SPR RESEARCH PROGRAM
SECOND-STAGE PROPOSAL SUMMARY

PROBLEM NUMBER AND TITLE
25-62: Building a Data Fusion Model to Develop Volume Estimates from Probe Data

PROBLEM SUMMARY
FHWA deployed the MIRE FDE program to ensure that state DOTs apply data-driven safety approaches on all public roads. This requires States to report on data that historically have NOT been collected; AADT being one. For ODOT to be prepared for the MIRE FDE requirements in 2026, accurate data and methods for traffic volume estimation is needed. ODOT staff are frequently asked by transportation data firms to purchase their data which include an estimate of AADT; but we have not purchased these data due to cost and low trust in their quality. A potential solution is for ODOT to construct our own traffic volume estimation toolkit. Using probe count and other data, accurate and transparent validation of traffic volume estimations can be achieved within ODOT, internally.

ODOT OBJECTIVES
This project would create a modeling framework and an accurate AADT traffic volume estimation toolkit that uses available probe data along with other important data, e.g. roadway, weather, population and employment for Oregon roads. This project would scope the technical feasibility of validating data fusion models, and also analyze the fiscal feasibility of maintaining AADT estimation models within ODOT. An agency operational capacity feasibility assessment would also be conducted.

BENEFITS
This project will provide a framework for estimating AADT that can be transparently validated for improved accuracy. In addition to establishing the technical feasibility of building accurate data fusion models, this project would analyze the fiscal feasibility of developing and maintaining probe-based AADT estimation models within ODOT. This would compare costs of traditional traffic monitoring approaches with newer approaches, while weighing differences in data quality, accuracy, and other potential sources of uncertainty, such as privacy regulation changes that could negatively affect probe data modeling, among others.

SCHEDULE, BUDGET AND AGENCY SUPPORT
Estimated Project Length: 24 months.
Estimated Project Budget: $395,000
ODOT Support: Chi Mai, Transportation System Analysis Engineer, chi.mai@odot.oregon.gov , Other ODOT staff in support of RAC review: Chris Melson, ODOT-TPAU Manager, Sylvan Hoover, ODOT Climate Office, Erik Havig, ODOT PDAD Manager, Don Crownover, ODOT-TCE, and Ali Jafarnejad, TMC-ODOT.

FOR MORE INFORMATION
For additional detail, please see the complete STAGE 2 RESEARCH PROBLEM STATEMENT online at:
https://www.oregon.gov/odot/Programs/ResearchDocuments/25-62.pdf

SPR RESEARCH PROGRAM
SECOND-STAGE PROBLEM STATEMENT
FY 2025

PROBLEM NUMBER AND TITLE
25-62: Building a Data Fusion Model to Develop Volume Estimates from Probe Data

RESEARCH PROBLEM STATEMENT
FHWA is deploying the MIRE-FDE program to ensure that each state DOT will apply data-driven safety approaches on ***all*** public roads. For ODOT to be prepared for these mandates by 2026, data sources and methods for estimating volumes on most Oregon roads will be necessary. Count data coverage is currently very poor on non-NHS roads. ODOT will be required to report on data that historically have NOT been collected; Annual Average Daily Traffic (AADT) being one. ODOT staff are frequently asked by data firms to purchase transportation data, including estimates of AADT derived from probe data, but ODOT has not purchased them due to their high cost, low trust in quality, and our inability to validate their accuracy. A second issue of these third-party data packages is that the methods employed are considered proprietary and are not fully disclosed to the purchasing agency. This issue is highly problematic because ODOT, as a public agency, must strive for a reasonable amount of transparency in critical, operational analytics, like those that produce traffic volume estimates. Not being able to disclose methods may sow distrust between ODOT and the public that decisions are likely to affect. A potential solution is for ODOT to construct its own traffic volume estimation toolkit. With available probe and other data, transparently validating our traffic volume estimations can be optimized by having agency staff and trusted partners develop open, transparent, and validated data fusion models for accurate traffic volume estimation.

RESEARCH OBJECTIVES
This research will develop: **1)** a comprehensive modeling framework for reliable traffic volume estimation, which utilizes probe data and other essential sources for improved accuracy; **2)** an in-depth analysis of the fiscal feasibility of developing and maintaining probe-based, AADT estimation models within ODOT. This aspect of the research would compare the costs of traditional traffic monitoring approaches used by ODOT with the cost of newer approaches, while weighing differences in data quality, accuracy, and other potential sources of uncertainty such as changes in privacy regulation that could eliminate the market for probe data. And **3)** an agency capacity assessment that develops a set of scalable implementation strategies for data fusion modeling of traffic volume estimations within a resource constrained agency, ODOT.

WORK TASKS, COST ESTIMATE AND DURATION
Task 1: Traffic Volume Estimation State-of-Practice [5 months]
The objective of this task is to review and document current practices in traffic volume estimation. The review will include current methods applied by other DOTs, including their application success, strengths, and limitations. The review will also include new and innovative approaches that have been applied to traffic volume estimation. Emphasis will be placed on applications that rely on probe data. The research team will also evaluate probe data packages that best fit ODOT's modeling framework needs. This task will identify gaps in the current state of practice regarding both data and method concerns related to accurate traffic volume estimation.

Task 2: Data Collection and Inventory [3 months]
The objective of this task is to collect and inventory ***all*** relevant data related to traffic volume prediction. This will include high-resolution probe data (e.g., speed, probe counts, etc.), ODOT traffic volume data, Oregon network data, roadway geometry, sociodemographic data, and other related data that may help estimate traffic volume. Findings from Task 1 will help guide the data collection process. Data integration will be evaluated for comprehensive fusion and model estimation comparison.

Task 3: Data Fusion and Descriptive Analysis [4 months]
This task will use the collected and inventoried data and apply data fusion techniques to (a) prepare the data for traffic volume estimation and (b) to test various traffic volume estimation models for comparative accuracy and precision later in this investigation. To integrate probe vehicle data with the ODOT road network, a network conflation process developed by the research team will be applied. After developing a data fusion methodology through a transparent and open communication process with research partners and ODOT, a descriptive analysis will be conducted. The goal of the descriptive analysis is to identify and summarize key trends in the probe, volume, network, sociodemographic, and other datasets. The descriptive analysis will identify key data characteristics to be considered for traffic volume estimation.

Task 4: Traffic Volume Estimation Models [4 months]
This task will take the fused data from Task 3 and apply a variety of estimation methods to estimate ODOT traffic volumes. The research team will apply traditional statistical and econometric methods, advanced statistical and econometric methods, and artificial intelligence (AI) models to determine the methodological approach that best estimates traffic volumes in Oregon. The research team will conduct in-depth investigation into volume estimation results and identify whether/how estimation accuracy varies by geographic areas and roadway classifications. This task will identify important information that should be included in traffic volume estimation methods to aid in future data collection efforts. All proposed methodologies will be applied transparently, and methods documented to be replicated by ODOT personnel.

Task 5: Feasibility of Developing and Maintaining a Probe-Based Traffic Volume Estimation Model [3 months]
Based on the results from Task 4, the research team will summarize and identify the feasibility of maintaining such approaches for ODOT from two perspectives. First, the research team will assess the feasibility of maintaining such models from a technical and methodological perspective (e.g., can the proposed methods be implemented year-to-year to produce traffic volume estimations based on toolkits developed through this research). Second, the research team will assess the fiscal feasibility by assessing the tradeoff between subscription-based probe data services and traditional traffic monitoring approaches. As part of this assessment, the research team will consider costs related to the proposed approach identified through this research, data quality, data reliability, and other potential sources of uncertainty. A full SWOT analysis will outline the pros and cons of each volume estimation model approach along with clearly explained recommendations and trade-offs (for each).

Task 6: Toolkit for Estimating Traffic Volume in Oregon [5 months]
This task will summarize and provide an overview of how to apply the proposed traffic volume estimation approach. This may include step-by-step tutorials or training materials developed by the research team. The goal of this task is to ensure the proposed method(s) are transparently reproducible by ODOT personnel to estimate traffic volumes for long-term and future operations.

Key Deliverables:

1) a comprehensive modeling framework for traffic volume estimation, which utilizes probe data and essential sources for improved accuracy for ODOT;
2) an in-depth analysis of the fiscal feasibility of developing and maintaining probe-based AADT estimation models for internal agency validation;
3) an agency capacity assessment that develops a set of scalable implementation strategies for data fusion modeling of traffic volume estimations within a resource constrained agency, ODOT.

Estimated Project Length: 24 months. | ***Estimated Project Budget:*** 395,000

IMPLEMENTATION

The plan for implementation into ODOT practical operations will be to use the developed ODOT toolkit to more accurately estimate traffic volumes through optimal modeling techniques, along with a comprehensive analysis of the fiscal feasibility in maintaining probe-based AADT estimation by comparing costs and benefits of various modeling approaches.

POTENTIAL BENEFITS

The potential benefits include addressing the persistent problem of relying on private third-party and/or proprietary products to estimate traffic volume by means of probe data and other sources of information. Though the techniques used by private third-party data firms has proven reliable in independent evaluations, challenges remain for state DOTs when using these products. The first challenge relates to the issue of ongoing evaluation of these traffic estimates since many of the third-party firms are able to access a state DOTs data directly from traffic count databases used by the DOTs (like MS2) or from public data sources equally accessible to the private firms. When private firms utilize these data they improve their products but then make it impossible for DOTs to independently validate the products because the same data used for evaluation was used in the model training and development. A second issue with using third-party data products is that the methods employed are considered proprietary and are not fully disclosed to the purchasing agency. This second issue is highly problematic because DOTs are public agencies and should strive for a reasonable amount of transparency for important analytic methods like those that produce traffic volume estimates. Not being able to disclose analytic methods could sow distrust between agencies and the public in which decisions are likely to impact.

PEOPLE

ODOT champion(s):

Chi Mai, P.E., Oregon Department of Transportation (Original S1 Submitter) Transportation System Analysis Engineer, Transportation Planning Analysis Unit (TPAU)

Problem Statement Contributors:

Chi Mai, P.E., Oregon Department of Transportation (Original Submitter)
Jason C. Anderson, Ph.D., Portland State University
Xu Zhang, Ph.D., Kentucky Transportation Center
Matt Bagwell, Ph.D., Oregon Department of Transportation

References have been omitted to meet written proposal 3-page limit.

STAFF REVIEW PAGE

Literature Check

TRID&RIP

☒ A review of TRID & RIP databases found no existing research that answers the research question.

Technology & Data assessment

☒ No Identified T&D output

☐ At the end of this project, the implementing unit(s) within ODOT will need to coordinate the adoption of new technology or data in order to realize the full potential of this research.

However, this proposal seeks to optimally validate and calibrate existing data and technologies currently in use by ODOT for continuous AADT Traffic Volume Estimation improvement to enhance system accuracy and reliability in our traffic modeling and ongoing communications with the public through systems, such as ODOT's Intelligent Transportation Systems (ITS), which at times communicates semi-autonomically with roadway users (drivers).

Cross-agency stakeholders

- List ODOT partners or impacted units: PDAD, TPAU, Traffic Counts, Climate, Equity, Communications and Information dissemination to the public (senior ODOT leadership, statewide implications).

- Identify any issues of concern raised by an ODOT partners. Note expected mitigation that addresses these concerns: The major concerns from ODOT staff were that if we do not examine these operational features, then we may not be ready for the MIRE-FDE requirements in 2026. Moreover, we have an agency need to ensure that we are communicating the optimally-most accurate (or reliable) information to the public for continuous operational improvement to promote an increased level of public trust in regard to traffic volume estimations that our Intelligent Transportation Systems rely on to inform the public on traffic conditions that are reported in real time.