Calculation of Measurement Uncertainty for Stiffness Modulus of Asphalt Mixture

Mieczysław Słowik and Mikołaj Bartkowiak

Institute of Civil Engineering, Poznan University of Technology, Poznań 60-965, Poland

Abstract: Asphalt mixture is a highly heterogeneous material, which is one of the reasons for high measurements uncertainty when subjected to tests. The results of such tests are often unreliable, which may lead to making bad professional judgments. They can be avoided by carrying out reliable analyses of measurement uncertainty adequate for the research methods used and conducted before the actual research is done. This paper presents the calculation of measurements uncertainty using as an example—the determination of the stiffness modulus of the asphalt mixture, which, in turn, was accomplished using the indirect tension method. The paper also shows the employment of the basic methods of statistical analysis, such as testing two mean values and conformity tests. Essential concepts in measurements uncertainty have been compiled and the determination of the stiffness module parameters are discussed. It has been demonstrated that the biggest source of error in the stiffness modulus measuring process is the displacement measure. The aim of the research was to find the measurement uncertainty for stiffness modulus by an indirect tensile test and the presentation of examples of the used statistical methods.

Key words: Measurement uncertainty, asphalt mixture, pooled experimental standard deviation, normality tests, indirect tensile test, stiffness modulus.

1. Introduction

Measurements uncertainty is an important indicator that enables the evaluation of measurements' credibility [1-3]. As claimed in Ref. [1], the result of the measurement is only an approximation of the value of the measured property and, therefore, every measurement result should be accompanied by uncertainty, which results from its chance variation. Results should not be presented as a single number but similarly to a continuous random variable so as an interval called the confidence interval. Thus, with specified probability that is assigned to this interval, it can be assumed that the resulting measurement value is contained within this interval, providing that the measurement and the accuracy analysis have been properly conducted. Calculating measurement uncertainty is related to such notions as tolerance and

required measurability. Tolerance is a denominate number that describes the difference between the maximal and minimal values of a certain property. It determines the interval, in which the real values of the specified property for the individual exemplars of the produced items should be contained. If the result of the measurement is to be useful [1], acceptable measurement uncertainty has to be established and that is understood as a limiting value at which the result is still a subject to evaluation or judgment [1]. Thus it may be justifiable to reduce the interval of measurement uncertainty. However, this will cause an increase in costs understood, for example, as financial means or time spent on carrying out the test. Therefore, an acceptable required measurements uncertainty has to be determined rationally. Such a determination is described in Refs. [4, 5] and called the required measurability. This required measurability has to be determined before attempting to indicate the value of the tested property. Arendarski [1] wrote that, for the measurement of machine parts, the most often assumed

---

Corresponding author: Mikołaj Bartkowiak, M.Sc.,
research fields: asphalt mixture, stiffness modulus and
measurement uncertainty. E-mail:
mikolaj.bartkowiak@put.poznan.pl.

required measurability should amount to $0.1T$, where T is measurement tolerance. In civil engineering area, especially in road engineering, reaching the required measurability is very difficult or almost impossible. This is due to the fact that such a material as asphalt mixture is characterized by much greater variability of its physical properties than, for example, steel and that, in turn, is caused by its greater heterogeneity.

An important question which has to be answered prior to any tests is that of a sample size n for a given determination. According to Refs. [3, 6], a sample may be considered:

- very small, when $n \leq 10$;
- small, when $10 < n \leq 30$;
- large, when $n > 30$.

A larger sample size allows to obtain better estimation of the standard variation for the results of the determination, but results in costs increase. According to Ref. [1], a standard deviation is a theoretical parameter whose credible value can be estimated on a sufficiently large ($n > 30$) series of measurements of a certain magnitude value. It is assumed [1] that a large series of measurements can be recommended when the testing aim is to indicate the so-called PESD (pooled experimental standard deviation), which would be later regarded as the known standard uncertainty of a single measurement taken in the same conditions. Designing the pooled experimental standard deviation for every employed testing procedure appears to be a compromise between the need to properly estimate a parameter and a cost of the testing itself, and it also allows for the estimation of required measurability for a given testing procedure.

There are two cases possible when calculating expanded uncertainty for an average of n measurements [1]:

- The variation of measurements distribution is known but equals the pooled standard uncertainty, obtained on the basis of a large series of earlier measurements (PESD);
- The variation of measurements distribution is

unknown and, therefore, the coverage factor is indicated employing t -student distribution.

If the standard uncertainty is known and we perform a series n_1 of new measurements using the same measuring method, in the same conditions but for a different material, the standard uncertainty $u(\bar{X})$ is calculated with the use of the following formula [1]:

$$u(\bar{X}) = \frac{\sigma}{\sqrt{n_1}} = \frac{u_p(X)}{\sqrt{n_1}}$$

This dependence allows for decreasing of standard uncertainty. If four tests are performed this decreases the standard uncertainty twofold. For the threefold decrease, nine tests have to be performed. It has to be pointed out that Arendarski [1] clearly states that these determinations should be new, excluding the possibility to use the same results for indicating the standard uncertainty and decreasing it as is the case with calculating measurements uncertainty using student's t -distribution [7].

2. Stiffness Modulus

Pavement design and evaluation for purposes of construction and rehabilitation require careful evaluation of a number of factors. Assuredly, material properties are one of the most significant factors in the design of pavement. The stiffness modulus and Poisson's ratio are two fundamental material properties that can be used to predict the response of the material to a repeated impulse or moving loads, such as those imposed by vehicle tires on the pavement surface. These two properties are commonly used in mechanistic models for designing pavements. If we assume that the material is isotropic and linear elastic, then it can be described using the two above mentioned properties which can be determined with the use of a material test. Other elastic constants can be calculated on the basis of these two material properties, using equations from the theory of elasticity.

The most popular methods for determining the stiffness modulus are:

- test applying indirect tension to cylindrical

specimens (IT-CY);

- two point bending test on trapezoidal specimens (2PB-TR);
- four point bending test on prismatic specimens (4PB-PR).

These different tests have different tension states, which are applied to the samples, hence the values obtained from these tests are also different. The stiffness modulus of asphalt concrete determined from bending (4PB-PR) and also from an indirect tensile test are usually smaller than those obtained from a compression test [8].

3. Description of Tested Material

The tested material was asphalt concrete for the binder course with granulation of 0~16 mm formed with the use of basalt aggregate. The target binder content was 4.6% (m/m) and the air void content was 4.3%, 35/50 penetration grade bitumen was used as the binder. Stiffness modulus determinations were conducted for three Marshall specimens made of the same mixture. For each sample, six stiffness modulus determinations were conducted using the indirect tension test. For the purpose of a statistical analysis, adjusted stiffness modulus values of diameters were taken, resulting in 36 values of the obtained stiffness modulus.

4. Indirect Tensile Test

The indirect tensile test was developed simultaneously but independently in Brazil and Japan. The test has been used to determine the stiffness modulus of asphalt mixture using a cylindrical sample. The cylindrical samples should have thickness of 30~75 mm. The thickness of the sample cannot be too small because of the three-dimensional effects and cannot be too high because the compaction of sample could vary. The testing system (Figs. 1 and 2) includes a loading apparatus, displacement measurement devices (including LVDTs (linear variable differential transformers)) and a data recording system. Over the

years, two basic types of loading have been used in laboratory tests: controlled-strain and controlled stress. In this report, we used the controlled-strain method. Repetitive load is applied to induce a constant displacement. The displacement is $7 \pm 2 \mu\text{m}$ for a sample with a nominal diameter of 150 mm and $5 \pm 2 \mu\text{m}$ for a sample with a nominal diameter of 100 mm. The repetitive load and displacement are measured and substituted into the formula [9]:

$$S_m = \frac{F \cdot (v + 0.27)}{z \cdot h}$$

where:

S_m —measured stiffness modulus (MPa);

F —applied vertical load (N);

v —Poisson's ratio;

z —horizontal displacement (mm);

h —average thickness of sample (mm).

The measured stiffness modulus should be modified using the following formula [9]:

$$S'_m = S_m \cdot (1 - 0.322) \cdot (\log S_m - 1.82) \cdot (0.60 - k)$$

where:

S'_m —adjusted stiffness modulus (MPa);

k —load area factor.

The load area factor (k) refers to the area under the haversine load curve (between the beginning of the pulse and the maximum load) divided by the area of the enclosing rectangle (which is the rise-time multiplied by the maximum load). The area under the load pulse is the amount of energy, which is applied to a specimen, and should be the same for each load. The load area factor is used to correct the value of the stiffness modulus (according to the formula) if the load curve is different than expected. The ITT (indirect tensile test) has been identified as an economic and practical means for determining the stiffness modulus of an asphalt mixture that the attractiveness of the test lies in the possibility to use core samples, instead of the samples with more complex shapes which are often associated with other stiffness modulus tests (bending beam). Moreover, we can also use the laboratory compacted specimens, like the Marshall ones.

Due to the geometry of the measuring system, the stress pattern in a sample is highly nonlinear and biaxial. The relationship between the vertically applied load and horizontally measured deformation is not straightforward (as would be the case, for example, in a uniaxial test). The simplest approach to relate these two quantities is to assume that the material is homogeneous, isotropic and linear elastic [10]. According to Ref. [8], the test simulates the state of stress in the lower position of the asphalt layers.

Many scientific papers [8, 11] describe repeatability of designing stiffness modulus results as low. Scattering of the results obtained from the indirect tensile test can be estimated from 5% to 10% [8]. This is one of the lowest variabilities that can be achieved for many ways of determining the stiffness modulus. In addition, the value of the resulting stiffness modulus depends on other testing parameters, such as, for example, a rise time [12]. A multitude of testing procedures and a possibility to change testing parameters by an operator implies small variations in the state of tensions that are present in samples, as well as other important testing conditions. However, practice shows [8] that scattering of the results caused by varying testing conditions is lower than scattering caused by material heterogeneity and instrumental errors. Therefore, to ensure that the results of stiffness modulus determinations are useful, an analysis of

Fig. 1 Sample ready for testing.

Fig. 2 General view of the test bed.

Table 1 Stiffness modulus testing parameters according to IT-CY.

Test conditions	Value	Tolerance according to European Standard	Resolution of the measuring instruments
Test temperature (°C)	10.0	±0.5	0.1
Target specimen diameter (mm)	101.6	±0.1	0.01
Target specimen thickness (mm)	63.5	±0.1	0.01
Poisson's ratio (assumed)	0.35	-	-
Target rise-time (loading time) (ms)	124	±4	1
Target horizontal displacement (µm)	5.0	±2	1 (European Standard); 0.1 (measuring device)
Load area factor	0.6	±0.1	0.01
Pulse repetition time (s)	3.0	±0.1	0.001
Vertical force (N)	-	±2%	10
Number of conditioning pulses (pcs.)	10	-	-
Number of test pulses (pcs.)	10	-	-

measurement uncertainty for this determination has been carried out. Stiffness modulus determinations have been conducted according to the procedure described in Ref. [9]. The testing parameters used are described in Table 1. The tests have been done on the Marshall samples. The load curve had a haversine shape.

5. Discarding of Results

Three values were rejected on the basis of a test for the outliers. These values were rejected by the Hampel elimination test which is described in Ref. [2]. One of the values was rejected because of the inadequate testing conditions (load area factor and rise-time). Altogether four values were rejected. It can be seen (in Table 2) that all of the discarded values are characterized by abnormal conditions of the assay, such as a non-valid shape of the displacement curve (Fig. 3), non-valid values of the horizontal displacement, nonstandard values of the load area factor and rise-time. Three outliers obtained from the Hampel test are characterized by the values of the

horizontal displacement which are not included in the range $z = 5.0 \pm 0.2 \mu\text{m}$ in 10°C . Horizontal displacements recorded for the other values are contained within this interval.

6. Results and Discussion

6.1 Results Presentation

For further analysis, 32 values were used. The obtained values are shown on the chart presented in Fig. 4. Average values, standard deviations and the ranges for the obtained modulus before and after excluding outliers are presented in Tables 3 and 4.

6.2 Verification of the Normality of Distribution

For the proper assessment of the measurement uncertainty, it is required to determine which probability distribution the tested property has. Preliminary information on a possible distribution of the examined population can be acquired employing a histogram. The intervals of the histogram should have the same width and its number k should be less than $1/4$ the sample size ($k \leq n/4$).

Table 2 Rejected values (averages of five loads).

Rejected value (MPa)	Horizontal displacement (μm)	Load area factor	Rise-time (ms)
18,908	4.5	0.61	124.0
16,686	4.9	0.64	131.4
19,285	4.6	0.61	122.8
18,682	4.5	0.61	124.2

Fig. 3 Abnormal shape of the displacement curve during one of the determination.

Fig. 4 Adjusted stiffness modulus for three specimens with rejected values.

Table 3 Results before discarding.

Sample	Number of results	Mean (MPa)	Standard deviation (MPa)	Range (MPa)
S1	12	16,689	1,328	3,108
S2	12	16,279	932	3,510
S3	12	15,713	735	2,580
All	36	16,190	1,063	4,946

Table 4 Results after discarding.

Sample	Number of results	Mean (MPa)	Standard deviation (MPa)	Range (MPa)
S1	9	16,006	679	2,278
S2	11	16,060	571	1,517
S3	12	15,713	735	2,580
All	32	15,915	665	3,154

Fig. 5 Histogram of the stiffness modulus for $k = 7$.

Based on the presented histogram (Fig. 5), it was not possible to unequivocally accept or reject an

assumption concerning normality of the distribution. Eventually, the distribution of the examined population

has been determined by the tests called “distribution tests”. A distribution test refers to a test which verifies a simple or complex hypothesis concerning conformance between the distribution of a set of values in a sample and a theoretical distribution [13]. The most well-known distribution tests are: the χ^2 -Pearson’s, the λ -Kolmogorov’s and the Shapiro-Wilk’s tests. The latter is designed for testing normality of the distribution. Such tests as the χ^2 -Pearson’s and the λ -Kolmogorov’s are more universal, but they require a much greater number of samples. A variable value distribution analysis should be performed for adequate number of resulting values. It is assumed that, for determining the distribution, there should be at least 15 values, however, it is recommended to perform a distribution test on a large sample ($n > 30$). If there are no other premises, it can be assumed that the distribution of populace will be a normal distribution (Gauss’ distribution), because it is the most commonly occurring probability distribution. It is presumed that the population will have normal distribution, especially if the value of the examined property is determined by independent factors. To check the normality of distribution, the Shapiro-Wilk test was chosen, because it is the most powerful of the normality tests. The test was conducted according to the rules described in Ref. [14], calculating the random variable W according to the formula:

$$W = \frac{\left(\sum_{i=1}^{n/2} a_i(n) \cdot X_{n-i+1} - X_i \right)^2}{\sum_{j=1}^n \left(X_j - \bar{X} \right)^2}$$

where:

$a_i(n)$ —constants dependent on a sample size;

X —value of the variable;

$\bar{X}$ —arithmetic mean.

For the sample size $n = 32$, the calculated statistics value is $W = 0.986$, for which the probability value is p -value = 92.9%. If the probability value p -value is greater than 5%, as in this case, there is no reason to reject the hypothesis that the distribution of the stiffness modulus for the analyzed asphalt mixture

shows normal distribution. Moreover, the additionally employed Lilliefors test confirmed normality of the distribution.

7. Uncertainty Assessment

After performing calculations confirming normality of the distribution, it can be assumed that the average $m = 15,915$ MPa and the standard deviation $\sigma = 665$ MPa are reliable estimators of the analyzed sample. The analyzed sample consists of $n = 32$ elements of the populace, so the standard deviation can be taken as the pooled experimental standard deviation $\sigma(\text{PESD}) = 665$ MPa (RSD (relative standard deviation) = 4.18% for the determination of the stiffness modulus at 10 °C. The permissible difference between resulting parallel determinations of the stiffness modulus, for the determined PESD should not exceed $2 \cdot \text{PESD} = 1,330$ MPa ($p = 95\%$). The tested asphalt mixture has the value of the stiffness modulus (assuming coverage factor $k = 2$):

$$S_m = 15,915 \pm 1,330 \text{ MPa}$$

$$U(S_m) = 1,330 \text{ MPa (8.36\%)}$$

Assuming that the general populace shows normal distribution, the measurement uncertainty has been calculated for each of the three samples, employing the t -student distribution for the calculation. The results are presented in the Table 5.

8. Test for Two Averages

For the obtained measurement results, it has been checked if the difference between the results is statistically significant. If this was the case, it would indicate that the properties of a single sample have a significant influence on the stiffness modulus determination for the asphalt mixture, which could indicate that the testing was not properly conducted. In

Table 5 Results of stiffness modulus determination.

Specimens	Results
Specimen 1	16,006 ± 522 MPa
Specimen 2	16,060 ± 384 MPa
Specimen 3	15,713 ± 467 MPa
All values	15,915 ± 240 MPa

order to test this hypothesis, the result for the second sample, for which an average value of the stiffness modulus determination is the largest, has been compared with the third sample, for which the corresponding value is the smallest. The results have been compared using a test for two averages that is described in Ref. [13] assuming that the difference between the results is statistically insignificant. In order to verify this hypothesis, a test based on statistics has been employed:

$$U = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{n_1 \cdot s_1^2 + n_2 \cdot s_2^2}{n_1 + n_2 - 2} \cdot \left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{16,060 - 15,713}{\sqrt{\frac{11 \times 544.54^2 + 12 \times 703.98^2}{11 + 12 - 2} \times \left(\frac{1}{11} + \frac{1}{12}\right)}} = 1,258$$

where:

$\bar{X}_1, \bar{X}_2$ —arithmetic means of values;

n_1, n_2 —sizes of samples;

s_1, s_2 —standard deviations calculated using n method.

Comparing the calculated statistic value $U = 1,258$ with the value of the t -students variable with the significance level $\alpha = 0.05$ that amounts to $t_\alpha = 2,080$, it can be stated that it is not included in the critical region which is determined in this case by inequality $t \geq t_\alpha$. Therefore, there is no evidence allowing us to reject the hypothesis of equality of the compared average stiffness moduli. The assumed hypothesis could be discarded with a probability equal to 78%.

9. Calculating Uncertainty with Method B

The most frequently method used to calculate uncertainty is a statistical method called Method A. But if statistical analyses are too burdensome or not cost-efficient, the Method B which is considered equivalent can be used, instead. This term refers to all other than statistical methods for acquiring information about uncertainty [1]. In order to calculate uncertainty utilizing Method B, one can utilize calibration uncertainties of measuring instruments. Determination

of the stiffness modulus is an indirect measurement and it is assumed that directly measured quantities are mutually independent. So, in this case, the standard complex uncertainty can be calculated according to the method described in Ref. [1]. It has to be emphasized, however, that the resolutions of the measuring instruments have been substituted for the standard uncertainties of the property measurements taken directly. The calculations performed with an average of all determinations are presented below:

$$\frac{U(S_m)}{S_m} = \sqrt{\left(\frac{U(F)}{F}\right)^2 + \left(\frac{U(z)}{z}\right)^2 + \left(\frac{U(h)}{h}\right)^2} = \sqrt{\left(\frac{10}{8,150}\right)^2 + \left(\frac{0.0001}{0.005}\right)^2 + \left(\frac{0.1}{63.5}\right)^2} = \sqrt{0.12^2 + 2.00^2 + 0.16^2} = 2.01\%$$

Hence, the expanded uncertainty $U(S_m)$ amounts to:

$$U(S_m) = 0.0201 \times 15,915 \times 2 = 640 \text{ MPa}$$

$$S_m = 15,915 \pm 640 \text{ MPa}$$

Taking into account the done calculations, it can be noted that the most important component factor of the measurement uncertainty is the displacement measurement. When registering displacement values approaching 5 μm , the way in which the sample is secured is of great importance. According to the authors' experience, the asymmetrical or non-centric fastening with so small a displacement can lead to erroneous results. It should also be noted that reducing values of the target displacement, especially with the tests performed in low temperatures, could result in a significant rise of uncertainty and obtaining useful results which could be rendered difficult.

10. Conclusions

The measurement uncertainty distribution for the stiffness modulus determined at 10 °C is a normal distribution. The pooled experimental standard deviation determined for an average of five loads amounts to $\sigma(\text{PESD}) = 665 \text{ MPa}$ (RSD = 4.18%). The value of measurement uncertainty calculated using a formula for the standard complex uncertainty

($U(S_m) = 4.02\%$) is more than twice as small as the value of uncertainty determined using a statistical method ($U(S_m) = 8.36\%$). It is caused by the fact that the calibration uncertainties for the measuring instruments, instead of the standard uncertainties, were used for the calculations employing standard complex uncertainty. Moreover, the standard complex uncertainty calculated this way cannot take into account other kinds of errors, like random errors. For the purpose of the calculations, the LVDT sensor resolution amounting to $0.1 \mu\text{m}$ has been assumed, but if normative measurement accuracy amounting to $1 \mu\text{m}$ were to be assumed, the extended uncertainty would be 10 times larger, amounting to $U(S_m) = 40.0\%$. For such a large measurement uncertainty, the results obtained would be useless. Based on the authors' experience, it can be said that, for such small displacement (in the range of $5 \mu\text{m}$), securing of the sample and a proper contact between the sample surface and the LVDT sensor is of great importance. The calculated measurement uncertainty for indirect tensile tests along with the other proposals for conducting the test may be useful for laboratories which analyze asphalt mixtures using this method. The statistical methods presented in the paper are very useful for the analysis of the test results.

References

- [1] Arendarski, J. 2013. *Uncertainty in Measurements*. Warszawa: Publishing House of Warsaw University of Technology. (in Polish)
- [2] Słowik, M. 2013. *Selected Problems Concerning Viscoelasticity of Road Modified Bitumens Containing SBS Elastomer*. Poznań: Publishing House of Poznan University of Technology. (in Polish)
- [3] Słowik, M. 2010. "Principles Underlying Calculation of Test Uncertainty of Bituminous Binder's Selected Properties." In *Archives of Institute of Civil Engineering 7/2010*. Poznań: Publishing House of Poznan University of Technology, 129-44. (in Polish)
- [4] ISO (International Standardization Organization). 2000. *ISO 9001:2000, Quality Management Systems—Requirements*. Geneva: ISO.
- [5] European Co-operation for Accreditation. 2003. *Guidelines on the Expression of Uncertainty in Quantitative Testing*. Paris: European Co-operation for Accreditation.
- [6] Szydlowski, H. 2001. *Uncertainty in Measurements, International Standards in Practice*. Poznań: Publishing House of Adam Mickiewicz University. (in Polish)
- [7] Greń, J. 1984. *Mathematical Statistics: Models and Tasks*. Warszawa: Państwowe Wydawnictwo Naukowe. (in Polish)
- [8] Barksdale, R., Alba, J., Khsola, P. N., Kim, R., Lambe, P. C., and Rahman, M. S. 1997. *Laboratory Determination of Resilient Modulus for Flexible Pavement Design*. NCHRP (National Cooperative Highway Research Council) Web Document 14. Washington, D.C.: NCHRP.
- [9] European Standards. *EN 12697-26:2007, Bituminous Mixtures—Test Methods for Hot Mix Asphalt—Part 26: Stiffness*. Brussels: European Standards.
- [10] Collop, C., and Airey, G. 2003. "Determination of Viscoelastic Properties from the Indirect Tensile Stiffness Modulus (ITSM) Test." Presented at 6th RILIEM Symposium, Zurich, Switzerland.
- [11] Carbonneau, X., Le Gal, Y., and Bense, P. 2003. "Evaluation of the Indirect Tensile Stiffness Modulus Test." Presented at 6th RILIEM Symposium, Zurich, Switzerland.
- [12] Shalaby, A., Liske, T., and Kavussi, A. 2004. *Comparing Back-Calculated and Laboratory Resilient Moduli of Bituminous Paving Mixtures*. Ottawa: NRC Research Press.
- [13] Kryszicki, W., and Włodarski W. 2007. *Probability and Mathematical Statistics in Tasks*. Warszawa: Wydawnictwo Naukowe PWN. (in Polish)
- [14] Razali, N. M., and Wah, Y. B. 2011. "Power Comparisons of Shapiro-Wilk, Kolmogorov-Smirnov, Lilliefors and Anderson-Darling Test." *Journal of Statistical Modeling and Analytics* 2 (1): 21-33.