Geospatial analysis and machine learning of the U.S. EPD database for asphalt mixtures to quantify strategies for carbon emission reductions.

Kun Zhang[1], Liangkai Zhang[2], You Zhou[3]*

[1]California State University, Chico, CA

2

3

*corresponding author:

Abstract

Keywords

1. Introduction

The development and use of environmental product declaration (EPD) of asphalt mixtures for pavements have been of increasing interest in recent years, which are driven by Green Public Procurement (GPP) (ref) and the passing of 'Buy Clean' legislation, such as Buy Clean California Act in 2017, Buy Clean Colorado Act in 2021, and Buy Clean Oregon in 2022 (ref.). The EPD is a Type III environmental label, which could provide a quantitative and transparent assessment of the environmental impacts of a product using the life cycle assessment (LCA)(ref.). Thus, EPDs can assist in communicating the environmental performance of asphalt mixtures and identifying effective measures to reduce carbon emissions (ref).

A few studies have used EPDs to communicate the cradle-to-gate LCA results of asphalt mixtures with different additives and assess various strategies to reduce environmental impacts.

The beneficial use of RAP to reduce environmental impacts based on LCA have been assessed by … Other studies also investigate the plant fuel type, hauling distance, and the use of bio-binder

Mattinzioli et al. (2022) performed 160 LCA simulations to investigate the influence of asphalt mixture classes, including dense-graded, gap-graded, open-graded, porous, and stone mastic, on environmental impact. It was found the dense-grade and stone mastic mixtures have higher impacts per lane-km than gap-graded mixtures, followed by porous mixture and open-graded mixtures due to mixture density and ingredients. They also performed a sensitivity analysis to study the impacts of the use of reclaimed asphalt pavement (RAP), aggregate haul distance, and the use of natural gas as plant heating fuel to substitute fuel oil on carbon emissions, which helps optimize mix design and plant operations to improve sustainability. Moretti et al. (2017) compared the environmental performance of hot mix asphalt (HMA) produced by two companies in Italy. The results highlighted the importance of fuel sources in the contribution of environmental impacts, as natural gas has a lower contribution to carbon emission compared with other fossil fuels. Mukherjee (2016) performed a sensitivity analysis of LCA to evaluate the beneficial use of RAP and/or recycled asphalt shingles (RAS) to replace virgin binders on the reduction in global warming potential (GWP) based on two mixtures. Mattinzioli (2021) performed LCA for five mixtures to assess the uses of RAP, crumb rubber, bio-binders, and lower production temperatures as mitigation strategies to reduce environmental impacts. Mantalovas and Mino (2020) compared the environmental and engineering performance of four asphalt mixtures with RAP contents varying from 0%, 30%, 60%, and 90% as case studies. The environmental sustainability and circularity indicator was developed and used to access cradle-to-gate environmental impacts of asphalt mixtures. It was found that asphalt mixtures with higher percentages of RAP had lower environmental impacts. Harvey et al. (2020) also performed case studies to quantify the use of RAP in asphalt mixtures as one of the effective strategies to reduce carbon emissions. It was found that the use of RAP at virgin binder replacement ratios of 20%, 35%, and 42% could reduce 0.7%, 5.2%, and 6.2% greenhouse gas (GHG) emissions, respectively. The use of bio-based rejuvenating agents in these RAP mixtures can further reduce environmental impacts. Gruber and Hofko (2023) reported mean values of GHG emissions for virgin aggregate (2.51±0.49 kg CO_2e/t), asphalt binder (365±142 kg CO_2e/t), and RAP (0.37±0.03 kg CO_2e/t) from various resources. Adding every 10% more RAP in an asphalt mixture could help save 1.3 kg CO_2e per ton of asphalt mixture. This study also highlighted the importance of reducing moisture content in aggregate on the reduction of GHG emissions.

Commented [KZ1]: Rangelov, M., Dylla, H., Mukherjee, A., & Sivaneswaran, N. (2021). Use of environmental product declarations (EPDs) of pavement materials in the United States of America (USA) to ensure environmental impact reductions. *Journal of Cleaner Production*, *283*, 124619.

Commented [KZ2]: Senseney, C. T., Harvey, J., Butt, A. A., & Meijer, J. (2023). Recommendations for cradle-to-gate environmental product declarations (EPD) in 'Buy Clean'procurement based on CDOT's experience. *Environmental Research: Infrastructure and Sustainability*, *3*(3), 035004.

Commented [KZ3]: Rangelov, M., Dylla, H., Mukherjee, A., & Sivaneswaran, N. (2021). Use of environmental product declarations (EPDs) of pavement materials in the United States of America (USA) to ensure environmental impact reductions. *Journal of Cleaner Production*, *283*, 124619.

Commented [KZ4]: Strömberg, L., Hintze, S., Al-Qadi, I. A. Q., & Okte, E. (2020). Assessment of asphalt concrete EPDs in Scandinavia and the United States

Commented [KZ5]: Samieadel, A., Schimmel, K., & Fini, E. H. (2018). Comparative life cycle assessment (LCA) of bio-modified binder and conventional aspha

Commented [KZ6]: Mattinzioli, T., Sol-Sanchez, M., Moreno-Navarro, F., Rubio-Gamez, M. C., & Martinez, G. (2022). Benchmarking the embodied environmental impacts of the design parameters for asphalt

Commented [KZ7]: Moretti, L., Mandrone, V., D'Andrea, A., & Caro, S. (2017). Comparative "from cradle to gate" life cycle assessments of Hot Mix Asphalt (HMA) materials. *Sustainability*, *9*(3), 400.

Commented [KZ8]: Mukherjee, A. (2016). Life cycle assessment of asphalt mixtures in support of an environmental product declaration. *National Asphalt Pavement Institute: Lanham, MD, USA*.

Commented [KZ9]: Mattinzioli, T., Sol-Sanchez, M., del Barco Carrion, A. J., Moreno-Navarro, F., del Carmen Rubio-Gamez, M., & Martinez, G. (2021). Analysis of the GHG savings and cost-effectiveness

Commented [KZ10]: Mantalovas, K., & Di Mino, G. (2020). Integrating circularity in the sustainability assessment of asphalt mixtures. *Sustainability*, *12*(2), 594.

Commented [KZ11]: Harvey, J. T., Butt, A. A., Saboori, A., Lozano, M. T., Kim, C., & Kendall, A. (2020). Life Cycle Assessment and Life Cycle Cost

Commented [KZ12]: Gruber, M. R., & Hofko, B. (2023). Life cycle assessment of greenhouse gas emissions from recycled asphalt pavement production. *Sustainability*, *15*(5), 4629.

These previous works have exhibited the effective use of EPD as a quantitative tool to measure and communicate the environmental impacts of asphalt mixtures and assess various strategies to reduce environmental impacts. However, these works and conclusions are made based on case studies with a limited number of asphalt mixtures or case simulations. An industry-wide study with various asphalt producers and types of asphalt mixtures shall be conducted to establish a repository of EPDs for asphalt mixtures. In the United States, the National Asphalt Pavement Association (NAPA) led an industry-wide effort to develop the Emerald Eco-Label EPD program and determine the benchmark for asphalt mixtures produced in the U.S. (Miller et al. 2024). Based on the published EPD of 1070 asphalt mixtures from 335 asphalt plants, the top 20 limit, top 40 limit, and industry average GWPs per metric ton of asphalt mixture from cradle-to-gate (A1-A3) were 55.4 kg CO_2e/t, 64.8 kg CO_2e/t, and 72.6 kg CO_2e/t, respectively. Shacat et al. (2022) estimated the total emissions for asphalt mixture produced in the U.S. in 2019 was 21.7 million metric tons CO_2e, which accounted for 0.3% of total U.S. GHG emissions in the same year. The use of RAP and recycled asphalt shingles (RAS) contributed to 85% of avoided emissions in asphalt mixture production, which were equivalent to 2.9 million metric tons CO_2e.

Commented [KZ13]: Miller, L., Ciavola, B., and Mukherjee, A. 2024. EPD Benchmark for Asphalt Mixtures. NAPA SIP-108 Report.

Commented [KZ14]: Shacat, J., Willis, J. R., Ciavola, B. 2022. GHG Emissions Inventory for Asphalt Mix Production in the United States: Current Industry Practices and Opportunities to Reduce Future Emissions. NAPA SIP 106.

Despite significant efforts being made to simplify the process of creating EPD of asphalt mixture, it is still a comprehensive, time-consuming, and costly process for digitalization of EPD for asphalt (Stromberg et al. 2020). For example, the Emerald Eco-Label tool requires the collection of primary input data, such as produced total asphalt mixture, transportation distances, fuel, and electricity used at the plant for the mix production, over a 12-month period. When a new mix design or technology is adopted, it may require recollecting input data and take a long time to analyze the environmental impacts. Thus, this study aims to develop a tool based on the US EPD database and machine learning model that could analyze and predict the environmental impacts of asphalt mixtures rapidly, which will facilitate the use of alternative mix designs with lower carbon emissions. The machine learning model will analyze the most important contributors, such as RAP and WMA, that could reduce carbon emissions based on the national database rather than case studies, which provide more reliable evidence to advocate for the adoption of these technologies. In addition, the work conducted the geospatial hotspot analysis to reveal the geospatial features of carbon emissions for the production of asphalt mixtures across the US, which assists in the forensic analysis of environmental impacts for asphalt mixtures.

Commented [KZ15]: Strömberg, L., Wendel, M., Berglund, M., & Lindgren, Å. (2020, May). Digitalization of EPDs for asphalt–experience from Sweden and input from Norway. In *publishing in proceeding of 7th Eurasphalt & Eurobitume Congress, Madrid* (pp. 12-14).

2. Data Collection

The data used in this study was collected from the published EPDs by the NAPA Emerald Eco-Label EPD program. The data was extracted on March 18, 2024. Table 1 shows the extracted variables used in the study. A total of 2211 EPDs across 44 states in the U.S. were involved.

The EPD covers the Cradle-to-Gate LCA analysis, including module A1 for raw materials extraction and manufacturing, module A2 for transportation of raw materials to the asphalt plant, and module A3 for plant production of asphalt mixtures. The sum of A1, A2, and A3 is the total amount of emissions based on the declared unit of 1 metric tonne of an asphalt mixture. The environmental impacts from an EPD report include global warming potential (GWP-100), ozone depletion potential (ODP), eutrophication potential (EP), acidification potential (AP), and photochemical ozone creation potential (POCP). This work selected GWP-100 as the primary environmental indicator as the exemplary work.

Table 1. Description of variables

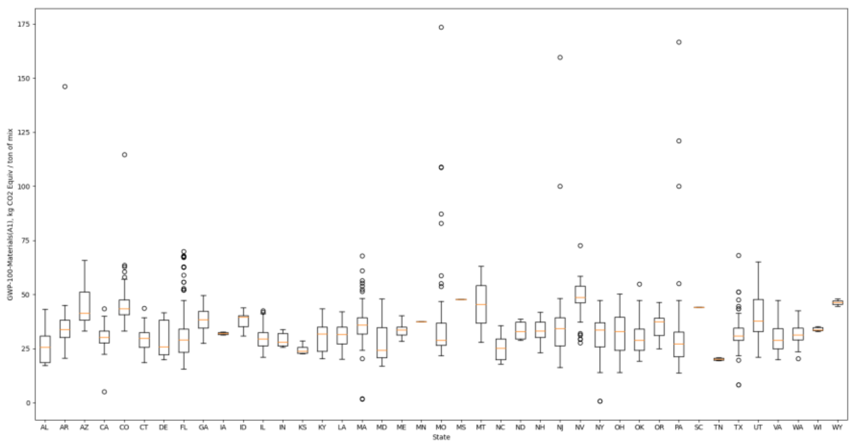
Variable		Data Type	Unit/Description
Independent variable	State Name	Categorical	Name of State in the US
	Gradation Type	Categorical	Dense, Gap, Open, Porous, Other
	Mix Design Method	Categorical	Superpave, Hveem, Marshall, Other
	Nominal Maximum Aggregate Size (NMAS)	Numeric	Inch
	Performance Grade (PG)	Categorical	PG of asphalt binder
	Binder Type	Categorical	Unmodified, GTR, PPA, SBS
	Mix Type	Categorical	HMA, WMA
	RAP Content	Numeric	%
	RAS Content	Numeric	%
	Binder Content	Numeric	%
	Lime Content	Numeric	%
	Aggregate Content	Numeric	%
Dependent variable	GWP-100-A1(Materials)	Numeric	kg CO_2e/ton of mix
	GWP-100-A2 (Transport)	Numeric	kg CO_2e/ton of mix
	GWP-100-A3 (Production)	Numeric	kg CO_2e/ton of mix
	GWP-100-Total (A1+A2+A3)	Numeric	kg CO_2e/ton of mix

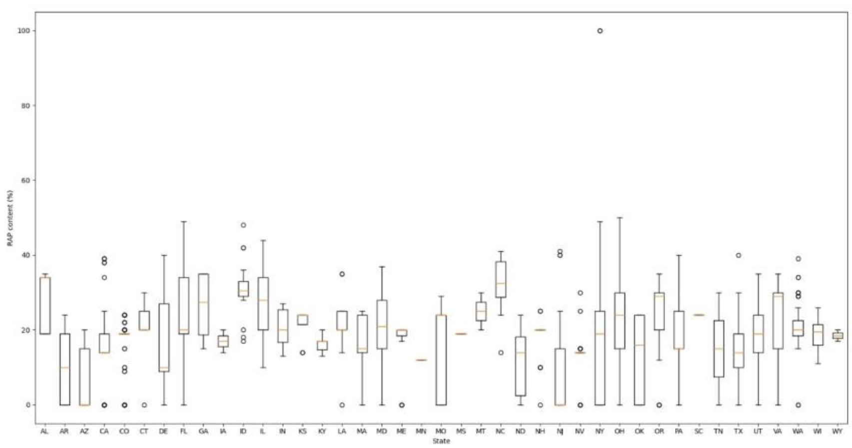
Commented [KZ16]: Or Numerical?

3. Analysis Methods, Results and Discussion (combined here first, will separate later on)

3.1 Descriptive Statistical Analysis

(1) Box and Whisker Plot for each state and all (mention number of EPDs after state and all), A1, A2, A3, and total, RAP, and Binder content

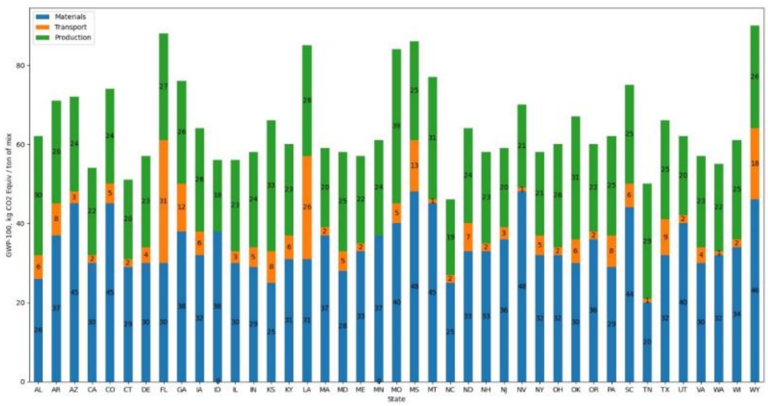


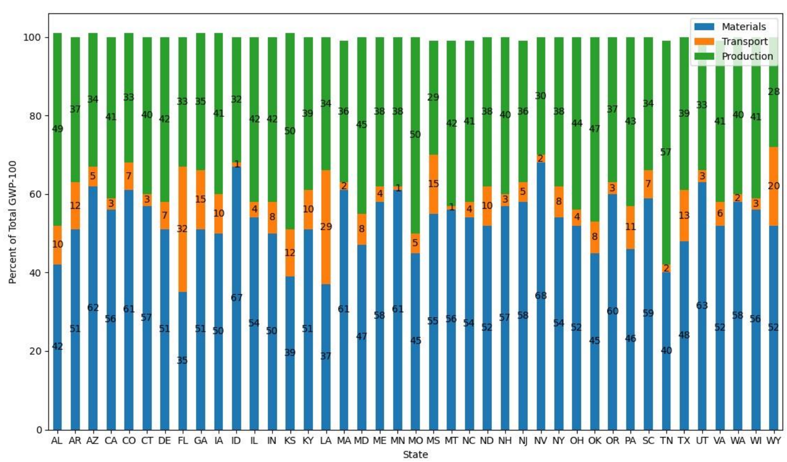


…

(2) Benchmarking Analysis using Industry-wide Average EPD values

Bar chart plot, e.g. for average results and 50-percentile results for all and each state.





Moretti et al. (2017), 39%-43% of GWP for A1 stage, 20-22% for A2 stage, and 35%-42% for A3 stages.

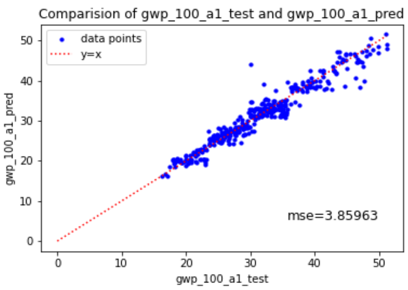
Commented [KZ17]: Moretti, L., Mandrone, V., D'Andrea, A., & Caro, S. (2017). Comparative "from cradle to gate" life cycle assessments of Hot Mix Asphalt (HMA) materials. *Sustainability*, *9*(3), 400.

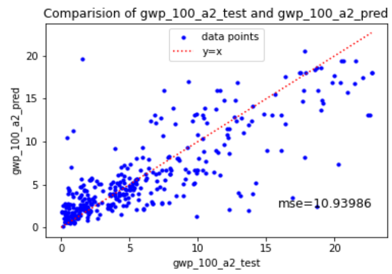
For RAP, the average results will be also compared with NAPA RAP survey data for states and national average.

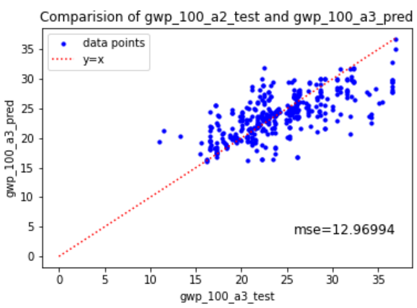
3.1 Machine Learning Models

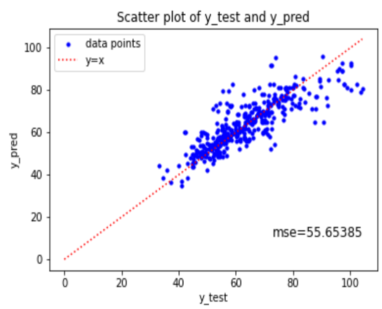
Data cleaning

Random forest analysis to predict A1, A2, A3, and Total









Rank the most important factors to affect A1, A2, A3, and Total

gwp_100_a1	
factors	feature importances
Binder_content	0.677681775
lime	0.125058481
binder type_SBS	0.096240563
state_AZ	0.018254706
Agg_content	0.010928976

gwp_100_a2	
factors	feature importances
state_FL	0.51501542
Agg_content	0.07906033
RAP_content	0.04829256
state_TX	0.04037757
NMAS	0.03263378

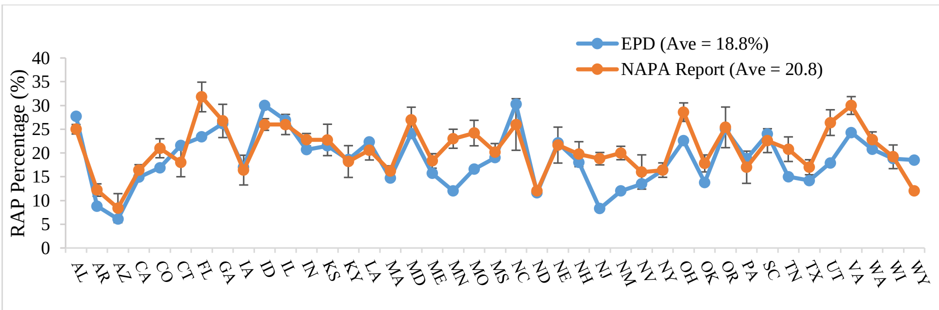
gwp_100_a3	
factors	feature importances
state_FL	0.134418513
Agg_content	0.097606564
RAP_content	0.090793899
state_OK	0.064369136
NMAS	0.059421476

gwp_100_total	
factors	feature importances
Binder_content	0.22611246
state_FL	0.2045301
Agg_content	0.07305677
RAP_content	0.07245063
lime	0.06052107

Example of Random forest analysis to predict GWP-100-Total

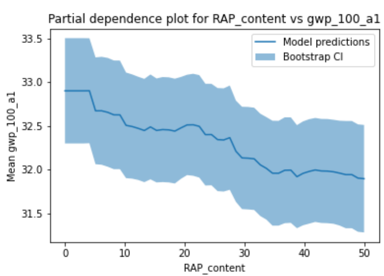
Assessment of RAP and WMA technology on Carbon Reductions

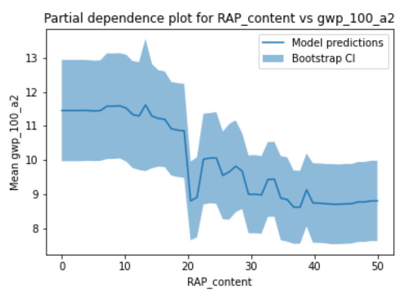
The impact of RAP on GWP-100,

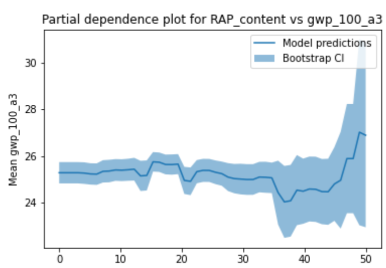


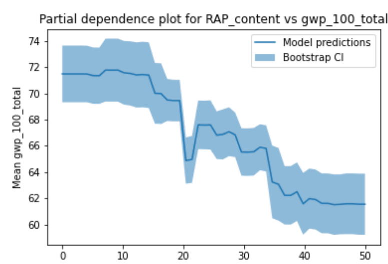
A few states have lower RAP repored by EPD because of lower published EPD, such as MN (n=1), MS (n=1). The EPD database could decently represent average RAP percentages in each state.

PDP plot for RAP vs GWP_100 (all data, and selected states (1-2))









With the increase of RAP-content usage, the declines in A1 and A2 can be witnessed. For A3, there is a fluctuation in A3 before RAP content usage up to 40%; after that, A3 rises slowly. In Total shows a sharp decline trend. Very interesting findings for RAP vs. A3.

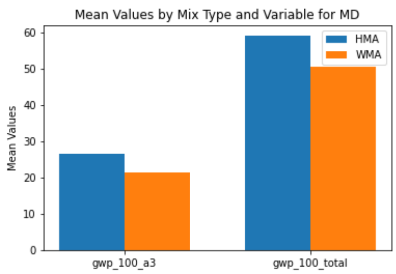
Linear regression results for selected state (slope and R2)

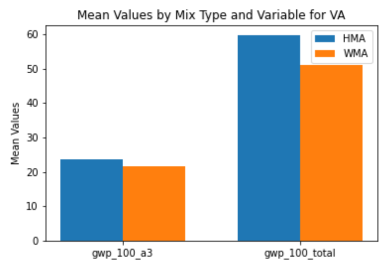
It seems the use of RAP has more beneficial for states which has higher GWP-A2.

WMA:

Select states (1-2) with similar amounts of WMA and HMA

t-test? Bar chart plot to show the difference.





MD	**t-statistics value**	**p-statistics value**
gwp_100_a3	5.990065066	1.76909E-07
gwp_100_total	2.538850727	0.01403621

VA	**t-statistics value**	**p-statistics value**
gwp_100_a3	2.329790074	0.026082437
gwp_100_total	3.283984583	0.002427362

In MD (HMA=29, WMA=28) and VA (HMA=23, WMA=24), A3 and Total are higher in the HMA group than in the WMA group.

Geospatial Hot Spot Analysis

The GIS-based hot spot analysis was used to identify geospatial characteristics of GWP-100-Total across the United States. A hot/cold spot is defined as a location in a cluster with its neighbor having significant high/low values (Environmental Systems Research Institute (Esri) 2023a). Zhang and Wang (2023) implemented the hot spot analysis to investigate the geospatial features of pavement distresses across the United States and Canada. Equation * is used to calculate Getis-Ord G_i^* to determine hot/cold spots (Ord and Getis 1995). For example, a hot spot with a confidence level of 99% is when the G_i^* > 2.58, while a cold spot with a confidence level

Commented [KZ18]: Environmental Systems Research Institute (Esri). (2023a). How Hot Spot Analysis (Getis-Ord Gi*) works. https://pro.arcgis.com/en/pro-app/latest/tool-reference/spatial-statistics/h-how-hot-spot-analysis-getis-ord-gi-spatial-stati.htm

Commented [KZ19]: Zhang, K., & Wang, Z. (2023). LTPP data-based investigation on asphalt pavement performance using geospatial hot spot analysis and decision tree models. *International Journal of Transportation Science and Technology*, *12*(2), 606-627.

Commented [KZ20]: Ord, J. K., and Getis, A. (1995). Local spatial autocorrelation statistics: distributional issues and an application. Geographical analysis, 27(4), 286-306.

of 95% is when the G_i^* value is between -2.58 and -1.96. When the G_i^* value is between -1.65 and 1.65, the location is a not significant spot. The ESRI ArcGIS Pro 3.1 was used in this work to perform the optimized hot spot analysis (OHSA). The optimal distance band to cluster pavement sections based on the average distance to 30 nearest neighbors is 88,706 meters (55.1 miles).

$$G_i^* = \frac{\sum_j w_{i,j} x_j - \bar{X} \sum_j w_{i,j}}{S \sqrt{\frac{n \sum_j w_{i,j}^2 - (\sum_j w_{i,j})^2}{n-1}}} \quad (*)$$

with $\bar{X} = \frac{\sum_j x_j}{n}$ and $S = \sqrt{\frac{\sum_j x_j^2}{n} - (\bar{X})^2}$

where $w_{i,j}$ is the spatial weight between location i (i = 1, 2, …, n) and j (j = 1, 2, …, n); x_j is the attribute value in location j; n is the total number of observations.

Geospatial Outlier Analysis

The optimized outlier analysis was conducted using the ESRI ArcGIS Pro 3.1 to calculate the local Moran's *I* value, as shown in Equation * (Anselin 1995, and Esri 2023 b). The z-score of local Moran's I value is calculated to identify statistically significant clusters of high (HH cluster) or low (LL cluster) values, as well as outliers that a low value surrounded by high values (LH outlier) or a high value surrounded by low values (HL outlier). The outlier analysis results are used to perform the forensic study to identify the asphalt mixtures with low emissions in the hot spot areas (LH outlier) or asphalt mixtures with high emissions in the cold spot areas (HL outlier). This helps to identify strategies to reduce emissions in hot spot areas or avoid high emissions in cold spot areas.

$$I_i = \frac{x_i - \bar{X}}{S_i^2} \sum_{j=1, j \neq i}^{n} w_{i,j} \, (x_j - \bar{X}) \quad (*)$$

with $S_i^2 = \frac{\sum_{j=1, j\neq i}^{n} (x_j - \bar{X})^2}{n-1}$

Commented [KZ21]: https://pro.arcgis.com/en/pro-app/latest/tool-reference/spatial-statistics/h-how-cluster-and-outlier-analysis-anselin-local-m.htm

Anselin, L. (1995). Local indicators of spatial association—LISA. *Geographical analysis*, *27*(2), 93-115.

Optimized Hot Spot Analysis

Figure * shows the hot spot map of GWP-100-Total of asphalt mixtures across the United States. There are four major hot spot areas identified for asphalt mixtures produced in states of Colorado, Florida, Louisiana, and Missouri. There are different reasons for these states to

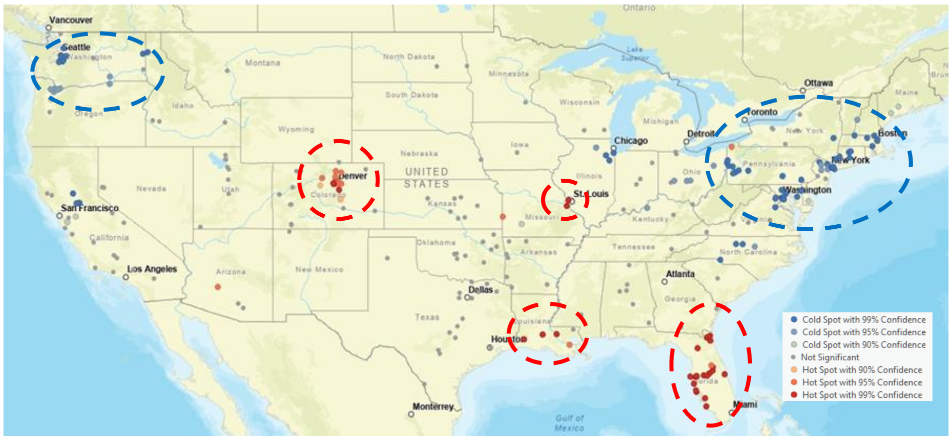
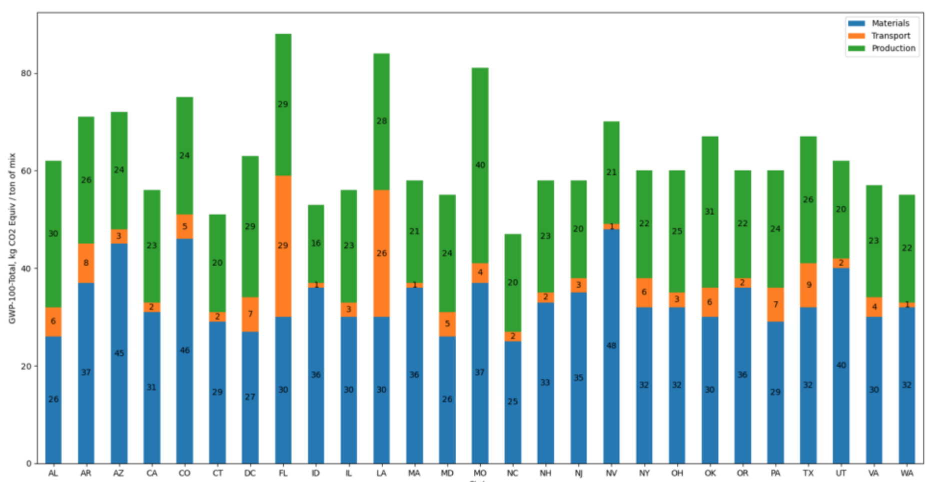


Figure * Hot spot map of GWP-100-Total of asphalt mixtures across the United States
Florida and Louisiana (A2). Denver, Colorado (A1 with lime?)



Optimized Hot Spot Analysis
Distance band = 88706 meters (55.1 miles) with 30 neighbours

Optimized Outlier Analysis
The optimal fixed distance band is based on the average distance to the 30 nearest neighbors: 88706.0 Meters (55.1 miles)

Low-High Outlier in Florida (ID819-824), asphalt mixes with 34%RAP, 39%RAP, and 29%RAP, which had lower GWP-100-A2.
ID534-536, Asphalt Mixes with 39%RAP.
Colorado: ID257 (22%RAP, 1%lime), ID266-268 (19%-24%RAP, with 1%lime or no lime), ID320 (3/4" mix with 19%RAP)

Cluster and Outlier Analysis

Neibourhood search threshold was 343,235 meter (213.3 miles)

Conclusion

This work only considered A1-A3, the full LCA analysis also includes construction, XX, XX, depending on the performance of mixtures.

Dynamic database, newer materials and additives, production changes,

NAPA, dynamic,

Requires analysis to track the carbon emissions of asphalt mixtures

References