



REFERENCE SAMPLE PROGRAM

**METHOD OF TEST FOR
DETERMINING THEORETICAL
MAXIMUM SPECIFIC GRAVITY AND
DENSITY OF HOT MIX ASPHALT**

2011 PROFICIENCY TEST RESULTS

State of California Department of Transportation
Office of Roadway Materials Testing
Division of Engineering Services
Materials Engineering and Testing Services-MS #5
5900 Folsom Blvd
Sacramento, California 95819-4612

November 11, 2010

TABLE OF CONTENTS

| | | |
|------------|---------------------------------------|----------|
| 1.0 | OVERVIEW | 1 |
| 2.0 | ANALYSIS OF TEST RESULTS | 2 |
| 2.1 | EVALUATION CRITERIA | 2 |
| 2.2 | INITIAL TEST | 2 |
| 2.3 | RETEST | 4 |
| 2.4 | COMBINED RESULTS | 4 |
| 2.5 | OBSERVATIONS..... | 4 |
| 3.0 | SUMMARY | 6 |
| 4.0 | REFERENCES | 6 |

Appendix A – Test Results from Initial Test

Appendix B – Test Results from Retest

Appendix C – Combined Final Scores from both Initial Test and Retest

LIST OF TABLES

| | |
|--|---|
| Table 1. Evaluation Criteria..... | 2 |
| Table 2. Summary of Individual Initial Test Results (1200 grams) | 2 |
| Table 3. Summary of Individual Initial Test Results (2400 grams) | 3 |
| Table 4. Labs and Test Results Considered as Outliers..... | 3 |
| Table 5. Summary of Initial Combined Test Results..... | 3 |
| Table 6. Summary of Retest Results..... | 4 |
| Table 7. Summary of Combined Test Results | 4 |

REFERENCE SAMPLE PROGRAM

DETERMINING THEORETICAL MAXIMUM SPECIFIC GRAVITY AND DENSITY OF HOT MIX ASPHALT

2011 PROFICIENCY TEST RESULTS

1.0 OVERVIEW

In late 2010, the proficiency test for determining the theoretical maximum specific gravity and density of hot mix asphalt was initiated. The proficiency test was based on the California Test (CT) 309, "Method of Test for Determining Theoretical Maximum (RICE) Specific Gravity and Density of Hot Mix Asphalt." Laboratory mixed samples with a known asphalt content and aggregate gradation was used to minimize sample variability.

The Rice specific gravity test determines values for percent air voids in compacted HMA, establishes target values for compacting HMA, and determines the amount of binder absorbed by the individual aggregate particles in HMA.

The laboratory mixed samples were prepared in accordance to CT 304. The target sample mass was 1200 grams using ½-inch nominal maximum aggregate size with 5% binder asphalt content. The materials used for the HMA samples (asphalt binder and aggregate) originated from a single source.

Based on previous proficiency test using CT 309 it was deemed that the 1200 gram sample was adequate to provide accurate results. However, during the initial sample distribution some laboratories expressed concern that the 1200 gram sample may not be enough to produce acceptable results. It was also pointed out that the latest test method specifies a minimum sample mass of 2400 grams for a ½-inch mix. To mitigate this issue, the participants were provided with two options. The first option would be to conduct the test on the 1200 gram sample and use these results in the statistical analysis. Should these results fail; an additional 2400 gram sample will be provided for a retest without penalty. It should be noted that the first option was considered due to a limited number of samples available. The second option was to provide participating laboratories with an additional 1200 gram sample to be combined with the first 1200 gram sample to satisfy the minimum sample size requirement of 2400 grams. A total of 56 laboratories opted for option 1 and 48 laboratories opted for option 2.

The premise of the RICE specific gravity proficiency test was to determine the participating laboratories proficiency in conducting the test and their ability to achieve statistically satisfactory results.

The test results are comprised of two data sets which are based on the 1200 and 2400 gram sample mass. Statistical analysis was conducted on each individual data sets as well as the combined data set.

The analysis for the combined set was used as the basis for the proficiency scores. It should also be noted that the analysis of the two data sets based on sample mass provided an opportunity to investigate if sample mass significantly affects the theoretical maximum specific (RICE) gravity of HMA samples.

2.0 ANALYSIS OF TEST RESULTS

2.1 EVALUATION CRITERIA

Test results were analyzed using a statistical evaluation system in which the mean (X) and standard deviation (s) was calculated for each test parameter. A rating score was then given to the test result based on the criteria shown in Table 1. A test result with a score of 3 or greater was considered acceptable. A test result with a score of 2 or less was considered unacceptable and a retest was required.

Table 1: Evaluation Criteria

| Test Result | Rating | Interpretation of Results | Acceptance |
|--------------|--------|---------------------------|--------------|
| $X \pm 1.0s$ | 5 | Very Good | Acceptable |
| $X \pm 1.5s$ | 4 | Good | |
| $X \pm 2.0s$ | 3 | Fair | |
| $X \pm 2.5s$ | 2 | Poor | Unacceptable |
| $X \pm 3.0s$ | 1 | Very Poor | |

2.2 INITIAL TEST

A total of 104 laboratories participated in the initial test. Out of the 104 laboratories that participated, 56 laboratories submitted results based on the 1200 gram sample mass. The remaining 48 laboratories submitted results based on the 2400 gram sample mass. Analyses for outliers in accordance with ASTM E 178 were conducted for both the 1200 and 2400 gram samples. The individual analysis for each data set yielded a single outlier for the 1200 gram data set and none for the 2400 gram data set. This outlier is summarized in Table 2. Table 3 summarized the analysis excluding the outlier.

Table 2: Summary of Initial Test Results

CT – 309 RICE specific gravity

| (1200 gram sample size) | | | | | | | | |
|-------------------------|-------|---------|--------------------|----------------------------------|---|---|---|----|
| Item | # Lab | Average | Standard Deviation | Number of Labs Achieved Score of | | | | |
| | | | | 5 | 4 | 3 | 2 | 1 |
| CT 309 | 56 | 2.559 | 0.022 | 47 | 4 | 4 | 0 | 1 |
| % of Total | | | | 84 | 7 | 7 | 0 | 2* |

*Lab #133 was considered as an outlier and was not included in the final analysis

Table 3: Summary of Initial Test Results

CT 309 – RICE specific gravity

| (2400 gram sample size) | | | | | | | | |
|--------------------------------|-------|---------|--------------------|----------------------------------|----|---|---|---|
| Item | # Lab | Average | Standard Deviation | Number of Labs Achieved Score of | | | | |
| | | | | 5 | 4 | 3 | 2 | 1 |
| CT 309 | 48 | 2.557 | 0.010 | 32 | 10 | 4 | 1 | 1 |
| % of Total | | | | 67 | 21 | 8 | 2 | 2 |

Table 4: Labs and Test Results Considered as Outliers

| Item | Data Set | # of Outliers | Lab ID |
|--------|----------------|---------------|--------|
| CT 309 | 1 (1200 grams) | 1 | 133 |
| | 2 (2400 grams) | --- | --- |

After excluding the outlier from data set 1, the mean and standard deviation for each test parameter were re-calculated to determine the score for the respective test parameter. The analysis results are presented in Table 5. Based on the analysis of the initial results, it was determined that rating scores would be based on the combined (1200 g and 2400 g sample size) test result data, since each individual data set are statistically similar regardless of sample size. Detailed test results are provided in Appendix A.

Table 5: Summary of Initial Test Results*

| (1200 g sample size) | | | | | | | | |
|--|-------|---------|--------------------|----------------------------------|----|---|---|---|
| Item | # Lab | Average | Standard Deviation | Number of Labs Achieved Score of | | | | |
| | | | | 5 | 4 | 3 | 2 | 1 |
| CT 309 | 55 | 2.557 | 0.016 | 43 | 5 | 3 | 3 | 1 |
| % of Total | | | | 78 | 10 | 5 | 5 | 2 |
| (2400 g sample size) | | | | | | | | |
| Item | # Lab | Average | Standard Deviation | Number of Labs Achieved Score of | | | | |
| | | | | 5 | 4 | 3 | 2 | 1 |
| CT 309 | 48 | 2.557 | 0.010 | 32 | 10 | 4 | 1 | 1 |
| % of Total | | | | 67 | 21 | 8 | 2 | 2 |
| (combined results from 1200 g and 2400 g) | | | | | | | | |
| Item | # Lab | Average | Standard Deviation | Number of Labs Achieved Score of | | | | |
| | | | | 5 | 4 | 3 | 2 | 1 |
| CT 309 | 103 | 2.557 | 0.014 | 77 | 15 | 4 | 3 | 4 |
| % of Total | | | | 74 | 15 | 4 | 3 | 4 |

*Outlier excluded

2.3 RETEST

In the initial test, 8 laboratories (including outlier) did not receive an acceptable score. Samples for a retest were sent to these laboratories in January 2011. Retest samples consisted of 2400 g of materials as indicated in the CT 309 test method. These laboratories were: 161, 482, 71, 380, 465, 118, 133, and 310. Their results were included in the analysis of the retest results.

The outlier analysis was performed following ASTM E-178. It was determined that there was a single outlier value on the retest results. The same laboratory (Lab # 133) exhibited unsatisfactory results. It should be noted that a Caltrans IA staff witnessed the retest process for Lab # 133. The retest score for each laboratory was determined by comparing the retest result with the rating range from the initial test. Table 6 presents the mean and standard deviation from the retest. Detailed test results and scores are provided in Appendix B.

Table 6: Summary of Retest Results

| Item | # Lab | Average | Standard Deviation | Number of Labs Achieved Score of | | | | |
|------------------------|-------|---------|--------------------|----------------------------------|------|------|---|------|
| | | | | 5 | 4 | 3 | 2 | 1 |
| Relative Compaction, % | 8 | 2.548 | 0.031 | 5 | 1 | 1 | 0 | 1 |
| % of Total | | | | 62.5 | 12.5 | 12.5 | 0 | 12.5 |

2.4 COMBINED RESULTS

A total of 104 laboratories participated in the reference sample program. Eight laboratories participated in both the initial test and the retest with a single laboratory failing to produce an acceptable result. Table 7 shows combined scores from both the initial test and the retest. The final combined scores are provided in Appendix C.

Table 7: Summary of Combined Test Results

| Test Method | Total # Labs | Number of Labs Achieved Score of | | | | |
|-------------|--------------|----------------------------------|----|---|---|---|
| | | 5 | 4 | 3 | 2 | 1 |
| CT 309 | 104 | 81 | 17 | 5 | 0 | 1 |
| | % of Total | 78 | 16 | 5 | 0 | 1 |

2.5 OBSERVATIONS

There were eight laboratories that failed the initial test. A retest was conducted by these laboratories and each achieved acceptable results except for one. Lab # 133 failed to produce acceptable results for the retest and was disqualified from performing CT 309. The following observations were noted by Caltrans IA during the retest witness for Lab # 133. Possible cause of the initial failure maybe attributed to the following:

- Sample size – based on initial observation, it was thought that sample size might affect the test results, however, the results from both the 1200 and 2400 gram sample are statistically identical. Sample size did not affect the actual results of the test.
- Equipment:
 - Manometer – (monitors vacuum pressure) if not calibrated or not in a proper working order may not hold the specified vacuum pressure which in turn may require constant adjustment and create inconsistent negative pressure
 - Pycnometer – (sample container) if not calibrated or not in a proper working condition will develop air leaks and may not achieve or hold correct vacuum pressure.
 - Bleeder valve – (controls vacuum pressure in pycnometer) if not in proper working order will develop air leaks resulting in inconsistent negative pressure
- Temperature:
 - Water – if below or above the specified limits indicated in the test method may produce inaccurate results (corrections has to be applied in relation to the water temperature reading)
- Not following proper test procedure/s or best practice, i.e., constant vacuum adjustment, poor weighing procedures, poor sample preparation (not removing uncoated aggregates, sample not in a free flowing condition – clumping, etc).
- Arithmetic and round-off errors. It was observed that some laboratories rounded up their values and while other laboratories performed in-correct calculations (these values were corrected prior to the analysis and corrected values were used.)

3.0 SUMMARY

- CT 309 – In the initial round of testing, 104 laboratories participated, 8 of which did not achieve an acceptable score. The 8 laboratories with poor scores were given an additional sample to conduct a retest. After the retest was completed, 7 of 8 laboratories submitted satisfactory results with 1 laboratory failing.

4.0 REFERENCES

ASTM, “Standard Practice for Dealing with Outlying Observations,” Designation E 178 – 80.

Caltrans, “Independent Assurance Manual,” Sacramento, July 2005.

APPENDIX – A

Test Results from Initial Test

| Lab ID# | Results, G _{mm} | Score |
|---------|--------------------------|-------|
| 3 | 2.543 | 4 |
| 8 | 2.543 | 4 |
| 9 | 2.539 | 4 |
| 10 | 2.547 | 5 |
| 11 | 2.552 | 5 |
| 12 | 2.556 | 5 |
| 16 | 2.572 | 4 |
| 20 | 2.533 | 3 |
| 24 | 2.545 | 5 |
| 29 | 2.553 | 5 |
| 31 | 2.562 | 5 |
| 32 | 2.550 | 5 |
| 34 | 2.575 | 4 |
| 42 | 2.568 | 5 |
| 43 | 2.557 | 5 |
| 47 | 2.557 | 5 |
| 48 | 2.569 | 5 |
| 49 | 2.575 | 4 |
| 53 | 2.560 | 5 |
| 57 | 2.555 | 5 |
| 59 | 2.580 | 3 |
| 62 | 2.558 | 5 |
| 63 | 2.559 | 5 |
| 67 | 2.557 | 5 |
| 68 | 2.556 | 5 |
| 69 | 2.552 | 5 |
| 71 | 2.517 | 1 |
| 73 | 2.564 | 5 |
| 75 | 2.568 | 5 |
| 79 | 2.573 | 4 |
| 84 | 2.566 | 5 |
| 88 | 2.554 | 5 |
| 90 | 2.547 | 5 |
| 94 | 2.566 | 5 |
| 110 | 2.531 | 3 |
| 114 | 2.573 | 4 |
| 118 | 2.516 | 1 |
| 119 | 2.558 | 5 |
| 120 | 2.560 | 5 |
| 122 | 2.548 | 5 |
| 123 | 2.559 | 5 |

| Lab ID# | Results, G _{mm} | Score |
|---------|--------------------------|-------|
| 131 | 2.546 | 5 |
| 133 | 2.668 | 0 |
| 135 | 2.554 | 5 |
| 140 | 2.562 | 5 |
| 144 | 2.581 | 3 |
| 147 | 2.554 | 5 |
| 154 | 2.545 | 5 |
| 155 | 2.546 | 5 |
| 156 | 2.543 | 4 |
| 158 | 2.560 | 5 |
| 160 | 2.554 | 5 |
| 161 | 2.598 | 1 |
| 163 | 2.564 | 5 |
| 164 | 2.563 | 5 |
| 173 | 2.553 | 5 |
| 176 | 2.568 | 5 |
| 177 | 2.569 | 5 |
| 183 | 2.566 | 5 |
| 193 | 2.564 | 5 |
| 200 | 2.568 | 5 |
| 206 | 2.578 | 4 |
| 210 | 2.558 | 5 |
| 244 | 2.542 | 4 |
| 246 | 2.566 | 5 |
| 248 | 2.571 | 5 |
| 249 | 2.557 | 5 |
| 255 | 2.571 | 4 |
| 256 | 2.575 | 4 |
| 257 | 2.556 | 5 |
| 263 | 2.562 | 5 |
| 265 | 2.551 | 5 |
| 281 | 2.561 | 5 |
| 293 | 2.567 | 5 |
| 294 | 2.577 | 4 |
| 295 | 2.553 | 5 |
| 300 | 2.559 | 5 |
| 309 | 2.558 | 5 |
| 310 | 2.517 | 1 |
| 316 | 2.552 | 5 |
| | | |
| | | |

| Lab ID# | Results, G _{mm} | Score |
|---------|--------------------------|-------|
| 323 | 2.563 | 5 |
| 354 | 2.561 | 5 |
| 361 | 2.543 | 5 |
| 364 | 2.555 | 5 |
| 370 | 2.556 | 5 |
| 377 | 2.550 | 5 |
| 380 | 2.586 | 2 |
| 393 | 2.551 | 5 |
| 395 | 2.546 | 5 |
| 398 | 2.537 | 4 |
| 457 | 2.546 | 5 |
| 464 | 2.546 | 5 |
| 465 | 2.529 | 2 |
| 467 | 2.564 | 5 |
| 479 | 2.555 | 5 |
| 482 | 2.585 | 2 |
| 483 | 2.555 | 5 |
| 560 | 2.546 | 5 |
| 562 | 2.558 | 5 |
| 565 | 2.551 | 5 |
| 571 | 2.558 | 5 |
| 581 | 2.549 | 5 |
| 582 | 2.553 | 5 |
| 605 | 2.564 | 5 |

Legend:

| | |
|-----|--------------------|
| 1,2 | Unacceptable Score |
| 0 | Outlier |

APPENDIX - B

Test Results from Retest

| Lab ID# | Results, G _{mm} | Score |
|---------|--------------------------|-------|
| 71 | 2.582 | 3 |
| 118 | 2.546 | 5 |
| 133 | 2.479 | 1 |
| 161 | 2.558 | 5 |
| 310 | 2.540 | 4 |
| 380 | 2.571 | 5 |
| 465 | 2.551 | 5 |
| 482 | 2.557 | 5 |

APPENDIX - C

Combined Final Scores from both Initial Test and Retest

| Lab No. | Results,G _{mm} | Score |
|---------|-------------------------|-------|
| 3 | 2.543 | 4 |
| 8 | 2.543 | 4 |
| 9 | 2.539 | 4 |
| 10 | 2.547 | 5 |
| 11 | 2.552 | 5 |
| 12 | 2.556 | 5 |
| 16 | 2.572 | 4 |
| 20 | 2.533 | 3 |
| 24 | 2.545 | 5 |
| 29 | 2.553 | 5 |
| 31 | 2.562 | 5 |
| 32 | 2.550 | 5 |
| 34 | 2.575 | 4 |
| 42 | 2.568 | 5 |
| 43 | 2.557 | 5 |
| 47 | 2.557 | 5 |
| 48 | 2.569 | 5 |
| 49 | 2.575 | 4 |
| 53 | 2.560 | 5 |
| 57 | 2.555 | 5 |
| 59 | 2.580 | 3 |
| 62 | 2.558 | 5 |
| 63 | 2.559 | 5 |
| 67 | 2.557 | 5 |
| 68 | 2.556 | 5 |
| 69 | 2.552 | 5 |
| 71 | 2.517 | 3 |
| 73 | 2.564 | 5 |
| 75 | 2.568 | 5 |
| 79 | 2.573 | 4 |
| 84 | 2.566 | 5 |
| 88 | 2.554 | 5 |
| 90 | 2.547 | 5 |
| 94 | 2.566 | 5 |
| 110 | 2.531 | 3 |
| 114 | 2.573 | 4 |
| 118 | 2.516 | 5 |
| 119 | 2.558 | 5 |
| 120 | 2.560 | 5 |
| 122 | 2.548 | 5 |
| 123 | 2.559 | 5 |

| Lab No. | Results,G _{mm} | Score |
|---------|-------------------------|-------|
| 131 | 2.546 | 5 |
| 133 | 2.668 | 1 |
| 135 | 2.554 | 5 |
| 140 | 2.562 | 5 |
| 144 | 2.581 | 3 |
| 147 | 2.554 | 5 |
| 154 | 2.545 | 5 |
| 155 | 2.546 | 5 |
| 156 | 2.543 | 4 |
| 158 | 2.560 | 5 |
| 160 | 2.554 | 5 |
| 161 | 2.598 | 5 |
| 163 | 2.564 | 5 |
| 164 | 2.563 | 5 |
| 173 | 2.553 | 5 |
| 176 | 2.568 | 5 |
| 177 | 2.569 | 5 |
| 183 | 2.566 | 5 |
| 193 | 2.564 | 5 |
| 200 | 2.568 | 5 |
| 206 | 2.578 | 4 |
| 210 | 2.558 | 5 |
| 244 | 2.542 | 4 |
| 246 | 2.566 | 5 |
| 248 | 2.571 | 5 |
| 249 | 2.557 | 5 |
| 255 | 2.571 | 4 |
| 256 | 2.575 | 4 |
| 257 | 2.556 | 5 |
| 263 | 2.562 | 5 |
| 265 | 2.551 | 5 |
| 281 | 2.561 | 5 |
| 293 | 2.567 | 5 |
| 294 | 2.577 | 4 |
| 295 | 2.553 | 5 |
| 300 | 2.559 | 5 |
| 309 | 2.558 | 5 |
| 310 | 2.517 | 4 |
| 316 | 2.552 | 5 |
| 323 | 2.563 | 5 |
| 354 | 2.561 | 5 |

| Lab ID# | Results, G _{mm} | Score |
|---------|--------------------------|-------|
| 361 | 2.543 | 5 |
| 364 | 2.555 | 5 |
| 370 | 2.556 | 5 |
| 377 | 2.550 | 5 |
| 380 | 2.586 | 5 |
| 393 | 2.551 | 5 |
| 395 | 2.546 | 5 |
| 398 | 2.537 | 4 |
| 457 | 2.546 | 5 |
| 464 | 2.546 | 5 |
| 465 | 2.529 | 5 |
| 467 | 2.564 | 5 |
| 479 | 2.555 | 5 |
| 482 | 2.585 | 5 |
| 483 | 2.555 | 5 |
| 560 | 2.546 | 5 |
| 562 | 2.558 | 5 |
| 565 | 2.551 | 5 |
| 571 | 2.558 | 5 |
| 581 | 2.549 | 5 |
| 582 | 2.553 | 5 |
| 605 | 2.564 | 5 |