DE GRUYTER Journal of Applied Geodesy 2014; 8(4): 283–293

Research Article

Yin Zhang* and Ingo Neumann

Utility Theory as a Method to Minimise the Risk in Deformation Analysis Decisions

Abstract: Deformation monitoring usually focuses on the detection of whether the monitored objects satisfy the given properties (e.g. being stable or not), and makes further decisions to minimise the risks, for example, the consequences and costs in case of collapse of artificial objects and/or natural hazards. With this intention, a methodology relying on hypothesis testing and utility theory is reviewed in this paper. The main idea of utility theory is to judge each possible outcome with a utility value. The presented methodology makes it possible to minimise the risk of an individual monitoring project by considering the costs and consequences of overall possible situations within the decision process. It is not the danger that the monitored object may collapse that can be reduced. The risk (based on the utility values multiplied by the danger) can be described more appropriately and therefore more valuable decisions can be made. Especially, the opportunity for the measurement process to minimise the risk is an important key issue. In this paper, application of the methodology to two of the classical cases in hypothesis testing will be discussed in detail: 1) both probability density functions (pdfs) of tested objects under null and alternative hypotheses are known; 2) only the pdf under the null hypothesis is known and the alternative hypothesis is treated as the pure negation of the null hypothesis. Afterwards, a practical example in deformation monitoring is introduced and analysed. Additionally, the way in which the magnitudes of utility values (consequences of a decision) influence the decision will be considered and discussed at the end.

Keywords: deformation monitoring, decision making, utility theory, hypothesis testing, risk analysis

DOI 10.1515/jag-2014-0012
Received June 03, 2014; accepted October 23, 2014.

***Corresponding Author: Yin Zhang:** Geodetic Institute Leibniz University of Hannover, Germany, E-mail: zhang@gih.uni-hannover.de
Ingo Neumann: Geodetic Institute Leibniz University of Hannover, Germany

1 Introduction

Deformation monitoring usually focuses on the detection of whether the monitored objects satisfy the given properties, for example, whether a bridge is stable or not, and furthermore, it makes decisions in order to meet the need of society to minimise negative environmental impacts or risks, for example, the consequences and costs in case of collapse of artificial objects and/or natural hazards. An optimal configuration for the measurement setup and all other decisions should therefore review and rate the risks of an individual monitoring project. Nowadays, the methodology in many engineering disciplines and mathematically founded decisions is usually based on probabilities and significance levels but not on the risk (consequences or costs) itself.

In the classical case, hypothesis testing serves to check the available information on the (unknown) parameters [3]. In linear models, two decisions are possible. The two assumptions are typically called the null hypothesis (stable object) and the alternative hypothesis (unstable object), respectively. The result of the test is the acceptance or rejection of the predefined hypotheses. A typical example is the detection of significant movements of a monitored object. The choice of the null or alternative hypothesis is based solely on probabilities, which have more or less no reference to practical applications. In cases in which the same probabilities appear in the acceptance and rejection regions, wrong decisions can be made and each decision may lead to dramatically different consequences.

One proper way to consider consequences within the decision process is through the so-called "utility theory" in decision making. It allows the consideration of the consequences or costs of decisions in order to meet the real requirements [4]. In this case, possible decisions are evaluated with cost functions for Type I and Type II errors, and the final choice of the most beneficial one leads to the minimum costs or consequences. The methodology was applied to a slide slope monitoring project in one of our previous researches, in a case where no statistical information about the behaviour of the object was available and the assessment for a critical situation had to be based on experience and/or expert knowledge [6].

As a continuation of the previous research, methodology-embedded hypothesis testing and utility theory is applied to the other two classical cases: 1) both probability density functions (pdfs) of the tested objects under null and alternative hypotheses are known; 2) only the pdf under the null hypothesis is known and the alternative hypothesis is the pure negation of the null hypothesis. Finally, an example of monitoring of a steel part will be given.

2 Mathematical definition of the situations

2.1 Possible situations in deformation monitoring

The availability of initial information on a monitored object, for example a bridge or slide slope, can differ. In accordance with the physical model or material parameters of a specific monitored object, there exist three main cases. Due to the lack of space, and for a better understanding, only these three main cases are considered within this paper.

The most optimal situation is the case when the physical model [e.g. Finite Element Model (FEM)] or other information is well known to a certain degree. Then the statistical behaviour of the monitored object under the null and alternative hypotheses can be statistically depicted (CASE I).

If a physical model can be provided, but only (statistical) statement about the behaviour of the stable object is available (e.g. the objects under the null hypothesis are prevailing), the non-stable objects are modelled as pure negation of the statistically defined null hypothesis in this case (CASE II). When the monitored targets are older objects, usually no physical model is available because the material parameters and exact object geometry are unknown. Therefore no statistical information about the behaviour of the object is available and the assessment for a critical situation has to be based on experience and/or expert knowledge (CASE III). This situation should generally be strictly avoided.

The uncertainty in the loads, geometry, material parameters, and so on, leads to the case in which the determined stresses, internal forces, or displacements are uncertain values too. Generally speaking, the uncertainty in the assessment of the behaviour of the monitored objects and of the geodetic measurements is taken into consideration. Therefore, the null and alternative hypotheses are specified with uncertainty. Both of the above-mentioned uncertainty components are mainly modelled with the aid of pdfs ($\rho(x)$ with $x \in \mathbb{R}$). Among the three mentioned cases, only the null and alternative hypotheses in CASE III are modelled with intervals. Since this situation was discussed in detail in our previous papers, for example [6] and [14], this case will not be discussed further; the interest of this paper is in the first two cases. For more detailed discussion of CASE III, please refer to [7]. Strategies for constructing fuzzy numbers or fuzzy intervals based on expert knowledge are given by [7] and [10].

2.2 Determination of Probabilities

For each possible situation mentioned above, defining the probability of its occurrence is the prerequisite for any further decision making and/or analysis steps. In general, this requires a multidimensional treatment of a structure and the corresponding situations. But usually this multidimensional case is mapped to a one-dimensional decision with the above-mentioned two alternatives (e.g. [3]). Therefore, the paper shows only the one-dimensional treatment of all methods. For all three cases, the uncertainty of the measured situation is described by the pdf: $T \sim \rho_T(x)$. But due to the space limitation, the probability determination for CASE III will not be elaborated in this paper. The related strategy was discussed in [9] and [5].

CASE I

This optimal situation is based on the richness of the knowledge of, for example, the physical model or other information. Take a steel part as a monitored object. It is assumed that the target object can be manufactured by either of two machines, the null hypothesis is that the monitored object (the steel part) is manufactured by machine I; the alternative hypothesis is that it is manufactured by machine II. Then the statistical behaviour of the monitored object under the null and alternative hypotheses can be statistically delineated. Due to the uncertainty of the input parameters, the knowledge of the behaviour of an object is also uncertain. Therefore the definitions of the null and alternative hypotheses must be uncertain values. Within the paper this uncertainty should be provided by the pdf, where $\rho_{H_0}(T)$ and $\rho_{H_1}(T)$ stand for the pdfs of objects that satisfy the null and alternative hypotheses, respectively. Then the probability $P(T|H_0)$ that the measured $T \sim \rho_T(x)$ belongs to the null hypothesis is

$$P(T|H_0) = \frac{\int_{\mathbb{R}} \rho_{H_0}(x)\rho_T(x)\,dx}{\int_{\mathbb{R}} \rho_{H_0}(x)\,dx + \int_{\mathbb{R}} \rho_{H_1}(x)\,dx}, \tag{1}$$

and similarly, the probability $P(T|H_1)$ that the measured $T \sim \rho_T(x)$ belongs to the alternative hypothesis is

$$P(T|H_1) = \frac{\int_{\mathbb{R}} \rho_{H_1}(x)\rho_T(x)\,dx}{\int_{\mathbb{R}} \rho_{H_0}(x)\,dx + \int_{\mathbb{R}} \rho_{H_1}(x)\,dx} = 1 - P(T|H_0). \tag{2}$$

CASE II

The existence of a physical model of the monitored object allows the behaviour of the object under specific loads and so on to be predicted. Regarding the uncertainty, for reasons similar to those introduced in CASE I, $\rho_{H_0}(T)$ is also provided as the pdf of the object under the null hypothesis. If no further information is available, the probability $P(T|H_0)$ that the measured $T \sim \rho_T(x)$ belongs to the null hypothesis is

$$P(T|H_0) = \frac{\int_{\mathbb{R}} \rho_{H_0}(x)\rho_T(x)\,dx}{\int_{\mathbb{R}} \rho_{H_0}(x)\,dx}. \tag{3}$$

When no knowledge about the alternative hypothesis is available, it is generally treated as the pure negation of the null hypothesis and the probability $P(T|H_1)$ that $T \sim \rho_T(x)$ belongs to the alternative hypothesis is

$$P(T|H_1) = 1 - P(T|H_0). \tag{4}$$

From the computational aspect, the integral for probability is hard to solve. Later in the paper, convolution formulas will be treated as a solution for the example.

3 Test with the Consideration of Costs

3.1 General Idea of Utility Theory

Within the paper, the so-called utility theory is introduced for the current methodology. The simplest case in decision making is where the situation resulting from each decision is known exactly, and the preferred situation can be decided through simple comparisons. But generally only the probabilities of situations can be predicted. The idea of utility theory allows the consideration of consequences or costs in decision making in order to meet the real requirements (e.g. [4]). In this case, possible decisions are evaluated with cost functions for Type I and Type II errors. Finally, the decision leading to the minimum costs or consequences is chosen as the most beneficial one.

The classical case of decision making in geodesy is so-called hypotheses testing with two possible alternatives (acceptance or rejection of the null hypothesis). Four possible alternatives according to the test hypotheses (H_0 or H_1) are displayed in Table 1.

The utility value is a measure of the consequences that occur as a result of a decision. It is well known that those decisions leading to a wrong assessment of the real situation lead to higher costs and therefore to lower utility values. Take bridge deformation monitoring as an instance. If a bridge is classified as unstable but in reality it is stable (Type I error), then the costs of, for example, the dynamic stabilisation of the bridge are determined as the corresponding utility values. There are four utility values corresponding to four possible test situations:

– U_{00}: Utility of a correct choice of the null hypothesis;
– U_{01}: Utility of an incorrect choice of the alternative hypothesis (Type I error);
– U_{11}: Utility of a correct choice of the alternative hypothesis;
– U_{10}: Utility of an incorrect choice of the null hypothesis (Type II error).

Considering that the correct decision is always better than the incorrect one, it can be derived that $U_{00} > U_{10}$ and $U_{11} > U_{01}$.

3.2 Test Strategy with the Aid of Cost Function

This section presents the mathematical procedure to identify the most beneficial decision with the aid of utility theory. In general, the case is that situations are not always explicit, and in practice, often only the probability of each situation occurring can be predicted. Then the decision making cannot be derived by simple comparison and picking of the preferred situation. Assume that $\rho_0(T)$ and $\rho_1(T)$ are probability densities of a test value T, for objects satisfying the null and alternative hypotheses, respectively. The probability $p_0(T) = P(H_0|T)$ for T which satisfies the null hypothesis can be determined by Bayes' Theorem [2]:

$$\begin{aligned} p_0(T) &= \frac{P(T|H_0)\cdot P(H_0)}{P(T|H_0)\cdot P(H_0) + P(T|H_1)\cdot P(H_1)} \\ &= \frac{\rho_0(T)\cdot P(H_0)}{\rho_0(T)\cdot P(H_0) + \rho_1(T)\cdot P(H_1)}. \end{aligned} \tag{5}$$

Table 1. Possible situations resulting from the test decision.

Situations	Acceptance of H_0	Rejection of H_0
H_0 is true	Correct choice of the null hypothesis	Incorrect choice of the alternative hypothesis **(Type I error)**
H_0 is false (H_1 is true)	Incorrect choice of the null hypothesis **(Type II error)**	Correct choice of the alternative hypothesis

where $P(H_0)$ and $P(H_1)$ represent the probabilities of a randomly chosen object satisfying the null or alternative hypothesis.

Meanwhile, probability $p_1(T) = P(H_1 T)$ that T satisfies the alternative hypothesis can be determined as

$$p_1(T) = 1 - \frac{\rho_0(T) \cdot P(H_0)}{\rho_0(T) \cdot P(H_0) + \rho_1(T) \cdot P(H_1)} = 1 - p_0(T). \quad (6)$$

According to [4], the expected utilities of the null and alternative hypotheses, written as K_0 and K_1, can be calculated with the aid of probabilities and utility values:

$$\begin{aligned} K_0 &= p_0(T)U_{00} + p_1(T)U_{10} = p_0(T)(U_{00} - U_{10}) + U_{10}, \\ K_1 &= p_0(T)U_{01} + p_1(T)U_{11} = p_0(T)(U_{01} - U_{11}) + U_{11}. \end{aligned} \quad (7)$$

Since the final decision that will result in the largest expected utility (the minimum costs) of the hypothesis is made, the null hypothesis is chosen if

$$p_0(T)U_{00} + p_1(T)U_{10} \geq p_0(T)U_{01} + p_1(T)U_{11} \quad (8)$$

holds. It is known from the last section that $U_{00} > U_{01}$ and $U_{11} > U_{10}$; then Equation (8) can be simplified to

$$\frac{p_0(T)}{p_1(T)} = \frac{U_{11} - U_{10}}{U_{00} - U_{01}}. \quad (9)$$

Substituting the Bayes equation (5) leads to

$$\frac{\rho_0(T)}{\rho_1(T)} \geq r_0 = \frac{(U_{11} - U_{10})\, p_1(T)}{(U_{00} - U_{01})\, p_0(T)}. \quad (10)$$

When the right hand side of Equation (10) is known, the decision is made by comparison with the existing threshold r_0. Then, the so called Neyman-Pearson criterion can be summarised as follows:

– the null hypothesis is selected if $\frac{\rho_0(T)}{\rho_1(T)} \geq r_0$ holds,
– otherwise, the alternative hypothesis is selected if $\frac{\rho_0(T)}{\rho_1(T)} < r_0$ holds.

In the monitoring concept, regulatory thresholds for critical movements play a key role nowadays. When tolerances are given, a production or inspection process can be checked by measurements with detection of deviation between the actual and nominal dimensions of an object. The nominal dimension is defined by lower and upper bounds, which are known as regulatory thresholds. Using Equation (7), Equation (8) can also be written as

$$p_0(T)(U_{00} - U_{10}) + U_{10} \geq p_0(T)(U_{01} - U_{11}) + U_{11}. \quad (11)$$

The rearranged form is

$$p_0(T) \geq p_{0,critical} = \frac{U_{11} - U_{10}}{U_{00} - U_{10} - U_{01} + U_{11}}. \quad (12)$$

If the probability $p_0(T)$ is larger than or equal to the critical probability $p_{0,critical}$, the null hypothesis is selected.

The decisions for regular thresholds can be extended to the linguistic imprecision or fuzziness of the formulated hypotheses, when decision making deals with reasoning that is more approximate rather than fixed and exact. It is also possible to consider non-stochastic uncertainties, such as systematic measurement errors. For more information on the strategy for non-stochastic measurement uncertainties and linguistic uncertainty for regulatory thresholds, the reader can refer to [9] and [1].

For geodetic monitoring with regular thresholds, detailed methodology and examples have been discussed, for example in [6] and [14]. In the following section, an example regarding the monitoring of steel parts will be introduced.

4 Example: Monitoring of a Machine Part

Suppose we have a steel part which is used in a machine, and the standard length of the steel part is assumed to be 100 mm. Now we need to decide, based on its length, whether the steel part is suitable for use in the machine or whether it may destroy the machine slowly and cause more risk later. In this case, there are four possible situations.

When we make a correct decision that the part is unsuitable (costs for U_{11}), we abandon it and buy or produce a new part. But if the abandonment is unnecessary, this decision will bring in additional costs for U_{01}. On the other hand, when we correctly classify it as suitable (costs for U_{00}), the test part will be used in the machine. In the same decision, there still exists a possibility that the machine may be destroyed little by little and the costs of reparation can be much higher (costs for U_{10}). Therefore, the optimal approach is to make a decision based on the expected total utilities which allows us to derive the minimum costs.

As mentioned in Section 2, the emphasis of this paper is on CASE I and CASE II situations. Thus, all the examples will be discussed in relation to these two situations. For each case, we will also show the results with a single test value and a random test value, respectively.

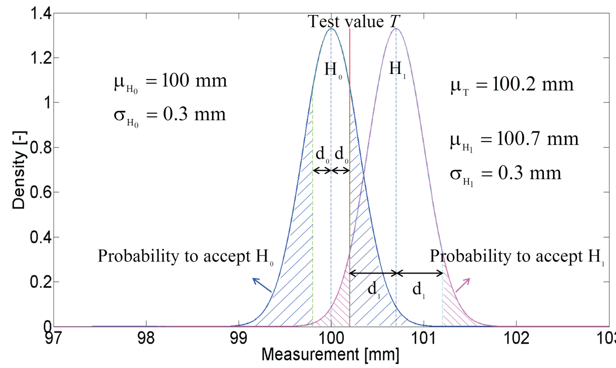

Fig. 1. Probability that H_0/H_1 can be accepted in CASE I with single test value.

4.1 Strategy Applying to the Example

CASE I

This is the most optimal case in which the physical model or other information is well known to a certain degree. Then the statistical behaviour of the monitored object under the null and alternative hypotheses can be statistically depicted. In this practical example, the behaviour of the steel part under both null and alternative hypotheses can be statistically modelled. The distribution of the lengths of the steel part which can be used in the machine properly is modelled mathematically (H_0). Also, the distribution of the test values under alternative hypothesis H_1, which means the distribution of the length of the steel part that will damage the machine or may work properly in another machine, is also known.

(i) Single test value

Assuming that H_0 follows the normal distribution, when there is a single test value T, the distance between the test value and the mean (μ_{H_0}) of the pdf can be used to decide how the test value differs from the given hypothesis. Therefore, the probability that a normal random variable d_0 is no bigger than $|\mu_{H_0} - T|$ denotes the probability that H_0 can be rejected. Similarly, under the alternative hypothesis (which also has a normal distribution with a mean of μ_{H_1}), the probability that H_1 can be rejected is equivalent to the probability that a normal random variable d_1 is no bigger than $|\mu_{H_1} - T|$. Figure 1 illustrates the two-sided situation.

Then, the probabilities that the null and alternative hypotheses can be accepted with the consideration of standardisation are

$$P(H_0|\,T) = \frac{1 - P(d_0 \le |\mu_{H_0} - T\,|)}{[1 - P(d_0 \le |\mu_{H_0} - T\,|)]+[1 - P(d_1 \le |\mu_{H_1} - T\,|)]}, \tag{13}$$

$$P(H_1|\,T) = \frac{1 - P(d_1 \le |\mu_{H_1} - T\,|)}{[1 - P(d_0 \le |\mu_{H_0} - T\,|)]+[1 - P(d_1 \le |\mu_{H_1} - T\,|)]} \tag{14}$$

(ii) Random test value

In practice, it is impossible to find the "true length" of a steel part without uncertainty. Advanced measuring equipment and analysis methods can only help to avoid gross errors and reduce systematic errors, but the random errors in measurements cannot be totally eliminated. With repeated measurements, the approximation of the "true length" of a steel part can be illustrated mathematically, usually by a normal distribution with a certain mean value and standard deviation as uncertainty.

When there is a random test value, the probabilities that the null and alternative hypotheses can be accepted are more complicated, as shown in Figure 2. The general equation for computing the probability is

$$P(H_0\,|T) = p(d_0) = \int_{-\infty}^{+\infty} pdf_{H_0}(d_0)pdf_T(x - d_0)dd_0, \tag{15}$$

which is known as convolution. This equation is in general not analytically solvable and therefore numeric solutions are needed. If the distribution of the random test value also has a normal distribution, the convolution of the test value and the normally distributed hypothesis (H_0 or H_1) can be analytically treated.

Assume that in general two different normal distributions are described by two probability density functions

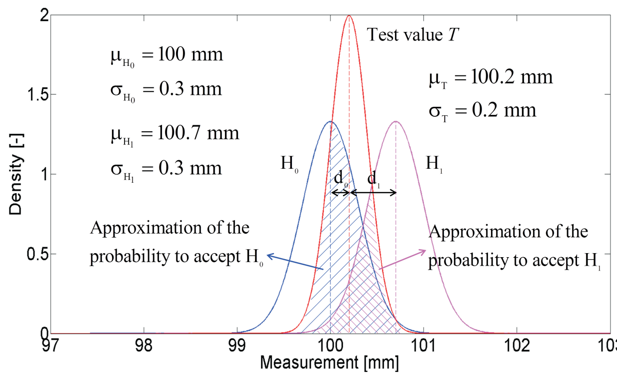

Fig. 2. Probability that H_0/H_1 can be accepted in CASE I with random test value.

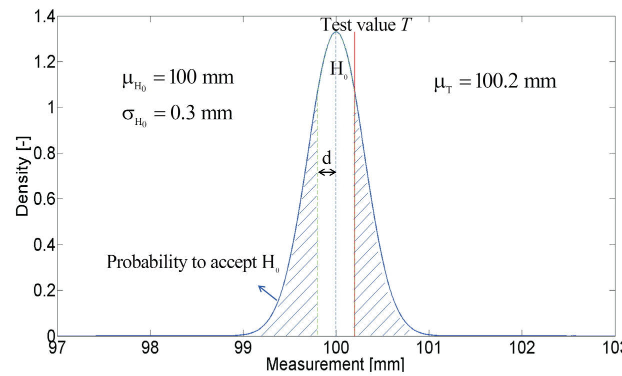

Fig. 3. Probability that H_0/H_1 can be accepted in CASE II with single test value.

$f_1(x)$ and $f_2(x)$:

$$f_1(x) = f_N^{(\mu_1,\sigma_1^2)}(x), \quad f_2(x) = f_N^{(\mu_2,\sigma_2^2)}(x), \tag{16}$$

with means of μ_1 and μ_2 and variances of σ_1^2 and σ_2^2, respectively. Then the convolution of the two functions is

$$f_Y(y) = f_N^{(\mu_1+\mu_2,\sigma_1^2+\sigma_2^2)}(y). \tag{17}$$

Due to the space limitation, the derivation of the above equation will not be discussed. For more information, please refer to [1], for example. In the specific example, the new derived distribution is still normal with a mean of $(\mu_{H_0} - \mu_T)$ and variance of $(\sigma_{H_0}^2 + \sigma_T^2)$. The convolution of the test value and the given alternative hypothesis is executed in the same way as shown above. Considering standardisation, the probabilities that the null and alternative hypotheses can be accepted are based on

$$P(H_0|\,T) = \frac{1 - P(d_0 \le |\mu_{H_0} - \mu_T|)}{[1 - P(d_0 \le |\mu_{H_0} - \mu_T|)]+[1 - P(d_1 \le |\mu_{H_1} - \mu_T|)]}, \tag{18}$$

$$P(H_1|\,T) = \frac{1 - P(d_1 \le |\mu_{H_1} - \mu_T|)}{[1 - P(d_0 \le |\mu_{H_0} - \mu_T|)]+[1 - P(d_1 \le |\mu_{H_1} - \mu_T|)]}, \tag{19}$$

$$\text{with} \quad d_0/\sqrt{(\sigma_{H_0}^2 + \sigma_T^2)} \sim N(0, 1)$$

$$d_1/\sqrt{(\sigma_{H_1}^2 + \sigma_T^2)} \sim N(0, 1).$$

CASE II

In this case, if a physical model or other information is provided, then the (statistical) statement about the behaviour of the object under H_0 is available. Applying this to this practical example, the behaviour of the steel part under H_0, which means the tested steel part is suitable for use in the specific machine, can be statistically modelled. And the alternative hypothesis (H_1), which means the steel part is unsuitable for the specific machine (it may be an unqualified steel part or one for another machine), is derived from a pure negation of the defined null hypothesis.

(i) Single test value

Similarly to CASE I, the null hypothesis H_0 is deemed to have a normal distribution, the test value T has only a single value, and the distance between the test value and the mean (μ_{H_0}) of the null hypothesis pdf can be used to decide how the test value differs from the given hypothesis. Therefore, the probability that a normal random variable d is no bigger than $|\mu_{H_0} - T|$ denotes the probability that H_0 can be rejected (see Figure 3). But the alternative hypothesis cannot be mathematically modelled and is treated as a pure negation of the null hypothesis in this case.

The probability that the null hypothesis can be accepted and the probability that it can be rejected are, respectively,

$$P(\,H_0|\,T) = 1 - P(d \le |\mu_{H_0} - T|), \tag{20}$$

$$P(\,H_1|\,T) = 1 - P(\,H_0|\,T) = P(d \le |\mu_{H_0} - T|). \tag{21}$$

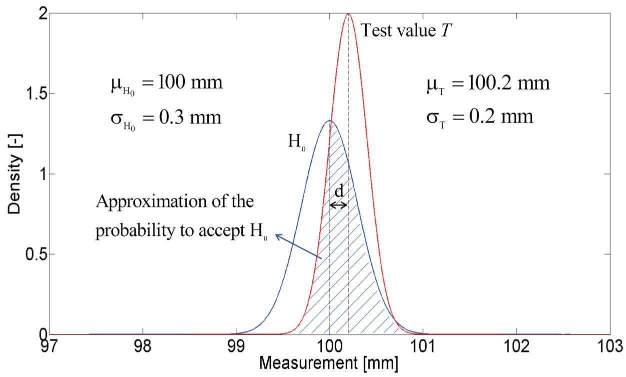

Fig. 4. Probability that H_0/H_1 can be accepted in CASE II with random test value.

(ii) Random test value

For CASE II, when the random test value follows the normal distribution with a mean of μ_{H_0} and a variance of $\sigma^2_{H_0}$, it is necessary to perform the convolution operation as in CASE I in order to obtain the probability that H_0 can be accepted (see Figure 4).

In the example, the new distribution derived from convolution also follows a normal distribution with a mean of $(\mu_{H_0} - \mu_T)$ and a variance of $(\sigma^2_{H_0} + \sigma^2_T)$. Since the probability that the test value is under the alternative hypothesis is a negation of the null hypothesis, the probability that the null alternative hypothesis can be accepted and the probability that it can be rejected are, respectively,

$$P(H_0|T) = 1 - P(d \le |\mu_{H_0} - \mu_T|), \quad \text{with} \quad d/\sqrt{(\sigma^2_{H_0} + \sigma^2_T)} \sim N(0,1); \tag{22}$$

$$P(H_1|T) = 1 - P(H_0|T) = P(d \le |\mu_{H_0} - \mu_T|), \quad \text{with} \quad d/\sqrt{(\sigma^2_{H_0} + \sigma^2_T)} \sim N(0,1) \tag{23}$$

4.2 Results and analysis

According to the strategies discussed above, the probabilities that the test value is under the null or alternative hypothesis can be calculated. Besides probabilities, the utilities (see Section 3) are also considered to make optimal decisions and to minimise the total costs. The utility values of four different situations are given in Table 2. The minus sign denotes the meaning of expenditure and the specific values should be given by relevant experts.

Table 2. Utility values for four situations.

Utilities	U_{00}	U_{01}	U_{11}	U_{10}
Currency	−2000	−3000	−3000	−10000

Table 3. Results for the single test value in Case I.

Hypothesis	Probability	Expected Utility	Decision
H_0	0.841	−3273.21	Reject
H_1	0.159	−3000.00	Accept

As mentioned in Section III, there is a numerical relation between the four values. From the practical point of view, the expense of correctly choosing the null hypothesis (U_{00}) could be the fixed costs of producing the steel part and assembling it in the machine. When we make a decision of choosing the alternative hypothesis (U_{01} and U_{11}), which means the steel part is considered to be unsuitable irrespective of whether or not this is true, the costs increase (here by 1000) due to the need to reproduce or repair it. But when the wrong decision causes a Type II error, an unsuitable steel part may damage the whole machine, thereby leading to more risk and costs in future. In this case, the costs of a Type II error (U_{10}) could be much higher (7000–8000) than in the other three situations.

Results and decisions for the example in the different cases are given below.

CASE I

The parameters of the null and alternative hypotheses are given, and in both cases they follow the normal distribution. For the null hypothesis, the mean μ_{H_0} = 100 mm and the standard deviation σ_{H_0} = 0.3 mm; for the alternative hypothesis, the mean μ_{H_1} = 100.7 mm and the standard deviation σ_{H_1} = 0.3 mm.

(i) Single test value

When there is a single test value of T = 100.2 mm, the results are as shown in Table 3.

(ii) Random test value

When the random test value obeys the normal distribution with the mean μ_T = 100.2 mm and the standard deviation σ_T = 0.2 mm, the results according to Equations (11), (18) and (19) are as shown in Table 4.

Table 4. Results for the random test value in Case I.

Hypothesis	Probability	Expected Utility	Decision
H_0	0.778	−3778.29	Reject
H_1	0.222	−3000.00	Accept

Table 5. Results for the single test value in Case II.

Hypothesis	Probability	Expected Utility	Decision
H_0	0.505	−5960.12	Reject
H_1	0.495	−3000.00	Accept

CASE II

If only the distribution of the null hypothesis is known, its parameters are the same as in CASE I: the mean μ_{H_0} = 100 mm and the standard deviation σ_{H_0} = 0.3 mm. The probability of the alternative hypothesis is obtained from the negation of the null hypothesis. For the purpose of comparison, the parameters of the test values remain the same as in CASE I.

(i) Single test value
The single test value remains T = 100.2 mm. The results are shown in Table 5.

(ii) Random test value
The random test value remains the same with the mean μ_T = 100.2 mm and the standard deviation σ_T = 0.2 mm. The results are shown in Table 6.

For all the results displayed above, the decision based only on probability and the decision with consideration of utility are consistent most of the time. But the results for CASE II appear to be different; for example, in the example with a single test value, the probabilities that the test value is under the null and alternative hypotheses are quite close to each other; here, they are 50.5 and 49.5%, respectively, and a decision based only on probability may lead to a dramatic situation. When the focus of the decision is more on the risk itself, that is, when the utilities or costs for each situation are also considered, the decision will result in the minimum risk or costs. In this example, the difference in utility of two decisions can be greater than 2000. Hence one can see the necessity of considering the utility in decision making in the field of deformation monitoring.

Table 6. Results for the random test value in Case II.

Hypothesis	Probability	Expected Utility	Decision
H_0	0.579	−5367.20	Reject
H_1	0.421	−3000.00	Accept

As discussed above in this paper, the Type II error generally costs the most among all four situations. Therefore, this value (U_{10}) plays an important role in the decision procedure. Figure 5 shows how it influences the magnitude of the expected utility of the two hypotheses as well as the final decision.

It is shown that the null hypothesis has a better chance of being selected when a Type II error costs less. The expected utilities of the null and alternative hypotheses are equivalent, reaching −3000 when $U_{10} \approx -4020.14$.

4.3 Estimation of Significance Level

As mentioned at the beginning of the paper, the methodology in many engineering disciplines and mathematically founded decisions is usually based on probabilities and significance levels. With the figures given above, it is not easy to find a clear and practical meaning for a specific example. But the utility values in four concrete situations are more comprehensible and can be given more easily by relevant experts; furthermore, they are necessary for the evaluation of the risk itself through the expected costs. From this point of view, we should find a way to calculate the significance levels based on given utility values in order to compare the results with the standard methodology.

According to Equation (12), we could determine the critical probability with fixed utility values. And if the relation between the observed variable and its probability distribution is clear, then decision making based on risk (consequences or costs) itself can be related to the observation directly. Take CASE II as an example. Figure 6 and Figure 7 show the variation tendency of the significance level with respect to different costs of Type II error for the single test value (according to Figure 3) and the random test (according to Figure 4), respectively.

For the single test value in Figure 6 we obtain the following:

- When $U_{10} = -15000$, if $p_0 \leq 0.92$ ($d \geq 0.03$ mm), H_0 can be rejected;
- When $U_{10} = -10000$, if $p_0 \leq 0.88$ ($d \geq 0.05$ mm), H_0 can be rejected;
- When $U_{10} = -5000$, if $p_0 \leq 0.67$ ($d \geq 0.13$ mm), H_0 can be rejected.

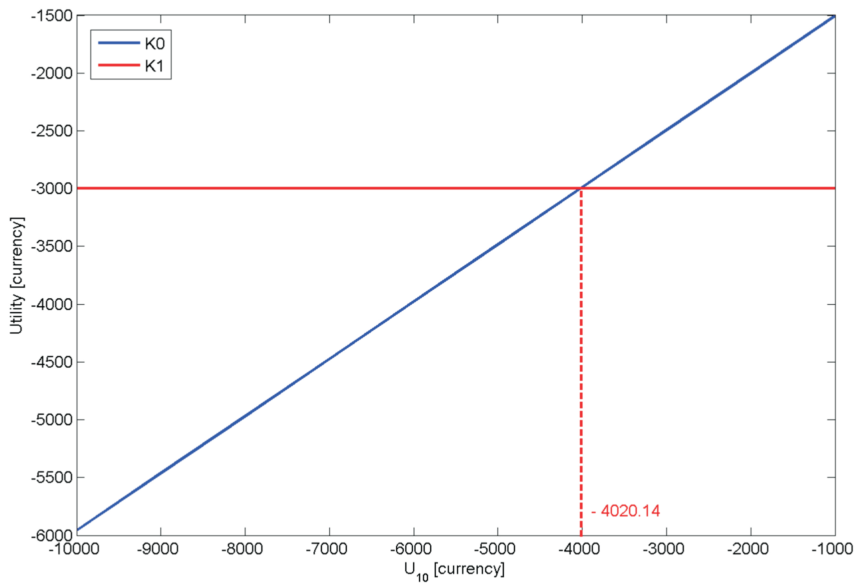

Fig. 5. Expected utilities of H_0 and H_1 with gradual changes in U_{10}.

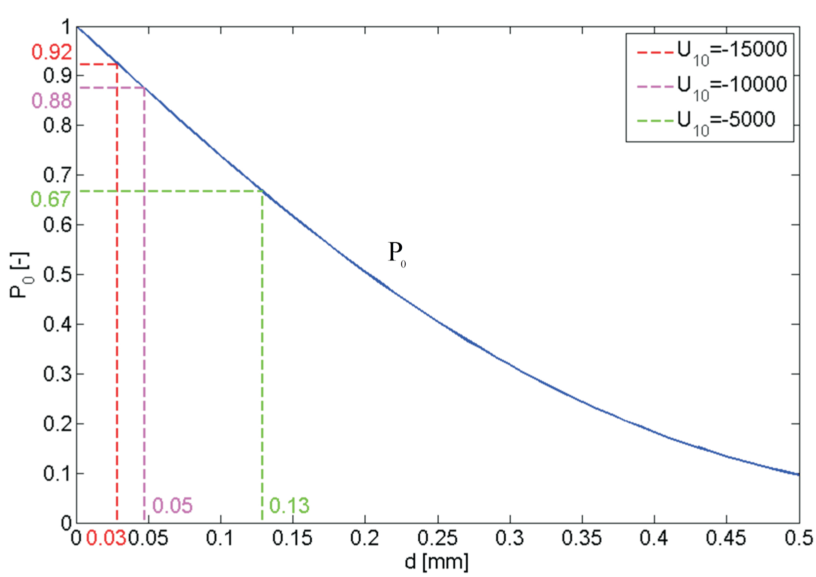

Fig. 6. Significance level under the influence of U_{10} for the single test value according to Figure 3.

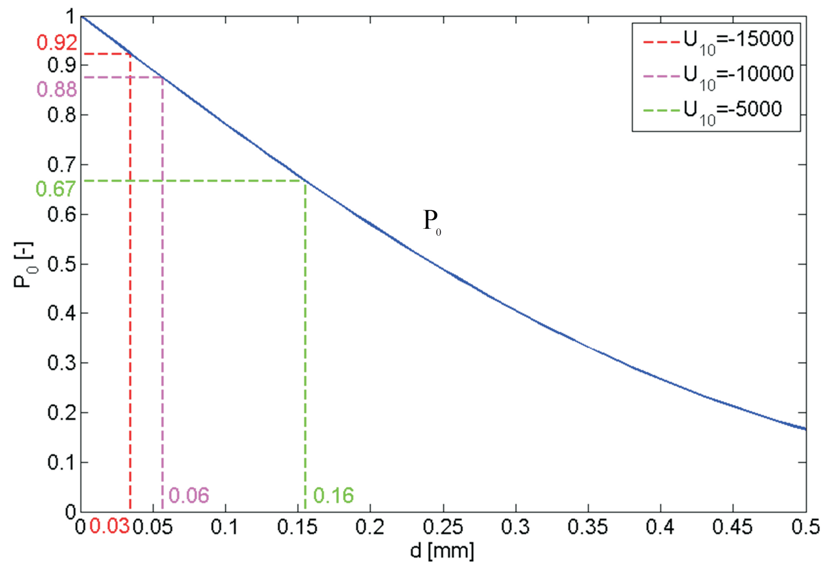

Fig. 7. Significance level under the influence of U_{10} for the random test value according to Figure 4.

For the random test value (see Figure 7):
- When $U_{10} = -15000$, if $p_0 \leq 0.92$ ($d \geq 0.03$ mm), H_0 can be rejected;
- When $U_{10} = -10000$, if $p_0 \leq 0.88$ ($d \geq 0.06$ mm), H_0 can be rejected;
- When $U_{10} = -5000$, if $p_0 \leq 0.67$ ($d \geq 0.16$ mm), H_0 can be rejected.

From the results, we can see that the significance level can be determined conversely through given utility values and described by the observation and the probability under their distributions. For large costs of Type II errors the null hypothesis can only be chosen if it is very probable that it is true. This is in full accordance with the theoretical assumptions, where one should obviously avoid the choice of the null hypothesis if it leads to a large risk (if the decision is not correct).

It should also be clearly stated here that this step is in general not necessary to evaluate the risk of a monitoring project. It is only worthwhile for a comparison of the new methodology with the classical approach.

5 Conclusion

Deformation monitoring concepts need to be more accurate and reliable when higher risks or costs result from the assumed deformation of an object, for example, under a load. However, decision making without consideration of costs or risks themselves therefore cannot be optimal for either artificial nor natural objects. In the case of a physical model, for example with the aid of an FEM, or when other information is available, the presented methodology can identify critical situations for an object and allows geodetic techniques to measure real deformations of the objects. Therefore the most beneficial decisions which minimise the consequences of actual behaviour of a monitoring project are identified.

In this paper, the presented methodology shows a concept in decision making with the consideration of costs or consequences in the case where at least one of the hypotheses can be modelled. The decisions are evaluated by extending the statistical hypothesis tests with cost functions for Type I and II errors. Finally, use of the methodology leads to the minimum costs or consequences for a deformation monitoring process.

In general, this methodology can also be applied to non-statistical uncertainties in the measurements and in the object's behaviour. For the civil engineering treatment, see for example [12] or [13], and for the geodetic measurements, see for example [8]. Furthermore, the uncertainty of the modelling of the regulatory thresholds based on different expert opinions can be considered [7].

Further work needs to extend the approach to multiple-criteria decisions with more than two alternatives. Additionally, it would also be meaningful to implement the strategy in a real project to numerically analyse and optimise the measurement process with respect to consequences. This is important in order to evaluate the influence of the uncertainty of the utility values on the decisions.

Acknowledgement: The presented paper shows results and new ideas developed during the research project NE 1453/3-1, "Steering and Optimisation of Measurement Processes in Consideration of Cost Functions", which is funded by the German Research Foundation (DFG). This is gratefully acknowledged by the authors.

References

[1] Grinstead C. M. and Snell J. L., Introduction to Probability: Second Revised Edition, American Mathematical Society, 1997.

[2] Klir G. J., *Uncertainty and Information: Foundations of Generalized Information Theory*, John Wiley & Sons, Hoboken – New Jersey, 2006.

[3] Koch K. R., *Parameter Estimation and Hypothesis Testing in Linear Models*, Springer, Berlin – Heidelberg – New York, 1999.

[4] Kreinovich V., Nguyen H. T., and Niwitpong S., Statistical hypothesis testing under interval uncertainty: an overview, *International Journal of Intelligent Technologies and Applied Statistics* 1(1) (2008), 1–33.

[5] Kutterer H. and Neumann I., Multidimensional statistical tests for imprecise data, in: P. Xu, J. Liu, and A. Dermanis (Eds.), *Proceedings of the 6th Hotine-Marussi-Symposium*, IAG Symposia 132 (2007), 232–237.

[6] Neumann I. and Zhang Y., Assessment of risk within geodetic decisions, *Joint International Symposium on Deformation Monitoring*, Hong Kong, China (CD-Proceedings), 2011.

[7] Neumann I., *On the analysis of an extended uncertainty budget in parameter estimation and hypothesis testing*, PhD Thesis, Leibniz University of Hannover, Germany, 2009 [in German].

[8] Neumann I. and Kutterer H., The probability of type I and type II errors in imprecise hypothesis testing with an application to geodetic deformation analysis, *Journal of Reliability and Safety* 3(1–3) (2009), 286–306.

[9] Neumann I. and Kutterer H., Optimal hypothesis testing in case of regulatory thresholds, in: N. Sneeuw, P. Novák, M. Crespi, and F. Sansò (Eds.), *Proceedings of the 7th Hotine-Marussi-Symposium, International Association of Geodesy Symposia*, Springer, Berlin – New York 137 (2011), 75–80, 2011. DOI: 10.1007/978-3-642-22078-4_11.

[10] Nguyen H. T. and Kreinovich V., Nested intervals and sets: concepts, relations to fuzzy sets, and applications, in: R. B. Kearfott and V. Kreinovich (Eds.), *Applications of Interval Computations*, Kluwer Publishers, Dordrecht, 1996, 245–290.

[11] Niwitpong S., Nguyen H. T., Kreinovich V., and Neumann I., Hypothesis testing with interval data: case of regulatory constraints, *International Journal of Intelligent Technology and Applied Statistics* 1(2) (2008), 19–41.

[12] Möller B. and Beer M., Fuzzy Randomness – Uncertainty in Civil Engineering and Computational Mechanics, Springer Verlag, Berlin and New York, 2004.

[13] Möller B., Beer M., Graf W., and Sickert J. U., Fuzzy finite element method and its application, in: *Trends in Computational Structural Mechanics*, CIMNE, Barcelona, 2001, 529–538.

[14] Zhang Y. and Neumann I., Risk assessment for slope monitoring, *Journal of Applied Geodesy* 7(3) (2013), 159–171.