# Machine strength grading – prediction limits – evaluation of a new method for derivation of settings

*R. Ziethén[1] & C. Bengtsson[2]*

## Abstract

This paper presents and analyses a new method for derivation of setting values for strength grading machines. The method uses a model for the relationship between indicating property (measured by a grading machine, IP) and the grade determining property (GDP). This model is of course not perfect and by summarising the errors in the model a confidence interval for the model can be calculated. The confidence interval is used for creating a prediction interval. The lower limit of the prediction interval, the prediction limit, is used to predict the GDP. The prediction limit method is analysed and evaluated by well defined input data. It is shown that fewer experimental data than required by the method in EN 14081-2 today is needed to determine reliable settings but the producer is awarded with less conservative settings with an increased number of experimental data. A weak correlation between IP and GDP or a high coefficient of variation also results in conservative settings. The settings are not dependent on average strength of the raw *material* used for deriving the settings.

## 1 Introduction

The results presented in this paper come from the European research project "GRADEWOOD". The main objective of the GRADEWOOD project is to enhance the use and to improve the reliability of wood as a structural material. One of the tasks in the project is to evaluate the European standard EN 14081 part 1 to 4 and to suggest possible improvements of the standard. The Standard EN 14081 consists of four parts:

| | |
|---|---|
| 14081-1 | General requirements for machine strength grading |
| 14081-2 | Derivation and supervision of settings using the machine control method |
| 14081-3 | Derivation and supervision of settings using the output control method |
| 14081-4 | Approved settings for existing grading machines |

This paper describes an alternative to the procedure described in EN 14081 part 2, the so called "Cost matrix method", Rouger 1997.

## 2 Background

### 2.1 Complexity

The method described in EN 14081-2 is complicated and it is also an issue of several interpretations of the text written. Because of this a special task group TG1 was established to interpret the standard and to assess reports with settings for strength grading machines based on these interpretations. An effect of this is an ongoing change of requested information and provisions for sampling and calculations. Today it

[1] Researcher, Email: rune.ziethen@sp.se
Trätek, Box 857, SE-501 15 Borås, Sweden
[2] Head of section. Email: charlotte.bengtsson@sp.se
SP Trätek, Box 857, SE-501 15 Borås, Sweden,
School of Engineering, Linnaeus University, SE-351 95 Växjö, Sweden

is not possible for anyone to have settings approved based on a report in accordance with the valid standard.

The determination of settings is based on the 5th-percentile for each grade in each grade combination. A result of this is several settings for each grade. Examples exist where one grade for one species from one geographic area has eight different settings, dependent on grade combination and the combination of countries covered. This leads to confusion. It is difficult to explain for producers and users of graded timber that different settings can be used depending on if the strongest material is included or if it is graded to higher grade.

The optimum grading in the standard is based on destructive bending or tension tests according to the European standards EN 384 and EN 408 and it serves as a “key” for the cost matrix analysis which is central for the determination of settings. The optimum grading is defined to be “the highest grade, of those for which settings are required, to which a piece of timber can be assigned, such that the grade determining properties of the graded sample will meet the values required for the grade”. This definition is unambiguous only if one grade and rejects are graded. When a combination of grades is graded a higher number of pieces in the highest class also results in a higher number of rejects. So far the optimum grading has been interpreted to be optimized from the highest grade and a higher number of rejects have been accepted although the yield from this grading has not been in line with the demands from the end-users. This interpretation can have a significant effect on the results from the cost-matrix analysis and thus a major effect on the determined settings.

### 2.2 Sensitivity

The standard requires a minimum sample of 900 pieces sampled in at least 8 sub-samples. For additional grades, origins or species the requirement is 450 pieces from at least 4 sub-samples. Although this high number of required test pieces the results are sensitive to small variations in properties of a few of the test pieces.

The minimum allowed setting for any class is based on a requirement of 0.5% or 5 pieces rejected from the samples used. For the most common commercial grades, C16, C18 and C24, this is generally the used requirement. For these grades more than 99% of the tested material is not at all used for the determination of the settings. For the lower tail of the distribution the measured values and also the relation between the machine value and the grade determining property is unstable. An element of random has a not negligible effect on settings determined by this requirement.

Also for other grades the settings are highly depending on a few observations. One example (based on real data) shows that a change of density of 2 to 4 kg/m³ for 7 out of 700 pieces in a sample can change the settings more than 5% and change the yield when grading with almost 10%. The economic effect of this change is large and can be the difference on the market between a competitive and a useless grading machine.

## 3 Theory

### 3.1 Linear regression

The examples above show some severe shortcomings of the present standard. An alternative to the method described in the standard is needed, Ziethén, R, Bengtsson C., 2009. A method that uses more of the information given by the tests, one alternative is linear regression, Figure 1. Linear regression is a simple well defined statistical tool, Jørgensen 1993.

Depending on the assumed distribution of the variables the regression can be used on linear or logarithmic values. Linear values imply a constant standard deviation for the variables and logarithmic values imply a constant coefficient of variation for the variables. The principles and equations are the same in both cases.

## 3.2 Confidence interval

### 3.2.1 Randomly distributed errors

The fitness of the regression model is often described by the coefficient of determination, $r^2$. A more informative method is to study the residual errors, the residual error $e_i$ is given by Equation (1)

$$e_i = Y_i - \left\lfloor \hat{\beta}_0 + \hat{\beta}_1 x_i \right\rfloor \quad (1)$$

These errors can have a number of different causes, as example:

- Measurement errors in any of the two variables
- Imperfection of the model
- Undetermined variables with an effect on the measured variables

The errors can be assumed to be normally distributed around the regression line, *Figure 1*

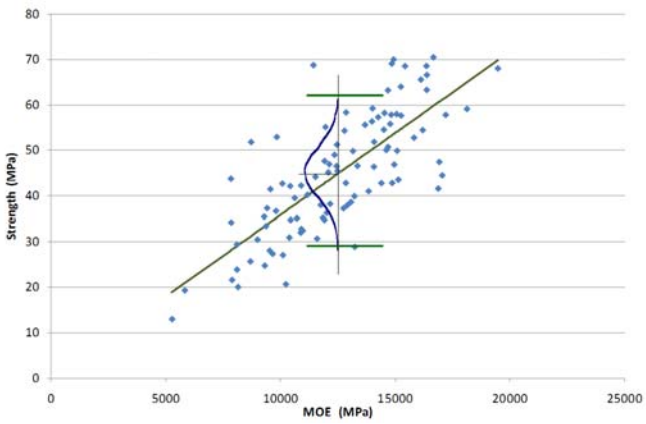


Figure 1 Linear regression with normally distributed residual errors

### 3.2.2 Calculation of confidence interval

From the estimated variance of the errors the variance of the coefficients $\hat{\beta}_0$ and $\hat{\beta}_1$ can be calculated. By adding the variances of the parameters and their covariance the confidence interval of the regression model, Equation (2), can be calculated.

$$Conf\ \mathrm{int}\left[\hat{y}(x)\right] \approx \pm t \cdot \sqrt{\hat{\sigma}^2 \left( \frac{(x-\bar{x})^2}{\sum_{i=1}^{n}(x_i-\bar{x})^2} + \frac{1}{n} \right)} \quad (2)$$

Where:
***t*** is taken from the student-t distribution with n-2 degrees of freedom.

I Figure 2 we can see the confidence interval for the regression line. Away from the mean value the interval is wider than close to the mean value. This is the effect of the error in the determination of the slope and the error in the determination of the intersection (the variance of the two coefficients).

### 3.3 Expanded prediction interval

Grading of timber shows another type of statistical challenge: to predict a future not yet observed observation. We will never know the true value but with increased sample size we can estimate the value with higher precision. The variance for the predicted value is given by Equation (3).

$$Var\left[y^*(x)\right] \approx \hat{\sigma}^2 \left( \frac{(x-\bar{x})^2}{\sum_{i=1}^{n}(x_i-\bar{x})^2} + \frac{1}{n} + 1 \right) \qquad (3)$$

From the variance we can calculate an interval for the predicted value according to Equation (4).

$$y^*(x) = \hat{\beta}_0 + \hat{\beta}_1 x \pm \hat{\sigma} \cdot t \cdot \sqrt{\frac{(x-\bar{x})^2}{\sum_{i=1}^{n}(x_i-\bar{x})^2} + \frac{1}{n} + 1} \qquad (4)$$

Where:
***t*** is taken from the student-t distribution with n-2 degrees of freedom.

The prediction interval shows big similarities with the confidence interval but it is expanded with a constant 1 under the root-sign. If the number of the observations used for the regression is large, the first two terms under the root-sign can be neglected. The equation for the prediction interval is then given by Equation (5).

$$y^*(x) = \hat{\beta}_0 + \hat{\beta}_1 x \pm \hat{\sigma} \cdot t \qquad (5)$$

The expanded prediction interval together with the confidence interval is shown in Figure 2

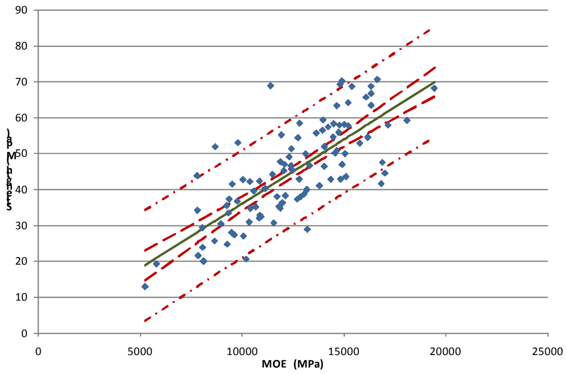


Figure 2 Regression line with confidence interval and the expanded prediction interval

## 4 Analysis of the method

### 4.1 General

Methods to determine settings for strength grading machines must give settings that are reliable but still not too conservative. The settings must also be robust and fair between different types and brands of grading machines. To evaluate these aspects of the proposed prediction limit method it was used on a number of simulated datasets with well defined properties. The use of simulated data can of course be questioned because of the absence of extreme values as well as out-layers.

## 4.2 Computer-generated data sets

Computer-generated data sets were used in the evaluation of the procedure, Ziethén, R, Bengtsson C., 2010. The properties varied are: mean value, coefficient of variation (COV) and coefficient of determination ($r^2$). Eight data sets with 20 000 observations were generated with assistance from the University of Ljubljana. The generated data sets were chosen to fit the distributions defined by COST Action E24 (JCSS probabilistic model code www.jcss.ethz.ch):

- MOR log-normal distribution
- MOE log-normal distribution
- Density normal distribution
- IP log-normal distribution

The properties of the master data set are based on experience from tests both regarding mean value, COV and correlation (mainly from GRADEWOOD). Similar data sets have been generated with different mean values but with constant COV, different COV but with constant mean value and correlations and with different correlations between MOR and IP but with constant mean value COV.

### 4.2.1 Sensitivity to data distribution

The distribution of data can be described by the mean value and the COV. The analyses in this paper are all based on the relation between IP and strength. MOE and density will of course give different numerical results but the principle results and conclusions will not be different.

In Figure 3 data sets with three different mean values are analysed using logarithmic values. From the figure it can be seen that for data with the same relation between strength and IP the mean value for the sample has no effect on the prediction limit.

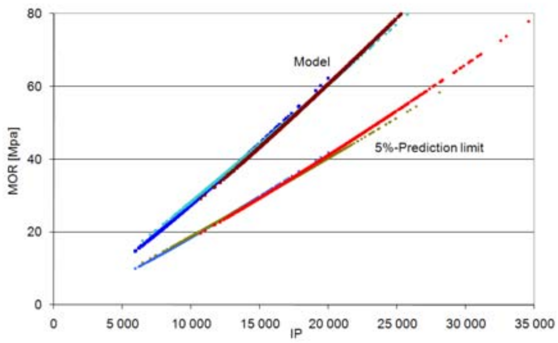


Figure 3 Linear regression line and lower 5%-prediction limit for datasets with different mean values

### 4.2.2 Sensitivity to model and correlation

The model and thus the correlation between the grading parameter (IP) and strength have a large impact on the 5%-prediction limit. A weaker correlation results in a more conservative limit, Figure 4.

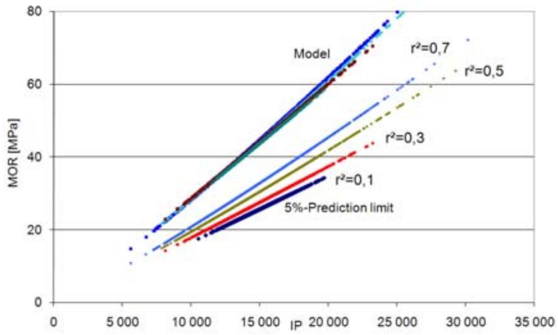


Figure 4 Linear regression line and lower 5%-prediction limit for datasets with four different correlation $r^2$

Using a linear approach instead of the logarithmic will result in more conservative settings for high strength grades and less conservative values for low strength grades. Earlier studies have also shown that the method is stable, not sensitive to out layers and not very sensitive to the number of observations, although less test pieces results in more conservative settings.

## 4.3 Evaluation based on test results

The evaluations based on the computer generated data sets have been supplemented with the results obtained from WP3 in Gradewood. In the tables below the settings for spruce in both bending and tension are summarized, all settings are presented in percent relative to the lowest setting for each grade. The settings calculated are based on the 5% prediction limit for strength and for density and the regression model for MOE, the required values are according to EN 338. These values are one possibility for the combination of calculation and requirements, it must not be regarded as the only possible choice of prediction level or requirements. The results obtained will follow the same pattern for any combination of prediction level and requirement.

### 4.3.1 Bending strength

In Table 1 the relative settings for four grades are presented. For some grades from some countries no value is given. It means that for one or more of the grade determining properties no yield was obtained and it was therefore not possible to determine settings for the grade. The property determining the settings are given in Table 2.

Table 1 Relative settings for bending based grades. Settings shown as a percentage of the lowest setting for each grade.

| Grade | Ukraine | Slovenia | Slovakia | Sweden | Romania | Poland |
|---|---|---|---|---|---|---|
| C 35 | 108 % | 103 % | 110 % | 100 % | 105 % | 100 % |
| C 30 | 110 % | 102 % | 108 % | 101 % | 110 % | 100 % |
| C 24 | 112 % | 103 % | 110 % | 102 % | 112 % | 100 % |
| C 18 | 107 % | 102 % | 106 % | 103 % | 107 % | 100 % |

Table 2 Critical property for different grades and counties for determination of settings

| Grade | Ukraine | Slovenia | Slovakia | Sweden | Romania | Poland |
|---|---|---|---|---|---|---|
| C 35 | MOR | MOR | MOR | MOR | MOE | MOR |
| C 30 | MOE | MOR | MOR | MOR | MOE | MOR |
| C 24 | MOE | MOE | MOE | MOR | MOE | MOR |
| C 18 | MOE | MOR | MOE | MOR | MOE | MOR |

From Table 1 it is difficult to find a systematic geographic reason for the differences of settings. But it can be seen, from Table 2, that MOE is more common to be the critical property for the countries with higher settings.

### 4.3.2 Tension strength

In an earlier part of Gradewood it was shown that similar settings could be used for large parts of Europe, that result is not confirmed here. In Table 3 the settings for four different grades based on tension strength is presented for different countries. The property determining the settings are given in Table 4.

Table 3 Relative settings for tension based grades. Settings shown as a percentage of the lowest setting for each grade. No value indicates no yield for the grade.

| Grade | Ukraine | Slovenia | Slovakia | Sweden | Romania | Poland | Switzer-land |
|---|---|---|---|---|---|---|---|
| L36 | | 120 % | | 107 % | | 100 % | 116 % |
| L30 | | 112 % | 112 % | 107 % | | 100 % | 121 % |
| L25 | | 108 % | 104 % | 101 % | 121 % | 100 % | 120 % |
| L17 | 106 % | 113 % | 108 % | 100 % | 111 % | 103 % | 115 % |

Table 4 Critical property for different grades and counties for determination of settings

| Grade | Ukraine | Slovenia | Slovakia | Sweden | Romania | Poland | Switzer-land |
|---|---|---|---|---|---|---|---|
| L36 | Density | Density | Density | Density | Density | Density | MOE |
| L30 | Density | Density | Density | Density | Density | MOR | MOE |
| L25 | Density | MOR | MOE | MOE | Density | MOR | MOE |
| L17 | MOR | MOR | MOR | MOR | MOR | MOR | MOE |

In Table 4 it can be see that density is the critical property for almost all data-sets for high grades. This raises the question whether the requirement profile really is the optimal.

If we only look at the strength there are settings for all classes. However, even for settings based on strength there are significantly high differences between different countries,

Table 5 Relative settings for tension based grades. Settings determined only based on strength and shown as a percentage of the lowest setting for each grade.

| Grade | Ukraine | Slovenia | Slovakia | Sweden | Romania | Poland | Switzer-land |
|---|---|---|---|---|---|---|---|
| L36 | 104 % | 106 % | 104 % | 101 % | 124 % | 100 % | 104 % |
| L30 | 104 % | 108 % | 104 % | 100 % | 121 % | 100 % | 104 % |
| L25 | 105 % | 110 % | 105 % | 100 % | 116 % | 102 % | 106 % |
| L17 | 106 % | 113 % | 108 % | 100 % | 111 % | 103 % | 107 % |

4.3.3 Splitted or combined growth areas

Today there is a focus on nation based settings. However, it is difficult to see a reason for a change of wood properties based on borders between nations. In this section it is shown how the differences between nations can be reflected in a calculation of an area with more than one country, Table 6. Sweden and Romania are chosen as an example. The number of observations is close to the same for both countries and there is a small difference for the settings. The calculations are based only on bending strength.

Table 6 Relative settings for bending based grades. Settings determined as a percentage of the lowest setting for each grade.

| Grade | Sweden | Romania | Combined Sweden and Romania |
|---|---|---|---|
| C 35 | 100 % | 101 % | 101 % |
| C 30 | 100 % | 102 % | 102 % |
| C 24 | 100 % | 103 % | 103 % |
| C 18 | 100 % | 105 % | 104 % |

We can see in this case that the conservative settings are dominating. This is however depending on a number of conditions such as the number of observations in the different data sets and the correlation between IP and the grade determining property for the combined data set compared to the correlations for each of the original data sets. It may be necessary to include requirements for data sets possible to combine into a common area.

The smallest area with the same settings allowed in the present standard is a country. When we evaluate possibilities to combine countries into larger areas it is important to look at the differences we must accept within a country when we are looking at restrictions for combined areas. As an example of this the settings for pine in bending are presented.

Table 7 Relative settings for bending based grades. Settings based only on bending strength are determined as a percentage of the lowest setting for each grade.

| Grade | Sweden | South Sweden | North Sweden | Poland | Poland, Swietjana | Poland, Murow |
|---|---|---|---|---|---|---|
| C 35 | 100 % | 105 % | 100 % | 112 % | 104 % | 119 % |
| C 30 | 100 % | 104 % | 100 % | 111 % | 106 % | 117 % |
| C 24 | 100 % | 103 % | 100 % | 112 % | 107 % | 116 % |
| C 18 | 100 % | 103 % | 102 % | 113 % | 110 % | 117 % |

A first look at settings shows a rather big difference between Sweden and Poland, app. 12 %. A closer look reveals that it is mainly one sample from Poland deviating from the rest with a difference of 15 – 20 % whilst the other Polish sample is close to the Swedish samples, Table 7. It is important that the requirements for different sub areas, when countries are combined, are of the same order as the differences that can be seen within the smallest allowed areas.

## 5 Conclusions

The evaluation of the proposed prediction limit method for calculation of settings has shown that:

- The calculated settings are not dependent on average strength of the raw material
- Weak correlation between indicating property and grade determining property results in more conservative settings
- Few observations for calculation of settings results in more conservative settings
- A combination of samples with slightly different regression models (from different areas) gives more conservative settings than for each of the areas alone.
- For a combination of samples were the differences in regression models are large the settings are less conservative than the most conservative settings for the sub-areas
- This indicates that special requirements on sample sizes and regression models can be needed if samples are combined.
- The differences within the present minimum acceptable grading area(country) are big. It is not reasonable to have harder restrictions when settings are determined for areas combining a number of countries than for the determination of settings for areas within a country.

## 6 References

Jørgensen, B.: The theory of linear models, CRC Press, 1993.

Rouger, F.: A new statistical method for the establishment of machine settings, CIB W18, proceedings paper 30-17-1, Vancouver, Canada, 1997.

EN 384 Structural timber – Determination of characteristic values of mechanical properties

EN 408 Structural timber – Test methods

EN 338 Structural timber – Strength classes

EN 14081 Timber structures – Strength graded structural timber with rectangular cross section, part 1-4

Ziethén, R, Bengtsson C.: Machine Strength Grading – a New Method for Derivation of Settings, CIB W18, proceedings paper 42-5-1, Dübendorf, Switzerland, 2009

Ziethén, R, Bengtsson C.: Machine strength grading – Prediction limits – evaluation of a new method for derivation of settings, WCTE, Riva de Garda, 2010