ISSN 1846-6168 (Print), ISSN 1848-5588 (Online)
https://doi.org/10.31803/tg-20250312153704

Original scientific paper
Received: 2025-03-12, Accepted: 2025-04-24

Risk Assessment Procedure for Calibration

Dubravka Božić*, Biserka Runje, Andrej Razumić

Abstract: In the product quality assessment process, it is determined whether the product meets the specified requirements. The assessment is based on measuring the specific characteristics of the product. Depending on the measurement uncertainty, the risk of accepting a product that does not meet specifications can be assessed, i.e., the consumer's risk. Conversely, rejecting a product that meets specifications poses a risk for the manufacturer. The standard method of risk assessment by points has been expanded planar and spatially and was applied for risk calculation during the calibration of the roundness measurement device on a moderate scale from −3 µm to 3 µm. A known observation of the dependent variables is used to calculate the producer's and consumer's risk of the corresponding explanatory variable. The risk of calibration was performed for the linearized tolerance interval. Measurement uncertainty of the calibrated value was determined from the regression line by utilizing the propagation of error.

Keywords: calibration; confusion matrix; consumer's risk; producer's risk; uncertainty

1 INTRODUCTION

1.1 Conformity Assessment, Basic Information

A product's quality assessment implies verification of its compliance with the required standards. The decision about conformity is based on the measured value of one or more characteristics of an item of interest and the measurement uncertainty associated with the provided measurement [1]. However, in the conformity assessment procedure, the wrong decisions can occur [2]. The product may be rejected as non-compliant despite meeting the specified requirements. This poses a risk to the producer and results in a loss of profit. On the other hand, accepting a product that does not conform to specifications is detrimental to the consumer [3].

In this study global producer's risk R_P and global consumer's risk R_C are calculated based on decision rules with a guard band [4]. A guard band of length w is placed between the tolerance interval [T_L, T_U] and the acceptance interval [A_L, A_U]. The tolerance interval represents the interval of permissible values for a particular property of an item of interest. It is assumed that the true value of the measurement should be within the tolerance range. Acceptance interval is the interval of permissible measured quantity values [5]. The symbols T_L and A_L refer to the lower limit of the tolerance interval and the acceptance interval, respectively. Analogously, the symbols T_U and A_U are used for the upper limit of the tolerance interval and the acceptance interval, respectively.

The global risk of producers and consumers was estimated using the Bayesian approach [6]. This approach combines prior knowledge of the quantity being measured with information obtained during the measurement process. The true value of measurement is assumed to be unknown. It is therefore modelled as a random variable with assigned the appropriate distribution g_0 with possible values η. The parameters of prior distribution g_0 are the best estimate of the measured quantity y_0 and the measurement uncertainty u_0 associated with the best estimate. The choice of prior depends on the measurand nature. Most often it is a normal distribution, but it can also be some other distribution [7-9]. Measurement data are modelled using the likelihood function h with taking the measurement uncertainty u_m obtained from the measurement process. The likelihood function usually is characterised by a normal distribution [5].

1.2 Regression vs Calibration Models for Risk Assessments

The calibration risk assessment models are based on regression risk assessment models. Depending on the method of calculating measurement uncertainty and the appearance of tolerance and acceptance curves, three models are used to estimate the global producer's and consumer's risk for regression [10]. Concerning the tolerance interval and acceptance interval, linearized and non-linearized models can be defined for risk assessment in regression, that is, calibration. In the linearized model, tolerance and acceptance lines are positioned parallel to the $y = x$ line. Also, tolerance and acceptance lines are placed symmetrically from both sides of the $y = x$ line. The same applies to the tolerance and acceptance curves for the non-linearized model, shown in Fig. 1.

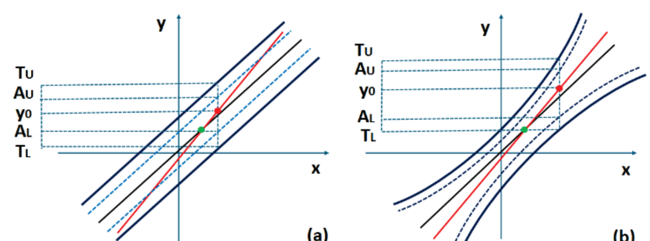

Figure 1 Tolerance and acceptance intervals: a) linearized model; b) non-linearized model [10]

Measurement uncertainty can be determined by using comprehensive measurement performance data or by considering a functional relationship obtained by linear regression analysis [10]. These measurement uncertainties are denoted with u_0^{GUM} and u_0^{LRA}, respectively. The measurement uncertainty u_0^{GUM} is the same for all points of the moderate scale x. On the other hand, measurement uncertainty u_0^{LRA} is calculated according to the propagation of error performed for the equation of regression line.

Therefore, the points of the moderate scale have different measurement uncertainties. The measurement uncertainties calculated in this manner are the largest for points located on the edges of the moderate scale, and the smallest in the middle of the moderate scale [11].

A trivial regression risk assessment model is a model with linearized acceptance and tolerance intervals combined with measurement uncertainty u_0^{GUM}. The non-linearized model is always used in combination with the measurement uncertainty u_0^{LRA}. In that case, the limits of tolerance and acceptance intervals depend on the measurement uncertainty u_0^{LRA} [10]. In this study, the risk assessment procedure for calibration is derived from a linearized regression model in combination with measurement uncertainty u_0^{LRA}.

1.3 Linearized Tolerance and Acceptance Intervals

For the linearized model, tolerance and acceptance lines are set parallel to the $y = x$ line. Let for the points of moderate scale x_i, $i = 1, 2, \ldots, n$ be valid $y_i' = x_i$, where n is the total number of points on a moderate scale, for which measurement was performed. When determining the regression line, the mean of m repeated measurements at the points of the moderate scale was used. The regression line has the form:

$$y = \alpha + \beta x, \tag{1}$$

where β is a slope and α is the y-intercept of the regression line. The values of the regression line y_i calculated at the points of the moderate scale x_i, $i = 1, 2, \ldots, n$ are obtained from the equation:

$$y_i = \alpha + \beta x_i,\ i = 1, 2, \ldots, n. \tag{2}$$

The upper tolerance line is set relative to the $y = x$, and it is given by the equation:

$$T_{\mathrm{U}i} = y_i' + \alpha_{\mathrm{U}},\ i = 1, 2, \ldots, n. \tag{3}$$

Analogously, the lower tolerance line is calculated from the equation:

$$T_{\mathrm{L}i} = y_i' - \alpha_{\mathrm{L}},\ i = 1, 2, \ldots, n. \tag{4}$$

The values α_{U} and α_{L} in Eq. (3) and (4) represent the y-intercept of the upper and lower tolerance lines, respectively. Due to symmetry, it is valid that $\alpha_{\mathrm{U}} = \alpha_{\mathrm{L}} > 0$. The tolerance interval has a range $\Delta T = \alpha_{\mathrm{U}} + \alpha_{\mathrm{L}} = 2\alpha_{\mathrm{U}}$. The acceptance interval may fall within the tolerance interval. If that is the case, then is valid that $[A_{\mathrm{L}}, A_{\mathrm{U}}] \subset [T_{\mathrm{L}}, T_{\mathrm{U}}]$, and holds $A_{\mathrm{L}} = T_{\mathrm{L}} + w$, $A_{\mathrm{U}} = T_{\mathrm{U}} - w$. Likewise, the acceptance interval can exceed the tolerance interval, i.e. it is valid that $[A_{\mathrm{L}}, A_{\mathrm{U}}] \supset [T_{\mathrm{L}}, T_{\mathrm{U}}]$. Then the acceptance interval is $[A_{\mathrm{L}}, A_{\mathrm{U}}] = [T_{\mathrm{L}} - w, T_{\mathrm{U}} + w]$. This is the widest interval along the guard band axis at which the risk curves can be observed. This interval encompasses both the above-mentioned interrelationships between the acceptance interval and tolerance interval, as well as the case when the limits of the tolerance and acceptance intervals coincide. The risk curves along the guard band axis will be shown for the subdivision nodes of the interval $[T_{\mathrm{L}} - w, T_{\mathrm{U}} + w]$. Also, behaviour of risk curves will be presented along a moderate scale.

Calibration risk assessment models are designated from the specified regression risk assessment models. The calibration and regression risks for the linearized model in combination with the measurement uncertainty u_0^{GUM} can be equal. This holds if the measurement uncertainty of the response variable in the regression model $u_{y_{\mathrm{r}}}^{GUM}$ is equal to the measurement uncertainty of the explanatory variable $u_{x_{\mathrm{e}}}^{GUM}$ in the calibration model. Also, the equality of risk in regression and calibration within this model arises from the model's geometry. It is straightforward to demonstrate that, for a given value of the response variable y_{r}, the value of the explanatory variable x_{e}, lies within the tolerance interval $[y_{\mathrm{r}} - \alpha_{\mathrm{U}}, y_{\mathrm{r}} + \alpha_{\mathrm{L}}]$ and holds $|\overline{AB}| = |\overline{CD}| = |\alpha_{\mathrm{U}} + \alpha_{\mathrm{L}}| = 2\alpha_{\mathrm{U}} = \Delta T$, as well as $|\overline{ET}| = |\overline{TF}| = |y_{\mathrm{r}} - x_{\mathrm{e}}|$, Fig. 2. This implies that x_{e} and y_{r} are symmetrically positioned within the segments $\overline{AB}$ and $\overline{CD}$ respectively, in terms of their distances from the endpoints of these segments. In this trivial model, the risks for regression and calibration differ if and only if $u_{y_{\mathrm{r}}}^{GUM} \neq u_{x_{\mathrm{e}}}^{GUM}$.

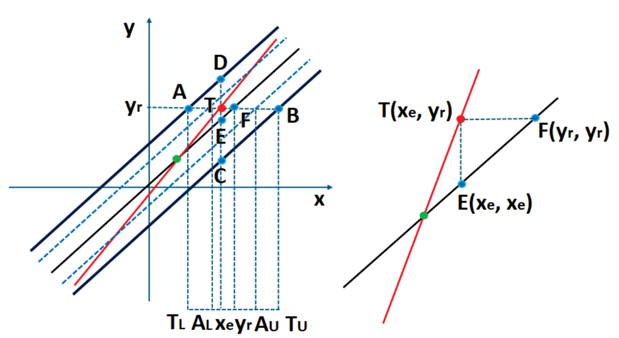

Figure 2 Linearized tolerance and acceptance interval for the explanatory variable

1.4 Measurement Uncertainty for Calibration Models

In the present study, a risk estimation model for calibration was examined, which is derived from a linearized risk estimation model for regression in combination with the measurement uncertainty u_0^{LRA}. The tolerance and acceptance intervals for risk estimation in calibration, in the case where the acceptance interval is contained within the tolerance interval, are modelled as described previously and shown in Fig. 2.

The fundamental difference in risk estimation between regression and calibration lies in the way the measurement uncertainty u_0^{LRA} is calculated. In risk estimation for regression, the measurement uncertainty u_0^{LRA} is computed for the values of the regression line y_i, $i = 1, 2, \ldots, n$ using the equation:

$$u_0^{LRA}(y_i) = \sqrt{u^2(\alpha) + x_i^2 u^2(\beta) + \beta^2 \sigma_x^2}. \tag{5}$$

The measurement uncertainty in the calibration model is calculated for the values of the explanatory variable, i.e., for the points of the moderate scale x_i, where $i = 1, 2, \ldots, n$ using the equation [11]:

$$u_0^{LRA}(x_i) = \sqrt{\frac{\sigma_y^2 + u^2(\alpha)}{\beta^2} + \frac{(y_i - \alpha)^2 u^2(\beta)}{\beta^4}}. \tag{6}$$

Eqs. (5) and (6) apply to a symmetric equidistant moderate scale centered around zero observed in this study. Details on how to calculate the measurement uncertainties $u(\alpha)$, $u(\beta)$ and the quantities σ_x and σ_y can be found in [10].

2 METHODS AND MATERIALS

2.1 Measurement Description

The risk assessment was performed during the calibration of the Mahr MMQ3 roundness measuring device. Data were collected throughout the calibration process of an inductive contact probe (dial indicator) on a universal length measuring machine (ULM). All measurements were conducted at the University of Zagreb in the Laboratory for Precise Measurement of Length at the Faculty of Mechanical Engineering and Naval Architecture. The risk assessment was carried for moderate scale in a range from –3 µm to 3 µm. For each of the $n = 13$ moderate scale points, $m = 3$ measurements were performed, and their mean values were considered [12]. The regression line is given by the equation:

$$y = -0.002784 + 1.008392x. \tag{7}$$

The remaining quantities necessary for calculating the measurement uncertainty u_0^{LRA} from Eq. (6) have the following values: $\sigma_y = 0.0182$ µm, $u(\alpha) = 0.005375$ µm and $u(\beta) = 0.003804$ µm. To overcome detected residuals autocorrelation from Ordinary Least Square regression (OLS), parameters α and β and their measurement uncertainties $u(\alpha)$ and $u(\beta)$ were determined using the Hildreth-Lu method for $\rho = 0.2$ [13, 14].

2.2 Risk Calculation

The global consumer risk R_C represents the probability where the measured value is within the acceptance interval, and the true value of an item of interest is outside the tolerance interval [5]. It is calculated as the sum of two double integrals, and holds $R_C = I_1 + I_2$, where:

$$I_1 = \int_{-\infty}^{T_L} \int_{A_L}^{A_U} g_0(\eta) h(\eta_m \mid \eta) \mathrm{d}\eta_m \mathrm{d}\eta, \tag{8}$$

$$I_1 = \int_{T_U}^{\infty} \int_{A_L}^{A_U} g_0(\eta) h(\eta_m \mid \eta) \mathrm{d}\eta_m \mathrm{d}\eta. \tag{9}$$

The global producer risk R_P represents the probability that the measured value is outside the acceptance interval, and that the true value is within the tolerance interval. It can also be expressed as the sum of two double integrals, and the following holds $R_p = I_3 + I_4$, where:

$$I_3 = \int_{-\infty}^{A_L} \int_{T_L}^{T_U} g_0(\eta) h(\eta_m \mid \eta) \mathrm{d}\eta_m \mathrm{d}\eta, \tag{10}$$

$$I_4 = \int_{A_U}^{\infty} \int_{T_L}^{T_U} g_0(\eta) h(\eta_m \mid \eta) \mathrm{d}\eta_m \mathrm{d}\eta. \tag{11}$$

For the prior g_0, a normal distribution with parameters y_0 and u_0 is assumed. Considering that the risk during calibration for the given values of the response variable y_r is calculated for the points of the moderate scale, this leads to role changes. The best estimates of y_0 in the risk assessment during calibration, for a given value of the response variable y_r, are calculated from the appropriate equation for the explanatory variable x_e:

$$y_0 \equiv x_{e_i} = \frac{y_{r_i} - \alpha}{\beta}, \; i = 1, 2, ..., n. \tag{12}$$

The response variable y_{r_i}, $i = 1, 2, \ldots, n$ takes on values in the interval of [–3.0279, 3.0224] µm.

The measurement uncertainty of the prior is given by Eq. (6). The measurement uncertainty u_m is the uncertainty of some future measurement process. This process is independent of the measurements from which the regression line from Eq. (7) was obtained. The expression for the likelihood function h contains the parameter u_m. In this work, it is assumed that $u_m = 0.5u_0$. The likelihood function is also modelled as a normal distribution.

The width of the tolerance interval is set so that it depends on the measurement uncertainty $u_0^{LRA}(x_i)$ of the points of the moderate scale and amounts to $\Delta T = 6\min(u_0^{LRA}(x_i))$, $i = 1, 2, \ldots, n$. This way, the linearized tolerance interval is derived from the measurement uncertainty u_0^{LRA}. Holds that $a_L = a_U = 3\min(u_0^{LRA}(x_i))$ for each point of moderate scale x_i, $i = 1, 2, \ldots, n$. The width of the acceptance interval ΔA when it holds that is $[A_L, A_U] \subset [T_L, T_U]$ amounts to $\Delta A = 0.8\Delta T$. For $[T_L, T_U] \subset [A_L, A_U]$, it holds that $\Delta A = 1.2\Delta T$. In both cases, the width of the guard band is equal to $w = 0.1\Delta T$. Detailed formulas for the calculation of the global risk for the producer and consumer, taking into account all the aforementioned parameters, can be found in [10]. These are omitted here for simplicity.

3 RESULTS AND DISCUSIONS

As well as in regression models, in risk assessment models for calibration the risk curves along the moderate scale are parabolas that are open upward, Fig. 3 [10]. The minimum is achieved for x_{min} = 0.3317 µm and is located at the point of intersection of the regression line from Eq. (7) and the line $y = x$. Since measurement uncertainty is always higher at the edges of the scale, the risk of calibrating is also higher at the edges. From Fig. 3, it can also be observed that the global risk for both the producer and the consumer is higher on the left side of the moderate scale. The global producer's risk values range from 0.05% to 15.1%, while the global consumer's risk values range from 0.2% to 5.3%.

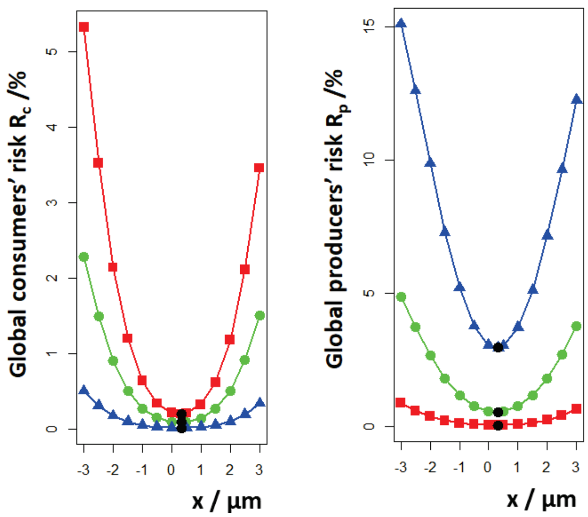

Figure 3 Risk curves along a moderate scale: —■— $[T_L, T_U] \subset [A_L, A_U]$, —●— $[T_L, T_U] = [A_L, A_U]$, —▲— $[T_L, T_U] \supset [A_L, A_U]$, ● Minimum point

Along the guard band axis, the global producer's risk increases, while the global consumer's risk decreases, as shown in Fig. 4.

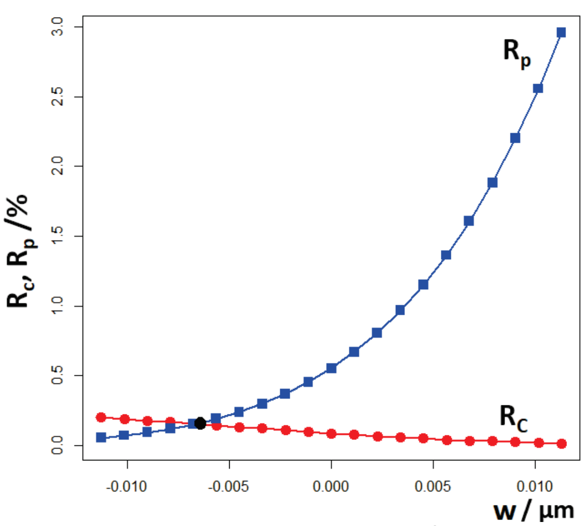

Figure 4 The risk curves along a guard band axis: —●— Global consumer's risk R_c, —■— Global producer's risk R_p, ● Intersection point

In Fig. 4, the values of global producer and consumer risk evaluated at the minimum point x_{min} are presented. Thus, these curves are commonly referred to as minimum curves. Negative values of the guard band width w indicate the situation where is $[T_L, T_U] \subset [A_L, A_U]$, while positive values correspond to the situation where is $[T_L, T_U] \supset [A_L, A_U]$. For $w = 0$, it holds that $[T_L, T_U] = [A_L, A_U]$. The risk curves shown in Fig. 4 intersect at $w \approx -0.0067$ µm, which corresponds to $\Delta A \approx 1.12\Delta T$. In this case, the global producer's and consumer's risk are equal, and both amount to $R_p = R_c \approx 0.15\%$.

Finally, the behaviours of the risk curves along the moderate scale and the guard band axis can be clearly discerned by observing the risk surfaces depicted in Fig. 5.

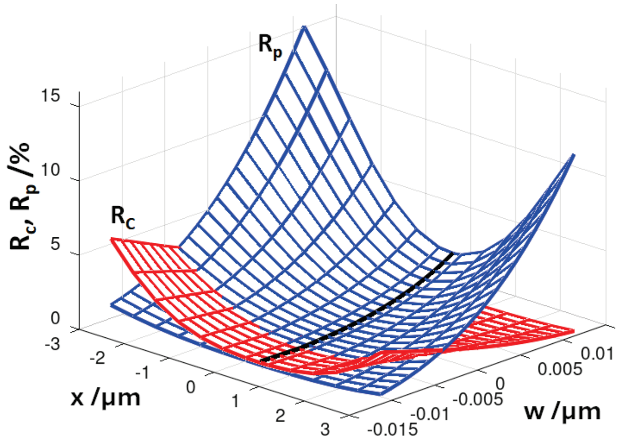

Figure 5 Risk surfaces: — Global consumer's risk R_c, — Global producer's risk R_p, — Curves of a minimum

The risk surfaces from Fig. 5 are presented on a discretized domain bounded by the intervals [–3, 3] µm on the moderate scale and by the interval [–0.011, 0.011] µm on the guard band axis. The highlighted lines on the risk surfaces in Fig. 5 correspond to the minimum curves presented in Fig. 4.

4 MODEL EVALUATION

The risk assessment model for calibration can be evaluated using confusion matrices by incorporating the conformance probability p_C in the calculations. The conformance probability indicates the degree of adherence to specified requirements [5]. When both the prior and likelihood functions follow a normal distribution, as noted in [5, 10], the conformance probability p_C for the values x_{e_i}, i = 1, 2, …, n is calculated using the equation:

$$p_{C_i} = \frac{1}{u_0^{LRA}\sqrt{2\pi}} \int_{T_{L_i}}^{T_{U_i}} \exp\left[-\frac{1}{2}\left(\frac{\eta - x_{e_i}}{u_0^{LRA}}\right)^2\right] \mathrm{d}\eta. \tag{13}$$

The behaviour of the conformance probability along the guard band axis is described by straight lines. These lines are parallel to the guard band axis. The line of maximum values follows the equation $p_C = 0.9972$ and is achieved at the points (x_{min}, w), where $w \in [-0.011, 0.011]$ µm. Along the moderate scale axis, the conformance probability curves are parabolas opening downward, with maxima at the moderate scale point x_{min}, as shown in Fig. 6.

If the true value of the characteristic of an item of interest falls within the tolerance interval, and the measured value falls within the acceptance interval, this constitutes a valid acceptance of an item of interest, i.e. a true positive value (*TP*). Conversely, if the true value of an item of interest is outside the tolerance interval and at the same time the measured value is outside the acceptance interval, it is about valid rejection, i.e. a true negative value (*TN*). The global producer's risk R_P corresponds to the probability of false rejection (FR), while the global consumer's risk R_C represents the probability of false acceptance (FA). The classes *TP*, *TN*, R_C, and R_P represent the elements of the confusion matrix. Holds that $TP = p_C - R_p$, $TN = 1 - p_C - R_C$ [7, 15].

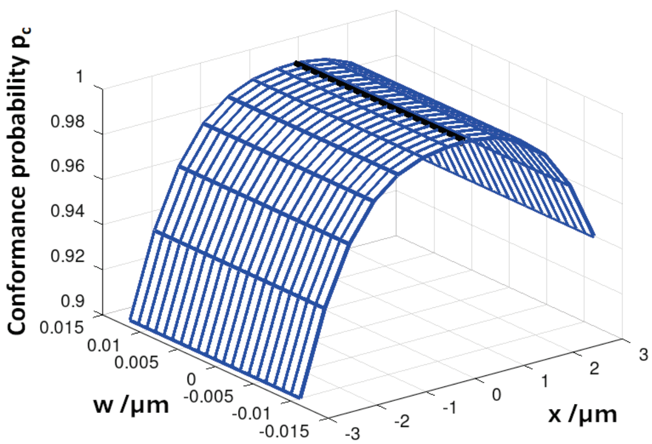

Figure 6 Conformance probability surface, — line of maximum values

Establishing a connection between conformance probability and classes of the confusion matrix allows the evaluation of models for risk assessment in calibration using confusion matrix-based metrics. The confusion matrix is defined based on the true value and the measured value, considering both the tolerance interval and the acceptance interval [7, 15]. The calibration of an inductive contact probe's risk assessment model was evaluated using standard metrics: accuracy, precision, recall, and F1 score, as illustrated in Fig. 7 [16].

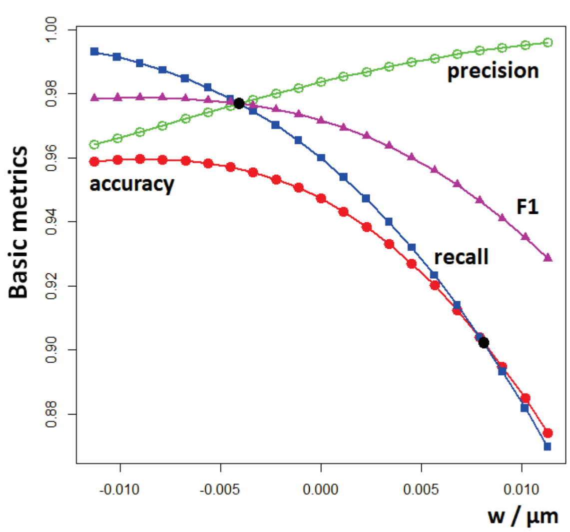

Figure 7 Basic metrics: —●— accuracy, —■— recall, —▲— F1, —○— precision

The metrics behaviours presented in Fig. 7 correspond to the case where $[T_L, T_U] \supset [A_L, A_U]$. This model is most considered in practice due to its consumer-favourable nature. In this scenario, the consumer's risk is minimized, while the producer's risk is maximized. Accuracy represents the ability of the model to classify *TP* and *TN* values concerning the total amount of data. This metric is not suitable for unbalanced data due to the prevalence of the *TP* class [17]. Recall is defined as the probability that true positive values fall within the tolerance interval. Due to its dependence on the global producer's risk, recall decreases along the guard band axis [15]. For $w \approx 0.008$ μm it holds that $accuracy = recall \approx 0.9$ and $\Delta A \approx 0.85\Delta T$. Precision measures a model's ability to identify true positive values that fall within the acceptance interval. Since precision is influenced by global consumer's risk, it diminishes along the guard band axis [15]. To avoid the trade-offs between precision and recall, the F1 score metric is considered. $F1$ score is the harmonic mean of precision and recall. It also measures model accuracy but is more suitable for imbalanced data [17]. $F1$ score shadowing standard accuracy, attaining higher values, as shown in Fig. 7. The precision, recall and F1 score intersect each other when it holds that $R_C = R_p$. This occurs when the width of the guard band is equal to $w \approx -0.004$ μm. Then holds $R_C = R_p \approx 2.17\%$ and $\Delta A \approx 1.07\Delta T$. Under these conditions, the metric values are $recall = precision = F1 = 0.9769$. The maximum values of accuracy metrics and $F1$ score metrics are achieved at the moderate scale point x_{min} where is $accuracy = 0.9961$ and $F1 = 0.9980$. Fig. 8 presents the surface of the $F1$ score metrics, highlighting the curve of maximum values.

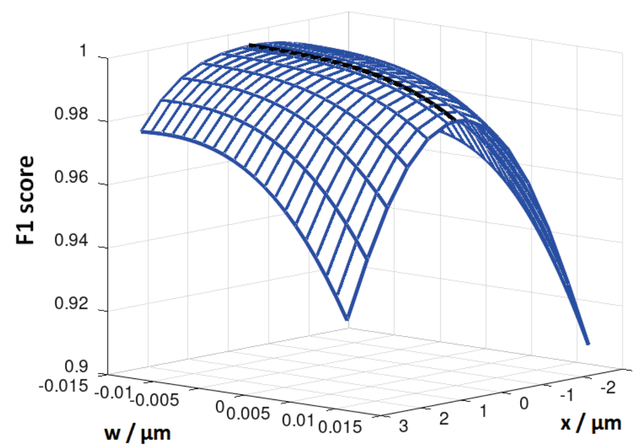

Figure 8 F1 score surface

All metrics on the domain exhibit high values, with probabilities ranging from 0.83 to 1. Additionally, the values for conformance probability are also high, ranging from 0.90 to 0.9972.

The outlined procedure for risk assessment in calibration identifies areas of highest and lowest risk, enabling a quantitative evaluation of the calibration process. It can be concluded that the greatest risk in calibrating the inductive contact probe occurred for the measurements obtained on the left side of the moderate scale in the model where $[T_L, T_U] \supset [A_L, A_U]$.

5 CONCLUSIONS

This study has described a method for assessing the risk of making wrong decisions during calibration. So far, such a risk assessment procedure for calibration in metrology has not yet been applied. The significance and practical effects of this risk assessment method for calibration should be viewed in the context of ensuring accurate and reliable measurements. A brief overview of risk assessment for the trivial, linearized model is provided, utilized alongside measurement uncertainty derived from extensive data on measurement performance. Emphasis is placed on the description of the linearized model where the measurement uncertainty is calculated from the available data about the calibration curve. It is demonstrated that this method can identify those areas of the domain where the risks incurred during the measurement are the highest. The evaluation of the model was carried out according to the confusion matrix-based metrics. The quality of the calibration procedure is indicated by the high values of standard metrics. Beyond metrology, this calibration risk assessment model could be applied in chemistry, medicine, economics, and other fields. For future work, it remains to describe a non-linearized risk assessment model for calibration.

6 REFERENCES

[1] Separovic, L., de Godoy Bertanha, M. L., Saviano, A. M., & Lourenço, F. R. (2020). Conformity decisions based on measurement uncertainty—a case study applied to agar diffusion microbiological assay. *Journal of Pharmaceutical Innovation*, *15*, 110-115. https://doi.org/10.1007/s12247-019-09374-8

[2] Runje, B., Horvatic Novak, A., Razumic, A., Piljek, P., Strbac, B., & Orosnjak, M. (2019, October). Evaluation of consumer and producer risk in conformity assessment decisions. In *Proceedings of the 30th DAAAM international symposium* (pp. 0054-0058). https://doi.org/10.2507/30th.daaam.proceedings.007

[3] Božić, D., Runje, B., & Razumić, A. The Cost of Wrong decisions. In *The 19th International Conference" Laboratory Competence-Brijuni 2024"*.

[4] Hibbert, D. B., Korte, E. H., & Örnemark, U. (2021). Metrological and quality concepts in analytical chemistry (IUPAC Recommendations 2021). *Pure and Applied Chemistry*, *93*(9), 997-1048. https://doi.org/10.1515/pac-2019-0819

[5] BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPAP, and OIML. Evaluation of measurement data — The role of measurement uncertainty in conformity assessment. Joint Committee for Guides in Metrology, JCGM 106:2012. https://doi.org/10.59161/jcgm106-2012

[6] Lira, I. (1999). A Bayesian approach to the consumer's and producer's risks in measurement. *Metrologia*, *36*(5), 397. https://doi.org/10.1088/0026-1394/36/5/1

[7] Božić, D., & Runje, B. (2022, October). Selection of an Appropriate Prior Distribution in Risk Assessment. In *Proceedings of the 33rd International DAAAM Virtual Symposium "Intelligent Manufacturing & Automation", Vienna, Austria* (pp. 26-27). https://doi.org/10.2507/33rd.daaam.proceedings.066

[8] Toczek, W., & Smulko, J. (2021). Risk Analysis by a Probabilistic Model of the Measurement Process. *Sensors*, *21*(6), 2053. https://doi.org/10.3390/s21062053

[9] Koucha, Y., Forbes, A., & Yang, Q. (2021). A Bayesian conformity and risk assessment adapted to a form error model. *Measurement: Sensors*, *18*, 100330. https://doi.org/10.1016/j.measen.2021.100330

[10] Božić, D., Runje, B., & Razumić, A. (2024). Risk Assessment for Linear Regression Models in Metrology. *Applied Sciences*, *14*(6), 2605. https://doi.org/10.3390/app14062605

[11] Heckert, N. A., Filliben, J. J., Croarkin, C. M., Hembree, B., Guthrie, W. F., Tobias, P., & Prinz, J. (2002). Handbook 151: Nist/sematech e-handbook of statistical methods. Retrieved from http://www.itl.nist.gov/div898/handbook/mpc/mpc.htm

[12] Bednjanec, F. (2016). Umjeravanje uređaja za mjerenje kružnosti (Diplomski rad). Zagreb: Sveučilište u Zagrebu, Fakultet strojarstva i brodogradnje. Retrived from https://urn.nsk.hr/urn:nbn:hr:235:701539 (in Croatian)

[13] Subhi, K. T., & Al Azkiya, A. (2022). Comparison of Cochrane-Orcutt and Hildreth-Lu methods to overcome autocorrelation in time series regression (case study of Gorontalo province HDI 2010-2021). *Parameter: Journal of Statistics*, *2*(2), 30-36. https://doi.org/10.22487/27765660.2022.v2.i2.15913

[14] Young, D. S. (2018). *Handbook of regression methods.* Chapman and Hall/CRC.

[15] Božić, D., Runje, B., Lisjak, D., & Kolar, D. (2023). Metrics related to confusion matrix as tools for conformity assessment decisions. *Applied Sciences*, *13*(14), 8187. https://doi.org/10.3390/app13148187

[16] Tharwat, A. (2021). Classification assessment methods. *Applied computing and informatics*, *17*(1), 168-192. Retrieved from https://www.emerald.com/insight/publication/issn/2634-1964/vol/17/iss/1

[17] Jeni, L. A., Cohn, J. F., & De La Torre, F. (2013, September). Facing imbalanced data--recommendations for the use of performance metrics. In *2013 Humaine association conference on affective computing and intelligent interaction* (pp. 245-251). IEEE. https://doi.org/10.1109/ACII.2013.47

Authors' contacts:

Dubravka Božić, mag. math.
(Corresponding author)
Faculty of Mechanical Engineering and Naval Architecture,
University of Zagreb,
Ivana Lučića 5, 10000 Zagreb, Croatia
dubravka.bozic@fsb.unizg.hr

Biserka Runje, PhD, Prof.
Faculty of Mechanical Engineering and Naval Architecture,
University of Zagreb,
Ivana Lučića 5, 10000 Zagreb, Croatia
biserka.runje@fsb.unizg.hr

Andrej Razumić, PhD
Department of Polytechnics,
Dr. Franjo Tuđman Defense and Security University,
Ilica 256b, 10000 Zagreb, Croatia
andrej.razumic@sois-ft.hr