

Article

Maneuver-Aware Residual Multi-Scale LSTM for Short-Term Vessel Trajectory Prediction in Port-Approach Waters

Xinyue Lin ¹, Xiaohan Zhang ², Wendong Bao ³, Yawen Duan ² and Jiansen Zhao ^{2,4,*}¹ College of Transport & Communications, Shanghai Maritime University, Shanghai 201306, China; 202310613077@stu.shmtu.edu.cn² Merchant Marine College, Shanghai Maritime University, Shanghai 201306, China; 202410111076@stu.shmtu.edu.cn (X.Z.); 202410111070@stu.shmtu.edu.cn (Y.D.)³ College of Information Engineering, Shanghai Maritime University, Shanghai 201306, China; 202310320072@stu.shmtu.edu.cn⁴ Hubei Key Laboratory of Inland Shipping Technology, School of Navigation, Wuhan University of Technology, Wuhan 430063, China

* Correspondence: jszhao@shmtu.edu.cn

Abstract

Accurate short-term vessel trajectory prediction is important for traffic monitoring and collision-risk screening in port-approach waters, where vessels frequently turn, accelerate, decelerate, and merge into traffic lanes. This study develops a maneuver-aware residual multi-scale long short-term memory (LSTM) framework for Automatic Identification System (AIS)-based 10–30 min trajectory prediction. The method predicts residual displacement relative to the last observed position, constructs maneuver-aware features from local displacement, course variation, speed variation, turning rate, and acceleration-like terms, and compares fixed and maneuver-guided multi-scale fusion strategies. Experiments are conducted on public AIS data from San Francisco Bay and adjacent approach waters using Maritime Mobile Service Identity (MMSI)-level train/validation/test splits and three random seeds. The largest observed gains come from residual prediction and maneuver-aware features. In the three-seed main evaluation, the fixed multi-scale LSTM (Fixed-MS-LSTM) provides the strongest 10 min accuracy, while the maneuver-guided multi-scale residual LSTM (MGMS-RLSTM) achieves lower average displacement error (ADE) at 20 and 30 min and learns distinct temporal-scale preferences across straight, turning, and speed-changing samples. Encounter-oriented closest point of approach (CPA) and time to closest point of approach (TCPA) evaluation further shows that the residual multi-scale models support more accurate CPA/TCPA-based high-risk screening under the evaluated benchmark. These findings indicate that maneuver-guided fusion can be characterized as a horizon-dependent scale-selection mechanism that complements the fixed multi-scale counterpart.

Keywords: AIS trajectory prediction; maneuver-guided fusion; multi-scale LSTM; residual learning; port-approach waters



Academic Editor: Haitong Xu

Received: 7 June 2026

Revised: 6 July 2026

Accepted: 17 July 2026

Published: 31 July 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

Maritime transportation supports global trade and port logistics, but dense traffic in coastal approach waters also increases the risk of close-quarters encounters. Compared with open-sea navigation, port-approach operations are affected by constrained waterways, fairway entrances, berth departures, pilotage areas, and crossing traffic streams. Vessels

therefore often change course or speed within a short time interval. For vessel traffic services (VTS), autonomous navigation, and collision-risk screening, 10–20 min forecasts are particularly useful because they provide near-term warning while remaining closely linked to the vessel's most recent motion state. A 30 min horizon is also evaluated as an extended stress case, where errors are expected to increase because route choice, traffic interaction, and waterway constraints become more influential.

Automatic Identification System (AIS) data are the main data source for this task. AIS records include vessel identity, timestamp, position, speed over ground (SOG), course over ground (COG), heading, and other navigation-related fields. Nevertheless, reliable AIS-based prediction in port-approach waters remains challenging. First, raw AIS trajectories contain duplicated messages, irregular sampling intervals, near-stationary or non-navigating records, physical jumps, and long temporal gaps. Second, directly predicting absolute latitude and longitude requires the model to learn regional spatial distribution and local motion evolution simultaneously. Third, the effective temporal scale varies with maneuvering state: stable straight sailing may benefit from a longer recent trend, whereas a turning or decelerating vessel may depend more strongly on recent course and speed variation.

Existing recurrent models, including long short-term memory (LSTM) and gated recurrent unit (GRU) networks, can capture temporal dependencies in AIS sequences [1–3]. Convolutional neural network (CNN)-LSTM variants further introduce convolutional layers to extract local motion patterns before recurrent encoding [4]. Multi-scale convolutional structures are also attractive because different kernel sizes can represent short-, medium-, and longer-range temporal patterns [5,6].

Recent AIS-based prediction studies have shown the value of enriching raw AIS variables and extracting trajectory patterns at multiple temporal scales. For example, LMCA [6] combines extended AIS features, parallel multi-scale convolution, LSTM encoding, and attention to represent both local trajectory changes and longer motion trends. This line of work suggests that multi-scale temporal representation is useful for ship trajectory prediction. In port-approach waters, a remaining question is how the contributions of multiple temporal scales should change when the recent AIS window indicates straight sailing, turning, acceleration, or deceleration.

This study therefore focuses on a narrower and application-oriented question: can maneuver statistics derived from AIS records guide temporal-scale fusion for short-term prediction in port-approach waters? It examines whether residual local-motion learning, maneuver-aware features, and maneuver-conditioned scale fusion can provide a lightweight and interpretable AIS-only predictor for 10–30 min forecasting.

The main contributions of this study are as follows. First, a residual displacement prediction formulation is developed for AIS-based short-term vessel trajectory prediction, reducing the burden of learning absolute regional coordinates. Second, a maneuver-aware feature representation is constructed from local displacement, course variation, speed variation, turning rate, and acceleration-like terms to describe port-approach motion changes. Third, a residual multi-scale LSTM framework is designed with both fixed and maneuver-guided fusion variants, enabling a direct comparison between stable fixed fusion and interpretable adaptive scale selection. Fourth, an encounter-oriented evaluation based on closest point of approach (CPA), time to closest point of approach (TCPA), and high-risk screening is introduced to connect pointwise trajectory accuracy with VTS-oriented collision-risk assessment.

2. Related Work

2.1. AIS-Based Vessel Trajectory Prediction

AIS data have been widely used in maritime surveillance, traffic analysis, anomaly detection, and trajectory prediction. The technical characteristics of AIS reporting are standardized by ITU-R Recommendation M.1371 [7], which makes AIS a practical source for large-scale trajectory modeling. Traditional prediction methods include constant-velocity extrapolation, support-vector regression, Markov models, and clustering-based approaches. These methods are usually transparent and computationally efficient, and they can perform well under stable motion conditions. However, their ability to model abrupt turns, speed changes, and route merging is limited, especially in port waters where vessel motion is strongly affected by local traffic and navigational constraints. Liu et al. [8], for example, combined AIS sensor data with an adaptive chaos differential evolution support-vector regression method. Such approaches remain valuable as interpretable baselines, but they are less suited to learning complex nonlinear motion patterns from long AIS sequences.

Deep learning methods have become increasingly common because AIS trajectories are naturally sequential. Recurrent networks, sequence-to-sequence models, graph neural networks, and Transformer variants have all been explored [2,9–12]. Feng et al. [13] proposed IS-STGCNN, which combines spatio-temporal graph convolution with a time-extrapolator CNN and ship-interaction modeling. Wang et al. [14] proposed TPTrans, a CNN–Transformer vessel trajectory predictor using AIS navigation features. Recent studies increasingly combine Transformer encoders, graph attention, route-network information, contextual maritime features, and probabilistic outputs to handle longer sequences, multi-vessel interactions, and uncertainty [15–19]. A recent review by Bi et al. [20] summarizes the transition from classical models to data-driven and deep learning methods in ship trajectory prediction. From the perspective of modeling assumptions, existing methods can be broadly grouped into classical motion models, deep sequence encoders, spatial or intention-enhanced predictors, interaction-aware models, and multi-scale temporal models. Table 1 summarizes how the present study is positioned relative to these recent model families. These directions address different aspects of the prediction problem, but many of them either rely on fixed temporal representations or require additional environmental, route, or intention information. The present study follows the AIS-only data-driven direction and focuses on a specific question: how to make short-term temporal-scale selection more consistent with recent maneuvering behavior.

2.2. LSTM-Based Sequence Modeling

LSTM networks were designed to model long-range dependencies in sequential data while mitigating vanishing gradients [1]. They have therefore been widely adopted for AIS trajectory prediction. Capobianco et al. [2] studied recurrent sequence-to-sequence vessel trajectory prediction and demonstrated the value of recurrent encoders for maritime tracks. Park et al. [3] used Bi-LSTM with spectral-clustered AIS data, and Volkova et al. [21] investigated neural-network-based AIS prediction.

Despite their usefulness, vanilla LSTM models can struggle when raw input features are dominated by absolute position. In short-term port prediction, future positions are closely related to recent local displacement, COG variation, and SOG variation. If the model is trained to predict absolute coordinates directly, it must learn regional location patterns and local maneuver dynamics at the same time. This motivates the use of residual displacement prediction and explicit maneuver-aware features.

Table 1. Methodological positioning relative to recent AIS-based trajectory prediction models.

Model Family	Representative Studies	Main Modeling Emphasis	Difference From this Study
Hybrid Transformer sequence models	TRFM-LS [12]; G-Trans [15]	Long-range temporal dependency modeling, often with improved positional encoding, recurrent blocks, or trajectory clustering	The proposed framework targets lightweight short-term local-motion prediction and uses residual multi-scale LSTM encoding instead of global self-attention as the main mechanism
Multi-vessel graph-Transformer models	ST-MGT [16]; GATransformer [17]	Spatial interaction modeling among vessels and, in some cases, waterway-network or intersection features	This study focuses on AIS-derived local-motion features and residual multi-scale temporal encoding, providing a lightweight short-term prediction setting complementary to context-rich or interaction-aware models
Probabilistic spatio-temporal graph models	STGTP [18]	Joint spatio-temporal prediction with uncertainty-aware probabilistic representation	The framework is deterministic and focuses on residual local-motion learning and interpretable temporal-scale analysis; probabilistic outputs are treated as future work
Context-aware Transformer models	FlashAttention-based context-aware Transformer [19]	Efficient long-sequence attention with environmental, structural, or maritime-spatial context features	The proposed framework emphasizes AIS-derived maneuver features and interpretable temporal-scale selection for short-term local-motion prediction
Multi-scale LSTM/CNN models	MP-LSTM [5]; LMCA [6]	Temporal feature extraction through multiple convolutional or attention-enhanced scales	Fixed-MS-LSTM is constructed as a strong fixed-fusion counterpart inspired by multi-scale LSTM designs; MGMS-RLSTM further makes scale fusion sample-dependent through window-level maneuver statistics

2.3. CNN-LSTM and Multi-Scale Temporal Modeling

CNN-LSTM models combine local convolutional feature extraction with recurrent temporal encoding. Wang and Xiao [4] proposed a CNN-LSTM-SE model for ship trajectory prediction and showed that convolution can help extract correlations among AIS variables before LSTM modeling. Multi-scale convolution extends this idea by using several kernel sizes to capture motion patterns over different temporal ranges. Gao et al. [5] proposed an MP-LSTM model for AIS trajectory prediction. More recently, Yin et al. [6] proposed LMCA, an LSTM model with extended AIS features, parallel multi-scale convolution, and an attention mechanism, further showing that temporal scale diversity and feature enhancement can improve AIS trajectory prediction.

LMCA demonstrates the effectiveness of extended AIS feature expansion, parallel multi-scale convolution, and LSTM-based temporal encoding for vessel trajectory prediction. Building on this line of work, this study constructs a residual multi-scale LSTM framework with two fusion variants. Fixed-MS-LSTM uses equal branch averaging as a stable fixed-fusion internal counterpart. MGMS-RLSTM retains the same residual prediction target, maneuver-aware features, multi-scale temporal branches, and LSTM encoder, but makes the fusion weights depend on recent maneuver intensity. The comparison examines when AIS-derived maneuver statistics provide useful and interpretable temporal-scale selection for near-term maneuvering scenarios.

2.4. Maneuver-Aware and Residual Learning

Maneuver-aware modeling recognizes that a vessel trajectory is not merely a position sequence. Course variation, speed variation, turning rate, and acceleration-like features help distinguish stable sailing from turning or speed-changing behavior. Residual learning, originally popularized in deep visual recognition [22], is also useful for short-term motion prediction: instead of learning the absolute future position directly, the network learns how the vessel moves relative to its latest observed state.

This paper combines maneuver-aware features, residual displacement prediction, and multi-scale LSTM encoding in a single AIS-only framework. The fixed and maneuver-guided variants are evaluated together so that accuracy gains from residual local-motion learning, maneuver-aware features, fixed multi-scale representation, and adaptive scale selection can be separated. The design keeps the evaluation centered on local motion history and maneuver-derived temporal-scale selection, providing a compact setting that can later be combined with richer contextual information.

3. Data and Preprocessing

3.1. Data Source and Study Area

The experiments use public AIS records distributed through Marine Cadastre AccessAIS by the NOAA Office for Coastal Management. Marine Cadastre provides coastal and marine planning data, including historical AIS records collected from terrestrial and satellite AIS receivers for U.S. waters. The AccessAIS download request covers 5 February 2023 to 7 April 2023 within the bounding box $37.0591\text{--}38.7663^\circ$ N and longitudes -123.5067° to -121.7873° . After preprocessing and 5 min resampling, the retained records cover San Francisco Bay and adjacent port-approach waters, including the Golden Gate, Central Bay, Oakland port area, and Pacific approaches. The original data contain 10,475,957 AIS records. The main fields include Maritime Mobile Service Identity (MMSI), timestamp, latitude, longitude, speed over ground (SOG), course over ground (COG), heading, vessel type, navigational status, vessel length, and vessel width. Figure 1 shows the spatial density of the retained AIS trajectories after resampling.

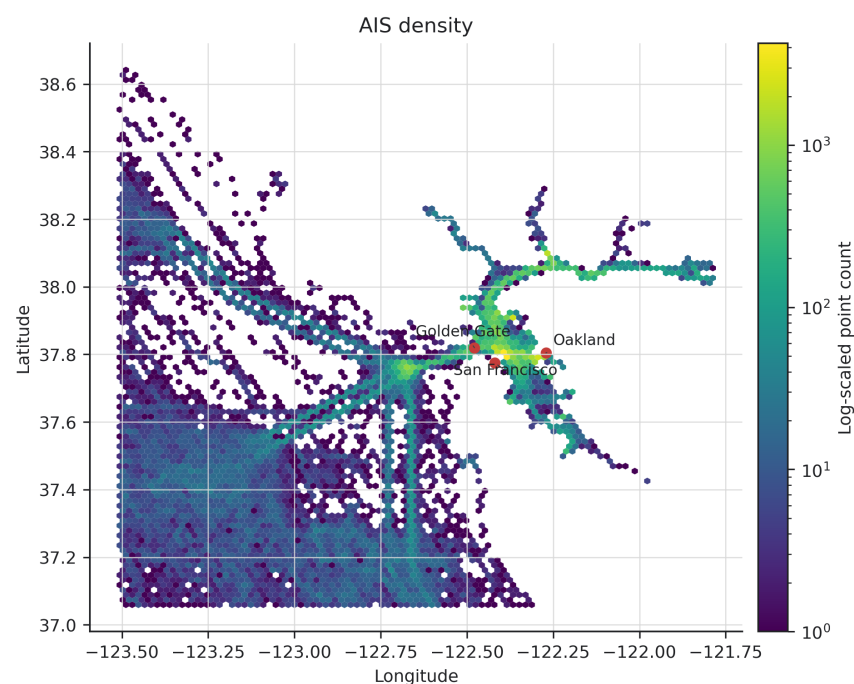


Figure 1. AIS trajectory density after resampling.

3.2. AIS Data Cleaning

All cleaning rules are applied before model training and uniformly to every vessel. An MMSI is retained only when it contains exactly nine digits, and timestamps must be parseable. Positions must be present and geographically valid; the AccessAIS request limits the source records to the study-area bounding box stated in Section 3.1. SOG must lie in the closed interval 0–50 knots, COG must lie in $[0, 360)$, and duplicate MMSI–timestamp records are removed. A record is also removed when a reported vessel length is below 3 m or a reported width is below 2 m; missing dimensions alone are retained. Navigational status is restricted to codes 0, 2, 3, 4, 7, and 8, corresponding to underway or active maneuvering states, and records with SOG below 0.5 knots are excluded from the moving-transit task. These rules retain deceleration during a continuous transit while excluding prolonged anchoring, mooring, and berthing dwell records.

The filtering defines a short-term moving-transit prediction task. Berthing dwell, anchorage waiting, and long near-stationary periods have different targets, such as departure time, berth operation, or anchorage behavior. Including these segments would increase the share of near-zero displacement targets and bias training toward stationary predictions.

Adjacent-point implied speed is computed from the Haversine distance. Short-gap jumps with implied speed above 60 knots are removed. The 60-knot threshold is used as a safety check for gross AIS position discontinuities. It is above the normal operating speed of commercial vessels in port-approach waters and affected only 32 rows in this dataset; therefore, this rule mainly removes extreme jumps and has limited influence on the sample distribution. Trajectories are split when adjacent records are separated by more than 30 min or when a remaining jump indicates a discontinuity. The 30 min value is a trajectory-continuity rule; forecast horizons are set separately during supervised sample construction. Table 2 lists the preprocessing criteria, and Table 3 reports the resulting record and sample counts.

Table 2. Rule-based AIS preprocessing criteria.

Rule Category	Criterion	Purpose
Identifier	Missing MMSI or a value other than nine digits	Remove non-trackable records
Time	Invalid timestamp or duplicate MMSI–timestamp	Ensure unique, ordered observations
Position	Missing/invalid coordinates; source request restricted to the study-area bounding box	Retain records in the requested region
Motion	SOG outside 0–50 knots or COG outside $[0, 360)$	Remove invalid motion records
Vessel dimensions	Reported length < 3 m or reported width < 2 m	Remove unreliable static records
Navigation status	Status codes 0, 2, 3, 4, 7, and 8 retained; SOG < 0.5 knots removed	Define the moving-transit task boundary
Physical jump	Adjacent-point implied speed > 60 knots	Remove gross position discontinuities
Long temporal gap	Adjacent records separated by > 30 min	Split discontinuous trajectories
Interpolation gap	Continuous missing span > 15 min is not bridged	Avoid interpolation across long gaps

Table 3. Data cleaning and sample construction summary.

Item	Value
Raw AIS records	10,475,957
Rows kept after rule filters	1,567,447
Duplicate MMSI–timestamp rows removed	90
Final cleaned rows before trajectory logic	1,567,357
Short-gap physical jump rows removed	32
Trajectory split boundaries	14,355
Trajectory segments written	14,989
Rows after trajectory construction	1,567,325
Valid 5 min resampled fragments	5,832
Resampled feature rows	312,747
Supervised train/validation/test samples	122,831/42,509/48,263

3.3. Resampling and Interpolation

Because raw AIS intervals are irregular, each trajectory segment is resampled to a 5 min interval. This interval is kept fixed for all horizons so that performance differences mainly reflect forecast length rather than a changed temporal discretization; it also gives the three convolution kernels stable interpretations of approximately 15, 25, and 35 min. The local planar coordinates, SOG, and the sine and cosine of COG are linearly interpolated. Direct interpolation of COG is avoided because course angles are circular. COG is converted to $\sin(\text{COG})$ and $\cos(\text{COG})$, interpolated in this representation, normalized, and then converted back when required. Continuous missing spans longer than 15 min are not bridged; resulting timestamp gaps split the trajectory so supervised samples do not cross missing intervals.

The 5 min interval supports the evaluated 10, 20, and 30 min horizons for near-term traffic monitoring and collision-risk screening. Raw AIS reports arrive at irregular intervals and can contain duplicate messages, gaps, and measurement noise. A fixed interval gives stable sequence lengths and temporal windows that can be interpreted consistently across samples. Sine–cosine interpolation reduces COG wrap-around errors near $0^\circ/360^\circ$. Accordingly, $d\text{COG}$ and turning_rate describe course change over each 5 min interval; they do not represent second-level turning dynamics. Very short-duration maneuvers may be smoothed by this temporal discretization, and finer sampling may be required when such dynamics are the prediction target.

Fragments shorter than 24 resampled points are removed because the supervised setting uses 12 input steps and up to 6 prediction steps, requiring at least 18 consecutive points for a single 30 min forecasting sample. The 24-point threshold ensures that each retained fragment can provide more than one sliding-window sample and a stable estimate of window-level maneuver statistics. It is a fixed-length sample-construction requirement for the present model.

3.4. Coordinate Transformation and Feature Construction

Latitude and longitude are converted into local meter coordinates using an equirectangular approximation around the data mean $(\text{lat}_0, \text{lon}_0) = (37.78905148, -122.50270701)$:

$$x = R \cos(\text{lat}_0)(\text{lon} - \text{lon}_0), \quad y = R(\text{lat} - \text{lat}_0), \quad (1)$$

where angles are in radians and $R = 6,371,000$ m. In this two-dimensional prediction setting, the local x and y axes are treated as the horizontal East and North components of a local East–North–Up (ENU) frame around the study-area reference point; the vertical Up component is not used.

For each resampled point, the sequence feature vector is

$$X_{\text{seq}} = [x, y, \text{SOG}, \sin(\text{COG}), \cos(\text{COG}), dx, dy, d\text{SOG}, d\text{COG}, \text{turning_rate}, \text{acceleration}], \quad (2)$$

where

$$\begin{aligned} dx_t &= x_t - x_{t-1}, & dy_t &= y_t - y_{t-1}, \\ d\text{SOG}_t &= \text{SOG}_t - \text{SOG}_{t-1}, & d\text{COG}_t &= ((\text{COG}_t - \text{COG}_{t-1} + 180) \bmod 360) - 180, \\ \text{turning_rate}_t &= \frac{d\text{COG}_t}{\Delta t}, & \text{acceleration}_t &= \frac{d\text{SOG}_t}{\Delta t}, \end{aligned} \quad (3)$$

for two adjacent resampled points at time steps $t - 1$ and t , with the fixed sampling interval $\Delta t = 5$ min. The variables x , y , dx , and dy are measured in meters; SOG and $d\text{SOG}$ in knots;

COG and $d\text{COG}$ in degrees; turning_rate in degrees per minute; and acceleration in knots per minute. The gate feature vector is computed over the input window:

$$X_{\text{gate}} = \begin{bmatrix} \text{mean}|d\text{COG}|, \max|d\text{COG}|, \text{std}(d\text{COG}), \\ \text{mean}|d\text{SOG}|, \max|d\text{SOG}|, \text{std}(d\text{SOG}), \\ \text{mean}(\text{SOG}), \text{std}(\text{SOG}) \end{bmatrix}. \quad (4)$$

Thus, X_{seq} describes pointwise motion evolution, while X_{gate} summarizes the maneuver intensity of the whole observation window.

3.5. Supervised Sample Construction

The final dataset uses a 5 min sampling interval, an input length of 12 steps (60 min), and a prediction length of 6 steps (30 min). Evaluation is reported at 2, 4, and 6 future steps, corresponding to 10, 20, and 30 min. The 10 and 20 min horizons are treated as the primary short-term prediction setting, while the 30 min horizon is retained as an extended test of whether recent maneuver information remains useful when uncertainty increases. The train, validation, and test sets are split by MMSI with a 70/15/15 ratio to reduce vessel-level leakage. The split contains 387, 82, and 84 MMSIs, respectively, and uses random seed 42. The random seed is a fixed initialization value used to make the MMSI-level train/validation/test split reproducible. Feature scalers and target scalers are fitted on the training split only and are then applied to the validation and test splits.

4. Methodology

The model combines residual trajectory prediction, maneuver-aware feature representation, multi-scale temporal convolution, maneuver-guided fusion, and recurrent encoding, as detailed below.

4.1. Problem Formulation

For each sample, let the observed sequence be

$$X_t = \{x_{t-L+1}, \dots, x_t\}, \quad L = 12, \quad (5)$$

where $L = 12$ is the observation length. Let the prediction horizon be $H \in \{2, 4, 6\}$. The target is defined as the cumulative residual displacement relative to the last observed position:

$$Y_t = \{\Delta p_{t+1}, \dots, \Delta p_{t+H}\}, \quad \Delta p_{t+k} = p_{t+k} - p_t. \quad (6)$$

The predicted absolute position is then recovered by

$$\hat{p}_{t+k} = p_t + \widehat{\Delta p}_{t+k}. \quad (7)$$

4.2. Residual Trajectory Prediction

Residual prediction is adopted because short-term vessel motion is locally continuous. A model that directly predicts absolute coordinates must learn both the regional location distribution and the displacement pattern of each vessel. In contrast, residual prediction focuses the learning task on local motion relative to the latest observed point. This formulation is also closer to the operational question faced by VTS operators: given the recent motion of a vessel, where is it likely to move next?

4.3. Maneuver-Aware Feature Representation

The feature representation separates position, local displacement, speed variation, and course variation. The terms dx and dy describe local spatial motion, $dSOG$ reflects acceleration or deceleration tendency, and $dCOG$ and turning rate describe course change. Sine and cosine COG terms are used to avoid the angular discontinuity between 359° and 0° . These features are particularly relevant in port-approach waters, where vessels frequently adjust course and speed over short time intervals.

4.4. Multi-Scale Temporal Convolution

The multi-scale module uses three one-dimensional convolution branches:

$$F_3 = \text{Conv}_{k=3}(X), \quad F_5 = \text{Conv}_{k=5}(X), \quad F_7 = \text{Conv}_{k=7}(X). \quad (8)$$

With a 5 min sampling interval, these kernels form a temporal filter bank rather than three manually assigned maneuver labels. The $k = 3$ branch has a nominal local receptive field of 15 min and emphasizes recent displacement, speed, and course changes. The $k = 5$ branch covers about 25 min and captures transition patterns lasting several AIS samples. The $k = 7$ branch has a nominal receptive field of 35 min within the 60 min observation window and represents the longest local trend among the three branches. The branches provide short-, medium-, and longer-range candidates for subsequent fusion without hard-coding straight, turning, or speed-changing behavior to a specific branch. Fixed-MS-LSTM uses the same residual prediction target, maneuver-aware input features, parallel temporal convolution branches, and LSTM encoder as MGMS-RLSTM, but fuses the three branches by equal averaging and therefore has no maneuver-guided scale gate.

4.5. Maneuver-Guided Multi-Scale Fusion

MGMS-RLSTM replaces fixed fusion with sample-dependent gate weights generated by a multilayer perceptron (MLP):

$$w = \text{softmax}(\text{MLP}(X_{\text{gate}})), \quad w = [w_3, w_5, w_7], \quad (9)$$

$$F = w_3 F_3 + w_5 F_5 + w_7 F_7. \quad (10)$$

The link between maneuver state and temporal scale is established through X_{gate} . Low values of $\text{mean}|dCOG|$ and $\text{mean}|dSOG|$ indicate approximately steady motion, for which recent displacement and a moderate local trend may be sufficient. High values indicate sustained course or speed changes across the input window and may require medium- or longer-range context to distinguish a continuing maneuver from a transient fluctuation. The gate therefore learns sample-dependent scale selection from maneuver statistics instead of imposing a manual rule, such as assigning all turning samples to the shortest branch. Its performance is evaluated together with the fixed-fusion counterpart across all prediction horizons.

4.6. Diagnostic Residual Fusion Variants

These variants are not part of the final proposed framework; they are diagnostic experiments used to test whether adaptive fusion can be stabilized by falling back to fixed fusion. Direct maneuver-guided fusion replaces the fixed multi-scale representation with a sample-dependent weighted combination of temporal branches. Although this design improves adaptability, it may also disturb the stable representation learned by fixed multi-scale fusion, especially for steady-motion or short-horizon samples. To test whether fixed fusion can be preserved as a stability prior, this study further evaluates a residual maneuver-guided fusion (RMGF) variant.

Specifically, the fixed multi-scale representation is first computed as

$$F_{fixed} = \frac{F_3 + F_5 + F_7}{3}. \quad (11)$$

The maneuver-guided representation is obtained by

$$F_{gate} = w_3 F_3 + w_5 F_5 + w_7 F_7, \quad (12)$$

where $w = [w_3, w_5, w_7]$ is generated from X_{gate} through the softmax gate. Instead of directly replacing the fixed representation, RMGF learns a residual correction:

$$F = F_{fixed} + \alpha(X_{gate})(F_{gate} - F_{fixed}), \quad (13)$$

where $\alpha(X_{gate}) \in [0, 1]$ is generated by a sigmoid MLP. When α approaches 0, the model degenerates to fixed multi-scale fusion; when α approaches 1, the model relies more heavily on maneuver-guided adaptive fusion. Thus, RMGF treats fixed fusion as a stability-preserving prior and learns maneuver-dependent corrections only when adaptive scale selection is beneficial. In this manuscript, RMGF is evaluated only as a diagnostic variant rather than as a final model component.

Because the first RMGF test shows that α may remain high even for straight samples, an additional maneuver-constrained RMGF-v2 diagnostic is evaluated. This variant keeps the same residual fusion form but multiplies the learned correction coefficient by an explicit maneuver-intensity mask:

$$M = \max\left(\frac{\text{mean}|dCOG|}{T_{COG}}, \frac{\text{mean}|dSOG|}{T_{SOG}}\right), \quad (14)$$

$$s = \sigma\left(\frac{M-1}{\tau}\right), \quad (15)$$

$$\alpha = s \cdot \sigma(\text{MLP}_\alpha(X_{gate}) + b), \quad (16)$$

where $T_{COG} = 30.34^\circ$, $T_{SOG} = 1.69$ knots, $\tau = 0.3$, and $b = -2$. The thresholds are the 75th-percentile maneuver-group thresholds computed from the training split. This constrained variant is designed to suppress adaptive correction for low-maneuver windows while allowing larger correction for turning or speed-changing samples.

Figure 2 summarizes the resulting MGMS-RLSTM architecture.

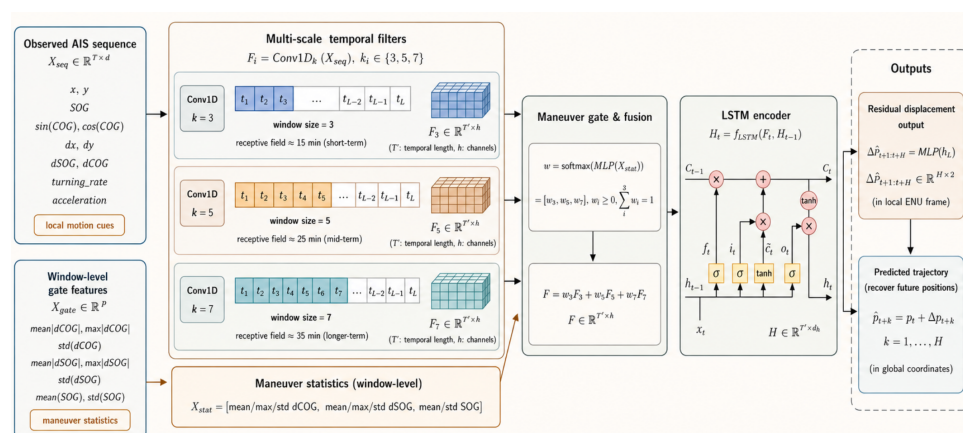


Figure 2. Architecture of MGMS-RLSTM.

4.7. LSTM Encoder and Prediction Head

The fused sequence F is passed to an LSTM encoder:

$$H = \text{LSTM}(F), \quad h_t = H_L. \quad (17)$$

The last hidden state is decoded by a multi-layer perceptron to produce $H \times 2$ residual outputs:

$$\hat{Y} = \text{MLP}(h_t). \quad (18)$$

The model is trained with mean squared error on residual displacement:

$$\mathcal{L} = \frac{1}{H} \sum_{k=1}^H \|\hat{\Delta p}_{t+k} - \Delta p_{t+k}\|^2. \quad (19)$$

All neural models are trained with the implementation protocol described in Section 5.3.

5. Experiments

5.1. Baseline Implementation Protocol

All compared models are evaluated under a unified AIS-only benchmark protocol. Each model is trained and tested on the same San Francisco Bay MMSI-level split, 5 min sampling interval, 12-step input window, 10/20/30 min forecast horizons, training scalers, and local-meter average displacement error (ADE), final displacement error (FDE), mean absolute error (MAE), and root mean square error (RMSE) metrics. This setting ensures that the comparison focuses on model behavior under the same data source, preprocessing pipeline, prediction target, training budget, and evaluation protocol.

IS-STGCNN [13] and TPTrans [14] are included as representative recent neural baselines and are configured according to their core architectural designs under the same benchmark setting. IS-STGCNN represents the spatio-temporal graph convolution family, while TPTrans represents CNN–Transformer trajectory prediction. LMCA [6] is discussed as a related multi-scale LSTM method. Fixed-MS-LSTM is constructed as a strong fixed-fusion counterpart, and MGMS-RLSTM further introduces sample-dependent maneuver-guided scale fusion.

5.2. Baseline Models

Seven models are compared to evaluate classical motion assumptions, standard neural sequence baselines, representative recent neural baselines, the fixed-fusion internal counterpart, and the maneuver-guided adaptive-fusion variant. The constant-velocity (CV) baseline extrapolates the latest local displacement. Vanilla LSTM uses a standard recurrent sequence encoder. CNN-LSTM adds a convolutional front end to the recurrent baseline. IS-STGCNN and TPTrans are included as representative recent neural baselines under the unified AIS-only benchmark. Fixed-MS-LSTM is included as a strong internal counterpart to isolate the effect of maneuver-guided fusion. MGMS-RLSTM uses the same multi-scale residual backbone as Fixed-MS-LSTM, but fuses the branches using maneuver-guided gate weights.

5.3. Implementation Details

All neural models are implemented in PyTorch and trained with the same train/validation/test split and the same optimization budget unless otherwise stated. The maximum number of epochs is 50, the batch size is 512, the optimizer is Adam [23] with learning rate 10^{-3} , and no explicit weight decay is used. The training objective is mean squared error on the standardized target. Validation loss is evaluated after every epoch,

the checkpoint with the lowest validation loss is retained, and early stopping is triggered when validation loss does not improve for 8 consecutive epochs. Gradients are clipped to a maximum norm of 1.0. The MMSI-level train/validation/test split is fixed across experiments, and neural training is repeated with random seeds 42, 2024, and 2025 for parameter initialization, data-loader shuffling, and NumPy/Torch random operations. Unless otherwise stated, the main tables report mean $\pm$ standard deviation over these three seeds.

The hidden dimension is 64 for all neural baselines. Vanilla LSTM, CNN-LSTM, Fixed-MS-LSTM, and MGMS-RLSTM use one LSTM layer. Dropout is set to 0.1 in the convolutional, recurrent-head, Transformer, and gate-related modules where applicable. Fixed-MS-LSTM and MGMS-RLSTM use three temporal convolution branches with kernel sizes 3, 5, and 7; each branch outputs 64 channels. The MGMS-RLSTM gate takes 8 standardized window-level maneuver statistics and uses a two-layer MLP with dimensions $8 \rightarrow 32 \rightarrow 3$, followed by softmax normalization over the three branches. The shared prediction head is an MLP with dimensions $64 \rightarrow 64 \rightarrow 2H$, where H is the prediction horizon in AIS steps. IS-STGCNN uses two temporal graph blocks and a two-layer temporal convolution decoder under the unified benchmark setting. TPTrans uses a 64-dimensional input projection, sinusoidal positional encoding, one local convolution layer, a four-layer Transformer encoder with eight attention heads, and separate x/y decoding heads under the unified benchmark setting. Table 4 summarizes the shared implementation settings.

Table 4. Main implementation settings.

Model	Main Structure	Training Budget
LSTM	One-layer LSTM, hidden size 64; MLP prediction head $64 \rightarrow 64 \rightarrow 2H$	Shared
CNN-LSTM	Two 1D convolution layers with 64 channels and $k = 3$; one-layer LSTM; MLP prediction head	Shared
IS-STGCNN	Spatio-temporal graph convolution blocks with temporal convolution decoder	Shared
TPTrans	CNN-Transformer encoder with positional encoding and separate x/y heads	Shared
Fixed-MS-LSTM	Three 1D branches with $k = 3, 5, 7$, 64 channels each; equal branch averaging; one-layer LSTM; MLP prediction head	Shared
MGMS-RLSTM	Same branches and LSTM as Fixed-MS-LSTM; gate MLP $8 \rightarrow 32 \rightarrow 3$ for adaptive branch fusion	Shared

Experiments are run on a Windows 11 workstation with a 13th Gen Intel Core i5-13500H CPU at 2.60 GHz and 16 GB RAM. The recorded experiment logs use Python 3.12.0 and PyTorch 2.11.0 with the CPU backend; eight CPU threads are used for Torch operations. CV is an analytic extrapolation baseline and therefore has no trainable parameters or training budget.

5.4. Ablation Settings

The ablation chain is constructed progressively using residual prediction (RP) and maneuver-aware features (MF): LSTM, LSTM + RP, LSTM + MF + RP, Fixed-MS-LSTM, and MGMS-RLSTM. Fixed-MS-LSTM is retained in the ablation because it compares the maneuver-guided gate directly with its fixed-fusion counterpart. This design prevents the overall improvement from being attributed solely to the gate and clarifies which components contribute most to prediction accuracy.

5.5. Evaluation Metrics

Metrics are computed in local x/y coordinates in meters. Average displacement error (ADE) and final displacement error (FDE) are defined as

$$ADE = \frac{1}{H} \sum_{k=1}^H \|\hat{p}_{t+k} - p_{t+k}\|_2, \quad FDE = \|\hat{p}_{t+H} - p_{t+H}\|_2. \quad (20)$$

Mean absolute error (MAE) and root mean square error (RMSE) are also reported over the x/y components. These metrics describe average geometric accuracy. To supplement mean-error evaluation, the experiments further report median ADE, 90th and 95th percentile ADE, large-error rates above 1 km and 2 km, parameter count, model size, training time, and batched CPU inference latency.

To make the evaluation more relevant to VTS collision-risk screening, this study computes encounter-oriented CPA/TCPA metrics from saved test-set predictions. For each prediction start time, test vessels observed at the same timestamp are paired when their current separation is greater than 1 nautical mile and no greater than 6 nautical miles. The evaluated horizons are 10–30 min, and this candidate-generation rule targets near-term encounter screening. The 1 nautical mile lower bound removes pairs already inside the high-risk threshold at prediction start. The 6 nautical mile upper bound limits the candidate set to pairs with plausible interaction within the evaluated horizons; a wider fixed radius includes more distant non-interacting pairs and increases class imbalance. For a vessel pair (i, j) , CPA and TCPA are computed over the piecewise-linear future relative trajectory:

$$CPA_{ij} = \min_{\tau \in [0, H\Delta t]} \|\hat{p}_i(\tau) - \hat{p}_j(\tau)\|_2, \quad (21)$$

$$TCPA_{ij} = \arg \min_{\tau \in [0, H\Delta t]} \|\hat{p}_i(\tau) - \hat{p}_j(\tau)\|_2, \quad (22)$$

where $\Delta t = 5$ min. The same computation is applied to the true future trajectories and to each model prediction. A future high-risk encounter is defined as a candidate pair whose true $CPA \leq 1$ nautical mile within the evaluated horizon. The application-oriented metrics are CPA mean absolute error over candidate pairs, TCPA mean absolute error over true high-risk pairs, and precision, recall, and F1 for high-risk encounter screening.

5.6. Maneuver-Group Evaluation

To examine whether the model behaves differently under different motion states, test samples are divided into Straight, Turning, and Speed-changing groups. Turning samples are those whose mean absolute $dCOG$ exceeds the 75th percentile of the training set. Speed-changing samples are those whose mean absolute $dSOG$ exceeds the 75th percentile of the training set. Straight samples exceed neither threshold. Turning and Speed-changing groups may overlap. In this dataset, the thresholds are $mean|dCOG| = 30.34^\circ$ and $mean|dSOG| = 1.69$ knots. This coarse grouping is used as a diagnostic stratification for examining whether scale selection and prediction error vary with maneuver intensity.

6. Results and Discussion

6.1. Overall Prediction Performance

The main three-seed performance results are reported in Table 5, and Figure 3 visualizes the ADE trends across horizons.

The comparison in Table 5 and Figure 3 includes a classical kinematic baseline (CV), standard neural sequence baselines (LSTM and CNN-LSTM), two representative recent neural baselines (IS-STGCNN and TPTrans), the fixed-fusion internal counterpart (Fixed-MS-LSTM), and the proposed maneuver-guided variant (MGMS-RLSTM). Values are reported as mean $\pm$ standard deviation over seeds 42, 2024, and 2025, and lower values are better. Overall, the largest observed gains come from residual prediction and maneuver-aware features, indicating that target formulation and local-motion feature construction are central to this task. Multi-scale temporal encoding further provides a strong lightweight backbone. Fixed-MS-LSTM remains slightly stronger at 10 min, whereas MGMS-RLSTM obtains lower ADE at 20 and 30 min in the three-seed main evaluation. This pattern

characterizes maneuver-guided fusion as a horizon-dependent scale-selection mechanism that complements the fixed multi-scale counterpart.

Table 5. Three-seed overall prediction performance.

Horizon	Model	MAE (m)	RMSE (m)	ADE (m)	FDE (m)
10 min	CV	692.9 ± 0.0	1375.0 ± 0.0	1094.7 ± 0.0	1579.6 ± 0.0
10 min	LSTM	876.1 ± 74.7	1265.7 ± 103.0	1374.3 ± 116.5	1659.3 ± 128.7
10 min	CNN-LSTM	1147.2 ± 116.1	1552.7 ± 127.9	1788.6 ± 179.3	1983.7 ± 224.6
10 min	Fixed-MS-LSTM	344.9 ± 3.3	659.0 ± 12.6	541.9 ± 6.2	744.4 ± 11.6
10 min	IS-STGCNN	742.8 ± 47.3	1100.1 ± 78.2	1170.0 ± 79.3	1341.9 ± 88.4
10 min	TPTrans	743.5 ± 17.3	1087.6 ± 45.5	1163.4 ± 25.1	1381.6 ± 27.1
10 min	MGMS-RLSTM	349.7 ± 7.9	673.0 ± 29.1	548.7 ± 12.4	752.8 ± 17.2
20 min	CV	1452.5 ± 0.0	2911.6 ± 0.0	2289.1 ± 0.0	4170.0 ± 0.0
20 min	LSTM	1212.6 ± 136.6	1827.0 ± 139.9	1884.7 ± 210.3	2727.8 ± 228.3
20 min	CNN-LSTM	1264.9 ± 60.4	1873.0 ± 98.4	1985.5 ± 91.8	2742.2 ± 174.0
20 min	Fixed-MS-LSTM	654.4 ± 20.5	1245.6 ± 38.5	1027.7 ± 34.8	1694.7 ± 61.7
20 min	IS-STGCNN	1112.5 ± 97.8	1672.3 ± 113.7	1746.8 ± 152.8	2399.9 ± 206.7
20 min	TPTrans	944.3 ± 39.3	1523.5 ± 50.8	1476.9 ± 62.7	2171.5 ± 116.8
20 min	MGMS-RLSTM	650.0 ± 6.7	1257.3 ± 10.6	1020.3 ± 11.4	1681.3 ± 15.7
30 min	CV	2305.1 ± 0.0	4592.1 ± 0.0	3626.7 ± 0.0	7010.6 ± 0.0
30 min	LSTM	1467.1 ± 97.2	2300.9 ± 130.2	2290.4 ± 143.0	3422.9 ± 157.3
30 min	CNN-LSTM	1487.7 ± 38.8	2240.3 ± 1.4	2324.4 ± 45.0	3252.9 ± 27.6
30 min	Fixed-MS-LSTM	969.0 ± 9.7	1827.6 ± 31.1	1519.4 ± 16.1	2601.4 ± 23.5
30 min	IS-STGCNN	1327.1 ± 16.7	2063.3 ± 31.0	2075.0 ± 30.0	3058.4 ± 61.1
30 min	TPTrans	1252.0 ± 102.3	2031.0 ± 129.2	1958.0 ± 162.1	2992.0 ± 250.9
30 min	MGMS-RLSTM	954.5 ± 9.1	1830.3 ± 30.3	1497.9 ± 17.7	2569.9 ± 35.6

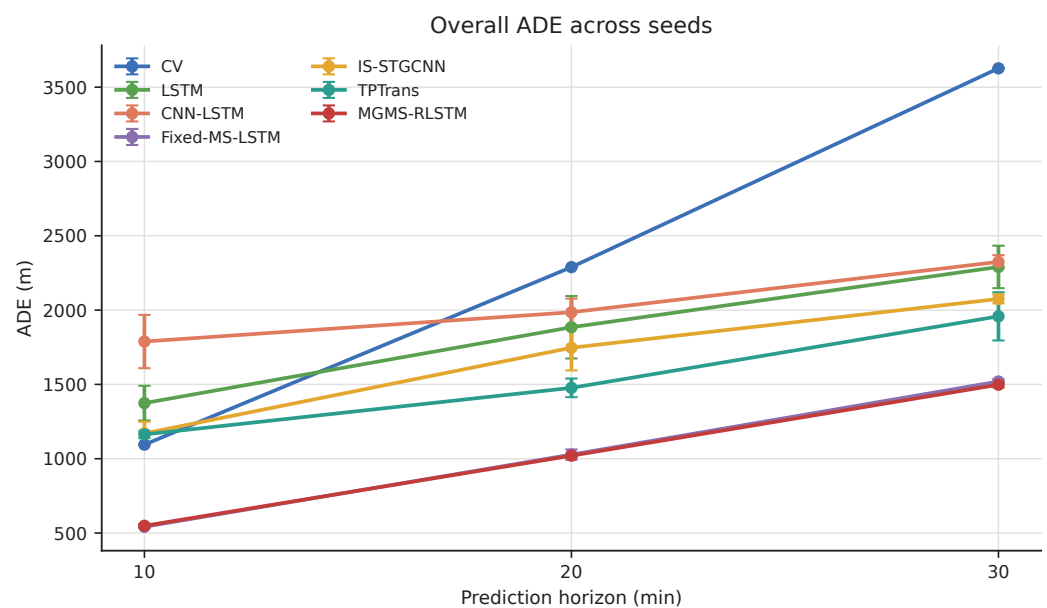


Figure 3. Overall ADE comparison across horizons.

6.2. Distributional Error and Efficiency

The distributional error statistics are summarized in Table 6, and Figure 4 shows the 30 min ADE distribution.

Table 6. Three-seed 30 min ADE distribution.

Model	Median (m)	P90 (m)	P95 (m)	ADE > 1 km (%)	ADE > 2 km (%)
CV	1882.1 ± 0.0	9945.3 ± 0.0	13003.5 ± 0.0	65.8 ± 0.0	48.3 ± 0.0
LSTM	1805.8 ± 124.1	4379.1 ± 258.7	5687.5 ± 357.8	77.8 ± 1.7	44.2 ± 3.8
CNN-LSTM	1884.8 ± 60.9	4233.3 ± 85.4	5430.3 ± 51.0	82.9 ± 0.7	46.3 ± 2.3
Fixed-MS-LSTM	1016.6 ± 27.3	3267.4 ± 93.4	4578.0 ± 170.7	50.7 ± 1.2	22.7 ± 0.7
IS-STGCNN	1624.2 ± 55.3	3952.1 ± 86.0	5113.8 ± 113.7	77.9 ± 2.6	36.8 ± 0.9
TPTrans	1491.7 ± 150.7	3803.1 ± 296.7	5091.9 ± 338.1	70.9 ± 4.1	33.9 ± 5.2
MGMS-RLSTM	981.3 ± 3.5	3280.4 ± 63.2	4625.7 ± 128.9	49.1 ± 0.2	21.9 ± 0.3

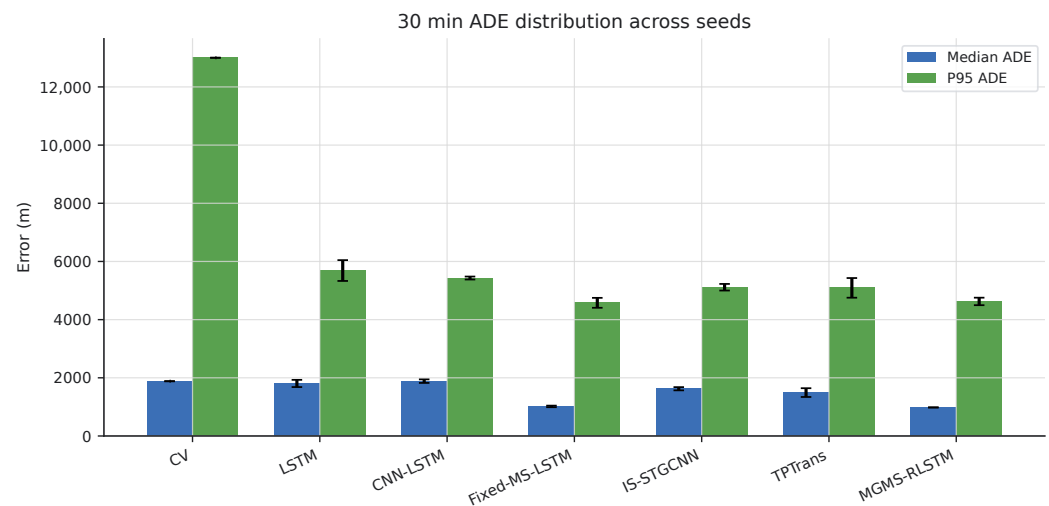


Figure 4. Thirty-minute ADE distribution.

The distributional results in Table 6 and Figure 4 show that the mean-error ranking is not driven by only a small number of easy samples. Values are reported as mean $\pm$ standard deviation over seeds 42, 2024, and 2025. At the 30 min horizon, MGMS-RLSTM achieves the lowest median ADE and the lowest large-error rates, while Fixed-MS-LSTM remains the closest residual multi-scale counterpart and has a slightly lower P95 ADE. IS-STGCNN and TPTrans reduce some large-error rates relative to weaker sequence baselines, but their high-percentile errors remain larger than those of the residual multi-scale models under the unified benchmark.

Table 7 compares model complexity and inference latency, and Figure 5 provides the corresponding visual comparison.

Table 7. Neural-model complexity and batched CPU latency.

Model	Parameters	Size (MB)	Inference (ms/Sample)
LSTM	22,860	0.091	0.0122 $\pm$ 0.0086
CNN-LSTM	51,404	0.201	0.0249 $\pm$ 0.0221
Fixed-MS-LSTM	48,972	0.192	0.0297 $\pm$ 0.0267
IS-STGCNN	76,172	0.335	0.0256 $\pm$ 0.0055
TPTrans	230,412	0.917	0.0723 $\pm$ 0.0229
MGMS-RLSTM	49,359	0.195	0.0192 $\pm$ 0.0057

Table 7 and Figure 5 compare neural-model complexity and 30 min batched CPU inference latency. Parameter counts are deterministic for each architecture, and latency is reported as mean $\pm$ standard deviation over seeds. The maneuver-guided gate adds only a small number of parameters relative to Fixed-MS-LSTM, while both residual multi-scale models remain lightweight. TPTrans has the largest parameter count among the compared neural baselines because of its Transformer encoder. All compared neural models remain below 0.1 ms per sample under batched CPU inference, indicating that the proposed framework is computationally lightweight for offline rolling prediction.

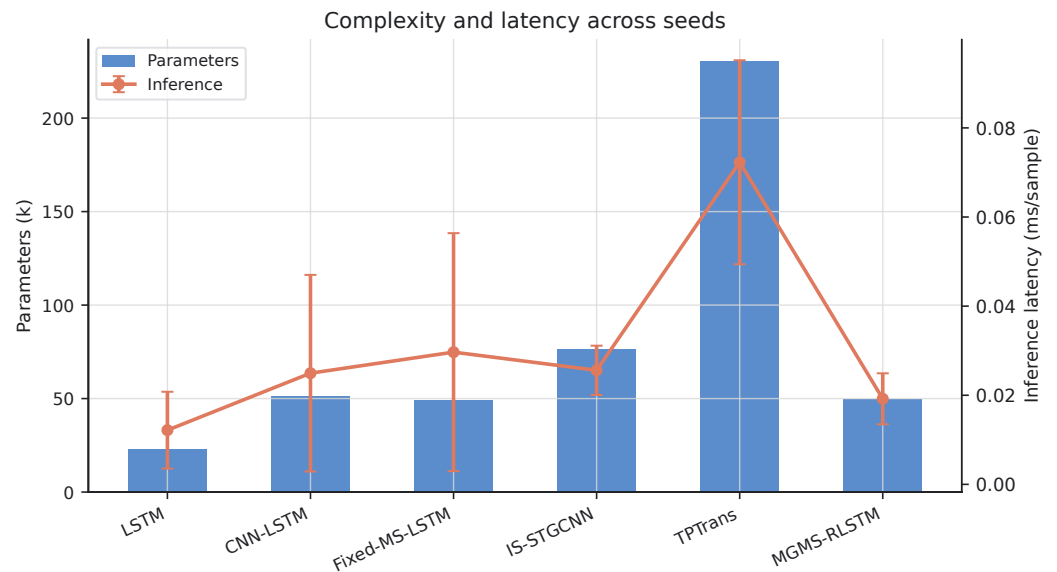


Figure 5. Model complexity and CPU inference latency.

6.3. Encounter-Oriented CPA/TCPA Evaluation

Table 8 summarizes the encounter-screening cases used for application-oriented evaluation. Candidate pairs are constructed only from vessels observed at the same prediction start time and currently separated by 1–6 nautical miles. This range focuses candidate generation on the 10–30 min screening task: the lower bound excludes pairs already within the high-risk threshold, and the upper bound limits the number of distant pairs unlikely to interact within the evaluated horizons. The number of future high-risk pairs increases with prediction horizon: 1776 pairs become high-risk within 10 min, 2905 within 20 min, and 3709 within 30 min. The median true TCPA also increases from 8.5 to 16.9 min, indicating that longer horizons expose additional potential encounters earlier but make accurate screening more difficult.

Table 8. Encounter-screening setup.

Horizon	Candidate Pairs	High-Risk Pairs	High-Risk Rate (%)	Median/P90 TCPA (min)
10 min	10,810	1776	16.4	8.5/10.0
20 min	10,810	2905	26.9	13.0/20.0
30 min	10,810	3709	34.3	16.9/30.0

The candidate-radius sensitivity check in Table 9 reuses the saved three-seed test predictions without model retraining. Expanding the upper bound increases the candidate set and reduces the high-risk rate because many added pairs remain distant over the evaluated horizon. The residual-model F1 ordering remains the same at each horizon: Fixed-MS-LSTM is slightly higher at 10 and 30 min, whereas MGMS-RLSTM is slightly higher at 20 min.

Table 9. Candidate-radius sensitivity for encounter screening.

Upper Bound (NM)	Candidate Pairs	High-Risk Rate (%) 10/20/30 min	Fixed-MS-LSTM F1 10/20/30 min	MGMS-RLSTM F1 10/20/30 min
6	10,810	16.43/26.87/34.31	0.766/0.706/0.680	0.761/0.710/0.672
8	13,902	13.27/22.81/30.20	0.767/0.697/0.666	0.762/0.702/0.658
10	17,027	10.98/19.66/26.43	0.770/0.695/0.662	0.764/0.701/0.655

The 6 nautical mile setting retains a larger high-risk share and slightly higher 20 and 30 min F1 than the wider settings, while the 10 min F1 changes by less than 0.01. The 6 NM threshold is a candidate-generation setting for this benchmark rather than a general navigational safety boundary; wider or speed-adaptive radii should be tested in future cross-port studies. The detailed CPA/TCPA screening results are reported in Table 10, and Figure 6 summarizes the high-risk screening F1 scores.

Table 10. CPA/TCPA error and high-risk encounter screening.

Horizon	Model	CPA-MAE (m)	TCPA-MAE (min)	Precision	Recall	F1
10 min	CV	807.7 ± 0.0	2.8 ± 0.0	0.796 ± 0.000	0.479 ± 0.000	0.598 ± 0.000
10 min	LSTM	789.4 ± 88.4	2.6 ± 0.3	0.767 ± 0.036	0.473 ± 0.054	0.585 ± 0.052
10 min	CNN-LSTM	925.2 ± 68.4	2.9 ± 0.1	0.704 ± 0.031	0.506 ± 0.031	0.589 ± 0.032
10 min	Fixed-MS-LSTM	478.0 ± 7.1	1.7 ± 0.1	0.876 ± 0.010	0.681 ± 0.011	0.766 ± 0.003
10 min	IS-STGCNN	695.3 ± 30.8	2.5 ± 0.2	0.819 ± 0.004	0.560 ± 0.021	0.665 ± 0.016
10 min	TPTrans	714.7 ± 11.9	2.5 ± 0.1	0.805 ± 0.018	0.540 ± 0.015	0.646 ± 0.017
10 min	MGMS-RLSTM	476.6 ± 13.6	1.8 ± 0.0	0.875 ± 0.004	0.674 ± 0.009	0.761 ± 0.006
20 min	CV	1446.1 ± 0.0	7.1 ± 0.0	0.772 ± 0.000	0.368 ± 0.000	0.498 ± 0.000
20 min	LSTM	1191.3 ± 54.9	5.1 ± 0.4	0.764 ± 0.007	0.445 ± 0.039	0.562 ± 0.032
20 min	CNN-LSTM	1253.0 ± 18.5	5.2 ± 0.2	0.711 ± 0.014	0.484 ± 0.009	0.576 ± 0.010
20 min	Fixed-MS-LSTM	848.3 ± 32.0	3.6 ± 0.1	0.834 ± 0.012	0.612 ± 0.030	0.706 ± 0.023
20 min	IS-STGCNN	1135.2 ± 57.6	5.0 ± 0.2	0.766 ± 0.012	0.492 ± 0.021	0.599 ± 0.018
20 min	TPTrans	1081.1 ± 21.6	4.7 ± 0.1	0.779 ± 0.007	0.510 ± 0.002	0.617 ± 0.003
20 min	MGMS-RLSTM	844.0 ± 25.7	3.6 ± 0.1	0.841 ± 0.006	0.614 ± 0.007	0.710 ± 0.003
30 min	CV	1866.2 ± 0.0	11.6 ± 0.0	0.786 ± 0.000	0.313 ± 0.000	0.447 ± 0.000
30 min	LSTM	1383.3 ± 50.0	7.2 ± 0.4	0.726 ± 0.017	0.439 ± 0.026	0.547 ± 0.024
30 min	CNN-LSTM	1354.5 ± 29.7	6.7 ± 0.1	0.696 ± 0.010	0.522 ± 0.013	0.596 ± 0.010
30 min	Fixed-MS-LSTM	1062.3 ± 19.8	5.3 ± 0.2	0.808 ± 0.006	0.587 ± 0.012	0.680 ± 0.010
30 min	IS-STGCNN	1294.6 ± 34.2	6.7 ± 0.1	0.728 ± 0.004	0.508 ± 0.014	0.598 ± 0.009
30 min	TPTrans	1264.8 ± 60.9	6.6 ± 0.5	0.744 ± 0.026	0.506 ± 0.021	0.602 ± 0.022
30 min	MGMS-RLSTM	1085.6 ± 16.7	5.7 ± 0.3	0.808 ± 0.001	0.575 ± 0.008	0.672 ± 0.006

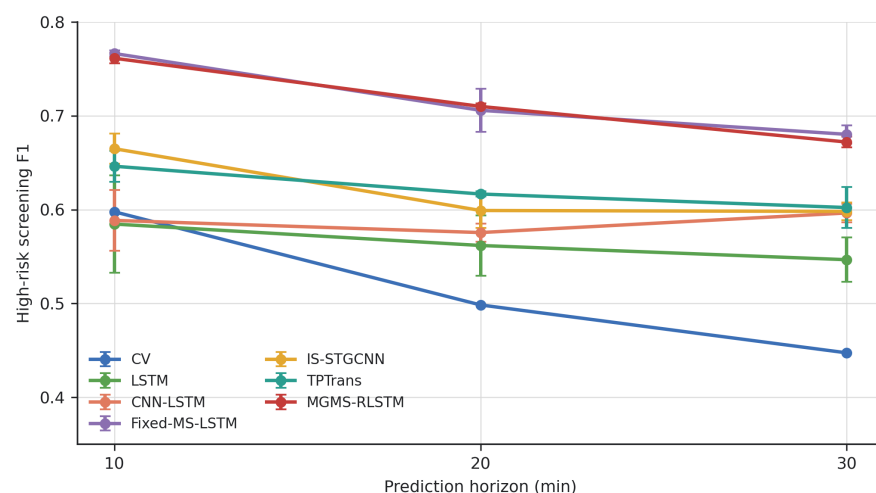


Figure 6. High-risk encounter screening F1.

The encounter-oriented results in Table 10 and Figure 6 provide a more operational interpretation of the geometric prediction errors. Values in Table 10 are reported as mean ± standard deviation over seeds 42, 2024, and 2025. The strongest models under ADE are also strong under CPA-based screening, but the encounter metrics are not simply duplicates of pointwise error. The encounter-oriented evaluation shows that both Fixed-MS-LSTM and MGMS-RLSTM support more accurate CPA/TCPA-based high-risk screening under the evaluated benchmark compared with classical, standard neural, and representative recent neural baselines. At 10 min, Fixed-MS-LSTM obtains the highest screening F1, while MGMS-RLSTM gives the lowest CPA-MAE and remains within 0.005 F1. At 20 min, MGMS-RLSTM gives the lowest CPA-MAE and highest F1. At 30 min, Fixed-MS-LSTM achieves the lowest CPA-MAE and highest F1 because it preserves recall slightly better under the high-risk threshold. MGMS-RLSTM provides adaptive scale in-

terpretation and competitive CPA/TCPA performance, while Fixed-MS-LSTM remains a strong operational counterpart when short-horizon stability or screening F1 is prioritized. These results indicate that the multi-scale residual models improve encounter-oriented screening under the evaluated benchmark, while also showing that longer horizons trade earlier warning opportunity for lower screening reliability.

6.4. Ablation Study

To separate component effects from the primary model comparison, Table 11 reports a seed-42 diagnostic ablation. In contrast, the main comparison in Table 5 reports mean $\pm$ standard deviation over three random seeds. Therefore, Table 11 is used to diagnose the contribution of residual prediction, maneuver-aware features, and multi-scale fusion, rather than to determine the main comparison between Fixed-MS-LSTM and MGMS-RLSTM.

Table 11. Seed-42 diagnostic ablation results.

Horizon	Variant	MAE (m)	RMSE (m)	ADE (m)	FDE (m)
10 min	LSTM	799.4	1146.6	1253.7	1502.0
10 min	LSTM + RP	375.7	701.3	588.4	814.7
10 min	LSTM + MF + RP	357.2	679.9	559.3	771.3
10 min	Fixed-MS-LSTM	347.6	676.1	545.8	754.8
10 min	MGMS-RLSTM	351.0	677.0	549.7	754.1
20 min	LSTM	1112.2	1703.6	1733.9	2533.5
20 min	LSTM + RP	720.0	1350.3	1125.7	1860.4
20 min	LSTM + MF + RP	670.0	1276.5	1050.9	1736.0
20 min	Fixed-MS-LSTM	633.0	1196.4	990.1	1630.6
20 min	MGMS-RLSTM	652.1	1271.1	1022.4	1683.6
30 min	LSTM	1392.7	2166.7	2165.4	3266.9
30 min	LSTM + RP	1032.6	1910.6	1609.3	2754.3
30 min	LSTM + MF + RP	978.9	1862.1	1530.5	2598.6
30 min	Fixed-MS-LSTM	978.3	1816.3	1534.5	2612.0
30 min	MGMS-RLSTM	959.0	1790.9	1501.1	2547.1

Bold values indicate the best performance within each prediction horizon.

Figure 7 visualizes the ablation ADE comparison.

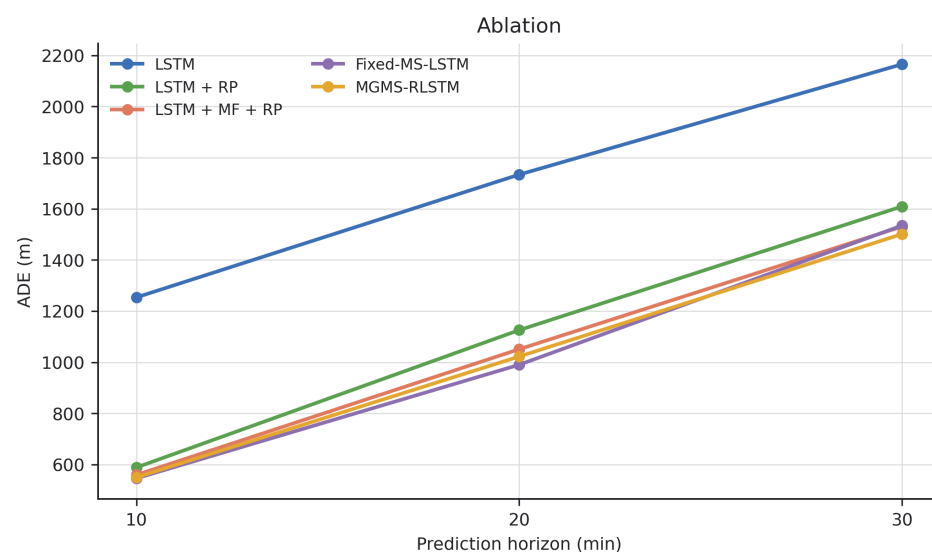


Figure 7. Ablation ADE curves.

The ablation results in Table 11 and Figure 7 clarify the contribution of each design component. Here, RP denotes residual prediction and MF denotes maneuver-aware features. The largest gain comes from residual prediction: at 10 min, ADE decreases from 1253.7 m for LSTM to 588.4 m for LSTM + RP. This indicates that short-term AIS prediction benefits strongly from learning local displacement rather than absolute coordinates. Maneuver-aware features provide further and stable improvements. LSTM + MF + RP outperforms LSTM + RP at all horizons, and the 30 min ADE decreases from 1609.3 m to 1530.5 m, suggesting that course and speed variation features provide useful information about port-approach maneuvering behavior.

In the seed-42 diagnostic ablation, Fixed-MS-LSTM obtains lower ADE than MGMS-RLSTM at 10 and 20 min, while MGMS-RLSTM obtains lower ADE at 30 min. This seed-specific result is reported separately from the three-seed main comparison in Table 5. Under the three-seed main evaluation, Fixed-MS-LSTM remains slightly better at 10 min, whereas MGMS-RLSTM obtains lower ADE at 20 and 30 min. Taken together, the results indicate that the largest observed gains come from residual prediction and maneuver-aware features, while the effect of maneuver-guided fusion varies across prediction horizons.

6.5. Diagnostic Residual Fusion Analysis

Table 12 reports a seed-42 diagnostic comparison between fixed fusion, direct maneuver-guided fusion, and residual maneuver-guided fusion. This experiment is used to examine whether residual correction stabilizes adaptive fusion under the same diagnostic protocol.

Table 12. Seed-42 residual-fusion diagnostic results.

Horizon	Model	MAE (m)	RMSE (m)	ADE (m)	FDE (m)
10 min	Fixed-MS-LSTM	341.1	644.6	534.9	731.0
10 min	Direct MGMS-RLSTM	349.6	657.6	548.7	751.6
10 min	RMGF-RLSTM	358.0	684.3	561.9	769.8
20 min	Fixed-MS-LSTM	650.7	1232.0	1020.8	1681.1
20 min	Direct MGMS-RLSTM	657.3	1267.5	1032.6	1692.6
20 min	RMGF-RLSTM	664.3	1249.2	1039.1	1690.6
30 min	Fixed-MS-LSTM	978.2	1845.1	1534.5	2628.2
30 min	Direct MGMS-RLSTM	944.0	1795.7	1477.5	2531.1
30 min	RMGF-RLSTM	970.1	1823.0	1519.1	2608.5

Bold values indicate the best performance within each prediction horizon.

The corresponding RMGF-RLSTM gate weights and correction coefficients are shown in Table 13.

Table 13. RMGF-RLSTM gate weights and correction coefficient.

Horizon	Maneuver Group	Mean w_3	Mean w_5	Mean w_7	Mean α
10 min	Straight	0.028	0.105	0.866	0.827
10 min	Turning	0.184	0.437	0.379	0.902
10 min	Speed-changing	0.240	0.505	0.254	0.932
20 min	Straight	0.016	0.327	0.657	0.932
20 min	Turning	0.515	0.271	0.214	0.991
20 min	Speed-changing	0.637	0.250	0.113	0.994
30 min	Straight	0.054	0.055	0.891	0.790
30 min	Turning	0.236	0.469	0.296	0.947
30 min	Speed-changing	0.263	0.565	0.173	0.968

Within the separate seed-42 diagnostic run in Tables 12 and 13, RMGF-RLSTM does not improve over the stronger of Fixed-MS-LSTM and Direct MGMS-RLSTM at any horizon. It is worse than Fixed-MS-LSTM at 10 and 20 min, and remains worse than Direct MGMS-RLSTM at 30 min. Therefore, the residual fusion family is retained as a diagnostic experiment and is not selected as the final main method. This diagnostic outcome is consistent with the seed-42 ablation in Table 11 and the three-seed main comparison in Table 5; the attempted residual-fusion stabilization yields no consistent additional accuracy gain in this run.

To address the high straight-sample α values, a maneuver-constrained RMGF-v2 variant is further tested. Table 14 extends the same separate seed-42 diagnostic run by adding the constrained residual-fusion variant. This experiment adds the explicit maneuver-intensity mask described in the methodology and keeps the same training settings and evaluation protocol.

Table 14. Seed-42 constrained RMGF-v2 diagnostic results.

Horizon	Model	MAE (m)	RMSE (m)	ADE (m)	FDE (m)
10 min	Fixed-MS-LSTM	341.1	644.6	534.9	731.0
10 min	Direct MGMS-RLSTM	349.6	657.6	548.7	751.6
10 min	RMGF-RLSTM	358.0	684.3	561.9	769.8
10 min	Constrained RMGF-RLSTM	346.2	661.1	544.8	746.5
20 min	Fixed-MS-LSTM	650.7	1232.0	1020.8	1681.1
20 min	Direct MGMS-RLSTM	657.3	1267.5	1032.6	1692.6
20 min	RMGF-RLSTM	664.3	1249.2	1039.1	1690.6
20 min	Constrained RMGF-RLSTM	663.3	1287.0	1043.9	1716.6
30 min	Fixed-MS-LSTM	978.2	1845.1	1534.5	2628.2
30 min	Direct MGMS-RLSTM	944.0	1795.7	1477.5	2531.1
30 min	RMGF-RLSTM	970.1	1823.0	1519.1	2608.5
30 min	Constrained RMGF-RLSTM	970.2	1846.8	1520.9	2601.7

Bold values indicate the best performance within each prediction horizon.

The constrained RMGF-RLSTM gate weights and correction coefficients are summarized in Table 15.

Table 15. Constrained RMGF-RLSTM gate weights and correction coefficient.

Horizon	Maneuver Group	Mean w_3	Mean w_5	Mean w_7	Mean α
10 min	Straight	0.024	0.015	0.960	0.113
10 min	Turning	0.143	0.431	0.426	0.711
10 min	Speed-changing	0.191	0.502	0.308	0.769
20 min	Straight	0.004	0.038	0.959	0.115
20 min	Turning	0.155	0.596	0.249	0.823
20 min	Speed-changing	0.198	0.644	0.159	0.851
30 min	Straight	0.193	0.008	0.798	0.109
30 min	Turning	0.020	0.644	0.336	0.639
30 min	Speed-changing	0.024	0.764	0.212	0.651

Table 15 shows that the maneuver constraint reduces the straight-group mean α to approximately 0.11 across horizons, whereas the turning and speed-changing group means remain much larger. The constrained variant improves the interpretability of the correction coefficient, but its accuracy gain remains inconsistent across horizons. At 10 min, Constrained RMGF-RLSTM improves over the unconstrained RMGF variant and slightly outperforms Direct MGMS-RLSTM, but it remains worse than Fixed-MS-LSTM. At 20 min, it is worse than both Fixed-MS-LSTM and Direct MGMS-RLSTM. At 30 min, it improves

over Fixed-MS-LSTM but remains worse than Direct MGMS-RLSTM. Therefore, the constrained variant is also kept as a diagnostic experiment, and the RMGF family is not selected as the final main method. The diagnostic analysis indicates improved gate interpretability alongside horizon-dependent accuracy.

6.6. Maneuver-Group Analysis

Table 16 reports the maneuver-group ADE comparison, and Figure 8 compares ADE across maneuver groups and horizons.

Table 16. Maneuver-group ADE comparison between Fixed-MS-LSTM and MGMS-RLSTM.

Horizon	Group	Samples	Fixed-MS ADE (m)	MGMS ADE (m)
10 min	Straight	21,087	247.6	247.1
10 min	Turning	22,073	805.6	813.0
10 min	Speed-changing	23,082	829.7	835.7
20 min	Straight	21,087	529.5	520.8
20 min	Turning	22,073	1355.5	1431.9
20 min	Speed-changing	23,082	1427.9	1494.7
30 min	Straight	21,087	886.9	856.9
30 min	Turning	22,073	2015.5	1970.8
30 min	Speed-changing	23,082	2158.5	2119.9

Bold values indicate the best performance within each prediction horizon and group.

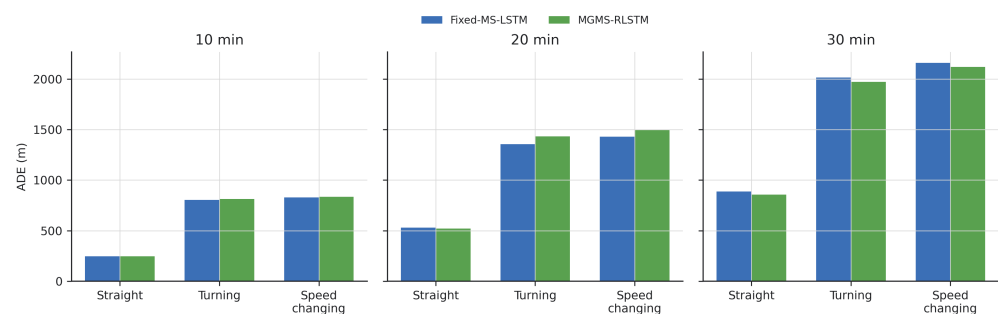


Figure 8. Maneuver-group ADE across horizons.

Figure 9 shows the mean gate weights by maneuver group.

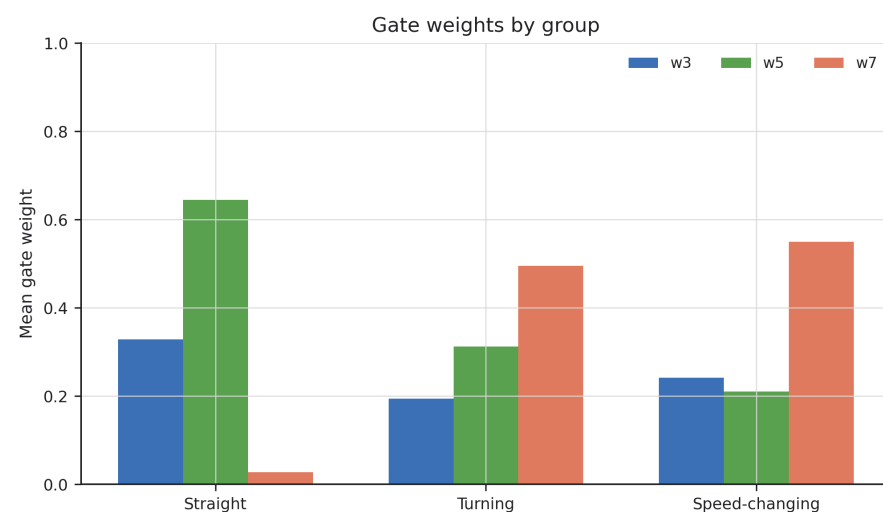


Figure 9. Mean gate weights by maneuver group.

The group evaluation in Table 16 and Figure 8 confirms that maneuvering samples are much more difficult than straight samples. At the 10 min horizon, turning and speed-

changing ADE values are more than three times the straight-group ADE. This supports the motivation for using maneuver-aware features. The relative ranking between Fixed-MS-LSTM and MGMS-RLSTM is horizon-dependent in this group-level view. MGMS-RLSTM is better for all three maneuver groups at 30 min, whereas Fixed-MS-LSTM is better for turning and speed-changing samples at 10 and 20 min within this stratified analysis. Thus, the maneuver gate changes model behavior in a meaningful way, but the resulting accuracy gain is not monotonic across horizons.

6.7. Gate-Weight and Case Analysis

Before interpreting the gate weights, the temporal persistence of the maneuver indicators is checked directly on the test windows. A point is marked as a high-turn point if its absolute $dCOG$ exceeds the 75th percentile of all training-window points (17.35°), and as a high-speed-change point if its absolute $dSOG$ exceeds the corresponding 75th percentile (1.04 knots). Table 17 shows that straight samples usually contain fewer high-change points, whereas turning and speed-changing samples contain multiple high-change points and multi-step runs within the 12-step input window. This supports the use of medium and longer kernels: many port-approach maneuvers in the data are sustained over several 5 min samples, not only one-step impulses.

Table 17. Temporal span of high-maneuver activity.

Group	Samples	High $dCOG$ Count	High $dCOG$ Run	High $dSOG$ Count	High $dSOG$ Run
Straight	21,087	1.56/1	1.16/1	1.24/0	0.92/0
Turning	22,073	6.69/7	3.73/3	6.78/7	4.63/4
Speed-changing	23,082	6.44/6	3.50/3	7.39/7	5.06/4

The mean gate weights and their correlations with maneuver statistics are summarized in Tables 18 and 19, respectively.

Table 18. Mean MGMS-RLSTM gate weights at 10 min.

Group	Samples	w_3	w_5	w_7
Straight	21,087	0.329	0.645	0.027
Turning	22,073	0.194	0.312	0.495
Speed-changing	23,082	0.241	0.210	0.550

For the correlation analysis, the gate weights are taken from the 10 min MGMS-RLSTM model on the test set. Increasing $mean|dCOG|$ is strongly negatively correlated with the short-scale weight w_3 and positively correlated with the longer-scale weight w_7 . Increasing $mean|dSOG|$ has an even stronger positive association with w_7 . In quartile analysis, w_7 increases from 0.013 to 0.506 from the lowest to highest $mean|dCOG|$ quartile, and from 0.014 to 0.671 from the lowest to highest $mean|dSOG|$ quartile. These statistics show that the gate is responding to course and speed variation rather than merely assigning fixed average weights.

Table 19. Correlations between maneuver statistics and gate weights.

Feature	w_3 r	w_3 ρ	w_5 r	w_5 ρ	w_7 r	w_7 ρ
$mean dCOG $	−0.772	−0.793	−0.365	−0.270	0.641	0.827
$mean dSOG $	−0.155	−0.313	−0.915	−0.747	0.926	0.947

The learned gate weights in Table 18 and Figure 9 provide direct evidence that the model does not fuse the three branches identically for all samples. In Table 19, r denotes

Pearson correlation and ρ denotes Spearman rank correlation. Straight samples mainly use the 5-step branch and almost ignore the 7-step branch. Turning and speed-changing samples place the largest average weight on the 7-step branch, with speed-changing samples showing the strongest shift away from the 5-step branch. This pattern is consistent with Table 17, where turning and speed-changing windows often contain sustained high-change runs, and it does not impose a rule that all turning samples require the shortest scale. The gate is therefore described as a learned adaptive fusion mechanism that maps maneuver intensity to temporal scale rather than a hand-coded maneuver classifier.

Figure 10 gives a broad multi-baseline view. To make the qualitative behavior more auditable, Figure 11 shows representative maneuver cases, Figure 12 gives direct Fixed-MS-LSTM and MGMS-RLSTM examples, and Figure 13 summarizes failure modes at the 30-min horizon.

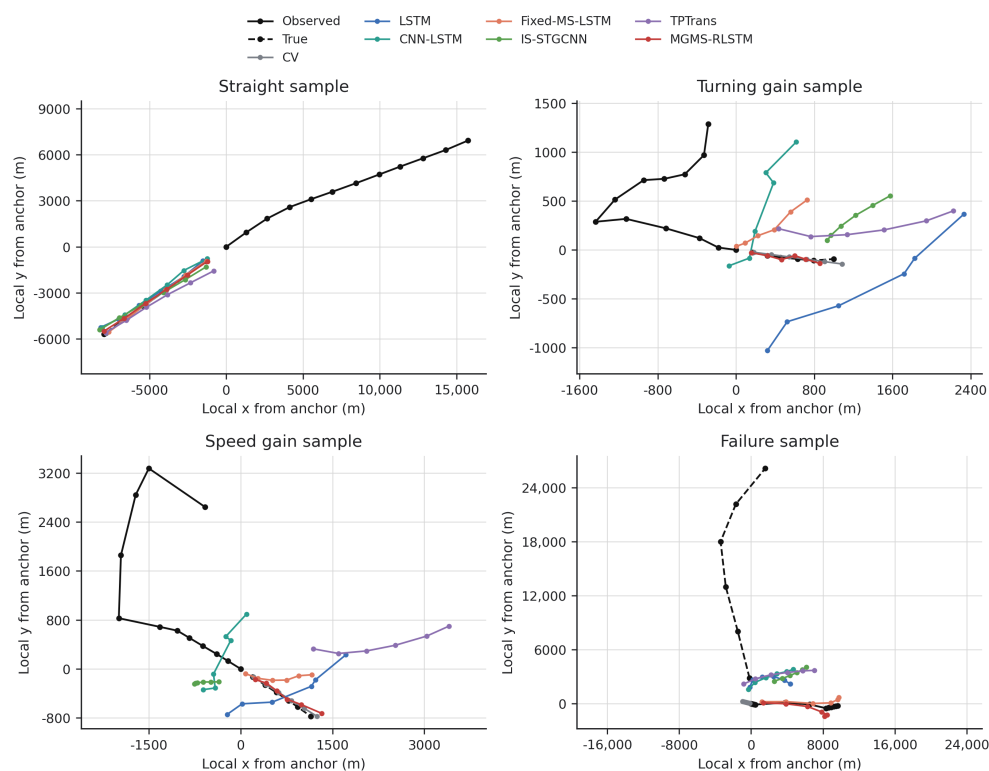


Figure 10. Representative 30-min prediction cases.

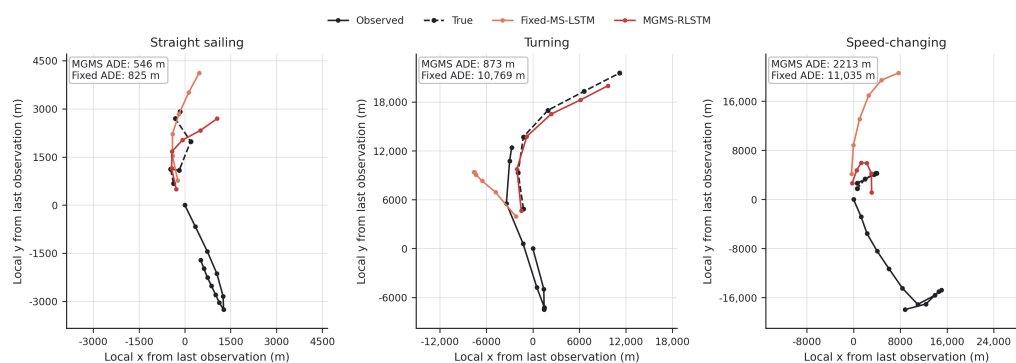


Figure 11. Representative cases by maneuver group.

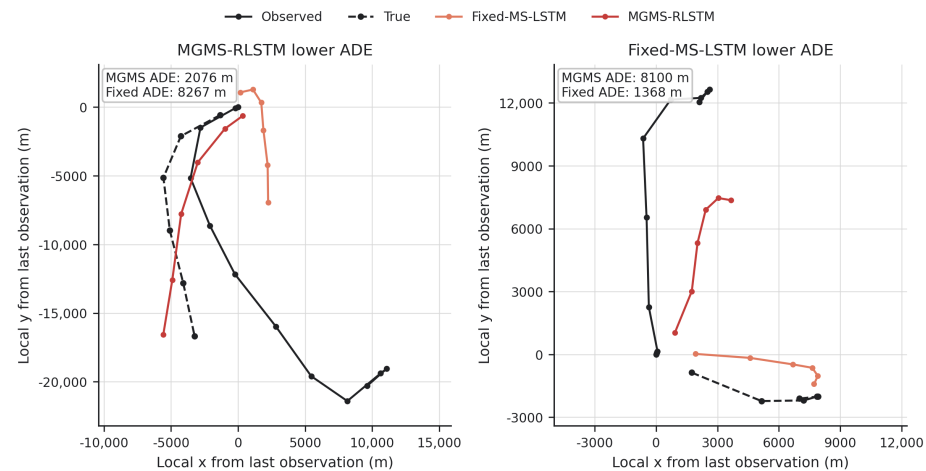


Figure 12. Fixed-MS-LSTM and MGMS-RLSTM case comparison.

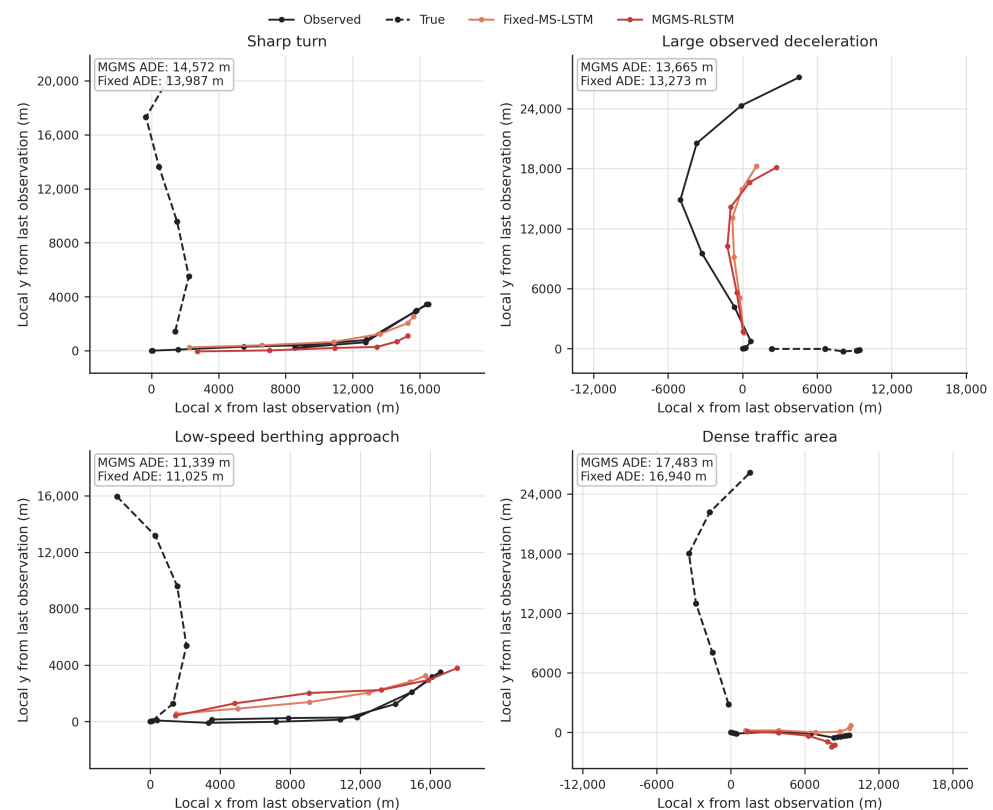


Figure 13. Failure-case taxonomy at the 30-min horizon.

The targeted cases help explain both the value and the limitation of the framework. In relatively regular motion, residual models follow the local displacement trend more closely than absolute-coordinate LSTM and simple extrapolation. In turning and speed-changing cases, multi-scale residual models can track the future path when the recent maneuver remains informative. However, Figure 12 illustrates both patterns: MGMS-RLSTM can correct a fixed-fusion miss, but fixed fusion can also be better when the learned gate emphasizes an unsuitable temporal scale. Figure 13 further shows that all AIS-only models can miss abrupt future route changes, boundary cases near berthing transitions, or motion changes in dense traffic. This supports the cautious interpretation used above: the framework improves short-term local motion modeling, but it does not replace route-intention, chart-constrained, or interaction-aware prediction.

Figures 14–16 further localize the large-error cases and the spatial differences between fixed and maneuver-guided fusion. Positive values in Figure 15 favor MGMS-RLSTM, whereas negative values favor Fixed-MS-LSTM.

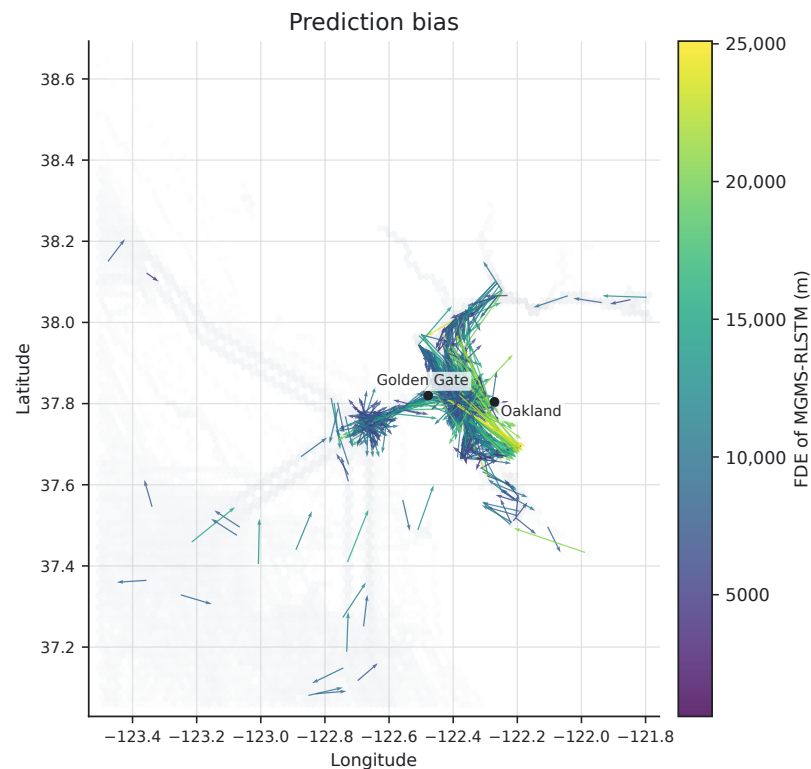


Figure 14. MGMS-RLSTM endpoint bias for top 10% ADE samples.

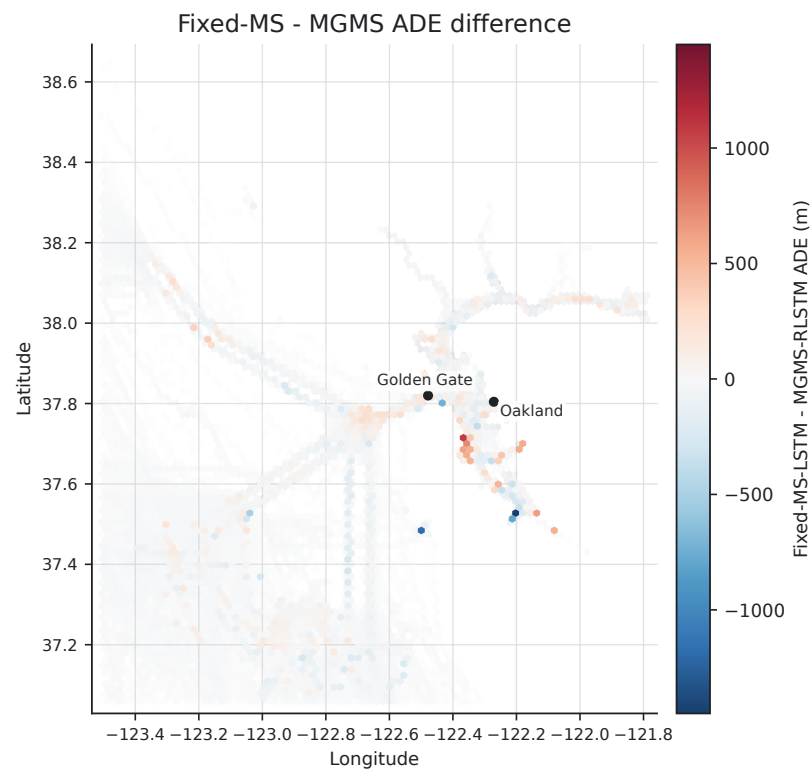


Figure 15. Spatial ADE difference between Fixed-MS-LSTM and MGMS-RLSTM.

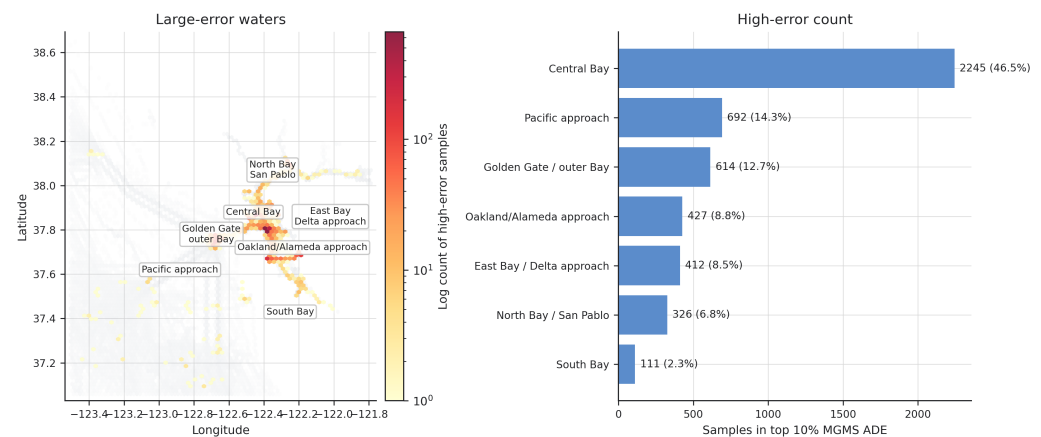


Figure 16. Water-area concentration of large MGMS-RLSTM errors.

The spatial diagnostics show that large errors are not uniformly distributed. Among the top 10% MGMS-RLSTM ADE samples, the largest share occurs in the Central Bay region (46.5%), followed by the Pacific approach (14.3%), the Golden Gate/outer Bay (12.7%), the Oakland/Alameda approach (8.8%), the East Bay/Delta approach (8.5%), the North Bay/San Pablo region (6.8%), and the South Bay (2.3%). These areas are characterized by route choices, turning behavior, merging traffic, or dense vessel movement. The Fixed-MS-LSTM minus MGMS-RLSTM ADE difference map shows a mixed spatial pattern, indicating that neither fusion rule has a uniform spatial advantage.

6.8. Application-Oriented Discussion

From an application perspective, the proposed framework is best understood as a rolling AIS-only predictor for near-term traffic monitoring. Given the most recent AIS window, it provides 10–30 min position estimates that can support preliminary CPA/TCPA-based screening of vessel pairs requiring closer attention. The model does not infer destination, planned route, or operator intention, and it is not designed to replace VTS decision-making. Instead, it provides a motion-level prediction component that can be updated with each new AIS window and used together with traffic rules, waterway knowledge, and operator judgment.

The encounter-oriented evaluation shows that both Fixed-MS-LSTM and MGMS-RLSTM support more accurate CPA/TCPA-based high-risk screening under the evaluated benchmark compared with classical, standard neural, and representative recent neural baselines. MGMS-RLSTM provides adaptive scale interpretation and competitive CPA/TCPA performance, while Fixed-MS-LSTM remains a strong operational counterpart when short-horizon stability or screening F1 is prioritized. In the three-seed main geometric evaluation, MGMS-RLSTM's ADE advantage appears at 20 and 30 min, while Fixed-MS-LSTM remains slightly stronger at 10 min. In the three-seed encounter evaluation, MGMS-RLSTM is best at 20 min F1, while Fixed-MS-LSTM is best at 10 and 30 min F1. These results indicate that the residual multi-scale framework can support rolling motion prediction and attention allocation, while longer horizons still require caution because route choice, traffic interaction, and waterway constraints become more influential.

6.9. Interpretability and Generalizability

The maneuver gate is driven by eight window-level statistics that summarize course variation, speed variation, and mean speed. Its three scale weights can therefore be analyzed across straight, turning, and speed-changing groups, as in Tables 18 and 19. Because a separate model is fitted for each forecast horizon, the learned mappings provide sample- and horizon-dependent temporal-scale selection. The results also retain Fixed-MS-LSTM as

a strong counterpart: fixed fusion provides stable short-horizon performance and achieves the highest encounter-screening F1 at 10 and 30 min, while the gated model offers sample-dependent scale analysis and the highest F1 at 20 min.

The experiments use San Francisco Bay and adjacent approach waters, and cross-port validation is needed for broader deployment. The 5 min interval is suitable for the present 10–30 min traffic-level prediction setting; finer sampling may be needed for second-level maneuver analysis. The present model is AIS-only and does not include waterway topology, destination intention, environmental conditions, or explicit multi-vessel interaction features. Future work may combine context-aware inputs with probabilistic outputs and speed-adaptive encounter candidate generation.

7. Conclusions

This study develops a maneuver-aware residual multi-scale LSTM framework for AIS-based short-term vessel trajectory prediction in port-approach waters. The largest observed gains come from residual displacement prediction and maneuver-aware features, confirming that local motion formulation is more suitable than direct absolute-coordinate prediction for short-term AIS forecasting. In the three-seed main evaluation, Fixed-MS-LSTM provides the strongest 10 min accuracy, while MGMS-RLSTM achieves lower ADE at 20 and 30 min. Diagnostic ablations further show that the main improvements come from residual prediction and maneuver-aware features. Maneuver-guided fusion can be characterized as a horizon-dependent scale-selection mechanism that complements the fixed multi-scale counterpart. Encounter-oriented CPA/TCPA evaluation further shows that residual multi-scale prediction supports more accurate CPA/TCPA-based high-risk screening under the evaluated benchmark. The residual-fusion diagnostics improve the interpretability of the correction coefficient, but their accuracy gains are inconsistent across horizons. Future work may incorporate spatial and cross-port validation, waterway topology, destination or route intention, traffic interaction, encounter-type-aware risk evaluation, stronger sequence and graph baselines, probabilistic forecasting, and separate status or intention models for prolonged anchorage, mooring, and berthing dwell states.

Author Contributions: Conceptualization, X.L. and J.Z.; methodology, X.L.; software, X.L. and W.B.; validation, X.L. and X.Z.; formal analysis, X.L., X.Z., W.B., Y.D. and J.Z.; data curation, X.L.; visualization, X.L. and Y.D.; writing—original draft preparation, X.L.; writing—review and editing, X.L., X.Z., W.B., Y.D. and J.Z.; project administration, J.Z.; funding acquisition, J.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported in part by the National Natural Science Foundation of China under Grant 52471379, in part by the Fund of Hubei Key Laboratory of Inland Shipping Technology under Grant NHHY2024005 and in part by the National Undergraduate Training Program for Innovation and Entrepreneurship for College Students under Grant G20260102.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The raw AIS data used in this study are publicly available from Marine Cadastre AccessAIS records distributed by the NOAA Office for Coastal Management. The data were requested from Marine Cadastre for the period from 5 February 2023 to 7 April 2023, within the spatial range of longitude -123.5066719499534 to -121.78728043447558 and latitude 37.059114568915376 to 38.76630184134356 . Processed datasets, MMSI split lists, configuration files, and scripts are available from the corresponding author upon reasonable request.

Conflicts of Interest: The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

References

- Hochreiter, S.; Schmidhuber, J. Long short-term memory. *Neural Comput.* **1997**, *9*, 1735–1780. [[CrossRef](#)] [[PubMed](#)]
- Capobianco, S.; Millefiori, L.M.; Forti, N.; Braca, P.; Willett, P. Deep learning methods for vessel trajectory prediction based on recurrent neural networks. *IEEE Trans. Aerosp. Electron. Syst.* **2021**, *57*, 4329–4346. [[CrossRef](#)]
- Park, J.; Jeong, J.; Park, Y. Ship trajectory prediction based on Bi-LSTM using spectral-clustered AIS data. *J. Mar. Sci. Eng.* **2021**, *9*, 1037. [[CrossRef](#)]
- Wang, X.; Xiao, Y. A deep learning model for ship trajectory prediction using Automatic Identification System (AIS) data. *Information* **2023**, *14*, 212. [[CrossRef](#)]
- Gao, D.; Zhu, Y.; Zhang, J.; He, Y.; Yan, K.; Yan, B. A novel MP-LSTM method for ship trajectory prediction based on AIS data. *Ocean Eng.* **2021**, *228*, 108956. [[CrossRef](#)]
- Yin, J.; Yu, Z.; Wu, H. Ship trajectory prediction based on LSTM model with multi-scale convolution and attention mechanism. *Ocean Eng.* **2025**, *338*, 122055. [[CrossRef](#)]
- International Telecommunication Union. *Technical Characteristics for an Automatic Identification System Using Time Division Multiple Access in the VHF Maritime Mobile Frequency Band*; Recommendation ITU-R M.1371-5; ITU: Geneva, Switzerland, 2014.
- Liu, J.; Shi, G.; Zhu, K. Vessel trajectory prediction model based on AIS sensor data and adaptive chaos differential evolution support vector regression (ACDE-SVR). *Appl. Sci.* **2019**, *9*, 2983. [[CrossRef](#)]
- Suo, Y.; Chen, W.; Claramunt, C.; Yang, S. A ship trajectory prediction framework based on a recurrent neural network. *Sensors* **2020**, *20*, 5133. [[CrossRef](#)] [[PubMed](#)]
- Chen, X.; Wei, C.; Zhou, G.; Wu, H.; Wang, Z.; Biancardo, S.A. Automatic Identification System (AIS) data supported ship trajectory prediction and analysis via a deep learning model. *J. Mar. Sci. Eng.* **2022**, *10*, 1314. [[CrossRef](#)]
- Bao, K.; Bi, J.; Gao, M.; Sun, Y.; Zhang, X.; Zhang, W. An improved ship trajectory prediction based on AIS data using MHA-BiGRU. *J. Mar. Sci. Eng.* **2022**, *10*, 804. [[CrossRef](#)]
- Zhang, Y.; Chen, H.; Zhou, F.; Wang, K.; Zhang, X. TRFM-LS: Transformer-based deep learning method for vessel trajectory prediction. *J. Mar. Sci. Eng.* **2023**, *11*, 880. [[CrossRef](#)]
- Feng, H.; Cao, G.; Xu, H.; Ge, S.S. IS-STGCNN: An improved social spatial-temporal graph convolutional neural network for ship trajectory prediction. *Ocean Eng.* **2022**, *266*, 112960. [[CrossRef](#)]
- Wang, W.; Xiong, W.; Ouyang, X.; Chen, L. TPTrans: Vessel trajectory prediction model based on Transformer using AIS data. *ISPRS Int. J. Geo-Inf.* **2024**, *13*, 400. [[CrossRef](#)]
- Xue, H.; Wang, S.; Xia, M.; Guo, S. G-Trans: A hierarchical approach to vessel trajectory prediction with GRU-based transformer. *Ocean Eng.* **2024**, *300*, 117431. [[CrossRef](#)]
- Liu, R.W.; Zheng, W.; Liang, M. Spatio-temporal multi-graph transformer network for joint prediction of multiple vessel trajectories. *Eng. Appl. Artif. Intell.* **2024**, *129*, 107625. [[CrossRef](#)]
- Yuan, H.; Liu, K.; Wu, X.; Yu, Y.; Xin, X.; Wang, W. GATransformer: A vessel trajectory prediction method based on attention algorithm in complex navigable waters. *Ocean Eng.* **2025**, *326*, 120902. [[CrossRef](#)]
- Gong, J.; Li, H.; Jiao, H.; Yang, Z. Uncertainty-aware ship trajectory prediction via Spatio-Temporal Graph Transformer. *Transp. Res. Part E Logist. Transp. Rev.* **2025**, *203*, 104315. [[CrossRef](#)]
- Esmaili, Y.; Motieyan, H.; Dezvareh, R. Context-aware vessel trajectory prediction using FlashAttention and maritime spatial features. *Ocean Eng.* **2026**, *344*, 123711. [[CrossRef](#)]
- Bi, J.; Cheng, H.; Zhang, W.; Bao, K.; Wang, P. Artificial intelligence in ship trajectory prediction. *J. Mar. Sci. Eng.* **2024**, *12*, 769. [[CrossRef](#)]
- Volkova, T.A.; Balykina, Y.E.; Bepalov, A. Predicting ship trajectory based on neural networks using AIS data. *J. Mar. Sci. Eng.* **2021**, *9*, 254. [[CrossRef](#)]
- He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 770–778.
- Kingma, D.P.; Ba, J. Adam: A method for stochastic optimization. In Proceedings of the International Conference on Learning Representations, San Diego, CA, USA, 7–9 May 2015.

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 MDPI and/or the editor(s). MDPI 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.