Zeitschrift: IABSE reports = Rapports AIPC = IVBH Berichte

Band: 74 (1996)

Artikel: Comparison of statistical evaluation models

Autor: Straalen, Ljsbrand van / Vrouwenvelder, Ton

DOI: https://doi.org/10.5169/seals-56083

Comparison of statistical evaluation models

IJsbrand VAN STRAALEN
TNO Building and Construction Research
Delft
The Netherlands

IJsbrand van Straalen, born 1963, got his civil engineering degree in 1988. He is a research engineer at TNO in the field of structural engineering and is working on his PhD thesis dealing with adhesive bonded joints.

Ton VROUWENVELDER
TNO Building and Construction Research
&
Delft University of Technology,
Department of Civil Engineering
Delft
The Netherlands

Ton Vrouwenvelder, born 1947, got his engineering degree in 1970. His is a senior research engineer at TNO in the field of reliability and dynamics and is professor at Delft University of Technology.

Summary

This paper deals with statistical evaluation models for resistance and material testing. It is shown that for a limited number of tests, say 1 to 4, as normal in daily practice, the model presented in Annex D 'Design Assisted by Testing' of Eurocode 1 [1] can result into unrealistic low design values. As an alternative, a sophisticated model which makes use of prior knowledge is presented. Also attention is payed to the evaluation on basis of a design model. The presented models are illustrated by examples.

1. Introduction

In most cases a structural engineer uses design formulae or data available in codes to establish design values of resistance properties of structural elements or materials. But in the following cases the engineer has to chose for a design based on experimental models:

- When no theoretical models or data are available, or the actual circumstances are not covered by existing models.
- When design formulae might give very conservative results and tests might lead to a more economic solution.
- To develop new design formulae.

When the choice is made for design by testing, the structural engineer is confronted with a lot of problems which has to be covered. In Annex D 'Design Assisted by Testing' of Eurocode 1 [1], the engineer can find guidelines which may be valuable for the planning and evaluation of tests. The evaluation model described in that document is based on a statistical analysis of test results and the partial safety factor design. One major issue the engineer has to deal with, is the fact that the number of tests should be sufficient for a valuable statistical interpretation. This implicates that the design by testing might be a very expensive and time consuming method. To study the possibility of using a smaller number of tests, TNO Building and Construction Research has carried out a review of a sophisticated statistical model. This so called Bayesian approach makes use of prior knowledge about the distribution of the test results. In this study also attention has been given to the evaluation on basis of a design model. The following chapters will give an overview of the available statistical evaluation models and will illustrate the possibilities with an example of a beam-column connection of a storage racking structure.

2. General considerations

2.1 Planning of test series

The planning of a test series is an important part of the design by testing, because correct choices have to be made to get valid results. To start with, the objective of the test series has to be formulated. Than a qualitative analysis has to be carried out in which e.g. the expected behaviour (parameters of influence, fail mechanism), boundary conditions, loading conditions, environmental conditions, time effects and differences between testing and reality are investigated. On basis of these results a relevant test arrangement has to be defined. This includes the specification of the type of specimen, the definition of the execution of the tests, the choice of environmental conditions, the method of observation and recording, the method of evaluation, the number of tests, the selection procedure of specimens and the design of the test rig. The development of the planning of a test series is not an easy task and requires appropriate theoretical knowledge, experience in testing and engineering judgement.

2.2 Execution of tests

After the planning of the test series has been worked out, the specimens have to be produced and selected, the test rig has to be build and the test programme has to be carried out. To ensure that the results are valid, the chosen measurement techniques should be in accordance with the required tolerances. One should be aware that the execution of tests is in accordance with the planning. If there is a discrepancy between the testing and the original planning, e.g. the occurrence of an unexpected failure mechanism, the whole planning of the test series has to be reconsidered. One should also be aware of uncontrolled reinforcements of e.g. the supports and unexpected environmental effects.

2.3 Evaluation of test results

After the tests are finished, the results have to be evaluated. The behaviour during loading and the failure mechanism of the tests have to be analyzed in general and the design values have to be determined. In the past several models to determine those design values were proposed, which are in many cases rules of thumb. E.g. according to the Dutch design recommendations of storage racking structures published in the seventies, the design strength of a beam-column connection as discussed in chapter 4, is equal to the factored value (0.67) of the lowest result of three tests. Nowadays it is generally excepted that a model based on the statistical theory is more in accordance with the partial safety factor design. A model based on the classical statistical theory is available, but also models based on the Bayesian theory which makes use of prior knowledge, are worked out for a single test series or a family of tests. In the following chapter a description of those models is given.

3. Description statistical evaluation models

3.1 Classical approach

According to the classical approach [2], [3] and [4], the design value of the resistance R is in case a normal distribution of the test results might be assumed, equal to:

$$R_d = \eta \frac{R_k}{\gamma_M} \tag{1}$$

where:
η is the conversion factor;
γ_M is the partial factor for the design;
R_k is the characteristic value based upon n results.

The conversion factor η takes into account the differences between testing conditions and actual ones. This factor is strongly dependent on the type of test and type of material. The value is mostly determined on basis of engineering judgement. The partial factor for the design γ_M is dependent on the field of application. The value should be taken from codes. The characteristic value R_k includes the statistical uncertainty. The value is determined by:

$$R_k = m_R - k_n s_R \tag{2}$$

where:
m_R is the mean value of the results;
k_n is the coefficient depending on the number of results n;
s_R is the standard deviation of the results.

For the classical approach the characteristic value is normally based on the 5 % fractile. If there is a complete lack of knowledge about the standard deviation, the value of k_n has to be taken from table 1 for the case that the standard deviation is unknown. If on the other hand, the standard deviation is fully known from prior knowledge, the value of k_n has to be taken from table 1 for the case that the standard deviation is known.

Table 1 - Values of k_n based on a 5 % fractile

standard deviation	n							
	3	4	6	8	10	20	30	∞
unknown	3.15	2.68	2.34	2.19	2.10	1.93	1.87	1.64
known	2.03	1.98	1.92	1.88	1.86	1.79	1.77	1.64

In the procedure given above a normal distribution of the test results is assumed. But in several applications other distributions are found, which leads to more economic design values. In case of a lognormal distribution the same procedure as given above can be followed if log values of the test results are used.

3.2 Bayesian approach

According to the Bayesian approach [2], [3] and [4], the design value of the resistance R in the case that a normal distribution of the test results might be assumed, equal to:

$$R_d = \eta \{m_R - t_\nu s_R \sqrt{1 + \frac{1}{n}}\} \quad (3)$$

where:
t_ν is the coefficient of the Student distribution.

The value of t_ν follows from table 2, where $\nu = n - 1$. The product $\alpha\beta$ corresponds to a fractile $P(\Phi)$ as indicated in table 2. The reliability index β is related to the failure probability for which a target is given by the code. The FORM weight factor α follows from a first order reliability method. In a design where the uncertainty of R is dominating, a value of $\alpha = 0.8$ should be used. Also other distributions of the test results than a normal distribution can be used.
Instead of using the direct method to determine the design value by equation (3), it is also possible to use the partial safety factor design as formulated with equation (1). The characteristic value R_k is then defined by equation (3) with $\alpha\beta = 1.64$. It is also possible to calculate the partial factor for design from $\gamma_M = R_k / R_d$.
It is known that the Bayesian approach is sensitive for the value of the standard deviation, specially if only a small number of test results is available. Too small or too large standard deviations might result into unsafe or uneconomic design values. An advantage of the Bayesian theory is that the prior knowledge can avoid unrealistic design values.

Table 2 - Values of t_ν

$\alpha\beta$	$P(\Phi)$	ν							
		2	3	5	7	9	19	29	∞
1.64	0.05	2.92	2.35	2.02	1.89	1.83	1.73	1.70	1.64
2.33	0.01	6.97	4.54	3.37	3.00	2.82	2.54	2.46	2.33
2.58	0.005	9.93	5.84	4.03	3.50	3.25	2.86	2.76	2.58
3.08	0.001	22.3	10.2	5.89	4.79	4.30	3.58	3.40	3.08

3.3 Prior knowledge

In literature [2], [3] and [4], the Bayesian approach which takes prior knowledge into account, is discussed. This approach establishes a prior distribution function for the unknown distribution parameters of the resistance R. On basis of this prior distribution in combination with the test results, a posterior distribution of the resistance R is derived. The prior (normal) distribution function can be represented by the following parameters:
$m(m_R)$ which is the mean value of the mean of the resistance R;
$s(m_R)$ which is the standard deviation of the mean of the resistance R;
$m(s_R)$ which is the mean value of the standard deviation of the resistance R;
$s(s_R)$ which is the standard deviation of the standard deviation of the resistance R.
It is noted here that if a lognormal distribution is chosen, the coefficient of variation V_R has to be used instead of the standard deviation s_R.

For practical applications it is important to know how the above given parameters of the prior distribution have to be determined. For many applications no prior knowledge about the mean of the resistance R is available. This implicates that $s(m_R)$ will have a large value and that the choice of the value for $m(m_R)$ is not relevant. On the other hand it is mostly possible to formulate prior knowledge about the standard deviation. This can be done by engineering judgement, but it is advised to determine the values for $m(s_R)$ and $s(s_R)$ of a group of comparable series of tests already available. The procedure which combines the prior distribution and the results of the considered tests, to determine the posterior distribution represented by the parameters m'', s'', ν'' and n'', is described in [4]. With these parameters the design value of the resistance R can be calculated with equation (3).

3.4 Evaluation on basis of a design model

It is also possible to evaluate tests on basis of a design model. More types of specimens with known varying parameters, e.g. plate thickness, beam height and yield strength, are included in the evaluation. These parameters might be deterministic or random. A mathematical relation (the 'design model') between those parameters has to be formulated. It must be kept in mind that the design model represents one failure mode that occurs in the tests. The result of the analysis is a design function for a given reliability level.
A description of the procedure that has to be followed is out of the scope of this paper. An overview is given in [4] and a detailed step by step procedure is described in [5]. The authors have added to this procedure the using of prior knowledge, as is reported in [6].

4. Example of a connection of a racking structure

4.1 Tests

To demonstrate the statistical evaluation models, an example is worked out. To optimize the economical solution of the design of a storage racking structure, design by testing of the components is preferred. The cantilever bending test on the beam-column connection is a standard test for which the planning and execution of the test is fully described in [7].

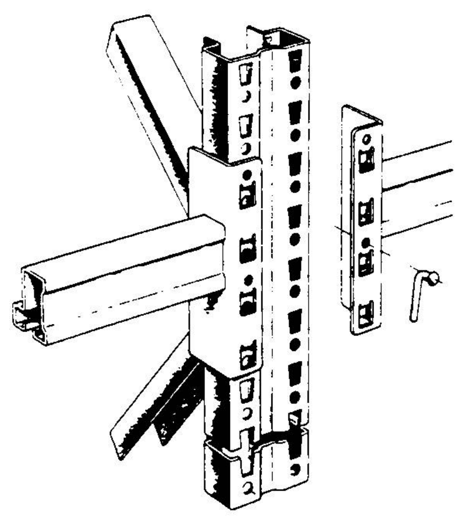

Figure 1 - Beam-column connection

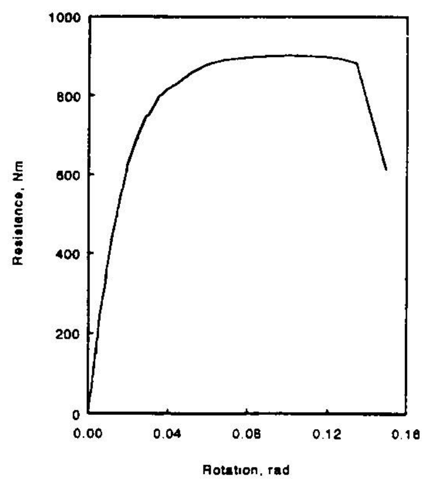

Figure 2 - Typical moment-rotation diagram

In figure 1 a beam-column connection is shown. The column is a cold-formed C-section with a continuing perforation pattern. The beam is also a cold-formed section. At the end of the beam a connector is welded, which has hooks or other devices which engage in the perforation. A typical moment-rotation diagram as a test result, is shown in figure 2.
For 6 types of specimen, A to F, with two plate thicknesses of the column and three beam heights, test series were carried out. The results are presented in table 3. It is assumed that the physical behaviour of the connections can be described by two parameters. One is the steel thickness t. The other one is the distance h, which is defined as the distance between the upper hook and the location of the connector where the beam rotates during loading (near bottom side of connected beam). Here it is assumed that the resistance moment of the connection is the maximum force in the hook times the distance h.

Table 3 - Overview measured resistances beam-column connections in Nm

Type of specimen	A	B	C	D	E	F
t, mm	2.0	2.0	2.0	2.5	2.5	2.5
h, mm	77	125	165	77	125	165
R_{test}, Nm	311	740	890	314	787	927
	353	740	820	323	837	950
	328	723	930	298	693	953
	337	693	773	310	870	1000
m_R, Nm	332	724	853	311	797	958
s_R, Nm	17.5	22.2	70.2	10.4	77.1	30.6
V_R	0.0528	0.0306	0.0823	0.0333	0.0968	0.0320

4.2 Results of interpretations

The test results given in table 3 are interpreted according to the statistical evaluation models. It is decided to assume a lognormal distribution, because the evaluation on basis of a design model is completely based on this type of distribution. For the interpretations according to each model, the following considerations have been made:

I Classical approach. It is assumed that there is a complete lack of prior knowledge. The characteristic value R_k is based on the 5 % fractile and the partial factor for design is taken equal to $\gamma_M = 1.25$.

II Bayesian approach without prior knowledge. In case of the determination of the design value R_d a reliability index of $\beta = 3.6$ is chosen and a FORM weight factor of $\alpha = 0.8$ is used. In case of the determination of the characteristic value R_k the product $\alpha\beta$ is chosen equal to 1.64, which corresponds to a fractile of 5 %.

III Bayesian approach with prior knowledge. The same considerations as mentioned for model II are used here. No prior knowledge for the mean value should be formulated, because significant differences between the resistances of the types of specimen (A to F, see table 3) might be expected. For the prior distribution function only the parameters of the coefficients of variation are determined on basis of the six series given in table 3: $m(V_R) = 0.0546$ and $V(V_R) = 0.524$.

IV Evaluation on basis of a design model without prior knowledge. The same considerations as mentioned for model II are used here. On basis of an engineering judgement it can be stated that the parameter h has a linear influence and the parameter t has a squire root influence (both parameters are assumed deterministic):

$$R = C \cdot h^{1.0} t^{0.5} \qquad (4)$$

The analysis determines the design value of the factor C.

V Evaluation on basis of a design model with prior knowledge. The same considerations and design model as mentioned for model IV are used here. For the prior distribution function the same parameters as given for model III are taken.

The calculations have been carried out with the program SCEPTRE developed by TNO [6]. The design values for the six considered series are graphically presented in figure 3.

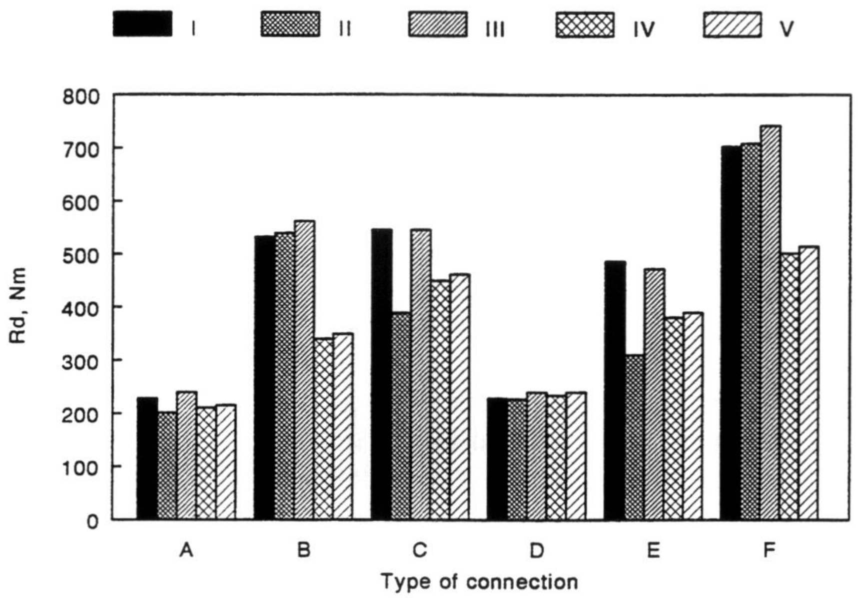

Figure 3 - Design values of types of specimen A to F according to models I to V

4.3 Remarks

The design values according to model II in a few cases significant lower (for type C and E) than those calculated by model I and III. The calculations indicate that the partial factor for design is for those two types is also very high. It can be stated that for this kind of test a total number of 4 results might give very conservative design values in case of model III. Model I is not effected by this lack of prior knowledge.

In case of the evaluation on basis of a design model it can be seen that the design values are significantly lower than those determined according to the other approaches. This indicates that the assumed design model according to equation (4) does not fully describe the physical behaviour. This discrepancy can be caused by several facts. E.g. the fact that several hooks are loaded is not taken into account. If more is known about the physical behaviour, the proposed design formulae can be reformulated. But it is noticed here that the prescribed formulae that can be used according to the theory of [5], are limited and that it might be impossible to give a correct description of the physical behaviour.

It is also marked that in case of the evaluation on basis of a design model no influence of prior knowledge can be seen. This is caused by the fact that the prior knowledge is based on exactly the same test results as those used in the analysis. This means that the prior knowledge is not independently from the evaluated test results.
Another item not discussed in this example is the choice of the lognormal distribution of the test results. From calculations not presented here, it is observed that in case of a normal distribution the design values are mostly lower and the differences between the approaches are more pronounced. So the chosen lognormal distribution leads to more economic design values.

5. Conclusions and recommendations

The following conclusions and recommendations are drawn:

- In Annex D of Eurocode 1 [1], the structural engineer can find valuable guidelines for the planning and evaluation of tests. The planning, the execution and the evaluation of tests require appropriate knowledge and experience.
- The well-known classical approach, the Bayesian approach and the evaluation on basis of a design model are discussed in this paper. For the last two prior knowledge can be incorporated. Also the possibility of using different distributions (normal and lognormal) of the test results is pointed out.
- In case of a small number of test results the Bayesian approach without using prior knowledge can give unrealistic design values. A more sophisticated model using prior knowledge might be a useful alternative. In case of the classical approach the use of a fixed partial factor of design is also a kind of prior knowledge.
- Prior knowledge should be formulated independently from the considered results.
- It is shown that the evaluation on basis of a design model is rather complicated. This is mainly due to the fact that a valid physical model have to formulated. It is noted that the possibilities of the prescribed formulae given in [5], are limited.

References

[1] ENV 1991-1, 'Eurocode 1 - Basis of Design and Actions on Structures - Part 1: Basis of Design', 1994.

[2] Raiffa, H., Schlaifer, R., 'Applied Statistical Decision Theory', Harvard University Press, 1961.

[3] Box, E.P., Tiao, G.C., 'Bayesian Inference in Statistical Analysis', Addison-Wesley Publishing Company, London, 1973.

[4] ISO TC 98/SC 2/WG 1, 'General Principles on Reliability for Structures - Revision of IS 2394 - Annex D: Design Based on Experimental Models', 6th draft, 1994.

[5] CEN/TC 250/SC 3, 'Eurocode 3 - Design of Steel Structures - Annex Z: Determination of Design Resistance from Tests', draft, 1994.

[6] Straalen, IJ.J. van, Vrouwenvelder, A.C.W.M., 'StatistiCal Evaluation Program for Test REsults (SCEPTRE) - Manual version 2.0', TNO-report 96-CON-R0130, TNO Building and Construction Research, The Netherlands, 1996.

[7] FEM 10.2.02, 'Recommendations for the Design of Steel Static Pallet Racking and Shelving', Federation Europeenne de la Manutention, to be published in 1995/1996.