12-2017

Using Historic Pavement Performance Data to Identify and Select Sites for Asphalt Mixture Design Studies

Allyson Richey

Using Historic Pavement Performance Data to Identify and Select Sites for Asphalt Mixture Design Studies

By

Allyson Richey

University of Arkansas, December 2017

1) Introduction

Since 1997 Arkansas asphalt mixtures have been designed using a procedure commonly known as 'Superpave'. In a traditional Superpave mix design, the design considers only traditional volumetric parameters; however, recent advances in mix design technology emphasize mixture's ability to *perform* in its environment [1]. There are numerous factors which affect the performance characteristics of an asphalt mix. Certainly, the asphalt binder's performance is a critical variable in this system. Asphalt binders are sensitive to temperature, so much so that their primary parameter on which they are recommended is a function of both temperature and the geographic latitude in which the pavement will ultimately reside. Additionally, binder becomes more brittle as it ages, which can be exacerbated if residing in an environment outside of the binder's recommended temperature range. Performance-related tests are included in the binder grading specification to account for this behavior.

Recent evolution of Superpave mix designs marry these concepts by recommending various types of physical (performance-related) testing in conjunction with the volumetric mix design. In order for a mix to be judged to be acceptable, it must pass these tests. The testing recommended in this system is related to the pavement's cracking and rutting potential.

Currently, the Arkansas Superpave system includes only a performance-related test for rutting; it does not have a defined procedure to judge cracking performance. TRC 1802 sets out to solve this problem by identifying tests in conjunction with volumetric mix designs that accurately identify early-age cracking susceptibility by recreating mix designs currently in the field. By applying cracking testing in these recreated mixes, the results can be compared to the recorded field survey data to determine how correct these early-age cracking tests are to mixes known to be or not to be cracked. These mixes in the field act as case studies and are representatives of many of the types of mixes all over the state. However, many mixes exist in the state and not all can be tested. Therefore, of the subset of mixes as provided by ARDOT, a handful must be chosen to be representative of pavements in various states of distress and these few will be fully tested. The site selection procedure outlined in this text will identify mixes to encapsulate the performance piece, bringing Arkansas into the modern era of asphalt mixture design. The site selection procedure developed in this project can serve as a template for future studies which propose to use field performance data to identify candidate materials/mixtures for additional study.

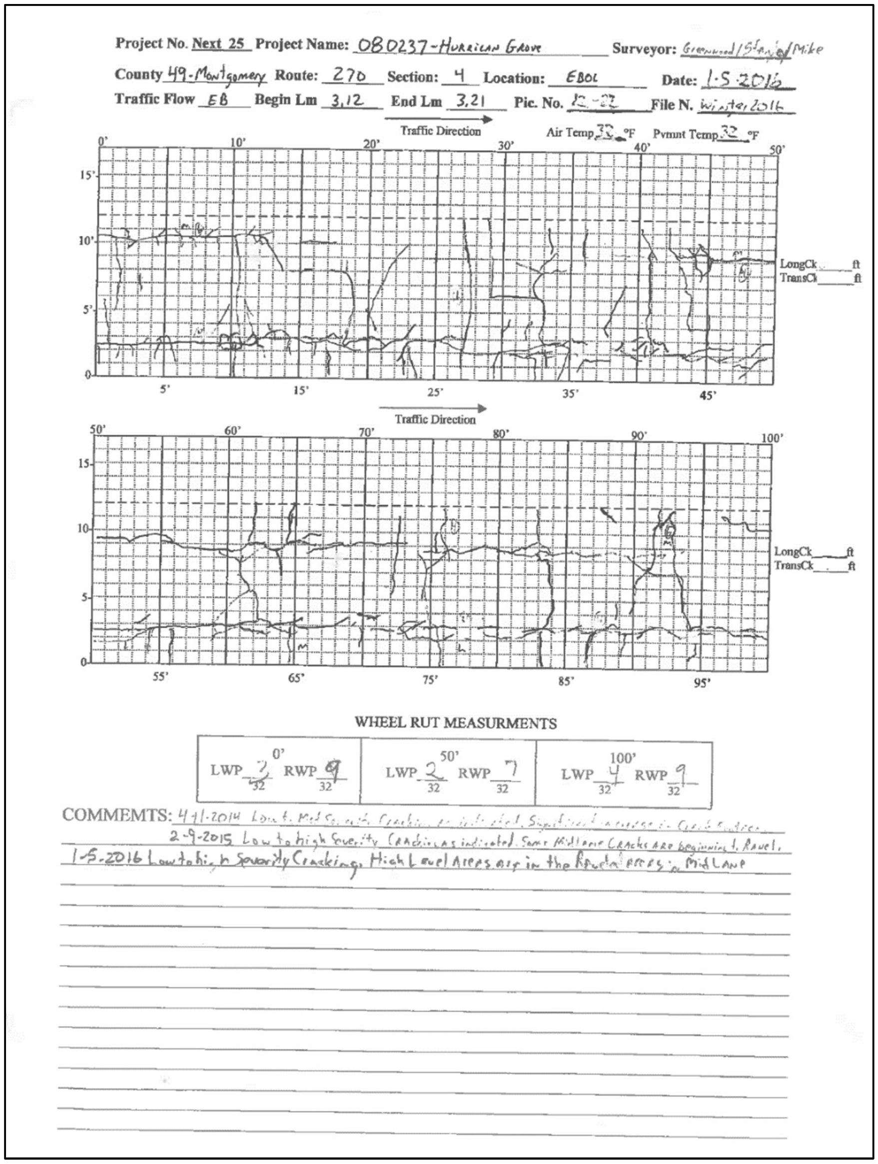
2) Data Compilation and Organization

The site selection piece of TRC 1802 is the first step in correlating volumetric and performance-based metrics. The Arkansas Department of Transportation (ARDOT) provided information on a subset of thirty projects composed of hot-mix asphalt pavements, in the form of a small database that was initially populated with identifying information useful to the state. The database was also accompanied by hard-copy scans of field data surveys taken at various points in time over the pavement's life. The thirty projects were all relatively new projects, the oldest being constructed no more than 15 years before this study. The pavements, per the field distress surveys, were usually evaluated for distresses between the months of December to March; the uniformity of this repetition varied from district to district within the state. ARDOT also provided construction plan sets for each job; these would be used later to help correlate distresses with the pavement structure for each site.

a. *Field Distress Surveys*

The most pertinent piece of data for the candidate projects was the field data surveys. The field data survey was a series of ten grids, each grid containing a total area of 850 square feet. The total area covered by a survey was 8,500 square feet. This same pavement was re-evaluated each time the pavement was surveyed. The location of the sample area was identified via log miles, a system widely used within ARDOT but typically not common with most state agencies today.

Figure 1 is exemplary of a typical field data survey received for this project. Often the field data surveys were filled out by various surveyors from year to year. This introduces an inconsistency in which the distresses were recorded—given the same pavement area and distresses, no two surveyors will record the exact same quantity and severity of pavement distresses. In addition to the visual representation of the cracking as well as the quantity, comments were often left to summarize the state of the pavement. However, these comments were often lost from year to year so the physical representation of the cracking on the grids was almost exclusively considered. If the quantity was not recorded by the surveyor, the quantity recorded was based on counting (to scale) the cracks on the grid itself. Once there was familiarity with the data initially given from ARDOT, data processing began. The first step in data processing was to create a database in which this data could be housed. Most importantly, a system needed to be created in which there was no need to reference the field data surveys themselves; the cracking data graphically available in them could be simplified and quantified into the database, expediting decision-making. Due to the numerous field data surveys and an interest in making sure this data was not aggregated, each job was given its own subset in the database. This allowed for each individual survey to be listed so the cracking could be looked at on a year by year basis. It was important to keep the data not aggregated because by being able to see the cracking progress on a year by year basis, the quality of the survey could be somewhat evaluated as well as helping identify what cracking mechanisms likely produced the cracks themselves. In the process of deciding this database structure, it became important to not only record how much cracking there was but what kind: longitudinal, transverse, or alligator cracking. These three types of cracking have different mechanisms that cause them; therefore by knowing what type of cracking the mechanism could potentially be identified. For example, if a pavement is showing an extreme amount of alligator cracking it could be symptomatic of a structural deficiency in the pavement more so than a material deficiency.


Project No. Next 25 Project Name: 0B0237-Hurrican Grove Surveyor: Greenwood/Stanley/Mike

County 49-Montgomery Route: 270 Section: 4 Location: EBOL Date: 1-5-2016

Traffic Flow EB Begin Lm 3,12 End Lm 3,21 Pic. No. 12-22 File N. Winter, 2016

Traffic Direction Air Temp 32 °F Pvmnt Temp 32 °F

WHEEL RUT MEASURMENTS

0'	50'	100'
LWP 2/32 RWP 9/32	LWP 2/32 RWP 7/32	LWP 4/32 RWP 9/32

COMMEMTS: 4-11-2014 Low to Mid Severity Cracking as indicated. Significant increase in Crack Features.
2-9-2015 Low to high Severity Cracking as indicated. Some Midlane Cracks are beginning to Ravel.
1-5-2016 Low to high Severity Cracking. High Level Areas are in the Raveled areas in Mid Lane

Figure 1. Typical Field Data Survey

As mentioned previously, the cracking data was pertinent for a multitude of reasons. Not only was the quantity of the cracks but type considered crucial for site identification and ultimately selection. It was also noted that the quality of the survey was in large part a function of the surveyor and their attention to detail and continuity. In totality 1,210 fifty feet by seventeen feet grids were evaluated for the site selection, or 121 field data surveys. Often the surveys were clean enough to be quantified with minimal effort, or there were no recorded distresses at all. In these cases, to populate amount of cracking per job was quick and was as simple as counting boxes with the mouse cursor.

This was not true for all surveys. For some sites, there were simply too many cracks to track accurately. For others, the presence of alligator cracking muddled the ability to quantify transverse and longitudinal cracking. Another approach needed to be taken. Therefore, on sites with large amounts of cracking, the surveys were annotated, first to categorize the cracks, and later counted to quantify them. **Figure 2** demonstrates this process. The purple lines represented areas of alligator cracking, the yellow lines represented longitudinal cracking, and the blue lines represented transverse cracking. A very small portion of the field data surveys required this level of detail for their analysis, roughly 12 percent.

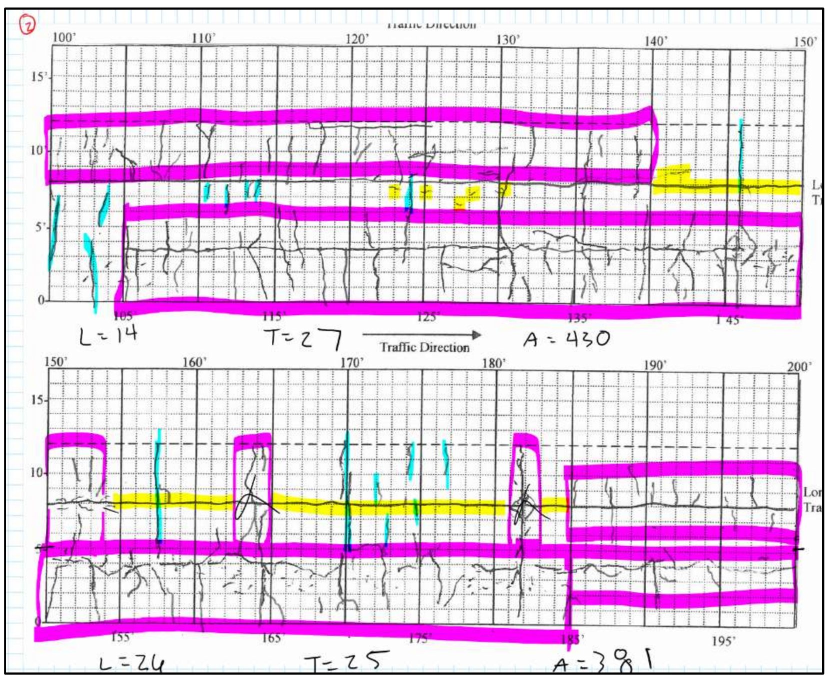


Figure 2. Typical Cracking Detail

An interesting observation from this exercise was a reflection on human behavior and perception. On sites with little to no cracking, surveyors were less inclined to record pavement distresses. Therefore, the longer a pavement stayed 'distress-free'-- especially in the earlier years of the pavement surveys – typically no additional distresses were recorded. Conversely, if a site showed cracking from the initial or early surveys, the increase in cracking as years progressed was generally more severe. Short of taking an ARAN truck through these sites, there is no way to eliminate this human bias but was an interesting take away from this analysis.

b. *Structural Cross-Sections*

Cracking was not the only parameter to be recorded in the database. As mentioned previously, the pavement structure itself can aid or exacerbate pavement distresses. The construction plan sets that accompanied the original dataset from the state became very important as they were the only document available that detailed the pavement structure down to the subgrade for each job. **Figure 3** is exemplary of a pavement cross-section and the extrapolated structural data pulled for each job in the dataset.

Pavement Structure	4" ACHM Surface Course
	3 " ACHM Binder Course
	4.5" ACHM Base Course
	6" Aggregate Base Course
	6" Lime Treated Subgrade
Total:	**23.5 Inches**

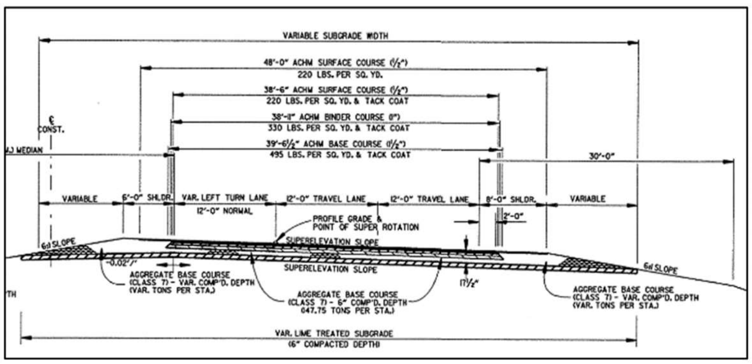


Figure 2. Typical Structural Cross-Section

c. *Database Creation*

The cracking data and the structural information were the critical components of the first database iteration. A few additional parameters would eventually be added as the project evolved and the sites evaluated, however this was all the information collected for all sites (as more parameters were added, in the interest of time not all information was collected for every site; rather, for a running shortlist). **Table 1** shows this database iteration and is representative of data that was collected for all thirty sites.

Table 1. Typical Database Entry by Job

ArDOT District	Date Constructed (year)	Full Depth Thickness (inches)	Total Pvmt Length (feet)	Pvmt Sample Length (feet)	Pvmt Sample Width (feet)	Pvmt Sample Area (sq. feet)	Date of Survey(s)
2	2004	26	12457	500	17	8500	12/7/2012
2	2004	26	12457	500	17	8500	4/7/2014
2	2004	26	12457	500	17	8500	1/12/2015
2	2004	26	12457	500	17	8500	3/2/2017
		Total Longitudinal Crack Length (feet)	**Total Tranverse Crack Length (feet)**	**Fatigue/Alligator Cracking Area (sq. feet)**	**Percent Longitudinal Cracked**	**Percent Transverse Cracked**	**Percent Fatigue/Alligator Cracked**
		32	0	0	6.40%	0.00%	0.00%
		747	5	0	149.40%	1.00%	0.00%
		297	11	334	59.40%	2.20%	3.93%
		308	13	601	61.60%	2.60%	7.07%
	Pavement Structure	2" ACHM Surface Course					
		3 " ACHM Binder Course					
		5" ACHM Base Course					
		10" Aggregate Base Course					
		6" Lime Treated Subgrade					
	Total:	**26 Inches**					

d. *Geospatial Representation*

After this initial database was created, it became obvious that the data needed to be represented in a different manner. Although the database located each job site by county and log mile, it was hard to assess how the sites were geographically distributed. It was ultimately proposed that these sites be mapped using geospatial software readily available to most people, Google Earth. Based on the site maps from the construction sets, each site was approximately

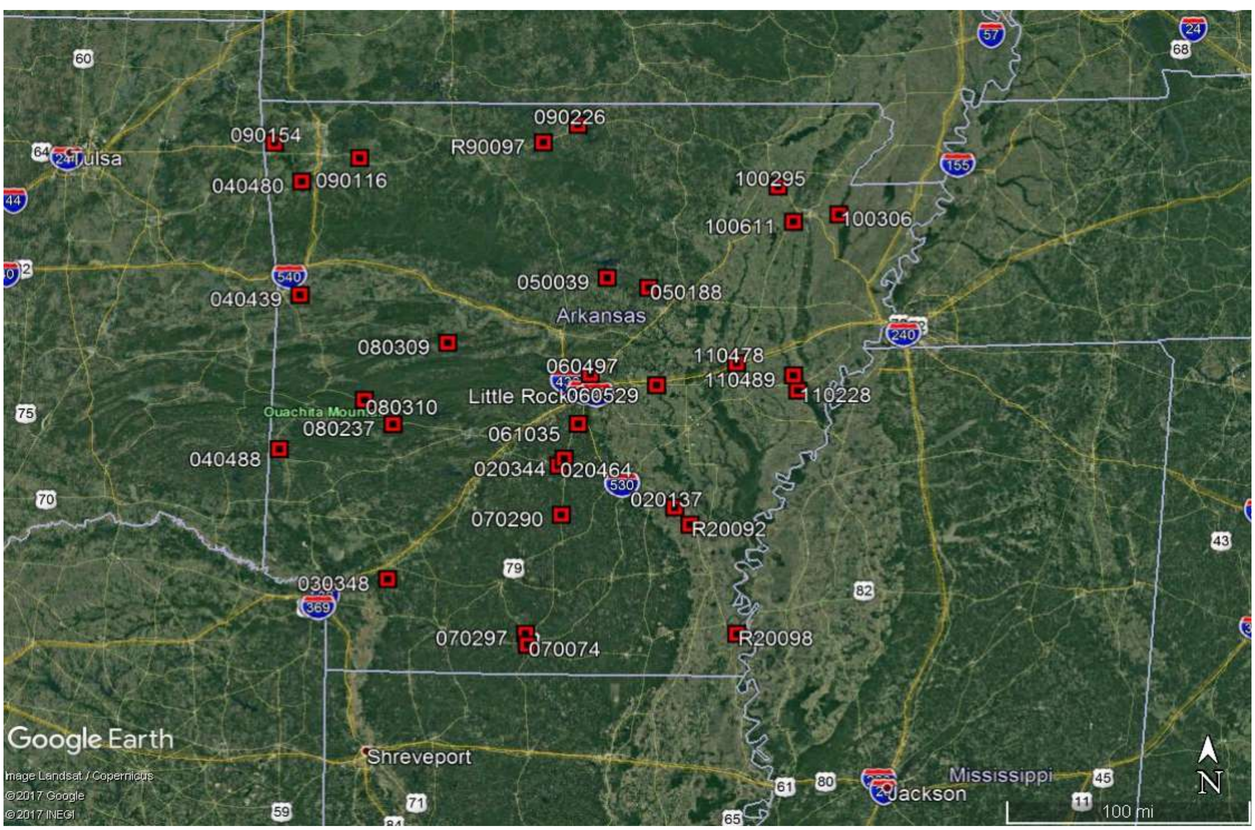


Figure 4. Geographic Distribution of original job sites

mapped to its real-world location. The word approximately is used as the pavement area is delineated by log mile, and the actual locations of these log miles was not made clear to those not working with each district in the state. So, the site location was approximated by consistently using the beginning of the job site as noted in the site map from the construction plans. **Figure 4** shows the resulting geographic distribution of all thirty sites throughout the state.

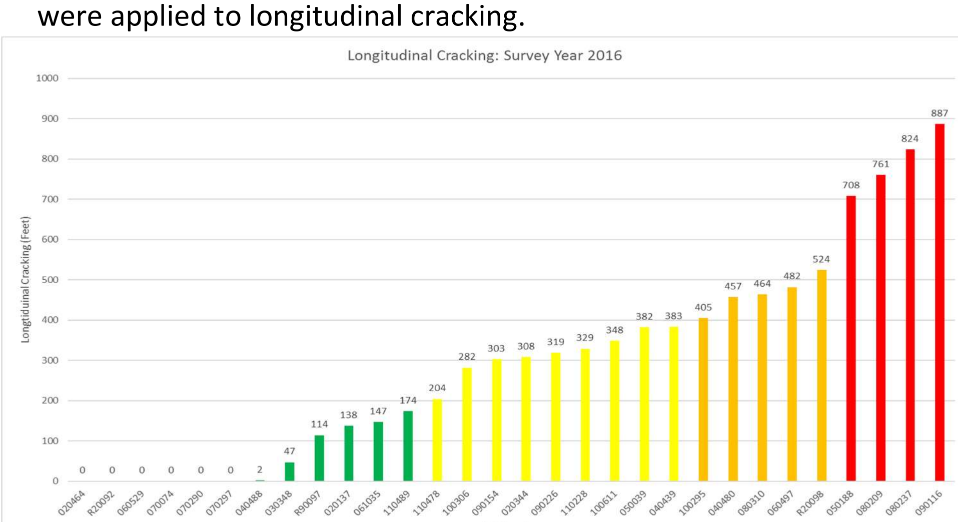
Cracking data quantification, database creation, and the creation of the geographic representation in Google Earth collectively took almost one month. Most of the analysis up to this point had not necessarily been on any one site, but rather generally in terms of the data set as a whole. Once this initial data collection and organization was complete, the database could be refined, shortlists of candidate sites could be made, and additional parameters could be discussed and considered.

3) Data Analysis and Refinement

a. *Quantification and Classification of Cracking Data*

The first level of refinement that was made was to categorize the amount of cracking by type, to enable the categorization of the sites from best to worst, as well as which general time period the cracking data was to represent.

The quantity of cracking greatly varied by the cracking type being considered. This somewhat goes back to the mechanisms that caused them, but ultimately became relevant due to the wide range in which these values existed. It did not appear reasonable to apply the same thresholds to transverse cracking as those in longitudinal cracking if the worst case in longitudinal cracking was 900 feet while the worst case in transverse cracking was 40 feet. The preliminary categories for classifying was a color gradient: green meant little to no cracking, yellow more so than green, orange more so than yellow, and red more so than orange. Different thresholds were applied to each type of cracking based on the natural breaks in the data. The natural breaks discussed simply refers to noting, by observation, bounds in which crack severity increased in a significant way such that ranges could be identified. **Figure 5** shows how these specific thresholds were applied to longitudinal cracking.



Thresholds (Linear Feet):
0-200
201-400
401-600
601+

Figure 5. Graphical Representation of categorical thresholds by Longitudinal Cracking

While categorizing each type of cracking into these four subsets, it became apparent that the time period over which the quantity was being evaluated could greatly affect which sites were considered good and poor. Two approaches were identified: (1) an average could be taken across all existing surveys and aggregated into a single number; or (2) only the most recent survey (which was 2017 for all sites) would be considered. Each approach had advantages and disadvantages.

If an average was taken, it could eliminate surveyor bias by putting each individual field data survey on an 'even playing field' as no singular survey would be more prevalent then another. However, by taking an average the progression of crack propagation would be lost. In addition, a crack that was initially considered longitudinal could propagate and connect with surrounding cracks[2]. By doing so, in the next field data survey, that same crack that was longitudinal is now considered alligator cracking. An average would not consider this natural progression of crack formation because it would essentially be counting that same count twice in two different categories of cracking. Additionally, taking the average greatly skewed values. If a singular pavement showed little to no distresses and in the most recent survey was recorded as greatly distressed with severe cracking, the magnitude of how severe the cracking developed was lost as it became a lesser value in the average.

If the latest survey was used to represent the data, crack progression would in turn be considered as a singular crack formation and would only be considered in one category. However, this methodology ignores prior field data surveys taken as they have no prevalence as only the latest is relevant. Ultimately, the integrity of crack formation and the underlying explanation that pointed to the mechanisms that caused them was too important to lose within an average. In addition to crack integrity, the most recent survey method was picked over the average method as the amount of field data surveys varied from project to project and there was not a clear way to fairly average a site with five surveys versus a site with two surveys. For this reason, the sites were categorized based on the most recent field data survey. A tacit assumption is included here: field surveys are designed to be 'cumulative' – that is, each subsequent survey should verify the results of the previous survey, and add to the data as new cracks (or expanding cracks) are identified. While this is generally the case in the field surveys supplied for this project, it could not be independently verified.

b. Identification of Cracking Mechanisms

Now that the time-related methodology was selected, the type of cracking to be considered needed to be chosen. Three types of cracking were quantified for every field data survey: Longitudinal cracking, transverse cracking, and alligator (fatigue) cracking. These three cracking distresses each have different mechanisms that cause them, so this distinction was critical to make. The overall objective of this study focuses on material-based characterization related to the susceptibility to pavement distresses and how to abate issues related to such. This clearly differentiates from structurally-based pavement distress mechanisms, as the issue is not with the pavement itself but rather with the underlying base and subgrade. For example, a pavement rehabilitation or repair designed to solve material-based pavement distresses like raveling would do no good against a structural-based deficiency such as a pothole[3]. So, the mechanisms that cause these three types of cracks were critical to understand.

Longitudinal cracking is often first found in the wheel path of a pavement as a symptom of fatigue cracking, or in seams in the pavement along joints[4]. Longitudinal cracks always run parallel to the laydown direction or centerline of a pavement. Typically, they are the first types of distresses to appear in accordance to fatigue cracking (although they are not always indicative of structural issues). Longitudinal cracking is typically referred to as a top-down crack, meaning the crack begins at the surface of the pavement and propagates downward. Traditionally, fatigue cracking is typically considered to initiate at the bottom of the asphalt layer, as the bottom of the asphalt layer experiences the most tensile strain. Therefore, traffic (dynamic loading) would cause this bottom layer to shear and crack starting at the bottom and moving upward. **Figure 6** illustrates how these forces work in tandem to create this type of cracking. However, top-down cracking would be the opposite of this; instead of the crack propagating from the bottom of the layer, it begins at the top and works downward through the layer. This is due to the materials in the surface interacting with both shear and tensile forces caused by tires. The reaction of the pavement surface is worsened when binder becomes brittle during the aging process, or by low stiffness in the upper layer of the pavement[5].

Transverse cracking, by contrast, it almost always related to material properties of the pavement and not typically considered indicative of any structural deficiency (barring reflection

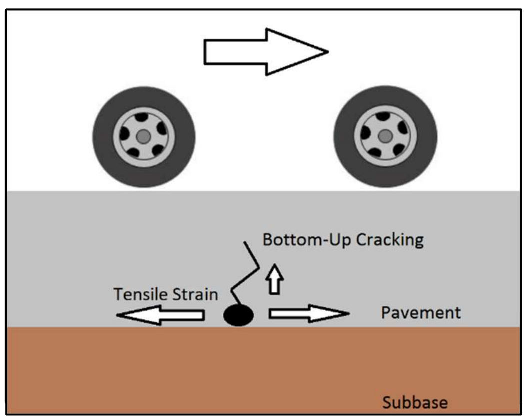


Figure 6. Bottom-Up Cracking

cracking)[6]. Most frequently, Transverse cracking is a result of shrinkage in the asphalt and typically runs perpendicular to the centerline. This shrinkage happens mostly within the binder, as stated previously, the binder is recommended almost exclusively based on predicted pavement temperatures and latitude of the project site.

Fatigue cracking, also known as alligator cracking, is when these longitudinal and transverse cracks connect and create a gridded appearance on the pavement surface. This gridded appearance is like the appearance of a back of an alligator, explaining why it is also called alligator cracking. Often this surface will deteriorate below the elevation of the surrounding pavement, doing so until it becomes a pothole. If fatigue cracking appears early in a pavement's life, it's typically indicative of a structural deficiency in the pavement structure. However, often over the course of a pavement's life wear-and-tear will cause fatigue cracking that eventually can turn into alligator cracking.

Understanding these mechanisms in light of the goals of TRC 1802 led to the decision to focus on choosing sites based on longitudinal cracking. Although the other types of cracking are not ignored, the focus was on longitudinal cracking. The intent of subsequent laboratory testing is to correlate sites with early-age fatigue cracking as related to pavement material deficiencies, not structural deficiencies. Sites with high alligator cracking could be more indicative of structural issues, which was not in the scope of the project to be considered. Transverse cracking quantities generally trended with the amount of longitudinal cracking, so recommendations between the two usually coincided. Therefore, site categorization and selection was based on longitudinal cracking the most recently available field data survey.

c. Additional Data

The research team sought additional data to create a more robust profile of each job site. A few important parameters typically considered in roadway and traffic engineering had not yet been considered; namely, what type of construction were these job sites? How similar was the pavement structure of these roadways? What was the traffic on these roadways? Some of this data was relatively easy to obtain, while others were not as readily available. The type of construction was available in the construction plans previously referenced to find the pavement structure. A large majority of the original thirty projects were either applying overlays, adding a travel lane, or both. Due to the difficulty of being able to specifically identify the sample area of the field data surveys, this factor was noted and projects of this type (either overlay or overlay and lane addition) were almost exclusively selected in order to keep the comparisons as similar as possible.

A very common and convenient method for comparing pavement structure is the Structural Number (SN). Structural Number is calculated in accordance with the 1993 AASHTO Pavement Design Guide[7], using the following equation:

$$SN = a_1D_1 + a_2D_2M_2 + a_3D_3M_3$$

Where a_n is the structural layer coefficient, D_n is the layer thickness, and M_n is the drainage coefficient. Structural layer coefficients (a_n) were taken from the ARDOT Roadway Plan Development Guide [8]. In Arkansas, a value of 1.0 is assumed for the drainage coefficient (M_n). [8}

The last parameter added was to consider the amount of traffic on these roadways. The most readily available source of such information was through traffic maps providing AADT, or average annual daily traffic. Not only was AADT collected, but percent truck traffic was also collected. This was a crucial factor as roadway degradation is expedited in the presence of increased truck traffic. Traffic loading done by various types of vehicles is compared by ESALs, or Equivalent Single Axle Loads. One ESAL is equivalent to 18,000 pounds on a single axle while a standard passenger car is about 2,000 pounds[9]. To put the damage to pavement caused by a semi-truck in perspective, a typical ESAL value for a semi-truck is 2.5; for a passenger car, the ESAL value is generally accepted as 0.0004. Thus, compared to passenger cars, a semi-truck can cause up to 10,000 times more damage to a pavement. So, a pavement with increased cracking could actually be performing better than a pavement with less cracking if that pavement has a higher percentage of truck traffic than the other pavement does. Having AADT and percent truck allows another perspective in the data analysis by being able to further explain and justify the type and quantity of pavement distress when combined with the field data survey information. **Figure 7** is an example of an AADT map from ARDOT from which the traffic data was sourced. In some cases the exact pavement length did not explicitly have an available truck traffic percentage associated with it. In this case, the road closest in proximity was identified and that truck traffic percentage was extrapolated to the pavement section of interest.

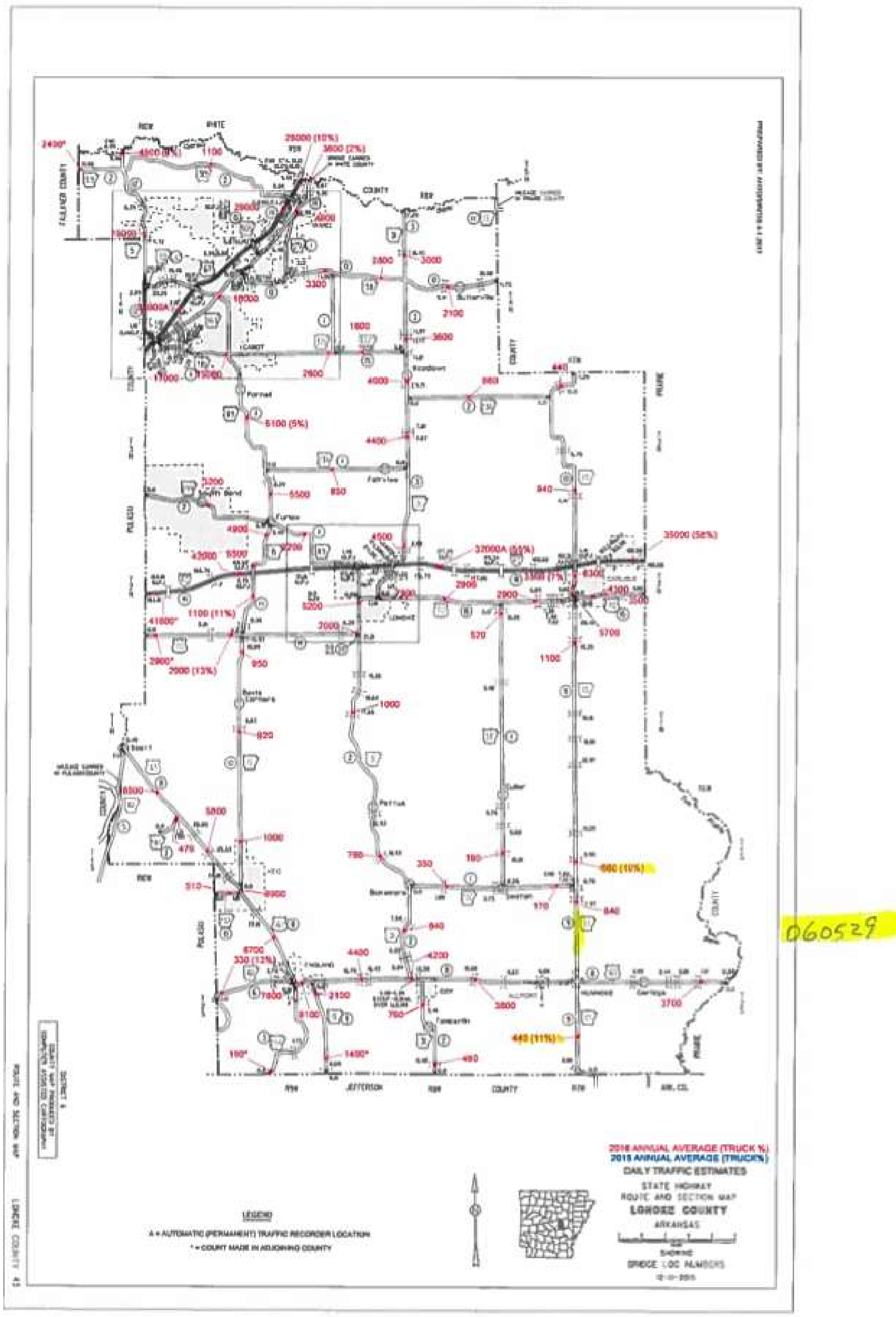


Figure 7. Annual Average Daily Traffic Map

4) Results and Conclusion

After the data collection, refinement, and analyses, it was obvious the first step to site selection would be creating a “shortlist” of the sites, based solely on amount of longitudinal cracking in the latest field data survey using the graph in **Figure 5**. The initial shortlist was then further refined with the consideration of pavement structure and traffic data, and mapped. **Figure 8** is the geographic distribution of this shortlist, color coded by categorized longitudinal cracking.

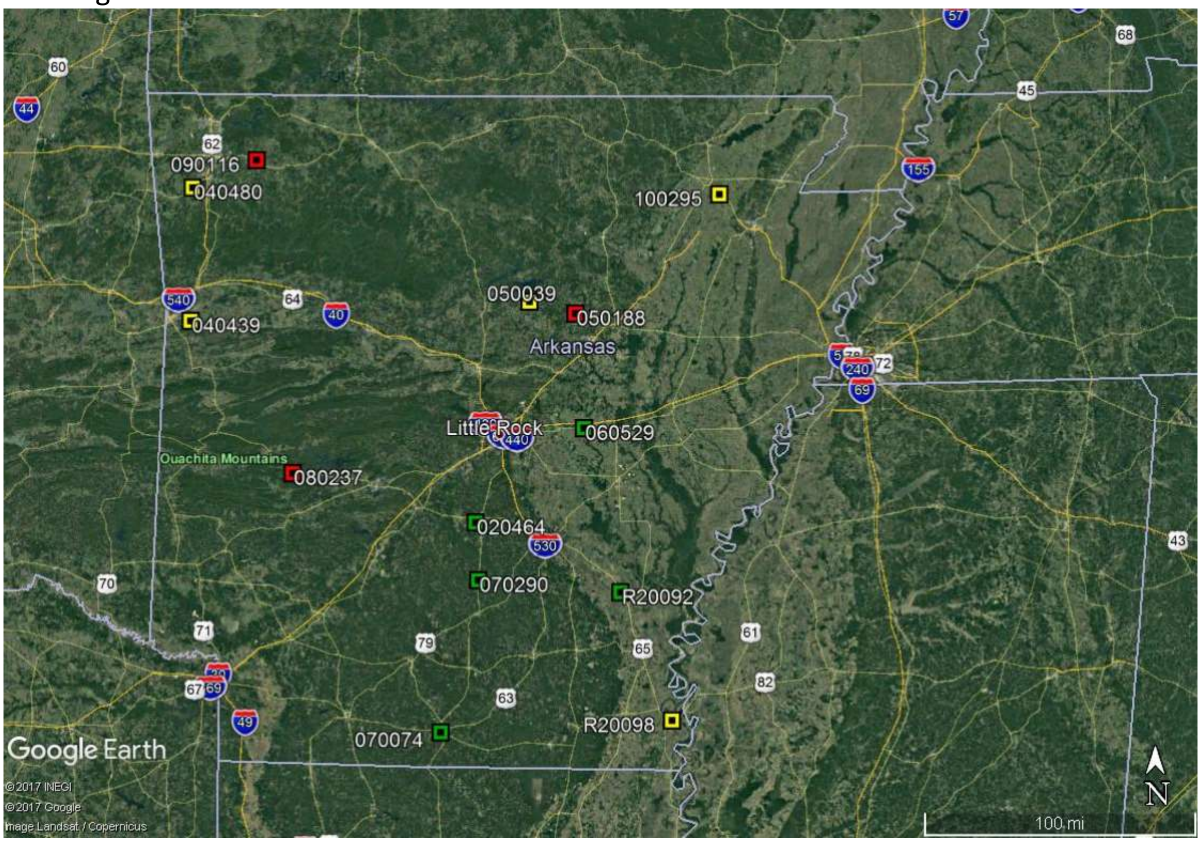


Figure 8. First shortlist job site geographic distribution

Observing the geographic distribution of these sites, although they were dispersed relatively throughout the state, many of the sites were close. Sites close to each other were not advisable due to the desire to isolate material behavior as it relates to pavement performance. Sites in geographic proximity to each other could feature very similar materials, and therefore essentially be the same mix with very similar material properties. Therefore, differences in early-age cracking behavior would not reflect the effect of materials, but rather a function of the traffic, the quality of the construction, and the pavement structure. So, these sites were evaluated and refined again to ensure an even geographic distribution as well as comparable pavement structure and traffic. **Figure 9** shows the final shortlist, bringing the proposed thirty sites down to six for testing and analysis.

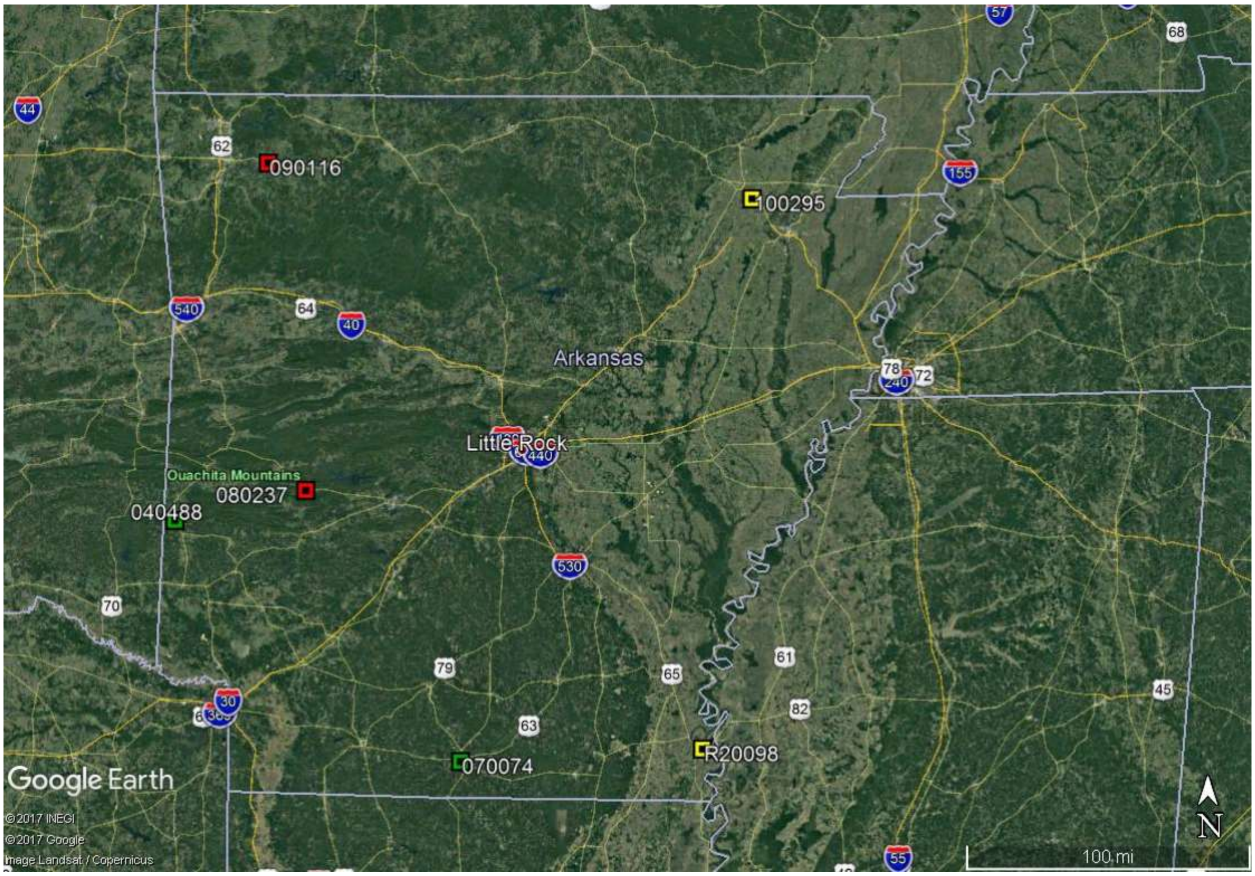


Figure 9. Geographic distribution for final six sites selected for analysis

Six sites were ultimately selected in part due to the ability for the lab team to recreate and test these mixes in a timely manner. The final six sites were distributed relatively evenly throughout the state, ensuring a diversity in quarried materials, mixes, and binders. The colored categories were simplified into Good, Fair, and Poor, with two of each classification being available in this subset. These six sites will have their mixes recreated and tested in terms of cracking to try to predict early-age cracking; this comparison will be done by comparing the lab data to observed field performance, as recorded on the field data surveys.

The site selection procedure outlined in this text is not exclusive to TRC 1802. Rather, the procedure will be a reference and a guideline in future projects in which situational site selection is called for. The procedure establishes a base line by which general metrics and parameters can be addressed and considered, while still offering flexibility to be modified to suit the needs of the project. The site selection to get to these six sites was iterative, reaching past the original dataset by consulting outside sources to get the most appropriate and robust dataset to make an informed decision. By doing so, site selection can become a more concrete and qualitative approach methodology based on the data itself, rather than personal interpretation.

5) References

[1]"Superpave Performance Grading." *Superpave Performance Grading | Pavement Interactive*, University of Washington, 2012, www.pavementinteractive.org/superpave-performance-grading/.

[2]"Fatigue Cracking." *Fatigue Cracking | Pavement Interactive*, University of Washington, 2012, www.pavementinteractive.org/fatigue-cracking/.

[3]"Stripping." *Stripping | Pavement Interactive*, University of Washington, 2012, www.pavementinteractive.org/stripping/.

[4]"Longitudinal Cracking." *Longitudinal Cracking | Pavement Interactive*, University of Washington, 2012, www.pavementinteractive.org/longitudinal-cracking/.

[5]"Top-Down Cracking." *Top-Down Cracking | Pavement Interactive*, University of Washington, 2012, www.pavementinteractive.org/top-down-cracking/.

[6]"Transverse Cracking." *Transverse Cracking | Pavement Interactive*, University of Washington, 2012, www.pavementinteractive.org/transverse-cracking/.

[7]"Appendix C: 1993 AASHTO Design Method" *Geotechnical Aspects of Pavements Reference Manual | Bridges and Structures,* Federal Highway Administration, 2017, www.fhwa.dot.gov/engineering/geotech/pubs/05037/ac.cfm/.

[8]"Appendix A: Pavement Design Criteria" *Roadway Design Plan Development Guidelines | Roadway Design Division,* Arkansas Department of Transportation, 2017, www.arkansashighways.com/roadway_design_division/Roadway%20Design%20Plan%20Development%20Guidelines.pdf.

[9]"Third Quarter Flexible Pavement Design Frequently Asked Questions" *Construction and Material Tips | Pavement Design Manual,* Texas Department of Transportation, 2005, https://ftp.dot.state.tx.us/pub/txdot-info/cst/tips/flex_pave_des_faq.pdf.