IDAHO TRANSPORTATION DEPARTMENT

RESEARCH REPORT

Statistical Analysis of 2018 HMA Production and Construction Data to Improve Quality Assurance and Acceptance Practices in Idaho

RP 275

By

Dr. Mojtaba Sadegh (PI), Dr. Deb Mishra (Co-PI),

Mostofa Najmus Sakib, and Arash Modaresi Rad

Boise State University and Oklahoma State University

In Cooperation with The

Idaho Transportation Department

ITD Research Program, Contracting Services

Highways Construction and Operations

Final July 2021

Disclaimer

This document is disseminated under the sponsorship of the Idaho Transportation Department and the United States Department of Transportation in the interest of information exchange. The state of Idaho and the United States Government assume no liability of its contents or use thereof.

Except for the Foreword, the contents of this report reflect the view of the authors, who are responsible for the facts and accuracy of the data presented herein. The contents do not necessarily reflect the official policies of the Idaho Transportation Department or the United States Department of Transportation.

The state of Idaho and the United States Government do not endorse products or manufacturers. Trademarks or manufacturers’ names appear herein only because they are considered essential to the object of this document.

This report does not constitute a standard, specification, or regulation.

Technical Report Documentation Page

1. Report No. 275
FHWA-ID-21-275

2. Government Accession No.

3. Recipient's Catalog No.

4. Title and Subtitle
Statistical Analysis of 2018 HMA Production and Construction Data to Improve Quality Assurance and Acceptance Practices in Idaho
Statistical Analysis of HMA Production Data

5. Publish Date
July 2021

6. Performing Organization Code

7. Author(s)
Mojtaba Sadegh, https://orcid.org/0000-0003-1775-5445
Deb Mishra, https://orcid.org/0000-0003-2354-1312
Mostofa Najmus Sakib,
Arash Modaresi Rad, https://orcid.org/0000-0002-6030-7923

8. Performing Organization Report No.

9. Performing Organization Name and Address
Boise State University
1910 W University Dr, Boise, ID, US, 83725

Oklahoma State University
248 Engineering North, Stillwater, OK, US, 74078

10. Work Unit No. (TRAIS)

11. Contract or Grant No.
BSU-18-03

12. Sponsoring Agency Name and Address
Idaho Transportation Department (SPR)
Highways Construction and Operations, Contracting Services, Research Program
PO Box 7129
Boise, ID 83707-7129

13. Type of Report and Period Covered
[Final or Interim] Report
[07/01/2018 – 06/30/2021]

14. Sponsoring Agency Code
RP-275

15. Supplementary Notes
Research project performed in cooperation with the Idaho Transportation Department and the Federal Highway Administration.

16. Abstract

Quality assurance procedure is adopted to ensure hot mix asphalts meet design performance. Robustness of QA reports, however, were challenged during 2017 and 2018 data reviews conducted on projects constructed over several previous years by the Idaho Transportation Department (ITD) and the Federal Highway Administration (FHWA). To further investigate the sources of discrepancies and unnatural trends between material testing reports produced by the prime contractors, third-party contractors, as well as ITD in HMA projects, this research study analyzed an audit trail dataset of submitted material testing results. The audit trail dataset showed multiple spreadsheet cell corrections before the final submitted test results. This research project analyzed this audit trail data, and categorized instances of multiple data corrections for certain parameters as Plausible Correction (P.C.) and Unexplained Corrections (U.C.). P.C. cases are those that can be attributed to typographical-type mistakes, and U.C. cases are those that cannot be simply explained as such. Results show that all 15 projects from year 2018 for which audit data was available contained data corrections that accumulated to several hundred instances of U.C., acknowledging that both the prime contractors and ITD reported data may be produced by third-party contractors hired by either party to act on their behalf. Further, the project analyzed the financial repercussions of U.C. cases from the prime contractors, third-party contractors, and ITD, which proved that an absolute majority of projects (i.e., 12 of the 15 projects that had enough audit trail data for financial analysis) indicated overpayments to the prime contractors due to U.C. cases. This research study highlights the importance of a rigorous oversight protocol for acceptance testing to be implemented by state and local highway agencies. Since 2018, ITD has ended its use of QC results for acceptance and payment, and has implemented several new policies in materials testing, specifications, and training to help eliminate instances of P.C. and U.C. and to ensure higher accuracy in reporting results from all parties conducting material acceptance testing.

17. Key Words Hot Mix Asphalt, Quality Control, Quality Assurance, Monetary Analysis, Data Reporting, Sensitivity Analysis	18. Distribution Statement Copies available from the ITD Research Program

19. Security Classification (of this report) Unclassified	20. Security Classification (of this page) Unclassified	21. No. of Pages 198	22. Price None

Acknowledgments

This study was sponsored by the Idaho Transportation Department (ITD) and FHWA. The authors would like to acknowledge the financial and technical support of these organizations. Several individuals from ITD contributed significantly towards ensuring successful completion of this research project; some of these individuals served officially as on the project Technical Advisory Committee (TAC), whereas others helped with the research work even though they were not officially a part of the study. Names of the TAC committee members are listed in the next section. In this section, the authors have attempted to list those individuals who volunteered their time to ensure success of this project; any omission is purely accidental. The authors greatly acknowledge the contributions of Jason Henscheid and Heather Miley of ITD, and Aisan Ranjbar Moshfeghi of Oklahoma State University.

Technical Advisory Committee

Each research project is overseen by a Technical Advisory Committee (TAC), which is led by an ITD project sponsor and project manager. The TAC is responsible for monitoring project progress, reviewing deliverables, ensuring that study objectives are met, and facilitating implementation of research recommendations, as appropriate. ITD's Research Program Manager appreciates the work of the following TAC members in guiding this research study.

Project Sponsor - John P. Bilderback, P.E.

Project Manager - Mike Copeland

TAC Members

Mike Santi, P.E. (retired)
Dan Harelson, P.E. (retired)
John Arambarri, P.E.
Eric Kokernak, P.E.
Steve Taylor

FHWA-Idaho Advisor - Kyle P. Holman, P.E.

Table of Contents

List of Tables

List of Figures

List of Abbreviations and Acronyms

AASHTO	American Association of State Highway and Transportation Officials
ACFE	Association of Certified Fraud Examiners
ANSI	American National Standards Institute
ASCE	American Society of Civil Engineers
23 CFR 637	Code of Federal Regulations, Title 23, Part 637
CPF	Composite Pay Factor
DOT	Department of Transportation
FHWA	Federal Highway Administration
G_{mb}	Bulk specific gravity
G_{mm}	Theoretical maximum specific gravity
HMA	Hot Mix Asphalt
IA	Independent Assurance
ISO	International Organization for Standardization
ITD	Idaho Transportation Department
JMF	Job Mix Formula
NaN	Not-A-Number
NAPA	National Asphalt Pavement Association
MEPDG	Mechanistic–Empirical Pavement Design Guide
OECD	Organization for Economic Co-operation and Development
Pa	Air voids
P_b	Asphalt binder content, percent by total mass of mixture
P.C.	Plausible Correction
PF	Pay Factor
PWL	Percent Within Limits
QC	Quality Control
QA	Quality Assurance

SSD	Saturated Surface Dry
U.C.	Unexplained Correction
US	United States
UW	Under-Water
VBA	Visual Basic for Application
VFA	Voids Filled with Asphalt
VMA	Voids in the Mineral Aggregate
WMA	Warm-Mix Asphalt

Foreword by the Idaho Transportation Department

This research report analyzed the 2018 Hot Mix Asphalt (HMA) production and construction data, including the audit trail data. Although the Idaho Transportation Department (ITD) did not specifically sanction the collection of audit data by ITD staff in the spreadsheet macros, as explained herein, this study does reference this data in work staff and Boise State University (BSU) researchers have completed herein. Since 2018, ITD has implemented many changes to its asphalt mix materials acceptance program due to this research and the FHWA Forensic Pavement Assessment (see Appendix A for the “FHWA Forensic Pavement Assessment for the Idaho Transportation Department (ITD), dated March 8, 2018”).

One of the most significant changes has been the elimination of the prime contractor’s or their third-party contractor’s Quality Control (QC) results for acceptance by ITD with a state acceptance program; thus, eliminating the statistical verification of contractor-reported test results. The following sections summarize some of the most significant changes.

Changes made from 2018 to 2020:

Quality Assurance

- State Acceptance only (either with Department forces or Consultants hired by the Department) for the 2020 construction season.
    - This was done in concurrence with the AGC.
- Developed procedures for verifying and calibrating asphalt plants.
    - To ensure that the mix that is being produced is the mix that is specified and designed (best practice).
- Pilot quality control plans for all statistically based pay items (Burati et al., 2003).
    - Better defining roles of parties in the quality assurance process (best practice).
      Future: Implement quality control plans for all items of work.

- Official meeting minutes prepared by the Department of the preconstruction and pre-paving meeting(s).
    - Standardization of processes. This was not being done in the past consistently. This was leading to conflict during construction.

Acceptance and Payment

- Added G_{se}, G_{mm}, and dust proportion for acceptance, and dust proportion for payment.
    - A change in Gse or Gmm from design are caused by one of two things. Either a change in mix or an issue with testing (Hughes et al., 2020).
    - These were items identified in the FHWA Forensic Report.

- Increase total possible incentive from 5% to 10% (plus smoothness incentive which was unchanged).
- Increase maximum incentive on density to 5% as a stand-alone incentive and implement the tier density pay equation used during the QIP Pilot in 2017 to incentivize consistent increased density.
    - Research has shown that a 1% increase of in-place density will add 10-15% in life to the road. Incentivizing this practice.
- Modify composite pay factor to include VMA, air voids, and dust to binder ratio for up to 5% incentive.
    - The addition of dust to binder ratio as a pay factor was based on FHWA's recommendation from the Forensic Report.

Construction Requirements

- Minimum lift thickness will not be less than 3.5 times the Nominal Maximum Aggregate Size (NMAS) (Blankenship et al., 2021)..
    - Recommendation from FHWA Forensic Report
- Formalized a nuclear density gauge procedure.
    - Standardization of a process. Correlation of nuclear density gauges were being done inconsistently across the ITD districts.
- Removed some of the more prescriptive requirements (e.g., the type of paver used or roller)
    - The specification was developed to be more focused on the end result and the contractor's plan to ensure quality.

Specific Gravity of Aggregates and RAP

- Department to witness sampling of RAP and aggregates for determining specific gravities.
    - FHWA requirement. Maintaining clear chain of custody on material samples that are used to establish acceptance.
- Determine specific gravities for each individual stockpile. Determine Gsb of the mix based on the aggregate blend in the design.
    - Previous practice was to receive a pre-blended sample from the mix designer and establish Gsb. That Gsb would be used during design and production regardless of blend changes.
- Monitor Gsb through production.
    - FHWA recommendation in Forensic Report. Using an incorrect Gsb during production can have a serious impact on quality (Hughes et al., 2020; NCAT, 2016).

Mix Design

- Formalized and streamlined the mix design review process which is now centralized and based on risk, paper review for mix designs.

  - Previously mix designs were verified by a Central Materials Lab confirmation. There was very little value added using the previous process regarding project quality.
- Hamburg and Ideal CT were added and removed the old APA and Immersion Compression (IC) tests.
  - FHWA Forensic Report identified the need for a cracking test and Ideal CT was recommended and is currently being used by 10 agencies nationally (Sias et al., 2020).
  - Hamburg is currently being used by 20 agencies nationally including surrounding states. It identifies both stripping and rutting in a single test and replaces APA and IC.
- Increased the dust to binder from 1.2 to 1.4 at mix design.
  - This was a concession to the AGC. National best practice is a maximum design dust to binder ratio of 1.2.
  - The concession was given with the intent of moving toward 1.2 in the future (Sreedhar 2019).

Test Strip

- Changed test strip process from multiple oil contents to a single oil content using the mix design.
  - Test strips were previously constructed with multiple sections with multiple oil content targets. Each oil target was essentially treated as its own test strip. This really increased workloads in both the District and Central Lab. The goal is for the Contractor to produce the mix as was designed.
- Determine the NCAT correction factor using plant produced material. (Idaho IR 157)
  - Historically there have been big changes in material properties of mix produced in the lab and mix produced through the plant during production paving. This was done to help better measure the properties of the plant produced mix and take into account the aggregate breakdown and dust generated during the production process.
- Established tolerances on G_{se} and G_{mm} from design.
  - This is to help ensure that the mix design is representative of the mix produced.
- All test strips evaluated and tested by the Central Materials Lab.
  - With the addition of performance testing and chemical extraction to the test strip testing process the Central Lab is currently the only Lab in the Department capable of performing all the required tests.
  - Having all test strips be tested in the Central Lab adds to the consistency and confidence in the testing during both test strip and production.
- Modified the test strip acceptance from average passing to Percent Within Limits (PWL).
  - Average passing method does not address variability in the production process. High variability during production will result in lower quality material.
- Implemented performance tests for the test strip on a pass or fail basis.
  - Previously the Department performed IC and APA testing during design or test strip to address stripping and rutting of the mix.
  - The addition of a cracking test (Ideal-CT) complements the rutting test (Hamburg).

  - Moving to Hamburg from APA and IC eliminated a test with questionable precision with a widely used method.
- Required the removal of all failing on-site test strips.
  - Previously failing test strips were sometimes allowed to be left in place at a reduced cost. Failing material increases maintenance costs and has lasting impact to the traveling public.
- After 3 failing test strips a new mix design is required.
  - Previously there have been up to 6 or 7 failing test strips on a project. The increased cost and time that it takes to develop a mix design is intended as an incentive for the Contractor to perform the proper quality control to ensure that they are ready to begin paving.

Production

- Removed daily control chart submittals by the contractor.
  - The QASP Web Portal currently being developed will in part perform this function.
    - This feature is a part of the 2nd phase of the QASP Web Portal development.
- Raised the upper density limit to 100.
  - To increase density the upper limit was raised to 100 as a statistical limit. (evaluated using percent within limits)
- Changed from allowing “slight adjustments” to a job mix formula to a well-defined adjustment table during production
  - Better defining and standardizing our specifications with national practice.
- Formalized the process for a Contractor to make adjustments to a job mix formula.
  - Past practice was for a Contractor to make adjustments without notifying the Department. Some adjustments appeared to be for cost savings and would not improve quality.
  - Goal is to let the Department know what adjustments are being made and why prior to making the adjustments.

New or Modified Idaho Reference Methods

- Acceptance Test Strip for Asphalt Mixtures (Idaho IR 125)
  - Modify procedure to increase clarity and standardization of the process.
  - Allow the Contractor more control of the amount of material to produce during the test strip. The idea is to allow the contractor to control their risk.
- Stratified Random Sampling (Idaho IR 148)
  - New: Defined procedure for the Department to perform stratified random sampling. Previously “stratified random sampling” was not clearly defined.
- Superpave Mix Design (Idaho IR 150)
  - Created and clarified a procedure that was previously in the specification. Modified to account for the shift towards paper review of mix designs and verifying during test strip.

- Superpave Mix Design Evaluation (Idaho IR 151)
  - New: Created a procedure for evaluating a mix design for a specific project using a risk based evaluation process.
  - Previously mix designs were verified and evaluated through testing in the Central Lab of lab produced (by mix designer) material. This new process is a paper review.
- HMA Quality Control Plan Development and Implementation (Idaho IR 152)
  - New: Created in conjunction with other changes within the specifications to better define the Contractor's and Department's individual roles and responsibilities within the quality assurance program.
- Split Sample Comparison (Idaho IR 153)
  - New: Previously this reference procedure was in Section 106 of the QASP. The intent of creating a stand-alone procedure is to give the Department a tool for evaluating all types of testing.
  - The future goal is to incorporate this into a proficiency program currently being developed.
- Nuclear Density Gauge Correlation (Idaho IR 154)
  - New: This was previously in IR-125. Gauge correlations occur at times other than test strip so it was pulled out into its own procedure and modified for clarity.
- Procedures for Checking Asphalt Drum Mix Plant Calibrations (Idaho IR 155)
  - New: The Department identified a risk that current testing and inspection is not able to address. This was developed to address the risk.
  - The procedure was developed based on similar procedures in other states, current state of practice, and discussions with industry.
- Method for Determining Rolling Gmm (Idaho IR 156)
  - New: this reference procedure was in the 405 specification. It was pulled out and put into a standalone method for clarity.
- NCAT Correction Factor (Idaho IR 157)
  - New: The Department identified a risk that current test procedures have been unable to address. This method was developed to address the changes in mix characteristics from design into production.
  - More of an end-result measurement. Measuring the asphalt content of the product as produced through the plant rather than a sample built in laboratory.
  - Similar methods are used in other states for determining asphalt content of mixes.
  - With the use of RAP, determining actual asphalt content using the previous method was inaccurate.
- Quality Control Plan Development and Implementation (Idaho IR 158)
  - New: Created in conjunction with other changes within the specifications to better define the Contractor's and Department's individual roles and responsibilities to ensure quality.

- Quality Control Plan Review Process (Idaho IR 159)
    - New: Created in conjunction with other changes within the specifications to better define the Contractor's and Department's individual roles and responsibilities to ensure quality.
- Evaluation and Approval of HMA Plants (Idaho IR 160)
    - New: Created in conjunction with other changes within the specifications to better define the Contractor's and Department's individual roles and responsibilities to ensure quality. It modifies AASHTO M 156 to be more in line with Idaho's state of practice.

Changes Made July 2020:

Mix Design

- Increased dust proportion range.
- Modified specification limit for Hamburg for SP-2 mixes.
- Ideal-CT for information only. Does not need to pass.
- Determining Asphalt Analyzer Offset
    - Used to determine if the Asphalt Analyzer is used during test strip to determine ignition furnace correction factor (Idaho IR 157) or if the previous method (Annex for AASHTO T 308) is used.
    - Pay for the offset of additional binder used during production if the Asphalt Analyzer is used.
- Modified the mixing plant calibration requirements.
    - Allow calibration using manufacture's recommendations and/or National Asphalt Pavement Association manuals and documented best practices.

Test Strip

- Limit size of test strips (including offsite test strips).
    - Required to be between 200 to 750 tons.
- Modified acceptance requirements.
    - Eliminated Gse and Gmm as acceptance criteria.
    - Changed performance tests (Hamburg, Ideal-CT) from pass/fail to information only.
- The Contractor will establish a Contractor's Job Mix Formula (C-JMF) after a passing test strip.
    - The contractor can make adjustments prior to notifying the Department within the adjustment table.
    - Adjustments made outside the limits of the adjustment table have to be approved by the Department prior to implementing. Previously, the adjustment table was the absolute limits for making adjustments.

Production Paving

- Acceptance for the first lot of production paving will be based on the test strip acceptance requirements except gradation.
- Modified acceptance requirements.
    - Lower limit on Air Voids reduced 17% for SP-5 mixes and 33% for SP-3 mixes.
    - Eliminated Gse and Gmm as acceptance criteria. Used for information only.
        - Requires a review of plant settings and test results if PWL < 40.

Acceptance and Payment

- Modified calculation of payment for lower quality material.
    - Less pay impact if a single quality characteristic is below 60 PWL than previous versions of specification.

Where Do the Specifications Need to Go Moving Forward?

Immediate Actions

- Continue moving towards performance-based mix design.
    - Require a passing cracking test in design.
- Allow mix designers to determine aggregate specific gravities for design.
    - Verify during test strip and monitor in production.
    - Put Gse and Gmm tolerances in place to address this risk.
- Implement performance testing and volumetrics for test strip acceptance.
    - This is national best practice.
- Reestablish controls that minimize the risks of changes in mix from design into production.
    - Put Gse and Gmm tolerances in place to address this risk.
    - Modify allowed mix blend adjustments to be more in line with national best practice.
- Web Portal (phase 2)
    - This will allow for program-wide data analysis and further refinement of specifications and business practices.
- The ITD Chief Operations Officer initiated two specific advisory groups beginning in 2020 with membership from ITD, FHWA, national experts, prime contractors, third-party material testing contractors, and a local highway partner. The advisory groups are directed by the Peer Advisory Review Group (PRAG) who are working on large initiatives and helping advise ITD on its future direction. A Technical Advisory Group (TAG), a subcommittee of the PRAG, is focused on HMA improvements and future direction.

Future Actions

- Move toward a combination of performance testing and volumetric properties for production acceptance.
    - This will require additional lab equipment and training to implement statewide.
- Implement a longitudinal joint density specification.

- Combine the Density incentive/disincentive with the smoothness incentive/disincentive as a single composite pay factor.
- Continue to address FHWA's recommendations from the Forensic Report and QA Stewardship Review.
- Continue to monitor and implement national best practice.

Executive Summary

Quality Assurance, QA, is a planned systematic approach to secure satisfactory performance of Hot Mix Asphalt (HMA) construction projects. Hundreds of millions of dollars are invested by government and state Departments of Transportation (DOTs) to construct large-scale HMA projects, requiring robust QA practices to ensure they meet quality standards and design life. QA is a statistical approach for checking the desired construction properties through independent testing, encouraged by the Federal Highway Administration (FHWA) since the mid 1960's. However, the standard, conventional QA practice is often criticized on how effective such statistical tests and how representative the reported material tests are. Material testing data correction in the HMA construction sector can render the QA practice ineffective and shadow the performance of asphalt pavements. There is approximately 18 billion tons of asphalt pavement on American roads, with more than $150 billion spent on highway projects in the United States only in 2013, highlighting the significance of QA.

Before 2018, the Idaho Transportation Department relied on contractor-produced Quality Control, QC, test results for the payment of the HMA pavement projects with ITD conducting the QA testing. In 2017, a case study by FHWA reviewed 13 ITD asphalt mix projects and found some unexpected trends where 74% of the ITD test results did not match with the contractor results (see Appendix A, "FHWA Forensic Pavement Assessment for the Idaho Transportation Department (ITD), dated March 8, 2018" for more information). ITD's approach to track down the accuracy of mix design and volumetric test dataset set the stage for this research to mark out instances of these unexpected results in asphalt pavement projects. Based on the findings of the 2017 FHWA study and this project, ITD has since changed its project acceptance and payment approach.

The first objective of this research study was to survey ITD and other state DOT employees' perception of material testing activities in HMA construction projects. The survey was distributed in the late 2019 to ITD staff and to each DOT member of the AASHTO Committees on Construction and Materials and Pavements, with the direct request to widely distribute the survey throughout their organizations. A total of 75 participants responded representing 48 ITD employees and 27 from several other DOTs (exact number of participating DOTs is unknown due to the anonymity of surveys). It's important to understand the survey limitation; specifically, the survey captured a very small portion of other state DOT employees as only 27 of many thousands of DOT employees throughout the country participated in this study. Furthermore, survey results convey participants' perception and cannot be considered as fact without further analyses and investigations. In general, a large portion (>60%) of ITD respondents believed that HMA projects usually do not meet their design life, with a perception that "deficient construction materials", "errors by contractor", and "climatic factors" are the three major causes for this underperformance. A majority of ITD respondents (>60%) believed that the reported mix design and volumetric test data may not be representative of the actual materials used in the field, but they (>70%) also reported that ITD is acting to investigate the reasons for the observed discrepancies. Further, they reported "pressure to affect payment factor in favor of the contractor", "avoiding scrutiny or conflict over results from the contractor", "unwillingness to reconduct the test", "strained workloads" and "avoiding scrutiny or conflict over results from the department" as potential descriptor for changes in the material test reports.

The second objective of this research study was to develop algorithmic logics to identify the patterns of corrections in agency- and contractor-produced QC/QA test results. This was possible with the audit trail dataset that was collected from 15 HMA projects in 2018. Audit trail is a chronological record that reconstructs and examines the sequence of activities surrounding or leading to a specific operation, procedure, or event in a security relevant transaction from inception to final result. This type of audit trail is commonly used in various systems and was even recommended as far back as 1989 in FHWA policy memos providing guidance on the use of source documents and electronic records (see Appendix B). The audit trail was incorporated into the HMA materials testing spreadsheet by the use of a spreadsheet macro that recorded all instances of data entry, and associated metadata, as well as their corrections for audited parameters. This is in accordance with the 1989 FHWA memorandum that stated "*the [computer] records must provide for the reconstruction of the chain of events that occurs on a project*" (see Appendix B). The researchers first manually analyzed several thousand entries in the audit trail data, and determined potential patterns that were categorized as Plausible Correction or Unexplained Correction. The term Plausible Correction (P.C.) refers to instances where a typographical error was likely made. The term Unexplained Correction (U.C.), on the other hand, has been used to refer to instances of data corrections which the project team, after exhausting all options, could not categorize as P.C. The research team then developed algorithmic logics to automatically categorize all instances of data corrections as P.C. or U.C. Research found that data reported by the prime contractors and ITD are both susceptible to data correction, noting that both the prime contractor and ITD may hire a third-party contractor to conduct material testing on their behalf. This research report refers to ITD and the prime contractor as Entity 1 and Entity 2, without any particular order, to avoid potential biases and misconceptions. Results show that a total of 595 and 316 unique parameters affecting prime contractor payment were changed 2,268 and 660 times that can be categorized as U.C. and P.C., respectively, in data reported by Entity 1. For data reported by Entity 2, a total of 387 and 280 unique parameters affecting prime contractor payment were changed 1,266 and 587 times that can be categorized as U.C. and P.C., respectively. Although a total of 15 HMA projects had audit trail data, only 12 projects had enough information (including both QC and QA data) to thoroughly research the impacts of data corrections. Even in the 12 projects, the audit trail data is usually not complete (i.e., the audit trail spreadsheet may not have been used throughout the entire project), but enough data is available to analyze, although the actual, full extent of changes is unknown.

The third objective of this research study was to evaluate whether a monetary impact had been incurred due to data corrections. The research team replicated ITD's procedure for HMA payment calculation, and quantified payment-related parameters and associated payment for each project for two cases: (1) when the first reported value for the parameters categorized as U.C. was used for payment calculation, and (2) when the last value for the parameters categorized as U.C. was used for payment calculation. All non-changed and P.C. parameters were kept at their final reported value. The premise behind this analysis is that the first reported value for a parameter that is categorized as U.C. is most likely the actual observed value from the test, and the last reported value represents the corrected value (in cases where the corrections could not simply be attributed to typographical errors). This analysis showed that there has been overpayment on a majority (i.e., 10 of the 12 projects) of analyzed construction projects across Idaho due to material testing data corrections. Overall, based on the available audit trail data, the research team found that the overpayments ranged between $14,000 to $360,000 in different projects with 2 projects showing a nominal underpayment of $-400 and $-3,000. Further analysis showed that corrections to each major material testing parameter's value can cause roughly $1,000 to $5,000 overpayment per parameter. Data corrections, however, did not always cause monetary gains. Other possible motives may include passing test verification and precision criteria.

The fourth and final objective of this research study was to evaluate the sensitivity of payment-related parameters in HMA projects to material testing parameters. A Monte Carlo study was conducted using a "leave-one-parameter-out" analysis approach to investigate the impacts of each material testing parameter on payment-related factors (VMA, P_a, G_{mm}, and G_{mb}) monitored by ITD. In this analysis, a range for each material testing parameter was determined from analyzing all available HMA projects from 2018, and Monte Carlo analysis with a normal distribution was conducted on this range. Sensitivity analysis shows that parameters in test procedures AASHTO T 209 and T 166 notably impact VMA, and similar test procedures have a conspicuous impact on Pa.

This research study coupled with other ITD initiatives has led to numerous implemented changes by ITD to improve the overall quality of asphalt material production as well as paving operations as further described in the Foreword.

Key Point:

The primary underlying reason(s) behind the significant number of data corrections is unclear. The objective of this research was not to identify who was responsible for the data corrections or the reason(s). Rather, the primary objective was to identify avenues for improvement in the data reporting and QA protocols.

1. Introduction

The development of any nation is greatly dependent on the quality of its physical infrastructure. Although widely accepted as the most developed economy in the world, the United States (US) has been facing significant challenges with respect to the quality of the country's physical infrastructure. The American Society of Civil Engineers (ASCE), in their annual report, continue to assign "failing" grades to the condition of the US infrastructure. For example, in their 2021 annual report card on America's infrastructure (Infrastructure Report Card Executive Summary 2021), the overall condition of the infrastructure was rated at C-, whereas the condition of roads was deemed to be at a "D" grade. Poor roads have resulted in US motorists spending approximately $130 billion every year in extra vehicle repairs and operating costs. The US currently has more than four million miles of roadways. Just to emphasize the magnitude of this number, that's equivalent to a 4-lane highway 40 times around the Earth. Out of the four million road miles, approximately 2.8 million lane miles have a paved surface.

In light of the magnitude and importance of the pavement infrastructure, it is imperative that special care be taken and appropriate measures be implemented to improve the quality of pavements by following "good engineering practice" starting from the project conception stage all the way through construction. The majority of paved surfaces in the US have been constructed using asphalt. Based on a 2018 report released by the US Department of Transportation, approximately 94% of these 2.8 million miles are surfaced with asphalt (U.S. Department of Transportation, FHWA 2018). Besides highway pavements, the use of asphalt is abundant in the construction of airfield pavements as well as parking lots. Based on a document released by the National Asphalt Pavement Association (NAPA), more than 2,650 runways in the Federal Aviation Administration's (FAA's) national airport system are surfaced with asphalt pavements (NAPA fast facts 2020).

The above statistics clearly establish the importance of the asphalt pavement network as well as the asphalt paving industry towards the growth of the country's economy. To ensure adequate quality of the paved surfaces, it is critical that strict measures be implemented to exercise control over the quality of asphalt mix being produced, as well as the construction practices. This can be accomplished through the implementation of well-developed Quality Assurance (QA) protocols. QA is defined as *"All those planned and systematic actions necessary to provide confidence that a product or facility will perform satisfactorily in service."* in AASHTO R-10, Standard Recommended Practice for Definition of Terms Related to Quality and Statistics. A QA program comprises the following six core elements: (1) contractor Quality Control (QC), (2) agency acceptance, (3) independent assurance (IA), (4) dispute resolution, (5) laboratory accreditation, and (6) personnel certification (Hughes et al., 2020). Out of these elements, some of the components fall under the category of 'agency responsibility', whereas some others fall under the category of 'contractor responsibility'. When it comes to the pavement construction, QC protocols are usually implemented by material producers and contractors, whereas agency acceptance is implemented by state and local highway agencies (Departments of Transportation or DOTs).

QC/QA Practices in the Asphalt Paving Industry

From the mid-1960's, the Federal Highway Administration (FHWA) started to demand for the exercise of statistics-based QC and QA specifications from contractors as well as state DOTs. These methods help the contractor assess whether the operations are producing an acceptable product. DOTs commonly use the QC and QA test results to determine the Pay Factors (PFs) for all the pay elements in a project. Early on, QC data was mostly omitted from the decision-making process, since there were always concerns regarding biased reporting by the contractors as compared to findings from the state transportation departments. However, an unpublished report in 1993, titled *"Limits of the Use of Contractor Performed Sampling and Testing"* emphasized on the utilization of contractor-produced QC data for better decision making in Hot Mix Asphalt (HMA) projects. Later in 1995, these recommendations were converted into federal law with the enactment of 23 CFR 637 (Office of the Federal Register, and Office of the Federal Register (US) Staff 2007). However, these regulations paved the way for the proper utilization of contractor-produced QC data based on the fulfillment of several prerequisites, including:

1. The technicians employed by the contractor must be qualified to perform sampling and testing.
2. Verification samples and testing of the material must be done independently of QC tests.
3. An independent assurance program must be used to examine the QC sampling and testing.

State DOTs have different HMA standard specifications, pay factors, tests to evaluate new pavement, payment plans, lot and sublot sizes and specification limits (Newcomb et al. 2017). Hand et al. (2020), through a survey of state DOTs in the US, reported that 22 out of 29 responding DOTs used contractor test results as a part of the acceptance procedure.

Project Background and Justification

From the mid-1990's to 2018, ITD relied on the prime contractor-produced QC results to determine pay factors for each HMA project. ITD did not pay the prime contractor based on QA verified results conducted by ITD or its hired third-party material testing contractor. Using this practice, ITD was effectually in noncompliance with the FHWA by paying the prime contractor without verifying the test results before 2018. Accordingly, in scenarios where discrepancies were observed between the prime contractor-produced and ITD-produced test results, the further steps taken to investigate the factors contributing to those discrepancies were futile due to the cumulative effect of various steps in the pre-2018 version of the QC/QA system and the lack of an avenue to pay off of ITD test results. This resulted in less accountability when testing deviations were found in the QC and QA test results. During 2017, ITD contacted FHWA, who undertook a forensic investigation of 13 projects to detect and analyze possible discrepancies between ITD-produced and the prime contractor-produced test results. Out of a total of 77 comparisons (based on comparing six key parameters) carried out, good agreement between ITD- and contractor-produced results were observed for only 20 (26%) of the cases; in 57 (74%) instances the two datasets did not “agree” with each other. This observation, combined with the fact that several pavement sections in the state of Idaho underwent premature failure, suggested a need to thoroughly examine and potentially modify ITD’s prevalent practices as far as quality assurance of HMA production and construction are concerned (see Appendix A, “FHWA Forensic Pavement Assessment for the Idaho Transportation Department (ITD), dated March 8, 2018” for more information). In-depth inspection of such discrepancies, quantification of their potential impacts on the pavement lifecycle costs, and recommendations for improvement would ultimately improve the overall condition of asphalt pavements in the state of Idaho. The current research study was undertaken under these premises, and had an overall objective of improving the HMA QA practices implemented by ITD.

Evidence of Correction in HMA QC/QA Data

This research study began with the goal of identifying why discrepancies existed as demonstrated in the FHWA Forensic Pavement Assessment report. During the course of this research study, the focus was changed to include the audit trail data indicating corrections were being made during asphalt quality control and acceptance testing. This audit trail data was collected using a spreadsheet macro in the ITD’s spreadsheet form 0777 titled *“Superpave Production Test Report”* (see Appendix C). This is a form that is used to report results during quality control and acceptance testing. The audit trail macro monitored all the data entries being done on this form. For example, if the user first input the number “1” in a particular cell of the form and later deletes it to enter “2”, the macro kept a log of both entries along with the corresponding metadata such as time stamps. Therefore, the reviewer of the data can get a chronological record of all the data entries made during the testing of a particular set of samples. It is important to note that the macro keeps track of all the data entries made into the cells within the spreadsheet form, and does not distinguish between instances where the user makes a typographical error, thus requiring them to change the number, compared to a case where the user may have intentionally changed a test result for some unknown reason. Careful judgement needs to be practiced to distinguish between these two instances. For the purpose of this research, instances where the data corrections can be easily attributed to typographical errors, have been labeled as *Plausible Correction* (P.C.). On the other hand, instances where no obvious justification for the data correction can be provided, after exhausting all options, have been labeled as *Unexplained Correction* (U.C.). The data analysis in this research project focused largely on differentiating between instances of P.C. and U.C., and subsequently quantifying the effect of these U.C.’s on the overall project costs.

Report Organization

This report documents findings from the research study titled *“Statistical Analysis of 2018 HMA Production and Construction Data to Improve Quality Assurance and Acceptance Practices in Idaho”* undertaken in collaboration between the Idaho Transportation Department (ITD), Boise State University, and Oklahoma State University. The primary objective of this research study was to identify sources of discrepancies/inconsistencies in HMA production and construction data through extensive statistical analysis.

The contents of this report have been divided into six chapters. Chapter 2 documents findings from an extensive review of published literature carried out on the topic of quality control and acceptance testing in the asphalt industry, and how the QA practices in Idaho as well as the rest of the country have evolved over the years. This chapter also documents findings from a national and state-wide (in the state of Idaho) survey of highway agency personnel associated with asphalt paving projects to assess the prevalence of data inconsistencies between the state DOT- and the contractor-produced test results.

Chapter 3 presents details about the data analysis effort, and describes the development of a logic-based framework that was used by the research team to differentiate instances of U.C. from P.C. The chapter also provides details about the functionality of this spreadsheet-based macro that was developed by the ITD staff to collect the data which was then provided to the research team. Chapter 4 analyzes the impact of the U.C. instances on the total project costs. Pay factors were compared for each paving project in the presence and absence of data corrections that could not be immediately attributed to user-related mistakes such as typographical errors. Quantifying the changes in project costs induced due to these U.C.s can give an idea about the quantity of additional taxpayer dollars being spent as a result of improperly implemented testing and QC/QA practice.

Chapter 5 presents findings from a sensitivity analysis carried out to assess how changes in different asphalt mix parameters affect the overall mix approval and associated costs. Chapter 6 presents a summary of the research findings, draws logical conclusions based on the findings, and makes recommendations for future research that should be pursued by ITD to improve the overall state of asphalt production, testing and paving practices in the state of Idaho.

It is important to note that asphalt concrete mix is often referred to as *H*ot *M*ix *A*sphalt or HMA due to the commonly used method of producing this mix (individual components are heated to an elevated temperature and mixed). Recently, different technologies have been developed that has reduced the production temperature for asphalt concrete mixes. In some cases, the term *W*arm *M*ix *A*sphalt (WMA) may be used to refer to a mix produced at lower than traditional temperatures. Nevertheless, the term HMA is often used to refer to an asphalt concrete mix irrespective of the exact temperature at which it is produced. *Accordingly, in this report, the term HMA has been used to refer to asphalt concrete in general, without making any implications regarding the temperature at which the mix has been produced.*

All the data analyzed, and the practices referred to in this research study correspond to ITD practices through the end of the 2018 calendar year. ITD has since made significant changes into the quality acceptance process as well as other practices related to HMA production and paving as further described in the Foreword.

2. Literature Review

Introduction

QA specifications have now become an integral part with the commitment to overall quality management. The quality management comprises of multiple components including but not limited to process control, independent assurance and acceptance of product (Buttlar and Harrell 1998). These specifications are designed to reward for increments of high quality. In general, DOT's are given the task to accept contractor-produced products that have reached the quality mark set by the highway agencies.

QC is defined as the continuous process of a contractor (i.e., pavement contractor in HMA projects) monitoring, assessing, and adjusting their production so that the final product meets the prespecified level of quality, whereas QA is defined as the process of ensuring the construction quality will satisfy the owner-specified requirements and is assigned to DOTs (TRB 2018). A third term, "Acceptance" refers to sampling and testing or inspection to determine the degree of compliance with contract requirements (TRB 2018 Glossary of Terms). It should be noted that the term QA *addresses the overall problem of obtaining the quality of a service, product, or facility in the most efficient, economical, and satisfactory manner possible. Within this broad context, QA involves continued evaluation of the activities of planning, design, development of plans and specifications, advertising and awarding of contracts, construction, and maintenance, and the interactions of these activities* (TRB 2018 Glossary of Terms)

From the earlier days, practices and methodologies used by federal and state highway agencies varied widely in purposes and concepts. Although these specifications are varied from state to state, in most cases they followed the statistical analysis concepts with the following three basic components:

i. Quality Control
ii. Acceptance Protocols (the DOT must have some method to verify the contractor's test results for acceptance and payment such as a QA process)
iii. Independent Assurance (IA)

All the processes of QC are typically conducted by the contractor, with different components including, but not limited to mixing, blending, sampling and handling of HMA materials. A well-developed and implemented QC plan has the following advantages (Caltrans 2020):

- Process can be controlled in the desired manner,
- Contractor can quickly search for alternatives in case the project missed the requirements,
- Contractor can properly respond to and correct any unwanted situation, and reroute the process to the correct track.

Acceptance protocol is in general maintained by the state highway agencies, although there are some cases where the contractor-produced test results are used in this step. The last step, Independent Assurance (IA), typically involves testing and evaluation of the samples by a third-party (other than the contractor and the DOT personnel associated with the project).

The prime contractor is responsible for inspection of transportation, production and placement of HMA, and the completion of the roadway. The production inspectors are generally required to inspect the HMA production plant before the first day of the work starts. If an HMA plant does not satisfy the requirements, the QC manager needs to take necessary action. At most of the HMA production plants, it is necessary for the production inspector to visit the asphalt production plant on a daily basis before the start of the production. There are varied guidelines provided by different state DOTs for plant inspection. Some common guideline to follow are (Caltrans 2020; Transportation Research Board 2018):

I. Carefully checking the overall plant to cross-check whether the prime contractor controls dust or smoke as required
II. Documentation of daily HMA production information
III. Testing the reclaimed asphalt pavement samples for moisture content
IV. Ensuring that the temperatures of the aggregate and asphalt binder are within the limit
V. Confirming that the batch size and the feed rates do not cross the mixing capacity range

Pay Factor (PF)

PF is used to calculate the payment to prime contractor based on the quality of their work. Akkinepally et al. (2006) displayed a flow chart for establishing a relationship between PF and prime contractors' quality of work. The PF is essential for finalizing the payment to the prime contractors. If the work of the prime contractor meets the limit of the available QC/QA specification they will receive the total one-hundred percent of the contract. But when the product quality fails to meet a certain quality threshold, some DOTs deduct an amount from the prime contractor in order to ensure contractors improve the quality to the level required by the state DOT. On the basis of pay adjustment factor, the payment of the prime contractor is generally recalculated. Most of the agencies follow the procedure of calculating individual PFs. Sometimes a composite PF is also calculated by multiplying the respective weight of each of the quality parameters. The PF in general varies from state to state, and agency to agency.

The schematic relation between QC, QA, and other major components of the statistical PFs is shown in Figure 2.1. According to the literature, pavement density and surface smoothness are frequently used to confirm construction adequacy. Often for purposes of QC, nondestructive tests using nuclear/non-nuclear gauges and for QA purposes, the destructive core density tests are used to determine pay-elements (Williams and Hall 2008, 150). Although alternative nondestructive measurement methods such as intelligent compaction can be used to avert induced cracking and reducing in pavement performance (Beainy 2011; Chang and Gallivan 2011). DOTs determine those pay-elements and a particular weight for each element according to its significance and correlation to the pavement performance and generally the more those pay-elements are balanced the longer HMA mixture can remain in service without any problems and the higher HMA mix quality would be. Particular weights are given to each element according to its importance and relevance to the pavement performance. This is done by the DOT using previous data and the assumption that these elements are related to the pavement performance. Furthermore, engineers' experiences and industry input are used to determine PF weights. (Al-Khayat et al. 2020, 146). DOT also determines the lot size (e.g., 5,000 t, 10,000 t, daily product, or any other definition) and sublot size (e.g., 500–1,500 t).

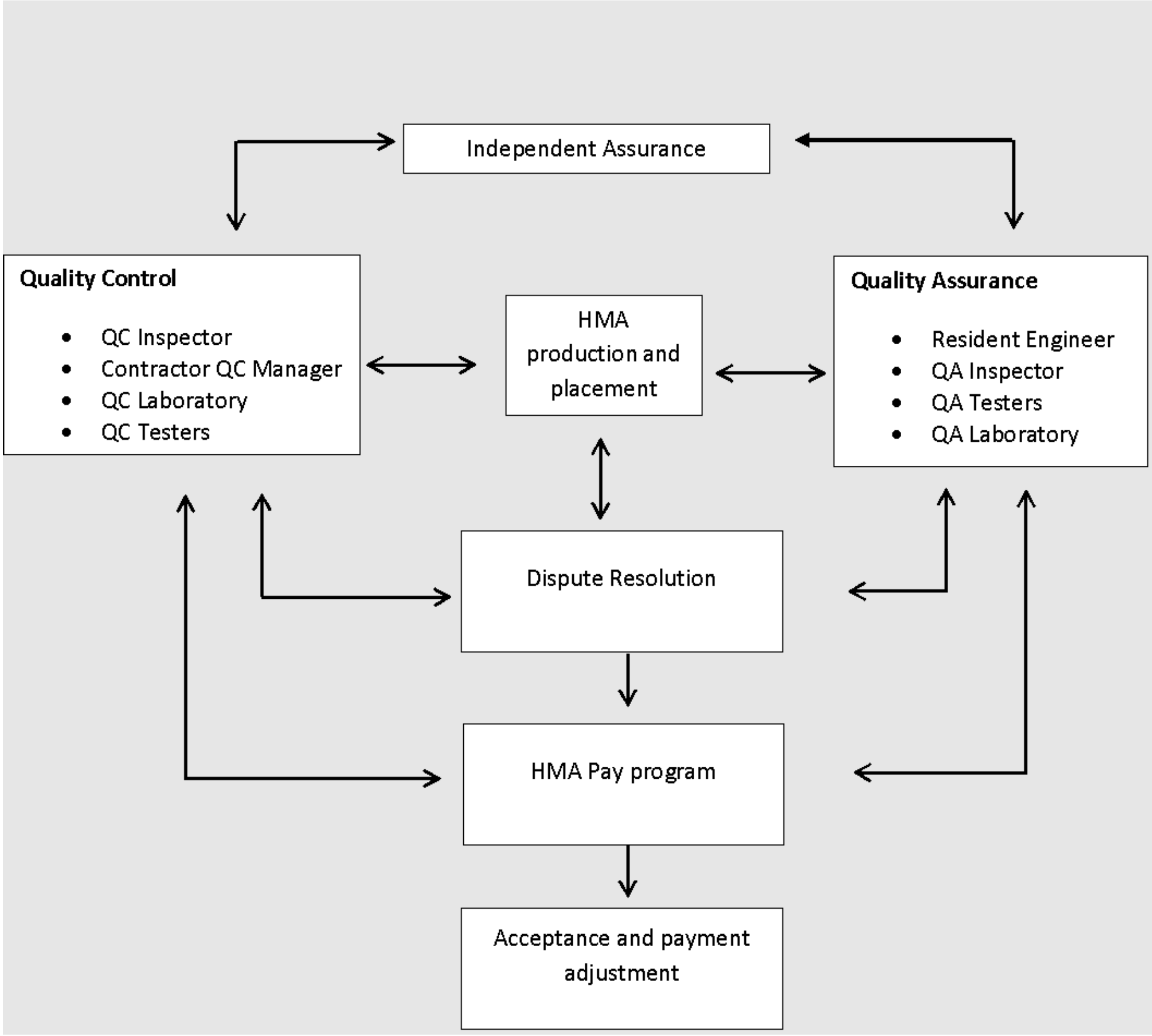


Figure 2.1 Relationship between QC/QA components

In general, prime contractors are responsible for the quality of the HMA produced and used in a paving project. It is one of the major responsibilities of the contractors to verify the quality of materials supplied by the subcontractors, suppliers and other external sources. QC/QA program is only effective when the prime contractor provides necessary quality control in accordance with the specifications. Financial incentives and disincentives for quality can help motivate the prime contractor to produce the best quality product they can.

Percent Within Limits (PWL) Specifications

DOTs, using previous data and based on assumptions, select PFs and their weighting factors. Furthermore, DOTs calculate the PWLs and PFs for each Lot. By considering the PWL specification, DOTs accept or reject part of the mixture for each lot. Moreover, according to the type of PF policy implemented, prime contractors may receive a financial incentive for high quality of mixture, or may be penalized for poor quality of the mixture, including having to remove and replace unacceptable asphalt. The PWL, by considering the sample average and standard deviation of each property of the HMA mix, estimates the percent of the materials that is within the specification limits. For instance, if for a specific lot the PWL of asphalt content is 87%, it means this lot has 87% of the mix within the limits of the asphalt content. The PWL is used to reward a producer that manufactures a product that is very close to the target value and has low variability. The low variability represents the consistency of the products in a project. From the ITD’s QA special provision, a PWL of 90 represents the full possible payment. Any prime contractor not receiving full payment may elect to remove defective material and replace it with new material on a lot basis, at no additional cost to ITD, to avoid a PWL of less than 90.

HMA projects’ different parameters including standard specifications, lot and sublot sizes, payment plans, PFs, tests to evaluate new pavement, and specification limits vary depending on the state DOT (Newcomb et al. 2017). These standards are monitored and are updated by the state DOT and usually, DOT accepts an HMA mixture when specification limits applied in PWL specification are within limits. However, these limits are dynamic and are developed by DOTs based on the variabilities in test methods and repeatable tests results. Such variabilities stem from different construction and testing techniques, technical errors, and the nature of the materials being used that can be difficult to control. These cause discrepancies that often lead to prime contractors not being able to achieve, for example, the exact target value of asphalt content or density. Accordingly, DOT defines specification limits for different parameters such as target value and materials falling within those limits would be accepted. Therefore, for each pay element specification, limits are developed according to statistical analysis based on their specific typical variability. Too wide or too narrow limits can be risky for the DOT and contractors and this can minimize or maximize the DOT payment to the prime contractor (Al-Khayat et al. 2020, 146).

Chronology of QC/QA in the US

DOT QA programs since the 1960s, had to adhere to 23 CFR 637 and be approved by FHWA (2007). The 23 CFR 637 indicates that the prime contractor test data can be used for acceptance of construction materials given that DOT validates the test data with independent test results. As a result, a key decision is made to determine whether the DOT should conduct the acceptance sampling and testing or use prime contractor test data for acceptance in the QA process. If it is decided by the DOT to conduct the acceptance sampling and testing, a combination of F- and t- hypothesis tests is commonly used to assess whether the DOT and prime contractor test results belong to the same statistical population/distribution or not. Subsequently if the DOT uses prime contractor test results in one or more of their acceptance decisions processes, then, QA program must meet certain requirements including DOTs are expected to use an appropriate process to validate the prime contractor's results. Before 1992, a survey conducted by AASHTO viewed that 8 states among 50 in the US, have made significant plans or already implemented the QC/QA specifications (Smith 1998). By 2005, around 46 states had already implemented the QC/QA specifications (Hughes 2005). A web-based survey of DOTs by Hand et al. (2020, 10) showed that 22 out 29 state highway agencies use prime contractor test results as part of the acceptance procedure. Therefore, each state highway authority should ideally develop a QA program that will ensure that the materials and workmanship incorporated into each federal-aid highway construction project on the national highway system are in conformity with the requirements of the approved plans and specifications, including approved changes. What is more, FHWA also conducts stewardship reviews to assess DOT QA program practices and procedures, as well as ascertain the status of DOT implementation of QA regulation (Al-Qadi et al. 2020). Literature review shows that erroneous acceptance and payment decisions can be contributed to deviations from implementation of statistical procedures described in the AASHTO including conducting inadequate number of DOT tests, elimination of the F-test, using inappropriate comparison methods (i.e., single DOT test method and the Difference Two-Sigma Limit method), using split rather than independent material samples, and performing unwarranted retests (Wani and Gharaibeh 2013, 67), whereas studies confirm that the use of F- and t-tests are effective tools for validating QC results (LaVassar et al. 2009). In this hypothesis testing the F-test and t-test identify possible differences in variances and means of the two datasets respectively. The overall work of this process is shown through a flowchart in Figure 2.2 (Hand and Epps 2006, 140).

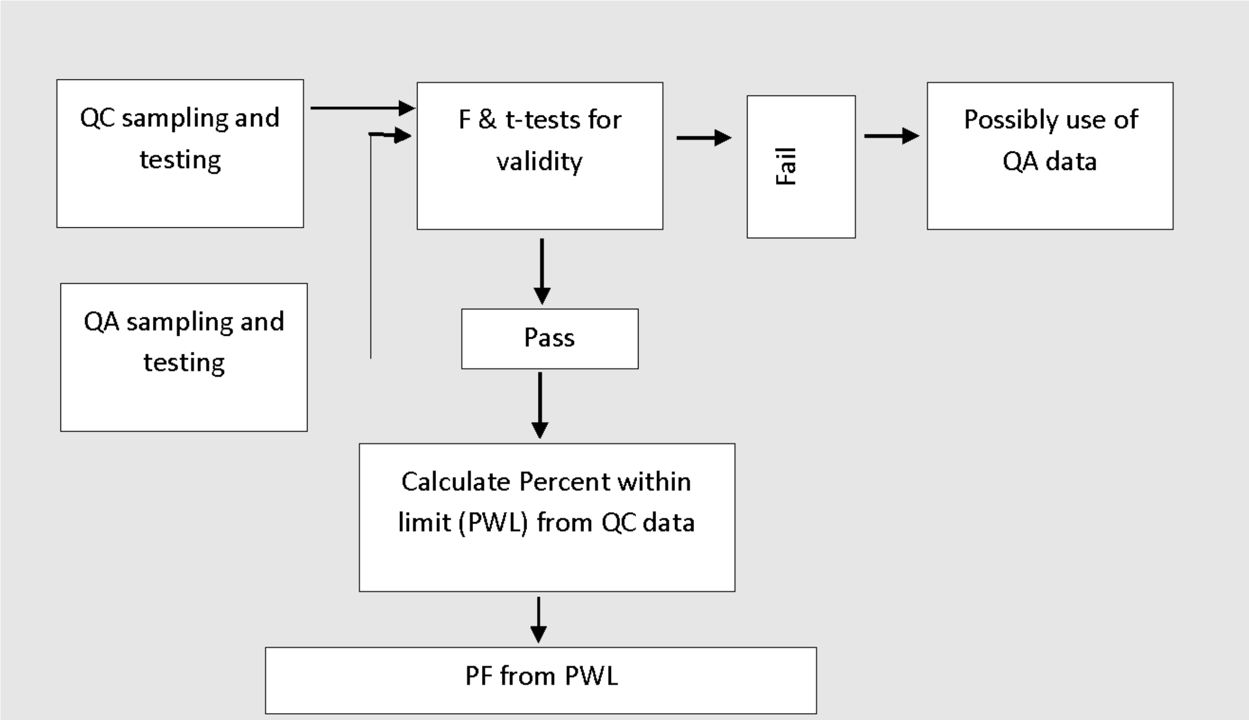


Figure Source: (Caltrans 2020)

Figure 2.2. Workflow of Statistically-Based QA Specifications

Per each lot, DOT performs a statistical analysis to calculate PWL value and based on PWL the amount of payment for the prime contractor is determined. Muench et al. (2001) defines PWL as the percentage of lot falling in between the lower and the upper specification limits. Similarly, percent defective was defined as the percentage of lot falling outside specification limits (Muench et al. 2001). To produce a uniform HMA mixture, each lot should have a certain job mix formula (JMF). The PF is then calculated by considering target value in the JMF in addition to specification limit (the upper limit) and the target value in the JMF subtracted from the specification limit (the lower limit). Once sufficient test results are gathered, certain percentage of quality limits (e.g., 90%) can be set to determine whether a prime contractor should receive the full payment or not (AASHTO 2016).

Different studies recently have been conducted to increase the robustness of the validation techniques for QA specifications. One example is a study by Schmitt et al. (2001, 86) where statistical parameters were proposed for either split or independent sampling representing the variation, risk, and size of the dataset to be used for finding mean differences between two datasets. Monte Carlo simulations are used to evaluate prime contractor test data based on F- and t-tests and detection of data alteration and expected pay. Using this method Wani and Gharaibeh (2013, 67) produced operation characteristic curves for identifying the probability of detection of data alteration. Their simulation results demonstrated that even at high sampling rates the power of the test was low such that the expected pay (for sample size 5 vs. 20) can be changed from 87% to 100% without detection in the most extreme cases, thus resulting in significant overpayment. They concluded that using larger lots and sample sizes to compare DOT test results to prime contractor test results and separating the project management team from the prime contractor quality management team can reduce detection error.

Other studies also suggested that accumulation of DOT and prime contractor test results on consecutive lots should be considered in an effort to increase sample size and improve power of statistical tests (Arambula and Gharaibeh 2014, 140). To this end, continuous cumulative and chain-lot sampling methods were used and the result of analysis showed that a chain-lot sampling method (with three accumulated lots) can significantly increase the power of F- and t-tests for detection of data corrections. On the other hand, the authors reported no significant change in PWL using this sampling method. Another validation technique is the use of risk-based multi-tiered verification, which ranks the acceptance quality characteristics of materials given its relative impact on performance to 1, 2, and 3 (where 1 is most and 3 is the least important). Thus, in case of category 1 AASHTO F- and t-tests is used for verification and in case of category 3 a simple review of control charts of prime contractor/inspection data would suffice. Assuming that data follows a normal distribution and both data have equal variance the pooled t-test is an alternative approach for purposes of hypothesis testing. To relax the assumption of having equal variance, Satterthwaite's method can be used that accounts for the discrepancy in variances by adjusting standard error and degrees of freedom (Satterthwaite 1946, 110). The nonparametric hypothesis tests are more robust since the assumption regarding normality, equality of variances, and independence of the populations are not required. One of the most popular and frequently used nonparametric hypothesis tests is the Mann-Whitney/Wilcoxon that tests the difference in medians of the populations or means of the populations if they are not symmetric (Wilcoxon 1945, 80). Other nonparametric methods are available in the literature for two-sample hypothesis testing including the Levene test and Fligner-Killeen test (Levene 1960, 278; Fligner and Killeen 1976, 210). Similar to nonparametric tests the bootstrap test does not require the assumptions of normality and quality of variance and by resampling with replacement from the sample data, they create bootstrap samples and approximate the sampling distribution of the statistic (Efron 1979, 569; Good 2005). These conclude that although F- and t-tests are powerful methods to capture difference in mean and variance of two datasets, extreme caution is to be made to ensure their assumptions are not violated (i.e., data not following a normal distribution) and in case those assumptions do not hold, then other nonparametric, permutation, and bootstrapping methods can be used to avoid misleading results (Moser and Stevens 1992, 19; Ruxton 2006, 688; Zimmerman and Zumbo 2009, 371; Kahler 2012).

Chronology of QC/QA in Idaho

ITD used Materials and Methods specifications until the late 1990's (1997 or so). Those specifications direct the prime contractor to use specified materials in definite proportions and specific types of equipment and methods to place the material. Each step was directed by a representative of the highway agency which meant ITD usually had someone at the hot-mix plant, someone taking weight tickets from the truck drivers and watching the dumping or windrowing of the mix, someone supervising the paver laydown operation, someone supervising the rollers and verifying the roller patterns. Finally, ITD representatives would also be involved in collecting density gauge readings, sampling and testing the mix (though sometimes these duties overlapped), etc. Accordingly, ITD maintained direct control over the entire process; QA tests were merely carried out to collect some information on the project. In 1994, ITD began its move toward QC/QA by statistically comparing the QC and QA results along with measuring variability of the materials used in ITD projects. One of the important factors in transition from materials and methods specifications to more testing of the final product was the high personnel requirements, along with the shrinking personnel resources of many highway agencies. This change also shifted the risk of production from ITD to the prime contractor. With ITD directing each activity, ITD was wholly responsible for the outcome.

According to FHWA 23 CFR Part 637B, the definitions for QC and acceptance differ from those in the TRB glossary. The FHWA 23 CFR 637B definition for QC is *"All Contractor/vendor operational techniques and activities that are performed or conducted to fulfill the contract requirements."* This definition was adapted from standard ANSI 90 and standard ISO 9000. Acceptance program is defined as "All factors that comprise DOTs determination of the quality of the product as specified in the contract requirements." Later in 1996 Quality Team was formed to implement and oversee QA measures in accordance with the CFR and to ensure quality of materials and construction on ITD’s roadways by partnering with prime contractors. The team was reestablished and renamed the QA Specification Team in 2003 where, the Division of Engineering Services Administrator serves as the team’s executive sponsor to *“provide continued development and improvement of the Department’s QA specifications, measures, and programs to assure quality materials are incorporated into department projects.”* (QA Manual Section 2110.00 QA Specification Team). Until 2018, ITD was exclusively relying on the prime contractor-produced QC results to determine PFs of all HMA projects. Despite the conductance of verification tests in ITD laboratories, the results would have no effect on the project payments. Further steps were taken to identify the factors contributing to the observed discrepancy between the prime contractor and ITD produced test results were futile due to cumulative effect of the QC/QA system used and the lack of an avenue to pay off of ITD test results.

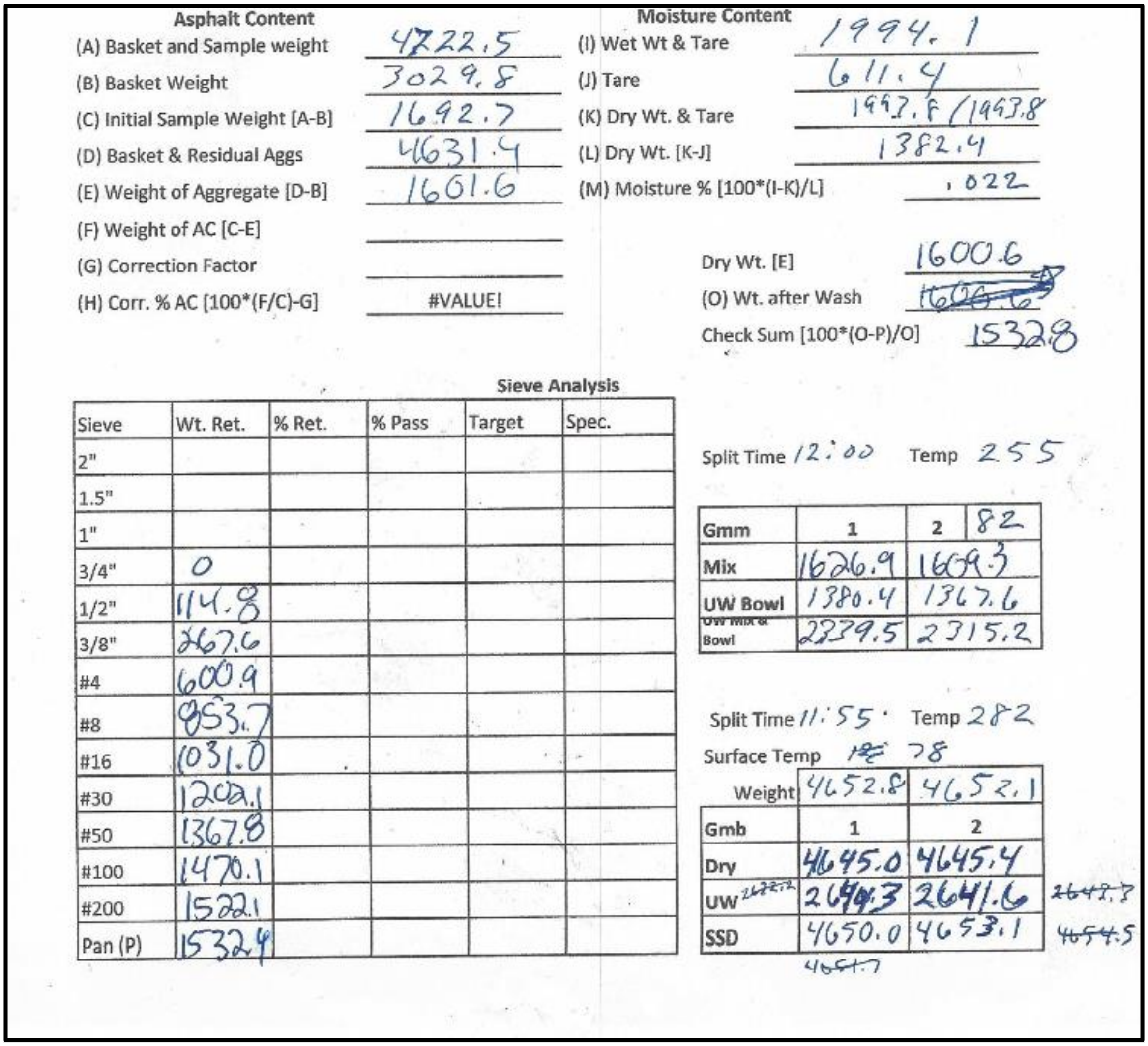
In 2018, ITD updated its QC/QA program to comply with 23CFR 637.207 to verify the prime contractor’s QC test results using a statistical verification process. Also, in 2018, ITD used three payment-related factors: Air voids, Voids in Mineral Aggregate (VMA), and Density for calculating the PWL and PF and finally, Composite Pay Factor (CPF) which includes both mix quality characteristics and density. It was during this time period that the research team analyzed the submitted test results including the audit trail data for this research.

In 2020, ITD stopped using prime contractor QC test results for acceptance and payment, and began utilizing only ITD or ITD’s hired third-party contractor test results for acceptance and payment. ITD also made some changes to its acceptance criteria by adding another quality characteristic, Dust Proportion, in the procedure of calculating composite pay factor (CPF) and payment adjustment. Thus, in 2020, the formulation of composite pay factor includes just quality characteristics and a new, tiered density pay factor incentivizing additional compaction is used in calculating the payment adjustment. Additional changes made by ITD can be reviewed in the Foreword.

Data Correction

Both the 2018 FHWA forensic report and anecdotal interviews with ITD and consulting engineers point to the possible existence of data corrections in HMA project reports (Dutton 2020; 2021). Errors and mistakes in material testing and reporting can result in lower than expected service life of HMA projects, higher maintenance costs, and in extreme cases even lower safety. ITD is investing ~$450 million (both federal and state funds) in Idaho highways in 2021 and a similar amount each year afterward by 2027; and material testing values have the potential to cost taxpayers millions of dollars if not done properly (ITD 2019).

Figure 2.3 shows an image of a laboratory datasheet submitted to ITD during one of the HMA projects reviewed by the research team. As seen from the datasheet, the values in several fields were corrected and over-written several times during the course of testing. This is particularly evident from the Under Water (UW) and Saturated Surface Dry (SSD) weights. Some of this can be attributed to the possibility that scale readings were affected by the testing environment such as excessive wind draft in the laboratory. However, repeated occurrence of such trends raises concerns about the quality of the test results or the competency of the tester. Moreover, such instances of changes were also observed in cases where the test data was entered into spreadsheet form and captured in the audit trail. This emphasizes the importance of studying the extent of such data inconsistencies in the reported values, and developing approaches to prevent future occurrences of poor/inaccurate testing and reporting practices.



Asphalt Content

(A) Basket and Sample weight 4722.5
(B) Basket Weight 3029.8
(C) Initial Sample Weight [A-B] 1692.7
(D) Basket & Residual Aggs 4631.4
(E) Weight of Aggregate [D-B] 1601.6
(F) Weight of AC [C-E]
(G) Correction Factor
(H) Corr. % AC [100*(F/C)-G] #VALUE!

Moisture Content

(I) Wet Wt & Tare 1994.1
(J) Tare 611.4
(K) Dry Wt. & Tare 1993.8/1993.8
(L) Dry Wt. [K-J] 1382.4
(M) Moisture % [100*(I-K)/L] .022

Dry Wt. [E] 1600.6
(O) Wt. after Wash ~~1606.6~~
Check Sum [100*(O-P)/O] 15328

Sieve Analysis

Sieve	Wt. Ret.	% Ret.	% Pass	Target	Spec.
2"					
1.5"					
1"					
3/4"	0				
1/2"	114.8				
3/8"	267.6				
#4	600.9				
#8	853.7				
#16	1031.0				
#30	1202.1				
#50	1367.8				
#100	1470.1				
#200	1522.1				
Pan (P)	1532.4				

Split Time 12:00 Temp 255

Gmm	1	2	82
Mix	1626.9	1609.3	
UW Bowl	1380.4	1367.6	
UW Mix & Bowl	2339.5	2315.2	

Split Time 11:55 Temp 282
Surface Temp 78

Weight	4652.8	4652.1	
Gmb	1	2	
Dry	4645.0	4645.4	
UW ~~2642.2~~	2649.3	2641.6	~~2647.7~~
SSD	4650.0 4651.7	4653.1	~~4654.5~~

Figure 2.3. Data Correction on a Paper Data Reporting Sheet

The 2018 forensic report by FHWA on 13 selected projects showed 4% average difference in incentive pay for asphalt mixture between the prime contractor and ITD results. If these projects are indeed representative of all projects in Idaho, extrapolating this would be about $4,300,000 extra in incentives in just one year. Research team acknowledges the inaccuracy of extrapolated calculations, but offer this estimate as a rough baseline for the scale of the issue. This mismatch not only can impact pavement projects' PFs, but also can have significant repercussions concerning the pavement service life and maintenance costs. Hence, reviewing and modifying ITD's current practices and policies regarding the QC/QA of HMA production and construction is necessary, and an in-depth inspection and analysis of the sources of these observed discrepancies between prime contractor and ITD test results can shed light on how construction costs, maintenance costs, and pavement life are altered as a result. Such inspection and analysis could ultimately help improve the overall condition of asphalt pavements in the state of Idaho.

Findings from National and State-Wide Survey

This research study further investigated the potential prevalence of and causes for the discrepancies in the material testing reports between prime contractor- and agency-generated material testing data and the actual materials used in the construction. Accordingly, a survey was designed and sent out to state DOTs through two AASHTO committees, the Committee on Construction and the Committee on Materials and Pavements in late 2019. The objective of the survey was to inquire about their state of practice for prevention and reduction of suspicious, and in extreme cases fraudulent, activities in construction and maintenance of HMA projects. This section provides an in-depth and analytical survey of the ITD and other state DOT engineers about the state of practice across Idaho and the nation, and engineers' perception of the prevalence of suspicious activities in ITD-sponsored and other DOT-sponsored projects. This chapter also provides information about the current practice at different ITD districts, and helps develop a consistent, rigorous quality assurance platform across the state. This survey will also create a baseline to design further training activities for ITD engineers. The conducted survey questionnaires can be found in Appendix D. These questions were designed by the research team through analyzing and synthesis of available literature, as well as consultation with engineers that work in the field of HMA construction. We also consulted with Dr. Steve Utych, a social scientist at Boise State University, to neutralize the language of the survey and ensure that question language does not have an impact on the respondents' choices.

Questionnaire Results

The questionnaire (see Appendix D) was sent out to different DOTs across the U.S. and the results were analyzed separately for Idaho and all other states. This allows intercomparing of methods, regulations and behavior of ITD and all other states. This section analyzes and discusses the answers given by the ITD employees and DOT employees in all other states. In the rest of this chapter, we refer to all DOTs except for ITD as state DOTs, and refer to ITD as ITD. Overall, 75 individuals responded to the survey with 48 ITD respondents and 27 respondents from several other DOTs. Number of DOTs that participated in this study is unknown, given the anonymous nature of the survey.

- Q4) Do you have experience working with asphalt pavements?

Figure 2.4, demonstrates asphalt pavement work experience of employees who participated in the survey. This figure shows that all employees from other states (state DOTs) that responded to this questionnaire had some working experience with asphalt pavement, and in Idaho 92% percent of individuals had work experience in asphalt pavement.

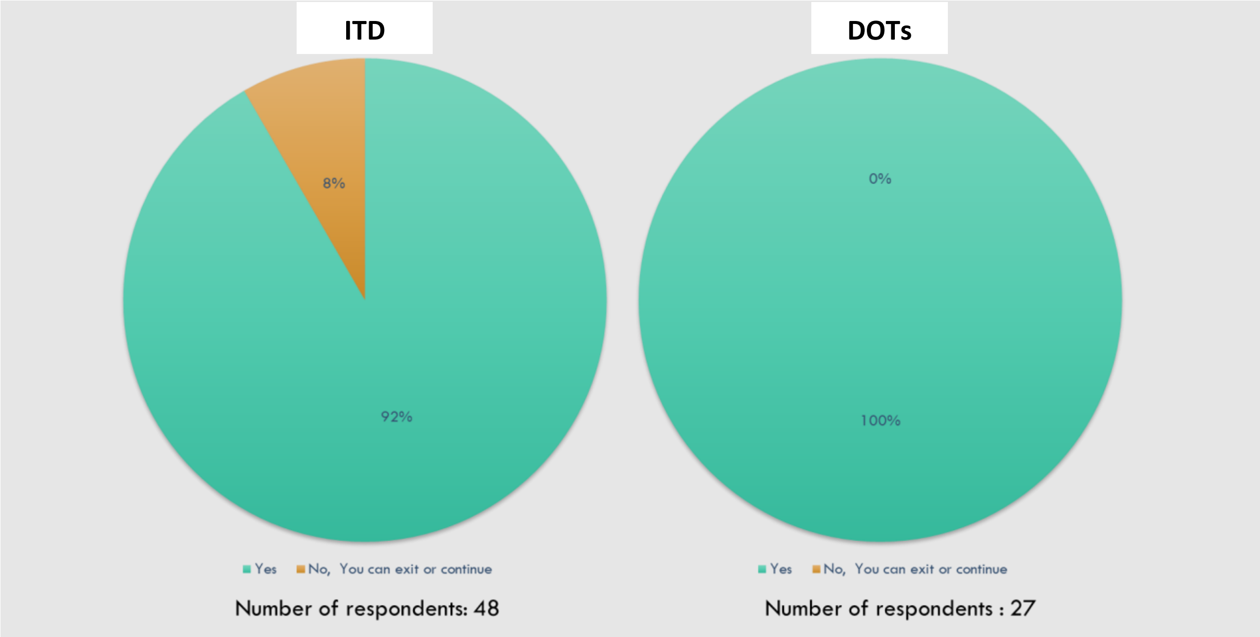


Figure 2.4 Work experience of survey participants with asphalt pavements

- Q5) Does your state Department of Transportation (DOT) implement any procedure to compare actual service life (before major rehabilitation efforts need to be undertaken) of asphalt pavements against the original design life?

Interestingly, the number of employees from state DOTs confirming that their agency compares actual service life of asphalt pavement to that of its original design life (40%) is almost identical to those who believe this procedure does not exist in their states (44%; Figure 2.5). A considerable percentage of participants (16%) responded with "Not sure" to this question. Similar pattern was also observed in Idaho but with a great percentage (33%) unaware/unsure that this procedure is being implemented or not. This indicates the need to emphasize the importance of QA procedures and further familiarizes DOT's employees with QA procedures and critical analysis of rehabilitation costs of constructed pavements during their lifetime.

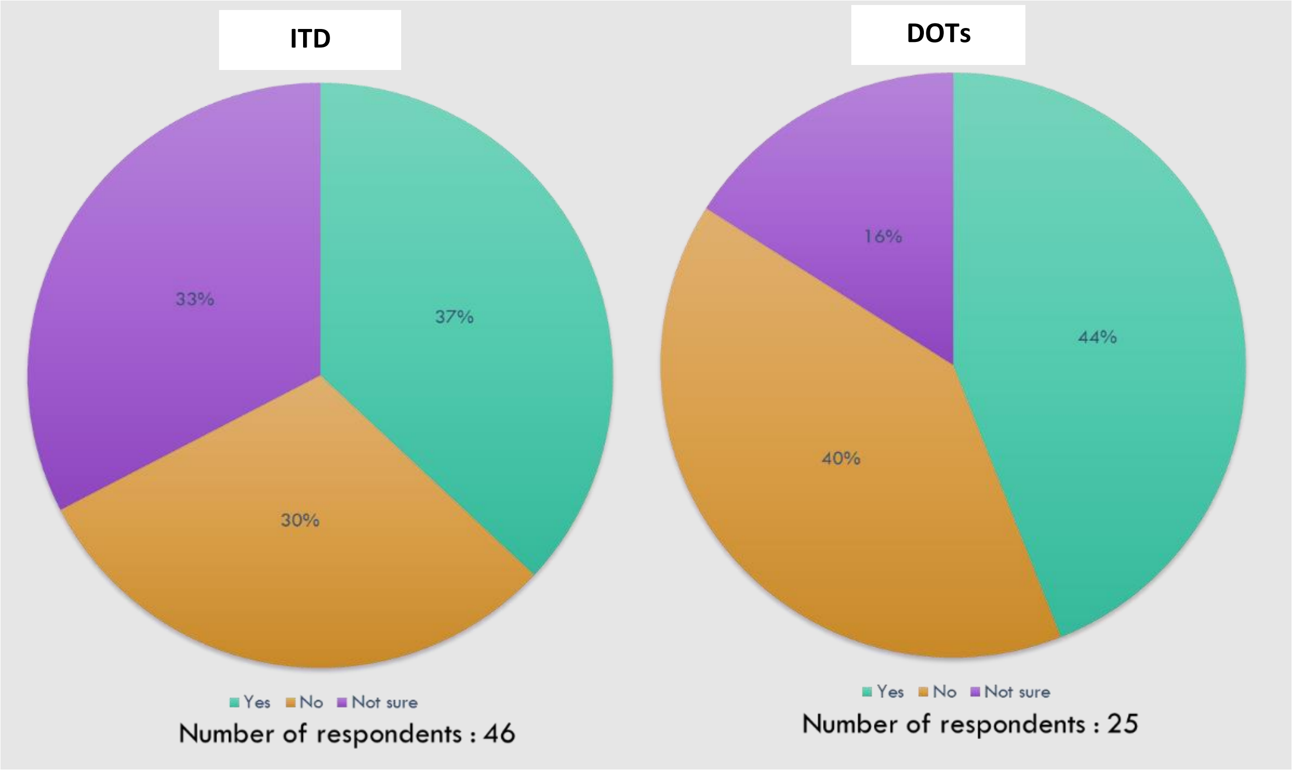


Figure 2.5 Implementation of procedures to compare actual service life of asphalt pavements against the original design life

- Q6) In your state, do asphalt pavements generally meet the original design life?

Among those that answered positively to the Q5, 64% of state DOT participants confirmed that asphalt pavements generally meet the original design life, while 28% asserted that asphalt pavements do not generally meet the original design life. However, in Idaho, a majority of participants (62%) stated that asphalt pavements do not meet the original design life, while only 10% believed that they do meet the design life (Figure 2.6). This highlights the need to further investigate the reasons why asphalt pavements do not meet the original design life and to investigate possible unnatural corrections and inconsistencies between the data that prime contractors reported and the data that ITD collected. We note that ITD engineers have participated in this study in larger numbers than that of other DOTs (47 in Idaho versus 25 in all other DOTs), which casts a question of whether or not other participants had a reservation to leave this question blank.

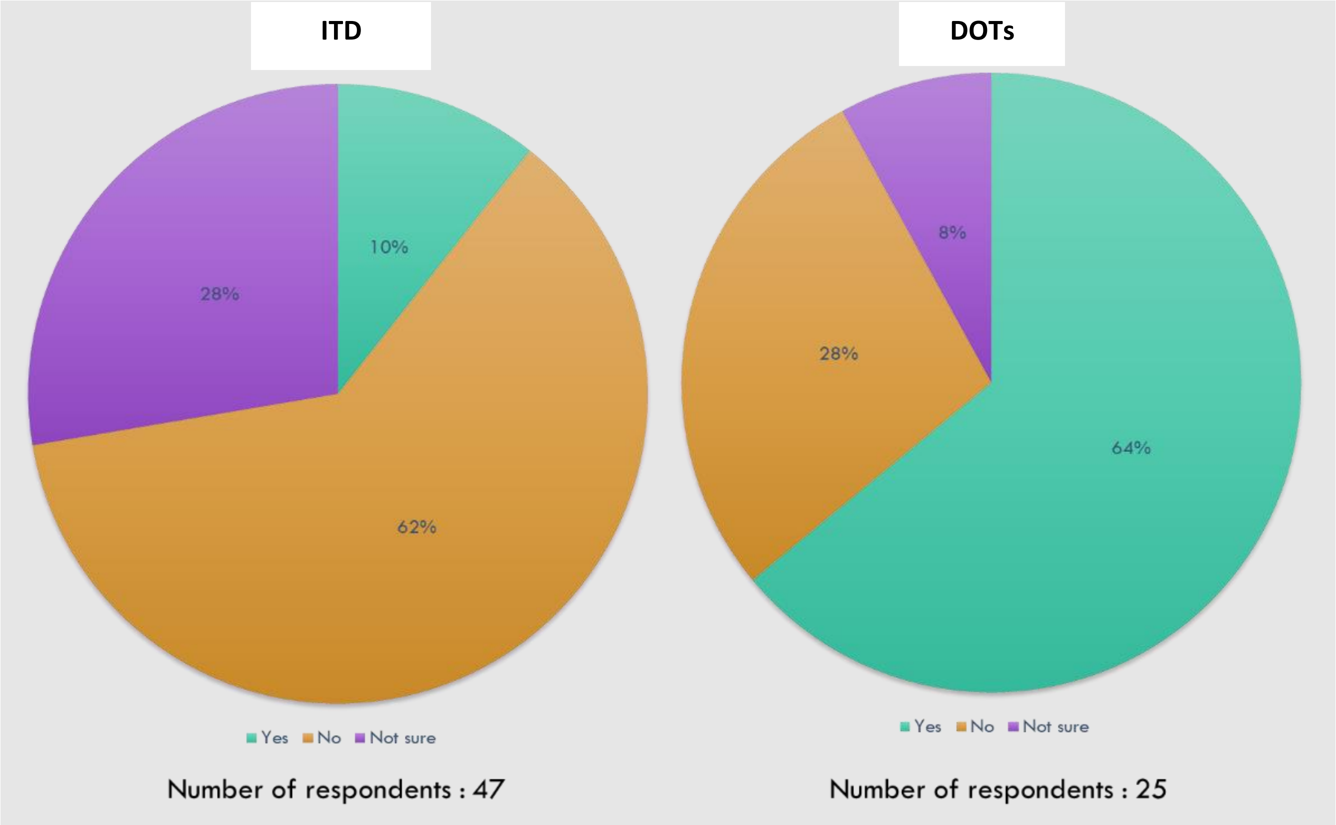


Figure 2.6 Answers to whether or not asphalt pavements generally meet the original design life

- Q7) Generally, what kind of discrepancies do you observe between the design life and service life of asphalt pavements in your state?

While the majority of answers (64%) from state DOTs show that service life is roughly equal or greater than the design life, only 6% of the answers from ITD participants correspond to having equal or greater service life. Breakdown of perception of service life of asphalt pavements as compared to their design life is shown in Figure 2.7. Similar to other questions, we note that a considerably lower number of participants from state DOTs participated in this survey, as compared to the ITD participants.

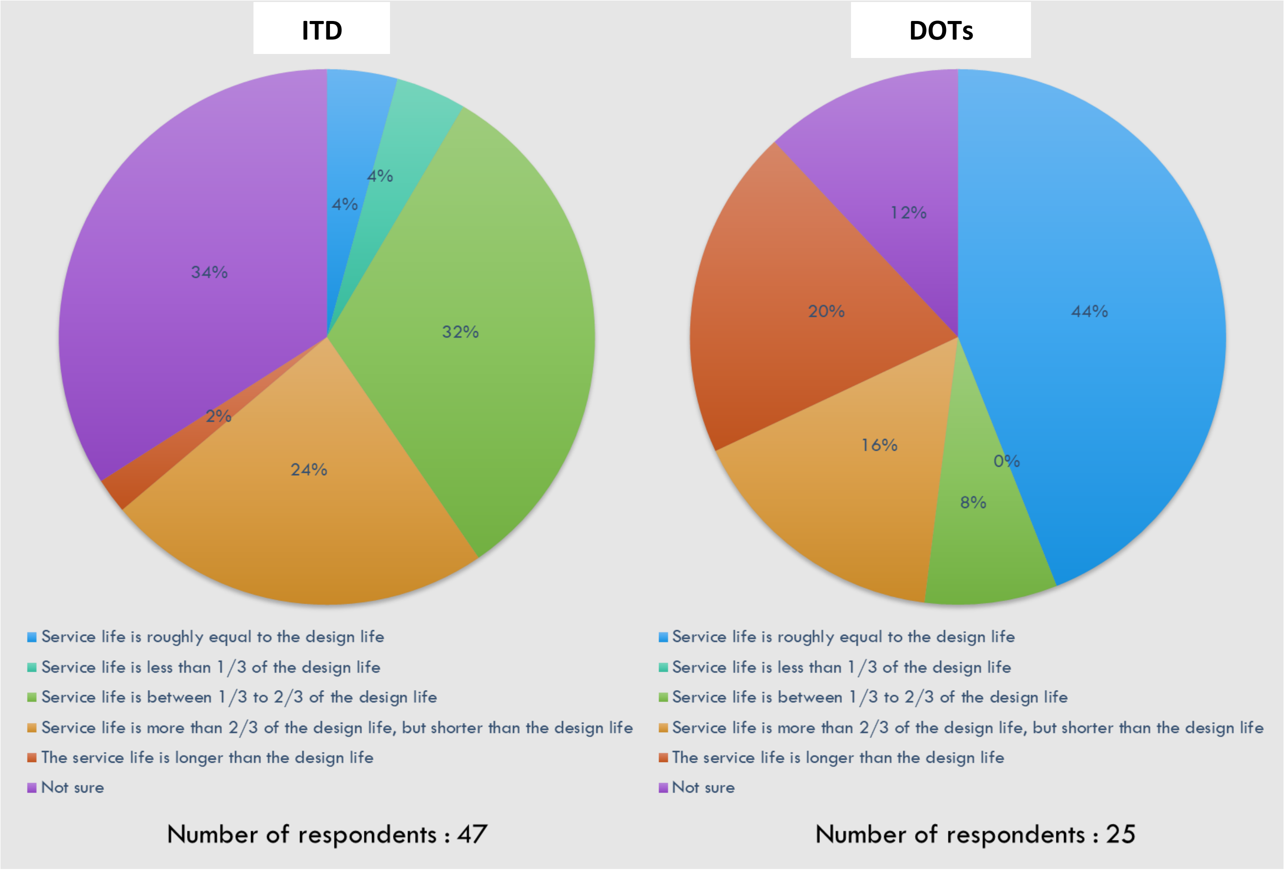


Figure 2.7 Type of discrepancies observed between the design life and service life of asphalt pavements

- Q8) What do you think is the cause for this discrepancy between the design life and the service life of asphalt pavements in your state? [mark all that apply]

Results from state DOTs show that the majority of the discrepancies are attributed to the traffic volume being underestimated during the design phase and using deficient construction materials, although other reasons were also selected in this question (Figure 2.8). Results from Idaho, however, indicate that observed discrepancies are mainly attributed, by ITD participants, to the usage of deficient construction materials.

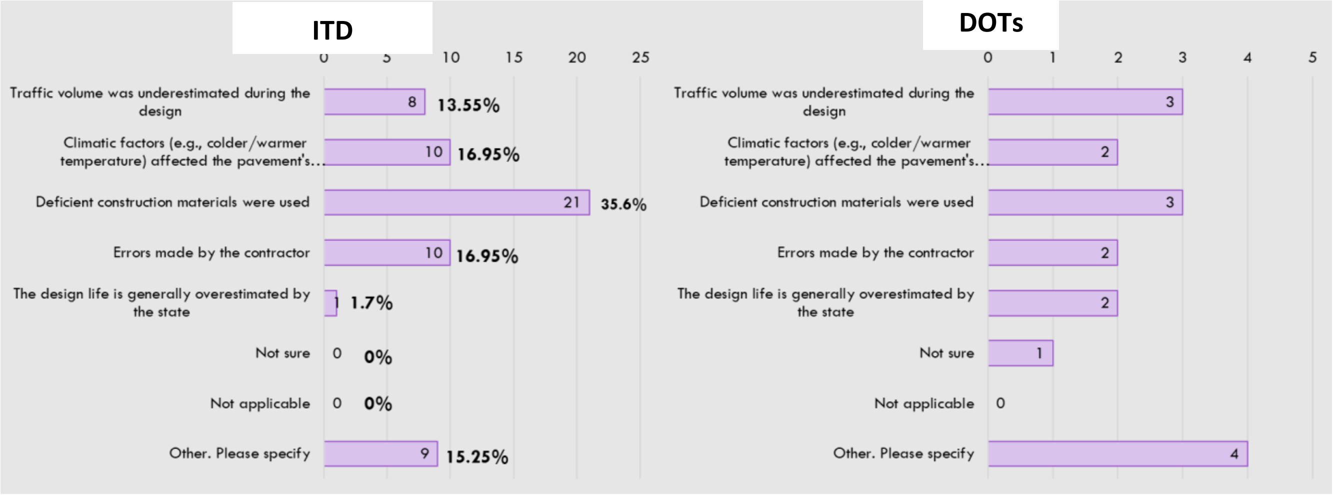


Figure 2.8 Cause for this discrepancy between the design life and the service life of asphalt pavements

- Q9) Who performs the acceptance testing (testing used for acceptance and payment) during the asphalt material production and paving? (select all that apply)

Figure 2.9 shows that the DOT and third party contracted by the DOT are the two major acceptance testing performers across the state DOTs. However, in Idaho the prime contractor and then a third party contracted by ITD are believed to be performing the tests. In Idaho, prime contractors and third party contracted by the prime contractor collectively account for 48% of acceptance testing cases, which is significantly higher compared to other state DOTs.

Figure 2.9 Party performing the acceptance testing during the asphalt material production and paving

- Q10) Is contractor QC test data used in your state DOT to determine contractor payment during the construction of asphalt pavements?

As demonstrated in Figure 2.10 more than half of answers from state DOTs and ITD (52% and 57%, respectively) indicate that the prime contractor QC test results are used to determine prime contractor payment but nevertheless, 44% and 30% of answers from state DOTs and ITD, respectively, show that QC results are not being used. Obviously, there is a discrepancy in engineers' perception in this case too, which points to a need for further training.

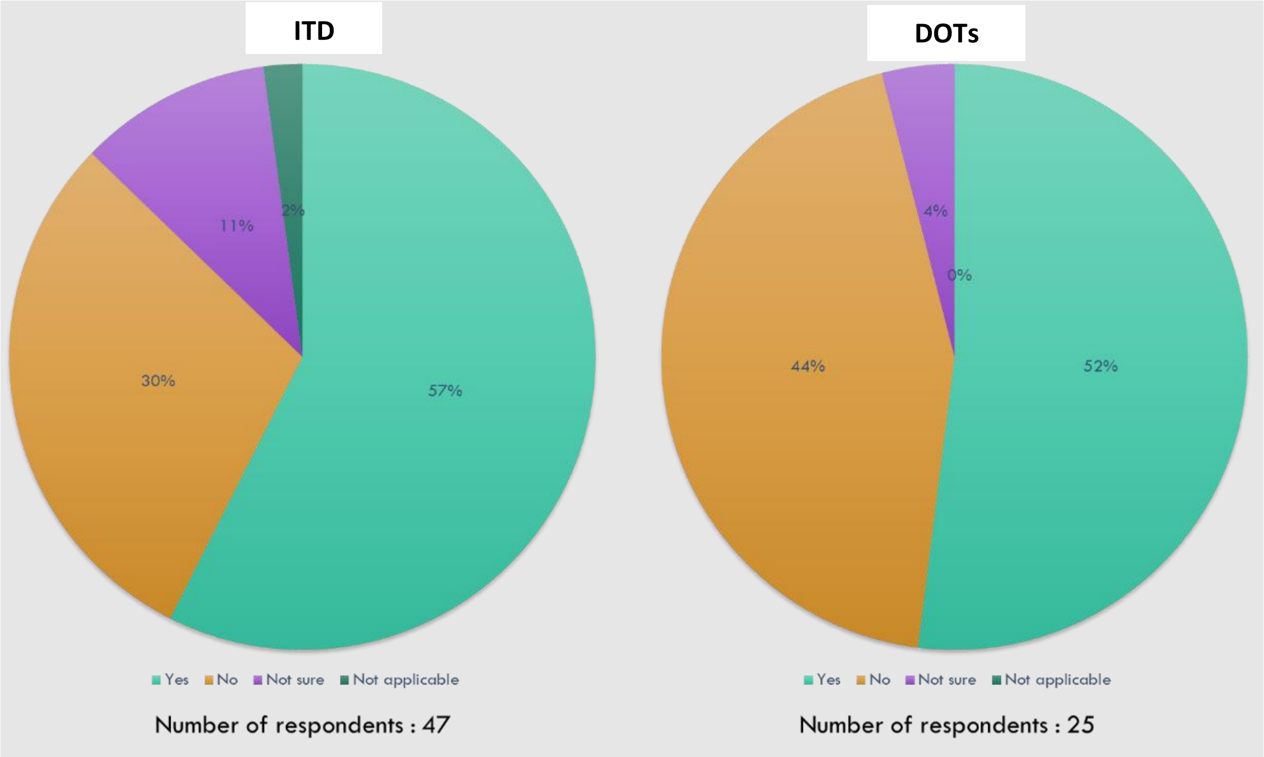


Figure 2.10 Whether or not contractor QC data is used to determine contractor payment during the construction of asphalt pavements

- Q11) Have you ever detected that the mix design and volumetric testing data reported prior to and during construction might not be representative of the actual material used in paving?

Results from state DOTs indicated that half of the participants believe the mix design and volumetric testing data reported prior to and during construction may not represent the actual material used in paving. This percentage is much larger in Idaho, where 63% of answers indicate reported data may not be representative of the actual materials used (Figure 2.11).

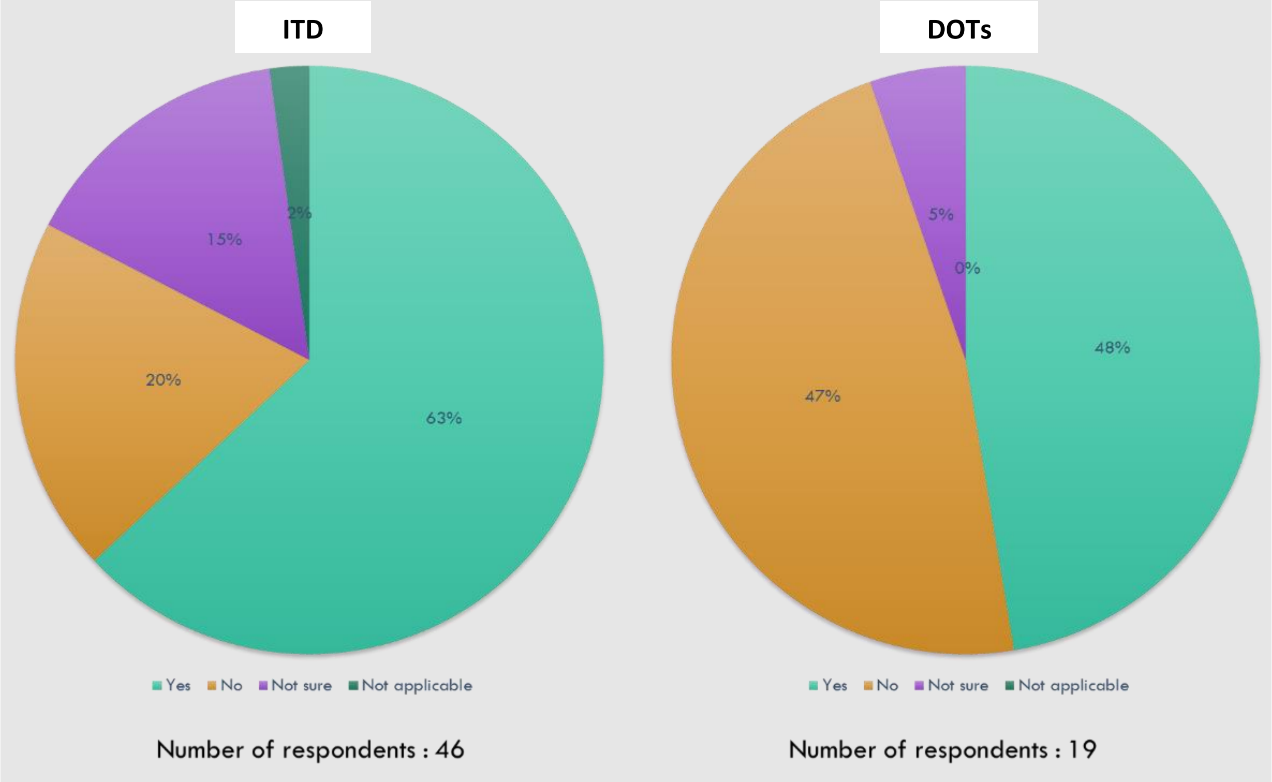


Figure 2.11 Whether or not employees ever detected that the mix design and volumetric testing data reported prior to and during construction might not be representative of the actual material used in paving

- Q12) What was the basis? [mark all that apply]

Both state DOTs and ITD results show that data analysis by the state DOT and personal observations were the basis for detection of mix design and volumetric testing data reported prior to and during construction not representing actual materials used in paving (Figure 2.12).

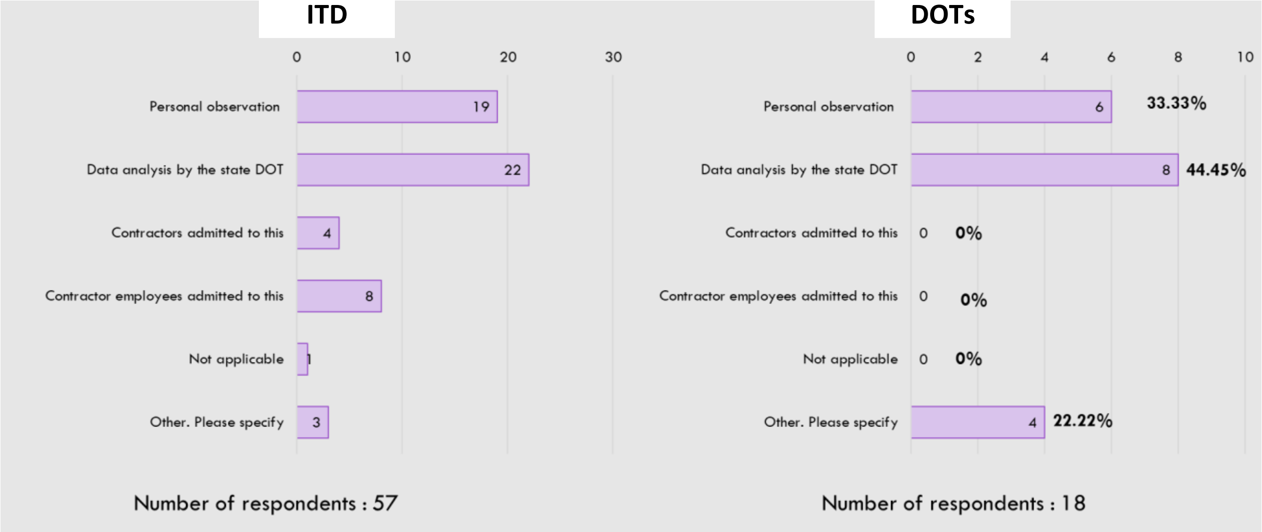


Figure 2.12 Basis for misrepresentation observed in Q11

- Q13) Which, if any, measures does your agency adopt to ensure the accuracy of mix design and volumetric testing data? [mark all that apply]

Both state DOTs and ITD results demonstrate that the use of independent assurance agents is a commonly used practice to ensure the accuracy of mix design and volumetric testing data (Figure 2.13). Also, state DOTs' results indicated a higher percentage of using "QA team works in conjunction with volumetric testing", as compared to the ITD results, to ensure the accuracy of mix design and volumetric testing data.

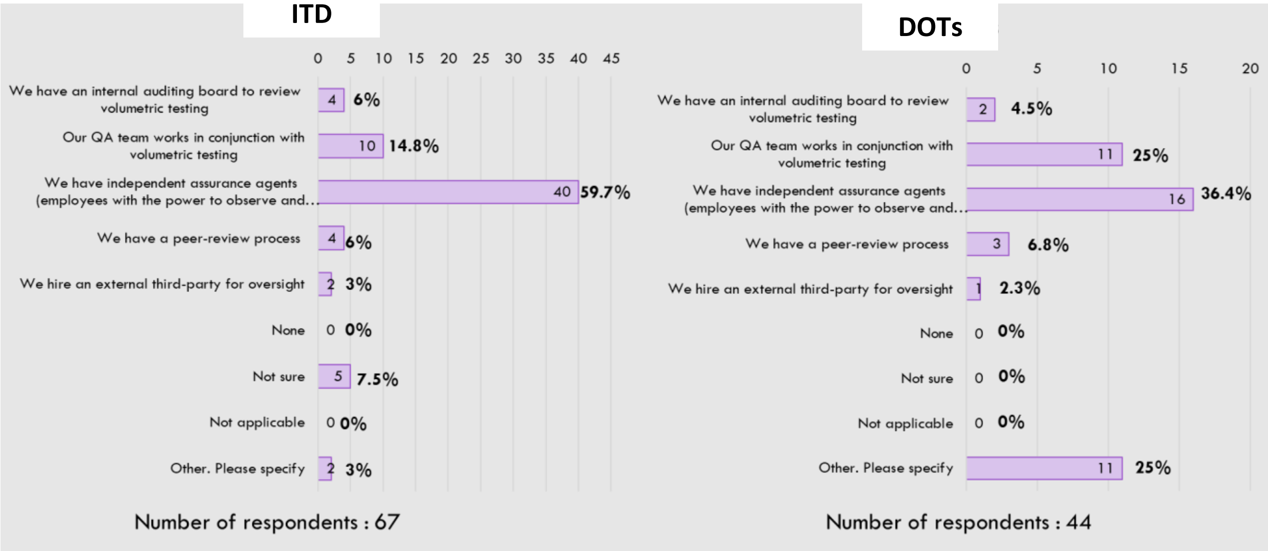


Figure 2.13 Measures adopted to ensure the accuracy of mix design and volumetric testing data

- Q14) Has your department ever attempted to detect/investigate potential manipulation of mix design and material testing data?

As demonstrated in Figure 2.14 half of the answers from the state DOTs indicated that they attempted to detect/investigate potential data alterations/corrections, whereas the majority (71%) of ITD participants believed that the agency attempted to detect/investigate potential data alterations/corrections. Idaho results are particularly interesting since only a small portion of answers (2%) show no action from ITD and the remaining participants are simply unaware. Interestingly, these findings show the same pattern to those in Q11 where the majority of answers showed that the mix design and volumetric testing data reported prior to and during the construction might not be representative of the actual material used in paving (Figure 2.11). This might point to a higher awareness of ITD engineers when it comes to detection of data alteration/correction in asphalt pavement projects.

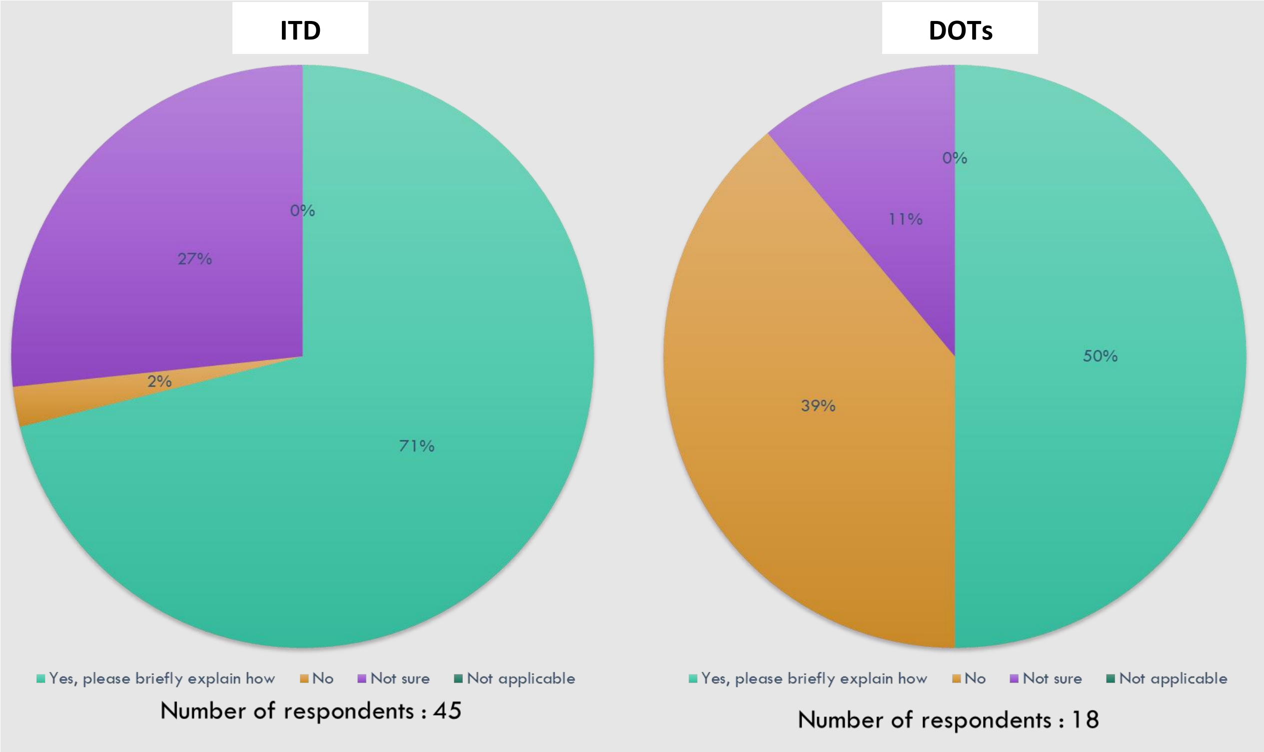


Figure 2.14 Whether or not agency attempted to detect / investigate potential manipulation of mix design and material testing data

- Q15) What are the parameters that are most likely to be manipulated while mix design and material testing data is reported for agency approval? Please note that the options presented below pertain to standard asphalt mix volumetric tests [mark all that apply]

This question tackles a very important issue of which parameters are being altered/corrected the most. Figure 2.15 shows that, for state DOTS, the percentage passing No. 200 sieve, submerged weight of puck in water, and submerged weight of bowl and sample are the most likely parameters to be altered/corrected, although a majority of participants believed there are other parameters that are frequently altered/corrected. For ITD, the submerged weight of bowl and sample, weight of puck at SSD condition, percent passing the No. 200 sieve are most likely parameters to be altered/corrected. Nevertheless, results from ITD and state DOTs indicate that all listed parameters are likely to be altered/corrected while material testing data is reported to the agency for approval and pavement.

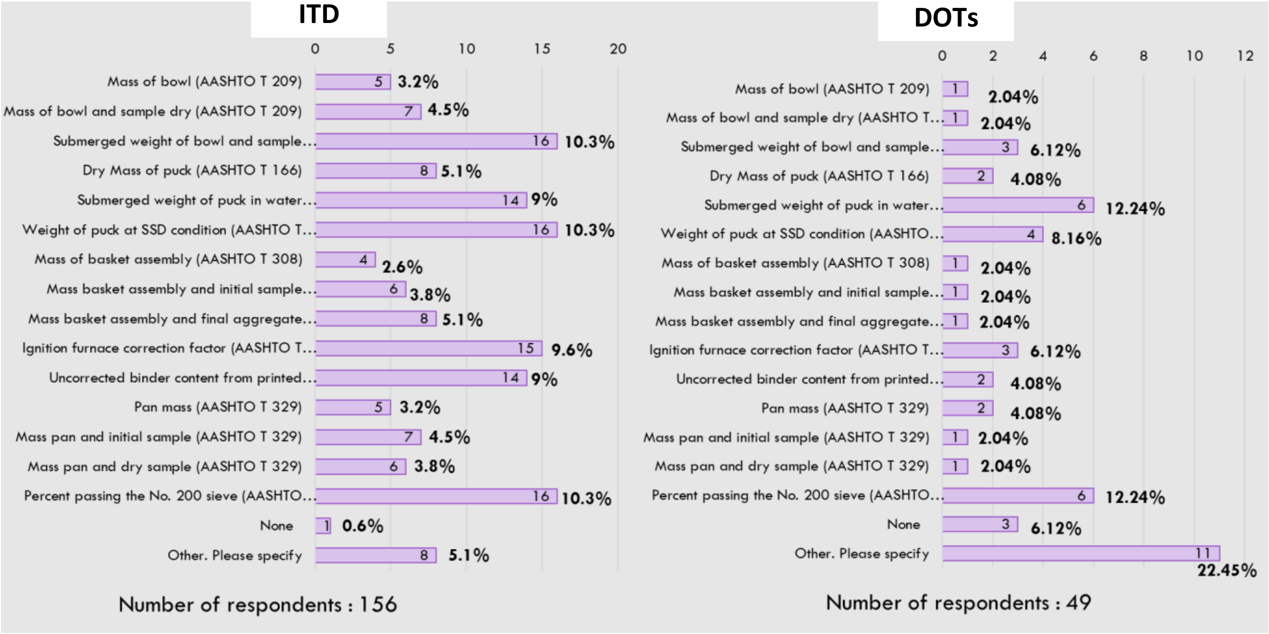


Figure 2.15 Most likely parameters to be altered in reported mix design and material testing data

- Q16) What reason do you think might explain possible manipulation of mix design and volumetric testing data? [mark all that apply]

Both ITD and state DOT results show that the most likely reason behind the data alteration/correction is the pressure to affect the payment factor in favor of the prime contractor (Figure 2.16). While in state DOT responses the next highly answered item is “not sure”, ITD participants are better informed of why data alteration/correction exists. The two other most likely reasons behind data alteration/correction in Idaho is to avoid scrutiny or conflict over results from the prime contractor and the unwillingness to reconduct the test. In other states, unwillingness to reconduct the test and “other reason” were most likely explanations for data alteration/correction. Least frequently selected reasons of why data is altered/corrected in both ITD and state DOT responses are “pressure to affect the payment factor in favor of the department” and “concerns regarding loss of testing qualifications”. The latter is an interesting finding given anecdotal evidence shows that lab technicians list losing qualification as a reason compelling them to alter/correct material test results.

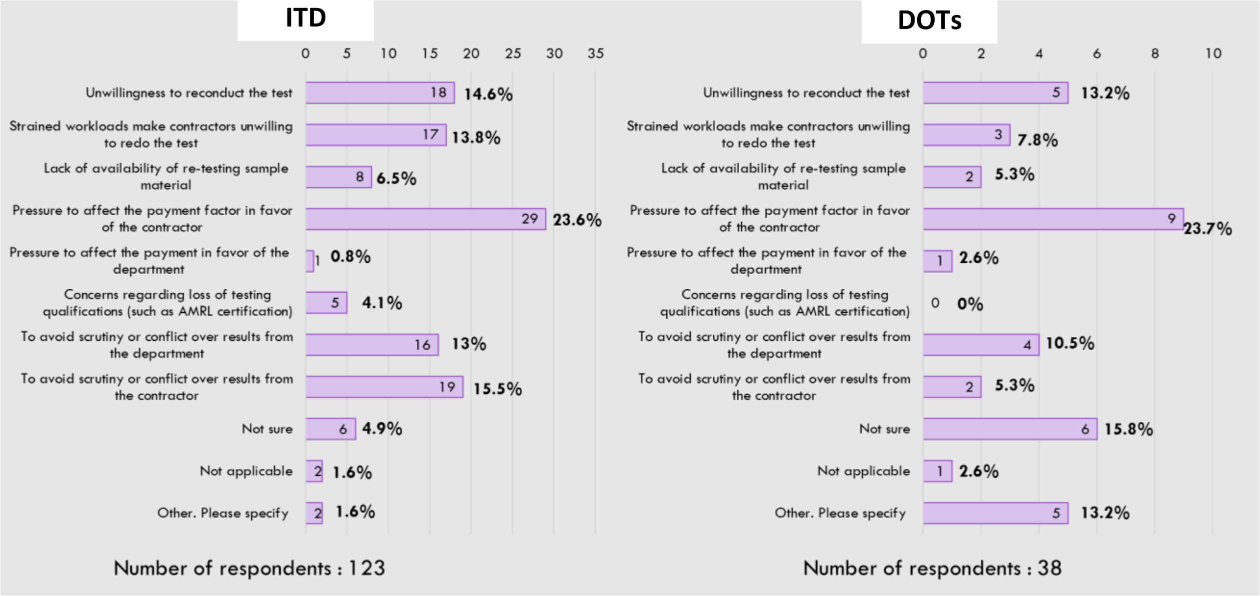


Figure 2.16 Reasons that employees think might explain possible alteration of mix design and volumetric testing data

- Q17) Does your agency have structured ethical frameworks (e.g., a Code of Conduct), and do they provide training to ensure employee comprehension and facilitate compliance?

Majority of the answers indicate that ITD and state DOTs do have a structured ethical framework and provide training to employees (Figure 2.17), however, the proportion of positive answers to this question in state DOTs is much higher than in Idaho (65% in other states compared to 43% in Idaho). This indicates the need for providing structured ethical frameworks (e.g., a Code of Conduct), and more frequent training for employees in Idaho to ensure their comprehension and facilitate compliance. We also point out that the number of ITD participants is twice as much as state DOTs.

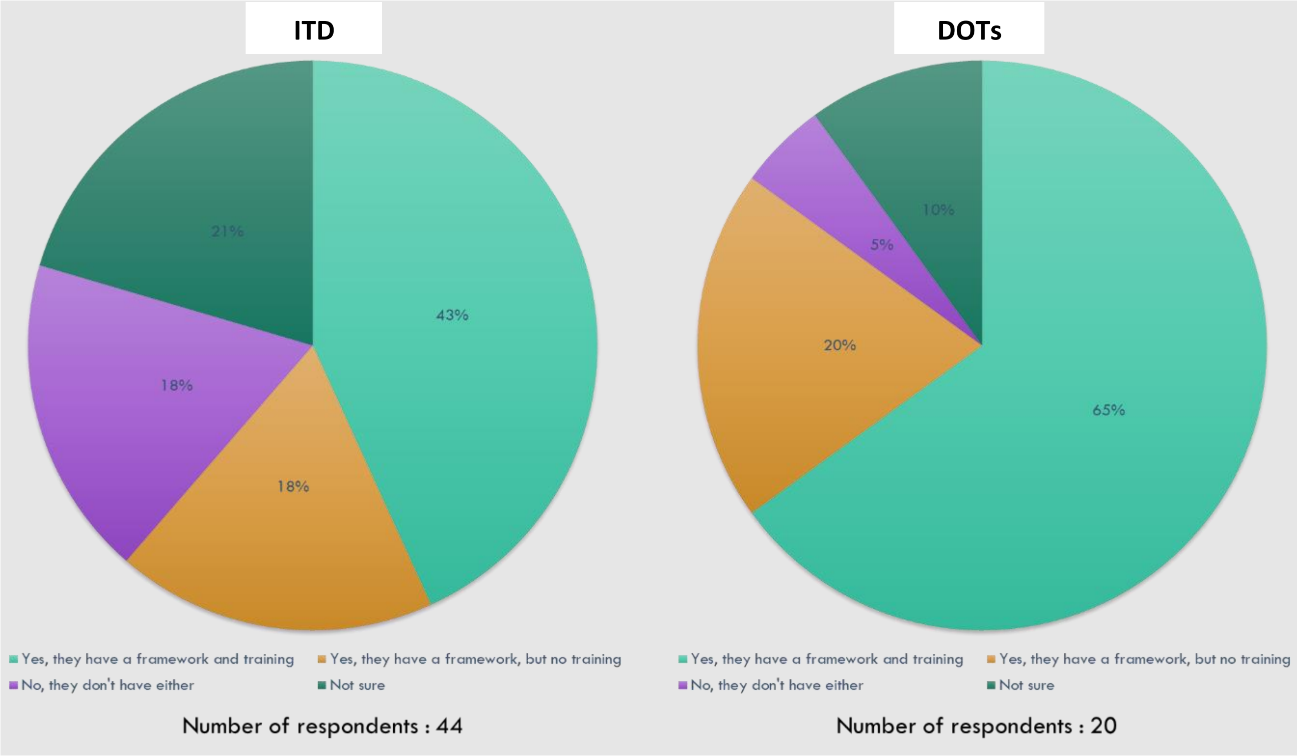


Figure 2.17 Whether or not agencies have structured ethical frameworks and provide training to ensure employee comprehension and facilitate compliance

- Q18) Do these ethical frameworks apply specifically to material testing? For instance, data or source material tampering, deviations from procedure while reporting written execution, purposely changing conditions between tests, etc.?

Of those who answered positively to Q17 (that their agencies have structured ethical frameworks and provide training to ensure employee comprehension and facilitate compliance), 74% of the ITD participants confirmed that the ethical frameworks do apply specifically to material testing while in state DOT responses a little more than half of answers (56%) indicated that the ethical frameworks do not specifically relate to material testing (Figure 2.18).

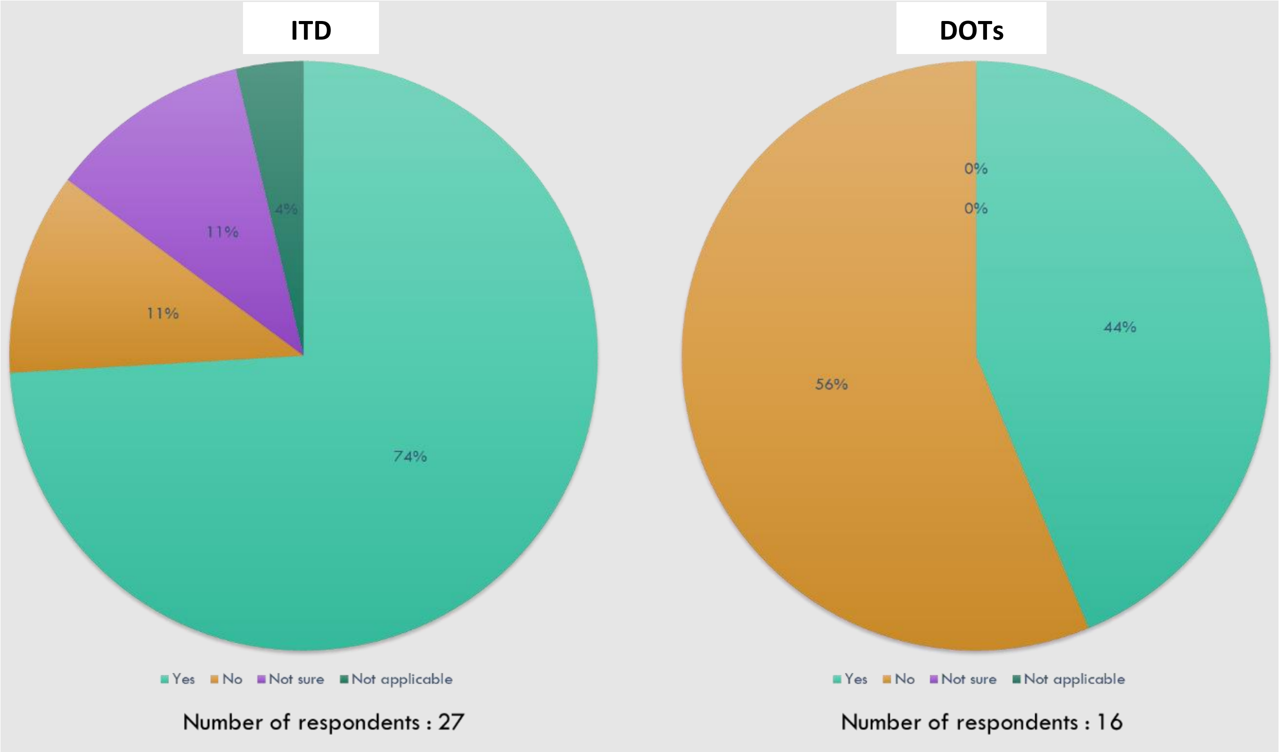


Figure 2.18 Answers to whether these ethical frameworks apply specifically to material testing or not

- Q19) If you ever observed an ethical violation in material testing, who did you report it to? [mark all that apply]

Results from state DOTs indicate that in case of observing ethical violation in material testing reports, employees most likely report to their manager or supervisor. In Idaho employees tend to report such instances to their supervisor, manager, or peer/colleagues as demonstrated in Figure 2.19.

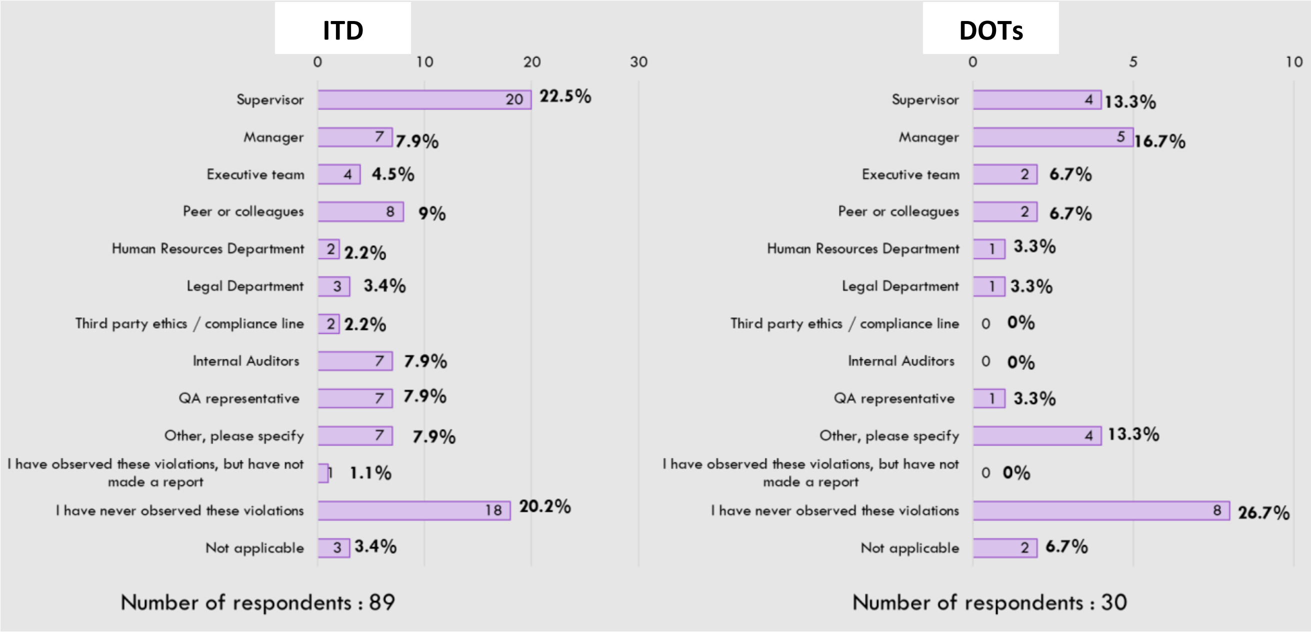


Figure 2.19 Parties to whom employees report the observed ethical violation in material testing

- Q20) If you observed but did not report ethical violations, why did you not make a report of the violations? [mark all that apply]

Only one employee from Idaho stated that he/she has observed an ethical violation in material testing and did not report this instance in Q19. Figure 2.20 shows this information, and demonstrates that the employee had other reasoning for not reporting (no additional information was provided).

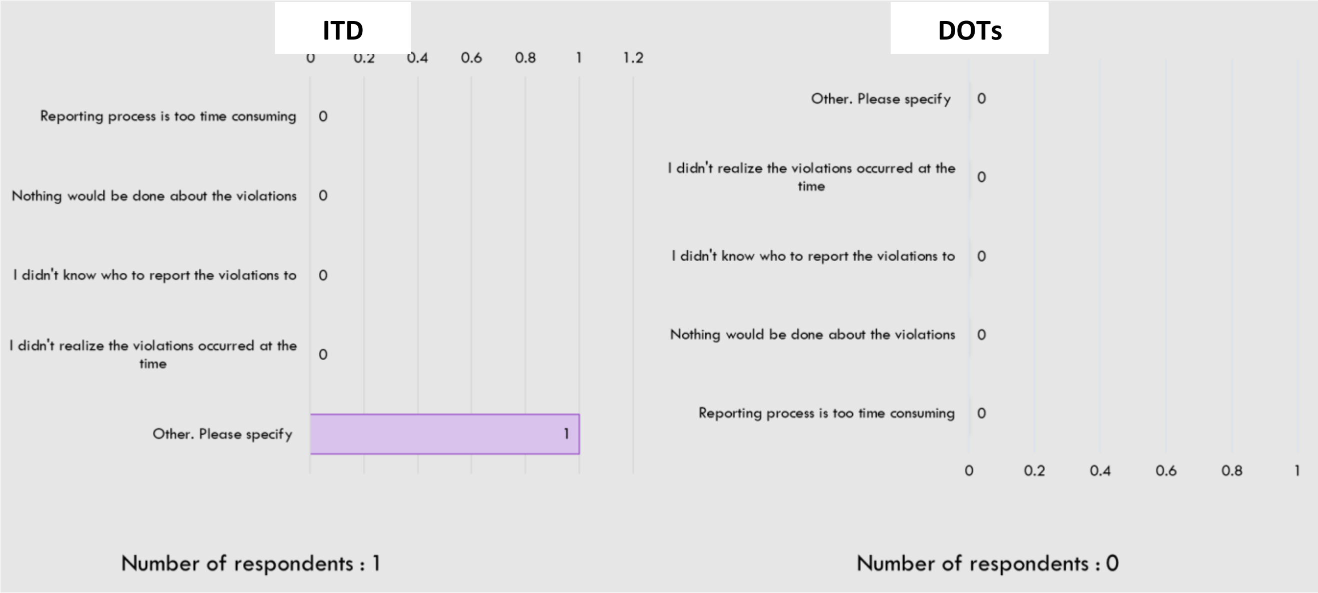


Figure 2.20 Reasons for not making a report of ethical violations

- Q21) Was your concerns taken seriously?

In Q19, if an employee selects one of the first 10 answers, this means that he/she had reported an instance of ethical violation in material testing and therefore are asked to respond to Q21. As shown in Figure 2.21, 80% of such reports were taken seriously in state DOTs, but in Idaho only 55% were taken seriously and around 14% were not considered serious by the person in charge. It is important to realize that these responses are associated with the perception of the participants, but nevertheless, this survey indicates that more emphasis and action should be taken to identify unexpected outcomes in Idaho.

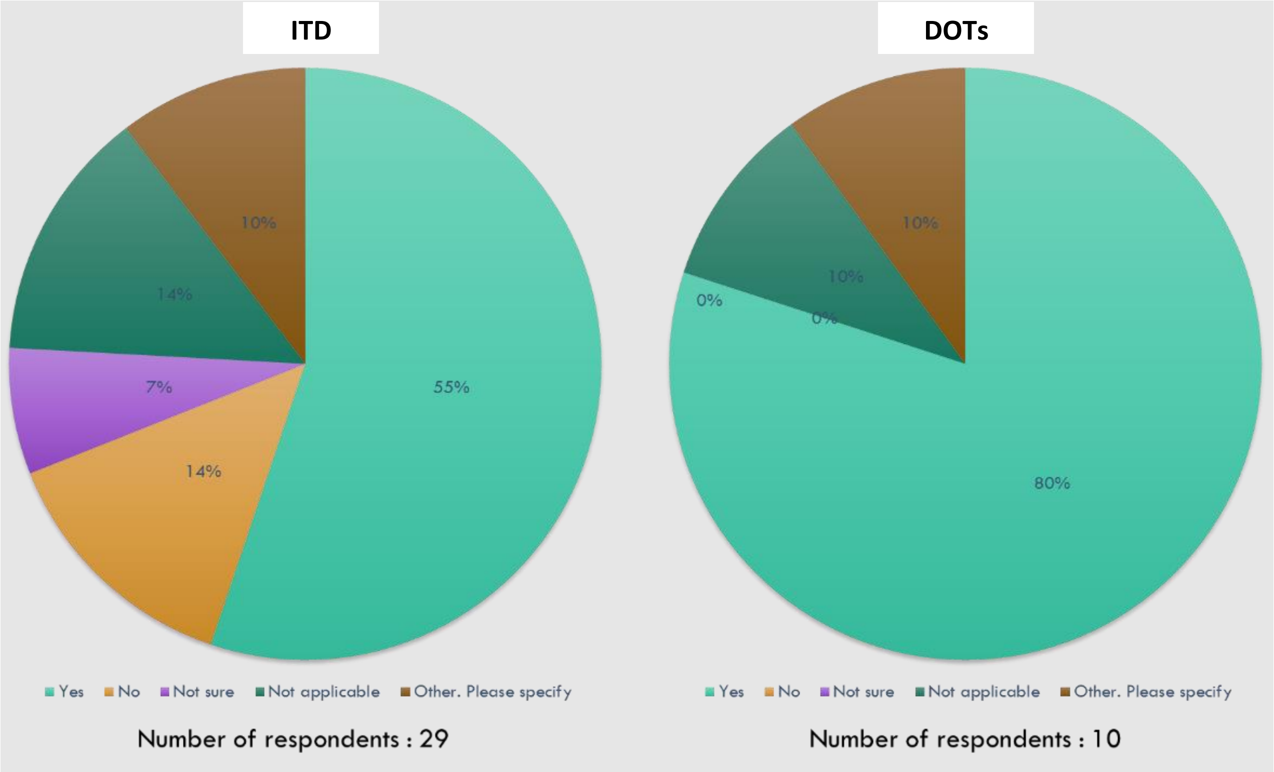


Figure 2.21 Whether or not reported employees' concerns were taken seriously

- Q22) Who was receptive to your concerns? [mark all that apply]

As shown in Figure 2.22, the answers here are almost identical to Q19 where, in state DOTs, employees most likely report such violations to their manager or supervisor and find them most receptive to the concerns; and in Idaho, employees tend to report such instances to their supervisor or peer/colleagues, who are most receptive to the employee concerns.

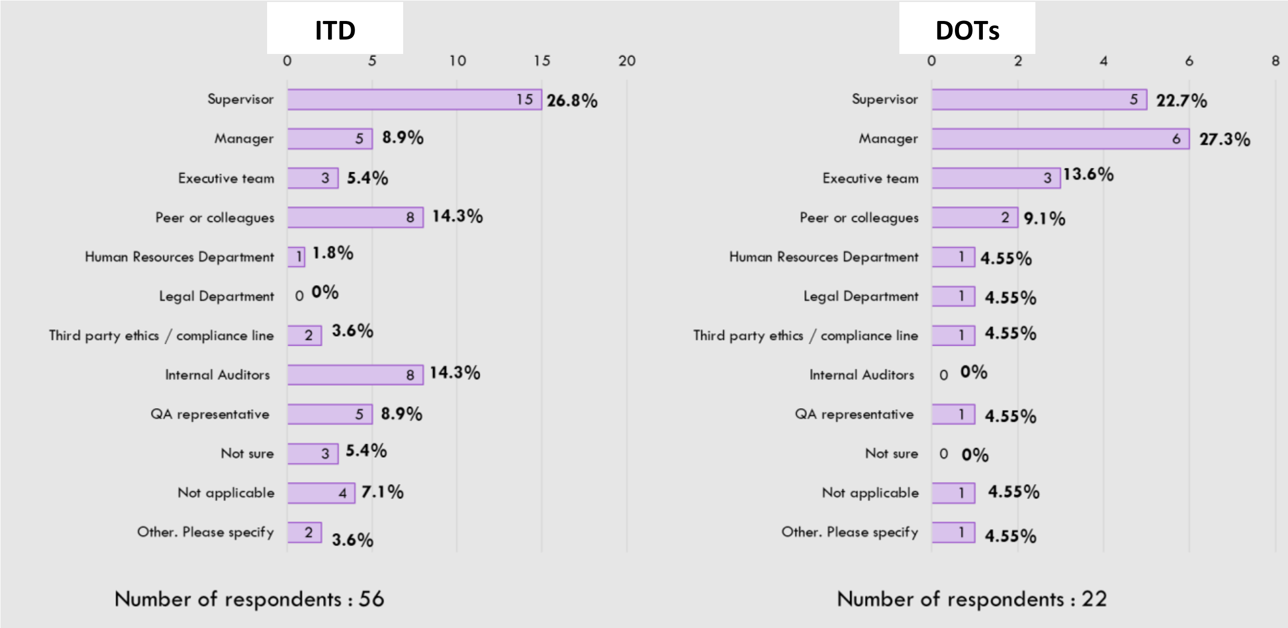


Figure 2.22 Parties receptive to employee concerns

- Q23) Irrespective to the testing laboratory (contractors and state), which option do you think is correct? [mark all that apply]

Interestingly when employees were asked about potential drivers for altering/correcting material testing data, a large proportion of state DOT employees were either unsure about the reason (37.5%) or they mentioned that this was not applicable to them (20.8%). Selection of these choices was much lower in Idaho (Figure 2.23) and the majority of ITD employees stated that either employees follow recommendation from the prime contractor to alter volumetric testing data or they do it on their own. Although in lower percentage, state DOT participants also selected these two choices.

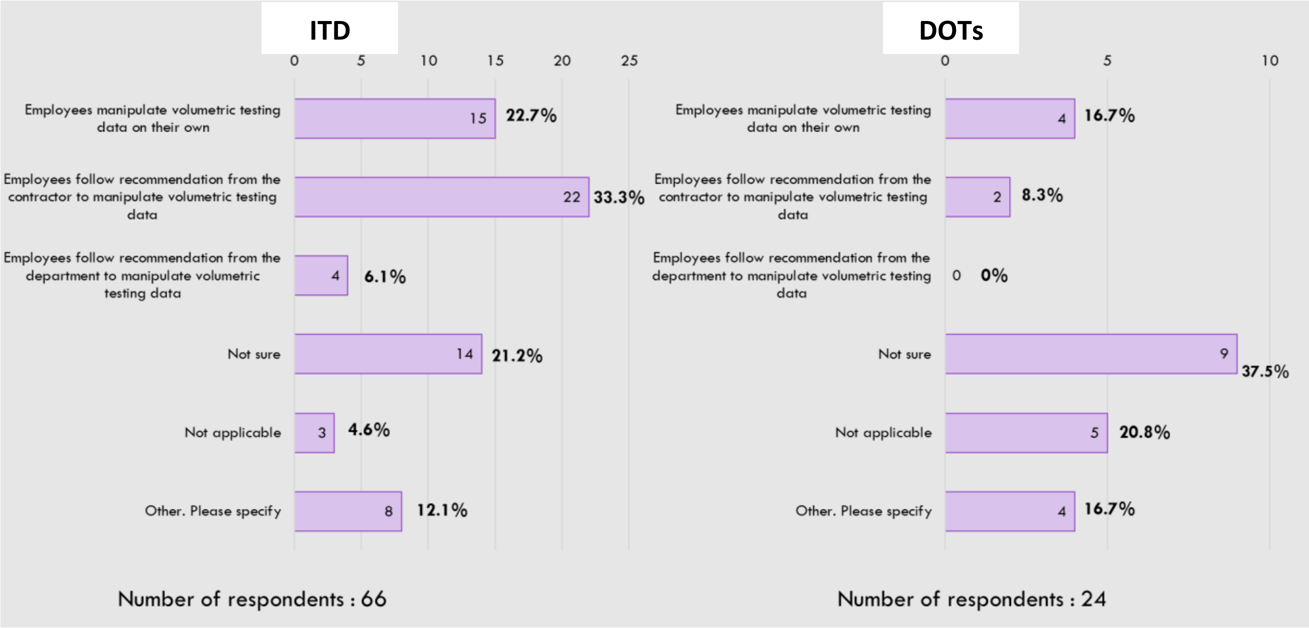


Figure 2.23 Answers to what employees think of how data alternation occurs

Discussions about ITD results and perception of ITD engineers

An in-depth and analytical survey of the ITD engineers and other DOT engineers about the state of practice and engineers' perception of the prevalence of suspicious activities in DOT-sponsored projects was conducted. Results of the survey demonstrated that ITD engineers' perception is that asphalt pavements most often do not meet their original design life in Idaho. More specifically, more than half ITD responses indicated that the service life of asphalt pavements in Idaho ranges between 1/3 of the original design life and shorter than the original design life. A majority of other state DOT engineers, however, indicate that their asphalt pavements do meet their original design life. We suggest caution in comparing the results between ITD survey and other state DOTs survey, as ITD engineers were most willing to participate in the study. We are certain that ITD engineers were more engaged in light of the recent federal investigation into materials testing activities in Idaho, among other reasons, but remain uncertain why other state DOTs did not participate to the level that can sustain confidence in the perceived trends. We note that this is a sensitive topic, and we can relate to the reservation of some engineers to share their beliefs. It is impossible for the research team to pinpoint the exact reason(s) for the corrections made.

The cause for the perceived discrepancy between the design life and the service life of asphalt pavements in Idaho is believed to have mainly stemmed from usage of deficient construction materials. In Idaho, survey results show that the prime contractor followed by a third party contracted by ITD performed the acceptance testing during the asphalt material production and paving. Results also indicated that the usage of prime contractor QC test data to determine prime contractor payment in Idaho is higher in comparison to other states. The use of prime contractor test results in Idaho, as opposed to the usage of transportation department test results in other states for acceptance testing, could have contributed to having shorter than the original design life. We note that ITD has changed its project payment and acceptance practice in 2020.

In both Idaho and other state DOTs, it was observed that the mix design and volumetric testing data reported prior to and during construction was not representative of the actual material used in paving (perception of participants and not necessarily facts). However, the proportion of positive answers (reported data not representing actual used materials) in Idaho was greater than other state DOTs. The bases of this belief in Idaho were data analysis by ITD and personal observations. This is in accordance with the basis of observation in all other states. To ensure the accuracy of mix design and volumetric testing data, different methods were used in both Idaho and all other states but the most common one was data analysis by the state DOT. We remain hesitant in comparing ITD results with other state DOTs, but survey results indicate potential higher occurrence of data alteration/correction cases in Idaho. We suspect that conducting more analyses and attempts to detect/investigate potential data alteration/correction cases in Idaho as compared to other states corresponds to the higher rates of perceived data alteration/correction cases in Idaho. This might be a matter of more awareness in Idaho, but further investigation is required to prove this theory right or wrong.

Survey results indicated that the submerged weight of bowl and sample, weight of puck at SSD condition, and percent passing the No. 200 sieve are the most likely parameters in Idaho to be altered/corrected when mix design and material testing data is reported for agency approval. Therefore, close scrutiny of these parameters by ITD can potentially identify data alteration/correction cases. In state DOT results, the percentage passing No. 200 sieve, submerged weight of puck in water, and submerged weight of bowl and sample are the most likely parameters to be altered/corrected, respectively. Both in Idaho and other states, the pressure to affect the payment factor in favor of the prime contractor was considered as the most likely reason to why data alteration/correction occurs. The two other most likely reasons behind data alteration/correction in Idaho was to “avoid scrutiny or conflict over results from the contractor” and the “unwillingness to reconduct the test”.

To ensure employee comprehension and facilitate compliance, both Idaho and other state DOTs have structured ethical frameworks (e.g. a Code of Conduct), and they provide training to employees. However, the proportion of employees responding positively to the existence of such frameworks and training in their department in other state DOTs is much higher than that in Idaho (65% in other state DOTs compared to 43% in Idaho). This emphasizes the need for more frequent training in Idaho to ensure employees comprehension and facilitate compliance. Further analysis shows that ITD ethical frameworks apply specifically to material testing as opposed to a majority of answers received from state DOTs indicating otherwise.

Results show that upon observation of ethical violations, almost all instances of violations are reported. Results from other state DOTs, indicate that in such cases, employees most likely report violations to their manager or supervisor. In Idaho, employees tend to report such instances to their supervisor or peers/colleagues. Interestingly, although a significant percentage of these cases (80% of such reports) were taken seriously in other state DOTs, in Idaho, only 55% were perceived to have been taken seriously and around 14% were perceived to have not been considered seriously by the person in charge. We caution against comparison of ITD results against those of other state DOTs, but these findings in Idaho alone provide a clear need for further training and developing structured frameworks nd protocols to address employee concerns. Also, a majority of ITD respondents stated that either employees follow recommendation from the prime contractor to alter/correct volumetric testing data or they do it on their own. These results provide in-depth information about the current practice at ITD, and help develop a consistent, rigorous QA platform across the state. This survey will also create a baseline to design further training activities for ITD engineers.

3. Prevalence of Data Changes in ITD HMA Paving Projects

Introduction

This chapter presents details on the development of a logic-based algorithms to learn the patterns in the audit trail of the material test results for several HMA projects that classify the observed data corrections into Plausible Correction (P.C.) and Unexplained Correction (U.C.) and presents and discusses the results all with the goal of increasing the confidence in the data reported. To avoid inherent bias, this report uses "Entity 1" and "Entity 2" to refer to ITD and prime contractor reported data, not necessarily in any particular order. In other words, it has not been disclosed to the reader whether Entity 1 represents data from the agency or prime contractor. The same is the case for Entity 2. It is important to note that each entity may hire a third-party contractor to do materials testing on its behalf.

Objectives

The objective of this research was to classify HMA construction audit trail data into green (P.C.) or red (U.C.) zone. Probable data changes may lead to not only financial losses but also poorly paved roadways. Alongside the classification task, the possible monetary loss associated with unexplained changes of material testing report data in multiple HMA projects across Idaho was calculated, the results of which are presented in Chapter 4.

Research goals in this chapter include:

i. Analysis of audit trail data for manual detection of patterns in material testing data reports
ii. Development of a logic-based algorithm to classify multiple data entries for each parameter in construction projects' audit trail data to P.C. and U.C.
iii. Summarizing the extent of data changes in projects with available audit trail data

Research tasks carried out to accomplish the above described goals were:

1. Data organization and cleaning: the spreadsheets including the audit trail data were sent to the research team from ITD. The audit trail recorded all instances for data entry from material testing results. The dataset was large in volume and needed proper "organizing" before the application of logic-based algorithms. The data was organized and imported to a MATLAB computer program for further analysis. Additionally, more data (test summary, lot information, volume of material, etc.) was organized or the later part of the analysis, which we will present in Chapter 4.
2. Development of algorithmic logics: At first, one project data was examined manually to untangle the general trend of data correction. This resulted in several cases of probable typing errors as well as some cases of unexplained changes. Subsequently, more projects were manually

analyzed to see if such patterns exist in different projects, and if there are other patterns in the corrected data. Later, these findings were converted to if/else cases and assembled to an algorithm to detect similar cases for all projects.

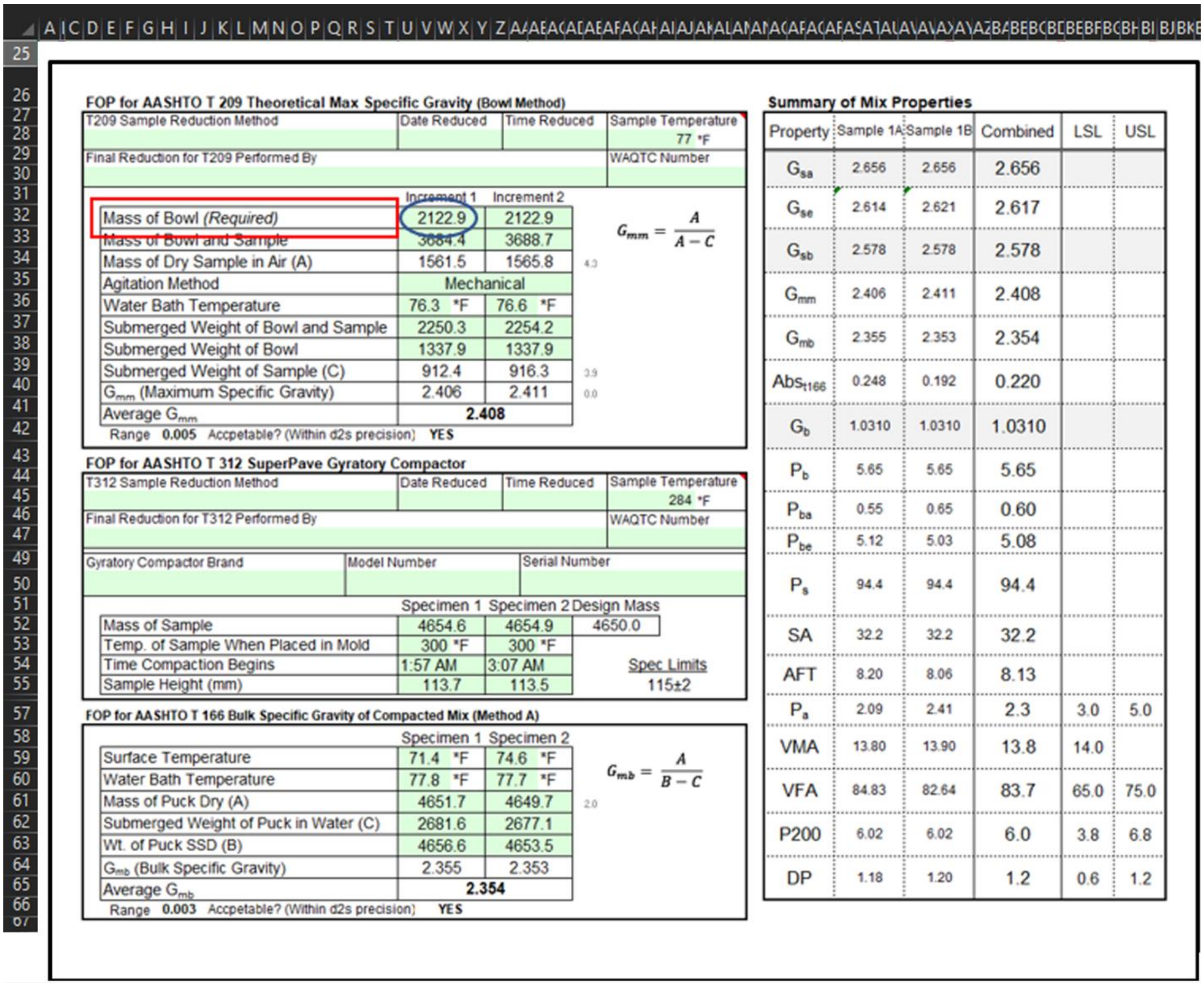
Audit Trail Data and Correction Monitoring

A forensic case study by FHWA published in 2018 highlighted inconsistencies between mix design parameter data reported by prime contractors as compared to those collected by ITD (see Appendix A). From this forensic study there were unexpected findings. This prompted ITD staff to determine why results were unexpected. The ITD staff, without specifically being sanctioned by ITD, encoded a VBA (Visual Basic for Applications) macro into the Excel materials testing reports to register all data entries for certain key parameters in each test. All material testing data were captured through these Excel files, which also recorded a sequence of data corrections for certain parameters. This practice provided extensive and invaluable information about data corrections in material testing reports.

This quality assurance audit trail data presented the opportunity for the research team to review all the corrections for any reported parameter. While various reasons can explain data corrections, and the research goal is not to investigate or to determine the reasons behind data corrections, but the worst-case scenario corresponds to the situation where data changes directly increased pay factors or helped pass substandard quality work.

Quality Assurance Audit Trail Data

The material testing spreadsheets including the audit trail data were acquired from ITD for 15 HMA projects completed in Idaho during the year 2018. These spreadsheets consisted of an internal audit trail macro to record the sequences of changing parameter values in the background (not visible to the operator). Figure 3.1 shows a screenshot of a typical data input file to record material testing data. For example, if an operator inputs the value (2122.9 in this case) for Mass of bowl (red box) for increment 1 (blue circle), that value is recorded under U32 (corresponding cell number for mass of bowl (increment 1) in the Excel file). If the operator deletes the value (2122.9) and registers a new value (2500 for example) both values are registered under U32 in the audit trail with the corresponding time stamp.



FOP for AASHTO T 209 Theoretical Max Specific Gravity (Bowl Method)

T209 Sample Reduction Method	Date Reduced	Time Reduced	Sample Temperature 77 *F
Final Reduction for T209 Performed By			WAQTC Number

	Increment 1	Increment 2	
Mass of Bowl *(Required)*	2122.9	2122.9	
Mass of Bowl and Sample	3684.4	3688.7	
Mass of Dry Sample in Air (A)	1561.5	1565.8	4.3
Agitation Method	Mechanical		
Water Bath Temperature	76.3 *F	76.6 *F	
Submerged Weight of Bowl and Sample	2250.3	2254.2	
Submerged Weight of Bowl	1337.9	1337.9	
Submerged Weight of Sample (C)	912.4	916.3	3.9
G_{mm} (Maximum Specific Gravity)	2.406	2.411	0.0
Average G_{mm}	**2.408**		

$$G_{mm} = \frac{A}{A-C}$$

Range **0.005** Accpetable? (Within d2s precision) **YES**

FOP for AASHTO T 312 SuperPave Gyratory Compactor

T312 Sample Reduction Method	Date Reduced	Time Reduced	Sample Temperature 284 *F
Final Reduction for T312 Performed By			WAQTC Number

Gyratory Compactor Brand	Model Number	Serial Number

	Specimen 1	Specimen 2	Design Mass
Mass of Sample	4654.6	4654.9	4650.0
Temp. of Sample When Placed in Mold	300 *F	300 *F	
Time Compaction Begins	1:57 AM	3:07 AM	Spec Limits
Sample Height (mm)	113.7	113.5	115±2

FOP for AASHTO T 166 Bulk Specific Gravity of Compacted Mix (Method A)

	Specimen 1	Specimen 2	
Surface Temperature	71.4 *F	74.6 *F	
Water Bath Temperature	77.8 *F	77.7 *F	
Mass of Puck Dry (A)	4651.7	4649.7	2.0
Submerged Weight of Puck in Water (C)	2681.6	2677.1	
Wt. of Puck SSD (B)	4656.6	4653.5	
G_{mb} (Bulk Specific Gravity)	2.355	2.353	
Average G_{mb}	**2.354**		

$$G_{mb} = \frac{A}{B-C}$$

Range **0.003** Accpetable? (Within d2s precision) **YES**

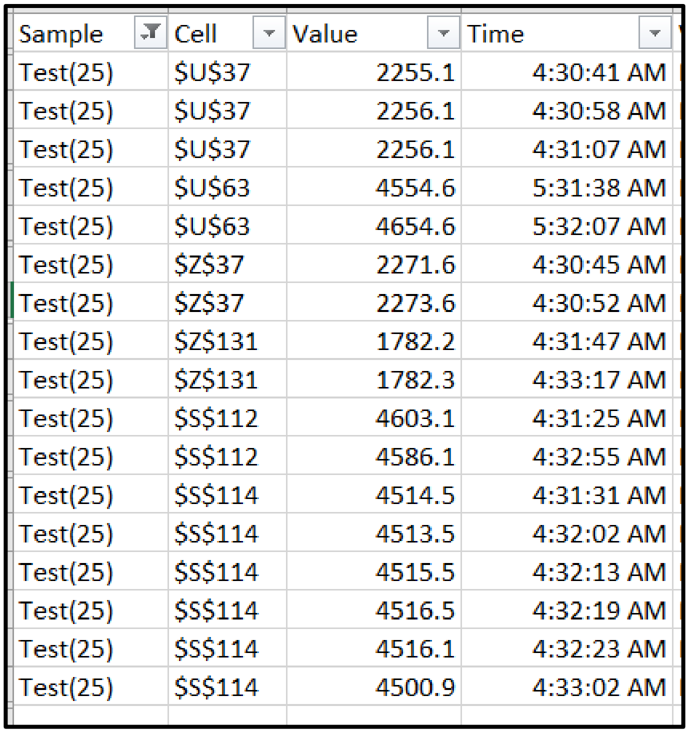
Summary of Mix Properties

Property	Sample 1A	Sample 1B	Combined	LSL	USL
G_{sa}	2.656	2.656	2.656		
G_{se}	2.614	2.621	2.617		
G_{sb}	2.578	2.578	2.578		
G_{mm}	2.406	2.411	2.408		
G_{mb}	2.355	2.353	2.354		
Abs_{t166}	0.248	0.192	0.220		
G_b	1.0310	1.0310	1.0310		
P_b	5.65	5.65	5.65		
P_{ba}	0.55	0.65	0.60		
P_{be}	5.12	5.03	5.08		
P_s	94.4	94.4	94.4		
SA	32.2	32.2	32.2		
AFT	8.20	8.06	8.13		
P_a	2.09	2.41	2.3	3.0	5.0
VMA	13.80	13.90	13.8	14.0	
VFA	84.83	82.64	83.7	65.0	75.0
P200	6.02	6.02	6.0	3.8	6.8
DP	1.18	1.20	1.2	0.6	1.2

Figure 3.1 Audit trail file to record material testing data

The dataset has several unique characteristics:

i. Material test reporting Excel files had a VBA script embedded, which had a unique ability to record each data entry typed in the Excel sheet. This develops a chronological record of all values entered into the spreadsheet in the form of an audit trail. Inspection of this audit trail can give a clear picture of how the test results were recorded. Figure 3.2 shows a screenshot of the audit trail file for one of the projects. Note that all identifying information, such as project name, test date, test time, testing lab, among others, have been removed from the figures in this manuscript to ensure the anonymity of the testing/reporting entities.



Sample	Cell	Value	Time
Test(25)	U37	2255.1	4:30:41 AM
Test(25)	U37	2256.1	4:30:58 AM
Test(25)	U37	2256.1	4:31:07 AM
Test(25)	U63	4554.6	5:31:38 AM
Test(25)	U63	4654.6	5:32:07 AM
Test(25)	Z37	2271.6	4:30:45 AM
Test(25)	Z37	2273.6	4:30:52 AM
Test(25)	Z131	1782.2	4:31:47 AM
Test(25)	Z131	1782.3	4:33:17 AM
Test(25)	S112	4603.1	4:31:25 AM
Test(25)	S112	4586.1	4:32:55 AM
Test(25)	S114	4514.5	4:31:31 AM
Test(25)	S114	4513.5	4:32:02 AM
Test(25)	S114	4515.5	4:32:13 AM
Test(25)	S114	4516.5	4:32:19 AM
Test(25)	S114	4516.1	4:32:23 AM
Test(25)	S114	4500.9	4:33:02 AM

Figure 3.2 Screenshot of the audit trail file showing data corrections in the spreadsheet

ii. Audit trail data was available for both QC as well as acceptance tests. In other words, records of data entries were available for certain projects irrespective of whether the tests were performed by the prime contractor (or a third-party testing laboratory hired by the prime contractor) or ITD (or a third-party testing laboratory hired by ITD). As already mentioned, the primary objective of the current research was to study the data correction patterns during HMA QC and acceptance testing. The discussions in this manuscript do not focus on whether the data corrections were carried out by representatives of the prime contractor or ITD.

iii. All parameters that would affect the payments of each project were also provided, which are listed in Table 3.1. There is a total of 27 different parameters that affect payment. They are categorized into three different categories (major/moderate/minor).

iv. Payment affecting parameters are similar for both department and prime contractor-reported data (Table 3.1). However, parameters that affect Density are only reported by the ITD data. Those parameters are enlisted in Table 3.2. These parameters are monitored by ITD to decide whether or not a particular asphalt mix meets specifications (VMA and Air Voids), and also whether or not a constructed pavement section has been adequately compacted (main line density). Reading 1 and 2 and Device Used are reported more than one time for each lot. So, if there are 2 tests in lot 1, then for reading 1, test 1 and 2 values would be registered in cell AC37 and AC38, respectively. Basically, there are only three parameters (Reading 1 and 2, Device used) in the density-related data.

v. Total number of material testing parameters (department/prime contractor/density) is summarized in Table 3.3.

Table 3.1 Material Testing Parameters and Their Impacts on Pay Factor Related Parameters

Cell Description	Voids in the Mineral Aggregate (VMA)	Air Voids	Density	Major/ Moderate/ Minor Effect
Mass of Bowl (Increment 1) (U32)	Yes	Yes	Yes	Major
Mass of Bowl and Sample Dry (Increment 1) (U33)	Yes	Yes	Yes	Major
Submerged Weight of Bowl and Sample (Increment 1) (U37)	Yes	Yes	Yes	Major
Submerged Weight of Bowl (Increment 1) (U38)	Yes	Yes	Yes	Major
Mass of Bowl (Increment 2) (Z32)	Yes	Yes	Yes	Major
Mass of Bowl and Sample Dry (Increment 2) (Z33)	Yes	Yes	Yes	Major
Submerged Weight of Bowl and Sample (Increment 2) (Z37)	Yes	Yes	Yes	Major
Submerged Weight of Bowl (Increment 2) (Z38)	Yes	Yes	Yes	Major
Mass of Puck Dry (Specimen 1) (U61)	Yes	Yes	No	Major
Submerged Weight of Puck in Water (Specimen 1) (U62)	Yes	Yes	No	Major
Weight of Puck SSD (Specimen 1) (U63)	Yes	Yes	No	Major
Mass of Puck Dry (Specimen 2) (Z61)	Yes	Yes	No	Major
Submerged Weight of Puck in Water (Specimen 2) (Z62)	Yes	Yes	No	Major
Weight of Puck SSD (Specimen 2) (Z63)	Yes	Yes	No	Major
Mass Basket Assembly (S111)	Yes	No	No	Moderate
Mass Basket Assembly & Initial Sample (S112)	Yes	No	No	Moderate
Mass Basket Assembly & Final Aggregate (S114)	Yes	No	No	Moderate
Ignition Furnace Correction Factor (S116)	Yes	No	No	Moderate
Calibration Factor (AP114)	Yes	No	No	Moderate
Uncorrected Binder Content (AP115)	Yes	No	No	Moderate
Pan Mass (N128)	Yes	No	No	Minor
Mass Pan and Initial Sample (N129)	Yes	No	No	Minor
Drying Cycle 1 Mass Pan and Sample (Z129)	Yes	No	No	Minor
Drying Cycle 2 Mass Pan and Sample (Z130)	Yes	No	No	Minor
Drying Cycle 3 Mass Pan and Sample (Z131)	Yes	No	No	Minor
Drying Cycle 4 Mass Pan and Sample (Z132)	Yes	No	No	Minor
Drying Cycle 5 Mass Pan and Sample (Z133)	Yes	No	No	Minor

Table 3.2 Material Testing Parameters (Density) and Their Impacts on Pay Factor Related Parameters

Cell Description	Voids in the Mineral Aggregate (VMA)	Air Voids	Density	Major/Minor Effect
Reading 1 (AC37-AC61)	No	No	Yes	Major
Reading 2 (AG37-AG61)	No	No	Yes	Major
Device Used (X37-X61)	No	No	Yes	Major

Table 3.3 Total Number of Material Testing Parameters

Parameter Type	**Total Number (Department and Prime contractor)**	**Total Number (Density)**
Parameters with Major Impact	14	3-75
Parameters with Moderate Impact	6	0
Parameters with Minor Impact	7	0

Plausible Corrections, P.C.

P.C. is defined as the incidents in which values were likely not changed deliberately. The most likely cause of such changes was mistyping while entering the data from paper reports into the Excel files. To identify P.C. incidents, 7 different scenarios where considered that are comprised of:

- Case 1: One digit may be pressed instead of a neighboring key

While typing a digit, there is always a chance that another digit is mistakenly pressed instead of the desired number. For this analysis, a keypad like that of Figure 3.3 was considered, because in most of the desktop computers the keypad has this format. Here, all the possible cases that can happen when typing a number was considered.

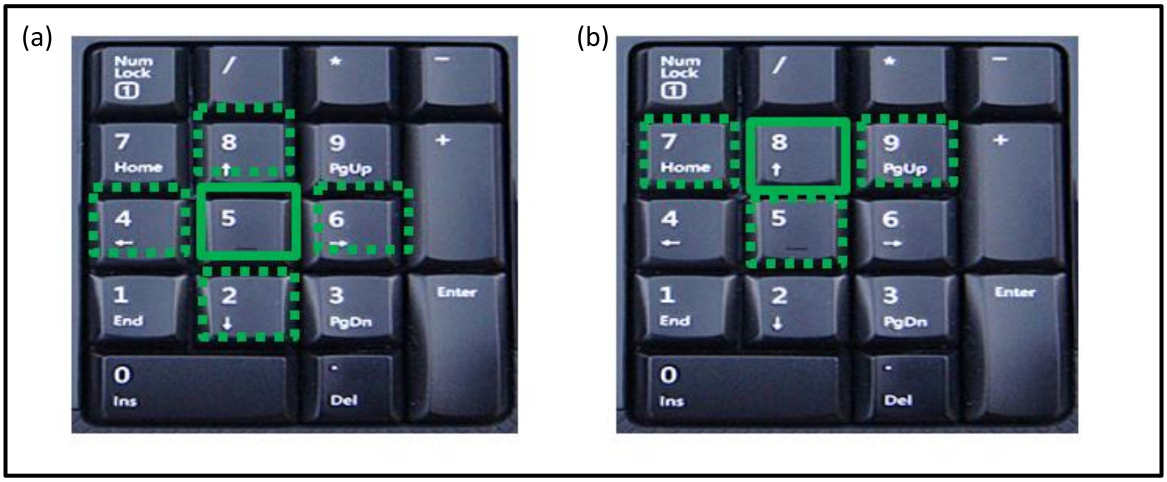
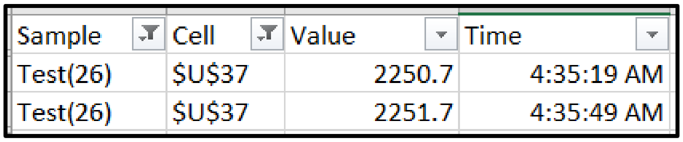


Figure 3.3 Plausible correction (case 1)

Usually, the neighboring keys surrounding a particular key have the highest probability of being mistakenly pressed. As in Figure 3.3a, if number 5 is considered, the closest buttons to 5 are 2, 4, 6, 8. It was assumed that the probability of mistakenly pressing any of these digits instead of 5 is the same. Similarly, for number 8 (Figure 3.3b), the closest keys are for numbers 5, 7, and 9. An algorithm was developed to label the repetition as P.C. if the number of repetitions is only one (there has been a change only from the 1st case to the 2nd case) and only one digit (at any position) is changed. This method is considered for all numbers from 0 to 9, and a series of neighboring possibilities are considered in each individual possible case.

The algorithm first separates each digit of a number. In the next step, the algorithm does an element by element comparison and tries to identify if the changed digit fits in the closest neighboring category. In Figure 3.4, for example, the number of changes/repetitions is only one and it is for one digit only (2250.7 versus 2251.7). The algorithm eliminates all the similar digits between the two entries except for the 4th digit. Then, a comparison is made for the unmatched digit, which is 0 versus 1, in this case. Since 1 fits in the adjacent neighboring rule of 0, this is considered a P.C.



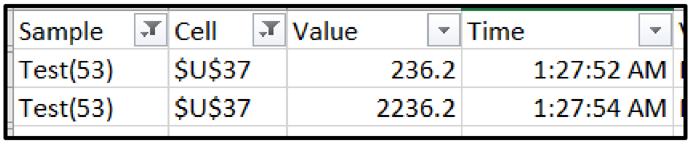
Sample	Cell	Value	Time
Test(26)	U37	2250.7	4:35:19 AM
Test(26)	U37	2251.7	4:35:49 AM

Figure 3.4 Plausible correction (case 1)—example

- Case 2: One or two digits were missed while typing

A very common scenario of P.C. is 1 or 2 digits were missed when trying to type quickly or simply because the desired digit was not pressed properly. An individual might want to press 123, but instead, he/she presses 13 and misses 2. This is a clear case of an honest mistake or P.C. Figure 3.5 shows an example of a missed digit case of a P.C. The typist tried to insert 2236.2, but instead, he/she initially typed 236.2 missing the digit 2, and later corrected it.

An element-wise comparison is simply not possible in this case because the missing number can be any digit at any place. The logic that was used here is that if the 2nd entry is smaller than 80% or larger than 120% of the first input, then it is a P.C. Usually, U.C.s are around the vicinity of the actual value, but are corrected to return a better result. When the two values are too far apart, it is most probably a P.C. case. Generally, if a number is missed, the first entry becomes much smaller than the final or corrected entry. Hence a percentage difference can help determine this case. However, there is no fixed percentage threshold that can be specified to accurately determine the missed number case, but through the manual analysis of data, the appropriate threshold was found to be 20% above or below the final entry. In this case, the change would be considered as a P.C. only if the number of repetitions is only one.

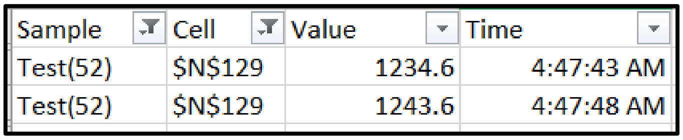


Sample	Cell	Value	Time
Test(53)	U37	236.2	1:27:52 AM
Test(53)	U37	2236.2	1:27:54 AM

Figure 3.5. Plausible correction (case 2)—example

- Case 3: Order of digits were reversed while typing

A very common case of P.C. is typing digits in the wrong order, for example, 34 instead of 43. Figure 3.6 depicts a case of order of digits being reversed when typing. The user wanted to type 1243.6, but instead, he/she typed 1234.6.

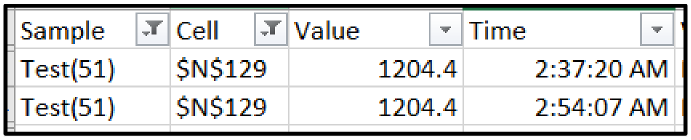


Sample	Cell	Value	Time
Test(52)	N129	1234.6	4:47:43 AM
Test(52)	N129	1243.6	4:47:48 AM

Figure 3.6 Plausible correction (case 3)—example

- Case 4: Exact same value was typed twice

The initial inspection of the dataset showed that, in some cases, the exact same value was entered twice for a single parameter. This happened quite often. A logic was added in this algorithm to identify this type of P.C., as in Figure 3.7.

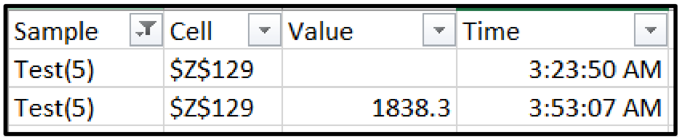


Sample	Cell	Value	Time
Test(51)	N129	1204.4	2:37:20 AM
Test(51)	N129	1204.4	2:54:07 AM

Figure 3.7 Plausible correction (case 4)—example

- Case 5: Cell was empty at first and was filled in the second entry

Manual inspection revealed some cases where the cell was empty at first, but it was filled later. A possible reason might be that the VBA script records everything, even a single click, as an input while nothing was actually entered. The user then inputted the actual entry, as demonstrated in Figure 3.8 which is considered as another case of P.C.

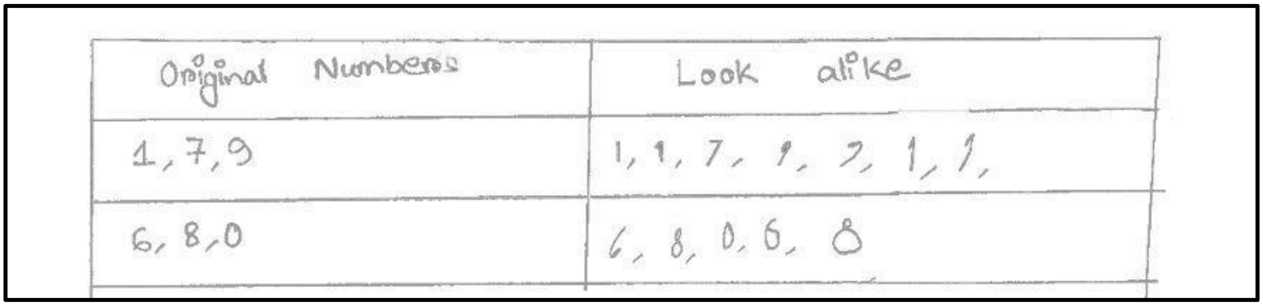


Sample	Cell	Value	Time
Test(5)	Z129		3:23:50 AM
Test(5)	Z129	1838.3	3:53:07 AM

Figure 3.8 Plausible correction (case 5)—example

- Case 6: Digits that are hand-written similarly

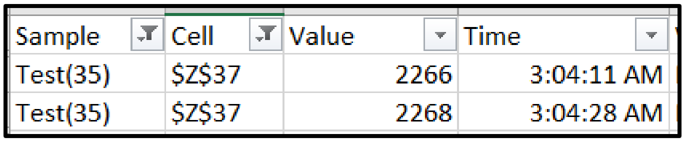
Another case of P.C. is the numbers that look alike in handwriting can be entered instead of one another. Test results are usually logged in a paper sheet and are later digitized into the ITD provided Excel file. It is evident that handwriting would not be similar for all people, and there is a possibility of typing a digit instead of the actual digit due to their similarity in handwriting. For instance, 1 might look like 7 or 9 in the handwriting of various people (Figure 3.9). Another combination can be 6/8/0. In any of these combinations, it is essential that the number of repetitions must be only one. If the number of repetitions is more than one, it is more likely to be an U.C. case.



Original Numbers	Look alike
1, 7, 9	1, 1, 7, 1, 7, 1, 1,
6, 8, 0	6, 8, 0, 6, 8

Figure 3.9 Look-wise case of plausible correction

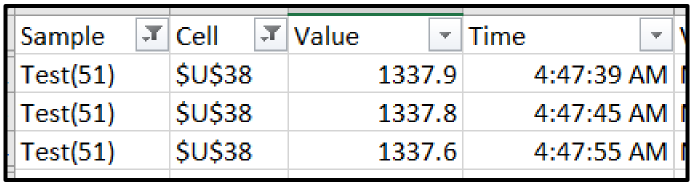
Figure 3.10 shows a change of digit from 6 to 8, which is most probably a P.C.



Sample	Cell	Value	Time
Test(35)	Z37	2266	3:04:11 AM
Test(35)	Z37	2268	3:04:28 AM

Figure 3.10 Plausible correction (case 6)—example

There is a point of argument here that this can fit in both cases, that the number was changed deliberately, or a simple look wise mistake has occurred. It is not possible to state with certainty that this is a P.C. or a U.C. case, since this is a subjective issue. It was concluded that if the number of changes is more than 1 (more than 1 repetition) the likelihood is higher toward U.C., whereas if the number of changes is only one, it aligns well with the P.C. case. Figure 3.11 shows a case where the changes could have been categorized as look wise error, but since the number of changes was more than one, this is no longer considered a P.C. case and it rather falls into a U.C. case.



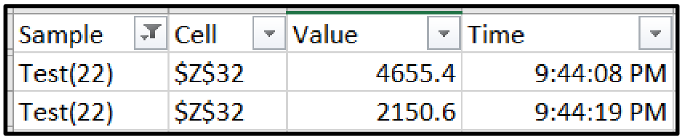
Sample	Cell	Value	Time
Test(51)	U38	1337.9	4:47:39 AM
Test(51)	U38	1337.8	4:47:45 AM
Test(51)	U38	1337.6	4:47:55 AM

Figure 3.11 Plausible correction (case 6)—example

- Case 7: Difference between two entries is too high

There have been some cases where the difference between two successive entries is too high. These incidents can also be differentiated through the percentage calculation. If the first entry is less than 80% or greater than 120% of the 2nd entry, then the change is likely a P.C. There might be several reasons for this P.C. case, including reporting a parameter value for another parameter or reporting the parameter value from one test/sample to another test/sample.

Figure 3.12 is a clear example of a large difference between successive entries, which can be considered as a P.C. Here the 2nd entry was less than 50 percent of the first case (4655.4 versus 2150.6), so this is most probably a P.C. case.



Sample	Cell	Value	Time
Test(22)	Z32	4655.4	9:44:08 PM
Test(22)	Z32	2150.6	9:44:19 PM

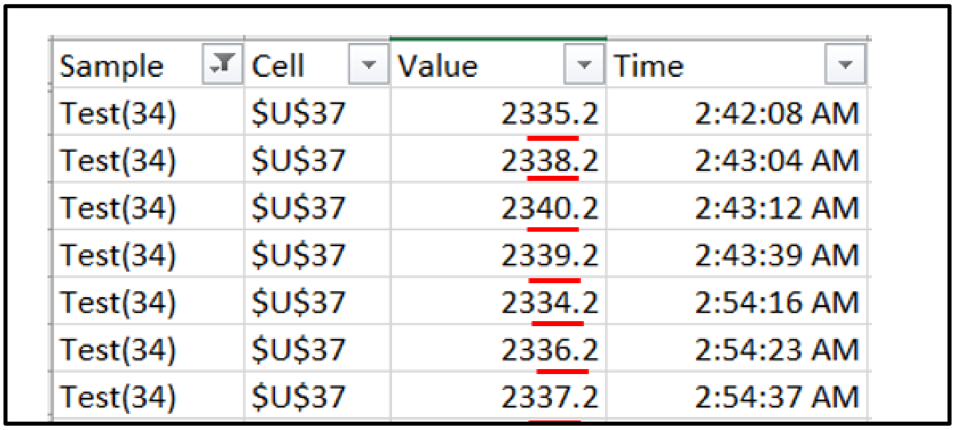
Figure 3.12 Plausible correction (case 7)—example

Unexplained Corrections, U.C.

U.C. is defined as the incidents of corrected values that could not be attributed to typographical and other cases of mistakes, after exhaustive consultation with advisors and engineers. Such corrections may have been done intentionally to reach a desired value, potentially change the payment, and/or modify a certain test outcome, although investigating the factual causes of such changes is beyond the scope of this research. To identify U.C. incidents, 4 different scenarios were considered that are comprised of:

- Case 1: Changing values in a pattern or following a combination

U.C. cases mostly followed a pattern of change. In most cases, the number of changes is more than one, and the values are changing by a value of 1/2/10 in the positive or negative direction. Figure 3.13 presents a clear indication of a U.C. case. Here, the total number of changes is 6 times. The value was increased in the first two cases, reduced on the 3rd and 4th corrections, but then in the final two incidents, it increased again.

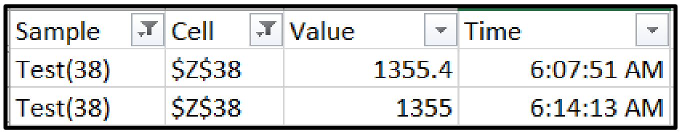


Sample	Cell	Value	Time
Test(34)	U37	2335.2	2:42:08 AM
Test(34)	U37	2338.2	2:43:04 AM
Test(34)	U37	2340.2	2:43:12 AM
Test(34)	U37	2339.2	2:43:39 AM
Test(34)	U37	2334.2	2:54:16 AM
Test(34)	U37	2336.2	2:54:23 AM
Test(34)	U37	2337.2	2:54:37 AM

Figure 3.13 Unexplained correction (case 1)—example

- Case 2: Decimal values are eliminated

In some U.C. cases, the digits after the decimal point are eliminated (e.g. Figure 3.14). In general, this might be a very small change, but even small changes in HMA samples can have high impacts. Therefore, these cases are also considered as U.C. in this algorithm.

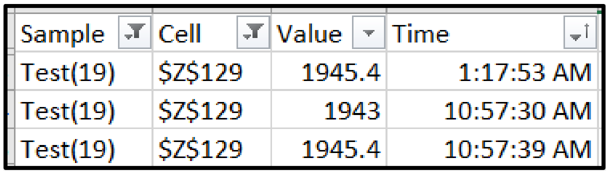


Sample	Cell	Value	Time
Test(38)	Z38	1355.4	6:07:51 AM
Test(38)	Z38	1355	6:14:13 AM

Figure 3.14 Unexplained correction (case 2)—example

- Case 3: Parameter values were changed but returned to the initial value

A clear case of correcting data is presented in Figure 3.15, where the values were changed but later returned to the original value. Here, initially, the value was entered as 1945.4, which was changed to 1943, but later brought back to 1945.4. Although the value did not change, this was considered as exploring values potentially for the wrong reasons and labeled it as U.C. We note, however, that this case does not leave a monetary impact. This has been taken into consideration when calculating monetary impacts of changes in Chapter 4.

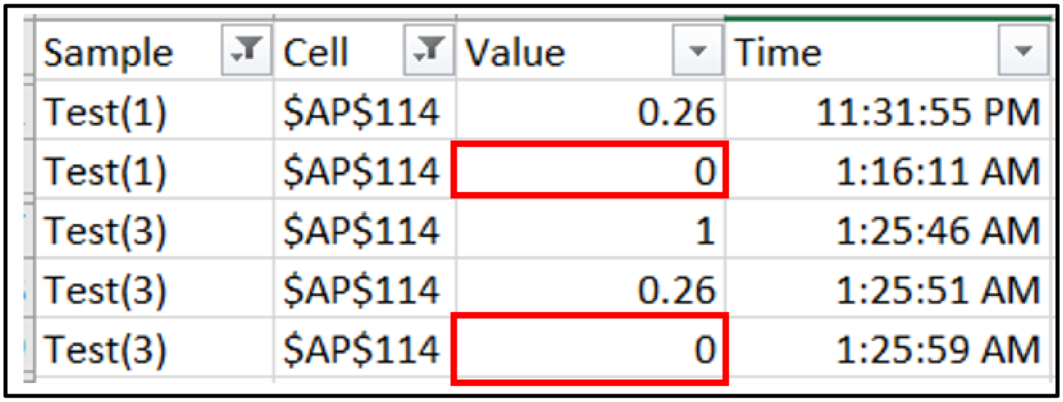


Sample	Cell	Value	Time
Test(19)	Z129	1945.4	1:17:53 AM
Test(19)	Z129	1943	10:57:30 AM
Test(19)	Z129	1945.4	10:57:39 AM

Figure 3.15 Unexplained correction (case 3)—example

- Case 4: Parameters were first assigned a value but finally changed to zero or removed entirely

There have been times, especially for parameters with small values, that the values were completely deleted or replaced with a value of zero. For example, in Figure 3.16, for sample Test (1) the value was set to 0.26 and replaced with zero. This change was considered as U.C.

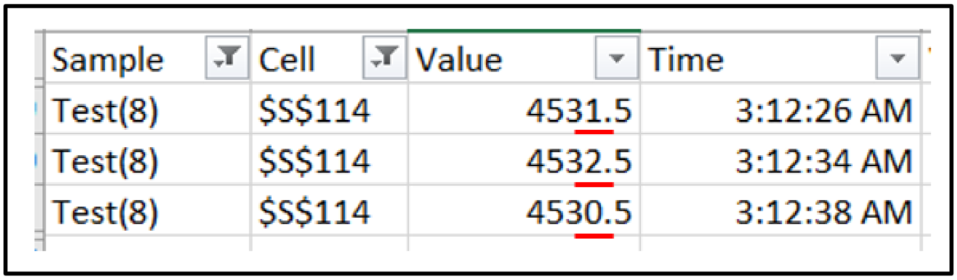


Sample	Cell	Value	Time
Test(1)	AP114	0.26	11:31:55 PM
Test(1)	AP114	0	1:16:11 AM
Test(3)	AP114	1	1:25:46 AM
Test(3)	AP114	0.26	1:25:51 AM
Test(3)	AP114	0	1:25:59 AM

Figure 3.16 Unexplained correction (case 4)—example

Uncertain Cases

Results indicated that there were incidents where the repetitions might fall in either U.C. or P.C. cases, an example of which is shown in Figure 3.17 (values changing from 4531.5 to 4532.5 and then to 4530.5). The first change was from 1 to 2, which might be considered P.C. In the second change, the digit 2 was replaced with 0, which is likely to be a U.C. However, there is enough room for argument to fit these cases in other categories. But the number of changes can be informative here. It is unlikely that both cases were a typo, hence this case is considered as U.C.

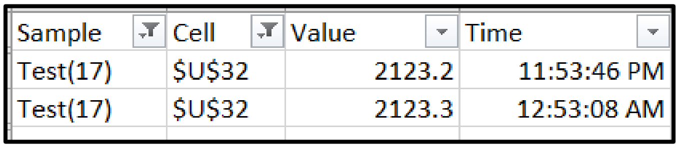


Sample	Cell	Value	Time
Test(8)	S114	4531.5	3:12:26 AM
Test(8)	S114	4532.5	3:12:34 AM
Test(8)	S114	4530.5	3:12:38 AM

Figure 3.17 Plausible correction /unexplained correction (case 1)—example

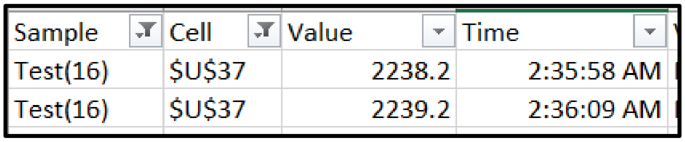
Impact of Time Stamp

Although U.C. cases generally occur in a relatively short period of time, a definite relationship between P.C./U.C. cases with time that can be explored in a computer code (Figures 3.18, 3.19, 3.20, and 3.21) could not be determined. Both categories have examples where a change occurred instantly or after some time.



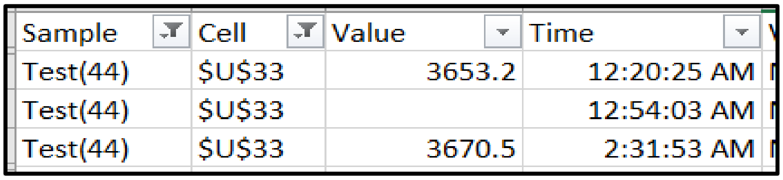
Sample	Cell	Value	Time
Test(17)	U32	2123.2	11:53:46 PM
Test(17)	U32	2123.3	12:53:08 AM

Figure 3.18 Plausible correction relationship with time—example 1



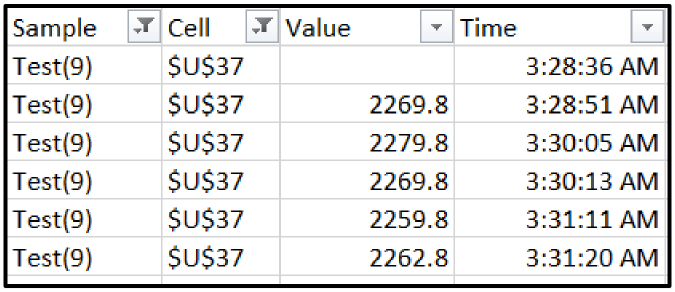
Sample	Cell	Value	Time
Test(16)	U37	2238.2	2:35:58 AM
Test(16)	U37	2239.2	2:36:09 AM

Figure 3.19 Plausible correction relationship with time—example 2



Sample	Cell	Value	Time
Test(44)	U33	3653.2	12:20:25 AM
Test(44)	U33		12:54:03 AM
Test(44)	U33	3670.5	2:31:53 AM

Figure 3.20 Unexplained correction relationship with time—example 1



Sample	Cell	Value	Time
Test(9)	U37		3:28:36 AM
Test(9)	U37	2269.8	3:28:51 AM
Test(9)	U37	2279.8	3:30:05 AM
Test(9)	U37	2269.8	3:30:13 AM
Test(9)	U37	2259.8	3:31:11 AM
Test(9)	U37	2262.8	3:31:20 AM

Figure 3.21 Unexplained correction relationship with time—example 2

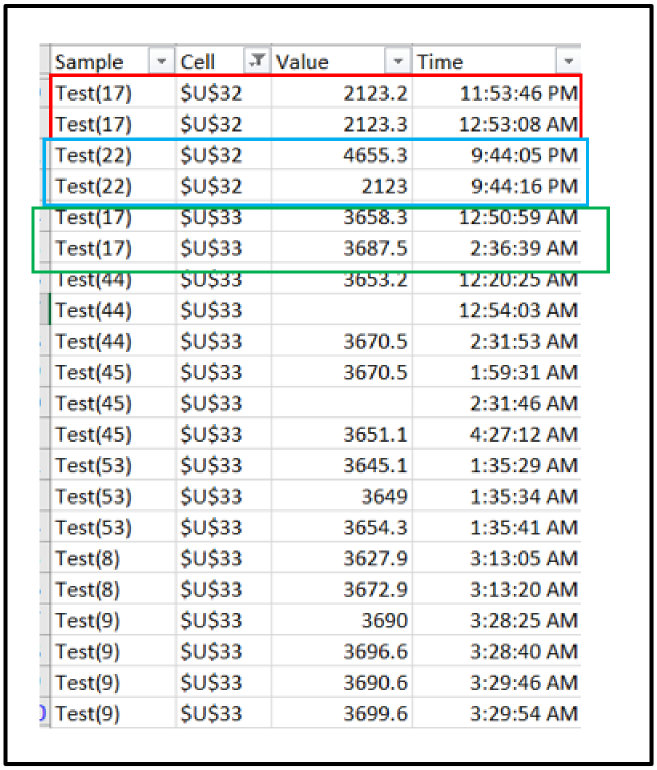
Categorization of Repeated Entries to P.C./U.C. Using Logic-Based Algorithms

This section describes the approach adopted to categorize the data corrections into two groups: (1) P.C. and (2) U.C., and to develop the logic-based algorithms. This was accomplished in several steps:

I. The first step was to separate the repeated data from the non-repeated incidents. Non-repeated data represent cases where no change in values was recorded for parameters in the input form.
II. The second step involved manual inspection of all the repeated (corrected) data to identify any existing patterns. Data corrections identified through this approach were categorized into P.C. and U.C.
III. The third step was to find general patterns in P.C. and U.C. cases.
IV. A total of 7 and 4 general patterns were found for the P.C. and U.C. categories, respectively.
V. Algorithmic logics were devised for each case, and computer codes were developed to automatically detect and categorize data value changes

Development of algorithms and code was accomplished in several steps:

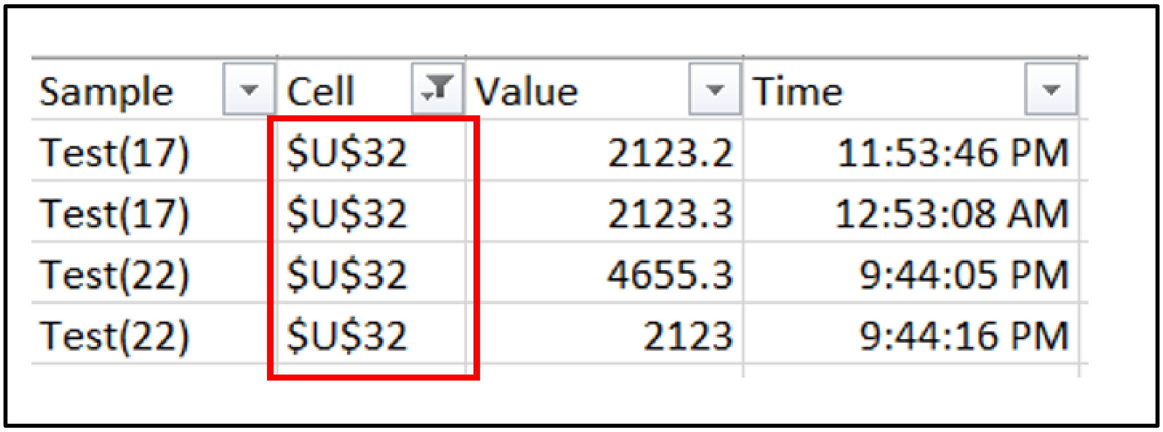
1. Initially, all cells with repeated values (more than one entry for each parameter) associated with the pay affecting parameters in each project were identified (Figure 3.22).



Sample	Cell	Value	Time
Test(17)	U32	2123.2	11:53:46 PM
Test(17)	U32	2123.3	12:53:08 AM
Test(22)	U32	4655.3	9:44:05 PM
Test(22)	U32	2123	9:44:16 PM
Test(17)	U33	3658.3	12:50:59 AM
Test(17)	U33	3687.5	2:36:39 AM
Test(44)	U33	3653.2	12:20:25 AM
Test(44)	U33		12:54:03 AM
Test(44)	U33	3670.5	2:31:53 AM
Test(45)	U33	3670.5	1:59:31 AM
Test(45)	U33		2:31:46 AM
Test(45)	U33	3651.1	4:27:12 AM
Test(53)	U33	3645.1	1:35:29 AM
Test(53)	U33	3649	1:35:34 AM
Test(53)	U33	3654.3	1:35:41 AM
Test(8)	U33	3627.9	3:13:05 AM
Test(8)	U33	3672.9	3:13:20 AM
Test(9)	U33	3690	3:28:25 AM
Test(9)	U33	3696.6	3:28:40 AM
Test(9)	U33	3690.6	3:29:46 AM
Test(9)	U33	3699.6	3:29:54 AM

Figure 3.22 Repeated data entry (third column) of pay affecting parameters (second column; e.g. U32) for tests in a project (first column; e.g. test (17)). Time of data entry is presented in column 4.

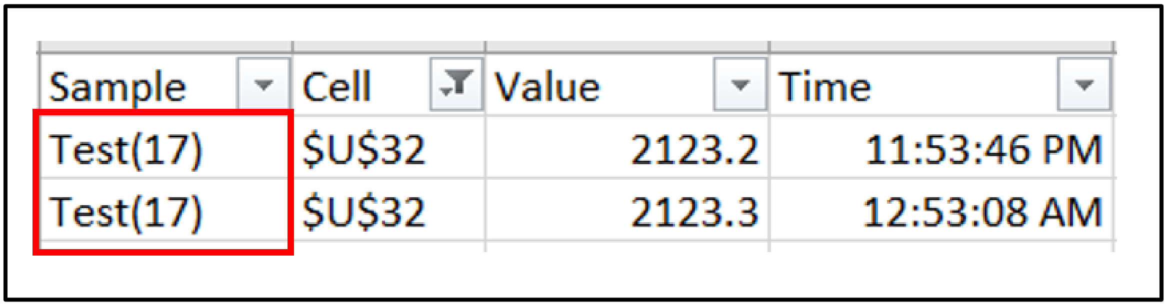
2. Repeated cells are then separated per parameter name. In Figure 3.23, for example, parameter U32 (mass of bowl) is separated.



Sample	Cell	Value	Time
Test(17)	U32	2123.2	11:53:46 PM
Test(17)	U32	2123.3	12:53:08 AM
Test(22)	U32	4655.3	9:44:05 PM
Test(22)	U32	2123	9:44:16 PM

Figure 3.23 Separation of cells based on parameter name

3. For each parameter, one set of samples (i.e., tests) is then considered at a time. Figure 3.23 had both Test (17) and Test (22), but in this step, each set of samples was considered separately, i.e., Test (17) (Figure 3.24).



Sample	Cell	Value	Time
Test(17)	U32	2123.2	11:53:46 PM
Test(17)	U32	2123.3	12:53:08 AM

Figure 3.24 Separation of cells based on test/sample

4. Cells are then run through a series of algorithms to determine cases of P.C. and U.C. based on the cases explained earlier.
5. When there are multiple repetitions for a single parameter, each two consecutive entries (for instance, 1st and 2nd entry of a series of corrections) are considered a pair, and these pairs are run through the algorithms to be labelled P.C. or U.C. This is repeated for all pairs (e.g. 2nd and 3rd, 3rd and 4th, and so on). Once the serial comparison is completed, the first and last entries are considered a pair, and a similar analysis is done. It is noted that there were some cases where the values were changed by a very small amount in every repetition, but this was done multiple times. In this case, each pair was labeled as P.C., but the comparison of first and last entries showed U.C. If the result is P.C. for all the pairs, the entire group is labeled as P.C. Upon detection of U.C., the entire group is labeled U.C. This procedure is visually represented in Figure 3.25.

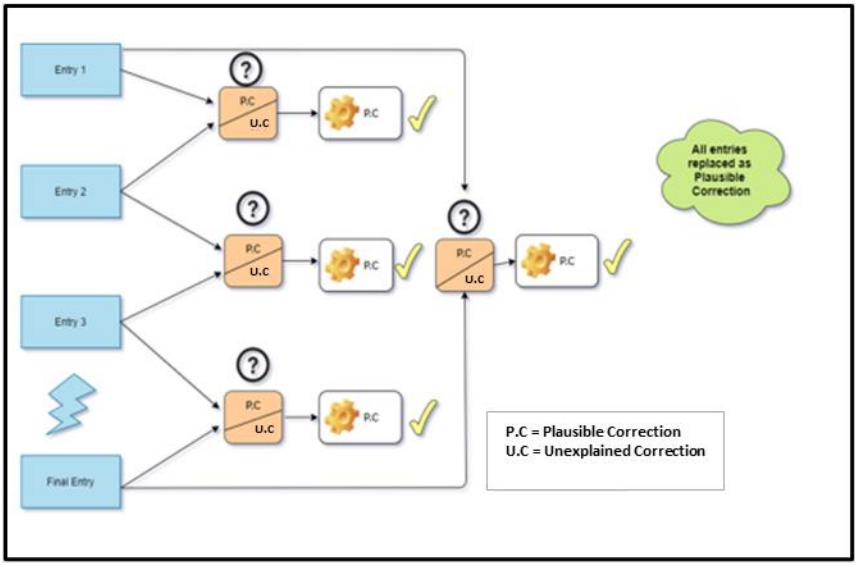


Figure 3.25 Methodology for plausible correction/unexplained correction categorization

Results: P.C. vs. U.C. for Entities 1 and 2

P.C./U.C. Classification

The algorithms explained earlier were applied to all audit trail data from the available projects' datasets (separately for entities 1 and 2) to determine P.C. and U.C. cases. For each project, the number of unique parameters that were corrected, and the total number of times those parameters were corrected were determined. Also, parameters with major/moderate/minor impacts on pay factor were separated to analyze whether or not one category might be more susceptible to correction than others. Figure 3.26 shows the total number of corrected parameters and frequency of corrections for project #1, as an example. Figure 3.26 shows the entity 1-reported statistics on the left side and the entity 2-reported statistics on the right side. In this project and for major parameters in entity 1-reported data, there were a total of 32 unique parameters that fell within the P.C. cases, and these parameters were changed a total of 66 times (an average of roughly one change per parameter). A greater number of U.C. cases was observed for the entity 1-reported major parameters, with a total of 58 parameters being changed 211 times (an average of roughly 2.5 changes per parameter). The higher average number of changes for major parameters in the case of U.C. compared to P.C. (2.5 versus 1) implies that there might be some effort potentially to tune the parameter values to obtain certain outcomes. For moderate parameters in the P.C. category, 11 unique parameters were changed 25 times (an average of roughly 1 change per parameter), and in the U.C. category, 18 unique parameters were changed 60 times (an average of roughly 3 changes per parameter). Finally, 18 unique minor parameters were changed 37 times for the P.C. category, and 24 parameters were changed 70 times for the U.C. category. It was observed that in this project, the number of changes for U.C. is roughly 2 times, or more, per unique parameter, whereas P.C. cases show roughly 1 change per unique parameter. While one may argue that this can be partly an artifact of the devised algorithms, careful manual investigation of P.C./U.C. categorized audit trail data confirmed that algorithms are performing accurately. We attribute this observation to the P.C. cases being unintentional, and if an error/mistake occurred, it is usually corrected in the second entry. This is, however, quite different for the U.C. cases due to the potentially intentional nature of the corrections as the operator might seek a certain outcome through changing parameter value. The parameters are indeed changed multiple (≥2) times, which resulted in a high number of changes for major/moderate/minor U.C. cases.

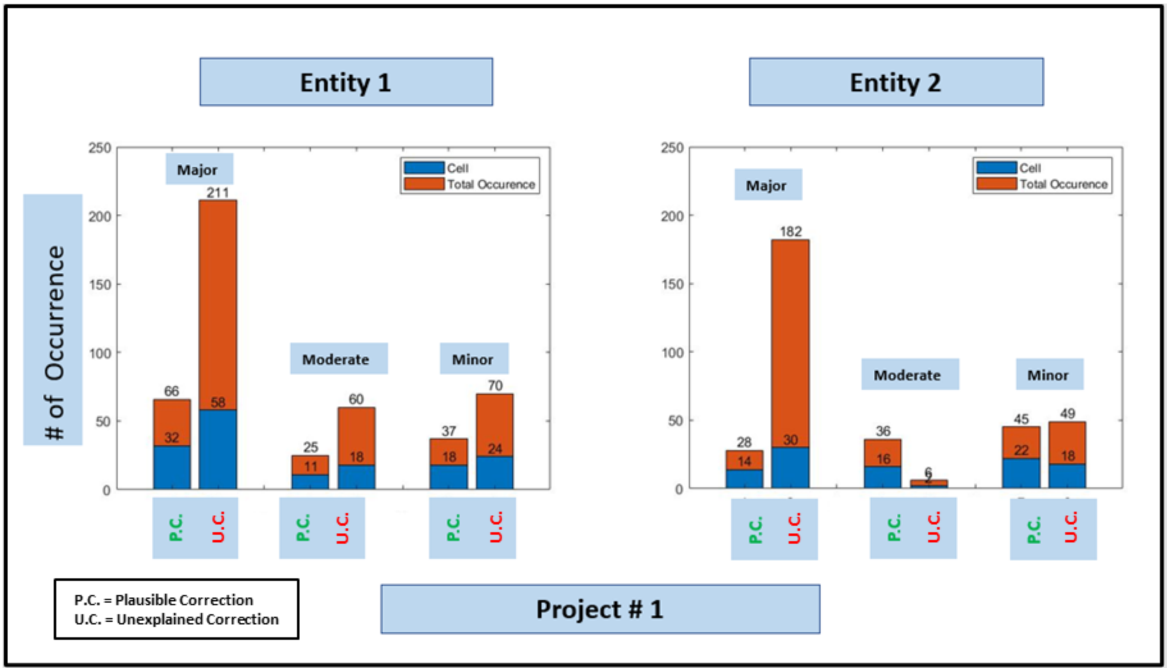


Figure 3.26 Number of occurrences of P.C./U.C. for project #1

A similar trend is observed in the entity 2-reported data for this project. A total of 14 major parameters in the P.C. category was changed 28 times, and 30 major parameters in the U.C. category were changed 182 times. In the case of moderate parameters in the P.C. category, 16 parameters were changed 36 times, whereas in the U.C. category, 2 parameters were corrected 6 times. For minor parameters in the P.C. category, 22 parameters were corrected 45 times, and in the U.C. category, 18 parameters were corrected 49 times. Data correction seems to be less pronounced in the entity 2 data compared to the entity 1-reported data.

This analysis was conducted on all available projects and the reports of their results are presented in Table 3.4.

Table 3.4 Unique and Total Number of Material Testing Parameter Changes

Project Number	Classification Category	Entity 1-Major-Unique Changes*	Entity 1-Major-All Changes**	Entity 1-Moderate-Unique Changes	Entity 1-Moderate-All Changes	Entity 1-Minor-Unique Changes	Entity 1-Minor-All Changes	Entity 2-Major-Unique Changes	Entity 2-Major-All Changes	Entity 2-Moderate-Unique Changes	Entity 2-Moderate-All Changes	Entity 2-Minor-Unique Changes	Entity 2-Minor-All Changes
Project 1	U.C.	58	211	18	60	24	70	30	182	2	6	18	49
Project 1	P.C.	32	66	11	25	18	37	14	28	16	36	22	45
Project 2	U.C.	94	404	18	53	26	81	26	66	0	0	0	0
Project 2	P.C.	29	64	14	30	17	41	11	22	0	0	0	0
Project 3	U.C.	2	6	2	8	0	0	0	0	0	0	0	0
Project 3	P.C.	10	20	3	6	0	0	0	0	0	0	0	0
Project 4	U.C.	31	96	5	13	25	62	93	276	4	22	8	21
Project 4	P.C.	9	18	4	8	2	4	15	30	6	12	8	16
Project 5	U.C.	19	52	2	5	2	6	39	87	6	12	9	22
Project 5	P.C.	10	20	1	2	5	10	48	98	18	37	30	63
Project 6	U.C.	25	73	3	9	0	0	0	0	0	0	0	0
Project 6	P.C.	11	23	1	2	1	2	0	0	0	0	0	0
Project 7	U.C.	63	303	9	43	9	23	37	151	1	7	7	16
Project 7	P.C.	17	36	6	12	11	22	16	39	3	7	3	6
Project 8	U.C.	33	138	1	7	6	19	19	77	2	5	5	17
Project 8	P.C.	13	26	7	14	7	14	13	26	3	6	6	15
Project 9	U.C.	1	2	1	2	1	2	1	3	2	5	2	4
Project 9	P.C.	5	10	1	2	1	2	2	4	7	14	2	4
Project 10	U.C.	7	28	2	17	0	0	0	0	0	0	0	0
Project 10	P.C.	0	0	0	0	0	0	0	0	0	0	0	0
Project 11	U.C.	8	17	1	3	2	4	17	60	10	37	5	19
Project 11	P.C.	8	16	11	22	4	9	13	28	5	10	4	9
Project 12	U.C.	7	17	2	6	1	3	26	56	5	10	3	7
Project 12	P.C.	4	8	3	7	0	0	3	6	2	4	5	10
Project 13	U.C.	0	0	1	2	1	2	0	0	0	0	0	0
Project 13	P.C.	1	2	0	0	0	0	0	0	0	0	0	0
Project 14	U.C.	66	334	7	27	5	14	0	0	0	0	0	0
Project 14	P.C.	20	41	11	22	5	11	0	0	0	0	0	0
Project 15	U.C.	3	21	3	23	1	2	10	49	0	0	0	0
Project 15	P.C.	0	0	2	4	1	2	3	8	2	4	0	0

* **Entity 1-Major-Unique Changes: Number of Unique Changes in parameters with Major impacts in Entity 1**

** **Entity 1-Major-Unique Changes: Number of All Changes in parameters with Major impacts in Entity 1**

Figures 3.27, 3.28, and 3.29 visually depict three example project results (projects #4, #7, #9).

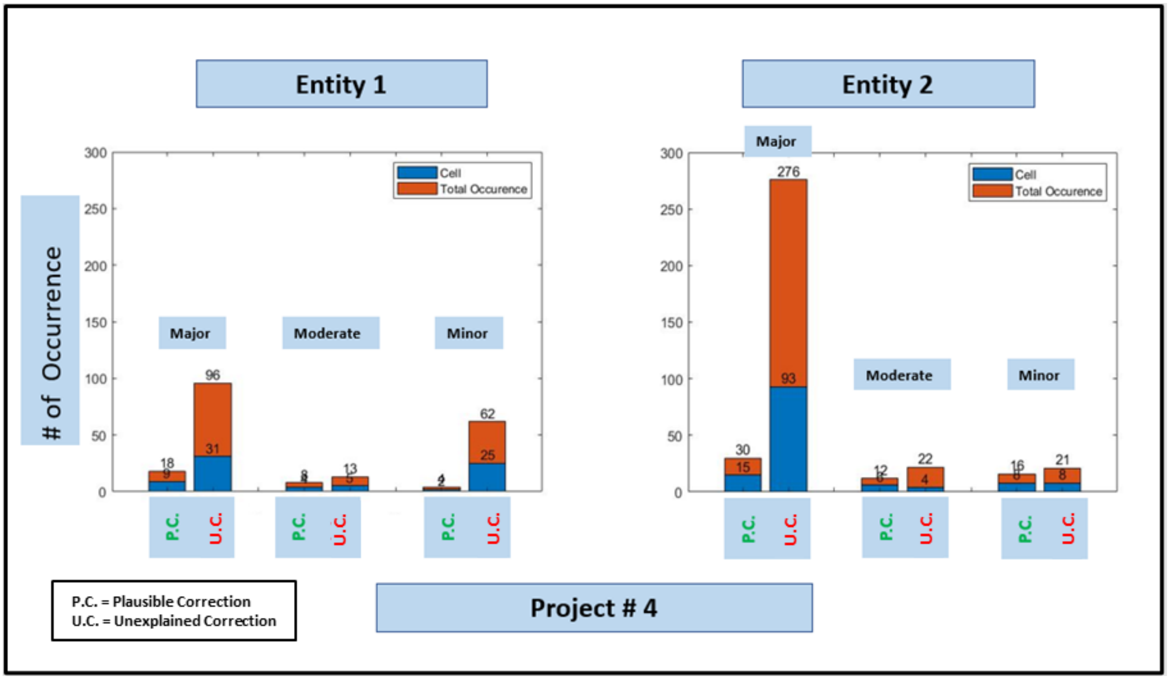


Figure 3.27 Number of occurrences of P.C./U.C. for project #4

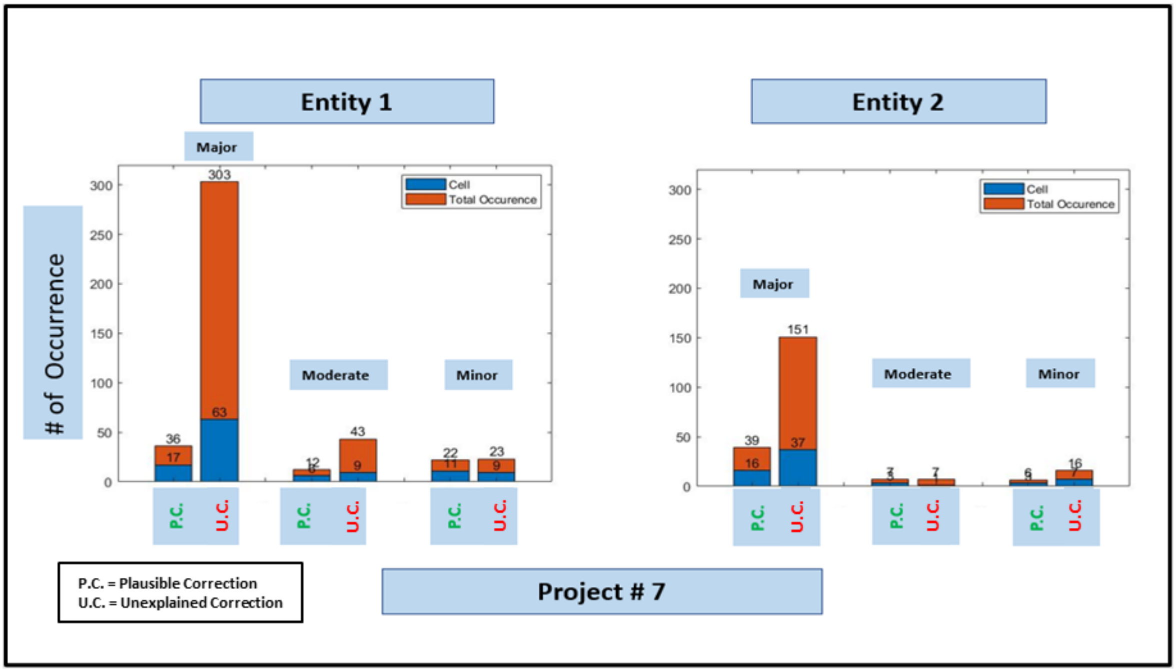


Figure 3.28 Number of occurrences of P.C./U.C. for project #7

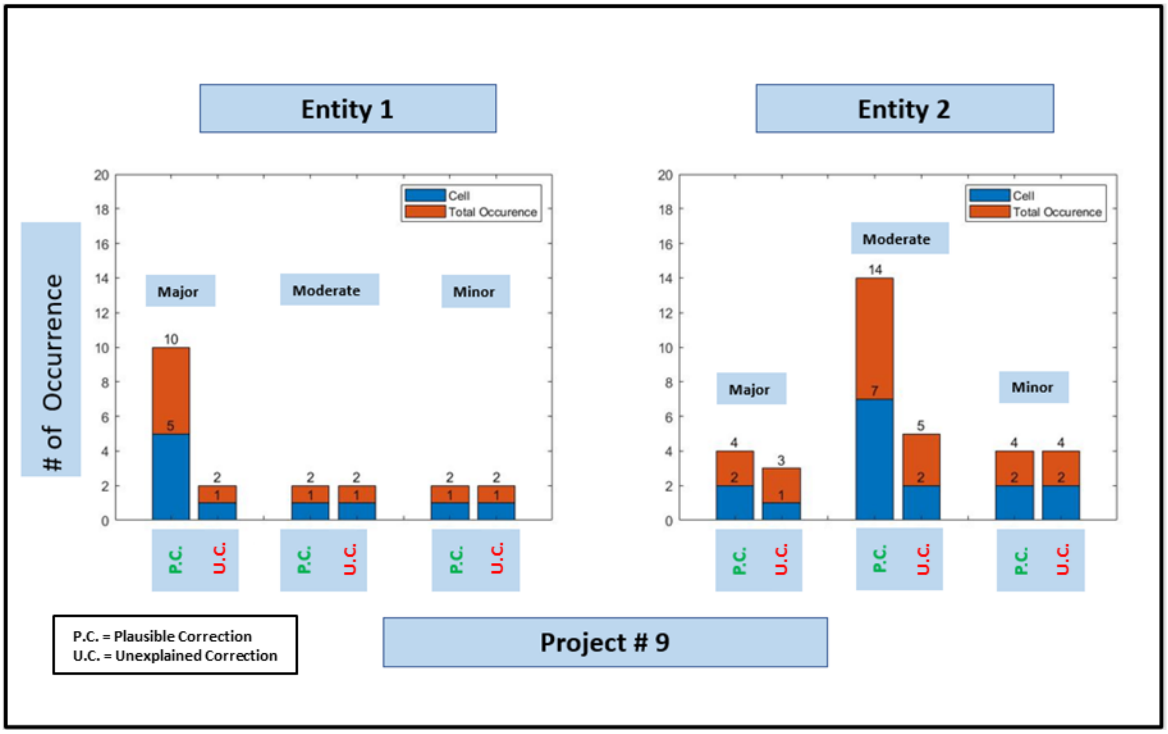


Figure 3.29 Number of occurrences of P.C./U.C. for project #9

Summary

In this research, we analyzed an audit trail dataset of material testing reports that registered all value entries in the submitted spreadsheets. Analyzing the series of changes in parameter values can provide important insights about the potential sources of discrepancies that are observed between the prime contractor test results and those of the ITD, as well as those of the original mix design. First, all the provided instances of changes in the parameter values were manually analyzed, and the general patterns in data reporting were determined. These instances were categorized into two general categories of Plausible Correction (P.C.) and Unexplained Correction (U.C.). Then, a logic-based computer algorithm was developed to automatically classify all instances of parameter value changes to P.C. and U.C. The performance of the automatic classification results from the computer algorithms were rigorously evaluated through manual inspection by various members of the research team. Results show that a total of 595 and 316 unique parameters were changed 2,268 and 660 times that can be categorized as U.C. and P.C., respectively, in entity 1-reported data. For entity 2-reported data, a total of 387 and 280 unique parameters were changed 1,266 and 587 times that can be categorized as U.C. and P.C., respectively. Furthermore, results indicated that major parameters were corrected more than two times on average per parameter, with some changing more than five or six times. Parameter values for P.C. cases were mostly changed only one time.

Findings emphasize the necessity of an advanced cumulative approach to improve QC/QA process. An improved approach is needed to remove probable unethical course of actions and bring more rigor to the QC/QA analysis. In recent years and informed by this research, ITD has taken several steps to bring more rigor to the quality assurance processes as described in the Foreword.

4. Analyzing the Financial Impact of Unexplained Corrections

Introduction

We acquired a unique dataset of material testing reports for HMA construction projects in Idaho that recorded every instance of data entry in the material testing spreadsheet. Audit trail recording was conducted in the background with a VBA code, and was not apparent to the data reporting personnel. This provided a series of data entry for some material testing parameters, and showed data corrections in many parameters. It is expected that each parameter be reported as observed, and hence being reported only once, although typographical errors may result in multiple entries for some parameters. The patterns observed in some parameters in the audit trail data, however, cannot be simply explained as typographical errors. As described in Chapter 3, a series of logic-based algorithms was applied to categorize all instances of multiple (more than 1) data entry as either P.C. or U.C. The aim of this chapter was to analyze the financial repercussions and impacts of U.C. instances. It is plausible that data corrections can occur for monetary benefit or personal/institutional advantage, although such investigation is beyond the scope of this study. U.C. may also have occurred to obtain bonus payments, avoid repetition of faulty tests and works, and pass substandard work, among other reasons. The focus of this chapter was mainly to determine whether or not U.C. instances had an impact on pay factors, and if so, what is the extent of that impact, without recourse to the potential reasons for U.C. occurrence. After identification of the U.C., in this chapter, the monetary payment calculation procedures followed by ITD were replicated and lot-wise payments for various lots of HMA projects prior to and after data corrections were calculated. Ultimately, the potential overpayments are calculated and analyzed for various data corrections.

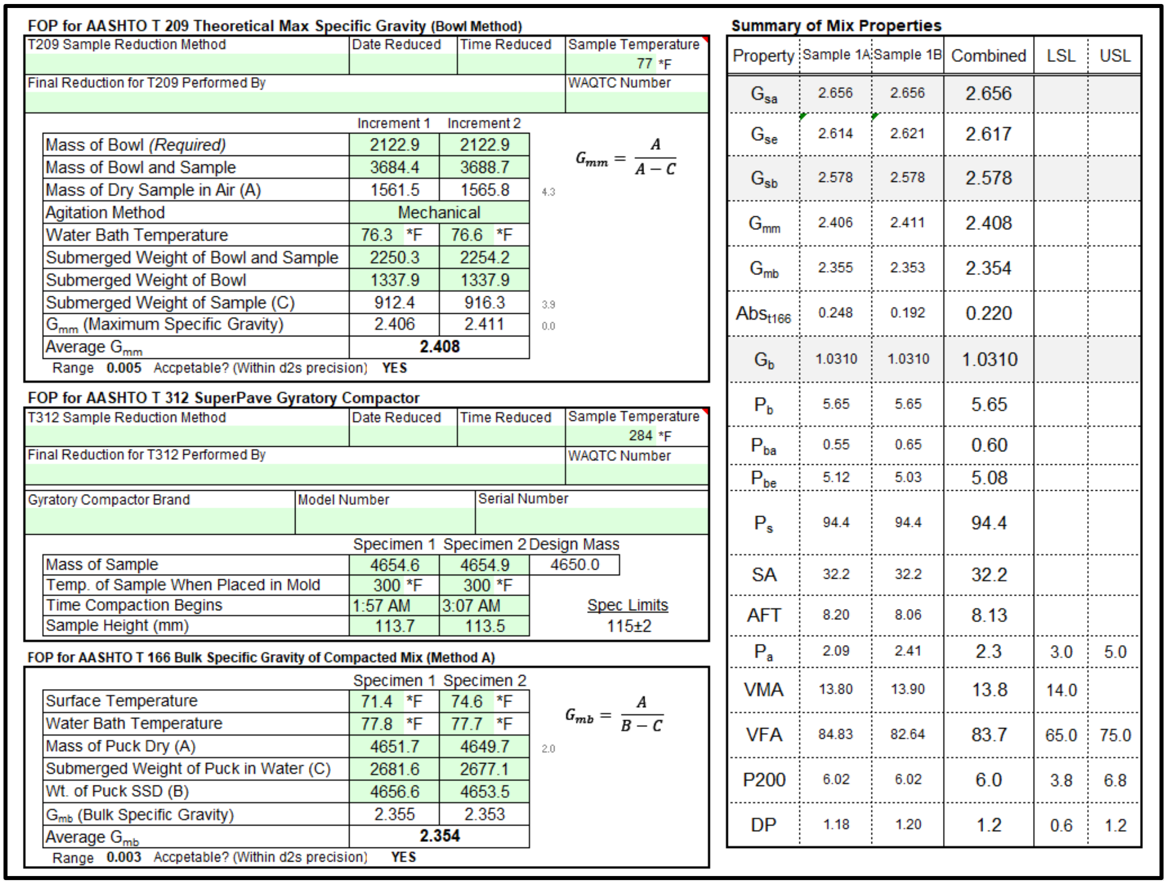
Objectives

The scope of the current chapter was to calculate the possible monetary impact (loss or gain for ITD) that may have occurred in HMA pavement projects due to these corrections in material testing reports. In the previous chapter, the U.C. instances were differentiated from the P.C. cases for multiple data entry values in volumetric testing reports. This chapter will demonstrate the possible range of economic impact of U.C. cases. The “required” payment to prime contractors were calculated if only the first acceptable instance of U.C. data entry was used. The assumption here is that the first U.C. data entry for each parameter represents that original measured value. The “required” payment was then compared to the payment based on the reported values (final U.C. instances). In all calculations, only the last entry for all P.C. instances were considered and the final reported values for the missing parameters were adopted. This is a conservative approach to calculating required payments, as missing values might also represent U.C. cases but had not been recorded in the excel file for a variety of reasons.

The basic procedure here was to go through the exact same calculation procedures followed by ITD for monetary calculation, quantify the payment-related parameters and associated payment for each lot in each project. To avoid inherent bias during the monetary analysis, this chapter uses “Entity 1” and “Entity 2” to refer to agency and prime contractor data, not necessarily in the same order. In other words, it has not been disclosed to the reader whether Entity 1 represents data from ITD or prime contractor. The same is the case for Entity 2.

Monetary Calculation

ITD follows a certain set of rules to determine the prime contractor payments for an HMA project. Several input parameters, like Mass of Bowl, Mass Pan and Initial Sample, and Calibration factor, are calculated while performing an HMA project. Once a test is completed, test results are grouped as lots based on certain tonnage of asphalt. Payment is finally calculated per lot. The required input parameters are translated into a group of asphalt mix design properties such as theoretical maximum specific gravity (G_{mm}), bulk specific gravity (G_{mb}), air voids (Pa), voids in the mineral aggregate (VMA) and voids filled with asphalt (VFA), among others. These mix design properties are then used as acceptance criteria at the start of the production. Out of these calculated mix design properties, three variables, namely Pa, VMA, and Mainline Density (percent compaction), are used for final payment calculation.
Figure 4.1 illustrates the overall representation of the generic input parameters tested in the lab/plant and later converted to mix design properties. These spreadsheets are identified as “ITD-0777” form (see Appendix C). The input parameters are shown on the left-hand side, and the calculated mix design properties are located on the right-hand side. Generally, these calculations are done for two increments (Sample 1A and Sample 1B), which are then averaged, and the combined values of Air voids, VMA, and Mainline Density (percent compaction) are used for payment calculation.



FOP for AASHTO T 209 Theoretical Max Specific Gravity (Bowl Method)

T209 Sample Reduction Method	Date Reduced	Time Reduced	Sample Temperature
			77 *F
Final Reduction for T209 Performed By			WAQTC Number

	Increment 1	Increment 2	
Mass of Bowl *(Required)*	2122.9	2122.9	
Mass of Bowl and Sample	3684.4	3688.7	
Mass of Dry Sample in Air (A)	1561.5	1565.8	4.3
Agitation Method	Mechanical		
Water Bath Temperature	76.3 *F	76.6 *F	
Submerged Weight of Bowl and Sample	2250.3	2254.2	
Submerged Weight of Bowl	1337.9	1337.9	
Submerged Weight of Sample (C)	912.4	916.3	3.9
G_{mm} (Maximum Specific Gravity)	2.406	2.411	0.0
Average G_{mm}	**2.408**		

$$G_{mm} = \frac{A}{A-C}$$

Range **0.005** Accpetable? (Within d2s precision) **YES**

FOP for AASHTO T 312 SuperPave Gyratory Compactor

T312 Sample Reduction Method	Date Reduced	Time Reduced	Sample Temperature
			284 *F
Final Reduction for T312 Performed By			WAQTC Number

Gyratory Compactor Brand	Model Number	Serial Number

	Specimen 1	Specimen 2	Design Mass
Mass of Sample	4654.6	4654.9	4650.0
Temp. of Sample When Placed in Mold	300 *F	300 *F	
Time Compaction Begins	1:57 AM	3:07 AM	Spec Limits
Sample Height (mm)	113.7	113.5	115±2

FOP for AASHTO T 166 Bulk Specific Gravity of Compacted Mix (Method A)

	Specimen 1	Specimen 2	
Surface Temperature	71.4 *F	74.6 *F	
Water Bath Temperature	77.8 *F	77.7 *F	
Mass of Puck Dry (A)	4651.7	4649.7	2.0
Submerged Weight of Puck in Water (C)	2681.6	2677.1	
Wt. of Puck SSD (B)	4656.6	4653.5	
G_{mb} (Bulk Specific Gravity)	2.355	2.353	
Average G_{mb}	**2.354**		

$$G_{mb} = \frac{A}{B-C}$$

Range **0.003** Accpetable? (Within d2s precision) **YES**

Summary of Mix Properties

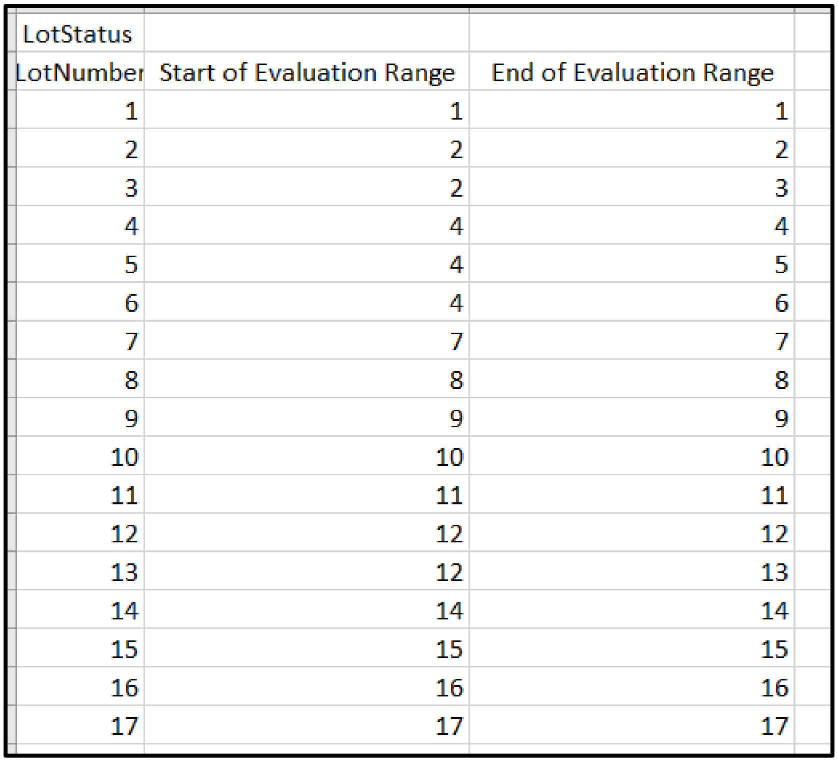
Property	Sample 1A	Sample 1B	Combined	LSL	USL
G_{sa}	2.656	2.656	2.656		
G_{se}	2.614	2.621	2.617		
G_{sb}	2.578	2.578	2.578		
G_{mm}	2.406	2.411	2.408		
G_{mb}	2.355	2.353	2.354		
Abs_{t166}	0.248	0.192	0.220		
G_b	1.0310	1.0310	1.0310		
P_b	5.65	5.65	5.65		
P_{ba}	0.55	0.65	0.60		
P_{be}	5.12	5.03	5.08		
P_s	94.4	94.4	94.4		
SA	32.2	32.2	32.2		
AFT	8.20	8.06	8.13		
P_a	2.09	2.41	2.3	3.0	5.0
VMA	13.80	13.90	13.8	14.0	
VFA	84.83	82.64	83.7	65.0	75.0
P200	6.02	6.02	6.0	3.8	6.8
DP	1.18	1.20	1.2	0.6	1.2

Figure 4.1 Typical data input file for asphalt pavement projects

Once the payment-related parameters are calculated for each test, tests are grouped to form a lot, and payments are calculated based on statistical tests on the lot data (which will be discussed later in this chapter).

Lot Grouping

Payment factors are calculated for each lot, but based on F and t tests from a group of tests that might include several lots. Grouping is done to enhance the diagnostic power of F and t tests. Statistical test results through this grouping practice determine whose test results should be used for payment purposes. For each lot, a few parameters define payment related calculations including “start of evaluation range” (lot number from where the evaluation would start) and “end of evaluation range” (lot number for which the payment would be calculated). For example, in Figure 4.2, for lot 2, the evaluation range started from lot 2 and also ended at 2. So, for this lot, no other lot is grouped for payment calculation. For lot 6, the evaluation started at lot 4 and ended at 6. So, all the tests from lots 4, 5, and 6 would be grouped together for payment calculation of lot 6.



LotStatus		
LotNumber	Start of Evaluation Range	End of Evaluation Range
1	1	1
2	2	2
3	2	3
4	4	4
5	4	5
6	4	6
7	7	7
8	8	8
9	9	9
10	10	10
11	11	11
12	12	12
13	12	13
14	14	14
15	15	15
16	16	16
17	17	17

Figure 4.2 Lot evaluation range for payment calculation

Test Statistics

Mean and standard deviation values for Air Voids, VMA, and Mainline Density of a lot group both from ITD and prime contractor-reported data are calculated first. From those values, a pass/fail test check is done using F and t tests. If p-values for both Air Voids and VMA and for both F and t tests (four combinations) are below 0.05, then they pass the test. If data is passed based on both F and t tests for both Air Voids and VMA, then the project lot gets a green signal, and prime contractor data is selected for payment. If in any of these tests, p-value exceeds 0.05 (rejected), then the test fails, and the prime contractor's data is unverified for the entire lot and is not used for acceptance and payment calculations; instead, the ITD-reported data is selected for calculating payment factor. As mentioned before, we refrain from pointing to prime contractor or ITD, and hence use entity 1 and 2, which might represent either of them.

Percent Within Limits (PWL)

The next step of the calculation of payment factor is the determination of PWL values. The lot average Pa, VMA, and Mainline Density values are considered, and through a series of calculations, PWL values are measured. The final payment factor for all three payment affecting parameters is computed through:

$$PF = \frac{55 + 0.5 \times PWL}{100}$$

Where:

$PF = Pay\ factor$
$PWL = Percent\ Within\ Limit$

The final payment value is then computed for the lot, using:

$$Payment = \ Composite\ pay\ factor \times Quantity\ represented\ by\ the\ lot \times Contract\ unit\ price$$

Where:

Composite pay factor = the weighted average of the individual pay factors

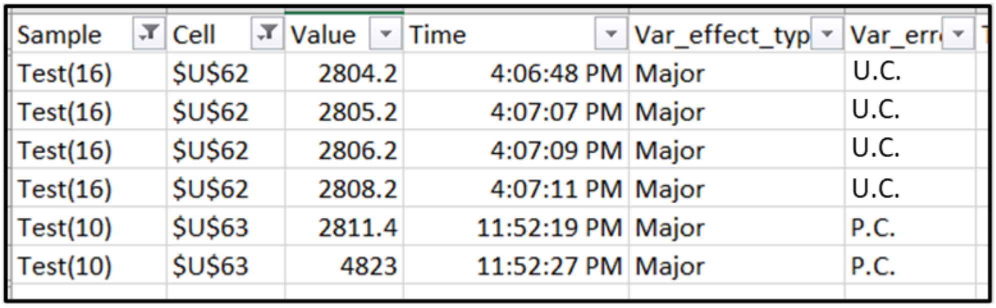
Where:

$Payment = Payment\ in\ USD$
"Quantity represented by lot" = Total volume of asphalt pavement produced in the lot
"Contract unit price" = Unit price to be paid to the prime contractor.

Formation of Input Data for Monetary Calculation

Two sets of data were created: first and last reported U.C. value, which was subsequently used for monetary impact analysis. With the hypothesis that the first "acceptable" U.C. value being the original value that was measured for a certain parameter, and the last value being the final reported value after corrections, the difference in payment calculations for these two cases is assumed to be the monetary impact of data correction in the material testing reports. As a reminder, three types of data are considered in this study: non-repeated data (one value is reported) and repeating data with P.C. and U.C. categorization (multiple data entry were recorded for each parameter). Since only the U.C. data can be held responsible for any sort of economic impact, first and last entry of the U.C. cases were selected. The P.C. and non-repeated cases do not have any influence on the monetary value, so they adopted their reported values. Also, any missing parameter value is assigned its reported value.

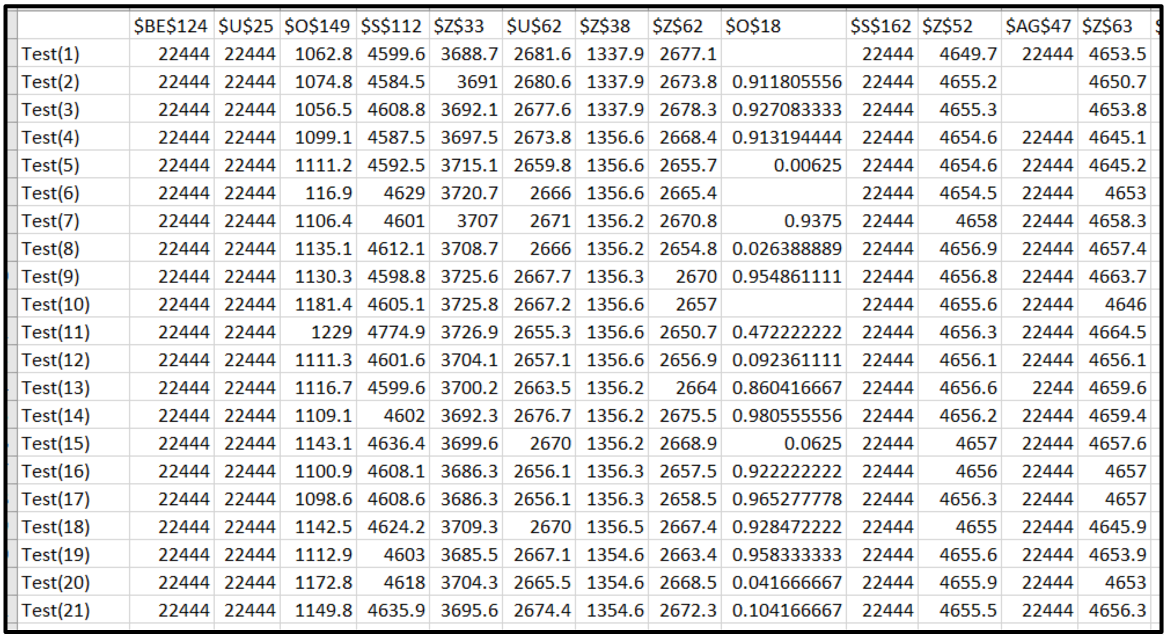
As an example, in Figure 4.3, cell U62 (Submerged weight of puck in water; specimen 1) from Test (16) has three repetitions with a total of four values and falls in the U.C. category. Hence, the first value of 2804.2 was selected for the first dataset (that was used for original payment calculation) and the last value of 2808.2 was selected for the second dataset (that was used for payment calculation after corrections). Cell U63 (Weight of puck SSD; specimen 1) from Test (10) falls in the P.C. category. So, the final value of 4823 was picked for both datasets. For non-repeated cells, the single corresponding value was kept for both datasets.



Sample	Cell	Value	Time	Var_effect_typ	Var_err
Test(16)	U62	2804.2	4:06:48 PM	Major	U.C.
Test(16)	U62	2805.2	4:07:07 PM	Major	U.C.
Test(16)	U62	2806.2	4:07:09 PM	Major	U.C.
Test(16)	U62	2808.2	4:07:11 PM	Major	U.C.
Test(10)	U63	2811.4	11:52:19 PM	Major	P.C.
Test(10)	U63	4823	11:52:27 PM	Major	P.C.

Figure 4.3 Sample classified plausible correction (P.C.) and unexplained correction (U.C.) data

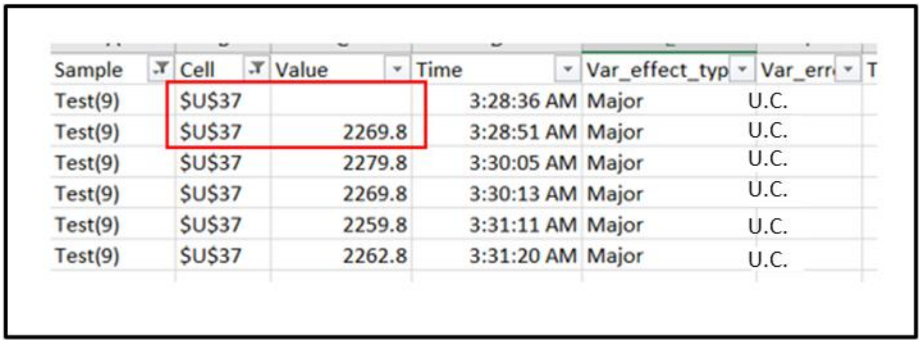
A Python code was generated to accomplish these steps. The code is designed to adopt the first and last values of U.C. and to take the last value of P.C. from the previously categorized audit trail data, and to take the final reported value for all non-repeating and missing variables. A sample of the newly generated dataset (which was subsequently used for payment calculation) is presented in Figure 4.4, where tests are presented in the rows and parameters/cells associated with each test in the columns.



	BE124	U25	O149	S112	Z33	U62	Z38	Z62	O18	S162	Z52	AG47	Z63
Test(1)	22444	22444	1062.8	4599.6	3688.7	2681.6	1337.9	2677.1		22444	4649.7	22444	4653.5
Test(2)	22444	22444	1074.8	4584.5	3691	2680.6	1337.9	2673.8	0.911805556	22444	4655.2		4650.7
Test(3)	22444	22444	1056.5	4608.8	3692.1	2677.6	1337.9	2678.3	0.927083333	22444	4655.3		4653.8
Test(4)	22444	22444	1099.1	4587.5	3697.5	2673.8	1356.6	2668.4	0.913194444	22444	4654.6	22444	4645.1
Test(5)	22444	22444	1111.2	4592.5	3715.1	2659.8	1356.6	2655.7	0.00625	22444	4654.6	22444	4645.2
Test(6)	22444	22444	116.9	4629	3720.7	2666	1356.6	2665.4		22444	4654.5	22444	4653
Test(7)	22444	22444	1106.4	4601	3707	2671	1356.2	2670.8	0.9375	22444	4658	22444	4658.3
Test(8)	22444	22444	1135.1	4612.1	3708.7	2666	1356.2	2654.8	0.026388889	22444	4656.9	22444	4657.4
Test(9)	22444	22444	1130.3	4598.8	3725.6	2667.7	1356.3	2670	0.954861111	22444	4656.8	22444	4663.7
Test(10)	22444	22444	1181.4	4605.1	3725.8	2667.2	1356.6	2657		22444	4655.6	22444	4646
Test(11)	22444	22444	1229	4774.9	3726.9	2655.3	1356.6	2650.7	0.472222222	22444	4656.3	22444	4664.5
Test(12)	22444	22444	1111.3	4601.6	3704.1	2657.1	1356.6	2656.9	0.092361111	22444	4656.1	22444	4656.1
Test(13)	22444	22444	1116.7	4599.6	3700.2	2663.5	1356.2	2664	0.860416667	22444	4656.6	2244	4659.6
Test(14)	22444	22444	1109.1	4602	3692.3	2676.7	1356.2	2675.5	0.980555556	22444	4656.2	22444	4659.4
Test(15)	22444	22444	1143.1	4636.4	3699.6	2670	1356.2	2668.9	0.0625	22444	4657	22444	4657.6
Test(16)	22444	22444	1100.9	4608.1	3686.3	2656.1	1356.3	2657.5	0.922222222	22444	4656	22444	4657
Test(17)	22444	22444	1098.6	4608.6	3686.3	2656.1	1356.3	2658.5	0.965277778	22444	4656.3	22444	4657
Test(18)	22444	22444	1142.5	4624.2	3709.3	2670	1356.5	2667.4	0.928472222	22444	4655	22444	4645.9
Test(19)	22444	22444	1112.9	4603	3685.5	2667.1	1354.6	2663.4	0.958333333	22444	4655.6	22444	4653.9
Test(20)	22444	22444	1172.8	4618	3704.3	2665.5	1354.6	2668.5	0.041666667	22444	4655.9	22444	4653
Test(21)	22444	22444	1149.8	4635.9	3695.6	2674.4	1354.6	2672.3	0.104166667	22444	4655.5	22444	4656.3

Figure 4.4 Input dataset for monetary calculation: rows show test number and columns represent parameter values associated with each test

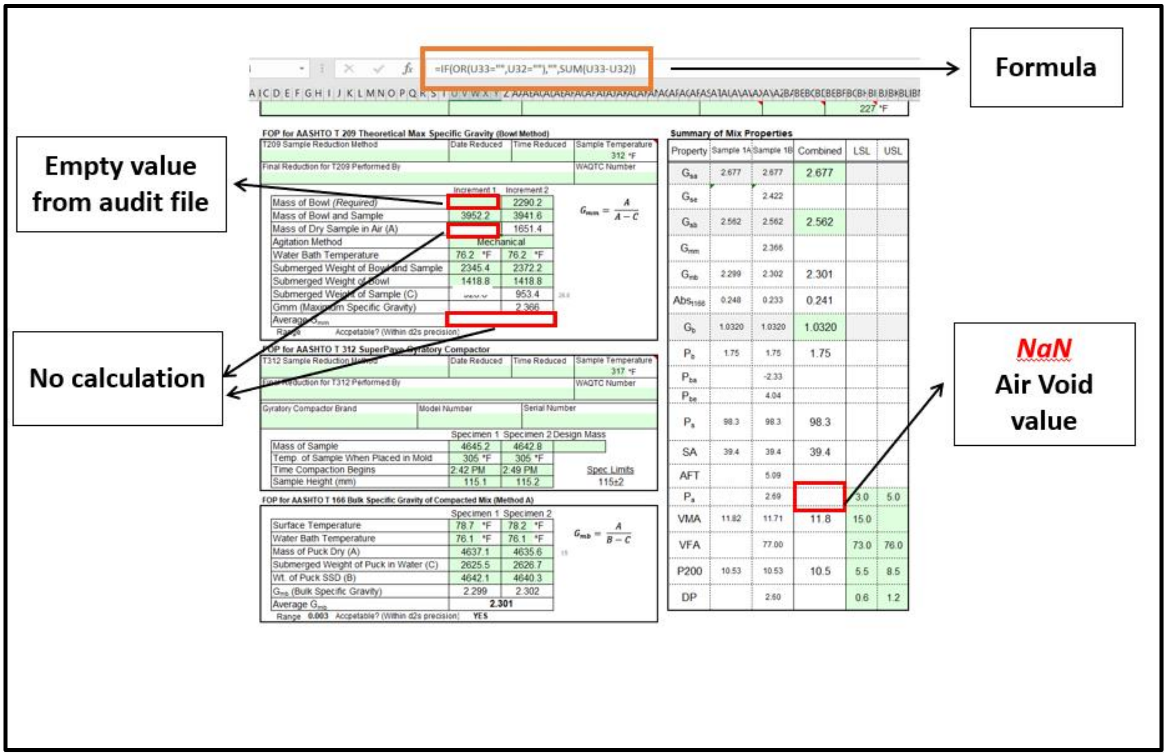
This step, however, was associated with some challenges. There were instances in which the first or last U.C. and last P.C. data had an empty cell, which precluded us from calculating monetary values. These empty cells created unreasonably large, negative or NaN values for the target parameters (Air Voids/VMA/Mainline Density). Hence, some strategies were devised to fill empty values. For the first entry, if the value was empty, the second cell value was selected; if the second was empty, the algorithm looks for the next one and continued until it found a value. A similar process was done for obtaining the value of the last cell but in a reverse order. The cell value before the last cell was selected if the last one was empty. These steps were continued from the last cell backwards until the algorithm found a value. Figure 4.5 demonstrates a missing first entry for cell U37 (Submerged weight of bowl and sample; increment 1), for which the next value was adopted.



Sample	Cell	Value	Time	Var_effect_typ	Var_err	T
Test(9)	U37		3:28:36 AM	Major	U.C.	
Test(9)	U37	2269.8	3:28:51 AM	Major	U.C.	
Test(9)	U37	2279.8	3:30:05 AM	Major	U.C.	
Test(9)	U37	2269.8	3:30:13 AM	Major	U.C.	
Test(9)	U37	2259.8	3:31:11 AM	Major	U.C.	
Test(9)	U37	2262.8	3:31:20 AM	Major	U.C.	

Figure 4.5 Empty cell for some parameters

Figure 4.6 demonstrates an example problem associated with having a NaN value for an input parameter. Because there were missing values for one of the cells, several calculations were not possible and resulted in NaN value (and/or empty cell) for Air Voids. Since secondary parameter values (payment-related parameters) depend on various primary parameters, lack of primary parameter values will preclude calculation of secondary values.



`=IF(OR(U33="",U32=""),"",SUM(U33-U32))`

FOP for AASHTO T 209 Theoretical Max Specific Gravity (Bowl Method)

T209 Sample Reduction Method	Date Reduced	Time Reduced	Sample Temperature 312 °F
Final Reduction for T209 Performed By			WAQTC Number

	Increment 1	Increment 2
Mass of Bowl *(Required)*		2290.2
Mass of Bowl and Sample	3952.2	3941.6
Mass of Dry Sample in Air (A)		1651.4
Agitation Method	Mechanical	
Water Bath Temperature	76.2 °F	76.2 °F
Submerged Weight of Bowl and Sample	2345.4	2372.2
Submerged Weight of Bowl	1418.8	1418.8
Submerged Weight of Sample (C)	[illegible]	953.4
Gmm (Maximum Specific Gravity)		2.366
Average G_{mm}		

Range Accpetable? (Within d2s precision)

$$G_{mm} = \frac{A}{A-C}$$

FOP for AASHTO T 312 SuperPave Gyratory Compactor

T312 Sample Reduction Method	Date Reduced	Time Reduced	Sample Temperature 317 °F
Final Reduction for T312 Performed By			WAQTC Number

Gyratory Compactor Brand	Model Number	Serial Number

	Specimen 1	Specimen 2	Design Mass
Mass of Sample	4645.2	4642.8	
Temp. of Sample When Placed in Mold	305 °F	305 °F	
Time Compaction Begins	2:42 PM	2:49 PM	
Sample Height (mm)	115.1	115.2	

Spec Limits
115±2

FOP for AASHTO T 166 Bulk Specific Gravity of Compacted Mix (Method A)

	Specimen 1	Specimen 2
Surface Temperature	78.7 °F	78.2 °F
Water Bath Temperature	76.1 °F	76.1 °F
Mass of Puck Dry (A)	4637.1	4635.6
Submerged Weight of Puck in Water (C)	2625.5	2626.7
Wt. of Puck SSD (B)	4642.1	4640.3
G_{mb} (Bulk Specific Gravity)	2.299	2.302
Average G_{mb}	2.301	

Range 0.003 Accpetable? (Within d2s precision) YES

$$G_{mb} = \frac{A}{B-C}$$

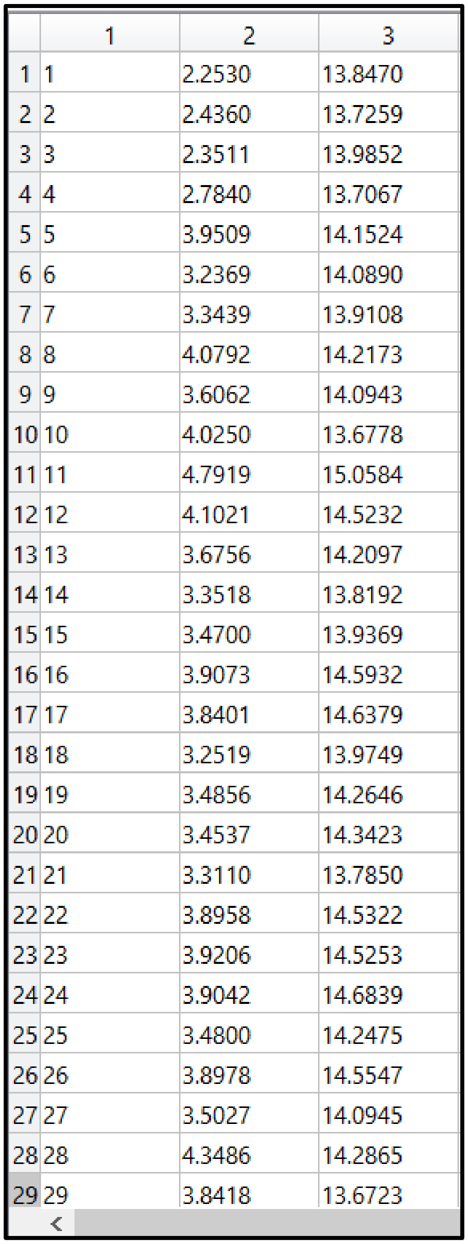
227 °F

Summary of Mix Properties

Property	Sample 1A	Sample 1B	Combined	LSL	USL
G_{sa}	2.677	2.677	2.677		
G_{se}		2.422			
G_{sb}	2.562	2.562	2.562		
G_{mm}		2.366			
G_{mb}	2.299	2.302	2.301		
Abs_{1166}	0.248	0.233	0.241		
G_b	1.0320	1.0320	1.0320		
P_b	1.75	1.75	1.75		
P_{ba}		-2.33			
P_{be}		4.04			
P_s	98.3	98.3	98.3		
SA	39.4	39.4	39.4		
AFT		5.09			
P_a		2.69		3.0	5.0
VMA	11.82	11.71	11.8	15.0	
VFA		77.00		73.0	76.0
P200	10.53	10.53	10.5	5.5	8.5
DP		2.60		0.6	1.2

Figure 4.6 Effect of empty/NaN cells on calculated payment-related parameters

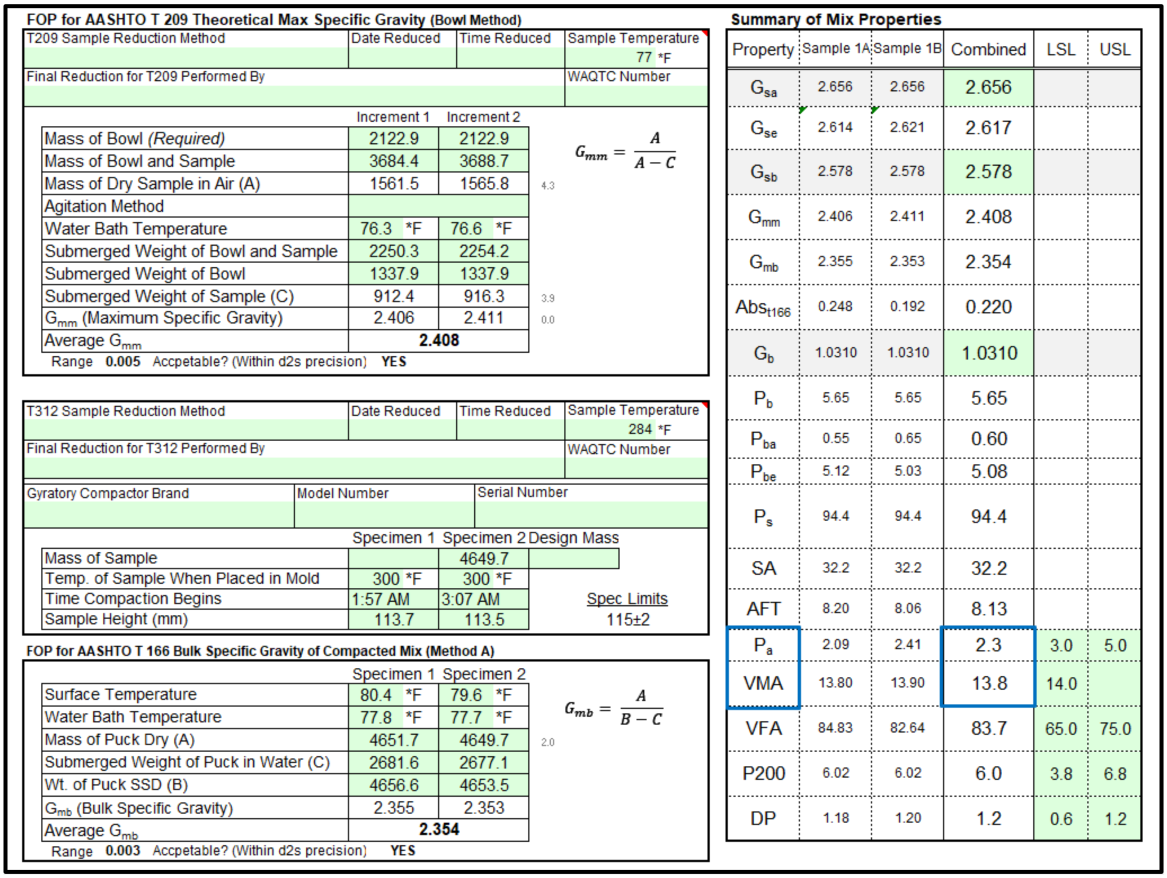
This final dataset was used to test and apply the formulas from “ITD-0777” (see Appendix C) to evaluate the monetary values (payments based on first and last U.C. values). A code was prepared in MATLAB in this step, which replicated the original calculation flow of the “ITD-0777” file and calculated Air Voids, VMA, and Mainline Density values. To ensure the accuracy of calculations, another code was prepared at this step to plug in the parameter values directly into the “ITD-0777” file. This enabled calculation of the parameter values from the coded program and from the “ITD-0777” file. Several projects were cross-checked to ensure the calculated monetary values through the developed code and “ITD-0777” file were exactly matching. The direct monetary calculation through the developed code was much faster as it could automatically produce all the test parameters of a project. Figure 4.7 demonstrates the calculation of Air voids and VMA values for each test of a sample project through the direct calculation in the code. Similar values were obtained from the “ITD-0777” file.



	1	2	3
1	1	2.2530	13.8470
2	2	2.4360	13.7259
3	3	2.3511	13.9852
4	4	2.7840	13.7067
5	5	3.9509	14.1524
6	6	3.2369	14.0890
7	7	3.3439	13.9108
8	8	4.0792	14.2173
9	9	3.6062	14.0943
10	10	4.0250	13.6778
11	11	4.7919	15.0584
12	12	4.1021	14.5232
13	13	3.6756	14.2097
14	14	3.3518	13.8192
15	15	3.4700	13.9369
16	16	3.9073	14.5932
17	17	3.8401	14.6379
18	18	3.2519	13.9749
19	19	3.4856	14.2646
20	20	3.4537	14.3423
21	21	3.3110	13.7850
22	22	3.8958	14.5322
23	23	3.9206	14.5253
24	24	3.9042	14.6839
25	25	3.4800	14.2475
26	26	3.8978	14.5547
27	27	3.5027	14.0945
28	28	4.3486	14.2865
29	29	3.8418	13.6723

Figure 4.7 Calculated Air Voids (column 2) and VMA (column 3) for an example project

Figure 4.8 shows an example “ITD-0777” file where all the input values have been inserted, and calculations were done by the internal formulas of this sheet. Since this procedure is lengthy and can only be done for one test at a time, the developed code that replicates “ITD-0777” file was used for the remainder of the analysis. However, 3 tests were randomly selected from each project to cross-check individual test results with the previously discussed code produced results.



FOP for AASHTO T 209 Theoretical Max Specific Gravity (Bowl Method)

T209 Sample Reduction Method	Date Reduced	Time Reduced	Sample Temperature
			77 *F
Final Reduction for T209 Performed By			WAQTC Number

$$G_{mm} = \frac{A}{A - C}$$

	Increment 1	Increment 2	
Mass of Bowl *(Required)*	2122.9	2122.9	
Mass of Bowl and Sample	3684.4	3688.7	
Mass of Dry Sample in Air (A)	1561.5	1565.8	4.3
Agitation Method			
Water Bath Temperature	76.3 *F	76.6 *F	
Submerged Weight of Bowl and Sample	2250.3	2254.2	
Submerged Weight of Bowl	1337.9	1337.9	
Submerged Weight of Sample (C)	912.4	916.3	3.9
G_{mm} (Maximum Specific Gravity)	2.406	2.411	0.0
Average G_{mm}	2.408		

Range **0.005** Accpetable? (Within d2s precision) **YES**

T312 Sample Reduction Method	Date Reduced	Time Reduced	Sample Temperature
			284 *F
Final Reduction for T312 Performed By			WAQTC Number

Gyratory Compactor Brand	Model Number	Serial Number

	Specimen 1	Specimen 2	Design Mass
Mass of Sample		4649.7	
Temp. of Sample When Placed in Mold	300 *F	300 *F	
Time Compaction Begins	1:57 AM	3:07 AM	Spec Limits
Sample Height (mm)	113.7	113.5	115±2

FOP for AASHTO T 166 Bulk Specific Gravity of Compacted Mix (Method A)

$$G_{mb} = \frac{A}{B - C}$$

	Specimen 1	Specimen 2	
Surface Temperature	80.4 *F	79.6 *F	
Water Bath Temperature	77.8 *F	77.7 *F	
Mass of Puck Dry (A)	4651.7	4649.7	2.0
Submerged Weight of Puck in Water (C)	2681.6	2677.1	
Wt. of Puck SSD (B)	4656.6	4653.5	
G_{mb} (Bulk Specific Gravity)	2.355	2.353	
Average G_{mb}	2.354		

Range **0.003** Accpetable? (Within d2s precision) **YES**

Summary of Mix Properties

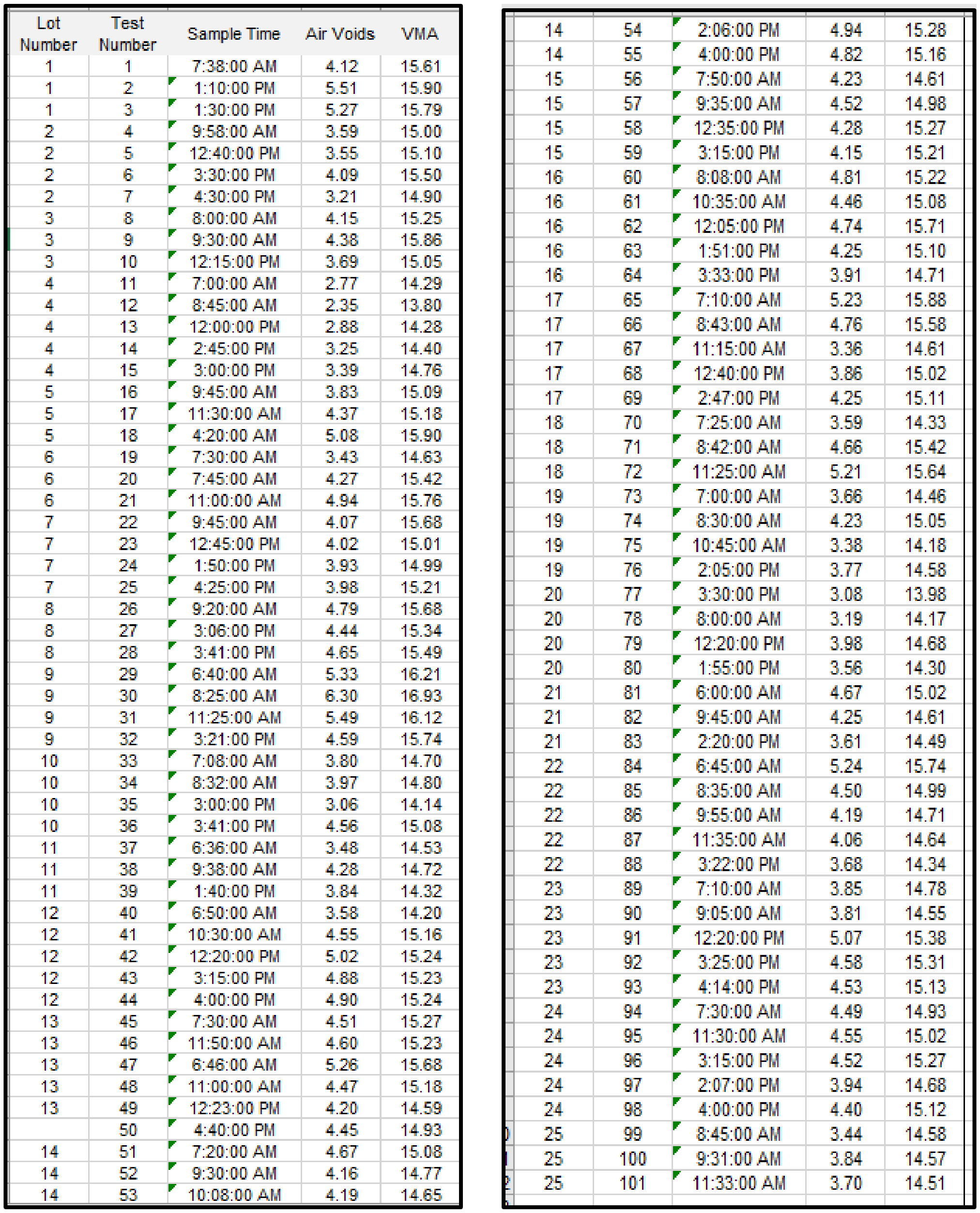
Property	Sample 1A	Sample 1B	Combined	LSL	USL
G_{sa}	2.656	2.656	2.656		
G_{se}	2.614	2.621	2.617		
G_{sb}	2.578	2.578	2.578		
G_{mm}	2.406	2.411	2.408		
G_{mb}	2.355	2.353	2.354		
Abs_{t166}	0.248	0.192	0.220		
G_b	1.0310	1.0310	1.0310		
P_b	5.65	5.65	5.65		
P_{ba}	0.55	0.65	0.60		
P_{be}	5.12	5.03	5.08		
P_s	94.4	94.4	94.4		
SA	32.2	32.2	32.2		
AFT	8.20	8.06	8.13		
P_a	2.09	2.41	2.3	3.0	5.0
VMA	13.80	13.90	13.8	14.0	
VFA	84.83	82.64	83.7	65.0	75.0
P200	6.02	6.02	6.0	3.8	6.8
DP	1.18	1.20	1.2	0.6	1.2

Figure 4.8 Calculation of air voids and VMA through ITD-0777 file

Unavailability of Audit Trail Files

Unfortunately, not all projects included audit trail files. Further, on many occasions, the audit trail files did not have the recorded values for all the tests of a project. Sometimes there were no audit trail files for neither entity 1 nor entity 2-reported data. Since data from both entities are needed for monetary calculation, the reported values from the project files that missed audit trail values were considered.

Project #1 shown in Figure 4.9, for example, has a total of 101 tests from the entity 1-reported data, while in the audit trail file only provided data for 52 tests (Figure 4.10). All tests in the audit trail file from Test (1) to Test (50) were missing except for Test (47). For the monetary calculations, the reported values for the missing tests were used. The reported values were exactly the same in both input datasets, so they did not induce any monetary difference. But the available tests in the audit trail file showed a significant difference in the monetary values (shown later).



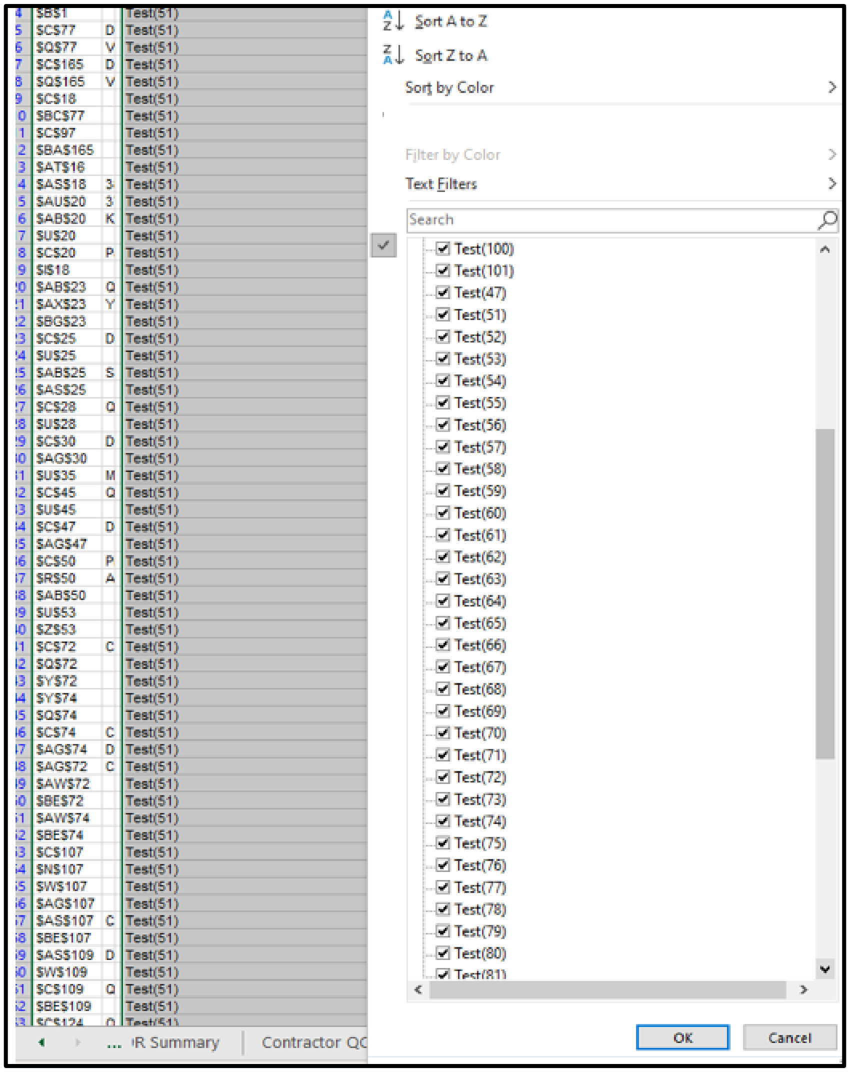
Lot Number	Test Number	Sample Time	Air Voids	VMA
1	1	7:38:00 AM	4.12	15.61
1	2	1:10:00 PM	5.51	15.90
1	3	1:30:00 PM	5.27	15.79
2	4	9:58:00 AM	3.59	15.00
2	5	12:40:00 PM	3.55	15.10
2	6	3:30:00 PM	4.09	15.50
2	7	4:30:00 PM	3.21	14.90
3	8	8:00:00 AM	4.15	15.25
3	9	9:30:00 AM	4.38	15.86
3	10	12:15:00 PM	3.69	15.05
4	11	7:00:00 AM	2.77	14.29
4	12	8:45:00 AM	2.35	13.80
4	13	12:00:00 PM	2.88	14.28
4	14	2:45:00 PM	3.25	14.40
4	15	3:00:00 PM	3.39	14.76
5	16	9:45:00 AM	3.83	15.09
5	17	11:30:00 AM	4.37	15.18
5	18	4:20:00 AM	5.08	15.90
6	19	7:30:00 AM	3.43	14.63
6	20	7:45:00 AM	4.27	15.42
6	21	11:00:00 AM	4.94	15.76
7	22	9:45:00 AM	4.07	15.68
7	23	12:45:00 PM	4.02	15.01
7	24	1:50:00 PM	3.93	14.99
7	25	4:25:00 PM	3.98	15.21
8	26	9:20:00 AM	4.79	15.68
8	27	3:06:00 PM	4.44	15.34
8	28	3:41:00 PM	4.65	15.49
9	29	6:40:00 AM	5.33	16.21
9	30	8:25:00 AM	6.30	16.93
9	31	11:25:00 AM	5.49	16.12
9	32	3:21:00 PM	4.59	15.74
10	33	7:08:00 AM	3.80	14.70
10	34	8:32:00 AM	3.97	14.80
10	35	3:00:00 PM	3.06	14.14
10	36	3:41:00 PM	4.56	15.08
11	37	6:36:00 AM	3.48	14.53
11	38	9:38:00 AM	4.28	14.72
11	39	1:40:00 PM	3.84	14.32
12	40	6:50:00 AM	3.58	14.20
12	41	10:30:00 AM	4.55	15.16
12	42	12:20:00 PM	5.02	15.24
12	43	3:15:00 PM	4.88	15.23
12	44	4:00:00 PM	4.90	15.24
13	45	7:30:00 AM	4.51	15.27
13	46	11:50:00 AM	4.60	15.23
13	47	6:46:00 AM	5.26	15.68
13	48	11:00:00 AM	4.47	15.18
13	49	12:23:00 PM	4.20	14.59
	50	4:40:00 PM	4.45	14.93
14	51	7:20:00 AM	4.67	15.08
14	52	9:30:00 AM	4.16	14.77
14	53	10:08:00 AM	4.19	14.65

(a)

14	54	2:06:00 PM	4.94	15.28
14	55	4:00:00 PM	4.82	15.16
15	56	7:50:00 AM	4.23	14.61
15	57	9:35:00 AM	4.52	14.98
15	58	12:35:00 PM	4.28	15.27
15	59	3:15:00 PM	4.15	15.21
16	60	8:08:00 AM	4.81	15.22
16	61	10:35:00 AM	4.46	15.08
16	62	12:05:00 PM	4.74	15.71
16	63	1:51:00 PM	4.25	15.10
16	64	3:33:00 PM	3.91	14.71
17	65	7:10:00 AM	5.23	15.88
17	66	8:43:00 AM	4.76	15.58
17	67	11:15:00 AM	3.36	14.61
17	68	12:40:00 PM	3.86	15.02
17	69	2:47:00 PM	4.25	15.11
18	70	7:25:00 AM	3.59	14.33
18	71	8:42:00 AM	4.66	15.42
18	72	11:25:00 AM	5.21	15.64
19	73	7:00:00 AM	3.66	14.46
19	74	8:30:00 AM	4.23	15.05
19	75	10:45:00 AM	3.38	14.18
19	76	2:05:00 PM	3.77	14.58
20	77	3:30:00 PM	3.08	13.98
20	78	8:00:00 AM	3.19	14.17
20	79	12:20:00 PM	3.98	14.68
20	80	1:55:00 PM	3.56	14.30
21	81	6:00:00 AM	4.67	15.02
21	82	9:45:00 AM	4.25	14.61
21	83	2:20:00 PM	3.61	14.49
22	84	6:45:00 AM	5.24	15.74
22	85	8:35:00 AM	4.50	14.99
22	86	9:55:00 AM	4.19	14.71
22	87	11:35:00 AM	4.06	14.64
22	88	3:22:00 PM	3.68	14.34
23	89	7:10:00 AM	3.85	14.78
23	90	9:05:00 AM	3.81	14.55
23	91	12:20:00 PM	5.07	15.38
23	92	3:25:00 PM	4.58	15.31
23	93	4:14:00 PM	4.53	15.13
24	94	7:30:00 AM	4.49	14.93
24	95	11:30:00 AM	4.55	15.02
24	96	3:15:00 PM	4.52	15.27
24	97	2:07:00 PM	3.94	14.68
24	98	4:00:00 PM	4.40	15.12
25	99	8:45:00 AM	3.44	14.58
25	100	9:31:00 AM	3.84	14.57
25	101	11:33:00 AM	3.70	14.51

(b)

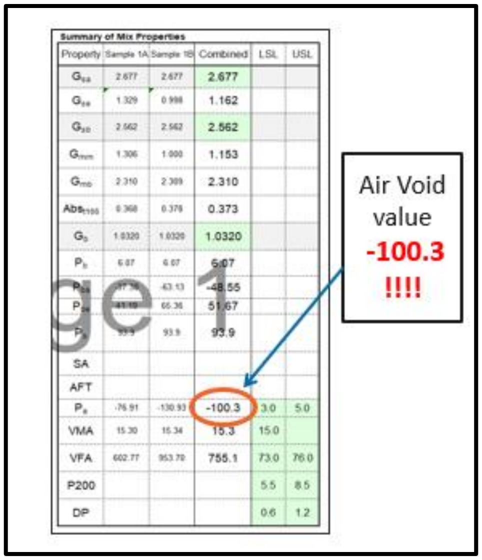
Figure 4.9 Total number of tests done for an example project (a) Test numbers 1-53 (b) Test numbers 54-101 (project 1)



B1		Test(51)
C77	D	Test(51)
Q77	V	Test(51)
C165	D	Test(51)
Q165	V	Test(51)
C18		Test(51)
BC77		Test(51)
C97		Test(51)
BA165		Test(51)
AT16		Test(51)
AS18	3	Test(51)
AU20	3	Test(51)
AB20	K	Test(51)
U20		Test(51)
C20	P	Test(51)
I18		Test(51)
AB23	Q	Test(51)
AX23	Y	Test(51)
BG23		Test(51)
C25	D	Test(51)
U25		Test(51)
AB25	S	Test(51)
AS25		Test(51)
C28	Q	Test(51)
U28		Test(51)
C30	D	Test(51)
AG30		Test(51)
U35	M	Test(51)
C45	Q	Test(51)
U45		Test(51)
C47	D	Test(51)
AG47		Test(51)
C50	P	Test(51)
R50	A	Test(51)
AB50		Test(51)
U53		Test(51)
Z53		Test(51)
C72	C	Test(51)
Q72		Test(51)
Y72		Test(51)
Y74		Test(51)
Q74		Test(51)
C74	C	Test(51)
AG74	D	Test(51)
AG72	C	Test(51)
AW72		Test(51)
BE72		Test(51)
AW74		Test(51)
BE74		Test(51)
C107		Test(51)
N107		Test(51)
W107		Test(51)
AG107		Test(51)
AS107	C	Test(51)
BE107		Test(51)
AS109	D	Test(51)
W109		Test(51)
C109	Q	Test(51)
BE109		Test(51)
C124	Q	Test(51)

Figure 4.10 Available tests in the audit trail file for an example project

Although the empty cells and missing audit trail values were successfully filled with the reported ones, we encountered some issues while trying to calculate the payment-related parameters. Issues included negative and unreasonably large secondary parameter values calculated based on the first U.C. primary parameters. Figure 4.11 shows an example of attempted monetary parameter calculation, where large negative Air Voids values was observed even after removing all the empty cells from the input parameter set.



Summary of Mix Properties

Property	Sample 1A	Sample 1B	Combined	LSL	USL
G_{sa}	2.677	2.677	2.677		
G_{se}	1.329	0.998	1.162		
G_{sb}	2.562	2.562	2.562		
G_{mm}	1.306	1.000	1.153		
G_{mb}	2.310	2.309	2.310		
Abs_{t100}	0.368	0.378	0.373		
G_b	1.0320	1.0320	1.0320		
P_b	6.07	6.07	6.07		
P_{ba}	[illegible]	-63.13	-48.55		
P_{be}	[illegible]	65.36	51.67		
P_s	93.9	93.9	93.9		
SA					
AFT					
P_a	-76.91	-130.93	-100.3	3.0	5.0
VMA	15.30	15.34	15.3	15.0	
VFA	602.77	953.70	755.1	73.0	76.0
P200				5.5	8.5
DP				0.6	1.2

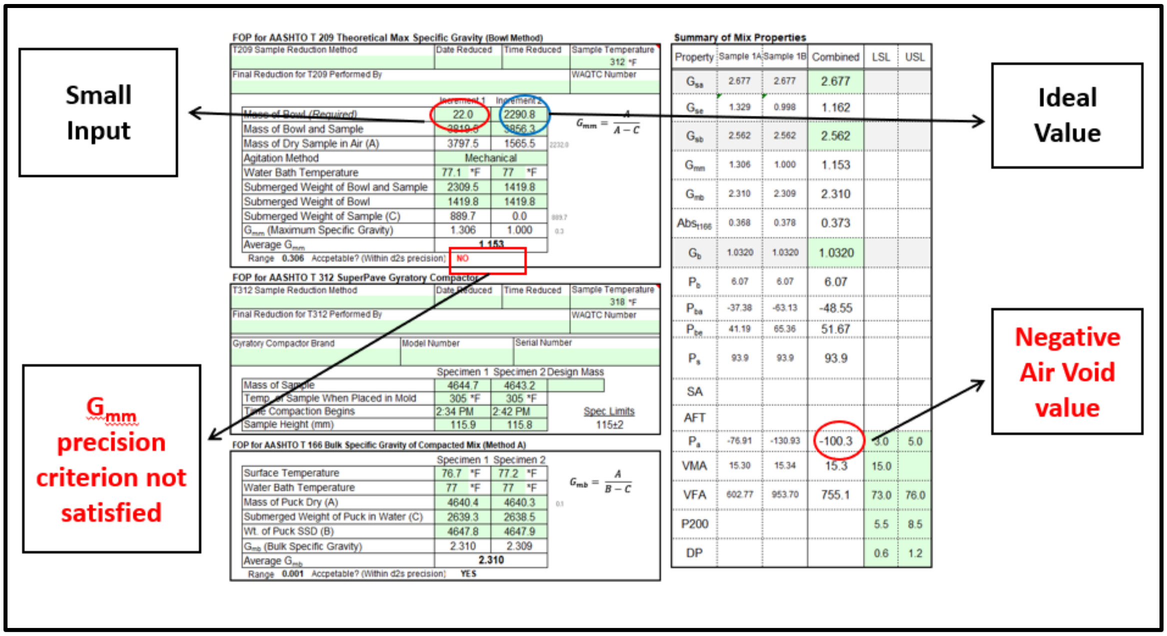
Air Void value **-100.3 !!!!**

Figure 4.11 Calculated negative air voids value

We investigated the sources of those negative and unreasonably large values by referring back to the ITD-0777 source file. Through trial and error, the reasons for getting those unusual values were discovered, which are presented under different cases below.

- Case1

The first case was an input that was unreasonably smaller than an ideal/original value for a parameter (Figures 4.12, and 4.13). Figure 4.12 shows that the mass of bowl for increment 1 has a value that was far lower than its original value, whereas the value for increment 2 was much closer to its ideal value. The smaller input resulted in a large negative Air voids value. Similarly, on other occasions, with inputs that have values much smaller than original values, positive, but unreasonably large, Air voids values were observed.



FOP for AASHTO T 209 Theoretical Max Specific Gravity (Bowl Method)

T209 Sample Reduction Method	Date Reduced	Time Reduced	Sample Temperature
			312 °F

Final Reduction for T209 Performed By	WAQTC Number

	Increment 1	Increment 2
Mass of Bowl *(Required)*	22.0	2290.8
Mass of Bowl and Sample	3819.5	3856.3
Mass of Dry Sample in Air (A)	3797.5	1565.5
Agitation Method	Mechanical	
Water Bath Temperature	77.1 °F	77 °F
Submerged Weight of Bowl and Sample	2309.5	1419.8
Submerged Weight of Bowl	1419.8	1419.8
Submerged Weight of Sample (C)	889.7	0.0
G_{mm} (Maximum Specific Gravity)	1.306	1.000
Average G_{mm}	1.153	

$G_{mm} = \frac{A}{A-C}$

Range 0.306 Accpetable? (Within d2s precision) NO

FOP for AASHTO T 312 SuperPave Gyratory Compactor

T312 Sample Reduction Method	Date Reduced	Time Reduced	Sample Temperature
			318 °F

Final Reduction for T312 Performed By	WAQTC Number

Gyratory Compactor Brand	Model Number	Serial Number

	Specimen 1	Specimen 2	Design Mass
Mass of Sample	4644.7	4643.2	
Temp. of Sample When Placed in Mold	305 °F	305 °F	
Time Compaction Begins	2:34 PM	2:42 PM	
Sample Height (mm)	115.9	115.8	

Spec Limits
115±2

FOP for AASHTO T 166 Bulk Specific Gravity of Compacted Mix (Method A)

	Specimen 1	Specimen 2
Surface Temperature	76.7 °F	77.2 °F
Water Bath Temperature	77 °F	77 °F
Mass of Puck Dry (A)	4640.4	4640.3
Submerged Weight of Puck in Water (C)	2639.3	2638.5
Wt. of Puck SSD (B)	4647.8	4647.9
G_{mb} (Bulk Specific Gravity)	2.310	2.309
Average G_{mb}	2.310	

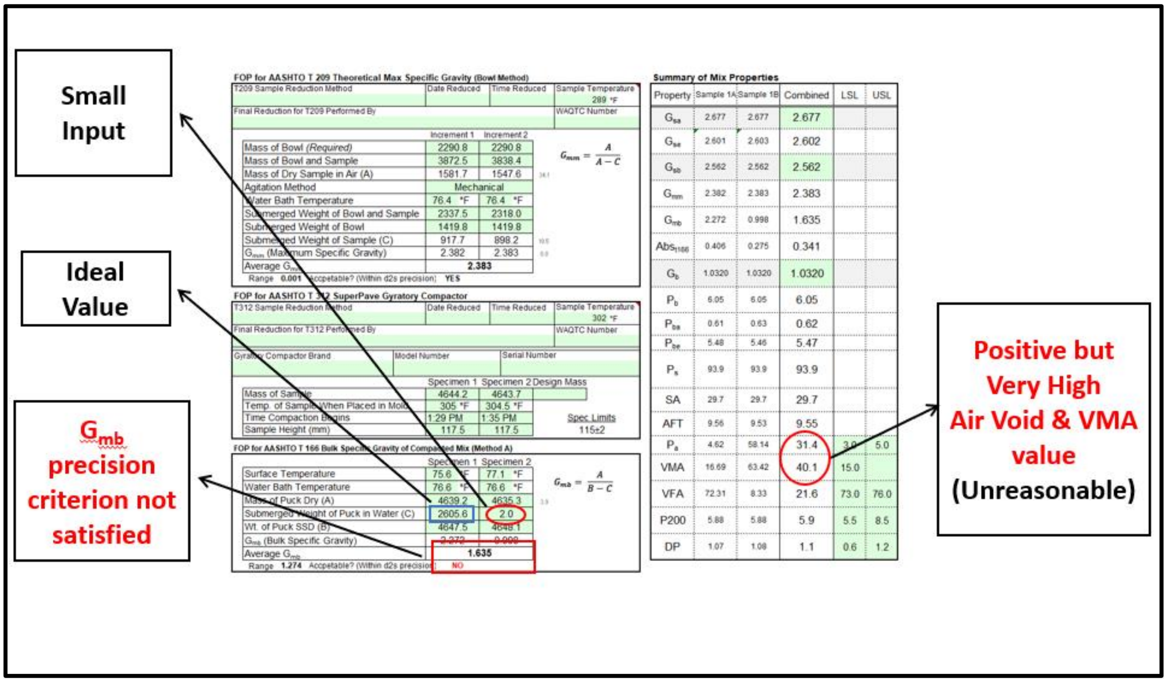
$G_{mb} = \frac{A}{B-C}$

Range 0.001 Accpetable? (Within d2s precision) YES

Summary of Mix Properties

Property	Sample 1A	Sample 1B	Combined	LSL	USL
G_{sa}	2.677	2.677	2.677		
G_{se}	1.329	0.998	1.162		
G_{sb}	2.562	2.562	2.562		
G_{mm}	1.306	1.000	1.153		
G_{mb}	2.310	2.309	2.310		
Abs_{t166}	0.368	0.378	0.373		
G_b	1.0320	1.0320	1.0320		
P_b	6.07	6.07	6.07		
P_{ba}	-37.38	-63.13	-48.55		
P_{be}	41.19	65.36	51.67		
P_s	93.9	93.9	93.9		
SA					
AFT					
P_a	-76.91	-130.93	-100.3	3.0	5.0
VMA	15.30	15.34	15.3	15.0	
VFA	602.77	953.70	755.1	73.0	76.0
P200				5.5	8.5
DP				0.6	1.2

Figure 4.12 Calculated negative air voids value due to unreasonably small primary parameter



FOP for AASHTO T 209 Theoretical Max Specific Gravity (Bowl Method)

T209 Sample Reduction Method	Date Reduced	Time Reduced	Sample Temperature
			289 °F

Final Reduction for T209 Performed By	WAQTC Number

	Increment 1	Increment 2
Mass of Bowl *(Required)*	2290.8	2290.8
Mass of Bowl and Sample	3872.5	3838.4
Mass of Dry Sample in Air (A)	1581.7	1547.6
Agitation Method	Mechanical	
Water Bath Temperature	76.4 °F	76.4 °F
Submerged Weight of Bowl and Sample	2337.5	2318.0
Submerged Weight of Bowl	1419.8	1419.8
Submerged Weight of Sample (C)	917.7	898.2
G_{mm} (Maximum Specific Gravity)	2.382	2.383
Average G_{mm}	2.383	

$G_{mm} = \frac{A}{A-C}$

Range 0.001 Accpetable? (Within d2s precision) YES

FOP for AASHTO T 312 SuperPave Gyratory Compactor

T312 Sample Reduction Method	Date Reduced	Time Reduced	Sample Temperature
			302 °F

Final Reduction for T312 Performed By	WAQTC Number

Gyratory Compactor Brand	Model Number	Serial Number

	Specimen 1	Specimen 2	Design Mass
Mass of Sample	4644.2	4643.7	
Temp. of Sample When Placed in Mold	305 °F	304.5 °F	
Time Compaction Begins	1:29 PM	1:35 PM	
Sample Height (mm)	117.5	117.5	

Spec Limits
115±2

FOP for AASHTO T 166 Bulk Specific Gravity of Compacted Mix (Method A)

	Specimen 1	Specimen 2
Surface Temperature	75.6 °F	77.1 °F
Water Bath Temperature	76.6 °F	76.6 °F
Mass of Puck Dry (A)	4639.2	4635.3
Submerged Weight of Puck in Water (C)	2605.6	2.0
Wt. of Puck SSD (B)	4647.5	4648.1
G_{mb} (Bulk Specific Gravity)	2.272	0.998
Average G_{mb}	1.635	

$G_{mb} = \frac{A}{B-C}$

Range 1.274 Accpetable? (Within d2s precision) NO

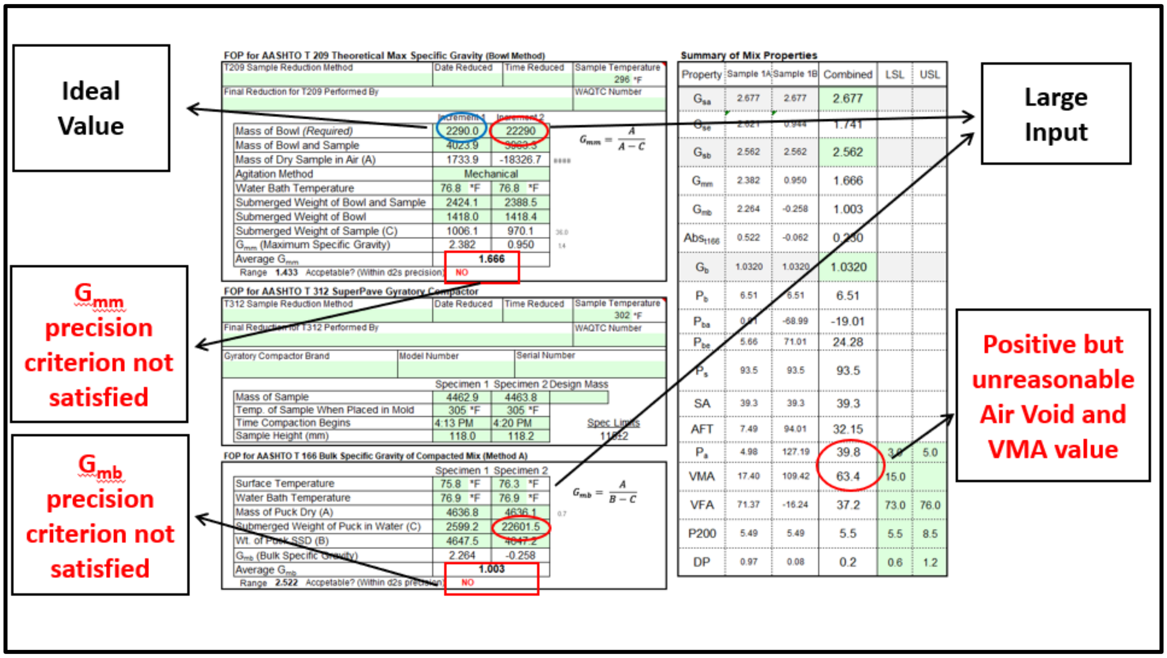
Summary of Mix Properties

Property	Sample 1A	Sample 1B	Combined	LSL	USL
G_{sa}	2.677	2.677	2.677		
G_{se}	2.601	2.603	2.602		
G_{sb}	2.562	2.562	2.562		
G_{mm}	2.382	2.383	2.383		
G_{mb}	2.272	0.998	1.635		
Abs_{t166}	0.406	0.275	0.341		
G_b	1.0320	1.0320	1.0320		
P_b	6.05	6.05	6.05		
P_{ba}	0.61	0.63	0.62		
P_{be}	5.48	5.46	5.47		
P_s	93.9	93.9	93.9		
SA	29.7	29.7	29.7		
AFT	9.56	9.53	9.55		
P_a	4.62	58.14	31.4	3.0	5.0
VMA	16.69	63.42	40.1	15.0	
VFA	72.31	8.33	21.6	73.0	76.0
P200	5.88	5.88	5.9	5.5	8.5
DP	1.07	1.08	1.1	0.6	1.2

Figure 4.13 Calculated unreasonable air voids and VMA values from unreasonably small input

- Case2

In some tests, unreasonably large primary parameter values were observed producing unreasonable secondary parameters (Figure 4.14). For example, the mass of bowl for increment 2 was 22290, which was much higher than the ideal value (2290). This directly affected the Air Voids calculation, which took a value that was much higher than expected. The value of 22290 was a typing error value, which in this case, was the last typing error value. The audit trail file recorded this value as the final reported value, which obviously cannot be used for monetary calculation. In this case, either the previous/succeeding reasonable parameter value from the audit trail file was adopted, or if this was not possible, the final reported value for this parameter was used.



FOP for AASHTO T 209 Theoretical Max Specific Gravity (Bowl Method)

T209 Sample Reduction Method	Date Reduced	Time Reduced	Sample Temperature
			296 °F

Final Reduction for T209 Performed By
WAQTC Number

	Increment 1	Increment 2
Mass of Bowl *(Required)*	2290.0	22290
Mass of Bowl and Sample	4023.9	3963.3
Mass of Dry Sample in Air (A)	1733.9	-18326.7
Agitation Method	Mechanical	
Water Bath Temperature	76.8 °F	76.8 °F
Submerged Weight of Bowl and Sample	2424.1	2388.5
Submerged Weight of Bowl	1418.0	1418.4
Submerged Weight of Sample (C)	1006.1	970.1
G_{mm} (Maximum Specific Gravity)	2.382	0.950
Average G_{mm}	1.666	

Range **1.433** Accpetable? (Within d2s precision) NO

FOP for AASHTO T 312 SuperPave Gyratory Compactor

T312 Sample Reduction Method	Date Reduced	Time Reduced	Sample Temperature
			302 °F

Final Reduction for T312 Performed By
WAQTC Number
Gyratory Compactor Brand
Model Number
Serial Number

	Specimen 1	Specimen 2	Design Mass
Mass of Sample	4462.9	4463.8	
Temp. of Sample When Placed in Mold	305 °F	305 °F	
Time Compaction Begins	4:13 PM	4:20 PM	
Sample Height (mm)	118.0	118.2	

Spec Limits
115±2

FOP for AASHTO T 166 Bulk Specific Gravity of Compacted Mix (Method A)

	Specimen 1	Specimen 2
Surface Temperature	75.8 °F	76.3 °F
Water Bath Temperature	76.9 °F	76.9 °F
Mass of Puck Dry (A)	4636.8	4636.1
Submerged Weight of Puck in Water (C)	2599.2	22601.5
Wt. of Puck SSD (B)	4647.5	4647.2
G_{mb} (Bulk Specific Gravity)	2.264	-0.258
Average G_{mb}	1.003	

Range **2.522** Accpetable? (Within d2s precision) NO

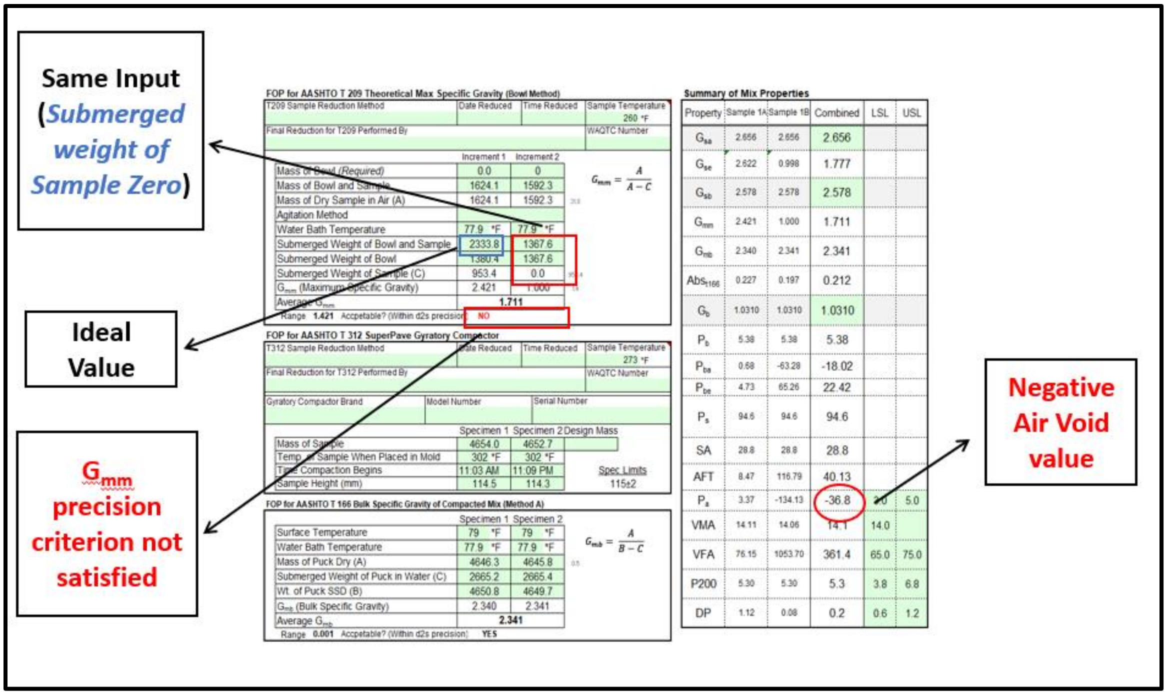
Summary of Mix Properties

Property	Sample 1A	Sample 1B	Combined	LSL	USL
G_{sa}	2.677	2.677	2.677		
G_{se}	2.621	0.944	1.741		
G_{sb}	2.562	2.562	2.562		
G_{mm}	2.382	0.950	1.666		
G_{mb}	2.264	-0.258	1.003		
Abs_{t166}	0.522	-0.062	0.230		
G_b	1.0320	1.0320	1.0320		
P_b	6.51	6.51	6.51		
P_{ba}	0.01	-68.99	-19.01		
P_{be}	5.66	71.01	24.28		
P_s	93.5	93.5	93.5		
SA	39.3	39.3	39.3		
AFT	7.49	94.01	32.15		
P_a	4.98	127.19	39.8	3.0	5.0
VMA	17.40	109.42	63.4	15.0	
VFA	71.37	-16.24	37.2	73.0	76.0
P200	5.49	5.49	5.5	5.5	8.5
DP	0.97	0.08	0.2	0.6	1.2

Figure 4.14 Calculated unreasonable air voids and VMA value from a large input value

- Case 3

Some audit trail values were exactly the same for multiple cells/parameters (Figure 4.15). This was probably due to the wrong input by a data entry person. A possible explanation can be that while the operator was trying to insert the values for a cell, they probably put the value in an adjacent cell. For example, the "submerged weight of bowl and sample" and the "submerged weight of bowl" both were set as 1367.6, which resulted in a value of 0 for the weight of sample.



FOP for AASHTO T 209 Theoretical Max Specific Gravity (Bowl Method)

T209 Sample Reduction Method	Date Reduced	Time Reduced	Sample Temperature 260 °F
Final Reduction for T209 Performed By			WAQTC Number

	Increment 1	Increment 2
Mass of Bowl (Required)	0.0	0
Mass of Bowl and Sample	1624.1	1592.3
Mass of Dry Sample in Air (A)	1624.1	1592.3
Agitation Method		
Water Bath Temperature	77.9 °F	77.9 °F
Submerged Weight of Bowl and Sample	2333.8	1367.6
Submerged Weight of Bowl	1380.4	1367.6
Submerged Weight of Sample (C)	953.4	0.0
G_{mm} (Maximum Specific Gravity)	2.421	1.000
Average G_{mm}	1.711	

$$G_{mm} = \frac{A}{A - C}$$

Range **1.421** Accpetable? (Within d2s precision) NO

FOP for AASHTO T 312 SuperPave Gyratory Compactor

T312 Sample Reduction Method	Date Reduced	Time Reduced	Sample Temperature 273 °F
Final Reduction for T312 Performed By			WAQTC Number
Gyratory Compactor Brand	Model Number	Serial Number	

	Specimen 1	Specimen 2	Design Mass
Mass of Sample	4654.0	4652.7	
Temp. of Sample When Placed in Mold	302 °F	302 °F	
Time Compaction Begins	11:03 AM	11:09 PM	Spec Limits
Sample Height (mm)	114.5	114.3	115±2

FOP for AASHTO T 166 Bulk Specific Gravity of Compacted Mix (Method A)

	Specimen 1	Specimen 2
Surface Temperature	79 °F	79 °F
Water Bath Temperature	77.9 °F	77.9 °F
Mass of Puck Dry (A)	4646.3	4645.8
Submerged Weight of Puck in Water (C)	2665.2	2665.4
Wt. of Puck SSD (B)	4650.8	4649.7
G_{mb} (Bulk Specific Gravity)	2.340	2.341
Average G_{mb}	2.341	

$$G_{mb} = \frac{A}{B - C}$$

Range **0.001** Accpetable? (Within d2s precision) YES

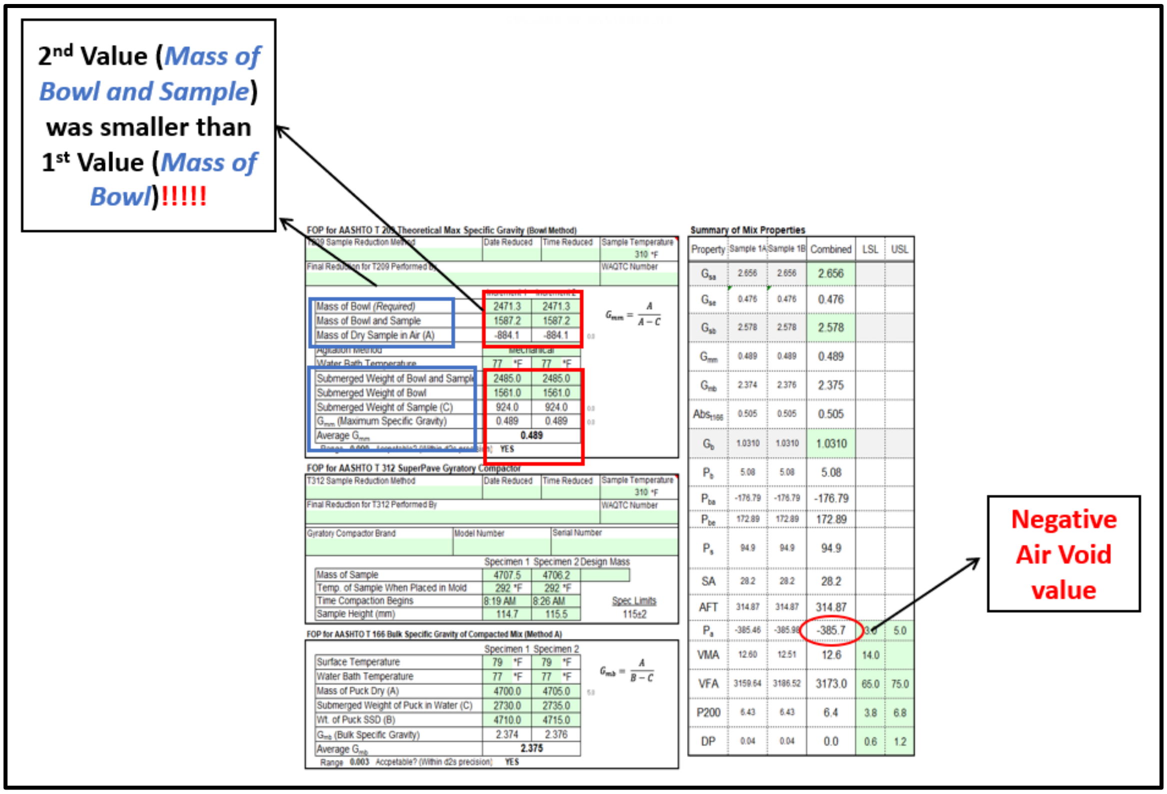
Summary of Mix Properties

Property	Sample 1A	Sample 1B	Combined	LSL	USL
G_{sa}	2.656	2.656	2.656		
G_{se}	2.622	0.998	1.777		
G_{sb}	2.578	2.578	2.578		
G_{mm}	2.421	1.000	1.711		
G_{mb}	2.340	2.341	2.341		
Abs_{t166}	0.227	0.197	0.212		
G_b	1.0310	1.0310	1.0310		
P_b	5.38	5.38	5.38		
P_{ba}	0.68	-63.28	-18.02		
P_{be}	4.73	65.26	22.42		
P_s	94.6	94.6	94.6		
SA	28.8	28.8	28.8		
AFT	8.47	116.79	40.13		
P_a	3.37	-134.13	-36.8	2.0	5.0
VMA	14.11	14.06	14.1	14.0	
VFA	76.15	1053.70	361.4	65.0	75.0
P200	5.30	5.30	5.3	3.8	6.8
DP	1.12	0.08	0.2	0.6	1.2

Figure 4.15 Calculated negative air voids value due to similar values inserted for adjacent cells

- Case 4

In some occasions, the later value (e.g., mass of bowl and sample) was smaller than the first value (e.g., mass of bowl), which is obviously not reasonable. Figure 4.16 shows such an example for which a test had a “mass of bowl” value higher than the “mass of bowl and sample”, which resulted in a large negative Air Voids value.



2nd Value (*Mass of Bowl and Sample*) was smaller than 1st Value (*Mass of Bowl*)!!!!!

Negative Air Void value

FOP for AASHTO T 209 Theoretical Max Specific Gravity (Bowl Method)

T209 Sample Reduction Method	Date Reduced	Time Reduced	Sample Temperature 310 °F
Final Reduction for T209 Performed By			WAQTC Number

	Increment 1	Increment 2
Mass of Bowl *(Required)*	2471.3	2471.3
Mass of Bowl and Sample	1587.2	1587.2
Mass of Dry Sample in Air (A)	-884.1	-884.1
Agitation Method	Mechanical	
Water Bath Temperature	77 °F	77 °F
Submerged Weight of Bowl and Sample	2485.0	2485.0
Submerged Weight of Bowl	1561.0	1561.0
Submerged Weight of Sample (C)	924.0	924.0
G_{mm} (Maximum Specific Gravity)	0.489	0.489
Average G_{mm}	0.489	

Range 0.000 Acceptable? (Within d2s precision) YES

$$G_{mm} = \frac{A}{A - C}$$

FOP for AASHTO T 312 SuperPave Gyratory Compactor

T312 Sample Reduction Method	Date Reduced	Time Reduced	Sample Temperature 310 °F
Final Reduction for T312 Performed By			WAQTC Number
Gyratory Compactor Brand	Model Number	Serial Number	

	Specimen 1	Specimen 2	Design Mass
Mass of Sample	4707.5	4706.2	
Temp. of Sample When Placed in Mold	292 °F	292 °F	
Time Compaction Begins	8:19 AM	8:26 AM	Spec Limits
Sample Height (mm)	114.7	115.5	115±2

FOP for AASHTO T 166 Bulk Specific Gravity of Compacted Mix (Method A)

	Specimen 1	Specimen 2
Surface Temperature	79 °F	79 °F
Water Bath Temperature	77 °F	77 °F
Mass of Puck Dry (A)	4700.0	4705.0
Submerged Weight of Puck in Water (C)	2730.0	2735.0
Wt. of Puck SSD (B)	4710.0	4715.0
G_{mb} (Bulk Specific Gravity)	2.374	2.376
Average G_{mb}	2.375	

Range 0.003 Accpetable? (Within d2s precision) YES

$$G_{mb} = \frac{A}{B - C}$$

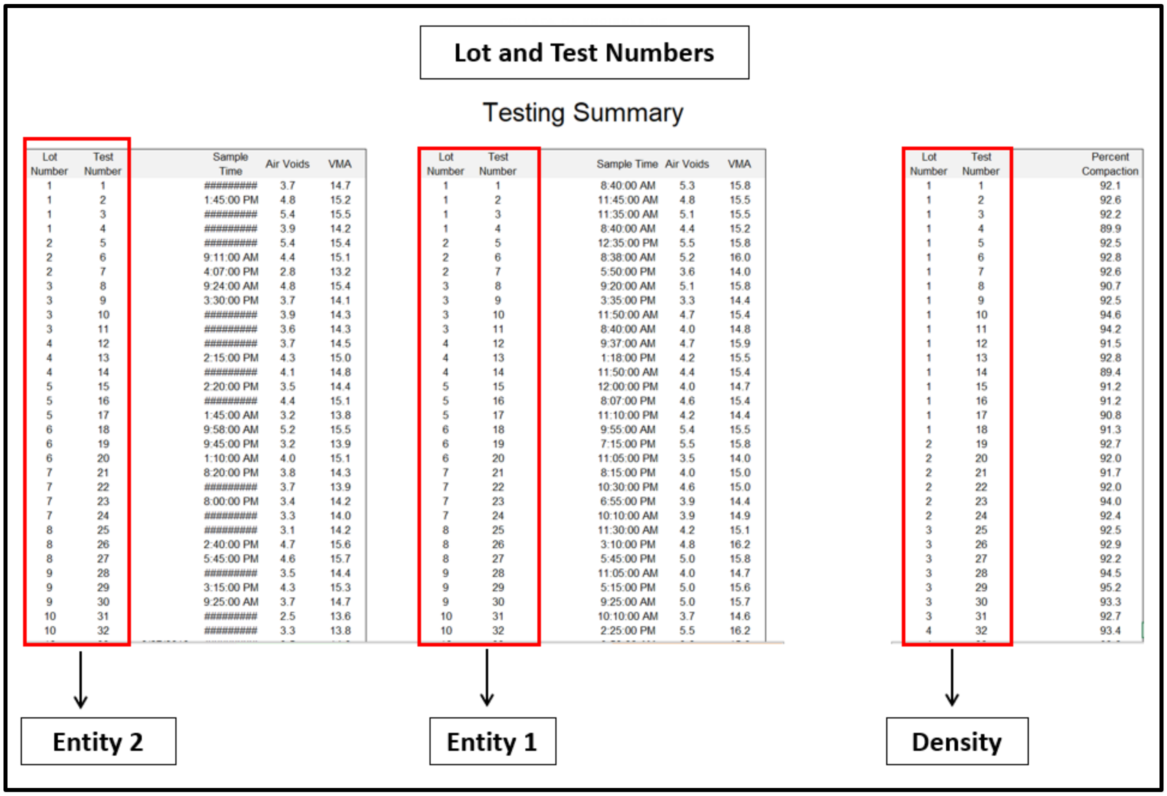
Summary of Mix Properties

Property	Sample 1A	Sample 1B	Combined	LSL	USL
G_{sa}	2.656	2.656	2.656		
G_{se}	0.476	0.476	0.476		
G_{sb}	2.578	2.578	2.578		
G_{mm}	0.489	0.489	0.489		
G_{mb}	2.374	2.376	2.375		
Abs_{T166}	0.505	0.505	0.505		
G_b	1.0310	1.0310	1.0310		
P_b	5.08	5.08	5.08		
P_{ba}	-176.79	-176.79	-176.79		
P_{be}	172.89	172.89	172.89		
P_s	94.9	94.9	94.9		
SA	28.2	28.2	28.2		
AFT	314.87	314.87	314.87		
P_a	-385.46	-385.98	-385.7	3.0	5.0
VMA	12.60	12.51	12.6	14.0	
VFA	3159.64	3186.52	3173.0	65.0	75.0
P200	6.43	6.43	6.4	3.8	6.8
DP	0.04	0.04	0.0	0.6	1.2

Figure 4.16 Calculated negative air voids value due to "mass of bowl and sample" being less than "mass of bowl"

Test and Lot Information

For the purpose of calculating the monetary value as well as removing unreasonable values, the Test and Lot information are required. From the "Testing Summary" sheet of ITD-0777 file (reported material testing data), all the Test and Lot information for each project were retrieved (Figure 4.17).



Lot Number	Test Number	Sample Time	Air Voids	VMA
1	1	##########	3.7	14.7
1	2	1:45:00 PM	4.8	15.2
1	3	##########	5.4	15.5
1	4	##########	3.9	14.2
2	5	##########	5.4	15.4
2	6	9:11:00 AM	4.4	15.1
2	7	4:07:00 PM	2.8	13.2
3	8	9:24:00 AM	4.8	15.4
3	9	3:30:00 PM	3.7	14.1
3	10	##########	3.9	14.3
3	11	##########	3.6	14.3
4	12	##########	3.7	14.5
4	13	2:15:00 PM	4.3	15.0
4	14	##########	4.1	14.8
5	15	2:20:00 PM	3.5	14.4
5	16	##########	4.4	15.1
5	17	1:45:00 AM	3.2	13.8
6	18	9:58:00 AM	5.2	15.5
6	19	9:45:00 PM	3.2	13.9
6	20	1:10:00 AM	4.0	15.1
7	21	8:20:00 PM	3.8	14.3
7	22	##########	3.7	13.9
7	23	8:00:00 PM	3.4	14.2
7	24	##########	3.3	14.0
8	25	##########	3.1	14.2
8	26	2:40:00 PM	4.7	15.6
8	27	5:45:00 PM	4.6	15.7
9	28	##########	3.5	14.4
9	29	3:15:00 PM	4.3	15.3
9	30	9:25:00 AM	3.7	14.7
10	31	##########	2.5	13.6
10	32	##########	3.3	13.8

Lot Number	Test Number	Sample Time	Air Voids	VMA
1	1	8:40:00 AM	5.3	15.8
1	2	11:45:00 AM	4.8	15.5
1	3	11:35:00 AM	5.1	15.5
1	4	8:40:00 AM	4.4	15.2
2	5	12:35:00 PM	5.5	15.8
2	6	8:38:00 AM	5.2	16.0
2	7	5:50:00 PM	3.6	14.0
3	8	9:20:00 AM	5.1	15.8
3	9	3:35:00 PM	3.3	14.4
3	10	11:50:00 AM	4.7	15.4
3	11	8:40:00 AM	4.0	14.8
4	12	9:37:00 AM	4.7	15.9
4	13	1:18:00 PM	4.2	15.5
4	14	11:50:00 AM	4.4	15.4
5	15	12:00:00 PM	4.0	14.7
5	16	8:07:00 PM	4.6	15.4
5	17	11:10:00 PM	4.2	14.4
6	18	9:55:00 AM	5.4	15.5
6	19	7:15:00 PM	5.5	15.8
6	20	11:05:00 PM	3.5	14.0
7	21	8:15:00 PM	4.0	15.0
7	22	10:30:00 PM	4.6	15.0
7	23	6:55:00 PM	3.9	14.4
7	24	10:10:00 AM	3.9	14.9
8	25	11:30:00 AM	4.2	15.1
8	26	3:10:00 PM	4.8	16.2
8	27	5:45:00 PM	5.0	15.8
9	28	11:05:00 AM	4.0	14.7
9	29	5:15:00 PM	5.0	15.6
9	30	9:25:00 AM	5.0	15.7
10	31	10:10:00 AM	3.7	14.6
10	32	2:25:00 PM	5.5	16.2

Lot Number	Test Number	Percent Compaction
1	1	92.1
1	2	92.6
1	3	92.2
1	4	89.9
1	5	92.5
1	6	92.8
1	7	92.6
1	8	90.7
1	9	92.5
1	10	94.6
1	11	94.2
1	12	91.5
1	13	92.8
1	14	89.4
1	15	91.2
1	16	91.2
1	17	90.8
1	18	91.3
2	19	92.7
2	20	92.0
2	21	91.7
2	22	92.0
2	23	94.0
2	24	92.4
3	25	92.5
3	26	92.9
3	27	92.2
3	28	94.5
3	29	95.2
3	30	93.3
3	31	92.7
4	32	93.4

Figure 4.17 Lot and test information for a test project

Parameter Values for Missing Tests

There were several tests with missing values in the audit trail file. For the sake of the monetary analysis, values from the audit trail file are required for all tests of a project. Hence, all the missing values were replaced with recorded values prior or after the missing value in the audit trail file, or if not available, with the final reported values. More often than not, final reported values (those that were formally used for payment calculation) were used to replace missing values. This is a conservative, and the only possible, approach to monetary impact assessment of unexplained data corrections in HMA projects.

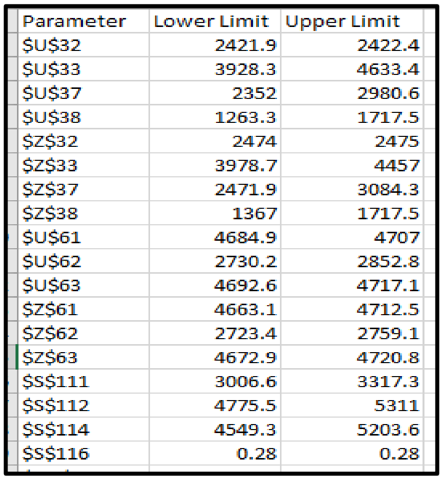
Removing Unreasonable Parameter Values

Occasionally, the first and last entry for U.C. and the last entry of P.C. from audit trail files had unreasonable values, potentially due to incomplete records of audit trail data. In order to remove them and only select reasonable values, we enforced multiple conditions through the following steps:

i. All the reported and audit trail values were taken for a parameter. For example, all values U32 (mass of bowl) for a project was considered as a list.

ii. Missing values from the list were removed at the first step.

3. Outlier data was removed from each list. For example, from the aforementioned case 1, a value of 22 was unreasonable for U32 (mass of bowl). This outlier value was removed using the Matlab's "rmoutlier" function. This removed any value that was outside three standard deviations range from the median.
4. A second step for removal of unreasonably high or low values was devised. It was noticed that "rmoutlier" did not remove all the unreasonable values, hence, if a value was greater than 1.2×mean or lower than 0.8×mean, it was removed. This threshold is set by expert opinion, and was manually checked for all tests in all projects to ensure its validity.
5. Some reasonable values, however, were removed through the process of step 4. In order to reintroduce the reasonable values to the list, the range of final reported values for each parameter across all tests was checked (Figure 4.18). If a removed parameter value fell within this range, it was reintroduced in the final list.

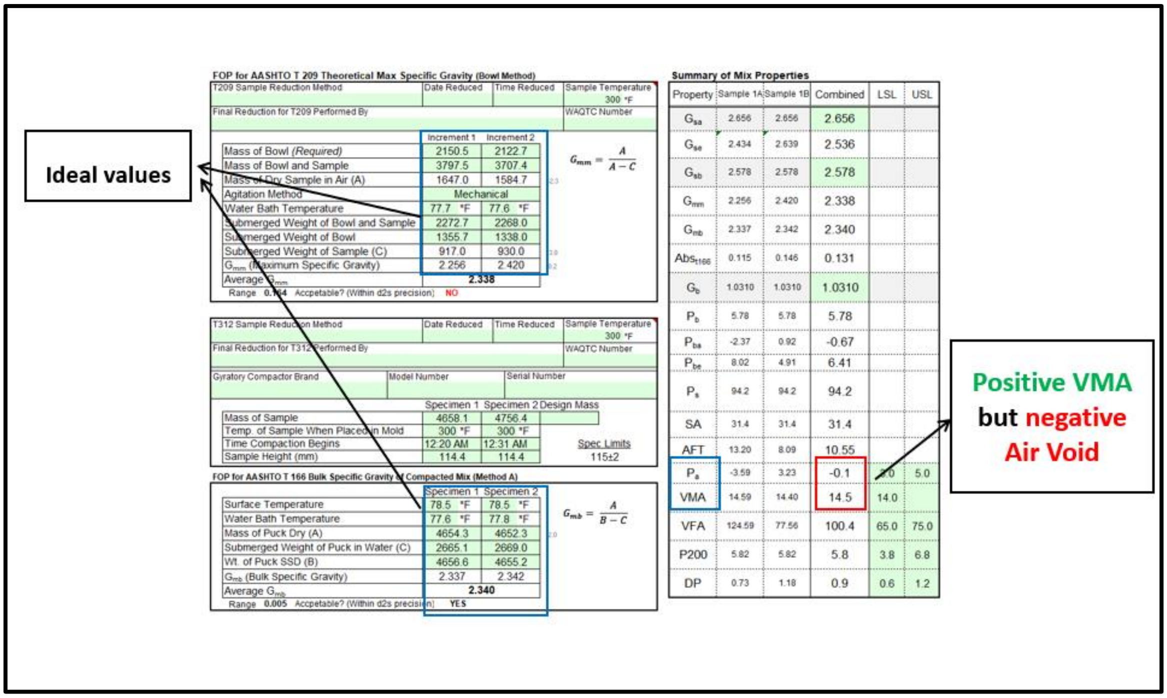


Parameter	Lower Limit	Upper Limit
U32	2421.9	2422.4
U33	3928.3	4633.4
U37	2352	2980.6
U38	1263.3	1717.5
Z32	2474	2475
Z33	3978.7	4457
Z37	2471.9	3084.3
Z38	1367	1717.5
U61	4684.9	4707
U62	2730.2	2852.8
U63	4692.6	4717.1
Z61	4663.1	4712.5
Z62	2723.4	2759.1
Z63	4672.9	4720.8
S111	3006.6	3317.3
S112	4775.5	5311
S114	4549.3	5203.6
S116	0.28	0.28

Figure 4.18 Lower and upper limit value for parameters

After completing all these steps, the desired dataset (two sets of parameter values, i.e., first and last U.C. values, for all tests) was finally ready for calculating the secondary parameters (Air voids/VMA/Mainline Density) that are used for monetary analysis.

Although the unreasonable values were removed, there is still the possibility of getting illogical (smaller/larger than expected or negative) secondary parameter values for first U.C. entry. This might be another reason why the data was corrected in the first place (to match with the ideal ranges for Air Voids—2-4, and VMA—12-16). In these cases, unreasonable values are either replaced by preceding/succeeding value in the audit trail file, if available, or the reported value for the considered parameter. Figures 4.19, 4.20, and 4.21 show cases in which even seemingly reasonable values of primary parameters resulted in secondary parameter values that do not fall in the acceptable range.



FOP for AASHTO T 209 Theoretical Max Specific Gravity (Bowl Method)

T209 Sample Reduction Method	Date Reduced	Time Reduced	Sample Temperature 300 °F
Final Reduction for T209 Performed By			WAQTC Number

$$G_{mm} = \frac{A}{A - C}$$

	Increment 1	Increment 2
Mass of Bowl *(Required)*	2150.5	2122.7
Mass of Bowl and Sample	3797.5	3707.4
Mass of Dry Sample in Air (A)	1647.0	1584.7
Agitation Method	Mechanical	
Water Bath Temperature	77.7 °F	77.6 °F
Submerged Weight of Bowl and Sample	2272.7	2268.0
Submerged Weight of Bowl	1355.7	1338.0
Submerged Weight of Sample (C)	917.0	930.0
G_{mm} (Maximum Specific Gravity)	2.256	2.420
Average G_{mm}	2.338	

Range 0.164 Accpetable? (Within d2s precision) NO

T312 Sample Reduction Method	Date Reduced	Time Reduced	Sample Temperature 300 °F
Final Reduction for T312 Performed By			WAQTC Number

Gyratory Compactor Brand	Model Number	Serial Number

	Specimen 1	Specimen 2	Design Mass
Mass of Sample	4658.1	4756.4	
Temp. of Sample When Placed in Mold	300 °F	300 °F	
Time Compaction Begins	12:20 AM	12:31 AM	Spec Limits
Sample Height (mm)	114.4	114.4	115±2

FOP for AASHTO T 166 Bulk Specific Gravity of Compacted Mix (Method A)

$$G_{mb} = \frac{A}{B - C}$$

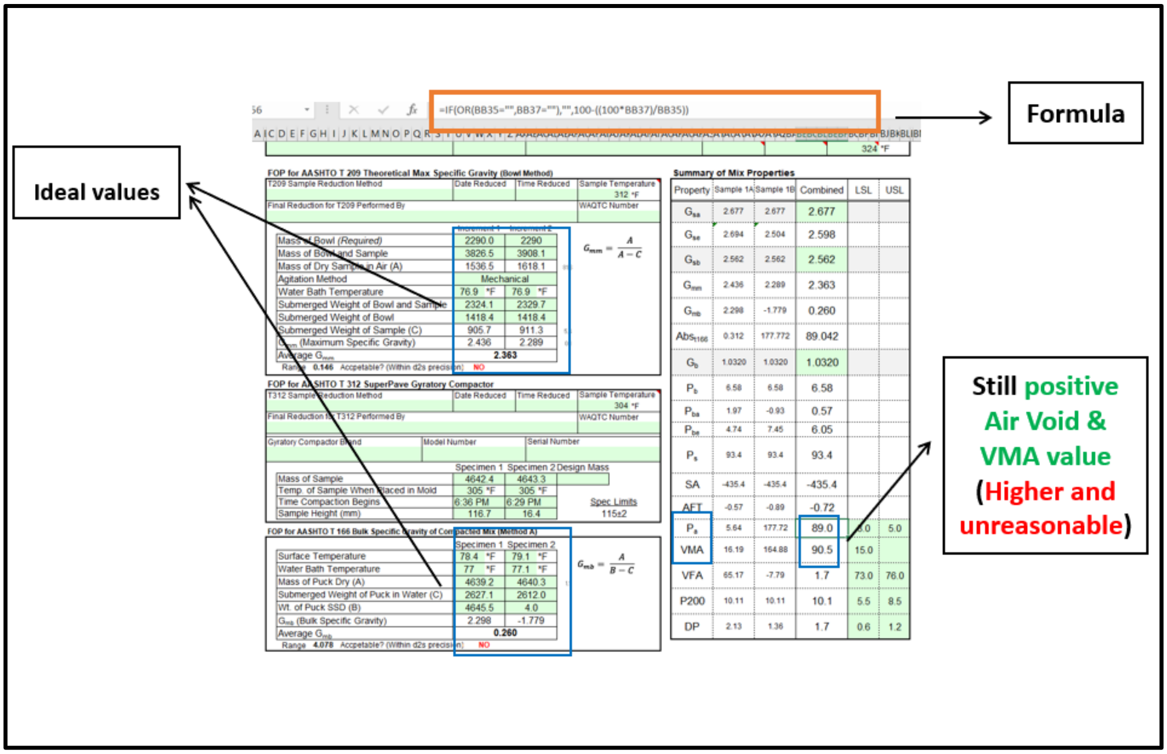
	Specimen 1	Specimen 2
Surface Temperature	78.5 °F	78.5 °F
Water Bath Temperature	77.6 °F	77.8 °F
Mass of Puck Dry (A)	4654.3	4652.3
Submerged Weight of Puck in Water (C)	2665.1	2669.0
Wt. of Puck SSD (B)	4656.6	4655.2
G_{mb} (Bulk Specific Gravity)	2.337	2.342
Average G_{mb}	2.340	

Range 0.005 Accpetable? (Within d2s precision) YES

Summary of Mix Properties

Property	Sample 1A	Sample 1B	Combined	LSL	USL
G_{sa}	2.656	2.656	2.656		
G_{se}	2.434	2.639	2.536		
G_{sb}	2.578	2.578	2.578		
G_{mm}	2.256	2.420	2.338		
G_{mb}	2.337	2.342	2.340		
Abs_{t166}	0.115	0.146	0.131		
G_b	1.0310	1.0310	1.0310		
P_b	5.78	5.78	5.78		
P_{ba}	-2.37	0.92	-0.67		
P_{be}	8.02	4.91	6.41		
P_s	94.2	94.2	94.2		
SA	31.4	31.4	31.4		
AFT	13.20	8.09	10.55		
P_a	-3.59	3.23	-0.1	2.0	5.0
VMA	14.59	14.40	14.5	14.0	
VFA	124.59	77.56	100.4	65.0	75.0
P200	5.82	5.82	5.8	3.8	6.8
DP	0.73	1.18	0.9	0.6	1.2

Figure 4.19 Unreasonable calculated air voids and VMA with seemingly reasonable primary parameter values



Formula

Ideal values

Still positive Air Void & VMA value (Higher and unreasonable)

```
=IF(OR(BB35="",BB37=""),"",100-((100*BB37)/BB35))
```

324 °F

FOP for AASHTO T 209 Theoretical Max Specific Gravity (Bowl Method)

T209 Sample Reduction Method	Date Reduced	Time Reduced	Sample Temperature
			312 °F

Final Reduction for T209 Performed By	WAQTC Number

	Increment 1	Increment 2
Mass of Bowl *(Required)*	2290.0	2290
Mass of Bowl and Sample	3826.5	3908.1
Mass of Dry Sample in Air (A)	1536.5	1618.1
Agitation Method	Mechanical	
Water Bath Temperature	76.9 °F	76.9 °F
Submerged Weight of Bowl and Sample	2324.1	2329.7
Submerged Weight of Bowl	1418.4	1418.4
Submerged Weight of Sample (C)	905.7	911.3
G_{mm} (Maximum Specific Gravity)	2.436	2.289
Average G_{mm}	2.363	

$$G_{mm} = \frac{A}{A - C}$$

Range **0.146** Accpetable? (Within d2s precision) NO

FOP for AASHTO T 312 SuperPave Gyratory Compactor

T312 Sample Reduction Method	Date Reduced	Time Reduced	Sample Temperature
			304 °F

Final Reduction for T312 Performed By	WAQTC Number

Gyratory Compactor Brand	Model Number	Serial Number

	Specimen 1	Specimen 2	Design Mass
Mass of Sample	4642.4	4643.3	
Temp. of Sample When Placed in Mold	305 °F	305 °F	
Time Compaction Begins	6:36 PM	6:29 PM	Spec Limits
Sample Height (mm)	116.7	16.4	115±2

FOP for AASHTO T 166 Bulk Specific Gravity of Compacted Mix (Method A)

	Specimen 1	Specimen 2
Surface Temperature	78.4 °F	79.1 °F
Water Bath Temperature	77 °F	77.1 °F
Mass of Puck Dry (A)	4639.2	4640.3
Submerged Weight of Puck in Water (C)	2627.1	2612.0
Wt. of Puck SSD (B)	4645.5	4.0
G_{mb} (Bulk Specific Gravity)	2.298	-1.779
Average G_{mb}	0.260	

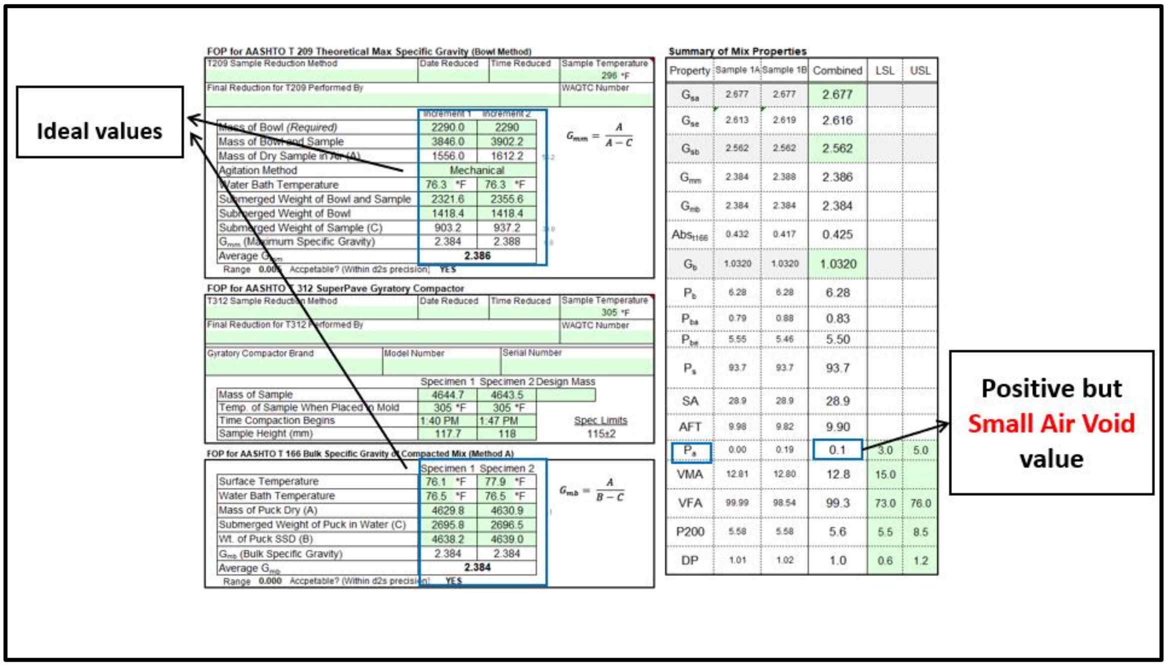
$$G_{mb} = \frac{A}{B - C}$$

Range **4.078** Accpetable? (Within d2s precision) NO

Summary of Mix Properties

Property	Sample 1A	Sample 1B	Combined	LSL	USL
G_{sa}	2.677	2.677	2.677		
G_{se}	2.694	2.504	2.598		
G_{sb}	2.562	2.562	2.562		
G_{mm}	2.436	2.289	2.363		
G_{mb}	2.298	-1.779	0.260		
Abs_{t166}	0.312	177.772	89.042		
G_b	1.0320	1.0320	1.0320		
P_b	6.58	6.58	6.58		
P_{ba}	1.97	-0.93	0.57		
P_{be}	4.74	7.45	6.05		
P_s	93.4	93.4	93.4		
SA	-435.4	-435.4	-435.4		
AFT	-0.57	-0.89	-0.72		
P_a	5.64	177.72	89.0	3.0	5.0
VMA	16.19	164.88	90.5	15.0	
VFA	65.17	-7.79	1.7	73.0	76.0
P200	10.11	10.11	10.1	5.5	8.5
DP	2.13	1.36	1.7	0.6	1.2

Figure 4.20 Unreasonable calculated air voids and VMA with seemingly reasonable primary parameter values



FOP for AASHTO T 209 Theoretical Max Specific Gravity (Bowl Method)

T209 Sample Reduction Method	Date Reduced	Time Reduced	Sample Temperature 296 °F
Final Reduction for T209 Performed By			WAQTC Number

	Increment 1	Increment 2
Mass of Bowl *(Required)*	2290.0	2290
Mass of Bowl and Sample	3846.0	3902.2
Mass of Dry Sample in Air (A)	1556.0	1612.2
Agitation Method	Mechanical	
Water Bath Temperature	76.3 °F	76.3 °F
Submerged Weight of Bowl and Sample	2321.6	2355.6
Submerged Weight of Bowl	1418.4	1418.4
Submerged Weight of Sample (C)	903.2	937.2
G_{mm} (Maximum Specific Gravity)	2.384	2.388
Average G_{mm}	2.386	

Range 0.00[illegible] Accpetable? (Within d2s precision) YES

$$G_{mm} = \frac{A}{A - C}$$

FOP for AASHTO T 312 SuperPave Gyratory Compactor

T312 Sample Reduction Method	Date Reduced	Time Reduced	Sample Temperature 305 °F
Final Reduction for T312 Performed By			WAQTC Number
Gyratory Compactor Brand	Model Number	Serial Number	

	Specimen 1	Specimen 2	Design Mass
Mass of Sample	4644.7	4643.5	
Temp. of Sample When Placed in Mold	305 °F	305 °F	
Time Compaction Begins	1:40 PM	1:47 PM	Spec Limits
Sample Height (mm)	117.7	118	115±2

FOP for AASHTO T 166 Bulk Specific Gravity of Compacted Mix (Method A)

	Specimen 1	Specimen 2
Surface Temperature	76.1 °F	77.9 °F
Water Bath Temperature	76.5 °F	76.5 °F
Mass of Puck Dry (A)	4629.8	4630.9
Submerged Weight of Puck in Water (C)	2695.8	2696.5
Wt. of Puck SSD (B)	4638.2	4639.0
G_{mb} (Bulk Specific Gravity)	2.384	2.384
Average G_{mb}	2.384	

Range 0.000 Accpetable? (Within d2s precision) YES

$$G_{mb} = \frac{A}{B - C}$$

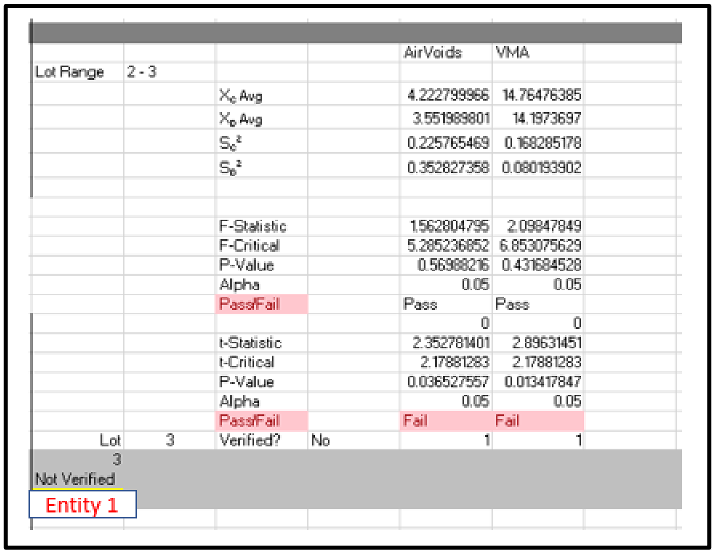
Summary of Mix Properties

Property	Sample 1A	Sample 1B	Combined	LSL	USL
G_{sa}	2.677	2.677	2.677		
G_{se}	2.613	2.619	2.616		
G_{sb}	2.562	2.562	2.562		
G_{mm}	2.384	2.388	2.386		
G_{mb}	2.384	2.384	2.384		
Abs_{t166}	0.432	0.417	0.425		
G_b	1.0320	1.0320	1.0320		
P_b	6.28	6.28	6.28		
P_{ba}	0.79	0.88	0.83		
P_{be}	5.55	5.46	5.50		
P_s	93.7	93.7	93.7		
SA	28.9	28.9	28.9		
AFT	9.98	9.82	9.90		
P_a	0.00	0.19	0.1	3.0	5.0
VMA	12.81	12.80	12.8	15.0	
VFA	99.99	98.54	99.3	73.0	76.0
P200	5.58	5.58	5.6	5.5	8.5
DP	1.01	1.02	1.0	0.6	1.2

Figure 4.21 Unreasonable calculated air voids and VMA with seemingly reasonable primary parameter values

Lot Grouping Changes

Based on the calculated Air Voids/VMA/Mainline Density parameter values for the first U.C. data, several lot groupings, that were originally used for monetary calculations, should have been changed, and many tests should have been rejected in the first place (Figure 4.22). However, it's not possible during these analysis steps to ask for a redo of the tests in the field and recalculate the secondary parameters, so the reported lot grouping was considered. This again is a conservative, and the only possible, approach for impact assessment of suspicious activities on HMA project payments.



				AirVoids	VMA
Lot Range	2 - 3				
		X_c Avg		4.222799966	14.76476385
		X_p Avg		3.551989801	14.1973697
		S_c^2		0.225765469	0.168285178
		S_p^2		0.352827358	0.080193902
		F-Statistic		1.562804795	2.09847849
		F-Critical		5.285236852	6.853075629
		P-Value		0.56988216	0.431684528
		Alpha		0.05	0.05
		Pass/Fail		Pass	Pass
				0	0
		t-Statistic		2.352781401	2.89631451
		t-Critical		2.17881283	2.17881283
		P-Value		0.036527557	0.013417847
		Alpha		0.05	0.05
		Pass/Fail		Fail	Fail
Lot	3	Verified?	No	1	1
3 Not Verified					

Figure 4.22 An example case of lot calculated parameters failing the statistical tests

Results of Monetary Analysis

The final payment-related parameter values were calculated for all tests of each project and for all projects. Detailed results and plots for project #1 are described in this section, and summary results for all projects are presented in a table format. Figure 4.23, for example, presents the number of unique cells that were changed in each lot for project #1 in data reported by Entity 1. The graph shows data for three separate categories of major/moderate/minor parameters for both P.C./U.C. instances. Lot 3, for example, has 5 instances of U.C. and 2 instances of P.C. for major parameters. Note that this graph presents the unique number of cells/parameters that were affected, not the number of times these cells were changed. The total number of times these cells were changed was much higher because each cell was changed multiple times.

The maximum number of U.C. instances for major parameters was observed in lot 15 in the data reported by Entity 2 for project 1 (Figure 4.24). It will be shown later that frequency of changes in U.C. parameters do not necessarily have a monotonic relationship with payment; and changes might be due to a variety of reasons including passing PWL or precision criteria. No direct relationship between the number of P.C. or U.C. changes in entity 1- reported versus the entity 2-reported data was observed either. Both datasets are prone to having multiple parameter value changes.

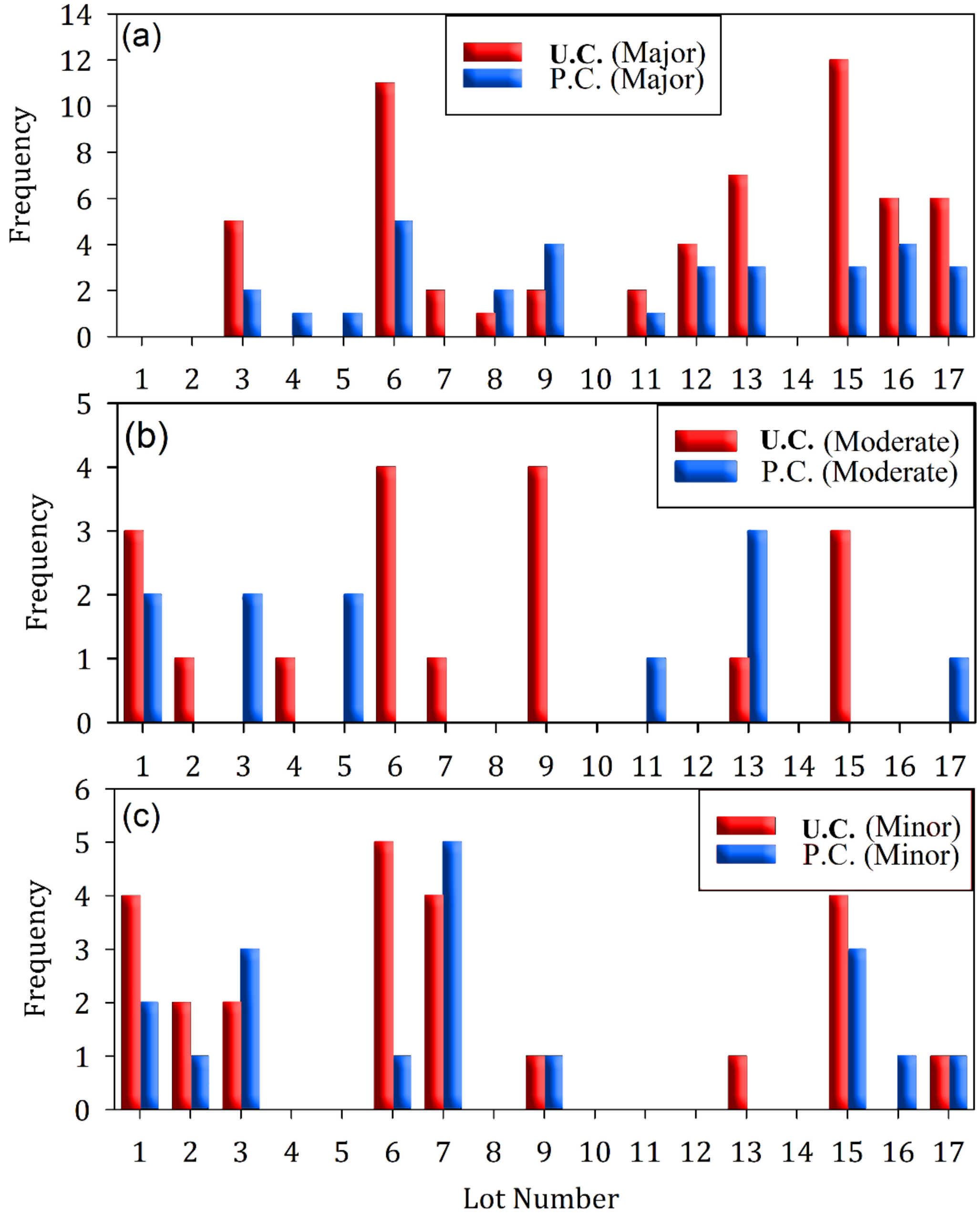


Figure 4.23 Number of unique P.C./U.C. parameter changes for each lot and each parameter type for the Entity 1-reported data for project #1

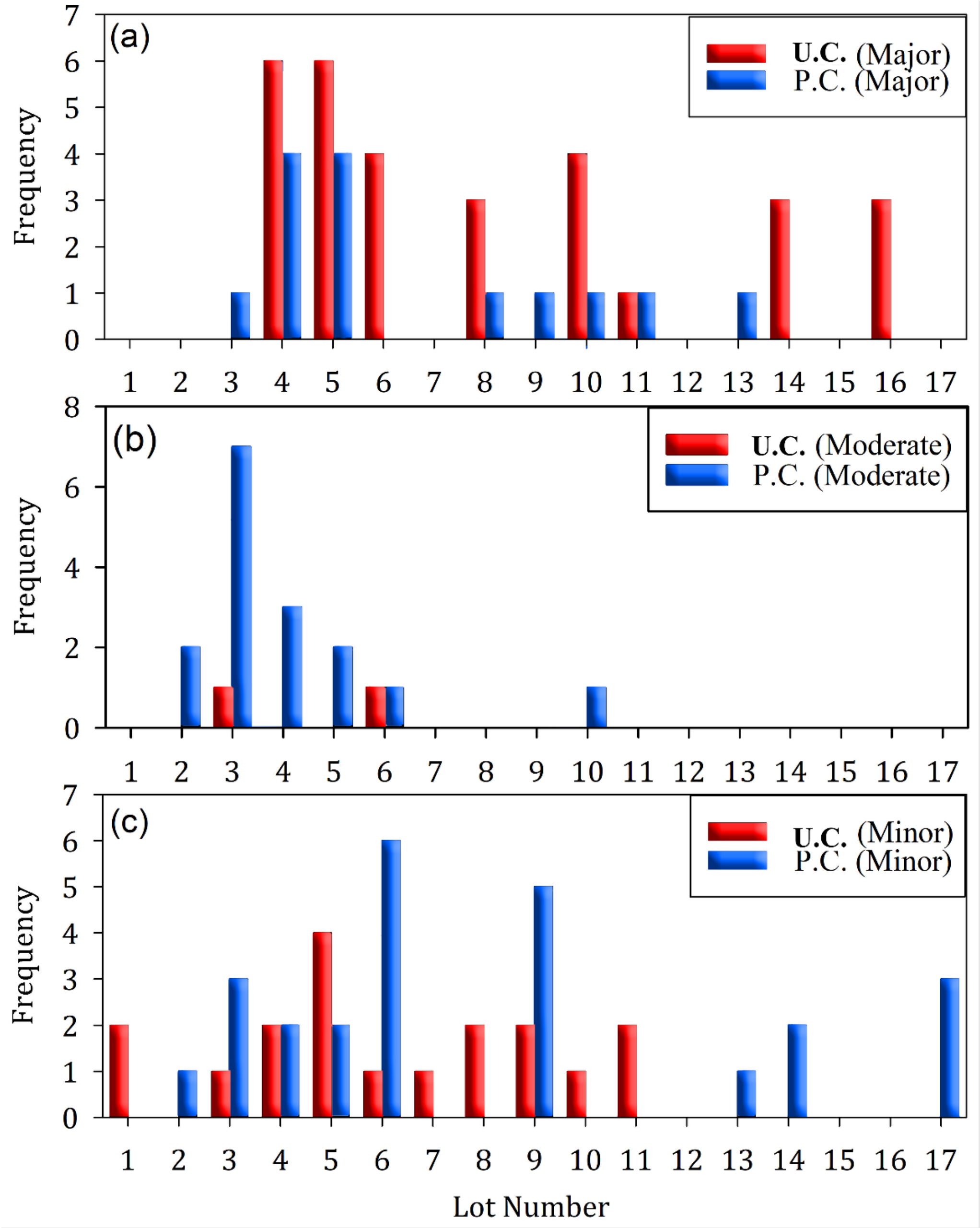
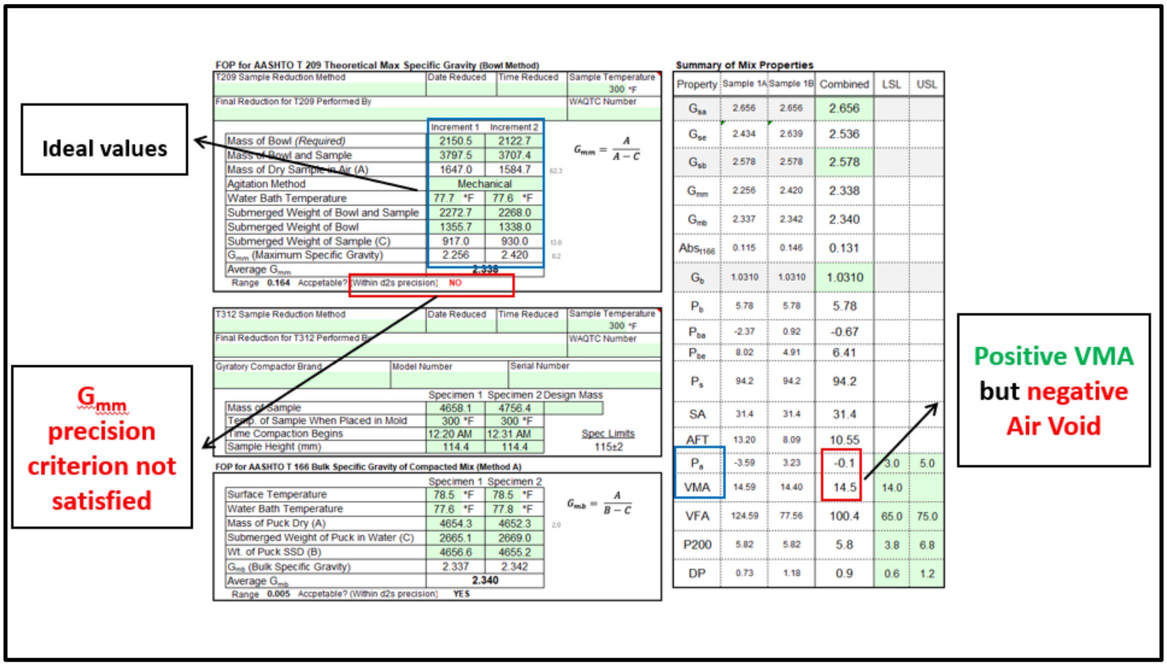


Figure 4.24 Number of unique P.C/U.C. parameters for Entity 2 tests for project #1

Before performing the monetary analysis, these primary parameters are checked for precision level in G_{mm}, G_{mb}, and P_b parameters. One of the precision checks is shown in Figure 4.25, where G_{mm} precision did not pass (result was "No") for this example test. For project #1, precision results for each test for entity 1 and entity 2 reported data are presented in Figure 4.26 (green: pas – red: fail). We observed that multiple tests did not pass the precision test.



FOP for AASHTO T 209 Theoretical Max Specific Gravity (Bowl Method)

T209 Sample Reduction Method	Date Reduced	Time Reduced	Sample Temperature
			300 °F
Final Reduction for T209 Performed By			WAQTC Number

	Increment 1	Increment 2
Mass of Bowl *(Required)*	2150.5	2122.7
Mass of Bowl and Sample	3797.5	3707.4
Mass of Dry Sample in Air (A)	1647.0	1584.7
Agitation Method	Mechanical	
Water Bath Temperature	77.7 °F	77.6 °F
Submerged Weight of Bowl and Sample	2272.7	2268.0
Submerged Weight of Bowl	1355.7	1338.0
Submerged Weight of Sample (C)	917.0	930.0
G_{mm} (Maximum Specific Gravity)	2.256	2.420
Average G_{mm}	2.338	

Range **0.164** Accpetable? (Within d2s precision) NO

T312 Sample Reduction Method	Date Reduced	Time Reduced	Sample Temperature
			300 °F
Final Reduction for T312 Performed By			WAQTC Number
Gyratory Compactor Brand	Model Number	Serial Number	

	Specimen 1	Specimen 2	Design Mass
Mass of Sample	4658.1	4756.4	
Temp. of Sample When Placed in Mold	300 °F	300 °F	
Time Compaction Begins	12:20 AM	12:31 AM	Spec Limits
Sample Height (mm)	114.4	114.4	115±2

FOP for AASHTO T 166 Bulk Specific Gravity of Compacted Mix (Method A)

	Specimen 1	Specimen 2
Surface Temperature	78.5 °F	78.5 °F
Water Bath Temperature	77.6 °F	77.8 °F
Mass of Puck Dry (A)	4654.3	4652.3
Submerged Weight of Puck in Water (C)	2665.1	2669.0
Wt. of Puck SSD (B)	4656.6	4655.2
G_{mb} (Bulk Specific Gravity)	2.337	2.342
Average G_{mb}	2.340	

Range **0.005** Accpetable? (Within d2s precision) YES

Summary of Mix Properties

Property	Sample 1A	Sample 1B	Combined	LSL	USL
G_{sa}	2.656	2.656	2.656		
G_{se}	2.434	2.639	2.536		
G_{sb}	2.578	2.578	2.578		
G_{mm}	2.256	2.420	2.338		
G_{mb}	2.337	2.342	2.340		
Abs_{t166}	0.115	0.146	0.131		
G_b	1.0310	1.0310	1.0310		
P_b	5.78	5.78	5.78		
P_{ba}	-2.37	0.92	-0.67		
P_{be}	8.02	4.91	6.41		
P_s	94.2	94.2	94.2		
SA	31.4	31.4	31.4		
AFT	13.20	8.09	10.55		
P_a	-3.59	3.23	-0.1	3.0	5.0
VMA	14.59	14.40	14.5	14.0	
VFA	124.59	77.56	100.4	65.0	75.0
P200	5.82	5.82	5.8	3.8	6.8
DP	0.73	1.18	0.9	0.6	1.2

Figure 4.25 Precision criterion not satisfied for an example project

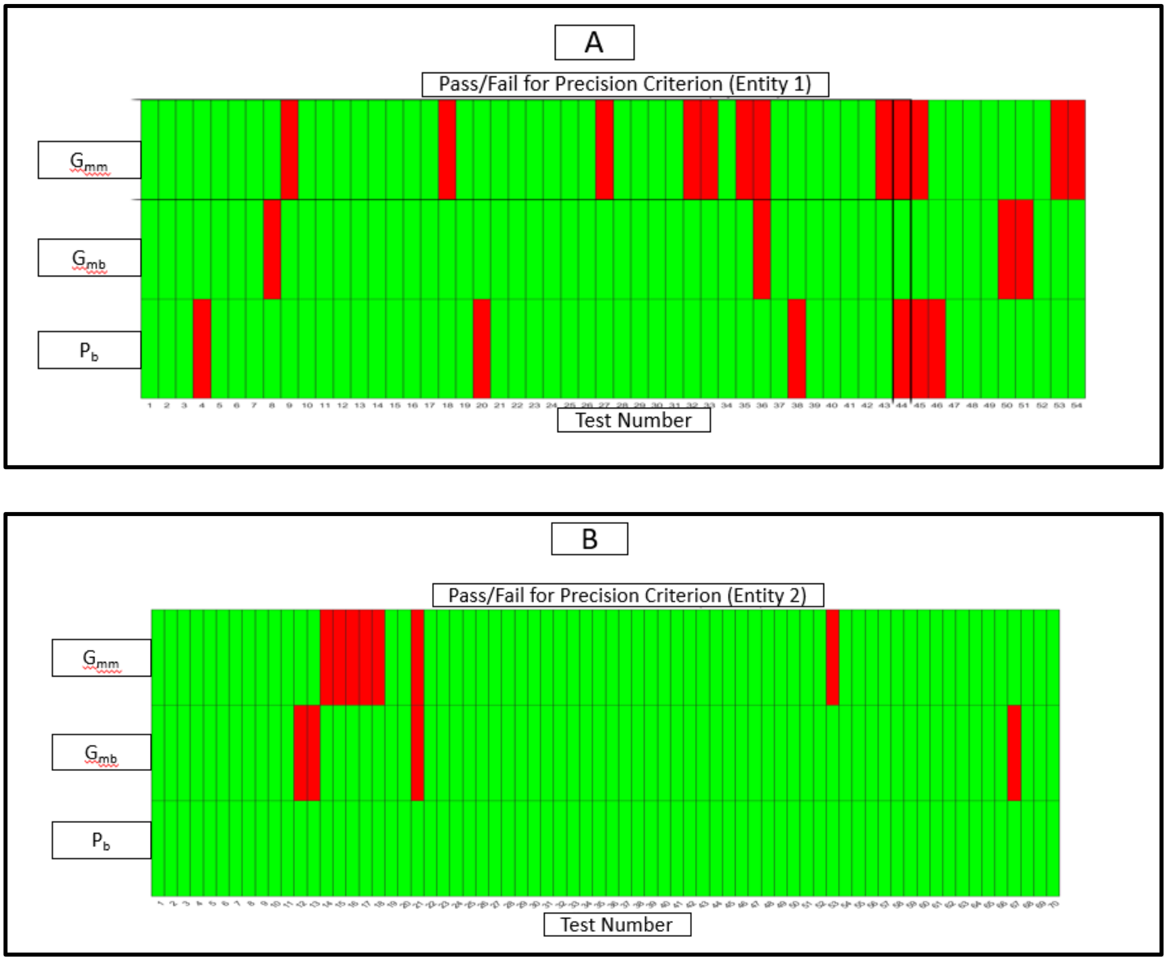


Figure 4.26 Precision criterion for each test of project #1 (upper: entity 1, lower: Entity 2). Green shows pass and red represents fail.

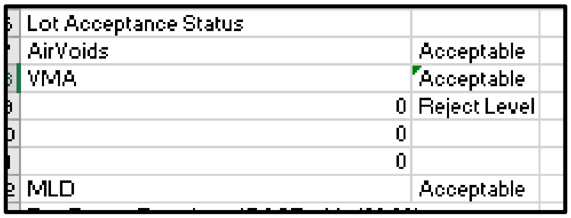
Acceptance Check

Monetary analysis starts with two statistical tests (F and t tests) to determine whether prime contractor-reported data should be used, or the ITD-reported data is to be used. Then the selected data goes through the "quality level analysis" for Air Voids/VMA/Mainline Density which subsequently determines whether or not the lot is at an acceptable level. Figure 4.27 shows an example graph with Accept (green)/Reject (red)/Stop Production (black) levels for PWL for Air Voids, VMA and Mainline Density for project #1. These checks were done for the first U.C. entry cases to see that if the first value was considered for payment, how many lots should have been rejected. This analysis indicates that even before considering payment, several lots might have been rejected. Usually, for the three PF related parameters, this acceptability check is done with the following generic value check.

$$PWL_{Air\ Void/\ VMA/\ Density} > 60 = \ Acceptable$$

$$PWL_{Air\ Void/\ VMA/\ Density} > 40 = \ Stop\ Production, Action\ Needed$$

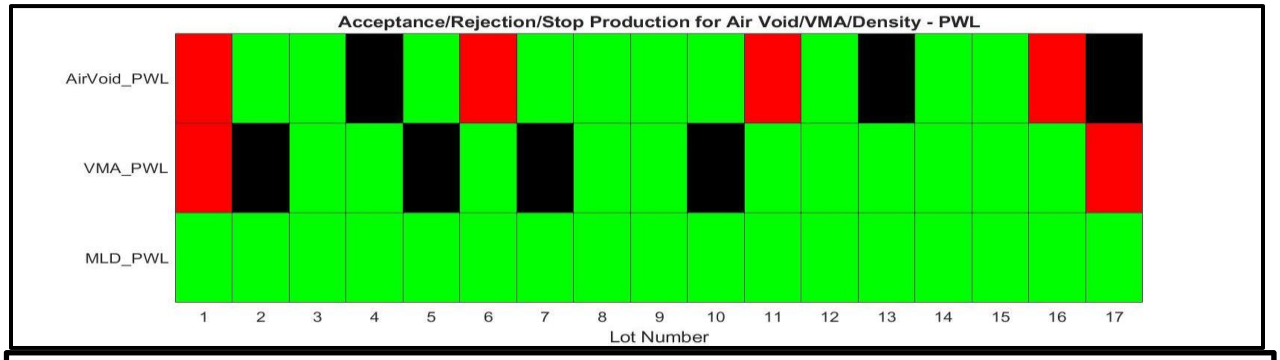
$$PWL_{Air\ Void/\ VMA/\ Density} < 40 = \ Reject\ Level$$



Lot Acceptance Status	
AirVoids	Acceptable
VMA	Acceptable
0	Reject Level
0	
0	
MLD	Acceptable

Figure 4.27 Acceptance check for payment related parameters

Figure 4.28 shows that multiple lots should have been rejected based on the PWL check for the first U.C. parameter value in project #1. The first row presents results for Air Voids, the second row is for VMA, and the last row is for Mainline Density. Five lots out of the total 17 got rejected in the parameter's quality level analysis check, and only 6 lots out of the 17 were at an acceptable level.



-- Acceptable Level, -- Reject Level, -- Stop Production, Action Needed

Figure 4.28 Lot-wise acceptance/rejection/stop production according to PWL for project #1

Now the focus changes to the monetary analysis of data corrections, based on the first and last acceptable entry for U.C. cases. In Figure 4.29, the Green bars show calculated monetary value for the first acceptable U.C. parameter values. As discussed earlier, for the unchanged parameter values (no correction) and for P.C. cases, the reported value and last P.C. value were selected for monetary analysis, respectively. The red bar shows calculated payment based on the last entry for U.C.

parameters. Yellow bars present the original reported payment. These payment levels are calculated for each lot separately.

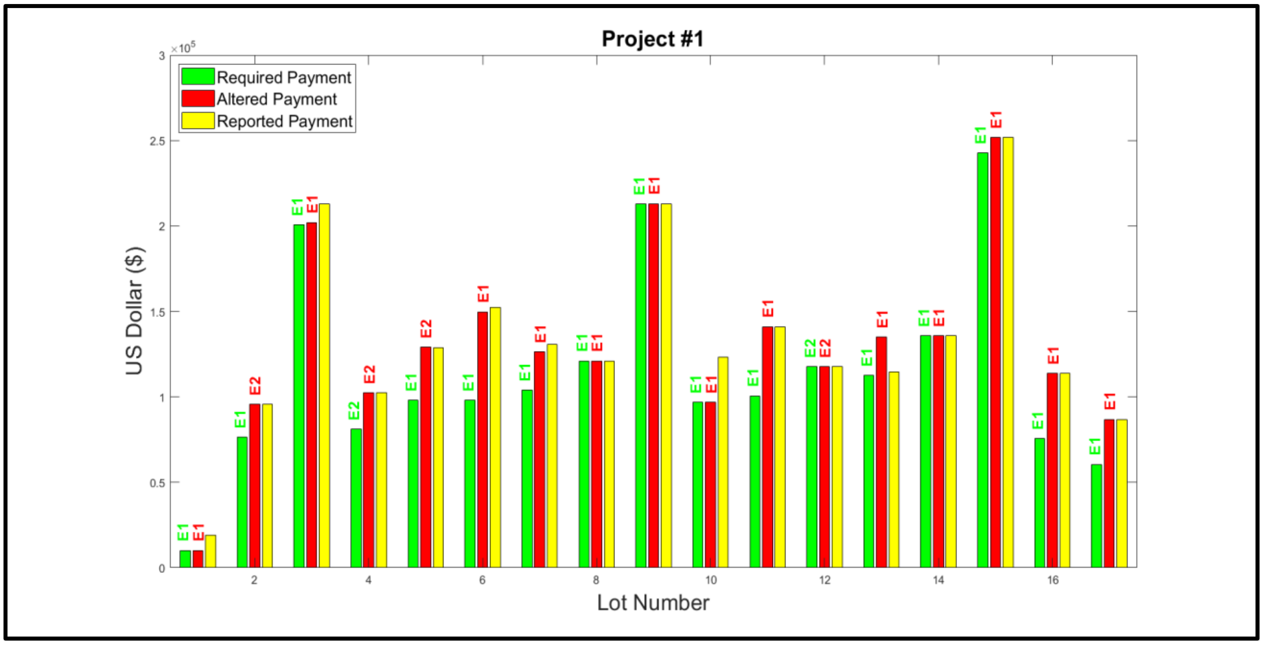


Figure 4.29 Lot-wise payment for project #1. Green bars show payment based on the first U.C. parameter values, red bars present payment based on last U.C. parameter values, and yellow bars show the actual payment formally made. E1 and E2 represent Entity 1 and Entity 2, respectively.

There were some lots for which the final calculated value did not match the reported formal value from the projects. There are two reasons for this observation:

i. Some of the lots had “dispute resolution” status, which was resolved by collecting data by a third party. However, no audit trail data from the third party was available. So, the herein calculated value was marginally different from the originally reported payments.

ii. As discussed earlier, audit trail files did not necessarily record all data entry, meaning that final reported parameter value might not be included in the audit trail file. Some instances were observed that the last value recorded in the audit trail file was not equal to the reported one. Because of the irregularity of the data in the audit trail file for some lots, the calculations did not match the exact reported value in a few cases.

Bars in Figure 4.29 are labeled as E1 and E2, which represent Entity 1 and Entity 2, respectively. This shows which reported data was chosen for payment analysis based on the F and t tests. For lot 2, for example, if the initially reported values were considered, Entity 1-reported data should have been used for payment, whereas due to correction, entity 2 data were used for payment. This resulted in an overpayment of around 20,000 dollars (+20%) for this lot. It is evident in Figure 4.29 that for several lots, payments should have been less if the initial U.C. entry value for parameters was chosen.

There were originally 15 projects obtained from ITD that had some sort of audit trail file included. Out of the 15 projects, however, 3 either were missing audit trail files or reported values were unavailable. Hence the focus is on the 12 projects for which payments could be calculated.

Table 4.1 shows cumulative monetary value for each project based on the first and last U.C. parameter values and also the final/formal reported payment. This table includes all the available number of audit tests from entity 1 and entity 2 as well as the cumulative monetary values for the projects. In most projects, there was a notable amount of overpayment.

Table 4.1 Calculated Payments for First and Last U.C. Parameter Values, and the Formally Paid Amount for Each Project. Table Also Enlists Statistics of Total Number of Lots and Available Audit Trail Files from the Entity 1 and Entity 2 Reports

Project Number	Total Lot	Total Test (E2)	AA* (E2)	Total Test (E1)	AA (E1)	Total Sheet (Dens)	AA (Dens)	Payment (First U.C.)	Payment (Last U.C.)	Formal Payment (Reported with Dispute Resolution)
Project 1	17	70	70	54	54	21	21	$1,945,217	$2,228,807	$2,260,795
Project 2	14	67	67	67	67	15	14	$2,492,391	$2,853,563	$3,215,331
Project 3	5	16	NA	16	16	12	5	$568,890	$583,246	$579,831
Project 4	25	101	51	101	52	27	13	$3,962,182	$4,082,441	$4,217,759
Project 5	50	241	84	150	12	57	5	$9,860,811	$9,906,251	$9,897,883
Project 6	21	79	NA	74	50	50	33	$1,976,327	$2,030,917	$2,040,929
Project 7	5	16	16	16	16	10	10	$762,583	$989,563	$989,797
Project 8	4	14	14	13	11	8	6	$586,866	$709,034	$709,243
Project 9	3	9	9	9	9	3	3	$195,573	$192,578	$212,967
Project 10	11	51	NA	33	4	13	3	$1,756,489	$1,878,476	$1,952,210
Project 11	13	25	25	42	42	13	13	$1,117,583	$1,142,740	$1,525,770
Project 12	17	54	14	51	11	19	6	$1,907,322	$1,906,912	$2,306,717

AA—Available Audit

Table 4.2 summarizes all PWL results for all projects. This table provides details about the number of lots in each project, number of lots for which audit trail files were available, and number of lots for which audit trail files were available for both entity 1 and entity 2. This table also enlists the number of lots that should have been rejected (based on at least one parameter, i.e. Air Voids, VMA, or Mainline Density), accepted or was at the stop production level. Projects #8 and #9 did not even have a single lot that could have been accepted (Table 4.2), whereas project #5 had the highest fraction of accepted lots (90% of all lots). On average, 8-50% of the lots should have been stopped and reformed the lot/redid the test, which indicates a considerable proportion of lots could have been rejected if data correction had not happened.

Table 4.2 Summary of Acceptance/Rejection and Stop Production for PWL Analysis for Each Project

Project	Total Lot	No of Lot (E1)	Missed Lot (E1)	No of Lot (E2)	Missed Lot (E2)	TA	TR	Total Stop Production , Action Needed	No of Lots *	Accepted *	Rejected *	Stop Production *
Project 1	17	15	10, 14	15	12, 15	6 (35%)	5 (29%)	7 (41%)	13	3 (23%)	5 (38%)	6 (46%)
Project 2	14	14	NA	14	NA	6 (43%)	7 (50%)	2 (14%)	14	6 (43%)	7 (50%)	2 (14%)
Project 3	5	4	1	No data	No data	4 (80%)	0 (0%)	1 (20%)	0	0 (0%)	0 (0%)	0 (0%)
Project 4	25	12	2 to 13, 21	14 to 25	1 to 13	20 (80%)	2(8%)	3 (12%)	11	9 (82%)	1 (9%)	1 (9%)
Project 5	50	5	2 to 46	34 to 50	1 to 33	45 (90%)	1 (2%)	4 (8%)	4	3 (75%)	1 (25%)	0 (0%)
Project 6	21	12	1 to 7, 11, 19	No data	No data	16 (76%)	2 (10%)	3 (14%)	0	0 (0%)	0 (0%)	0 (0%)
Project 7	5	5	0	5	0	1 (20%)	2 (40%)	2 (40%)	5	1 (20%)	2 (40%)	2 (40%)
Project 8	4	4	0	4	0	0 (0%)	3 (75%)	2 (50%)	4	0 (0%)	3 (75%)	2 (50%)
Project 9	3	3	0	2	1	0 (0%)	3 (100%)	1 (33%)	2	0 (0%)	2 (100%)	1 (50%)
Project 10	11	1	1 to 10	No data	No data	2 (18%)	2 (18%)	1 (9%)	0	0 (0%)	0 (0%)	0 (0%)
Project 11	13	10	4, 5, 9	8	9 to 13	3 (23%)	9 (69%)	3 (23%)	6	1 (17%)	4 (67%)	2 (33%)
Project 12	17	4	5 to 17	5	6 to 17	7 (41%)	6 (35%)	6 (35%)	4	0 (0%)	4 (100%)	2 (50%)

*—Mutual between E1 & E2
TA—Total Accepted
TR—Total Rejected

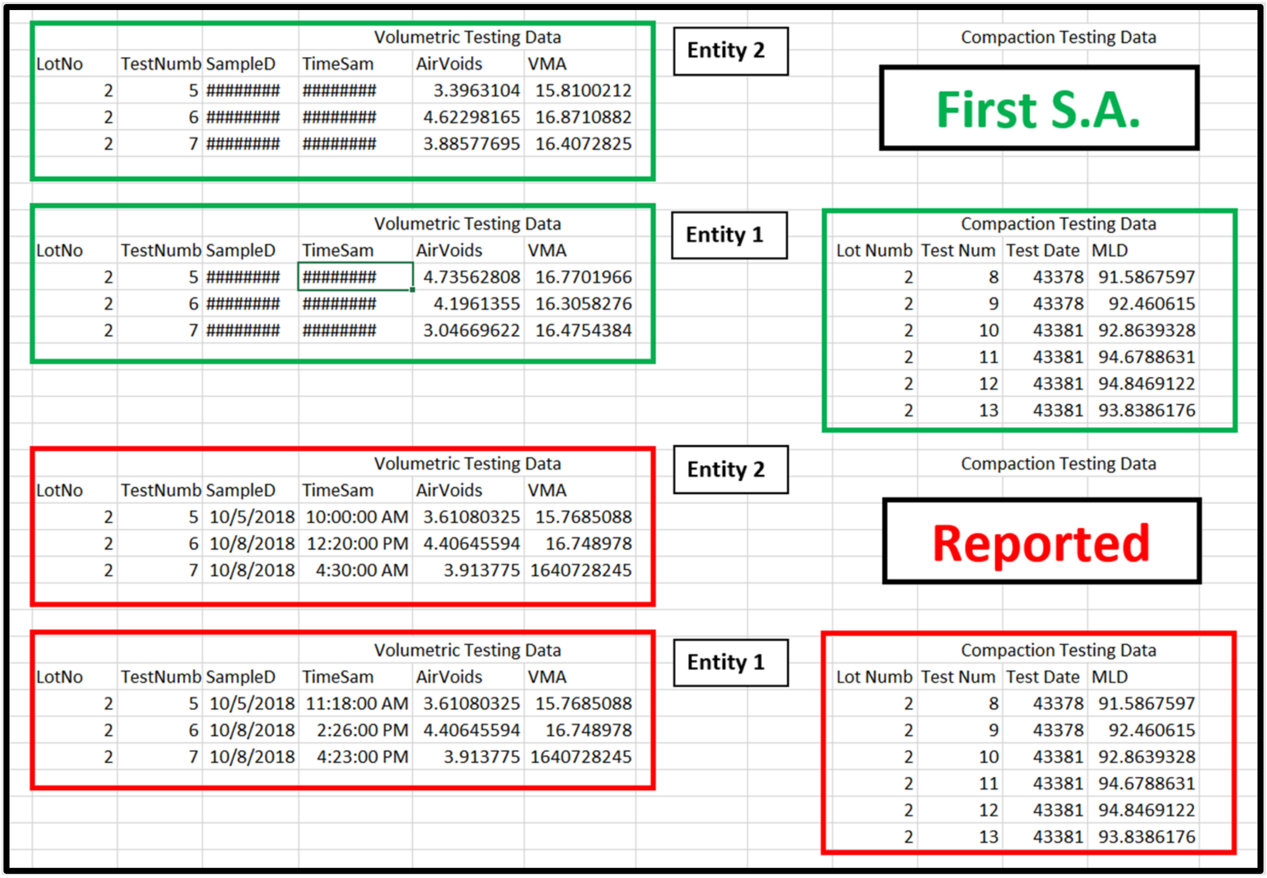
Table 4.3 summarizes the calculated overpayment for each project, as well as the average extra payment per unique parameter changed. In this table, the total major U.C. unique parameters and total U.C. unique parameters represent either entity 1 or entity 2 based on which of them were selected for monetary analysis. For example, in lot 1 of a project, material testing report either from entity 1 or 2 is selected for payment based on the statistical test results. If entity 1 is selected, then the number of major and total number of unique U.C. parameters are considered. Therefore, the final value of the number of unique major U.C. and total U.C. parameters that are presented in this table are summations of all the lots of a project from either entity 1 or entity 2 based on which of them were selected on each individual lot. The maximum amount of extra payment was seen on project #2, where more than $361,000 was overpaid. In this project, 94 major and a total of 138 parameters were corrected. The high number of corrections resulted in a notable monetary change in this project. A majority of the analyzed projects had a significant amount of overpayment. For some projects (9 and 12) a reduction in payment was seen, although the sheer value of reduction is marginal. It is noteworthy that there were also some lots in different projects for which detected U.C. values resulted in minor decrease in payment, but for the entire project, the summation of all lots resulted in overpayment. It is also interesting to observe in this table that each U.C. parameter change resulted in roughly $1,000-$5,000 extra payment in each project. The audit trail files did not necessarily capture all changes in reported parameter values, and it is expected that if those are factored in, the change in payment can be even higher.

Table 4.3 Summary of Payment Change, and Number of Unique U.C. Parameters Involved for Each Project

Project Number	Total Lot	Total payment change ($) (first and last U.C.)	Total major U.C. unique parameters	Total payment changes per unique major U.C. parameter ($/parameter)	Total U.C. unique parameters	Total payment changes per unique U.C. parameter ($/parameter)
Project 1	17	$283,590	60	4,727	103	2,753
Project 2	14	$361,172	94	3,842	138	2,617
Project 3	5	$14,356	0	NA	0	NA
Project 4	25	$120,258	38	3,165	64	1,879
Project 5	50	$45,440	33	1,377	45	1,010
Project 6	21	$54,590	20	2,729	22	2,481
Project 7	5	$226,980	47	4,829	66	3,439
Project 8	4	$122,168	36	3,394	45	2,715
Project 9	3	$-2,995	1	-2,995	5	-599
Project 10	11	$121,987	7	17,427	9	13,554
Project 11	13	$25,158	14	1,797	23	1,094
Project 12	17	$-409	7	-58	10	-41

Relationship between U.C. Instances and Payment

An essential question is whether or not data correction always translates to financial impacts. A simple answer is “No”. It was observed that data corrections did not necessarily translate into monetary changes all the time. Through in-depth analysis, we investigated the potential reasons for this observation. An overall comparison of the monetary-related parameter (Air Voids/VMA/Mainline Density) values from the primary parameters for first U.C. parameter entry and final reported parameter is shown in Figure 4.30. The upper part (green) and lower part (red) of Figure 4.30 present all the test values for first U.C. entry and final reported entry for a particular lot in project #7. Looking closely, most of the test values are different between the two cases (green versus red).



Volumetric Testing Data

LotNo	TestNumb	SampleD	TimeSam	AirVoids	VMA
2	5	########	########	3.3963104	15.8100212
2	6	########	########	4.62298165	16.8710882
2	7	########	########	3.88577695	16.4072825

Entity 2

Compaction Testing Data

First S.A.

Volumetric Testing Data

LotNo	TestNumb	SampleD	TimeSam	AirVoids	VMA
2	5	########	########	4.73562808	16.7701966
2	6	########	########	4.1961355	16.3058276
2	7	########	########	3.04669622	16.4754384

Entity 1

Compaction Testing Data

Lot Numb	Test Num	Test Date	MLD
2	8	43378	91.5867597
2	9	43378	92.460615
2	10	43381	92.8639328
2	11	43381	94.6788631
2	12	43381	94.8469122
2	13	43381	93.8386176

Volumetric Testing Data

LotNo	TestNumb	SampleD	TimeSam	AirVoids	VMA
2	5	10/5/2018	10:00:00 AM	3.61080325	15.7685088
2	6	10/8/2018	12:20:00 PM	4.40645594	16.748978
2	7	10/8/2018	4:30:00 AM	3.913775	1640728245

Entity 2

Compaction Testing Data

Reported

Volumetric Testing Data

LotNo	TestNumb	SampleD	TimeSam	AirVoids	VMA
2	5	10/5/2018	11:18:00 AM	3.61080325	15.7685088
2	6	10/8/2018	2:26:00 PM	4.40645594	16.748978
2	7	10/8/2018	4:23:00 PM	3.913775	1640728245

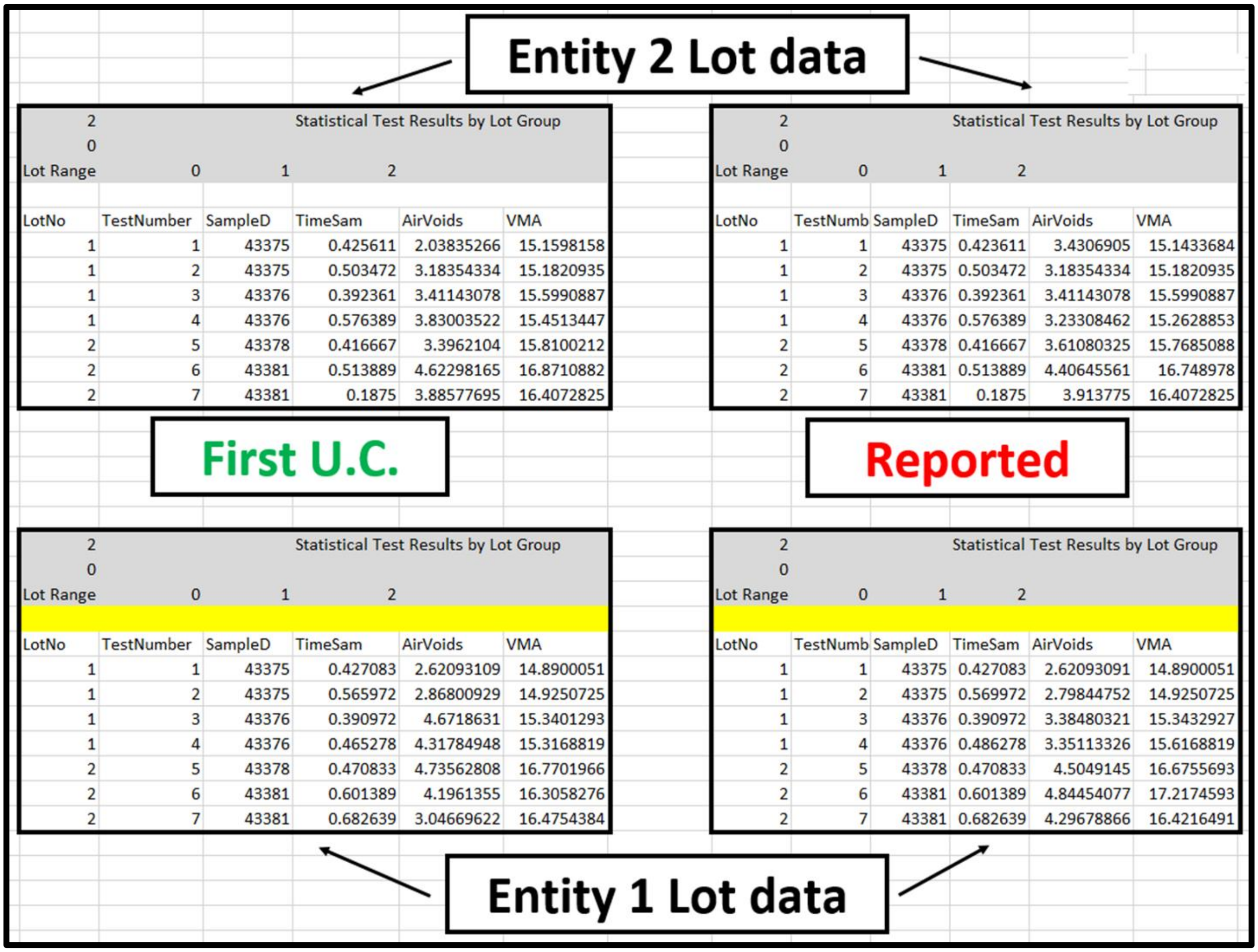
Entity 1

Compaction Testing Data

Lot Numb	Test Num	Test Date	MLD
2	8	43378	91.5867597
2	9	43378	92.460615
2	10	43381	92.8639328
2	11	43381	94.6788631
2	12	43381	94.8469122
2	13	43381	93.8386176

Figure 4.30 Lot-wise calculated air voids/VMA/mainline density parameters based on first U.C. and final reported parameter values (project #7)

But this is not all that is needed for monetary calculation. If the prime contractor material testing data pass statistical tests, the prime contractor data will be used for payment purposes. Sometimes corrections are observed in ITD's data, but not in the contractor data, or vice versa. If corrections only occur in the ITD data, and they result in using the prime contractor data for payment purposes, the corrections do not lead to direct monetary impact that can be identified in our analysis. Another step is to form the lot groups, which impacts the F and t tests used to determine whose reported data should be used for payment. Data corrections can change the lot formation (e.g., the lot shown in Figure 4.31), and thereby impact payments. However, since we followed reported lot formations, such impacts cannot be seen in our analysis results.



2			Statistical Test Results by Lot Group		
0					
Lot Range	0	1	2		
LotNo	TestNumber	SampleD	TimeSam	AirVoids	VMA
1	1	43375	0.425611	2.03835266	15.1598158
1	2	43375	0.503472	3.18354334	15.1820935
1	3	43376	0.392361	3.41143078	15.5990887
1	4	43376	0.576389	3.83003522	15.4513447
2	5	43378	0.416667	3.3962104	15.8100212
2	6	43381	0.513889	4.62298165	16.8710882
2	7	43381	0.1875	3.88577695	16.4072825

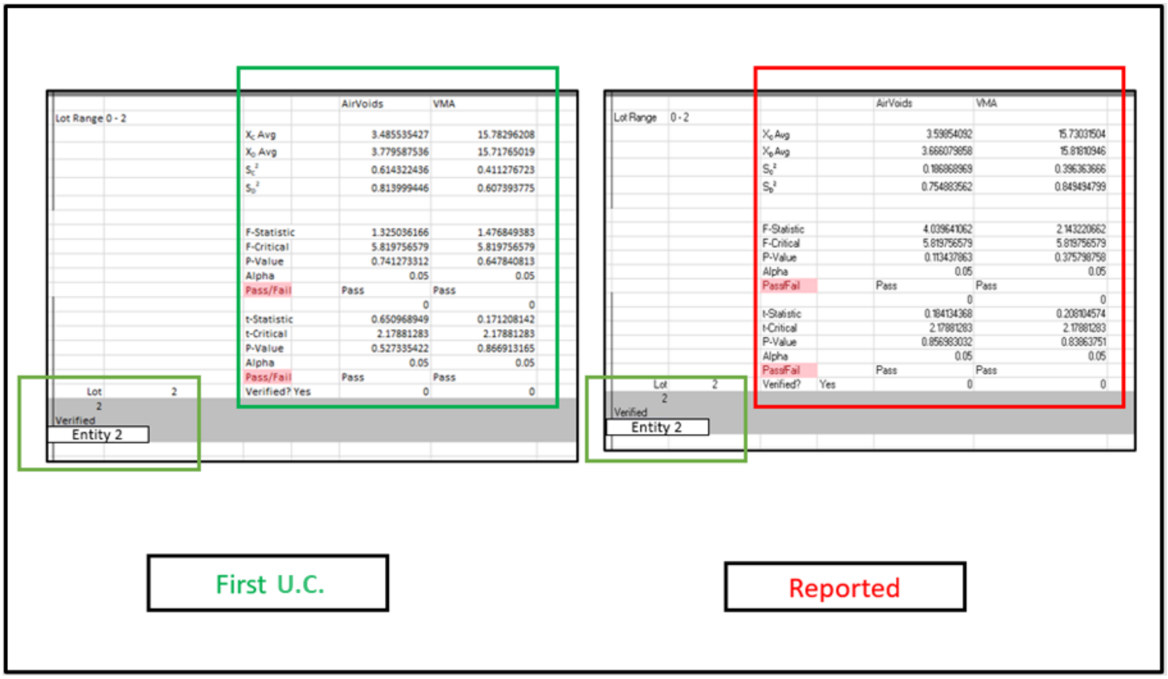
2			Statistical Test Results by Lot Group		
0					
Lot Range	0	1	2		
LotNo	TestNumb	SampleD	TimeSam	AirVoids	VMA
1	1	43375	0.423611	3.4306905	15.1433684
1	2	43375	0.503472	3.18354334	15.1820935
1	3	43376	0.392361	3.41143078	15.5990887
1	4	43376	0.576389	3.23308462	15.2628853
2	5	43378	0.416667	3.61080325	15.7685088
2	6	43381	0.513889	4.40645561	16.748978
2	7	43381	0.1875	3.913775	16.4072825

2			Statistical Test Results by Lot Group		
0					
Lot Range	0	1	2		
LotNo	TestNumber	SampleD	TimeSam	AirVoids	VMA
1	1	43375	0.427083	2.62093109	14.8900051
1	2	43375	0.565972	2.86800929	14.9250725
1	3	43376	0.390972	4.6718631	15.3401293
1	4	43376	0.465278	4.31784948	15.3168819
2	5	43378	0.470833	4.73562808	16.7701966
2	6	43381	0.601389	4.1961355	16.3058276
2	7	43381	0.682639	3.04669622	16.4754384

2			Statistical Test Results by Lot Group		
0					
Lot Range	0	1	2		
LotNo	TestNumb	SampleD	TimeSam	AirVoids	VMA
1	1	43375	0.427083	2.62093091	14.8900051
1	2	43375	0.569972	2.79844752	14.9250725
1	3	43376	0.390972	3.38480321	15.3432927
1	4	43376	0.486278	3.35113326	15.6168819
2	5	43378	0.470833	4.5049145	16.6755693
2	6	43381	0.601389	4.84454077	17.2174593
2	7	43381	0.682639	4.29678866	16.4216491

Figure 4.31 Formation of lot group (project #7)

As described earlier, the second step of the financial analysis is to check the acceptability of the entity 1/entity 2 data through the F and t tests (Figure 4.32). Generally, prime contractor data should be used for payment calculations, if they pass F and t tests; and ITD data should be used otherwise. Now, if entity 1 changed their data to ensure the data from the other party is used for payment calculation, changes in entity 1 data does not show in the payment calculation. Same rule applies to entity 2 data.



Lot Range 0 - 2		AirVoids	VMA
X_C Avg		3.485535427	15.78296208
X_D Avg		3.779587536	15.71765019
S_C^2		0.614322436	0.411276723
S_D^2		0.813999446	0.607393775
F-Statistic		1.325036166	1.476849383
F-Critical		5.819756579	5.819756579
P-Value		0.741273312	0.647840813
Alpha		0.05	0.05
Pass/Fail		Pass	Pass
		0	0
t-Statistic		0.650968949	0.171208142
t-Critical		2.17881283	2.17881283
P-Value		0.527335422	0.866913165
Alpha		0.05	0.05
Pass/Fail		Pass	Pass
Verified?	Yes	0	0

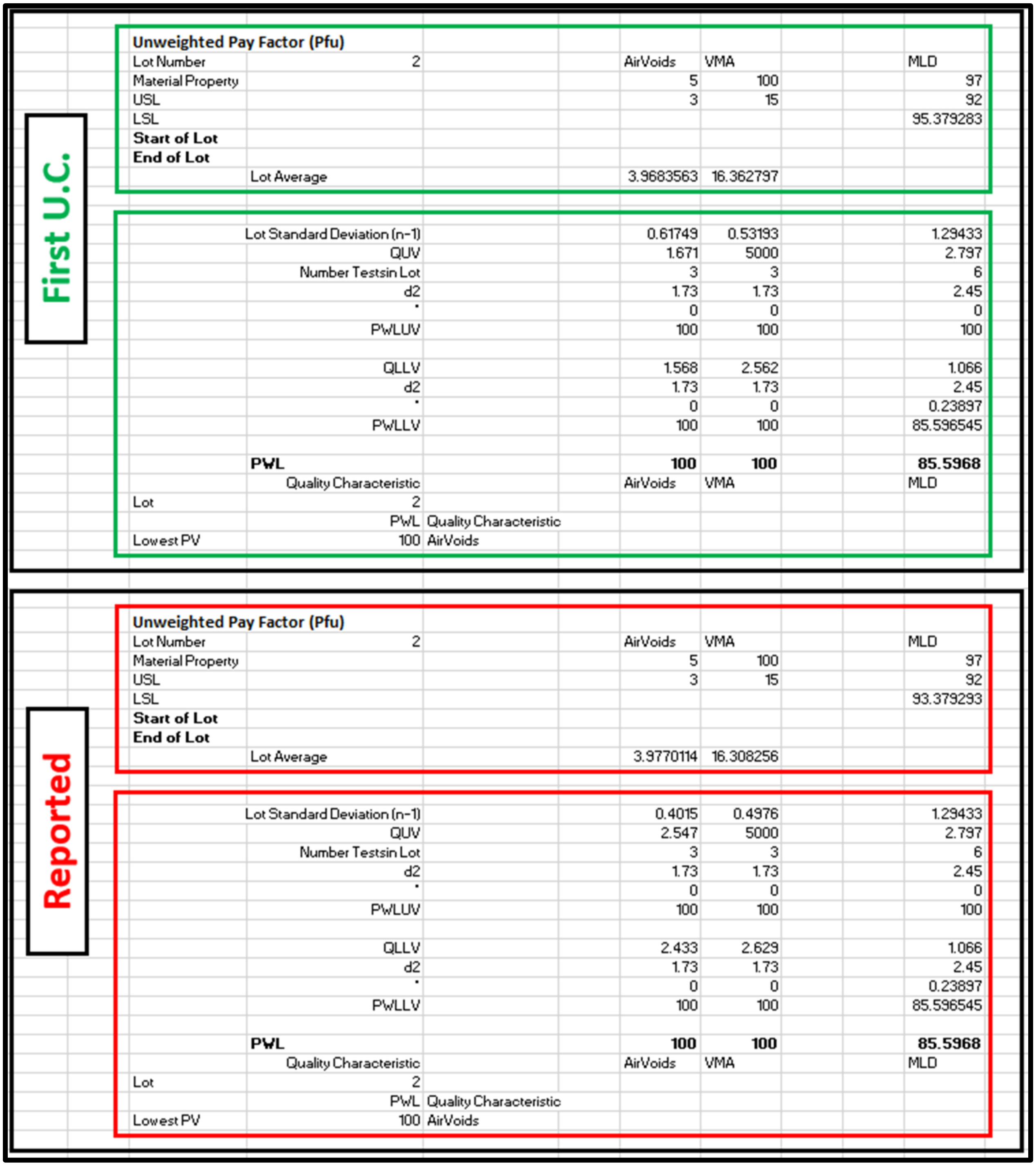
Lot 2
2
Verified
Entity 2

Lot Range 0 - 2		AirVoids	VMA
X_C Avg		3.59854092	15.73031504
X_D Avg		3.666079858	15.81810946
S_C^2		0.186868969	0.396363666
S_D^2		0.754883562	0.849494799
F-Statistic		4.039641062	2.143220662
F-Critical		5.819756579	5.819756579
P-Value		0.113437863	0.375798758
Alpha		0.05	0.05
Pass/Fail		Pass	Pass
		0	0
t-Statistic		0.184134368	0.208104574
t-Critical		2.17881283	2.17881283
P-Value		0.856983032	0.83863751
Alpha		0.05	0.05
Pass/Fail		Pass	Pass
Verified?	Yes	0	0

Lot 2
2
Verified
Entity 2

Figure 4.32 Selection of Entity 1/Entity 2 test result based on F and t tests (project #7)

The next target was to calculate the unweighted pay factor. In this step, the average value of Air Voids, VMA, and Mainline Density is used for each test. This average value, often, can compensate for the test value change, hence not resulting in payment change although test results might have been corrected. Some reported test values were lower than the first U.C. instances, and some were higher. Since a mean value is taken, there is a possibility that the average of corrected test data was close to that of un-corrected data. For instance, the average Air Voids value was 3.96 from the first U.C. calculation, whereas it was 3.97 in the reported section. Similarly, the average VMA value came up as 16.36 from the first U.C. calculation, and it was reported as 16.30. Despite all the clear corrections done on the earlier steps, averaged monetary-related parameters can take values very close to the original values (Figure 4.33).

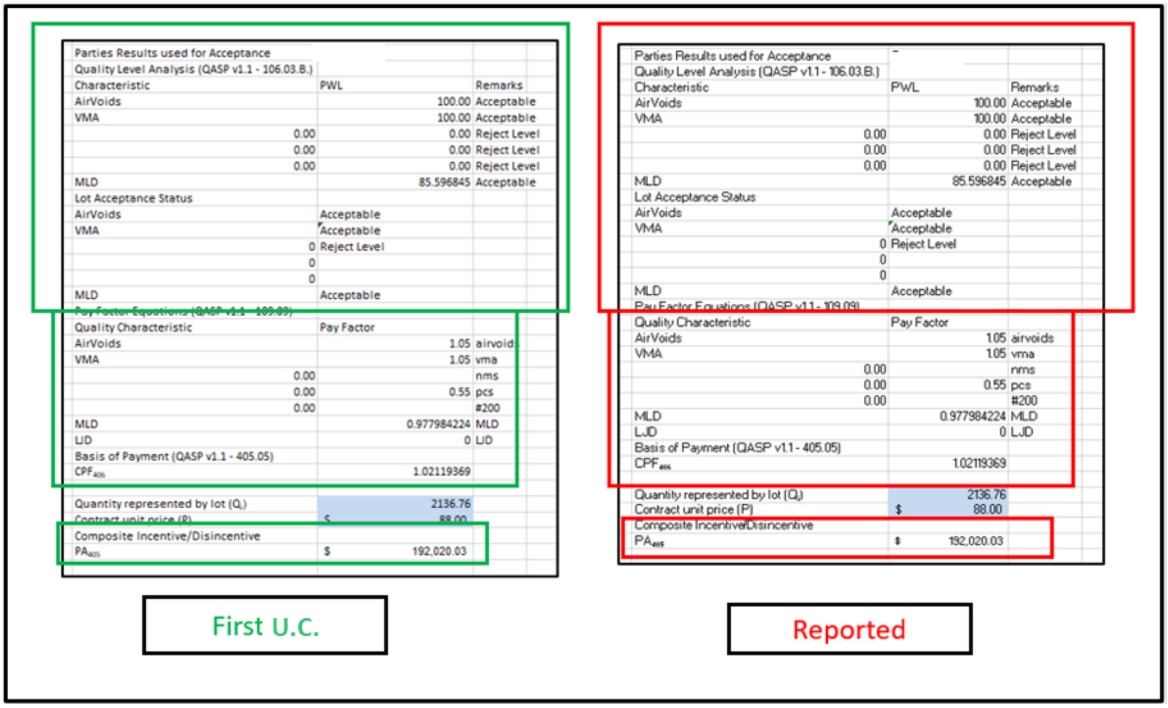


Unweighted Pay Factor (Pfu)						
Lot Number		2		AirVoids	VMA	MLD
Material Property				5	100	97
USL				3	15	92
LSL						95.379283
Start of Lot						
End of Lot						
	Lot Average			3.9683563	16.362797	
		Lot Standard Deviation (n-1)		0.61749	0.53193	1.29433
		QUV		1.671	5000	2.797
		Number Testsin Lot		3	3	6
		d2		1.73	1.73	2.45
		·		0	0	0
		PWLUV		100	100	100
		QLLV		1.568	2.562	1.066
		d2		1.73	1.73	2.45
		·		0	0	0.23897
		PWLLV		100	100	85.596545
	PWL			**100**	**100**	**85.5968**
		Quality Characteristic		AirVoids	VMA	MLD
Lot		2				
		PWL	Quality Characteristic			
Lowest PV		100	AirVoids			

Unweighted Pay Factor (Pfu)						
Lot Number		2		AirVoids	VMA	MLD
Material Property				5	100	97
USL				3	15	92
LSL						93.379293
Start of Lot						
End of Lot						
	Lot Average			3.9770114	16.308256	
		Lot Standard Deviation (n-1)		0.4015	0.4976	1.29433
		QUV		2.547	5000	2.797
		Number Testsin Lot		3	3	6
		d2		1.73	1.73	2.45
		·		0	0	0
		PWLUV		100	100	100
		QLLV		2.433	2.629	1.066
		d2		1.73	1.73	2.45
		·		0	0	0.23897
		PWLLV		100	100	85.596545
	PWL			**100**	**100**	**85.5968**
		Quality Characteristic		AirVoids	VMA	MLD
Lot		2				
		PWL	Quality Characteristic			
Lowest PV		100	AirVoids			

Figure 4.33 Calculation of unweighted pay factor (project #7)

The last step was to determine the PWL value and calculate the monetary values (Figure 4.34). It was seen in Figures 4.32 and 4.33 that because the average value of the secondary parameters was almost equal; the PWL value came precisely the same for these specific tests. The end result was, hence, an identical payment value for both scenarios. It is argued that for some cases no matter how many times data correction had been done, there might still be zero payment impact. Obviously, this does not apply to all projects and tests. As shown in Tables 4.2 and 4.1, data correction has often resulted in overpayment.



Parties Results used for Acceptance		
Quality Level Analysis (QASP v1.1 - 106.03.B.)		
Characteristic	PWL	Remarks
AirVoids	100.00	Acceptable
VMA	100.00	Acceptable
0.00	0.00	Reject Level
0.00	0.00	Reject Level
0.00	0.00	Reject Level
MLD	85.596845	Acceptable
Lot Acceptance Status		
AirVoids	Acceptable	
VMA	Acceptable	
0	Reject Level	
0		
0		
MLD	Acceptable	
Pay Factor Equations (QASP v1.1 - 109.09)		
Quality Characteristic	Pay Factor	
AirVoids	1.05	airvoids
VMA	1.05	vma
0.00		nms
0.00	0.55	pcs
0.00		#200
MLD	0.977984224	MLD
LJD	0	LJD
Basis of Payment (QASP v1.1 - 405.05)		
CPF_{405}	1.02119369	
Quantity represented by lot (Q_L)	2136.76	
Contract unit price (P)	$ 88.00	
Composite Incentive/Disincentive		
PA_{405}	$ 192,020.03	

Parties Results used for Acceptance	-	
Quality Level Analysis (QASP v1.1 - 106.03.B.)		
Characteristic	PWL	Remarks
AirVoids	100.00	Acceptable
VMA	100.00	Acceptable
0.00	0.00	Reject Level
0.00	0.00	Reject Level
0.00	0.00	Reject Level
MLD	85.596845	Acceptable
Lot Acceptance Status		
AirVoids	Acceptable	
VMA	Acceptable	
0	Reject Level	
0		
0		
MLD	Acceptable	
Pay Factor Equations (QASP v1.1 - 109.09)		
Quality Characteristic	Pay Factor	
AirVoids	1.05	airvoids
VMA	1.05	vma
0.00		nms
0.00	0.55	pcs
0.00		#200
MLD	0.977984224	MLD
LJD	0	LJD
Basis of Payment (QASP v1.1 - 405.05)		
CPF_{405}	1.02119369	
Quantity represented by lot (Q_L)	2136.76	
Contract unit price (P)	$ 88.00	
Composite Incentive/Disincentive		
PA_{405}	$ 192,020.03	

Figure 4.34 Calculation of PWL and monetary value (project #7)

Summary

Construction projects are generally performed in a complex dynamic environment and are highly sensitive to data corrections. Failure to take adequate measures to protect these sensitive tasks against data corrections results in higher costs and time overruns. This research leveraged the availability of a unique audit trail dataset (recording sequence of all entered parameter values in a material testing form) to calculate possible monetary impacts of potential unexplained corrections of material testing reports. Such claim of data corrections upholds the necessity for reformation of traditional QC/QA practice which seems to be vulnerable to intentional or unintentional data inconsistencies and can cause loss in monetary values. The monetary payment calculation procedures followed by ITD were successfully replicated and lot-wise payments for various lots of 12 HMA projects prior to and after data corrections were calculated. The majority of the projects indicate possible overpayment, even with the researchers' conservative approach that was utilized for monetary calculations. Detailed analysis showed that corrections in each major parameter value may result in $1,000-$5,000 overpayment. Further, a great fraction of the analyzed lots did not pass either the precision test or the PWL thresholds – i.e. were at an acceptable level – raising questions about the efficacy of QC/QA practices. We note that ITD changed its QC/QA practice in 2020 in an attempt to remove some of the deficiencies in this process and to bring more clarity and rigor to material testing in Idaho.

5. Sensitivity Analysis of HMA Test Parameters

Introduction

Sensitivity analysis is an efficient and effective approach to identify the most influential parameters in any modeling effort. Sensitivity analysis is particularly important for the HMA projects which are very complex and are characterized by a large number of parameters. Importantly, all the parameters do not equally contribute to the payment related secondary parameters, acceptance or rejection of lot materials, and/or final HMA quality. Sensitivity analysis is particularly useful to identify the most influential testing parameters on HMA outcomes, and can inform potential further scrutiny of specific sections of the reported test results. This can lead to a better understanding of HMA production and can potentially enhance HMA production quality. Lack of this knowledge, on the contrary, may result in unnecessary work and evaluation of insensitive parameters that is time consuming. This lack of understanding may dilute efforts that otherwise could have led to conclusive findings about potential data corrections.

Sensitivity analysis can be conducted for a variety of purposes including, determination of input parameters that contribute most to variability of output asphalt mix design properties, locating the optimal regions within the parameter space in future calibration studies, increasing knowledge of parameter behavior to reduce the uncertainty of asphalt mix design properties, and identifying which parameters are insensitive and can possibly be held constant in the material testing report. Studies with similar purpose have been done to provide design guides for HMA using sensitivity analysis (ARA 2004; Schwartz et al. 2013, 12). For example, the MEPDG from AASHTO defines what is a feasible design scenario and contains pavement analysis and performance predictions for such scenarios (MEPDG 2008; AASHTOWare Pavement Me Design 2011). In these studies, a large quantity of parameters, characterizing the pavement materials, layers, design features, and condition, were used to obtain MEPDG performance predictions for the anticipated climatic and traffic conditions (MEPDG 2008). In one particular, but probably most informative, type of sensitivity analysis, majority of the parameters are kept constant, and only one parameter value is changed to make inferences regarding the significance of input parameters in calculation of asphalt mix design properties (El-Basyouny and Witczak 2004; Graves and Mahboub 2006, 122).

Sensitivity Analysis Method

In this chapter, we adopt a one-parameter-at-a-time sensitivity analysis approach. In doing so, we have adopted a “reasonable” parameter value for all parameters of the ITD-0777 form (see Appendix C). This value is selected based on the average reported values in all HMA project testing reports from year 2017 available to us. Further, we have determined the minimum and maximum reported value for each parameter, which we later used as the possible range for the associated parameter. “Reasonable” parameter values and their ranges are reported in Table 5.1. For each parameter, we have assumed a normal distribution centered at the average reported parameter and selected a standard deviation that equals to one-sixth of the full parameter range. In a normal distribution, 3 standard deviation divergence to each side of the mean would cover 99% of all possible perturbations. Hence, the entire variability range for each parameter would be roughly equal to 6 standard deviations. Further, we have set all other parameter values at their constant level, and using a Monte Carlo approach, we generated 10,000 random samples from the parameter of interest using the normal distribution that was described earlier. For each perturbation, we have calculated secondary parameter values including VMA, Gmb, Gmm and Pa. In other words, we are estimating the relative impact of each input parameter on VMA, Gmb, Gmm and Pa, if all other parameters are constant. We report results in terms of plots and coefficient of variation (ratio of standard deviation to mean).

Results

A total of 55 different input parameters from material tests were considered for sensitivity analysis, the impacts of which were analyzed on four major asphalt mix design properties (i.e., Gmm, Gmb, Pa, and VMA). These 55 parameters consist of values reported from summary of mix properties, Bowl properties, and Ignition furnace, among others. For all 55 parameters a mean and standard deviation was calculated based on the reported data for all ITD projects in 2017 available to us. Here we used the most common form of sensitivity analysis namely the independent parameter perturbation (Ferreira et al. 1995, 493) combined with the Monte Carlo simulation (Shaffer et al. 1988, 1782). These parameters were allowed to vary around their mean values independently following a normal distribution, and the sensitivity of asphalt mix design properties to these parameters were analyzed. Results indicated that some parameters only influence one mix design parameter (Gmm, Gmb, Pa, and VMA), some impact multiple and others do not have any significant impact on any of the mix design properties. Their level of influence also varies widely from one input parameter to another, which will be the focus of this chapter.

In this chapter, we provide some example figures and provide an overview of the results for all parameters, and we refer the reader to Appendix E for detailed results. Figure 5.1 shows that the changes in mass of bowl sample increment 1 significantly influences Pa and Gmm, whereas it has no influence on VMA and Gmb. This marks the mass of bowl sample increment 1 as a variable that can be potentially corrected to obtain satisfactory HMA test results. Note that it is likely to obtain negative values both in VMA and Pa, which are physically not possible, due to random sampling of parameter values from a normal of distribution, where for example the ratio of Gmb and Gmm is more than one (see equations for VMA and Pa).

$$VMA = 100 - (\frac{G_{mb} . P_s}{G_{sb}})$$

and

$$P_a = 100 - (\frac{100 \times G_{mb}}{G_{mm}})$$

Where:

VMA = Voids in mineral aggregate
P_a= Air voids in compacted mixture
G_{sb} = Bulk specific gravity of the aggregate
G_{mb} = Bulk specific gravity of compacted mixture.
P_s = Aggregate, percent by total dry weight of mixture

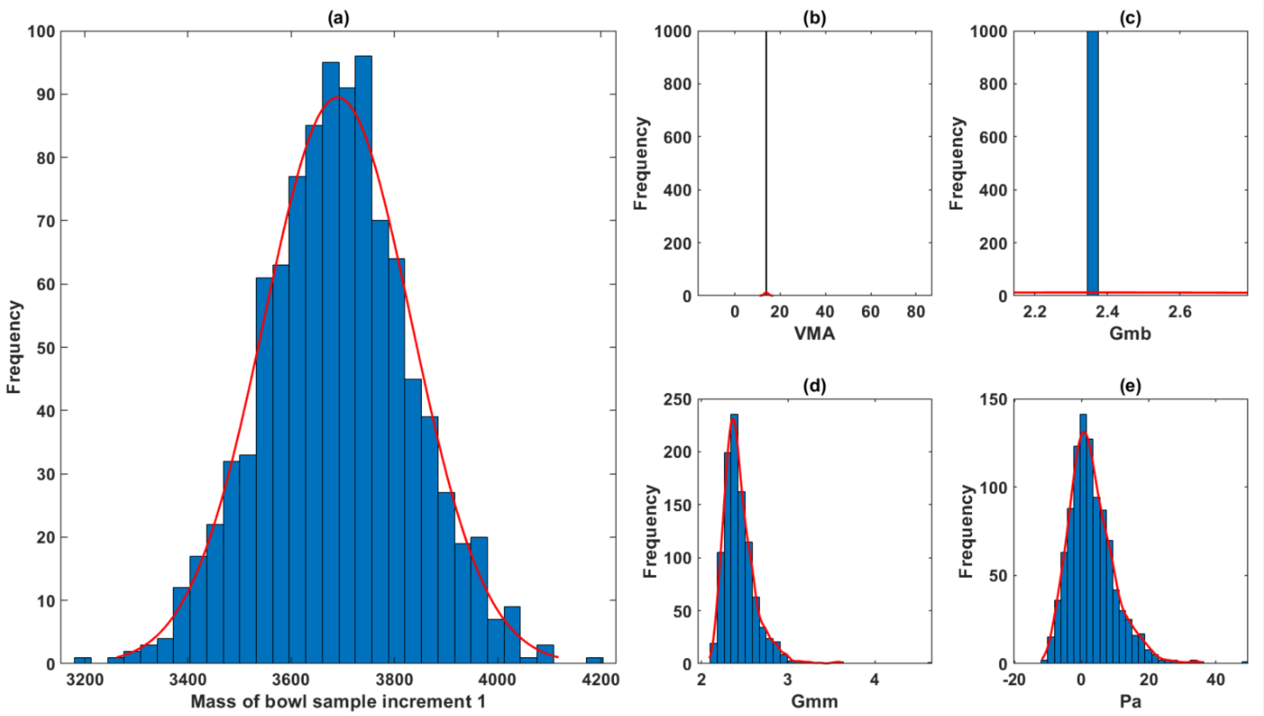


Figure 5.1 Sensitivity of VMA, Gmb, Gmm, and Pa parameters (sub plots b, c, d and e, respectively) with respect to changes in mass of bowl sample increment 1.

Contrary to the mass of bowl sample increment 1, it can be seen in Figure 5.2 that changes in ignition furnace correction factor has no significant influence on the four major asphalt mix design properties and therefore it is an unlikely candidate for analysis of data correction cases.

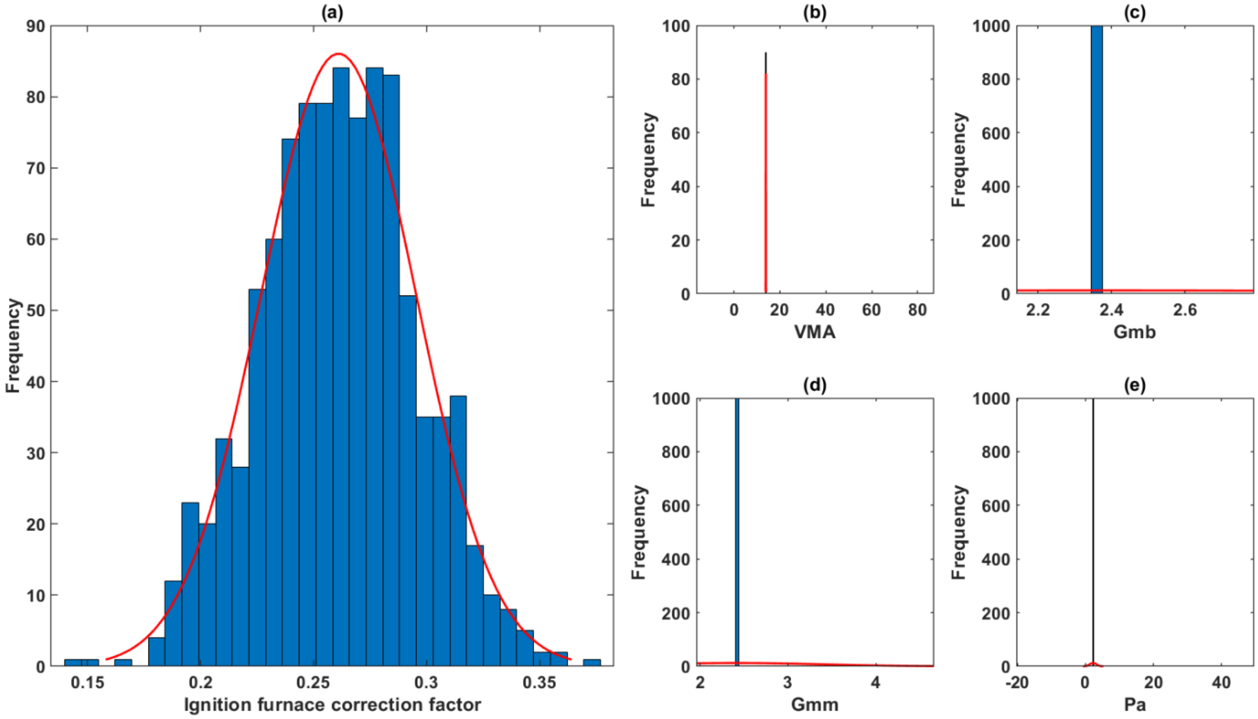


Figure 5.2 Sensitivity of VMA, Gmb, Gmm, and Pa parameters (sub plots b, c, d and e, respectively) with respect to changes in ignition furnace correction factor.

Mass of puck dry specimen 1 is another important parameter that influences three mix design properties, namely VMA, Pa, and Gmb. However, it has weaker influence on Pa and VMA values as compared to Gmb and therefore is more suited for studying data correction cases that are rooted in changes in Gmb (Figure 5.3).

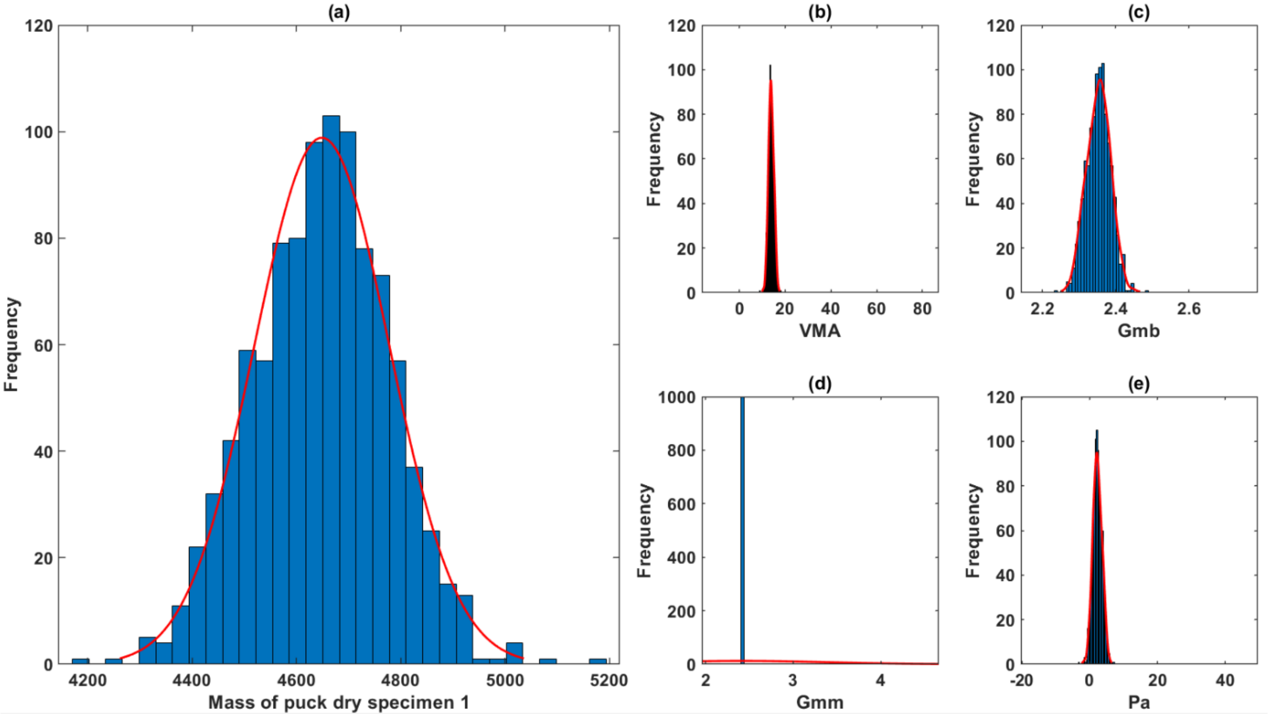


Figure 5.3 Sensitivity of VMA, Gmb, Gmm, and Pa parameters (sub plots b, c, d and e, respectively) with respect to changes in mass of puck dry specimen 1.

A few parameters, such as chamber set point, did not vary across ITD projects and therefore were kept as constants in the analysis (Figure 5.4). The sensitivity analysis results for the remaining 51 parameters are presented in Appendix E.

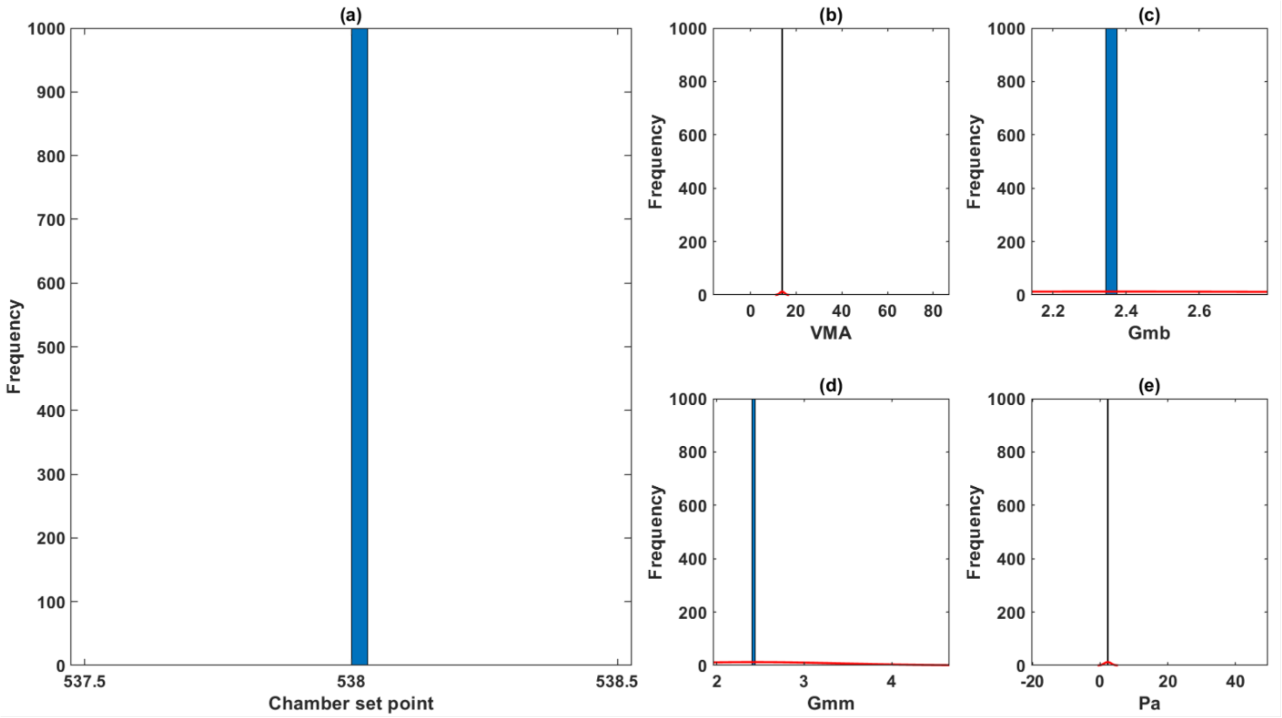


Figure 5.4 Sensitivity of VMA, Gmb, Gmm, and Pa parameters (sub plots b, c, d and e, respectively) with respect to changes in chamber set point.

Summary and Conclusions

As shown in the sensitivity analysis results, the degree to which these 55 parameters influence four major asphalt mix design properties vary significantly. In order to outline the overall impact of input parameters on mix design properties, we summarize the sensitivity analysis results in Table 5.1. This table enlists the coefficient of variation (standard deviation divided by mean) for asphalt mix design properties with respect to changes in each parameter. Higher coefficient of variations is associated with larger influence of the input parameters on the mix design properties. A coefficient of variation of zero represents no impact of the input parameter on the mix design property. The most important parameters that significantly control the values of four asphalt mix design properties are summarized below:

- The parameters that significantly influence VMA are mass pan and initial sample, mass pan and dry sample (AASHTO T 329), mass basket assembly & initial sample, mass basket assembly & final aggregate (AASHTO T 308), submerged weight of puck in water specimen 1 & 2, and weight of puck ssd specimen 1 & 2 (AASHTO T 166). The rest of the parameters either have small or no impact on the VMA value.

- In case of Pa, the influential parameters are submerged weight of bowl inc 1 & 2, submerged weight of bowl and sample inc 1 & 2, mass of bowl increment 1 & 2, mass of bowl sample increment 1 & 2 (AASHTO T 209), weight of puck ssd specimen 1 & 2, submerged weight of puck in water specimen 1 & 2, and mass of puck dry specimen 1 & 2 (AASHTO 166).
- The parameters that control changes in Gmm, with a relatively similar impact, include mass of bowl increment 1 & 2, mass of bowl sample increment 1 & 2, submerged weight of bowl inc 1 & 2, submerged weight of bowl and sample inc 1 & 2 (AASHTO 209).
- The parameters that control Gmb include weight of puck ssd specimen 1 & 2, submerged weight of puck in water specimen 1 & 2, and mass of puck dry specimen 1 & 2 (AASHTO 166).

The attention, when analyzing QA results, should be focused on the parameters that have the most significant impact on the asphalt mix design properties, as they are ultimately crucial for quality of HMA projects. A deep understanding of the sensitivity of HMA projects' quality to the input parameters will allow ITD to focus attention on monitoring the right parameters and will permit the best use of data in search for data correction cases. For example, experience has shown that Gmm is one of the primary derived factors that is changed through unexplained corrections. If the Gmm values are artificially corrected, the Gse values tend to follow the trend of the binder content in a mix. However, this is not expected as the Gse value should stay relatively constant for a given aggregate-binder combination. Therefore, variations in Gse values can serve as a surrogate indicator of unexplained corrections with respect to Gmm. Focusing on such details during material testing data validation will facilitate the implementation of a robust quality assurance program.

Table 5.1 Descriptive statistics of mix design properties in response to changes in input parameters

Parameters	Mean	Min	Max	VMA	Pa	Gmm	Gmb
Mass of Bowl Increment 1 (AASHTO T 209)	2123	1697.3	2603.2	NA	2.1268	0.0713	NA
Mass of Bowl Increment 2 (AASHTO T 209)	2122.9	1783.7	2484.6	NA	1.6505	0.0541	NA
Mass of Bowl Sample Increment 1 (AASHTO T 209)	3684.4	3192.9	4189.2	NA	2.2552	0.0815	NA
Mass of Bowl Sample Increment 2 (AASHTO T 209)	3688.7	3287.5	4176.3	NA	2.2008	0.0774	NA
Water Bath Temp Inc 1 (AASHTO T 209)	76.3	74.8	78.9	NA	NA	NA	NA
Water Bath Temp Inc 2 (AASHTO T 209)	76.6	74.3	78.4	NA	NA	NA	NA
Submerged Weight of Bowl and Sample Inc1 (AASHTO T 209)	2250.3	2004.4	2479.2	NA	2.2797	0.0631	NA
Submerged Weight of Bowl and Sample Inc2 (AASHTO T 209)	2254.2	2041.5	2485.6	NA	2.0578	0.0558	NA
Submerged Weight of Bowl Inc1 (AASHTO T 209)	1337.9	994.8	1653.5	NA	2.3832	0.0913	NA
Submerged Weight of Bowl Inc2 (AASHTO T 209)	1337.9	1015.3	1726.1	NA	3.1699	0.0932	NA
Mass of Sample Specimen 1 (AASHTO T 312)	4654.6	4187.6	5103.2	NA	NA	NA	NA
Mass of Sample Specimen 2 (AASHTO T 312)	4654.9	4285.4	5105.8	NA	NA	NA	NA
Mass of Sample Design Mass (AASHTO T 312)	4650	4277.5	5184.6	NA	NA	NA	NA

Temp of Sample When Placed in Mold Specimen 1 (AASHTO T 312)	300	108	423	NA	NA	NA	NA
Temp of Sample When Placed in Mold Specimen 2 (AASHTO T 312)	300	170	427	NA	NA	NA	NA
Time Compaction Begins Specimen 1 (AASHTO T 312)	10	0	25	NA	NA	NA	NA
Time Compaction Begins Specimen 2 (AASHTO T 312)	10	0	21	NA	NA	NA	NA
Sample Height Specimen 1 (AASHTO T 312)	113.7	111.3	115.5	NA	NA	NA	NA
Sample Height Specimen 2 (AASHTO T 312)	113.5	112.3	115.2	NA	NA	NA	NA
Surface Temperature Specimen 1 (AASHTO T 166)	71.4	64.5	78.1	NA	NA	NA	NA
Surface Temperature Specimen 2 (AASHTO T 166)	74.6	69.6	80.4	NA	NA	NA	NA
Water Bath Temp Specimen 1 (AASHTO T 166)	77.8	76.5	79.4	NA	NA	NA	NA
Water Bath Temp Specimen 2 (AASHTO T 166)	77.7	76.4	78.9	NA	NA	NA	NA
Mass of Puck Dry Specimen 1 (AASHTO T 166)	4651.7	4183.3	5197.8	0.0862	0.5935	NA	0.0139
Mass of Puck Dry Specimen 2 (AASHTO T 166)	4649.7	4191.3	5134.4	0.0966	0.661	NA	0.0156
Submerged Weight of Puck in Water Specimen 1 (AASHTO T 166)	2681.6	2285.7	3123.6	0.2096	1.5764	NA	0.0332
Submerged Weight of Puck in Water Specimen 2 (AASHTO T 166)	2677.1	2217.9	3198.3	0.2347	1.8243	NA	0.037
Weight of Puck ssd Specimen 1 (AASHTO T 166)	4656.6	4334.2	4997.6	0.1663	1.2671	NA	0.0263
Weight of Puck ssd Specimen 2 (AASHTO T 166)	4653.5	4353.8	5033.8	0.1673	1.2408	NA	0.0266
Mass Basket Assembly (AASHTO T 308)	3060.9	2210.5	3864.7	0.0655	NA	NA	NA
Mass Basket Assembly & Initial Sample (AASHTO T 308)	4599.6	4193.5	4957.2	0.5823	NA	NA	NA
Mass Basket Assembly & Final Aggregate (AASHTO T 308)	4509.3	4258.3	4804.9	0.3809	NA	NA	NA
Ignition Furnace Correction Factor (AASHTO T 308)	0.26	0.15	0.36	0.0023	NA	NA	NA
Elapsed Time (AASHTO T 308)	46	23	70	NA	NA	NA	NA
Temperature Compensation Factor (AASHTO T 308)	0.19	NA	0.3	NA	NA	NA	NA
Chamber Set Point (AASHTO T 308)	538	538	538	NA	NA	NA	NA
Calibration Factor (AASHTO T 308)	NA	NA	0.05	NA	NA	NA	NA
Uncorrected Binder Content (AASHTO T 308)	5.94	5.39	6.57	0.0042	NA	NA	NA
Oven Temperature (AASHTO T 329)	320	315	326	NA	NA	NA	NA
Pan Mass (AASHTO T 329)	683.4	512.1	874.7	0.0001	NA	NA	NA
Mass Pan and Initial Sample (AASHTO T 329)	1892.7	1385.9	2484.8	0.935	NA	NA	NA
Mass Pan and Dry Sample (AASHTO T 329)	1892.3	1657.1	2127.6	0.4325	NA	NA	NA

6. Summary, Conclusions, and Recommendations for Future Research

This project tackled an important, challenging and controversial topic focused on evaluating the robustness of the quality assurance process for HMA projects. Materials testing is the backbone to ensure HMA projects meet their design life; implementation of a robust materials testing program can save millions of dollars through warranting the quality of asphalt pavements. Robustness of this process, however, can be undermined by potential corrections of material testing data. Instances of data corrections observed in previous material testing reports have been observed in Idaho and other states, which prompted us to investigate the perception of Idaho Transportation Department and other state Department of Transportation engineers about data correction cases, the prevalence of such instances in material testing reports and the potential impacts of such cases on the final payment of the projects.

In a survey that was sent in an online and a pdf format to ITD and other state DOT engineers, we asked their perception regarding the prevalence of unexplained corrections in HMA projects. We received a much larger participation from Idaho than that of all other states combined. We understand that due to the sensitive nature of this survey, some engineers may not be inclined to participate in this study. However, due to recent investigation of data corrections in HMA projects in Idaho, its media coverage and engagement of various districts, we observed good participation from ITD participants. A total of 75 participants responded representing 48 ITD employees and 27 from several other DOTs. Generally, there is a perception that HMA projects do not meet their design life, in part due to deficient construction materials and due to unrepresentativeness of material testing reports. Climatic factors, errors made by the prime contractor, and underestimation of traffic volume were also implicated as causes of the HMA projects not meeting their design life. While there is evidence of a perception that material testing reports may be corrected for potential prime contractor gains, strenuous working hours and avoiding conflict with prime contractor and state DOTs are also implicated as potential factors that played into material testing data correction. Overall, although concerns are prevalent, engineers in Idaho trust that ITD is investigating this issue and will make amends. Indeed, ITD revolutionized its quality assurance approach and project acceptance procedures since 2018 with more to be done as described in the Foreword to address these raised concerns.

This project provided an unprecedented opportunity to investigate the source of discrepancies and unexplained trends in materials testing data reported by ITD and prime contractors. An audit trail macro/algorithm that was built into the ITD-0777 form recorded every instance of data entry to the form. This provided a series of values for each parameter, as well as some other statistics such as time of entry. Some parameters were only reported once, which is expected, but others were reported multiple times. Through an extensive manual investigation of the audit trail data that affects acceptance and payment, the research team determined patterns that could be determined as Plausible Corrections (PC) associated with typographical mistakes. But there were other instances that we could not categorize as such, after exhausting all possibilities, and we categorized them as Unexplained Corrections (UC). We then devised algorithmic logics to automatically categorize every instance of change in parameter values (that affect payment) into PC and UC categories. This algorithm marked all changes with proper categories, and we manually checked several instances from each project to ensure the results are robust. We conducted this analysis over 15 available projects that were constructed in 2018, and determined that data reported by the prime contractor as well as ITD have seen changes in reported parameter values. We note that both entities may have used independent labs for testing and reporting. In this research project, to ensure objectivity of the data analysis process as well as to maintain anonymity of the reporting entity (ITD or prime contractor), the two entities were referred to as 'Entity 1' and 'Entity 2', not necessarily in the same order. Analysis of the audit trail data reported by Entity 1 showed that a total of 595 unique parameters affecting prime contractor payment were changed 2,268 times, with the changes being categorized as U.C. On the other hand, a total of 316 unique parameters affecting payment were changed 660 times, with the changes being categorized as P.C. For entity 2-reported data, a total of 387 and 280 unique parameters that affect prime contractor payment were changed 1,266 and 587 times, with the changes being categorized as U.C. and P.C., respectively. Furthermore, results indicated that parameters with major payment impact were corrected more than two times on average per parameter, with some changing more than five or six times. Parameter values for PC cases were mostly changed only one time. We note that algorithmic categorization of parameter value changes cannot be interpreted as facts. Further, investigation of the causes of data corrections is beyond the scope of this study.

We furthered the analysis by testing whether or not changes in parameter values resulted in financial impact either benefiting ITD or the prime contractors. For this purpose, we replicated all procedures from ITD-0777 form to calculate secondary payment-related parameters and translated them to lot-based pay factors and payments. We hypothesize that the first "acceptable" value of parameters that are categorized as UC represents the original value that was observed from the test. We calculated lot-wise payments for first and last parameter values that were labeled as UC, and attributed the changes in the project payment to financial repercussions of suspicious activities in material testing reports. All other parameters, i.e., those that were not changed and those that were categorized as PC, were kept as their final reported values. This analysis was conducted over 12 of the 15 available projects, which provided all the required data for payment analysis. Our conservative analysis shows that 10 of the 12 studied projects show an overpayment, ranging between $14,000 to $360,000, with 2 projects showing a nominal decrease of $-400 and $-3,000. Results show that change in each major parameter may translate to $1,000-$5,000 extra payment. Further, results show that changes in parameter values do not necessarily translate to a financial outcome, and many parameter changes translated to passing statistical (F- and t-) tests; these parameters would have failed the standard quality tests if original values were used. This was observed for percent-within-limits values and precision tests.

Finally, it is paramount to study the sensitivity of payment-related parameters to each material testing parameters, as this can inform future investigations and preventive actions. We accomplished sensitivity analysis through 'leave-one-parameter-out Monte Carlo analysis'. We define domain range for each parameter of the ITD-0777 form based on the values reported in the 15 available projects from 2018. We then randomly draw values for one parameter from a normal distribution that encapsulate the entire parameter range and fix all other parameters to their "expected" values, and compute the payment-related secondary parameters accordingly. Results show that parameters that notably influence VMA are **"**mass of pan and initial sample", "mass of pan and dry sample", "mass of basket assembly & initial sample", "mass of basket assembly & final aggregate", "submerged weight of puck in water specimen 1 & 2", and "weight of puck in SSD specimen 1 & 2" (AASHTO T 209 for Gmm and T 166 for Gmb). The rest of the parameters either have small or no impact on the VMA value. In case of Pa, the influential parameters are "submerged weight of bowl inc 1 & 2", "submerged weight of bowl and sample inc 1 & 2", "mass of bowl increment 1 & 2", "mass of bowl sample increment 1 & 2", "weight of puck ssd specimen 1 & 2", "submerged weight of puck in water specimen 1 & 2", and "mass of puck dry specimen 1 & 2" (AASHTO T 209 for Gmm and T 166 for Gmb). The parameters that control changes in Gmm, with a relatively similar impact, include "mass of bowl increment 1 & 2", "mass of bowl sample increment 1 & 2", "submerged weight of bowl inc 1 & 2", "submerged weight of bowl and sample inc 1 & 2" (AASHTO T 209). The parameters that control Gmb include "weight of puck ssd specimen 1 & 2", "submerged weight of puck in water specimen 1 & 2", and "mass of puck dry specimen 1 & 2" (AASHTO T 166).

The attention, when analyzing QC/QA results, should be focused on the parameters that have the most significant impact on the asphalt mix design properties, as they are ultimately crucial for quality of HMA projects. Understanding the sensitivity of HMA parameters will allow ITD to focus attention on monitoring the most sensitive parameters that influence prime contractor payment.

Recommendations for the Future

i. Additional oversight including more quality assurance controls should be implemented to ensure data corrections, and the number of data corrections, are reasonable.
ii. Rigorous training of field engineers and technicians (both prime contractor and state DOT) involved in HMA production, quality control, and acceptance testing is crucial to avoid potential mistakes and suspicious activities. Emphasis should be on the importance of test accuracy and repeatability, and how they affect the end product.
iii. Long-term monitoring of the projects that are identified in this project to be associated with potential data corrections can provide significant insights about how these suspicious activity instances translated to performance of asphalt pavements.
iv. Extensive review of agency-adopted specifications related to HMA mix design and construction is recommended. Special care should be taken to ensure the specifications and tolerances are developed based on materials commonly used in the region. Setting "unreasonable" targets for material quality will ultimately lead to undesirable practices and inferior pavement performance.
v. The lot reformation was not possible in this analysis. If there is another way of lot reformation after the project has been completed, payments can be recalculated.

vi. A forensic analysis which involves a socioeconomical and cultural analysis is recommended to shed further light on the reasoning behind suspicious activities in the HMA construction projects.

vii. A detailed analysis of whether certain prime contractors or certain individuals are more prone to suspicious activities can be conducted. A further analysis can focus on the reasons prompting or enabling them to act suspiciously.

viii. Using modern technology, online apps, and streamlined process that promote automated data collection and minimize human interference in reporting can help minimize unexplained activities, and can be the subject of further research.

7. Cited Works

AASHTO. 2016. “Standard practice for developing a quality assurance plan for hot mix asphalt (HMA).” R 42-06. Washington, DC: AASHTO. 2016.

AASHTO. 2008. “Mechanistic–Empirical Pavement Design Guide, Interim Edition: A Manual of Practice.” AASHTO, Washington, D.C., 2008.

AASHTO R-10. 2016. Standard Recommended Practice for Definition of Terms Related to Quality and Statistics, American Association of State Highway and Transportation Officials, Washington, D.C., 2016.

AASHTOWare Pavement ME Design. AASHTO, 2011. http://www.aashtoware.org/Pavement/Pages/default.aspx.

Akkinepally, R., and Attoh-Okine, N. 2006. “Quality control and quality assurance of hot mix asphalt construction in Delaware.” Delaware Department of Transportation.

Al-Khayat, Haydar, Charles Gurganus, David E. 2020. Newcomb, and Maryam S. Sakhaeifar. "Developing Specification Limits for Hot Mix Asphalt Properties and Impact on Pay Factors." Journal of Transportation Engineering, Part B: Pavements 146, no. 3 (2020): 05020002. https://doi.org/10.1061/JPEODX.0000185

Al-Qadi, Imad L., José J. Rivera-Perez, Hasan Ozer, Watheq Sayeh, Javier J. García Mainieri, Hadi Meidani, John Huang, and Adam Hand. 2020. “Data Trends and Variability in Quality Control for Performance and Pay for Performance Specifications: Statistical Analysis.” Illinois Center for Transportation/Illinois Department of Transportation, 2020. https://doi.org/10.36501/0197-9191/20-006

“America's Infrastructure Report Card” ASCE's 2017 Infrastructure Report Card, American Society of Engineers, 2017. https://www.infrastructurereportcard.org/wp-content/uploads/2017/01/Roads-Final.pdf

ARA, Inc., ERES Consultants Division. 2004. “Guide for Mechanistic–Empirical Design of New and Rehabilitated Pavement Structures.” Final report, NCHRP Project 1-37A. Transportation Research Board of the National Academies, Washington, D.C., 2004. http://www.trb. org/mepdg/guide.htm.

Arambula, Edith, and Nasir Gharaibeh. 2014. "Methods for Accumulating Construction and Material Quality Test Results and Their Effect on Acceptance Decisions." Journal of Construction Engineering and Management 140, no. 8 (2014): 06014006. https://doi.org/10.1061/(ASCE)CO.1943-7862.0000867

Association of Certified Fraud Examiners (ACFE). 2010. “Report to the nations on occupational fraud and abuse.” Retrieved on August 1, 2020. Accessed at: https://www.acfe.com/uploadedFiles/ACFE_Website/Content/documents/rttn-2010.pdf

Atwood, Joseph A., James F. Robison-Cox, and Saleem Shaik. 2006. "Estimating the prevalence and cost of yield-switching fraud in the federal crop insurance program." American journal of agricultural economics 88, no. 2 (2006): 365-381.

Bagli, C. V. 2018. “Investigators Eye Possible $100 Million Construction Fraud.” The New York Times. Retrieved on August 1, 2020. Accessed at: https://www.nytimes.com/2018/02/26/nyregion/bloomberg-interior-construction-fraud.html

Beainy, Fares. "Non-contact sensor for the real-time measurement of stiffness of asphalt pavements during compaction." (2011).

Beainy, Fares, Sesh Commuri, and Musharraf Zaman. 2012. "Quality assurance of hot mix asphalt pavements using the intelligent asphalt compaction analyzer." Journal of Construction Engineering and Management 138, no. 2 (2012): 178-187. https://doi.org/10.1061/(ASCE)CO.1943-7862.0000420

Blankenship Ph. 2021. “Improving pavement density.” The Magazine of the Asphalt Institute, 2021. Accessed at: http://asphaltmagazine.com/density/

Bowen, Paul, Akintola Akintoye, Robert Pearl, and Peter J. Edwards. 2007. "Ethical behaviour in the South African construction industry." Construction Management and Economics 25, no. 6 (2007): 631-648. https://doi.org/10.1080/01446190701225707

Burati J.L., Weed R. M., Hughes C. S., and Hill H. S., “Optimal Procedures for Quality Assurance Specifications.” Federal Highway Administration Report No. FHWA-RD-02-095, 2003, p.4. Accessed at: https://www.fhwa.dot.gov/publications/research/infrastructure/pavements/pccp/02095/index.cfm

Buttlar WG, Harrell M. 1998. “Development of end-result and performance-related specifications for asphalt pavement construction in Illinois.” InCrossroads 2000 Iowa State University and Iowa Department of Transportation 1998.

Caltrans. 2020. “QUALITY CONTROL MANUAL FOR HOT MIX ASPHALT USING STATISTICAL PAY FACTORS.” link: https://dot.ca.gov/-/media/dot-media/programs/construction/documents/construction-standards/hma-intelligent-compaction-construction/202002-quality-control-manual-for-hma-using-spf.pdf

Chang G, Gallivan VL. 2011. “Accelerated implementation of intelligent compaction technology for embankment subgrade soils, aggregate base, and asphalt pavement materials.” United States. Federal Highway Administration; 2011 Jan 1.

Craig, G. 2019. “Three contractors accused of fraud in billion-dollar school modernization project.” Democrat and Chronicle. Retrieved on August 1, 2020. Accessed at: https://www.democratandchronicle.com/story/news/2019/08/27/three-contractors-accused-fraud-construction-plan-reach-deal-feds/2131956001/

De Jong, Maarten, William P. Henry, and Neill Stansbury. 2009. "Eliminating corruption in our engineering/construction industry." Leadership and Management in Engineering 9, no. 3 (2009): 105-111.

Dutton, A. 2020. “Idaho highway contractors changed records hundreds of times. Then, they got bonuses.” Idaho Statesman. Retrieved on August 1, 2020. Accessed at: https://www.idahostatesman.com/news/local/investigations/watchdog/article239690033.html

Dutton, 2021, “Idaho overpaid road builders millions for shoddy work. State says it won’t happen anymore”, Idaho Capital Sun. Retrieved on July 1, 2021. Accessed at: https://idahocapitalsun.com/2021/05/18/idaho-overpaid-road-builders/

Efron, Bradley. 1992. "Bootstrap methods: another look at the jackknife." In Breakthroughs in statistics, pp. 569-593. Springer, New York, NY, 1992. https://doi.org/10.1007/978-1-4612-4380-9_41

El-Basyouny, M., and M. Witczak. 2004. “Calibration of the Alligator Fatigue Cracking Model for the 2002 Design Guide.” Dissertation. University of Arizona, May 2004.

Ferreira, V. A., G. A. Weesies, D. C. Yoder, G. R. Foster, and K. G.Renard. 1995. “The site and condition specific nature of sensitivity analysis.” J. Soil & Water Conserv. 1995; 50(5): 493-497.

Fligner, Michael A., and Timothy J. Killeen. 1976. "Distribution-free two-sample tests for scale." Journal of the American Statistical Association 71, no. 353 (1976): 210-213. https://doi.org/10.1080%2F01621459.1976.10481517

Good, Phillip I. 2006. “Permutation, parametric, and bootstrap tests of hypotheses.” Springer Science & Business Media, 2006.

Graves RC, Mahboub KC. 2006. “Pilot study in sampling-based sensitivity analysis of NCHRP design guide for flexible pavements.” Transportation research record. 2006;1947(1):122-35.

Gunduz, Murat, and Oytun Önder. 2013. "Corruption and internal fraud in the Turkish construction industry." Science and Engineering Ethics 19, no. 2 (2013): 505-528. https://doi.org/10.1007/s11948-012-9356-9

Hand, ADAM JT, and J. A. Epps. 2006. "Fundamentals of Percent within Limits and Quality Control–Quality Assurance Compaction Specifications." Transportation Research E-Circular No. EC105 (2006): 140-162.

Hand, Adam JT, Mohamed A. Nimeri, Elie Y. Hajj, Peter E. Sebaaly, Randy C. West, Michael A. Heitzman, Fan Yin, Charles S. Hughes, and Shiraz Tayabji. 2020. “Procedures and Guidelines for Validating Contractor Test Data.” No. Project 10-100. 2020.

Harvey, I. 2019. “Forensic investigation into $80M Bondfield fraud digs deeper.” Daily Commercial News by ConstructConnect. Retrieved on August 1, 2020. Accessed at: https://canada.constructconnect.com/dcn/news/government/2019/12/forensic-investigation-into-80m-bondfield-fraud-digs-deeper

Hewitt, D. 2015. “China To Demolish New Skyscraper As Part Of Anti-Corruption Campaign.” International Business Times. Retrieved on August 1, 2020. Accessed at: https://www.ibtimes.com/china-demolish-new-skyscraper-part-anti-corruption-campaign-2225814

Hughes, Charles S. 2005. “State construction quality assurance programs.” Vol. 346. Transportation Research Board, 2005.

Hughes Ch., Hand A., Aschenbrener T. 2020. “Asphalt Materials Quality Assurance Practices.” Federal Highway Administration Report No. FHWA-HIF-20-033, 2020. Accessed at: https://scholarworks.unr.edu/handle/11714/7535

ITD. 2019. “FY20 to FY26: Idaho Transportation Investment Program (ITIP).” Retrieved on August 1, 2020. Accessed at: https://apps.itd.idaho.gov/Apps/Fund/itip2020/draft/FY20_Draft_itip.pdf

Jain, Arushi, and Sarvesh Shinde. 2019. "A Comprehensive Study of Data Mining-based Financial Fraud Detection Research." In 2019 IEEE 5th International Conference for Convergence in Technology (I2CT), pp. 1-4. IEEE, 2019. https://doi.org/10.1109/I2CT45611.2019.9033767

Kahler, D. L. 2012. “Why Statistical Validation of Materials Can Be Misleading.” 2012. https://www.researchgate.net/publication/271964767_Why_Statistical_Validation_Of_Materials_Can_Be_Misleading

Kyriacou, Andreas P., Leonel Muinelo-Gallo, and Oriol Roca-Sagalés. 2015. "Construction corrupts: empirical evidence from a panel of 42 countries." Public Choice 165, no. 1-2 (2015): 123-145.

LaVassar, Colin J., Joe P. Mahoney, Kim A. Willoughby, and Kim Willoughby. 2009. "Statistical assessment of quality assurance-quality control data for hot mix asphalt." Report No. WA-RD 686 (2009).

Le, Yun, Ming Shan, Albert PC Chan, and Yi Hu. 2014. "Investigating the causal relationships between causes of and vulnerabilities to corruption in the Chinese public construction sector." Journal of

Construction Engineering and Management 140, no. 9 (2014): 05014007. https://doi.org/10.1061/(ASCE)CO.1943-7862.0000886

Levene, Howard. 1960. "Contributions to probability and statistics." Essays in honor of Harold Hotelling (1960): 278-292.

Locatelli, Giorgio, Giacomo Mariani, Tristano Sainati, and Marco Greco. 2017. "Corruption in public projects and megaprojects: There is an elephant in the room!." International Journal of Project Management 35, no. 3 (2017): 252-268. https://doi.org/10.1016/j.ijproman.2016.09.010

Manning, R. 2019. “Portland Public Schools Respond To Fraudulent $3 Million Construction Payment.” Oregon Public Broadcasting. Retrieved on August 1, 2020. Accessed at: https://www.opb.org/news/article/portland-public-schools-fraud-3-million-construction-payment/

Markelevich, Ariel, and Rebecca L. Rosner. 2013. "Auditor fees and fraud firms." Contemporary Accounting Research 30, no. 4 (2013): 1590-1625.

Moghaddam, Taher Baghaee, Mehrtash Soltani, Mohamed Rehan Karim, and Hassan Baaj. 2015. "Optimization of asphalt and modifier contents for polyethylene terephthalate modified asphalt mixtures using response surface methodology." Measurement 74 (2015): 159-169.

Moser, Barry K., and Gary R. Stevens. 1992. "Homogeneity of variance in the two-sample means test." The American Statistician 46, no. 1 (1992): 19-21. https://doi.org/10.1080/00031305.1992.10475839

Muench, Stephen T., Joe P. Mahoney, and Jim Walter. 2001. “A Quantification and Evaluation of WSDOT's Hot Mix Asphalt Concrete Statistical Acceptance Specification.” No. WA-RD 517.1,. Washington State Department of Transportation, 2001.

NAPA fast facts. 2020. https://www.asphaltpavement.org/uploads/documents/NAPA_Fast_Facts_Sept_2020_Final_Version.pdf

NCAT. 2016. “National Center for Asphalt Technology. Is Your Gsb Correct?” 2016. Accessed at: http://www.eng.auburn.edu/research/centers/ncat/newsroom/2017-spring/gsb.html

Newcomb, D., H. Al-Khayat, C. Gurganus, Maryam Sakhaeifar, and J. Epps. 2017. "Review of Oregon Depatment of Trasportation Asphalt Mix Specification, Phase II." (2017).

Ngai, Eric WT, Yong Hu, Yiu Hing Wong, Yijun Chen, and Xin Sun. 2011. "The application of data mining techniques in financial fraud detection: A classification framework and an academic review of literature." Decision support systems 50, no. 3 (2011): 559-569. https://doi.org/10.1016/j.dss.2010.08.006

Nordin, Rumaizah Mohd, Roshana Takim, and Abdul Hadi Nawawi. 2013. "Behavioural factors of corruption in the construction industry." Procedia-Social and Behavioral Sciences 105 (2013): 64-74. https://doi.org/10.1016/j.sbspro.2013.11.008

Office of the Federal Register, and Office of the Federal Register (US) Staff. 2007. “Code of federal regulations.” Vol. 5. Office of the Federal Register (2007).

Olkin, Ingram. 1960. “Contributions to probability and statistics: essays in honor of Harold Hotelling.” Stanford University Press, 1960.

Perols, Johan. 2011. "Financial statement fraud detection: An analysis of statistical and machine learning algorithms." Auditing: A Journal of Practice & Theory 30, no. 2 (2011): 19-50.

Robertson, L. 2014. “Scale of international bribery laid bare by new OECD report.” Organisation for Economic Co-operation and Development. Retrieved on August 1, 2020. Accessed at: https://www.oecd.org/daf/anti-bribery/scale-of-international-bribery-laid-bare-by-new-oecd-report.htm

Rodriguez, D., Waite, G., & Wolfe, T. 2005. “Global corruption report 2005. Transportation International.” https://images.transparencycdn.org/images/2005_GCR_Construction_EN.pdf

Ruxton, Graeme D. 2006. "The unequal variance t-test is an underused alternative to Student's t-test and the Mann–Whitney U test." Behavioral Ecology 17, no. 4 (2006): 688-690. https://doi.org/10.1093/beheco/ark016

Satterthwaite, Franklin E. 1946. "An approximate distribution of estimates of variance components." Biometrics bulletin 2, no. 6 (1946): 110-114. https://doi.org/10.2307/3002019

Sreedhar Sh., Coleri E., Sogol S. H., Ihsan A.O. 2019. “Developing Performance-Based Specifications to Improve the Fatigue Life of Asphalt Pavements in Oregon.” DISSERTATION. 2019. Accessed at: https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=&ved=2ahUKEwjKsdC2sJbsAhVYnp4KHWk8BR4QFjAGegQIBRAC&url=https%3A%2F%2Fir.library.oregonstate.edu%2Fdownloads%2Ftm70n1717%3Flocale%3Den&usg=AOvVaw1s0nBx53CvZXMynya2yRim

Sias J.E., Dave E.V., McCarthy L.M. 2020. “Practices for Fabricating Asphalt Specimens for Performance Testing in Laboratories.” Transportation Research Board, no. 552 (2020): 0547-5570. Accessed at: https://trid.trb.org/view/1715599

El Sayed, Mohamed. 2012. "Effect of changing theoretical maximum specific gravity on asphalt mixture design." Engineering Journal 16.4 (2012): 137-148.

Schmitt, Robert L., Awad S. Hanna, Jeffrey S. Russell, and Erik V. Nordheim. 2001. "Statistically based methods for verification testing." Transportation research record 1761, no. 1 (2001): 86-92. https://doi.org/10.3141/1761-11

Schwartz, C.W., Li, R., Ceylan, H., Kim, S. and Gopalakrishnan, K. 2013. “Global Sensitivity Analysis of Mechanistic–Empirical Performance Predictions for Flexible Pavements.” Transportation research record, 2368(1):12-23.

Shaffer, M. J. 1988. “Estimating confidence bands for soil-cropsimulation models.” Soil Sci. Soc. Am. J. 1988; 52(6): 1782-1789.

Sohail, M., and Sue Cavill. 2008. "Accountability to prevent corruption in construction projects." Journal of Construction Engineering and management 134, no. 9 (2008): 729-738.

Srivastava, Abhinav, Amlan Kundu, Shamik Sural, and Arun Majumdar. 2008. "Credit card fraud detection using hidden Markov model." IEEE Transactions on dependable and secure computing 5, no. 1 (2008): 37-48.

Smith, G. 1998. “State DOT Management Techniques for Materials and Construction Acceptance.” NCHRP Synthesis 263, Transportation Research Board, National Academy Press, Washington D.C., 1998.

Transportation Research Board (TRB). 2016. “Optimal Acceptance Standards for Statistical Construction Specifications” https://www.fhwa.dot.gov/publications/research/infrastructure/pavements/pccp/02095/01.cfm

Transportation Research Board (TRB). 2018. “Glossary of Transportation Construction Quality Assurance Terms.” Seventh Edition. Transportation Research Circular. No. E-C235, Washington, DC. http://www.trb.org/Main/Blurbs/178085.aspx

U.S. Department of Transportation, Federal Highway Administration. 2018. “Table HM-12 - Highway Statistics 2017.” https://www.fhwa.dot.gov/policyinformation/statistics/2017/hm12.cfm

Wani, Sujay S., and Nasir G. Gharaibeh. 2013. "Evaluating the Use of Contractor-Performed Test Results in Highway Construction and Material Acceptance Decisions." Journal of infrastructure systems 19, no. 1 (2013): 67-73. https://doi.org/10.1061/(ASCE)IS.1943-555X.0000113

West, Jarrod, Maumita Bhattacharya, and Rafiqul Islam. 2014. "Intelligent financial fraud detection practices: an investigation." In International Conference on Security and Privacy in Communication Networks, pp. 186-203. Springer, Cham, 2014.

Wilcoxon, F. 1945. "Individual comparisons by ranking methods. ", Biom. Bull. 1: 80–83. (1945).

Wilcoxon, Frank. 1992. "Individual comparisons by ranking methods." In Breakthroughs in statistics, pp. 196-202. Springer, New York, NY, 1992. https://doi.org/10.1007/978-1-4612-4380-9_16

Williams, Stacy G. 2008. “Non-nuclear methods for HMA density measurements: final report”, June 2008. No. MBTC 2075. University of Arkansas, Fayetteville. Dept. of Civil Engineering, 2008.

Williams, Stacy G., and Kevin D. Hall. 2008. "Critical Factors Affecting Field Determination of Hot-Mix Asphalt Density Using Nonnuclear Devices." Transportation Research Record 2081, no. 1 (2008): 150-157. https://doi.org/10.3141/2081-16

Zarkada-Fraser, Anna, and Martin Skitmore. 2000. "Decisions with moral content: collusion." Construction Management & Economics 18, no. 1 (2000): 101-111.

Zimmerman, Donald W., and Bruno D. Zumbo. 2009. "Hazards in choosing between pooled and separate-variances t tests." Psicológica 30, no. 2 (2009): 371-390.

Appendix A. 2018 FHWA report

FHWA Forensic Pavement Assessment for the Idaho Transportation Department (ITD)

Report Date: March 8, 2018

Executive Summary of Recommendations:

These recommendations should be acted upon immediately by ITD:

1. ITD should witness sampling from aggregate and RAP stockpiles for their asphalt mixture design verification.
2. ITD should monitor the G_{sb}, G_{se}, G_{sa} and their relationship throughout the asphalt mixture design and production process.
3. ITD should use a tested value of G_{sb} and G_{sa} for RAP aggregate when the asphalt binder replacement from the RAP is greater than 17 percent.
4. ITD should monitor the asphalt content, G_{mm}, and G_{se} throughout the asphalt mixture design and production for consistency.
5. ITD should update their specifications to address the general constructability issues with longitudinal joint density, in-place density, segregation and the use of the appropriate lift thickness-to-nominal maximum aggregate size.
6. ITD and the contractors' technicians should conduct check testing prior to the start of the project.
7. Project testers for ITD and the contractor should participate in a pre-testing meeting prior to the start of construction.
8. F & t-testing should be conducted on the actual test results used to calculate the volumetric properties: G_{mm}, G_{mb}, asphalt content, and G_{se}.
9. A tolerance should be created on the weight of the samples prepared for gyratory compaction.

These recommendations should be acted upon in the near future by ITD:

10. ITD should develop a standard asphalt mixture design submittal format which would include all properties required.
11. A tolerance between corrections factors for different ignition furnaces for the same mix should be developed. The current process for submitting and distributing the correction factor samples should be continued for use. This process should be evaluated in a year or two for effectiveness.
12. ITD should examine the CoreLok procedure used to determine the fine aggregate Gsb and consider replacing it with AASHTO T 84.
13. ITD should include dust proportion as a production pay factor.
14. ITD should consider adding a cracking test to their asphalt mixture requirements for design and acceptance.

15. To maintain independence between quality control tests and verification tests, the state should not share test results with the contractor until the lot has been completed and the contractor has submitted their test results to ITD.

Field review performed: November 6 thru November 9, 2017
Asphalt mixture information gathering and analysis: November 2017 thru January 2018
Meetings with ITD and AGC: February 6 thru February 8, 2018

1. Scope

The objective of this assessment and information sharing was to conduct a forensic assessment of asphalt pavements for the Idaho Transportation Department (ITD). A forensic review of 13 asphalt pavements was conducted. The team conducting the review and the primary individuals consulted during the review are as follows.

FHWA Report and Review Team:
Tim Aschenbrener, Senior Asphalt Pavement Engineer, Office of Asset Management, Pavements, and Construction
Dennis Dvorak, Pavement & Materials Engineer, Resource Center Office Kyle Holman, Operations Engineer, Idaho Division Office

FHWA Division Office Contacts:
Peter Hartman, Division Administrator, Idaho Division Office
Gus Shanine, Assistant Division Administrator, Idaho Division Office
John Perry, Field Operations Team Leader, Idaho Division Office

ITD Contacts:
Kimbol Allen, ITD Chief Engineer
John Bilderback, ITD Construction and Materials Engineer
Jason Brinkman, ITD Engineering Manager
Mike Copeland, ITD Materials
Bob Engelmann, ITD District 3
Travis McGrath, ITD Chief Operations Officer
Laura Meyer, ITD Facilitator
Blake Rindlisbacher, ITD Engineering Services Administrator
Mike Santi, ITD Materials
Mark Wheeler, ITD Pavement Engineer

Other Contacts:
Taj Anderson, Poe Asphalt

Darver Arnold, Knife River
Tim Bentley, Idaho Materials
Chase Camberlango, Western Construction
Scott Cron, Strata
Justin Drye, Knife River
Paul Franz, Old Castle Materials
Chris Hartman, Knife River
Marv Kerbs, Allwest
Pat McEntee, Central Paving
Terry McEntee, Central Paving
Greg Mitchell, Knife River
Ryan Russell, Idaho Materials
Ron Shippy, Old Castle Materials
Josh Smith, Knife River

2. Itinerary

November 6, 2017 –

- Opening meeting with the Division Office in Boise.
- Opening meeting with ITD, contractors and consultants at the ITD Headquarters in Boise.
- Review of video logs at the ITD Headquarters in Boise of US-91 at Yellowstone (ID 5), US-95 at Lewiston Hill (ID 9), US-95 at Worley (ID 2), SH-8 at Dreary (ID 10), and I-86 at the Salt Lake City Interchange (ID 1).

November 7, 2017 –

- Field review of projects in Boise and north of Boise including US 20/26 (ID 11), SH-16 at Emmett (ID 7), and US 95 overlays (ID 8).

November 8, 2017 –

- Field review of projects east of Boise including I-84 at Eisenman (ID 3), US-91 in Pocatello (ID 5), I-86 at the Salt Lake City Interchange (ID 1), I-84 at MP 228 (ID 4).

November 9, 2017 –

- Tour of ITD Central Materials Laboratory in Boise.
- Meeting to summarize findings and next steps with ITD, FHWA, contractors and consultants at the ITD Headquarters in Boise.

February 6, 2018 –

- Opening meeting with ITD and FHWA at the ITD Headquarters in Boise.

February 6 and 7, 2018 –

- Interviews and review of recommendations with ITD, FHWA, contractors and consultants at the ITD Headquarters in Boise.

February 8, 2018 –

- Closeout meeting with ITD and FHWA staff at the ITD Headquarter in Boise.

3. Description of Projects Visited

A total of 13 projects were identified for evaluation. The projects and general information are shown in Table 1. Ten of the projects were reviewed visually. Seven of these were done in person on the field trip and three were done by video only. The field performance ranged from "good" to "moderate." During the review, there were no catastrophic pavement failures observed.

Table 1. Projects Reviewed and General Information.

ID	Year Built	District	Key Number	Project Name	Google Map Link	Contractor	Mix Type
1	2012	4	6521	I-86, SLC IC to Raft River	6521	Knife River	3/4" SP-6
2	2015	1	12212	US-95, Worley North	12212	Interstate	3/4" SP-5
3	2013	3	12352	I-84, Eisenmenn IC to Exit 70	12352	Central Paving	1/2" (3/4") SP-6
4	2014	4	12390	I84, Jct IC228 to IC245 WB (towards SLC)	12390	Western Construction	1/2" SP-6
5	2014	5	12416	US -91, Pocatello (Yellowstone Ave, Alameda	12416	Jack B. Parson Co.	1/2" SP-5

				to Flandro)			
6	2014	6	12467	US20, Rexburg IC #332 to SF Teton River Br	12467	H K Contractors	3/4&1/2" SP-5
7	2012	3	13021	SH-16, Jct SH-44 to Emmett	13021	Central Paving	
8	2014	3	13361	US-95 Overlays & Alpine to Council	13361	Knife River	1/2" SP-3
9	2015	2	13435	US-95, Lewiston Hill	13435	Knife River	1/2" SP-5
10	2016	2	13875	SH -8, Deary to Bovill	13875	Poe Asphalt	1/2" SP-3
11	2016	3	13928	US20 /26 Branstetter to Jct I-184	13928	Idaho Materials and Construction	1/2" SP-3
12	2017	2	19187	US12 Arrow Bridge to Big Canyon Cr Br		Knife River	3/4" SP-6
13	2013	6	11478	US-20, Bellin Road to Yellowstone Hwy		HK Contractors	1/2" SP-5

Legend:

Blue – Projects visited in person

Yellow – Projects with “moderate” level of distress Green – Projects with “good” performance

Unshaded – Field performance not evaluated

The “Mix Type” in Table 1 references the NMAS (3/4” or ½”) and the Superpave mixture requirement that includes the number of design gyrations (e.g., SP-3 is 75 gyrations, SP-5 is 100 gyrations).

During the field review of the pavements, it was identified that a more detailed review of the overall asphalt mixture design process was needed because of some anomalies in the asphalt mixture design and production data that could lead to reduced

performance. The subjective levels of distress observed did not always relate to the anomalies identified.

Additional data and analysis was performed in the following areas:

- Relationship of aggregate specific gravities
- Theoretical maximum specific gravity (G_{mm})
- Mixture design acceptance results
- Ignition furnace correction factors
- Reclaimed Asphalt Pavement (RAP): asphalt content and gradation
- Field adjustments of asphalt mixture design

From November through January, additional information on the contractor's asphalt mixture design, ITD's mixture verification, and the field acceptance results were gathered, analyzed and reviewed. Based on the analysis and review, a series of questions were developed that needed further clarification. ITD and the contractors elected to have the discussion as a group. On February 6 and 7, these questions and the team's observations and draft recommendations were discussed with ITD and the contractors utilizing an ITD provided facilitator. The draft recommendations were modified based on the results of the discussions with the group.

4. Observations and Recommendations

The following observations and recommendations were divided into three groups as described below:

- Group 1 - Reducing Risk of Project Acceptance
- Group 2 – Improving Testing Consistency
- Group 3 – Future Considerations

FHWA's Office of Asset Management, Pavements, and Construction and the FHWA Resource Center are available to provide technical support in addressing the recommendations.

Group 1 – Reducing Risk of Project Acceptance

1. **Observation**. The contractor is responsible for developing the asphalt mixture design that meets ITD requirements. ITD then conducts a verification of the asphalt mixture design in the laboratory. The contractor provides aggregate and RAP samples to ITD. ITD does not have a chain of custody of

these samples to ensure that they have been sampled from the stockpiles that will be used for the project.

Recommendation. ITD should witness sampling from aggregate and RAP stockpiles for their asphalt mixture design verification.

2. **Observation**. The test results for the aggregate bulk specific gravity (G_{sb}), aggregate effective specific gravity (G_{se}), and aggregate apparent specific gravity (G_{sa}) are important in calculating the volumetric properties of the asphalt mixture. There is a unique relationship with these specific gravities such that $G_{sb} < G_{se} < G_{sa}$. This is a fundamental truth based on the math of how these variables are calculated; such that

   when the results violate this relationship it means there must have been an error or use of incorrect data somewhere in a test result or calculation. In three of the 13 projects reviewed this relationship was violated, the G_{se} was greater than the G_{sa}. An example is shown in Figure 1.

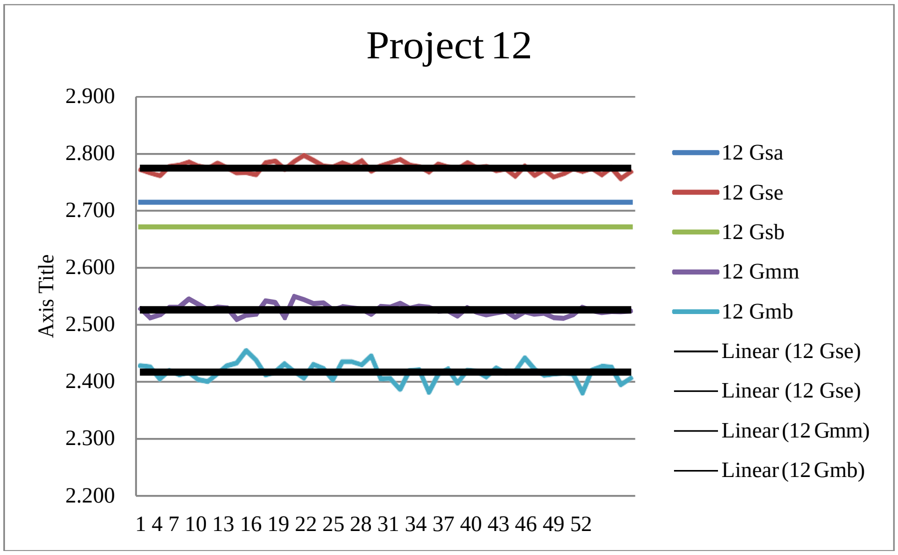


Figure 1. Unexpected relationship of G_{sb}, G_{se}, and G_{sa}.

Recommendation. ITD should monitor the G_{sb}, G_{se}, G_{sa} and their relationship throughout the asphalt mixture design and production process to ensure that they do not violate the rule that $G_{sb} < G_{se} < G_{sa}$.

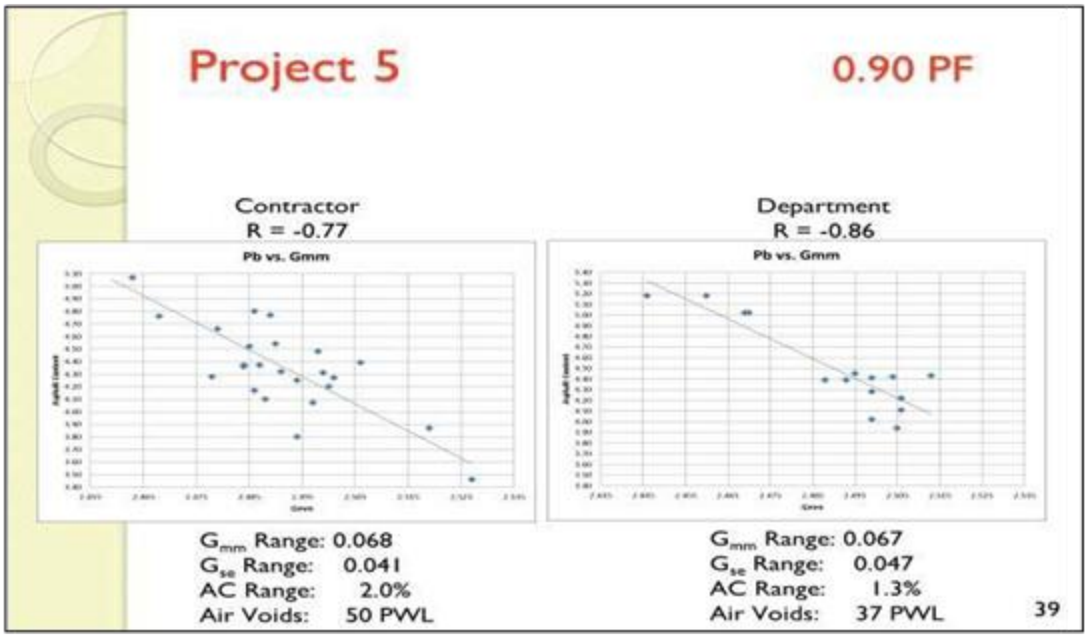
3. **Observation**. When RAP is used in an asphalt mixture, ITD estimates the G_{sa} of the RAP. The G_{se} of the RAP is measured through testing and then the G_{se} is adjusted by a portion of the absorption to estimate the G_{sa} of the RAP. Through the interview process with ITD and contractors, it was desired that G_{sa} be measured instead of estimated.

Recommendation. ITD should use a tested value (e.g., using aggregate from the ignition furnace) of G_{sb} and G_{sa} of RAP aggregate when the asphalt binder replacement from the RAP is greater than 17 percent. The 17 percent was based on current thresholds in the ITD specification. When there is less than 17 percent asphalt binder replacement from RAP in the asphalt mixture design, the estimated G_{sb} and G_{sa} should be acceptable.

4. **Observation**. The asphalt content is measured with the ignition furnace. There should be a linear relationship between the asphalt content and G_{mm} which are also related to the G_{se}. The test procedures for both the asphalt content and G_{mm} are relatively quick, simple and repeatable. A strong linear relationship was observed on 9 of the 13 projects. An example is shown in Figure 2.

   On 4 of the 13 projects there was an unexpected amount of scatter in the results from both the contractor and ITD. An example is shown in Figure 3.

   Recommendation. ITD should monitor the asphalt content, G_{mm} and G_{se} throughout the asphalt mixture design and production for consistency. A tolerance should be applied to determine if the results are acceptable or if the materials have changed enough to require a new asphalt mixture design.



Project 5

0.90 PF

Contractor
R = -0.77

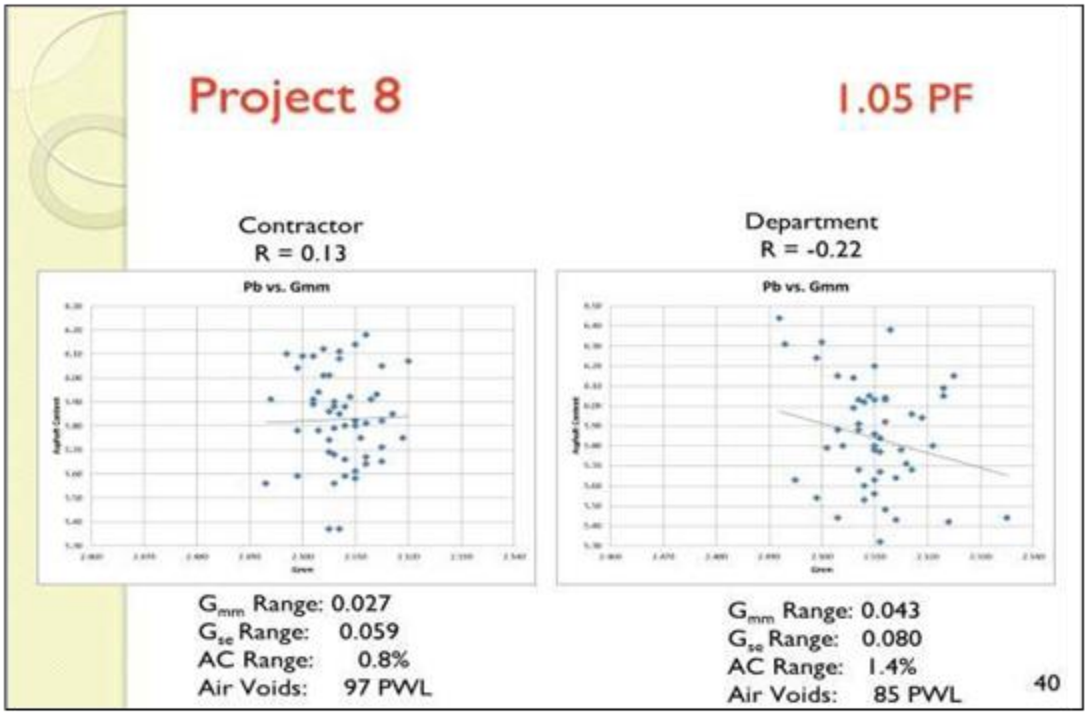
Department
R = -0.86

G_{mm} Range: 0.068
G_{se} Range: 0.041
AC Range: 2.0%
Air Voids: 50 PWL

G_{mm} Range: 0.067
G_{se} Range: 0.047
AC Range: 1.3%
Air Voids: 37 PWL

39

Figure 2. Expected Relationship of Asphalt Content and G_{mm}.



Project 8

1.05 PF

Contractor
R = 0.13

Department
R = -0.22

G_{mm} Range: 0.027
G_{se} Range: 0.059
AC Range: 0.8%
Air Voids: 97 PWL

G_{mm} Range: 0.043
G_{se} Range: 0.080
AC Range: 1.4%
Air Voids: 85 PWL

40

Figure 3. Unexpected Relationship of Asphalt Content and G_{mm}.

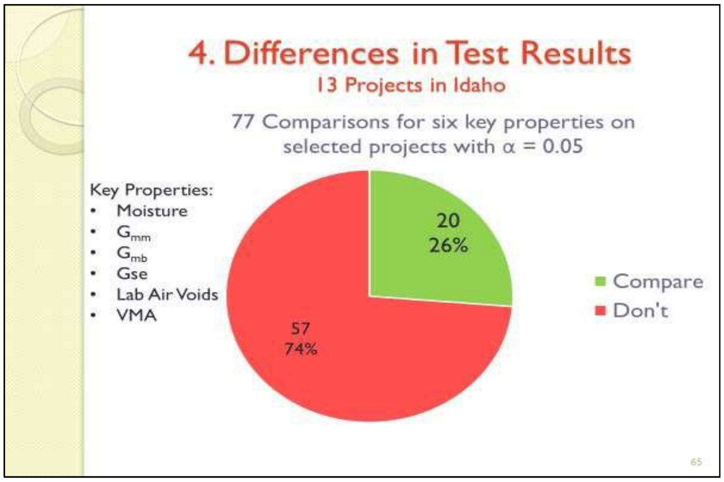
5. **Observation**. During the field review of projects in November, some general areas for improvement related to constructability were identified. Some of the pavement distresses were related to these constructability issues which included: longitudinal joint density, in-place density, segregation and the use of the appropriate lift thickness- to-nominal maximum aggregate size.

   Recommendation. ITD should update their specifications to address the general constructability issues with longitudinal joint density, in-place density, segregation and the use of the appropriate lift thickness-to-nominal maximum aggregate size.

Group 2 – Improving Testing Consistency

6. **Observation**. For the 13 projects reviewed, six key properties of the component aggregates and combined asphalt mixture were identified: moisture content, G_{mm}, G_{mb} (bulk specific gravity of the mixture), G_{se}, air voids, and voids in the mineral aggregate (VMA). For these six properties, 77 comparisons were made between ITD and the contractor's results using the statistical F-test (to compare sample variances) and t- test (to compare sample means). Seventy-four percent (57 out of 77) of the comparisons failed either the F-test, the t-test or both as shown in Figure 4.

ITD’s Quality Assurance Special Provision (QASP) that describes ITD’s statistical (F&t) based quality assurance process does not comply with 23 CFR 637.207(a)(1)(ii)(B). This is due to the QASP allowing contractor quality control tests that have not been validated which shows that they are from the same population of test results to be used in the acceptance decision and pay factor calculation. This would typically result in a recommendation, however ITD has implemented the necessary steps to improve their statistical based acceptance of contractor’s quality control tests in their QASP version 1.1 that has been incorporated into 2018 ITD contracts. ITD should continue to monitor the QASP v1.1.



4. Differences in Test Results

13 Projects in Idaho

77 Comparisons for six key properties on selected projects with $\alpha = 0.05$

Key Properties:
- Moisture
- G_{mm}
- G_{mb}
- Gse
- Lab Air Voids
- VMA

Figure 4. Results from 77 Comparisons Between Results from ITD and Contractors for Six Key Properties.

Recommendation. In an effort to make improvements to the materials testing process and achieve consistent results (as measured by passing the F & t-tests), ITD and the contractors’ technicians should conduct check testing prior to the start of the project.

The six key properties identified to have different results should be checked: moisture content, G_{mm}, G_{mb}, G_{se}, air voids, and VMA. Also, gradation and asphalt content should be checked since they are important factors in measuring the six key properties.

7. **Observation**. During the discussions with ITD and the Contractors on February 6 and 7, it was clearly communicated that there was a need for improving test reproducibility. Some of the discussions related to the test procedures, such as the method for developing the correction factor for the ignition furnace. Some of the discussions related to the processes prior to the start of the testing procedures, such as identification of the sampling location and curing times.

Recommendation. Project testers for ITD and the contractor should participate in a pre-testing meeting prior to the start of construction. An agenda should be developed for statewide use which includes the important aspects that would influence testing results. This meeting is a separate meeting from the pre-construction meeting and will typically follow the preconstruction meeting.

8. **Observation**. Currently, F & t-testing is done with volumetric properties of air voids and VMA. These properties are calculated from several underlying test results which can have various impacts on the calculated volumetric properties.

   Recommendation. F & t-testing should be conducted on the actual test results used to calculate the volumetric properties. This should include: G_{mm}, G_{mb}, asphalt content, and G_{se}. This will allow for more efficient trouble shooting.

9. **Observation.** There is a relationship between the G_{mb} and the height of the gyratory compacted sample. Also, the height and weight of each sample are closely correlated. For each project, a comparison of the G_{mb} and height was made for the first compacted sample as well as the second compacted sample. Based on a typical distribution of data, the same trend would be expected for both the first and second compacted sample. The correlation was virtually the same for the first and second compacted sample on all 13 projects with ITD data. With the contractor's data, two of the 13 projects had unexpected results. There was a desire to improve testing consistency.

   Recommendation. A tolerance should be created on the weight of the samples prepared for gyratory compaction. The tolerance should vary based on the nominal maximum aggregate size.

Group 3 – Future Considerations

10. **Observation**. When analyzing the asphalt mixture designs, it was not always possible to find the necessary information. Each contractor's asphalt mixture design submittal used a different format and contained different information.

    Recommendation. ITD should develop a standard asphalt mixture design submittal format which would include all properties required. This would assist the contractors with the submittals and assist ITD with their review and future analysis. An electronic format, such as an Excel spreadsheet, would be most effective.

11. **Observation**. A correction factor is developed for each ignition furnace. Contractors prepare the loose asphalt mixture and deliver 16 samples to ITD. ITD then distributes samples to the contractor, the district laboratory and the central laboratory for use in developing the correction factor for each of the ignition furnaces.

    When comparing correction factors from different ignition furnaces for the same asphalt mixture, there were often similarities. In three of the 13 projects the difference between the ignition furnace correction factors for ITD and the contractor were greater than 0.25, which is a significant difference. These three projects also had significant differences between other ITD and contractor test results. However, these projects did not use the current process utilizing the 16 samples delivered to ITD.

Recommendation. A tolerance between correction factors for different ignition furnaces for the same mix should be developed to determine if the correction factors are acceptable to address the anomalies identified. The current process for submitting and distributing the correction factor samples should be continued for use. This process should be evaluated in a year or two for effectiveness. The desired process is for ITD to have full control, such that each lab should mix their own correction factor samples from the component asphalt binder and aggregate materials.

12. **Observation**. When determining the aggregate bulk specific gravity, ITD uses AASHTO T 85 for the coarse aggregate. For the fine aggregate, ITD has created their own test procedure using the CoreLok device. The CoreLok device provides higher G_{sb} than AASHTO T 84 which would result in artificially high VMA results.

    Recommendation. ITD should examine the CoreLok procedure used to determine the fine aggregate G_{sb} and consider replacing it with AASHTO T 84. Excellent guidance is provided in the following two research reports:
    - NCHRP Report 805, Improved Test Methods for Specific Gravity and Absorption of Course and Fine Aggregate, confirmed AASHTO T 84 and 85 to be the best available, but recommended incremental improvements to the test.
    - FHWA Tech Brief, Review of Aggregate and Asphalt Mixture Specific Gravity Measurements and Their Impacts on Asphalt Mix Design Properties and Mix Acceptance, supported AASHTO T 84 and 85 as other methods provide different results.

13. **Observation**. The dust proportion can be a maximum of 1.6 for coarse gradations in ITD Specifications. The dust proportion is only a requirement for the asphalt mixture design. High dust proportion in the asphalt mixture design along with increases in dust proportion during production can negatively impact pavement performance. In three of the 13 projects, the dust proportion had significant increases from the mixture design to production as shown in Table 2.

Table 2. Increases in Dust Proportion from Asphalt Mixture Design to Production.

Project	Design	Production
2	1.38	1.61
4	1.25	1.42
8	1.22	1.55

Recommendation. ITD should include dust proportion as a production pay factor.

14. **Observation**. Volumetric properties are used for the asphalt mixture design and field acceptance. Volumetric properties work well, but can be limited in effectiveness with all the new materials, binder additives, and recycled materials increasingly being used in today's asphalt mixtures. In Idaho, contractors often submit asphalt mixture designs with high quantities of RAP. Four of the 13 projects had 50 percent and 8 of the 13 projects had 30 percent RAP by asphalt binder replacement. Volumetric properties alone may not be adequate to predict performance.

    Recommendation. ITD should consider adding a cracking test to their asphalt mixture requirements for design and acceptance. This would be a better indicator of performance for high-RAP mixtures. It could also be an indicator of changes in the RAP content.

15. **Observation**. ITD has received requests from contractors to provide the ITD's verification test results within 6 hours of completion and no later than the next morning before the contractor begins paving for the day. This request has been made by contractors so they know how to adjust their operation based on the previous day's test results. Contractors control their quality control tests and have those results immediately for the purpose of controlling their operations.

    Recommendations. To maintain independence between quality control tests and verification tests, the state should not share test results with the contractor until the lot has been completed and the contractor has submitted their test results to ITD. Typically, DOT's using similar processes provide test results to the contractor a minimum of 1 to 2 working days of lot completion.

Appendix B. Federal Highway Administration Memorandum on Electronic Security Issues

U.S. Department of
Transportation
Federal Highway Administration

MEMORANDUM

Subject: **ACTION**: Electronic Security Issues - Kansas Department of Transportation (KDOT)
Construction Management System (CMS)

From: Chief, Construction and Maintenance Division
Office of Engineering

To: Mr. Volmer K. Jensen
Regional Federal Highway Administrator (HRA-07)
Kansas City, MO

Mr. Eric B. White's June 8 memorandum requesting advice has been forwarded to the Construction and Maintenance Division for a response. Your request pertains to the Kansas Division's review and approval of KDOT's CMS. After a review of the attached materials we offer the following comments.

The KDOT appears to be basically establishing a computerized construction project information management system. While this type of system is currently being established in several states across the country, the distinction in Kansas is that they propose to take their system one step further and go totally paperless with the use of electronic signature technology.

Any computerized project record keeping system must meet certain criteria to ensure that the legal and financial interests of the Federal Government are protected. Such a system must be established so that the collection and retention of construction records are acceptable from an engineering, audit, and legal standpoint. These requirements should be no more stringent than they were for the hard copy system that the computerized system is replacing. In either case, the records must provide for the reconstruction of the chain of events that occurs on a project.

It appears, from the documentation provided by the Division Office's report, that the proposed KDOT CMS is acceptable from an engineering standpoint. It is replacing a paper system, maintaining essentially the same structure and audit trail that was acceptable and logical.

However, the KDOT's CMS, from an auditing and legal standpoint, appears highly suspect. The Division's report notes that within KDOT's various programs in the CMS there is an inability to verify who makes actual approvals via electronic signature. This is comparable to not being able to determine who signed the paper version of an approval letter or not being able to verify a persons handwriting with all the resulting legal implications.

The GAO opinion that the Division referred to as only applying to electronic contracting was only one of two decisions that established the safeguards for electronic signatures. In the other decision (B-238449), the Comptroller focused on the general use of an electronic signature rather than its specific use in contractual obligations. Both opinions were rather clear that the use of electronic signature technology had to be unique to the signer, under the control of the signer, have the capacity to be verified, and be a system of acceptable integrity. It appears that the KDOT's CMS does not meet this established criteria. A copy of both GAO decisions is attached for your information.

What is important in the use of electronic signature technology, is not so much the technology itself (i.e., hardware/software), but how the technology is used and more importantly how that use is controlled. These are the things one takes for granted in the uniqueness of a handwritten signature or more simply in a person's handwritten documentation. The computerized versions cannot be compromised any less, whether the documentation is created for contractual obligations directly or indirectly. If the documentation is something that is normally signed or is a record of information that needs to be identified with a specific user (inspector/project engineer, etc.), it is "official documentation" that must be supportable in a court of law.

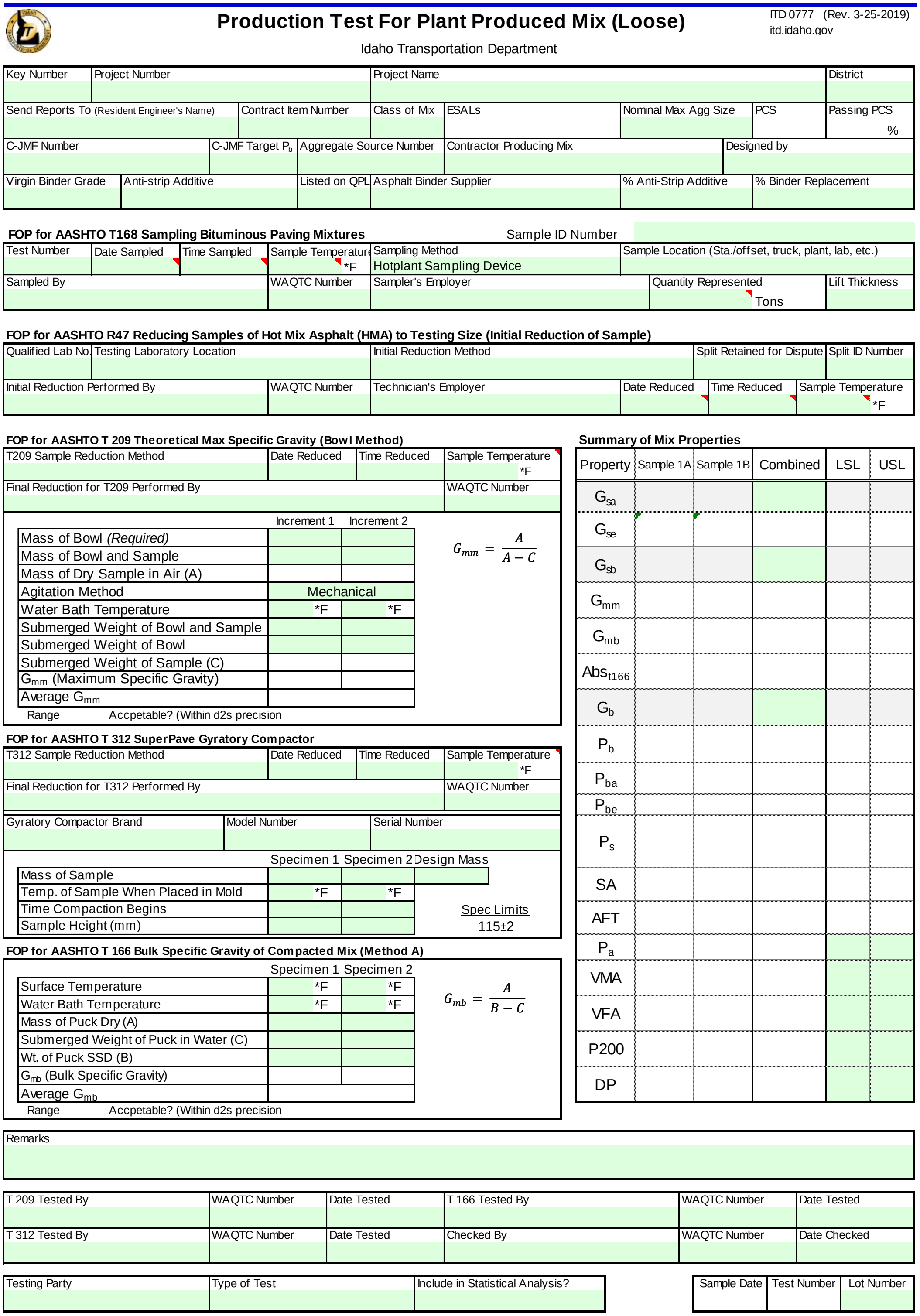
The solution to the problem, as we see it, is relatively simple, implement "tighter computer security. This means control of user access in terms of ID's and passwords that are unique to approved users. Specifically, ID's and passwords should not be shared among a group of users, and the security codes must be changed periodically. Security of such computerized information systems cannot be compromised for the sake of convenience. We recommend the necessary security changes be implemented immediately.

The other important aspect of computer security is the adequacy of data and program system backups. The KDOT appears to have adequately addressed this issue with the procedures and routines required to backup their proposed system.

We trust that the above guidance satisfactorily addresses your concerns. If you have further questions or concerns please contact Mr. Robert S. Wright of my staff at (202) 366-1558.

/s/ original signed by
William A. Weseman
Attachment

Appendix C. Idaho Transportation Department Form 0777



Production Test For Plant Produced Mix (Loose)

Idaho Transportation Department

ITD 0777 (Rev. 3-25-2019)
itd.idaho.gov

Key Number	Project Number	Project Name					District
Send Reports To (Resident Engineer's Name)	Contract Item Number	Class of Mix	ESALs	Nominal Max Agg Size	PCS	Passing PCS %	
C-JMF Number	C-JMF Target P_b	Aggregate Source Number	Contractor Producing Mix	Designed by			
Virgin Binder Grade	Anti-strip Additive	Listed on QPL	Asphalt Binder Supplier	% Anti-Strip Additive	% Binder Replacement		

FOP for AASHTO T168 Sampling Bituminous Paving Mixtures Sample ID Number

Test Number	Date Sampled	Time Sampled	Sample Temperature *F	Sampling Method Hotplant Sampling Device	Sample Location (Sta./offset, truck, plant, lab, etc.)	
Sampled By	WAQTC Number	Sampler's Employer	Quantity Represented Tons	Lift Thickness		

FOP for AASHTO R47 Reducing Samples of Hot Mix Asphalt (HMA) to Testing Size (Initial Reduction of Sample)

Qualified Lab No.	Testing Laboratory Location	Initial Reduction Method	Split Retained for Dispute	Split ID Number	
Initial Reduction Performed By	WAQTC Number	Technician's Employer	Date Reduced	Time Reduced	Sample Temperature *F

FOP for AASHTO T 209 Theoretical Max Specific Gravity (Bowl Method)

T209 Sample Reduction Method	Date Reduced	Time Reduced	Sample Temperature *F
Final Reduction for T209 Performed By			WAQTC Number

	Increment 1	Increment 2
Mass of Bowl *(Required)*		
Mass of Bowl and Sample		
Mass of Dry Sample in Air (A)		
Agitation Method	Mechanical	
Water Bath Temperature	*F	*F
Submerged Weight of Bowl and Sample		
Submerged Weight of Bowl		
Submerged Weight of Sample (C)		
G_{mm} (Maximum Specific Gravity)		
Average G_{mm}		

$$G_{mm} = \frac{A}{A-C}$$

Range Accpetable? (Within d2s precision

FOP for AASHTO T 312 SuperPave Gyratory Compactor

T312 Sample Reduction Method	Date Reduced	Time Reduced	Sample Temperature *F
Final Reduction for T312 Performed By			WAQTC Number
Gyratory Compactor Brand	Model Number	Serial Number	

	Specimen 1	Specimen 2	Design Mass
Mass of Sample			
Temp. of Sample When Placed in Mold	*F	*F	
Time Compaction Begins			
Sample Height (mm)			

Spec Limits
115±2

FOP for AASHTO T 166 Bulk Specific Gravity of Compacted Mix (Method A)

	Specimen 1	Specimen 2
Surface Temperature	*F	*F
Water Bath Temperature	*F	*F
Mass of Puck Dry (A)		
Submerged Weight of Puck in Water (C)		
Wt. of Puck SSD (B)		
G_{mb} (Bulk Specific Gravity)		
Average G_{mb}		

$$G_{mb} = \frac{A}{B-C}$$

Range Accpetable? (Within d2s precision

Summary of Mix Properties

Property	Sample 1A	Sample 1B	Combined	LSL	USL
G_{sa}					
G_{se}					
G_{sb}					
G_{mm}					
G_{mb}					
Abs_{t166}					
G_b					
P_b					
P_{ba}					
P_{be}					
P_s					
SA					
AFT					
P_a					
VMA					
VFA					
P200					
DP					

Remarks

T 209 Tested By	WAQTC Number	Date Tested	T 166 Tested By	WAQTC Number	Date Tested
T 312 Tested By	WAQTC Number	Date Tested	Checked By	WAQTC Number	Date Checked

Testing Party	Type of Test	Include in Statistical Analysis?

Sample Date	Test Number	Lot Number

Production Test For Plant Produced Mix (Loose)

Idaho Transportation Department

ITD 0777 (Rev. 3-25-2019)
itd.idaho.gov

Key Number	Project Number	Project Name				District
Send Reports To (Resident Engineer's Name)	Contract Item Number	Class of Mix	ESALs	Nominal Max Agg Size	PCS	Passing PCS %
C-JMF Number	C-JMF Target P_b	Aggregate Source Number	Contractor Producing Mix	Designed by		
Virgin Binder Grade	Anti-strip Additive	Listed on QPL	Asphalt Binder Supplier	% Anti-Strip Additive	% Binder Replacement	

FOP for AASHTO T168 Sampling Bituminous Paving Mixtures Sample ID Number (for lab use)

Test Number	Date Sampled	Time Sampled	Sample Temperatur *F	Sampling Method Hotplant Sampling Device	Sample Location (Sta./offset, truck, plant, lab, etc.)
Sampled By		WAQTC Number	Sampler's Employer	Quantity Represented Tons	Lift Thickness

FOP for AASHTO R47 Reducing Samples of Hot Mix Asphalt (HMA) to Testing Size (Initial Reduction of Sample)

Qualified Lab No.	Testing Laboratory Location	Initial Reduction Method		Split Retained for Dispute	Split ID Number
Initial Reduction Performed By	WAQTC Number	Technician's Employer	Date Reduced	Time Reduced	Sample Temperature *F

FOP for AASHTO T308 Asphalt Content by Ignition Method

Furnace Serial Number	Lift Test Result grams	Date of Last Air Flow Check	Date of Last Internal Balance Chec	Checks Performed By	WAQTC Number
T308 Sample Reduction Method	Date Reduced	Time Reduced	Sample Temperatur *F	Performed By	WAQTC Number

Data From External Balance	
Mass Basket Assembly	
Mass Basket Assembly & Initial Sample	
Mass Initial Sample (M_i)	
Mass Basket Assembly & Final Aggregat	
Mass Residual Aggregate (M_f)	
Ignition Furnace Correction Factor (C_f)	
Moisture Content from AASHTO T 329 (M	
P_b (Method B)	

Data from Ignition Furnace Printed Ticket (Attach Ticket)	
Elapsed Time	
Temperature Compensation Factor	
Chamber Set Point	*C
Calibration Factor	
Uncorrected Binder Content	
Ignition Furnace Correction Factor (C_f)	
Moisture Content from AASHTO T 329 (MC)	
P_b (Method A)	

$$P_b = \frac{M_i - M_f}{M_i} \times 100 - MC - C_f$$

P_b =

Difference between Method A and Method B P_b within 0.1

FOP for AASHTO T329 Moisture Content of Bituminous Mixes

T329 Sample Reduction Method	Date Reduced	Time Reduced	Sample Temperatur *F	Performed By	WAQTC Number

Moisture Content Calculations	
Oven Temperature	*F
Pan Mass	
Mass Pan and Inital Sample	
Mass Initial Sample (M_i)	
Mass Pan and Dry Sample	
Final Dry Mass (M_f)	
Moisture Content (MC)	

Constant Mass Calculations

Drying Cycle	Mass Pan and Sample	Mass Sample	%Change
M_i			
M_{p1}			
M_{p2}			
M_{p3}			
M_{p4}			
M_{p5}			

$$\%Change = \frac{M_p - M_n}{M_p} \times 100$$

$$MC = \frac{M_i - M_f}{M_i} \times 100$$

FOP for AASHTO T30 Sieve Analysis

Mass Dry Sample Before Wash M_{T30} Mass Dry Sample After Wash M_w

	Cumulative Mass Retained **CMR**	Cumulative % **CPR**	Calculated % Passing **PP**	Aggregate Correction **$C_{f(Agg)}$**	Reported % Passing **RPP**	Spec Limits **LSL**	Spec Limits **USL**
2" (50mm) Sieve							
1 1/2" (37.5mm) Sieve							
1" (25mm) Sieve							
3/4" (19mm) Sieve							
1/2" (12.5mm) Sieve							
3/8" (9.5mm) Sieve							
No. 4 (4.75mm) Sieve							
No. 8 (2.36mm) Sieve							
No. 16 (1.18mm) Sieve							
No. 30 (0.600mm) Sieve							
No. 50 (0.300mm) Sieve							
No. 100 (0.150mm) Sieve							
No. 200 (0.075mm) Sieve							
Pan (-0.075mm) Sieve		**M_s**	ggregate Surface Are **(SA)**				

$$Mass\ Verification = \frac{M_{f\,(T308)} - M_{(T30)}}{M_{f\,(t308)}} = \quad \%$$

$$CPR = \frac{CMR}{M_{T30}} \times 100$$

$$PP = 100 - CPR$$

$$RPP = PP + C_{f(Agg)}$$

$$Check\ Sum = \frac{M_w - M_s}{M_w} \times 100 = \quad \%$$

Remarks

T 308 Tested By	WAQTC Number	Date	T 329 Tested By	WAQTC Number	Date
T 30 Tested By	WAQTC Number	Date	Checked By	WAQTC Number	Date

Testing Party	Type of Test	Include in Statistical Analysis?

Sample Date	Test Number	Lot Number

Appendix D. Survey Questionnaire

Questionnaire

Study Information Sheet and Consent Form

Title: Analyzing possible discrepancies between laboratory-reported mix design results and the actual material used in asphalt pavement construction

The study is being conducted by Dr. Deb Mishra, associate professor of Civil Engineering at Oklahoma State University and Dr. Mojtaba Sadegh, assistant professor of Civil Engineering at Boise State University. In coordination with Idaho Transportation Department (ITD), we are trying to identify prevalence of cases where discrepancies are observed between laboratory reported asphalt mix design results and the actual materials used in construction.

The purpose of this study is strictly for research purposes. During the study, you will answer some survey questions and read a brief text. This survey should take you less than 15 minutes to complete. You must be at least 18 years old to complete this survey.

This study involves no foreseeable risks. You may discontinue the study at any time.

For this research project, we are requesting demographic information. The researchers will make every effort to protect your confidentiality. However, if you are uncomfortable answering any of these questions, you may leave them blank.

CONTACTS FOR QUESTIONS OR PROBLEMS

Contact Information: If you should have any questions about this research study, please contact
1. Deb Mishra at 405-744-3332 or deb.mishra@okstate.edu, or
2. Mojtaba Sadegh at 208-426-3774 or mojtabasadegh@boisestate.edu.

In consideration of all of the above, I give my consent to participate in this research study.
In consideration of all of the above, I DO NOT consent to participate in this research study.
[EXIT IF NO CONSEN]

1. What state department of transportation do you work for?
   Drop down box with all the states…

2. How long have you worked for the state department of transportation in years? [text entry, numeric]

3. What is your role at the department of transportation? [dropdown list of relevant roles]
   - I. Engineer manager
   - II. Field Staff
   - III. State materials engineer
   - IV. District / Division materials engineer
   - V. Project engineer
   - VI. Other. Please specify__________________

4. Do you have experience working with asphalt pavements?
   - I. Yes
   - II. No – You can exit or continue

5. Does your state department of transportation (DOT) implement any procedure to compare actual service life (before major rehabilitation efforts need to be undertaken) of asphalt pavements against the original design life?
   - I. Yes
   - II. No
   - III. Not sure

6. In your state, do asphalt pavements generally meet the original design life?
   - I. Yes
   - II. No
   - III. Not sure

7. Generally, what kind of discrepancies do you observe between the design life and service life of asphalt pavements in your state?
   - I. Service life is roughly equal to the design life
   - II. Service life is less than 1/3 of the design life
   - III. Service life is between 1/3 to 2/3 of the design life
   - IV. Service life is more than 2/3 of the design life, but shorter than the design life
   - V. The service life is longer than the design life
   - VI. Not sure

8. [If II-IV to Q7] What do you think is the cause for this discrepancy between the design life and the service life of asphalt pavements in your state? [mark all that apply]
   I. Traffic volume was underestimated during the design
   II. Climatic factors (e.g., colder/warmer temperature) affected the pavement's service life
   III. Deficient construction materials were used
   IV. Errors made by the contractor
   V. The design life is generally overestimated by the state
   VI. Not sure
   VII. Not applicable
   VIII. Other. Please specify__________________

9. Who performs the acceptance testing (testing used for acceptance and payment) during the asphalt material production and paving? (select all that apply)
   I. Contractor
   II. Third party contracted by the contractor
   III. Transportation department
   IV. Third party contracted by the transportation department
   V. Not Sure
   VI. Not applicable
   VII. Other. Please specify__________________

10. Is contractor QC test data used in your state DOT to determine contractor payment during the construction of asphalt pavements?
   I. Yes
   II. No
   III. Not sure
   IV. Not applicable

11. Have you ever detected that the mix design and volumetric testing data reported prior to and during construction might not be representative of the actual material used in paving?
   I. Yes
   II. No
   III. Not sure
   IV. Not applicable

12. [If yes to Q11] What was the basis? [mark all that apply]
   I. Personal observation
   II. Data analysis by the state DOT

III. Contractors admitted to this
IV. Contractor employees admitted to this
V. Not applicable
VI. Other. Please specify__________________

13. Which, if any, measures does your agency adopt to ensure the accuracy of mix design and volumetric testing data? [mark all that apply]
    I. We have an internal auditing board to review volumetric testing
    II. Our QA team works in conjunction with volumetric testing
    III. We have independent assurance agents (employees with the power to observe and review technicians and perform replication testing)
    IV. We have a peer-review process
    V. We hire an external third-party for oversight
    VI. None
    VII. Not sure
    VIII. Not applicable
    IX. Other. Please specify__________________

14. Has your department ever attempted to detect / investigate potential manipulation of mix design and material testing data?
    I. Yes, please briefly explain how_________________
    II. No
    III. Not sure
    IV. Not applicable

15. What are the parameters that are most likely to be manipulated while mix design and material testing data is reported for agency approval? Please mark that the options presented below pertain to standard asphalt mix design volumetric tests [mark all that apply]
    I. Mass of bowl (AASHTO T 209)
    II. Mass of bowl and sample dry (AASHTO T 209)
    III. Submerged weight of bowl and sample (AASHTO T 209)
    IV. Dry Mass of puck (AASHTO T 166)
    V. Submerged weight of puck in water (AASHTO T 166)
    VI. Weight of puck at SSD condition (AASHTO T 166)
    VII. Mass of basket assembly (AASHTO T 308)
    VIII. Mass basket assembly and initial sample (AASHTO T 308)
    IX. Mass basket assembly and final aggregate (AASHTO T 308)
    X. Ignition furnace correction factor (AASHTO T 308)
    XI. Uncorrected binder content from printed ticket (AASHTO T 308)
    XII. Pan mass (AASHTO T 329)

XIII. Mass pan and initial sample (AASHTO T 329)
XIV. Mass pan and dry sample (AASHTO T 329)
XV. Percent passing the No. 200 sieve (AASHTO T 30)
XVI. None
XVII. Other. Please specify__________________

16. What reason do you think might explain possible manipulation of mix design and volumetric testing data? [mark all that apply]
    I. Unwillingness to reconduct the test
    II. Strained workloads make contractors unwilling to redo the test
    III. Lack of availability of re-testing sample material
    IV. Pressure to affect the payment factor in favor of the contractor
    V. Pressure to affect the payment in favor of the department
    VI. Concerns regarding loss of testing qualifications (such as AMRL certification)
    VII. To avoid scrutiny or conflict over results from the department
    VIII. To avoid scrutiny or conflict over results from the contractor
    IX. Not sure
    X. Not applicable
    XI. Other. Please specify__________________

17. Does your agency have structured ethical frameworks (e.g. a Code of Conduct), and do they provide training to ensure employee comprehension and facilitate compliance?
    I. Yes, they have a framework and training
    II. Yes, they have a framework, but no training
    III. No, they don't have either
    IV. Not sure

18. [If any yes (I-II) to Q17] Do these ethical frameworks apply specifically to material testing? For instance, data or source material tampering, deviations from procedure while reporting written execution, purposely changing conditions between tests, etc.
    I. Yes
    II. No
    III. Not sure
    IV. Not applicable

19. If you ever observed an ethical violation in material testing, who did you report it to? [mark all that apply]
    I. Supervisor
    II. Manager

III. Executive team
IV. Peer or colleagues
V. Human Resources Department
VI. Legal Department
VII. Third party ethics / compliance line
VIII. Internal Auditors
IX. QA representative
X. Other, please specify____________________
XI. I have observed these violations, but have not made a report
XII. I have never observed these violations
XIII. Not applicable

20. [If answered XI to Q19] Why did you not make a report of these violations? [mark all that apply]

I. Reporting process is too time consuming
II. Nothing would be done about the violations
III. I didn’t know who to report the violations to
IV. I didn't realize the violations occurred at the time
V. Other. Please specify ____________

21. [If any yes (I-X) to Q19] Was your concerns taken seriously?

I. Yes
II. No
III. Not sure
IV. Not applicable

22. [If any yes (I-X) to Q19] Who was receptive to your concerns? [mark all that apply]

I. Supervisor
II. Manager
III. Executive team
IV. Peer or colleagues
V. Human Resources Department
VI. Legal Department
VII. Third party ethics / compliance line
VIII. Internal Auditors
IX. QA representative
X. Not sure
XI. Not applicable
XII. Other. Please specify ____________

23. Irrespective to the testing laboratory (as in, both contractors and state), which option do you think is correct? [mark all that apply]
   I. Employees manipulate volumetric testing data on their own
   II. Employees follow recommendation from the contractor to manipulate volumetric testing data
   III. Employees follow recommendation from the department to manipulate volumetric testing data
   IV. Not sure
   V. Not applicable
   VI. Other. Please specify ____________

24. Anything else you would like to share at this time?

25. If you would like to be interviewed and provide more detailed information, please provide your contact information and availability time OR please contact the research team at deb.mishra@okstate.edu or mojtabasadegh@boisestate.edu.

26. If you want to share your email address please share it.

Appendix E. Sensitivity Analysis Results

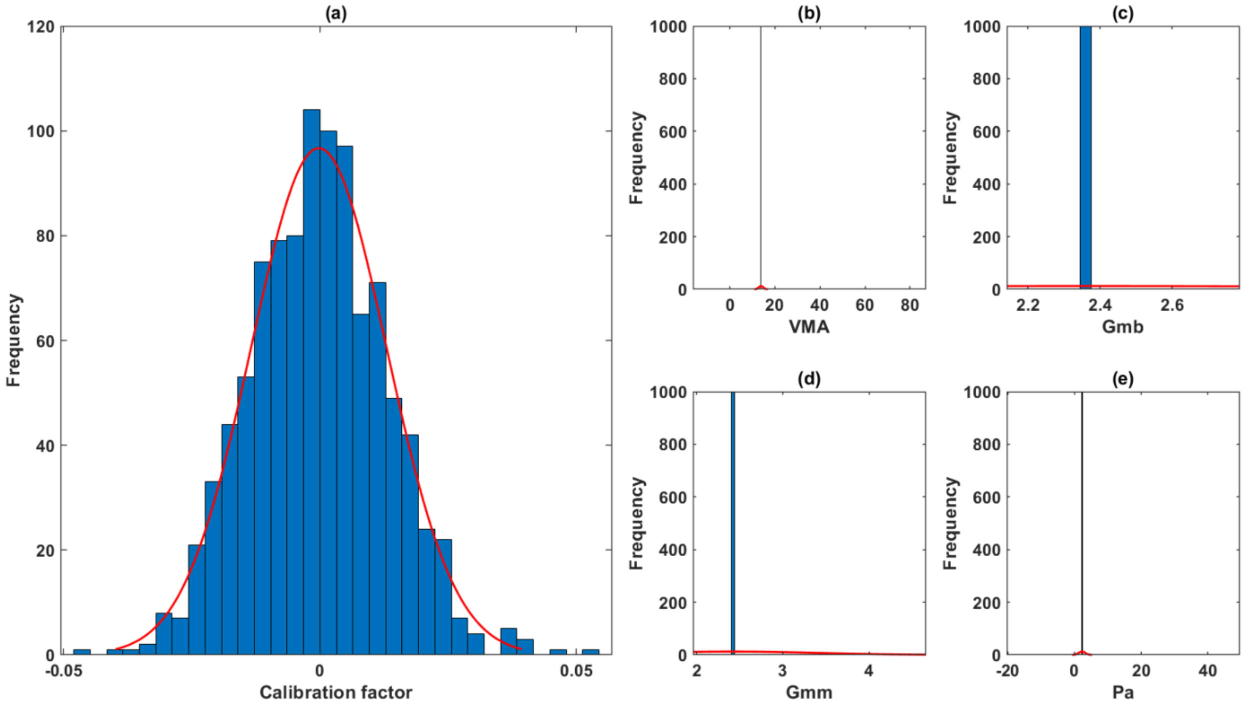


Figure 1. Sensitivity of VMA, Gmb, Gmm, and Pa parameters (sub plots b, c, d and e respectively) with respect to changes in calibration factor.

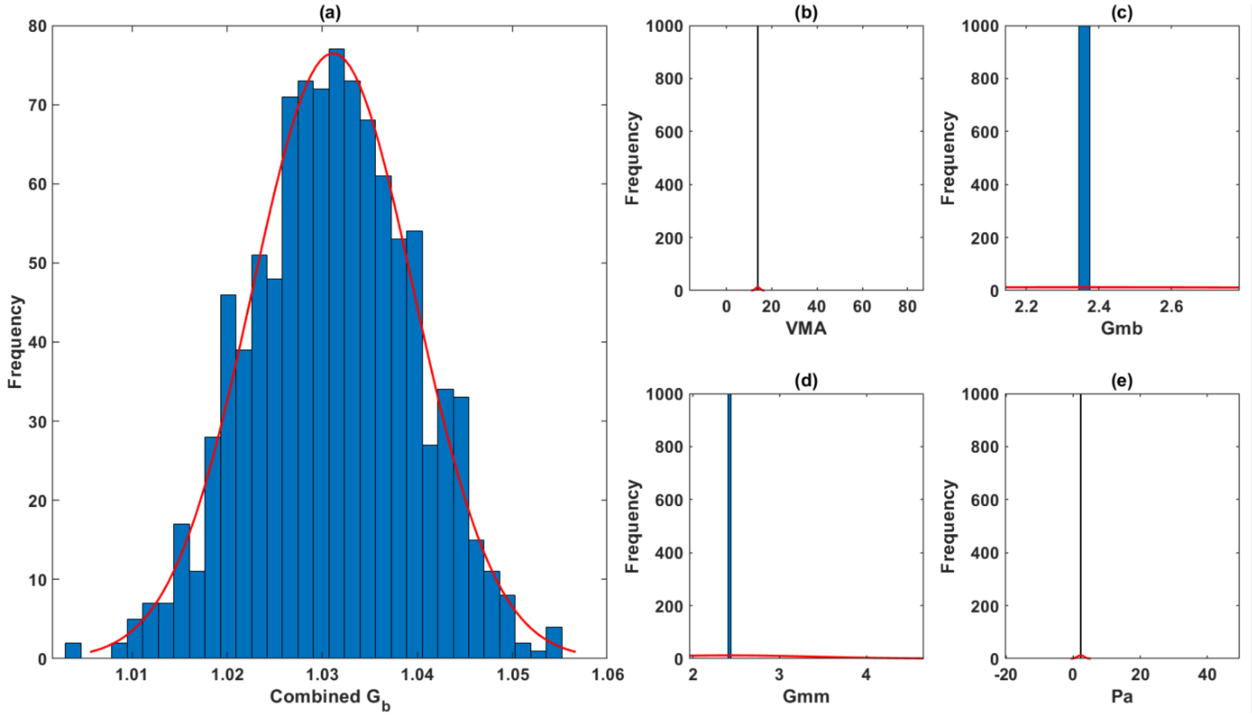


Figure 2. Sensitivity of VMA, Gmb, Gmm, and Pa parameters (sub plots b, c, d and e respectively) with respect to changes in combined Gb.

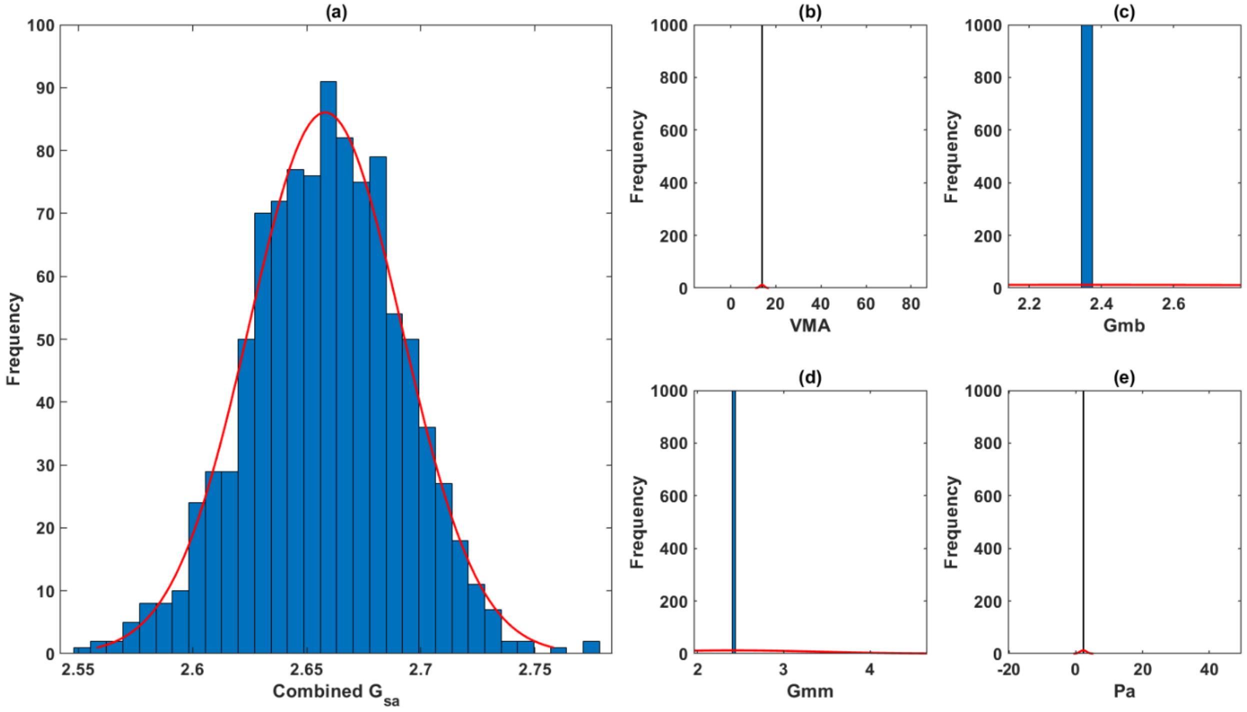


Figure 3. Sensitivity of VMA, Gmb, Gmm, and Pa parameters (sub plots b, c, d and e respectively) with respect to changes in combined Gsa.

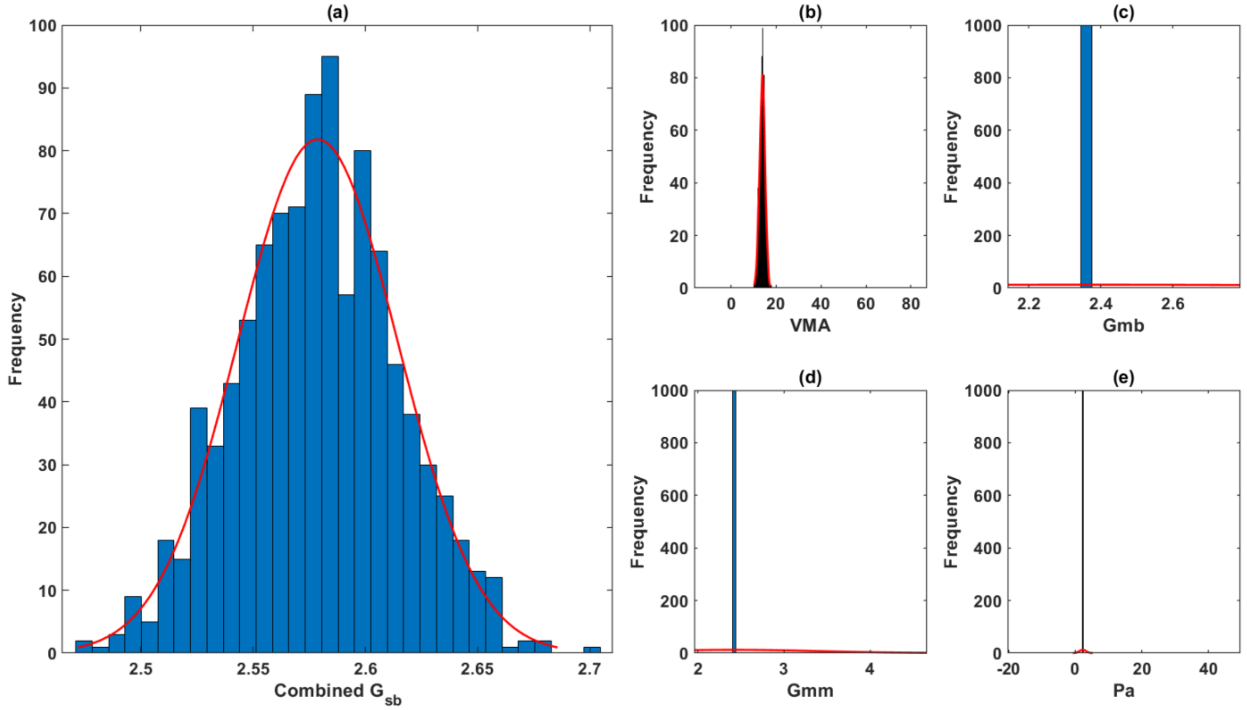


Figure 4. Sensitivity of VMA, Gmb, Gmm, and Pa parameters (sub plots b, c, d and e respectively) with respect to changes in combined Gsb.

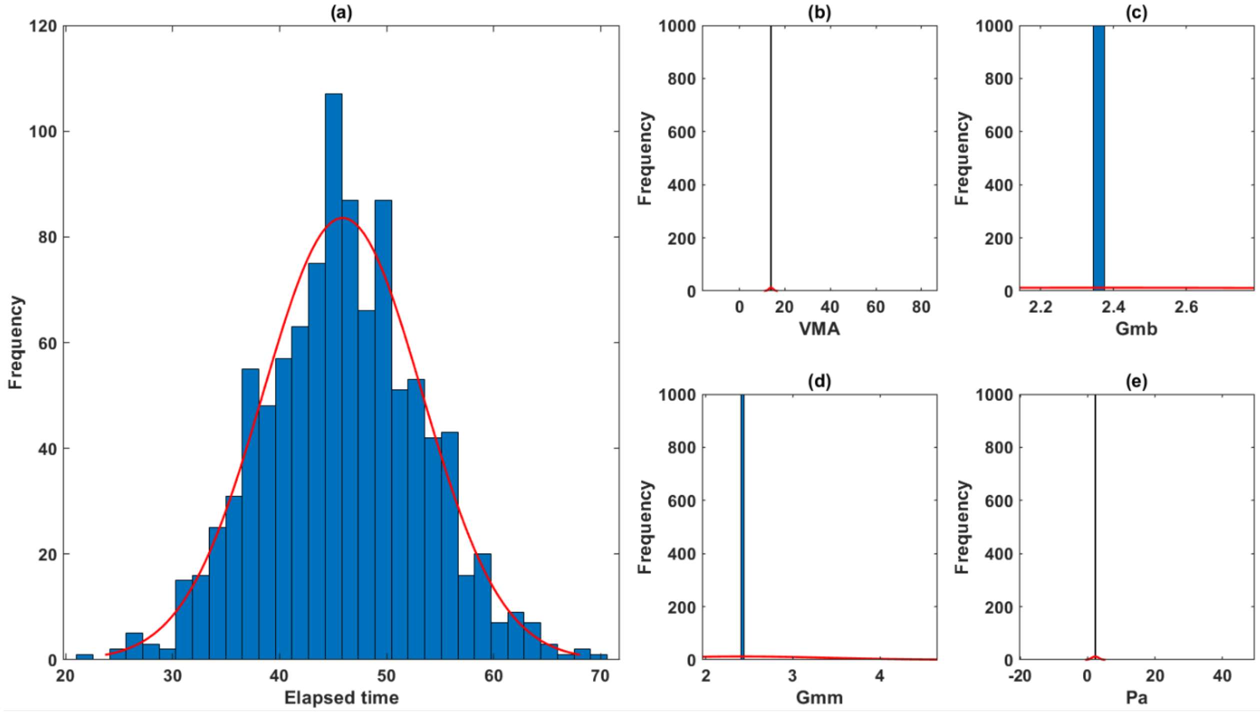


Figure 5. Sensitivity of VMA, Gmb, Gmm, and Pa parameters (sub plots b, c, d and e respectively) with respect to changes in elapsed time.

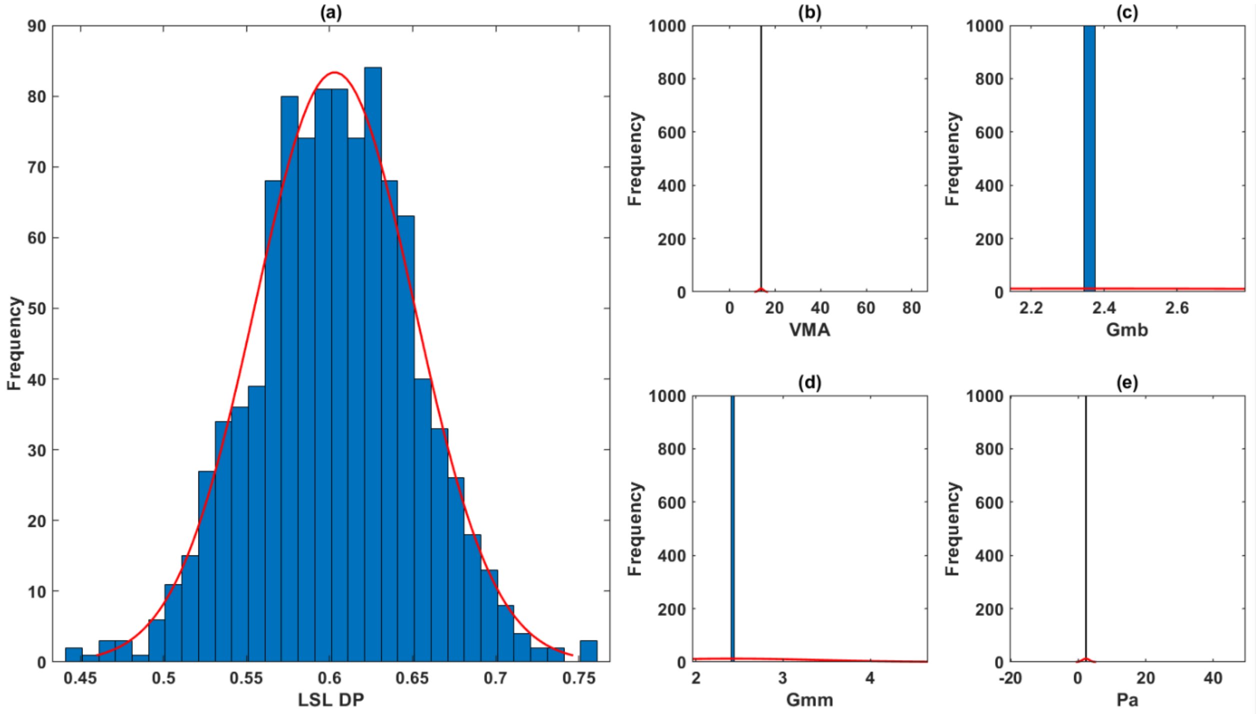


Figure 6. Sensitivity of VMA, Gmb, Gmm, and Pa parameters (sub plots b, c, d and e respectively) with respect to changes in LSL DP.

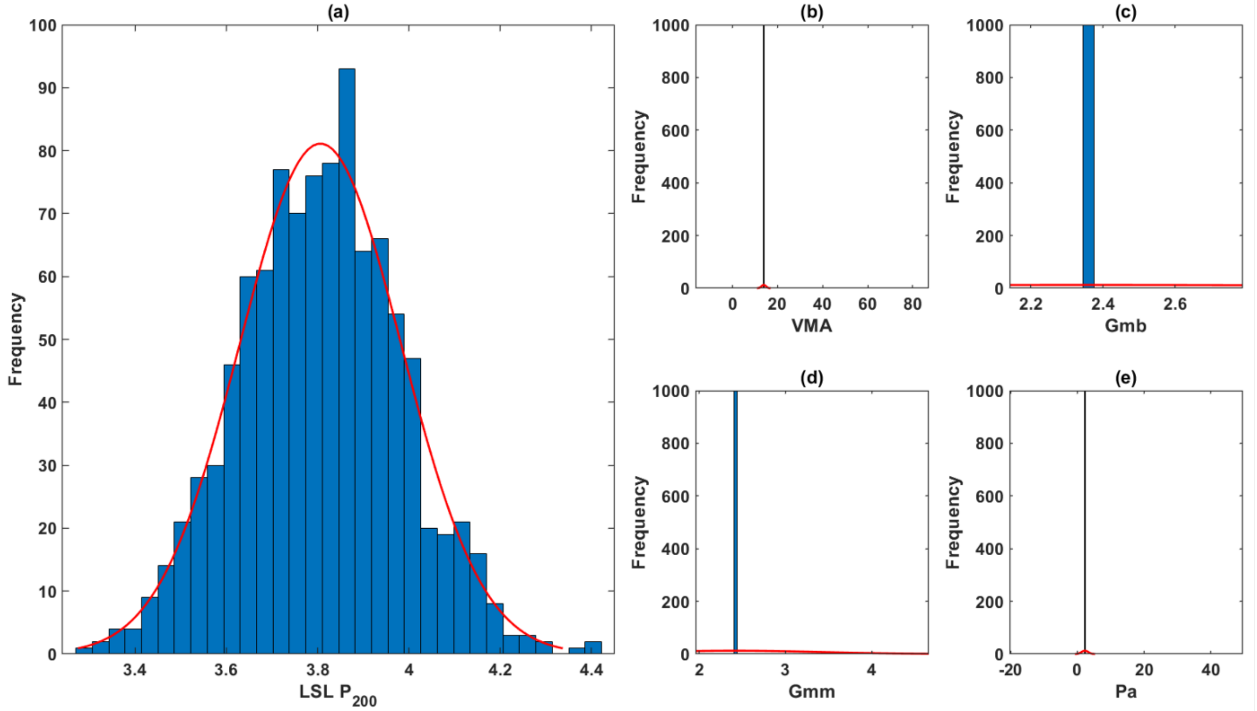


Figure 7. Sensitivity of VMA, Gmb, Gmm, and Pa parameters (sub plots b, c, d and e respectively) with respect to changes in LSL P200.

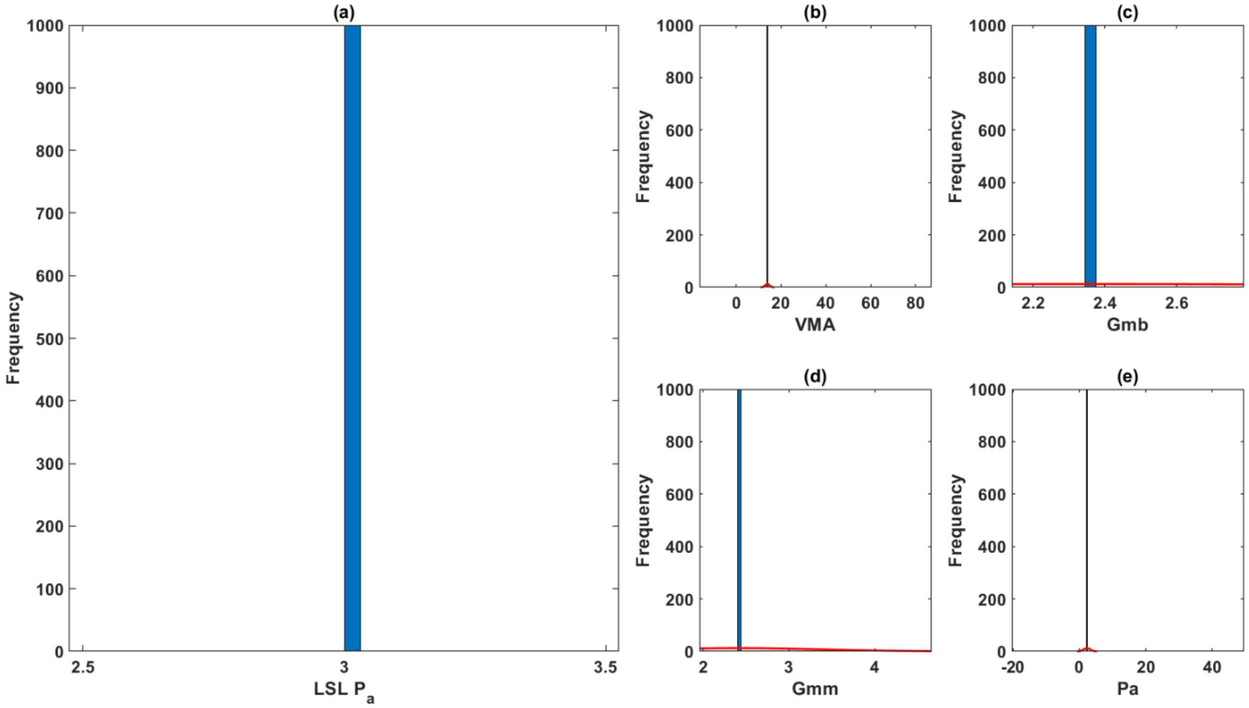


Figure 8. Sensitivity of VMA, Gmb, Gmm, and Pa parameters (sub plots b, c, d and e respectively) with respect to changes in LSL Pa.

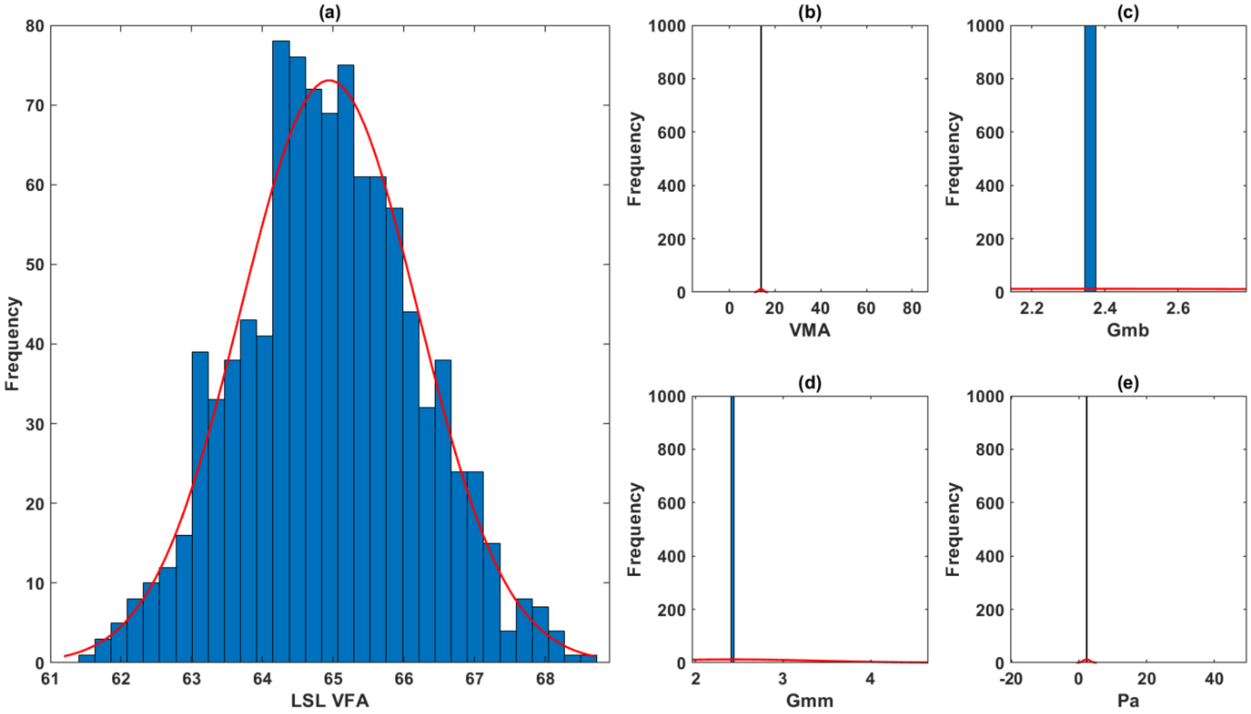


Figure 9. Sensitivity of VMA, Gmb, Gmm, and Pa parameters (sub plots b, c, d and e respectively) with respect to changes in LSL VFA.

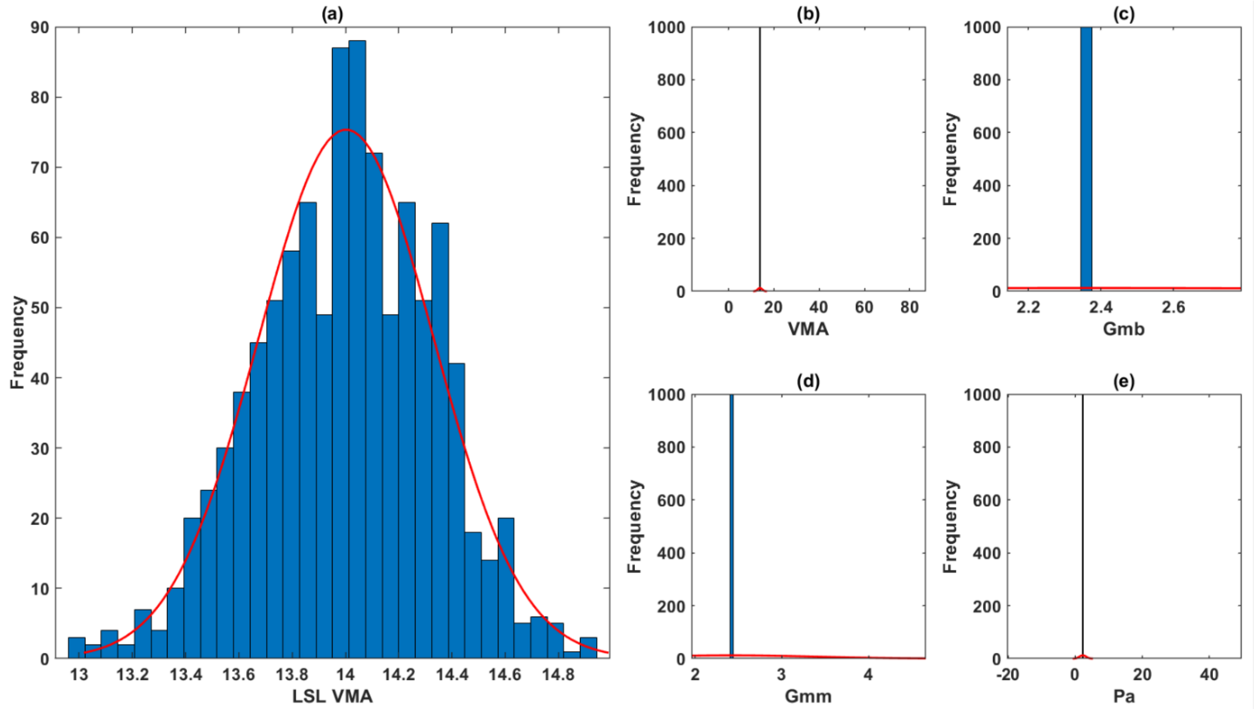


Figure 10. Sensitivity of VMA, Gmb, Gmm, and Pa parameters (sub plots b, c, d and e respectively) with respect to changes in LSL VMA.

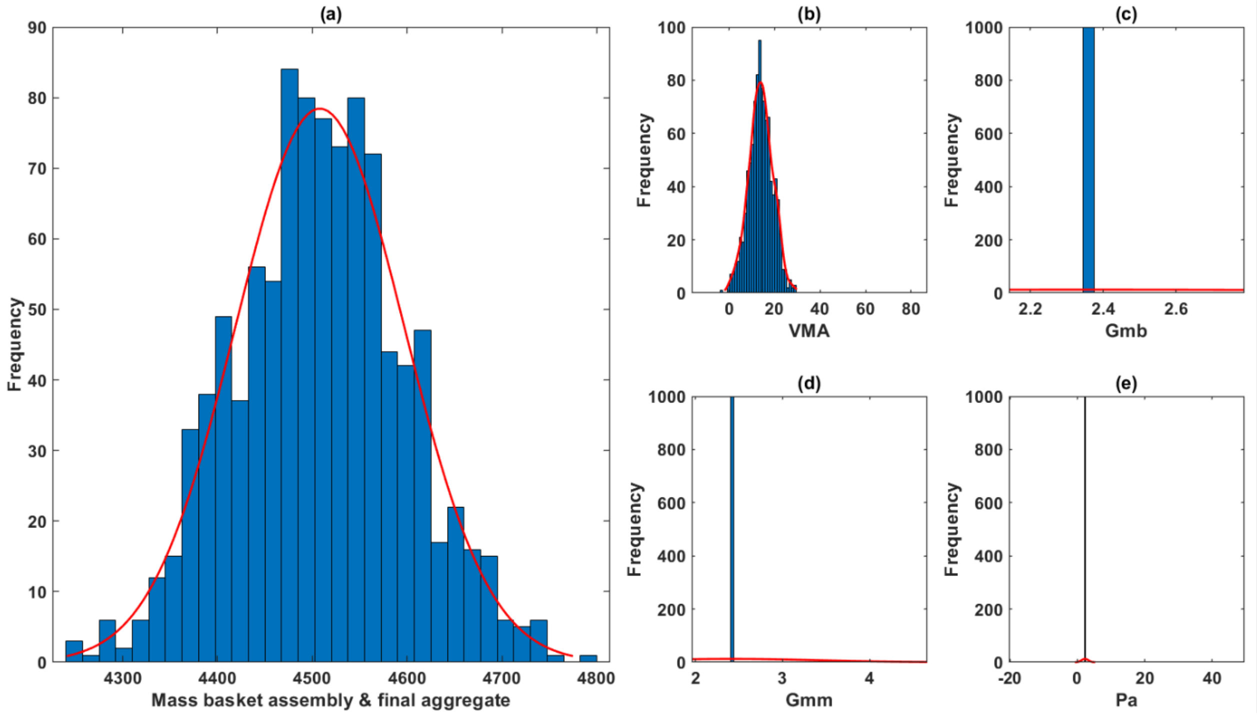


Figure 11. Sensitivity of VMA, Gmb, Gmm, and Pa parameters (sub plots b, c, d and e respectively) with respect to changes in mass basket assembly & final aggregate.

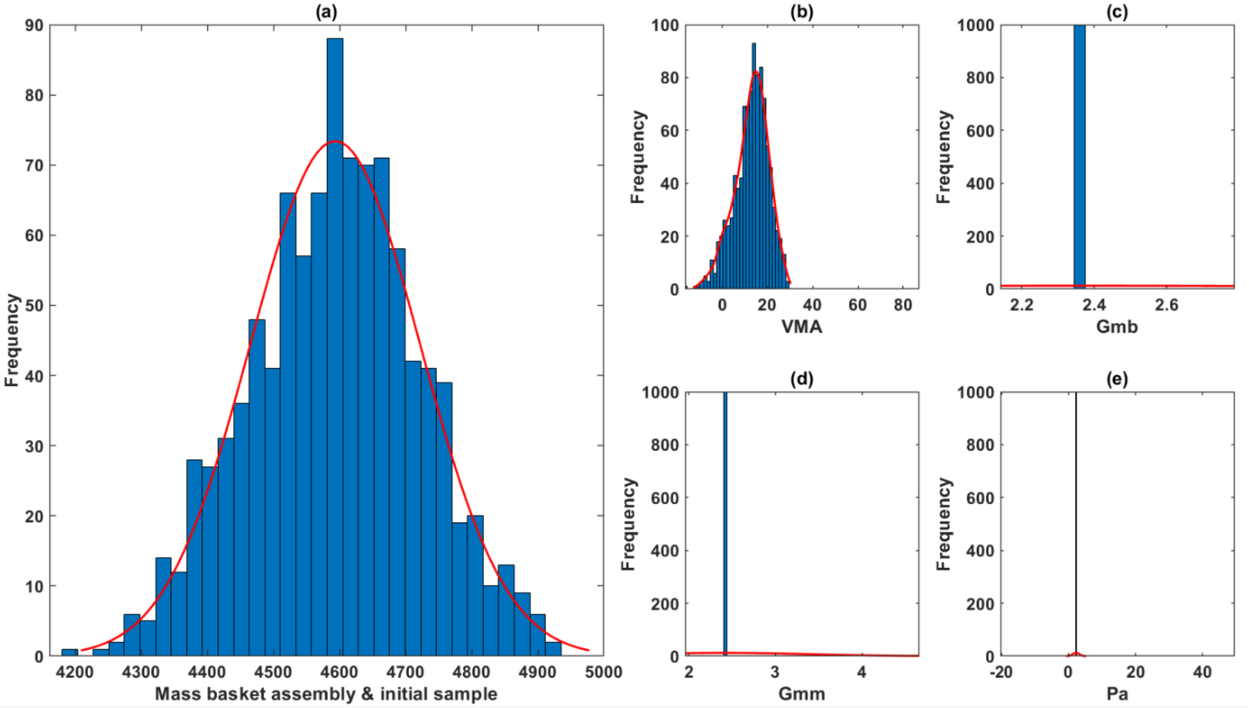


Figure 12. Sensitivity of VMA, Gmb, Gmm, and Pa parameters (sub plots b, c, d and e respectively) with respect to changes in mass basket assembly & initial sample.

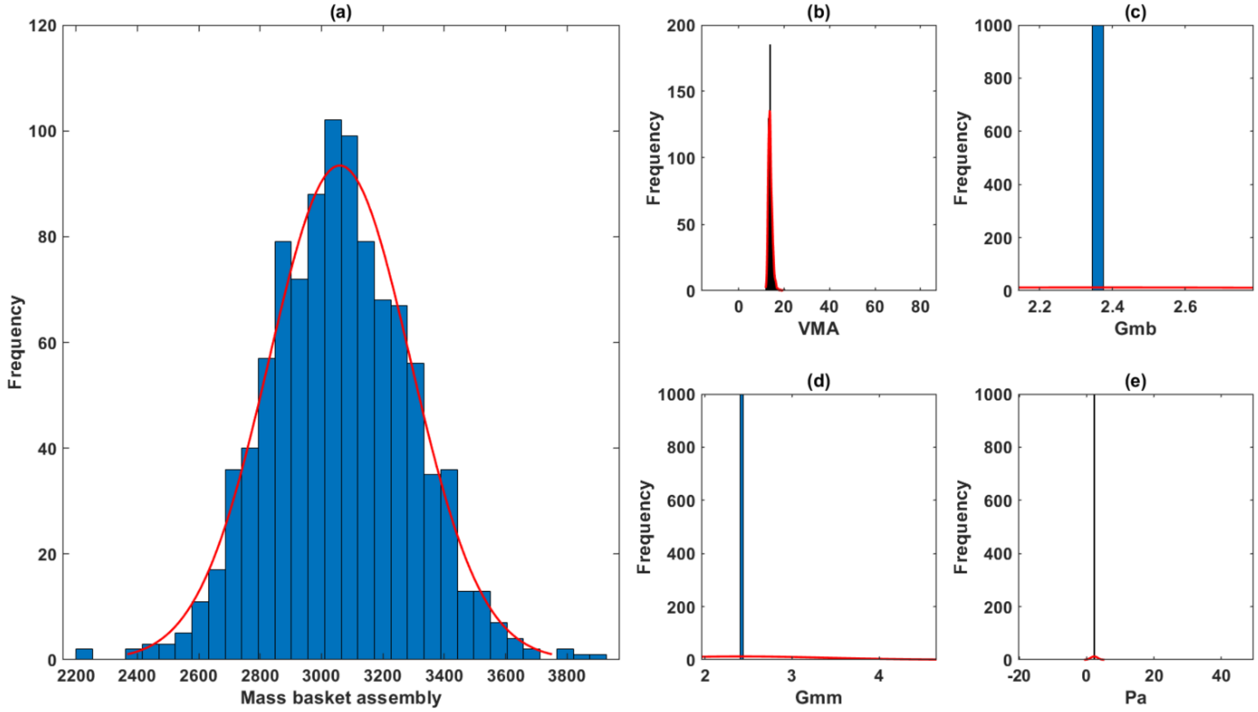


Figure 13. Sensitivity of VMA, Gmb, Gmm, and Pa parameters (sub plots b, c, d and e respectively) with respect to changes in mass basket assembly.

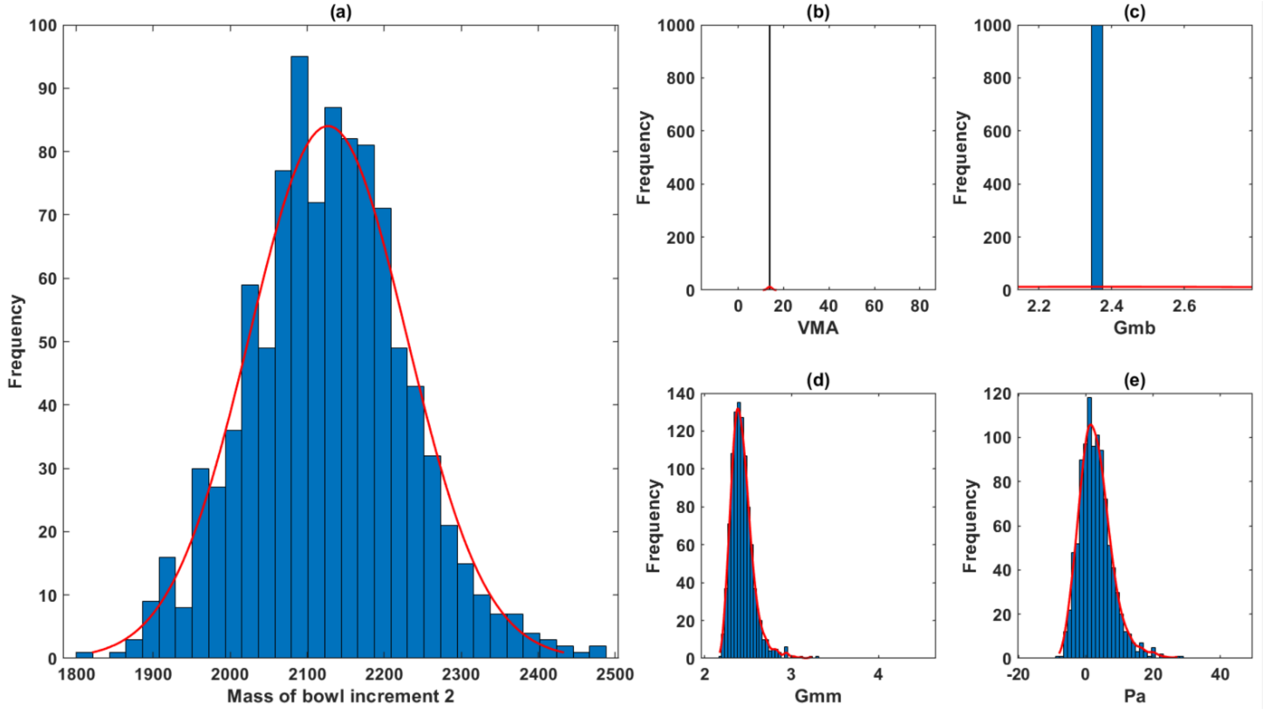


Figure 14. Sensitivity of VMA, Gmb, Gmm, and Pa parameters (sub plots b, c, d and e respectively) with respect to changes in mass of bowl increment 2.

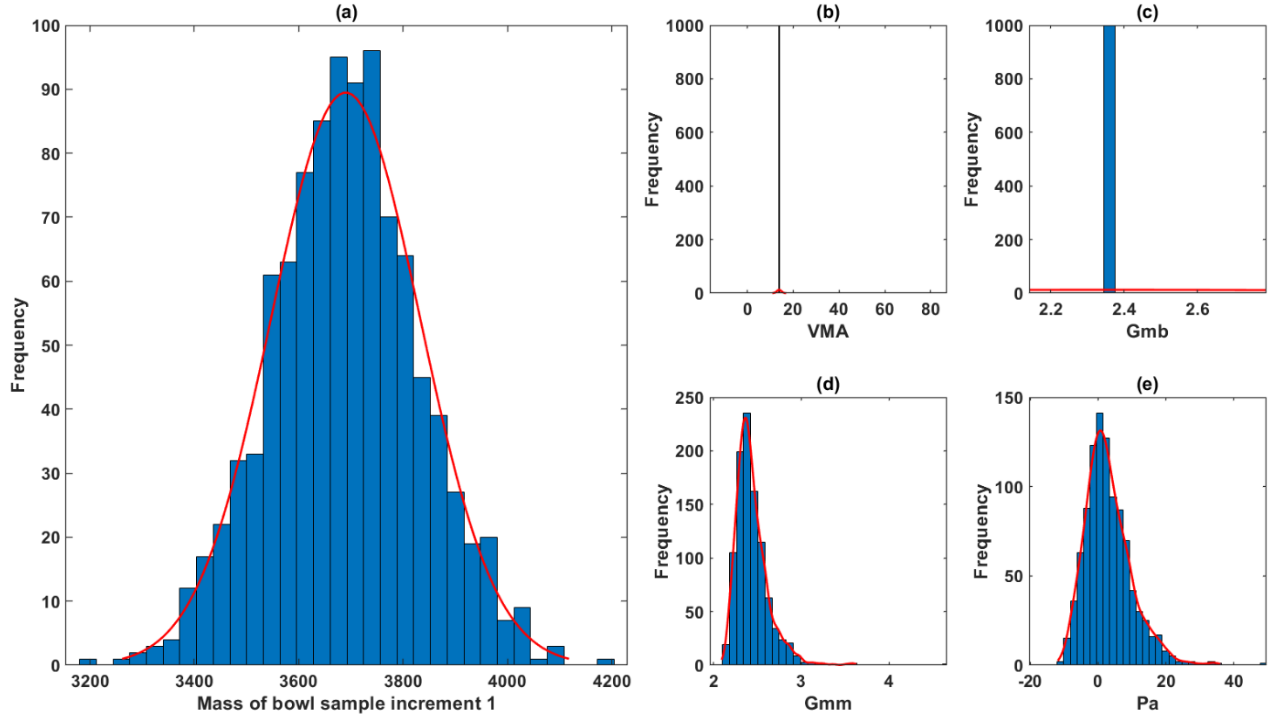


Figure 15. Sensitivity of VMA, Gmb, Gmm, and Pa parameters (sub plots b, c, d and e respectively) with respect to changes in mass of bowl sample increment 1.

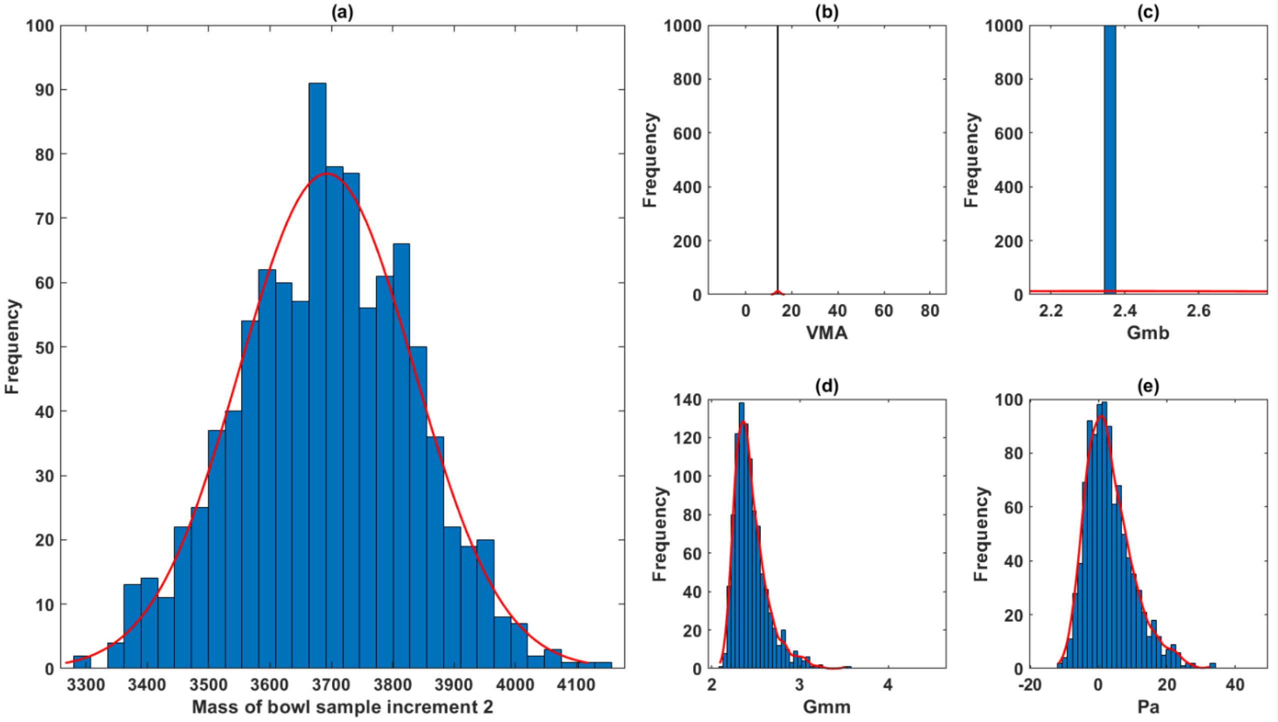


Figure 16. Sensitivity of VMA, Gmb, Gmm, and Pa parameters (sub plots b, c, d and e respectively) with respect to changes in mass of bowl sample increment 2.

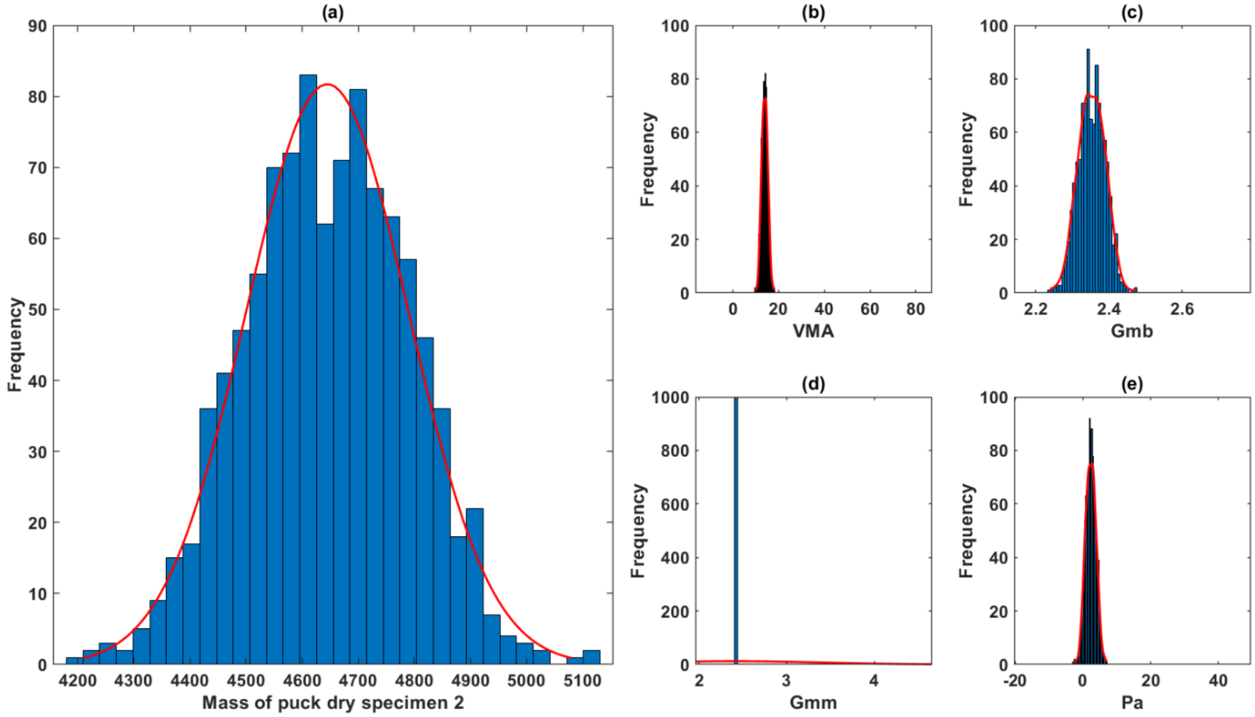


Figure 17. Sensitivity of VMA, Gmb, Gmm, and Pa parameters (sub plots b, c, d and e respectively) with respect to changes in mass of puck dry specimen 1.

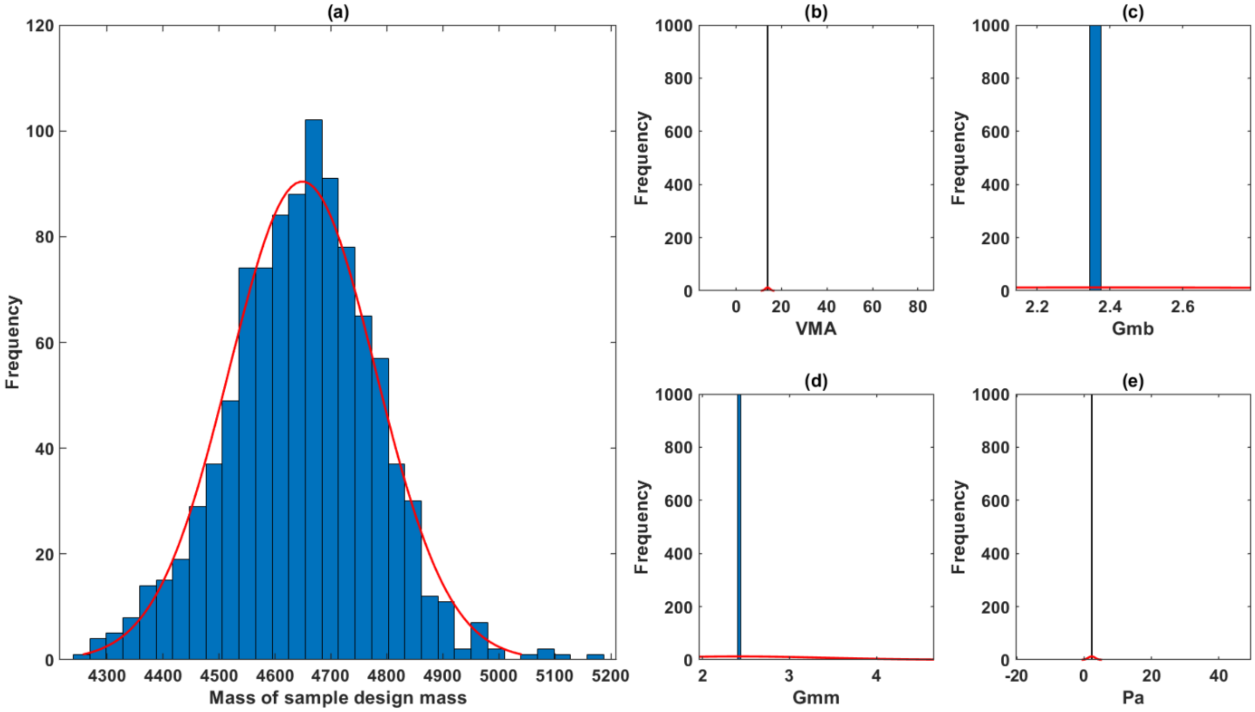


Figure 18. Sensitivity of VMA, Gmb, Gmm, and Pa parameters (sub plots b, c, d and e respectively) with respect to changes in mass of sample design mass.

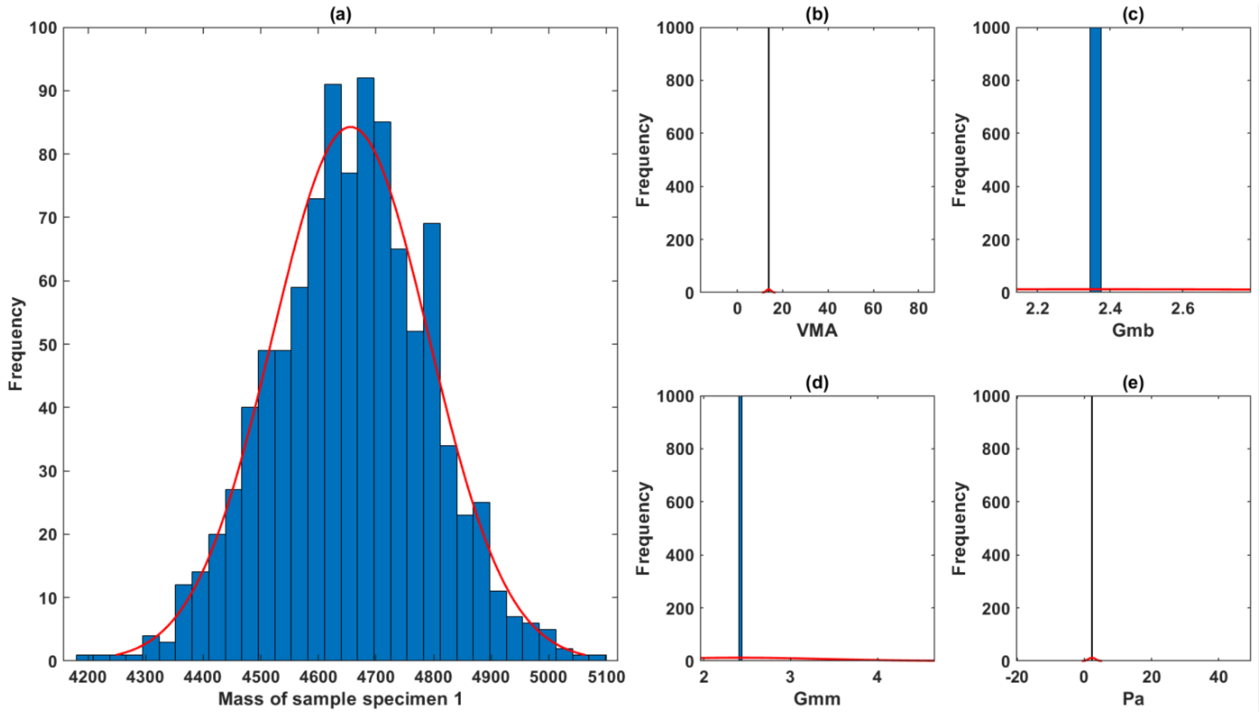


Figure 19. Sensitivity of VMA, Gmb, Gmm, and Pa parameters (sub plots b, c, d and e respectively) with respect to changes in mass of sample specimen 1.

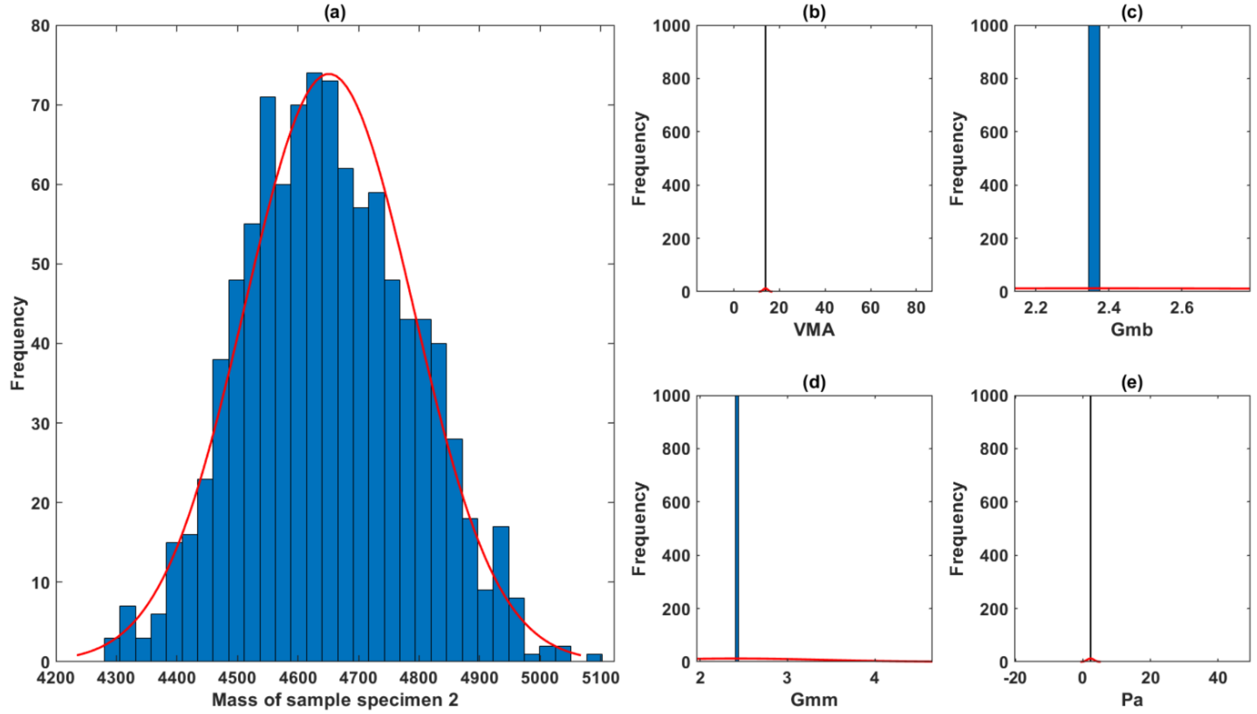


Figure 20. Sensitivity of VMA, Gmb, Gmm, and Pa parameters (sub plots b, c, d and e respectively) with respect to changes in mass of sample specimen 2.

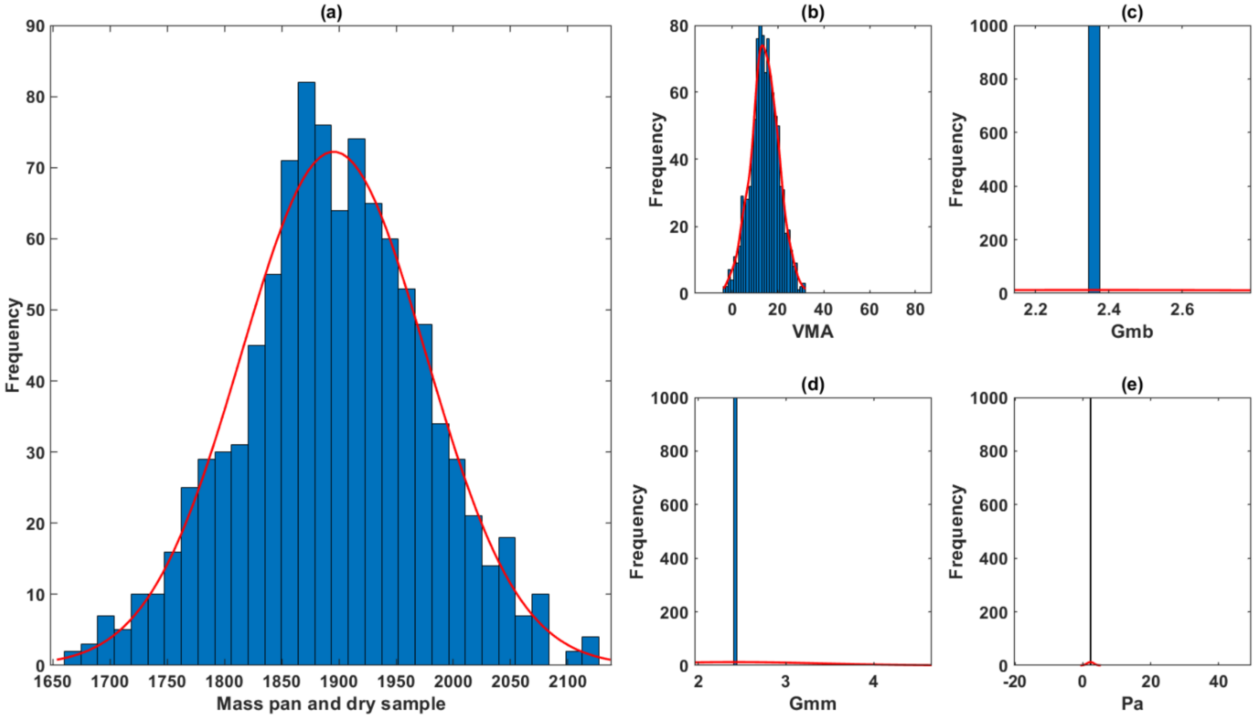


Figure 21. Sensitivity of VMA, Gmb, Gmm, and Pa parameters (sub plots b, c, d and e respectively) with respect to changes in mass pan and dry sample.

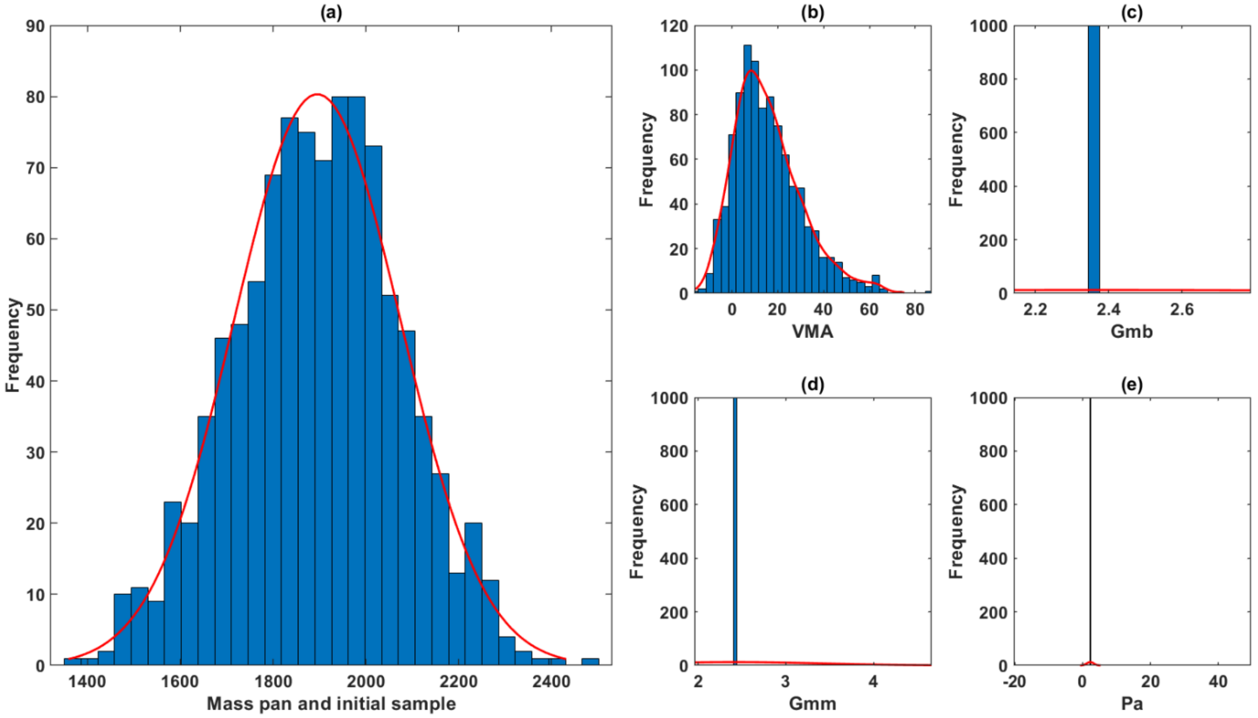


Figure 22. Sensitivity of VMA, Gmb, Gmm, and Pa parameters (sub plots b, c, d and e respectively) with respect to changes in mass pan and initial sample.

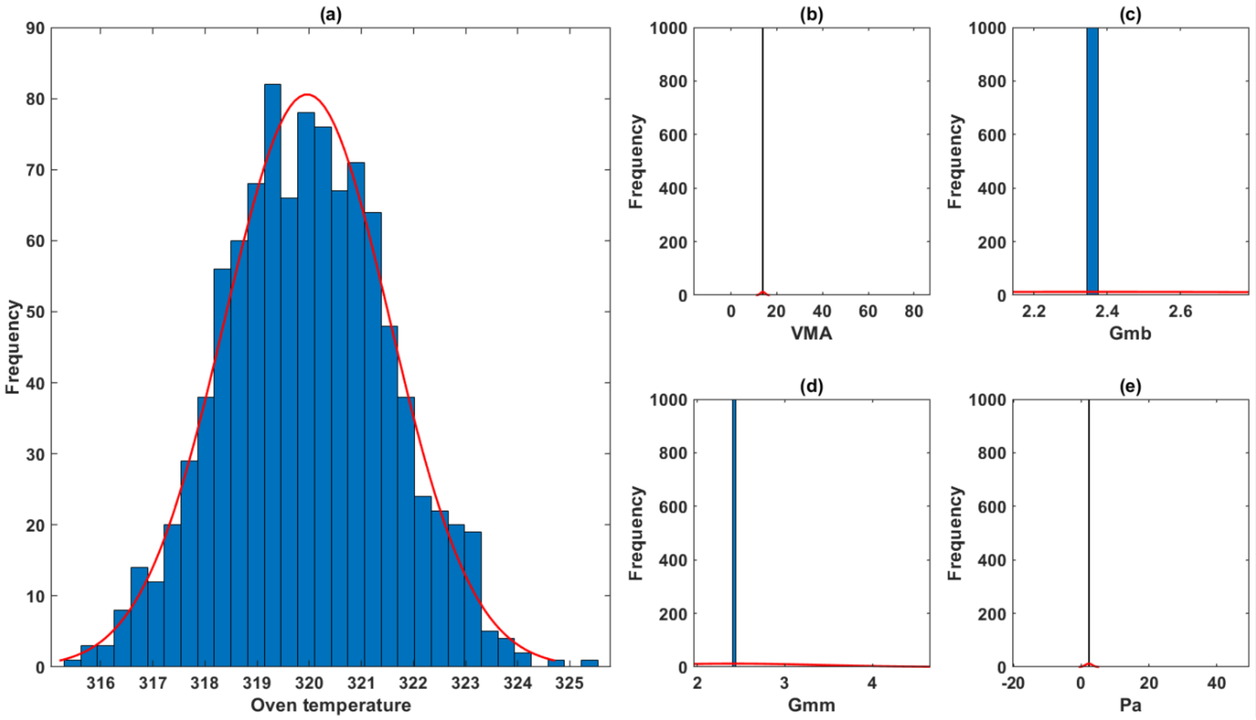


Figure 23. Sensitivity of VMA, Gmb, Gmm, and Pa parameters (sub plots b, c, d and e respectively) with respect to changes in oven temperature.

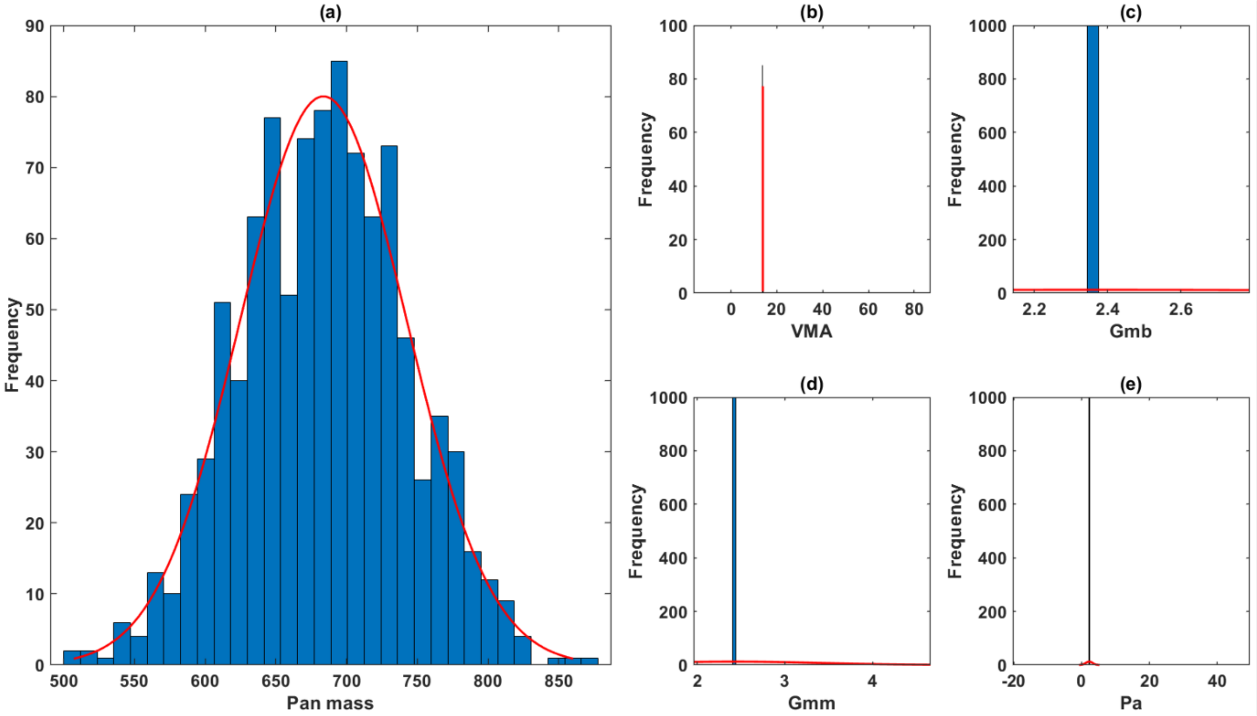


Figure 24. Sensitivity of VMA, Gmb, Gmm, and Pa parameters (sub plots b, c, d and e respectively) with respect to changes in pan mass.

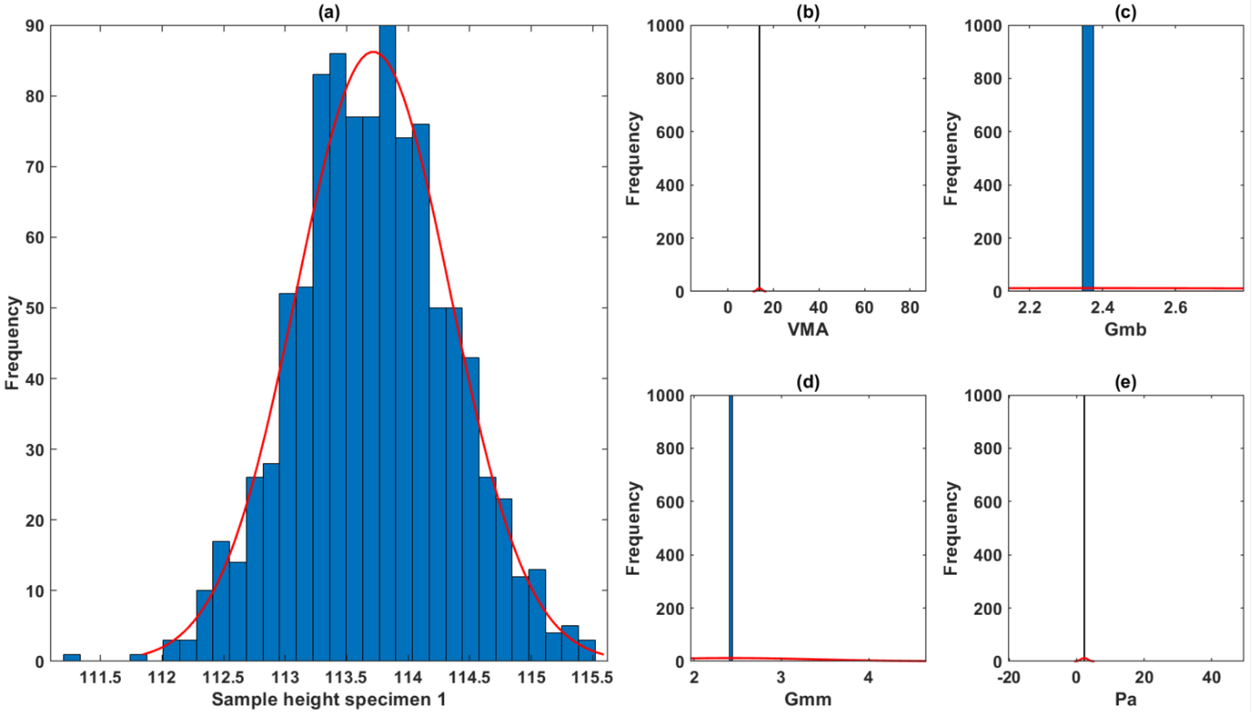


Figure 25. Sensitivity of VMA, Gmb, Gmm, and Pa parameters (sub plots b, c, d and e respectively) with respect to changes in sample height specimen 1.

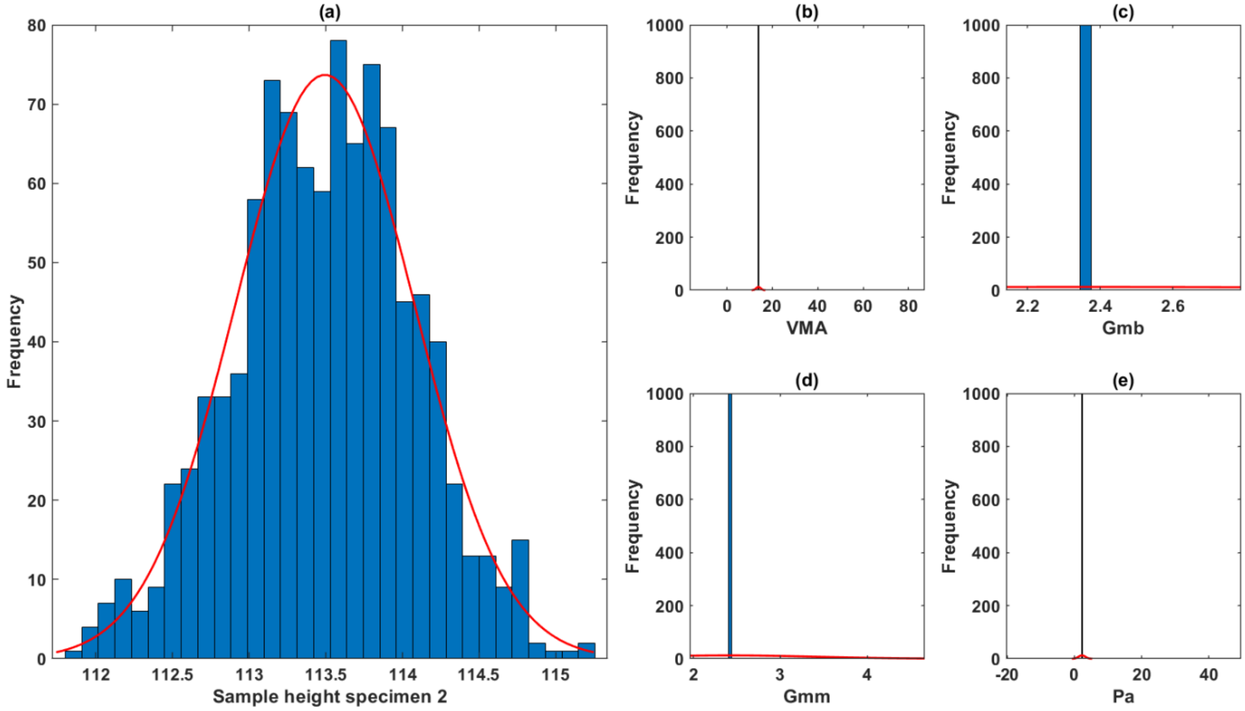


Figure 26. Sensitivity of VMA, Gmb, Gmm, and Pa parameters (sub plots b, c, d and e respectively) with respect to changes in sample height specimen 2.

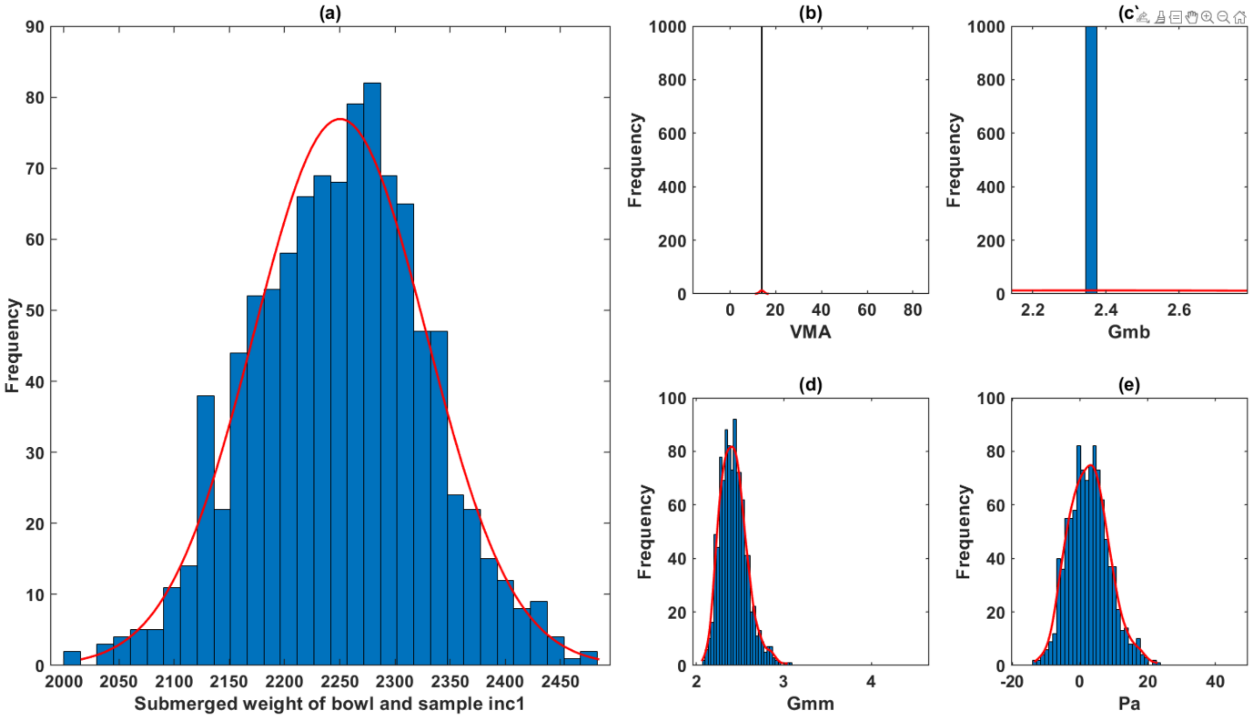


Figure 27. Sensitivity of VMA, Gmb, Gmm, and Pa parameters (sub plots b, c, d and e respectively) with respect to changes in submerged weight of bowl and sample inc1.

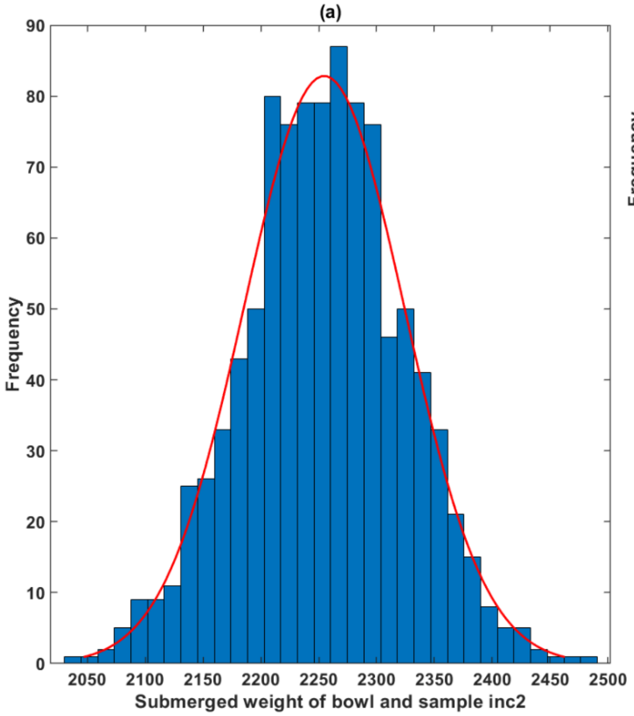


Figure 28. Sensitivity of VMA, Gmb, Gmm, and Pa parameters (sub plots b, c, d and e respectively) with respect to changes in submerged weight of bowl and sample inc2.

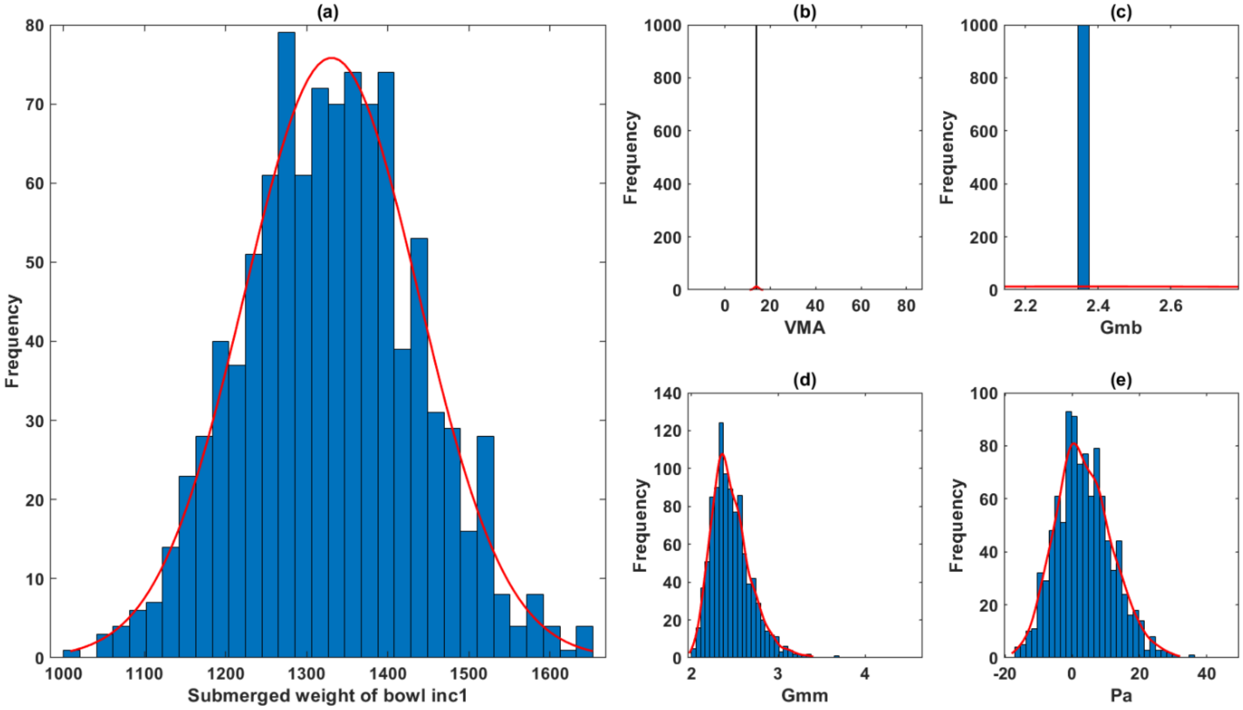


Figure 29. Sensitivity of VMA, Gmb, Gmm, and Pa parameters (sub plots b, c, d and e respectively) with respect to changes in submerged weight of bowl inc1.

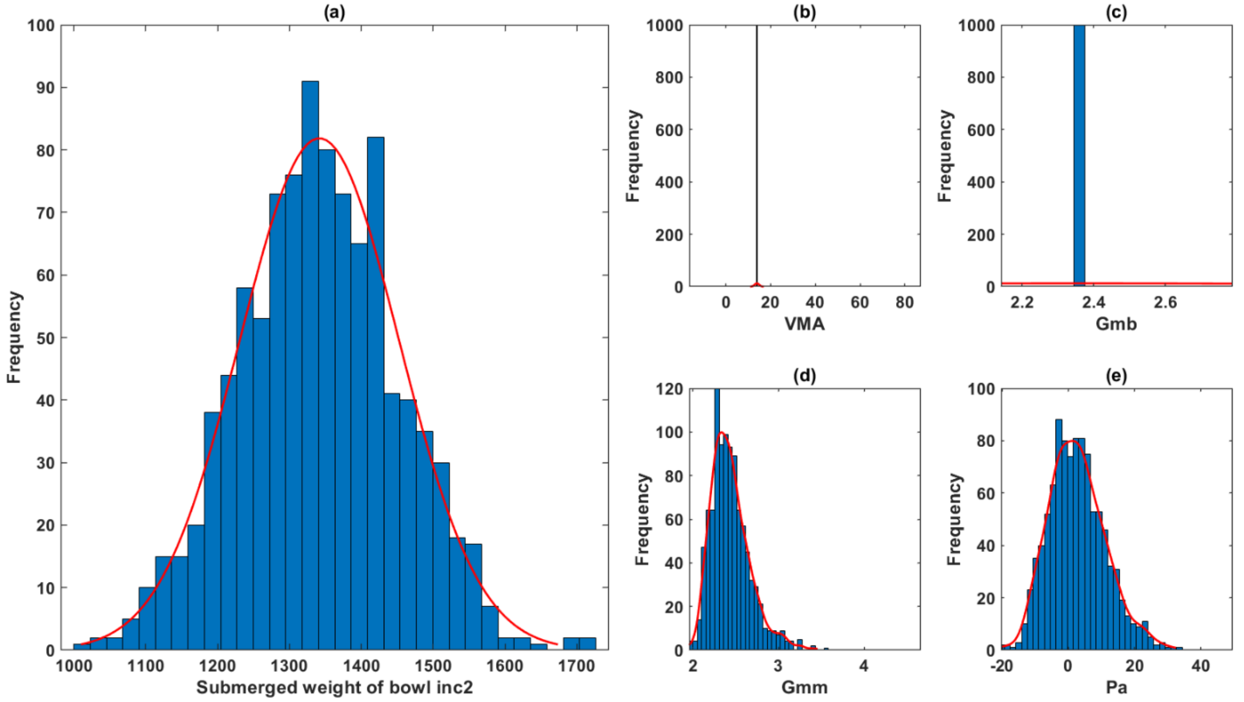


Figure 30. Sensitivity of VMA, Gmb, Gmm, and Pa parameters (sub plots b, c, d and e respectively) with respect to changes in submerged weight of bowl inc2.

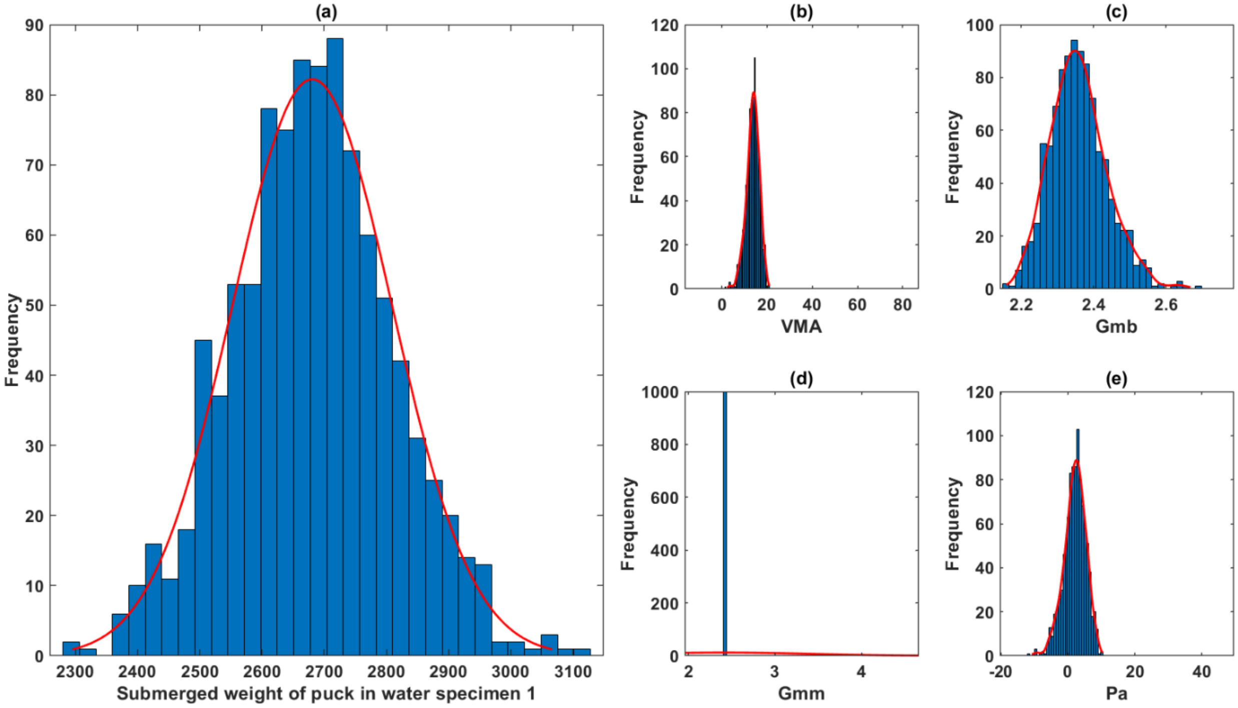


Figure 31. Sensitivity of VMA, Gmb, Gmm, and Pa parameters (sub plots b, c, d and e respectively) with respect to changes in submerged weight of puck in water specimen 1.

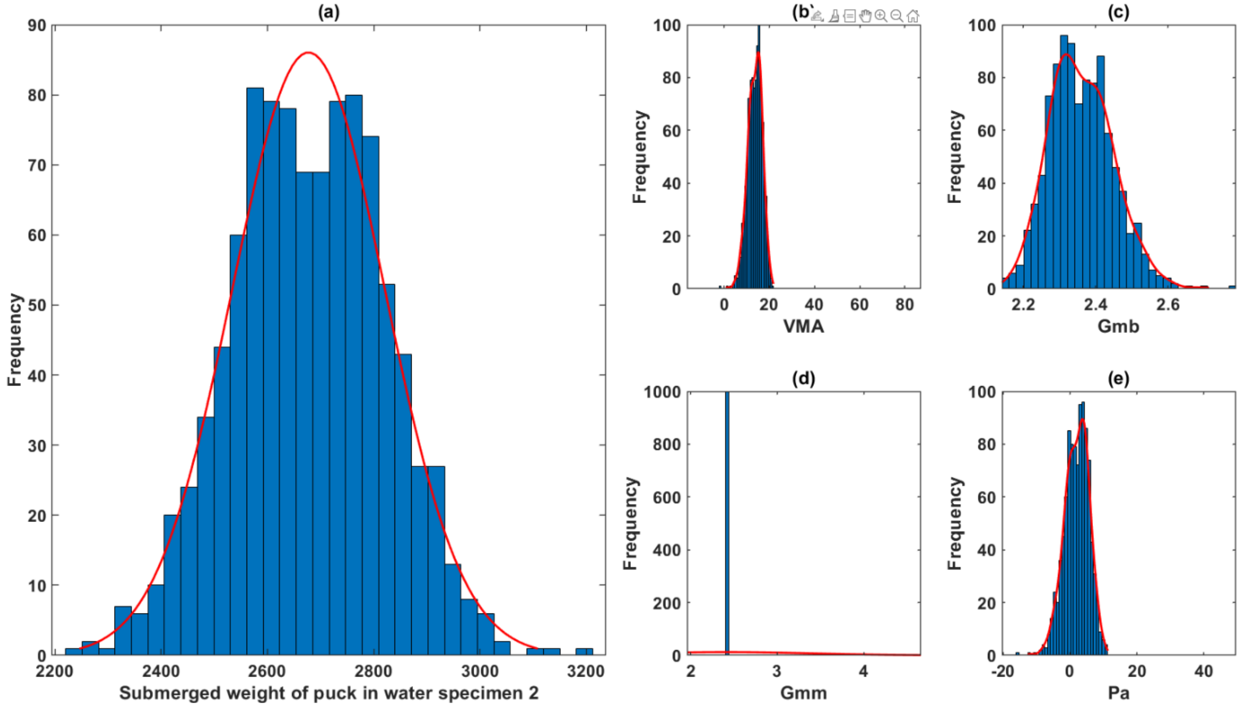


Figure 32. Sensitivity of VMA, Gmb, Gmm, and Pa parameters (sub plots b, c, d and e respectively) with respect to changes in submerged weight of puck in water specimen 2.

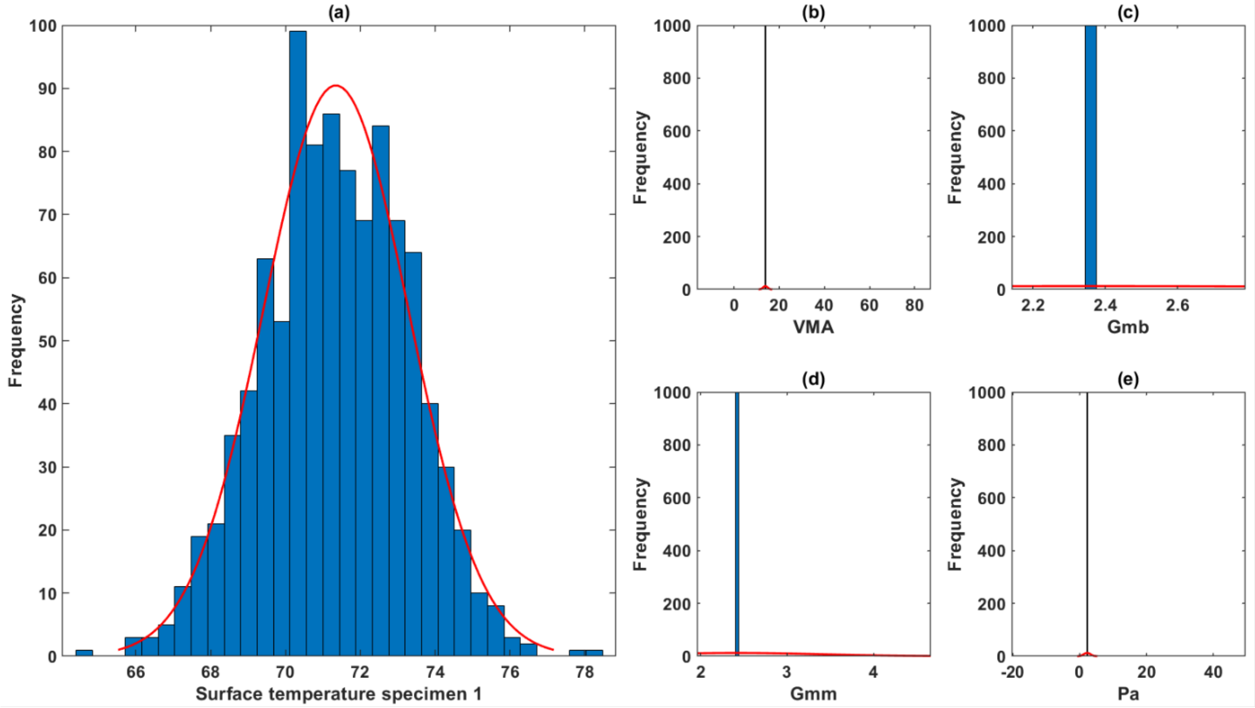


Figure 33. Sensitivity of VMA, Gmb, Gmm, and Pa parameters (sub plots b, c, d and e respectively) with respect to changes in surface temperature specimen 1.

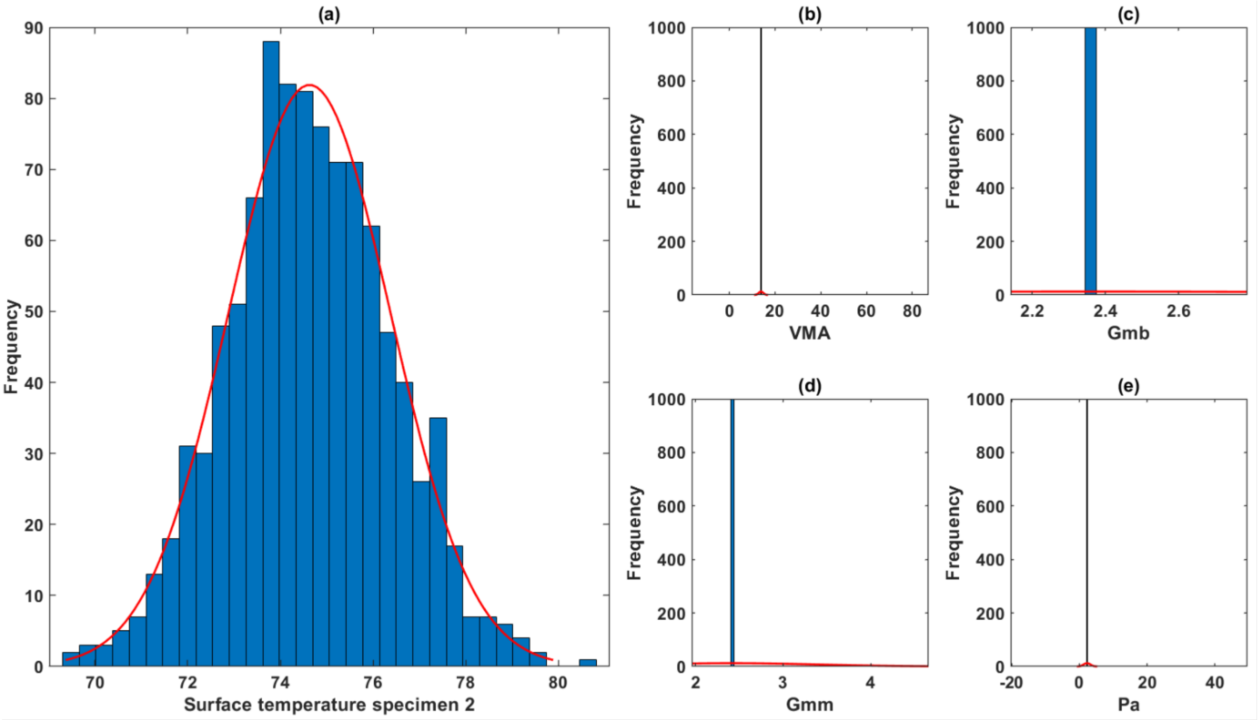


Figure 34. Sensitivity of VMA, Gmb, Gmm, and Pa parameters (sub plots b, c, d and e respectively) with respect to changes in surface temperature specimen 2.

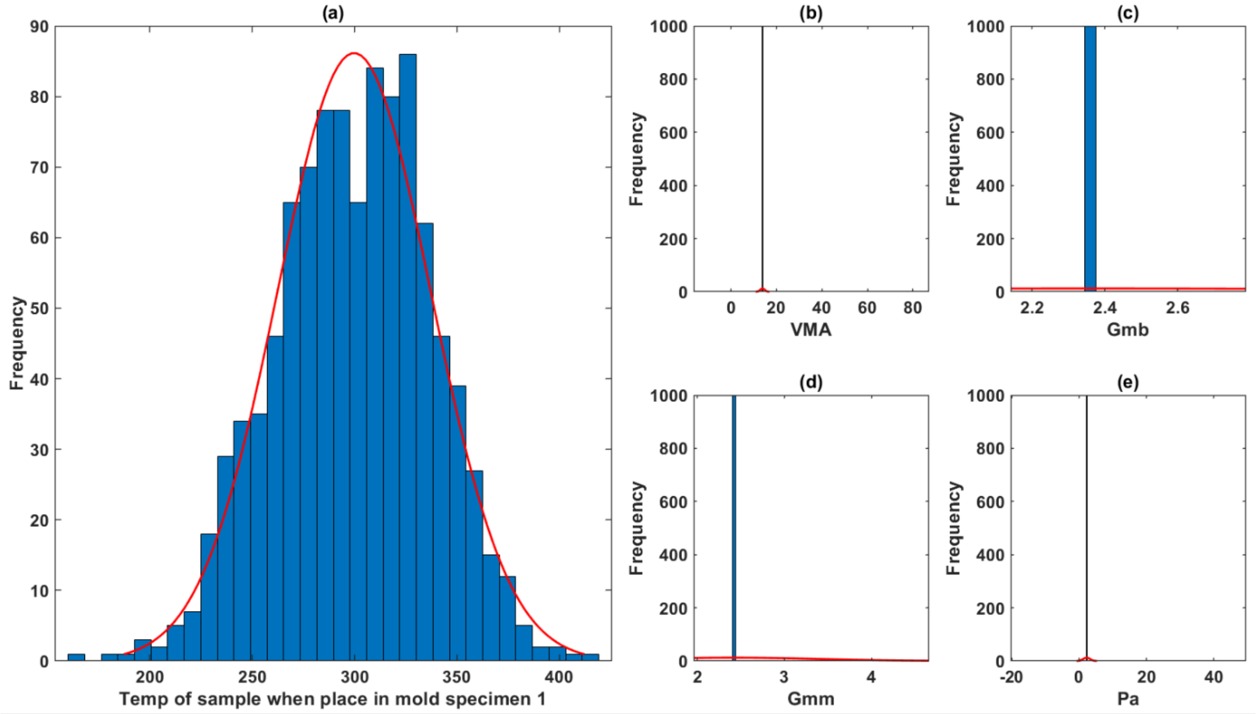


Figure 35. Sensitivity of VMA, Gmb, Gmm, and Pa parameters (sub plots b, c, d and e respectively) with respect to changes in temp of sample when placed in mold specimen 1.

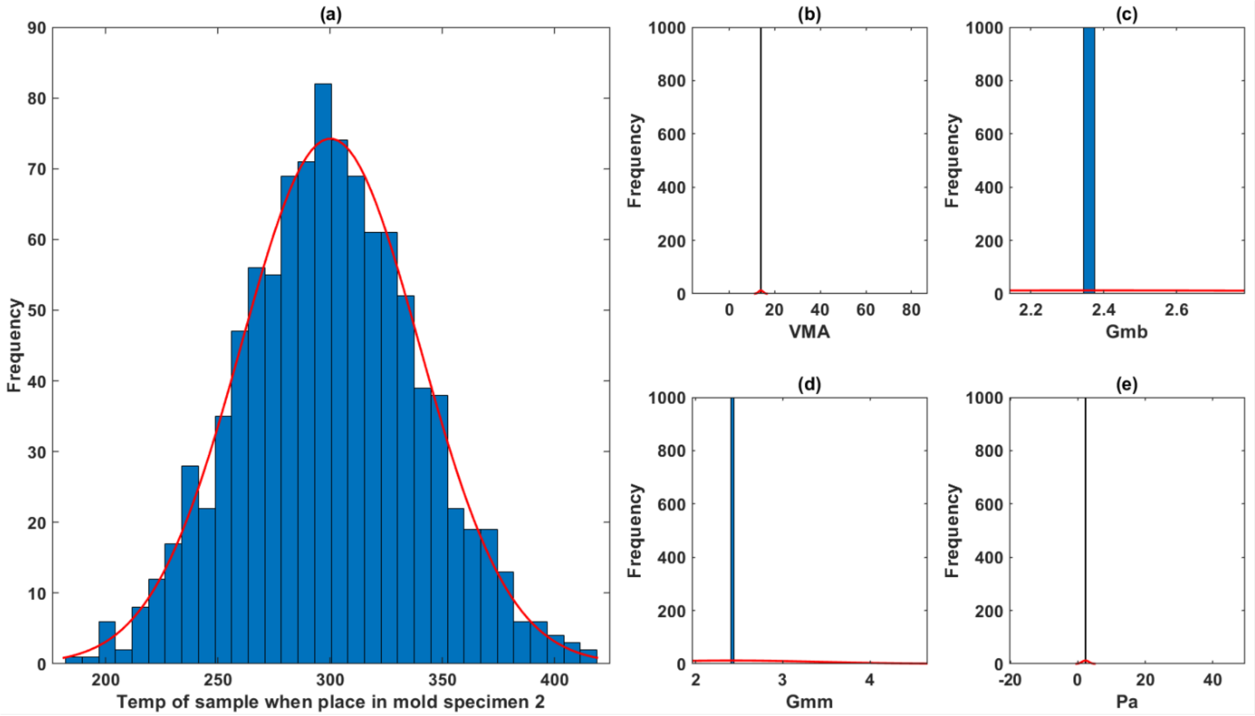


Figure 36. Sensitivity of VMA, Gmb, Gmm, and Pa parameters (sub plots b, c, d and e respectively) with respect to changes in temp of sample when placed in mold specimen 2.

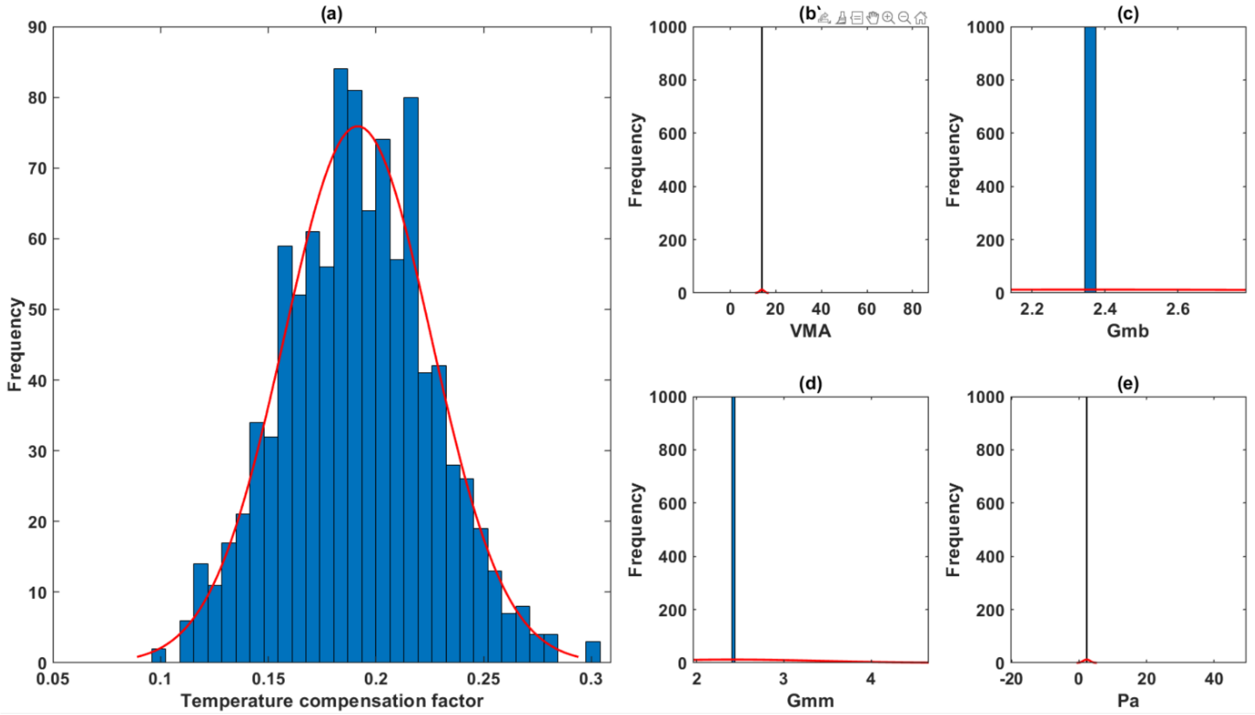


Figure 37. Sensitivity of VMA, Gmb, Gmm, and Pa parameters (sub plots b, c, d and e respectively) with respect to changes in temperature compensation.

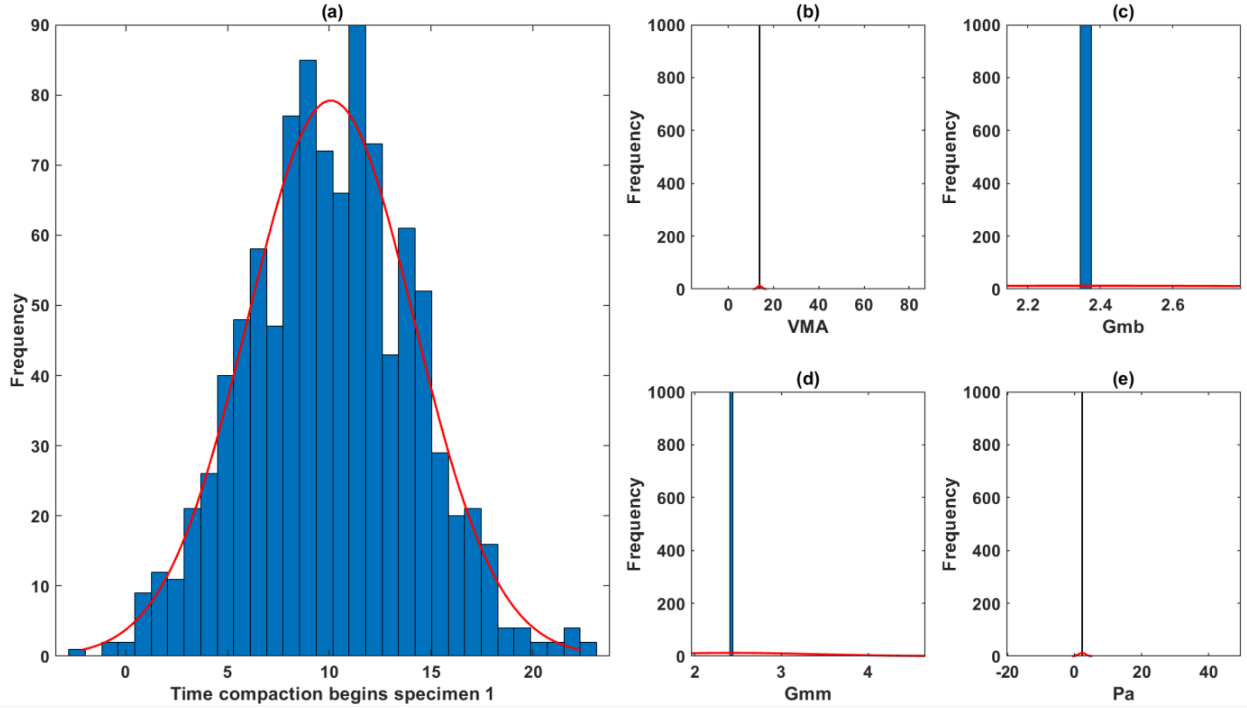


Figure 38. Sensitivity of VMA, Gmb, Gmm, and Pa parameters (sub plots b, c, d and e respectively) with respect to time compaction begins specimen 1.

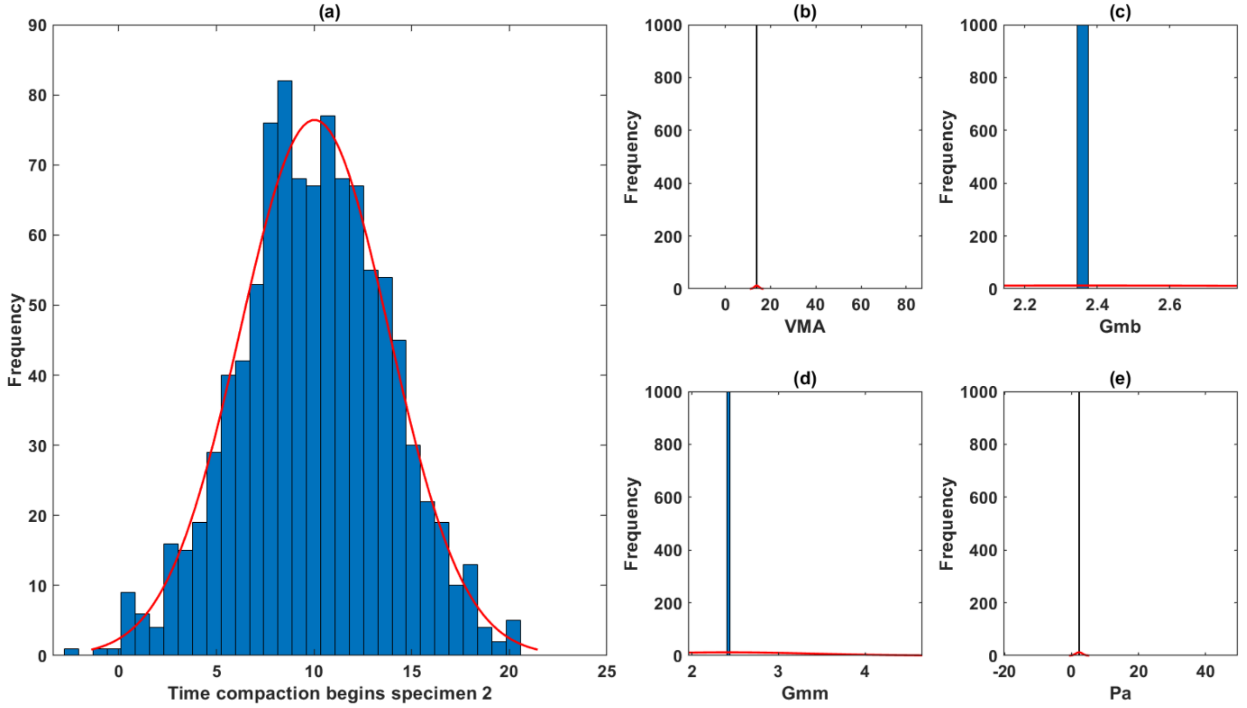


Figure 39. Sensitivity of VMA, Gmb, Gmm, and Pa parameters (sub plots b, c, d and e respectively) with respect to time compaction begins specimen 2.

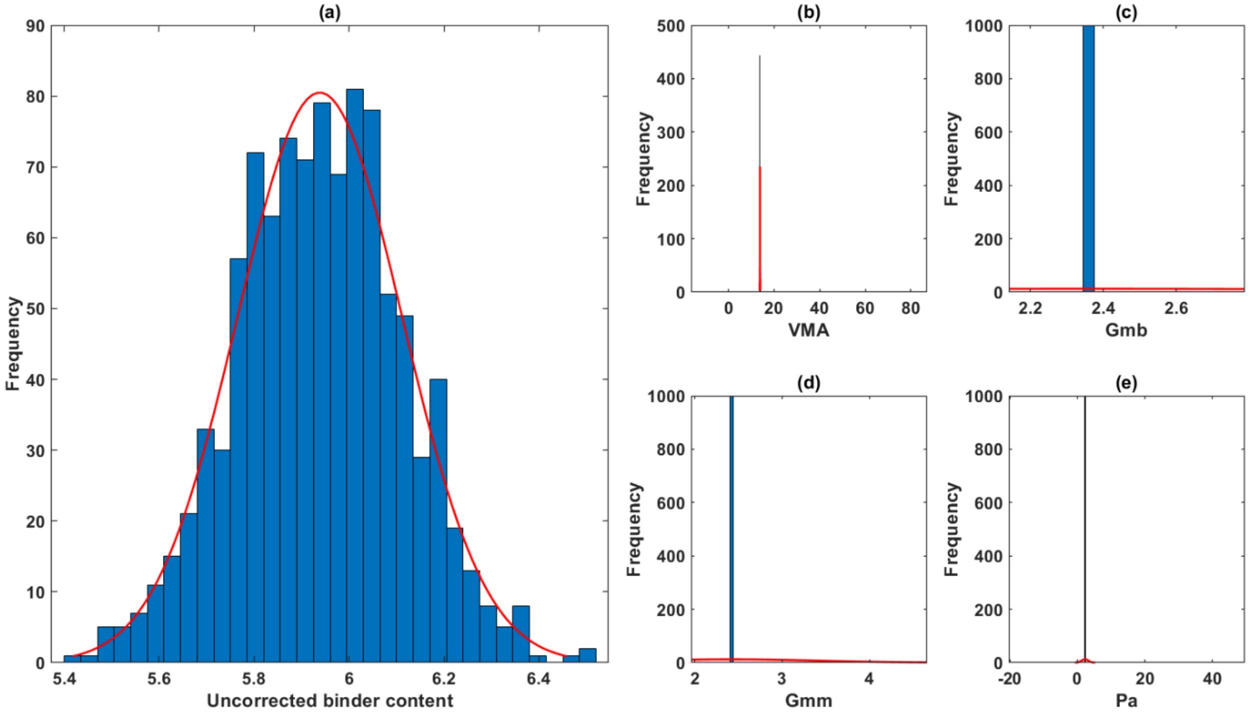


Figure 40. Sensitivity of VMA, Gmb, Gmm, and Pa parameters (sub plots b, c, d and e respectively) with respect to uncorrected binder content.

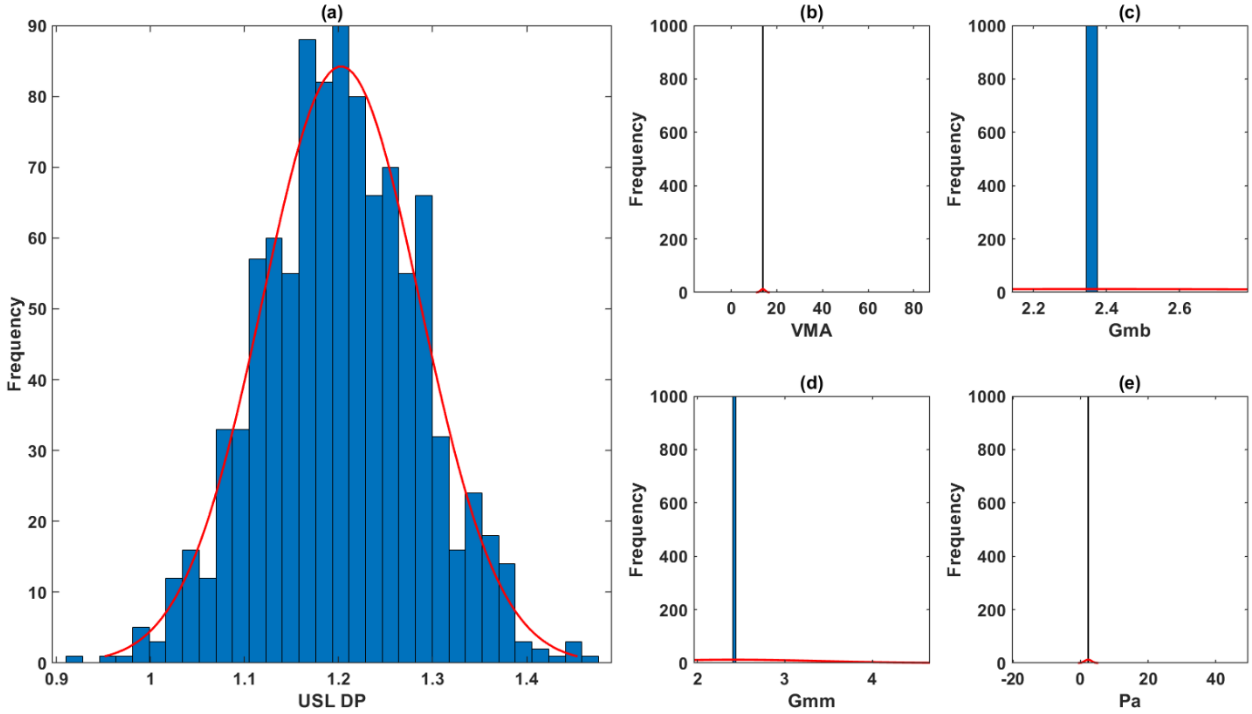


Figure 41. Sensitivity of VMA, Gmb, Gmm, and Pa parameters (sub plots b, c, d and e respectively) with respect to USL DP.

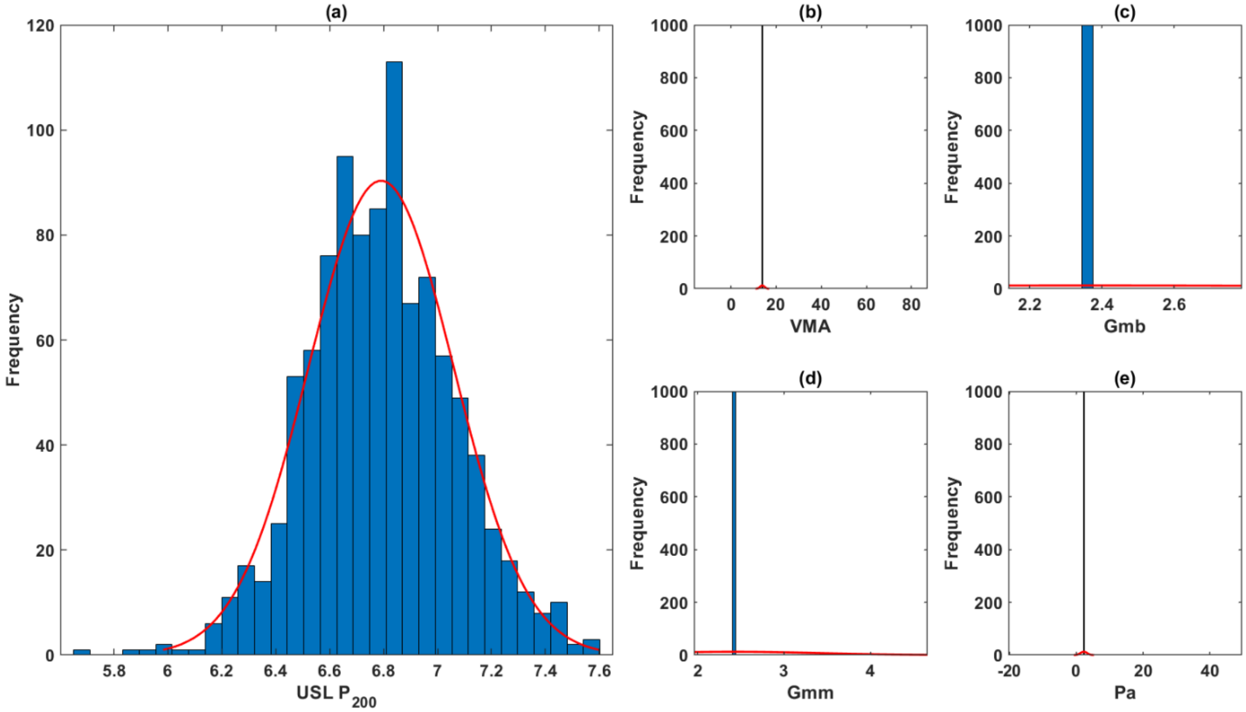


Figure 42. Sensitivity of VMA, Gmb, Gmm, and Pa parameters (sub plots b, c, d and e respectively) with respect to USL P200.

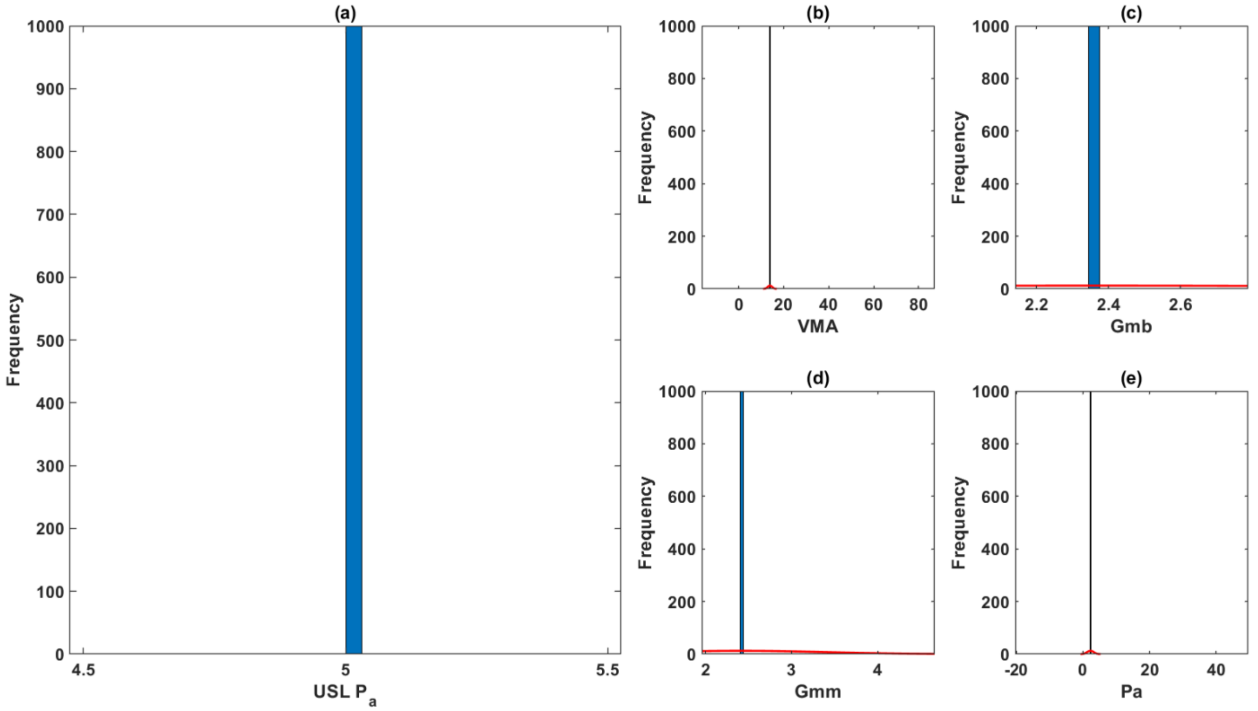


Figure 43. Sensitivity of VMA, Gmb, Gmm, and Pa parameters (sub plots b, c, d and e respectively) with respect to USL Pa.

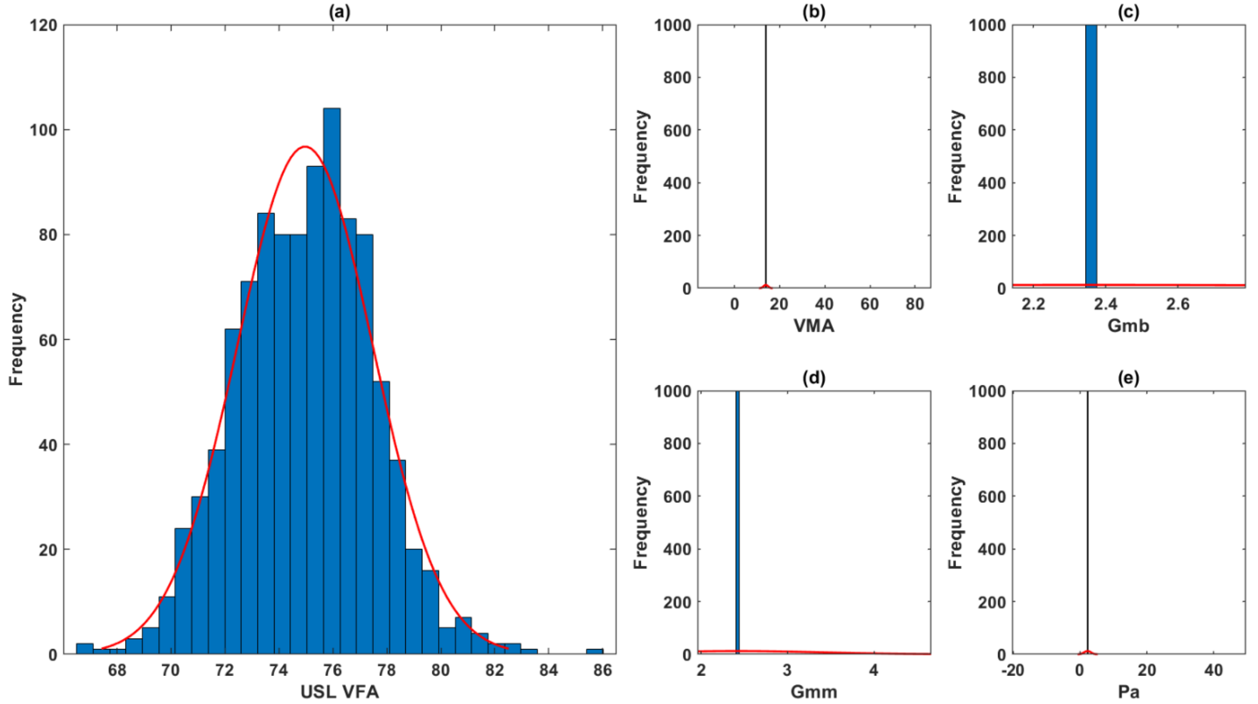


Figure 44. Sensitivity of VMA, Gmb, Gmm, and Pa parameters (sub plots b, c, d and e respectively) with respect to USL VFA.

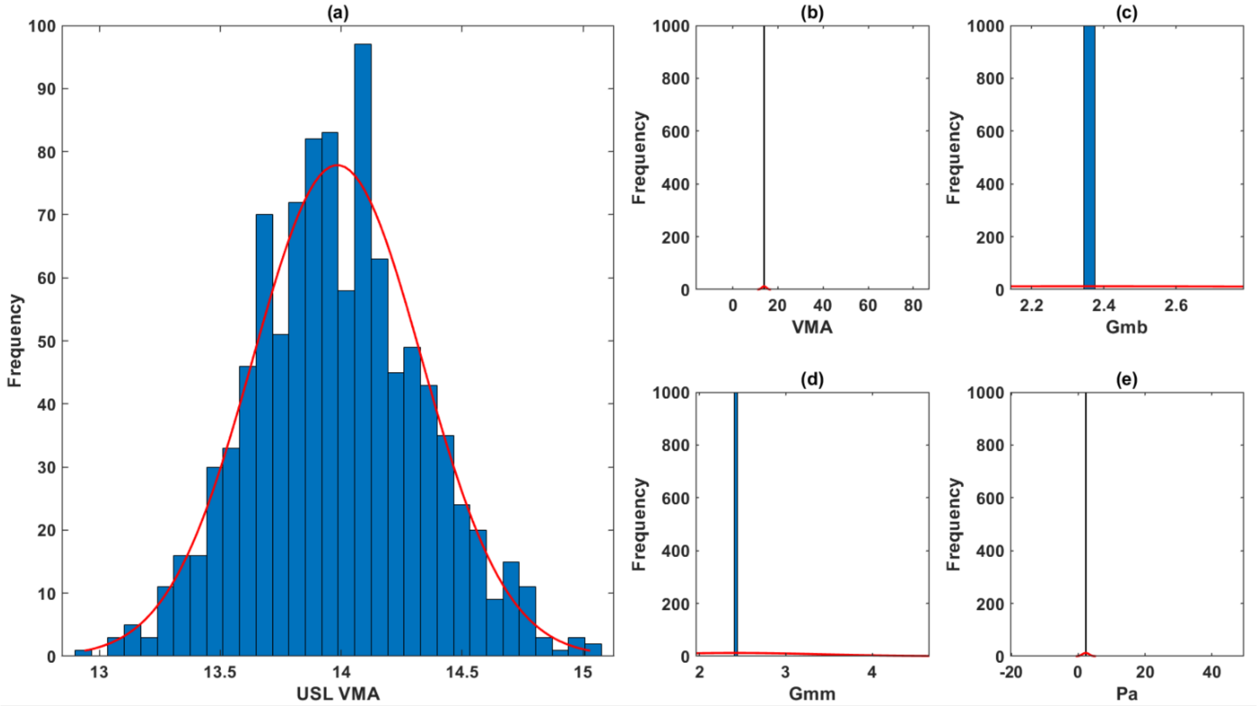


Figure 45. Sensitivity of VMA, Gmb, Gmm, and Pa parameters (sub plots b, c, d and e respectively) with respect to USL VMA.

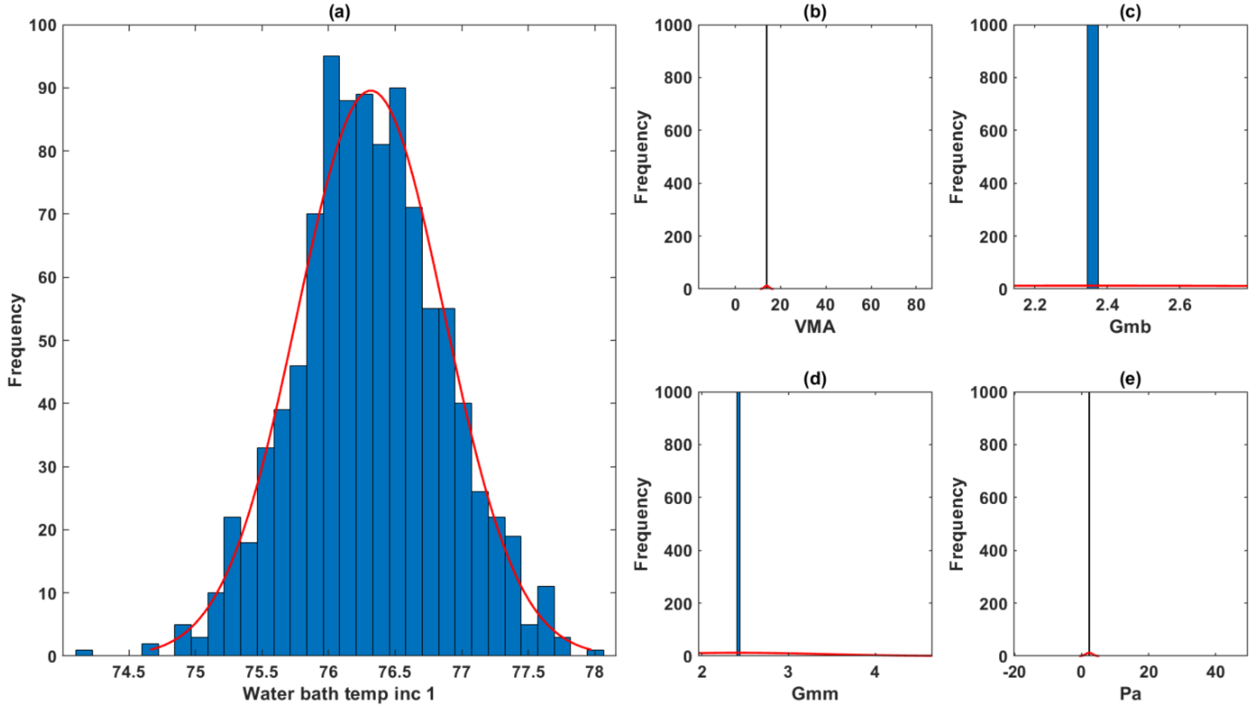


Figure 46. Sensitivity of VMA, Gmb, Gmm, and Pa parameters (sub plots b, c, d and e respectively) with respect to water bath temp inc 1.

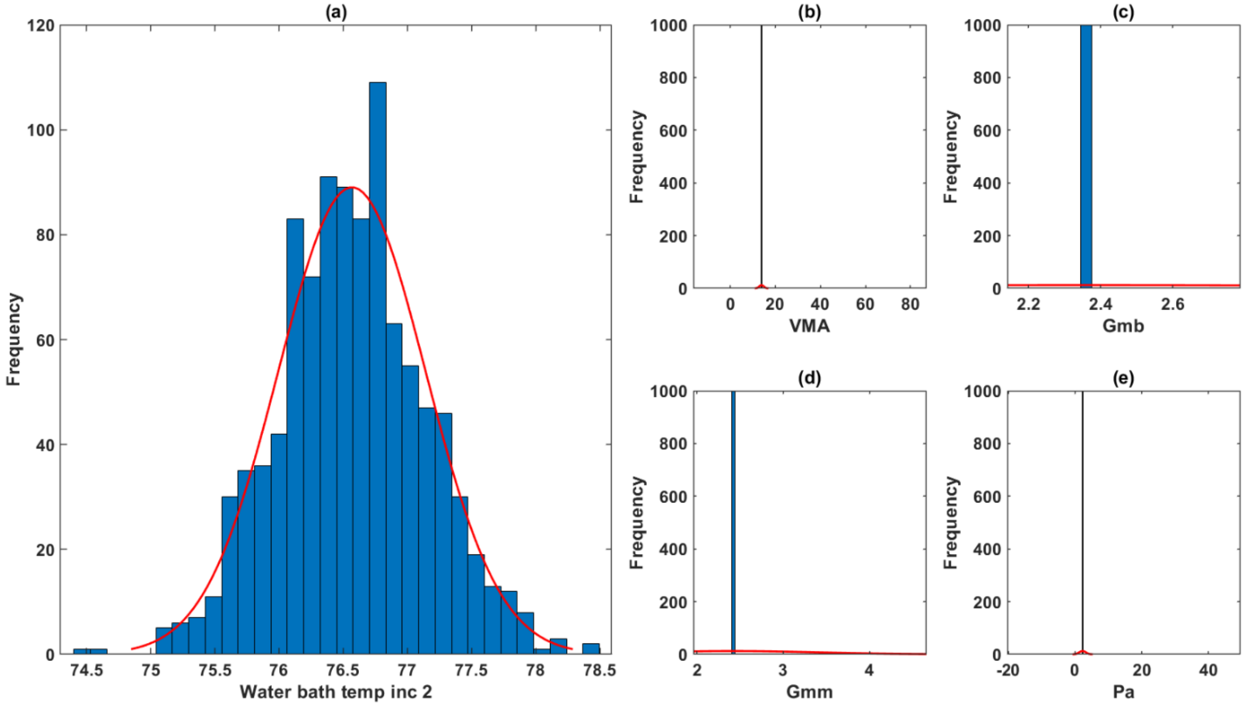


Figure 47. Sensitivity of VMA, Gmb, Gmm, and Pa parameters (sub plots b, c, d and e respectively) with respect to water bath temp inc 2.

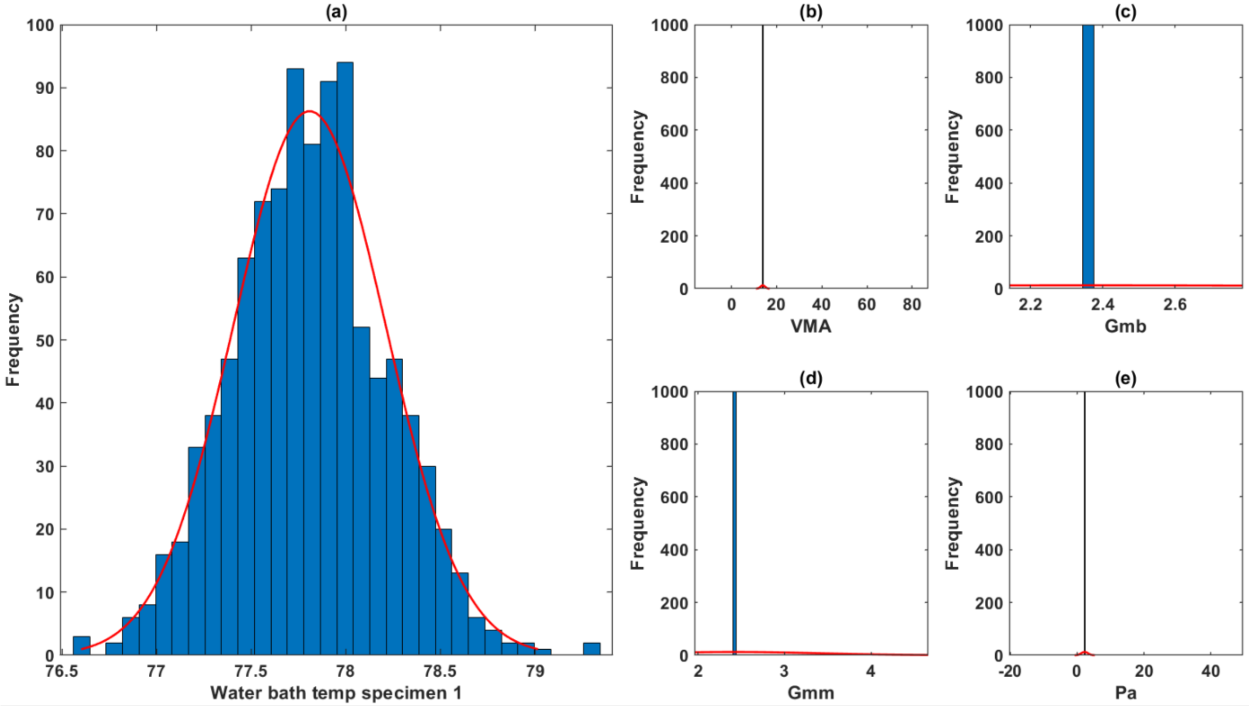


Figure 48. Sensitivity of VMA, Gmb, Gmm, and Pa parameters (sub plots b, c, d and e respectively) with respect to water bath temp specimen 1.

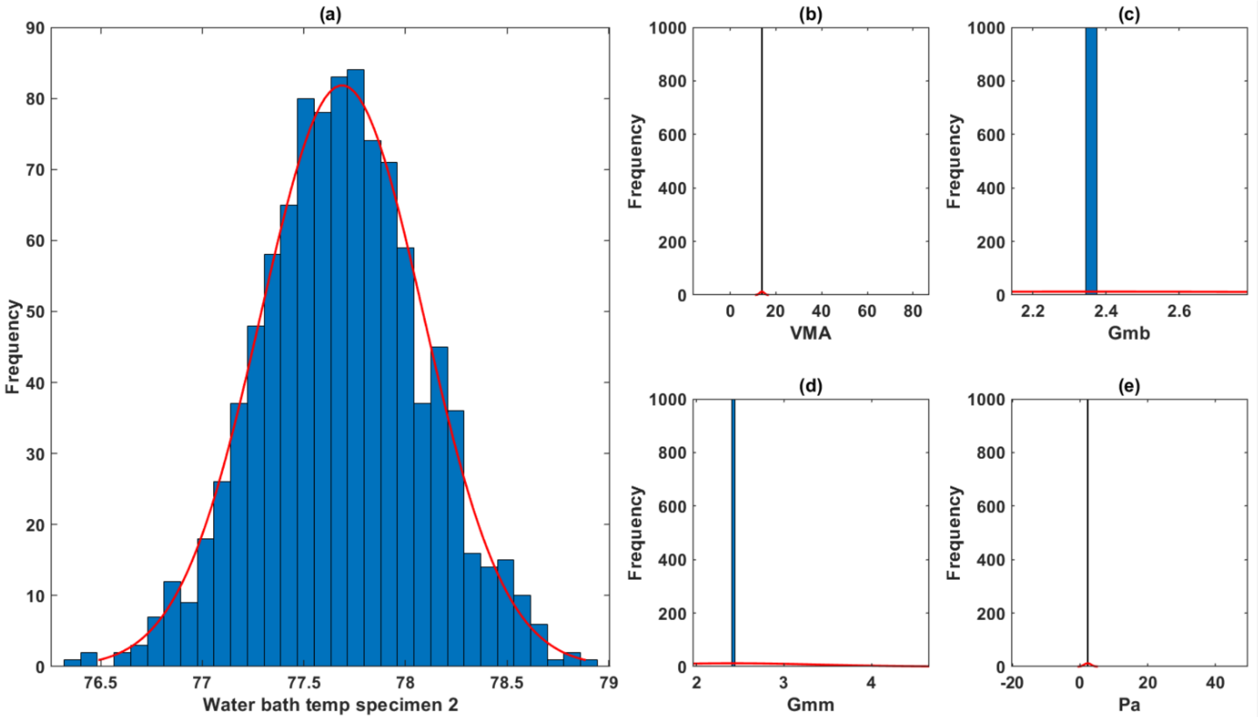


Figure 49. Sensitivity of VMA, Gmb, Gmm, and Pa parameters (sub plots b, c, d and e respectively) with respect to water bath temp specimen 2.

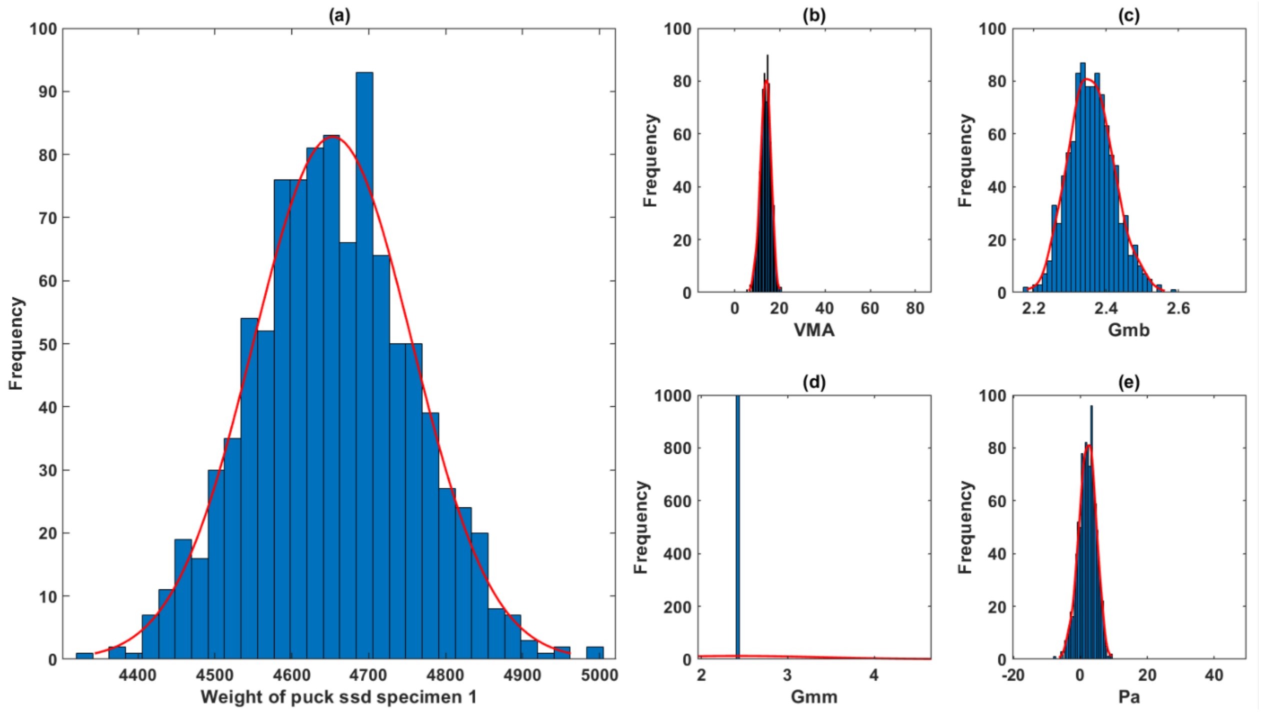


Figure 50. Sensitivity of VMA, Gmb, Gmm, and Pa parameters (sub plots b, c, d and e respectively) with respect to weight of puck ssd specimen 1.

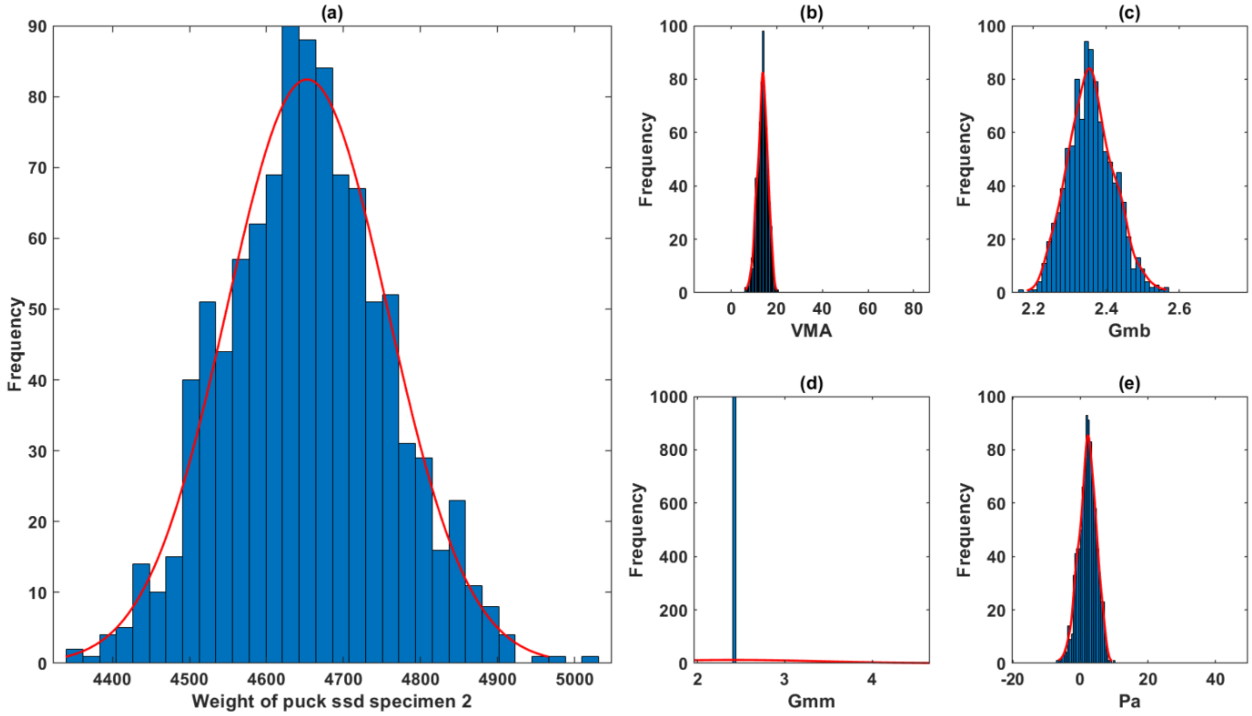


Figure 51. Sensitivity of VMA, Gmb, Gmm, and Pa parameters (sub plots b, c, d and e respectively) with respect to weight of puck ssd specimen 2.