, peak, and warning behavior

The horizon profiles show when the route increments emerge over the 6–72 h trajectory. In the Thames–Southern North Sea, Lagged pair coupling and Directed-delay propagation remain below the local-reference RMSE at several leads. Gulf/Galveston shows a broad Regional-forcing advantage over early and intermediate leads. Tokyo–Tokai and Los Angeles Bight retain mean improvements with smaller upper-tail effects, showing that trajectory reconstruction and event magnitude can rank the same spatial route differently. These profiles are shown in Figure 7.

Detection and timing add two further distinctions. Route-specific changes in POD and F1 score were evaluated at q95 and q99, and the q95-leading route advanced the first threshold crossing by 4.47 h in Gulf/Galveston and 3.21 h in New York Bight. The corresponding changes were −1.33 h in Charleston, −6.92 h in Southeast Australia, and −2.90 h for Lagged pair coupling in the Thames–Southern North Sea. Event-magnitude error and warning lead follow separate route orderings. Detection, peak timing, and warning lead measure different parts of the high-water trajectory and provide distinct evidence for warning-oriented evaluation; the route profiles are shown in Figure 8 and Supplementary Fig. S3.

Horizon, detection, timing, and event-level distributions for all seven systems are provided in Supplementary Figs. S1–S6.

3.4 Event-level evidence and targeted component checks

In the Thames–Southern North Sea panel, the local reference follows the repeated oscillations well and the Lagged pair line stays close through the ordinary portions of the horizon. Its clearest correction appears on the rising limb and around the final high-water crest, where the local trajectory is slightly low before the peak. The route keeps the peak phase close to the observation while adjusting the level across the rise and crest.

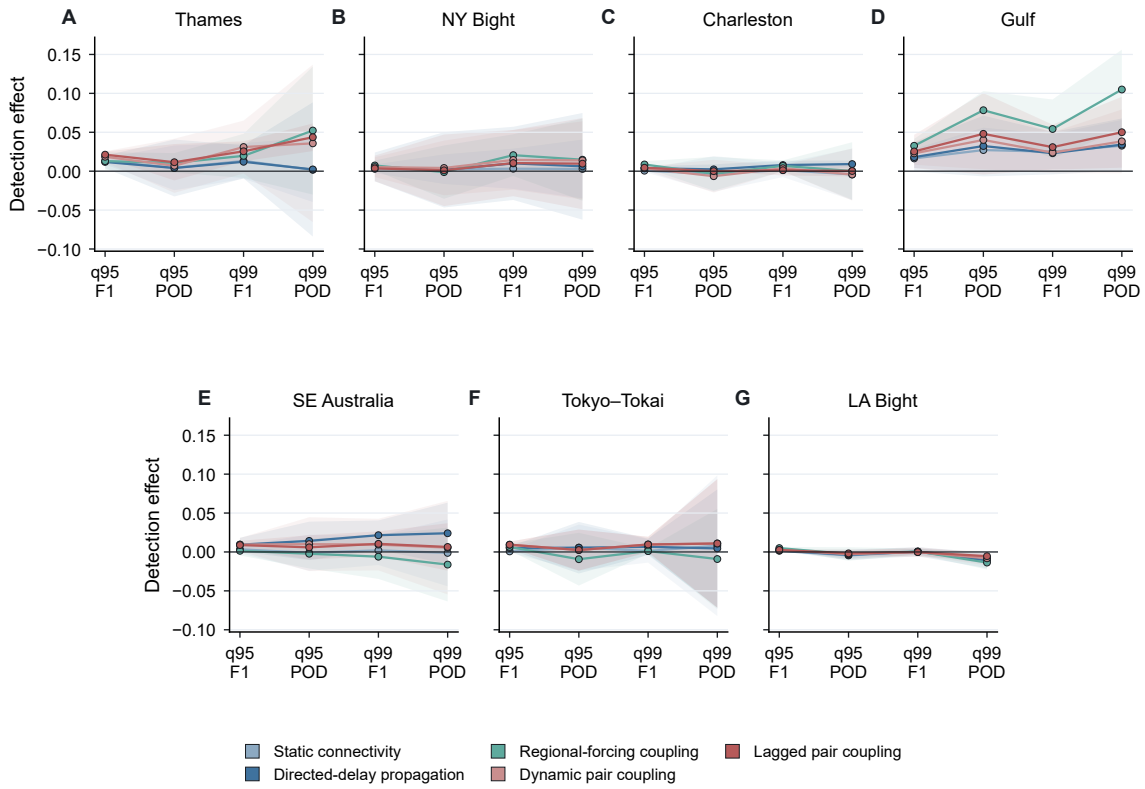


Figure 8: Detection profiles for q95 and q99 thresholds. Panel titles identify the seven coastal systems. Lines give route-minus-local-reference changes in q95 F1, q95 POD, q99 F1, and q99 POD. Supplementary Figs. S4 and S5 report the full detection matrices.

Table 4: Southeast Australia route dissection. Effects use the frozen Local-GRU reference B and temporal state h. Negative tail and RMSE changes denote lower candidate error.

Route	RMSE change	q95 effect [95% interval]	q99 effect [95% interval]	q95 F1 change	Peak-amplitude MAE change (m)	Warning-lead change (h)
Static connectivity	-1.32%	-0.0006 [-0.0020, +0.0005]	-0.0011 [-0.0044, +0.0018]	+0.003	-0.0004	-7.56
Directed-delay propagation	-3.63%	-0.0038 [-0.0066, -0.0016]	-0.0047 [-0.0107, +0.0018]	+0.009	-0.0029	-6.92
Regional-forcing coupling	-0.82%	-0.0009 [-0.0041, +0.0011]	+0.0003 [-0.0039, +0.0045]	+0.002	+0.0007	-1.00
Dynamic pair coupling	-3.58%	-0.0026 [-0.0050, -0.0005]	-0.0020 [-0.0075, +0.0031]	+0.010	-0.0024	-5.81
Lagged pair coupling	-3.38%	-0.0019 [-0.0039, +0.0001]	-0.0013 [-0.0066, +0.0038]	+0.009	-0.0024	-8.64

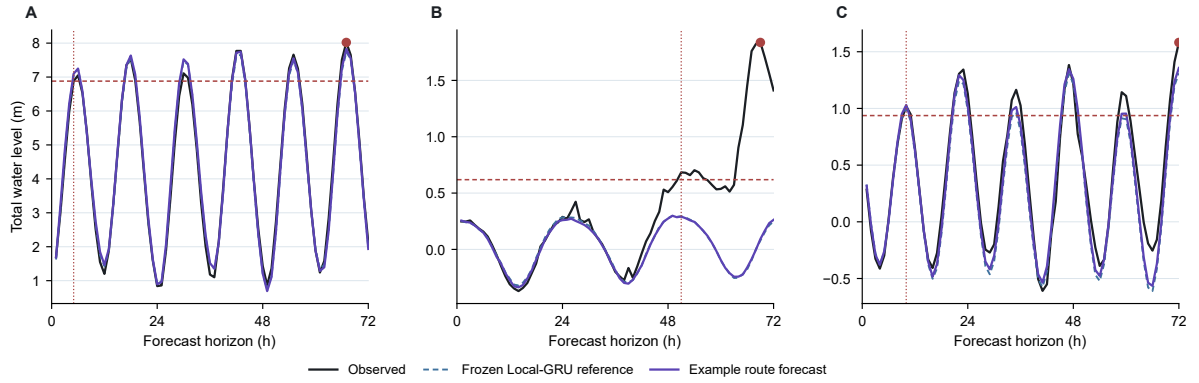


Figure 9: High-water trajectories from three regional route configurations. Panel A shows the Thames–Southern North Sea with Lagged pair coupling, Panel B Gulf/Galveston with Regional-forcing coupling, and Panel C Charleston with Directed-delay propagation. For each region, the station–cycle window with the largest observed q95 exceedance among the test-set trajectories was selected. Curves show observed total water level, the frozen Local-GRU reference, the example route forecast, and the station-specific q95 threshold; vertical and point markers identify the first threshold crossing and observed peak.

The Gulf/Galveston panel shows a broad Regional-forcing correction. The route changes the level and curvature through the approach to the threshold and over the surrounding falling limb. The late observed crest remains larger than both forecast curves in this window, illustrating why the regional q95 effect is estimated across the full event catalogue. The first-crossing and peak markers keep timing separate from amplitude.

In Charleston, Directed-delay propagation follows the rising and falling portions of the successive oscillations, with the largest differences appearing near the threshold crossings and the later peak. The route changes phase and amplitude together, while the peak remains close in time to the observed maximum. Across the three cases, route corrections appear on the rising limb, crest, or falling limb. The event curves connect those trajectory changes to the separate magnitude, detection, peak, and warning metrics used in the seven-system comparison (Figure 9).

The targeted component checks in Supplementary Fig. S8 examine individual route ingredients. Thames–Southern North Sea includes no-delay, no-pair-forcing, and no-common-mode routes. New York Bight includes no-delay and shuffled-delay routes. Gulf/Galveston includes no-pair-forcing, and Southeast Australia includes no-delay, no-direction, no-pair-forcing, and no-common-mode routes. These comparisons use the same local forecast and temporal state while changing one declared operation at a time, showing how the regional effect changes when delay, directional asymmetry, pair forcing, or common-mode centring is removed.

Table 4 reports the Southeast Australia dissection. Supplementary Fig. S7 summarizes the ordered pair-forcing variation across physical edges, and Supplementary Fig. S9 provides the corresponding topology summary. Detailed numerical results, cluster counts, component-check results, and model settings are provided in Supplementary Tables S5–S8.

4 Discussion

4.1 Regional value of spatial information

The seven systems separate the general mean-trajectory value of spatial exchange from the regional value of a particular route. Lagged pair coupling improves mean trajectory accuracy in every system, while q95 and q99 distinguish where each route contributes most. Thames–Southern North Sea has the largest delayed-pair q95 gain, Gulf/Galveston shows a forcing-conditioned profile across mean and upper-tail outcomes, and Southeast Australia provides the strongest interval-supported Directed-delay result. The remaining systems show smaller or metric-dependent tail increments. Graph-coupling value consequently varies across regional gauge networks and forecast outcomes.

The station-forcing control sharpens this comparison. Station-resolved wind and pressure improve broad trajectory reconstruction in most systems, while the graph-based configurations produce larger high-water gains in Thames–Southern North Sea and Gulf/Galveston. Across the five spatial models, the forecast effects distinguish stable geographic weighting, delayed sender access, receiver conditioning, simultaneous pair contrast, and lagged pair contrast. Comparing these information pathways across seven systems shows how their value changes with regional setting.

4.2 Coastal interpretation of the regional profiles

The seven route profiles align with the diversity of coastal response represented by the station networks. Thames–Southern North Sea combines an estuarine entrance, adjacent shelf stations, pronounced tidal modulation, and exposure to regional atmospheric forcing. Tide–surge phase interaction, estuarine storage, shelf-scale forcing, and station ordering can all affect a 72 h trajectory [22, 24–28]. Lagged pair coupling represents delayed sender states together with ordered atmospheric contrasts. Its q95 advantage and the event trajectory in Figure 9 place the largest corrections around the rising limb and crest.

New York Bight, Charleston, and Southeast Australia place Directed-delay propagation first in the q95 point comparison. Their networks span connected embayments, open-coast segments, or extended alongshore arrangements in which related water-level changes can occur at different times. Southeast Australia provides the clearest evidence for this profile, with the q95 interval below zero and the same route leading RMSE. Charleston and New York Bight show the corresponding point rankings, with wider intervals in Charleston. The performance pattern is consistent with useful temporal ordering in the station records, which can reflect regional forcing, coastal storage, and spatial differences in local response.

Gulf/Galveston shows a different regional structure. Regional-forcing coupling leads RMSE, q95, and q99 and advances the first threshold crossing. The Gulf station network experiences coherent wind and pressure systems together with spatial contrasts across bays, barrier coasts, and open-shelf gauges [14, 15, 23]. Receiver forcing augments the station state, and delayed exchange places that state in the surrounding network context. The station-forcing control shows the value of local atmospheric input, while the larger graph-route tail effect is associated with the combined atmospheric and cross-station terms in that route.

Tokyo–Tokai and Los Angeles Bight retain mean improvements with smaller tail increments. Tokyo–Tokai has eleven stations and 110 directed edges, and its q95, q99, and RMSE rankings favour different routes. Los Angeles Bight has six stations and shows a small q95 change under its leading route. Across the seven systems, tail effects vary alongside gauge spacing, coastal orientation, forcing contrast, local forecast quality, network size, and the number of independent high-water clusters. Figure 1 and Supplementary Fig. S9 place these profiles alongside the station coordinates, edge counts, and water-path distances.

4.3 Forecasting use and scope

The route profiles link forecast objectives to observed performance. RMSE and horizon behaviour describe routine trajectory reconstruction, while high-water forecasting also requires q95 and q99 magnitude, threshold detection, peak timing, and warning lead. The seven-system results show that these mea-

sures can favour different route configurations, and warning lead indicates whether a lower event error is accompanied by an earlier actionable signal.

The objective-aligned profiles provide engineering guidance for subsequent regional evaluation. A trajectory product, an extreme-level product, and an early-warning product may emphasize different evidence. Candidate spatial models can then be evaluated prospectively with the input definition and metric set used here.

The study uses historical tide-gauge observations and issue-time ERA5 reanalysis within fixed station networks. The deterministic tide sequence is supplied over the forecast horizon. Operational deployment will replace the issue-time reanalysis field with the atmospheric forecast available at the issue time and incorporate data latency, quality control, and missing-station monitoring. The seven-system hindcast provides the regional route profiles under a common forecasting task for future operational evaluation.

5 Conclusions

Across seven coastal water-level networks, spatial coupling had a consistent value for mean trajectory reconstruction and a regional value for high-water prediction. Lagged pair coupling reduced total-water RMSE in every system, while the leading high-water route changed with the network. The Thames–Southern North Sea showed its largest q95 correction under Lagged pair coupling, Gulf/Galveston showed a coherent q95 and q99 improvement under Regional-forcing coupling, and Southeast Australia showed the clearest interval-supported q95 gain under Directed-delay propagation; its q99 point effect was also negative, with a wider interval that included zero. The remaining networks contributed smaller or metric-dependent tail profiles. Route rankings also changed across event detection, peak behaviour, and hindcast-derived warning timing. Coastal graph value therefore depends on the cross-station information carried by a regional network and on the forecast objective used to evaluate it.

The innovation of CoastGATE lies in making that relation measurable across systems. Five spatial information pathways were attached to one local temporal forecast and evaluated across seven independent gauge networks. Each pathway was measured by its change relative to the same local forecast. Stable connectivity, time-ordered station access, receiver atmospheric conditioning, and pairwise forcing could then be compared using the same error and event criteria. This design makes regional variation directly comparable and reveals how the pathway favoured by ordinary trajectory reconstruction can differ from the pathway favoured by high-water magnitude or warning timing.

The engineering implication is direct. A trajectory product can prioritise RMSE and horizon behaviour; a high-water product can prioritise q95 and q99 magnitude together with event detection; a warning system must consider first-crossing lead and peak timing. The regional profiles provide a basis for prospective spatial-route evaluation in fixed gauge networks. The present hindcast uses historical observations and issue-time ERA5 reanalysis; operational use will require atmospheric forecasts available at issue time together with explicit handling of data latency and gauge availability. Coastal graph forecasting becomes more informative when spatial structure is evaluated as a set of regional information pathways, with each pathway judged against the engineering outcome it is intended to support.

Funding

This work was partially supported by the Zhejiang Provincial Department of Education under grant Y202456288, the Wenzhou Science and Technology Program under grant GC20250005, and the 2023 Wenzhou–Kean University Research Start-up Project under grant ISRG2023023.

Data and code availability

Water-level observations, atmospheric forcing, bathymetry, and coastline vectors were obtained from the provider archives cited in the Methods and remain subject to their respective terms of use. Metric def-

initions and numerical results are reported in the tables and figures. The processed data products and analysis code are available from the corresponding authors upon reasonable request, subject to the terms of the original data providers.

Competing interests

The authors declare no competing interests.

Author contributions

Jia Rong and Guanchao Tong contributed equally. Both authors contributed to study design, methodology, analysis, visualization, and manuscript preparation. Both authors reviewed and approved the manuscript.

References

[1] S. Hochreiter, J. Schmidhuber, Long short-term memory, Neural Computation 9 (8) (1997) 1735–1780. https://doi.org/10.1162/neco.1997.9.8.1735.

[2] K. Cho, B. van Merrienboer, C. Gulcehre, D. Bahdanau, F. Bougares, H. Schwenk, Y. Bengio, Learning phrase representations using RNN encoder–decoder for statistical machine translation, in: Proceedings of EMNLP, 2014, pp. 1724–1734. https://doi.org/10.3115/v1/D14-1179.

[3] Y. Li, R. Yu, C. Shahabi, Y. Liu, Diffusion convolutional recurrent neural network: Data-driven traffic forecasting, International Conference on Learning Representations, 2018. https://openreview.net/forum?id=SJiHXGWAZ.

[4] B. Yu, H. Yin, Z. Zhu, Spatio-temporal graph convolutional networks: A deep learning framework for traffic forecasting, in: IJCAI, 2018, pp. 3634–3640. https://doi.org/10.24963/ijcai.2018/505.

[5] P. Veličković, G. Cucurull, A. Casanova, A. Romero, P. Lio, Y. Bengio, Graph attention networks, International Conference on Learning Representations, 2018. https://openreview.net/forum?id=rJXMpikCZ.

[6] Z. Wu, S. Pan, G. Long, J. Jiang, C. Zhang, Graph WaveNet for deep spatial–temporal graph modeling, in: IJCAI, 2019, pp. 1907–1913. https://doi.org/10.24963/ijcai.2019/264.

[7] T. N. Kipf, M. Welling, Semi-supervised classification with graph convolutional networks, International Conference on Learning Representations, 2017. https://arxiv.org/abs/1609.02907.

[8] Z. Wu, S. Pan, G. Long, J. Jiang, X. Chang, C. Zhang, Connecting the dots: Multivariate time series forecasting with graph neural networks, in: KDD, 2020, pp. 753–763. https://doi.org/10.1145/3394486.3403118.

[9] Z. Wu, S. Pan, F. Chen, G. Long, C. Zhang, P. S. Yu, A comprehensive survey on graph neural networks, IEEE Transactions on Neural Networks and Learning Systems 32 (2021) 4–24. https://doi.org/10.1109/TNNLS.2020.2978386.

[10] M. Jin et al., A survey on graph neural networks for time series: Forecasting, classification, imputation, and anomaly detection, IEEE Transactions on Pattern Analysis and Machine Intelligence 46 (2024) 10466–10485. https://doi.org/10.1109/TPAMI.2024.3443141.

[11] A. Cini, I. Marisca, D. Zambon, C. Alippi, Graph deep learning for time series forecasting, ACM Computing Surveys 57 (2025) 1–34. https://doi.org/10.1145/3742784.

[12] X. Shi, P. Chen, Z. Ye, X. Zhang, W. Wang, Tide level prediction during typhoons based on variable topology in graph convolution recurrent neural networks, Ocean Engineering 312 (2024) 119228. https://doi.org/10.1016/j.oceaneng.2024.119228.

[13] W. Jiang, J. Zhang, Y. Li, D. Zhang, G. Hu, H. Gao, Z. Duan, Advancing storm surge forecasting from scarce observation data: A causal-inference based spatio-temporal graph neural network approach, Coastal Engineering 190 (2024) 104512. https://doi.org/10.1016/j.coastaleng.2024.104512.

[14] M. Tadesse, T. Wahl, A. Cid, Data-driven modeling of global storm surges, Frontiers in Marine Science 7 (2020) 260. https://doi.org/10.3389/fmars.2020.00260.

[15] M. G. Tadesse, T. Wahl, A database of global storm surge reconstructions, Scientific Data 8 (2021) 125. https://doi.org/10.1038/s41597-021-00906-x.

[16] Z. Zhang et al., Short-term prediction of storm surges in estuarine and coastal waters via a multipoint deep learning neural network with limited training samples, International Journal of Digital Earth 18 (2025). https://doi.org/10.1080/17538947.2025.2536074.

[17] J. Yang et al., A ConvLSTM nearshore water level prediction model with integrated attention mechanism, Frontiers in Marine Science 11 (2024) 1470320. https://doi.org/10.3389/fmars.2024.1470320.

[18] M. Nagaraj, A. Rodriguez, T. Wahl, Regional modeling of storm surges using localized features and transfer learning, Journal of Geophysical Research: Machine Learning and Computation 2 (2025). https://doi.org/10.1029/2025JH000650.

[19] T. H. J. Hermans et al., Computing extreme storm surges in Europe using neural networks, Natural Hazards and Earth System Sciences 25 (2025) 4593–4612. https://doi.org/10.5194/nhess-25-4593-2025.

[20] Z. Huang et al., Prediction of the typhoon-induced tidal level at multiple stations using dynamic graph neural network, Ocean Engineering 349 (2026) 124180. https://doi.org/10.1016/j.oceaneng.2026.124180.

[21] J. Rong, G. Tong, Complementary Lag-Aligned Gauge Information and Pairwise Atmospheric Forcing for Coastal Water-Level Forecasting, Research Square, 2026. https://doi.org/10.21203/rs.3.rs-10634756/v1.

[22] K. J. Horsburgh, C. Wilson, Tide-surge interaction and its role in the distribution of surge residuals in the North Sea, Journal of Geophysical Research: Oceans 112 (2007) C08003. https://doi.org/10.1029/2006JC004033.

[23] S. Muis, M. Verlaan, H. C. Winsemius, J. C. J. H. Aerts, P. J. Ward, A global reanalysis of storm surges and extreme sea levels, Nature Communications 7 (2016) 11969. https://doi.org/10.1038/ncomms11969.

[24] S. Barbot, L. Pineau-Guillou, J.-M. Delouis, Extreme storm surge events and associated dynamics in the North Atlantic, Journal of Geophysical Research: Oceans 129 (2024) e2023JC020772. https://doi.org/10.1029/2023JC020772.

[25] D. Pugh, P. Woodworth, Sea-Level Science: Understanding Tides, Surges, Tsunamis and Mean Sea-Level Changes, Cambridge University Press, 2014. https://doi.org/10.1017/CBO9781139235778.

[26] F. M. Calafat et al., Trends in Europe storm surge extremes match the rate of sea-level rise, Nature 603 (2022) 841–845. https://doi.org/10.1038/s41586-022-04426-5.

[27] K. A. Serafin, P. Ruggiero, Simulating extreme total water levels using a time-varying extreme value approach, Journal of Geophysical Research: Oceans 119 (2014) 6305–6329. https://doi.org/10.1002/2014JC010092.

[28] T. Wahl et al., Understanding extreme sea levels for broad-scale coastal impact and adaptation analysis, Nature Communications 8 (2017) 16075. https://doi.org/10.1038/ncomms16075.

[29] M. I. Vousdoukas et al., Global probabilistic projections of extreme sea levels show intensification of coastal flood hazard, Nature Communications 9 (2018) 2360. https://doi.org/10.1038/s41467-018-04692-w.

[30] S. Vitousek et al., Doubling of coastal flooding frequency within decades due to sea-level rise, Scientific Reports 7 (2017) 1399. https://doi.org/10.1038/s41598-017-01362-7.

[31] C. Tebaldi et al., Extreme sea levels at different global warming levels, Nature Climate Change 11 (2021) 746–751. https://doi.org/10.1038/s41558-021-01127-1.

[32] R. Almar et al., A global analysis of extreme coastal water levels with implications for potential coastal overtopping, Nature Communications 12 (2021) 3775. https://doi.org/10.1038/s41467-021-24008-9.

[33] C. W. J. Granger, Investigating causal relations by econometric models and cross-spectral methods, Econometrica 37 (1969) 424–438. https://doi.org/10.2307/1912791.

[34] J. Runge et al., Inferring causation from time series in earth system sciences, Nature Communications 10 (2019) 2553. https://doi.org/10.1038/s41467-019-10105-3.

[35] J. Runge, P. Nowack, M. Kretschmer, S. Flaxman, D. Sejdinovic, Detecting and quantifying causal associations in large nonlinear time series datasets, Science Advances 5 (2019) eaau4996. https://doi.org/10.1126/sciadv.aau4996.

[36] NOAA Center for Operational Oceanographic Products and Services, CO-OPS metadata API and tide-gauge station services, 2026. https://api.tidesandcurrents.noaa.gov/mdapi/prod/.

[37] British Oceanographic Data Centre, UK National Tide Gauge Network and National Tide and Sea Level Facility data services, 2026. https://www.bodc.ac.uk/data/online_delivery/ntslf/.

[38] Global Extreme Sea Level Analysis, GESLA data release and download archive, version 4.0, 2026. https://www.gesla.org/.

[39] Australian Bureau of Meteorology, Australian Baseline Sea Level Monitoring Project: hourly sea-level and meteorological data, 2026. https://www.bom.gov.au/oceanography/projects/abslmp/data/.

[40] Japan Oceanographic Data Center, Tide-gauge and sea-level data services, 2026. https://www.jodc.go.jp/jodcweb/JDOSS/index.html.

[41] H. Hersbach et al., The ERA5 global reanalysis, Quarterly Journal of the Royal Meteorological Society 146 (2020) 1999–2049. https://doi.org/10.1002/qj.3803.

[42] Copernicus Climate Change Service, ERA5 hourly data on single levels from 1940 to present, 2024. https://doi.org/10.24381/cds.adbb2d47.

[43] GEBCO Compilation Group, The GEBCO 2026 Grid, British Oceanographic Data Centre, NERC EDS, 2026. https://doi.org/10.5285/4f68d5c7-45eb-f999-e063-7086abc036fa.

[44] Natural Earth, Natural Earth 1:10m physical vectors: land and coastline, version 5.1.1, 2023. https://www.naturalearthdata.com/downloads/10m-physical-vectors/.

[45] R. J. Hyndman, A. B. Koehler, Another look at measures of forecast accuracy, International Journal of Forecasting 22 (2006) 679–688. https://doi.org/10.1016/j.ijforecast.2006.03.001.

[46] T. Gneiting, Making and evaluating point forecasts, Journal of the American Statistical Association 106 (2011) 746–762. https://doi.org/10.1198/jasa.2011.r10138.

[47] S. Coles, An Introduction to Statistical Modeling of Extreme Values, Springer, 2001. https://doi.org/10.1007/978-1-4471-3675-0.

[48] H. R. Künsch, The jackknife and the bootstrap for general stationary observations, The Annals of Statistics 17 (1989) 1217–1241. https://doi.org/10.1214/aos/1176347265.

[49] W. Härdle, J. Horowitz, J.-P. Kreiss, Bootstrap methods for time series, International Statistical Review 71 (2003) 435–459. https://doi.org/10.1111/j.1751-5823.2003.tb00485.x.

Additional material

Supplementary Tables S1–S11 report station order, route-level metrics, initialisation results, forcing variation, component checks, event summaries, model settings, the shared local forecast, the persistence reference, and residual-tail effects. Supplementary Figs. S1–S10 show the corresponding horizon, tail, detection, timing, event, forcing, and topology summaries. Route order and colours follow the five route configurations defined in the main article and shown in the corresponding figures.

Table S1: Station lists

System	Station order used in the regional dataset
Thames–Southern North Sea	Cromer; Newhaven; Immingham; Sheerness; Portsmouth; Dover
New York Bight	The Battery; Kings Point; Sandy Hook; Montauk; Atlantic City; Cape May; Ship John Shoal; Lewes
Charleston	Charleston; Fort Pulaski; Springmaid Pier; Wilmington; Wrightsville Beach; Mayport; Trident Pier
Gulf/Galveston	Galveston Pier 21; Eagle Point; Packery Channel; Port Isabel; Grand Isle; Galveston Bay Entrance; Matagorda City; Morgans Point
Southeast Australia	Fort Denison; Jervis Bay; Port Macquarie; Ulladulla; Bermagui; Spring Bay; Portland
Tokyo–Tokai	Akabane; Maisaka; Omaezaki; Shimizu Minato; Uchiura; Odawara; Tokyo; Chiba; Yokohama Shinko; Yokosuka; Mera
Los Angeles Bight	San Diego; La Jolla; Los Angeles; Santa Monica; Santa Barbara; Port San Luis

Table S2: Aggregate metrics. The q95 and q99 MAE columns are pooled over all valid threshold-exceeding forecast cells. They are descriptive aggregate metrics; the equal-event-weighted effects use the separate event-level estimator defined in the main article.

System	Route	RMSE	MAE	R^2	q95 MAE	q99 MAE
Thames–Southern North Sea	Frozen Local-GRU reference	0.225	0.170	0.980	0.155	0.165
Thames–Southern North Sea	Static connectivity	0.211	0.158	0.982	0.149	0.161
Thames–Southern North Sea	Directed-delay propagation	0.210	0.158	0.982	0.148	0.160
Thames–Southern North Sea	Regional-forcing coupling	0.213	0.160	0.982	0.148	0.150
Thames–Southern North Sea	Dynamic pair coupling	0.208	0.156	0.983	0.144	0.150
Thames–Southern North Sea	Lagged pair coupling	0.209	0.156	0.983	0.142	0.147
New York Bight	Frozen Local-GRU reference	0.148	0.105	0.929	0.132	0.170
New York Bight	Static connectivity	0.147	0.104	0.930	0.131	0.170
New York Bight	Directed-delay propagation	0.146	0.103	0.931	0.130	0.168
New York Bight	Regional-forcing coupling	0.142	0.102	0.935	0.130	0.164
New York Bight	Dynamic pair coupling	0.145	0.103	0.932	0.130	0.167
New York Bight	Lagged pair coupling	0.145	0.104	0.931	0.131	0.168
Charleston	Frozen Local-GRU reference	0.114	0.084	0.957	0.081	0.096
Charleston	Static connectivity	0.113	0.083	0.958	0.081	0.097
Charleston	Directed-delay propagation	0.112	0.082	0.958	0.080	0.095
Charleston	Regional-forcing coupling	0.110	0.080	0.960	0.081	0.098

System	Route	RMSE	MAE	R^2	q95 MAE	q99 MAE
Charleston	Dynamic pair coupling	0.112	0.082	0.959	0.082	0.099
Charleston	Lagged pair coupling	0.112	0.082	0.959	0.081	0.099
Gulf/Galveston	Frozen Local-GRU reference	0.118	0.078	0.740	0.131	0.205
Gulf/Galveston	Static connectivity	0.116	0.076	0.750	0.125	0.195
Gulf/Galveston	Directed-delay propagation	0.116	0.076	0.750	0.125	0.195
Gulf/Galveston	Regional-forcing coupling	0.111	0.074	0.768	0.118	0.179
Gulf/Galveston	Dynamic pair coupling	0.115	0.076	0.753	0.124	0.194
Gulf/Galveston	Lagged pair coupling	0.115	0.076	0.754	0.122	0.191
Southeast Australia	Frozen Local-GRU reference	0.086	0.065	0.977	0.075	0.091
Southeast Australia	Static connectivity	0.085	0.064	0.978	0.074	0.090
Southeast Australia	Directed-delay propagation	0.083	0.063	0.979	0.072	0.086
Southeast Australia	Regional-forcing coupling	0.086	0.065	0.977	0.075	0.092
Southeast Australia	Dynamic pair coupling	0.083	0.063	0.979	0.073	0.089
Southeast Australia	Lagged pair coupling	0.083	0.063	0.978	0.073	0.090
Tokyo–Tokai	Frozen Local-GRU reference	0.079	0.061	0.990	0.054	0.060
Tokyo–Tokai	Static connectivity	0.079	0.061	0.990	0.054	0.059
Tokyo–Tokai	Directed-delay propagation	0.077	0.060	0.990	0.053	0.058
Tokyo–Tokai	Regional-forcing coupling	0.077	0.059	0.991	0.054	0.059
Tokyo–Tokai	Dynamic pair coupling	0.075	0.058	0.991	0.053	0.057
Tokyo–Tokai	Lagged pair coupling	0.075	0.058	0.991	0.053	0.057
Los Angeles Bight	Frozen Local-GRU reference	0.044	0.033	0.993	0.032	0.032
Los Angeles Bight	Static connectivity	0.043	0.032	0.993	0.032	0.032
Los Angeles Bight	Directed-delay propagation	0.043	0.032	0.993	0.031	0.031
Los Angeles Bight	Regional-forcing coupling	0.042	0.032	0.993	0.031	0.032
Los Angeles Bight	Dynamic pair coupling	0.043	0.032	0.993	0.031	0.031
Los Angeles Bight	Lagged pair coupling	0.043	0.032	0.993	0.031	0.031

Table S3: Tail and event-detection metrics

System	Route	$\Delta q95$	$\Delta q99$	q95 F1	q99 F1	Warning Δ (h)
Thames–Southern North Sea	Static connectivity	-0.0091	-0.0014	0.787	0.616	+0.72
Thames–Southern North Sea	Directed-delay propagation	-0.0112	-0.0020	0.786	0.615	+1.49
Thames–Southern North Sea	Regional-forcing coupling	-0.0177	-0.0067	0.786	0.623	-0.31
Thames–Southern North Sea	Dynamic pair coupling	-0.0185	-0.0089	0.792	0.634	-1.59
Thames–Southern North Sea	Lagged pair coupling	-0.0213	-0.0107	0.795	0.629	-2.90

System	Route	$\Delta q95$	$\Delta q99$	q95 F1	q99 F1	Warning Δ (h)
New York Bight	Static connectivity	+0.0006	-0.0004	0.707	0.554	+3.73
New York Bight	Directed-delay propagation	-0.0020	-0.0035	0.710	0.561	+3.21
New York Bight	Regional-forcing coupling	+0.0007	-0.0035	0.711	0.571	+3.12
New York Bight	Dynamic pair coupling	-0.0008	-0.0031	0.709	0.565	+4.06
New York Bight	Lagged pair coupling	-0.0005	-0.0025	0.708	0.561	+0.91
Charleston	Static connectivity	+0.0013	+0.0001	0.775	0.683	-1.88
Charleston	Directed-delay propagation	-0.0007	-0.0022	0.778	0.690	-1.33
Charleston	Regional-forcing coupling	+0.0025	+0.0027	0.783	0.689	-1.67
Charleston	Dynamic pair coupling	+0.0003	+0.0019	0.778	0.685	-1.15
Charleston	Lagged pair coupling	-0.0007	+0.0005	0.778	0.683	-0.61
Gulf/Galveston	Static connectivity	-0.0035	-0.0127	0.668	0.499	+0.60
Gulf/Galveston	Directed-delay propagation	-0.0018	-0.0103	0.670	0.497	+1.00
Gulf/Galveston	Regional-forcing coupling	-0.0107	-0.0349	0.684	0.528	+4.47
Gulf/Galveston	Dynamic pair coupling	-0.0033	-0.0137	0.675	0.498	+0.60
Gulf/Galveston	Lagged pair coupling	-0.0049	-0.0172	0.677	0.505	+0.62
Southeast Australia	Static connectivity	-0.0006	-0.0011	0.778	0.614	-7.56
Southeast Australia	Directed-delay propagation	-0.0038	-0.0047	0.784	0.634	-6.92
Southeast Australia	Regional-forcing coupling	-0.0009	+0.0003	0.776	0.606	-1.00
Southeast Australia	Dynamic pair coupling	-0.0026	-0.0020	0.784	0.621	-5.81
Southeast Australia	Lagged pair coupling	-0.0019	-0.0013	0.783	0.622	-8.64
Tokyo–Tokai	Static connectivity	+0.0003	-0.0017	0.761	0.594	-2.70
Tokyo–Tokai	Directed-delay propagation	-0.0003	-0.0028	0.765	0.599	-0.48
Tokyo–Tokai	Regional-forcing coupling	+0.0010	-0.0031	0.766	0.594	+2.74
Tokyo–Tokai	Dynamic pair coupling	-0.0018	-0.0057	0.769	0.602	+2.74
Tokyo–Tokai	Lagged pair coupling	-0.0018	-0.0057	0.769	0.602	+2.74
Los Angeles Bight	Static connectivity	-0.0001	+0.0002	0.919	0.861	+0.48
Los Angeles Bight	Directed-delay propagation	-0.0004	+0.0004	0.920	0.861	-0.54
Los Angeles Bight	Regional-forcing coupling	-0.0006	-0.0004	0.923	0.861	-1.81
Los Angeles Bight	Dynamic pair coupling	-0.0003	-0.0012	0.920	0.861	-0.70
Los Angeles Bight	Lagged pair coupling	-0.0003	-0.0014	0.920	0.861	-1.02

Table S4: Lagged pair coupling robustness across three independent initialisations

System	Run	RMSE change	q95 effect	q99 effect	Warning change (h)
Thames–Southern North Sea	1	-6.92%	-0.0209	-0.0113	-4.15
Thames–Southern North Sea	2	-6.91%	-0.0193	-0.0145	-5.00
Thames–Southern North Sea	3	-7.12%	-0.0236	-0.0062	+0.46
New York Bight	1	-0.78%	+0.0008	-0.0007	-0.55
New York Bight	2	-2.10%	-0.0058	-0.0088	+4.73
New York Bight	3	-1.35%	+0.0036	+0.0019	-1.45
Charleston	1	-1.68%	-0.0005	+0.0008	-0.18
Charleston	2	-2.10%	-0.0016	-0.0004	-3.27
Charleston	3	-1.89%	+0.0001	+0.0010	+1.64
Gulf/Galveston	1	-2.65%	-0.0049	-0.0211	+0.33
Gulf/Galveston	2	-2.50%	-0.0059	-0.0159	+0.80
Gulf/Galveston	3	-2.98%	-0.0038	-0.0145	+0.73
Southeast Australia	1	-3.45%	-0.0017	-0.0013	-8.50
Southeast Australia	2	-2.92%	-0.0024	-0.0016	-8.00
Southeast Australia	3	-3.76%	-0.0017	-0.0010	-9.42

System	Run	RMSE change	q95 effect	q99 effect	Warning change (h)
Tokyo–Tokai	1	-5.22%	-0.0021	-0.0061	+2.33
Tokyo–Tokai	2	-5.51%	-0.0013	-0.0030	+3.00
Tokyo–Tokai	3	-3.96%	-0.0019	-0.0080	+2.89
Los Angeles Bight	1	-2.11%	-0.0005	-0.0008	-3.52
Los Angeles Bight	2	-2.00%	-0.0002	-0.0020	-1.14
Los Angeles Bight	3	-1.62%	-0.0003	-0.0013	+1.62

Table S5: Ordered pair-forcing variation summary. The forcing columns report standard deviations (SD) of receiver–sender differences. Wind components are in m s^{-1}, pressure in Pa, and wind stress in Pa.

System	u_{10} SD (m s^{-1})	v_{10} SD (m s^{-1})	Pressure SD (Pa)	Stress SD (Pa)	Edges	q95 clusters
Thames–Southern North Sea	2.431	2.172	191.038	0.056	30	13
New York Bight	1.826	1.640	139.805	0.036	56	11
Charleston	2.563	2.410	180.595	0.051	42	11
Gulf/Galveston	2.104	2.374	83.008	0.048	56	15
Southeast Australia	2.571	3.261	157.818	0.063	30	12
Tokyo–Tokai	3.857	3.321	141.213	0.077	110	18
Los Angeles Bight	1.522	2.068	101.287	0.016	30	21

Table S6: Targeted route component checks

System	Component check	RMSE change	q95 effect	q99 effect	Warning change (h)
Thames–Southern North Sea	Frozen station-forcing control	-2.32%	-0.0118	-0.0069	-2.46
Thames–Southern North Sea	No common mode	-5.66%	-0.0180	+0.0002	-3.18
Thames–Southern North Sea	No delay	-5.63%	-0.0182	-0.0042	+3.95
Thames–Southern North Sea	No pair forcing	-6.56%	-0.0112	-0.0020	+1.49
New York Bight	Frozen station-forcing control	-0.30%	+0.0034	+0.0036	-3.39
New York Bight	No delay	-1.53%	-0.0014	-0.0033	+2.15
New York Bight	Shuffled delay	-0.98%	-0.0002	-0.0011	+2.00
Charleston	Frozen station-forcing control	-2.64%	+0.0037	+0.0040	-2.36
Charleston	No pair forcing	-1.63%	-0.0007	-0.0022	-1.33
Gulf/Galveston	Frozen station-forcing control	-3.02%	-0.0083	-0.0229	+3.07
Gulf/Galveston	No pair forcing	-1.89%	-0.0018	-0.0103	+1.00
Southeast Australia	Frozen station-forcing control	-0.19%	-0.0003	-0.0004	-1.00
Southeast Australia	No common mode	-1.82%	-0.0025	-0.0035	+1.36
Southeast Australia	No delay	-3.07%	-0.0020	-0.0021	-7.86
Southeast Australia	No direction	-2.82%	-0.0031	-0.0052	+0.78
Southeast Australia	No pair forcing	-3.63%	-0.0038	-0.0047	-6.92
Tokyo–Tokai	Frozen station-forcing control	-1.52%	-0.0018	-0.0031	+3.04
Los Angeles Bight	Frozen station-forcing control	-1.98%	-0.0001	-0.0002	-2.48

Table S8: Model settings

Item	Value
Random initialisations	Three independent random initialisations were evaluated for each principal route.
Local reference	The frozen Local-GRU forecast B and state h are shared by every matched spatial route and the station-forcing control.
Encoder	Shared Local-GRU width 96 with a learned station-availability embedding; 96 h history.
Local forecast head	Linear map from the 96-dimensional state and 72 h future tide to a 72 h residual trajectory.
Graph projections	Query, key, and value dimensions 96; graph update implemented with a 96-dimensional gated recurrent unit cell.
Pair encoder	Ordered 12-variable forcing feature mapped through $12 \rightarrow 96 \rightarrow 1$ with a Gaussian error linear unit activation.
Receiver-forcing encoder	Four issue-time forcing variables mapped through $4 \rightarrow 96 \rightarrow 96$ with a Gaussian error linear unit activation.
Correction head	Concatenated local state, graph state, and local trajectory mapped through $264 \rightarrow 96 \rightarrow 72$ with Gaussian error linear unit and hyperbolic-tangent output.
Regularization	Dropout 0.1; bounded correction scale set to the station- and lead-specific 0.90 quantile of local residual error.
Optimization	AdamW; learning rate 3×10^{-4}; weight decay 10^{-4}; maximum 300 epochs; early-stopping patience 30; tail-loss weight $\lambda_{95} = 0.1$; batch size 512.

Table S9: Shared local forecast across the seven systems. Each spatial route used the system-specific Local-GRU forecast, which remained fixed during route fitting and test scoring.

System	Local forecast	Route fitting	Test scoring
Thames–Southern North Sea	Shared Local-GRU	Yes	Yes
New York Bight	Shared Local-GRU	Yes	Yes
Charleston	Shared Local-GRU	Yes	Yes
Gulf/Galveston	Shared Local-GRU	Yes	Yes
Southeast Australia	Shared Local-GRU	Yes	Yes
Tokyo–Tokai	Shared Local-GRU	Yes	Yes
Los Angeles Bight	Shared Local-GRU	Yes	Yes

Table S7: Event-cluster counts

System	q95 clusters	q99 clusters	Test issue cycles	Lagged-pair q95 effect	Lagged-pair q99 effect
Thames–Southern North Sea	13	7	688	-0.0213	-0.0107
New York Bight	11	20	1385	-0.0005	-0.0025
Charleston	11	14	1385	-0.0007	+0.0005
Gulf/Galveston	15	13	1268	-0.0049	-0.0172
Southeast Australia	12	17	1278	-0.0019	-0.0013
Tokyo–Tokai	18	15	1211	-0.0018	-0.0057
Los Angeles Bight	21	15	1447	-0.0003	-0.0014

Table S10: Persistence reference for the 72 h task. The last available non-tidal residual at each issue time was carried through the forecast window and combined with the deterministic tide. Values are total-water test metrics.

System	RMSE	MAE	q95 MAE	q99 MAE
Thames–Southern North Sea	0.4552	0.3524	0.3178	0.3014
New York Bight	0.2384	0.1804	0.1828	0.2071
Charleston	0.2146	0.1657	0.1670	0.1821
Gulf/Galveston	0.1372	0.0900	0.1286	0.1784
Southeast Australia	0.1309	0.1022	0.1108	0.1187
Tokyo–Tokai	0.1350	0.1058	0.0985	0.1016
Los Angeles Bight	0.1471	0.1166	0.1246	0.1327

Table S11: Residual and total-water event effects from the same test predictions. Values are route minus the common Local-GRU reference in metres, averaged across the three initialisations. The main tail intervals use 2,000 equal-event resamples for each initialisation.

System	Route	Residual q95	Residual q99	Total q95	Total q99
Thames–Southern North Sea	Local-GRU reference	0.0000	0.0000	0.0000	0.0000
	Static connectivity	-0.0467	-0.0183	-0.0091	-0.0014
	Directed-delay propagation	-0.0500	-0.0262	-0.0112	-0.0020
	Regional-forcing coupling	-0.0324	-0.0427	-0.0177	-0.0067
	Dynamic pair coupling	-0.0488	-0.0347	-0.0185	-0.0089
	Lagged pair coupling	-0.0568	-0.0475	-0.0213	-0.0107
	Frozen station-forcing control	+0.0022	-0.0128	-0.0118	-0.0069
New York Bight	Local-GRU reference	0.0000	0.0000	0.0000	0.0000
	Static connectivity	-0.0016	-0.0033	+0.0006	-0.0004
	Directed-delay propagation	-0.0002	+0.0005	-0.0020	-0.0035
	Regional-forcing coupling	-0.0072	-0.0152	+0.0007	-0.0035
	Dynamic pair coupling	-0.0016	-0.0017	-0.0008	-0.0031
	Lagged pair coupling	-0.0013	-0.0013	-0.0005	-0.0025
	Frozen station-forcing control	+0.0050	+0.0039	+0.0034	+0.0036
Charleston	Local-GRU reference	0.0000	0.0000	0.0000	0.0000
	Static connectivity	+0.0010	+0.0014	+0.0013	+0.0001
	Directed-delay propagation	+0.0022	-0.0001	-0.0007	-0.0022
	Regional-forcing coupling	-0.0020	-0.0022	+0.0025	+0.0027
	Dynamic pair coupling	+0.0033	+0.0001	+0.0003	+0.0019
	Lagged pair coupling	+0.0004	-0.0034	-0.0007	+0.0005
	Frozen station-forcing control	+0.0005	-0.0047	+0.0037	+0.0040
Gulf/Galveston	Local-GRU reference	0.0000	0.0000	0.0000	0.0000
	Static connectivity	-0.0038	-0.0119	-0.0035	-0.0127
	Directed-delay propagation	-0.0002	-0.0057	-0.0018	-0.0103
	Regional-forcing coupling	-0.0218	-0.0347	-0.0107	-0.0349
	Dynamic pair coupling	-0.0029	-0.0110	-0.0033	-0.0137
	Lagged pair coupling	-0.0060	-0.0145	-0.0049	-0.0172
	Frozen station-forcing control	-0.0203	-0.0228	-0.0083	-0.0229
Southeast Australia	Local-GRU reference	0.0000	0.0000	0.0000	0.0000
	Static connectivity	+0.0003	-0.0011	-0.0006	-0.0011
	Directed-delay propagation	-0.0075	-0.0097	-0.0038	-0.0047
	Regional-forcing coupling	+0.0012	+0.0022	-0.0009	+0.0003
	Dynamic pair coupling	-0.0057	-0.0097	-0.0026	-0.0020
	Lagged pair coupling	-0.0042	-0.0056	-0.0019	-0.0013

Table S11: Residual and total-water event effects (continued).

System	Route	Residual q95	Residual q99	Total q95	Total q99
	Frozen station-forcing control	+0.0003	+0.0003	-0.0003	-0.0004
Tokyo–Tokai	Local-GRU reference	0.0000	0.0000	0.0000	0.0000
	Static connectivity	-0.0007	-0.0005	+0.0003	-0.0017
	Directed-delay propagation	-0.0045	-0.0040	-0.0003	-0.0028
	Regional-forcing coupling	-0.0024	-0.0030	+0.0010	-0.0031
	Dynamic pair coupling	-0.0120	-0.0114	-0.0018	-0.0057
	Lagged pair coupling	-0.0120	-0.0114	-0.0018	-0.0057
	Frozen station-forcing control	-0.0042	-0.0038	-0.0018	-0.0031
Los Angeles Bight	Local-GRU reference	0.0000	0.0000	0.0000	0.0000
	Static connectivity	+0.0015	+0.0012	-0.0001	+0.0002
	Directed-delay propagation	+0.0014	+0.0017	-0.0004	+0.0004
	Regional-forcing coupling	+0.0007	0.0000	-0.0006	-0.0004
	Dynamic pair coupling	+0.0020	+0.0032	-0.0003	-0.0012
	Lagged pair coupling	+0.0016	+0.0025	-0.0003	-0.0014
	Frozen station-forcing control	+0.0006	+0.0022	-0.0001	-0.0002

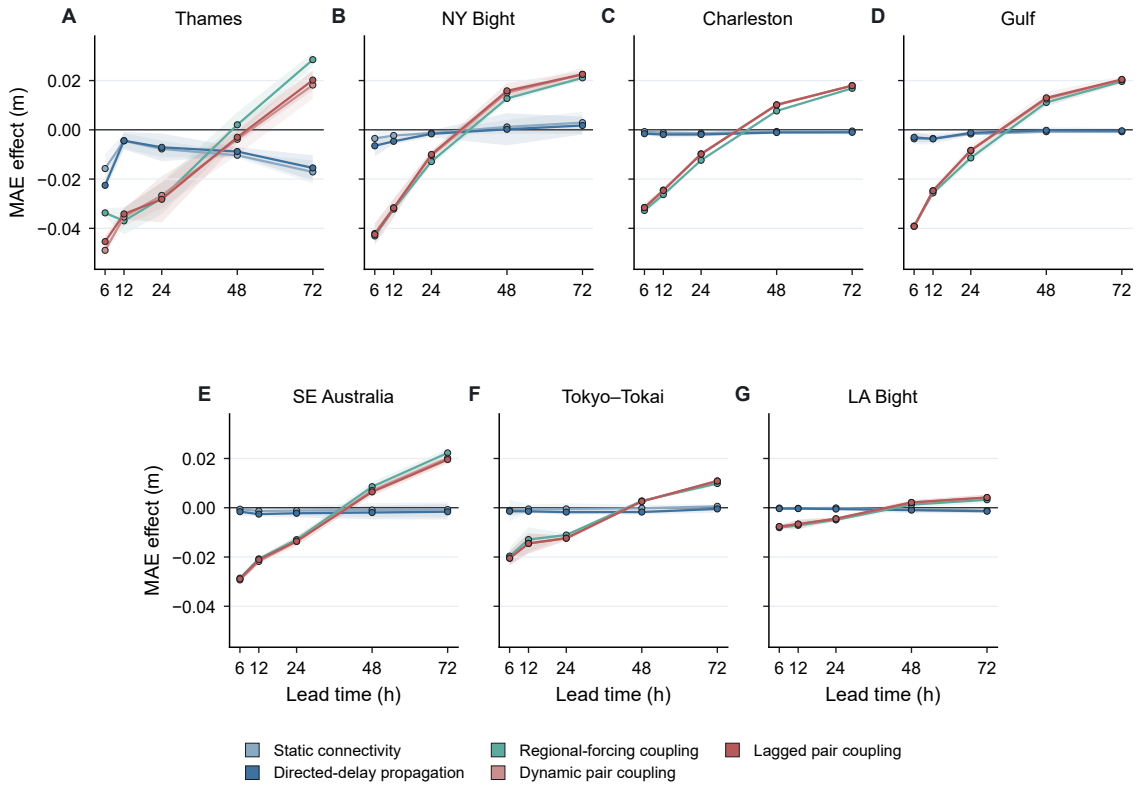


Figure S1: Horizon-wise MAE profiles. Panels A–G show the Thames–Southern North Sea, New York Bight, Charleston, Gulf/Galveston, Southeast Australia, Tokyo–Tokai, and Los Angeles Bight, respectively. Curves give the five spatial routes relative to the frozen Local-GRU reference at 6, 12, 24, 48, and 72 h.

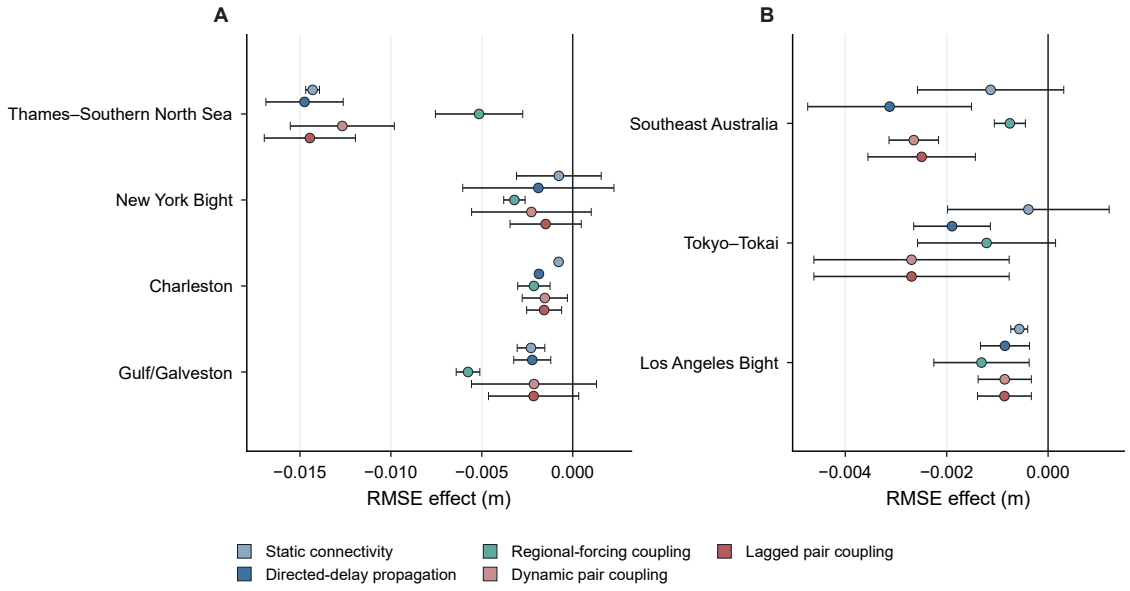


Figure S2: Aggregate RMSE effects for the five spatial routes. Panel A contains the Thames–Southern North Sea, New York Bight, Charleston, and Gulf/Galveston systems; Panel B contains Southeast Australia, Tokyo–Tokai, and Los Angeles Bight. Coloured points identify the five spatial models, and horizontal intervals show initialisation variation relative to the frozen Local-GRU reference.

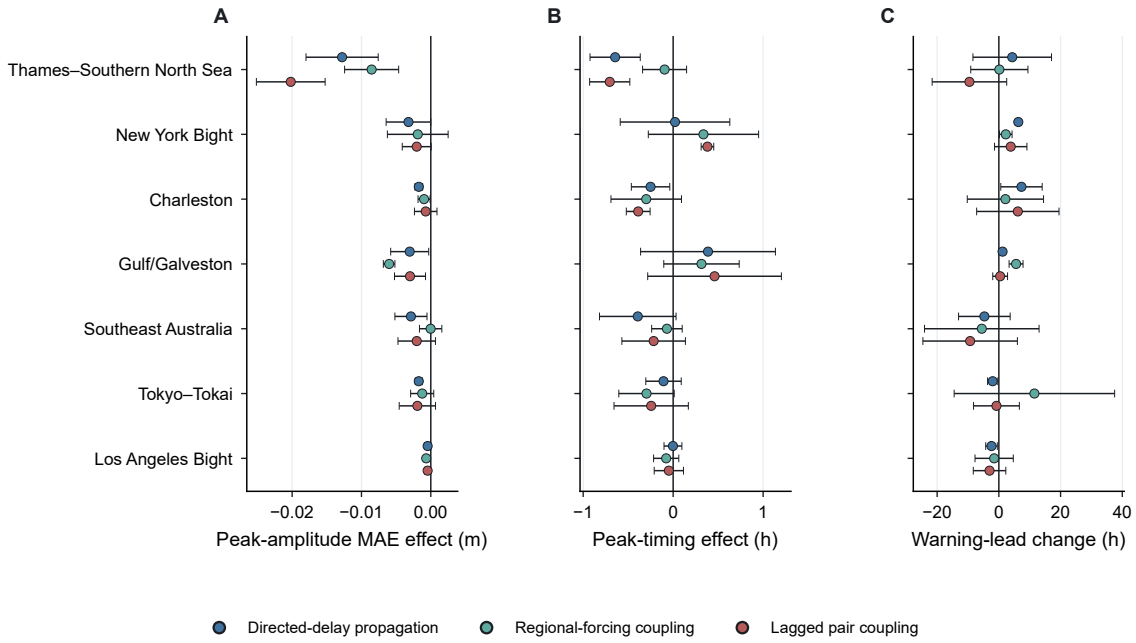


Figure S3: Peak-amplitude, peak-timing, and warning-lead profiles relative to the frozen Local-GRU reference. Panel A shows peak-amplitude error, Panel B peak-timing error, and Panel C warning-lead change; each panel contains all seven systems and the three displayed spatial routes.

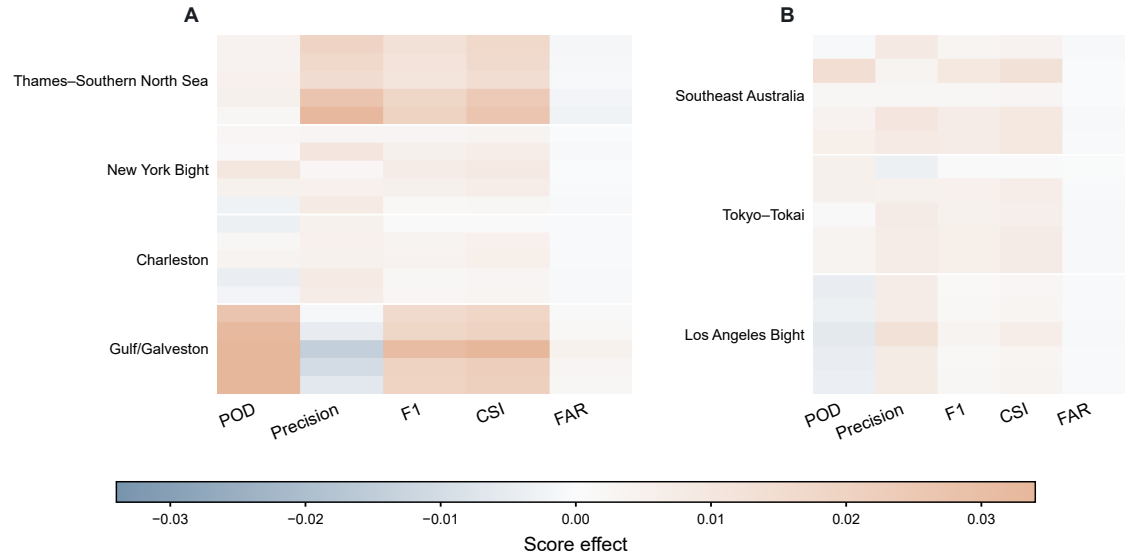


Figure S4: q95 detection metrics across systems and spatial routes. Panel A contains the Thames–Southern North Sea, New York Bight, Charleston, and Gulf/Galveston systems; Panel B contains Southeast Australia, Tokyo–Tokai, and Los Angeles Bight. Within each system block, the five consecutive rows correspond to Static connectivity, Directed-delay propagation, Regional-forcing coupling, Dynamic pair coupling, and Lagged pair coupling. Columns show POD, precision, F1, CSI, and FAR; cell colours show route effects relative to the frozen Local-GRU reference.

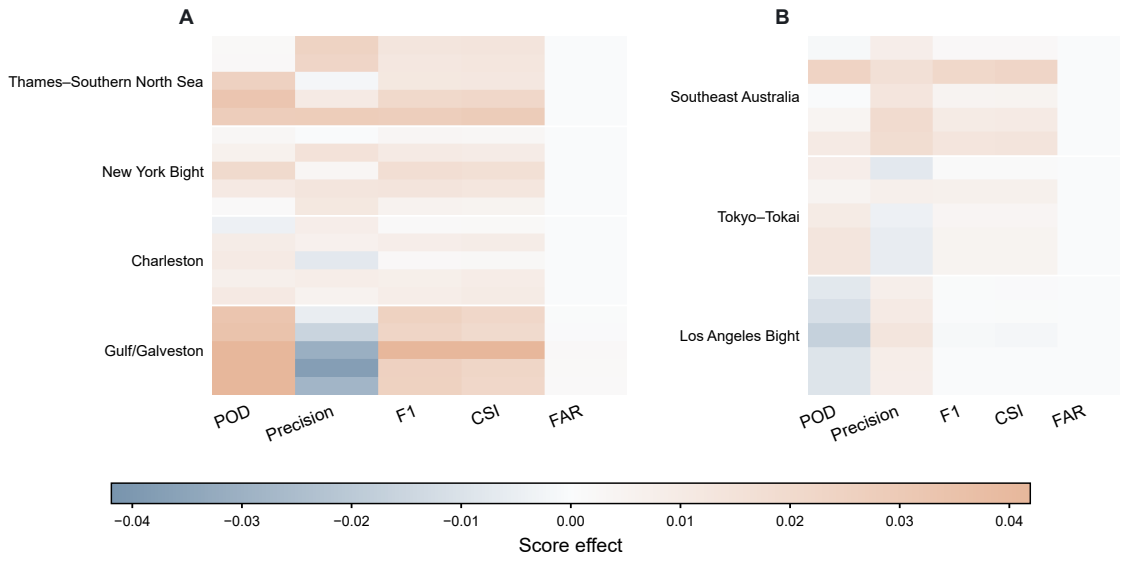


Figure S5: q99 detection metrics across systems and spatial routes. Panel A contains the Thames–Southern North Sea, New York Bight, Charleston, and Gulf/Galveston systems; Panel B contains Southeast Australia, Tokyo–Tokai, and Los Angeles Bight. Within each system block, the five consecutive rows correspond to Static connectivity, Directed-delay propagation, Regional-forcing coupling, Dynamic pair coupling, and Lagged pair coupling. Columns show POD, precision, F1, CSI, and FAR; cell colours show route effects relative to the frozen Local-GRU reference.

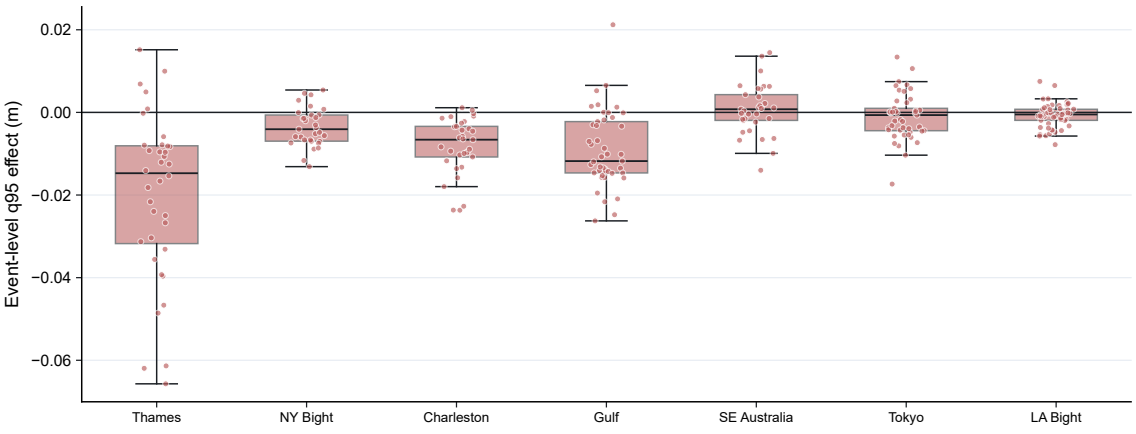


Figure S6: Event-level q95 effects of Lagged pair coupling across the seven systems. Values are route minus the frozen Local-GRU reference. Boxes show the interquartile range, central lines show medians, and points show individual event clusters.

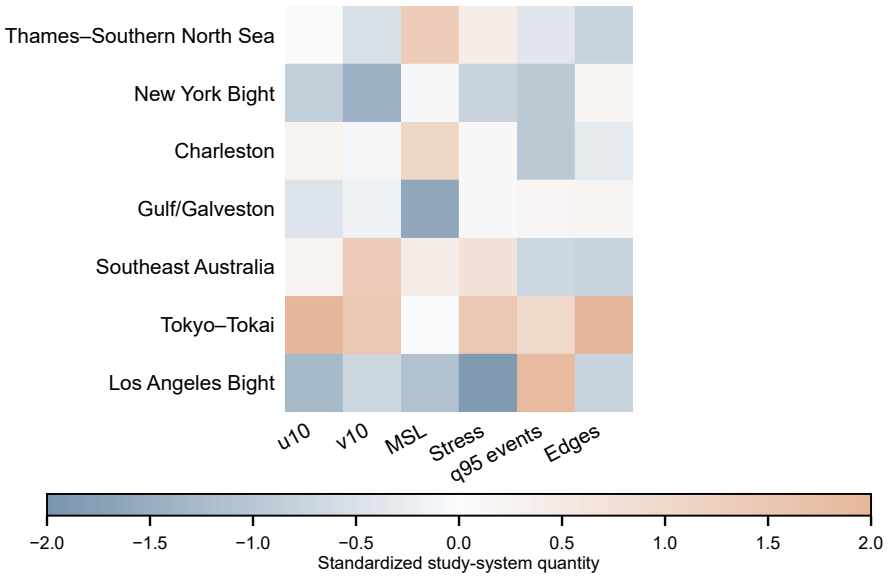


Figure S7: Ordered pair-forcing variation across the seven systems. The heatmap shows standardized system-level quantities for the four station-resolved meteorological contrasts, the q95 event-cluster count, and the directed physical-edge count.

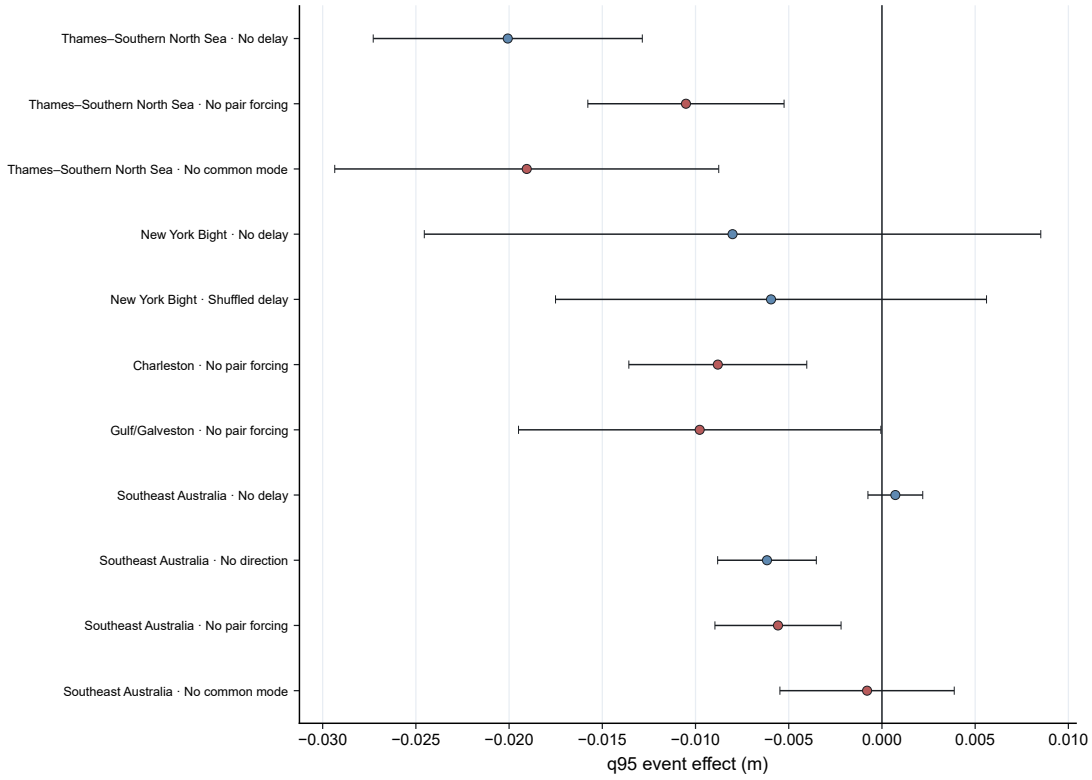


Figure S8: Targeted route component checks for delay, direction, pair forcing, and common-mode operations. The forest plot shows q95 event effects relative to the frozen Local-GRU reference for the targeted system–route pairs included in the study design; horizontal intervals show the corresponding uncertainty.

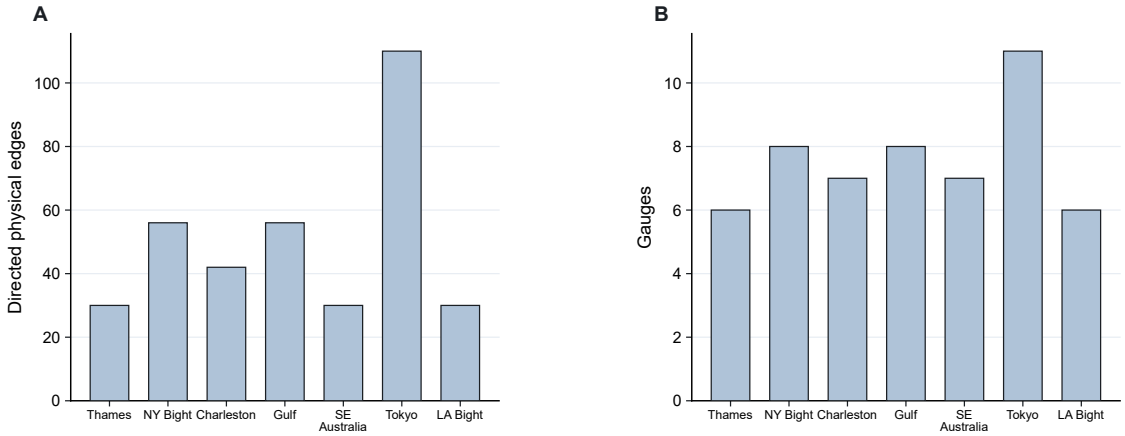


Figure S9: Physical topology summary for the seven study systems. Panel A reports the number of ordered physical edges for each system, and Panel B reports the number of gauges.

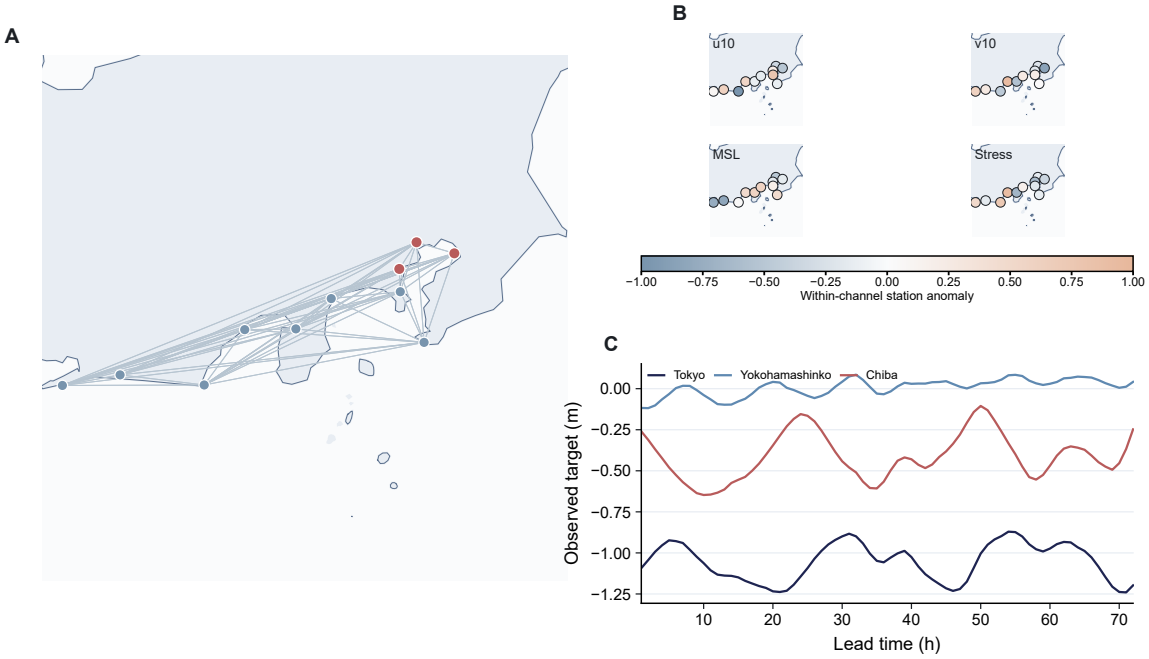


Figure S10: Example of the 72 h forecasting task for a Tokyo–Tokai q99 test event. Panel A shows the station geometry and physical mask; highlighted stations are used in Panel C. Panel B maps the four station-resolved ERA5 forcing inputs at the forecast issue time, including mean-sea-level pressure (MSL). Colours show within-channel station anomalies. Panel C shows observed 72 h target trajectories from the same event window.