PURPOSE IN THE MACHINE: DO TRAFFIC SIMULATORS PRODUCE DISTRIBUTIONALLY EQUIVALENT OUTCOMES FOR REINFORCEMENT LEARNING APPLICATIONS?

Rex Chen
Kathleen M. Carley
Fei Fang
Norman Sadeh

School of Computer Science
Carnegie Mellon University
4665 Forbes Avenue
Pittsburgh, PA 15213, USA

ABSTRACT

Traffic simulators are used to generate data for learning in intelligent transportation systems (ITSs). A key question is to what extent their modelling assumptions affect the capabilities of ITSs to adapt to various scenarios when deployed in the real world. This work focuses on two simulators commonly used to train reinforcement learning (RL) agents for traffic applications, CityFlow and SUMO. A controlled virtual experiment varying driver behavior and simulation scale finds evidence against distributional equivalence in RL-relevant measures from these simulators, with the root mean squared error and KL divergence being significantly greater than 0 for all assessed measures. While granular real-world validation generally remains infeasible, these findings suggest that traffic simulators are not a deus ex machina for RL training: understanding the impacts of inter-simulator differences is necessary to train and deploy RL-based ITSs.

1 INTRODUCTION

Transportation efficiency is becoming an increasingly critical challenge due to continual growth in the volume of people and objects that need to be transported. The *2021 Urban Mobility Report* (Schrank et al. 2021) projected that, while the COVID-19 pandemic alleviated congestion, traffic levels in the US will quickly rebound in areas with expanding populations and job markets to produce the most rapid congestion growth since 1982. The increased traffic will stress existing infrastructure and result in social, economic, and environmental costs (Schrank et al. 2021), thus making the development and deployment of intelligent transportation systems (ITSs) a critical priority. At the same time, advances in computational algorithms and roadway infrastructure made in response to these challenges provide opportunities to enhance ITS learning. For example, novel traffic signal control technologies based on reinforcement learning (RL), which learn adaptive signaling policies from simulations generated using real-world traffic data, have already achieved performance on par with and even exceeding traditional control methods (Chen et al. 2020).

However, collecting data for ITS learning remains a nontrivial task. First, many state-of-the-art deep learning approaches for ITSs inherently require large quantities of data for training (Garg et al. 2019), but they also need to be able to learn from a heterogeneous set of experiences so that they can achieve robust performance in the real world. Second, once they are trained, ITSs will need to be able to interact with many human users, including both end-users (i.e., drivers) and decision-makers (i.e., traffic engineers and city planners). End-users have unique needs and make unique choices that must be addressed individually and equitably. Decision-makers also have complex and potentially conflicting requirements (Goh and Love 2012) that necessitate repeated iteration of the design of ITSs. Thus, data collection for ITS learning must be

large-scale and continuously-occurring. However, performing data collection from real-world transportation systems in this manner may result in significant efficiency and safety costs (Garg et al. 2019).

Due to these issues, traffic simulators are commonly used to substitute or complement real-world data collection. They serve as safe sandboxes that can generate realistic data for both ITS learning (Barceló et al. 2004) and real-world decision-making (Alvarez Lopez et al. 2018; Fries et al. 2007). This work narrowly focuses on open-source traffic simulators that have been developed for the training of ITSs based on RL methods, including systems for traffic signal control, autonomous driving, ramp metering, and tolling (Haydari and Yilmaz 2022). RL algorithms are particularly demanding in terms of the number of environmental interactions that they require, so traffic simulators for RL applications must be able to quickly generate high-fidelity, high-granularity data (Zhang et al. 2019).

Released in 2001, SUMO (Alvarez Lopez et al. 2018) is the traffic simulator that has been most commonly used to train deep RL-based ITSs (Noaeen et al. 2022). It supports an expansive framework for simulation definitions, as well as efficient programmatic API access. However, it is single-threaded and thus scales less well to very large road networks. Motivated by this limitation, Zhang et al. (2019) introduced CityFlow, a traffic simulator designed for applications of RL to traffic signal control. It is multi-threaded, and consequently was reported to achieve a speedup of >20x over SUMO. However, its framework for simulation definitions is more restricted and focuses on aspects which are essential for RL training. Its adoption is limited but increasing (Noaeen et al. 2022).

If traffic simulators are to serve as training environments for RL algorithms, their modelling assumptions must be sufficiently realistic that RL-based ITSs can learn to adapt to a variety of scenarios during deployment. Thus, real-world validation is crucial. However, granular validation with real-world data is usually not possible due to the aforementioned challenges of data collection. Comparisons between simulators take a partial step towards this goal by verifying that different simulators lead to equivalent outcomes. This work compares CityFlow against the more granular SUMO. Zhang et al. (2019) included a preliminary comparison of CityFlow to SUMO using the expected travel time of vehicles in a road network. However, travel time is a long-term, system-level outcome, whereas RL algorithms use instantaneous, local features for learning. A more comprehensive comparison was performed to answer the following research questions:

RQ1   Do the low-level simulation outcomes of CityFlow and SUMO have a statistically significant level of distributional equivalence?

RQ2   How is this distributional equivalence affected by incorporating variation between vehicles in terms of different driver behavioral models?

RQ3   How is this distributional equivalence affected by the scale of the simulation, in terms of traffic demand and road network size?

2   RELATED WORK

2.1 Validating Traffic Simulators

Validation is an important yet challenging aspect of the development of traffic simulators that ensures their fidelity to the real world. A multitude of road networks have been used to validate SUMO itself (Bedogni et al. 2015; Alvarez Lopez et al. 2018) and to calibrate its car-following models (Krajzewicz et al. 2002) in comparison to detector data. For instance, Lobo et al. (2020) compared traces for vehicle counts and crossing times between SUMO and detector data for a road network in Ingolstadt, Germany; their simulation is used in this work. Meanwhile, CityFlow was validated by comparing average travel times under various traffic volumes to that of SUMO (Zhang et al. 2019).

A number of studies have compared outcomes from multiple simulators; this work falls in this setting. Maciejewski (2010) compared vehicle counts under different traffic demand and driver behavior settings for SUMO, the commercial simulator VISSIM, and TRANSIMS. Several studies involved SUMO and the commercial simulator AIMSUN. Leksono and Andriyana (2012) compared travel times and queue lengths in these simulators under different traffic management interventions for a roundabout in Norrköping, Sweden.

Ronaldo and Ismail (2012) used t-tests to compare flows and speeds in these simulators to observations from a highway interchange in Stockholm, Sweden. Baza-Solares et al. (2022) applied t-tests to vehicle counts from the two simulators for a road network from Bucaramanga, Colombia. This work differs from prior approaches in that: (1) the measures used relate to the distribution of outcomes across the network, not just at individual points; and (2) these measures were evaluated across multiple road network scales.

2.2 Modelling Driver Behavior

There has been a significant body of literature on building realistic models of driver behavior. Car-following behavior was the first to be modelled, with the early Gazis-Herman-Rothery model being over 50 years old. Subsequent work has yielded optimal velocity, fuzzy logic, collision avoidance, action point, and cellular automaton models (Li and Sun 2012; Saifuzzaman and Zheng 2014). Collision avoidance and action point models were implemented for this work. Lane-changing behavior has been modelled by a newer, separate line of work (Moridpour et al. 2010; Zheng 2014), which has produced models based on rules, discrete choice, game theory, and cellular automata. Some work has built unified models of car-following, lane-changing, and other driver behavior (Toledo et al. 2007; Markkula et al. 2012). Here, car-following and lane-changing models were considered separately but co-varied in virtual experiments to elucidate the effects of their interactions. The most relevant prior work is Capela Dias et al. (2013), who analyzed the impact of driver behavior on system-level travel time in SUMO; by contrast, this work focuses on distributional equivalence between two simulators in terms of lower-level outcomes.

3 TRAFFIC SIMULATORS FOR RL

This section reviews the use of traffic simulators for RL-based ITS training.

3.1 RL for Transportation

RL algorithms learn policies to maximize their expected cumulative reward through repeated interactions with an environment, which for ITSs is the traffic simulator. Examples of cumulative reward objectives include travel time minimization for traffic signal control and speed limit control (Noaeen et al. 2022; Wu et al. 2020), and revenue maximization for tolling (Pandey et al. 2020). However, cumulative rewards are not useful for learning because they are system-level measures calculated over longer timespans. Instead, RL algorithms observe instantaneous vehicle-level state features, use them to select actions, and then receive instantaneous rewards (Haydari and Yilmaz 2022).

Formulations of the observation and instantaneous reward spaces vary between papers, but they are rarely aligned directly with the cumulative reward. Observations usually consist of counts (e.g., queue lengths), positions, speeds, or other aggregated properties of individual vehicles (Noaeen et al. 2022; Pandey et al. 2020; Wu et al. 2020). Rewards are more problem-dependent. For traffic signal control, queue length, waiting time, and speed are common factors (Koohy et al. 2022); speed limit control may use proxies for travel time, crash probability, or emissions (Wu et al. 2020); and tolling may optimize a combination of vehicle count, travel time on segments, and collected tolls (Pandey et al. 2020). Regardless of these different formulations, it is evident that system-level measures involving the movement of vehicles through the whole road network are not directly used for RL training.

3.2 Comparing CityFlow and SUMO

Microscopic traffic simulators are most often used for RL in transportation (Haydari and Yilmaz 2022). In contrast to macroscopic and mesoscopic simulators, microscopic simulators are primarily agent-based, as they model the behavior of individual vehicles (Barceló 2010). This is helpful for gathering the types of vehicle-level observations introduced in Section 3.1. Both CityFlow and SUMO are microscopic. They represent road networks as graphs, with intersections as nodes ("junctions" in SUMO) and roads as edges

between nodes. Vehicles are generated by flows; all vehicles within a flow share similar attributes, including behavioral parameters and routes. Routes are defined as sequences of edges. Where edges are joined at an intersection, vehicles select a pair of lanes ("roadlinks" in CityFlow, "connections" in SUMO) to traverse the intersection. However, key differences exist in finer details. Unlike CityFlow, SUMO supports definitions for more complex types of lanes (e.g., sidewalks, bicycle lanes, and ramps) and vehicle classes (e.g., public transit and emergency vehicles, pedestrians, and cyclists). It also supports stochastic generation of routes and vehicle types, and more sophisticated traffic signalling, including yellow lights.

Where possible, the virtual experiments in this work were designed to control or co-vary differences between the simulators. In particular, flows in both CityFlow and SUMO were converted to be single-vehicle and deterministic, and to share the same routes. However, differences remain, particularly aspects of driver behavior (Section 4) driven by randomness in SUMO; by contrast, CityFlow only uses random generation for vehicle priority. Such uncontrolled factors may account for their differences.

4 VARYING DRIVER BEHAVIOR

Addressing RQ2 requires distributional equivalence to be assessed under different variations of driver behavior; as established in Section 2.2, these are implemented through car-following and lane-changing models in traffic simulators. In both CityFlow and SUMO, car-following logic is used to maintain a safe gap to the leading vehicle, while lane changes are used to switch vehicles between lanes so that they can take appropriate roadlinks at intersections to continue their routes. However, in CityFlow, car-following and lane-changing are handled in separate threads; in SUMO, they are handled in sequence, with car-following logic being executed before lane-changing logic.

4.1 Car-Following Models

Car-following models determine the speeds at which vehicles travel unobstructed (free speed), follow behind a lead vehicle (following speed), and stop at an obstacle (stopping speed). Two types of numerical integration can be used to calculate vehicle speeds: a Euler update, which solves for the speed at discrete timesteps, and a ballistic update, which solves for the acceleration at discrete timesteps and applies it to the speed. Both simulators were controlled to use ballistic updates.

Table 1: Parameter settings for six aggressiveness types based on Capela Dias et al. (2013). Maximum emergency deceleration was uniformly set to $-9.0\mathrm{m/s^2}$, as they did not specify this parameter.

Type	Max. accel.	Max. decel.	Max. emerg. decel.	Min. gap	Min. headway
Aggressive young	3.1 $\mathrm{m/s^2}$	-5.5 $\mathrm{m/s^2}$	-9.0 $\mathrm{m/s^2}$	1.2 m	1.0 s
Courteous young	2.5 $\mathrm{m/s^2}$	-4.5 $\mathrm{m/s^2}$	-9.0 $\mathrm{m/s^2}$	2.5 m	1.0 s
Aggressive middle-aged	2.9 $\mathrm{m/s^2}$	-5.0 $\mathrm{m/s^2}$	-9.0 $\mathrm{m/s^2}$	2.0 m	1.3 s
Courteous middle-aged	2.4 $\mathrm{m/s^2}$	-4.1 $\mathrm{m/s^2}$	-9.0 $\mathrm{m/s^2}$	2.5 m	1.5 s
Aggressive old	2.6 $\mathrm{m/s^2}$	-4.5 $\mathrm{m/s^2}$	-9.0 $\mathrm{m/s^2}$	2.0 m	1.7 s
Courteous old	2.3 $\mathrm{m/s^2}$	-3.8 $\mathrm{m/s^2}$	-9.0 $\mathrm{m/s^2}$	2.5 m	1.9 s

Several shared parameters have varying effects on different car-following models. CityFlow assumes that vehicles have usual and maximum accelerations and decelerations; SUMO assumes that vehicles have a maximum possible acceleration and deceleration (which are used as the usual values), and a maximum emergency deceleration. The latter formulation was used here, although this may have resulted in more aggressive behavior than if lower usual accelerations and decelerations were used. Additionally, both simulators model vehicles as having minimum desired following distances in terms of space (i.e.,

the minimum gap) and time (i.e., the minimum headway). Capela Dias et al. (2013) co-varied these parameters to model the effect of driver aggressiveness on travel time. This work adopts their taxonomy of aggressiveness types, but excludes gender effects due to conflicting conclusions regarding their significance in the literature (Taieb-Maimon and Shinar 2001; Tasca 2000). This gave six parameter settings (Table 1).

The default car-following models in CityFlow and SUMO are both modified from the collision avoidance model of Krauß (1998). For free speed, CityFlow's implementation uses the maximum speed, while SUMO's implementation modulates this by the visible lookahead distance. For stopping speed, both simulators solve somewhat different quadratic equations to determine the deceleration needed to stop within a fixed distance. For following speed, a target distance is maintained, which is computed by the desired minimum headway as well as the speed and maximum deceleration of the lead vehicle. In this work, SUMO's variant of the Krauß model was added to CityFlow. SUMO implements several other car-following models, which were also re-implemented in CityFlow to introduce variation in driver behavior:

- The collision avoidance/action point model of Wagner (2008), which probabilistically combines Krauß's model with the action point model of Todosiev (1963).
- The collision avoidance model of Wiedemann (1974), which is a behavioral model that varies between free, approaching, following, and emergency modes based on the gap to the lead vehicle.
- The adaptive cruise control (ACC) model of Milanés and Shladover (2014), which determines acceleration using a "speed control" method if the gap to the lead vehicle is large and using a "gap control" method if the gap is small.

4.2 Lane-Changing Models

In both CityFlow and SUMO, lane changes are initiated by signaling the lead and lag vehicles on the destination lane, and the vehicle is instantly moved to the other lane upon completion of the lane change. (SUMO also supports a sublane model that explicitly models lateral movement, which was excluded as a control.) However, the implementation details of lane-changing likewise differ between these simulators. Vehicles in CityFlow only perform lane changes to follow their routes, and lane-changing is based on the explicit insertion of a copy of the vehicle (the "shadow vehicle") on the target lane that the lag vehicle will follow. Vehicles in SUMO follow a parameterized hierarchy of motivations, which includes strategic lane changes to follow routes as well as tactical lane changes for overtaking.

One of SUMO's parameters, the gap tolerance factor, was introduced to CityFlow in this work. When vehicles are changing lanes, the minimum gap on the destination lane required for a vehicle to initiate a lane change is given by the necessary gap for collision avoidance divided by this constant factor. A factor of 1.0 represents the default behavior. The virtual experiments in this work varied this factor between 0.5, 0.82, 1.0, 1.18, and 1.5, representing decreases/increases of 18 % and 50 % in tolerance. This is based on Sun and Kondyli (2010), who observed the mean gaps for forced, cooperative, and free lane change maneuvers to be 45 ft, 53 ft, and 109 ft in a video dataset.

5 EXPERIMENTAL SETUP

To address the research questions, two virtual experiments were designed. Both address RQ1 and RQ2 by assessing CityFlow and SUMO's distributional equivalence under different settings of driver behavior. The two experiments also address separate aspects of RQ3: Experiment 1 assesses distributional equivalence while varying the simulation scale by traffic demand, and Experiment 2 while varying it by road network size. This was accomplished using road networks from the benchmark dataset of Ault and Sharon (2021), as described in the following subsections. All of these networks were initially defined using SUMO syntax; the SUMO-to-CityFlow network converter of Zhang et al. (2019) was used to generate their CityFlow counterparts, and a flow converter was created to map between vehicular flows in the two simulators.

Four independent variables were considered in both virtual experiments: the car-following model (Section 4.1, 5 levels), the car-following aggressiveness parameters (Section 4.1, 6 levels), the lane-

changing gap tolerance (Section 4.2, 5 levels), and the road network (2 levels). For the car-following model, the default and SUMO-based implementations of the Krauß model in CityFlow were compared to Krauß model in SUMO, while the other models' CityFlow implementations were compared to their respective SUMO counterparts. The fundamental lane-changing models for the two simulators were held constant, as was the traffic signal program – a simple fixed-time program retained from the original dataset.

Eight instantiations of two types of measures were used to assess distributional equivalence. First, the root mean squared error (RMSE) quantifies the point-to-point difference in individual outcome measures. This was computed as the mean RMSE of the total travel time and waiting time (defined as time that a vehicle spends queued with a speed < 0.1 m/s) over all vehicles and timesteps; the mean RMSE of per-lane total vehicle counts and queued vehicle counts over all lanes and timesteps; and the mean RMSE of the speed and acceleration over all individual vehicles and timesteps. Second, the Kullback-Liebler (KL) divergence measures the difference in the distribution of outcomes over the entire road network. This was computed for the distributions of vehicle counts and queued vehicle counts as the mean over all timesteps.

5.1 Experiment 1: Traffic Demand

For this experiment, the road networks arterial4x4 (Ma and Wu 2020) and grid4x4 (Chen et al. 2020) (Figure 1) from Ault and Sharon (2021)'s RESCO benchmark were used. Both are synthetic grid networks with similar topologies, but they vary in the level of congestion. grid4x4 consists of six-lane roads with a uniformly distributed demand of 1,473 vehicles. arterial4x4 consists of major (four-lane) and minor (two-lane) roads, and a demand of 2,484 vehicles that alternates between major and minor roads. arterial4x4's traffic pattern leads to congestion and degraded RL performance in simulations (Ault and Sharon 2021).
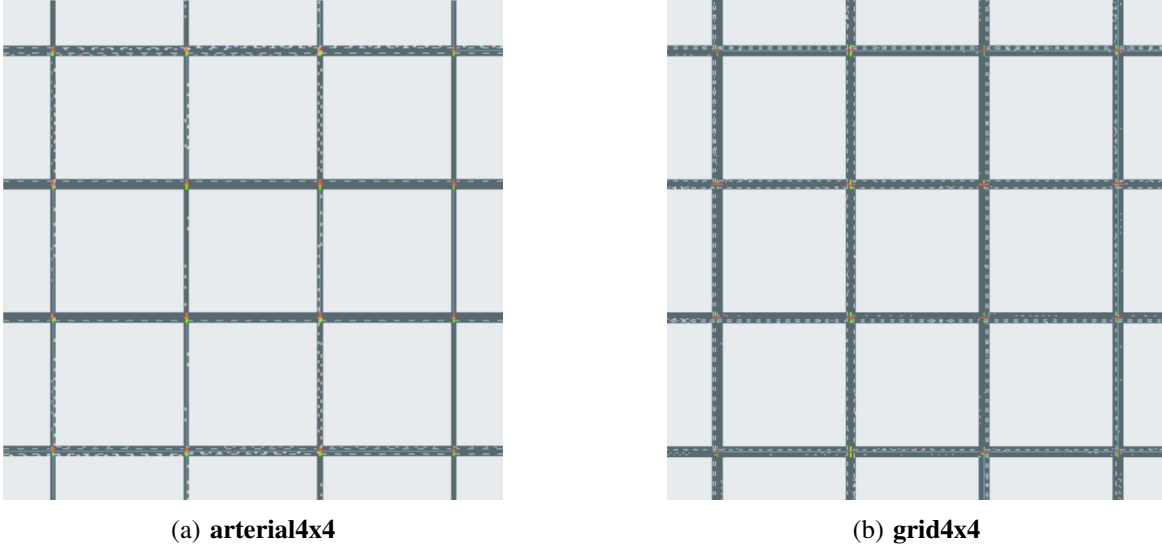

(a) **arterial4x4**　　　　　　　　　　　　　　　　(b) **grid4x4**

Figure 1: Screenshots in CityFlow of the arterial4x4 and grid4x4 road networks.

The following power analysis uses these variable names: *C* for car-following models, *A* for car-following aggressiveness, *L* for lane-changing gap tolerance, and *R* for road network. The second-order linear multiple regression included 5 (*C*) + 6 (*A*) + 1 (*L*) + 1 (L^2) + 2 (*R*) + 10 (*C·R*) + 12 (*A·R*) + 2 (*L·R*) + 30 (*C·A*) + 5 (*C·L*) + 6 (*A·L*) = 80 variables. Using G*Power 3.1.9.7's power calculation for ordinary linear multiple regression with a fixed model and R^2 increase, a small effect size of 0.02, and $\alpha = \beta = 0.95$, the total necessary sample size was computed as 2,646. Divided by the number of cells, $5 \cdot 6 \cdot 5 \cdot 2 = 300$, the number of replications per cell was computed as $\lceil \frac{2,646}{300} \rceil = 9$. All replications were executed using Python 3.9.16,

SUMO 1.12.0, and a modified version of CityFlow 0.1 on a shared server with four cores, two 4.2GHz Intel i7-7700K processors per core, and 62 GiB of RAM.

5.2 Experiment 2: Network Scale

For this experiment, the road networks ingolstadt1 and ingolstadt7 (Figure 2) from Ault and Sharon (2021)'s RESCO benchmark were used. Both are subsets of the Ingolstadt road network that was simulated by Lobo et al. (2020). They respectively contain 1 and 7 signalized intersections, representing a single busy intersection and a larger arterial road; ingolstadt7 is a superset of ingolstadt1. The total demands of the two road networks are respectively 1,716 and 3,031 vehicles.

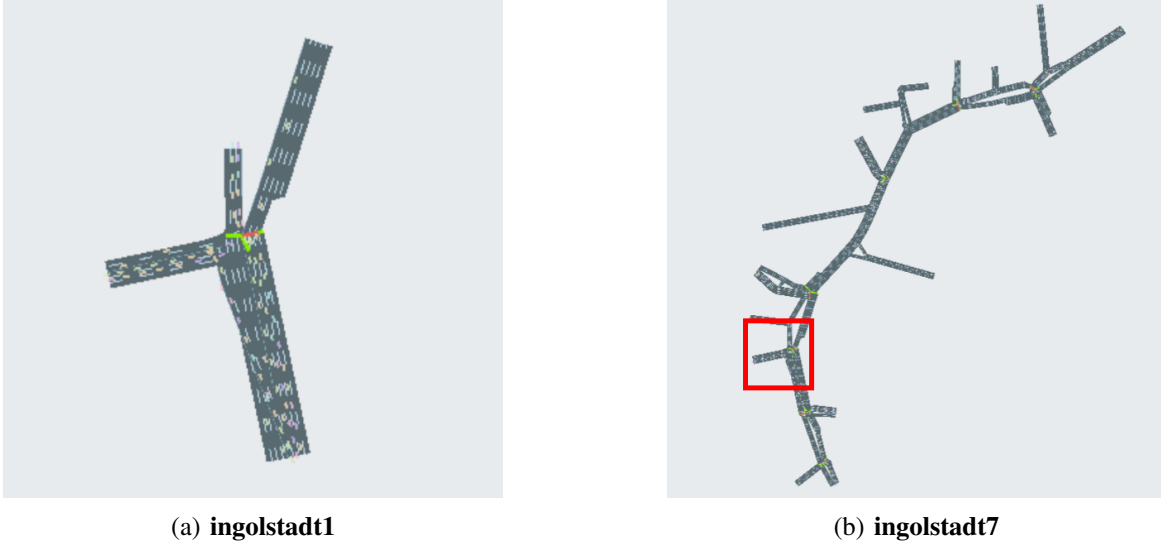

(a) **ingolstadt1**   (b) **ingolstadt7**

Figure 2: Screenshots in CityFlow of the ingolstadt1 and ingolstadt7 road networks.

Based on the analysis in Section 5.1, the number of replications per cell was computed as $\lceil \frac{2,646}{300} \rceil = 9$. All replications were executed using Python 3.9.16, SUMO 1.12.0, and a modified version of CityFlow 0.1 on a shared server with four cores, two 4.2 GHz Intel i7-7700K processors per core, and 62 GiB of RAM.

6 EXPERIMENTAL RESULTS

One-sample t-tests indicated that all of the RMSE and KL divergence measures were significantly different from 0, with a p-value $\ll 0.001$ for all cells in Experiments 1 and 2. This suggests a lack of distributional equivalence between CityFlow and SUMO. The following subsections explore the results of second-order linear multiple regression for each of the measures. Notably, coefficients involving the difference between the SUMO and CityFlow Krauß model implementations were generally not significant.

6.1 Experiment 1: Traffic Demand

For total time and waiting time RMSEs, the road network, car-following model, and aggressiveness generally had significant effects, as did various pairwise interactions between them. The uncongested grid4x4 network had significantly lower RMSEs (coefficients: $-890.61/-1035.46$) than the congested arterial4x4 network (intercepts: 1718.61/1815.46), suggesting that congestion worsened the distributional equivalence of these measures. In arterial4x4, the Wagner (coefficients relative to SUMO Krauß: 546.8/466.65) and Wiedemann models (coefficients: 100.96/208.99) had significantly higher RMSEs; these differences were smaller for grid4x4 (coefficients relative to SUMO Krauß: 49.5/33.75 for Wagner; 48.91/34.85 for Wiedemann).

For total and queued vehicle count RMSEs and KL divergences, the road network, car-following model, aggressiveness, and gap tolerance generally had significant effects, as did various pairwise interactions between them. The RMSEs and KL divergences showed distinct patterns: the RMSEs were much lower for grid4x4 than arterial4x4 (coefficients: $-5.15/-4.93$), but the KL divergences had less variation (coefficients: $-0.621/-0.013$). Yet, the KL divergences had low enough standard deviations that the road network's effects remained significant. High aggressiveness in arterial4x4 generally yielded higher RMSEs (coefficients for aggressive young relative to aggressive middle-aged: 2.16/2.46) and KL divergences (coefficients: 0.228/0.139). While the same was true for the RMSEs in grid4x4, its KL divergences were lower for more aggressive settings (coefficients: $-0.204/-0.122$). Despite higher time measures, the Wagner model led to lower measures for total vehicle count (coefficients relative to SUMO Krauß: $-1.16/-0.193$).

For vehicle speed and acceleration RMSEs, the road network, car-following model, and aggressiveness generally had significant effects, as did various pairwise interactions between them. Greater equivalence in vehicle distributions did not always correspond to more similar vehicle-level measures. Both speed and acceleration RMSEs increased for grid4x4 (coefficients: 2.62/0.418) even though the other measures were lower on average. Also, unlike its vehicle count measures but like its time measures, the Wagner model had significantly higher speed and acceleration RMSEs (coefficients relative to SUMO Krauß: 2.71/1.26).

6.2 Experiment 2: Network Scale

For total time and waiting time RMSEs, the aggressiveness and its interactions generally had significant effects, along with the Wiedemann model and its interactions. The best-fitting model for total time did not include a road network-gap tolerance interaction, whereas the model for waiting time did. The smaller ingolstadt1 network had lower but more variable RMSEs (intercepts: 899.29/1170.3), while the larger ingolstadt7 network had significantly higher but more uniform RMSEs (coefficients: 1264.59/1046.07). For ingolstadt1, the most aggressive parameter settings led to significantly higher RMSEs (coefficients of aggressive young relative to aggressive middle-aged: 408.36/568.65). Likewise, the Wiedemann model had significantly higher RMSEs than other car-following models in ingolstadt1 (coefficients relative to SUMO Krauß: 192.29/1917.31), but this effect was reversed for ingolstadt7 (coefficients: $-349.13/-464.13$).

For total and queued vehicle count RMSEs and KL divergences, the road network and aggressiveness generally had significant effects, as did various pairwise interactions between them and with the gap tolerance. Unlike Experiment 1, the RMSE measures were lower in ingolstadt7 than in ingolstadt1 (coefficients: $-2/-1.67$), but the KL divergence measures were higher (coefficients: 0.461/2.48). More aggressive parameter settings again led to significant increases in the RMSEs and KL divergences, with a larger increase in RMSEs (coefficients of aggressive young relative to aggressive middle-aged: 1.35/2.29) than in KL divergences (coefficients: 0.595/0.717). However, the increase in both measures was smaller in ingolstadt7 (coefficients: 0.764/1.507 for RMSE, 0.283/0.413 for KL divergences). Both the Wagner and ACC models had RMSEs that significantly increased with gap tolerance (coefficients relative to SUMO Krauß per unit of gap tolerance: 0.151/0.056 for Wagner, 0.165/0.092 for ACC), but KL divergences that significantly decreased with it (coefficients: $-0.285/-0.511$ for Wagner, $-0.354/-0.642$ for ACC).

For vehicle speed and acceleration RMSEs, the Wagner and Wiedemann models along with the aggressiveness had significant effects, as did various pairwise interactions of the road network with the car-following models and aggressiveness. Again, the road network had significant effects on both speed and acceleration RMSEs, with these measures being higher for ingolstadt7 (coefficients: 5.51/0.173). The Wagner car-following model had significantly higher RMSEs (coefficients relative to SUMO Krauß: 0.363/0.461), whereas the Wiedemann model had significantly lower RMSEs (coefficients: $-1.94/-0.957$).

6.3 Parameter Validity

Parameter validation was conducted by comparing the parameter settings for car-following and lane-changing models used in the virtual experiments to prior driving simulator and real-world studies. Overall,

the dependence of these parameters on external factors such as traffic density and speed suggests that the settings used in traffic simulators should be calibrated to specific road networks and conditions. However, the settings used in this work remain largely reasonable considering the variation reported in the literature.

For acceleration and deceleration, the settings used in the virtual experiments were based on Capela Dias et al. (2013), with smaller values for older and less aggressive drivers (Section 4.1); however, they considered gender effects to be negligible. Similar values have been reported in prior work (Nakagawa et al. 2006; Fadhloun et al. 2015; Körber et al. 2016). However, among driving simulator studies, Körber et al. (2016) demonstrated an age effect opposite to that assumed by Capela Dias et al. Among real-world studies, Moon and Yi (2008) found a dependence of the 95th percentile of braking decelerations on speed, and Nakagawa et al. (2006) reported a significant interaction between age and gender.

For minimum gap and headway time, the settings used in the virtual experiments also followed Capela Dias et al. (2013), with larger values for older and less aggressive drivers (Section 4.1). Similar values have again been reported in prior work (Michael et al. 2000; Vogel 2002). In particular, real-world studies for which reported values are influenced by a similar age effect include Delorme and Song (2001), Taieb-Maimon and Shinar (2001), and Moon and Yi (2008). Taieb-Maimon and Shinar (2001) and Shangguan et al. (2019) reported larger minimum gaps and smaller minimum headways, but their measurements were made for higher speeds (respectively over 50 $^{km}/_h$ and 100 $^{km}/_h$).

For lane-changing gap tolerance, this work used gap widths in empirical lane-changing behavior to approximately quantify its level of variation, specifically using the mean gaps for forced, cooperative, and free lane changes from Sun and Kondyli (2010)'s real-world study (Section 4.2). The settings in this work are consistent with standard deviations in lead and lag gaps as reported in prior studies (Moridpour et al. 2008; Balal et al. 2014). Ali et al. (2020)'s driving simulator study identified significant factors that impact gap tolerance: relative to the average male middle-aged driver, gaps are smaller for younger drivers, larger for female drivers, and smaller as speed increases. Likewise, Hill et al. (2015)'s real-world study reported that the mean and standard deviation of lag gaps depended on congestion. Future work could use these factors to create a taxonomy of lane-changing behavior similar to Capela Dias et al. (2013).

7 CONCLUSION

In this work, virtual experiments compared the low-level simulation outcomes of two traffic simulators, CityFlow and SUMO. To capture the effects of modelling real-world heterogeneity, various parameters of driver behavior and road network scale were varied. The results indicate a lack of distributional equivalence between the simulators, with certain parameter settings worsening distributional equivalence.

However, as noted in Section 3, this work is insufficient to provide a complete characterization of what the critical differences between these simulators are. Many aspects of CityFlow and SUMO that were not controlled – simulation control-flow, other aspects of driver behavior, the effects of traffic signals, and randomness – could all have contributed to the observed discrepancies. Thus, future work should perform more comprehensive, controlled evaluations of these two simulators.

Regardless, researchers in RL for transportation must not take traffic simulators for granted as a deus ex machina for training, and must recognize that they may not be interchangeable. Which simulator, then, should be chosen? This work does not aim to answer this question, but some observations can be made:

- **SUMO** provides a detailed simulation that models real-world heterogeneity, and captures additional aspects of traffic management and driver behavior
- **CityFlow** provides an efficient simulation that abstracts out and homogenizes various details, reducing the number of parameters that need to be tuned

Therefore, as with many other problems in simulation, the core trade-off between these two simulators (and others) involves veridicality and efficiency. There is no one best simulator; researchers must decide whether using a coarser abstraction of the environment is acceptable in exchange for faster training.

But how exactly should researchers make this decision? Crucially, RL-based ITSs may not necessarily perform better when they are trained with more granular simulators. Prior work has shown that both introducing unnecessary complexity (Zheng et al. 2019) and removing needed complexity (Genders and Razavi 2018) in the observation space may harm performance. Researchers should compare training results from different simulators and use them to design RL formulations in a principled way. As a first step, Goel et al. (2023) report training results on both SUMO and CityFlow, but they do not use the same road networks for both simulators. One strategy may be to train a baseline using an efficient simulator (e.g., CityFlow), and then to finetune it by further training with a veridical one (e.g., SUMO).

Ultimately, the true goal is to ensure that RL-based ITSs can perform well under real-world traffic conditions, which requires that their training and validation environments are close to reality. As previously noted, distributional equivalence between simulators is a proxy measure of this goal. Unfortunately, a chicken-and-egg problem exists in that traffic simulators are intended to replace real-world data collection, yet cannot be validated without it. For now, simulations should still be developed in collaboration with stakeholders to ensure that they meet acceptable standards of fidelity. However, the future holds promise for traffic simulations that are both veridical and efficient: the increasing prevalence of connected vehicles (Mathew et al. 2021) means that collection of granular real-world data for validation may be within reach.

ACKNOWLEDGMENTS

REFERENCES

Ali, Y., Z. Zheng, M. Haque, M. Yildirimoglu, and S. Washington. 2020. "Understanding the Discretionary Lane-changing Behaviour in the Connected Environment". *Accident Analysis & Prevention* 137:105463.

Alvarez Lopez, P., M. Behrisch, L. Bieker-Walz, J. Erdmann, Y.-P. Flötteröd, R. Hilbrich, L. Lücken, J. Rummel, P. Wagner, and E. Wießner. 2018. "Microscopic Traffic Simulation using SUMO". In *Proceedings of the 21st International Conference on Intelligent Transportation Systems*, ITSC '18, 2575–2582. Piscataway: Institute of Electrical and Electronics Engineers.

Ault, J., and G. Sharon. 2021. "Reinforcement Learning Benchmarks for Traffic Signal Control". In *Proceedings of the 35th Conference on Neural Information Processing Systems, Track on Datasets and Benchmarks*, NeurIPS '21, 1–11. Neural Information Processing Systems.

Balal, E., R. L. Cheu, T. Gyan-Sarkodie, and J. Miramontes. 2014. "Analysis of Discretionary Lane Changing Parameters on Freeways". *International Journal of Transportation Science and Technology* 3(3):277–296.

Barceló, J. 2010. "Models, Traffic Models, Simulation, and Traffic Simulation". In *Fundamentals of Traffic Simulation*, edited by J. Barceló, 1–62. New York: Springer.

Barceló, J., E. Codina, J. Casas, J. Ferrer, and D. García. 2004. "Microscopic Traffic Simulation: A Tool for the Design, Analysis and Evaluation of Intelligent Transport Systems". *Journal of Intelligent and Robotic Systems* 41:173–203.

Baza-Solares, N., R. Velasquez-Martínez, C. Torres-Bohórquez, Y. Martínez-Estupiñán, and C. Poliziani. 2022. "Traffic Simulation with Open-Source and Commercial Traffic Microsimulators: A Case Study". *Communications — Scientific Letters of the University of Zilina* 24(2):E49–E62.

Bedogni, L., M. Gramaglia, A. Vesco, M. Fiore, J. Härri, and F. Ferrero. 2015. "The Bologna Ringway Dataset: Improving Road Network Conversion in SUMO and Validating Urban Mobility via Navigation Services". *IEEE Transactions on Vehicular Technology* 64(12):5464–5476.

Capela Dias, J. A., P. Machado, and F. Câmara Pereira. 2013. "Simulating the Impact of Drivers' Personality on City Transit". In *Proceedings of the 13th World Conference on Transport Research*. July 15th–18th, Rio de Janeiro, Brasilia.

Chen, C., H. Wei, N. Xu, G. Zheng, M. Yang, Y. Xiong, K. Xu, and Z. Li. 2020. "Toward A Thousand Lights: Decentralized Deep Reinforcement Learning for Large-Scale Traffic Signal Control". In *Proceedings of the 34th AAAI Conference on Artificial Intelligence*, AAAI '20, 3414–3421. Washington, DC: Association for the Advancement of Artificial Intelligence.

Delorme, D., and B. Song. 2001. "Human Driver Model for SmartAHS". Technical Report UCB-ITS-PRR-2001-12, Institute of Transportation Studies, University of California, Berkeley, Berkeley.

Fadhloun, K., H. Rakha, A. Loulizi, and A. Abdelkef. 2015. "Vehicle Dynamics Model for Estimating Typical Vehicle Accelerations". *Transportation Research Record* 2491:61–71.

Fries, R., I. Inamdar, M. Chowdhury, K. Taaffe, and K. Ozbay. 2007. "Feasibility of Traffic Simulation for Decision Support in Real-Time Regional Traffic Management". *Transportation Research Record* 2035:169–176.

Garg, D., M. Chli, and G. Vogiatzis. 2019. "Traffic3D: A Rich 3D-Traffic Environment to Train Intelligent Agents". In *Proceedings of the 19th International Conference on Computational Science*, ICCS '19, 749–755. New York: Springer.

Genders, W., and S. Razavi. 2018. "Evaluating Reinforcement Learning State Representations for Adaptive Traffic Signal Control". *Procedia Computer Science* 130:26–33.

Goel, H., Y. Zhang, M. Damani, and G. Sartoretti. 2023. "SocialLight: Distributed Cooperation Learning towards Network-Wide Traffic Signal Control". In *Proceedings of the 22nd International Conference on Autonomous Agents and MultiAgent Systems*, AAMAS '23, 1–10. London: International Foundation for Autonomous Agents and Multiagent Systems.

Goh, Y. M., and P. E. Love. 2012. "Methodological Application of System Dynamics for Evaluating Traffic Safety Policy". *Safety Science* 50(7):1594–1605.

Haydari, A., and Y. Yilmaz. 2022. "Deep Reinforcement Learning for Intelligent Transportation Systems: A Survey". *IEEE Transactions on Intelligent Transportation Systems* 23(1):11–32.

Hill, C., L. Elefteriadou, and A. Kondyli. 2015. "Exploratory Analysis of Lane Changing on Freeways Based on Driver Behavior". *Journal of Transportation Engineering* 141(4):1–11.

Koohy, B., S. Stein, E. Gerding, and G. Manla. 2022. "Reward Function Design in Multi-Agent Reinforcement Learning for Traffic Signal Control". In *Proceedings of the 12th International Workshop on Agents in Traffic and Transportation*. July 23rd–25th, Vienna, Austria, 14–31.

Körber, M., C. Gold, D. Lechner, and K. Bengler. 2016. "The Influence of Age on the Take-over of Vehicle Control in Highly Automated Driving". *Transportation Research Part F: Traffic Psychology and Behaviour* 39:19–32.

Krajzewicz, D., G. Hertkorn, P. Wagner, and C. Rössel. 2002. "An Example of Microscopic Car Models Validation Using the Open Source Traffic Simulation SUMO". In *Proceedings of the 14th European Simulation Symposium*, ESS '02, 318–322. Caserta: European Council for Modelling and Simulation.

Krauß, S. 1998. *Microscopic Modeling of Traffic Flow: Investigation of Collision Free Vehicle Dynamics*. Ph. D. thesis, German Aerospace Center, Cologne.

Leksono, C. Y., and T. Andriyana. 2012. "Roundabout Microsimulation Using SUMO: A Case Study in Idrottsparken Roundabout, Norrköping, Sweden". Master's thesis, Linköping University, Linköping.

Li, Y., and D. Sun. 2012. "Microscopic Car-following Model for the Traffic Flow: The State of the Art". *Journal of Control Theory and Applications* 10:133–143.

Lobo, S. C., S. Neumeier, E. M. G. Fernandez, and C. Facchi. 2020. "InTAS – The Ingolstadt Traffic Scenario for SUMO". In *Proceedings of the 2020 SUMO User Conference*, SUMO '20, 1–20. Cologne: German Aerospace Center.

Ma, J., and F. Wu. 2020. "Feudal Multi-Agent Deep Reinforcement Learning for Traffic Signal Control". In *Proceedings of the 19th International Conference on Autonomous Agents and MultiAgent Systems*, AAMAS '20, 816–824. Auckland: International Foundation for Autonomous Agents and Multiagent Systems.

Maciejewski, M. 2010. "A Comparison of Microscopic Traffic Flow Simulation Systems for an Urban Area". *Transport Problems* 5(4):27–38.

Markkula, G., O. Benderius, K. Wolff, and M. Wahde. 2012. "A Review of Near-Collision Driver Behavior Models". *Human Factors* 54(6):1117–1143.

Mathew, J., J. Desai, R. S. Sakhare, W. Kim, H. Li, and D. M. Bullock. 2021. "Big Data Applications for Managing Roadways". *Institute of Transportation Engineers Journal* 91(2):28–35.

Michael, P. G., F. C. Leeming, and W. O. Dwyer. 2000. "Headway on Urban Streets: Observational Data and an Intervention to Decrease Tailgating". *Transportation Research Part F: Traffic Psychology and Behaviour* 3:55–64.

Milanés, V., and S. E. Shladover. 2014. "Modeling Cooperative and Autonomous Adaptive Cruise Control Dynamic Responses Using Experimental Data". *Transportation Research Part C: Emerging Technologies* 48:285–300.

Moon, S., and K. Yi. 2008. "Human Driving Data-based Design of a Vehicle Adaptive Cruise Control Algorithm". *Vehicle System Dynamics* 46(8):661–690.

Moridpour, S., M. Sarvi, and G. Rose. 2010. "Lane Changing Models: A Critical Review". *Transportation Letters* 2(3):157–173.

Moridpour, S., M. Sarvi, G. Rose, and E. Ramsay. 2008. "Variables Influencing Lane Changing Behaviour of Heavy Vehicles". In *Proceedings of the 31st Australasian Transport Research Forum*, ATRF '08, 1–15. Canberra: Australasian Transport Research Forum.

Nakagawa, S., D. Kriellaars, C. Blais, J. Montufar, and M. M. Porter. 2006. "Speed and Acceleration Patterns of Younger and Older Drivers". Technical report, University of Manitoba, Winnipeg.

Noaeen, M., A. Naik, L. Goodman, J. Crebo, T. Abrar, Z. S. H. Abad, A. L. Bazzan, and B. Far. 2022. "Reinforcement Learning in Urban Network Traffic Signal Control: A Systematic Literature Review". *Expert Systems with Applications* 199:116830.

Pandey, V., E. Wang, and S. D. Boyles. 2020. "Deep Reinforcement Learning Algorithm for Dynamic Pricing of Express Lanes with Multiple Access Locations". *Transportation Research Part C: Emerging Technologies* 119:102715.

Ronaldo, A., and M. T. Ismail. 2012. "Comparison of the two Micro-Simulation Software AIMSUN & SUMO for Highway Traffic Modelling". Master's thesis, Linköping University, Linköping.

Saifuzzaman, M., and Z. Zheng. 2014. "Incorporating Human-Factors in Car-following Models: A Review of Recent Developments and Research Needs". *Transportation Research Part C: Emerging Technologies* 48:379–403.

Schrank, D., L. Albert, B. Eisele, and T. Lomax. 2021. *2021 Urban Mobility Report*. Technical Report, Texas A&M Transportation Institute, College Station.

Shangguan, Q., X. Wang, S. Liu, and J. Wang. 2019. "Car Following Behavior under Foggy Conditions with Different Road Alignments – A Driving Simulator Based Study". In *Proceedings of the 5th International Conference on Transportation Information and Safety*, ICTS '19, 127–135. Piscataway: Institute of Electrical and Electronics Engineers.

Sun, D. J., and A. Kondyli. 2010. "Modeling Vehicle Interactions during Lane-Changing Behavior on Arterial Streets". *Computer-Aided Civil and Infrastructure Engineering* 25:557–571.

Taieb-Maimon, M., and D. Shinar. 2001. "Minimum and Comfortable Driving Headways: Reality Versus Perception". *Human Factors* 43(1):159–172.

Tasca, L. 2000. "A Review of the Literature on Aggressive Driving Research". Technical report, Ontario Ministry of Transportation, Toronto.

Todosiev, E. P. 1963. *The Action Point Model of the Vehicle-Driver System*. Ph. D. thesis, The Ohio State University, Columbus.

Toledo, T., H. N. Koutsopoulos, and M. Ben-Akiva. 2007. "Integrated Driving Behavior Modeling". *Transportation Research Part C: Emerging Technologies* 15(2):96–112.

Vogel, K. 2002. "What Characterizes a "Free Vehicle" in an Urban Area?". *Transportation Research Part F: Traffic Psychology and Behaviour* 5:15–29.

Wagner, P. 2008. "Action Point Models of Human Driving Behaviour". In *Proceedings of the 2008 Traffic Simulation Workshop*. June 30th–July 2nd, Graz, Austria.

Wiedemann, R. 1974. "Simulation des Straßenverkehrsflusses". *Publications of the Institute for Transportation, University of Karlsruhe* 8:1–42.

Wu, Y., H. Tan, L. Qin, and B. Ran. 2020. "Differential Variable Speed Limits Control for Freeway Recurrent Bottlenecks via Deep Actor-critic Algorithm". *Transportation Research Part C: Emerging Technologies* 117:102649.

Zhang, H., S. Feng, C. Liu, Y. Ding, Y. Zhu, Z. Zhou, W. Zhang, Y. Yu, H. Jin, and Z. Li. 2019. "CityFlow: A Multi-Agent Reinforcement Learning Environment for Large Scale City Traffic Scenario". In *Proceedings of the 2019 World Wide Web Conference*, WWW '19, 3620–3624. New York: Association for Computing Machinery.

Zheng, G., X. Zang, N. Xu, H. Wei, Z. Yu, V. Gayah, K. Xu, and Z. Li. 2019. "Diagnosing Reinforcement Learning for Traffic Signal Control". *arXiv preprint* 1905.04716:1–10.

Zheng, Z. 2014. "Recent Developments and Research Needs in Modeling Lane Changing". *Transportation Research Part B: Methodological* 60:16–32.

AUTHOR BIOGRAPHIES

REX CHEN (email rexc@cmu.edu) is a third-year PhD candidate in the Societal Computing program of the Software and Societal Systems Department in the School of Computer Science at Carnegie Mellon University. His research focuses on applications of reinforcement learning and other AI techniques to socially impactful domains such as transportation. His work aims to evaluate the impact of real-world uncertainty and user behavior on AI algorithms designed for these domains.

KATHLEEN M. CARLEY (email kathleen.carley@cs.cmu.edu) (H.D. University of Zurich, Ph.D. Harvard, S.B. MIT) is a Professor of Societal Computing, in the School of Computer Science's Software and Societal Systems Department at Carnegie Mellon University; and the Director of the Center for Computational Analysis of Social and Organizational Systems (CASOS), and the Center for Informed Democracy and Social Cybersecurity (IDeaS). She and her teams have developed infrastructure tools for analyzing large-scale dynamic networks, bot hunters, and various agent-based simulation systems.

FEI FANG (email feifang@cmu.edu) is the Leonardo Assistant Professor in the Software and Societal Systems Department in the School of Computer Science at Carnegie Mellon University. Her research interests lie in the area of artificial intelligence and multi-agent systems, focusing on the integration of computational game theory and machine learning to address real-world challenges in critical domains such as security, sustainability, and mobility.

NORMAN SADEH (email sadeh@cs.cmu.edu) is a professor in the School of Computer Science at CMU, where he is affiliated with the Software and Societal Systems Department, the Human-Computer Interaction Institute, and the CyLab Security and Privacy Institute. His research interests span mobile computing, the IoT, cybersecurity, privacy, machine learning, AI, and related public policy issues. His past work includes deployed planning and scheduling technologies for commercial systems.