Vessel Trajectory Prediction Using Robust AIS Preprocessing and Dual-Self-Attention GRU

Marilena Sinni, Dimitris M. Kyriazanos

National Centre for Scientific Research "Demokritos", Institute of Informatics and Telecommunications (IIT)

Workshop on AI-driven Data Engineering and Reusability for Earth and Space Sciences (DARES'25), co-located with the 28th

European Conference on Artificial Intelligence (ECAI 2025), Bologna, Italy, October 25, 2025.

Received funding under project FaRADAI, Grant No 101103386

Motivation

- Precise forecasting of vessel trajectories
- Why? For safe and efficient maritime traffic management
- Difficulties
  - Maritime data suffer from inherent errors
  - Continuous forecasting for hours if challenging

Goal & Challenges

Prediction Goal

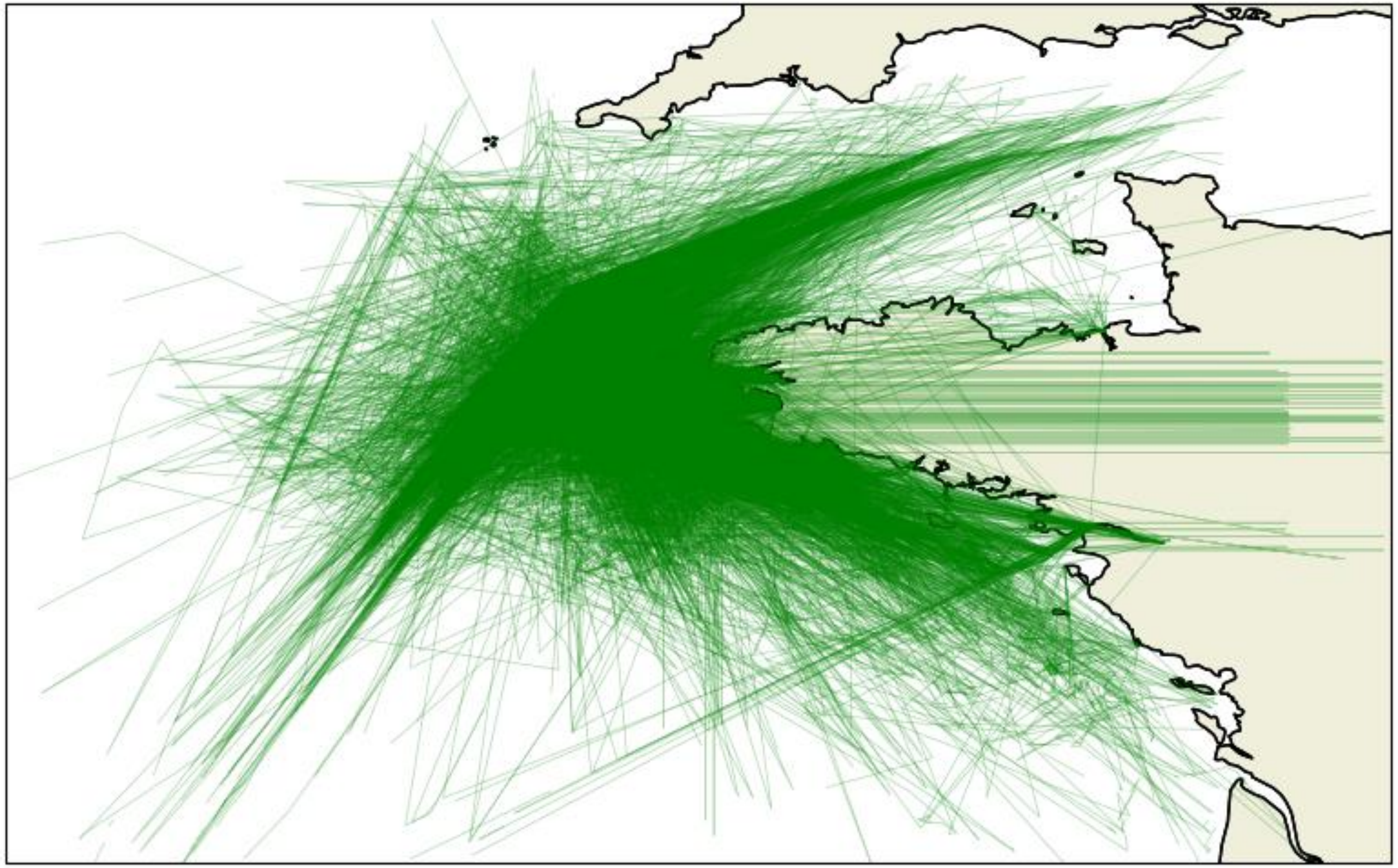
- Predict latitude & longitude, Speed over Ground (SOG) and Course over Ground (COG)
- Prediction over a multi-step time horizon spanning several hours.

Key Challenge

- Bridge the gap between complex Automatic Identification System (AIS) data preprocessing and accurate future trajectory forecasting

Raw AIS Data Inaccuracies

- **Irregular sampling & gaps:** messages arrive unevenly (seconds → hours).
- **Noise and anomalies:** includes position spikes on land, duplicated or missing reports, and physically impossible speed values.
- **Heterogeneous vessel behavior:** different ship types and manoeuvres.



Our Contributions

- Comprehensive AIS preprocessing pipeline ensuring noise-free, consistent trajectories that enhance prediction performance.

Our Contributions

- Comprehensive AIS preprocessing pipeline ensuring noise-free, consistent trajectories that enhance prediction performance.
- A robust sequence forecasting model, combining a single-layer GRU with a dual-head self-attention mechanism.

Our Contributions

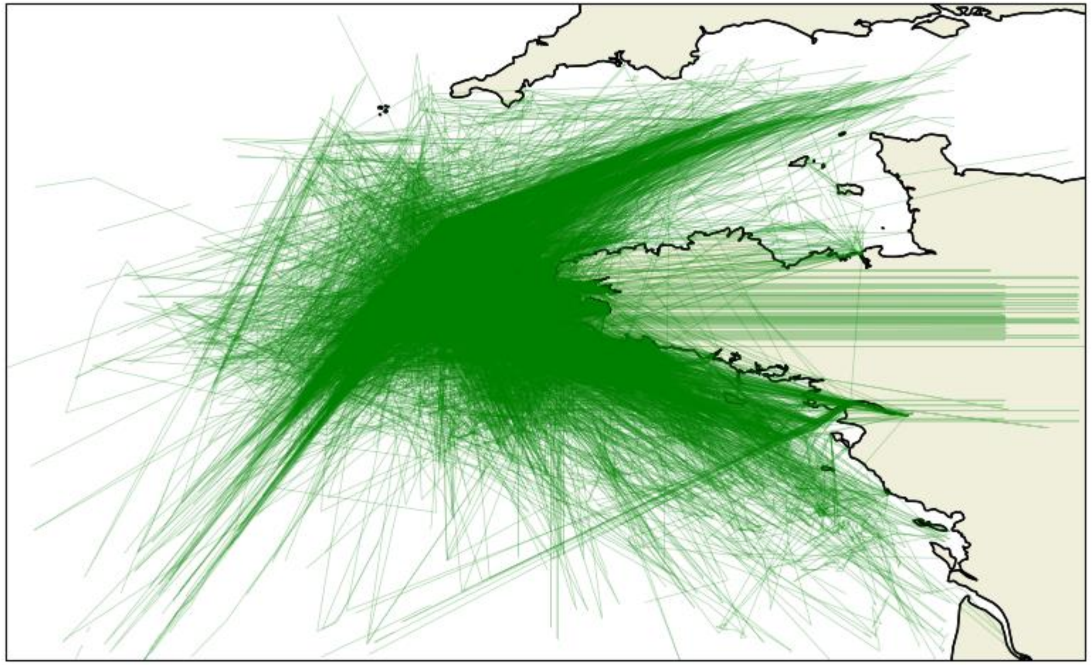
- Comprehensive AIS preprocessing pipeline ensuring noise-free, consistent trajectories that enhance prediction performance.
- A robust sequence forecasting model, combining a single-layer GRU with a dual-head self-attention mechanism.
- Joint prediction of future vessel position, Speed Over Ground (SOG), and Course Over Ground (COG) up to 7 hours ahead, significantly outperforming LSTM and standard GRU baselines.

Methodology (1/2): Preprocessing

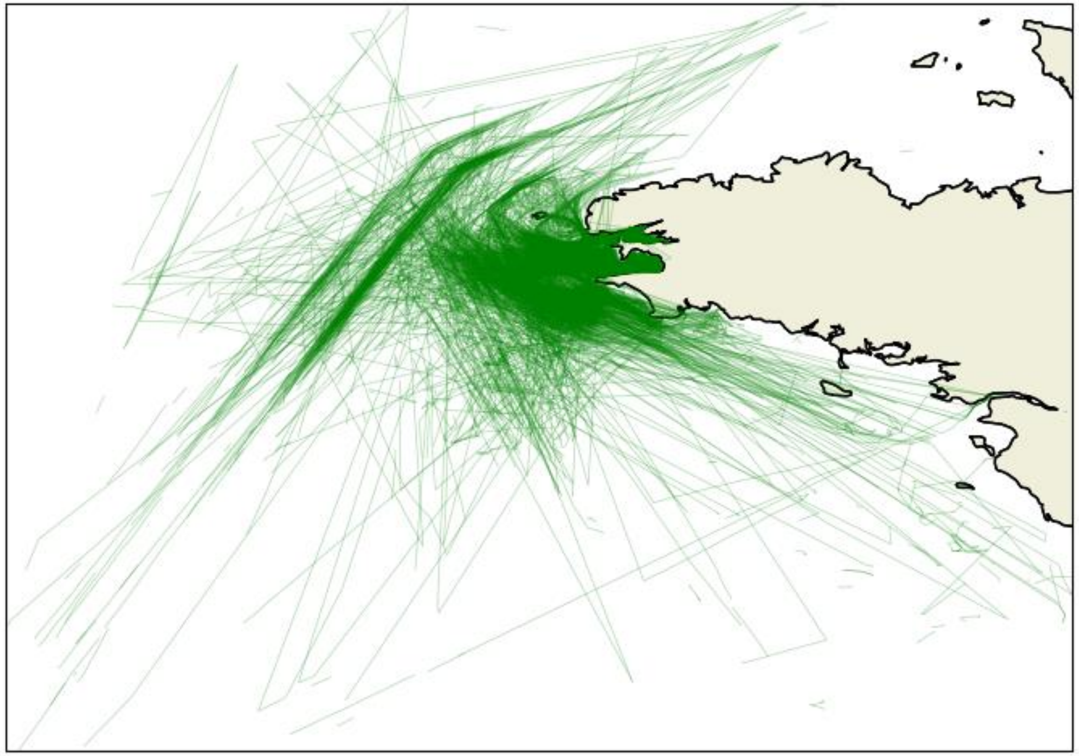
AIS data preprocessing

- Cleaning
  - Remove implausible values
  - Deduplication
  - Remove stationary vessels
- Trajectory extraction
  - Split voyages
  - Trajectory segmentation
- SOG error correction
  - Prevent unrealistic jumps
- Position error correction
  - Smooths unrealistic displacements
- Temporal resampling to 1-minute fixed interval

Raw AIS data



Preprocessed AIS data



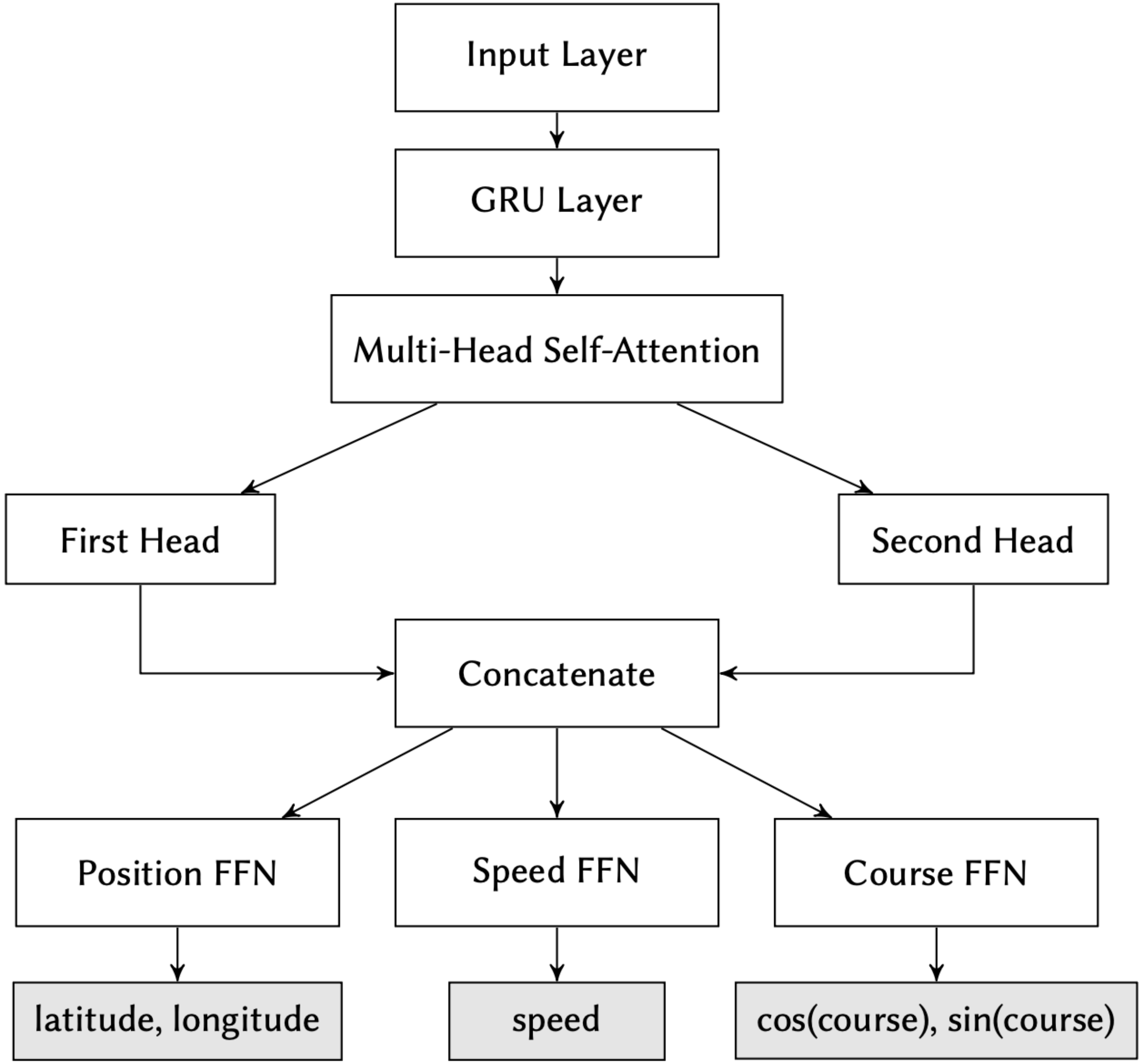
Preprocessing results

Correction Type	Points Updated	Updated (%)
Speed anomaly correction	1 106 345	13.56%
Position anomaly correction	1 620 114	19.85%

Stage	Points	Removed (%)	Vessels / Trajectories
Initial	19 035 630	—	5 055 vessels
After Deduplication	18 026 004	2.93%	5 055 vessels
After COG Filtering	18 025 812	0.00%	5 055 vessels
After SOG Filtering	9 389 527	45.37%	4 982 vessels
After Ship-Type Filtering	8 617 489	4.06%	4 891 vessels
Extracted Trajectories	8 159 865	5.31%	31 136 traj./ 1 681 vessels
After Stationary Removal	6 947 719	14.86%	25 710 trajectories
After 1-min Resampling	1 737 625	75.00%	25 660 traj./ 1 661 vessels

Methodology (2/2): Vessel trajectory prediction framework

Dual-Self-Attention GRU Prediction Model Architecture



Input layer:

Raw trajectory point at time t:

$$p_t = (x_t, y_t, v_t, \theta_t)$$

- x, y: longitude/latitude
- v: Speed Over Ground (SOG)
- θ: Course Over Ground (COG)

Handle course circularity:

$$p'_t = (x_t, y_t, v_t, \cos\psi_t, \sin\psi_t)$$

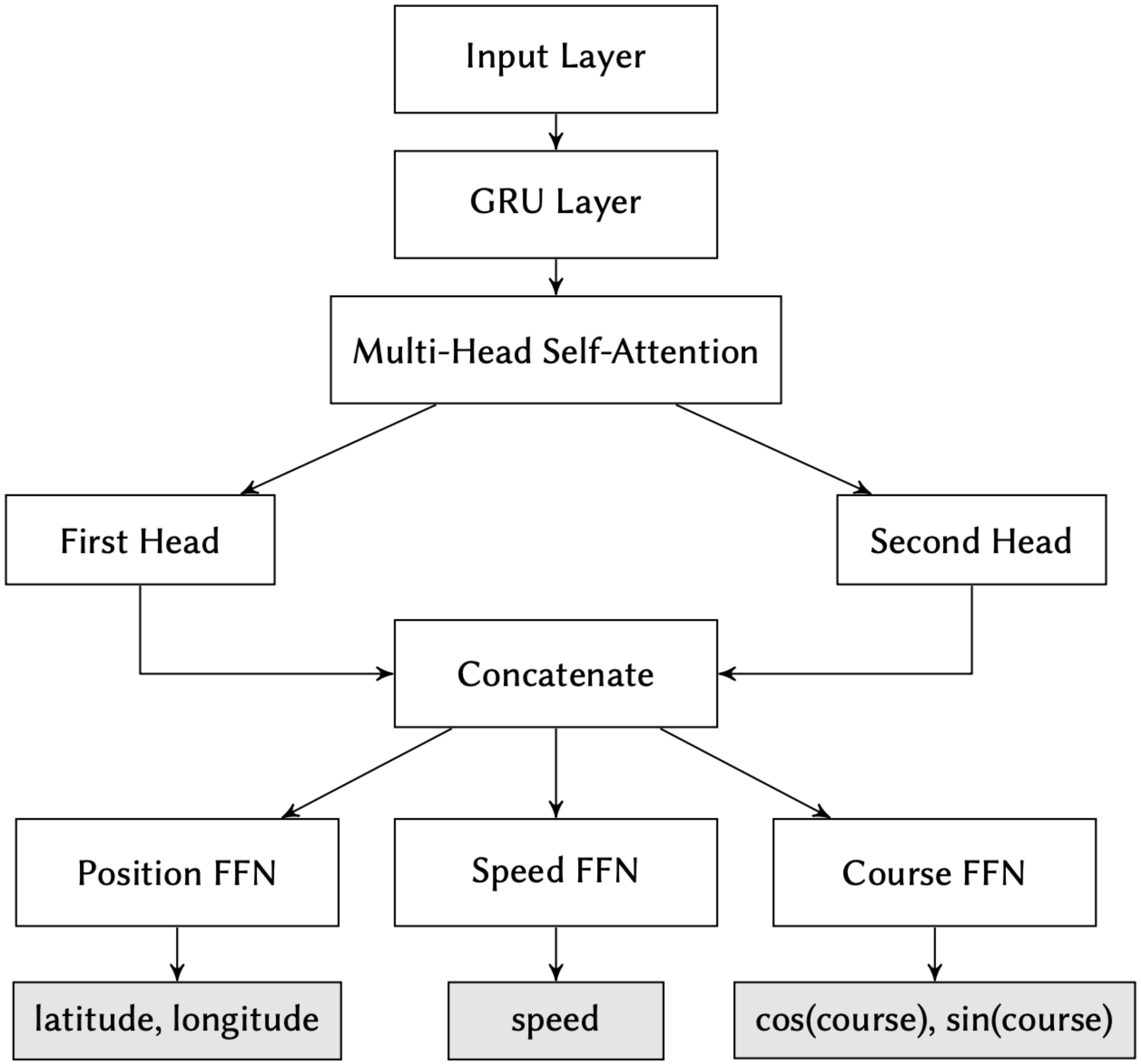
where COG is converted to radians $\psi = \frac{\pi}{180}\vartheta$ and mapped to sine-cosine components

- Model input

$$Xi = \{\tilde{p}_1, \tilde{p}_{i+1}, \tilde{p}_{i+2}, \tilde{p}_{i+3}, \tilde{p}_{i+4}\}$$

- Prediction Target

$$yi = \tilde{p}_{i+5}$$



GRU Layer:

- Hidden size = 256
- Generates full sequences of hidden states $H = \{h_1, h_2, \ldots, h_T\} \in \mathbb{R}^{TX256}$

Dual-head self-attention (two parallel heads):

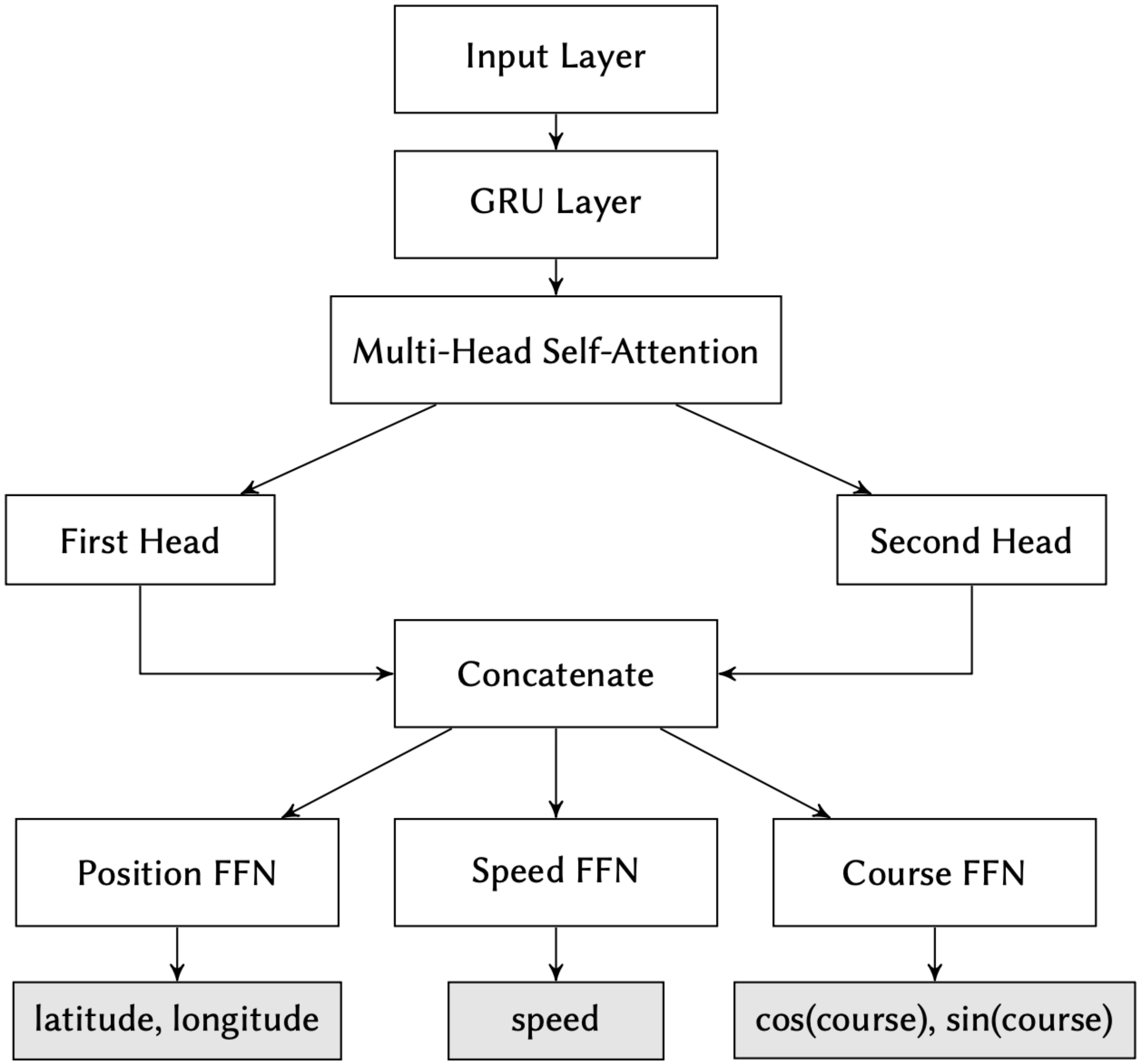
For each head $k \epsilon \mathbb{R}^{TX256}$,

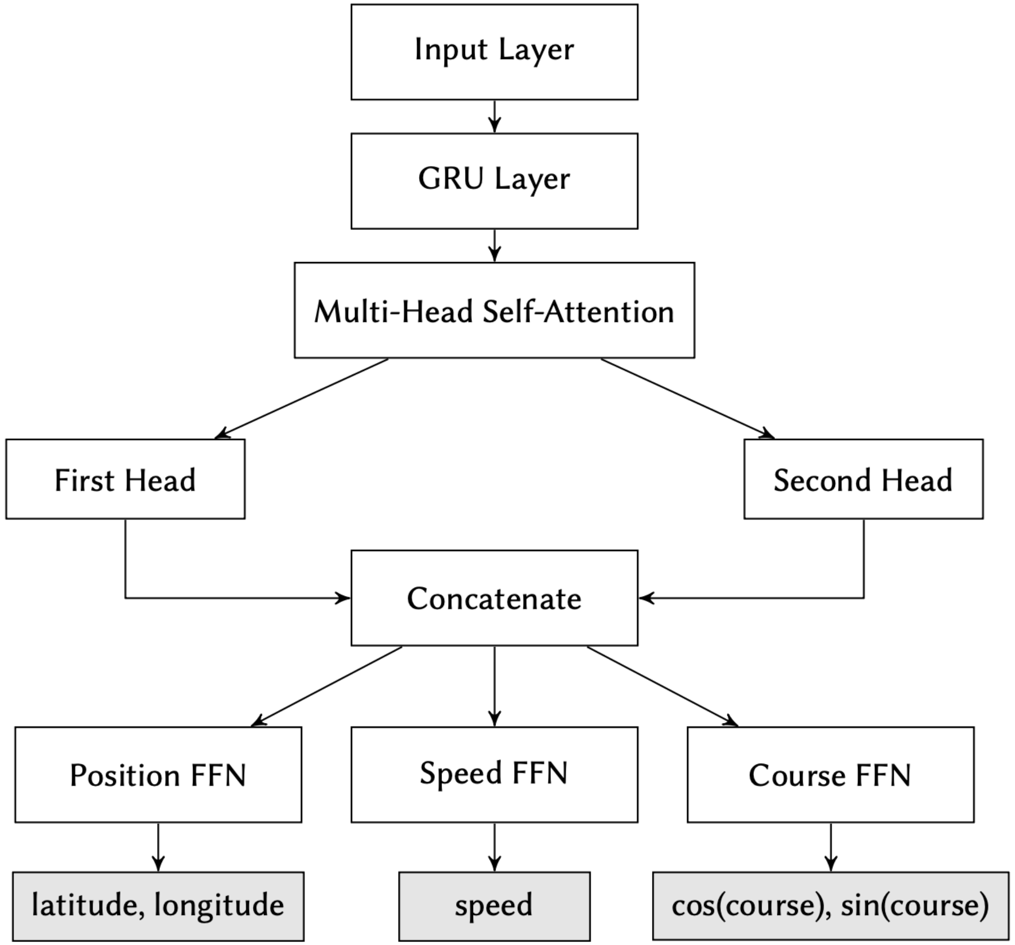
$$\mathrm{head}^{(k)} = \mathrm{softmax}\left(\frac{Q^{(k)}(K^{(k)})^T}{\sqrt{d_k}}\right)V^{(k)}$$

Concatenate heads and project:

$$\mathrm{AttOut} = [\mathrm{head}^{(1)}; \mathrm{head}^{(2)}]W_o + b_o$$

$$H_{out} = \mathrm{LayerNorm}(H + \mathrm{AttOut})$$





Output heads:
One MLP for each modality
- Position (latitude, longitude)
- Speed
- Course $(\cos\theta, \sin\theta)$

Loss Function: $$\mathcal{L} = \frac{1}{N}\sum_{i=1}^{N}\left[(x_i - \hat{x}_i)^2 + (y_i - \hat{y}_i)^2 + (v_i - \hat{v}_i)^2 + \left(1 - \left(\cos\vartheta_{\mathrm{i}}\,\cos\tilde{\vartheta}_i + \sin\vartheta_{\mathrm{i}} \cdot \sin\tilde{\vartheta}_i\right)\right)\right]$$

Experimental results

Model performance comparison

Model	Total Loss	Median Error (m)	Avg Error (m)	Max Error (m)
LSTM	0.0158	98.5	154.0	7341.6
GRU	0.0157	86.3	131.8	8796.3
GRU+Single-head Attention Layer	0.0157	63.1	97.4	7198.9
GRU+Dual-head Attention Layer	**0.0156**	**58.5**	**90.4**	**6327.0**

Cargo vessel trajectory prediction

- 49 points

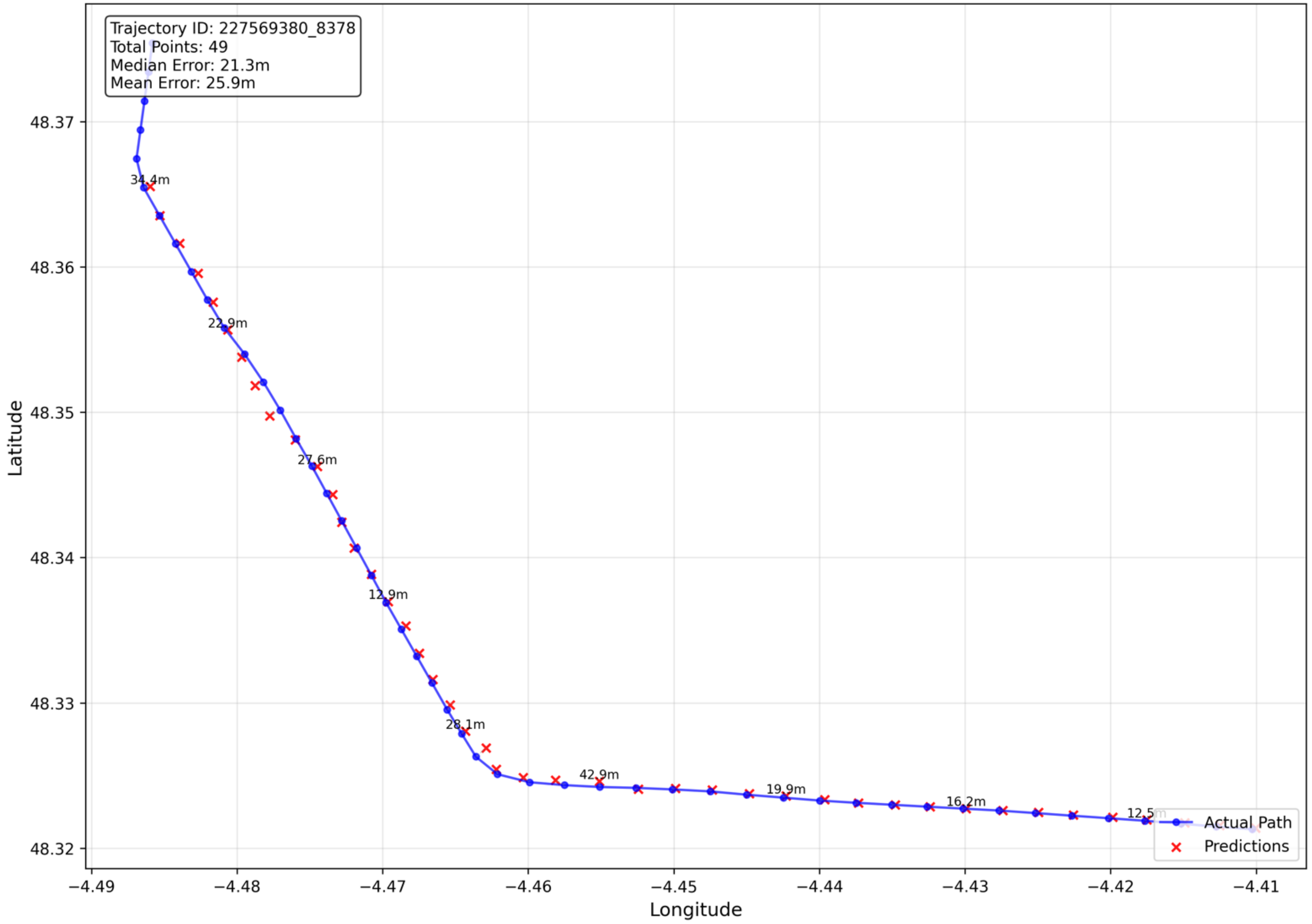


Figure 1a: latitude and longitude prediction

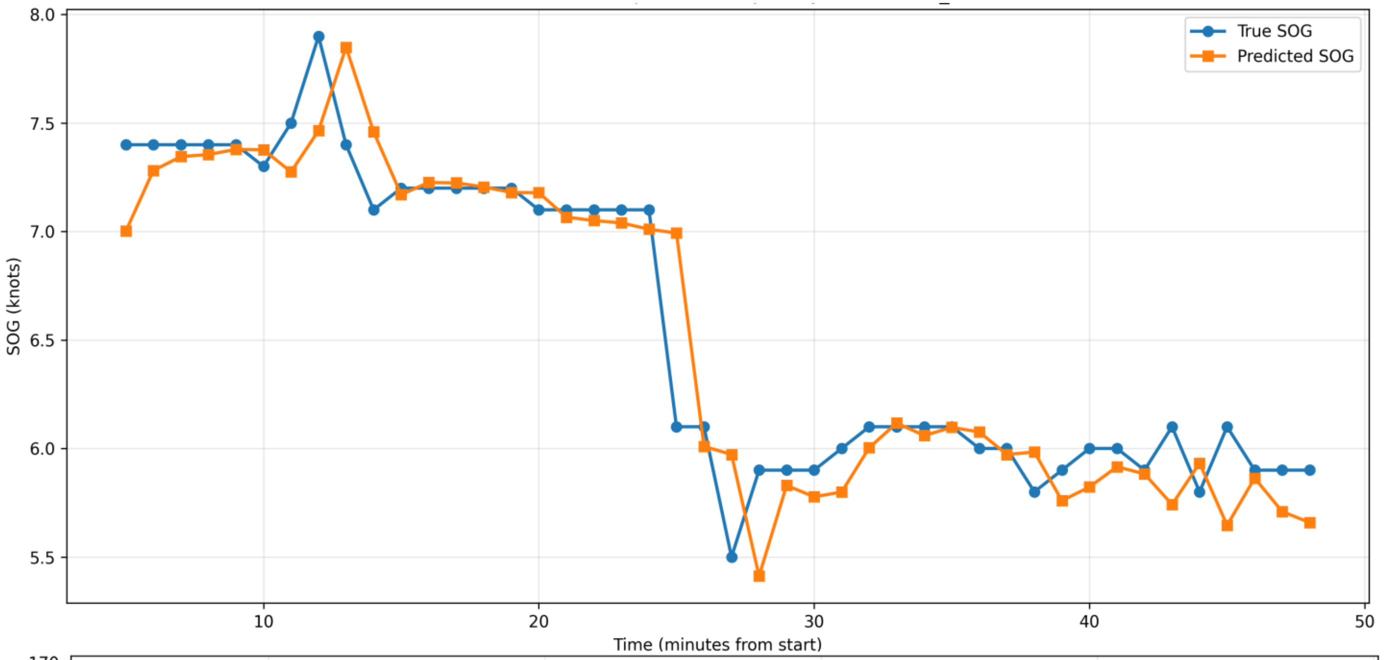


Figure 1.b: SOG prediction

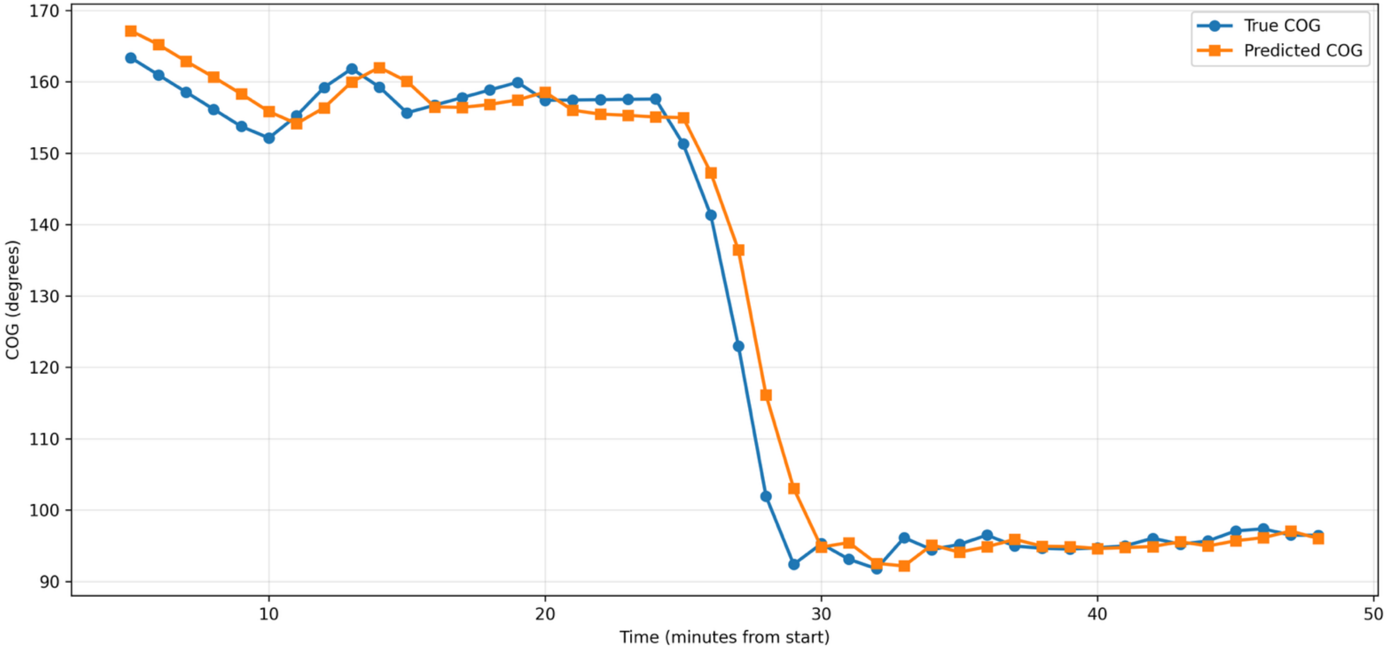


Figure 1.c: COG prediction

Tug vessel trajectory prediction

- 208 points

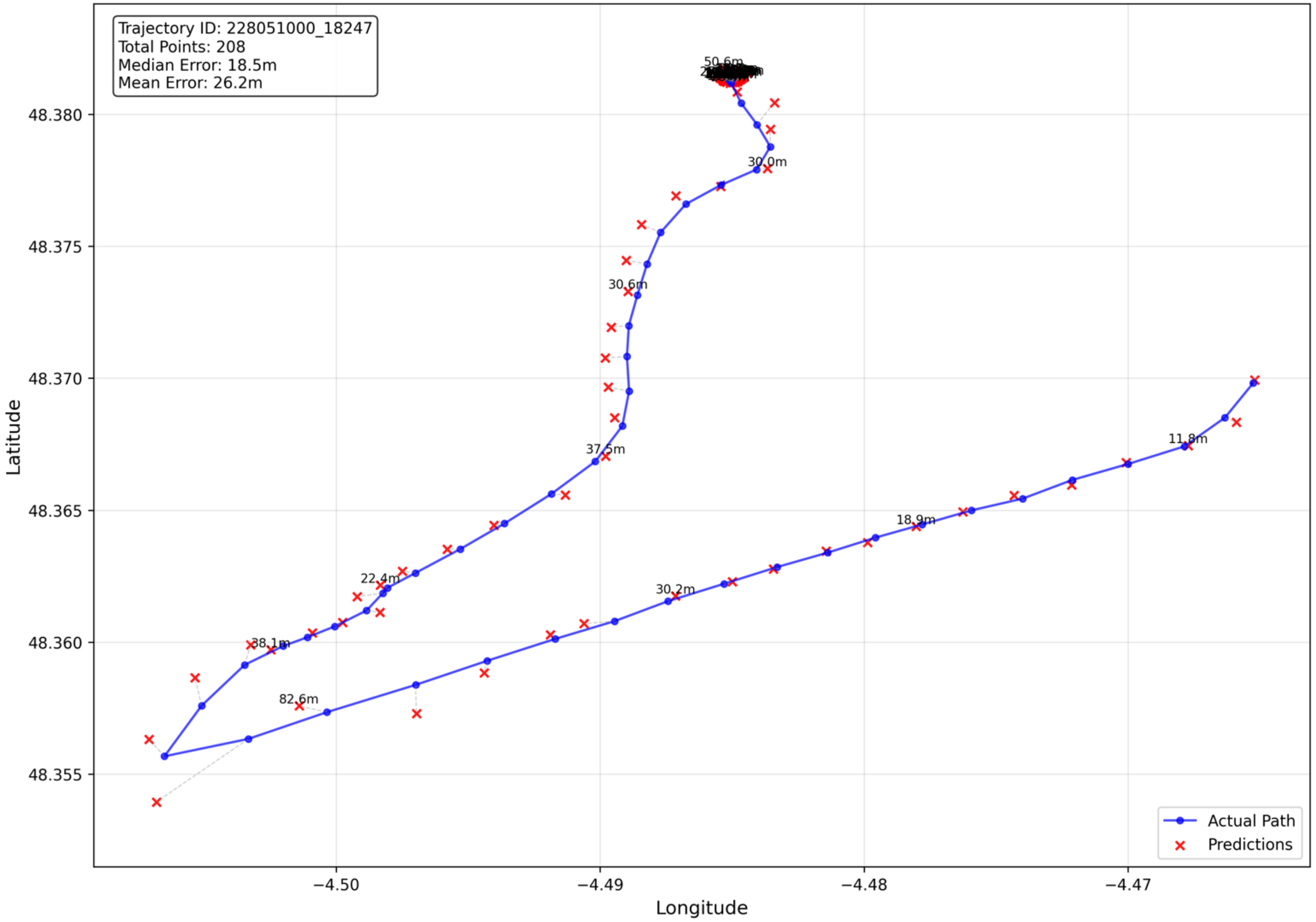
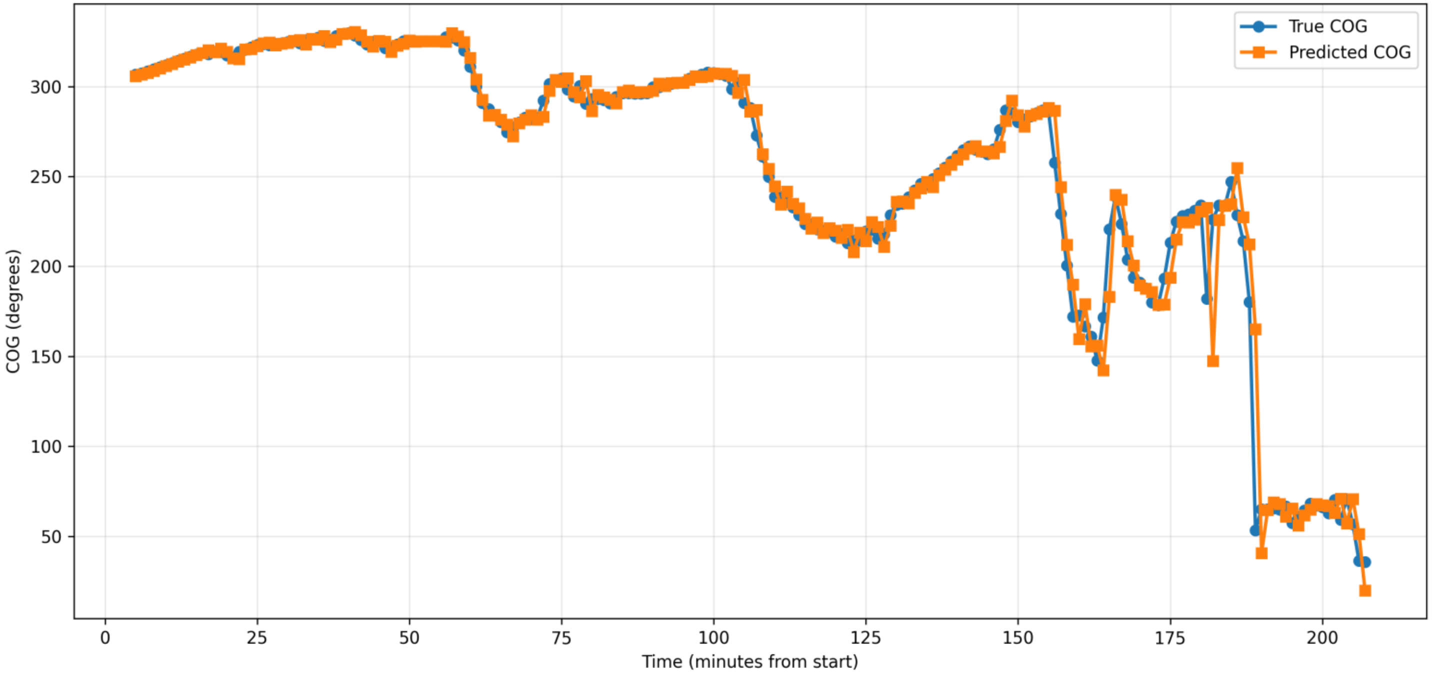


Figure 2a: latitude and longitude prediction



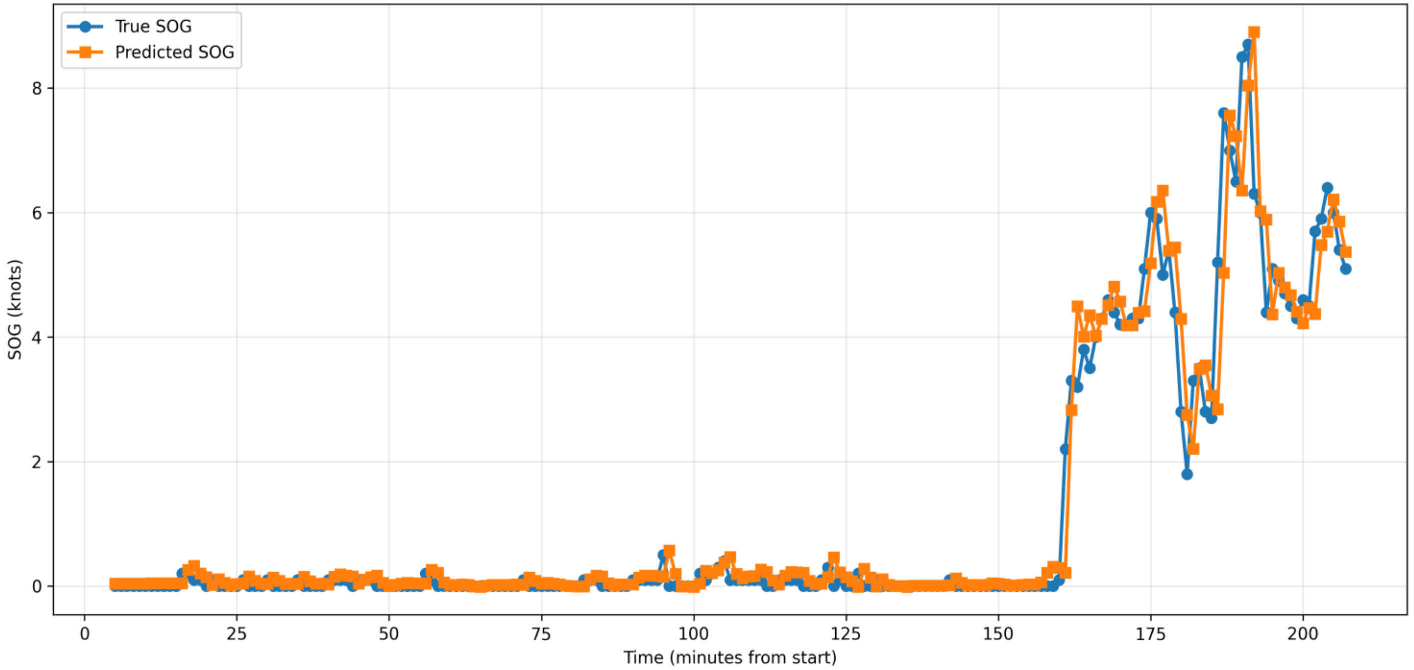


Figure 2.b: SOG prediction

Figure 2.c: COG prediction

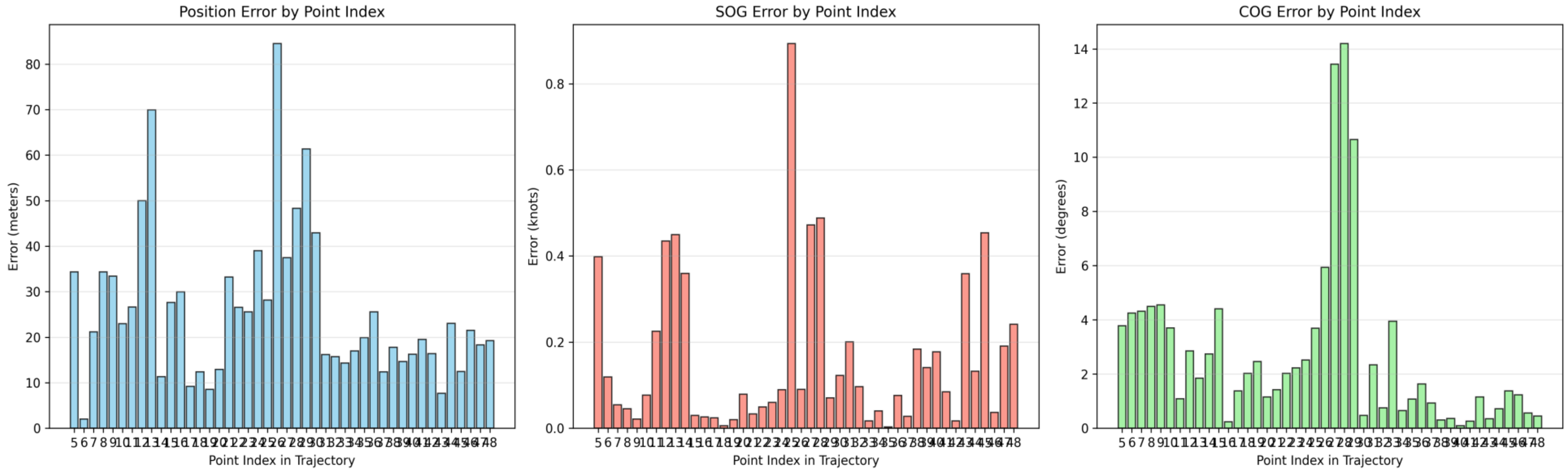


Figure 2.a: Tug vessel error distribution

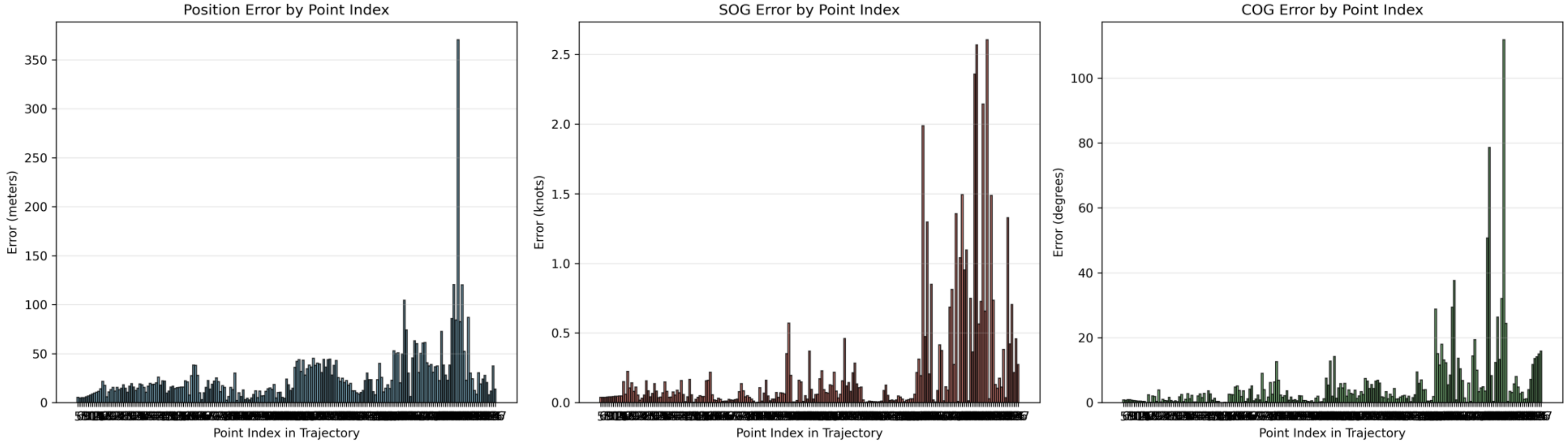


Figure 2.b: Cargo vessel error distribution

Conclusions

- **Proposed preprocessing pipeline** effectively cleans and structures AIS trajectories, making them suitable for prediction and other downstream tasks.
- **Self-attention** dynamically re-weights past time steps, capturing long-range dependencies and abrupt maneuvers that recurrent layers may miss.
- **GRU + dual-head self-attention architecture** accurately forecasts future vessel positions, speeds, and courses for prediction horizons of up to 7 hours.
- **Dual-head attention** outperforms single-head and standard RNN/GRU baselines, demonstrating clear performance gains from learning complementary temporal contexts.

Limitations & future work

- The heterogeneity of our dataset enables better generalization across various dataset types - need to be validated.
- Benchmark our model's strong performance against additional state-of-the-art methods, such as transformers and bidirectional RNNs.
- Evaluate the preprocessed dataset on anomaly detection tasks.

Thank you!

Questions?