Learning Eco-driving Strategies that Generalize

Vindula Jayawardana[a*], Baptiste Freydt[b], Ao Qu[a], Cameron Hickert[a], Edgar Sanchez[a], Catherine Tang[a], Mark Taylor[c], Blaine Leonard[c] and Cathy Wu[a]

[a] Massachusetts Institute of Technology, Cambridge, USA
{vindula, qua, chickert, edgarrs, cattang cathywu}@mit.edu
[b] ETH Zürich, Zürich, Switzerland
bfreydt@student.ethz.ch
[c] Utah Department of Transportation, Utah, USA
bleonard@utah.gov, marktaylor@utah.gov
[*] Corresponding author

Extended abstract submitted for presentation at the Conference in Emerging Technologies in Transportation Systems (TRC-30)
September 02-03, 2024, Crete, Greece

April 29, 2024

Keywords: eco-driving, reinforcement learning, generalization, scenario modeling

1 INTRODUCTION

Recent years have seen a surge in interest in cooperative eco-driving at signalized intersections, driven by its potential to contribute to climate change mitigation goals Huang (2018). In cooperative eco-driving, a fraction of vehicles in the fleet are coordinately controlled to reduce the fleet level emission. The rise of autonomous vehicles opens up new avenues for implementing these technologies. Various studies have introduced methods for developing eco-driving controllers, spanning from heuristic approaches to model-based and model-free techniques.

Optimizing eco-driving at the traffic network scale can yield the highest benefits, but it is a significantly hard problem. Network decompositions are often performed to make the problem tractable. Many works then study eco-driving at a single-intersection or two-intersection scope Mintsis (2020). Such decomposition strategies remain reasonable provided intersections and traffic scenarios are modeled with consideration of eco-driving factors, regulated intersection throughputs to avoid overflow and bottleneck spillback, and the distribution of signalized intersections is used in evaluation to prevent method overfitting to specific intersections.

However, many recent studies that use network decomposition overlook these desired requirements Xu *et al.* (2021). They often focus on devising methods at a few select intersections. An alternative is to test eco-driving in the real world, but this is unlikely to scale. Consequently, the devised methods risk failing to generalize across diverse intersections, and we may fail to observe the (prospective) impact of eco-driving when executed at the network scale.

In this work, considering these requirements, we aim to leverage network decomposition to devise eco-driving controllers that generalize. We then analyze insights that are derived from applying them at diverse signalized intersections. Accordingly, we first model three major US cities (Atlanta, Los Angeles, and San Francisco), each at the intersection scope capturing major factors that affect eco-driving benefits. In total, we model 6000 signalized intersections in simulations. Then, we use model-free multi-task deep reinforcement learning to learn eco-driving strategies that generalize. Last, while ensuring the controlled throughput at intersections, we evaluate the benefits of learned policies.

We find that our eco-driving policies result in, on average, up to 14% emission benefits. We further find interesting insights that were previously overlooked in smaller-scale analyses. First, the impact of eco-driving factors (e.g., lane length and vehicle inflow) changes with the eco-driving adoption level. Second, 70% of total emission benefits come from 20% of intersections at every adoption level. However, the 20% of intersections that yield the 70% benefits change with eco-driving adoption, calling for further research on how to deploy eco-driving gradually.

2 METHODOLOGY

Our method involves two main steps. First, we model traffic scenarios. Each traffic scenario stems as a permutation of a set of factors (e.g., intersection topology, temperature, humidity, etc.) that affect eco-driving benefits. We identify 33 influential eco-driving factors based on previous literature. Each traffic scenario is then modeled within a high-fidelity agent-based traffic simulator, replicating real-world intersections and calibrating for real-world conditions. Due to the limited space, we omit the details and provide a summary in Figure 1.

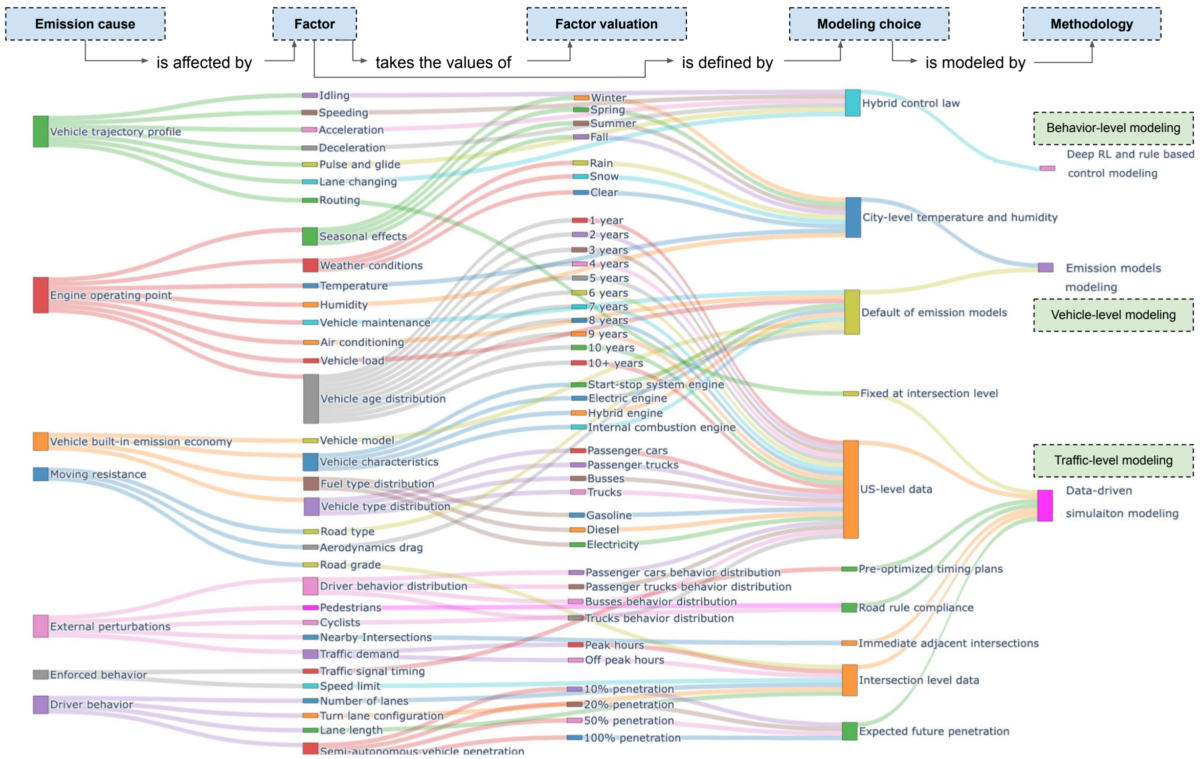

Figure 1 – ***A visual illustration of eco-driving factors and how they affect emissions modeling.*** *The legend on the top indicates how each column is connected.* ***Emission cause:*** *Factors stem from different emission causes (only a known predominant cause is illustrated).* ***Factor****: around 33 major factors affect vehicle emissions.* ***Factor valuations****: A list of factor values we consider as each factor can take multiple values.* ***Modeling choice****: Indicates what knowledge or data sources inform each factor/factor value.* ***Methodology****: Indicates how we translate the modeling choice into our simulation-based analysis.*

We use three levels of modeling of eco-driving factors. For vehicle-level, we capture the impact of each factor by making them input to known vehicle emission models Sanchez *et al.* (2022). In traffic-level modeling, we calibrate micro-simulations to reflect the effects of the factors. Last, at behavior-level, we use deep reinforcement learning, as discussed next.

In behavior-level modeling, we look at learning eco-driving strategies that generalize across scenarios. Each modeled traffic scenario originates as a multi-vehicle control problem and can be formulated as a Decentralized Partially Observable Markov Decision Process (Dec-POMDP). Then, the eco-driving problem can be formalized as a Contextual Markov Decision Process, which manifests as the collection of these Dec-POMDPs. We seek a policy that can solve the eco-driving CMDP using multi-task deep reinforcement learning.

We design a custom multi-task training scheme. First, we partition the scenario space into 28 partitions based on their scenario similarity. For each partition, we train a separate multi-task deep reinforcement learning policy to solve the scenarios in that partition. During evaluations, we also evaluate the zero-shot transfer performance of learned policies across partitions.

In every scenario, each eco-driving vehicle assesses features related to its own status, surrounding vehicles in the ego lane and adjacent lanes, and Signal Phase and Timing (SPaT) messages from the traffic signal. The learned policy governs the vehicle's acceleration, and to adapt to diverse environments, a context vector representing the intersection configuration is also appended to each vehicle observation. We formulate the reward function as $r_t^i = v_t^i - \alpha 1_{v_t^i < \tau} - \beta e_t^i$, where r_t^i, v_t^i, and e_t^i denote the reward, velocity, and CO_2 emissions of vehicle i at time t. Here, α, β, and τ are hyperparameters. The term v_t^i preserves low travel time, the indicator function $1_{v_t^i < \tau}$ identifies vehicle idling, and the penalty on e_t^i encourages reduced emissions.

As discussed earlier, regulating intersection throughputs is a crucial aspect of network decomposition. One approach is to maintain the same throughput as status quo human driving with new eco-driving strategies. However, this poses a challenge, as it transforms the reinforcement learning problem into a constrained optimization problem. To address this, we control eco-driving vehicles on incoming approaches using our learned controllers. Upon leaving, we revert to typical human driving behaviors, preventing vehicles from exploiting rewards by exiting the intersection too quickly. Practically, this approach effectively regulates intersection throughput.

Last, while our learned policy controls the longitudinal accelerations of eco-driving vehicles, we do not restrict them from performing lane-changing maneuvers. We implement a hierarchical approach: first, a pre-defined rule-based lane-changing controller assesses lane-changing suitability for the current time step, and if needed, it takes control of both longitudinal and lateral movements of the vehicle. Otherwise, the default learned policy governs longitudinal movements.

3 RESULTS

We assess the generalizability of learned policies by evaluating them in each of the three cities. In analysis, we restrict improvements in intersection throughput in the range $0\% - 1\%$. That is, if learned policies do not meet this criterion, we default to the status quo human driving baseline. This prevents overestimation of benefits due to reduced throughput. Given the lengthy incoming approaches and non-saturation vehicle inflows in the cities under consideration, a minor 1% improvement in throughput is unlikely to trigger a spillback effect. This conservative approach prevents underestimation of benefits due to overly strict throughput constraints.

Figure 2a illustrates the annual average emission benefits in comparison to a calibrated Intelligent Driver Model as the status quo human driving baseline. Learned policies exhibit up to 14% benefits, showing their generalization capacity. Interestingly, we note non-linear benefit scaling, where even a 10% adoption of eco-driving could result in a significant total benefit. Figure 2b indicates that all these benefits are attained with less than a 0.6% increase in throughput.

In Figure 2c, we look at the correlation between controllable eco-driving factors and emission benefits. Through Pearson correlation analysis, we find that at lower eco-driving adoption, vehicle inflow, lane count, and traffic signal time affect the benefits the most. As adoption increases, lane length and vehicle inflows become more influential. This highlights the need for further research to optimize intersection compatibility for eco-driving.

To further analyze the benefit dynamics and the role of different intersections, in Figure 2d, we show a Pareto chart for Los Angeles intersections revealing that 70% of emission benefits can be achieved by implementing eco-driving in just 20% of intersections at every adoption level.

However, the Venn diagram in Figure 2e illustrates the specific 20% of intersections delivering the 70% emission benefits vary with adoption level. Intersections effective at lower eco-driving adoption may not maintain the same efficacy as the adoption increases, further emphasizing the need for careful deployment planning.

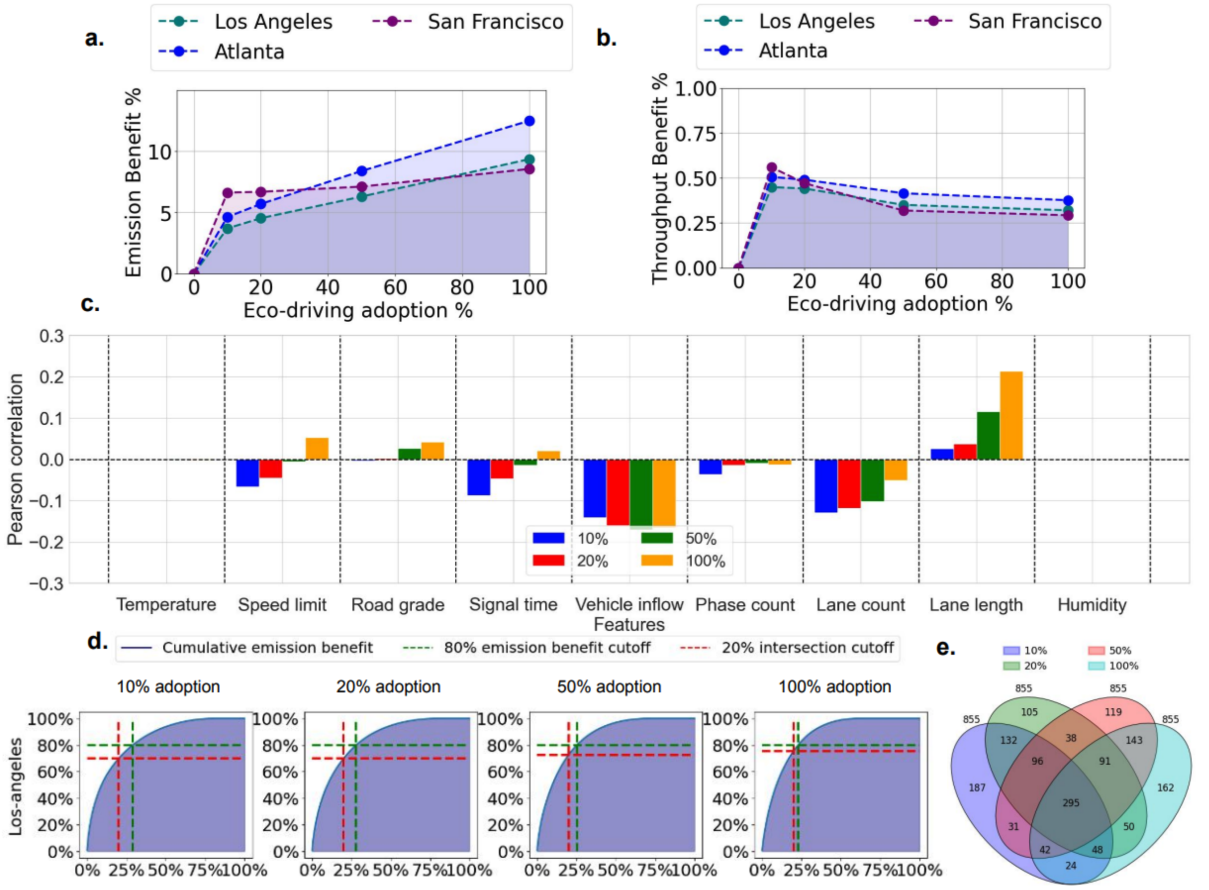

Figure 2 – ***a-b*** *Annual average emission and throughput benefits.* ***c****. Correlation analysis of how eco-driving factors affect benefits.* ***d****. Pareto charts showing emission reduction across Los Angeles intersections.* ***e****. Venn diagram that shows the distribution of 20% of intersections that yield 70% of emission under each adoption level.*

4 DISCUSSION

In this work, we devised eco-driving controllers that generalize using network decomposition as a modeling framework. Our evaluation across three major US cities unveiled novel insights previously overlooked in eco-driving research. Future efforts will expand this analysis to more cities, considering their unique intersection distributions to uncover additional potential learning challenges and resultant emission benefits.

References

Huang, Yuhan et al. 2018. Eco-driving technology for sustainable road transport: A review. *Renewable and Sustainable Energy Reviews*, **93**, 596–609.

Mintsis, Evangelos et al. 2020. Dynamic eco-driving near signalized intersections: Systematic review and future research directions. *Journal of Transportation Engineering, Part A: Systems*, **146**(4), 04020018.

Sanchez, Edgar Ramirez, Tang, Catherine, Jayawardana, Vindula, & Wu, Cathy. 2022. Learning surrogates for diverse emission models. *Tackling Climate Change with Machine Learning, NeurIPS*.

Xu, Nan, Li, Xiaohan, Liu, Qiao, & Zhao, Di. 2021. An overview of eco-driving theory, capability evaluation, and training applications. *Sensors*, **21**(19), 6547.