# VTrans Research

## Technical Memo – Hamburg Wheel Tracker Preliminary Data Analysis

To: Andy Willette, Troy Lawson, Aaron Schwartz, Nick Van Den Berg, Vermont Agency of Transportation

From: Dave Novak, UVM

Date: October 8, 2018 (amended January 28, 2019)

## Project Scope

This memo documents the results of initial data analysis efforts by the UVM research team for pavement testing using the Hamburg Wheel Tracker (HWT) test by the Vermont Agency of Transportation's Materials Testing & Certification Lab.

The ultimate objective of the Hamburg Wheel Tracker project is to develop a method for calculating pay factors for use by the Agency in paying contractors for hot mix asphalt (HMA) mixture durability characteristics in accordance with AASHTO T324. The Agency's intent is to use pay factors to create incentives for HMA contractors to furnish, mix, haul, and place high performing HMA mixtures. The initial component of the project involves cleaning, sorting, and analyzing the data the Agency currently has related to the testing of different HMA mixtures using the HWT test.

## Data Analysis Process and Findings

I initially met with Andy, Troy, and Aaron on Thursday, July 18th to discuss the project. I received the data files early the following week. I had several email exchanges with Troy and Aaron in the following months regarding questions I had related to the data.

I examined a total of 176 HWT files as part of the project. Each file is a separate Excel workbook and it was necessary to open and inspect each file individually. Each file consists of two samples. Ten of the 176 HWT files used testing criteria of 30,000 passes while the remaining 166 used a criterion of 20,000 passes. Based on my understanding, the 30,000 pass tests were largely test cases and 20,000 passes is the official number of HWT test passes. Any sample that successfully completes all 20,000 passes with a rut depth of less than 12.5 mm "passes" the HWT test. If a sample hits or exceeds a rut depth of 12.5 mm at any point during the test, it "fails" the HWT test. I will deliver the "organized" files to VTrans on a USB key.

The overall goal of this analysis was to become familiar with the HWT data and consider how the data can be used to develop a pay factor schedule and then determine statistical compliance using the well-established Percent-Within-Limits (PWL) method. One of the sub-goals of the analysis was to separate the data into three categories: 1) pass, 2) fail, and 3) invalid. The categories allow the analyst to group each file according to whether it passed or failed the HWT. The initial criterion for invalid was a difference greater than or equal to 6 mm rut depth between the two samples.

It is important to point out that there are a number of inconsistencies in the data files. These inconsistencies are quite prevalent are definitely going to create some challenges with respect to the design and implementation of a pay factor schedule. The inconsistencies affect the categorization of some of the files and it is possible that there are even more invalid tests than the ones labeled as “Invalid”. I created the additional categories of Pass/Invalid and Fail/Invalid to store some of the files that appear (in my opinion) to either clearly pass or fail, but have notable data inconsistencies. The Agency can also use these categorizations to see examples of some of the problems that were encountered during this project. *All of the files that are in some way classified as “Invalid” (including Pass/Invalid and Fail/Invalid need to be revisited by the Agency)*. Questions/concerns related to data inconsistencies are documented in this memo.

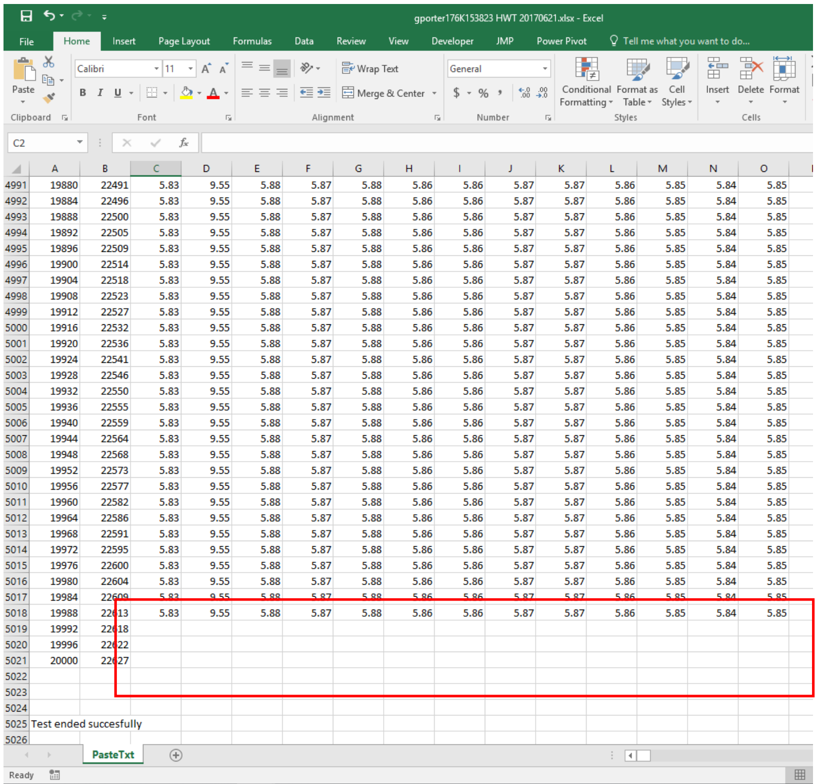
Of the 176 files examined, there were 90 failures (23 of those are Fail/Invalid) and 68 passes (4 of those are Pass/Invalid). The remaining 18 files are classified as “Invalid”. The files are organized by category in the Summary.xlsx spreadsheet (attached).

## Initial Data Manipulation and Data Inconsistencies

To develop a pay factor schedule, it is necessary to be able to filter the data according to Mix Type and PG Grade and calculate summary statistics (mean, median, and standard deviation) and conduct comparative analyses for the performance measure being tested (i.e. the rut depth). To do this, it is important to “clean” the data and to remove any biased or invalid data from the data set.

It is important to note is that the data are <u>not</u> organized in a way that is conducive to calculating summary statistics or conducting comparative analyses. I initially began separating files based on the test summary information provided at the bottom of each Excel workbook file. This was based on the text “test ended successfully” and “the maximum rut depth has been exceeded” for tests that pass and fail the HWT test respectively. Unfortunately, the test summary information is inconsistent and not always correct. Therefore, the tests cannot be sorted using the summary information. There were also mistakes in the alignment between rows in columns in somewhere between one-third to one-half of *all* HWT data files. This meant that I had to redo my initial classification and go back through every file row-by-row and column-by-column. Examples of conflicting output include gporter175G075800, gporter176K153823, jglover176D063539, and jglover175U062239 (among others). Looking at the detailed data in these files, one can see that the test clearly exceeds maximum rut depth in many locations; however, the HWT test completes all 20,000 passes and reports “test ended successfully”. This raises concerns regarding testing / reporting consistency and accuracy. Short of manually going through each file line-by-line, how can the output be trusted? If you cannot trust the summary output related to whether or not the HWT was passed or not, you cannot develop a programmatic QA/QC test using those data. Because of the time required to do a manual inspection, I performed only a quick visual scan of all files. Thus, it is important to state up front that while I identified many issues and data inconsistencies, my visual inspection of the files was not thorough or exhaustive.

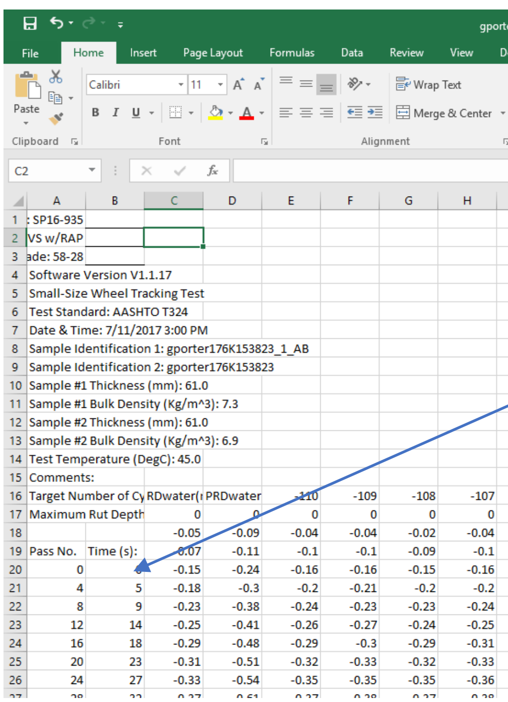
It is also important to note that I had to manually fix all of the files that had mistakes in alignment between the rows and columns. Examples are shown in Figures 1 and 2.




| | A | B | C | D | E | F | G | H | I | J | K | L | M | N | O |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 4991 | 19880 | 22491 | 5.83 | 9.55 | 5.88 | 5.87 | 5.88 | 5.86 | 5.86 | 5.87 | 5.87 | 5.86 | 5.85 | 5.84 | 5.85 |
| 4992 | 19884 | 22496 | 5.83 | 9.55 | 5.88 | 5.87 | 5.88 | 5.86 | 5.86 | 5.87 | 5.87 | 5.86 | 5.85 | 5.84 | 5.85 |
| 4993 | 19888 | 22500 | 5.83 | 9.55 | 5.88 | 5.87 | 5.88 | 5.86 | 5.86 | 5.87 | 5.87 | 5.86 | 5.85 | 5.84 | 5.85 |
| 4994 | 19892 | 22505 | 5.83 | 9.55 | 5.88 | 5.87 | 5.88 | 5.86 | 5.86 | 5.87 | 5.87 | 5.86 | 5.85 | 5.84 | 5.85 |
| 4995 | 19896 | 22509 | 5.83 | 9.55 | 5.88 | 5.87 | 5.88 | 5.86 | 5.86 | 5.87 | 5.87 | 5.86 | 5.85 | 5.84 | 5.85 |
| 4996 | 19900 | 22514 | 5.83 | 9.55 | 5.88 | 5.87 | 5.88 | 5.86 | 5.86 | 5.87 | 5.87 | 5.86 | 5.85 | 5.84 | 5.85 |
| 4997 | 19904 | 22518 | 5.83 | 9.55 | 5.88 | 5.87 | 5.88 | 5.86 | 5.86 | 5.87 | 5.87 | 5.86 | 5.85 | 5.84 | 5.85 |
| 4998 | 19908 | 22523 | 5.83 | 9.55 | 5.88 | 5.87 | 5.88 | 5.86 | 5.86 | 5.87 | 5.87 | 5.86 | 5.85 | 5.84 | 5.85 |
| 4999 | 19912 | 22527 | 5.83 | 9.55 | 5.88 | 5.87 | 5.88 | 5.86 | 5.86 | 5.87 | 5.87 | 5.86 | 5.85 | 5.84 | 5.85 |
| 5000 | 19916 | 22532 | 5.83 | 9.55 | 5.88 | 5.87 | 5.88 | 5.86 | 5.86 | 5.87 | 5.87 | 5.86 | 5.85 | 5.84 | 5.85 |
| 5001 | 19920 | 22536 | 5.83 | 9.55 | 5.88 | 5.87 | 5.88 | 5.86 | 5.86 | 5.87 | 5.87 | 5.86 | 5.85 | 5.84 | 5.85 |
| 5002 | 19924 | 22541 | 5.83 | 9.55 | 5.88 | 5.87 | 5.88 | 5.86 | 5.86 | 5.87 | 5.87 | 5.86 | 5.85 | 5.84 | 5.85 |
| 5003 | 19928 | 22546 | 5.83 | 9.55 | 5.88 | 5.87 | 5.88 | 5.86 | 5.86 | 5.87 | 5.87 | 5.86 | 5.85 | 5.84 | 5.85 |
| 5004 | 19932 | 22550 | 5.83 | 9.55 | 5.88 | 5.87 | 5.88 | 5.86 | 5.86 | 5.87 | 5.87 | 5.86 | 5.85 | 5.84 | 5.85 |
| 5005 | 19936 | 22555 | 5.83 | 9.55 | 5.88 | 5.87 | 5.88 | 5.86 | 5.86 | 5.87 | 5.87 | 5.86 | 5.85 | 5.84 | 5.85 |
| 5006 | 19940 | 22559 | 5.83 | 9.55 | 5.88 | 5.87 | 5.88 | 5.86 | 5.86 | 5.87 | 5.87 | 5.86 | 5.85 | 5.84 | 5.85 |
| 5007 | 19944 | 22564 | 5.83 | 9.55 | 5.88 | 5.87 | 5.88 | 5.86 | 5.86 | 5.87 | 5.87 | 5.86 | 5.85 | 5.84 | 5.85 |
| 5008 | 19948 | 22568 | 5.83 | 9.55 | 5.88 | 5.87 | 5.88 | 5.86 | 5.86 | 5.87 | 5.87 | 5.86 | 5.85 | 5.84 | 5.85 |
| 5009 | 19952 | 22573 | 5.83 | 9.55 | 5.88 | 5.87 | 5.88 | 5.86 | 5.86 | 5.87 | 5.87 | 5.86 | 5.85 | 5.84 | 5.85 |
| 5010 | 19956 | 22577 | 5.83 | 9.55 | 5.88 | 5.87 | 5.88 | 5.86 | 5.86 | 5.87 | 5.87 | 5.86 | 5.85 | 5.84 | 5.85 |
| 5011 | 19960 | 22582 | 5.83 | 9.55 | 5.88 | 5.87 | 5.88 | 5.86 | 5.86 | 5.87 | 5.87 | 5.86 | 5.85 | 5.84 | 5.85 |
| 5012 | 19964 | 22586 | 5.83 | 9.55 | 5.88 | 5.87 | 5.88 | 5.86 | 5.86 | 5.87 | 5.87 | 5.86 | 5.85 | 5.84 | 5.85 |
| 5013 | 19968 | 22591 | 5.83 | 9.55 | 5.88 | 5.87 | 5.88 | 5.86 | 5.86 | 5.87 | 5.87 | 5.86 | 5.85 | 5.84 | 5.85 |
| 5014 | 19972 | 22595 | 5.83 | 9.55 | 5.88 | 5.87 | 5.88 | 5.86 | 5.86 | 5.87 | 5.87 | 5.86 | 5.85 | 5.84 | 5.85 |
| 5015 | 19976 | 22600 | 5.83 | 9.55 | 5.88 | 5.87 | 5.88 | 5.86 | 5.86 | 5.87 | 5.87 | 5.86 | 5.85 | 5.84 | 5.85 |
| 5016 | 19980 | 22604 | 5.83 | 9.55 | 5.88 | 5.87 | 5.88 | 5.86 | 5.86 | 5.87 | 5.87 | 5.86 | 5.85 | 5.84 | 5.85 |
| 5017 | 19984 | 22609 | 5.83 | 9.55 | 5.88 | 5.87 | 5.88 | 5.86 | 5.86 | 5.87 | 5.87 | 5.86 | 5.85 | 5.84 | 5.85 |
| 5018 | 19988 | 22613 | 5.83 | 9.55 | 5.88 | 5.87 | 5.88 | 5.86 | 5.86 | 5.87 | 5.87 | 5.86 | 5.85 | 5.84 | 5.85 |
| 5019 | 19992 | 22618 | | | | | | | | | | | | | |
| 5020 | 19996 | 22622 | | | | | | | | | | | | | |
| 5021 | 20000 | 22627 | | | | | | | | | | | | | |
| 5022 | | | | | | | | | | | | | | | |
| 5023 | | | | | | | | | | | | | | | |
| 5024 | | | | | | | | | | | | | | | |
| 5025 | Test ended succesfully | | | | | | | | | | | | | | |
| 5026 | | | | | | | | | | | | | | | |

I initially thought this meant the test of invalid – summary information says the test ends successfully, but the test appeared to cut off early

Figure 1. Column / Row Offset Mistake in the Raw Data (bottom of file)



| | A | B | C | D | E | F | G | H |
|---|---|---|---|---|---|---|---|---|
| 1 | : SP16-935 | | | | | | | |
| 2 | VS w/RAP | | | | | | | |
| 3 | ade: 58-28 | | | | | | | |
| 4 | Software Version V1.1.17 | | | | | | | |
| 5 | Small-Size Wheel Tracking Test | | | | | | | |
| 6 | Test Standard: AASHTO T324 | | | | | | | |
| 7 | Date & Time: 7/11/2017 3:00 PM | | | | | | | |
| 8 | Sample Identification 1: gporter176K153823_1_AB | | | | | | | |
| 9 | Sample Identification 2: gporter176K153823 | | | | | | | |
| 10 | Sample #1 Thickness (mm): 61.0 | | | | | | | |
| 11 | Sample #1 Bulk Density (Kg/m^3): 7.3 | | | | | | | |
| 12 | Sample #2 Thickness (mm): 61.0 | | | | | | | |
| 13 | Sample #2 Bulk Density (Kg/m^3): 6.9 | | | | | | | |
| 14 | Test Temperature (DegC): 45.0 | | | | | | | |
| 15 | Comments: | | | | | | | |
| 16 | Target Number of Cy | | RDwater(r | PRDwater | -110 | -109 | -108 | -107 |
| 17 | Maximum Rut Depth | | 0 | 0 | 0 | 0 | 0 | 0 |
| 18 | | | -0.05 | -0.09 | -0.04 | -0.04 | -0.02 | -0.04 |
| 19 | Pass No. | Time (s): | -0.07 | -0.11 | -0.1 | -0.1 | -0.09 | -0.1 |
| 20 | 0 | [illegible] | -0.15 | -0.24 | -0.16 | -0.16 | -0.15 | -0.16 |
| 21 | 4 | 5 | -0.18 | -0.3 | -0.2 | -0.21 | -0.2 | -0.2 |
| 22 | 8 | 9 | -0.23 | -0.38 | -0.24 | -0.23 | -0.23 | -0.24 |
| 23 | 12 | 14 | -0.25 | -0.41 | -0.26 | -0.27 | -0.24 | -0.25 |
| 24 | 16 | 18 | -0.29 | -0.48 | -0.29 | -0.3 | -0.29 | -0.31 |
| 25 | 20 | 23 | -0.31 | -0.51 | -0.32 | -0.33 | -0.32 | -0.33 |
| 26 | 24 | 27 | -0.33 | -0.54 | -0.35 | -0.35 | -0.35 | -0.36 |

I discovered that this was due to mistakes in the Excel file. I inserted three rows at row 16, then cut text from old cells A16 and A17 (target # and Max Rut Depth, then paste under Comments. Then, Cut all Pass No. and Time (s) in columns A and B and paste to align with columns C, D, E, etc.

This realigns everything properly

Figure 2. Column / Row Offset Mistake in the Raw Data (top of file)

To conduct comparative analyses, all data should ideally be contained in a single source file. Currently, there are 166 separate files for all of the 20K HWT tests. Each test file consists of a paired sample (two separate samples). The current format includes 454 different fields (the attribute columns). There are 221 separate positions associated with each sample, ranging from a position of -110 mm to 110 mm where a separate data element (each column x row intersection) is recorded for each pass for each sample for single every mm change in position. Data associated with every fourth pass are recorded. An example is provided in Figure 3. Each successful test has 2,270,454 separate data elements (5,001 rows x 454 columns).

Each test is a separate file and there are currently 166 separate files. Thus, we are talking about working with more than 350 million data elements. This is extremely unwieldly and not particularly useful for any type of comparative analysis.

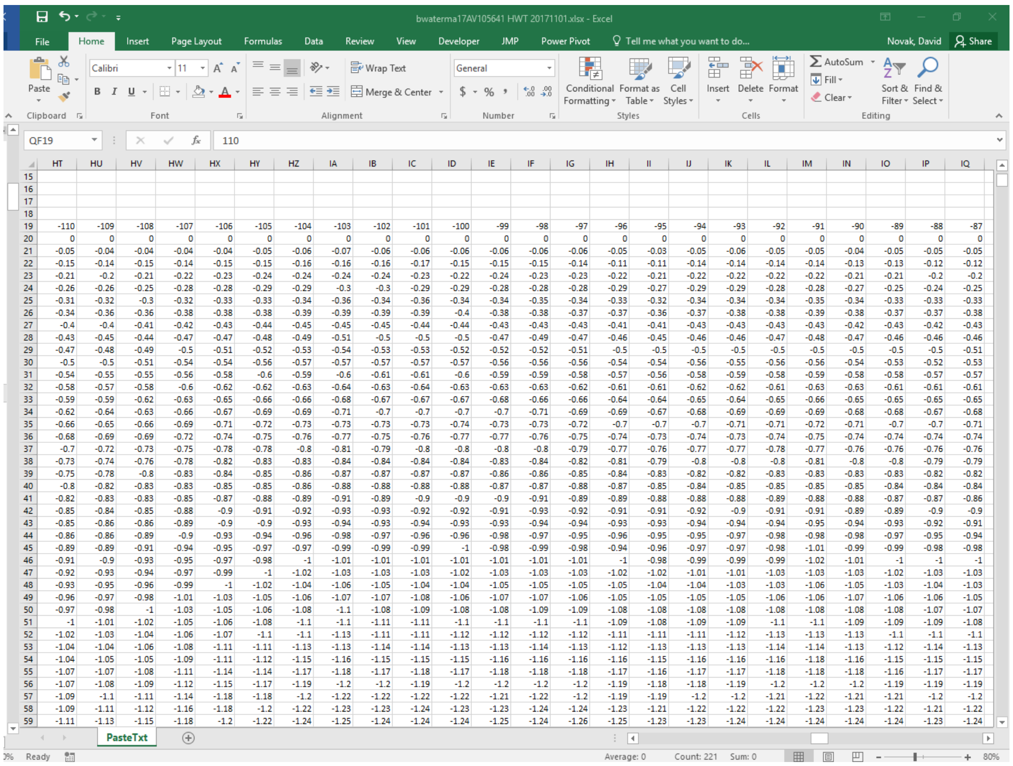


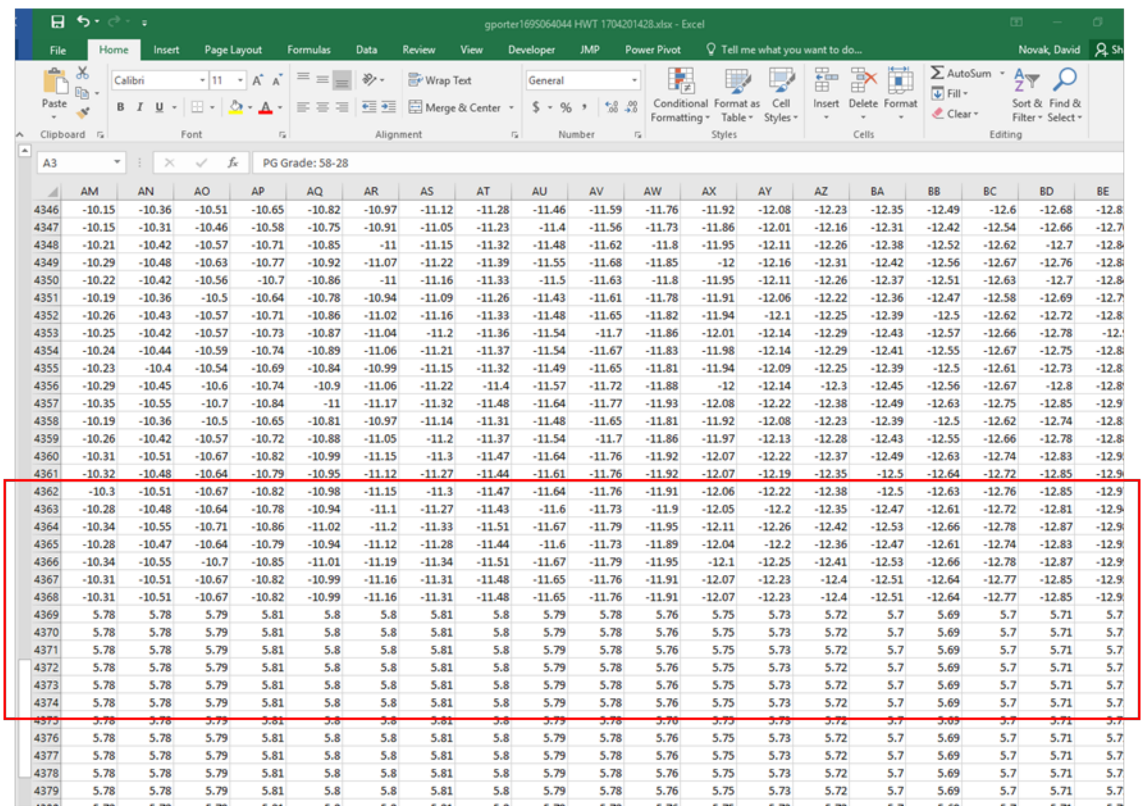
Figure 3. Example of Data Format

While Microsoft Excel is not the software of choice for this type of analysis, I was hesitant to use software specifically designed for handling large data sets such as SAS (https://www.sas.com/en_us/home.html) as using software that the Agency does not have a site license for is counterproductive. It is worthwhile to note that Excel's data storage capabilities would be exceeded by combining all data elements from all files into a single source (this is not even a possibility in Excel). I spent some time experimenting with reducing the amount of data by capturing only every 1,000$^{th}$ pass and limiting the positioning for each test to five separate values (-110, -60, 0, 60, 110). I wrote some Excel macros for reducing the dimensionality of each file; however, it was not clear to me which of the columns provided unique valued added information and were essential. Likewise, it was not clear if collecting data related to every 1,000$^{th}$ pass helped with analysis, although it did greatly reduce dimensionality. If the data are aggregated by every 1,000$^{th}$ pass, important information related to failures and data inconsistencies is lost.

Furthermore, the HWT results do not indicate exactly where (the exact column / row intersection) the maximum rut depth is first exceeded or where anomalies such as a sudden shift from a negative value like -10.45 (where we expect the following records to be less than or equal to -10.45) to some positive value like 5 occur. Currently, the analyst must manually look for these points. It is also important to note that, in many cases, the HWT test continues to run long after the rut depth is exceeded and continues to run despite anomalies. This is hugely problematic if the Agency wants to compare outcomes from different samples or outcomes from different test files. For example, answering a question such as “on average at what pass and what position(s) do failed samples of Mix Type IVS w/RAPT exceed the maximum rut depth” are not possible to answer without spending a lot more time cleaning the data, determining exactly what results need to be collected, determining where to put these results (spreadsheet versus database) and then coding a collection mechanism. These are a few things that the Agency should discuss moving forward.

During the testing process, it appears that it is very common for asphalt material to fall back into the wheel tracker rut. This results in a notable increase in the rut depth during the next pass. Depending on the amount of material, this may affect the rut depth for *many or all subsequent passes*. Therefore, it is not clear if all tests where this occurs should be treated as invalid, or whether a reasonable failure determination can be made even though this occurs. For this project, I did the best I could to determine whether or not this type of anomaly impacted the test results and attempted to document this in the Comments field in the Summary.xlsx file. For example, if material falls back into the wheel tracker rut and the maximum rut depth is still exceeded (the sample still fails the HWT test), we can be certain that the sample failed the test. However, we cannot use the data to determine exactly when and where the sample failed because the data are corrupt. This is important, as the ultimate goal is the use the data to develop a pay factor schedule and then to assess statistical compliance. For example, note that the last pass in Figure 1 has a rut depth of nearly 6 for the wheel tracker positions shown in the figure. Clearly, this is not a valid rut depth.

At this point in time, I’m not sure what to do about these situations, as having material fall back into the HWT test biases the results. How would one compute an accurate and meaningful mean and standard deviation for the rut depth if material artificially falls back into the existing rut? An example is shown in Figure 4.



| | AM | AN | AO | AP | AQ | AR | AS | AT | AU | AV | AW | AX | AY | AZ | BA | BB | BC | BD | BE |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 4346 | -10.15 | -10.36 | -10.51 | -10.65 | -10.82 | -10.97 | -11.12 | -11.28 | -11.46 | -11.59 | -11.76 | -11.92 | -12.08 | -12.23 | -12.35 | -12.49 | -12.6 | -12.68 | -12.8 |
| 4347 | -10.15 | -10.31 | -10.46 | -10.58 | -10.75 | -10.91 | -11.05 | -11.23 | -11.4 | -11.56 | -11.73 | -11.86 | -12.01 | -12.16 | -12.31 | -12.42 | -12.54 | -12.66 | -12.7 |
| 4348 | -10.21 | -10.42 | -10.57 | -10.71 | -10.85 | -11 | -11.15 | -11.32 | -11.48 | -11.62 | -11.8 | -11.95 | -12.11 | -12.26 | -12.38 | -12.52 | -12.62 | -12.7 | -12.8 |
| 4349 | -10.29 | -10.48 | -10.63 | -10.77 | -10.92 | -11.07 | -11.22 | -11.39 | -11.55 | -11.68 | -11.85 | -12 | -12.16 | -12.31 | -12.42 | -12.56 | -12.67 | -12.76 | -12.8 |
| 4350 | -10.22 | -10.42 | -10.56 | -10.7 | -10.86 | -11 | -11.16 | -11.33 | -11.5 | -11.63 | -11.8 | -11.95 | -12.11 | -12.26 | -12.37 | -12.51 | -12.63 | -12.7 | -12.8 |
| 4351 | -10.19 | -10.36 | -10.5 | -10.64 | -10.78 | -10.94 | -11.09 | -11.26 | -11.43 | -11.61 | -11.78 | -11.91 | -12.06 | -12.22 | -12.36 | -12.47 | -12.58 | -12.69 | -12.7 |
| 4352 | -10.26 | -10.43 | -10.57 | -10.71 | -10.86 | -11.02 | -11.16 | -11.33 | -11.48 | -11.65 | -11.82 | -11.94 | -12.1 | -12.25 | -12.39 | -12.5 | -12.62 | -12.72 | -12.8 |
| 4353 | -10.25 | -10.42 | -10.57 | -10.73 | -10.87 | -11.04 | -11.2 | -11.36 | -11.54 | -11.7 | -11.86 | -12.01 | -12.14 | -12.29 | -12.43 | -12.57 | -12.66 | -12.78 | -12. |
| 4354 | -10.24 | -10.44 | -10.59 | -10.74 | -10.89 | -11.06 | -11.21 | -11.37 | -11.54 | -11.67 | -11.83 | -11.98 | -12.14 | -12.29 | -12.41 | -12.55 | -12.67 | -12.75 | -12.8 |
| 4355 | -10.23 | -10.4 | -10.54 | -10.69 | -10.84 | -10.99 | -11.15 | -11.32 | -11.49 | -11.65 | -11.81 | -11.94 | -12.09 | -12.25 | -12.39 | -12.5 | -12.61 | -12.73 | -12.8 |
| 4356 | -10.29 | -10.45 | -10.6 | -10.74 | -10.9 | -11.06 | -11.22 | -11.4 | -11.57 | -11.72 | -11.88 | -12 | -12.14 | -12.3 | -12.45 | -12.56 | -12.67 | -12.8 | -12.8 |
| 4357 | -10.35 | -10.55 | -10.7 | -10.84 | -11 | -11.17 | -11.32 | -11.48 | -11.64 | -11.77 | -11.93 | -12.08 | -12.22 | -12.38 | -12.49 | -12.63 | -12.75 | -12.85 | -12.9 |
| 4358 | -10.19 | -10.36 | -10.5 | -10.65 | -10.81 | -10.97 | -11.14 | -11.31 | -11.48 | -11.65 | -11.81 | -11.92 | -12.08 | -12.23 | -12.39 | -12.5 | -12.62 | -12.74 | -12.8 |
| 4359 | -10.26 | -10.42 | -10.57 | -10.72 | -10.88 | -11.05 | -11.2 | -11.37 | -11.54 | -11.7 | -11.86 | -11.97 | -12.13 | -12.28 | -12.43 | -12.55 | -12.66 | -12.78 | -12.8 |
| 4360 | -10.31 | -10.51 | -10.67 | -10.82 | -10.99 | -11.15 | -11.3 | -11.47 | -11.64 | -11.76 | -11.92 | -12.07 | -12.22 | -12.37 | -12.49 | -12.63 | -12.74 | -12.83 | -12.9 |
| 4361 | -10.32 | -10.48 | -10.64 | -10.79 | -10.95 | -11.12 | -11.27 | -11.44 | -11.61 | -11.76 | -11.92 | -12.07 | -12.19 | -12.35 | -12.5 | -12.64 | -12.72 | -12.85 | -12.9 |
| 4362 | -10.3 | -10.51 | -10.67 | -10.82 | -10.98 | -11.15 | -11.3 | -11.47 | -11.64 | -11.76 | -11.91 | -12.06 | -12.22 | -12.38 | -12.5 | -12.63 | -12.76 | -12.85 | -12.9 |
| 4363 | -10.28 | -10.48 | -10.64 | -10.78 | -10.94 | -11.1 | -11.27 | -11.43 | -11.6 | -11.73 | -11.9 | -12.05 | -12.2 | -12.35 | -12.47 | -12.61 | -12.72 | -12.81 | -12.9 |
| 4364 | -10.34 | -10.55 | -10.71 | -10.86 | -11.02 | -11.2 | -11.33 | -11.51 | -11.67 | -11.79 | -11.95 | -12.11 | -12.26 | -12.42 | -12.53 | -12.66 | -12.78 | -12.87 | -12.9 |
| 4365 | -10.28 | -10.47 | -10.64 | -10.79 | -10.94 | -11.12 | -11.28 | -11.44 | -11.6 | -11.73 | -11.89 | -12.04 | -12.2 | -12.36 | -12.47 | -12.61 | -12.74 | -12.83 | -12.9 |
| 4366 | -10.34 | -10.55 | -10.7 | -10.85 | -11.01 | -11.19 | -11.34 | -11.51 | -11.67 | -11.79 | -11.95 | -12.1 | -12.25 | -12.41 | -12.53 | -12.66 | -12.78 | -12.87 | -12.9 |
| 4367 | -10.31 | -10.51 | -10.67 | -10.82 | -10.99 | -11.16 | -11.31 | -11.48 | -11.65 | -11.76 | -11.91 | -12.07 | -12.23 | -12.4 | -12.51 | -12.64 | -12.77 | -12.85 | -12.9 |
| 4368 | -10.31 | -10.51 | -10.67 | -10.82 | -10.99 | -11.16 | -11.31 | -11.48 | -11.65 | -11.76 | -11.91 | -12.07 | -12.23 | -12.4 | -12.51 | -12.64 | -12.77 | -12.85 | -12.9 |
| 4369 | 5.78 | 5.78 | 5.79 | 5.81 | 5.8 | 5.8 | 5.81 | 5.8 | 5.79 | 5.78 | 5.76 | 5.75 | 5.73 | 5.72 | 5.7 | 5.69 | 5.7 | 5.71 | 5.7 |
| 4370 | 5.78 | 5.78 | 5.79 | 5.81 | 5.8 | 5.8 | 5.81 | 5.8 | 5.79 | 5.78 | 5.76 | 5.75 | 5.73 | 5.72 | 5.7 | 5.69 | 5.7 | 5.71 | 5.7 |
| 4371 | 5.78 | 5.78 | 5.79 | 5.81 | 5.8 | 5.8 | 5.81 | 5.8 | 5.79 | 5.78 | 5.76 | 5.75 | 5.73 | 5.72 | 5.7 | 5.69 | 5.7 | 5.71 | 5.7 |
| 4372 | 5.78 | 5.78 | 5.79 | 5.81 | 5.8 | 5.8 | 5.81 | 5.8 | 5.79 | 5.78 | 5.76 | 5.75 | 5.73 | 5.72 | 5.7 | 5.69 | 5.7 | 5.71 | 5.7 |
| 4373 | 5.78 | 5.78 | 5.79 | 5.81 | 5.8 | 5.8 | 5.81 | 5.8 | 5.79 | 5.78 | 5.76 | 5.75 | 5.73 | 5.72 | 5.7 | 5.69 | 5.7 | 5.71 | 5.7 |
| 4374 | 5.78 | 5.78 | 5.79 | 5.81 | 5.8 | 5.8 | 5.81 | 5.8 | 5.79 | 5.78 | 5.76 | 5.75 | 5.73 | 5.72 | 5.7 | 5.69 | 5.7 | 5.71 | 5.7 |
| 4375 | 5.78 | 5.78 | 5.79 | 5.81 | 5.8 | 5.8 | 5.81 | 5.8 | 5.79 | 5.78 | 5.76 | 5.75 | 5.73 | 5.72 | 5.7 | 5.69 | 5.7 | 5.71 | 5.7 |
| 4376 | 5.78 | 5.78 | 5.79 | 5.81 | 5.8 | 5.8 | 5.81 | 5.8 | 5.79 | 5.78 | 5.76 | 5.75 | 5.73 | 5.72 | 5.7 | 5.69 | 5.7 | 5.71 | 5.7 |
| 4377 | 5.78 | 5.78 | 5.79 | 5.81 | 5.8 | 5.8 | 5.81 | 5.8 | 5.79 | 5.78 | 5.76 | 5.75 | 5.73 | 5.72 | 5.7 | 5.69 | 5.7 | 5.71 | 5.7 |
| 4378 | 5.78 | 5.78 | 5.79 | 5.81 | 5.8 | 5.8 | 5.81 | 5.8 | 5.79 | 5.78 | 5.76 | 5.75 | 5.73 | 5.72 | 5.7 | 5.69 | 5.7 | 5.71 | 5.7 |
| 4379 | 5.78 | 5.78 | 5.79 | 5.81 | 5.8 | 5.8 | 5.81 | 5.8 | 5.79 | 5.78 | 5.76 | 5.75 | 5.73 | 5.72 | 5.7 | 5.69 | 5.7 | 5.71 | 5.7 |

Figure 4. Example of Positive Values where Material Falls into HWT Test

It is not clear why the HWT test continues to run in cases where the maximum rut depth is exceeded, or what the criteria for ending the test are. In Figures 5, the maximum rut depth is hit in at least one location around pass 18,000 and multiple locations soon after that, yet the test continues to run until pass 19,738. To accurately record failures (or any data associated with the HWT test), it is important to know the pass where the failure first occurs with consistent granularity. Why would the test continue to run for nearly 2,000 passes after failure? Based on the reporting, one would assume that the test fails at 19,738, but that is not the case. To get accurate failure results, I had to go through the file cell by cell to identify the exact first point of failure. See file jgerhrig1769051309 HWT1769051309 for example.

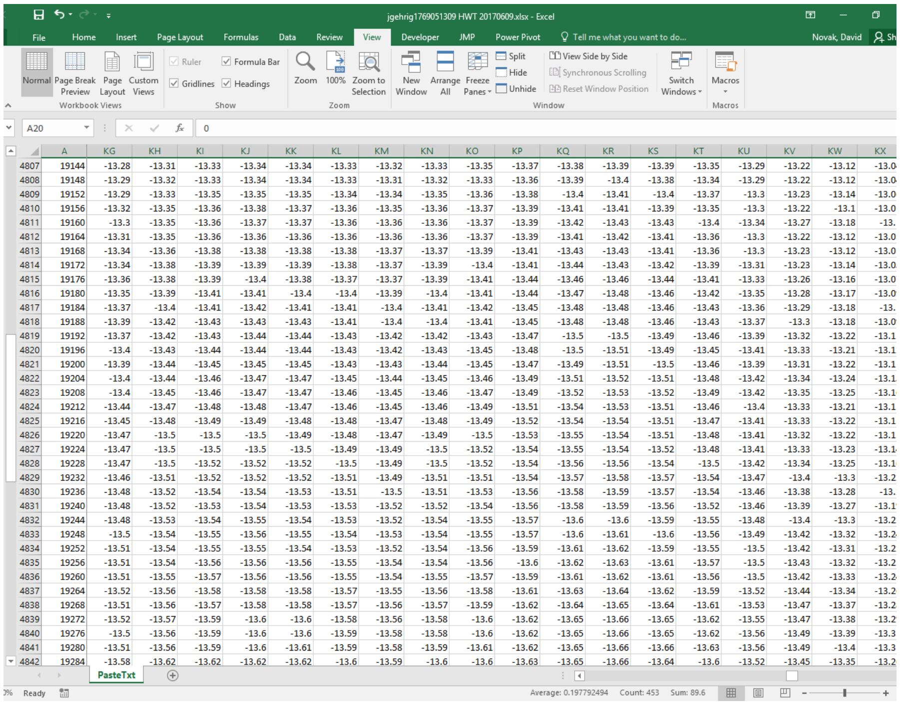



Figure 5. HWT Test Continuing to Run After Failure

Figure 6 provides another illustration of a situation where the max rut depth is clearly exceeded, yet the HWT test continues to run.

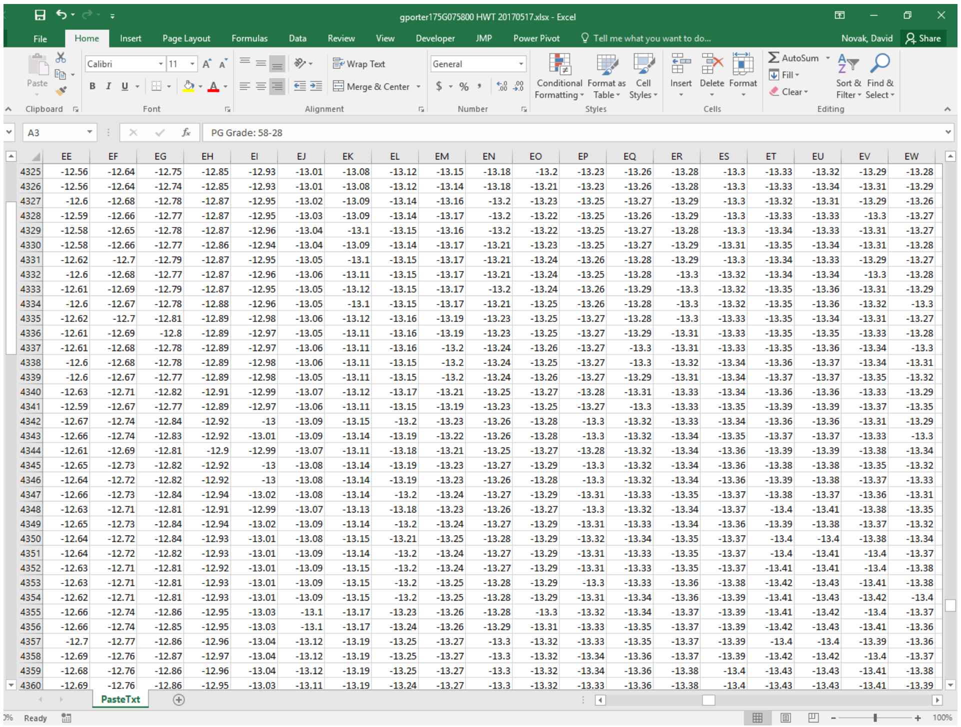
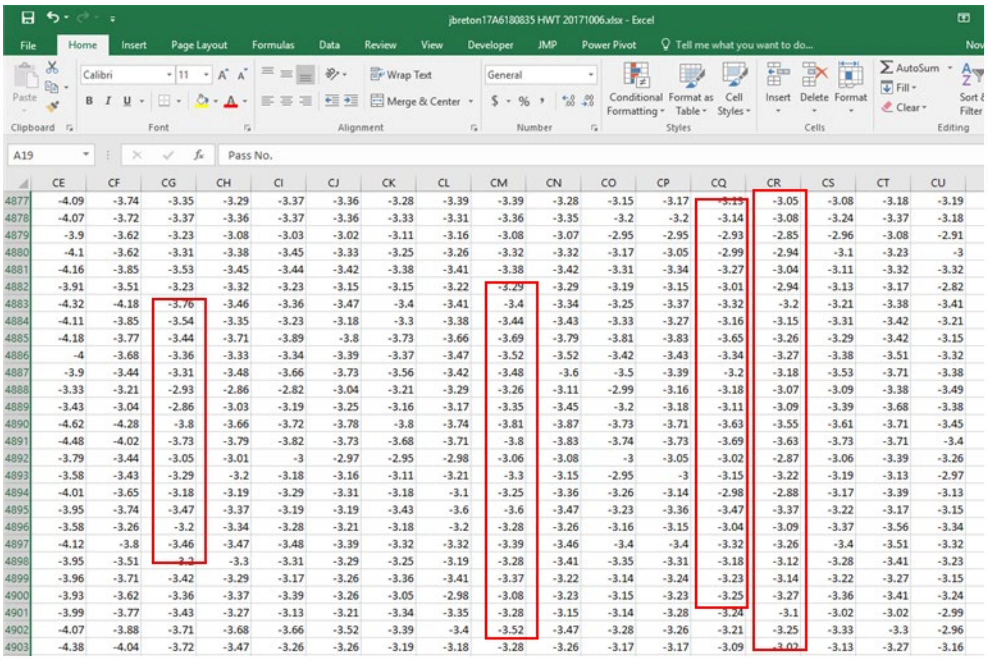



Figure 6. Max Rut Depth Exceeded – HWT Test Continues to Run

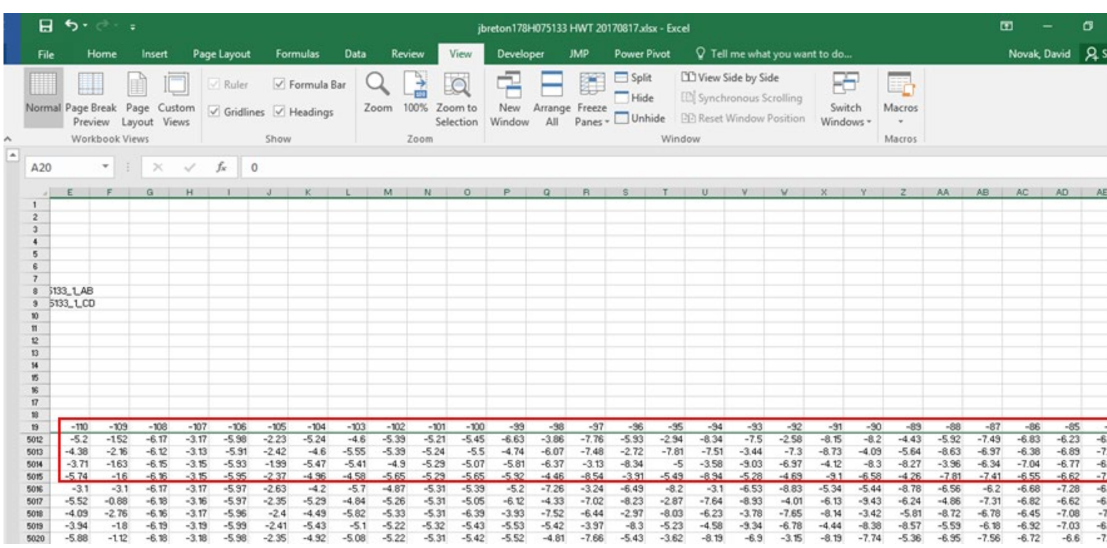
It is also clearly problematic to have values increasing, then decreasing, then increasing, then decreasing again. A reasonable question is what happens if this type of activity occurs right around a rut depth of -12.5mm? For example, if a test ends “successfully” and we see this type of jumping around (let’s say from -12.25, to -12.55, to -12.3, to -12.53, to -12.4 over the last 5 passes or so), what does this mean? I assume this is the jitter in the data that Troy and Aaron initially discussed with me. Has the issue been resolved? Is there a margin of error that is applied to jitter in every pass? This is unclear, but important in terms of analysis. An example of inconsistent data elements is provided in Figure 7. The rut depth should be less than or equal to the rut depth from the same position in subsequent passes.




| | CE | CF | CG | CH | CI | CJ | CK | CL | CM | CN | CO | CP | CQ | CR | CS | CT | CU |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 4877 | -4.09 | -3.74 | -3.35 | -3.29 | -3.37 | -3.36 | -3.28 | -3.39 | -3.39 | -3.28 | -3.15 | -3.17 | -3.15 | -3.05 | -3.08 | -3.18 | -3.19 |
| 4878 | -4.07 | -3.72 | -3.37 | -3.36 | -3.37 | -3.36 | -3.33 | -3.31 | -3.36 | -3.35 | -3.2 | -3.2 | -3.14 | -3.08 | -3.24 | -3.37 | -3.18 |
| 4879 | -3.9 | -3.62 | -3.23 | -3.08 | -3.03 | -3.02 | -3.11 | -3.16 | -3.08 | -3.07 | -2.95 | -2.95 | -2.93 | -2.85 | -2.96 | -3.08 | -2.91 |
| 4880 | -4.1 | -3.62 | -3.31 | -3.38 | -3.45 | -3.33 | -3.25 | -3.26 | -3.32 | -3.32 | -3.17 | -3.05 | -2.99 | -2.94 | -3.1 | -3.23 | -3 |
| 4881 | -4.16 | -3.85 | -3.53 | -3.45 | -3.44 | -3.42 | -3.38 | -3.41 | -3.38 | -3.42 | -3.31 | -3.34 | -3.27 | -3.04 | -3.11 | -3.32 | -3.32 |
| 4882 | -3.91 | -3.51 | -3.23 | -3.32 | -3.23 | -3.15 | -3.15 | -3.22 | -3.29 | -3.29 | -3.19 | -3.15 | -3.01 | -2.94 | -3.13 | -3.17 | -2.82 |
| 4883 | -4.32 | -4.18 | -3.76 | -3.46 | -3.36 | -3.47 | -3.4 | -3.41 | -3.4 | -3.34 | -3.25 | -3.37 | -3.32 | -3.2 | -3.21 | -3.38 | -3.41 |
| 4884 | -4.11 | -3.85 | -3.54 | -3.35 | -3.23 | -3.18 | -3.3 | -3.38 | -3.44 | -3.43 | -3.33 | -3.27 | -3.16 | -3.15 | -3.31 | -3.42 | -3.21 |
| 4885 | -4.18 | -3.77 | -3.44 | -3.71 | -3.89 | -3.8 | -3.73 | -3.66 | -3.69 | -3.79 | -3.81 | -3.83 | -3.65 | -3.26 | -3.29 | -3.42 | -3.15 |
| 4886 | -4 | -3.68 | -3.36 | -3.33 | -3.34 | -3.39 | -3.37 | -3.47 | -3.52 | -3.52 | -3.42 | -3.43 | -3.34 | -3.27 | -3.38 | -3.51 | -3.32 |
| 4887 | -3.9 | -3.44 | -3.31 | -3.48 | -3.66 | -3.73 | -3.56 | -3.42 | -3.48 | -3.6 | -3.5 | -3.39 | -3.2 | -3.18 | -3.53 | -3.71 | -3.38 |
| 4888 | -3.33 | -3.21 | -2.93 | -2.86 | -2.82 | -3.04 | -3.21 | -3.29 | -3.26 | -3.11 | -2.99 | -3.16 | -3.18 | -3.07 | -3.09 | -3.38 | -3.49 |
| 4889 | -3.43 | -3.04 | -2.86 | -3.03 | -3.19 | -3.25 | -3.16 | -3.17 | -3.35 | -3.45 | -3.2 | -3.18 | -3.11 | -3.09 | -3.39 | -3.68 | -3.38 |
| 4890 | -4.62 | -4.28 | -3.8 | -3.66 | -3.72 | -3.78 | -3.8 | -3.74 | -3.81 | -3.87 | -3.73 | -3.71 | -3.63 | -3.55 | -3.61 | -3.71 | -3.45 |
| 4891 | -4.48 | -4.02 | -3.73 | -3.79 | -3.82 | -3.73 | -3.68 | -3.71 | -3.8 | -3.83 | -3.74 | -3.73 | -3.69 | -3.63 | -3.73 | -3.71 | -3.4 |
| 4892 | -3.79 | -3.44 | -3.05 | -3.01 | -3 | -2.97 | -2.95 | -2.98 | -3.06 | -3.08 | -3 | -3.05 | -3.02 | -2.87 | -3.06 | -3.39 | -3.26 |
| 4893 | -3.58 | -3.43 | -3.29 | -3.2 | -3.18 | -3.16 | -3.11 | -3.21 | -3.3 | -3.15 | -2.95 | -3 | -3.15 | -3.22 | -3.19 | -3.13 | -2.97 |
| 4894 | -4.01 | -3.65 | -3.18 | -3.19 | -3.29 | -3.31 | -3.18 | -3.1 | -3.25 | -3.36 | -3.26 | -3.14 | -2.98 | -2.88 | -3.17 | -3.39 | -3.13 |
| 4895 | -3.95 | -3.74 | -3.47 | -3.37 | -3.19 | -3.19 | -3.43 | -3.6 | -3.6 | -3.47 | -3.23 | -3.36 | -3.47 | -3.37 | -3.22 | -3.17 | -3.15 |
| 4896 | -3.58 | -3.26 | -3.2 | -3.34 | -3.28 | -3.21 | -3.18 | -3.2 | -3.28 | -3.26 | -3.16 | -3.15 | -3.04 | -3.09 | -3.37 | -3.56 | -3.34 |
| 4897 | -4.12 | -3.8 | -3.46 | -3.47 | -3.48 | -3.39 | -3.32 | -3.32 | -3.39 | -3.46 | -3.4 | -3.4 | -3.32 | -3.26 | -3.4 | -3.51 | -3.32 |
| 4898 | -3.95 | -3.51 | -3.2 | -3.3 | -3.31 | -3.29 | -3.25 | -3.19 | -3.28 | -3.41 | -3.35 | -3.31 | -3.18 | -3.12 | -3.28 | -3.41 | -3.23 |
| 4899 | -3.96 | -3.71 | -3.42 | -3.29 | -3.17 | -3.26 | -3.36 | -3.41 | -3.37 | -3.22 | -3.14 | -3.24 | -3.23 | -3.14 | -3.22 | -3.27 | -3.15 |
| 4900 | -3.93 | -3.62 | -3.36 | -3.37 | -3.39 | -3.26 | -3.05 | -2.98 | -3.08 | -3.23 | -3.15 | -3.23 | -3.25 | -3.27 | -3.36 | -3.41 | -3.24 |
| 4901 | -3.99 | -3.77 | -3.43 | -3.27 | -3.13 | -3.21 | -3.34 | -3.35 | -3.28 | -3.15 | -3.14 | -3.28 | -3.24 | -3.1 | -3.02 | -3.02 | -2.99 |
| 4902 | -4.07 | -3.88 | -3.71 | -3.68 | -3.66 | -3.52 | -3.39 | -3.4 | -3.52 | -3.47 | -3.28 | -3.26 | -3.21 | -3.25 | -3.33 | -3.3 | -2.96 |
| 4903 | -4.38 | -4.04 | -3.72 | -3.47 | -3.26 | -3.26 | -3.19 | -3.18 | -3.28 | -3.26 | -3.17 | -3.17 | -3.09 | -3.02 | -3.13 | -3.27 | -3.16 |

Figure 7. Inconsistent Data Elements –Row-to-Row

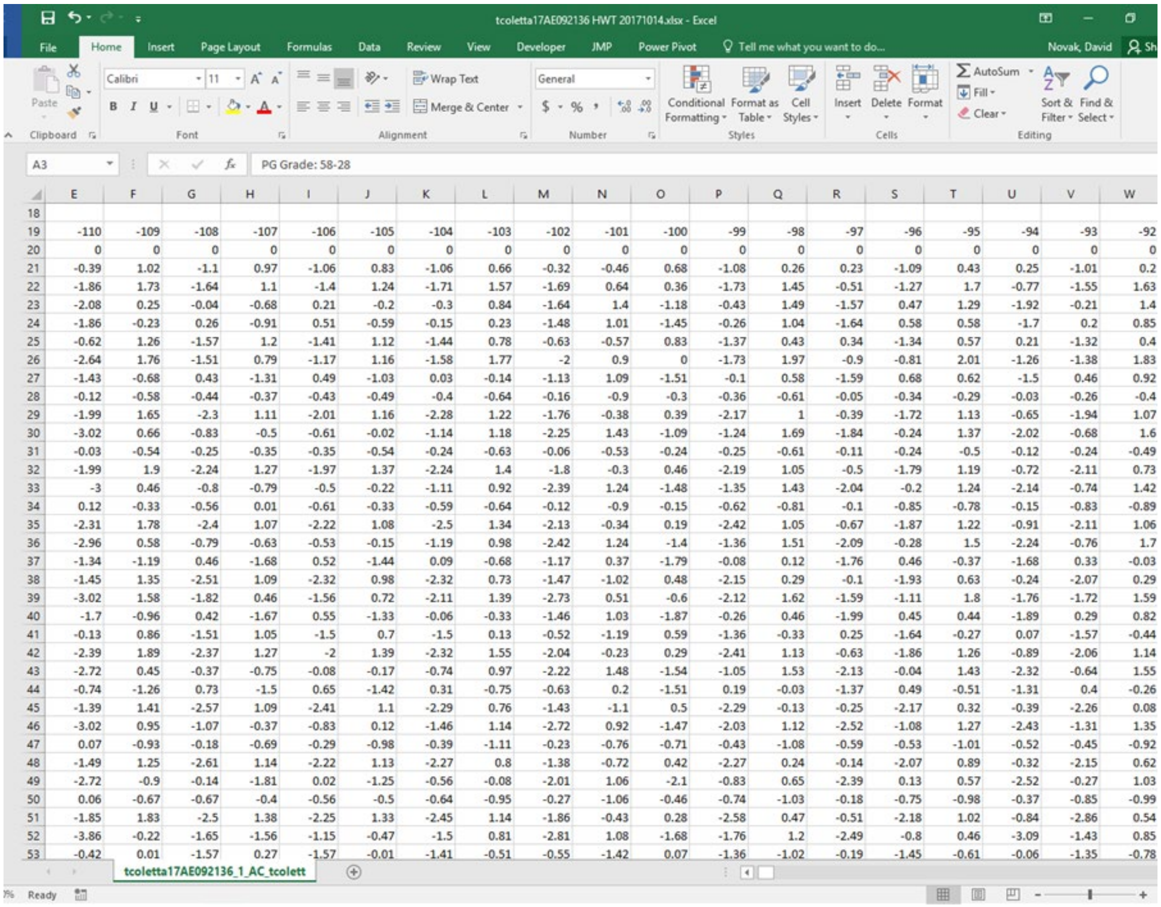
There are also numerous examples where the HWT test seems to bounce around between positive and negative values in adjacent positions by more than 2mm. This does not seem to be logically consistent. For example, given the parabolic nature of the HWT test, how can position -110 have a value of -5.88mm, while position -109 have a value of -1.12mm, position -108 have a value of -6.18, and position -107 have a value of -3.18. There should be relative continuity between the adjacent positions (see jbreton178H075133 HWT20170817 or gporter178E070851 HWT20170814 and others).




| 19 | -110 | -109 | -108 | -107 | -106 | -105 | -104 | -103 | -102 | -101 | -100 | -99 | -98 | -97 | -96 | -95 | -94 | -93 | -92 | -91 | -90 | -89 | -88 | -87 | -86 | -85 | -8 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 5012 | -5.2 | -1.52 | -6.17 | -3.17 | -5.98 | -2.23 | -5.24 | -4.6 | -5.39 | -5.21 | -5.45 | -6.63 | -3.86 | -7.76 | -5.93 | -2.94 | -8.34 | -7.5 | -2.58 | -8.15 | -8.2 | -4.43 | -5.92 | -7.49 | -6.83 | -6.23 | -6.7 |
| 5013 | -4.38 | -2.16 | -6.12 | -3.13 | -5.91 | -2.42 | -4.6 | -5.55 | -5.39 | -5.24 | -5.5 | -4.74 | -6.07 | -7.48 | -2.72 | -7.81 | -7.51 | -3.44 | -7.3 | -8.73 | -4.09 | -5.64 | -8.63 | -6.97 | -6.38 | -6.89 | -7.4 |
| 5014 | -3.71 | -1.63 | -6.15 | -3.15 | -5.93 | -1.99 | -5.47 | -5.41 | -4.9 | -5.29 | -5.07 | -5.81 | -6.37 | -3.13 | -8.34 | -5 | -3.58 | -9.03 | -6.97 | -4.12 | -8.3 | -8.27 | -3.96 | -6.34 | -7.04 | -6.77 | -6.5 |
| 5015 | -5.74 | -1.6 | -6.16 | -3.15 | -5.95 | -2.37 | -4.96 | -4.58 | -5.65 | -5.29 | -5.65 | -5.92 | -4.46 | -8.54 | -3.91 | -5.49 | -8.94 | -5.28 | -4.69 | -9.1 | -6.58 | -4.26 | -7.81 | -7.41 | -6.55 | -6.62 | -7.3 |
| 5016 | -3.1 | -3.1 | -6.17 | -3.17 | -5.97 | -2.63 | -4.2 | -5.7 | -4.87 | -5.31 | -5.39 | -5.2 | -7.26 | -3.24 | -6.49 | -8.2 | -3.1 | -6.53 | -8.83 | -5.34 | -5.44 | -8.78 | -6.56 | -6.2 | -6.68 | -7.28 | -6.7 |
| 5017 | -5.52 | -0.88 | -6.18 | -3.16 | -5.97 | -2.35 | -5.29 | -4.84 | -5.26 | -5.31 | -5.05 | -6.12 | -4.33 | -7.02 | -8.23 | -2.87 | -7.64 | -8.93 | -4.01 | -6.13 | -9.43 | -6.24 | -4.86 | -7.31 | -6.82 | -6.62 | -6.6 |
| 5018 | -4.09 | -2.76 | -6.16 | -3.17 | -5.96 | -2.4 | -4.49 | -5.82 | -5.33 | -5.31 | -6.39 | -3.93 | -7.52 | -6.44 | -2.97 | -8.03 | -6.23 | -3.78 | -7.65 | -8.14 | -3.42 | -5.81 | -8.72 | -6.78 | -6.45 | -7.08 | -7.1 |
| 5019 | -3.94 | -1.8 | -6.19 | -3.19 | -5.99 | -2.41 | -5.43 | -5.1 | -5.22 | -5.32 | -5.43 | -5.53 | -5.42 | -3.97 | -8.3 | -5.23 | -4.58 | -9.34 | -6.78 | -4.44 | -8.38 | -8.57 | -5.59 | -6.18 | -6.92 | -7.03 | -6.7 |
| 5020 | -5.88 | -1.12 | -6.18 | -3.18 | -5.98 | -2.35 | -4.92 | -5.08 | -5.22 | -5.31 | -5.42 | -5.52 | -4.81 | -7.66 | -5.43 | -3.62 | -8.19 | -6.9 | -3.15 | -8.19 | -7.74 | -5.36 | -6.95 | -7.56 | -6.72 | -6.6 | -7.0 |

Figure 8. Inconsistent Data Elements – Column-to- Column

There are also files that have notable inconsistencies between adjacent columns as well as from row to row. For example, see tcoletta17AE092136 HWT20171014.



tcoletta17AE092136 HWT 20171014.xlsx - Excel

A3 | PG Grade: 58-28

| | E | F | G | H | I | J | K | L | M | N | O | P | Q | R | S | T | U | V | W |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 18 | | | | | | | | | | | | | | | | | | | |
| 19 | -110 | -109 | -108 | -107 | -106 | -105 | -104 | -103 | -102 | -101 | -100 | -99 | -98 | -97 | -96 | -95 | -94 | -93 | -92 |
| 20 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 21 | -0.39 | 1.02 | -1.1 | 0.97 | -1.06 | 0.83 | -1.06 | 0.66 | -0.32 | -0.46 | 0.68 | -1.08 | 0.26 | 0.23 | -1.09 | 0.43 | 0.25 | -1.01 | 0.2 |
| 22 | -1.86 | 1.73 | -1.64 | 1.1 | -1.4 | 1.24 | -1.71 | 1.57 | -1.69 | 0.64 | 0.36 | -1.73 | 1.45 | -0.51 | -1.27 | 1.7 | -0.77 | -1.55 | 1.63 |
| 23 | -2.08 | 0.25 | -0.04 | -0.68 | 0.21 | -0.2 | -0.3 | 0.84 | -1.64 | 1.4 | -1.18 | -0.43 | 1.49 | -1.57 | 0.47 | 1.29 | -1.92 | -0.21 | 1.4 |
| 24 | -1.86 | -0.23 | 0.26 | -0.91 | 0.51 | -0.59 | -0.15 | 0.23 | -1.48 | 1.01 | -1.45 | -0.26 | 1.04 | -1.64 | 0.58 | 0.58 | -1.7 | 0.2 | 0.85 |
| 25 | -0.62 | 1.26 | -1.57 | 1.2 | -1.41 | 1.12 | -1.44 | 0.78 | -0.63 | -0.57 | 0.83 | -1.37 | 0.43 | 0.34 | -1.34 | 0.57 | 0.21 | -1.32 | 0.4 |
| 26 | -2.64 | 1.76 | -1.51 | 0.79 | -1.17 | 1.16 | -1.58 | 1.77 | -2 | 0.9 | 0 | -1.73 | 1.97 | -0.9 | -0.81 | 2.01 | -1.26 | -1.38 | 1.83 |
| 27 | -1.43 | -0.68 | 0.43 | -1.31 | 0.49 | -1.03 | 0.03 | -0.14 | -1.13 | 1.09 | -1.51 | -0.1 | 0.58 | -1.59 | 0.68 | 0.62 | -1.5 | 0.46 | 0.92 |
| 28 | -0.12 | -0.58 | -0.44 | -0.37 | -0.43 | -0.49 | -0.4 | -0.64 | -0.16 | -0.9 | -0.3 | -0.36 | -0.61 | -0.05 | -0.34 | -0.29 | -0.03 | -0.26 | -0.4 |
| 29 | -1.99 | 1.65 | -2.3 | 1.11 | -2.01 | 1.16 | -2.28 | 1.22 | -1.76 | -0.38 | 0.39 | -2.17 | 1 | -0.39 | -1.72 | 1.13 | -0.65 | -1.94 | 1.07 |
| 30 | -3.02 | 0.66 | -0.83 | -0.5 | -0.61 | -0.02 | -1.14 | 1.18 | -2.25 | 1.43 | -1.09 | -1.24 | 1.69 | -1.84 | -0.24 | 1.37 | -2.02 | -0.68 | 1.6 |
| 31 | -0.03 | -0.54 | -0.25 | -0.35 | -0.35 | -0.54 | -0.24 | -0.63 | -0.06 | -0.53 | -0.24 | -0.25 | -0.61 | -0.11 | -0.24 | -0.5 | -0.12 | -0.24 | -0.49 |
| 32 | -1.99 | 1.9 | -2.24 | 1.27 | -1.97 | 1.37 | -2.24 | 1.4 | -1.8 | -0.3 | 0.46 | -2.19 | 1.05 | -0.5 | -1.79 | 1.19 | -0.72 | -2.11 | 0.73 |
| 33 | -3 | 0.46 | -0.8 | -0.79 | -0.5 | -0.22 | -1.11 | 0.92 | -2.39 | 1.24 | -1.48 | -1.35 | 1.43 | -2.04 | -0.2 | 1.24 | -2.14 | -0.74 | 1.42 |
| 34 | 0.12 | -0.33 | -0.56 | 0.01 | -0.61 | -0.33 | -0.59 | -0.64 | -0.12 | -0.9 | -0.15 | -0.62 | -0.81 | -0.1 | -0.85 | -0.78 | -0.15 | -0.83 | -0.89 |
| 35 | -2.31 | 1.78 | -2.4 | 1.07 | -2.22 | 1.08 | -2.5 | 1.34 | -2.13 | -0.34 | 0.19 | -2.42 | 1.05 | -0.67 | -1.87 | 1.22 | -0.91 | -2.11 | 1.06 |
| 36 | -2.96 | 0.58 | -0.79 | -0.63 | -0.53 | -0.15 | -1.19 | 0.98 | -2.42 | 1.24 | -1.4 | -1.36 | 1.51 | -2.09 | -0.28 | 1.5 | -2.24 | -0.76 | 1.7 |
| 37 | -1.34 | -1.19 | 0.46 | -1.68 | 0.52 | -1.44 | 0.09 | -0.68 | -1.17 | 0.37 | -1.79 | -0.08 | 0.12 | -1.76 | 0.46 | -0.37 | -1.68 | 0.33 | -0.03 |
| 38 | -1.45 | 1.35 | -2.51 | 1.09 | -2.32 | 0.98 | -2.32 | 0.73 | -1.47 | -1.02 | 0.48 | -2.15 | 0.29 | -0.1 | -1.93 | 0.63 | -0.24 | -2.07 | 0.29 |
| 39 | -3.02 | 1.58 | -1.82 | 0.46 | -1.56 | 0.72 | -2.11 | 1.39 | -2.73 | 0.51 | -0.6 | -2.12 | 1.62 | -1.59 | -1.11 | 1.8 | -1.76 | -1.72 | 1.59 |
| 40 | -1.7 | -0.96 | 0.42 | -1.67 | 0.55 | -1.33 | -0.06 | -0.33 | -1.46 | 1.03 | -1.87 | -0.26 | 0.46 | -1.99 | 0.45 | 0.44 | -1.89 | 0.29 | 0.82 |
| 41 | -0.13 | 0.86 | -1.51 | 1.05 | -1.5 | 0.7 | -1.5 | 0.13 | -0.52 | -1.19 | 0.59 | -1.36 | -0.33 | 0.25 | -1.64 | -0.27 | 0.07 | -1.57 | -0.44 |
| 42 | -2.39 | 1.89 | -2.37 | 1.27 | -2 | 1.39 | -2.32 | 1.55 | -2.04 | -0.23 | 0.29 | -2.41 | 1.13 | -0.63 | -1.86 | 1.26 | -0.89 | -2.06 | 1.14 |
| 43 | -2.72 | 0.45 | -0.37 | -0.75 | -0.08 | -0.17 | -0.74 | 0.97 | -2.22 | 1.48 | -1.54 | -1.05 | 1.53 | -2.13 | -0.04 | 1.43 | -2.32 | -0.64 | 1.55 |
| 44 | -0.74 | -1.26 | 0.73 | -1.5 | 0.65 | -1.42 | 0.31 | -0.75 | -0.63 | 0.2 | -1.51 | 0.19 | -0.03 | -1.37 | 0.49 | -0.51 | -1.31 | 0.4 | -0.26 |
| 45 | -1.39 | 1.41 | -2.57 | 1.09 | -2.41 | 1.1 | -2.29 | 0.76 | -1.43 | -1.1 | 0.5 | -2.29 | -0.13 | -0.25 | -2.17 | 0.32 | -0.39 | -2.26 | 0.08 |
| 46 | -3.02 | 0.95 | -1.07 | -0.37 | -0.83 | 0.12 | -1.46 | 1.14 | -2.72 | 0.92 | -1.47 | -2.03 | 1.12 | -2.52 | -1.08 | 1.27 | -2.43 | -1.31 | 1.35 |
| 47 | 0.07 | -0.93 | -0.18 | -0.69 | -0.29 | -0.98 | -0.39 | -1.11 | -0.23 | -0.76 | -0.71 | -0.43 | -1.08 | -0.59 | -0.53 | -1.01 | -0.52 | -0.45 | -0.92 |
| 48 | -1.49 | 1.25 | -2.61 | 1.14 | -2.22 | 1.13 | -2.27 | 0.8 | -1.38 | -0.72 | 0.42 | -2.27 | 0.24 | -0.14 | -2.07 | 0.89 | -0.32 | -2.15 | 0.62 |
| 49 | -2.72 | -0.9 | -0.14 | -1.81 | 0.02 | -1.25 | -0.56 | -0.08 | -2.01 | 1.06 | -2.1 | -0.83 | 0.65 | -2.39 | 0.13 | 0.57 | -2.52 | -0.27 | 1.03 |
| 50 | 0.06 | -0.67 | -0.67 | -0.4 | -0.56 | -0.5 | -0.64 | -0.95 | -0.27 | -1.06 | -0.46 | -0.74 | -1.03 | -0.18 | -0.75 | -0.98 | -0.37 | -0.85 | -0.99 |
| 51 | -1.85 | 1.83 | -2.5 | 1.38 | -2.25 | 1.33 | -2.45 | 1.14 | -1.86 | -0.43 | 0.28 | -2.58 | 0.47 | -0.51 | -2.18 | 1.02 | -0.84 | -2.86 | 0.54 |
| 52 | -3.86 | -0.22 | -1.65 | -1.56 | -1.15 | -0.47 | -1.5 | 0.81 | -2.81 | 1.08 | -1.68 | -1.76 | 1.2 | -2.49 | -0.8 | 0.46 | -3.09 | -1.43 | 0.85 |
| 53 | -0.42 | 0.01 | -1.57 | 0.27 | -1.57 | -0.01 | -1.41 | -0.51 | -0.55 | -1.42 | 0.07 | -1.36 | -1.02 | -0.19 | -1.45 | -0.61 | -0.06 | -1.35 | -0.78 |

tcoletta17AE092136_1_AC_tcolett

Figure 9.  Data Inconsistencies – Both Columns and Rows

Finally, in some tests the samples have the exact same sample identification values. In other cases, the samples have slightly different sample identification values. Is it important to distinguish between the tests that have different versus identical sample identification values? This is not clear.

## Recommendation for Moving Forward

Based on an initial review of the HWT data, I have more questions than answers. The data inconsistencies mentioned above are quite serious with respect to undermining the goal of developing a programmatic QA/QC program for asphalt, and must be addressed. I suggest that the Agency develop and implement data management plan for the HWT test. This would specifically include a plan for determining: 1) exactly what data need to be collected, 2) how those data should be organized and stored (a database?), 3) the development and implementation of data integrity constraints, 4) clear guidance for determining when a particular test is invalid, 5) the ability to clearly explain how the test works and how the results are interpreted to the contactors. I have some thoughts / concerns bulleted below.

- It is not clear to me what types of analyses the Agency would like to be able to conduct with the HWT data. The current data collection and reporting format appears unwieldy. It seems that the current approach allows the Agency to examine the results from the individual data files, but nothing beyond that. The data do not allow comparative analysis.
- Excel may not be an adequate software platform for this particular data management task.
- If the data are going to be used to develop pay factors and then to conduct compliance testing, they **MUST** be accurate (error free and validated), complete (not missing any parts), and consistent

(reconcilable and unbiased). At this point, they are definitely not. The Agency will need to clean the existing data and be able to guarantee that the data are reliable. Imagine a situation where a sample is rejected, the contractor then asks to see the test results, and the test file states "test concluded successfully". The contractor should rightly question the validity of the test (and maybe the entire QA/QC process) – even if the max rut depth of -12.5 mm is clearly exceeded. No test where the maximum rut depth is exceeded should state, "test concluded successfully".

- The Agency needs accurate summary statistics for the HWT data and needs to be able to compare one sample to another as well as one file (set of samples) to another. The wheel tracker seems to generate a parabolic rut in most cases. Which positions for each sample are most important (-110, 0, 110)? Others? How are these values to be calculated – point values? An average? For a comparison, in-place bridge concrete has a well-defined compression test that occurs at 28 days. There is one statistic of interest for each sample – the concrete compressive strength at 28 days. What statistic or set of statistics will be used for the HWT test? Are there going to be multiple data points of interest for each sample? Before this can be determined, we need to be confident that the data are accurate, complete, and consistent.
- What other HWT data fields are important in the testing process? There was no focus on any of the other fields in our discussions. What about the other fields?
- What pass resolution is needed – every 100$^{th}$ pass, every 1,000$^{th}$ pass? Identifying the failure point seems important, how will the Agency do this if we only collect data from every 100$^{th}$ or 1,000$^{th}$ pass – manually inspect each file?
- What is the process for handling samples that have positive values or inconsistencies between rows and or columns? Are all of these samples invalid?
- The Agency should establish baseline data integrity conditions from column to column as well as from row to row and then implement integrity constraints for each set of samples, so that they do not have to visually inspect every row and every column for every file. This will require some type of automation (coding) to implement these conditions once the Agency determines exactly what they should be.
- Are there other state agencies that have used HWT data to develop pay factors? If so, looking into exactly how they manage the data would be very useful. It is essential that the Agency develop clear guidance for determining when a particular test is invalid so that they can explain this to the contractors. Since the contractors won't have the ability to conduct their own HWT tests, any inconsistencies or data failures on the part of the Agency will create problems with respect to the use of pay factors.

## Description of Additional Work – Following up on October 26, 2018 meeting

I only worked with the 20K pass files. My goal was to put together a spreadsheet that VTrans can use to filter HWT test results as they see fit. The summary-revised spreadsheet consists only of passes: 5,000, 10,000, 15,000, 20,000 (where they exist) and columns: 46, 23, 0, -23, -46 for both the left and right samples. Only some passes were recorded for each failed test. For example, gporter168B174144 HWT 1609081034.xlsx stops at pass 18,970 and only includes data for passes 5,000, 10,000, 15,000.

The following files are excluded from the analysis: ahicks15A2062822 HWT 20151002, gporter164S083448 HWT 1605101047.xlsx, gporter 166B124302 HWT 1606300920.xlsx, gporter165D112152 1606271131.xlsx, irothlon1666070713 HWT 1606271618.xlsx, irothlon1669153000 HWT 1606281514.xlsx, tcoletta1659094018 HWT 1606240913.xlsx as positions 46, 23, -23, -46 do not exist in raw data files that I received.

From summary-revised.xlsx "Both Sides" tab (441 total records)

Left side sample value is positive (10 occurrences): rows 9, 30, 43, 52, 56, 60, 84, 147, 151, 185

Right side sample value is positive (18 occurrences): rows 33, 55, 63, 67, 74, 94, 106, 121, 124, 130, 134, 137, 139, 140, 141, 144, 177, 181.

I removed all samples where one side has positive values in "Both Clean" tab to test to see if mean and variance are statistically different between left side and right side samples. Data for tests are in "Tests" tab, where I use JMP from SAS to compare left and right values for each of the 5 positions for each side.

I conducted two separate tests to compare the means of the right and left side samples: 1) Nonparametric Comparisons using Wilcoxon Method where $H_0$ = means are equal, and 2) Tukey-Kramer HSD where a positive value indicates pairs of means that are significantly different.

| Position | -46 | -23 | 0 | 23 | 46 |
|---|---|---|---|---|---|
| Wilcoxon<br>p-value (α = 0.05) | 0.1798<br>fail to reject | 0.4944<br>fail to reject | 0.7009<br>fail to reject | 0.5113<br>fail to reject | 0.2187<br>fail to reject |
| Tukey-Kramer HSD (α = 0.05) | L/R -0.0813<br>no difference | L/R -0.1289<br>no difference | L/R -0.2221<br>no difference | L/R -0.2038<br>no difference | L/R -0.1552<br>no difference |

Based on the results of both tests (which are consistent for all positions), we can conclude that the depth values associated with the right and left samples are not significantly different from one another. Thus, we could work with just one sample (L or R), or even take the average between the two to create a single sample for each position (-46, -23, 0, 23, 46) for each pass 5,000, 10,000, 15,000, 20,000 for each test. This simplifies future analysis.

The only Mix Type that has a large enough sample to work with is Type IVS w/RAP (without adjusting for PG Grade). At pass 20,000 there are 31 observations that include both passed and failed tests (if the test made it that far). Pass 20,000 data are the most important data with respect to constructing pay factors (similar to the 28 day CCS for in-place concrete). All data for Mix Type IVS w/RAP is contained on the spreadsheet tab with the same label. VTrans can filter these (or any data) as needed. Summary values for each of the four positions is provided below for Mix Type IVS w/RAP.

| 5000 | L (-46) | L (-23) | L (0) | L (23) | L (46) | R (-46) | R (-23) | R (0) | R (23) | R (46) |
|---|---|---|---|---|---|---|---|---|---|---|
| Min | -12.01 | -11.63 | -12.3 | -13.29 | -12.42 | -10.64 | -9.23 | -9.04 | -8.77 | -7.91 |
| Max | -1.25 | -1.4 | -1.48 | -1.47 | -0.65 | -1.26 | -1.27 | -1.34 | -1.27 | -1.32 |
| Mean | -3.513 | -3.753 | -4.094 | -3.911 | -3.495 | -3.656 | -3.764 | -3.905 | -3.774 | -3.509 |
| Std | 1.736 | 1.773 | 1.916 | 1.874 | 1.811 | 1.737 | 1.638 | 1.610 | 1.579 | 1.459 |

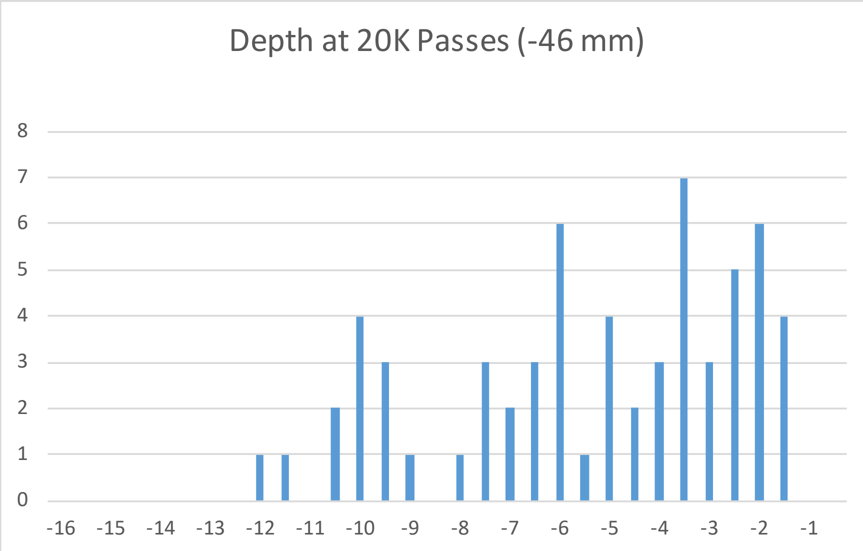
| 10000 | L (-46) | L (-23) | L (0) | L (23) | L (46) | R (-46) | R (-23) | R (0) | R (23) | R (46) |
|---|---|---|---|---|---|---|---|---|---|---|
| Min | -10.61 | -12.3 | -11.23 | -12.85 | -13.73 | -12.04 | -12.29 | -11.71 | -11.91 | -12.16 |
| Max | -1.55 | -1.76 | -1.83 | -1.74 | -1.61 | -1.55 | -1.57 | -1.6 | -1.58 | -1.56 |
| Mean | -4.410 | -4.482 | -4.741 | -4.645 | -4.423 | -4.860 | -4.895 | -4.853 | -4.949 | -4.916 |
| Std | 2.066 | 2.115 | 2.073 | 2.098 | 2.232 | 2.538 | 2.510 | 2.336 | 2.524 | 2.684 |

| 15000 | L (-46) | L (-23) | L (0) | L (23) | L (46) | R (-46) | R (-23) | R (0) | R (23) | R (46) |
|---|---|---|---|---|---|---|---|---|---|---|
| Min | -13.06 | -13.22 | -12.44 | -11.83 | -11.69 | -12.14 | -12.52 | -12 | -12.24 | -12.4 |
| Max | -1.67 | -1.95 | -2.07 | -2.03 | -1.82 | -1.7 | -1.73 | -1.8 | -1.8 | -1.74 |
| Mean | -5.381 | -5.314 | -5.270 | -5.305 | -5.278 | -5.819 | -5.705 | -5.297 | -5.436 | -5.609 |
| Std | 2.854 | 2.659 | 2.409 | 2.527 | 2.706 | 2.686 | 2.556 | 2.066 | 2.374 | 2.825 |

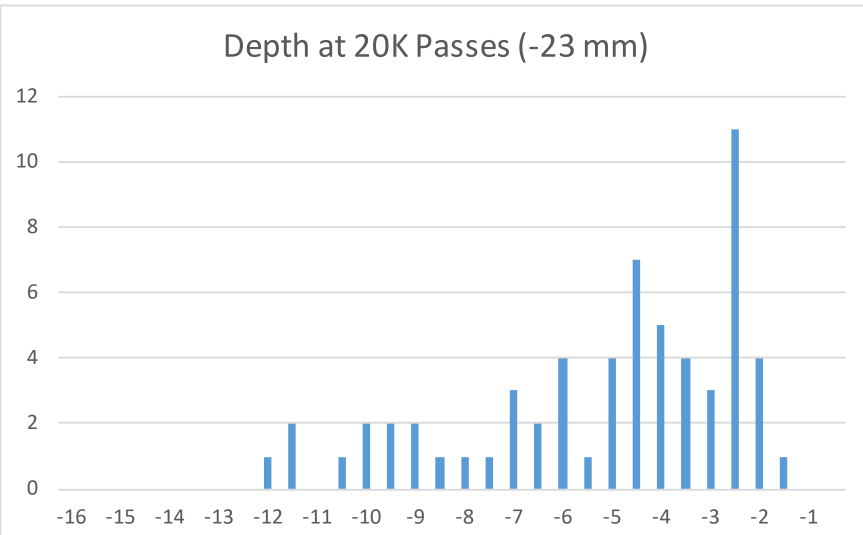
| 20000 | L (-46) | L (-23) | L (0) | L (23) | L (46) | R (-46) | R (-23) | R (0) | R (23) | R (46) |
|---|---|---|---|---|---|---|---|---|---|---|
| Min | -11.65 | -10.36 | -11.99 | -13.71 | -13.63 | -12.24 | -12.31 | -12.31 | -15.04 | -15.2 |
| Max | -1.79 | -2.16 | -2.27 | -2.21 | -1.98 | -1.8 | -1.88 | -1.96 | -2.04 | -1.82 |
| Mean | -5.196 | -5.055 | -4.965 | -5.136 | -5.247 | -6.063 | -5.859 | -5.454 | -5.696 | -5.795 |
| Std | 2.686 | 2.493 | 2.245 | 2.580 | 2.935 | 3.173 | 3.071 | 2.644 | 3.094 | 3.448 |

To be thorough, I went back through and compared the means from the R and L sides for just the 20,000 pass values. The results were the same as those for the entire sample consisting of all passes. For the nonparametric Wilcoxon test, all p-values were sufficiently large to fail to reject the null hypothesis that the means are equal. For Tukey-Kramer HSD, all test values were negative indicating no statistically significant difference between the two sides.

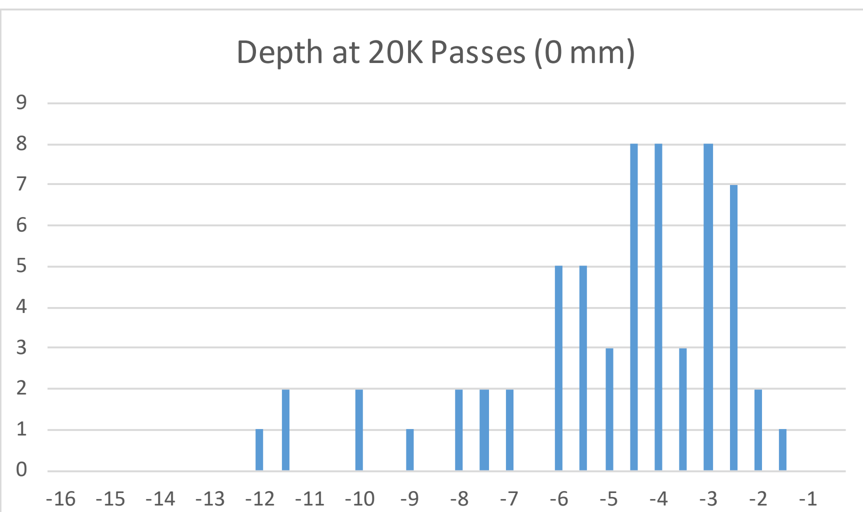
Histograms and summary statistics are provided for pass 20,000 for Mix Type IVS w/RAP for all PG Binder types combined in tab “20000 analysis”.



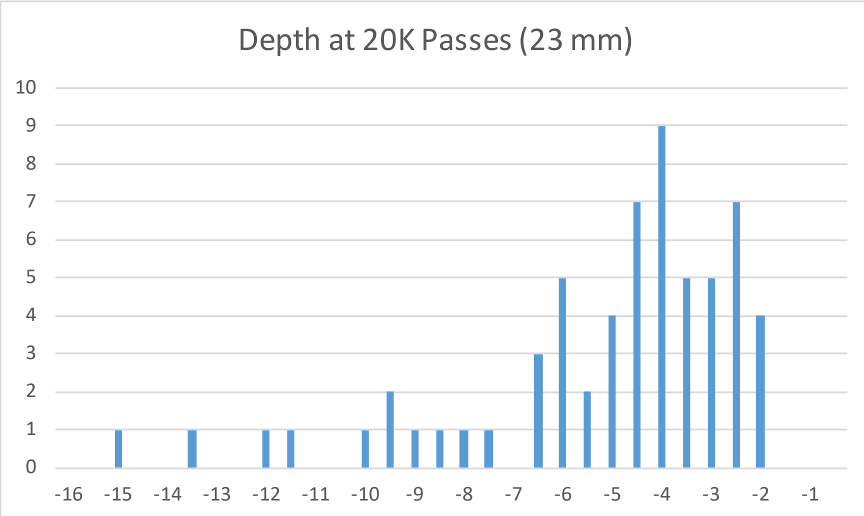

| *Pass 20K (-46 mm)* | |
|---|---|
| Mean | -5.62935 |
| Standard Error | 0.374441 |
| Median | -5.13 |
| Mode | #N/A |
| Standard Deviation | 2.948352 |
| Sample Variance | 8.69278 |
| Kurtosis | -0.8264 |
| Skewness | -0.5734 |
| Range | 10.45 |
| Minimum | -12.24 |
| Maximum | -1.79 |
| Sum | -349.02 |
| Count | 62 |
| Confidence Level(95.0%) | 0.748741 |



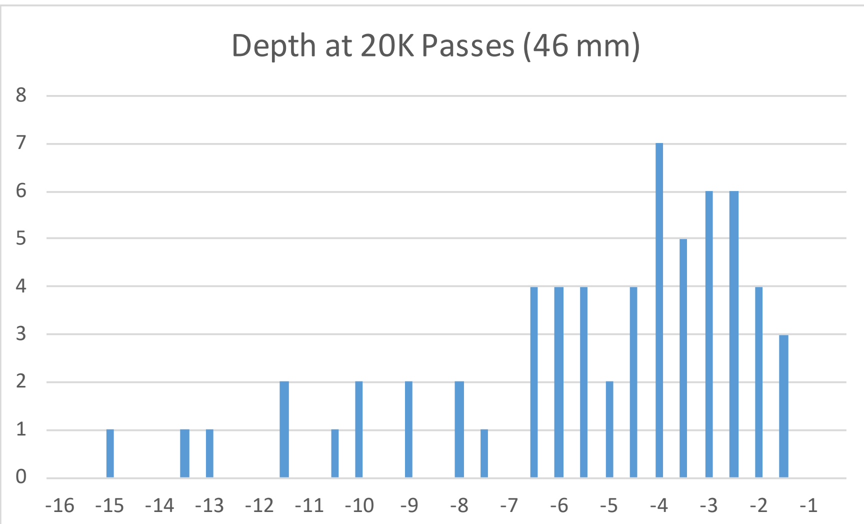

| *Pass 20K (-23 mm)* | |
|---|---|
| Mean | -5.45694 |
| Standard Error | 0.356026 |
| Median | -4.755 |
| Mode | -4.95 |
| Standard Deviation | 2.80335 |
| Sample Variance | 7.858769 |
| Kurtosis | -0.29316 |
| Skewness | -0.83594 |
| Range | 10.43 |
| Minimum | -12.31 |
| Maximum | -1.88 |
| Sum | -338.33 |
| Count | 62 |
| Confidence Level(95.0%) | 0.711918 |




| *20K passes (0 mm)* | |
|---|---|
| Mean | -5.20952 |
| Standard Error | 0.3105 |
| Median | -4.695 |
| Mode | #N/A |
| Standard Deviation | 2.444876 |
| Sample Variance | 5.977418 |
| Kurtosis | 1.374436 |
| Skewness | -1.28389 |
| Range | 10.35 |
| Minimum | -12.31 |
| Maximum | -1.96 |
| Sum | -322.99 |
| Count | 62 |
| Confidence Level(95.0%) | 0.620882 |




| *20K Passes (23 mm)* | |
|---|---|
| | |
| Mean | -5.41597 |
| Standard Error | 0.360595 |
| Median | -4.72 |
| Mode | -4.34 |
| Standard Deviation | 2.839332 |
| Sample Variance | 8.061805 |
| Kurtosis | 2.191319 |
| Skewness | -1.49877 |
| Range | 13 |
| Minimum | -15.04 |
| Maximum | -2.04 |
| Sum | -335.79 |
| Count | 62 |
| Confidence Level(95.0%) | 0.721055 |




| *20K Passes (46 mm)* | |
|---|---|
| | |
| Mean | -5.52129 |
| Standard Error | 0.404797 |
| Median | -4.47 |
| Mode | -6.37 |
| Standard Deviation | 3.187371 |
| Sample Variance | 10.15934 |
| Kurtosis | 1.051116 |
| Skewness | -1.26686 |
| Range | 13.38 |
| Minimum | -15.2 |
| Maximum | -1.82 |
| Sum | -342.32 |
| Count | 62 |
| Confidence Level(95.0%) | 0.809441 |