

Estimating the Uncertainty of the Result for Tests of Resistance to Environmental Conditions on the Example of the Method of Resistance to Neutral Salt Mist According to EN ISO 9227 / ASTM B117

Marcin GARBACZ¹

Summary

Environmental laboratory tests are one of the most frequently performed tests to evaluate materials used, among others, for the construction of rail vehicles. The requirements of the EN ISO/IEC 17025 standard for research laboratories, particularly when evaluating the compliance of materials with the specified requirements, impose on laboratories the need to consider the results of final measurements along with the uncertainties of these results. Due to the complexity of the physical and chemical processes occurring during environmental tests, determining the sources of uncertainty of the measurement result can be very complicated. The article presents one of the methods of estimating the complex uncertainty for environmental tests on the example of corrosion tests using the NORDTEST TR 537 concept of uncertainty estimation. The article presents an exemplary method of uncertainty estimation based on a set of empirical data obtained in an accredited Laboratory for Testing Materials and Structural Elements of the Railway Institute with the use of within-laboratory reproducibility and method bias. Examples of uncertainty estimation depending on the type of tested objects (metal details and paint coatings) and the method of their evaluation after corrosion tests (quantitative and qualitative methods) are presented. The article also briefly presents the possibilities of interpreting and processing the obtained data as part of the control carried out inside the laboratory on the basis of a simple statistical tool such as Shewhart control charts and the Ishikawa diagram for the method of determining corrosivity in salt chambers, identifying important factors influencing the measurement uncertainty and at the same time showing the complexity the entire research process.

Keywords: neutral salt spray, control charts, within-laboratory reproducibility, inter-laboratory reproducibility, bias, metal workpieces, paint coatings

1. Introduction

The environmental resistance tests carried out in the laboratory are designed to reflect the natural working conditions of the product along with factors that may adversely affect its performance. Various physical and chemical factors can adversely affect the product's usability, shelf life and safety. The purpose of the performed environmental tests is to assure the manufacturer that their product is constructed in accordance with the relevant standards/legal requirements and in such a way that it is safe for use. One particularly sensitive area concerning material safety is rail transport. Future designs of rolling stock involve creating new structural solutions based on lightweight materials while ensuring maximum safety. The vehicle's body must be protected against corrosion, sound, heat, and

vibration, and the materials used for these purposes must be resistant to various environmental factors. The environmental resistance tests of materials carried out in the laboratory can reflect many naturally occurring ageing mechanisms, such as resistance to corrosion, resistance to solar radiation, resistance to climatic conditions including humidity, resistance to thermal conditions, i.e. flammability properties, etc. The selection of the preferred ageing test for the material should be tailored to the everyday exposure it will experience in the train's construction.

In most cases, laboratory tests based on environmental tests are standardised and provide a reasonably high repeatability/reproducibility, as well as the possibility of creating a specific quality ranking of materials under the same ageing conditions carried out in the laboratory. However, one often overlooked issue is that

¹ Ms.C. Eng.; Railway Research Institute, Materials & Structure Laboratory; e-mail: mgarbacz@ikolej.pl.

environmental tests are burdened with a certain degree of uncertainty in the obtained results, and attempting to determine this uncertainty is a complex matter. The accreditation of research laboratories, in light of the requirements of the EN ISO/IEC 17025 [1] standard, as well as the evolving approach of ILAC to assessing these requirements, forces research laboratories to take increasing responsibility in reporting results along with determining the uncertainty of those results. One of the key changes introduced by the standard concerns the principle of decision-making. This principle describes how measurement uncertainty is taken into account when determining compliance with a specified requirement, especially when meeting specific legal regulations is necessary. Furthermore, each laboratory should implement procedures describing how uncertainty is estimated and specify the identified sources of uncertainty for each method, even when this uncertainty is not reported in the final reports (the information is irrelevant to the Client). The article presents an example of an approach to estimating measurement uncertainty of environmental tests, using corrosion tests as an illustrative case.

Accelerated corrosion testing is carried out in special chambers allowing testing in an artificial corrosive atmosphere. The applied atmosphere in these chambers intensifies the corrosion processes of metals, their alloys, and other materials used for corrosion protection, such as paint systems. The intensification of corrosion processes is achieved by exposing the materials to factors such as temperature, relative humidity, moisture condensation, pH, the presence and concentration of corrosive components (e.g., acids, chlorides, SO₂, H₂S), etc.

The article describes the issues of estimating the complex uncertainty of corrosion tests and uncertainty assessment of samples after testing in a salt spray chamber in accordance with the requirements of the EN ISO 9227 [2] standard and its American equivalent ASTM B117 [3]. Estimating the uncertainty of the results is challenging due to the complexity of the test, a significant number of partial uncertainty components involved in the research process, numerous instruments used in auxiliary measurements, as well as various types of tested samples and different methods of their assessment.

The exemplary approach to estimating the measurement uncertainty of salt chamber tests, as well as selected methods for assessing samples after this type of testing (complex uncertainty), is based on experimental results obtained at the accredited Materials & Structure Laboratory of the Railway Research Institute.

1.1. Salt spray corrosion resistance test method

The most widely used accelerated corrosion testing method worldwide is the salt spray test method

with appropriately defined process parameters. It has been employed for many decades to determine the corrosion resistance of metals and protective coatings (metallic, conversion, oxide, and paint coatings). Salt spray corrosion tests are used to compare materials exposed to the same agent in the form of a uniformly atomised mist of sodium chloride solution at a fixed concentration, temperature and suitably adjusted pH of the spray. The primary purpose of these tests is to determine the time required to initiate corrosion under established conditions and classify the materials' resistance based on the chosen method parameters (solely for comparative purposes). However, the results may (but do not necessarily have to) also be indicative of the service life of these elements in real operating conditions.

Worldwide, the most commonly used international standard for assessing material corrosiveness is EN ISO 9227 [2], as well as the American standard ASTM B117 [3]. The standards are almost identical, with the main difference being that the American counterpart does not include the possibility of using acidic salt spray (AASS) or copper-accelerated acetic acid salt spray (CASS) as described in the ISO standard (these methods are described in separate American standards). They also differ in the use of different reference materials and the approach to chamber calibration, which is briefly characterized in the later part of the article. Both standards contain detailed guidelines for conducting corrosion resistance tests, and the most important requirements include the supervision of process parameters such as:

- chamber operating temperature,
- salt spray collection rate (in measuring cylinders with funnels, in ml/h),
- concentration and pH of the sprayed salt solution,
- appropriate purity of water and reagents for preparing salt solutions,
- method of spraying (particle size distribution), etc,
- humidity (the compressed air that is used to create the salt spray is heated and saturated with moisture in a pressurised humidifier).

Research in a salt spray chamber (depending on the type) can be conducted using three different types of salt spray, as presented in Table 1. Test 1 is described by the previously mentioned ISO and ASTM standards, while Tests 2 and 3 are only described by EN ISO 9227 [2].

Testing in neutral salt spray (NSS) is particularly useful for studying: metals and their alloys, metallic coatings (anodic and cathodic), conversion coatings, anodic oxide coatings, and organic coatings on metal materials. On the other hand, testing in acetic acid salt spray (AASS) and copper accelerated salt spray (CASS) is useful for studying decorative coatings such

Table 1

Standardised corrosion resistance tests using salt spray tests

Test number	1	2	3
Test type:	Neutral Salt Spray (NSS) tests	Acetic Acid Salt Spray (AASS) tests	Copper Accelerated Salt Spray (CASS) tests
Solution composition:	NaCl 50 ± 5 g/l	NaCl 50 ± 5 g/l, CH ₃ COOH ¹	NaCl 50 ± 5 g/l CuCl ₂ ·2H ₂ O 0.26 g/l ± 0.02 g/l, CH ₃ COOH ¹
Temperature	35 ± 2°C		50 ± 2°C
pH of the condensate at 25°C	6.5÷7.2	3.1÷3.3	
The average condensate collection rate for a surface of 80cm ²	1.5 ml/h ± 0.5 ml/h		

¹ CH₃COOH used for acidifying the solution and adjusting it to the appropriate pH. [Authors' own elaboration].

as copper + nickel + chromium or nickel + chromium. It has also been found that such tests are suitable for studying anodic and organic coatings on aluminum substrates [4]. In the case of rail transport, where organic coatings on steel and aluminum are primarily used for corrosion protection, the most common application involves testing with neutral salt spray.

The salt spray corrosion test is a quality control examination commonly used to assess the quality of a specific technological process, for example, for elements coated with anti-corrosive layers. It tests elements that are typically capable of withstanding, for instance, 480 hours without visible signs of corrosion on the surface. If a batch of material starts corroding after 96 hours, it indicates an error in the technological process and the need to rectify this error in the anti-corrosive layer manufacturing process [4].

The results of salt spray corrosion tests cannot be directly converted into operating time under real-world operating conditions. Particularly large discrepancies between natural ageing and accelerated ageing in salt spray were observed for zinc coatings on steel substrates and organic coatings pigmented with zinc phosphate and another active inhibitor. The reasons for these observed differences lie in the absence of "dry periods" during salt spray ageing, which are necessary for the formation and reformation of the passivation layer on zinc, as well as the hydrolysis of the inhibitor and surface passivation. As a result of the observed discrepancies between the effects of natural and accelerated ageing, cyclic corrosion tests (CCT) have been developed in recent years, mainly driven by the automotive industry. These tests aim to better replicate the corrosive damage that may occur under natural conditions. Cyclic tests are based on intermittent exposure of samples in salt spray, with alternating

periods of varying humidity, which sufficiently approximates the real-world corrosive conditions to which the samples are exposed in their natural environment [4].

1.2. Concepts for determining the uncertainty of salt spray corrosion tests

Building the uncertainty budget for tests using salt spray chambers (or environmental testing in general) can be based on two different concepts.

The first concept involves estimating the uncertainty of the method based on constructing the uncertainty budget, considering each activity performed during the research process, and then determining the complex uncertainty, following guidelines such as the GUM [5].

Estimating the uncertainty budget for environmental tests based on this concept can be problematic and highly complex due to the large number of variables and possible intercorrelations of individual components, which may lead to inaccurate estimation of sensitivity coefficients in the complex uncertainty. Such an approach proves to be very useful when quantifying individual components of uncertainty. However, in some cases, this methodology can also result in underestimating the measurement uncertainty, partially because it is challenging to consider all possible contributions to uncertainty, as is the case in the considered method for corrosion resistance testing using salt spray chambers [6]. Figure 1 presents a simplified cause-and-effect diagram ("fishbone" diagram otherwise known as the Ishikawa diagram) for the method of determining corrosion resistance in salt spray chambers according to EN ISO 9227 [2] and ASTM B117 [3].

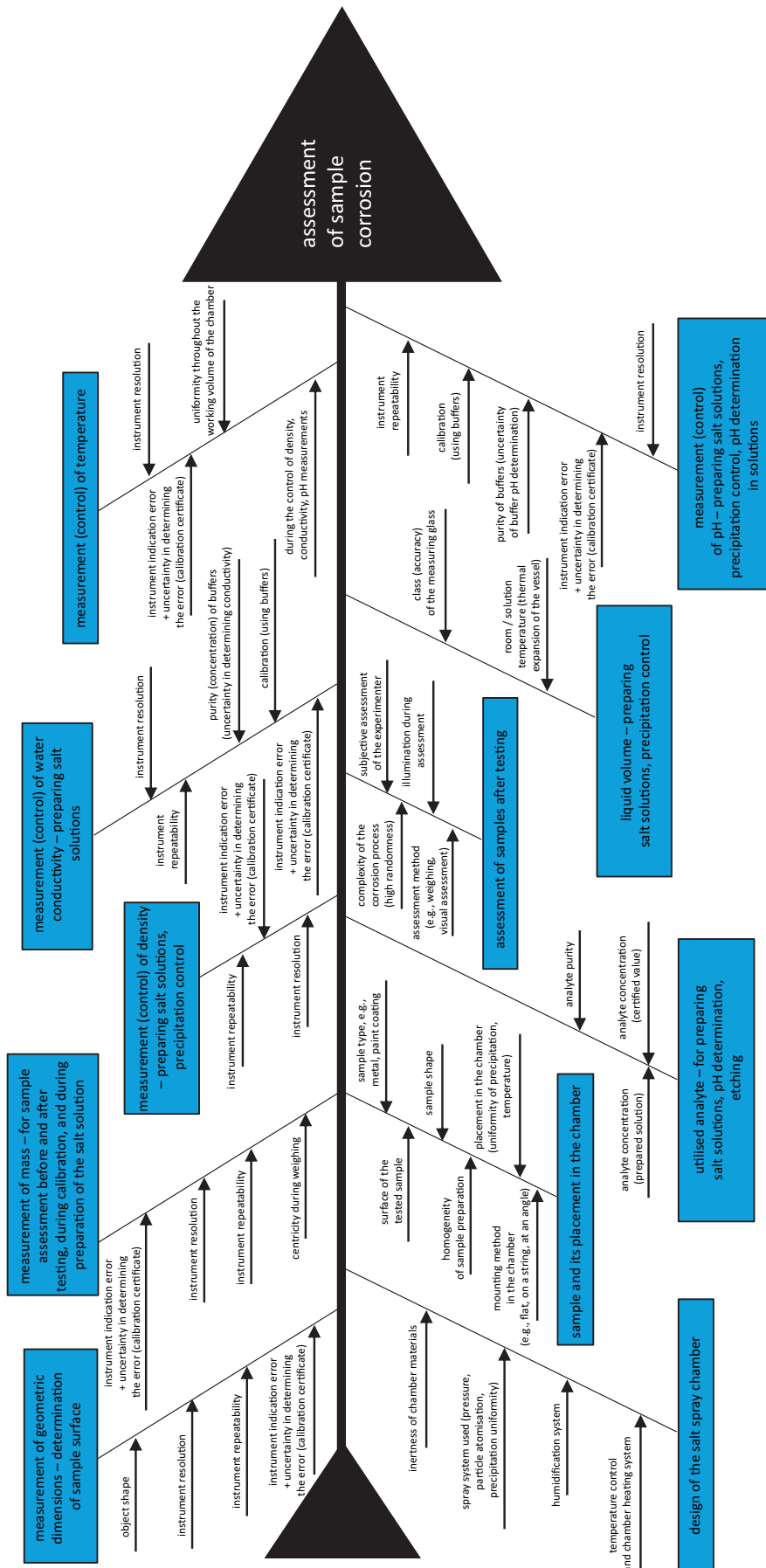


Fig. 1. Cause-and-effect diagram (commonly known as the “fishbone” diagram) for the method of determining corrosion resistance in salt spray chambers according to the respective EN ISO 9227 [2] and ASTM B117 [3] standards, illustrating selected factors influencing measurement uncertainty and the final result [Authors’ own elaboration]

The second concept concerns the calculation of measurement uncertainty adopted in environmental laboratories based on the document NORDTEST TR 537 [7]. This method has also been described in Polish in the Information Bulletin of the Pollab Club 2/51/2008 [8]. In the proposed model for calculating uncertainty, both within-laboratory reproducibility (R_w) and the estimation of laboratory and method bias together (using CRM, interlaboratory comparisons, recovery studies) are taken into account, or the estimated inter-laboratory reproducibility (S_R – standard deviation of reproducibility) is used for this purpose, e.g., based on results from different rounds of proficiency testing for a given method.

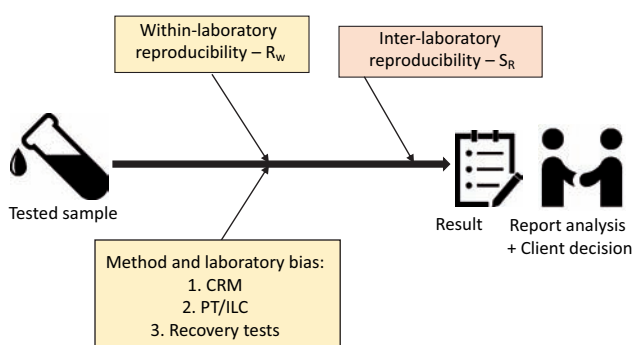


Fig. 2. Uncertainty model of measurement as depicted in the fishbone diagram [Author’s own elaboration]

The article provides a more detailed description of the method using the concept of calculating measurement uncertainty adopted in environmental laboratories, based on the document NORDTEST TR 537 [7], utilising within-laboratory reproducibility along with estimating the laboratory and method bias using reference materials (RMs).

It should be strongly emphasised that the estimation of each method’s uncertainty should encompass the entire research process, including the assessment of the sample after the completion of the test (regardless of whether the method is qualitative or quantitative). In cases where data will be reported only for the method used for the corrosion test (without considering the ageing test itself), such information must always be included in the test reports for the Clients.

1.2.1. Flow chart for estimating uncertainty for tests conducted in a salt spray chamber

According to the methodology described in NORDTEST TR 537 [7], as well as in the translation of the Information Bulletin of the Pollab Club 2/51/2008 [8], the results of uncertainty calculations should be based on a flow chart consisting of 6 specified stages (steps). This flow chart is presented in Table 2.

Table 2

Flow chart for determining uncertainty based on within-laboratory control and laboratory bias for a method in accordance with the concept of NORDTEST TR 537 [7] for environmental tests

Step	Action
1	Determine the measured value
2	Determine quantitatively (relatively) the within-laboratory reproducibility, $u(R_w)$, for: A: control sample B: any steps not covered by the control sample
3	Determine quantitatively (relatively) the components of laboratory and method bias, $u(\text{bias})$
4	Transform the components into standard uncertainty
5	Calculate the complex standard uncertainty, u_c
6	Calculate the expanded uncertainty, U

1.2.2. Within-laboratory reproducibility – R_w

Estimating the within-laboratory reproducibility uncertainty $u(R_w)$ must cover all stages of the research process, and the measurement results of the control sample should be conducted exactly in the same way as the tested samples. When estimating the within-laboratory reproducibility uncertainty $u(R_w)$, it is important to consider the long-term variations of the most significant components of uncertainty occurring within the laboratory, e.g., caused by different standard solutions, new batches of reagents, different instruments, or different laboratories conducting the tests.

For tests conducted in salt spray chambers, the entire research process is based solely on placing the sample (workpiece) into the chamber and then continuously monitoring the proper operation of the salt spray chamber (often on a daily basis) by observation of the temperature inside the chamber, the amount of condensate in the measuring cylinders (with a funnel of 80 cm² surface area), pH, and density of the sprayed solution. Additionally, during the test, the salt solution tanks are replenished with a 5% sodium chloride solution with the appropriate pH and density. The solution must be prepared from suitably pure reagents – with certificates of analysis (p.a. grade) – and water with a conductivity of less than < 20 μS/cm or < 5 μS/cm (depending on the standard), obtained through distillation or deionisation processes.

To analyse the within-laboratory reproducibility for an environmental method, it would be best to base it on the principle of determining the method’s *robustness*, which assesses the impact of small, unintentional changes (conditions) in the measurement method on the final result. Robustness testing is carried out based on significance tests or analysis of variance (ANOVA). It involves detecting significant differences between

variances and mean values of measurement series generated for slightly different measurement conditions, such as variations in temperature, pH of salt solutions, or purity of reagents used in the analysis. The aim of such tests is to identify which parameters of the measurement method should be strictly adhered to and which parameters, if varied randomly, have an insignificant impact on the test result. However, this approach requires a significant amount of work and research, as well as familiarity with more advanced statistical techniques.

Also, a useful and less time-consuming tool for ongoing quality control of measurements in the laboratory is the use of control charts. They make it possible to monitor variability in a controlled process or sequence of analyses and to determine whether deviations of measurement results from accepted normative values are random in nature. The most commonly used control charts are Shewhart charts and CuSum (*cumulative summing chart*) charts.

The fundamental element of Shewhart charts is the set of control lines. It consists of a centre line (CL), with two warning limits drawn parallel to it, also known as action limits (upper and lower warning limits UWL and LWL), and two more distant control limits known as alarm limits (upper and lower control limits UCL and LCL). To determine the position of lines on the Shewhart chart, it is necessary to know the estimator of the expected value of the parameter that characterises the controlled process and the estimator of the measure of dispersion (variance) of the value of this parameter around its mean value. In practice, the estimator of the expected value is typically an arbitrarily imposed nominal value, a known value for a certified reference material, or the mean from a pilot sample. The estimator of dispersion can be an acceptable or assumed level of precision, the standard deviation value of a population of results, or the standard deviation of a pilot sample. Depending on the statistical implementation of the quality control program, the Shewhart chart can be constructed by plotting the following values: the arithmetic mean of measurement results (multiple or duplicate analysis), the standard deviation (multiple analysis), the range (multiple or duplicate analysis), the moving range (individual analysis), or individual measurement values [9].

The CuSum chart, on the other hand, is based on the differences between the means of consecutive measurement series and the target value (reference). If the process of performing the analysis is correct, the points plotted on the chart will cluster around zero along the horizontal axis. If, on the other hand, a line is formed from the points that is inclined relative to the horizontal axis, this indicates a systematic error [9].

Section 3 of this article, which contains the experimental results for the within-laboratory

reproducibility analysis $u(R_w)$, utilises the Shewhart control chart.

1.2.3. Laboratory/method bias

When determining the source of laboratory and method bias $u(\text{bias})$, reference materials (RM/CRM), results from inter-laboratory comparisons (proficiency tests) or recovery studies of the addition of a standard to the sample (for typical chemical analyses) are used. Bias sources should be eliminated as much as possible. According to the GUM [5] guide, the measurement result should always be corrected, if the bias is significant, based on reliable data, preferably such as CRM. However, even if the bias is zero, it must be estimated and treated as a component of uncertainty.

Although, unlike exposure in natural atmospheric conditions, laboratory testing in salt sprays is conducted with a reduced number of controlled variables, often leading to more repeatable results, estimating laboratory or method bias is questionable due to the highly random nature of the corrosion process itself (e.g., random formation of protective coatings on steel test panels under the influence of corrosion products, irregular alloy composition, significant randomness during the test duration, etc.). In this article, it is determined on the basis of a reference material (calibration steel plates).

1.2.4. Inter-laboratory reproducibility – S_R

In the case of obtaining satisfactory results from inter-laboratory comparisons for estimating method uncertainty, the reproducibility index (usually provided as the standard deviation of measurement reproducibility) can be used, provided that the comparison includes all uncertainty components and estimation steps (stages). In this scenario, the estimation of expanded uncertainty is possible by directly utilising the standard deviation parameter from inter-laboratory comparisons. However, using this method can lead to significant overestimation of uncertainty, depending on the quality of the laboratory's work [10].

2. Research material, methodology and apparatus used

For many years, the Materials & Structure Laboratory of the Railway Research Institute has been conducting research using *neutral salt spray* (NSS) in accordance with the methodology described in EN ISO 9227 [2] and ASTM B117 [3] standards, for which the Laboratory is accredited by the Polish Centre for

Accreditation (AB 369). The Materials & Structure Laboratory of the Railway Research Institute conducts corrosion resistance tests on:

- metallic elements with diverse chemical compositions (aluminum, steel, copper, and others) as well as metal alloys used in various industrial sectors, primarily in the railway industry,
- paint coatings commonly used in the transportation industry.

During the examination of the ageing effects of metal elements in salt spray, the Laboratory determines the percentage of the sample's surface affected by corrosion and its type, e.g., according to EN ISO 10289 [11] standard. Additionally, when possible, the Laboratory measures the mass loss of the samples using gravimetric methods.

The effects of salt spray ageing of paint coatings are usually assessed using the EN ISO 4628 series of standards (for which the Laboratory is also accredited):

- EN ISO 4628-1 General introduction [12],
- EN ISO 4628-2 Assessment of degree of blistering [13],
- EN ISO 4628-3 Assessment of degree of rusting [14],
- EN ISO 4628-4 Assessment of degree of cracking [15],
- EN ISO 4628-5 Assessment of degree of flaking [16],
- EN ISO 4628-8 Assessment of degree of delamination and corrosion around a scribe [17],

and other selected methods to assess changes in physical and chemical properties, e.g. assessment of adhesion or hardness of the coating.

The article also presents a method for estimating the corrosion uncertainty for selected assessment methods of tested objects (quantitative and qualitative-quantitative methods) based on long-term experimental results from the resources of the Materials & Structure Laboratory.

2.1. Standardised salt spray test plates (calibration)

In order to ensure repeatability and reproducibility of test results according to EN ISO 9227 [2] and ASTM B117[3] standards, it is recommended to periodically "calibrate" the salt spray chamber by placing steel reference samples inside it (evenly distributed over the entire surface used for testing) and estimating the mass loss after a specified exposure time to salt spray. In this context, the term "calibration" is understood as checking the proper operation of the salt spray chamber, including the correct execution of individual steps during its operation in accordance with the requirements of the standard (e.g., concentration and density of the salt solution, correction of the pH of the solutions when required, purity of the water

used for their preparation, proper spraying of the salt spray, etc.). Calibration results that significantly deviate from the acceptable criteria suggest that the chamber is not working properly or the entire standardised testing process is being conducted incorrectly. The obtained calibration results will not make it possible to adjust the instrument, nor do they provide precise information on the cause of an incorrect result, but they are a valuable source of information on the quality of the tests carried out.

2.1.1. Calibration in accordance with EN ISO 9227

The requirements of EN ISO 9227 [2] regarding the calibration of the chamber are precisely described in detail in the ISO/TC156/WG7 [18] report referred to in the standard, which was created by a specialist technical group responsible for developing standards in the field of corrosion of metals and alloys. This report served as the basis for subsequent updates of the standard with appropriate provisions concerning the verification of the proper functioning of salt spray chambers, and the established assumptions for the entire calibration procedure were developed based on the results of research conducted by 18 laboratories from 9 countries. The mentioned report defines an appropriate control sample, which is a CR4 steel plate with an exact composition specified in ISO 3574 [19], and also describes the precise methodology for conducting salt spray chamber calibration. According to the guidelines of the report [18] and standard [2], calibration should be performed on a minimum of four, preferably six, of the mentioned steel plates or clean zinc plates. The dimensions of the plates are (150 × 70 × 1) mm for steel and (50 × 100 × 1) mm for zinc plates, evenly distributed throughout the chamber at an angle of $20 \pm 5^\circ$ relative to the vertical axis. These plates should be prepared in advance by degreasing and applying adhesive tape to their back and edges, ensuring that only one side of the reference sample is exposed to the salt spray.

The duration of the test under the specified conditions is 24 or 48 hours, depending on the chosen variant of the salt spray test. Immediately after removing the samples from the salt spray chamber (which is crucial, as corrosion continues to progress), the samples are cleaned using hydrochloric acid with an inhibitor in the form of hexamethylenetetramine or diammonium citrate for steel plates, and a saturated solution of glycine for zinc samples. The specific mass loss of the reference plates caused by corrosion, measured with an accuracy of 1 mg, is then converted to mass loss relative to the exposed surface area in the salt spray test. Table 3 shows the permissible mass loss of the reference steel and zinc plates, as defined by standard PN-EN ISO 9227 [2], while Figure 3 illustrates the

appearance of steel plates before and after exposure following the etching process.

It is worth mentioning at this point that while steel corrodes much faster than zinc under natural conditions (approximately 14–25 times, depending on the atmosphere), under NSS test conditions it corrodes only 1.4 times faster, indicating an uneven acceleration of corrosion of steel and zinc under the conditions found in salt chambers [20].

2.1.2. Calibration in accordance with ASTM B117

The requirements of ASTM B117 [3] regarding chamber calibration are described in Chapter X3 of the standard (detailed information is provided in the inter-laboratory comparison report RR:G01-1003 [21]). According to the provisions of this annex, the required test panels should have dimensions of (76 × 127 × 0.8) mm, i.e., (3.0 × 5.0 × 0.0315) inches, and should be made of cold-rolled SAE 1008 carbon steel (UNS G10080). These plates must be prepared in advance by degreasing, although it is not necessary to tape the back and edges of the plates as for the calibration described in EN ISO 9227 [2]. Calibration should be carried out for a minimum of 2 test panels placed at an angle of 30° relative to the vertical axis, and cylinders with funnels for measuring salt solution precipitation should be placed near the panels. The tested panels can be subjected to 3 different exposure periods, namely 48, 96, and 168 hours. After removing the panels, they should be rinsed with water and cleaned by etching in the same way as described in the methodology according to the EN ISO 9227 standard [2]. The test determines the mass loss of the reference plates due to corrosion with an accuracy of 1 mg (there is also no need to determine the mass loss per surface as in EN ISO 9227 [2]).

The above-described calibration practice has been validated, and the relevant standard includes information about the method's repeatability and reproducibility, as well as established criteria for accepting the obtained results depending on the test duration. The allowable mass loss of the reference steel plates, specified by ASTM B117 [3], is presented in Table 4, while Figure 4 shows an exemplary appearance of the steel plates before and after exposure following the etching process.

Calibration according to EN ISO 9227 [2] and ASTM B117 [3] can be carried out in parallel, as all the parameters of the salt chamber processes are the same. Simultaneous calibration on both types of test

Table 3
Permissible mass loss of reference plates during chamber calibration according to EN ISO 9227 [2]

Test method	Exposure time [h]	Permissible mass loss of the zinc reference plate [g/m ²]	Permissible mass loss of the steel reference plate [g/m ²]
NSS	48	50 ± 25	70 ± 20
AASS	24	30 ± 15	40 ± 10
CASS	24	50 ± 20	55 ± 15



Fig. 3. CR4 steel calibration plates before and 48h after testing (etched) [Author's own elaboration]

Table 4
Permissible mass loss of reference plates (specified repeatability and reproducibility limits) during chamber calibration according to ASTM B117 [3]

Test method	Exposure time [h]	Average sample mass loss ¹ [g]	Mean standard deviation of repeatability S_r^1 [g]	Mean standard deviation of reproducibility S_R^1 [g]	Repeatability limit r^1 [g] (P = 95%)	Reproducibility limit R^1 [g] (P = 95%)
NSS	48	0.8174	0.0559	0.1094	0.1565	0.4352
	96	1.5054	0.0982	0.1655	0.2750	0.4633
	168	2.5968	0.2227	0.3510	0.6234	0.9828

¹ Average values for 3 types of UNS G10080 panel sets (designations QP1, AP, QP2) and 3 laboratories (for all data presented in the standard).

panels contributes to a more accurate determination of the proper functioning of the salt spray chamber.

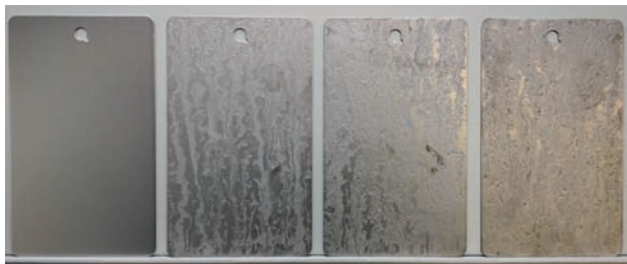


Fig. 4. SAE 1008 grade steel calibration plates before and after exposure in the salt chamber (etched) after exposure times of 48h, 96h and 168h (from left to right, respectively) [Author's own elaboration]

3. Test results and their interpretation

As mentioned at the beginning of the article, measurement uncertainties should include the entire test process, i.e. the complex uncertainty, i.e. in this case the uncertainty resulting from the corrosion test and the uncertainty resulting from the assessment of the test samples after the corrosion test. The method of estimating uncertainties for corrosion tests can vary significantly depending on the type of tested object and the assessment method. In this regard, each laboratory should introduce its own procedures for estimating uncertainties based on the applied methods or generalise them to at least distinguish between quantitative assessment (hard numerical data) and qualitative assessment (often highly subjective results) Table 5 presents the most important remarks and the method of estimating uncertainties for selected example tests of corrosion resistance conducted in the Materials & Structure Laboratory according to the adopted concept of NORDTEST TR 537 [7]. The assessment covers both qualitative tests (gravimetric methods) and qualitative-quantitative tests (paint coatings/metal workpieces).

3.1. Estimation of the complex uncertainty for quantitative assessment methods (gravimetric method) for salt chamber tests

The estimation of the reproducibility uncertainty $u(R_w)$ of the test method was based on the results of the salt chamber calibration using the reference plates as RM. Testing of the calibration plates described in sections 2.1.1 and 2.1.2 of this article was carried out at the Materials & Structure Laboratory of the Railway Research Institute in accordance with the requirements of EN ISO 9227 [2] and ASTM B117 [3].

The analysis of data for calibration on RM according to EN ISO 9227 [2] was based on Shewhart chart

for individual measurements using arbitrarily provided normative values. Based on [18], to estimate the reproducibility for the NSS method according to EN ISO 9227 [2], it was assumed that the expected value is 69.1 g/m^2 , while the standard deviation of the population is 10 g/m^2 . The necessary data used for the calculations are presented in Table 6. The data includes test results from 3 years, conducted by different researchers. Figure 5 presents the cumulative corrosiveness of 6 calibration samples determined by 3 laboratory technicians from August 2018 to October 2021. Figure 6 shows the Shewhart control chart along with the developed statistics. Statistical tests for gross error and systematic error for the collected data did not show significantly deviating values at the adopted significance level of 5%. Additionally, a test was conducted to check whether the introduced results (sample size $n = 36$) belong to a population with a normal distribution to be able to estimate the appropriate coverage factors for expanded uncertainty. The applied statistics of the test (construction of the so-called normal probability plot) at the adopted significance level $\alpha = 5\%$ did not provide evidence to reject the hypothesis that the distribution is normal [critical value of the test statistic r_{crit} (95%, 36): 0.9686, experimental value of the test statistic (correlation coefficient of the normal plot) r_{exp} : 0.9692].

In the case of calibration data in accordance with ASTM B117 [3], the analysis was based on the mass loss graph of reference plates as a function of time, with the determination of appropriate trend lines forced to intersect at the origin of the coordinate system. The collected data is presented in Table 7 (the data includes test results from 3 years, conducted by different researchers), and the graphical presentation of the results is shown in Figure 7. Also, in this case, a test was conducted to check whether the introduced results belong to a population with a normal distribution, in order to estimate the appropriate coverage factors for expanded uncertainty. The applied statistics of the test (construction of the so-called normal probability plot) at the adopted significance level $\alpha = 5\%$ did not provide evidence to reject the hypothesis that the distribution is normal [critical value of the test statistic r_{crit} (95%, 12): 0.9267, experimental value of the test statistic for 48h: r_{exp} : 0.9771 (after excluding two outliers), 96h: r_{exp} : 0.9437, 168h: r_{exp} : 0.9636].

Figure 8 shows the distribution of samples in the Ascott Premium CC1000iP salt spray chamber together with a summary of the mean results and standard deviations for the reference material (RM) tested, as summarised in Tables 6 and 7. Analysis of such data additionally makes it possible to determine whether there is a stable and uniform distribution of corrosion in the chamber throughout its working space. Table 8 presents a comprehensive summary of

Table 5
Flow chart for determining uncertainty based on within-laboratory control and laboratory bias (including indirectly the method bias) according to the NORDETEST TR 537 [7] concept for conducted corrosion tests

Step	Action	Notes for the corrosivity determination method – depending on the test object and the assessment method adopted:
1	<p>Determine the measured value</p>	<ol style="list-style-type: none"> 1. Metal mass loss due to the corrosive action of salt (gravimetric method – quantitative). 2. Assessment of the quality condition of metal workpieces and organic paint coatings (visual methods – qualitative-quantitative): <ul style="list-style-type: none"> • percentage surface involvement of red/white corrosion on metal workpieces in relation to the whole or separated area of the object under examination, • qualitative assessment of defects in organic coatings, such as blistering, rusting, cracking, flaking.
2	<p>Determine quantitatively (relatively) the within-laboratory reproducibility, $u(R_w)$, also known as intermediate precision, for:</p> <p style="text-align: center;">A: control sample</p> <p style="text-align: center;">B: any steps not covered by the control sample</p>	<ol style="list-style-type: none"> 1. Gravimetric method (quantitative method): $u(R_w)$ determined based on the mass loss (g/m² or g) due to the corrosive environment for reference plates (RM) according to ISO 9227 [2] and/or ASTM B11 [3] – described in sections 2.1.1 and 2.1.2 (assume a constant sample etching time for removing corrosion products in hydrochloric acid with successive repetitions – a very important parameter significantly affecting the result). Data acquisition for $u(R_w)$ analysis can be based, for example, on data plotted on control charts, such as Shewhart charts. The data should be sourced over a longer period of time and should include relevant factors influencing the measurement uncertainty (e.g. different salt chambers, different lab technicians, different batches of reagents, etc.). 2. Visual assessment - $u(R_w)$ determined by: <ul style="list-style-type: none"> • When assessing the surface involvement of red/white corrosion on metal workpieces, it is possible to use a specific uncertainty $u(R_w)$ derived from the data for the gravimetric method for RM (similar material - metal). It is also necessary to take into account the uncertainty arising from the experimenter's judgement, due to the highly subjective nature of the assessment. • For the qualitative-quantitative assessment of defects in organic coatings such as blistering, rusting, cracking or flaking, it is possible to use data from numerous tests carried out in the everyday work of the laboratory (assessment of the influence of variation in the result due to corrosion testing, also indirectly including non-uniformity in the execution of the test samples). It is also necessary to take into account the uncertainty arising from the experimenter's judgement, due to the highly subjective nature of the assessment. <p>A: $u(R_w) = \frac{s}{ \bar{x} } \cdot 100 [\%]$ (otherwise known as the coefficient of variation),</p> <p>s – standard deviation of the sample, \bar{x} – sample mean,</p> <p>B: The analysis of the control sample includes all stages (steps) of the analytical process after sample collection (this stage may be omitted).</p> <p>The relative standard uncertainty of the method for the metal reference samples determined according to EN ISO 9227 [2] is: 14.5% (estimation based on data from [18]) while for ASTM B117 [3] it is: 12.7%. The data concern the average coefficients of variation for conducted inter-laboratory comparisons to determine appropriate calibration criteria in the respective standards (based on the standard deviation of reproducibility and mean reproducibility).</p> <p>Note: for estimating within-laboratory variability, it is best to use root-mean-square, which is less sensitive to outliers than the arithmetic mean and disregards the significance of the sign of the obtained result.</p>

Step	Action	Notes for the corrosivity determination method – depending on the test object and the assessment method adopted:
3	<p>Determine quantitatively (relatively) the components of laboratory and method bias, u(bias)</p>	<p>The complex uncertainty of the bias u(bias) is estimated based on the formula:</p> $u(\text{bias}) = \sqrt{\text{RMS}_{\text{bias}}^2 + u(\text{Cref})^2}$ <p>where:</p> $\text{RMS}_{\text{bias}} = \sqrt{\frac{\sum (\text{bias}_i)^2}{n_{\text{RM}}}} \quad (\text{laboratory/method bias})$ <p>bias_i is defined as the difference in % from the nominal/certified value, n_{RM} – the number of reference materials used to estimate bias,</p> $u(\text{cref}) = \frac{S_R}{\sqrt{n_{\text{lab}}}}$ <p>(uncertainty in the nominal value of the RM used),</p> <p>u(cref) - estimated in this case on the basis of the data specified in the standards (inter-laboratory reproducibility S_R expressed as a relative value and the number of laboratories that determined the adopted reference value).</p> <ol style="list-style-type: none"> 1. Gravimetric method: Bias can be determined with respect to the accepted nominal values of RM, i.e., 69.1 g/, based on [18] for the method according to EN ISO 9227 [2], and in the case of ASTM B117 [3], the estimation can be based on average values for a given exposure time, as presented, among others, in Table 4 2. Visual assessment: Correctly estimating bias in this case is unlikely and can be omitted (lack of reference material). Based on analyses and long-term observations, it is assumed that the uncertainty arising from the experimenter's judgment outweighs the uncertainty related to differences in defects resulting from the corrosion test (usually very high repeatability and reproducibility). It is also possible to include own stable reference materials showing significant surface defects and perform periodic calibration as in the case of steel reference plates.
4	<p>Transform the components into standard uncertainty</p>	<p>When using data from calibration certificates or other reference documents, the determined confidence intervals can be converted to a standard uncertainty. In the adopted procedure model, the components of standard uncertainty include:</p> $u_1 = u(R_w),$ $u_2 = u(\text{bias}) = \sqrt{\text{RMS}_{\text{bias}}^2 + u(\text{Cref})^2},$ $u_3 = u(\text{opb}) \quad (\text{uncertainty of assessment after corrosion testing}),$ $u_n \quad (\text{other relevant factors influencing the final uncertainty of the measurement}).$
5	<p>Calculate the complex standard uncertainty, u_C</p>	$u_C = \sqrt{u_1^2 + u_2^2 + \dots + u_n^2},$
6	<p>Calculate the expanded uncertainty, U</p>	$U = k \cdot u_C.$ <p>In this case, appropriate statistical tests can be performed to determine the distribution of the empirical data in order to adopt the appropriate coverage factor. For simplicity, in this article, it is assumed with a confidence level of 1 - α = 95% that the coverage factor is equal to k = 1.96, which corresponds to a normal distribution.</p>

[Authors' own elaboration].

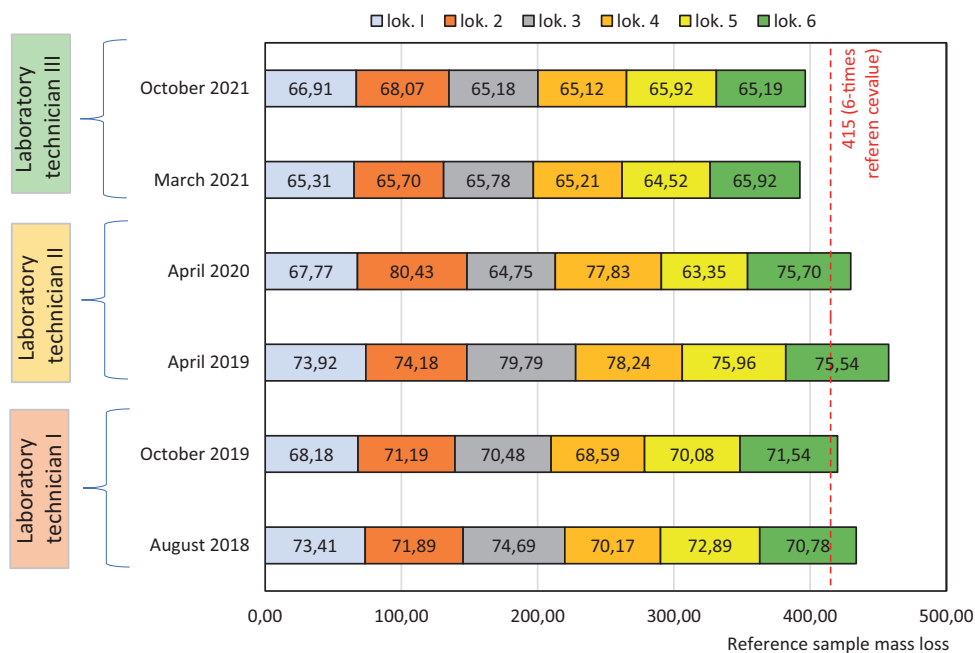
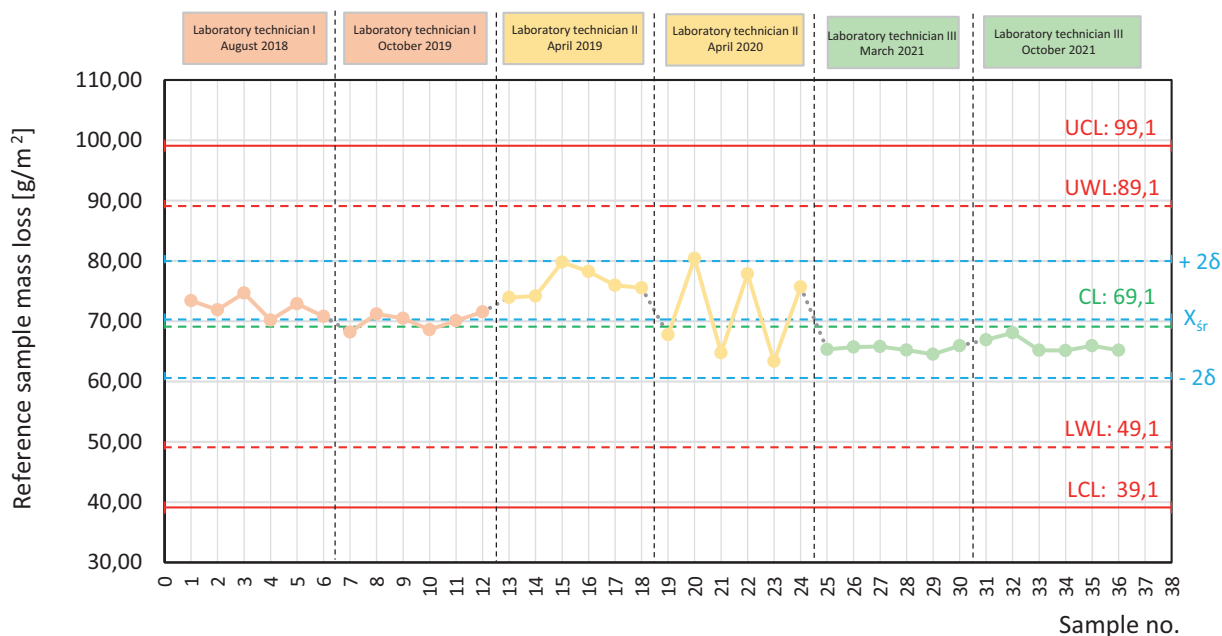


Fig. 5. Summary of the total corrosivity (6 calibration samples), determined by three laboratory technicians over the time period August 2018 – October 2021, in accordance with EN ISO 9227 [2] [Author’s own elaboration]



Statistical analysis of the series		Statistical parameters of the series
Gross error test (Grubbs): (P = 95%) no outliers $T_{exp}: 2.09$ $T_{crit}(95\%, 36): 2.99$	Bias test (Student's t-test, two-sided): (P = 95%) mean value of the series is not significantly different from the expected value $t_{exp}: 1.46$ $t_{crit}(95\%, 35): 2.03$ P: 84.78%	•max: 80.43 •min: 63.35 •mean: 70.28 •std. dev.: 4.85 •coe. of var.: 6.90%

Fig. 6. Shewhart chart with statistics created for single measurements on the basis of arbitrarily provided normative values for calibration results according to EN ISO 9227 [2]; [Author’s own elaboration]

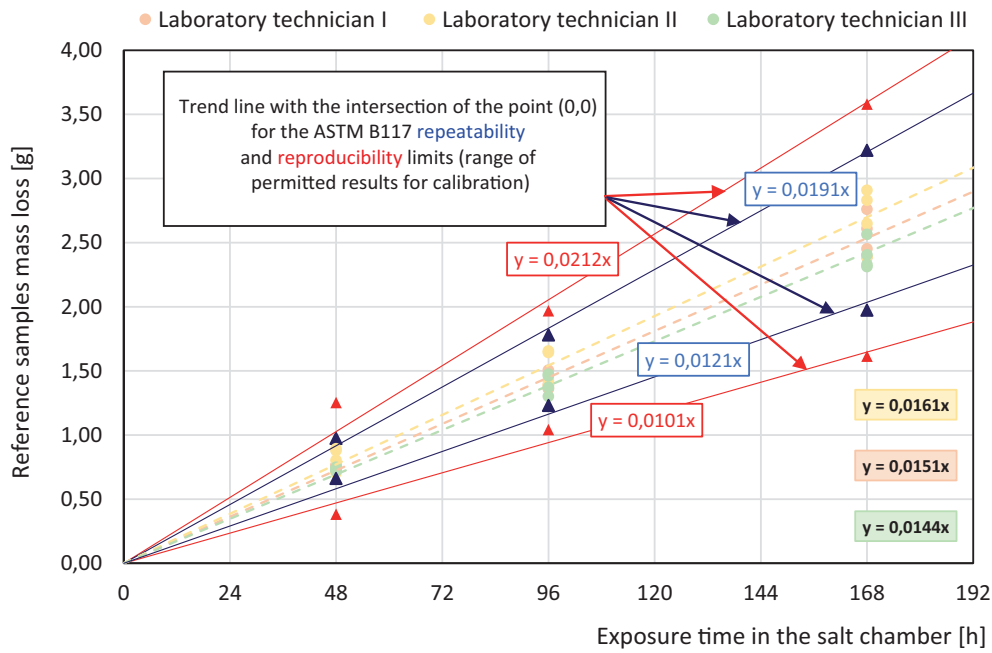


Fig. 7. Graphical presentation of the results of the salt spray chamber calibration performed by 3 laboratory technicians for the time period August 2018 – October 2021 complying with ASTM B117 [3]; [Authors' own elaboration]

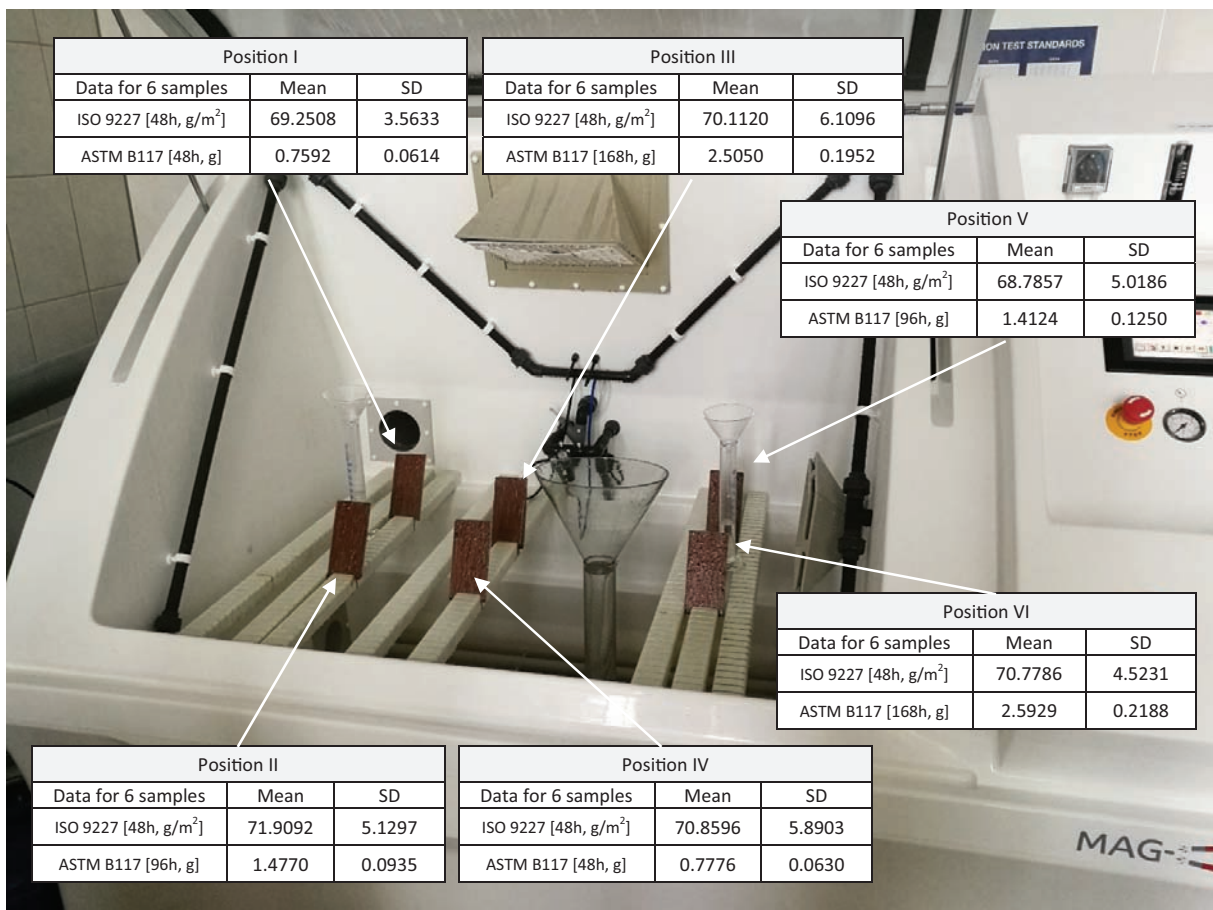


Fig. 8. Distribution of samples in the Ascott Premium CC1000iP salt spray chamber, together with a summary of the mean results and standard deviations for the reference material (RM) tested, as summarised in Tables 6 and 7 [Author's own elaboration]

Table 6

Salt chamber calibration results determined by 3 laboratory technicians during the time period August 2018 – October 2021 in accordance with EN ISO 9227 [2], to determine within-laboratory reproducibility

Calibration date	08.2018	10.2019	04.2019	04.2020	03.2021	10.2021	Mean	Standard deviation
Laboratory technician	I		II		III			
Mass loss [g/m ²] depending on sample location according to Figure 8								
Position I	73.41	68.18	73.92	67.77	65.31	66.91	69.25	3.56
Position II	71.89	71.19	74.18	80.43	65.70	68.07	71.91	5.13
Position III	74.69	70.48	79.79	64.75	65.78	65.18	70.11	6.11
Position IV	70.17	68.59	78.24	77.83	65.21	65.12	70.86	5.89
Position V	72.89	70.08	75.96	63.35	64.52	65.92	68.79	5.02
Position VI	70.78	71.54	75.54	75.70	65.92	65.19	70.78	4.52
Mean	72.30	70.01	76.27	71.64	65.41	66.07	-	
Standard deviation	1.69	1.37	2.31	7.26	0.51	1.20		

[Authors' own elaboration].

Table 7

Salt chamber calibration results determined by 3 laboratory technicians during the time period August 2018 – October 2021 in accordance with ASTM B117 [3], to determine within-laboratory reproducibility

Calibration date	08.2018	10.2019	04.2019	04.2020	03.2021	10.2021	Mean	Standard deviation
Laboratory technician	I		II		III			
Mass loss [g] as a function of exposure time to neutral salt spray								
48h	0.7603	0.7426	0.8777	0.7039	0.7484	0.7225	0.7592	0.0614
	0.7652	0.7505	0.8920	0.8034	0.7189	0.7357	0.7776	0.0630
96h	1.3933	1.5053	1.3863	1.6425	1.4602	1.4746	1.4770	0.0935
	1.3628	1.3603	1.4238	1.6548	1.3009	1.3718	1.4124	0.1250
168h	2.3842	2.4515	2.6476	2.8300	2.4038	2.3128	2.5050	0.1952
	2.6091	2.7591	2.9079	2.3893	2.5634	2.3285	2.5929	0.2188
Mean	48h	0.7547		0.8193		0.7314		-
	96h	1.4054		1.5269		1.4019		
	168h	2.5510		2.6937		2.4021		
Standard deviation	48h	0.0101		0.0801		0.0862		
	96h	0.0683		0.0632		0.1416		
	168h	0.1677		0.1923		0.2304		

[Authors' own elaboration].

the estimated complex uncertainty for quantitative assessment methods (gravimetric method) for tests conducted in the salt spray chamber, according to the concept of uncertainty estimation for environmental testing NORDTEST TR 537 [7].

Analysis of the Shewhart chart presented in Figure 6 and the graph presented in Figure 7, in which the collected calibration data is plotted together with the determined limits of repeatability and reproducibility,

makes it very easy to draw conclusions about the correctness and quality of the tests conducted in the salt spray chamber, also indirectly reflected in the daily laboratory tests. All results are within the accepted normative criteria and in their middle ranges, which testifies to a precise and reproducible research process. After a detailed analysis of all the data collected, it was found that laboratory technician I obtained results closest to the expected results set by EN ISO 9227 [2] and ASTM

Table 8

Estimated uncertainty of corrosion tests for quantitative assessment (gravimetric method) according to NORDTEST TR 537 concept [7]

No.	1	2	3
1	Standard →	EN ISO 9227	ASTM B-117
2	Relative within-laboratory reproducibility $u(R_w)$	$u(R_w) = 6,9\%$	$u(R_w) = 8,1\%$ (for all the results collected, over the entire 48h, 96h and 168h time periods)
3	Relative laboratory bias (including indirectly the method)	$RMS_{bias_RM1} = 7,1\%$	$RMS_{bias_RM2} = 8,4\%$
4	Components of standard uncertainty	$u(R_w) = 6,9\%$ $RMS_{bias_RM2} = 8,4\%$ (calculated for publication data [18]) ¹ $RMS_{bias_RM1} = 7,1\%$ $u(bias_RM1) = 7,9\%$	$u(R_w) = 8,1\%$ $u(Cref_RM2) = 4,2\%$ (calculated for data in Table 4) ¹ $RMS_{bias_RM2} = 8,4\%$ $u(bias_RM2) = 9,4\%$
5	Complex standard uncertainty	$u_c = 10,5\%$	$u_c = 12,4\%$
6	Expanded uncertainty	$P = 95\%$ and $k = 1.96$ (normal distribution) $U = 20,6\%$	$P = 95\%$ and $k = 1.96$ (normal distribution) $U = 24,3\%$
Average expanded uncertainty determined from two different RMs: 22.5% ~ 23%			

¹ For EN ISO 9227 [2], $s = 10 \text{ g/m}^2$ with mean $x = 69.1 \text{ g/m}^2$ and number of measurements/laboratories $n = 19$, while for ASTM B117 [3] the data from Table 4 (relative mean for 3 times 48h, 96h, 168h and for method reproducibility data) with mean sample size $n = 9$ were used. [Authors' own elaboration].

B117 [3]. Laboratory technician III received results with values slightly below the expected mean value. In the case of laboratory technician II, for the April 2020 calibration result, a significantly deviating variability (data dispersion) is apparent in relation to the rest of the pool of collected data (standard deviation for RM from CR4 steel was 7.26 g/m^2 ; for the Cochran test: $C_{crit}(95\%, 5, 6): 0.44$, $C_{exp}: 0.82$), even though the results of the calibration itself in terms of the sample mass loss are correct and within the allowed ranges set by the standard. Also, inflated average results of the sample mass loss, as well as increased standard deviations, are observed by this laboratory scientist during calibration carried out on samples in accordance with ASTM B117 [3] (Table 7). Data obtained in this way are indicative of a different way of working for this laboratory technician, rather than of incorrect operation of the chamber itself. After in-depth analysis, it became apparent that the differences in variability were due to a different way of etching the samples. This example shows how the Shewhart charts serve as a good tool to control the work inside the laboratory. The ideal mapping of the equation of the straight line determined on the basis of data from the American standard and Table 4 corresponds to the equation $y = 0.0156x$ (an average agreement of 97%, 97%, 92% was obtained for laboratory technician I, II and III, respectively).

The determined corrosivity of the chamber throughout its working interior, for the control

samples (Fig. 8), shows no significantly deviating values. The Cochran test for the concordance of variances of multiple series based on the data in Table 6 ($C_{crit}(95\%, 5, 6): 0.44$; $C_{exp}: 0.24$) showed that the variances are equal to each other. The data in Table 7 were not analysed for concordance of variance, due to the low number of results for the same level (calibration time) of sample mass loss (apparent trend of increasing standard deviation with the duration of calibration time).

3.2. Estimation of the uncertainty of the visual assessment method for tests conducted in a salt spray chamber

In the case of corrosion tests for the assessment of coatings and/or paint systems protecting metal on test objects and/or test panels according to the EN ISO 4628 Part II, III, IV and V series of standards [13, 14, 15, 16], the determination of within-laboratory reproducibility $u(R_w)$ is significantly hampered by the lack of standardised general reference material available. In this case, it is possible to create one's own reference materials and test them over a fixed, sufficiently long period of time, taking into account all relevant variations for the corrosion test method mentioned in the article. For the purposes of this article, the relative within-laboratory reproducibility $u(R_w)$ for corrosion tests of paint coatings was estimated based on a review of

data collected over several years. These concerned the repeatability and reproducibility of tests carried out for Clients in the Materials & Structure Laboratory of the Railway Research Institute. Control tests on the same products, carried out on average every 3 years (at least 3 samples of a given system/coating were customarily tested), showed an average variation in repeatability and reproducibility of ± 0.5 assessment units for both the number and size of defects assessed according to the listed EN ISO 4628 series of standards. It is not possible to estimate the bias of a method or laboratory for this type of test sample/object due to the lack of a certified

(reference) value. However, based on many years of observing the variability of results for test samples/objects protected with a coating or paint system, it was concluded that the uncertainty in the experimenter's assessment determined after the test significantly exceeds the uncertainty caused by the method/laboratory bias (usually the results are highly reproducible) and can be ignored in this case. The estimated expanded uncertainty of corrosion tests conducted for coatings and/or paint systems protecting metal on test objects and/or test panels according to the NORDTEST TR 537 [7] concept is presented in Table 9.

Table 9

Estimated uncertainty of corrosion tests for qualitative-quantitative assessment according to the NORDTEST TR 537 concept [7]

No.	1	2	3
1	Standard →	EN ISO 4628-2 [13] – blistering EN ISO 4628-3 [14] – rusting EN ISO 4628-4 [15] – cracking EN ISO 4628-5 [16] – flaking	EN ISO 10289 [11] Red/white corrosion occupation A [%]
2	Relative within-laboratory reproducibility $u(R_w)$	$u(R_w) = 8,3\%$ (0.5 unit of assessment for both the number and size of defects for all methods assessed)	$u(R_w) = 7,5\%$ – root mean square for EN ISO 9227 and ASTM B117 methods extracted from Table 8
3	Relative laboratory bias (including indirectly the method)	not estimated (no reliable reference material)	$u(\text{bias}) = 8,7\%$ – root mean square for EN ISO 9227 and ASTM B117 methods extracted from Table 8
4	Components of standard uncertainty	$u(R_w) = 8,3\%$ $u(\text{opb})_{\text{PN-EN ISO 4628-2}} = 6,6\%$ (I) i 7,5% (R) $u(\text{opb})_{\text{PN-EN ISO 4628-3}} = 8,5\%$ (I) i 5,6% (R) $u(\text{opb})_{\text{PN-EN ISO 4628-4}} = 6,6\%$ (I) i 9,3% (R) $u(\text{opb})_{\text{PN-EN ISO 4628-5}} = 8,2\%$ (I) i 5,0% (R) Approximate uncertainty values resulting from the assessment – a root mean square i.e. 7.5% (I) and 7.1% (R) were used for further calculations. Key: $u(\text{opb})$ – post-test assessment uncertainty; I – intensity, R – size	$u(R_w) = 7,5\%$ $u(\text{bias}) = 8,7\%$ $u(\text{opb})_{\text{KC}} = 4\%$ (uncertainty of post-test assessment, red corrosion) $u(\text{opb})_{\text{KB}} = 8\%$ (uncertainty of post-test assessment, white corrosion)
5	Complex standard uncertainty u_c	$u_c(I) = \sqrt{u(R_w)^2 + u(\text{opb})^2}$ $u_c(I) = 11,2\%$ $u_c(R) = \sqrt{u(R_w)^2 + u(\text{opb})^2}$ $u_c(R) = 10,9\%$	$u_c(KC) = \sqrt{u(R_w)^2 + u(\text{bias})^2 + u(\text{opb})_{\text{KC}}^2}$ $u_c(KC) = 12,2\%$ $u_c(KB) = \sqrt{u(R_w)^2 + u(\text{bias})^2 + u(\text{opb})_{\text{KB}}^2}$ $u_c(KB) = 14,0\%$
6	Expanded uncertainty U	$P = 95\%$ i $k = 1,96$ $U(I) = 22,0\%$ $U(R) = 21,4\%$ Calculated on the 6-point rating scale used: $U(I) = 22,0\% \cdot 6 \sim \pm 1,3$ (intensity) $U(R) = 21,4\% \cdot 6 \sim \pm 1,3$ (size) The standards recommend rounding off the results to halves, therefore the expanded uncertainty of the assessment after the corrosion test is: ± 1.5 grade for the intensity and size of the defect according to the EN ISO 4628 part II, III, IV and V series of standards.	$P = 95\%$ and $k = 1.96$ $U(KC) = 23,9\% \sim 24\%$ (red corrosion) $U(KB) = 27,5\% \sim 28\%$ (white corrosion)

Authors' own elaboration based on [7].

In corrosion tests on metal workpieces (red/white corrosion assessment), in order to determine the within-laboratory reproducibility $u(R_w)$, it is possible to use approximate data for the gravimetric method, due to the similar nature of the test object, i.e. a workpiece made of metal. In this case, however, it should be borne in mind that the shape of the test object can have a significant impact on the final result and thus on the uncertainty of the method (the relative within-laboratory reproducibility determined here concerned rectangular flat test panels). Also in the case of galvanised samples/objects, the uncertainty can reach much higher uncertainty values (underestimation), and this conclusion was drawn on the basis of data in the publication [18], in which galvanised samples were also tested (the Railway Research Institute laboratory does not perform chamber calibrations on this type of sample). To estimate the total uncertainty of the corrosion test for metal workpieces/samples (in general, not differentiating by metal type), data obtained from calibrations carried out on standard steel reference plates (RM material) required by industry standards were used, and the calculations are also presented in Table 9.

The expanded uncertainty of the corrosion tests together with the post-test assessment of coatings and/or paint systems protecting the metal workpiece/sample was estimated at an assumed significance level of $\alpha = 5\%$ with an assumed coverage factor of $k = 1.96$ (normal distribution was assumed) in order not to overcomplicate the calculations. There is likely to be a mixed distribution in the form of: normal (in-chamber testing, based on RM-based within-laboratory reproducibility data) and additionally rectangular (post-test visual assessment). In this case, an analytical method based on the convolution of the input distributions can be used to estimate the coefficient accurately, and the calculation scheme can be carried out according to the publication [22] (the proposed method does not require numerical calculations).

The calculation of the uncertainty, resulting from the method of assessing defects after tests conducted in a salt spray chamber according to the EN ISO 4628 (paint coatings [13, 14, 15, 16]) and EN ISO 10289 [11] series of standards (metal workpieces), was based on an estimation of the standard deviation of the experimental results carried out. For this purpose, 15 images were prepared for each type of defect to be assessed and assigned to selected persons for assessment as described in the mentioned standards (3 persons with experience, working in the laboratory on a daily basis, and 3 persons who had never been exposed to such assessment and testing were selected). It was decided to take this step of estimating the uncertainty of the assessment because the Railway Research Institute

laboratory did not have such a large number of samples with defects to assess, especially when it was customary in laboratory practice to test defects for blistering. To estimate the uncertainty of the expanded method, the results of the qualified persons assessing the test pieces on a daily basis at the Laboratory were used. The experimental data collected, together with the estimated statistics used in further calculations, are summarised in Table 10.

The experiment described above was conducted to test whether experience in assessing samples significantly affects the way samples are assessed. To assess this hypothesis, a test was used to compare the values of two pairwise series (pairs of mean results of experienced and non-experienced first-time assessors). The obtained results are presented in Table 11.

The hypotheses of the pairwise value comparison test assume: $H_0: \mu_d = 0$, $H_a: \mu_d \neq 0$ (two-sided variant), $H_a: \mu_d > 0$ or $\mu_d < 0$ (one-sided variant). Such a test can be used when comparing two methods or analytical studies with multiple chemical samples, differing, for example, in the amount of the substance to be determined or, as in this case, in the intensity and size of the coating defect, as reflected in the formula:

$$t_{\text{eksp}} = \frac{|\bar{x}_d|}{s_d} \sqrt{n},$$

where:

- \bar{x}_d – the mean value of the set of differences within each pair,
- s_d – the standard deviation of the set of differences within each pair,
- n – number of pairwise measurements.

The null hypothesis of this test is that the mean value of the series consisting of the differences of the pairs under consideration (taken from the population of results with a true value μ_d) is equal to zero, confirming the thesis that the assessment made by qualified and non-qualified persons does not differ significantly. The null hypothesis is rejected when the critical value of the parameter t , t_{crit} , derived from the Student's distribution for the appropriate number of degrees of freedom ($n_d - 1$, n_d – the size of the series of differences of the pairs of results) and the given confidence level, is less than the t_{exp} value calculated for the series of differences of the considered pairs of results.

In the case of the EN ISO 4628-2 [13] (blistering), EN ISO 4628-3 [14] (rusting), EN ISO 4628-4 [15] (cracking), EN ISO 4628-5 [16] (flaking) methods, the assessment is carried out on the basis of a 6-grade scale (the standard permits half-assessment). The relative uncertainty of the results obtained was calculated in relation to the maximum assessment range of the method, i.e. 6 units (scale 0–5).

Table 10

Summary statistics of the results obtained by experienced and non-experienced persons for the assessment methods of blistering, rusting, cracking and flaking of paint systems, as well as white and red corrosion of metal workpieces

Assessment method ↓ / laboratory technician No. →	Assessment of qualified persons (with experience)			Assessment of non-qualified persons (without any experience)			Relative average pooled standard deviation [%] ² (a: qualified persons, b: non-qualified persons c: all persons)
	1	2	3	4	5	6	
EN ISO 4628-2 ¹ [13]	3.2 / 3.3	3.3 / 3.7	3.3 / 3.6	3.1 / 3.4	2.8 / 3.7	3.6 / 3.1	a: 6.6 / 7.5 b: 10.2 / 9.7 c: 8.0 / 8.2
Statistics ³ :	Mean: 3.4 / 3.7 Standard deviation: 0.4 / 0.4			Mean: 3.3 / 3.6 Standard deviation: 0.6 / 0.6			
EN ISO 4628-3 ¹ [14]	3.9 / 4.5	3.7 / 4.5	3.5 / 4.6	3.5 / 4.3	4.5 / 4.3	3.2 / 3.1	a: 8.6 / 5.6 b: 13.4 / 14.5 c: 10.5 / 11.7
Statistics ³ :	Mean: 3.8 / 4.6 Standard deviation: 0.5 / 0.3			Mean: 3.8 / 4.0 Standard deviation: 0.8 / 0.9			
EN ISO 4628-4 ¹ [15]	4.5 / 4.0	4.3 / 3.8	4.1 / 3.9	3.5 / 3.1	3.9 / 4.1	3.8 / 3.4	a: 6.6 / 9.3 b: 9.3 / 15.1 c: 10.1 / 11.4
Statistics ³ :	Mean: 4.4 / 4.0 Standard deviation: 0.4 / 0.6			Mean: 3.9 / 3.6 Standard deviation: 0.6 / 0.8			
EN ISO 4628-5 ¹ [16]	4.7 / 4.5	4.5 / 4.4	4.2 / 4.5	3.5 / 4.5	3.8 / 4.4	3.8 / 3.4	a: 8.2 / 5.0 b: 14.1 / 12.9 c: 13.5 / 10.1
Statistics ³ :	Mean: 4.5 / 4.4 Standard deviation: 0.5 / 0.3			Mean: 3.8 / 4.1 Standard deviation: 0.8 / 0.8			
EN ISO 10289 [11] red/white corrosion [A%]	20 / 29	17 / 25	19 / 23	27 / 43	28 / 56	34 / 35	a: 4 / 8 b: 13 / 16 c: 13 / 17
R _p (protection factor) ⁴	2 / 2	2 / 2	2 / 2	2 / 1	2 / 1	1 / 1	
Statistics ³ :	Mean: 29 / 32 Standard deviation: 4 / 8			Mean: 41 / 52 Standard deviation: 13 / 16			

¹ The assessment was carried out in accordance with the subject standards, where the first value relates to the quantity of the defect and the second to its size (a scale with a 6-grade span in the assessment range of 0÷5). For example, 3.9 / 4.5 according to EN ISO 4628-2 for the assessment of blistering means a blistering defect quantity of 4 and defect size of 5 in relation to the entire assessed surface.

² Calculated from the root mean square (root of the sum of squares divided by the number of measurements) of the mean results and mean standard deviations for all 15 samples analysed. Scales with a 6-point spread were adopted for the relative assessment of defect size and quantity.

³ Refers to the root mean square of all mean results and mean standard deviations for the 15 samples analysed.

⁴ Protection factor determined from the formula in the standard $R_p = 3(2 - \log A)$, where A is a number expressing the percentage of the surface on which corrosion of the substrate has occurred. The R_p value was rounded to the nearest whole number. R_p value given for information, where A value was used for statistical calculations.

[Authors' own elaboration].

Based on the results of the pairwise comparison test of the values of the two series (e-stat module calculations – Table 11) for the assessment methods of blistering, rusting, sintering and flaking at the assumed confidence level of $1 - \alpha = 95\%$, in most cases significant differences were observed in the assessment of experienced and non-experienced persons. Only when assessing the amount and size of blistering (and the amount of rusting) was the null hypothesis not rejected in favour of the alternative hypothesis assuming that the mean value of the series consisting of the differences of the pairs considered is significantly different from

zero. This result may be affected by the detailed and well-characterised description of the assessment in the subject standard (lots of image patterns). When considering the results of the test comparing the values of two pairwise series for the method of assessing the percentage occupation of the workpiece surface by red and white corrosion at the assumed confidence level of $1 - \alpha = 95\%$, significantly larger and more significant differences were observed in the assessment of experienced versus non-experienced persons. Particularly noticeable are the large discrepancies in the results for the parameter of white corrosion, which is a characteristic

Table 11

Results of the pairwise comparison test (assessment by experienced vs. non-experienced persons) for the assessment methods of blistering, rusting, cracking and flaking as well as red and white corrosion (e-stat module calculations)

The hypotheses of the test were: $H_0: \mu_d = 0$, $H_a: \mu_d \neq 0$, Critical value of the $t_{crit}(95\%, 14)$ test statistic: 2.15		
Assessment method ↓ / assessment nature →	Intensity	Size
EN ISO 4628-2 [13]	$H_0: \mu_d = 0$ Experimental value of test statistics t_{exp} : 1.00 Confidence level of the test P: 66.6%	$H_0: \mu_d = 0$ Experimental value of test statistics t_{exp} : 1.47 Confidence level of the test P: 83.6%
EN ISO 4628-3 [14]	$H_0: \mu_d = 0$ Experimental value of test statistics t_{exp} : 0.43 Confidence level of the test P: 32.98%	$H_a: \mu_d \neq 0$ Experimental value of test statistics t_{exp} : 5.29 Confidence level of the test P: 99.99%
EN ISO 4628-4 [15]	$H_a: \mu_d \neq 0$ Experimental value of test statistics t_{exp} : 4.18 Confidence level of the test P: 99.91%	$H_a: \mu_d \neq 0$ Experimental value of test statistics t_{exp} : 4.00 Confidence level of the test P: 99.89%
EN ISO 4628-5 [16]	$H_a: \mu_d \neq 0$ Experimental value of test statistics t_{exp} : 4.58 Confidence level of the test P: 99.96%	$H_a: \mu_d \neq 0$ Experimental value of test statistics t_{exp} : 2.45 Confidence level of the test P: 97.19%
Assessment method ↓ / assessment nature →	Red corrosion	White corrosion
EN ISO 10289 [11]	$H_a: \mu_d \neq 0$ Experimental value of test statistics t_{exp} : 2.97 Confidence level of the test P: 98.99%	$H_a: \mu_d \neq 0$ Experimental value of test statistics t_{exp} : 5.35 Confidence level of the test P: 99.99%

[Authors' own elaboration].

product of zinc corrosion. The impact on such an assessment may be due to the fact that in the case of red corrosion there is a greater contrast to the metal than in the case of white corrosion, especially bearing in mind that the silvery-glossy zinc coating tarnishes and turns grey after testing in the salt spray chamber through the formation of a natural patina and can be confused with white corrosion, which is a white powder deposit.

4. Summary and conclusion

This article presents one approach to estimating the uncertainty of environmental tests, using the example of corrosion tests conducted in a salt spray chamber. This estimation example concerned, among other things, anti-corrosive materials used in rail transport, such as paint coatings. The Railway Research Institute certifies such materials according to its own developed

normative document, where the uncertainty of the obtained result is also taken into account when assessing and stating compliance with the specified criteria described in the aforementioned document. This approach is based on the PN-EN ISO/IEC 17025 quality system implemented in the laboratory and the Laboratory's awareness of the importance of the uncertainty of the result for a given test method.

The estimated expanded uncertainty according to the NORDTEST TR 537 concept for corrosion tests conducted in neutral salt spray together with post-test assessment of the samples (composite uncertainty) based on data from the Materials & Structure Laboratory is:

- for the gravimetric method for metal workpieces: $\pm 23\%$,
- for the qualitative-quantitative method in the form of assessment of coating defects assessed in accordance with the PN-EN ISO 4628 series of standards: ± 1.5 for the assessment of the quantity and size of the observed defect,

- for the qualitative-quantitative method in the form of assessment of the red corrosion occupancy of metal workpieces: $\pm 24\%$,
- for the qualitative-quantitative method in the form of an assessment of the white corrosion occupancy of metal workpieces: $\pm 28\%$.

Analysis of the data obtained by cyclically checking the corrosivity of the chamber with reference materials provides valuable information on the quality and correctness of the test conducted in the salt spray chambers and enables simple estimation of within-laboratory reproducibility. A simple tool such as the Shewhart chart, as well as other graphical and statistical methods of presenting the results, provide ongoing control of the corrosion process on the basis of trend analysis of the results obtained and give confidence in the quality of the results obtained.

Bearing in mind that quantitative-qualitative tests are the most frequently performed tests in salt spray chambers and carried out on a daily basis in laboratories (visual assessment of corrosion of workpieces, defects in protective coatings), the impact of such assessment on the final uncertainty of the result was determined. In this case, it was shown that the qualitative-quantitative assessment has a significant contribution to the final uncertainty of the result and comes largely from the subjective judgement of the experimenter, and therefore must necessarily be included in the uncertainty of the final result. For this purpose, it was decided to carry out a test to check whether the experience of the laboratory technician has a significant impact on the assessment of the samples after testing. The test of comparing the values of two pairwise series in almost every case showed significant differences in the assessment of persons with and without experience, except in the case of the assessment of blistering for paint coatings, which is probably due to the fact that a lot of image patterns are included in the standard to classify a given defect to its level. Large differences were also observed in the assessment of white corrosion, which can be confused with the passivation layer of zinc by persons without experience.

References

1. EN ISO/IEC 17025:2017: General requirements for the competence of testing and calibration laboratories.
2. EN ISO 9227:2017: Corrosion tests in artificial atmospheres – Salt spray tests.
3. ASTM B117-19: Standard Practice for Operating Salt Spray (Fog) Apparatus.
4. Garbacz M.: *Szacowanie niepewności przyspieszonych badań korozyjnych w obojętnej mgłę solnej wg ISO 9227 / ASTM B117 w kontekście badań powłok lakierowych oraz detali metalowych realizowanych w akredytowanym Laboratorium Badań Materiałów i Elementów Konstrukcji Instytutu Kolejnictwa* [Estimation of the uncertainty of accelerated corrosion tests in neutral salt spray according to ISO 9227 / ASTM B117 in the context of testing varnish coatings and metal details carried out in the accredited Materials and Structures Testing Laboratory of the Railway Institute], Praca dyplomowa w zakresie „Metrologii Chemicznej”, Uniwersytet Warszawski, Wydział Chemii, Warszawa, maj 2019.
5. Guide to the Expression of Uncertainty in Measurement (GUM), BIPM, IEC, IFCC, ISO, IUPAC, IUPAP, OIML. International Organization of Standardization, Geneva Switzerland, 1st Edition 1993, Corrected and reprinted 1995. Wydanie polskie: Wyrażanie niepewności pomiarów (Guide to the Expression of Uncertainty in Measurements) – przewodnik, wydanie polskie, Główny Urząd Miar, 1999, www.gum.gov.pl.
6. Measurement Uncertainty in Testing. Eurolab Technical Report No. 1/2002.
7. Magnusson B. et al.: *Handbook for calculation of measurement uncertainty in environmental laboratories*, Nordtest Report TR 537 (ed. 4) 2017.
8. Biuletyn Informacyjny Klubu POLLAB 2/51/2008, Podręcznik obliczania niepewności pomiaru w laboratoriach środowiskowych, ISSN 1428-6009 (tłumaczenie Nordtest Report TR 537, wersja 3 styczeń 2008).
9. Hyk W., Stojek Z.: *Analiza statystyczna w laboratorium*, wydanie I, Wydawnictwo naukowe PWN, Warszawa, 2016,
10. ISO/TS 21748:2004: Guidance to the use of repeatability, reproducibility and trueness estimates in measurement uncertainty estimation.
11. EN ISO 10289:2001: Methods for corrosion testing of metallic and other inorganic coatings on metallic substrates – Rating of test specimens and manufactured articles subjected to corrosion tests.
12. EN ISO 4628-1:2016: Paints and varnishes – Evaluation of degradation of coatings – Designation of quantity and size of defects, and of intensity of uniform changes in appearance – Part 1: General introduction and designation system.
13. EN ISO 4628-2:2016: Paints and varnishes – Evaluation of degradation of coatings – Designation of quantity and size of defects, and of intensity of uniform changes in appearance – Part 2: Assessment of degree of blistering.
14. EN ISO 4628-3:2016: Paints and varnishes – Evaluation of degradation of coatings – Designation of quantity and size of defects, and of intensity of uniform changes in appearance – Part 3: Assessment of degree of rusting.

15. EN ISO 4628-4:2016: Paints and varnishes – Evaluation of degradation of coatings – Designation of quantity and size of defects, and of intensity of uniform changes in appearance – Part 4: Assessment of degree of cracking.
16. EN ISO 4628-5:2016: Paints and varnishes – Evaluation of degradation of coatings – Designation of quantity and size of defects, and of intensity of uniform changes in appearance – Part 5: Assessment of degree of flaking.
17. EN ISO 4628-8:2012: Paints and varnishes – Evaluation of degradation of coatings – Designation of quantity and size of defects, and of intensity of uniform changes in appearance – Part 8: Assessment of degree of delamination and corrosion around a scribe or other artificial defect
18. Shigeru Suga & Shigeo Suga. Report on the results from the ISO/TC156/WG7 International Round Robin Test Programme on ISO 9927 Salt spray tests., J. Surface Finish. Soc. Japan. 2005, 56 p. 28, DOI:10.4139/sfj.56.28.
19. ISO 3574:2012: Cold-reduced carbon steel sheet of commercial and drawing qualities.
20. Augustyński Ł.: *Badania przyspieszone – komora solna i komora wilgotnościowa*, Materiały szkoleniowe TUV Rheinland Polska.
21. ASTM Research Report RR:A01-1003.
22. Fotowicz P.: *Obliczanie niepewności rozszerzonej metodą analityczną opartą na splocie rozkładów wielkości wejściowych* [Calculation of Expanded Uncertainty by Analytical Method Based on Convolution of Distributions of Input Sizes], *Pomiary Automatyka Robotyka*, 1/2005.