Available online at www.sciencedirect.com
ScienceDirect
Procedia CIRP 27 (2015) 143 – 148
www.elsevier.com/locate/procedia

13th CIRP conference on Computer Aided Tolerancing

A criterion for comparing measurement results and determining conformity with specifications

Hao Ding, Paul J. Scott* and X. Jiang

EPSRC Innovative Manufacure Research Centre in Advanced Metrology, Centre for Precision Technologies, School of Computing and Engineering, University of Huddersfield, Queensgate, Huddersfield, HD1 3DH, UK

* Corresponding author. Tel.: +44-01484-473953; E-mail address: p.j.scott@hud.ac.uk.

Abstract

In this paper a new criterion for comparing measurement results and determining conformity with specifications is proposed, which essentially is a strategy of estimating the empirical relationships of objects. Comparing with traditional methods given in GUM: 2008 and ISO 14253-1, this criterion improves the resolution of comparison by reducing the sizes of the coverage intervals to be compared. Interval order (a binary relation) is used for comparing the coverage intervals of the measurand and represents the empirical relations. The systematic effects of measurement are classified into two types: monotonic and non-monotonic effects, so that, without correcting the monotonic effects, a biased measurand can be specified to represent the empirical relations. Thereby the uncertainty components arising from the monotonic effects can be removed from the combined uncertainty. A strategy is given for determining the relationships among measurement results and specification limits. An example is given to demonstrate the application of the criterion.

Peer-review under responsibility of the organizing committee of 13th CIRP conference on Computer Aided Tolerancing

Keywords: Conformity with specifications; measurement uncertainty; interval order; systematic effects; montonic effects

1. Introduction

It is well known that the objective of measurement is to obtain the values of the measurand (the quantity to be measured) [1]. According to the representational theory of measurement [2], the values of the measurand and their numerical relations are used to represent the (measured) objects and their empirical relations. Hence the implicit objective of obtaining the values of the measurand is to compare the objects and estimate the empirical relation.

Most researchers in metrology are focused on the former objective (i.e. to obtain the values of the measurand), since the latter objective (i.e. to estimate the empirical relation) can be achieved by using the measurement results. However, to get the best estimations of the measurand, the calculation can be quite complicated, and the measurement uncertainties can be very large (e.g. larger than 30% of the size of the tolerance). In this paper, we focus on the latter objective, and investigate whether it can be achieved with small uncertainties and simpler calculation.

An important reason of achieving the former objective is we need to compare the measurement results with the specifications and determine the conformity. ISO 14253-1 [3] together with GUM [1] provides a method for determining conformity with specifications when uncertainty is involved. In that method, a *complete measurement result* is expressed as $Y = y \pm U$, where y is the estimate of the value of measurand, U is the expanded uncertainty with a stated level of confidence (e.g. 95%). Here Y is taken as a *coverage interval* (*CI*), which is an interval containing the value of a measurand with a stated probability [4]. The conformity with specification is determined by the relation of the CI and specification limits.

However, if we take the specification limits as the values of the measurand of some objects, called *limit samples*, determining the conformity with specification is the same as estimating the empirical relations between the measured objects and the limit samples, which is consistent with the latter objective. Hence the criterion given in this paper is

Peer-review under responsibility of the organizing committee of 13th CIRP conference on Computer Aided Tolerancing
doi:10.1016/j.procir.2015.04.057

essentially a strategy of estimating the empirical relations of the set objects including the limit samples.

Nomenclature

x_i	An influence quantity
y	The measurand
y'	The biased measurand
f	The functional relational between the influence quantities and the measurand
$u(y)$	The standard uncertainty of the measurand
CI(y)	The coverage interval of the value of the measurand
$E(x)$	The expected value of a random variable x
$\prec$	Interval order, strictly smaller
$\sim$	Indifference relation

2. Relation between coverage intervals

To estimate the empirical relation of several objects, we need to define the relation of their measurement results, which is the relation of the corresponding CIs (i.e. intervals on the real axis) when the measurement uncertainty is significant. The relation between single real numbers, named simple order $\leq$, is not suitable to define the relation between intervals.

When the CIs A, B of two values a, b intersect with each other (i.e. $A \cap B \neq \emptyset$), both $a \leq b$ and $b \leq a$ are possible. In this situation A, B can be considered as 'being same up to a small error', which is a kind of relation called indifference [5], written as $A \sim B$. This relation is intransitive ($A \sim B$ and $B \sim C$, do not imply $A \sim C$) and symmetric ($A \sim B$ implies $B \sim A$). The intransitive property makes $\sim$ different from the equivalent relation, =.

If A, B do not intersect, then one interval is strictly smaller than the other. This relation can be defined by a binary relation on the set of CIs, named interval order [6]. Interval order $\prec$ is irreflexive (for any interval A, not $A \prec A$) and satisfies the defining property: if $A \prec B$ and $C \prec D$ then $A \prec D$ or $C \prec B$. For an interval order, the indifference relation is necessary to be defined together: if neither $A \prec B$ nor $A \prec B$, then $A \sim B$.

Base on the interval order of the CIs, the relation of the values of the measurand, which is a representation of the empirical relation of the measured objects, can be clearly estimated under a stated probability. So the interval order of the complete measurement results is also an estimation of the empirical relation.

For example, let A, B, C be the CIs of three objects α, β, γ respectively. If, under 95% confidence level, A, B, C are distributed as shown in fig.1, their relation can be written as $A \sim B \sim C$, $A \prec C$.

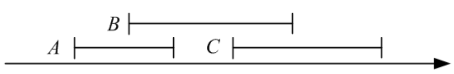

Fig. 1. The distribution of CIs *A, B, C*

$A \sim B$ means, under 95% confidence level, no inference can be given on the empirical relation between α, β. And $A \prec C$ means, with 95% confidence, α is strictly smaller than γ.

3. Principle of estimating of the empirical relations

3.1. Resolution of comparison

It is obvious that the size of CIs (the length of intervals) may affect the interval order, and thus affect the estimation of the empirical relation. Consider some CIs of an identical size, if this size is large (e.g. 30% of the maximum difference of their estimated values of the measurand), the adjacent CIs are quite likely to intersect with each other. When two CIs intersect, the empirical relation will be considered as indifference, i.e. not able to be identified. So, for estimating empirical relations, the sizes of CIs are preferred to be smaller.

For this reason, we call the average size of the CIs of a set of measurement results as the *resolution of comparison*. A main objective of this paper is to improve the resolution of comparison.

3.2. Monotonic systematic effects

Measurement uncertainties can be either classified according to their evaluation methods (statistical or non-statistic) into Type A and Type B, or by the sources of the uncertainties. The latter way classifies uncertainties as the following two types.

- Random uncertainty components: the uncertainties arise from the random effects;
- Systematic uncertainty components: the uncertainties arise from incomplete knowledge of the systematic effects.

According to GUM [1], a random effect is the effect of stochastic or unpredictable variations of influence quantities; and a systematic effect is a recognized effect of an influence quantity on a measurement result. Hence both effects cause some deviation of the measured value from the value of measurand. Each effect can be taken as a function in terms of

$$\eta(y, x_i) = y', \tag{1}$$

where y is the measurand, x_i is a related influence quantity, y' is a quantity deviated from the measurand due to the random/ systematic effect.

In practice, it can be difficult to distinguish the two types of effects very clearly. But, since in replicate measurements the systematic error arise from systematic effect remains constant or varies in a predictable manner [4], a systematic effect itself is a *deterministic function*. In contrast, a random effect is a random function, and the related x_i is always a random variable.

Moreover, under the repeatability conditions given in GUM ([1] B.2.15), some influence quantities of systemic effects are always *fixed* during the measurements of all the measured objects. These quantities are constants, although the exact values are unknown due to incomplete knowledge. Hence, for a fixed influence quantity x_i, equation (1) can be taken as a function of measurand, denoted as $\eta_{xi}(y) = y'$, or $\eta(y) = y'$, if it is clear what x_i is.

In most cases, $\eta(y)$ is an increasing function of the measurand. For instance, the effect of imperfect calibration of a gauge can be written as $\eta_c(y) = y + c$, where c is a constant (but unknown) offset error. The effect of incomplete knowledge of the sensitivity of the instrument, which gives rise to the sensitivity error, is in the form of $\eta_a(y) = ay$, where a is unknown constant close to 1. The effect of resolution or digital rounding is in the form of $\eta_b(y) = 10^{-b}\lfloor 10^b y + 0.5 \rfloor$, where b is a integer, $\lfloor\,\rfloor$ is the floor function. The above functions of systematic effects are all (monotonically) increasing. We define this type of systematic effects as monotonic effects.

Definition: A systematic effect $\eta(y, x_i)$ is called a *monotonic effect*, if it is an increasing function of the measurand y, and x_i is fixed as a constant in the measurements of all the objects.

That means for any monotonic effect $\eta(y, x_i)$, we have

$$y_1 \le y_2 \Rightarrow \eta_{xi}(y_1) \le \eta_{xi}(y_2), \quad (2)$$

where y_1, y_2 are two arbitrary values of the measurand.

The relation between y_1 and y_2 is a presentation of the empirical relation of the corresponding objects. Monotonic effects preserve the relation of y_1 and y_2, thus they also preserve empirical relations.

With this definition, systematic effects are classifies into two types: monotonic effects and non-monotonic effects. A monotonic effect may become non-monotonic when the measurement method changes. For example, if the temperature of the objects is an influence quantity, it may be fixed or changing depends on the measurement environment. So to classify the effects, the actual situation of measurement should be fully understood.

Correspondingly, the uncertainty components arise from these effects can be further classified according to their sources as shown in fig. 2. For example, monotonic uncertainty components are the uncertainties arise from monotonic effects.

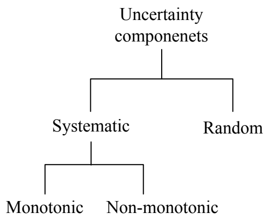

Fig. 2 Classification of uncertainty components by the sources

3.3. Monotonic uncertainty components

To estimate the value of measurand, all the systematic effects should be corrected from the observed data. But for estimating the empirical relation, it's not necessary to correct the monotonic effects, because, monotonic effects preserve empirical relations. As shown in equation (2), although $\eta(y_1)$ and $\eta(y_2)$ consists the systematic error rise from the monotonic effect, they still reflect the relation of y_1 and y_2. Thus, without correcting the monotonic (systematic) effects, we can find a quantity, $y' = \eta(y)$, named the *biased measurand*, to estimate the empirical relation.

This is also true when the empirical relation is represented by the interval order of CIs. For example, let η be a monotonic effect, $\eta(y) = ay + b$, where *a, b* are positive real numbers, and let the relation of the CIs Y_1, Y_2 and Y_3 of three measurement results be $Y_3 \sim Y_2, Y_2 \prec Y_1, Y_3 \prec Y_1$. As shown in figure 3, $\eta(Y_3) \sim \eta(Y_2), \eta(Y_2) \prec \eta(Y_1), \eta(Y_3) \prec \eta(Y_1)$, η does not change the relation of the CIs.

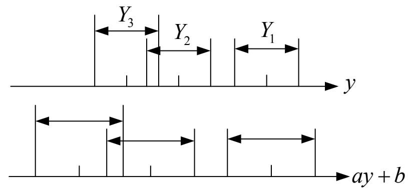

Fig. 3 The relations of the CIs with and without monotonic effects

Proposition: Let Y_1, Y_2 be the CIs of y_1 and y_2, if $\eta : y \to y'$ is a monotonic effect, then

$$\eta(Y_1) \prec \eta(Y_2) \Rightarrow Y_1 \prec Y_2 .$$

See appendix A for the proof.

That means the interval order of the CIs of the biased measurand can be used to estimate the interval order of the CIs of the measurand, and thus estimate the empirical relation. The CIs of the biased measurand is smaller in size than the CIs of the measurand, because the monotonic uncertainty components are not included in the former CIs. So the resolution of comparison is improved by using the CIs of the biased measurand.

It can be proved that $\eta(Y_1) \Box \eta(Y_2)$ does not imply $Y_1 \Box Y_2$, both $Y_1 \prec Y_2$ and $Y_2 \prec Y_1$ are possible. But similar to $Y_1 \Box Y_2$, it means no inference on the empirical relation can be given under the confidence level.

3.4. Strategy of estimating empirical relations

To determine the conformity with a specification, we need to compare the measurement results with the specification limits. Traditionally, the measurement results should be corrected for all the recognized systematic effects before the comparison (see fig. 4). Conversely, without correcting the monotonic effects, we can specify a biased measurand y', and estimate the CIs of y' of the limit samples according to the monotonic effects and their uncertainties; and then compare the CIs of the limit samples with the CIs of measurement results (see fig. 5).

Figure 5 demonstrates the principle of improving the resolution of comparison: due to the order-preserving property of monotonic effects, we can use the biased measurand instead of the measurand to estimate the empirical relation, so that the sizes of the CIs to be compared can be reduced.

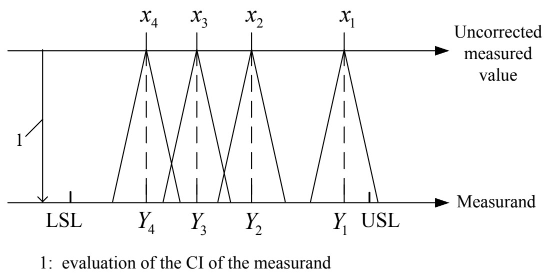

Fig. 4 The traditional way of determining the conformity with a spec.

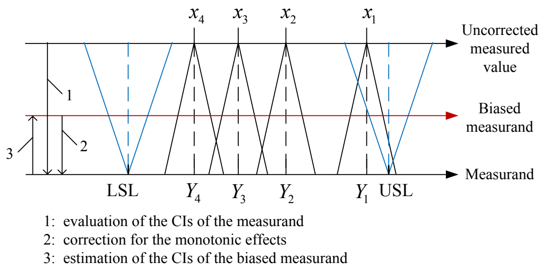

Fig. 5 The amended way of determining the conformity with a spec.

Following this principle, the strategy of estimating empirical relations is summarized as following.

1. Express the measurand in terms of a function of the influence quantities, such as
$$y = f(x_1, x_2, ..., x_n)\,. \tag{3}$$
All the significant errors and corrections should be included in the function.
2. According to the actual situation of the measurement, sort out the influence quantities which are fixed as a constant in the replicate measurements of all the objects.
3. Move the fixed influence quantities to the LHS of the equation (3), and get a new equation. Specify a biased measurand y' with the LHS of the new equation, which should consist only of the fixed influence quantities.
4. For each measured object, evaluate the expected value and the expanded uncertainty of the biased measurand with the RHS of the new equation.
5. For the specification limits, take them as the values of the measurand, and use the LHS of the new equation to estimate the expected values and the expanded uncertainties of the biased measurand.
6. Use interval order to describe the relation of all the CIs of the biased measurand, and according the interval order to estimate the empirical relation and decide the conformity with the specification.

This strategy together with the concept of using interval order to describe the relation of complete measurement results is the criterion to be proposed in this paper.

4. An example of measuring end gauges

End gauge calibration is an example of uncertainty evaluation given in GUM ([1] H.1). Here three end gauges, named *a, b, c*, are of the same specification: 50mm +0.001/−0 mm at 20°C. They are measured to determine the conformity with the specification and to find out their ordered relation in length. This example demonstrates how to implement the proposed criterion to a dimensional measurement.

The end gauges are measured by comparing them with a calibrated standard gauge of the same nominal length. The difference of length d is measured by a comparator. As shown in the example in GUM, with the effect of thermal expansion, the measurand, i.e. length of the end gauges at 20°C, can be expressed as the following function:

$$l = f(l_S, d, \alpha_S, \theta, \delta\alpha, \delta\theta) = l_S + d - l_S(\delta\alpha \cdot \theta + \alpha_S \cdot \delta\theta)$$

where l is the measurand; l_S is the length of the standard gauge given in its calibration certificate; d is the difference of length; α and α_S are the thermal expansion coefficients of the end gauge and the standard gauge respectively, and $\delta\alpha = \alpha - \alpha_S$; θ and θ_S are the deviations in temperature from 20°C, respectively, of the end gauge and the standard gauge, and $\delta\theta = \theta - \theta_S$.

The arithmetic mean of the readings of the comparator $\bar{d}$ and the actual difference d can be related by the following equation.

$$\bar{d} = d + d_1 + d_2$$

where d_1 and d_2 are quantities describing, respectively, the random and the systematic effects of the comparator. From the above two equations, we obtain

$$l = l_S + \bar{d} - d_1 - d_2 - l_S(\delta\alpha \cdot \theta + \alpha_S \cdot \delta\theta) \tag{4}$$

All the expected values, uncertainties and probability distributions of the influence quantities of l are known and given in table 1. For comparing with the classical method, we use the data given in the example of GUM. And for simplicity, the degrees of freedom of the Type B uncertainty components are assumed to be infinite.

The values of l_S and α_S are always fixed, since there is only one standard gauge in the measurements. We can also assume that the systematic error of the comparator, d_2 is fixed during the measurements. $\delta\alpha, \theta$ and $\delta\theta$ are related to systematic effects, but they are not fixed. θ and $\delta\theta$ vary with time; $\delta\alpha$ can be different for different end gauges. So l_S, d_2 and α_S are related to monotonic effects, where α_S is in a nonlinear term, it cannot be moved to the LHS of (4) alone. By moving l_S and d_2, we obtain a biased measurand l'.

$$l' = l - l_S + d_2 = \bar{d} - d_1 - l_S(\delta\alpha \cdot \theta + \alpha_S \cdot \delta\theta) \tag{5}$$

Since the expected values of d_1, $\delta\alpha$ and $\delta\theta$ are 0, from (5), we get

$$E(l') = \bar{d}\,, \tag{6}$$

where $E(l')$ is the expected value of l'.

x_i	Expected/ mean value	$u(x_i)$	Source of uncertainty	Value of standard uncertainty	Probability distribution	$c_i \equiv \partial f / \partial x_i$	Degree of freedom
l_S	50.000623 mm	$u(l_S)$	Monotonic systematic effect	25 nm	Normal	1	∞
$\bar{d}$	215 nm	$u(\bar{d})$	Random effect	5.8 nm	Normal	1	9
d_1	0 nm	$u(d_1)$	Random effect	3.9 nm	Normal	1	∞
d_2	0 nm	$u(d_2)$	Monotonic systematic effect	6.7 nm	Normal	1	∞
α_S	11.5 x10^{-6} °C^{-1}	$u(\alpha_S)$	Monotonic systematic effect	1.2x10^{-6} °C^{-1}	Rectangular	0	∞
θ	-0.1°C	$u(\theta)$	Non-monotonic systematic effect	0.41°C	Rectangular	0	∞
$\delta\alpha$	0°C^{-1}	$u(\delta\alpha)$	Non-monotonic systematic effect	0.58 x10^{-6} °C^{-1}	Triangular	$-l_S\theta$	∞
$\delta\theta$	0°C	$u(\delta\theta)$	Non-monotonic systematic effect	0.029°C	Triangular	$-l_S\alpha_S$	∞

Table 1 Summary of standard uncertainty components

Based on a first-order Taylor series approximation of equation (5), the combined standard uncertainty of l' can be evaluation by the following equation (refer to GUM for the detail of the evaluation method).

$$u_C^2(l') = u^2(\bar{d}) + u^2(d_1) + (l_S\theta)^2 u(\delta\alpha) + (l_S\alpha_S)^2 u(\delta\theta)$$
$$= u^2(\bar{d}) + 299.2\text{nm}^2$$

Due to the nonlinear of (5), the following second-order terms in the Taylor series of (5) are significant, which should be added to $u_C^2(l')$.

$$l_S^2 u^2(\delta\alpha) u^2(\theta) + l_S^2 u^2(\alpha_S) u^2(\delta\theta) = 139.8\text{nm}^2$$

So we have

$$u_C^2(l') = u^2(\bar{d}) + 299.2\text{nm}^2 + 139.8\text{nm}^2 . \quad (7)$$

Ten replicate measurements are taken for each end gauge. The values of $\bar{d}$ and $u(\bar{d})$ of the measured gauges are listed as following.

Gauge	$\bar{d}$	$u(\bar{d})$
a	215 nm	5.8 nm
b	91 nm	13.4 nm
c	254 nm	9.3 nm

Substitute the data of $\bar{d}$ and $u(\bar{d})$ into (6) and (7), we obtain the following results.

Gauge	$E(l')$	$u_C(l')$
a	215 nm	22 nm
b	91 nm	25 nm
c	254 nm	23 nm

Comparing with the result of combined standard uncertainty of the measurand obtained in GUM $u_C(l) = 34\text{nm}$, $u_C(l')$ of gauge *a* (with the same value of every influence quantity), 22nm is much smaller.

The effective degree of freedom of $u_C(l')$, $v_{\text{eff}}(l')$ can be obtained from the following Welch-Satterthwaite formula [7].

$$v_{\text{eff}}(l') = \frac{u_C^4(y)}{\sum_{i=1}^{n} \frac{u_i^4(y)}{v_i}} = \frac{u_C^4(l')}{\frac{u^4(\bar{d})}{9} + 0 + 0 + 0} = 9\frac{u_C^4(l')}{u^4(\bar{d})}$$

The values of $v_{\text{eff}}(l')$ of the three end gauges are all above 100, so we can take the coverage factor k = 2, providing a coverage probability of approximately 95%. The CIs of the three end gauges can be stated as below.

$$\text{CI}(l')_a = (215 \pm 44)\text{nm}$$
$$\text{CI}(l')_b = (91 \pm 50)\text{nm}$$
$$\text{CI}(l')_c = (254 \pm 46)\text{nm}$$

Moreover, the expected values and the standard uncertainties of the upper and lower specification limits (USL & LSL) can be evaluated as following.

$$E(l')_{USL} = E(l - l_S + d_2) = 377\text{nm}$$
$$E(l')_{LSL} = E(l - l_S + d_2) = -623\text{nm}$$
$$u(l')_{USL} = u(l')_{LSL} = (u^2(l_S) + u^2(d_2))^{\frac{1}{2}} = 26\text{nm}$$

Take k = 2 as the coverage factor, we have

$$\text{CI}(l')_{USL} = (377 \pm 52)\text{nm}$$
$$\text{CI}(l')_{LSL} = (-623 \pm 52)\text{nm}$$

Put the CIs of the three end gauges and the specification limits together. Their relation can be observed from the graph below.

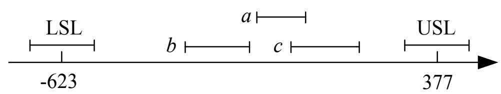

Fig. 6 The relation of the end gauges and spec. limits (not to scale)

So the relation of the end gauges and the specification limits can be stated with interval order as LSL≺*b*≺*a*≺USL, *a* ~ *c* and LSL≺*b*≺*c*≺USL. The three end gauges are all within specification.

4. Conclusion

The criterion proposed in this paper is designed for estimating the empirical relation of measured objects and determining the relation of the specification and the measured objects when measurement uncertainty is significant. It provides a method of defining the relation between complete measurement results by taking measurement results as coverage intervals. Moreover, it provides a strategy to reduce the size of the intervals by ignoring uncertain type of uncertainty components, which makes the estimated relation more meaningful without introducing any bias. The principle of ignoring uncertain type of uncertainty components is explained by introducing a concept called monotonic effect, which further classified the concepts of systematic effects and systematic uncertainty components.

This criterion can be quite useful for the following situations: the measurement uncertainty is very significant or too large such that the measurement results are not very meaningful; the specification is given by some standard samples instead of numbers. It is a universal method, and can be applied to many areas of metrology, such as to classify objects into different classes (e.g. A, B, C, D) according to the measurand.

Acknowledgements

The authors gratefully acknowledge the UK's Engineering and Physical Sciences Research Council (EPSRC) funding of the EPSRC centre for Innovative Manufacturing in Advanced Metrology (Ref: EP/I033424/1) and the funding of the EPSRC Fellowship in Manufacturing (Ref: EP/K037374/1).

Appendix A. Proof of the proposition

Let $Y_1 = [a,b], Y_2 = [c,d]$, where a,b,c,d are some real constants, then

$$Y_1 \prec Y_2 \Leftrightarrow b < c \,. \tag{8}$$

By definition, η is an increasing function, so $\eta(Y_1) = [\eta(a),\eta(b)]$ and $\eta(Y_2) = [\eta(c),\eta(d)]$, hence

$$\eta(Y_1) \prec \eta(Y_2) \Leftrightarrow \eta(b) < \eta(c) \,. \tag{9}$$

Since η is increasing, $b < c$ implies $\eta(b) \le \eta(c)$, and $b \ge c$ implies $\eta(b) \ge \eta(c)$, and either $b \ge c$ or $b < c$, thus

$$\eta(b) < \eta(c) \Rightarrow b < c \,. \tag{10}$$

By (8), (9) and (10), we obtain

$$\eta(Y_1) \prec \eta(Y_2) \Leftrightarrow \eta(b) < \eta(c) \Rightarrow b < c \Leftrightarrow Y_1 \prec Y_2 \,.$$

So the proposition is proved.

References

[1] JCGM 100. Evaluation of measurement data — Guide to the expression of uncertainty in measurement (GUM). 1st ed. 2008.

[2] Krantz DH, Luce RD, Suppes P, Tversky A. Foundations of measurement. 2nd ed. Dover Publication, Inc.; 1999.

[3] ISO 14253-1. Geometrical product specifications (GPS) — Inspection by measurement of workpieces and measuring equipment Part 1 : Decision rules for proving conformity or nonconformity with specifications. BSI Standards Publication; 2013.

[4] JCGM 200. International vocabulary of metrology — Basic and general concepts and associated terms (VIM). 3rd ed. JCGM; 2008.

[5] Luce RD. Semiorders and a Theory of Utility Discrimination. Econometrica 1956;24:178–91.

[6] Fishburn PC. Intransitive indifference with unequal indifference intervals. J Math Psychol 1970;149:144–9.

[7] Welch BL. The Generalization of `Student's' Problem when Several Different Population Variances are Involved. Biometrika 1947;34:28–35. doi:10.2307/2332510.