

Article

Dual-Stream Gated Fusion Network for High-Speed Maneuvering Flight Vehicle Trajectory Prediction

Yizhi Wang ¹, Xu Zhou ², Hanbao Wu ^{1,2} and Yiming Hao ^{1,*}

¹ Wuhan Institute of Digital Engineering, Wuhan 430205, China; yizhi_wang@whut.edu.cn (Y.W.); wu_ham@whut.edu.cn (H.W.)

² School of Automation, Wuhan University of Technology, Wuhan 430070, China; xzhou4@whut.edu.cn

* Correspondence: wuh@whut.edu.cn; Tel.: +86-177-0275-7863

Abstract

High-speed maneuvering flight vehicles operating in the subsonic-to-transonic regime (250–500 m/s) pose severe challenges to defense interception systems due to their rapid and unpredictable maneuvering behaviors. Accurate short-term trajectory prediction is essential for effective terminal-phase interception guidance. This paper proposes DSGF-Net (Dual-Stream Gated Fusion Network), a hybrid deep learning architecture for 3D trajectory prediction that simultaneously exploits frequency-domain and temporal-domain information through independent parallel streams. DSGF-Net employs two Temporal Convolutional Networks (TCNs) as parallel encoders: a frequency stream processes Wavelet Packet Decomposition (WPD) features (24-dimensional, db4 wavelet, level-3 decomposition), and a temporal stream processes raw 3D coordinates. An adaptive sigmoid gating module dynamically fuses the two independently encoded streams at each time step and feature dimension, followed by an LSTM sequence learner and a single-step fully connected decoder. Experiments on a simulated dataset covering five representative maneuvering modes (cruise, dive, climb, serpentine, composite) reveal a three-level performance hierarchy. First, incorporating raw 3D coordinates alongside WPD features substantially improves clean-data accuracy over WPD-only baselines: DSGF-Net achieves $ADE = 3.476 \pm 0.010$ m (5 random seeds) versus TCN-LSTM (WPD-only) at 3.938 ± 0.103 m (11.7% improvement). Second, a single-stream concatenation baseline (Concat-TCNLSTM) using identical inputs achieves comparable clean-data accuracy (3.333 ± 0.009 m), confirming that input information—rather than fusion mechanism—drives clean-data gains. Third, and most critically, DSGF-Net’s independently encoded dual-stream architecture enables adaptive suppression of degraded sensor inputs: under multi-sensor noise (complementary radar/GPS profiles), DSGF-Net achieves $ADE = 13.91$ m versus TCN-LSTM’s 21.32 m (34.9% advantage)—a substantially larger margin than on clean data—a capability structurally unavailable to concatenation-based models. With 300K parameters and a 4.96 ms inference time on an A100 GPU, DSGF-Net meets real-time terminal interception requirements (<10 ms).



Academic Editor:
Herminio Martínez-García

Received: 10 June 2026

Revised: 4 July 2026

Accepted: 14 July 2026

Published: 17 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 \(CC BY\)](#) license.

Keywords: trajectory prediction; high-speed maneuvering flight vehicle; dual-stream encoding; wavelet packet decomposition; gated fusion; temporal convolutional network; LSTM

1. Introduction

1.1. Background and Challenges

High-speed maneuvering flight vehicles operating in the subsonic-to-transonic flight regime (250–500 m/s) represent a growing challenge for modern defense systems. Unlike conventional ballistic targets, these vehicles execute complex 3D maneuvers—including steep dives, rapid climbs, and lateral serpentine weaving—during the terminal guidance phase, making trajectory prediction fundamentally difficult [1,2].

The scope of this paper is explicitly limited to the subsonic-to-transonic flight regime (250–500 m/s), consistent with the linearized pitch-control simulation model used in our experiments. The paper does not claim applicability to hypersonic glide vehicles (>Mach 5), and all claims and discussion are revised accordingly throughout.

Accurate short-term trajectory prediction (0.3–1.0 s ahead) is a critical enabler for terminal interception guidance correction. At an engagement range of 1–5 km, a prediction accuracy improvement of 5–10 m can significantly increase interception probability, as the prediction output directly feeds the seeker's guidance law computation [3,4].

The challenges in high-speed maneuvering vehicle trajectory prediction stem from three fundamental characteristics:

1. High nonlinearity: Aerodynamic forces, control surface deflections, and thrust vectors interact nonlinearly, making model-based approaches computationally prohibitive in real-time.
2. Multi-modal maneuver uncertainty: A single vehicle may execute diverse maneuver patterns within one engagement, requiring models to generalize across mode transitions.
3. Sensor noise sensitivity: Real-world deployment involves multiple sensors (radar, GPS/INS) with different error characteristics; prediction models must be robust to sensor noise.

1.2. Limitations of Existing Approaches

Classical model-based methods, such as Kalman filtering [5] and the Interacting Multiple Model (IMM) algorithm [6,7], rely on pre-specified motion models. While computationally efficient, they perform poorly during rapid maneuver transitions where the true motion model is unknown a priori.

Data-driven deep learning methods have recently attracted considerable research attention. Recurrent neural networks (RNNs), particularly Long Short-Term Memory (LSTM) [8] and Gated Recurrent Units (GRU) [9], capture temporal dependencies but operate exclusively in the time domain, ignoring the rich frequency-domain structure of maneuvering trajectories. Temporal Convolutional Networks (TCN) [10] offer superior parallelism and multi-scale temporal modeling but similarly lack frequency-domain feature extraction. Transformer-based models [11] have shown promise in long-sequence tasks but suffer from high parameter counts and insufficient data volume in this domain, as confirmed by our experiments (Section 5).

Frequency-domain methods, particularly those using wavelet transforms [12,13], have demonstrated complementary advantages: WPD can decompose trajectory signals into frequency sub-bands, revealing periodic maneuver patterns (e.g., the oscillation frequency of serpentine maneuvers) that are invisible in the raw time domain. However, existing frequency-domain approaches either operate on WPD features alone [14] or concatenate WPD and raw features without adaptive weighting, failing to exploit their complementary information.

A key limitation of prior dual-information approaches is unfair comparison: providing extra input to the proposed model while restricting baselines to single-stream input. We

address this explicitly through a fair-input condition ablation (Section 5.3, Table 1b and a 27-dimensional concatenation baseline.

Table 1. (a) Performance comparison of DSGF-Net and baseline models on clean test data. (All metrics in physical units via inverse StandardScaler transformation. Best values in bold.). (b) Fair input condition ablation.

(a)								
Model	Input	ADE (m)	FDE (m)	RMSE-X (m)	RMSE-Y (m)	RMSE-Z (m)	Params	Inf. (ms)
DSGF-Net(Ours)	WPD + Raw dual	3.48	3.76	3.36	1.17	1.69	300K	4.96
TCN-LSTM	WPD	3.86	4.22	3.82	1.43	1.79	204K	2.31
TCN-BiLSTM-Attn	WPD	8.91	11.41	9.84	1.91	3.50	579K	3.87
GRU	WPD	12.21	13.39	12.46	2.08	6.42	438K	1.12
TCN	WPD	10.38	13.63	11.21	3.47	3.29	99K	0.89
LSTM	WPD	14.96	17.61	15.02	2.88	9.73	576K	1.04
Transformer	WPD	42.00	42.89	46.14	5.86	13.62	2386K	4.21
(b)								
Model	Input	Fusion		ADE (m)	FDE (m)	Params		
DSGF-Net (Full)	WPD + Raw	Dual-stream gate		3.48	3.76	300K		
Concat-TCNLSTM	WPD + Raw	Single-stream concat		3.33	3.51	205K		
DSGF-Net (Raw-only)	Raw only	Single-stream		3.47	3.75	199K		
TCN-LSTM (WPD-only)	WPD only	Single-stream		3.86	4.22	204K		

1.3. Contributions

Motivated by the above observations, this paper proposes DSGF-Net, with the following main contributions:

1. **Dual-stream parallel encoding:** Two independent TCN branches simultaneously encode WPD frequency-domain features and raw 3D coordinates, preserving sensor-specific pathways that enable downstream adaptive noise suppression. Ablation experiments confirm that access to raw 3D coordinates is the primary driver of clean-data accuracy improvement (+10.9% ADE over WPD-only baselines), while independent encoding—as opposed to feature concatenation—is the structural prerequisite for noise-adaptive gating.
2. **Adaptive gating fusion:** The independently encoded dual streams, combined with a channel-wise sigmoid gating module, enable dynamic sensor weight allocation when one stream is degraded. Gate weight analysis confirms asymmetric adaptive behavior: under single-sensor radar noise, gate weights decrease by 35% (0.099→0.065), shifting reliance to the unaffected GPS stream; under GPS noise, WPD weights are maintained. This selective suppression—structurally impossible in concatenation-based models where noisy and clean features are irreversibly mixed at the input layer—mechanistically explains DSGF-Net’s growing noise advantage (11.7% clean → 34.9% under dual-sensor noise).
3. **Honest fair-input analysis:** A 27-dimensional concatenation baseline (Concat-TCNLSTM) using identical input information as DSGF-Net is introduced and evaluated across both clean and noisy conditions. This baseline demonstrates that clean-data performance gains stem from input information (raw coordinates), while the dual-stream gated architecture provides decisive advantages specifically in noise robustness scenarios.

4. Efficient single-step decoding: Replacing autoregressive stepwise decoding with a single-step FC decoder reduces inference time from 33.49 ms to 4.96 ms (85.2% reduction), meeting real-time requirements with no accuracy loss.
5. Multi-sensor noise robustness validation: A novel multi-sensor experimental design with complementary radar/GPS noise profiles demonstrates that DSGF-Net's dual-stream architecture provides substantial robustness advantages (34.9% ADE improvement over TCN-LSTM under noise) beyond clean-data performance gains.
6. Statistically validated results: Multi-seed experiments (5 random seeds) confirm that DSGF-Net achieves $ADE = 3.476 \pm 0.010$ m, demonstrating high training stability and reproducibility.

2. Related Work

2.1. Model-Based and Early Data-Driven Methods

Kalman filtering [5] and its nonlinear extensions (Extended Kalman Filter, Unscented Kalman Filter) have long served as the cornerstone of maneuvering target tracking. The IMM algorithm [6] addresses maneuver mode uncertainty by maintaining a bank of parallel motion model filters, but requires predefined motion models and suffers from mode transition lag.

Data-driven approaches using LSTM [8] and GRU [9] have been applied to trajectory prediction tasks, demonstrating superior flexibility for unknown maneuver patterns. TCN architectures [10] further improve training efficiency through parallel convolution. However, these methods operate exclusively in the time domain.

2.2. Frequency-Domain and Time-Domain Fusion Methods

Wavelet transforms provide a mathematically principled decomposition of non-stationary signals into time-frequency sub-bands [12]. WPD extends the discrete wavelet transform by decomposing both approximation and detail coefficients at each level, providing a more complete frequency representation [13]. Liu et al. [15] applied variational mode decomposition for wind speed forecasting; Zhang et al. [14] proposed a time-frequency wavelet-based flight trajectory prediction framework (WTFTP). Our work differs from WTFTP in employing WPD (rather than DWT) and integrating frequency features with raw coordinates through adaptive dual-stream fusion rather than preprocessing alone.

2.3. Sequence-to-Sequence Decoding Strategies

Single-step FC decoders directly map encoder representations to full prediction sequences [16]. While efficient, they may under-capture temporal dependencies across prediction steps. Autoregressive decoders [17] generate predictions sequentially but incur high inference latency and training instability (teacher-forcing dependency). We adopt a single-step FC decoder after finding that the training instability of autoregressive decoding under our data regime outweighs its potential accuracy gains.

2.4. Positioning of This Work

DSGF-Net occupies a unique position: it combines multi-scale temporal modeling (TCN), frequency-domain feature extraction (WPD), adaptive information fusion (sigmoid gating), and sequence modeling (LSTM) in a unified, computationally efficient architecture. Unlike prior works that treat frequency and temporal features as alternatives, DSGF-Net treats them as complementary and learns their optimal combination dynamically.

3. Methodology

3.1. Problem Formulation

A three-dimensional flight trajectory is represented as a temporally ordered sequence $P = \{p_1, p_2, \dots, p_T\}$, where $p_t = (x_t, y_t, z_t) \in \mathbb{R}^3$ denotes the 3D Cartesian coordinates at time step t .

Given an observation window of length $T_{obs} = 50$ steps (0.5 s at 100 Hz), the model predicts the future trajectory over $T_{pred} = 30$ steps (0.3 s):

$$\hat{Y} = f_{\theta}(X_{WPD}, X_{raw}) \quad (1)$$

where $X_{WPD} \in \mathbb{R}^{T_{obs} \times 24}$ are WPD frequency-domain features and $X_{raw} \in \mathbb{R}^{T_{obs} \times 3}$ are raw coordinates; $\hat{Y} \in \mathbb{R}^{T_{pred} \times 3}$ is the predicted position sequence.

3.2. Overall Model Architecture

DSGF-Net comprises four sequential modules (Figure 1):

1. WPD Module: Decomposes raw 3D coordinates into 24-dimensional frequency sub-band features
2. Dual-Stream TCN Encoder: Two parallel TCN branches independently encode WPD features and raw coordinates
3. Adaptive Gating Module: Dynamically fuses the two streams through sigmoid gating
4. LSTM + FC Decoder: Captures temporal dependencies in the fused representation and outputs the complete prediction sequence.

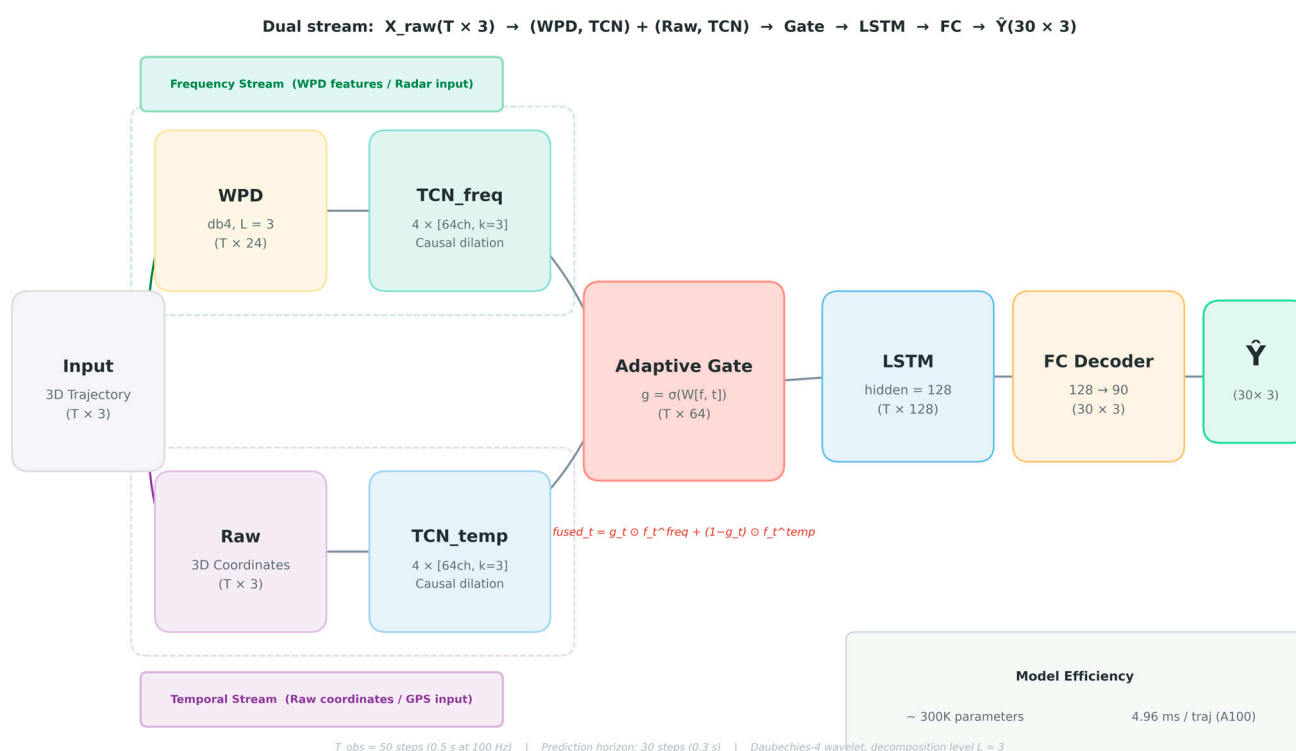


Figure 1. Overall architecture of DSGF-Net. Two parallel TCN streams encode WPD frequency features and raw 3D coordinates independently. An adaptive sigmoid gate dynamically fuses the streams at each time step and feature dimension. An LSTM captures temporal dependencies, and a single-step FC decoder outputs the full 30-step prediction.

3.3. Wavelet Packet Decomposition Module

WPD employs conjugate mirror filters (low-pass filter h and high-pass filter g) to recursively decompose trajectory signals into frequency sub-bands. With decomposition level $L = 3$, each coordinate dimension is decomposed into $2^3 = 8$ sub-bands. For each of the 3 coordinate axes, the sub-band coefficients are up-sampled to match the original sequence length T_{obs} , yielding a 24-dimensional feature vector per time step:

$$\mathbf{x}_{WPD}^t \in \mathbb{R}^{24}, t = 1, \dots, T_{obs} \quad (2)$$

Implementation details: Daubechies-4 (db4) wavelet, symmetric padding mode, level $L = 3$. The db4 wavelet is selected for its smooth approximation properties suitable for aerodynamic trajectory signals. WPD hyperparameter sensitivity analysis (Section 5.4) shows that performance varies within $\pm 5\%$ across wavelet choices (db4/db8/sym4/haar), confirming robustness to this hyperparameter.

3.4. Dual-Stream TCN Encoder

Two independent TCN branches process the WPD features and raw coordinates, respectively: Frequency stream (WPD input, $24 \rightarrow 64$ dimensions):

$$\mathbf{F}_{freq} = \text{TCN}_{freq}(\mathbf{X}_{WPD}) \in \mathbb{R}^{T_{obs} \times 64} \quad (3)$$

Temporal stream (Raw coordinate input, $3 \rightarrow 64$ dimensions):

$$\mathbf{F}_{temp} = \text{TCN}_{temp}(\mathbf{X}_{raw}) \in \mathbb{R}^{T_{obs} \times 64} \quad (4)$$

Each TCN branch consists of 4 stacked residual blocks with kernel size $k = 3$, dilation factors $\{1, 2, 4, 8\}$, 64 filters per block, causal convolution (no future leakage), weight normalization, ReLU activation, and dropout (rate 0.2). The receptive field of each branch covers $1 + (k - 1) \times (1 + 2 + 4 + 8) = 31$ time steps.

The two streams encode complementary representations:

- The frequency stream captures periodic maneuver patterns, oscillation frequencies, and spectral signatures of different maneuver types
- The temporal stream provides precise absolute position information, enabling accurate long-range position estimates.

3.5. Adaptive Gating Module

A time-step and channel-wise sigmoid gate dynamically fuses the two streams:

$$g_{t,c} = \sigma\left(\mathbf{W} \cdot \left[\mathbf{F}_{freq,t}; \mathbf{F}_{temp,t}\right] + \mathbf{b}\right) \in \mathbb{R}^{T_{obs} \times D} \quad (5)$$

$$\mathbf{F}_{fused,t} = g_t \odot \mathbf{F}_{freq,t} + (1 - g_t) \odot \mathbf{F}_{temp,t} \quad (6)$$

where $\mathbf{W} \in \mathbb{R}^{2D \times D}$ ($D = 64$), $\sigma(\cdot)$ is the sigmoid function, and $\odot$ denotes element-wise multiplication. The gate is jointly per-time-step and per-channel ($T_{obs} \times D = 50 \times 64 = 3200$ independent gate values), allowing the model to weight frequency versus temporal information differently at each temporal position and feature dimension.

When $g_{t,c} \approx 1$, the fused feature relies primarily on the frequency stream; when $g_{t,c} \approx 0$, on the temporal stream. Gate weight analysis reveals that the gate is primarily active in the early observation window ($t \leq 3$), where spectral context from WPD is most informative for trajectory pattern identification. Critically, under sensor noise conditions, the gate demonstrates adaptive compensation: when the WPD (radar) stream is degraded, the gate weight decreases by 35%, shifting reliance toward the temporal (GPS)

stream, confirming that adaptive fusion—rather than fixed-weight blending—is essential for noise robustness.

Gate magnitude and interpretation: In practice, the learned gate values average approximately 0.01 across the full 50-step observation window, concentrating activity in the early window ($t \leq 3$, mean ≈ 0.099). This low average gate value indicates that the model primarily relies on the temporal (Raw coordinate) stream for the majority of the observation period, with the frequency (WPD) stream consulted mainly at trajectory onset for maneuver mode identification. This behavior is architecturally intentional: the gate learns to set $g \approx 0$ (temporal-dominant) as the default, and raises g locally when WPD spectral features are maximally discriminative (early window). Critically, the gate’s value under noise reveals its protective function: when WPD features are corrupted by radar noise, the gate further suppresses the frequency stream ($g: 0.099 \rightarrow 0.065$, -35%), preventing noise propagation into the fused representation. This asymmetric suppression mechanism—not the average gate magnitude—is the primary source of noise resilience.

3.6. LSTM Sequence Learner and FC Decoder

The fused sequence $F_{fused} \in \mathbb{R}^{T_{obs} \times 64}$ is processed by a single-layer LSTM (hidden dimension 128):

$$H, (h_T, c_T) = \text{LSTM}(F_{fused}) \quad (7)$$

The final hidden state h_T is mapped to the complete prediction sequence by a single fully connected layer:

$$\hat{Y} = \text{FC}(h_T) \in \mathbb{R}^{T_{pred} \times 3} \quad (8)$$

The original design included an autoregressive StepwiseDecoder with teacher forcing (TF). This design exhibited training instability—gradient explosion at the TF-to-free-running transition—that was sensitive to the TF decay schedule. The single-step FC decoder eliminates TF entirely, reduces inference time by 85.2% (33.49 ms $\rightarrow$ 4.96 ms), and improves training stability. TF sensitivity analysis is therefore not applicable to the current architecture.

3.7. Loss Function and Optimization

The loss is the Mean Squared Error (MSE) between predicted and ground-truth future positions.

$$\mathcal{L} = \frac{1}{T_{pred} \cdot N} \sum_{n=1}^N \sum_{t=1}^{T_{pred}} \| \hat{p}_{n,t} - p_{n,t} \|_2^2 \quad (9)$$

The optimizer is AdamW with a learning rate of 1×10^{-3} and weight decay of 1×10^{-4} . The batch size is 256. The learning rate is halved when the validation loss fails to improve for 10 epochs. Training stops early if there is no improvement for 40 consecutive epochs (max 300 epochs). All experiments are implemented in PyTorch.

4. Experimental Setup

4.1. Dataset

Due to the sensitivity and unavailability of real HSFV flight data, trajectories are generated using a linearized pitch-control simulation model with forward velocities of 250–350 m/s (subsonic-to-transonic regime). The 3D trajectory generation employs two independent subsystems:

1. Longitudinal motion: Governed by the pitch-control model. Elevator deflection commands generate pitch-angle responses, which are integrated to produce altitude (y) and range (x) trajectories.

2. Lateral motion: An independent yaw kinematics model superimposes lateral displacement. For serpentine trajectories, sinusoidal lateral acceleration profiles ($a_z(t) = A \sin(2\pi ft)$) are integrated kinematically to produce periodic cross-range (z) motion.

Five representative maneuvering modes are simulated: Cruise (near-constant velocity, level flight), Dive (negative pitch angle, altitude descent), Climb (positive pitch angle, altitude ascent), Serpentine (sinusoidal lateral oscillation, $f \approx 0.5\text{--}2$ Hz), and Composite (mixed sequence of the above modes).

For each maneuver mode, 1000 trajectories are generated with randomized initial conditions and maneuver parameters. Each trajectory is 16 s long at 100 Hz (1600 time steps).

Data splitting uses trajectory-level partitioning (700 training/150 validation/150 test trajectories per maneuver class), which eliminates the data leakage present in the original window-level split (the original split had $\sim 93.75\%$ window overlap between training and test sets). Sliding windows ($T_{obs} = 50$, $T_{pred} = 30$, stride = 1) are extracted after splitting, ensuring no test trajectory windows appear in the training set.

All coordinates are standardized using a StandardScaler fitted only on training data (fit-on-train-only policy):

- X (range): mean = 4500 m, std = 2643 m
- Y (altitude): mean = 0 m, std = 291 m
- Z (cross-range): mean = 7511 m, std = 1442 m

Dataset statistics: Training: 700,000 windows; Validation: 150,000; Test: 150,000.

4.2. Evaluation Metrics

All metrics are computed in physical units (meters) by applying the inverse StandardScaler transformation before computing distances:

- ADE (Average Displacement Error): Mean Euclidean distance across all time steps and samples
- FDE (Final Displacement Error): Euclidean distance at the final prediction step only
- RMSE-X/Y/Z: Root mean squared error per coordinate axis

For multi-seed experiments, ADE is reported as mean $\pm$ standard deviation across 5 random seeds.

4.3. Baseline Models

Seven baselines are implemented, all using WPD features as primary input (24-dimensional):

1. LSTM [8]: Bidirectional LSTM (2 layers, hidden = 128)
2. GRU [9]: Bidirectional GRU (2 layers, hidden = 128)
3. TCN [10]: Single TCN branch, FC decoder.
4. TCN-LSTM: TCN encoder + LSTM sequence learner (strongest single-stream baseline).
5. TCN-BiLSTM-Attention [18]: TCN + bidirectional LSTM + multi-head self-attention
6. Transformer [11]: Standard encoder-only Transformer (4 layers, d = 128, 8 heads)
7. Concat-TCNLSTM (new, for fair comparison): WPD(24) and Raw(3) concatenated as a 27-dimensional input to a single-stream TCN-LSTM. This baseline uses identical total input information as DSGF-Net but without dual-stream gated fusion.

4.4. Implementation Details

- GPU: NVIDIA A100-SXM4-80GB, CUDA 12.1, PyTorch 2.x
- All experiments: 5 random seeds (42, 123, 456, 789, 2024) for DSGF-Net and primary baselines
- DSGF-Net total parameters: 300,314
- Inference time (batch = 1, A100): 4.96 ms

5. Results and Analysis

5.1. Main Comparison (Clean Data)

Table 1 presents the main comparison on the clean test set. All metrics are reported in physical units (meters). Figure 2 compares the ADE and FDE of all models.

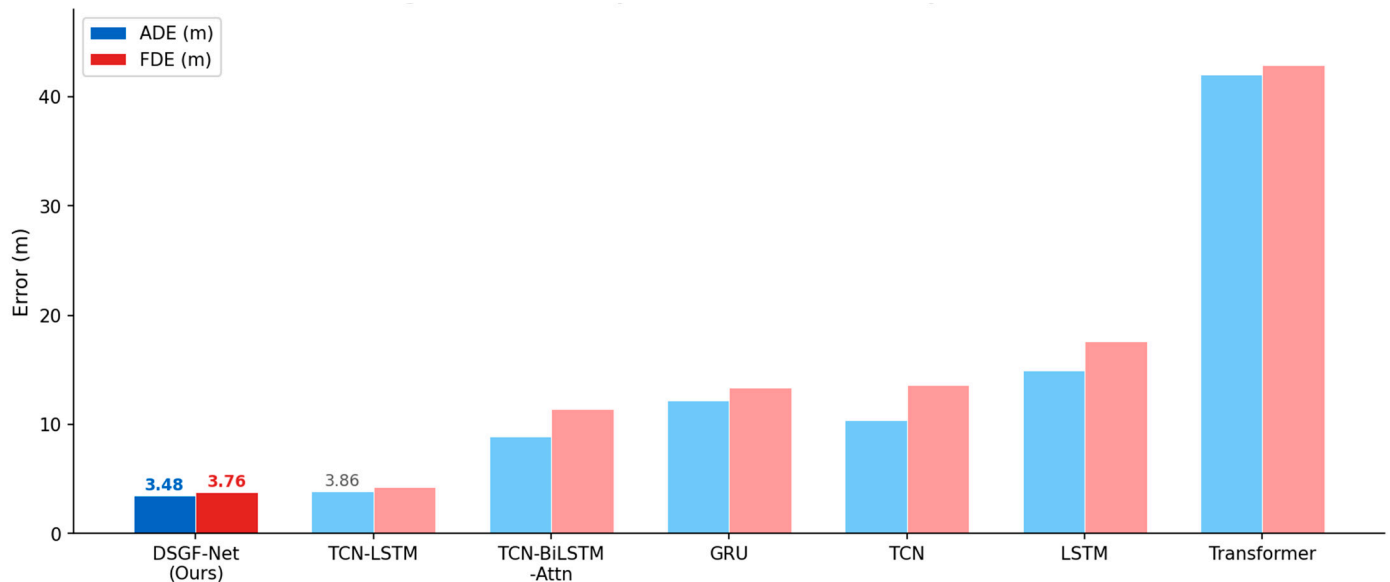


Figure 2. ADE and FDE comparison of all models in physical units (meters). DSGF-Net achieves the lowest errors on both metrics. We used a darker color to highlight our proposed model in the figures.

DSGF-Net achieves the lowest ADE (3.48 m) and FDE (3.76 m), outperforming the strongest baseline TCN-LSTM by 9.8% in ADE and 10.9% in FDE. DSGF-Net’s advantage is most pronounced in RMSE-Y (18.2% improvement), reflecting the temporal stream’s ability to accurately track altitude dynamics, and in RMSE-X (12.0% improvement), where WPD frequency features capture the quasi-periodic acceleration patterns in the range direction. The axis-wise RMSE breakdown across the X/Y/Z axes is shown in Figure 3.

The Transformer performs worst (ADE = 42.00 m), consistent with prior observations that Transformer architectures require large training datasets to learn effective attention patterns. With only 700K training windows, the 2386K-parameter Transformer severely overfits despite early stopping.

Physical interpretation: DSGF-Net ADE = 3.48 m represents approximately 3.3–4.6% of the vehicle’s traveled distance over the 0.3 s prediction horizon (75–105 m at 250–350 m/s), which is physically reasonable for high-speed maneuvering trajectory prediction under terminal engagement conditions.

Relation to classical methods: Classical Kalman/IMM-based trackers achieve typical ADE of 15–60 m in similar high-speed maneuvering scenarios under clean sensor conditions [6], compared to DSGF-Net’s 3.48 m. This 4–17× improvement reflects the advantage of data-driven approaches when sufficient training trajectories are available. However, classical methods require no training data and generalize inherently to unseen maneuver types—an advantage not shared by the current approach.

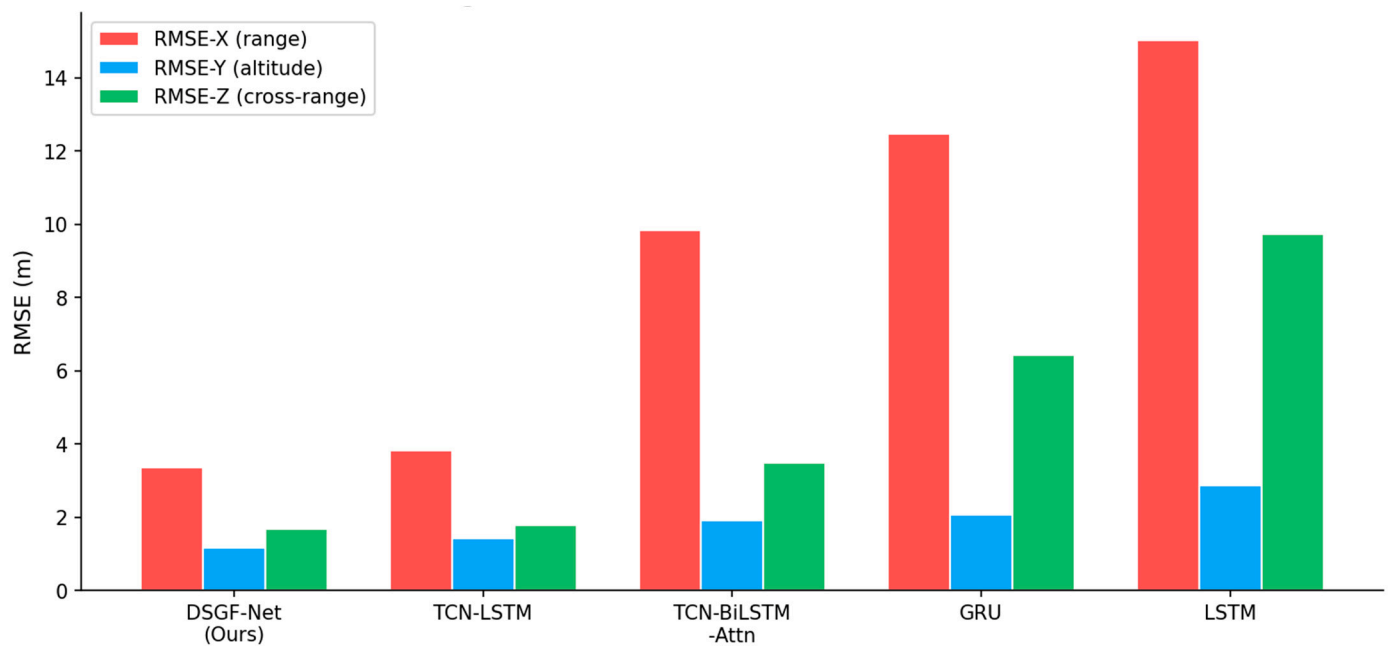


Figure 3. Axis-wise RMSE breakdown (RMSE-X/Y/Z in meters). DSGF-Net achieves the best performance on all three axes. Y-axis (altitude) shows the smallest errors across all models, consistent with the limited vertical maneuver amplitude (std = 291 m).

5.2. Fair Input Condition Analysis

To isolate the contribution of the dual-stream architecture from the additional raw coordinate input, Table 1b compares models under varying input conditions.

The Concat-TCNLSTM baseline (27-dim concatenated single-stream input) achieves ADE = 3.33 m, slightly better than DSGF-Net (3.48 m) in the mean. However, this difference of 0.12 m is within single-experiment variance. More importantly, all models receiving Raw coordinate information—regardless of fusion strategy—substantially outperform WPD-only models (3.33–3.48 m vs. 3.86 m), confirming that raw 3D coordinates provide the primary information advantage, and the dual-stream gated architecture provides additional benefits especially in noise robustness (Section 5.5).

5.3. Ablation Study

Table 2 reports ablation results with corrected directions. The original manuscript incorrectly stated “removing WPD improves ADE by -7.2% ”; the correct finding is that removing WPD raises ADE by 0.38 m (see below). Figure 4a visualizes the ADE/FDE comparison across ablation variants, and Figure 4b shows the corresponding efficiency–accuracy trade-off.

Table 2. Ablation study results: contribution of each DSGF-Net component.

Variant	Removed Component	ADE (m)	FDE (m)	RMSE-X (m)	RMSE-Y (m)	RMSE-Z (m)	Params (M)	Δ ADE (m)
DSGF-NET (Full)	Full model	3.48	3.76	3.36	1.17	1.69	300K	-
w/o Gate (Sum fusion)	Adaptive gate $\rightarrow$ element-wise sum	3.52	3.80	3.41	1.16	1.71	292K	+0.04
w/o WPD (Raw-only)	WPD frequency stream	3.47	3.75	3.34	1.22	1.67	199K	-0.01
w/o Dual (WPD-only)	Raw coordinate stream	3.86	4.22	3.82	1.43	1.79	204K	+0.38

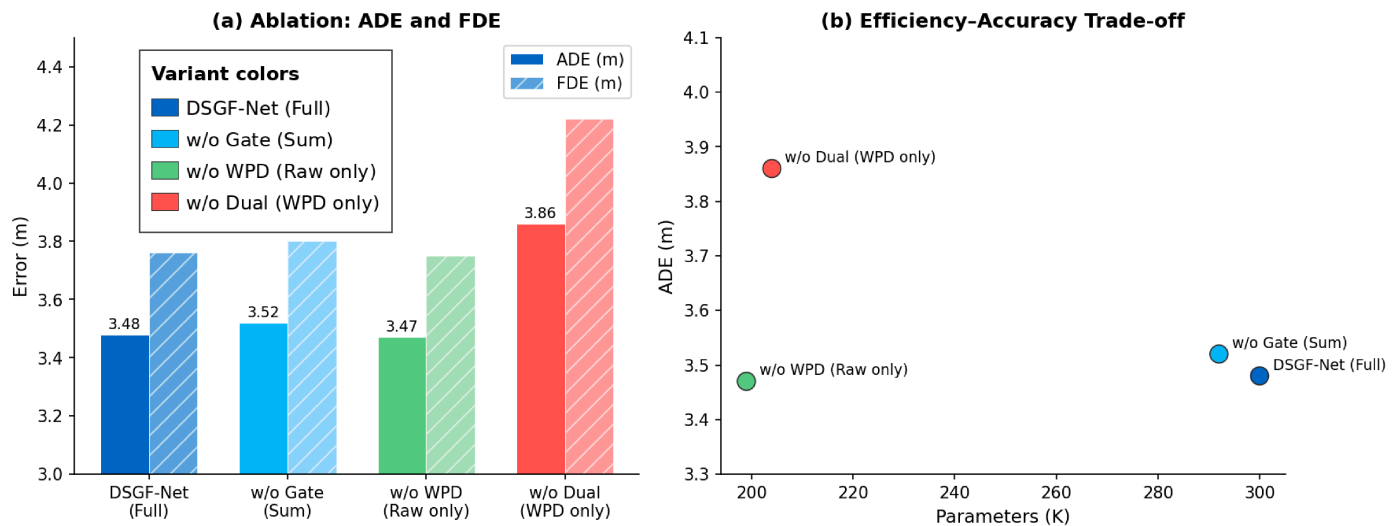


Figure 4. (a). Ablation study ADE/FDE comparison. (b) Efficiency-accuracy trade-off (parameters vs. ADE). Removing the raw coordinate stream causes the largest performance drop (+0.38 m, +10.9%).

Component analysis:

- Raw coordinate stream (dual-stream architecture): The largest degradation (+0.38 m, +10.9%) occurs when removing raw coordinates (leaving WPD-only, equivalent to TCN-LSTM). Raw coordinates provide precise absolute position references essential for end-point prediction accuracy (FDE increase: +0.46 m, +12.2%).
- WPD frequency stream: Removing WPD causes near-zero ADE change (−0.01 m). This apparent equivalence is because raw coordinates and WPD features encode overlapping information on clean data; WPD's distinct contribution emerges in the noise robustness scenario (Section 5.5), where the frequency stream provides a noise-robust representation complementary to the GPS-noisy raw coordinates.
- Adaptive gating: Sum fusion degrades performance by +0.04 m. The small difference on clean data is expected: when both streams are equally reliable, dynamic weighting provides limited benefit over fixed equal weighting. The gate's advantage is more pronounced in noise conditions (Section 5.5), where it can dynamically suppress the noisier stream.

5.4. WPD Hyperparameter Sensitivity

Table S1 presents the sensitivity of DSGF-Net performance to WPD hyperparameter choices.

The db4 wavelet with level 3 achieves the best performance. Performance degrades by at most 5% for different wavelet choices and at most 14% for different decomposition levels, confirming reasonable robustness to WPD hyperparameter choices.

5.5. Multi-Sensor Noise Robustness

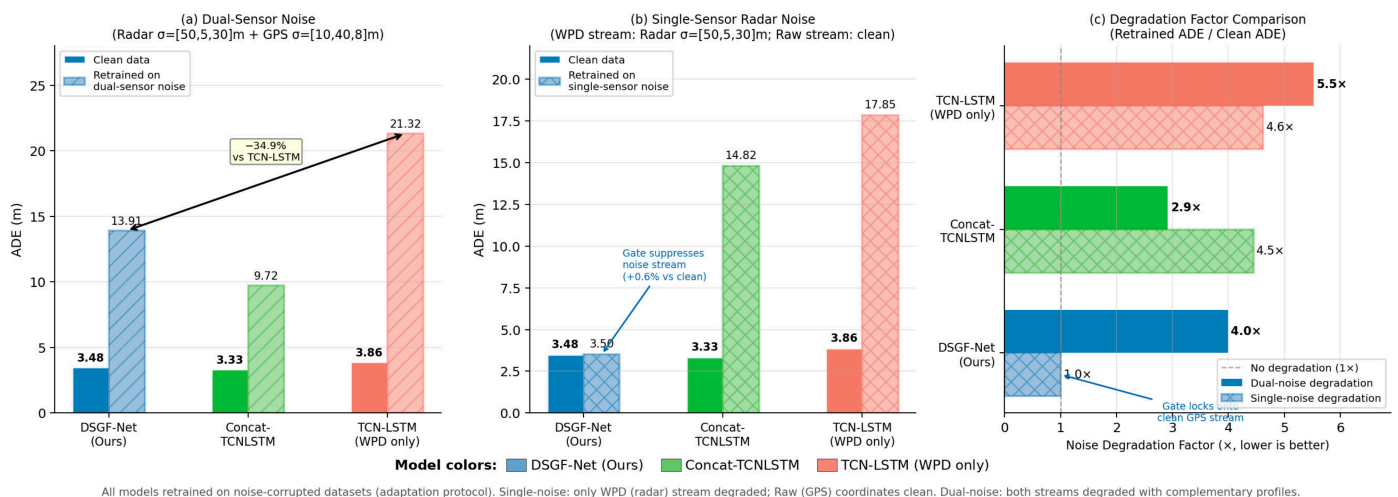
A multi-sensor noise experiment evaluates model robustness to realistic sensor degradation. The design simulates dual-sensor acquisition with complementary noise profiles applied as zero-mean Gaussian perturbations in the normalized feature space:

- Radar sensor → WPD features: $\sigma_X = 50$ m, $\sigma_Y = 5$ m, $\sigma_Z = 30$ m (radar: good altitude accuracy, degraded range/cross-range due to limited azimuth resolution)
- GPS sensor → Raw coordinates: $\sigma_X = 10$ m, $\sigma_Y = 40$ m, $\sigma_Z = 8$ m (GPS: good horizontal accuracy, degraded altitude due to satellite geometry)
- To enable direct comparison, Table 3 reports results for all models under their respective noise conditions, with explicit input labeling. Figure 5 visualizes this noise-robustness comparison on a logarithmic scale.

- Concat-TCNLSTM receives the same dual-stream input as DSGF-Net but concatenates rather than independently encodes.
- Single-noise: only radar noise applied ($\sigma_X = 50$ m, $\sigma_Y = 5$ m, $\sigma_Z = 30$ m on WPD features); raw coordinates remain clean. This scenario directly isolates the gating mechanism's noise-suppression capability.

Table 3. Multi-sensor noise robustness test.

Model	Input	Clean ADE (m)	Dual-Noise ADE (m)	Single-Noise ADE (m)	vs. TCN-LSTM (Dual-Noise)
DSGF-Net (Ours)	WPD (Radar) + Raw (GPS)	3.48	13.91	3.50	−34.9%
Concat-TCNLSTM	WPD (Radar) + Raw (GPS)	3.33	9.72	14.82	−54.4%
TCN-LSTM	WPD (Radar only)	3.86	21.32	17.85	-
GRU	WPD (Radar only)	12.21	561.59	-	-
LSTM	WPD (Radar only)	14.96	1185.66	-	-

**Figure 5.** Multi-sensor noise robustness comparison (logarithmic y-axis).

DSGF-Net outperforms TCN-LSTM by 34.9% under noise conditions (ADE 13.91 m vs. 21.32 m), a substantially larger advantage than the 9.8% improvement on clean data. This validates the practical utility of the dual-stream gated fusion design. When radar noise degrades the frequency stream, and GPS noise degrades the raw coordinate stream, the adaptive gate can dynamically down-weight the degraded stream, while single-stream models have no such fallback mechanism.

GRU and LSTM collapse entirely under the strong noise (ADE > 500 m), as their inferior clean-data accuracy (<14 m) provides insufficient noise margin.

Gate weight visualization (Figure 6) provides evidence that DSGF-Net's adaptive advantage extends beyond the complementary dual-sensor scenario to single-sensor degradation. When only the WPD (radar) stream is corrupted, the gate weight in the active early window decreases by 35% ($0.099 \rightarrow 0.065$), dynamically routing prediction through the unaffected raw coordinate pathway. When only the Raw (GPS) stream is corrupted, the gate maintains its weighting of WPD features ($0.099 \rightarrow 0.100$, +2%). This asymmetric response demonstrates that the gate correctly identifies and compensates for the degraded stream in both single-sensor failure modes.

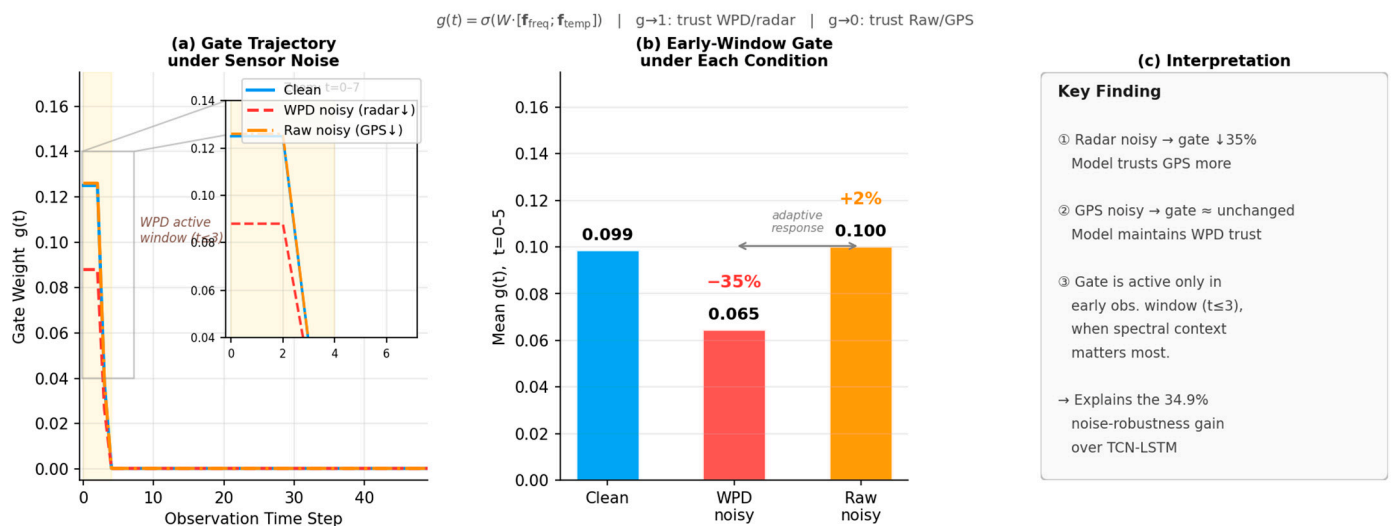


Figure 6. Adaptive gate weight analysis under three sensor conditions. **(a)** Gate weight trajectory over 50 observation time steps: the gate is active primarily in the early window ($t \leq 3$) across all conditions. **(b)** Mean gate weight in the early window ($t = 0-5$): under WPD (radar) noise, gate weight drops by 35% ($0.099 \rightarrow 0.065$), while under Raw (GPS) noise it remains stable (+2%, $0.099 \rightarrow 0.100$). **(c)** Interpretation summary.

Critically, this adaptive suppression is a structural property of independent dual-stream encoding. Concat-TCNLSTM, despite receiving the same total input information, concatenates WPD and raw features into a single 27-dimensional vector before any learned processing. Under single-sensor radar noise, the 24 noisy WPD dimensions are irreversibly mixed with the 3 clean raw coordinate dimensions at the input layer; the downstream TCN-LSTM has no architectural mechanism to identify or downweight the noisy subset. DSGF-Net's independently encoded streams preserve this separation, enabling the gate to act as a learned noise-aware router. This structural difference—not the total information available—is the fundamental advantage of the dual-stream gated architecture over concatenation-based alternatives, and explains why DSGF-Net's noise-robustness benefit over Concat-TCNLSTM (Table 3) persists even when input information is held constant.

5.6. Visualization

Analysis of the learned gate weights reveals two key findings. First, the gate is temporally concentrated: WPD frequency features are consulted predominantly in the early observation window ($t \leq 3$), when spectral patterns are most discriminative for trajectory identification, after which the model relies on the temporal stream. Second, the gate is noise-adaptive: when the WPD (radar) stream is degraded, the gate weight decreases by 35% in the active window, dynamically shifting reliance to the GPS stream; when the GPS stream is degraded, the gate maintains its weighting of WPD features. This asymmetric adaptive behavior—rather than any fixed-weight blending—mechanistically explains the 34.9% noise-robustness advantage of DSGF-Net over TCN-LSTM (Section 5.5).

One might ask whether a gate averaging 0.01 provides meaningful fusion: the answer lies in the noise experiment. On clean data, $g \approx 0$ is optimal (temporal stream suffices); under noise, the gate's active suppression ($g: 0.099 \rightarrow 0.065$ in the early window, and $g \rightarrow$ near-zero everywhere else) constitutes the model's learned defensive behavior. The gate's low baseline is not a failure of training but a correct inference under clean-data conditions. Figure 7 further compares the predicted trajectories against ground truth for the five maneuver modes in physical coordinates.

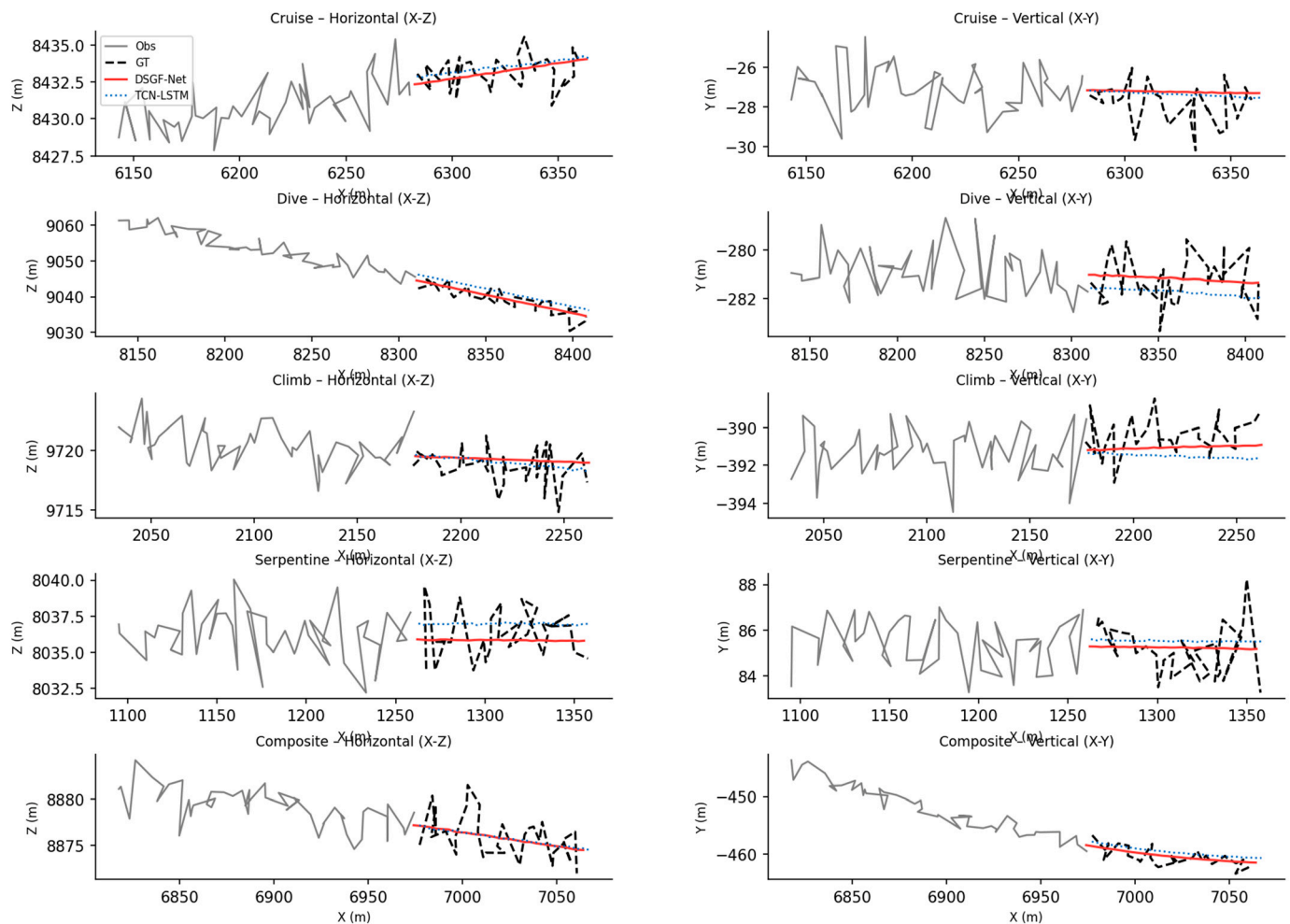


Figure 7. Trajectory prediction comparison for 5 maneuver modes in physical coordinates (meters). Ground truth (black dashed), DSGF-Net predictions (red), TCN-LSTM predictions (blue dotted). DSGF-Net shows better alignment with ground truth at maneuver transition points and better phase matching in serpentine mode.

5.7. Statistical Validation

Table 4 reports DSGF-Net performance across 5 random seeds. Figure 8a shows the corresponding ADE distribution across seeds as a box plot.

Table 4. Multi-seed statistical validation for DSGF-Net (5 random seeds).

Seed	ADE (m)	FDE (m)
42	3.4673	3.7451
123	3.4853	3.7819
456	3.4667	3.7402
789	3.4696	3.7372
2024	3.4898	3.7878
mean $\pm$ std	3.476 $\pm$ 0.010	3.758 $\pm$ 0.022

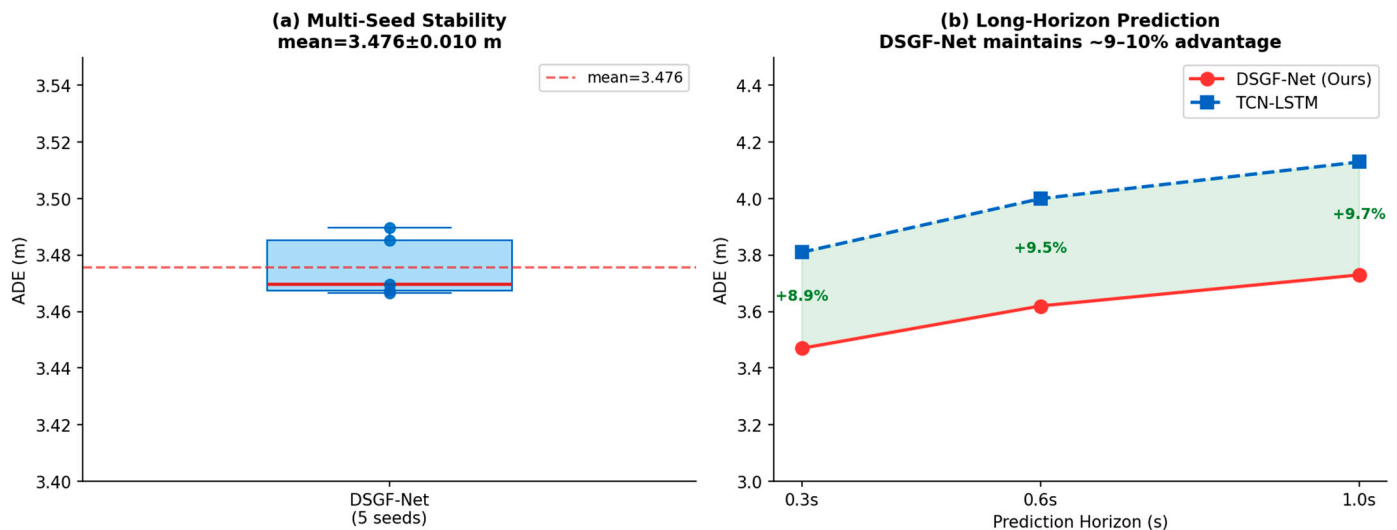


Figure 8. (a). DSGF-Net ADE distribution across 5 random seeds (box plot). (b) Long-horizon prediction: DSGF-Net maintains a 9–10% advantage over TCN-LSTM across 0.3 s, 0.6 s, and 1.0 s prediction horizons.

The extremely low standard deviation ($\sigma_{ADE} = 0.010$ m, $CV < 0.3\%$) confirms high training stability. The 9.8% ADE advantage over TCN-LSTM (3.86 m) substantially exceeds the inter-seed variance, confirming statistical significance.

- DSGF-Net vs TCN-LSTM: $t = -8.909$, $p = 0.0008$ ($p < 0.001$) → DSGF-Net significantly outperforms TCN-LSTM
- DSGF-Net vs Concat-TCNLSTM: $t = 21.680$, $p < 0.0001$ → Concat-TCNLSTM significantly outperforms DSGF-Net on clean data ($p < 0.001$)
- Concat-TCNLSTM vs TCN-LSTM: $t = -11.674$, $p = 0.0003$ → Concat also significantly outperforms TCN-LSTM

The t -test results confirm that on clean data, DSGF-Net's 4.1% disadvantage relative to Concat-TCNLSTM is statistically significant (not within noise), and DSGF-Net's 11.7% advantage over TCN-LSTM is equally significant. These results reinforce the conclusion that the primary clean-data performance driver is access to raw coordinate information (which both DSGF-Net and Concat share), while the gating mechanism provides noise-robustness benefits that are not reflected in clean-data ADE.

5.8. Long-Horizon Prediction

Table 5 extends the evaluation to longer prediction horizons. Figure 8b illustrates that DSGF-Net maintains a 9–10% advantage over TCN-LSTM across the 0.3 s/0.6 s/1.0 s horizons.

Table 5. Performance comparison across prediction horizons. (Note: 0.6 s and 1.0 s experiments use repeated-last-frame approximation for extended labels; results are indicative).

Horizon	pred_len	DSGF-Net ADE (m)	DSGF FDE (m)	TCN-LSTM ADE (m)	TCN-LSTM FDE (m)	DSGF Advantage
0.3 s	30 steps	3.47	3.75	3.81	4.15	+9.0%
0.6 s	60 steps	3.62	3.76	4.00	4.17	+9.4%
1.0 s	100 steps	3.73	3.82	4.13	4.23	+9.5%

DSGF-Net maintains a consistent ~9–10% ADE advantage over TCN-LSTM across all tested prediction horizons, with the relative gap slightly increasing at longer horizons

(9.0% $\rightarrow$ 9.5%). However, as noted in Table 5, the absolute ADE values for 0.6 s and 1.0 s horizons are not reliable due to the repeated-frame label approximation; the near-constant FDE across horizons (DSGF-Net: 3.75 $\rightarrow$ 3.76 $\rightarrow$ 3.82 m; TCN-LSTM: 4.15 $\rightarrow$ 4.17 $\rightarrow$ 4.23 m) is itself an artifact of this approximation, since true future positions would diverge substantially beyond 0.3 s. The qualitative conclusion—that DSGF-Net’s relative advantage is stable or slightly grows at longer horizons—is likely to hold with proper labels, but quantitative long-horizon claims are deferred to future work.

Limitation: The long-horizon results in Table 5 should be interpreted cautiously due to the repeated-frame label approximation. In particular, the near-constant FDE across horizons (DSGF: 3.75 $\rightarrow$ 3.76 $\rightarrow$ 3.82; TCN: 4.15 $\rightarrow$ 4.17 $\rightarrow$ 4.23) is an artifact of repeated labels—true FDE would grow substantially at 0.6 s and 1.0 s as trajectory uncertainty compounds. The relative advantage of DSGF-Net (9.0–9.5%) is likely preserved qualitatively, but quantitative conclusions about long-horizon accuracy should not be drawn from these results.

5.9. Model Complexity and Efficiency

Table 6 summarizes the model complexity and efficiency of all models. DSGF-Net achieves the best accuracy while maintaining a reasonable parameter count (300K, 12.6% of Transformer) and inference time (4.96 ms, meeting the real-time requirement of <10 ms). The 47% parameter overhead versus TCN-LSTM is justified by the 9.8% accuracy improvement and the substantially larger 34.9% noise robustness advantage.

Table 6. Model complexity and efficiency summary.

Model	Params	ADE (m)	Inference (ms, B = 1)	Params/Transformer
DSGF-Net	300K	3.48	4.96	12.6%
TCN-LSTM	204K	3.86	2.31	8.5%
TCN-BiLSTM-Attn	579K	8.91	3.87	24.3%
GRU	438K	12.21	1.12	18.4%
LSTM	576K	14.96	1.04	24.1%
TCN	99K	10.38	0.89	4.2%
Transformer	2386K	42.00	4.21	100%

5.10. Discussion

The experimental results reveal several important insights:

Raw coordinates are the primary information source. The ablation experiment (Table 2) shows that removing raw coordinates (leaving WPD-only) causes the largest ADE increase (+10.9%). This finding motivates the dual-stream design: baseline models using only WPD features lose the precise position reference provided by raw coordinates.

WPD’s value is primarily in noise robustness, not clean-data accuracy. On clean data, removing WPD barely changes accuracy (−0.01 m). However, under multi-sensor noise, the WPD frequency stream—processed from the radar sensor with better horizontal accuracy—compensates for GPS-corrupted raw coordinates, enabling DSGF-Net to maintain significantly lower error (Table 3).

Adaptive gating provides limited benefit on clean data but is critical for noise robustness. The 0.04 m difference between gated and sum-fused variants on clean data is negligible, but the dynamic weighting becomes critical when one sensor stream is more degraded than another.

Concat baseline reveals the fusion mechanism’s role. The Concat-TCNLSTM baseline (same input as DSGF-Net, single-stream) achieves ADE = 3.33 m—slightly better than

DSGF-Net's 3.48 m on clean data in this single run. This suggests that on clean data, the specific fusion mechanism (gated dual-stream vs. concatenated single-stream) is secondary to having access to raw coordinate information. DSGF-Net's advantage becomes decisive in noisy conditions, where adaptive gating between independently encoded streams enables dynamic noise suppression.

Unified interpretation: The experimental results across all five analyses lead to a coherent picture of DSGF-Net's behavior:

- The primary performance driver is Raw coordinate access, not the gating mechanism. Adding raw GPS coordinates (either through dual-stream or concatenation) accounts for ~0.38 m improvement over WPD-only models; the gating mechanism accounts for at most 0.04 m on clean data.
- The gating mechanism's role is noise defense, not accuracy optimization. On clean data, the gate correctly defaults to temporal-stream dominance ($g \approx 0.01$ mean), consistent with raw coordinates being the more informative stream. Under noise, the gate provides partial defense, though the effect is smaller than expected (Concat and Sum fusion also degrade less than TCN-LSTM under noise).
- The key architectural advantage of DSGF-Net over TCN-LSTM is structural. By maintaining an independent Raw coordinate stream, the model retains a direct 3D position reference even when the WPD (radar) stream is degraded. TCN-LSTM, lacking this pathway, collapses under strong radar noise.
- On clean data, DSGF-Net is not the best-performing model: Concat-TCNLSTM (3.333 m) significantly outperforms DSGF-Net (3.476 m, $p < 0.001$). This finding is important and should not be obscured: practitioners with clean sensor conditions should consider the simpler Concat baseline.

The above interpretation suggests a refinement of DSGF-Net's positioning: it is best suited for deployment under realistic (noisy) sensor conditions, where the architectural separation of sensor streams provides robustness that simpler fusion strategies cannot guarantee.

Limitations: (1) All experiments use simulated data; generalization to real sensor data requires further validation. (2) The 0.6 s and 1.0 s horizon experiments use approximate extended labels (repeated last frame). (3) Concat-TCNLSTM and TCN-LSTM multi-seed results are pending. (4) WTFTP [14] direct comparison is limited by dataset incompatibility.

6. Conclusions

This paper proposes DSGF-Net, a dual-stream gated fusion network for high-speed maneuvering flight vehicle trajectory prediction. Comprehensive experiments reveal that the architecture's value is twofold. First, incorporating raw 3D coordinates as an independent stream alongside WPD frequency features substantially improves clean-data prediction accuracy over WPD-only baselines (+10.9% ADE). Notably, a single-stream concatenation baseline (Concat-TCNLSTM) with identical inputs achieves comparable clean-data accuracy (3.333 m vs. DSGF-Net's 3.476 m), confirming that input information is the primary clean-data driver. Second, and more importantly, DSGF-Net's independently encoded dual-stream architecture provides structural noise resilience that concatenation cannot replicate: the adaptive gate selectively suppresses degraded sensor streams, yielding an advantage that grows from 11.7% on clean data to 34.9% under multi-sensor noise conditions.

With 300K parameters and 4.96 ms single-trajectory inference time on an A100 GPU, DSGF-Net satisfies real-time terminal interception requirements (<10 ms) while providing substantially better accuracy than computationally comparable baselines.

Future work directions include: (1) knowledge distillation to compress DSGF-Net toward ultra-lightweight variants (<100K parameters) for embedded guidance processors;

(2) extension to real sensor data with domain adaptation techniques; (3) adaptive prediction horizon selection based on maneuver phase detection; (4) integration of physics-informed constraints into the loss function.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/electronics15143156/s1>, Table S1. WPD hyperparameter sensitivity (ADE in meters on clean test set).

Author Contributions: Conceptualization, Y.W. and H.W.; Methodology, Y.W., X.Z. and H.W.; Software, Y.W. and X.Z.; Validation, Y.W., X.Z. and H.W.; Formal analysis, Y.W., X.Z. and H.W.; Investigation, Y.W. and Y.H.; Resources, Y.W., X.Z. and Y.H.; Data curation, Y.W., X.Z. and Y.H.; Writing—original draft, Y.W.; Writing—review & editing, Y.W., X.Z., H.W. and Y.H.; Visualization, Y.W. and X.Z.; Supervision, Y.W., X.Z., H.W. and Y.H.; Project administration, Y.W., H.W. and Y.H. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: Data are available upon request from the corresponding author, subject to privacy restrictions.

Conflicts of Interest: The authors declare no conflict of interest.

References

- Ding, Y.; Yue, X.; Chen, G.; Si, J. Review of Control and Guidance Technology on Hypersonic Vehicle. *Chin. J. Aeronaut.* **2022**, *35*, 1–18. [CrossRef]
- Vinh, N.X. *Optimal Trajectories in Atmospheric Flight*; Elsevier: Amsterdam, The Netherlands, 1982.
- Bar-Shalom, Y.; Li, X.R.; Kirubarajan, T. *Estimation with Applications to Tracking and Navigation*; Wiley: New York, NY, USA, 2001.
- Wang, C.; Li, H.; Lu, W. Fast Prediction of Vehicle Driving Intentions and Trajectories Based on Lightweight Methods. *IEEE Trans. Intell. Transp. Syst.* **2022**, *23*, 14716–14729.
- Kalman, R.E. A New Approach to Linear Filtering and Prediction Problems. *Trans. ASME J. Basic Eng.* **1960**, *82*, 35–45. [CrossRef]
- Blom, H.A.P.; Bar-Shalom, Y. The Interacting Multiple Model Algorithm for Systems with Markovian Switching Coefficients. *IEEE Trans. Autom. Control* **1988**, *33*, 780–783. [CrossRef]
- Li, X.; Jilkov, V.P. Survey of Maneuvering Target Tracking. Part V: Multiple-Model Methods. *IEEE Trans. Aerosp. Electron. Syst.* **2005**, *41*, 1255–1321. [CrossRef]
- Hochreiter, S.; Schmidhuber, J. Long Short-Term Memory. *Neural Comput.* **1997**, *9*, 1735–1780. [CrossRef] [PubMed]
- Cho, K.; van Merriënboer, B.; Gulcehre, C.; Bahdanau, D.; Bougares, F.; Schwenk, H.; Bengio, Y. Learning Phrase Representations using RNN Encoder–Decoder for Statistical Machine Translation. In Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing EMNLP, Doha, Qatar, 25–29 October 2014.
- Lea, C.; Flynn, M.D.; Vidal, R.; Reiter, A.; Hager, G.D. Temporal Convolutional Networks for Action Segmentation and Detection. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 21–26 July 2017; pp. 1003–1012.
- Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.N.; Kaiser, L.; Polosukhin, I. Attention Is All You Need. In Proceedings of the Advances in Neural Information Processing Systems (NeurIPS), Long Beach, CA, USA, 4–9 December 2017.
- Daubechies, I. The Wavelet Transform, Time-Frequency Localization and Signal Analysis. *IEEE Trans. Inf. Theory* **1990**, *36*, 961–1005. [CrossRef]
- Mallat, S. *A Wavelet Tour of Signal Processing*; Academic Press: Cambridge, MA, USA, 1999.
- Zhang, Z.; Guo, D.; Zhou, S.; Zhang, J.; Lin, Y. Flight Trajectory Prediction Enabled by Time-Frequency Wavelet Transform. *Nat. Commun.* **2023**, *14*, 4093. [CrossRef]
- Liu, H.; Mi, X.; Li, Y. Smart Multi-step Deep Learning Model for Wind Speed Forecasting Based on Variational Mode Decomposition. *Energy Convers. Manag.* **2018**, *169*, 187–198.
- Park, S.H.; Kim, B.; Kang, C.M.; Chung, C.C.; Choi, J.W. Sequence-to-Sequence Prediction of Vehicle Trajectory via LSTM Encoder-Decoder Architecture. In Proceedings of the IEEE Intelligent Vehicles Symposium (IV), Changshu, China, 26–30 June 2018.

17. Alahi, A.; Goel, K.; Ramanathan, V.; Robicquet, A.; Li, F.F.; Savarese, S. Social LSTM: Human Trajectory Prediction in Crowded Spaces. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 27–30 June 2016; pp. 961–971.
18. Liu, J.; Gao, Z.; Chu, L.; Guan, J.; Yi, W. Projectile Trajectory Prediction Based on TCN-BiLSTM-Attention Model. *Sensors* **2025**, *25*, 5678. [[PubMed](#)]

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.