INTRODUCTION

Recent advancements in probe-based vehicle data could provide a cost-effective solution to increase the observability of traffic volumes. There are numerous potential benefits to ubiquitous traffic volumes:

- Economic and energy assessment
- Estimate economic impact of congestion
- Traffic information for traveler and vehicle safety
- Traffic information for traveler and vehicle safety

Method

Recent advancements in probe-based vehicle data could provide a cost-effective solution to increase the observability of traffic volumes. There are numerous potential benefits to ubiquitous traffic volumes:

- Economic and energy assessment
- Estimate economic impact of congestion
- Traffic information for traveler and vehicle safety
- Traffic information for traveler and vehicle safety

Project Background

Probe-based data offers a cost-effective solution for increasing the observability of traffic volumes.

Weather Underground

- Temperature, precipitation, visibility, fog, rain, snow
- Probe counts (key ingredient) and average speed

This paper explores spatial transferability of volume estimation models from three states (Colorado, North Carolina, and Pennsylvania).

Ample ground truth traffic data is a prerequisite for developing robust volume estimation models.

Key Research Question

Are probe-based volume estimation models spatially transferable? If so, which factors ensure a good spatial transfer? What are the critical pitfalls to look for while spatially transferring volume estimation models?

METHODOLOGY

Model Specification

Average hourly volumes are modeled as a function of commercial probe vehicle data, weather information, and roadway characteristics. The models are designed to be chosen for the established spatial and accurate performance:

\[ f(x) = \sum_{i=1}^{n} a_i \text{feature}_i(x) + \text{loss function} \]

where \( a \) is a tree model and \( n \) is the set of the parameters for the \( k \)th tree model, in each iteration, XGBoost is trained in a forward stagewise manner to minimize the loss function specified as:

\[ L(y, \hat{y}) = \sum_{i=1}^{N} L(y_i, \hat{y}_i) \]

where \( L \) is the loss function, \( N \) is the number of training data points, \( y \) is the true value, and \( \Omega(h) \) is the regularization term.

Baseline Volume Estimation Models

The models use data from stations validated using the cross-validation technique, in each iteration, all but one continuous count station is used for training, the held-out station is used for validation. This is repeated for \( N \) stations.

Average hourly volumes are modeled as a function of commercial probe vehicle data, weather information, and roadway characteristics. The models are trained and validated using the cross-validation technique; in each iteration, all but one continuous count station is used for training. The held-out station is used for validation. This is repeated for \( N \) stations.

Model Training and Validation

The models are trained and validated using the cross-validation technique; in each iteration, all but one continuous count station is used for training. The held-out station is used for validation. This is repeated for \( N \) stations.

Table 1. Baseline Model Results.

<table>
<thead>
<tr>
<th>State</th>
<th>MAE (veh/hr)</th>
<th>EFFR (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Colorado</td>
<td>0.01</td>
<td>0.09</td>
</tr>
<tr>
<td>North Carolina</td>
<td>0.01</td>
<td>0.09</td>
</tr>
<tr>
<td>Pennsylvania</td>
<td>0.01</td>
<td>0.09</td>
</tr>
</tbody>
</table>

Figure 1. Modeling framework.

SPATIAL TRANSFERABILITY

Region-to-Region

Spatial transferability exercises were carried out by applying the model estimated in one state to predict hourly traffic volumes (on the same type of roadway segments) in a different state. The model performances between NC and PA spatial transfer is much better than that of CO-NC spatial transfer. Upper bounds of observed traffic volumes were the most impactful factor in poor transferability, moderate for PA transferability, and much for NC transferability.

CONCLUSIONS & FUTURE WORK

- When transferring a volume estimation model from one region to another, it is extremely important to have temporal consistency.
- Transferring a model across states with similar weather and traffic characteristics will result in a better spatial transfer.
- The model is designed to have the best model performance and is shown to overcome the bounding issues identified with spatial transfers.
- Future will consider a few identifying state typologies based on traffic volume patterns, weather information, and roadway characteristics to build meta-models.

REFERENCES


Figure 2. Distribution of (a) hourly road volumes and (b) penetration rates.

Figure 3. Distribution of (a) hourly road volumes and (b) penetration rates.

Figure 4. Observed vs. predicted volumes in the Denver baseline model.

Figure 5. Predicted versus observed hourly volumes for a typical week for the model trained on NC and tested on CO (Thus displayed are PA predicted volumes), and (b) the model trained on PA and tested on NC (Thus displayed are NC predicted volumes).

Figure 6. Predicted versus observed hourly volumes for a typical week for (a) the model trained on CO and tested on NC and (b) the model trained on CO and tested on PA (Thus displayed are PA predicted volumes).

Table 1. Results from spatial transferability exercises.

<table>
<thead>
<tr>
<th>Type of Model</th>
<th>Test Region</th>
<th>EMFR (%</th>
<th>MAE (veh/hr)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline Model (CO)</td>
<td>NC</td>
<td>0.78</td>
<td>2.9</td>
</tr>
<tr>
<td>Baseline Model (NC)</td>
<td>CO</td>
<td>0.81</td>
<td>2.0</td>
</tr>
<tr>
<td>PA Spatial Transferability</td>
<td>NC</td>
<td>0.78</td>
<td>2.9</td>
</tr>
<tr>
<td>PA Spatial Transferability</td>
<td>CO</td>
<td>0.81</td>
<td>2.0</td>
</tr>
</tbody>
</table>

Figure 7. Predicted versus observed hourly volumes for a typical week for the model trained on PA and tested on NC (Thus displayed are PA predicted volumes) and the meta model (PA+NC train) and test on NC. The displayed station and week are the same as Figure 5b.