Recommend Approval: Team Leader Date 1/19/12 / Division Chief Date 1/19/12	Maryland Department of Transportation State Highway Administration Office of Materials Technology MARYLAND STANDARD METHOD OF TESTS	
Approved: Director Date 03/09/12	**PROCEDURE FOR DETERMINING STATISTICAL OUTLIERS**	**MSMT 734**

SCOPE:

This procedure is used for determining outlying observations in sample test results and how to evaluate the statistical significance of them. The Administration or the Contractor may raise the question of whether an observation is an outlier on a project.

An outlying observation, or "outlier," is a test result that appears to differ significantly from the other sample test result values in the same population from which it was taken. When dealing with possible outliers for a project, two alternatives need to be considered.

1. The outlying observation may be an extreme value of the population caused by the random variability inherent in the data. When this is the case, the observation should be retained and used in the same manner as the other observations.

2. The outlying observation may be the result of gross deviation from the prescribed sampling and testing procedures or an error in calculating or recording the numerical value. When this is the case, the observation should be discarded.

REFERENCE DOCUMENT:

E 178 - Practice for Dealing with Outlying Observations.

PROCEDURE:

With this procedure one of two alternatives will be determined:

(a) The observation is not an outlier and should not be discarded, or

(b) The observation is an outlier and should be discarded.

1. Determine whether a testing related physical reason exists for the outlying test value. When a physical reason exists, the outlying test value is excluded from pay factor calculations. Normally, only the individual test value is excluded; the test results for the entire sample are only excluded when the physical reason for the outlying test value applies to the entire sample.

Physical reasons for excluding a test value are:

(a) Damaged test sample before testing.

(b) Gross deviation from prescribed test procedure. When it is determined that a gross deviation from the prescribed test procedure has occurred, the resulting observation should be discarded, whether or not it agrees with the rest of the data.

(c) Test equipment malfunction.

(d) Computational error was made. Errors discovered in computation may be corrected and the corrected value used as the test result.

(e) The test result is outside the range of possible results.

The following are examples of reasons that are ***not*** sufficient for excluding a test value:

(a) The sample was taken from a segregated area of the mat.

(b) The acceptance test results do not agree with the quality control results.

(c) The core had paint on it.

(d) The test result is larger/smaller than all the rest.

(e) Plant malfunctions - are an assignable cause for the test result being different, because the material is different. It is not a reason for discarding a sample or a test result.

2. When a physical reason cannot be determined for an apparent outlying value the following calculation procedure should be used to determine whether the test result meets statistical criteria as an outlying value.

CALCULATION PROCEDURE:

This procedure is based on a two-tailed t-test with a level of significance of 1 percent, adopted from E 178. The use of a two-tailed test means that the outlier may be either on the high or the low side of the average. The 1 percent level of significance means that if it is decided that the value is an outlier, there is only a 1 percent chance that it is not.

1. Determine whether there is an assignable cause for the apparent outlier. An assignable cause means that a reason exists for the material being different, for example:

    a) The sample was taken at the end of a truckload.

    b) There is visible segregation at that location in the mat.

    c) The paver wings were dumped at the sample location.

    d) The plant was having problems.

    e) The loader operator put the aggregate in the wrong bins.

    If there is an assignable cause, the sample should not be excluded and the analysis should not proceed.

2. Identify the sample set to be used in the statistical analysis. The statistical procedure being used bases its criteria on the assumption that the samples used in the analysis must be part of

the same population. Lots produced under different mix designs or when there have been significant changes to the mix are to be considered in different populations and should not be combined for the purpose of determination of statistical outliers. A target value change does not always indicate a significant change to the mix.

Compaction. For determination of statistical outliers in compaction lots, use all of the core results from the lot with the suspected outlier. Thus, n is normally 5 for the determination of compaction outliers.

Mix Properties. For determination of statistical outliers in mix properties, use all of the test results from the lot with the suspected outlier and the two previous lots. If either of these lots has been refereed, the referee data should be used.

If there are not two previous lots with the same mix design or it is the first or second lot in the project, the following lots should be used. For example, if the lot containing the suspected outlier is the first lot of a new mix design, use the two following lots in the analysis. If the lot containing the suspected outlier is the second lot of a new mix design, use the previous lot and the following lot in the analysis. If there are not three consecutive lots with the same mix design, the analysis is conducted using only the samples in one or two lots.

3. Calculate the sample average ($\bar{x}$) and standard deviation (s) of ALL of the samples in the sample set using the equations below. The suspected outlier is NOT excluded from these calculations.

$$\bar{x} = \frac{\sum x}{n} \qquad (1)$$

$$s = \left\{ \frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n-1} \right\}^{1/2} \qquad (2)$$

where:

x = sample test value

n = number of samples

Note: Round $\bar{x}$ to one decimal place more than the data used to calculate it and s to two more decimal places more than the data used to calculate it.

4. Determine the critical value for T from Table 1 using the total number of samples (n) in the sample set.

Table 1: Critical Values for T at 1% significance level (two-tailed test)

n	T	n	T
3	1.155	8	2.274
4	1.496	9	2.387
5	1.764	10	2.482
6	1.973	11	2.564
7	2.139	12	2.636

5. Determine the lower and upper outlier limits using the equations below. Round LO and UO to the same number of decimal places as the test values.

$$LO = \bar{x} - (T \times s) \qquad (3)$$

$$UO = \bar{x} + (T \times s) \qquad (4)$$

6. Provided there is no assignable cause for the occurrence of the test result in question, discard test data which falls outside of the lower and upper outlier limits calculated with equations 3 and 4. The entire sample is not discarded, only the outlying test result.

EXAMPLE CALCULATIONS:

EXAMPLE 1: Suspected Compaction Outlier

The following 10 core densities were obtained, is core number 4 an outlier? No physical reason or assignable cause could be identified for the low density.

Core	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Density (pcf)	141.5	141.8	142.3	138.3	141.6	142.0	141.6	141.7	141.0	141.2

where:

$n = 10$

$\bar{x} = 141.30$

$s = 1.117$

From Table 1, $T = 2.482$

$$LO = \bar{x} - (T \times s) = 141.30 - (2.482 \times 1.117) = 138.5$$

$$UO = \bar{x} + (T \times s) = 141.30 + (2.482 \times 1.117) = 144.1$$

Because core density #4 is below the lower outlier level (LO), core number 4 should be discarded and pay factor determinations should be made using the remaining 9 cores. Note that the calculated values for LO and UO are rounded to the same number of decimal places as the test data, in this case one decimal place.

EXAMPLE 2: Suspected Air Voids Outlier.

The following test results were obtained for three consecutive lots on a project. Are the air voids for Lot 3, sample 1 an outlier? No physical reason or assignable cause could be identified for the high air voids.

Lot 1 Results:

SAMPLE NO.	Bulk Density (pcf)	VOIDS (%)	Gmm*
1	151.8	4.2	2.540
2	152.1	5.8	2.586
3	152.1	4.1	2.540
4	153.2	4.7	2.577

Lot 2 Results:

SAMPLE NO.	Bulk Density (pcf)	VOIDS (%)	Gmm*
1	152.4	4.8	2.564
2	152.7	4.3	2.558
3	152.6	4.3	2.556
4	152.7	3.5	2.470

Lot 3 Results:

SAMPLE NO.	Bulk Density (pcf)	VOIDS (%)	Gmm*
1	149.5	7.5	2.590
2	151.7	5.0	2.559
3	151.9	4.5	2.550
4	151.5	4.9	2.553

* Maximum specific gravity

where:

$n = 12$

$\bar{x} = 4.80$

$s = 1.022$

From Table 1, $T = 2.636$

$$LO = \bar{x} - (T \times s) = 4.80 - (2.636 \times 1.022) = 2.1$$

$$UO = \bar{x} + (T \times s) = 4.80 + (2.636 \times 1.022) = 7.5$$

The voids for Lot 3, Sample 1 are equal to the UO, thus this value is not an outlier and shall be included in the pay factor determination. The value in question should be outside the lower and upper outlier limits to be considered an outlier. Note that the calculated values for LO and UO are rounded to the same number of decimal places as the test data, in this case one decimal place.

EXAMPLE 3: Suspected outlier in maximum specific gravity (Gmm), when it is used to calculate the compaction.

The data from example three is from a project where the compaction is calculated as a percentage of the Gmm. Is the Gmm for Lot 2, sample 4 an outlier? No physical reason or assignable cause could be identified for the low Gmm.

where:

$n = 12$

$\bar{x} = 2.554$

s = .031

From Table 1, $T = 2.636$

$$LO = \bar{x} - (T \times s) = 2.554 - (2.636 \times 0.031) = 2.472$$

$$UO = \bar{x} + (T \times s) = 2.554 + (2.636 \times 0.031) = 2.636$$

The Gmm for Lot 2, Sample 4 is below the lower outlier level (LO), thus the Gmm for this sample shall be discarded and the Gmm for Lot 2 shall be determined using the average of the remaining 3 Gmm. Note that the calculated values for LO and UO are rounded to the same number of decimal places as the test data, in this case three decimal places.