Learning Surrogates for Diverse Vehicle Emission Models

Edgar Ramirez Sanchez
Catherine Tang
Vindula Jayawardana
Prof. Cathy Wu
Motivation

● Emission modeling for transportation is crucial

● MOVES as the industry standard
  ○ Provided and enforced by EPA
  ○ Used by transportation offices and policy makers across the US.
  ○ Very comprehensive

● Emerging transportation strategies from advances in interconnectivity, decision-making systems, ML
What is the issue with current models?

Lack of a model that is:

1. Diverse (breadth of inputs)
2. Instantaneous (second by second emissions)
3. Programmatic (ease of use and running time)

<table>
<thead>
<tr>
<th>Models</th>
<th>Features</th>
<th>Fuel Variety</th>
<th>Road Grade</th>
<th>Vehicle Type</th>
<th>Vehicle Age</th>
<th>Instantaneous</th>
<th>Programmatic</th>
</tr>
</thead>
<tbody>
<tr>
<td>MOVES [16]</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>FastSIM [3]</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>x</td>
<td>✓</td>
</tr>
<tr>
<td>HBEFA [8]</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>x</td>
</tr>
<tr>
<td>PHEM [4]</td>
<td>✓</td>
<td>✓</td>
<td>*</td>
<td>✓</td>
<td>x</td>
<td>✓</td>
<td>x</td>
</tr>
<tr>
<td>MOVESTAR [18]</td>
<td>x</td>
<td>x</td>
<td>*</td>
<td>✓</td>
<td>x</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td><strong>Ours</strong></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>
Goal

Application #1

\[ f(s, a) \]

Evaluation

Application #2

\[ f(s, a) \]

Learning

Application #3

\[ f(s, a) \]

Ecodriving

\[ CO_2 \text{ Emissions} \]

Vehicle age, vehicle type, road grade

\[ s_i, a_i \rightarrow f(s, a) \]

Time

Speed

\[ a_i, s_i \]
Objective: Calculate CO2 Emission of a given trajectory

Driving behavior
- Links
- Driving cycles

Specifications:
- Vehicle Type and Age
- Fuel type

Environment:
- City
- Time of Year
- Hour of the Day

Output
- Emissions (mass/time)
MOVES is not suitable for some analyses

Impact Assessment tools

Roadway interventions [strategies/technologies]
- EV adoption

Emerging Behavioral-based
- Eco Driving
- Lagrangian Congestion mitigation

Infrastructure, technology, etc.
Why - Extend and improve MOVES-model use

1) MOVES hard to use for behavioral analyses: Default driving cycle, cumbersome program

2) Retrospective approach

3) Other instantaneous and programmatic approaches lack diversity

<table>
<thead>
<tr>
<th>Features</th>
<th>Fuel Variety</th>
<th>Road Grade</th>
<th>Vehicle Type</th>
<th>Vehicle Age</th>
<th>Instantaneous</th>
<th>Programmatic</th>
</tr>
</thead>
<tbody>
<tr>
<td>MOVES [16]</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>FastSIM [3]</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>HBEFA [8]</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>x</td>
</tr>
<tr>
<td>PHEM [4]</td>
<td>* ✓</td>
<td>✓</td>
<td>✓</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>MOVESTAR [18]</td>
<td>x</td>
<td>x</td>
<td>✓</td>
<td>x</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Ours</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

Contribution:

Diverse: fuel types, vehicle types, road grades, vehicle ages, and cities of interest.

Instantaneous: able to calculate the emissions for an action taken within a single time step

Lightway-programmatic: have an API or scripted-based queries and results can be returned quickly
Methodology: Reverse engineer MOVES

1. Trajectory Modeling
   - controlled trajectory
   - baseline trajectory

2. MOVES Modeling
   - Modeled trajectories
   - Emission data

3. Emission Extraction
   - controlled trajectory
   - baseline trajectory
   - \( T_{CO2} \)
   - \( B_{CO2} \)
   - per step = \( T_{CO2} - B_{CO2} \)
Fitting

- Fit a third order two-variable polynomial
- Utilized sklearn LinearRegression
- Also considered higher order polynomials and a neural network

<table>
<thead>
<tr>
<th>Speed (m/s)</th>
<th>Accel (m/s^2)</th>
<th>Emissions (g CO₂)</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>-4</td>
<td>1000.6</td>
</tr>
<tr>
<td>10</td>
<td>-2.5</td>
<td>1203.5</td>
</tr>
<tr>
<td>30</td>
<td>3.5</td>
<td>1788.2</td>
</tr>
</tbody>
</table>

\[ F(v, a) = \max \{ I, c_0 + c_1 v + c_2 a + c_3 v^2 + c_4 v a + c_5 a^2 + c_6 v^3 + c_7 (v^2) a + c_8 v(a^2) + c_9 a^3 \} \]
Figure 3: An emission model with low error: 10 year-old Light Commercial Trucks on 0% road grade.

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

1. Created ~1000 drive cycles

2. Calculated emissions of all 1000 drive cycles using MOVES (ground truth)

3. For each drive cycle, ran each inst. model on the drive cycle and summed up the second by second instantaneous emissions
Validation: Error

**Graphs:**
- Left graph: Error bars showing emission error for different trajectory types (Piecewise Constant, Random Walk, Sine w/ Noise, Eco-driving, IDM).
- Right graph: Frequency distribution of per unit length absolute percent emission error, with a red dashed line indicating the mean.
Outcomes

- Created a set of 1100 models encompassing:
  - 5 vehicle types
  - 20 vehicle ages
  - 11 road grades

- Models with second-by-second emission output: $f(s,a)$

- API - Programmatic format
  - A python script: choosing model + passing in a speed and acceleration value