Multi-Time Scale State Estimation of Vehicle Dynamics Using Physics-Informed Neural Process

Devin Hunter, Jared Augsburger, Dr. Sean O'Rourke

Ph.D. Student | AFRL Sensors Directorate | US Army Combat Capabilities Development Command Army Research Lab

INTRODUCTION

Reliable vehicle tracking from remote sensing data often requires robust state estimation across widely varying temporal resolutions. In this work, we propose a physics-informed attentive neural process (PI-AttNP) for multi-scale vehicle tracking that explicitly integrates simplified kinematic models as structural priors within a neural process framework. We benchmark this estimation approach on both real-world ship-tracking Automatic Identification System (AIS) and aeroplane Automatic Dependent Surveillance–Broadcast (ADS-B) datasets on various time scales spanning a queried time of two minutes to several hours. We compare our estimator with state-of-the-art baselines and obtain promising results.

PROBLEM STATEMENT

In this study, the objective is to learn nonlinear, coupled vector field f that represents time-varying vehicle dynamics in the following discrete-time transitional model:

$$x_{k+1} = f(\mathcal{X}, \mathcal{U}, \Delta, \Xi, \delta\mathcal{X})$$

We desire to approximate nominal dynamics f with learned representation f_Γ with neural process parameters Γ that uses *inferred, physics* model $g^*(\mathcal{X}, \mathcal{U}, \Delta) \in \mathcal{G} = \{g_1, g_2, \ldots, g_K\}$ through the following implicit prediction scheme:

$$f_\Gamma = g^*(\mathcal{X}, \mathcal{U}, \Delta) + NN(\mathcal{X}, \mathcal{U}, \Delta|\Gamma)$$

Such that $f_\Gamma \approx f$ and no knowledge of $\Xi, \delta\mathcal{X}$

Contextual State Vector: $\mathcal{X} = \{x_{k-C}, \ldots, x_{k-1}, x_k\}$

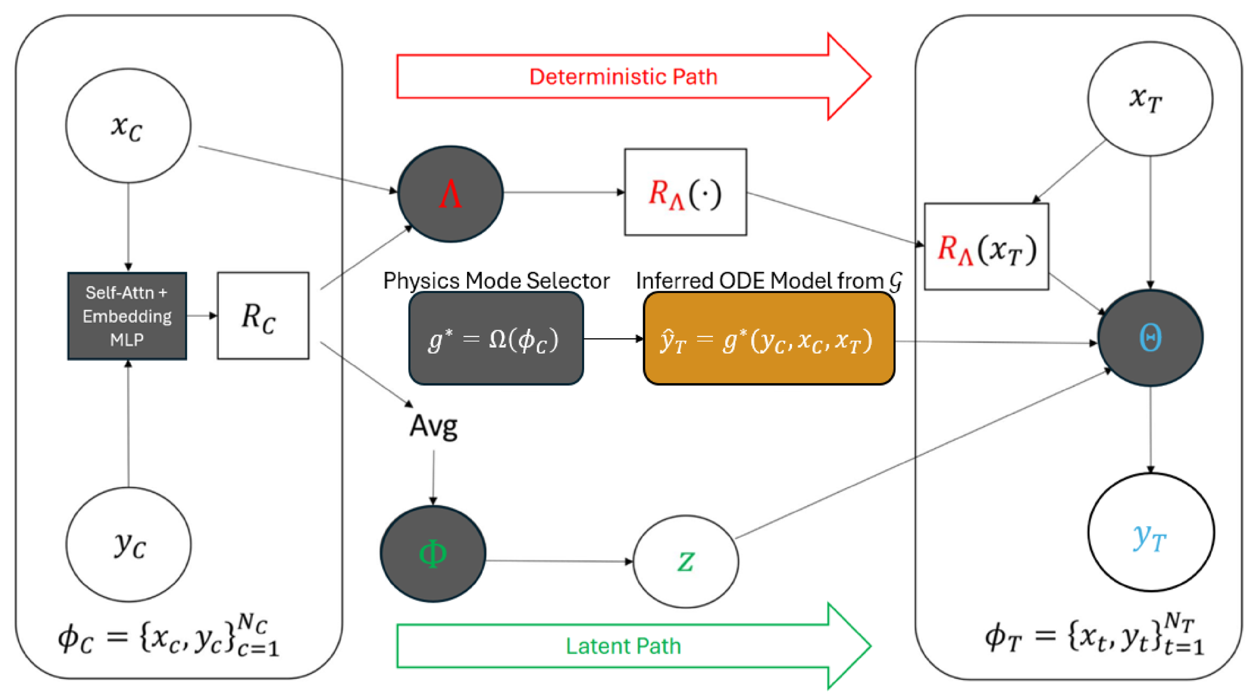
Contextual Input Vector: $\mathcal{U} = \{x_{k-C}, \ldots, x_{k-1}, x_k\}$

Time Step Vector: $\Delta = \{\Delta k - C, \ldots, \Delta k - 1, \Delta k\}$

Sensor Noise Vector: $\Xi = \{\xi_{k-C}, \ldots, \xi_{k-1}, \xi_k\}$

State Disturbance Vector: $\delta\mathcal{X} = \{\delta x_{k-C}, \ldots, \delta x_{k-1}, \delta x_k\}$

TECHNICAL APPROACH



Computational forward diagram of the physics-informed attentive neural process (PI-AttNP). Black circles denote utilized neural networks in the NP architecture

- With the inclusion of both self and cross attention mechanisms within the forward pass of the NP, the PI-AttNP closely mimics the structure of the previous AttNP model with both *latent* and *deterministic* paths
- The inferred ODE model $g^*(y_C, x_C, x_T) \in \mathcal{G}$ is utilized as a physics prior that outputs apriori estimate $\hat{y}$ given to decoder $\Theta(x_T, z, R_\Lambda, \hat{y})$
- With the inclusion of $\hat{y}$ in decoder distribution, the NP uses the following variational lower-bound optimization:

$$p^*_\Theta, q^*_\Phi, R^*_\Lambda = \max_\Gamma \mathcal{L}(p_\Theta, q_\Phi, R_\Lambda)$$

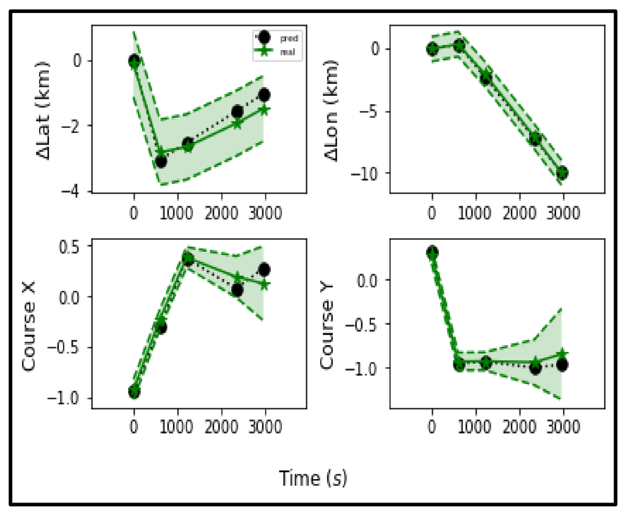
$$\log p(y_T|x_T, \phi_C) \geq \mathcal{L}(p_\Theta, q_\Phi, R_\Lambda)$$

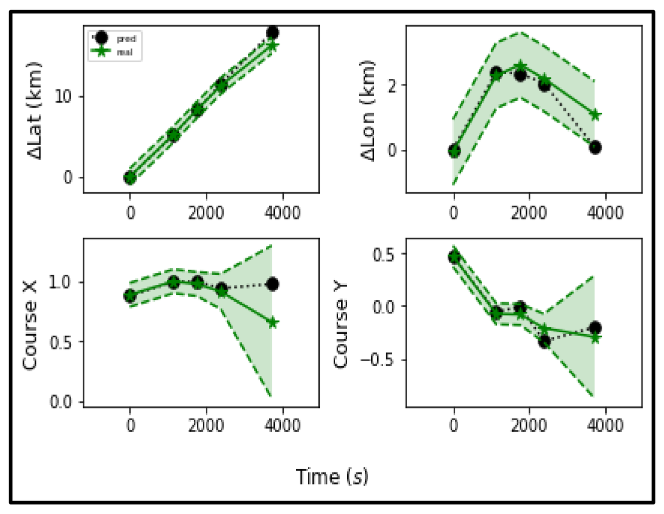
$$= \mathbb{E}_{z\sim q_\Phi} \log p_\Theta(y_T|x_T, z, R_\Lambda, \hat{y})$$

$$-D_{KL}\big(q_\Phi(z|\phi_T)||q_\Phi(z|\phi_C)\big)$$

For parameters $\Gamma = \{\Theta, \Phi, \Lambda\}$

TRAINING PROCEDURE





AIS Training Regime

NP Input: $x_T = \begin{bmatrix}\Delta k \\ v_k\end{bmatrix}$

NP Output: $y_T = \begin{bmatrix}\Delta Lat_{k+1} \\ \Delta Lon_{k+1} \\ \theta_x(k+1) \\ \theta_y(k+1)\end{bmatrix}$

Where v_k = speed over ground and θ_x, θ_y are x-y components of heading θ

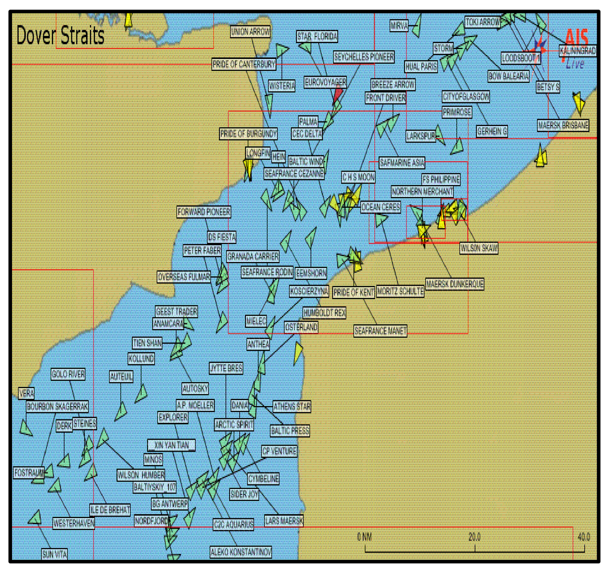
ADS-B Training Regime

NP Input: $x_T = \begin{bmatrix}\Delta k \\ v_k\end{bmatrix}$

NP Output: $y_T = \begin{bmatrix}\Delta Lat_{k+1} \\ \Delta Lon_{k+1} \\ \Delta h_{k+1} \\ \theta_x(k+1) \\ \theta_y(k+1)\end{bmatrix}$

Where v_k = longitudinal plane speed and h = altitude

PHYSICS MODELS USED IN $\mathcal{G}$



For AIS Dynamics

Within the AIS dataset, most dynamic transitions could be modeled as either a straight-line g_1, constant velocity model g_2, and stationary g_3 (i.e. $\mathcal{G} = \{g_1, g_2, g_3\}$). For g_1 and g_2, the following ellipsoid-aware dynamics are utilized:

$$\Delta\phi_{km} = v_k \Delta k \cdot c_{ellips} \cos(\phi_k)$$

$$\Delta\lambda_{km} = v_k \Delta k \cdot c_{ellips} \sin(\phi_k)$$

$$\Delta\theta_k = \hat{\omega}\Delta k$$

Where,

- $c_{ellips} = \frac{1852}{3600 \cdot 1000} \cdot \frac{180}{\pi} \cdot \frac{111}{R}$
- ϕ = latitude (deg)
- λ = longitude (deg)
- $\hat{\omega}$ = turning rate (rad/s)

$\hat{\omega}$ is inferred from turning patterns found in $\mathcal{X}$ within ϕ_C (for $g_1, \hat{\omega} \approx 0$)

For ADS-B Dynamics

We model aircraft motion using a three-mode physics prior:

$$\mathcal{G} = \{g_{ascend}, g_{level}, g_{descend}\}$$

Aircraft heading $\Delta\psi_i$ and position $\Delta\phi_{km}, \Delta\lambda_{km}$ is predicted using recent contextual detections:

Ground Truth: $\Delta\psi_i = \psi_i - \psi_{i-1}$

Predicted: $\widehat{\Delta\psi}_k = \sum w_i \Delta\psi_i$

$$\psi_{eff} = \psi_C + \frac{1}{2}\widehat{\Delta\psi}_k$$

$$\Delta\phi_{km} = \frac{v_k \cos(\psi_{eff})\Delta k}{M(\phi_C) + h_k}$$

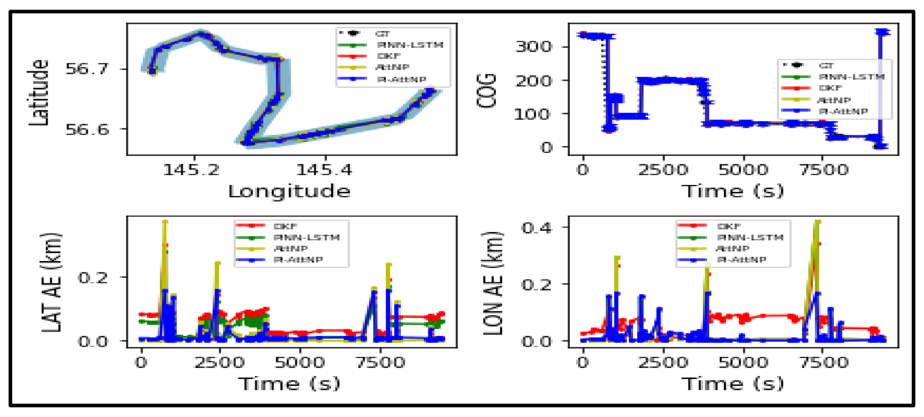
$$\Delta\lambda_{km} = \frac{v_k \sin(\psi_{eff})\Delta k}{(N(\phi_C) + h_k)\cos(\phi_k)}$$

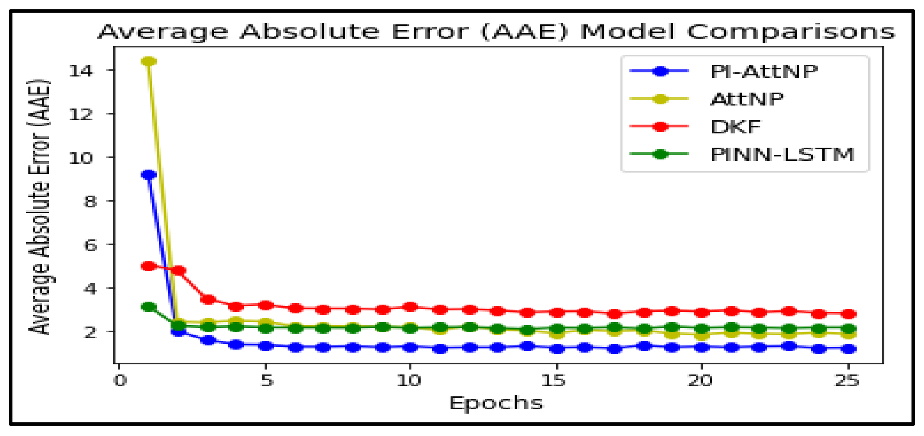
$$\hat{v}_z = \sum_{i=1}^{C} w_i \frac{\Delta h_i}{t_i - t_{i-1}}$$

where ψ_{eff} = midpoint heading, M, N = earth radii of ellipsoid, $\hat{v}$ = predicted vertical velocity

RESULTS & DISCUSSIONS

AIS Results





REFERENCES

1. Garnelo, M., Schwarz, J., Rosenbaum, D., Viola, F., Rezende, D. J., Eslami, S. M. A., & Teh, Y. W. (2018). Neural processes. *arXiv preprint arXiv:1807.01622*.
2. Kim, H., Mnih, A., Schwarz, J., Garnelo, M., Eslami, S. M. A., Rosenbaum, D., Vinyals, O., & Teh, Y. W. (2019). Attentive neural processes. *International Conference on Learning Representations*.
3. Krishnan, R. G., Shalit, U., & Sontag, D. (2017). *Deep Kalman filters. arXiv preprint arXiv:1511.05121*.