When Context Is Not Enough: Modeling Unexplained Variability in Car-Following Behavior

Chengyuan Zhang[a], Zhengbing He[b], Cathy Wu[b], and Lijun Sun[a,†] († lijun.sun@mcgill.ca)

a. Department of Civil Engineering, McGill University, Montreal, Quebec H3A 0C3, Canada.

b. Laboratory for Information & Decision Systems (LIDS), Massachusetts Institute of Technology, Cambridge, MA 02139, USA

Scan Me

Introduction

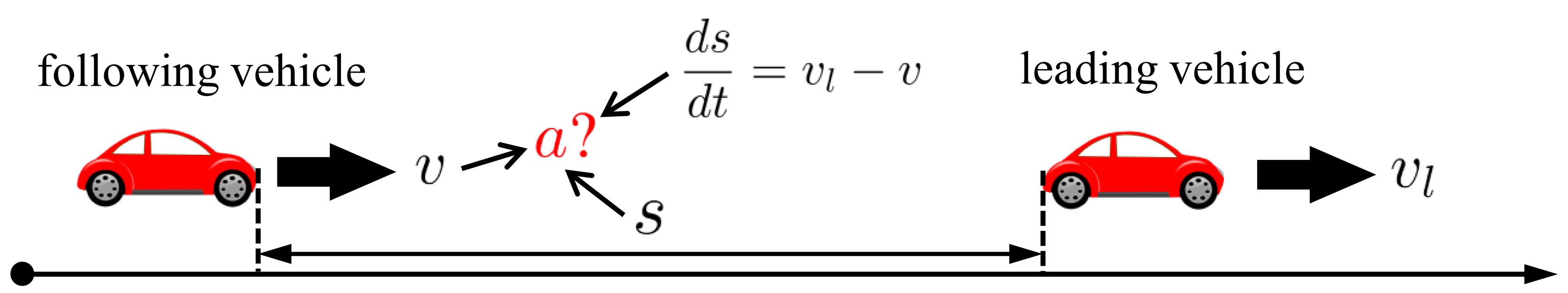


Fig. 1: Physical settings of a car-following scenario.

- **The Challenge**: Traditional car-following models are deterministic or use simplistic, uncorrelated noise. They fail to capture the "**stochasticity**" in human driving arising from latent intentions, perception errors, and memory effects.
- **The Gap**: Existing stochastic models assume **stationarity** (fixed noise structure), which cannot adapt to evolving traffic contexts (e.g., sudden braking).
- **Our Goal**: Develop an interpretable car-following framework that captures context-dependent temporal correlations.

We represent the action $a_t = a(\boldsymbol{x}_t, t)$ of a following vehicle at time t as

$$a_t = f_{\mathrm{CF}}(\boldsymbol{x}_t) + \delta(t) + \epsilon_t, \tag{1}$$

where $\boldsymbol{x}_t = [s_t, \Delta v_t, v_t]$ denotes the input covariates, $f_{\mathrm{CF}}(\boldsymbol{x}_t)$ as a function of $\boldsymbol{x}_t$ represents the mean car-following model, $\delta(t)$ accounts for temporal correlations, and $\epsilon_t \overset{i.i.d.}{\sim} \mathcal{N}(0, \sigma_0^2)$ is an *independent and identically distributed (i.i.d.)* noise term with variance σ_0^2.

Table 1: Modeling of temporal correlations in the literature, corresponds to Equation (1).

Reference	$f_{\mathrm{CF}}(\boldsymbol{x})$	$\delta(t)$	Nonstationary?
Treiber et al. (2006)	IDM	Ornstein-Uhlenbeck (OU) processes	✗
Hoogendoorn and Hoogendoorn (2010)	GHR/IDM	Cochrane-Orcutt correction (i.e., AR(1) process)	✗
Zhang and Sun (2024)	IDM	Gaussian processes	✗
Zhang et al. (2024b)	IDM	AR processes with higher orders	✗
This work	NN	nonstationary GPs	✓

☛ arXiv: https://arxiv.org/pdf/2507.07012 (Accepted at **ISTTT26** & **TR Part B**)

☛ Code will be released soon.

Summary

- **Hybrid Neural-GP Framework**: We introduce a novel stochastic model that integrates deep recurrent neural networks (DeepAR) with a nonstationary Gibbs kernel, explicitly capturing context-dependent temporal correlations in car-following behavior that traditional models overlook.
- **Interpretable Behavioral Dynamics**: The model offers explainable insights by learning dynamic kernel parameters: the lengthscale adapts to reflect driver reaction frequency (memory), while the variance captures evolving tolerance for behavioral heterogeneity.
- **Superior Simulation Fidelity**: Validated on the HighD dataset, our approach significantly outperforms both deterministic baselines and stationary GP variants, achieving lower simulation errors (RMSE) and more realistic, well-calibrated uncertainty quantification (CRPS/ES).

Methodology

We propose a hybrid formulation combining Deep Neural Networks with Nonstationary Gaussian Processes (GPs). The general model formulation:

$$a_t = \underbrace{a_{\mathrm{NN}}(\boldsymbol{h}_t;\boldsymbol{\theta})}_{\text{Mean Dynamics}} + \underbrace{\delta_{\mathrm{GP}}(t;\boldsymbol{\lambda})}_{\text{Temporal Correlation}} + \underbrace{\epsilon_t}_{\text{White Noise}}, \tag{2}$$

The Core Innovation: Scenario-Adaptive Gibbs Kernel

$$k_{\text{Gibbs}}(t,t';\boldsymbol{\lambda}) := \sigma(t)\sigma(t')\underbrace{\sqrt{\frac{2\ell(t)\ell(t')}{\ell(t)^2+\ell(t')^2}}\exp\left(-\frac{(t-t')^2}{\ell(t)^2+\ell(t')^2}\right)}_{k^\star_{\text{Gibbs}}(t,t';\boldsymbol{\lambda})}, \tag{3}$$

Notably, the Gibbs kernel simultaneously captures two critical properties: **Heteroskedasticity** via context-dependent variance $\sigma^2(\boldsymbol{x}_t)$, representing time-varying uncertainty; and **Nonstationary correlations** via context-adaptive lengthscale $\ell(t)$, encoding dynamic adaptation in temporal dependence.

Training Method: "Better Batch" strategy

$$\underbrace{\begin{bmatrix} a_1 \\ \vdots \\ a_{\Delta T} \end{bmatrix}}_{\boldsymbol{a}^{\text{batch}}} \sim \mathcal{N}\left(\underbrace{\begin{bmatrix} a_{\mathrm{NN}}(\boldsymbol{h}_1;\boldsymbol{\theta}) \\ \vdots \\ a_{\mathrm{NN}}(\boldsymbol{h}_{\Delta T};\boldsymbol{\theta}) \end{bmatrix}}_{\boldsymbol{a}_{\mathrm{NN}}^{\text{batch}}}, \boldsymbol{\Sigma}\right), \tag{4}$$

$$\Sigma = \underbrace{\begin{bmatrix} k(\Delta t, \Delta t;\boldsymbol{\lambda}) & \cdots & k(\Delta t, \Delta T\Delta t;\boldsymbol{\lambda}) \\ \vdots & \ddots & \vdots \\ k(\Delta T\Delta t, \Delta t;\boldsymbol{\lambda}) & \cdots & k(\Delta T\Delta t, \Delta T\Delta t;\boldsymbol{\lambda}) \end{bmatrix}}_{\boldsymbol{K}=\mathrm{diag}(\boldsymbol{\sigma}^{\text{batch}})\boldsymbol{K}^\star\mathrm{diag}(\boldsymbol{\sigma}^{\text{batch}})} + \sigma_0^2\boldsymbol{I}_{\Delta T} \tag{5}$$

where $\boldsymbol{I}_{\Delta T}$ denotes a $\Delta T \times \Delta T$ identity matrix, $\sigma_0^2\boldsymbol{I}_{\Delta T}$ captures homoskedastic observation noise, $\boldsymbol{\sigma}^{\text{batch}} = [\sigma(\Delta t), \ldots, \sigma(\Delta T\Delta t)]$, and $\boldsymbol{K}^\star = [k^\star_{\text{Gibbs}}(t,t';\boldsymbol{\lambda})]$ is the base kernel matrix prior to applying variance modulation.

Prediction: Given the neural predictions, the conditional likelihood of the observed acceleration sequence $\boldsymbol{a}^{\text{batch}}$ follows the multivariate Gaussian

$$\boldsymbol{a}^{\text{batch}} | \boldsymbol{a}_{\mathrm{NN}}^{\text{batch}}, \boldsymbol{\ell}_{\mathrm{NN}}^{\text{batch}}, \boldsymbol{\sigma}_{\mathrm{NN}}^{\text{batch}}, \boldsymbol{\theta} \sim \mathcal{N}(\boldsymbol{a}_{\mathrm{NN}}^{\text{batch}}, \boldsymbol{K} + \sigma_0^2\boldsymbol{I}_{\Delta T}), \tag{6}$$

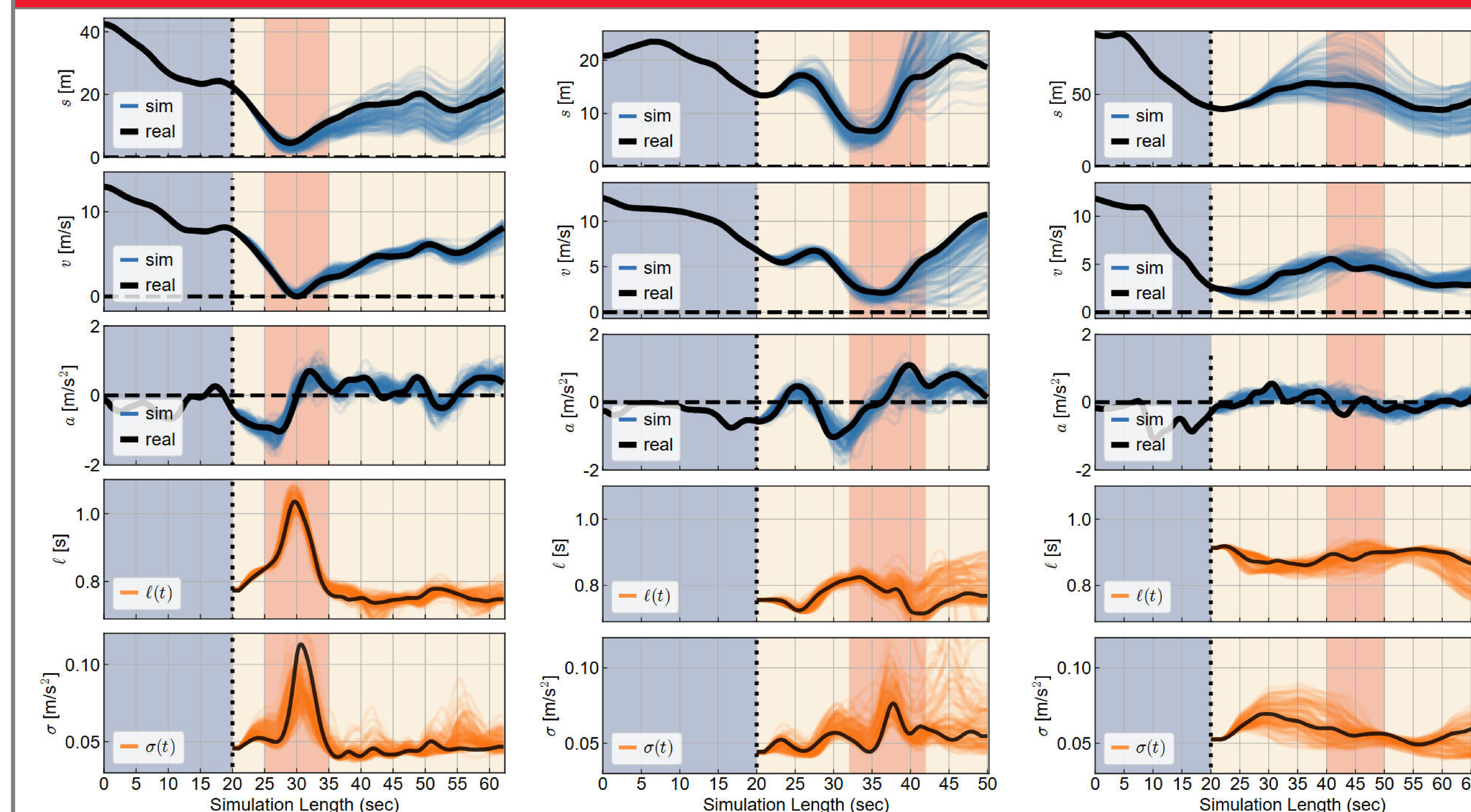
where $\sigma_0^2\boldsymbol{I}_{\Delta T}$ models homoskedastic observation noise. This formulation allows the model to output both predictions and temporally structured uncertainty, modulated by the GP kernel.

Optimization Problem:

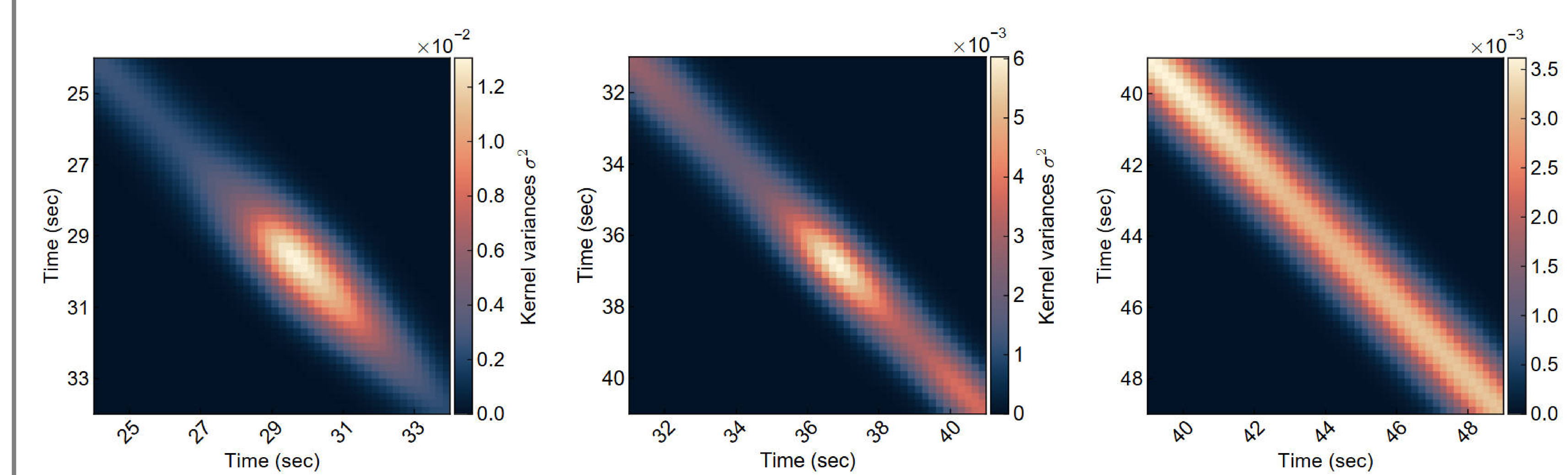
$$\boldsymbol{\theta}^\star = \arg\min_{\boldsymbol{\theta}} \frac{1}{2}\log|\boldsymbol{K}+\sigma_0^2\boldsymbol{I}_{\Delta T}| + \frac{1}{2}(\hat{\boldsymbol{a}}^{\text{batch}} - \boldsymbol{a}_{\mathrm{NN}}^{\text{batch}})^\top(\boldsymbol{K}+\sigma_0^2\boldsymbol{I}_{\Delta T})^{-1}(\hat{\boldsymbol{a}}^{\text{batch}} - \boldsymbol{a}_{\mathrm{NN}}^{\text{batch}}). \tag{7}$$

Here $\hat{\boldsymbol{a}}^{\text{batch}}$ denotes the observed accelerations, and $\boldsymbol{a}_{\mathrm{NN}}^{\text{batch}}$ are the corresponding model predictions. The constant term $n\log(2\pi)/2$ is omitted as it does not affect the optimization.

Stochastic Simulations



(a) Car-following Pair #1. (b) Car-following Pair #2. (c) Car-following Pair #3.



(d) Example $\boldsymbol{K}$ in Pair #1. (e) Example $\boldsymbol{K}$ in Pair #2. (f) Example $\boldsymbol{K}$ in Pair #3.

Fig. 2: Stochastic simulation results for three representative car-following cases. (a)-(c): Each example shows 100 predicted trajectories of spacing s, speed v, and acceleration a (blue), compared to the ground-truth (black). The dashed line marks the forecast start. Below each example, simulated context-dependent $\ell(\boldsymbol{x}_t)$ and $\sigma(\boldsymbol{x}_t)$ (orange) are compared to DeepAR conditioned on ground truth (black). (d)-(f): Bottom-row heatmaps visualize the kernel $\boldsymbol{K}$, revealing the evolving temporal correlation structure during the forecast horizon.

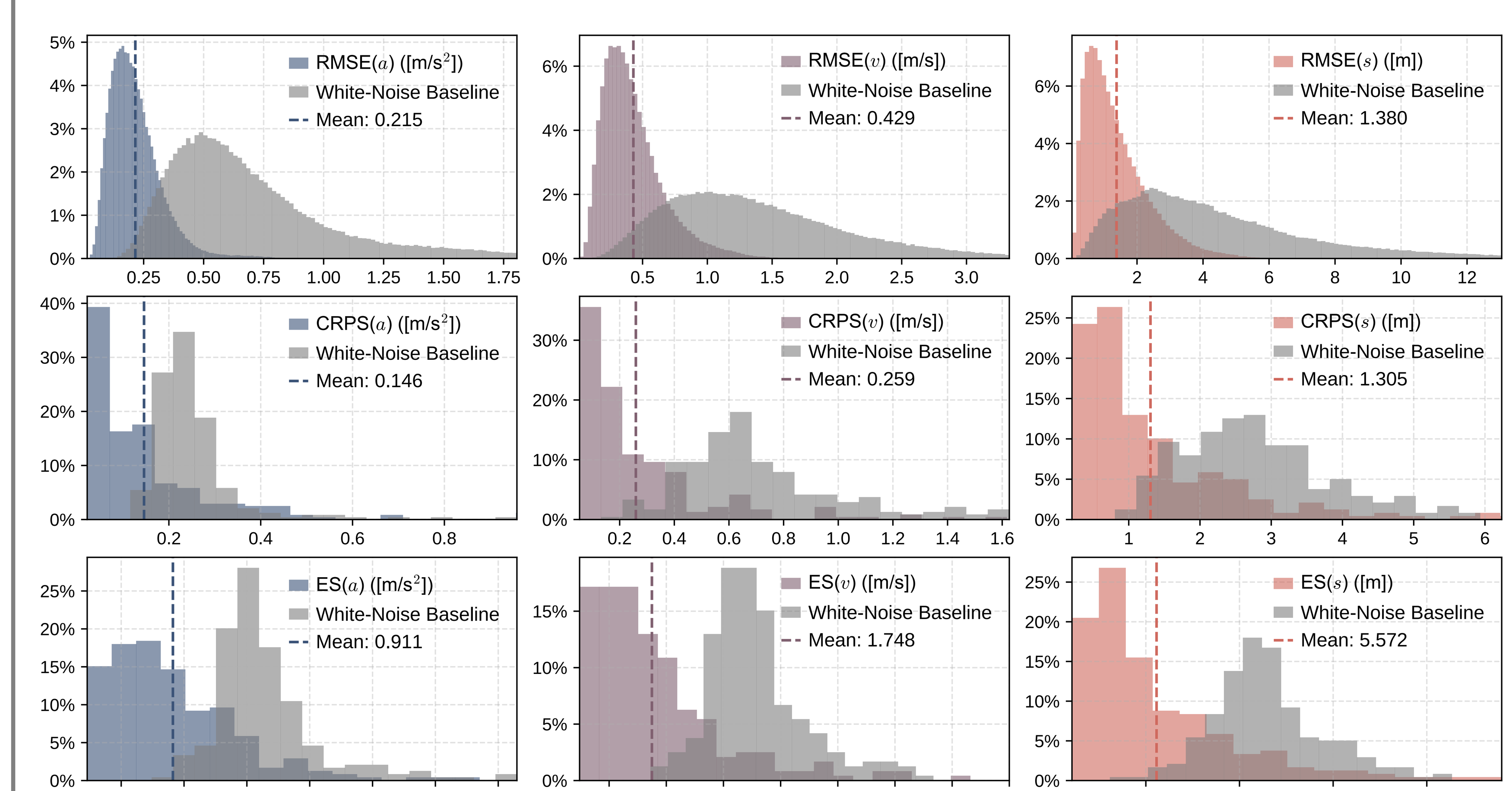


Fig. 3: Error distributions.