Non-Reactive Autonomous Vehicle Simulation

Daniel Dauner[1,2] Marcel Hallgarten[1,5] Tianyu Li[3] Xinshuo Weng[4] Zhiyu Huang[4,6] Zetong Yang[3] Hongyang Li[3] Igor Gilitschenski[7,8] Boris Ivanovic[4] Marco Pavone[4,9] Andreas Geiger[1,2] Kashyap Chitta[1,2]

[1]University of Tübingen [2]Tübingen AI Center [3]OpenDriveLab at Shanghai AI Lab [4]NVIDIA Research [5]Robert Bosch GmbH [6]Nanyang Technological University [7]University of Toronto [8]Vector Institute [9]Stanford University

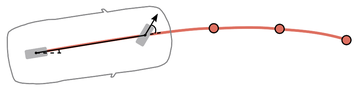
NAVSIM evaluates driving agents on real-world data.

Agent Input (with 1.5s history):

- 8 × surround-view cameras
- 5 × merged LiDAR
- Ego velocity & acceleration
- Navigation goal

Task: Predict short-term trajectory

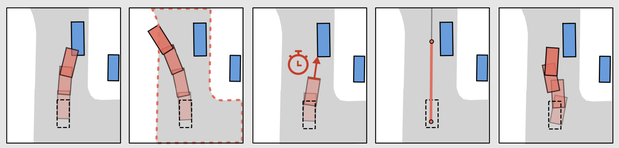
We simulate trajectories in non-reactive environments…



$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u},$$
$$J = \int_0^\infty (\mathbf{x}^\mathsf{T}\mathbf{Q} + \mathbf{u}^\mathsf{T}\mathbf{R}\mathbf{u})\,\mathrm{d}t$$

…with a *kinematic bicycle model* and an *LQR controller*.

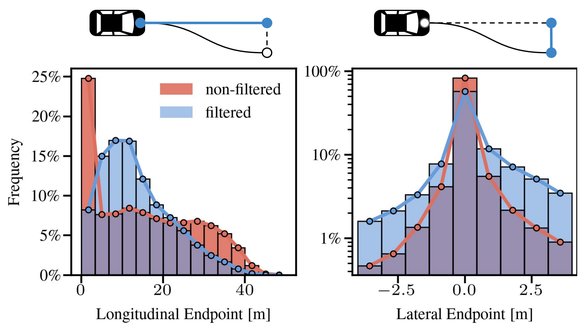
Our benchmark uses five simulation-based metrics…



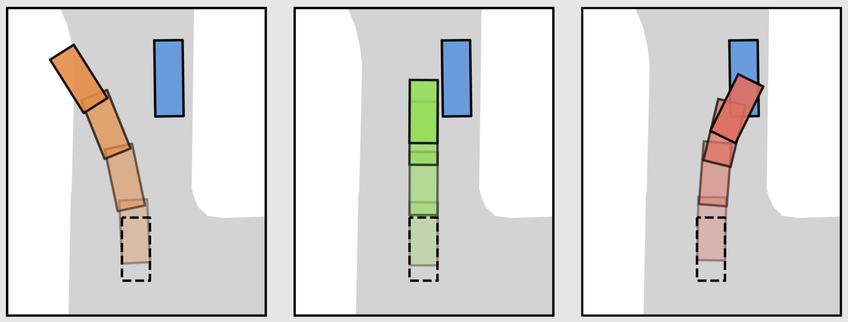
$$\text{PDMS} = \underbrace{\left(\prod_{m \in \{\text{NC,DAC}\}} \texttt{score}_m\right)}_{\text{penalties}} \times \underbrace{\left(\frac{\sum_{w \in \{\text{EP,TTC,C}\}} \texttt{weight}_w \times \texttt{score}_w}{\sum_{w \in \{\text{EP,TTC,C}\}} \texttt{weight}_w}\right)}_{\text{weighted average}}.$$

…summarized in the **Predictive Driver Model Score (PDMS)**.

We test agents on more diverse and challenging scenes.



Which trajectory is best?*



ADE: 2.3m | ADE: 1.1m | ADE: 1.0m

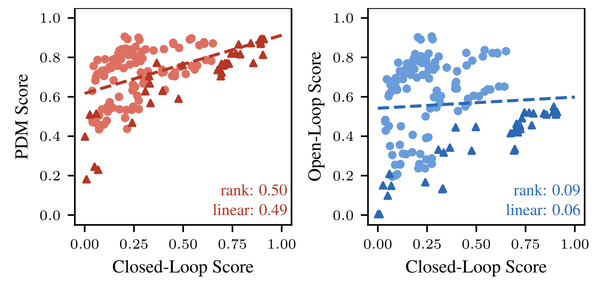
*Red has the lowest displacement error, but crashes.

Our benchmark measures what matters!

Does PDMS improve on displacement errors? Yes!

Metric ↑			
No Collision	1.0	1.0	**0.0**
Drivable Area Compl.	**0.0**	1.0	1.0
Time-to-Collision	1.0	1.0	0.0
Ego Progress	1.0	0.93	0.97
Comfort	0.0	1.0	1.0
PDMS	**0.0**	0.97	**0.0**

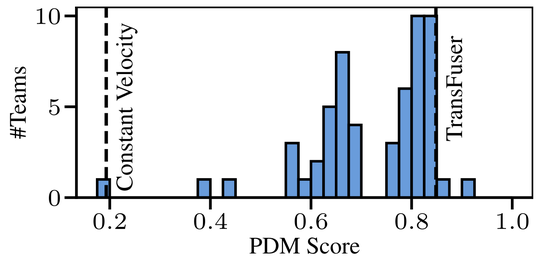
PDMS is better correlated to closed-loop testing.



Simple policies match recent large-scale models.

Method	Sensors	NC↑	DAC↑	TTC↑	EP↑	PDMS↑
UniAD	8 × Cams	98	92	93	79	83.8
PARA-Drive	8 × Cams	98	92	93	79	84.0
TransFuser	1 × Cam + LiDAR	98	93	93	79	84.0
LTF	1 × Cam	97	93	92	79	83.8
Human	-	*100*	*100*	*100*	*88*	*94.8*

143 Teams participated in our CVPR challenge.



Check out our GitHub & leaderboard page!