2022 IMEKO TC-4 International Conference on
Metrology for Archaeology and Cultural Heritage
University of Calabria, Italy, October 19-21, 2022

Evaluation of Intermediate Precision in Thermographic Measurements for Monitoring the Conservation State of Metallic Discoveries

Rosario Morello[1], Claudio De Capua[1], Laura Fabbiano[2]

[1] *Dept. DIIES, University Meditereanea of Reggio Calabria, Italy, rosario.morello@unirc.it*
[2] *DMMM, Politecnico di Bari, University, Bari, Italy, laura.fabbiano@poliba.it*

***Abstract* – This paper deals with the metrological issue of the reliability in measurements. The evaluation of intermediate precision is discussed to assure repeatable data over time. As an example, the monitoring process of archaeological or historic discoveries needs that data coming from different measurements and performed in different time instants are comparable and traceable. Thermography is used here to evaluate the conservation state of metallic artefacts and their decay over time. Data on the integrity of an ancient iron oil lamp have been collected at different points in time. Intermediate precision has been computed to evaluate the data reliability so to guarantee the comparability of data. The proposed approach can allow getting accurate and consistent information on the discovery decay over time. Intermediate precision evaluation provides quantitative information about the effectiveness of the monitoring process and the reliability of its results.**

***Keywords* – thermography; intermediate precision; conservation state; metal.**

I. INTRODUCTION

Thermographic images are today used in several application fields. In this paper, thermography is considered to investigate the conservation state of archaeological discoveries over time. Measurement reliability and repeatability are two important issues to be assured in order to make consistent decisions on data collected in different time instants. Often, measurement data are used to assess the state of a discovery in a specific time instant. The comparison among time series provides further information on possible degradation processes. However, when monitoring activities are carried out even the evaluation of intermediate precision is important to guarantee repeatable data over time. As an example, the monitoring process of archaeological or historic discoveries needs that data coming from different measurements and performed in different time instants are comparable and traceable. Thermography is used here to evaluate the conservation state of metallic artefacts and their decay over time. The issues of measurement repeatability and intermediate precision are here described in order to alert operators about the consistency of decisions based on time series data.

The Guide to the *Expression of Uncertainty in Measurement (GUM)* provides the approach to evaluate uncertainty when computations involve measured quantity values, [1]. The uncertainty is evaluated by a standard deviation of a probability distribution, interpreted not only in frequentist manner (*type A evaluation*), but also as a degree of belief (*type B evaluation*), quantified by means of a subjective probability distribution. ISO 5725 reports the definition of "Trueness", [2], it refers to the closeness of agreement between the arithmetic mean of a large number of test results and the true or accepted reference value. Otherwise, the term "Precision" refers to the closeness of agreement between different test results. These two contributions allow getting quantitative information on the accuracy of the measurement. The bias has systematic effects on measurements and its contribution can be estimated by using reference values and then subsequently corrected. Instead, the contributions to uncertainty in terms of standard deviations are due to random effects and they have to be considered and evaluated properly. Particular attention is here given to the random contribution associated with the reproducibility condition. Consequently, during the measurement process, the operator has to pay attention to the measurement procedure in order to assure the requirements of repeatability and intermediate precision.

The paper is organized as follows. In Section II, the theory of measurement reliability and measurement issues are described. Section III reports the experimental results. Finally, considerations and conclusions are outlined in Section IV.

II. INTERMEDIATE PRECISION IN THERMOGRAPHY

Depending on measurement conditions, precision can be evaluated at 3 different levels [3]. In detail, we can assert that:

1. to assure repeatability in repeated measurements it is

essential to assure a set of conditions that includes the same measurement procedure, same operator, same measuring system, same operating conditions and same location, and replicate measurements on the same or similar objects over a short period of time (VIM3 2.20);
2. to assure intermediate precision of measurement it is essential to assure a set of conditions that includes the same measurement procedure, same location, and replicate measurements on the same or similar objects over an extended period of time, but it may include other conditions involving changes (VIM3 2.22);
3. to assure reproducibility of measurement it is essential to assure a set of conditions that includes different locations, operators, measuring systems, and replicate measurements on the same or similar objects (VIM3 2.25);

As a consequence, repeatability is evaluated in a short time interval and it provides information on the contribution to precision evaluated under the same conditions (see the first point of the previous bullet list). Otherwise, intermediate precision provides information on the accuracy contribution of intermediate measurements made in the intermediate precision conditions (see the second point of the previous bullet list). Reproducibility refers to measurements performed in different locations with different measurement systems and by different operators assuring reproducibility conditions (see the third point of the previous bullet list).

In detail, we can note that intermediate precision is referred to measurements on the same measurand in different conditions but within the same laboratory. This occurs when repeated measurements on the same measurand are carried out for monitoring purpose or to check its value in different time instants. Reproducibility contribution comes out from measurements of the same measurand performed in different laboratories with similar conditions. It occurs in collaborative laboratories with the aim to prove the reproducibility of a measurement process.

On the basis of the previous definitions, intermediate precision is proposed in the following for monitoring the conservation state of metallic discoveries by using thermography. The same unit under test is investigated in different time instants at the same laboratory to evaluate changes in its conservation state. Consequently, we define measurements under conditions of intermediate precision those measurements carried out on the same discovery on different days. The associated uncertainty contributions are evaluated according to the *type A* procedure in the GUM. By assuming that *N* is the number of repeated measurements carried out under repeatability conditions, the uncertainty to be associated with the measurement is given by the best estimate of the standard deviation:

$$S_{\bar{x}} = \frac{\sqrt{\sum_{i=1}^{N} \frac{(x_i - \bar{x})^2}{N-1}}}{\sqrt{N}} = u_{intra} \quad (1)$$

This is the uncertainty contribution referred to the measurement performed in a specific time instant. In other words, this is the result of uncertainty associated to the conservation state assessment in a specific time instant. To evaluate the contribution to uncertainty due to measurements carried out in successive time instants we have to evaluate the uncertainty in conditions of intermediate precision. So if *M* is the number of replicated measurements, we obtain:

$$S'_{\bar{x}} = \frac{\sqrt{\sum_{j=1}^{M} (\bar{x}_j - \bar{\bar{x}})^2}}{\sqrt{M}} \quad (2)$$

where

$$\bar{\bar{x}} = \frac{1}{M} \sum_{j=1}^{M} \bar{x}_j \quad (3)$$

and

$$S''_{\bar{x}} = \frac{\sqrt{\frac{1}{M} \sum_{j=1}^{M} S_j^2}}{\sqrt{NM}} \quad (4)$$

where

$$S_j = \frac{\sqrt{\sum_{j=1}^{M} \sum_{i=1}^{N} \frac{(x_i - \bar{x}_j)^2}{N-1}}}{\sqrt{NM}} \quad (5)$$

The overall contribution to the uncertainty of the intermediate test is given by:

$$u_{inter} = \sqrt{S'^2_{\bar{x}} + S''^2_{\bar{x}}} \quad (6)$$

A. Measurement Procedure

Thermographic measurements are affected by several issues which require some precautions in order to guarantee the repeatability and reliability of measurement results, [4]-[9]. However, no standardized procedure has been defined at the moment. As a consequence, authors provide a brief guideline to be followed to avoid some common random and systematic errors during the measurement procedure. To this aim, it is important to distinguish some systematic errors concerning parameters settings from interference sources which could have a random effect on measurement results.

Primarily, operator has to pay attention to atmospheric parameters which have influence on thermographic measurements. In detail, relative humidity, environment temperature and reflected temperature are the main parameters which have to be set in the thermal camera settings utility. Measurements carried out at the *Advanced Thermography Center* of Reggio Calabria have shown that an underestimation of these parameters can be cause of errors of about $0.5°C$ in the measured absolute temperature. In addition, the object distance and the material emissivity can add significant errors in the evaluation of the measurand value. The operator must measure these parameters, keep them constant and set them in the thermal camera. In this way, the thermal camera converts the captured radiance pixel by pixel into temperature values compensating the influence of the previous parameters. The thermographic image is created by assigning to each temperature value a different colour shade.

During the measurement procedure, attention has been paid to other interference sources such as solar irradiation, external infrared sources, artificial lightning and background radiation. They can entail errors affecting the temperature measurements of about $1.5°C$. As a consequence, measurements have to be performed avoiding any external interference. Even the use of thermal conditioning systems for air conditioning must be avoided.

The best advice is to carry out the measurement procedure in a controlled environment, such as a laboratory.

B. Image Processing

The right evaluation of the measurand value and of the repeatability and intermediate precision associated to the measurement process needs to perform repeated measurements in the shortest possible time in order to evaluate the uncertainty contributions above reported. In fact, if measures are made in a brief time, it is possible to consider the measurand stable and constant. This assumption is basic for both the measurement processes carried out in different time instants to evaluate the conservation state of the discovery after months or years. We can assume that a short time is approximately equal to 1 second. So frame rates over 30 frames/s are suggested.

Finally, another important issue regards the choice of the Region of Interest (ROI). The image processing needs to define the pixels set or the image area to be processed. About the ROI choice, no rule exists. The only suggestion is to pay more attention to that areas which are more subjected to possible deterioration during time, having an early cracking process or more exposed to atmospheric agents. Depending on the case study, the best measurand is evaluated by considering the maximum temperature value in the ROI or the mean temperature associated to all pixels of the ROI. The maximum value allows evaluating the presence of cracks, non-uniformity of the material, thickness changes, delamination, oxidation or erosion processes. In this case, the operator must pay attention to the presence of defective pixels which can be cause of an altered maximum value which could be not representative of the true value of the measurand. Otherwise, the mean value allows to evaluate the occurrence of different responses in different time instants to the same external thermal excitation. This can be considered the effect of deterioration during time.

III. CASE STUDY AND EXPERIMENTAL RESULTS

The proposed case study refers to an ancient iron oil lamp, see Figure 1. The metallic artifact was discovered in the subsoil of a rural area. It is supposed that its origin is dated back to the last half of the 1800's. It was underground for several years before its discovery. The lamp looks in fair conditions, however, the exposure to high humidity levels has corroded its structure. In fact, the surface is covered by rust. No further surface defects are visible at naked eye. Table I reports the main characteristics of the examined artifact.

Fig. 1. The case study: an ancient iron oil lamp.

Table 1. Artifact Features

	Value/Description	*Measurement Unit*
Use	oil lamp for domestic lighting	-
Material	iron	-
Height	27	cm
Diameter	10.5	cm
Mass	495	gr
Main body layer tickness	0.5-1	mm
Handle tickness	3	mm
Emissivity	1	-
Discovery year	2010	-
Origin Age	1850-1940	-
Conservation State	visible rust formation	-

The measurement instrumentation includes a thermal camera FLIR x8400sc. The climatic conditions of the laboratory have been kept constant during experimentation. All thermal camera settings have been measured and set in the camera utility. The environment temperature was equal to $27 \pm 0.5\,°C$ with a relative humidity of 57 %. The estimate reflected temperature was equal to $27.2°C$, and the emissivity of the lamp was equal to 0.7. External interferences have been reduced by turning off the artificial lightning and by covering the background with a black cotton cloth.

The experimentation has been carried out to analyse the thickness irregularities and the conservation state of the test object by detecting possible surface cracks and defects.

Fig. 2. Lamp details.

The thermal response over time of the artifact has been recorded in two different time instants. Experimental results of a first observation (on 2017) have showed evident cracks or defects on the lamp surface, [10], [11]. The same defects were observed successively after 5 years (on 2022) in order to evaluate changes in its conservation state. The proposed approach has allowed us to get accurate and consistent information on the lamp decay over time. Intermediate precision evaluation has been used to provide quantitative information about the effectiveness of the monitoring process and the reliability of its results.

To evaluate the repeatability and intermediate precision contributions reported in (1) and (6), a frame rate of 32 frames/s has been used during the acquisition of N repeated measures for both measurement processes in different intermediate times. It is important to notice that the two measurement processes have been performed five years apart, therefore they are not repeated measurements but replicated ones. In fact, the last measurement process has been performed in a different intermediate time (after years). In order to evaluate the repeatability associated to the first measurement process as reported in (1), 32 repeated measures have been performed within 1 s. Figure 3 shows some thermographic images of the lamp details. Appropriate ROIs have been designed to delineate the areas of interest.

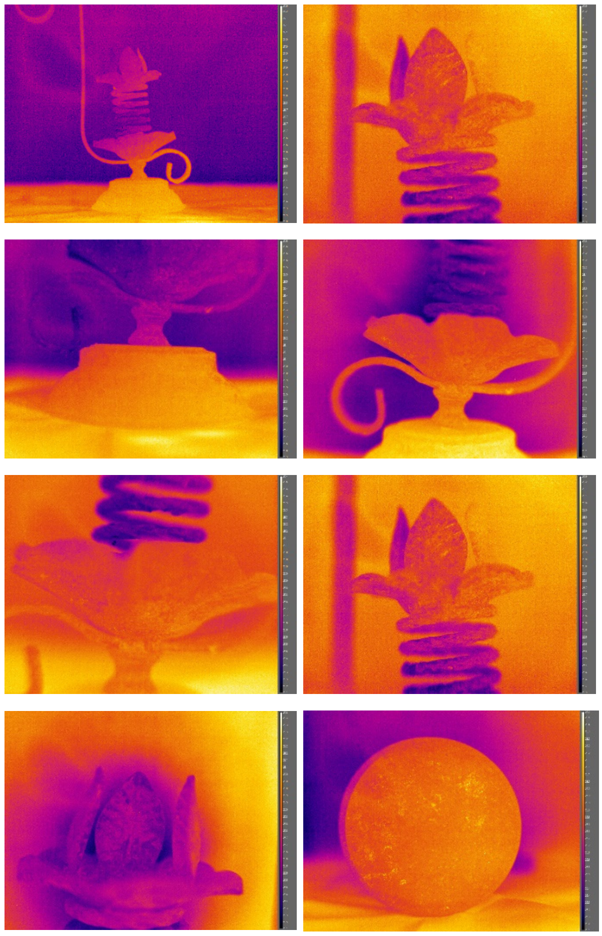

Fig. 3. Thermography images of lamp details.

The mean temperature of the pixels included in the ROIs has been evaluated as representative of the lamp. The measurements repeatability obtained in the first measurement process is equal to 0.00038 °C. The previous value allows us to assert that the measurement process has been properly carried out according to GUM and VIM recommendations. In fact, it is representative of a high level of repeatability of measurement. The intermediate precision has been evaluated by considering the results of both measurement processes. In this case, M is equal to 2 replicated measurements, where the number of repeated measures N is equal to 32 for each replicated measurement process. The evaluated intermediate precision in (6) is equal to 0.047 °C. This value shows as the replication of a measurement process on the same measurand after some time is inevitably affected by uncertainty which is greater than the repeatability contribution. This value allows us to understand if data coming from different measurement processes can be put in comparison for any purpose (conservation state monitoring, results comparison, crack progress evaluation …).

The aim of this paper is to highlight the relevance of this issue when monitoring applications over time are considered, so to alert the operator that the comparison between data coming from different experimentations is not taken for granted. The recommendations included in GUM and VIM can allow to improve the intermediate precision in order to make this comparison possible. Successively, by means of measurement repeatability and intermediate precision, user can get quantitative information on the possibility to use data coming from different experimentations over time.

IV. CONCLUSIONS

In this paper the authors have faced some metrological issues affecting the thermographic imaging. Reliability in measurements is an important aspect to assure in order to guarantee the consistency of the decisions based on measurement results. Often results are affected by significant uncertainty and errors. Repeatability and intermediate precision have been here proposed in order to guide users to improve the reliability of thermographic measurements.

It is appropriate to make clear that these two uncertainty contributions allow to obtain complementary information. Repeatability allows to evaluate the reliability associated to the results of a specific measurement process acquired in a short period of time. In other words, it provides information on the closeness of the measurement results obtained on the same measurand using the same measurement procedure, same operator, same measurement system, same operating conditions and same location. The smaller the quantity evaluated by equation (1), the higher the reliability of the measurement results.

Intermediate precision allows to understand if data from different measurement processes carried out in different time instants can be reliably put into comparison. In fact, it provides information about the capability to replicate reliably the same measurement process after a longer period of time. This uncertainty contribution includes more changes than repeatability which cannot be considered constant over a longer time period such as months or years. For this reason, the uncertainty contribution related to the intermediate precision is greater than one related to repeatability.

Monitoring the conservation state of ruins or discoveries needs to perform replicated measurements at a distance of time, after months or years, in order to check the evolution of cracks or to evaluate their maintenance over time. The comparison of data coming from different experimentations needs to collect data which must be coherent with previous measurement results. This is a requirement to assure in order to make comparison between data possible.

This issue becomes mandatory if thermographic images are considered. In fact, the variety of errors and interference sources which characterize thermographic imaging may make significant the previous uncertainty contributions so compromising the measurement reliability. In this case, it is discouraged any data comparison. For example, environment temperature, object emissivity, reflected temperature, relative humidity, distance are cause of systematic errors or independent biases. Whereas, external lightning, infrared sources or background interferences are cause of random errors. As a consequence, the user has to limit the previous interference sources. In such a circumstance, the repeatability and intermediate precision contributions can provide useful quantitative information about the measurement reliability assessment.

REFERENCES

[1] Guide to the expression of uncertainty in measurement (GUM) (BIPM, IEC, IFCC, ISO, IUPAC, IUPAP, and OIML) ISBN92-67-10188-9 (available at: www.bipm.org /utils/ common/documents/jcgm/ JCGM_100_2008_E.pdf) (accessed 3 May 2022).

[2] International Standardization Organization (ISO 5725-4). Accuracy (Trueness and Precision) of Measurement Methods and Results—Part 4: Basic Methods for the Determination of the Trueness of a Standard Measurement Method; ISO: Geneva, Switzerland, 2020.

[3] JCGM 200 2012 International vocabulary of metrology—basic and general concepts and associated terms, 3rd edn (VIM3) (available at: www.bipm.org/utils/common/documents/jcgm/JCGM_200_2012.pdf) (accessed 3 May 2022).

[4] Hamrelius, Tjorbjohn, “Accurate temperature measurement in thermography: an overview of relevant features, parameters, and definitions”, Thermosense XIII 1467 (1991), pp. 448-457.

[5] Usamentiaga, Rubén, et al. "Infrared thermography for temperature measurement and non-destructive testing." Sensors 14.7 (2014): 12305-12348.

[6] Dinardo, G., Fabbiano, L., Tamborrino, R., & Vacca, G, “Automatic defect detection and characterization by thermographic images based on damage classifiers evaluation”, Metrology and Measurement Systems, 27(2).

[7] Rosario Morello, Andrea M. Gennaro, Laura Fabbiano, An Emissivity-based Approach to Assess the Integrity State of Archaeological Discoveries by Thermography, 2022 IMEKO TC-4 International Conference on Metrology for Archaeology and Cultural Heritage, University of Calabria, Italy, October 19-21, 2022.

[8] Xiao, X., Gao, B., Woo, W. L., Tian, G. Y., & Xiao, X. T. “Spatial-time-state fusion algorithm for defect detection through eddy current pulsed thermography”. Infrared Physics & Technology, 90, 133-145, 2018.

[9] Benítez, H. D., Ibarra-Castanedo, C., Bendada, A., Maldague, X., Loaiza, H., & Caicedo, E. “Definition of a new thermal contrast and pulse correction for defect quantification in pulsed thermography”. Infrared Physics & Technology, 51(3), 160-167, 2008.

[10] Claudio De Capua, Rosario Morello, Ireneusz Jablonski, “Active and Eddy Current Pulsed Thermography to Detect Surface Crack and Defect in Historical and Archaeological Discoveries”, Measurement Journal, ISSN: 0263-2241, Vol. 116, pp. 676-684,

[11] R. Morello, C. De Capua, Thermography for non-Invasive Diagnosis of Conservation State of Archaeological Discoveries, 6th EnvImeko - IMEKO TC19 Symposium on Environmental Instrumentation and Measurements, June 24-25, 2016, Reggio Calabria, Italy, ISBN: 978-92-990075-2-5, pp. 26-29, 2016.