Learning Surrogates for Diverse Emission Models

Edgar Ramirez Sanchez[*], **Catherine Tang**[*], **Vindula Jayawardana, Cathy Wu**
Massachusetts Institute of Technology
{edgarrs, cattang, vindula, cathywu}@mit.edu

Abstract

Transportation plays a major role in global CO_2 emission levels, a factor that directly connects with climate change. Roadway interventions that reduce CO_2 emission levels have thus become a timely requirement. An integral need in assessing the impact of such roadway interventions is access to industry-standard programmatic and instantaneous emission models with various emission conditions such as fuel types, vehicle types, cities of interest, etc. However, currently, there is a lack of well-calibrated emission models with all these properties. Addressing these limitations, this paper presents 1100 programmatic and instantaneous vehicular CO_2 emission models with varying fuel types, vehicle types, road grades, vehicle ages, and cities of interest. We hope the presented emission models will facilitate future research in tackling transportation-related climate impact. The released version of the emission models can be found here.[*]

1 Introduction

Global CO_2 emission levels are steadily increasing and are considered a primary driving factor of climate change. Out of all sectors transportation is responsible for the largest share with 37% [1]. Efforts have been taken to reduce transportation-related emission levels including better roadways [2], optimizing traffic signals [7], electrical vehicles [20], among others. Emerging technologies like connected autonomous vehicles (CAVs) provide a substantial opportunity in optimizing driving patterns [10, 14, 6] and traffic flow [21, 17], which, in turn, can significantly reduce CO_2 emission.

An integral requirement in assessing the impact of these interventions is the access to calibrated CO_2 emission models. Emerging learning-based interventions like eco-driving [6], inherently require programmatic and instantaneous emission models that can be used to define a reward function in training agents [6]. For an emission model to be programmatic, it should have an API or scripted-based queries and results can be returned quickly. In addition, instantaneous means being able to calculate the emissions for an action taken within a single time step. Similarly, modern control methods like trajectory optimizations [12] share similar requirements in formulating objective functions. In general, any roadway intervention that changes the underlying vehicular dynamics should be properly assessed for climate impact using carefully calibrated CO_2 emission models under diverse conditions. Industry-standard CO_2 emission models are neither instantaneous nor programmatic [16, 3] and often limited to specific types of vehicles, one fuel type, and make simplifying assumptions such as zero road grades [8, 4]. Such limitations significantly hinder the progress of research and could even advance the field in a vacuum due to the lack of realism and variety [5, 19].

In this paper, we address these limitations with the hope of facilitating future research in roadway interventions for tackling climate change. In particular, we introduce a spectrum of programmatic and instantaneous vehicular CO_2 emission models that contains 1100 individual CO_2 emission models

[*]These authors contributed equally to this work.
[*]To preserve the anonymity of the submission, an anonymized version of the released emission model data set is presented with this submission.

Tackling Climate Change with Machine Learning: workshop at NeurIPS 2022.

under varying fuel types, vehicle types, road grades, vehicle ages, and cities of interest. We use the Motor Vehicle Emission Simulator (MOVES) [16] by the US Environmental Protection Agency as our emission data source and produce CO_2 emission models by an isolation methodology, MOVES processing and function fitting. We validate the accuracy of our emission models by using MOVES as ground truth. We have publicly released our emission models as a dataset and hope future research in roadway interventions for tackling climate change will benefit from it.

2 Related Work

A summary of related work is presented in Table 1. All the related works lack at least one feature compared to ours. The HBEFA model is the closest feature-wise, but it is designed to work specifically with SUMO traffic simulator and is also known to be less accurate, especially when vehicles decelerate. MOVES, on the other hand, has a comprehensive feature set. However, it is neither instantaneous nor programmatic and has high computational overhead making it undesirable for modern learning-based roadway interventions. However, since it is calibrated and well-known, we use MOVES data to create equivalent yet lightweight surrogate emission models that preserve accuracy.

Table 1: A comparison of five well-known emission models with our models across 6 features. The '✓' indicates that the feature is present, an 'x' that the feature is absent, and a '$*$' when the feature is present but with less options than within our models.

Models \ Features	Fuel Variety	Road Grade	Vehicle Type	Vehicle Age	Instantaneous	Programmatic
MOVES [16]	✓	✓	✓	✓	x	x
FastSIM [3]	✓	✓	✓	✓	x	✓
HBEFA [8]	✓	✓	✓	✓	✓	x
PHEM [4]	*	✓	*	x	x	x
MOVESTAR [18]	x	x	*	x	✓	✓
Ours	✓	✓	✓	✓	✓	✓

3 Methodology

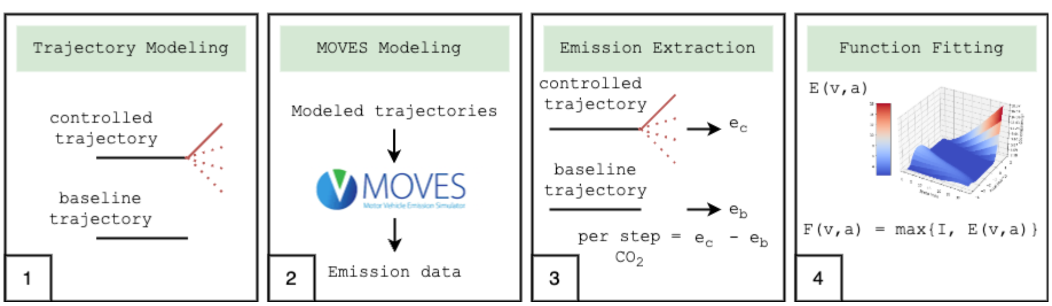
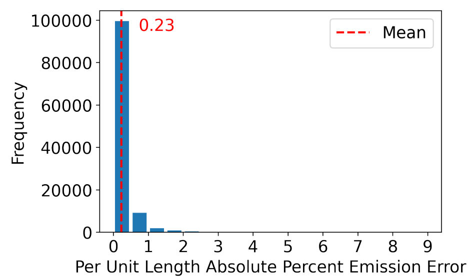


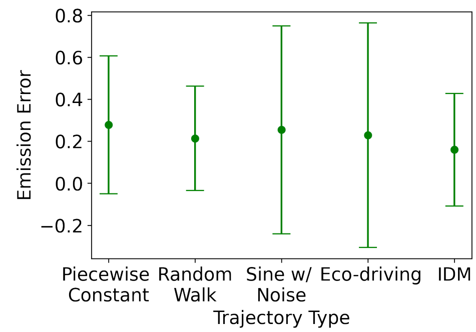
Figure 1: Overall methodology for learning surrogate emission models from MOVES

While being the industry standard for vehicle emission models, MOVES uses a retrospective approach that calculates the emission of a complete and already defined trajectory with a high computing overhead. Here, we present our method that translates the MOVES model to a fast, programmatic and instantaneous learning-based equivalent. Our overall method is depicted in Figure 1.

Based on a careful analysis of MOVES's internal computations (see Appendix for more details), we start our pipeline by constructing a collection of trajectories designed to extract emission data from MOVES. We define a trajectory $\tau = \{v_1 \dots v_n\}$ where v_i is the velocity of a vehicle at time step i. Next, a set of trajectory tuples of the form $< \tau_b, \tau_c >$ is generated with a baseline trajectory τ_b, and a controlled trajectory τ_c , of lengths n and $n + 1$ respectively. As in step 1 of Figure 1 illustrates,



(a) Distribution of emission error (e_i^τ) across all 1100 models evaluated on all 1132 trajectories



(b) Mean and standard deviation of emission error based on trajectory types

Figure 2: Emission error distribution: overall error in 2a and as grouped by trajectory type in 2b

the baseline τ_b is set to a constant speed, and the controlled trajectory τ_c appends an extra modeled action (i.e. an acceleration at a certain speed) maintaining $\tau_b[0:n] = \tau_c[0:n]$.

Next, every trajectory tuple $< \tau_b, \tau_c >$ along with all the generated input files that characterize the models' parameters (i.e. temperature, humidity, vehicle age distribution, fuel distribution and age distribution) are fed into MOVES (step 2 of Figure 1) to model per trajectory total CO_2 emissions. Let emission of controlled trajectory and baseline trajectory of a given trajectory tuple be e_c and e_b respectively. In step 3, the instantaneous emission is extracted by computing emission difference $\Delta e = e_c - e_b$. By repeating step 2 and 3 for all trajectory tuples, we create final emission dataset that conatins a mapping from speed and acceleration to instantaneous emissions.

Finally, we perform function fitting to generate our final emission models as in Step 4. Inspired by the Simulation of Urban Mobility [11] parameters and Lee et al.[9], we define a third order polynomial $E(a, v)$ with all possible terms of instantaneous velocity v and instantaneous acceleration a as our surrogate model for function fitting. Like other models [13], we bound the outputs of $E(v, a)$ to incorporate the idling (v = 0, a = 0) emission I and prevent negative emission from extrapolations or fitting errors. Equation 1 shows the structure of the final emission model.

$$F(v,a) = max\{I, c_0 + c_1 v + c_2 a + c_3 v^2 + c_4 va + c_5 a^2 + c_6 v^3 + c_7 (v^2) a + c_8 v(a^2) + c_9 a^3\} \quad (1)$$

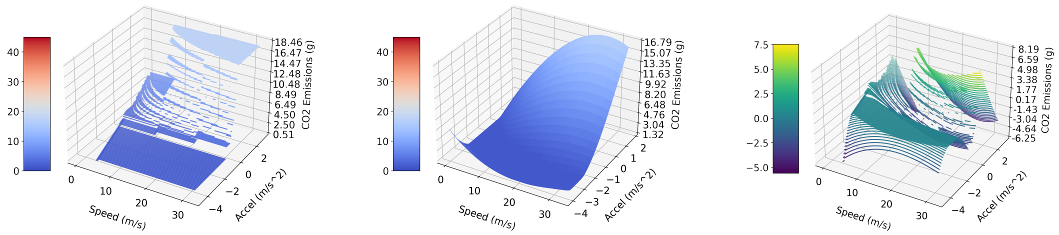
4 Validation

In this section, we evaluate the accuracy of the 1100 learned emission models by comparing the emission for vehicle trajectories against the ground truth from MOVES.

In evaluating emission models, we generate 1132 vehicle velocity trajectories T = $\{\tau_1, \ldots, \tau_{100}\}$ both randomly and as seen in previous works. In particular trajectories from 5 subsets: $T_1, T_2, ...T_5$ to capture a range of driving behaviors: A). random trajectories by sampling accelerations from a normal distribution $N(0,1)$, B). following sinusoidal generative model C). naturalistic human driving trajectories by simulating human drivers using Intelligent Driver Model (IDM) [15], D). eco-driving trajectories as seen in [6] and E) by following human compatible driving policies from [14].

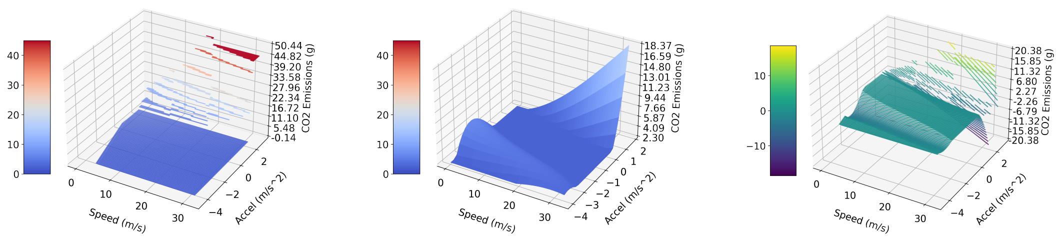
Given an emission model i and a vehicle trajectory τ, we define *per unit length absolute percent emission error* denoted by e_i^τ. We define e_i^τ as $e_i^\tau = \frac{|E_i^\tau - E_{MOVES}^\tau|}{E_{MOVES}^\tau \times \tau_l} * 100\%$, where E_j^τ and E_{MOVES}^τ are the total emissions resulting from trajectory τ measured based on emission model i and *MOVES* respectively. τ_l is the length of the trajectory τ measured in seconds. Given the set of vehicle trajectories T, *average emission error* e_{avg}^i of a given emission model i can therefore be defined as $e_{avg}^i = \mathbb{E}\{e\}_{\tau \sim T}$. In Figure 2a and Figure 2b, we analyze the error of our learned emission models.

As seen in Figure 2a, the distribution of emission errors is centered closely to 0 with a mean value of 0.2268% error/second. We further analyze this by comparing an emission model from the highest error bin and one from the lowest error bin in Figure 4b and Figure 3b respectively. These two fitted



(a) Ground truth emissions data (b) Fitted emission model (c) Function fitting error

Figure 3: An emission model with low error: 10 year-old Light Commercial Trucks on 0% road grade.



(a) Ground truth emissions data (b) Fitted emission model (c) Function fitting error

Figure 4: An emission model with high error: 6 year-old transit buses on -25% road grade

models show different shapes with one levelling off at large values of speed and acceleration. Worth noting, a large and middle section in Figure 4b was truncated to the lower bound: idle emission value, showing a fitting improvement opportunity that could be addressed with alternative fitting methods. Finally, the residuals don't show consistent trends across the different models.

The aggregated emission error per trajectory (Figure 2b) shows no statistically significant difference in the performance of the models across types of trajectories. This indicates that the models are able to behave consistently across different driving behaviors. The large standard deviation across all 5 categories indicates a high amount of variability for model performance within overall of model performance, independent of which trajectory type.

5 Conclusion

In this paper, we presented 1100 CO_2 emission models under varying fuel types, vehicle types, road grades, vehicle ages, and cities of interest. Our emission models are instantaneous and programmatic facilitating integration with state-of-the-art micro-simulators for impact assessments of roadway interventions in terms of climate impact. Future work of this study include extending the emission model datasets with more attributes such as weather conditions and seasonal variations. We hope the released emission models will facilitate impact assessments of future research in roadway interventions for tackling climate change.

References

[1] U.S. Energy Information Administration. Total energy, 2022. data retrieved from Environment (11.1-11.5), `https://www.eia.gov/totalenergy/data/monthly/`.

[2] Hessam AzariJafari, Jeremy Gregory, and Randolph Kirchain. Potential contribution of deflection-induced fuel consumption to u.s. greenhouse gas emissions. *Transp. Res. Rec.*, 2674(8):931–937, August 2020.

[3] A. Brooker, J. Gonder, L. Wang, E Wood, and et al. Fastsim: A model to estimate vehicle efficiency, cost, and performance, 2015.

[4] Karin Hirschmann, Michael Zallinger, Martin Fellendorf, and Stefan Hausberger. A new method to calculate emissions with simulated traffic conditions. In *13th International IEEE Conference on Intelligent Transportation Systems*, pages 33–38, 2010.

[5] Vindula Jayawaradana, Catherine Tang, Sirui Li, Dajiang Suo, and Cathy Wu. The impact of task underspecification in evaluating deep reinforcement learning. *Neural Information Processing Systems*, 2022.

[6] Vindula Jayawardana and Cathy Wu. Learning eco-driving strategies at signalized intersections. In *2022 European Control Conference (ECC)*, pages 383–390, 2022.

[7] Bálint Kővári, Lászlo Szőke, Tamás Bécsi, Szilárd Aradi, and Péter Gáspár. Traffic signal control via reinforcement learning for reducing global vehicle emission. *Sustainability*, 13(20):11254, October 2021.

[8] Daniel Krajzewicz, Michael Behrisch, Peter Wagner, Raphael Luz, and Mario Krumnow. Second generation of pollutant emission models for sumo. In Michael Behrisch and Melanie Weber, editors, *Modeling Mobility with Open Data*, pages 203–221, Cham, 2015. Springer International Publishing.

[9] Jonathan W Lee, George Gunter, Rabie Ramadan, Sulaiman Almatrudi, Paige Arnold, John Aquino, William Barbour, Rahul Bhadani, Joy Carpio, Fang-Chieh Chou, Marsalis Gibson, Xiaoqian Gong, Amaury Hayat, Nour Khoudari, Abdul Rahman Kreidieh, Maya Kumar, Nathan Lichtlé, Sean McQuade, Brian Nguyen, Megan Ross, Sydney Truong, Eugene Vinitsky, Yibo Zhao, Jonathan Sprinkle, Benedetto Piccoli, Alexandre M Bayen, Daniel B Work, and Benjamin Seibold. Integrated framework of vehicle dynamics, instabilities, energy models, and sparse flow smoothing controllers. In *Proceedings of the Workshop on Data-Driven and Intelligent Cyber-Physical Systems*, New York, NY, USA, May 2021. ACM.

[10] Nathan Lichtlé, Eugene Vinitsky, Matthew Nice, Benjamin Seibold, Dan Work, and Alexandre M. Bayen. Deploying traffic smoothing cruise controllers learned from trajectory data. In *2022 International Conference on Robotics and Automation (ICRA)*, pages 2884–2890, 2022.

[11] Pablo Alvarez Lopez, Michael Behrisch, Laura Bieker-Walz, Jakob Erdmann, Yun-Pang Flötteröd, Robert Hilbrich, Leonhard Lücken, Johannes Rummel, Peter Wagner, and Evamarie Wiessner. Microscopic traffic simulation using sumo. In *2018 21st International Conference on Intelligent Transportation Systems (ITSC)*, pages 2575–2582, 2018.

[12] Mehmet Fatih Ozkan and Yao Ma. Fuel-economical distributed model predictive control for heavy-duty truck platoon. In *2021 IEEE International Intelligent Transportation Systems Conference (ITSC)*, pages 1919–1926, 2021.

[13] Sangjun Park, Hesham Rakha, Kyoungho Ahn, and Kevin Moran. Virginia tech comprehensive power-based fuel consumption model (vt-cpfm): Model validation and calibration considerations. *International Journal of Transportation Science and Technology*, 2(4):317–336, 2013.

[14] Mayuri Sridhar and Cathy Wu. Piecewise constant policies for human-compatible congestion mitigation. In *2021 IEEE International Intelligent Transportation Systems Conference (ITSC)*, pages 2499–2505, 2021.

[15] Treiber, Hennecke, and Helbing. Congested traffic states in empirical observations and microscopic simulations. *Physical review. E*, 2000.

[16] USEPA. Motor Vehicle Emission Simulator: MOVES3, 8 2022.

[17] Eugene Vinitsky, Kanaad Parvate, Aboudy Kreidieh, Cathy Wu, and Alexandre Bayen. Lagrangian control through deep-rl: Applications to bottleneck decongestion. In *2018 21st International Conference on Intelligent Transportation Systems (ITSC)*, pages 759–765, 2018.

[18] Ziran Wang, Guoyuan Wu, and George Scora. Movestar: An open-source vehicle fuel and emission model based on usepa moves, 2020.

[19] Shimon Whiteson, Brian Tanner, Matthew E. Taylor, and Peter Stone. Protecting against evaluation overfitting in empirical reinforcement learning. In *2011 IEEE Symposium on Adaptive Dynamic Programming and Reinforcement Learning (ADPRL)*, pages 120–127, 2011.

[20] Maxwell Woody, Parth Vaishnav, Gregory A Keoleian, Robert De Kleine, Hyung Chul Kim, James E Anderson, and Timothy J Wallington. The role of pickup truck electrification in the decarbonization of light-duty vehicles. *Environmental Research Letters*, 17(3):034031, 2022.

[21] Cathy Wu, Aboudy Kreidieh, Kanaad Parvate, Eugene Vinitsky, and Alexandre M. Bayen. Flow: Architecture and benchmarking for reinforcement learning in traffic control. *IEEE Transactions on Robotics (T-RO)*, 2021.

Appendix

.1 Details on the Methodology

MOVES is the industry-standard emission model, provided, maintained and enforced by EPA's to be used for State Implementation Plans (SIP) and transportation conformity analyses by government offices [cite]. It's methodology and scope are the "state-of-the-science"[cite] and given that its results are the reference used for decision making, there's a natural interests in using it as the validated model to compute the emission of all sorts of transportation analyses. Unfortunately, ML-based approaches like eco-driving can't incorporate MOVES as part of their training pipeline

The first part of this work consisted in defining our space and determining an accurate way to get the data that we want of of MOVES. We explored MOVES computations, approaches, assumptions, and limitations through multiple sensitivity analyses, EPA's workshops, and communication with MOVES developers. Key findings include the linear scaling of emission with respect to volume, implying that MOVES doesn't model or account for congestion. Additionally, we found MOVES to be deterministic and that the region selection informs parameters like weather conditions and fleet characterization (i.e. Age, fuel, and type of vehicle distributions), which can all be overwritten manually to model a general case.

Out of the multiple input parameters that equate to an emission value, we chose speed and acceleration to be the model variables, and Vehicle type, fuel, vehicle age, weather conditions and road grade to be model parameters, for which we get a model per combination.

We generate, according to our space sampling, approach and resolution, millions of designed trajectories are modeled using MOVES, characterized by drive cycles (second-by-second vehicle speed) and scenario parameters. By using links (a segment of road defined by the user with specific properties) as independent experiments within a single run and MOVES's command line we are able us automate the process and run large batches of MOVES runs, making the data generation significantly faster but still in the order or weeks for a well-equipped personal computer.

In the third part of the pipeline lies the core component of our methodology, the extraction of the the instantaneous emission. In essence, we isolate the emission from a single second of a vehicle with a given speed and acceleration by comparing a designed trajectory to a baseline. Internally, MOVES uses finite differences to calculate the acceleration from a given driving cycle. Given that this calculation only looks one step backward, the emission from a unique second is independent to the shape of the driving cycle before or after, and can be extracted by modeling and analyzing a single step in the driving cycle. A number of MOVES considerations are taken in account to construct this approach.

For equating the difference, we need the baseline to be one second shorter than the controlled trajectory, however, the length of the driving cycle [$seconds$] can't be directly provided into MOVES. As a workaround we used the limit of the time window MOVES has for project scale analyses. Basically, we forced the driving cycle to be modeled as one hour and then scale back the emission to the desired length, considering the trajectory exactly one second more than the baseline. This scaling is possible because MOVES uses the driving cycle to compute a operating mode distribution, which is a percentage of time at each emission bin and then multiplies that by the total length of the driving cycle, which means it scales linearly with time and can be manipulated.

The final step to generate the emission models entails fitting the obtained results to a continuous function that can be directly used as a substitute for MOVES. We used a polynomial regression method from the machine learning python package: scikit learn, to fit two variable (speed and acceleration) degree three polynomials to the data. With a 80-20 split, we test the train and testing accuracy of each model.

.2 Choice of Machine Learning Model

For our fitting step, we fitted our data using polynomial regression with a third-order polynomial. We chose to use a polynomial function given the requirements of potential downstream applications we expect our model to be used for. Within control applications using reinforcement learning for example, it is important that the function is continuous and differentiable to determine the adjustments that need to be made to the behavior of an agent to achieve an optimal policy. Alternatives like decision trees or piecewise functions that capture the discretized nature of MOVES would not be compatible with downstream applications. On the other hand, the selection a lower 3rd order polynomial rather than other higher order polynomials is based on maintaining interpretability, and not incorporating unnecessary non-linearity and complexity to provide non-significant accuracy improvements: A 4, 5 and 6 order polynomials were tested, and even though the fitting accuracy increased by around 3x, 3.5x and 4x respectively, the validation performance didn't improve by more than 10 percent, showing overfitting.