EPD BENCHMARK
FOR ASPHALT MIXTURES

Prepared for NAPA by WAP Sustainability

Lianna Miller, Benjamin Ciavola, Amlan Mukherjee

TABLE OF CONTENTS

1. THIRD-PARTY REVIEW LETTER

JOHN BEATH ENVIRONMENTAL, LLC
Striving to make something better every day

888.777.4310
JBE@beath.us
www.beath.us

February 26, 2024

Joseph Shacat   National Asphalt Pavement Association
Director of Sustainable Pavements   6406 Ivy Lane, Suite 350
jshacat@asphaltpavement.org   Greenbelt, MD 20770-1441
301-731-4748

Dear Mr. Shacat,

The National Asphalt Pavement Association (NAPA) commissioned John Beath Environmental, LLC (JBE) to conduct an external independent verification of its development of a methodology for quantifying the industry average global warming potential (GWP) of asphalt mixtures as well as benchmark results for the industry average. The methodology and results were developed by WAP Sustainability (WAP) and are intended to support the Federal Highway Administration (FHWA) to determine and industry average calculation process in relation to Inflation Reduction Act (IRA) funding.

The methodology and results were provided to the reviewer in a report titled "EPD Benchmark for National Asphalt Pavement Association" authored by Lianna Miller, Benjamin Ciavola, and Amlan Mukherjee and dated February 9, 2024. The report was reviewed for conformance to the following requirements:

- Email titled "FHWA Process to Identify Industry Average Calculations" sent by LaToya Johnson on August 2, 2023
- ISO 21930:2017, Sustainability in buildings and civil engineering works — Core rules for environmental product declarations of construction products and services
    - Section 10, Project Report
- ISO 21678:2020, Sustainability in buildings and civil engineering works — Indicators and benchmarks — Principles, requirements and guidelines
    - Section 4.2.2, Reference Values

The scope of the review focused on the sections of the standards noted above. The review process did not include a comprehensive review for conformance with ISO 21678 and ISO 21930 and this review statement makes no representations of conformance with these standards.

The review took place after a draft report was provided to JBE. The report was reviewed for conformance to the requirements outlined above. Feedback was sent to WAP and NAPA via a comment matrix spreadsheet. Two rounds of feedback were conducted. Responses by the development team to each issue raised were addressed to the satisfaction of the reviewer.

The methodology and results in the provided report were found to be in conformance with the applicable requirements.

Sincerely,

Sevda Alanya Rosenbaum, Ph.D.
Senior Sustainability Consultant
John Beath Environmental

2. ABBREVIATED TERMS

AASHTO American Association of Highway and Transportation Officials

EPA Environmental Protection Agency

EPD Environmental Product Declaration

FHWA Federal Highway Administration

GSA General Services Administration

GWP Global Warming Potential

IRA Inflation Reduction Act

ISO International Organization for Standardization

LCA Life Cycle Assessment

LCI Life Cycle Inventory

LCIA Life Cycle Impact Assessment

NAPA National Asphalt Pavement Association

NETL National Energy Technology Laboratory

NSSGA National Stone, Sand & Gravel Association

PCR Product Category Rules

SBS Styrene-Butadiene-Styrene

TRACI Tool for Reduction and Assessment of Chemicals and Other Environmental Impacts

3. SUMMARY

The objective of this study is to evaluate the exogenous, uncontrollable factors that influence the Global Warming Potential (GWP) of asphalt mixture production and to propose a phase-by-phase method for regionalized benchmarking. This approach is designed to ensure contractors in all regions of the United States are fairly incentivized to transition to low carbon construction materials in accordance with the goals of the Inflation Reduction Act (IRA) (2022).

The proposed method uses a statistical assessment of GWP as reported in Environmental Product Declarations (EPDs) for each life cycle phase. Benchmarking by life cycle phase allows for the selection of lower carbon materials while ensuring equitable access to IRA funds despite factors outside a contractor's control, such as agency-specific material specifications, geology, and regional climate. This is a departure from the current practice of using total GWP reported in EPDs to create blanket national thresholds, which can unfairly burden some regions of the country while advantaging others.

This report develops a method that allows agencies to establish asphalt mixture GWP thresholds that account for regional variations and agency specifications by treating the life cycle phases (A1, A2, and A3) as independent components that, when combined, provide the relevant 20th percentile, 40th percentile, and average GWP values for similar products. The combined thresholds can provide a single embodied carbon number representing to the region and material specification.

This report recommends the following:

- **Use Impact Factors for A1 Impacts:** A1 impacts are deterministic and entirely dependent on mix design. Agencies can use GWP impact factors for

each mix ingredient type (e.g. RAP, virgin aggregate, or binder) to set 'baseline' A1 GWP values based on local mix specifications. The role of material choice in driving GWP is communicated with this method, and A2 and A3 impacts can then be combined with A1 to form a complete set of localized GWP thresholds.

- **Use Regional Trends for A2 Benchmarks:** The use of region-specific distribution-driven thresholds is recommended for the A2 life cycle phase. Impacts due to the transportation of materials are dictated by availability of materials locally, which in turn is a function of local geology. For example, states like Louisiana and Florida have significantly higher A2 impacts than any other regions.

- **Use Climate Regions for A3 Benchmarks:** Impacts due to production vary to a smaller but statistically significant degree across different AASHTO climate regions. While A3 impacts are within the control of material producers, the recommendation is to develop distribution-driven thresholds that account for climatic differences.

- **Use the Sum of A1, A2, and A3 Benchmarks for Procurement:** By combining the local benchmark value for each of A1 (deterministic), A2 (distribution), and A3 (distribution) an agency can identify a set of fair GWP thresholds that incentivize improvements in environmental performance by local contractors.

- **Improve Sampling:** The analysis recommends the continued development of more representative data sets when establishing thresholds for procurement, driven by an intentional sampling process that targets states and regions with limited participation.

This report presents the development, justification, and implementation of this framework. Data collected through the National Asphalt Pavement Association Emerald Eco-Label EPD tool's benchmarking feature is used to demonstrate the feasibility and utility of this approach.

4. GENERAL INFORMATION

This study was commissioned by the National Asphalt Pavement Association (NAPA). This study uses data derived from NAPA's Emerald Eco-Label EPD program, and a sample of facility and mixture specific EPD data collected independently for this study to develop a benchmark for asphalt mixture Environmental Product Declarations (EPD). There were 107 participating organizations representing 335 production facilities. A full list of study participants may be requested from NAPA.

This study is an evaluation of Asphalt Mixtures (UNSPSC 30111509) and guided by NAPA's Product Category Rules for Asphalt Mixtures v2.0, published April 2022 and valid through March 2027[1]. Asphalt mixtures are typically comprised of virgin or recycled construction aggregates (>90% by mass), virgin or recycled asphalt binders (typically 4% – 8% by mass), dry additives (<2% by mass), and wet additives (<1% by mass). Asphalt mixtures are typically incorporated as part of the structure of a roadway, parking lot, driveway, airfield, bike lane, pedestrian path, railroad track-bed, or recreational surface. Additional information about the environmental impacts of asphalt mixtures can be found as part of NAPA's recycling report[2] and Greenhouse Gas Inventory[3]. The study is aligned to meet the reporting requirements of ISO 21930 and ISO 21678. However, as this is a benchmark study and not a Product Category Rule or an Environmental Product Declaration, complete conformance with ISO 21930 is not guaranteed.

Third Party Review Statement

This study was subjected to third-party review by Sevda Rosenbaum of John Beath Environmental, LLC. It was reviewed to the "FHWA Expectations for Industry Averages for IRA 60506." The project report and underlying data were made available to the reviewer with the requirements on confidentiality stated in ISO 14025. The review was also conducted with reference to ISO 21678, though conformance to the standard was not reviewed. A single reviewer was used for this study instead of a panel because of FHWA's stated intention to perform a panel review after the next planned round of data collection.

---

[1] "PCR for Asphalt Mixtures."
https://www.asphaltpavement.org/uploads/documents/EPD_Program/NAPA_PCR_AsphaltMixtures_v2.pdf.
[2] National Asphalt Pavement Association, "Sustainability Resources: Recycling."
https://www.asphaltpavement.org/expertise/sustainability/sustainability-resources/recycling
[3] Joseph Shacat, Richard Willis, Ph.D., and Benjamin Ciavola, Ph.D., "GHG Emissions Inventory for Asphalt Mix Production in the United States."
https://www.asphaltpavement.org/uploads/documents/Sustainability/SIP-106_GHG_Emissions_Inventory_for_Asphalt_Mix_Production_in_the_US_%E2%80%93_NAPA_June_2022.pdf

5. GOAL AND SCOPE OF THE STUDY

This study was undertaken to establish a global warming potential (GWP) benchmarking methodology for asphalt mixtures that accounts for the sensitivity of GWP to factors outside a contractor's influence (such as climate, geology, and mixture specification requirements) and provide initial estimates of industry average GWPs for use in implementing Sections 60503 and 60506 of the IRA. The intended audience for this study is local, state, and Federal agencies that have set or are considering setting performance thresholds for asphalt mixture procurement, as well as asphalt industry professionals seeking to understand the methodological approach and benchmark their own products against relevant industry averages. FHWA and other Federal agencies can use the report and calculations to be referenced by recipients of the Low Carbon Transportation Materials Program to identify substantially low carbon materials.

Sections 60503 and 60506 of the IRA provide funding to the General Services Administration (GSA) and Federal Highway Administration (FHWA) to pay for the differential cost or incentives for

agencies to purchase construction materials with substantially lower embodied carbon as reported in EPDs than estimated industry averages. The material categories identified in the IRA and EPA Interim Determination include asphalt, concrete, flat glass, and steel. The EPA issued an interim determination in December 2022 establishing a cascading set of thresholds to define what it means to be "substantially lower" than industry averages. Under this rubric, the first threshold is materials that are the best performing 20% for GWP values. If not available locally, the next threshold is the best performing 40%. The EPA's determination goes on to state, "If materials/products in the Top 40 percent are not available in a project's location, then a material/product qualifies for funding... if its GWP is better than the estimated industry average."

In supporting the goal of the IRA to incentivize the transition to low carbon materials, this study assesses the extent to which exogenous factors (design and production parameters that producers cannot control) influence the GWP of asphalt mixtures, and proposes a phase-by-phase
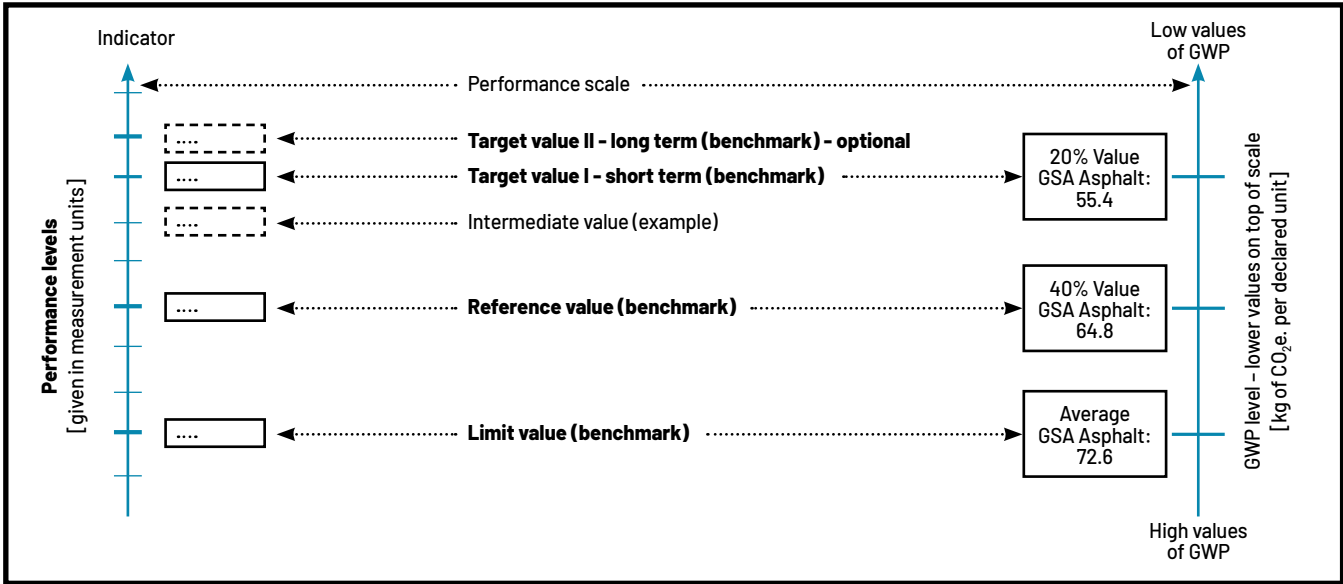

Figure 1. Illustration of Limit, Reference and Target Values as defined in ISO 21678 mapped onto EPA's Interim Determination Performance Levels and Reference Values set by GSA

regionalized benchmarking approach that will ensure that market participants in all regions are appropriately incentivized to realize the intentions of the programs set forth in the IRA.

This study follows the ISO 21678 standard, *Sustainability in buildings and civil engineering works: Indicators and benchmarks — Principles, requirements and guidelines*[4]. Additionally the protocols outlined in the PCR document *Product Category Rules (PCR) for Asphalt Mixtures* for the calculation of GWP are adopted, specifically with respect to impact assessment methodology, allocation procedures, and the use of upstream inventories. The scope of this study is limited to only the GWP midpoint indicator, even though the methods outlined here can be extended to any of the other indicators as well.

ISO 21678 establishes four key benchmark values for defining thresholding policies (Figure 1).

The thresholding technique adopted by EPA and GSA maps to these benchmark values in a straightforward fashion. The method we present in this document likewise maps to the benchmark types established in ISO 21678, but develops its benchmark values using a composite of regionalized benchmarks for each of A1, A2, and A3.

The declared unit for asphalt mixtures is one metric tonne (one short ton) of mix. Both units are provided since ISO 21930 requires the use of SI units, but asphalt mixtures are typically procured on the basis of short tons in the United States. The scope of this study is limited to cradle-to-gate and includes the life cycle phases and the processes within the system boundaries shown in Figure 2. This study follows all procedures for calculation of impacts detailed in the PCR for Asphalt Mixtures[1]. The PCR takes a prescriptive approach, fully describing the required upstream data sets, allocation procedures, cutoff criteria, and treatment of missing data,

Construction Works Assessment Information														
Construction Works Life Cycle Information Within the System Boundary														Optional supplementary information beyond the system boundary
A1–A3			A4–A5		B1–B7					C1–C4				D
Production Stage			Construction Stage		Use Stage					End-Of-Life Stage				
A1	A2	A3	A4	A5	B1	B2	B3	B4[a]	B5	C1	C2	C3	C4	
Extractional upstream production	Transport to factory	Manufacturing	Transport to site	Installation	Use	Maintenance (incl. production, transport, and disposal of necessary materials)	Repair (incl. production, transport, and disposal of necessary materials)	Replacement (incl. production, transport, and disposal of necessary materials)	Refurbishment (incl. production, transport, and disposal of necessary materials)	Deconstruction / Demolition	Transport to waste processing or disposal	Waste processing	Disposal of waste	Potential net benefits from reuse, recycling, and/or energy recovery beyond the system boundary
			Scenario	Scenario	Scenario	Scenario	Scenario	Scenario	Scenario	Scenario	Scenario	Scenario	Scenario	Scenario
					B6 Operational Energy Use Scenario									
					B7 Operational Water Use Scenario									

[a] Replacement information module (B4) not applicable at the product level

Figure 2. ISO 12930 system boundary declaration from the Asphalt Mixtures PCR

[4] ISO, "ISO 21678."

among other requirements. Additionally, this study uses the procedures detailed in the underlying LCA[5] for quantification of energy and material inputs and outputs, considering how plant-level data are allocated to the declared products, description of the application of cut-off criteria and assumptions, and list of excluded processes. A list of Masterformat Codes and relevant technical specifications can be found in the PCR.

The values for GWP are characterized using the TRACI 2.1 methodology. As with all midpoint indicators, the LCIA results given are relative expressions and do not predict impacts of category endpoints, exceedance of thresholds, safety margins, or risks[6].
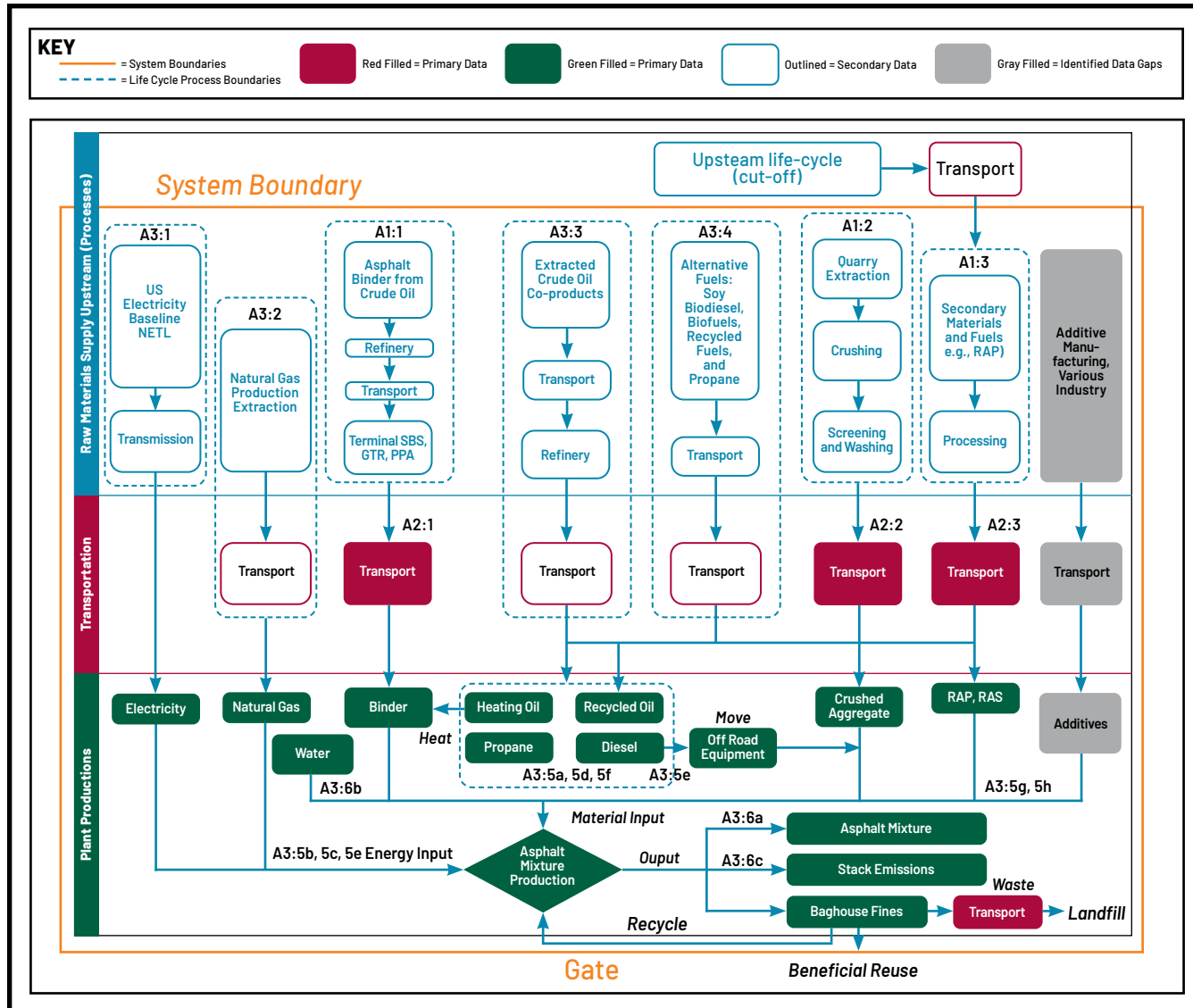
Figure 3. System boundary for asphalt mixture production.[5]

[5] ??

[6] Mukherjee, Amlan. "Update to the Life Cycle Assessment for Asphalt Mixtures in Support of the Emerald Eco Label Environmental Product Declaration Program." https://www.asphaltpavement.org/uploads/documents/Programs/Emerald_Eco-Label_EPD_Program/PCR_Public_Comment_Period/LCA_Asphalt_Mixtures_07_29_2021.pdf

6. BACKGROUND

The GSA's document *Interim IRA Low Embodied Carbon Material Requirements*, released May 16, 2023,[7] establishes thresholds for asphalt based on the methodology outlined in the EPA documents *Interim Determination on Low Carbon Materials under IRA and COVER MEMO – EPA's Interim Determination for GSA & DOT/FHWA on low greenhouse gas construction materials under IRA Sections 60503 and 60506* released December 22, 2022.[8,9] The GSA thresholds are shown in Figure 4.

An accompanying FAQ identifies the data sources used for these thresholds: "GSA's GWP limits were developed based on industry average EPDs and actual products publicly available EPD data, filtered by material type, PCR(s) specified in GSA's requirements, North American geographical scope and current validity."[10]

The EPA Interim Determination provides the following calculation guidance:

*materials/products qualify if their product-specific GWP is in the best performing 20 percent (Top 20 percent or lowest 20 percent in embodied greenhouse-gas emissions), when compared to similar materials/products (for example, materials/products within the same product category that meet the same functional requirements). If materials/products in the Top 20 percent are not available in a project's location, then a material/product qualifies per this determination if its GWP is in the Top 40 percent (lowest 40 percent in embodied greenhouse gas emissions). If materials/products in the Top 40 percent are not available in a project's location, then a material/product qualifies per this determination if its **GWP is better than the estimated industry average**. [emphasis added]*

GSA IRA Limits for Low Embodied Carbon Asphalt - May 16, 2023 (EPD-Reported GWPs, in kilograms of carbon dioxide equivalent per metric ton - $kgCO_2e/t$)		
Top 20% Limit	Top 40% Limit	Better Than Average Limit
55.4	64.8	72.6

Figure 4. GSA's GWP limits as published[7]

---

[7] Bare, J.C. (2012). Tool for the Reduction and Assessment of Chemical and Other Environmental Impacts (TRACI), Version 2.1 — User's Guide. Report No. EPA/600/R-12/554 2012. U.S. Environmental Protection Agency, Cincinnati, Ohio. https://nepis.epa.gov/Adobe/PDF/P100HN53.pdf

[8] "Interim Determination on Low Carbon Materials under IRA 60503 and 60506." https://www.epa.gov/system/files/documents/2023-01/2022.12.22%20Interim%20Determination%20on%20Low%20Carbon%20Materials%20under%20IRA%2060503%20and%2060506_508.pdf

[9] "COVER MEMO - EPA's Interim Determination for GSA & DOT/FHWA on low greenhouse gas construction materials under IRA Sections 60503 and 60506." US Environmental Protection Agency, December 22, 2022. https://www.epa.gov/system/files/documents/2023-01/2022.12.22%20COVER%20MEMO%20Interim%20Determination%20under%20IRA%20Sections%2060503%20and%2060506_508.pdf

[10] "U.S. General Services Administration Interim IRA Low Embodied Carbon Material Requirements." https://www.gsa.gov/system/files/Interim%20IRA%20LEC%20Material%20Requirements%20-%20used%20in%20Pilot%20May%202023%2005162023.pdf

The document then addresses data sets that may be used in threshold calculation:

Estimating the best performing 20 percent and 40 percent and industry averages.

Agencies shall estimate the GWP at the 20th and 40th percentiles and the industry average, as needed, for each material/product category using data from a verified source (e.g., an open source EPD database, industrywide EPDs or a 3rd party verified LCA developed using the relevant PCR). In addition, agencies shall disclose the GWPs, the methodology for determining the percentiles and averages, the source(s) used for each material/product, and the parameters (including performance specification) that can be used to set the GWP.

The Interim Determination – while an important and significant step in furthering sustainable public procurement – presents the following opportunities for further investigation:

■ First, the EPA determination does not describe a particular calculation method for determining the 20%, 40%, and estimated industry average. For example, it is unclear whether the average is intended as an arithmetic mean, or as the median.
■ In the absence of available industry averages, GSA has used publicly available databases as a source for EPDs. While these databases serve as a useful industry resource, they are not third-party verified for correctness or representativeness.

Using the above as points of departure, this study intends to analytically support the process that EPA and GSA have set forth, as follows:

■ Developing an analysis framework for characterizing the sensitivity of the GWP metric to different drivers of uncertainty using industry specific data. Asphalt mixture production is a distributed, local production process that is highly sensitive to uncertainties arising from exogenous regional factors through each of the A1, A2, and A3 life cycle phases such that using a nationwide, region-invariant benchmark to drive procurement can unfairly disadvantage some producers. Assessing the sensitivity of GWP to exogenous factors will support thresholds that do not unfairly limit contractors from accessing the IRA funding intended to help them reduce their emissions.
■ Establishing a set of statistical approaches that, while simple, provide a formal approach that can be applied consistently to generate thresholds. This includes but is not limited to clarifying the ramifications of using either arithmetic mean or median as alternative definitions of average.
■ Establishing separate categorization-driven benchmarking techniques for each of these phases as a candidate alternative to the existing single-valued cradle-to-gate approach covering A1-A3.

This study builds on the work done by EPA and GSA by establishing a larger dataset and examining multiple sources of uncertainty for each of A1, A2, and A3 phases of the asphalt mixture life cycle.

7. METHODOLOGICAL FRAMEWORK

The methodological framework is organized by sub-objectives that support the overall objective of developing industry benchmarks that account for the sensitivity of the GWP metric to various exogenous factors. The two primary sub-objectives are as follows:

O1: Develop a Data Collection and Sampling Methodology:
- Establish a representative dataset for analysis.
- Develop a data collection protocol that is easy to implement.
- Conduct a data quality assessment of the collected sample.

O2: Uncertainty Characterization:
- Identify the exogenous factors, or categories by which the GWP can be classified, for example, climate, geography, availability of aggregate, and material specification, among others.
- Characterize the extent to which GWP of asphalt mixtures is influenced by each of the categories and test for statistical significance.
- Establish a list of statistical tests used for making comparisons.

O3: Proposed Benchmark Calculation Method:
- Devise a method that agencies can use to assist decision-making.
- Establish a defensible set of coefficients/ formulas based in LCA outcomes to calculate benchmarks.

Each of these components is addressed in detail in the rest of Section 7 of this report.

a. O1: Data Collection and Sampling Methodology

From May to June of 2023, NAPA conducted an industry benchmarking survey to collect data from producers regarding their plant operations and material transport distances. An updated version of the Emerald Eco-Label tool was used to collect data spanning a 12-month period within the last 5 years. The objective of the data collection was to establish industry averages for transportation (A2) and mixture production (A3) life cycle phases. A reproduction of the modified input page (Figure 5) and full list of additional data types collected (Appendix 1) are included in this document. Participation was free to producers and performed on a voluntary basis. Pavement performance data was not collected, as pavement performance-related impacts are reported and influenced by factors outside the scope of the of the current asphalt mixture PCR. Furthermore, performance-related impacts are not currently included in the scope of the proposed or forthcoming regulations outlined in the EPA and GSA Interim Determinations.

Aggregates

This section aims to assess the impacts of aggregate transport. Over the course of the 12-month data reporting period, which two quarries or pits supplied the most aggregates to this plant?

For each of the top two aggregate sources report the total quantity supplied to this plant (tons) and the one-way transport distance and mode(s). The transport distance should be the entire distance from the quarry or pit where the aggregates were extracted to your plant. If your plant is co-located with the aggregate source, enter a distance of 0 miles.

For portable plants that sourced aggregates from the same quarry or pit while at multiple locations during the 12-month data collection period, enter the total tonnage sourced from the same quarry and the weighted average distance. Contact Joseph Shacat with questions.

Most Used Quarry/Pit

Approximate quantity purchased from this source*

	shtn

If you do not have exact numbers please estimate to the nearest 5000 tons.

Truck Distance (mi)	Train Distance (mi)	Barge Distance (mi)	Ocean Distance (mi)
miles	miles	miles	miles
Enter the truck distance in miles	Enter the train distance in miles	Enter the barge distance in miles	Enter the ocean distance in miles

Second-Most Used Quarry/Pit

Approximate quantity purchased from this source*

	shtn

If you do not have exact numbers please estimate to the nearest 5000 tons. Enter 0 if this plant only sourced aggregates from a single quarry/pit.

Truck Distance (mi)	Train Distance (mi)	Barge Distance (mi)	Ocean Distance (mi)
miles	miles	miles	miles
Enter the truck distance in miles	Enter the train distance in miles	Enter the barge distance in miles	Enter the ocean distance in miles

Asphalt Binder

Please provide the transportation distance and mode(s) for this plant's most-used asphalt binder supplier. Enter the one-way transport distance from the asphalt binder terminal to this asphalt plant.

Binder Truck Distance (mi)	Binder Train Distance (mi)	Binder Barge Distance (mi)	Binder Ocean Distance (mi)
miles	miles	miles	miles
Enter the truck distance in miles	Enter the train distance in miles	Enter the barge distance in miles	Enter the ocean distance in miles

Figure 5. A subset of inputs from the benchmarking survey tool

The data were evaluated for representativeness by the following criteria:

Time Period	Good	All data were from a 12-month period within the past 5 years.
Climate Region	Good	The number of participants in a climate region varied from 31 to 133.
State / Market Coverage	Adequate with Gaps	40 states and the District of Columbia were present in the benchmarking dataset. Since asphalt plants operate within a limited geographic region, state participation was used as a proxy for market coverage. States with no participating locations total about 9% of national production by tonnage. The total production of the participating locations is about 14% of the national production. 29 states had a sufficient degree of company participation (N>=3) to report state-level numbers.
Technology	Good	Batch plants (48), parallel flow plants (39), and counter flow plants (248) participated.

Geographic representation is evaluated for representativeness by both climate and political / market boundaries.

In total, 335 plants provided data for the benchmarking study.

b. O2: Uncertainty Characterization

Bhat & Mukherjee (2019)[11] discuss the different kinds of uncertainty that influence the GWP measure. The first is referred to as aleatory uncertainty, or uncertainty arising from factors that can be characterized as random variables. Aleatory uncertainty describes the stochasticity in or due to exogenous factors *from outside the system boundaries*. In the context of GWP, aleatory uncertainties influence the foreground LCA parameters that distinguish one facility and product specific EPD from another. As the source of these uncertainties are due to factors outside the system boundary, and beyond the control of the contractor, (e.g. climate), they can be modeled using random variables. In this document aleatory uncertainty is referred to as parametric uncertainty.

The second source of uncertainty is epistemic uncertainty, or uncertainty arising from gaps in knowledge about a system and/or data gaps. Epistemic uncertainty can arise due to incompleteness or limited quality of background LCI, or due to inconsistencies in LCA modeling approaches. As epistemic uncertainties are within the system boundaries, a purely statistical approach to characterizing their impact on GWP can be difficult, especially given the additional confounding

[11] Bhat, C. G., & Mukherjee, A. (2019) Sensitivity of Life-Cycle Assessment Outcomes to Parameter Uncertainty: Implications for Material Procurement Decision-Making, Transportation Research Record, Journal of the Transportation Research Board of the National Academies, 2673(3) 106:114.

influence of parametric uncertainty. Therefore, every attempt should be made to isolate epistemic uncertainties and evaluate them as alternative cases when investigating parametric uncertainties. For example, given two sets of EPDs where one set was calculated using background database A and the other with background database B, each would undergo a separate analysis of the parameters, controlling for the use of different databases. Such a conditional analysis of GWP would ensure that the benchmarking is a function of statistical parametric uncertainty and not choice of background data. Typically, a PCR can be written to manage for epistemic uncertainty by specifying background datasets, selection of proxies for data gaps, and modeling approaches. In the asphalt industry, epistemic uncertainties due to data gaps have been reduced by model and background data specification in the PCR.

It is important that the uncertainty being characterized can be associated with known factors. Hence, the first step is to classify the data into independent categories. Classification allows assessment of uncertainties arising from each category independently one at a time holding all the other categories constant, thus eliminating confounding interactions between multiple variables. For example, climate, an exogenous factor, impacts energy requirements during asphalt mixture production (A3), and geology, another exogenous factor, impacts material availability and therefore travel distance (A2). Hence, the sensitivity of GWP to travel distances should be assessed across each climate zone separately, requiring a classification of the EPD data by climate. The classifiers used in this study are as follows:

1. Environmental factors
   a. Variations in local climate
   b. Variations in local geology

2. Technological factors
   a. Plant efficiency
   b. Plant type
3. Calculation factors
   a. Data entry errors
   b. Modeling approaches
   c. Choice of life cycle inventory
4. Political / economic factors
   a. Local mixture specifications
   b. Variations in regional availability of fuel sources
   c. Variations in local power grids

It is important to recognize that each of the above factors impact the GWP in different ways, some through parametric uncertainty or epistemic uncertainty (e.g., items 3 and 4). Items 1 and 2 directly influence foreground parameters such as plant energy use and travel distances. Classifying the data provides the opportunity to individually characterize the impact of each factor and enable the estimation of context specific thresholds with a higher level of confidence. Specifically, our choice of classifiers focuses on factors that are outside the control of an asphalt producer.

Classification by Climate Regions
The AASHTO climate regions were used for a first order examination of climate effects. The continental United States is split into four regions, depending on rainfall and temperature. Wet regions receive an annual rainfall above 508mm (20"), and non-freeze regions have a freezing index above 83.3°C-days[12]. The resulting regions of Wet No Freeze (hot-wet), Dry No Freeze (hot-dry), Wet Freeze (cold-wet), and Dry Freeze (cold-dry) are shown in Figure 6.

Though more granular regions for climate are available, the number of participating sites limited the detail that could be examined with statistical relevance. In limited cases where the sub-population data for an individual state was great enough, the A3 data could also be examined by state.

---

[12] Jackson, N., & Puccinelli, J. (2006). Long-Term Pavement Performance (LTPP) Data Analysis Support: National Pooled Fund Study TPF-5(013) Effects of Multiple Freeze Cycles and Deep Frost Penetration on Pavement Performance and Cost (p. 69). Federal Highway Administration. https://www.fhwa.dot.gov/publications/research/infrastructure/pavements/ltpp/06121/06121.pdf
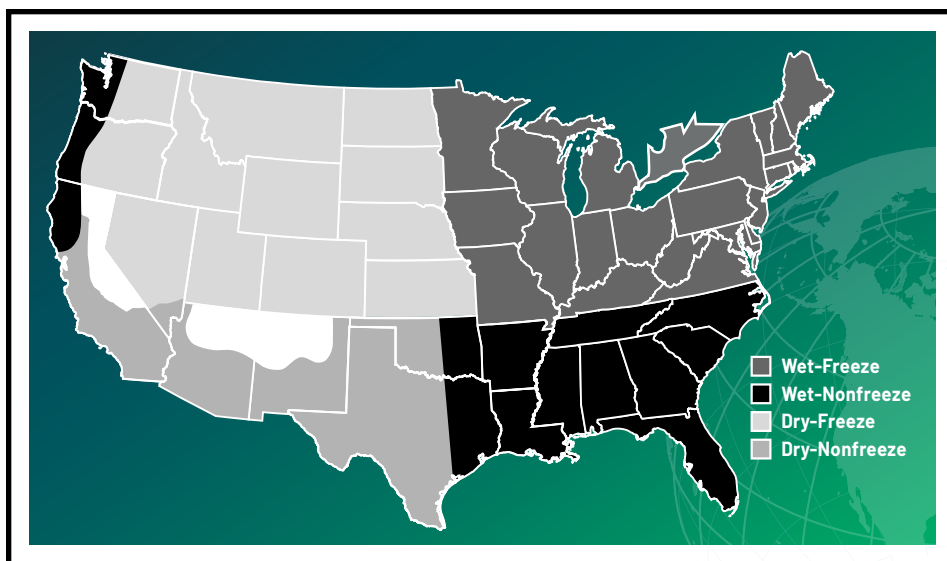
Figure 6. LTPP/AASHTO Climate Regions [13]

Legend: Wet-Freeze, Wet-Nonfreeze, Dry-Freeze, Dry-Nonfreeze

This prescriptive approach to background data sets eliminates the potential for uncertainty due to differences in background data set choices. All calculated GWP indicators in this study use the same background data sets, leaving foreground or parametric data as the main source of uncertainty. Any uncertainty or error in the LCI used to estimate the A1 impacts will uniformly impact the estimation of GWP for all asphalt mixtures and not advantage or disadvantage any mix evaluated under this system.

Epistemic Uncertainty

The PCR for Asphalt Mixtures prescribes all background LCI datasets using the USLCI except where there is a data gap. In addition, the underlying LCA modeling method is implemented in the Emerald Eco-Label tool and is the same for all asphalt mixture EPDs. This consistency has played a significant role in reducing the impact of epistemic uncertainties on GWP and benchmarking results. The data prescriptions can be found in Annex 1 of the Asphalt Mixtures PCR.[14] The only differences in background LCI are due to intentional choices that reflect regional differences.

Electricity data is drawn from the 2019 NETL dataset, which accounts for production, power sharing, and transmission to the user and provides at-user impact information at the balancing authority level. Based on the structure of this database, more than one balancing authority may be reported as servicing a particular plant location. In such a case, the balancing authority with the highest GWP value is used.

Consequently, in the current version of the PCR, the impacts from the A1 category are not influenced by any exogenous uncertainty and can therefore be treated as a deterministic quantity that can be calculated as a function of the mixture design. However, this situation may change in future as asphalt mixture EPDs incorporate supply chain specific EPDs for upstream materials like asphalt binder and aggregate. Even then, the impact associated with A1 will remain a deterministic quantity for a given design and supplier(s). However, with a large number of EPDs per mix ingredient type, the characterization for A1 impacts may require additional treatment to continue to ensure fair benchmarking. That said, at this time the characterization of uncertainty is limited only to A2 and A3 impacts that are influenced by exogenous factors.

---

[13] Chapter 7. Recommendations - Evaluation of Long-Term Pavement Performance (LTPP) Climatic Data for Use in Mechanistic-Empirical Pavement Design Guide (MEPDG) Calibration and Other Pavement Analysis , May 2015 - FHWA-HRT-15-019. (2015, May). Federal Highway Administration. https://www.fhwa.dot.gov/publications/research/infrastructure/pavements/ltpp/15019/008.cfm

[14] "Prescribed Upstream (Secondary) Data Sources." https://www.asphaltpavement.org/uploads/documents/EPD_Program/NAPA_PCR_AsphaltMixtures_v2_Annex1_v2.pdf

Uncertainty Assessment Method

The EPA's Interim Determination and the GSA's quintile-based threshold[7] are referenced in this study. As originally conceived, this method was intended to account for uncertainty in GWP reporting. The method proposed here better accounts for uncertainty through more accurate modeling of the asphalt mixture supply chain. This contrasts with the method originally proposed by the GSA that prescribed modification of GWP values reported in EPDs by applying an arbitrary "uncertainty factor" to the GWP.

In this study, we develop a metric for identifying "substantially lower" levels of embodied carbon through application of the EPA quantile method to each life cycle phase. This phase-by-phase application of the method allows the development of regionalized benchmarks such that materials are judged as "substantially lower" levels of embodied carbon as compared to those designed and produced under similar geographic and technological circumstances.

Following the system established by GSA, the average, median, 40th percentile, and 20th percentile GWP were calculated for each region or category. In all cases for A2 and A3, states or regions with fewer than 5 participating sites or 3 participating organizations were excluded from reporting and identified as high priority for additional data gathering. The 20% thresholds were calculated via the quintile measurement, such that if 100 values are reported, the lowest 20 data points would be included, and the value of the 21st number is the threshold. The 40% thresholds are then the 41st number in this example. The 50% thresholds reported follow this method and are also the mathematical median. Finally, arithmetic means and medians were both calculated and are reported as averages.

While it is unclear if the GSA threshold for the "Better Than Average Limit" is based on arithmetic mean or median, this study assumes that the GSA used the arithmetic mean. This report includes both mean and median values whenever presenting 20%, 40%, and "Better Than Average" benchmarks for ease of comparison.

To characterize the causes of uncertainty, the GWP values for A2 and A3 were calculated for the participating locations in the benchmarking data. Additionally, the set of EPDs published in the Emerald Eco-Label tool as of July 2023 were analyzed using the GSA's approach to serve as a basis of comparison to the proposed methodology. Numerous comparisons were performed within the benchmarking data, and against the published EPDs. The following population and sub-population comparisons were evaluated for statistical significance:

Published EPDs [A1 – Upstream Materials]

- A1 in Colorado vs A1 in all other states (chosen to highlight the role of hydrated lime in driving GWP because Colorado includes hydrated lime in many mix specifications)

Benchmarking data [A2 – Materials Transport]

- Each climate region A2 vs national average A2
- State level A2 vs national average A2 (performed on states where initial screening showed large variance from national average)
- National average A2 vs average of published EPDs A2

Benchmarking data [A3 – Production Data]

- Each climate region A3 vs national average A3
- Each climate region A3 vs other climate regions A3
- State level A3 vs climate region A3 (where sufficient state data existed)
- National average A3 vs average of published EPDs A3

The benchmarking survey did not address the full spectrum of upstream suppliers and mixture designs and were therefore not subject to analysis at the A1 level. However, this does not present a limitation due to the deterministic nature of the A1 impacts, easily calculable from a given mixture design or estimated based on agency specifications.

All comparisons of data used the two-tailed t-test to determine if differences between populations were of statistical significance. P values below 0.05 were deemed significant. This test is considered adequate for comparisons of means of non-normal distributions if the sample size is sufficiently large (n>35). All populations and sub-populations in this study met this requirement except for plants in the Dry Freeze region (n=30), and those in several individual states. The implications of low sample size for those populations are covered in Section 9: Discussion. A full listing of all comparisons and values is in Appendix 3: Statistical Test Outcomes.

c. O3: Proposed Benchmark Calculation Method

This report seeks to establish the use of a per-phase categorization and analysis framework based on the mathematical structure of EPD reporting:

$$GWP_{bench} = (x_1y_1 + x_2y_2 + \cdots + x_ny_n) + A2 + A3$$

Equation 1: Breaking out A1 from cradle-to-gate GWP calculation.

$$s.t. \sum_{i=1}^{n} x_i = 1 \, tonne$$

Equation 2: Mass-based constraint on Equation 1

Where:

- GWP_{bench} = Total reported A1-A3 life cycle impacts
- x_i = Total amount of ingredient i in tonnes per tonne of mix
- y_i = GWP intensity of one tonne of ingredient i
- A2 = Total A2 GWP
- A3 = Total A3 GWP

The separation of the benchmark by the three phases A1, A2 and A3, allows the individual characterization of uncertainty in phases A2 and A3, while calculating the impacts from A1 formulaically as a function of the mixture designs. The methodology for the calculation of GWP

from upstream materials in A1 is outlined in detail in the LCA supporting the PCR.[5]

The following classification system has been used to categorize mixture types and establish candidate regionalization schemes:

A1 categories were established for four nominal mixture types that correspond to virgin mixture design parameters that significantly influence GWP. These four mixtures are representative of mixture designs that are specified by agencies across the United States based on local preference and performance requirements.

> **A1 impacts are deterministic and depend on mix design.**
>
> Agencies can use GWP impact factors for each mix ingredient type (e.g. RAP, virgin aggregate, or binder) to set 'baseline' A1 GWP values based on local mix specifications. (Table 1)

The total asphalt binder content for the nominal mixture types is 5.5% based on the typical asphalt binder content for a 9.5-mm dense graded surface mix, one of the most common generic mixture types produced across the U.S. The mixture types are as follows:

- Virgin aggregate (94.5%) with neat asphalt binder (5.5%)
- Virgin aggregate (93.5%) with neat asphalt binder (5.5%) and hydrated lime (1%)
- Virgin aggregate (94.5%) with 3.5% SBS modified asphalt binder (5.5%)
- Virgin aggregate (93.5%) with 3.5% SBS modified asphalt binder (5.5%) and hydrated lime (1%)
- These same mixes with 21.87% reclaimed asphalt pavement (RAP) content[15]

For each mixture type, a baseline was set assuming no recycled material content in the mixture such that all binder and construction aggregate content

[15] Williams, B.A., J.R. Willis, & Shacat, J. (2022). *Annual Asphalt Pavement Industry Survey on Recycled Materials and Warm-Mix Asphalt Usage: 2021, 12th Annual Survey* (IS 138). National Asphalt Pavement Association, Greenbelt, Maryland. DOI:10.13140/RG.2.2.23149.26081

will be virgin. This represents a conservative baseline mixture design, as industry surveys and EPD data indicate a higher incidence of recycled content use. RAP-containing mixes in this study use the stated 21.87% RAP content. Similarly, RAP represents 21.9% of a mixture by mass as of NAPA's 2021 industry-wide survey[15] and a value of 19.6% is reflected in average EPD data reported by the Emerald Eco-Label tool and analyzed as part of this study. The impacts associated with A1 can be calculated as a deterministic value for each mixture case.

A2 categories were established by analyzing response data on a state-by-state basis followed by identification of clusters of dissimilar performance. The working hypothesis for this study is that regions with low availability of high-grade construction aggregates would have a significantly higher A2 GWP due to the need to import aggregates from distant producers. This hypothesis was tested statistically.

A2 GWP impacts were calculated based on the combined transport impacts of aggregate, binder, and RAP. Aggregate and binder transport distances were directly reported in the benchmarking tool. Average RAP transport distance was estimated at 6.74 miles based on industry-wide surveys conducted by NAPA.

A3 categories were established regionally using both state boundaries and the four AASHTO climate regions (Figure 6). State boundaries provide a more granular view into regional energy consumption patterns and electricity grid performance, while AASHTO climate regions allow a high-level evaluation of the role of climate on plant performance.

Results from the analysis and hypothesis testing are reviewed in Section 9.

d. Process for Future Updates

As data collection, processing, and reporting for both EPD and benchmarking data are performed using the Emerald Eco-Label tool, future updates to this benchmark can be made at reasonable intervals. At a minimum, this will be concurrent with PCR updates to reflect any significant changes in the PCR, which happens at least every five years. Updates can occur more frequently when market conditions allow or interim updates to the PCR drive the need for new revised data. Preferably, these updates will occur as often as possible while balancing the burden of data collection and reporting, with the goal of continually increasing participation. Updates may provide value to the industry by including more specific climate region data, updates to primary energy use, or reflect efforts by industry to reduce impacts.

8. INDUSTRY AVERAGE BENCHMARK

This study reports a single indicator, GWP. Results for each of A1, A2, and A3 follow, based on the methodology described. Unlike A2 and A3, impacts for A1 are entirely deterministic, not stochastic, and should not have thresholds or A1-A3 threshold contributions set based on regional GWP distributions. Instead, we provide adjustment factors that can be used alongside 'baseline' mix designs as presented in Section 9 to establish reasonable A1 GWP expectations based on the design envelope allowed by local mix design specifications. For example, if a state agency allows up to 20% RAP, then a reference 'best practice' 20% RAP mix design can be established and its GWP estimated using the factors in Table 1. A state could also use its reported average from the annual industry-wide survey from NAPA and FHWA[2]. A producer using less than 20% RAP would not realize the full possible GWP benefits of allowable RAP use. Additional detail on how baseline mix designs and adjustment factors should be used during threshold setting is provided in Section 10.

Adjustment factors for major materials were developed from the upstream data sets required by the PCR and Equation 1. Fundamental to the application of Equation 1 is the fact that mixture design and life cycle assessment is performed on a mass basis. That is, each ingredient represents a percentage of the total mixture by mass, and adjusting a mixture ingredient amount by, for example, +1% requires a corresponding reduction in another mixture ingredient or ingredients of –1%. Table 1 provides the list of mixture ingredients described earlier, the mixture ingredients they displace when their mass fraction increases, and a GWP adjustment factor for one metric tonne (short ton) of mixture given this adjustment. A more complete list of mixture ingredient GWP intensities can be found in Appendix 2 and applied using Equation 1 subject to the constraints of Equation 2. Discussion of increasing the binder mass fraction of SBS can be found in Section 9. Example reference mixture designs were developed using this technique and are presented in Table 2.

Table 1. Each component of an asphalt mixture has a different impact per unit mass, with some significantly higher than others.

A1 Material	Mass balanced with	GWP Intensity kg CO2e/tonne ingredient (*/shtn)	Adjustment factor for using ingredient for additional 1% of mixture by mass kg CO_2e/tonne mixture (*/shtn)
Neat Binder	Aggregate	631.51 (573.06)	+6.30 (+5.71)
3.5% SBS Modified Binder	Aggregate	758.71 (688.49)	+7.57 (+6.86)
Lime	Aggregate	1389.0 (1259.9)	+13.87 (+12.58)
RAP	Aggregate + Neat Binder	0.781 (0.710)	–0.357 (–0.325)
Aggregate (USLCI, prescribed)	Neat Binder	1.94 (1.761)	–6.30 (–5.71)

Table 2. Mix designs used for A1 study. Values are reported in % of total mixture by mass and kg CO_2e per tonne mix.

|	Aggregate (% mixture mass)	Neat Binder (% mixture mass)	Modified Binder (% mixture mass)	Lime (% mixture mass)	RAP (% mixture mass)	A1 GWP kg CO_2e / tonne
Virgin	94.5	5.5	0	0	0	36.57
With RAP	73.83	4.3	0	0	21.87	28.74
With Lime	93.5	5.5	0	1	0	50.44
With Lime, RAP	72.83	4.3	0	1	21.87	42.61
With 3.5% SBS	94.5	0	5.5	0	0	43.56
With SBS, RAP	73.83	0	4.3	0	21.87	34.21
With 3.5% SBS, Lime	93.5	0	5.5	1	0	57.43
With SBS, Lime, RAP	72.83	0	4.3	1	21.87	48.07
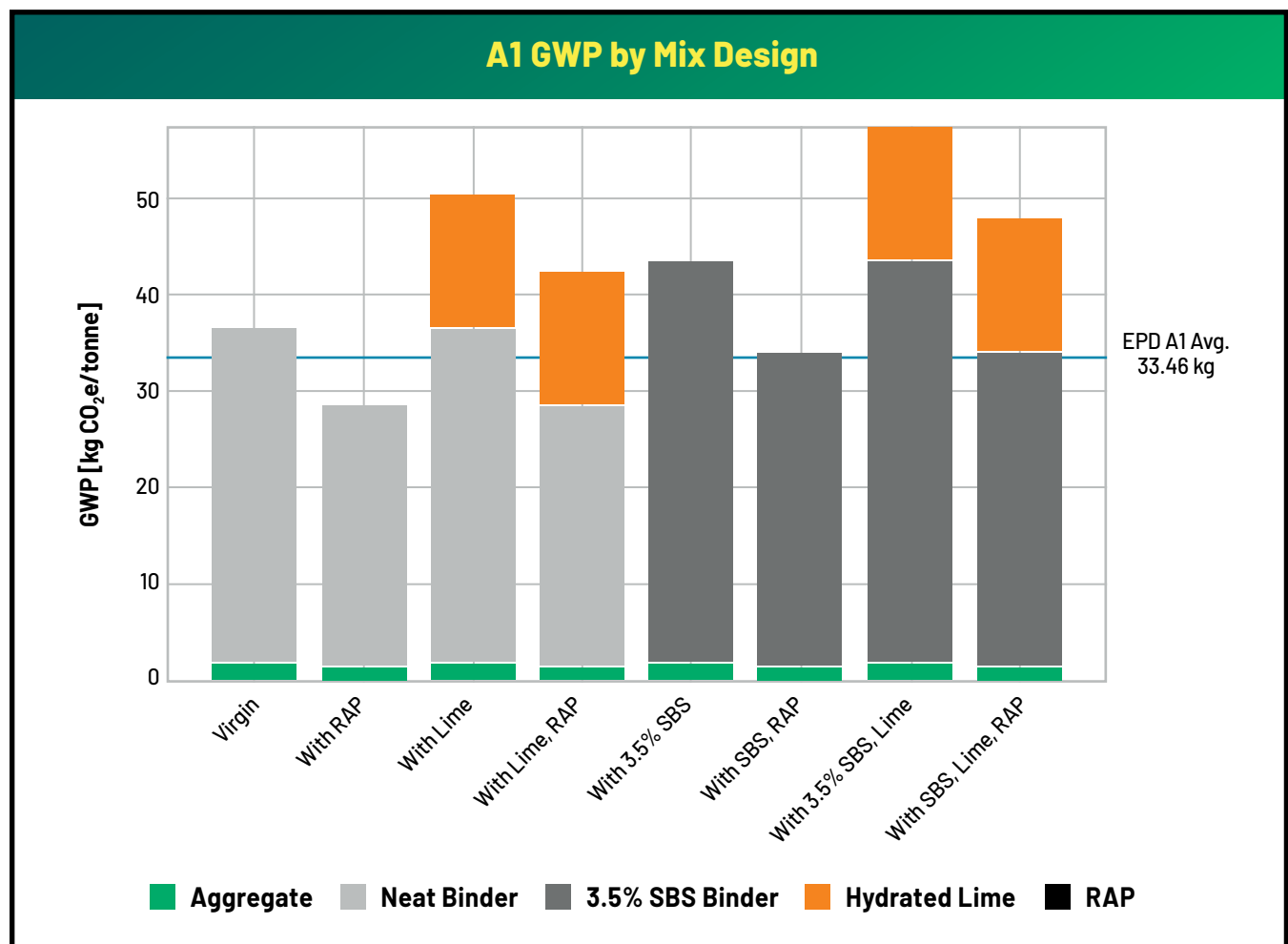
Figure 7. Modified asphalt mixtures consume a significant additional portion of a producer's carbon budget.

The following assumptions were used:

- Mixes were developed with a nominal final binder content of 5.5% by mass.
- Any RAP used in a mixture design was assumed to have a 5.5% total binder content by mass, and binder replacement was calculated assuming 100% availability of RAP-derived binder.
- RAP-containing mixes were developed using 21.87% RAP, the average value reported by producers during this benchmarking study.

This value agrees with the 2021 reported average RAP content of 21.9% and the general year-over-year trend of increasing RAP use.

- Mixes with 3.5% SBS used the USLCI inventory for terminal-blended 3.5% SBS binder.

The impacts for each of these mixture designs were estimated using impact factors derived from the NAPA Emerald Eco-Label EPD tool and USLCI in accordance with the PCR for Asphalt Mixtures. The relative GWP performance of each reference mix is shown in Figure 7.

Table 3 shows the A2 GWP impacts by state, also including the top 20% and 40% benchmarks, on a per metric tonne (per short ton) basis. If Florida and Louisiana are categorized together, their median A2 value is 30.4 kg CO_2 e./tonne, while the rest of the country's median A2 value shrinks to 3.0 $CO2$ e./tonne.

The A3 GWP impacts by AASHTO climate region, including top 20% and 40% benchmarks, are given in a per metric tonne (per short ton) basis in Table 4.

Table 3. Top quintiles and averages in the benchmarking study of A2 GWP by state, in kg CO_2 e. / tonne mix

A2 by State	Florida kg CO_2 e./tonne (kg CO_2 e./shtn)	Louisiana kg CO_2 e./tonne (kg CO_2 e./shtn)	All Others kg CO_2 e./tonne (kg CO_2 e./shtn)
20%	3.8 (3.5)	16.7 (15.1)	1.0 (0.9)
40%	19.5 (17.7)	24.6 (22.3)	2.1 (1.9)
50%	37.4 (34.0)	29.2 (26.5)	3.0 (2.7)
Average	42.0 (38.1)	29.7 (26.9)	4.6 (4.2)

Table 4. Top quintiles and averages in the benchmarking study of A3 GWP by climate region, in kg CO_2 e. / tonne mix

A3 by AASHTO Region	Wet No freeze kg CO_2e/tonne (kg CO_2e/shtn)	Wet Freeze kg CO_2e/tonne (kg CO_2e/shtn)	Dry No freeze kg CO_2e/tonne (kg CO_2e/shtn)	Dry Freeze kg CO_2e/tonne (kg CO_2e/shtn)
20%	23.2 (21.1)	20.9 (19.0)	17.5 (15.9)	21.8 (19.8)
40%	25.6 (23.2)	22.8 (20.7)	19.9 (18.1)	24.4 (22.1)
50%	26.1 (23.7)	23.6 (21.4)	21.5 (19.5)	26.3 (23.8)
Average	27.5 (24.9)	24.7 (22.4)	22.9 (20.8)	27.2 (24.7)

9. DISCUSSION

Applying the benchmarking process outlined in Section 7 reveals key insights into the variability and uncertainty in A1, A2, and A3 impacts for a given mix.

Across agency jurisdictions there is a high level of variation in mix specifications, driven by state and local DOT policy, leading to significant controllable variation in aggregate baseline A1 impacts. Our method allows a particular agency to control for local specifications and deterministically establish a fair target to provide for contractors for A1 emissions. It further provides agencies with insight into how specifications can influence emissions.

The data in this study shows significant variance in A2 impacts between a small subset of reporting US states and all others. Controlling this variability is possible by developing state-specific A2 benchmark distributions in cases where the state's A2 distribution differs significantly from the national distribution. This variability is largely uncontrollable by a given agency or contractor, as it is driven by local geology and the inability to change project location.

Climate zone is shown to have a small but statistically significant influence on A3 impacts. While variation in distribution parameters (the average and 50th, 40th, and 20th percentiles) between the particular climate region breakdown is smaller than state-by-state variation in A2, it does exist and further refinement of climate zone categories warrants investigation.

A1 (Upstream Materials)
Local specifications can drive carbon impacts by requiring the use of materials and mixture designs that significantly impact the embodied carbon of asphalt mixtures. To evaluate the effect of mixture design specifications, eight 'synthetic' mixture designs incorporating different uses of virgin material, recycled material, hydrated lime, and modified binder were developed and shown in Section 8. This study demonstrated that the use of polymer modified binders, RAP, and hydrated lime each drive A1 mix GWP.

Polymer modified binders used at the same volumetric values as neat binders incur higher GWP impacts. Binder that has been modified at the terminal to include 3.5% SBS has GWP impacts 20.1% higher than neat binder, according the USLCI data prescribed by the PCR for Asphalt Mixtures and acquired from the Asphalt Institute's 2019 LCA.[16] The impact of SBS use can be approximated by 1.675 kg of CO_2e / short ton per % SBS for the baseline mixture design.[5] This factor may require revision for additional mixture designs.
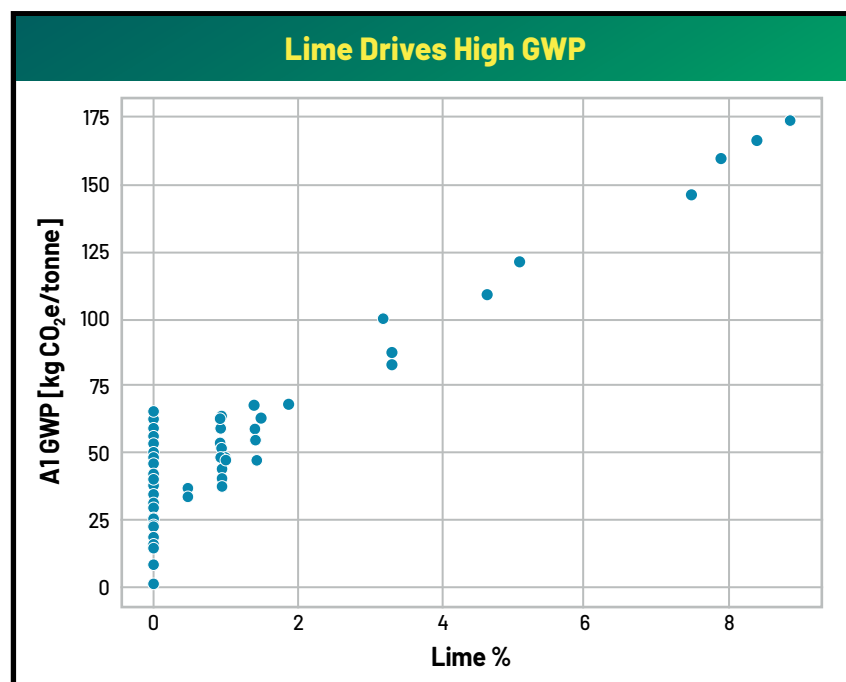

Figure 8. Correlation between high GWP and use of hydrated lime in asphalt mixtures.

---

[16] Wildnauer, Mulholland, and Liddie, Life Cycle Assessment of Asphalt Binder. https://trid.trb.org/view/1645171

The use of RAP influences overall mixture GWP by replacing both virgin aggregate and binder. While aggregate represents the majority of mass replaced, GWP reductions are primarily due to binder replacement. Given the use of a neat binder and some percentage of RAP, the amount of GWP reduction a mixture will see can be approximated as 6.52 kg CO_2e/tonne (5.92 kg CO_2e/short ton) per % asphalt binder replacement (% ABR).[5] This dynamic is borne out by this current study, as all mixes with RAP show a significant reduction in A1 GWP compared to their virgin counterparts.
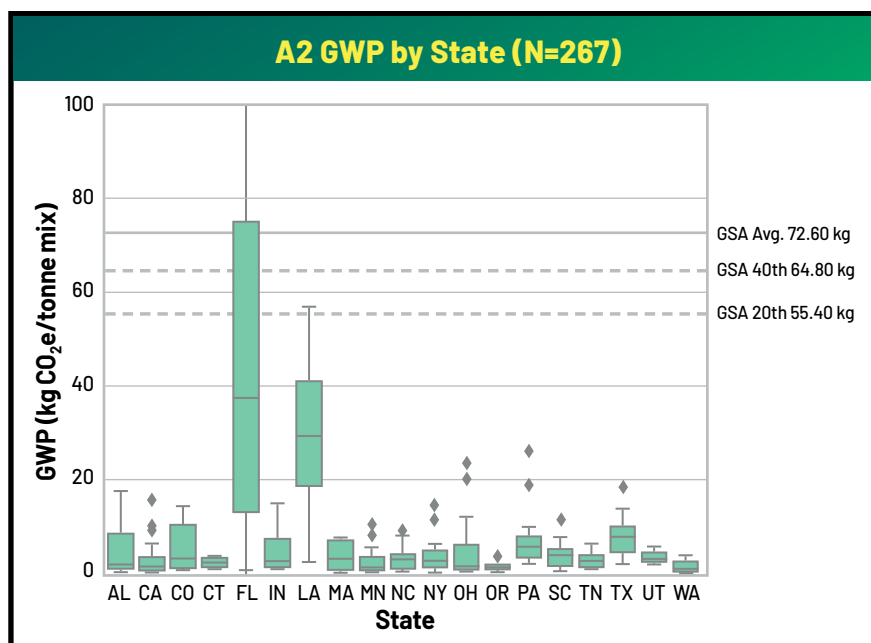

Figure 9. Aggregate-poor states such as Florida and Louisiana have very high A2 impacts.

When used in these mixture designs, hydrated lime directly displaces virgin aggregate on a 1:1 mass basis. While the lime fraction is low (1% of total mixture mass), lime has a GWP intensity 715x that of virgin aggregate (Table 1). An evaluation of high-A1-GWP mixes in the Emerald Eco-Label EPD tool shows that all mixes over 75 kg CO_2e /tonne A1 include large amounts of lime (Figure 8).

These three materials represent clear examples of how agency mixture design specifications (e.g.. those that specify the inclusion or limitation of hydrated lime, polymer modifiers, or RAP) have significant effects on both A1 and cradle-to-gate GWP. If a producer operates in a region specifying both polymer modification and lime, then it will be exceedingly difficult for them to meet any threshold based on national or regional cradle-to-gate averages.

A2 (Materials Transport)
Significant regional variability exists for A2, as some states have far greater transport impacts than the national average (Figure 9). Since the GSA threshold is set for the sum of A1 to A3 impacts, these cases of unavoidably high transport impacts are a strong argument for separate benchmarking values for the A2 phase.

Florida and Louisiana had A2 median values of 37.4 kg CO_2e/tonne and 29.2 kg CO_2e/tonne, respectively. The median of all other states was 3.0 kg CO_2e/tonne. In some cases, the transport phase impacts alone were enough to exceed even the most forgiving GSA threshold (72.6 kg CO2e/tonne). Industry consensus is that this extreme degree of transport is driven primarily by a lack of locally available road-quality aggregate and additionally in some cases by economic factors such as 1:1 local production:consumption in nearby regions, thus requiring acquisition from farther afield.

Some states with anecdotally high transport distances did not have enough data for this study to draw conclusions. It is recommended that these states (notably Alaska) be targeted in a second round of data gathering.

A3 (Production Phase)
A3 impacts are driven by the consumption of energy by the production facility. Energy sources include grid electricity and fuels combusted in the asphalt burner, oil heaters, and onsite equipment.

AASHTO climate regions had statistically significant effects on GWP per short ton of asphalt (Figure 10). The Wet No Freeze region had the highest average

value (27.5 kg CO_2e/tonne), and Dry No Freeze the lowest (22.9 kg CO_2e/tonne)(Table 4).

Examining the A3 distributions by state appears to reveal elevated impacts in the states of Alabama and Utah, as shown in Figure 11.

Note that Figure 11 displays a subset of the sample, n=267 of a total sample size of 335. To protect potentially sensitive data, A3 values from states with fewer than five unique producers participating are not displayed.

Analysis of the Alabama A3 data revealed a marked statistically significant increase above the average GWP for Wet No Freeze. The Alabama average was 31.2 kg CO_2e/tonne, while the Wet No Freeze average was 27.5 kg CO_2e/tonne.
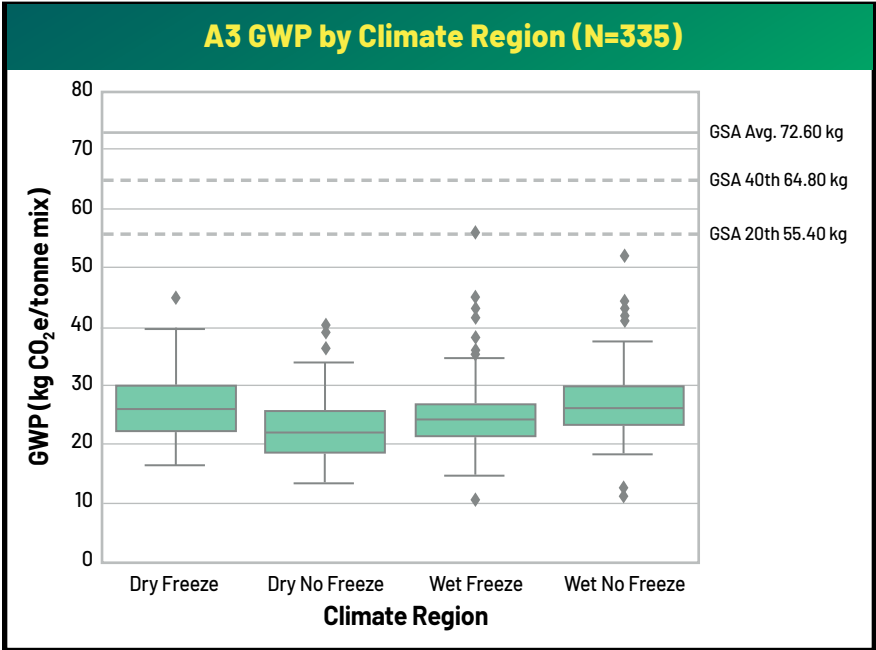
Figure 10. Dry Freeze and Wet No-Freeze regions exhibited higher average GWP values

The situation appeared even more stark in Utah with the average value of 31.8 kg CO_2e/tonne and a regional average of 22.9 kg CO_2e/tonne. But statistical tests do not show this difference to be significant. This may be because the number of reporting producers is low in Utah (n=7) and the sample size for the Dry Freeze region is the smallest of the climate regions by a wide margin (n=30). At these levels statistical variance is difficult to prove. It is recommended that states like Utah and other states in the Dry Freeze climate region be targeted during the second round of data collection.

Figure 11. A3 GWP by State

10. USING BENCHMARKS IN PUBLIC PROCUREMENT SCENARIOS

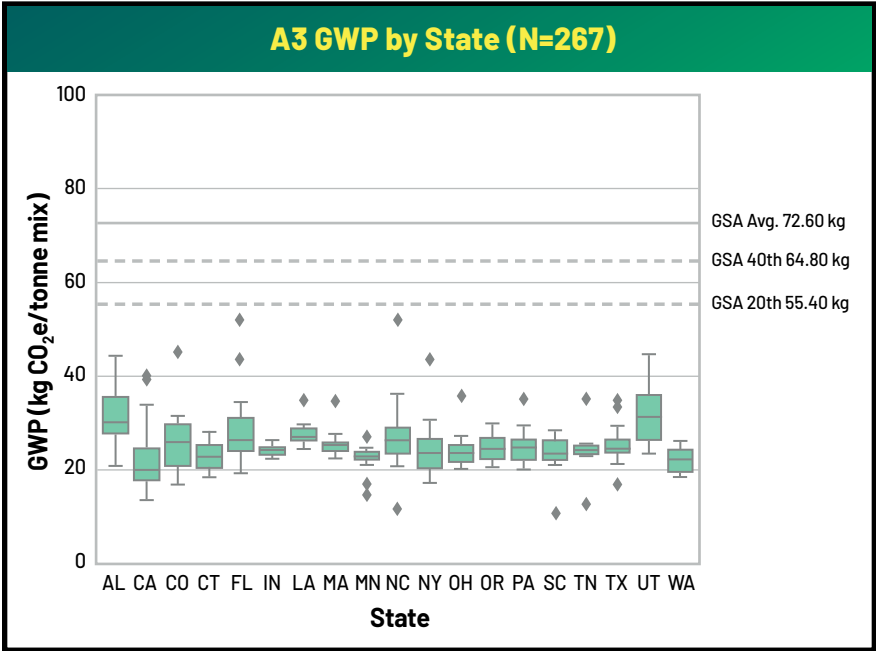The benchmarks in this study are intended to allow agencies to account for variabilities in climate, geology, and agency specifications when determining whether a specific asphalt mixture has substantially lower embodied carbon than estimated industry averages. A crucial takeaway of this study is that a national average benchmark cannot be used in procurement scenarios without accounting for differences in regional factors and practices that are beyond a contractor's control. Hence, the benchmarks are broken into life cycle phases to reduce confounding factors and can be summed to make a set of GWP values at the 20th percentile, 40th percentile, and average, that is regionally appropriate for agency-defined mix types.

Take the example of an agency in Colorado setting benchmarks for the 'virgin' asphalt mixture described in Table 4 (5.5% virgin binder and 94.5% virgin aggregate). This agency would use the baseline mixture values from Table 4 for A1, add the national average values for A2, and add the Dry Freeze regional values for A3. This approach is shown in Table 5. Each column provides the values used for each phase. Working from the left, the A1 column has no 20% / 40% / 50% because it is the calculated value for the 'virgin' mix. The next column are the national 20% / 40% / 50% benchmark values established in this study for A2, followed by the column showing the 20% / 40% / 50% benchmark values for the climate region that includes Colorado

Table 5. Benchmark for 'virgin' mixture in Colorado

[all values in kg CO_2 e. / tonne]	A1 (Baseline Mix)	A2 (National Benchmark)	A3 (Dry Freeze)	A1–A3 Total (Proposed Method)	Current A1–A3 GSA Thresholds
20%		1.0	21.8	**59.4**	55.4
40%	36.57	2.1	24.4	**63.0**	64.8
50%		3.0	26.3	**65.8**	x
Average		4.6	27.2	**68.5**	72.6

Table 6. Benchmark for 'virgin' mixture + 1% lime in Colorado

[all values in kg CO_2 e. / tonne]	A1 (Baseline Mix)	A2 (National Benchmark)	A3 (Dry Freeze)	A1–A3 Total (Proposed Method)	Current A1–A3 GSA Thresholds
20%		1.0	21.8	**73.3**	55.4
40%	50.44	2.1	24.4	**76.9**	64.8
50%		3.0	26.3	**79.7**	x
Average		4.6	27.2	**82.3**	72.6

Table 7. Benchmark for 'virgin' mixture in Florida

[all values in kg CO_2 e. / tonne]	A1 (Baseline Mix)	A2 (Florida Benchmark)	A3 (Wet Non Freeze)	A1–A3 Total (Proposed Method)	Current A1–A3 GSA Thresholds
20%		3.8	23.2	**63.6**	55.4
40%	36.57	19.5	25.6	**81.6**	64.8
50%		37.4	26.1	**100.1**	x
Average		42.0	27.5	**106.1**	72.6

(Dry Freeze). The next column, **in bold**, sums the prior columns, creating the final values for benchmarking this mixture in Colorado. The GSA thresholds are provided for clarity.

This method produces a lower GWP benchmark than the current GSA thresholds for this mix, bringing the average value 7% below the GSA average.

Now take the example of this Colorado agency setting benchmarks for the same mixture but with 1% lime by mass added (and 1% aggregate removed). Hydrated lime is a common additive that is specified by Colorado DOT and can also have significant effects on GWP. Referring to Table 1, the adjustment factor for adding 1% lime is 13.87 kg CO_2e/tonne. Adding this to the 'virgin' baseline A1 value of 36.57 kg CO_2e/tonne yields a new A1 factor of 50.44 kg CO_2e/tonne. The A2 and A3 benchmarks for Colorado remain unchanged, and new threshold for this mixture is shown in Table 6.

This results in benchmark values that are higher than the current GSA thresholds, with the average value 12% greater than the GSA average. This method accounts for the local specifications that are driving GWP (in this case, the addition of lime).

Florida allows for a different example. Because of the local geology, road quality aggregate in many parts of the state must be transported far greater distances than most other parts of the country. Florida is also in the Wet No Freeze climate region. Hence, benchmarks for that same 'virgin' mixture produced in Florida are calculated by adding the A1 baseline from Table 4, the A2 benchmarks for Florida, and the Wet No Freeze values for A3 (Table 7).

The transport distances in Florida increase the benchmarks above the current GSA thresholds – averaging 45% higher. This example is particularly telling because most of Florida would be unable to meet the GSA thresholds, and therefore prevented from accessing IRA funds, simply by the nature of its local geology if no alternate benchmarking scheme is employed.
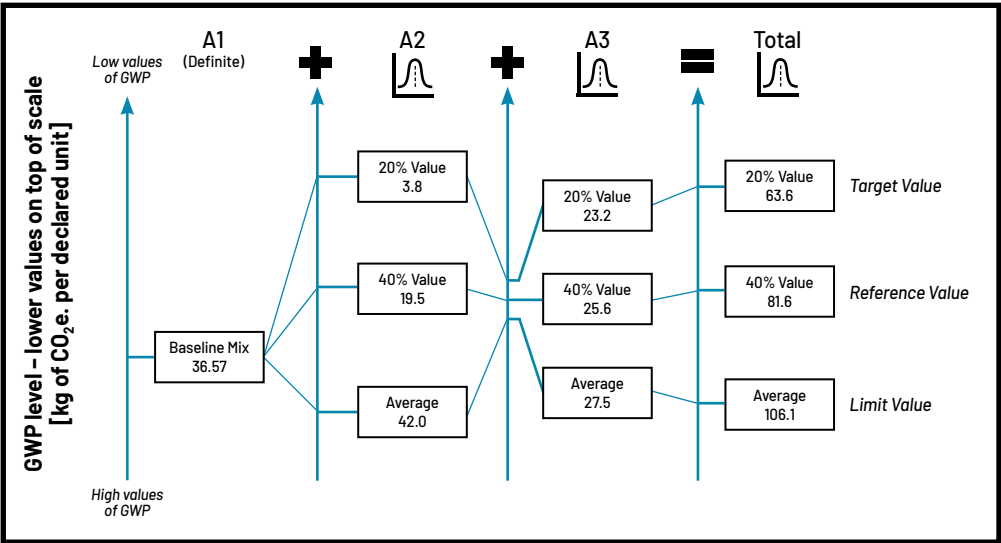

Figure 12. This technique and the values developed in this Florida case study map to the benchmarking approach established in ISO 21678.

11. SUMMARY AND RECOMMENDATIONS

Using the current set of published EPDs to create national thresholds without regard for regional differences or material specifications results in GWP limits that unfairly burden some regions of the country. This approach should be avoided until EPD creation has happened at a large enough scale to fully represent the industry.

Agencies looking to procure low embodied carbon asphalt mixtures under Sections 60503 and 60506 of the IRA should establish thresholds for each life cycle phase (A1, A2, and A3) that are appropriate for the agency's geographic location and material specifications. These thresholds can be added up for each agency mix type to give a single set of embodied carbon thresholds at the 20th percentile, 40th percentile, and average levels that is appropriate to the region, state, and mixture design specifications. Benchmarking by life cycle phase allows for selection of lower carbon materials while not preventing a region or state from accessing IRA funds because of the influence of factors such as agency specifications, geology, and climate, all of which are beyond the control of the contractor.

a. Limitations of this study

This study is a first attempt at establishing regionalized benchmarks for the asphalt mixture industry. The emphasis and primary takeaway from this report is the proposed underlying methodology that serves as a first step and is expected to evolve as EPD use matures in the industry and more insights are gained on their role in public procurement. Hence, the limitations discussed in this section are also possible next steps in improving the benchmarking process. In discussing the limitations, the underlying principle is articulated to provide context for future steps. The primary limitations can be discussed as follows:

- As a principle, representativeness of the data samples analyzed is critical to developing useful

benchmarks. At this time, all regions/states have not been represented equally or adequately. In Utah, this lack of representation is apparent in the analysis. However, the data reported from various other states while meeting the threshold of numbers may still fall short of representativeness due to voluntary reporting practices. While statistical tests can be conducted to assess the likely skew in the data, at present due to a lack of uniform adoption of EPDs it is difficult to establish the skew. That said, the analysis does reveal important findings such as the impact of geology on the A2 impacts reported due to a lack of material availability. In phase two of data collection, a deeper data pool can help validate (or rectify) current findings, assess skew in reporting, and/or shed new light on other relevant trends.

- In this methodology, A1 has been treated as a definite quantity because of the prescriptive nature of the current PCR and the predominant use of average LCI for upstream materials like asphalt binder. As EPDs for upstream suppliers become more available, this approach may need an amendment. As a principle, it will still be true that the upstream impacts (A1) are a definite and calculable quantity given the design and selection of materials. Indeed, for a given mixture design, a contractor can intentionally select between two asphalt binder producers based on their EPDs to deliver a mixture with target GWP. Similar selections can be made with respect to selection of additives. In effect, the A1 GWP component is likely to become more competitive than it is now. The future benchmarking process will have to account for availability of supply chain specific EPDs for different upstream materials to develop intervals for A1 benchmarks. These fundamental principles will still apply: (i) allowing contractors the ability to compete based on selection of supply chain partners and (ii) recognizing that A1 impacts are an outcome of choice that should not be treated as a random variable.

- Finally, the benchmarking process is driven by EPA's Interim Determination that only accounts for A1–A3 GWP values and does not explicitly account for performance outcomes. Clearly, the materials procured by a DOT will have to meet design requirements, negating the possibility of a race to the bottom and selection of mixtures that are competitive on GWP only. However, the current framework with its emphasis on selecting lower GWPs discourages contractors from using additives and other material science innovations that deliver significant performance improvement with potentially higher GWPs. Future work will have to account for a benchmarking framework that accounts for GWP in addition to mixture performance properties.

b. Conclusions and recommendations for future work

Based on the analysis this report recommends the following:

- Impacts of upstream materials are dependent on mixture design specifications that are often outside the control of asphalt mix producers. Hence, the use of a baseline mixture with constituent-specific adjustment factors is recommended for estimating A1 impacts. The impacts of material choices are clearly communicated using this method while also reflecting sensitivities of a mixture's GWP to locally specified mixture design.

- Impacts due to the transportation of materials are dictated by availability of materials locally which in turn is a function of local geology. For example, states like Louisiana and Florida have significantly higher A2 impacts than any other regions due to a lack of locally available aggregates with sufficient quality for use in asphalt mixtures. Hence, the use of region-specific thresholds is recommended for the A2 life cycle phase.

- Impacts due to production vary to a smaller but statistically significant degree across different AASHTO climate regions. While A3 impacts are within the control of material producers, the recommendation is to develop thresholds that account for climatic variations.

- The analysis recommends the use of a representative data set when establishing thresholds for procurement, including an intentional sampling process that targets states and regions with limited participation so far.

The contribution of the benchmarking framework discussed in this study is that it: (i) affords a web-based user-friendly data collection platform across different mixtures, (ii) can be scaled to include multiple regions, and (iii) can be updated over time to reflect changes in technology and other factors.

12. APPENDIX 1: ADDITIONAL DATA COLLECTION

Table 8. Additional data and metadata collection established during benchmarking activity.

Plant type
Whether the plant is portable
Rated Capacity
Drum diameter
Operating months per year
Operating days per week
Operating hours per day
Total operating hours
Whether operating hours is an estimate
Percent mixes modified with polymers
Short tons of mix modified with polymer
Average mix RAP %
Total RAP used
Quarry A: Total aggregate supplied
Quarry B: Total aggregate supplied
Quarry A: Total transport by mode
Quarry B: Total transport by mode
Electricity meter type
Electricity meter sharing by operation
Whether electricity meter data is from billing
Gas meter type
Gas meter sharing by operation
Whether gas meter data is from billing
Notes

13. APPENDIX 2: ADDITIONAL MATERIAL FACTORS

Data gaps for A1 materials are being filled at an exceptional rate as upstream suppliers work to quantify their own impacts. Below is an incomplete list of suppliers that have created product specific EPDs for inclusion in the calculation of the environmental impacts of asphalt mixtures and links to their EPDs.

Cargill: https://www.cargill.com/bioindustrial/anova-asphalt
Ingevity: https://www.ingevity.com/markets/pavement/evotherm/
Surface Tech: https://surface-tech.com/documents-all/

Table 9. Additional Material Factors for A1

A1 Material	GWP Intensity kg CO_2e/tonne (*/shtn)
Neat Binder	631.51 (573.06)
3.5% SBS Modified Binder	758.71 (688.49)
Lime	1388.4 (1259.9)
RAP	0.781 (0.710)
Aggregate (USLCI, prescribed)	1.94 (1.761)
Aggregate (forthcoming, explosive mining)	5.32 (4.83)
Aggregate (forthcoming, non-explosive mining)	3.76 (3.42)
Cargill Anova 1501 WMA Cargill Anova 1815 Rejuv.	1594 (1446) 1288 (1168)
Ingevity	5640 (5115)
Surface Tech	6134 (5563)
[Additional Future Ingredient types]	....

14. APPENDIX 3: STATISTICAL SUMMARIES & TEST OUTCOMES

Appendix 3 includes the typical summary statistics of mean, standard deviation, and confidence interval (on the mean) for each grouping discussed in this report. It also contains the outcomes of the statistical tests performed (not including initial screening level analysis). All comparisons were performed using a two-tailed t-test. P-values are reported in the tables below, where $p < 0.05$ was considered significant.

Table 10. Summary statistics of benchmarking A3 data, climate regions (all values in kg CO_2e/ metric tonne)

Category	Wet No Freeze	Wet Freeze	Dry No Freeze	Dry Freeze	Nationwide
Count (N)	134	127	44	30	335
Mean	27.49	24.73	22.89	27.24	25.82
Standard Deviation	6.10	5.99	6.30	7.42	6.41
Confidence Interval	1.04	1.05	1.91	2.77	0.69

Table 11. Summary statistics of benchmarking A2 data (all values in kg CO_2e/metric tonne)

Category	All	Minus FL + LA	Just FL	Just LA
Count (N)	335	299	30	6
Mean	8.4	4.6	42.0	29.7
Standard Deviation	15.8	4.9	33.9	19.4
Confidence Interval	1.7	0.6	12.6	20.3

Table 12. Summary statistics of benchmarking A3 data, Alabama and Utah only (all values in kg CO_2e/ metric tonne)

Category	Alabama A3	Utah A3
Count (N)	29	7
Mean	31.2	31.8
Standard Deviation	6.0	7.6
Confidence Interval	2.3	7.1

Table 13. T-test p values for benchmarking A3 data, climate regions vs each other and nationwide

Category for comparison	Wet No Freeze	Wet Freeze	Dry No Freeze	Dry Freeze	Nationwide
Count (N)	134	127	44	30	335
Wet No Freeze	n/a	0.0003	0.0001	0.7747	0.0077
Wet Freeze	....	n/a	0.0943	0.0937	0.0876
Dry No Freeze	....	....	n/a	0.0111	0.0053
Dry Freeze	....	....	....	n/a	0.3224

Table 14. T-test p values for benchmarking A3 data climate regions and nationwide vs. published EPDs A3 data

Category for comparison	Wet No Freeze	Wet Freeze	Dry No Freeze	Dry Freeze	Nationwide
Count (N)	134	127	44	30	335
All Published EPDs A3 (n=1070)	5.61E-06	0.6840	0.0612	0.0991	0.0193

Table 15. T-test p values for benchmarking A3 data, Alabama and Utah vs. local climate regions and nationwide

Category for comparison	Wet No Freeze	Dry Freeze	Nationwide
Count (N)	134	30	335
Alabama A3 (N=29)	0.0045	....	5.73E-05
Utah A3 (N=7)	....	0.1868	0.0847

Table 17. T-test p values for benchmarking A2 nationwide vs. published EPDs A2 data

Category for comparison	Nationwide A2
Count (N)	335
All Published EPDs A2 (n=1070)	0.11

Table 16. T-test p values for benchmarking A2 data, climate regions vs nationwide

Category for comparison	Wet No Freeze	Wet Freeze	Dry No Freeze	Dry Freeze
Count (N)	134	127	44	30
All Benchmark A2 (N=335)	0.0071	4.15E-05	0.0007	0.0098

Table 18. T-test p values for benchmarking A2 data, Florida, Louisiana, and all but FL & LA vs. nationwide

Category for comparison	All except Florida & Louisiana	Florida	Louisiana
Count (N)	299	30	6
All Benchmark A2 (n=331)	3.54E-05	7.85E-06	0.0435

Table 19. T-test p values for published EPD A1 data, Colorado vs nationwide

Category for comparison	Colorado EPDs A1
Count (N)	79
All Published EPDs A1 (n=1070)	3.29E-32