NTNU – Trondheim
Norwegian University of
Science and Technology

The Influence of Grading on the Reliability of Timber Structures

Ola Bakken Lindseth

Civil and Environmental Engineering (2 year)
Submission date: June 2015
Supervisor: Jochen Kohler, KT

Norwegian University of Science and Technology
Department of Structural Engineering

FOREWORD

The work on this thesis has been very interesting and educational, but, unfortunately, also filled with quite a few weeks of work that did not lead to useful results. Trying to find simplifications and approximations to expressions led me down different paths, but neither of them led me to the answer I was looking for. In addition, quite a lot of time was spent to learn the necessary skills in Bayesian data analysis in order to be able to perform a linear regression and create expressions based on this knowledge. The nature of this thesis also required quite a lot of work in MATLAB, and since I had practically no knowledge of this program from before (I had used it to a very small degree several years ago only), quite a lot of time was spent to learn the basics of this software. In addition, quite a lot of time was spent trying to debug code and eliminate errors that showed up as I was writing scripts and functions. Considering all this time spent, I now feel, with the finished thesis in front of me, that much of the work that has been put in is not necessarily shown in the final product. At the same time, however, I hope that the results that are presented in this paper are of high quality and usefulness, and that the reader finds them of interest.

Despite the hurdles along the way, I am very satisfied with the choice of topic for my thesis; it has been a very educational and rewarding process to work through this subject matter. The work would not have been possible without the insightful guidance of my supervisor, Jochen Köhler, who set me on the right path, guided me under way, and helped me reach the goal in the end. Big thanks to him! I would also like to thank my parents, Marit Bakken Lindseth and Iver Arvid Lindseth, for supporting me through my studies. Without their support, I would probably not have gotten to the point of even being able to begin with this thesis. Lastly, warm thanks to my dear Ekaterina Danilevskaya, who have put up with me through many late nights of working and has been positive the entire time, making the work feel like a breeze. To all other friends and family: I am not going to name you all here, but thank you for being there and supporting me!

Trondheim, June 2015 | Ola Bakken Lindseth

SUMMARY

In the present thesis, the influence of grading on the reliability of timber structures is analyzed. The thesis develops necessary expressions and explains methods for performing this type of analysis, and presents an analysis of a set of given data.

The analysis of the given data shows that changing grading machine settings can reduce the probability of failure of a structure in where the graded timber is to be used by up to approximately 25 % at lower grades, and up to approximately 15 % at higher grades. In addition, it shows that the influence of the accuracy of the grading machine has a significantly larger impact than that of the machine settings. Comparing a grading machine with a relatively high degree of accuracy to a machine with a relatively low degree of accuracy shows that the use of the grading machine with the lower accuracy gives a probability of failure that is up to 750 % higher than that of the more accurate one.

SAMMENDRAG

I denne masteroppgaven analyseres det hvordan gradering av tømmer påvirker sikkerheten og påliteligheten av konstruksjoner hvor det graderte tømmeret vil bli brukt. Oppgaven utleder nødvendige uttrykk og forklarer metoder for å gjennomføre en slik analyse, og presenterer en analyse av et gitt sett med data.

Analysen av det gitte datasettet viser at endringer av innstillingene for en graderingsmaskin kan redusere sannsynligheten for svikt i en konstruksjon hvor det graderte tømmeret er brukt med opptil cirka 25 % for lavere tømmerklasser, og opptil cirka 15 % for høyere klasser. I tillegg vises det at påvirkningen av en maskins nøyaktighet er vesentlig høyere enn effekten fra maskininnstillingene. Ved å sammenligne en graderingsmaskin med en relativt høy nøyaktighet med en maskin med relativt lav nøyaktighet, vises det at bruk av maskinen med den lavere nøyaktigheten gir en sannsynlighet for svikt som er opp til 750 % høyere enn for den mer nøyaktige maskinen.

TABLE OF CONTENTS

1 INTRODUCTION

The topic of this thesis is as follows:

> In the present Master Project existing control schemes for timber grading machines are analyzed and assessed in regard to their ability to reduce the variability of timber material properties. A benchmark study is performed on how different quality control schemes influence the reliability of timber structures.

Unlike structural materials like concrete or steel, timber is not produced from a "recipe" – timber specimens are gathered from nature, and these have grown under certain conditions (quality of the soil, amount of sunlight, density of trees, etc.), which influence the properties of the timber as a structural material. Because of this, the material properties, like strength, elasticity and density, cannot easily be determined without a relatively large uncertainty. For structural materials that are produced by man, the combination of ingredients and the process that makes the material is directly influencing the material properties. This causes the distribution of material properties to have a relatively low variation, since many of the factors that influence this variation is controlled by the producer of the material. For timber, this variation is very large, in comparison, since we do not produce the material in this manner.

In order to be able to determine the material properties with a higher degree of certainty, timber is divided into grades. Timber assigned to one of these grades have a smaller variation of the material properties than ungraded timber. There are different kinds of wood species that are used for structural timber, and a number of geographical areas in where these species are grown. It is therefore possible to divide the population of all timber into smaller populations, based on e.g. geographical area and species. This already gives a smaller variation of the material properties, but still the variation within each subpopulation is relatively large, and therefore the timber is subdivided further, into the aforementioned timber grades. Since each of these grades have a lower variation of the material properties than their parent population, a more optimal use of the timber can be achieved.

There are different ways of grading timber: visual grading, in where a qualified person inspects the timber specimen visually and based on this assigns a grade, or machine grading, where the timber specimens are run through a machine that assigns grades to the different specimens. Visual grading is based on a subjective, visual evaluation of things like knots and fissures in the wood, and it is very dependent on the person performing the grading – two different people could assign different grades to the same sample of timber specimens. Machine grading is based on nondestructive tests performed by a machine on the timber specimens. Based on these tests, indicating properties corresponding to the material properties (strength, elasticity, and density) are given. Based on these indicating properties, the timber specimens are assigned to a given grade (or they are rejected or accepted to a given grade).

For a given structural material, the uncertainty of the material properties influence the reliability of the structure in which the material is to be used. The reliability of a structure can be expressed in terms of the probability of failure. A higher degree of uncertainty of the material properties would give a different probability of failure than a lower degree of uncertainty.

1.1 SCOPE OF WORK

In this thesis, the grading process will be looked at, and we will consider how different ways of grading timber influences the reliability of structures in which the graded timber is to be used. Visual grading will not be looked at, only machine grading of timber will be analyzed. The different aspects of the grading process that will be looked at are the use of different grading machines, of varying degree of accuracy, and different settings of the individual machines. In the grading process, material properties which influence both the ultimate limit state (failure of structure) and the serviceability limit state (deflection, durability, vibration) are taken into account. In this thesis, only the ultimate limit state will be considered, and grading based on the strength of the timber only, since this is what influences the reliability, the probability of failure, of the structure. Serviceability failures will not be discussed.

In order to consider the influence of the grading process on the reliability of structures, a set of steps will be performed. Firstly, a regression analysis of the relation between the indicating property from the grading machine and the material property of the timber will be performed. This will tell something about the accuracy of the grading machines, relative to each other. In addition, it will show the result we are mostly interested in, how the probability distribution of the material property is, for a given value of the indicating property.

Once the regression has been performed, we will consider how different settings for a given grading machine can give the same timber grade, and how the different settings change the distribution of the material property within the same grade. The use of different grading machines also affects the distribution of the material property, and this effect will also be analyzed.

Lastly we will look at how the different probability distributions of the material property (material resistance – tensile strength) for a given grade influences the reliability of a general structure with a general load. In this way, we can see how the grading process (the use of different grading machines and different settings for the different machines) influences the reliability of the structure.

In chapters 3 and 4, we will present the formulas necessary to perform our analysis, and in chapter 5, we will present numbers and figures from a given dataset. The figures and numbers presented will have been calculated and created in MATLAB. The scripts and functions that created these figures and numbers will not be presented in the text itself, but can be found written in Appendix B. The raw data that has been used can be found in Appendix C.

2 NOTATION

Throughout this thesis, there will be presented a large number of formulas and symbols, and therefore we give an overview of these symbols here. The notation mainly follows that as used in (Gelman, et al., 2014) and (Schneider, 1997), with some additional symbols for our specific cases. We also choose to express vectors and matrices using bold characters; this differs from the notation used in (Gelman, et al., 2014).

2.1 SYMBOLS

Below is presented a table with an overview of symbols that will be used when working with probability distributions, regression and reliability calculations. A general description of the symbols is presented along with a specific description of the use of the figure in the context of timber grading and reliability calculations.

Symbol	General	Timber grading/Reliability
$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$	A column vector of n outcome variables.	y_i are the natural logarithm of measurements of the material property of interest.
$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$	A column vector of n explanatory variables	x_i are the natural logarithm of registered indicating properties, corresponding to the logarithm of the material properties, y_i.
$\mathbf{X} = \begin{bmatrix} \mathbf{x_1} & \mathbf{x_2} & \cdots & \mathbf{x_k} \end{bmatrix}$ $\mathbf{X} = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1k} \\ x_{21} & x_{22} & \cdots & x_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nk} \end{bmatrix}$	An $n \times k$ matrix of predictors. There are n observations (corresponding to the n outcome variables in $\mathbf{y}$) and k explanatory variables per observation.	An $n \times 2$ matrix where $x_{i1} = 1$ and x_{i2} are natural logarithms of measurements of the indicating property of interest.
$\boldsymbol{\beta} = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_k \end{bmatrix}$	A column vector of k regression parameters.	A column vector of 2 regression parameters.
$\sigma^2 = \mathrm{var}\left(y_i \middle\vert \boldsymbol{\beta}, \sigma, \mathbf{X}\right)$	The population variance of the outcome variables.	
$\boldsymbol{\varepsilon} = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$	A column vector of n realizations of the error of the regression model.	
s^2	An estimate of σ^2	

$\tilde{\mathbf{y}}$	Future values of the outcome variable, $\mathbf{y}$	
$\tilde{\mathbf{X}}$	Future values of the predictors, $\mathbf{X}$	
IP		Indicating property from grading machine
MP		Material property of the timber
μ	Mean value	
$\overline{x}$	Sample mean	
P_f		Probability of failure
$p(\cdot)$	Probability density function	
$f_R(\cdot)$		Probability density function of the material resistance
$f_S(\cdot)$		Probability density function of the applied stress
$F_R(\cdot)$		Cumulative distribution of the material resistance
$F_S(\cdot)$		Cumulative distribution of the applied stress
z		Structural factor, taking into account geometry, dimensions, etc.
R		Stochastic variable of the material resistance
S		Stochastic variable of the applied stress
r_k		Characteristic value of the material resistance
s_k		Characteristic value of the applied stress
γ_R		Safety factor for the material resistance
γ_S		Safety factor for the applied stress

Table 2.1: Symbols

2.2 PROBABILITY DISTRIBUTIONS

We will use some common probability distributions when performing our data analysis, and we will use the following notation: If we, for example, have a normally distributed variable, x, with mean μ and variance σ^2, we will write it as follows:

$$x \sim \mathrm{N}\left(\mu, \sigma^2\right).$$

We will use a similar notation when referring to other common distributions. To see the specific parameters used in the different distributions, see Appendix A.

2.3 MATLAB SPECIFIC CALCULATIONS

Throughout this paper, alternative formulations of expressions will sometimes be shown. These are specific to when calculating in MATLAB. These additional calculations and formulations have been presented in grey boxes to separate them from the main text, like the following:

> This is a MATLAB specific section.
>
> It explains how we can reformulate something in order to calculate it using MATLAB.

3 GRADING OF TIMBER

In this chapter, how the different ways of grading timber influence the probability distribution of the material property will be analyzed. As mentioned earlier, only machine grading will be considered, where the grading machines gives indicating properties for the corresponding material properties (strength, elasticity, density). In the context of reliability of structures, the strength of the material is of interest, and thus we choose to look at this material property only, combined with the corresponding indicating property from the grading machines. However, the formulas presented in this chapter can also be applied to other material properties and corresponding indicating properties.

The timber grades are given as the characteristic value of the strength. The characteristic value of the strength is defined as the 5 % fractile of the distribution of the strength. This means that a random specimen from timber grade C30 has a 5 % chance to have a tensile strength that is less than 30 MPa, and similarly for other grades. The timber grade says nothing of how the material property is distributed outside of this one value. To illustrate this, we present in Figure 3.1 a set of probability densities of the strength, and in Figure 3.2 a corresponding set of cumulative distributions, all with a 5 % fractile of 24 MPa, but with different distributions for the material property. From the figures, it can be seen that timber with a relatively high degree of uncertainty regarding the material property and timber with a relatively low degree of uncertainty can both belong to the same timber grade.

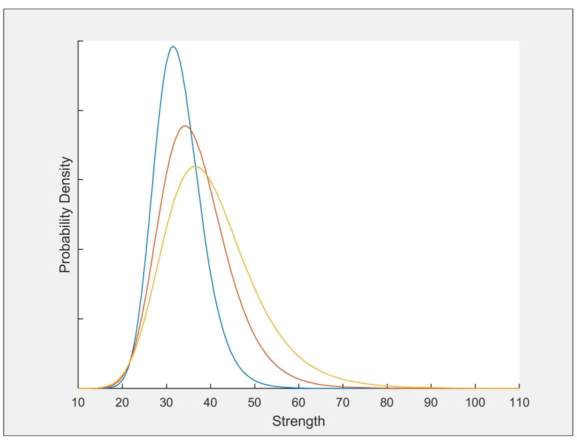


Figure 3.1: Probability densities – C24

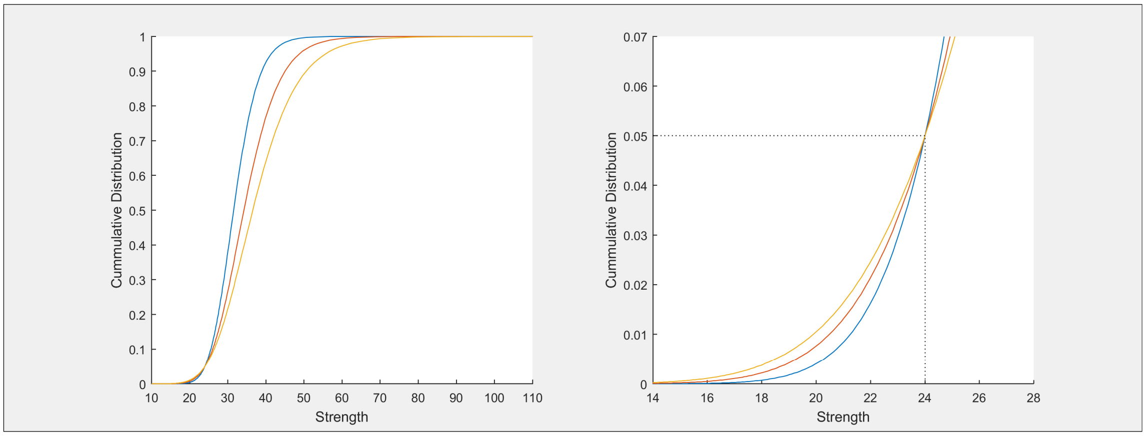


Figure 3.2: Cumulative distributions – C24

In light of this, we are interested in looking at how two different factors affect the distribution of the material property for a given timber grade:

- The effect of different settings for a given grading machine
- The effect of using different grading machines, with different levels of accuracy

The settings for a grading machine is understood as the range of indicating properties that accepts a timber specimen into the given grade. For example, say that the acceptance criteria for grade C40 for a given grading machine is indicating property between 473 and 831 (this is not numbers from a specific machine, just a random example). In this case, all timber specimens with an indicating property within this range will be accepted to the given grade, while all specimens with indicating property outside this range will be rejected. As mentioned above, there are normally more than one indicating property, and each of the indicating properties has to lie within the corresponding range for the given grade in order for the specimen to be accepted, but we will only consider one indicating property in this paper: the strength.

In the coming sections, in order to find the distribution of the material property, we will first have a look at the relation between the indicating property and the material property, and from that, we will make an expression for the distribution, given the machine setting.

3.1 LINEAR REGRESSION

We wish to analyze the relation between the indicating properties of the grading machine and the physical properties of the test specimens. We choose to perform a linear regression, using a Bayesian approach, as new knowledge, which may be obtained during the grading process, can be integrated into the model. See chapter 2 for an explanation of the notation used.

3.1.1 Transformation of data

A lognormal distribution is assumed (if $\ln(x)$ is normal distributed, then x is lognormal distributed), for both the explanatory variable (indicating property, IP), and the outcome variable (material property, MP). Therefore, before the regression is performed, the indicating properties and material properties are transformed as follows:

$$x = \ln(IP) \tag{3.1}$$

$$y = \ln(MP) \tag{3.2}$$

As the regression and further calculations are performed, these transformed values, x and y, will be used. The values will only be transformed back when presenting final values and when generating figures.

3.1.2 Basic Bayesian model with non-informative prior

The regression will be performed with a non-informative prior distribution. Prior data may be incorporated into the regression model, but this is not something that will be discussed in this thesis. For more information on Bayesian data analysis, (Gelman, et al., 2014) discusses this topic in detail. The expressions given in these sections are general for any linear regression using a Bayesian approach. The expressions and formulas are presented without formal proof, and an in-depth explanation of Bayesian data analysis is not given, but the expressions are presented to give a basic understanding of where the result we get comes from. A more in-depth discussion can be found in (Gelman, et al., 2014).

For our given observations of $\mathbf{y}$ and $\mathbf{X}$ we have the following:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$$

In words: the outcome variable is given as the explanatory variables multiplied by the regression parameters in $\boldsymbol{\beta}$, plus an error. The realizations of the error, given in $\boldsymbol{\varepsilon}$, tells us how far from the linear regression line the observed value lies.

The elements of $\boldsymbol{\varepsilon}$, ε_i, are assumed to be normal distributed with a mean equal to 0 and a population variance equal to σ^2:

$$\boldsymbol{\varepsilon} \sim \mathrm{N}\left(\mathbf{0}, \sigma^2\mathbf{I}\right)$$

where $\mathbf{0}$ is a column vector of n zeros, and $\mathbf{I}$ is an $n \times n$ identity matrix.

We can write this (ordinary linear regression) as follows:

$$\mathbf{y} \mid \boldsymbol{\beta}, \sigma, \mathbf{X} \sim \mathrm{N}\left(\mathbf{X}\boldsymbol{\beta}, \sigma^2\mathbf{I}\right) \tag{3.3}$$

where $\mathbf{I}$ is an $n \times n$ identity matrix. This states that, given the parameters $\boldsymbol{\beta}$ and σ^2 and predictors $\mathbf{X}$, $\mathbf{y}$ is normal distributed with mean $\mathbf{X}\boldsymbol{\beta}$ and variance σ^2.

3.1.3 The posterior distribution

The regression parameters, $\boldsymbol{\beta}$, given the variance, σ^2, is multivariate normal distributed:

$$\boldsymbol{\beta}|\sigma,\mathbf{y} \sim \mathrm{N}\left(\hat{\boldsymbol{\beta}},\mathbf{V}_{\boldsymbol{\beta}}\sigma^2\right) \tag{3.4}$$

The mean and variance is calculated as follows:

$$\mathrm{E}\left(\boldsymbol{\beta}|\sigma,\mathbf{y}\right) = \hat{\boldsymbol{\beta}} = \left(\mathbf{X}^{\mathbf{T}}\mathbf{X}\right)^{-1}\mathbf{X}^{\mathbf{T}}\mathbf{y} \tag{3.5}$$

$$\mathrm{var}\left(\boldsymbol{\beta}|\sigma,\mathbf{y}\right) = \mathbf{V}_{\boldsymbol{\beta}}\sigma^2 = \left(\mathbf{X}^{\mathbf{T}}\mathbf{X}\right)^{-1}\sigma^2 \tag{3.6}$$

When working with numerical computations, we want to avoid using inverse matrices. In MATLAB, one should use the built-in matrix division operator, as it is better, in terms of both execution time and numerical accuracy, than calculating the inverse matrix. We therefore (using the MATLAB matrix division operator, $\backslash$) reformulate the equation for $\hat{\boldsymbol{\beta}}$:

$$\begin{aligned}
&\hat{\boldsymbol{\beta}} = \left(\mathbf{X}^{\mathbf{T}}\mathbf{X}\right)^{-1}\mathbf{X}^{\mathbf{T}}\mathbf{y}\\
&\mathbf{X}^{\mathbf{T}}\mathbf{X}\hat{\boldsymbol{\beta}} = \mathbf{X}^{\mathbf{T}}\mathbf{y}\\
&\hat{\boldsymbol{\beta}} = \left(\mathbf{X}^{\mathbf{T}}\mathbf{X}\right)\backslash\left(\mathbf{X}^{\mathbf{T}}\mathbf{y}\right)
\end{aligned}$$

Similarly we can express the variance of $\boldsymbol{\beta}$ like this:

$$\begin{aligned}
&\mathrm{var}\left(\boldsymbol{\beta}|\sigma,\mathbf{y}\right) = \mathbf{V}_{\boldsymbol{\beta}}\sigma^2 = \left(\mathbf{X}^{\mathbf{T}}\mathbf{X}\right)^{-1}\sigma^2\\
&\mathbf{X}^{\mathbf{T}}\mathbf{X}\,\mathrm{var}\left(\boldsymbol{\beta}\,|\,\sigma,\mathbf{y}\right) = \sigma^2\mathbf{I}\\
&\mathrm{var}\left(\boldsymbol{\beta}|\sigma,\mathbf{y}\right) = \left(\mathbf{X}^{\mathbf{T}}\mathbf{X}\right)\backslash\left(\sigma^2\mathbf{I}\right)
\end{aligned}$$

where $\mathbf{I}$ is a $k \times k$ identity matrix.

An estimate of the variance, σ^2, can be found like this:

$$s^2 = \frac{1}{n-k}\left(\mathbf{y}-\mathbf{X}\hat{\boldsymbol{\beta}}\right)^T\left(\mathbf{y}-\mathbf{X}\hat{\boldsymbol{\beta}}\right) \tag{3.7}$$

The distribution of the variance, σ^2, has a scaled inverse chi-square form:

$$\sigma^2|\mathbf{y} \sim \text{Inv-}\chi^2\left(n-k,s^2\right) \tag{3.8}$$

To perform a simulation draw from the scaled inverse chi-square distribution, we can first draw a value, w, from the chi-square distribution with $n-k$ degrees of freedom and then let $\sigma^2 = \left(n-k\right)s^2/w$.

The classical, non-Bayesian, standard error estimate for $\boldsymbol{\beta}$ is obtained by setting $\sigma = s$ in (3.4).

3.1.4 Posterior predictive distribution, given a new set of observations

We are interested in finding the probability distribution of the material property for future specimens, given the indicating property from a grading machine. The posterior predictive distribution therefore has to be found.

For a new set of observations, $\tilde{\mathbf{X}}$, from which the outcomes, $\tilde{\mathbf{y}}$, should be predicted, it can be expressed in the following way:

$$\tilde{\mathbf{y}} \sim \mathrm{N}\left(\tilde{\mathbf{X}}\boldsymbol{\beta}, \sigma^2\mathbf{I}\right) \tag{3.9}$$

This, however, requires that both $\boldsymbol{\beta}$ and σ are known exactly. Our knowledge of these parameters are summarized by our posterior distribution, described above. Using simulation we can therefore first draw σ and $\boldsymbol{\beta}$ from (3.8) and (3.4), and then draw $\tilde{\mathbf{y}}$ from (3.9).

In the case of a normal linear model, we can also determine the posterior predictive distribution analytically. Given σ, we have the following:

$$\tilde{\mathbf{y}}|\sigma, \mathbf{y} \sim \mathrm{N}\left(\tilde{\mathbf{X}}\hat{\boldsymbol{\beta}}, \left(\mathbf{I} + \tilde{\mathbf{X}}\mathbf{V}_{\boldsymbol{\beta}}\tilde{\mathbf{X}}^{\mathbf{T}}\right)\sigma^2\right) \tag{3.10}$$

Again, when working in MATLAB, we want to avoid inverse matrices (instead using MATLAB's matrix division operator). Therefore, we choose to do the following with the variance of the posterior predictive distribution:

$$\begin{aligned}
&\operatorname{var}\left(\tilde{\mathbf{y}}|\sigma, \mathbf{y}\right) = \left(\mathbf{I} + \tilde{\mathbf{X}}\mathbf{V}_{\boldsymbol{\beta}}\tilde{\mathbf{X}}^{\mathbf{T}}\right)\sigma^2 \\
&\tilde{\mathbf{X}}\mathbf{V}_{\boldsymbol{\beta}}\tilde{\mathbf{X}}^{\mathbf{T}} = \tilde{\mathbf{X}}\mathbf{W} \\
&\mathbf{W} = \mathbf{V}_{\boldsymbol{\beta}}\tilde{\mathbf{X}}^{\mathbf{T}} = \left(\mathbf{X}^{\mathbf{T}}\mathbf{X}\right)^{-1}\tilde{\mathbf{X}}^{\mathbf{T}} \\
&\mathbf{X}^{\mathbf{T}}\mathbf{X}\mathbf{W} = \tilde{\mathbf{X}}^{\mathbf{T}} \\
&\mathbf{W} = \left(\mathbf{X}^{\mathbf{T}}\mathbf{X}\right) \backslash \tilde{\mathbf{X}}^{\mathbf{T}} \\
&\tilde{\mathbf{X}}\mathbf{V}_{\boldsymbol{\beta}}\tilde{\mathbf{X}}^{\mathbf{T}} = \tilde{\mathbf{X}}\left(\left(\mathbf{X}^{\mathbf{T}}\mathbf{X}\right) \backslash \tilde{\mathbf{X}}^{\mathbf{T}}\right) \\
&\operatorname{var}\left(\tilde{\mathbf{y}}|\sigma, \mathbf{y}\right) = \left(\mathbf{I} + \tilde{\mathbf{X}}\left(\left(\mathbf{X}^{\mathbf{T}}\mathbf{X}\right) \backslash \tilde{\mathbf{X}}^{\mathbf{T}}\right)\right)\sigma^2
\end{aligned}$$

Where, if $\tilde{\mathbf{X}}$ is an $m \times k$ matrix (m observations with k explanatory variables per observation), $\mathbf{W}$ would be a $k \times m$ matrix and $\mathbf{I}$ an $m \times m$ identity matrix (the variance-covariance matrix would also, naturally, be an $m \times m$ matrix).

Since σ is not known exactly, it must be averaged over the marginal posterior distribution of σ^2 given in (3.8), which gives a posterior predictive distribution of $\tilde{\mathbf{y}}$, given $\mathbf{y}$, which is multivariate t with location $\tilde{\mathbf{X}}\hat{\boldsymbol{\beta}}$, squared scale matrix $\left(\mathbf{I} + \tilde{\mathbf{X}}\mathbf{V}_{\boldsymbol{\beta}}\tilde{\mathbf{X}}^{\mathbf{T}}\right)s^2$, and $\nu = n - k$ degrees of freedom:

$$\tilde{\mathbf{y}}|\sigma, \mathbf{y} \sim \mathrm{t}_{n-k}\left(\tilde{\mathbf{X}}\hat{\boldsymbol{\beta}}, \left(\mathbf{I} + \tilde{\mathbf{X}}\mathbf{V}_{\boldsymbol{\beta}}\tilde{\mathbf{X}}^{\mathbf{T}}\right)s^2\right) \tag{3.11}$$

3.2 TIMBER GRADING

The expression in (3.11) can be used to express the distribution of the material property, given a single value of the indicating property. We are, however, interested in finding the distribution of the material property, given a timber grade. The result in (3.11) is therefore in itself not something we can apply directly to our data, but we will use it to find the distribution we are interested in: given that the indicating property of interest fulfills the acceptance criteria of the given grade, how will the corresponding material property be distributed?

We define our acceptance criteria for the indicating property as follows: $IP_L \leq IP \leq IP_H$. In words, the indicating property must lie between a lower and an upper limit.

3.2.1 Distribution of indicating property

The distribution of the material property depends on the distribution of the indicating property. Within the range of the indicating property for a given timber grade, not all values of IP have the same probability of occurrence. Because of this, the distribution of the material property will be skewed, compared to the expression we found in (3.11).

The distributions looks different for different grades, as the distribution of the indicating property will be different within the limits for the different grades. That is, if we are working on low grades, the values of indicating property we are looking at are on the left side of the distribution of IP , while if we are working with high grades, the IP stays on the right side. This means that our weighting function will have a different shape. We can see this by the illustration below:

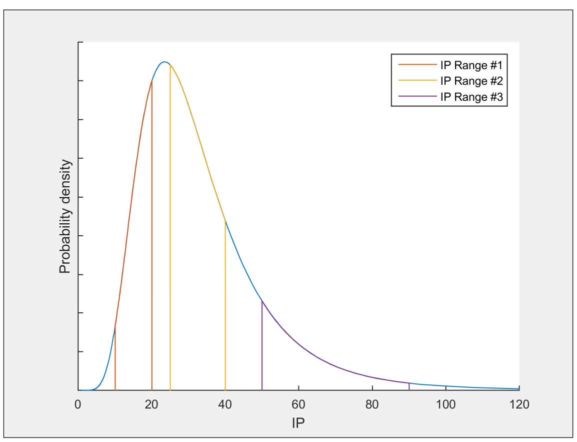


Figure 3.3: Example of distribution of IP

In the figure, the vertical lines indicate areas where the indicating property might lie for different grades, and we can observe that the weighting of the distribution of the material property, as a result, is quite different for the different grades (the vertical lines here are just illustrative examples, and do not represent settings for actual grades).

Because of this, before we can continue with finding the distribution of the material property, we need to find the distribution of the indicating property. We continue with the assumption of a lognormal distribution, as described in chapter 3.1.1:

$$\ln\left(IP\right) = x \sim \mathrm{N}\left(\mu_x, \sigma_x^{\ 2}\right) \tag{3.12}$$

We choose to use subscript x when working with the distribution of the indicating property, in order to distinguish it from the distribution we used in our linear regression above (which expressed the distribution of the material property and the distributions of parameters connected to that distribution).

If we do not know the parameters (population mean and variance) exact, we can use a sample from our population to find an estimate of the distribution of the parameters. In the example in chapter 5, the same sample is used for calculating the distribution of the indicating property as to perform the regression. This is not necessary, and the use of subscript x, as mentioned above, will help show this in the expressions below.

3.2.1.1 Normal data with a non-informative prior distribution

Given a sample of n_x observations from the population, we have the following distribution of the mean, given the population variance:

$$\mu_x \left| \sigma_x^2, \mathbf{x} \sim \mathrm{N}\left(\bar{x}, \sigma_x^2 / n_x\right)\right. \tag{3.13}$$

Estimates for the mean and variance can be calculated as follows:

$$\bar{x} = \tfrac{1}{n_x} \sum_{i=1}^{n_x} x_i$$

$$s_x^2 = \tfrac{1}{n_x - 1} \sum_{i=1}^{n_x} \left(x_i - \bar{x}\right)^2$$

The population variance has a scaled inverse chi-square distribution:

$$\sigma_x^2 \left| \mathbf{x} \sim \text{Inv-}\chi^2\left(n_x - 1, s_x^2\right)\right. \tag{3.14}$$

We can draw σ_x from (3.14) and then μ_x from (3.13), or we can use the following:

$$\mu_x \left| \mathbf{x} \sim \mathrm{t}_{\mathrm{n}_x\text{-}1}\left(\bar{x}, s_x^2 / n_x\right)\right.$$

Or, in words, the population mean is t distributed with location $\bar{x}$, scale $s_x \big/ \sqrt{n_x}$ and $n_x - 1$ degrees of freedom. Using the t distribution with center 0, we can formulate it like this:

$$\left.\frac{\mu_x - \bar{x}}{s_x \big/ \sqrt{n_x}}\right| \mathbf{x} \sim t_{n_x - 1}$$

3.2.1.2 Posterior predictive distribution

Again, the distribution of future observations of IP is of interest, and we need to use the posterior predictive distribution in order to find this. The posterior predictive distribution for a future observation, $\tilde{x}$ will be given as a t distribution with location $\bar{x}$, scale $\sqrt{1 + \tfrac{1}{n_x}} s_x$ and $n_x - 1$ degrees of freedom:

$$\tilde{x} \left| \mathbf{x} \sim \mathrm{t}_{\mathrm{n}_x\text{-}1}\left(\bar{x}, \left(1 + \tfrac{1}{n_x}\right) s_x^2\right)\right. \tag{3.15}$$

3.2.2 Distribution of material property

Now that both our regression and distribution of the indicating property has been found, we can start to look at the distribution of the material property, given the indicating property lies within the range specified by a given timber grade.

For a single value of the indicating property, IP_*, we have the following:

$$\ln\left(IP_*\right) = \tilde{x}_*$$
$$\tilde{\mathbf{X}}_* = \begin{bmatrix}1 & \tilde{x}_*\end{bmatrix}$$

From the linear regression, we have the posterior predictive distribution given in (3.9). Using this, for the given indicating property we have:

$$\tilde{y}\middle|\tilde{x}_* \sim \mathrm{N}\left(\tilde{\mathbf{X}}_*\boldsymbol{\beta}, \sigma^2\right) = \mathrm{N}\left(\beta_1 + \beta_2\tilde{x}_*, \sigma^2\right)$$

If the parameters $\boldsymbol{\beta}$ are not known, we can also use (3.10).

We wish to find the distribution of the material property, given indicating properties fulfill an acceptance criterion:

$$\tilde{y}\middle| x_L \le \tilde{x} \le x_H$$

where $x_L = \ln\left(IP_L\right)$ and $x_H = \ln\left(IP_H\right)$.

The easiest way to find this is by simulation, as described in the following section. It is also possible to formulate an expression for the distribution, which can be solved numerically. We come back to this in chapter 3.2.2.2.

3.2.2.1 Simulation

We can easily simulate the distribution of the material property, given a range of the indicating property, in the following way:

1. We draw σ_x^2 from (3.14) (or we use a known value)
2. We draw μ_x from (3.13) (or we use a known value)
3. We draw $\tilde{x}$ from (3.12) and discard any values outside the range of our acceptance criteria
4. We draw σ^2 from (3.8) (or we use a known value)
5. We draw $\boldsymbol{\beta}$ from (3.4) (or we use known values)
6. We draw $\tilde{y}$ from (3.9)

3.2.2.2 Integration

We will now formulate an expression that can be used to give the distribution of the material property. For now, let us assume that the parameters $\boldsymbol{\beta}$, σ, μ_x and σ_x are all known. We then have normal distributions, and if we write the expressions for the distributions, we get the following probability densities:

$$p\left(\tilde{y}\middle|\tilde{x}_*\right) = \frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{1}{2\sigma^2}\left(\tilde{y}-\left(\beta_1+\beta_2\tilde{x}_*\right)\right)^2\right)$$

$$p\left(\tilde{x}\right) = \frac{1}{\sqrt{2\pi}\sigma_x}\exp\left(-\frac{1}{2{\sigma_x}^2}\left(\tilde{x}-\mu_x\right)^2\right)$$

As mentioned before, we wish to find the distribution of the material property, given a range of indicating properties. We therefore want to integrate the density of the material property over the range of indicating properties. However, since the indicating properties are not uniformly distributed, we want to weight the expression with the distribution of the indicating property.

Going forward, in order to simplify the expressions, we omit the use of ~ above our x and y. It is understood that we are expressing the distributions of future observations, not the data used in our regression.

The unnormalized distribution of the material property, given range of indicating property, can be expressed as the distribution of MP, given a single value of IP, weighted by the distribution of IP, and integrated over the range of IP:

$$\begin{aligned}
p\left(y\middle|x_L \le x \le x_H\right) &\propto \int_{x_L}^{x_H} p\left(y\middle|x\right)p\left(x\right)dx \\
&= \int_{x_L}^{x_H} \tfrac{1}{\sqrt{2\pi}\sigma}\exp\left(-\tfrac{1}{2\sigma^2}\left(y-\left(\beta_1+\beta_2 x\right)\right)^2\right)\tfrac{1}{\sqrt{2\pi}\sigma_x}\exp\left(-\tfrac{1}{2{\sigma_x}^2}\left(x-\mu_x\right)^2\right)dx \\
&= \tfrac{1}{2\pi\sigma\sigma_x}\int_{x_L}^{x_H} \exp\left(-\tfrac{1}{2\sigma^2}\left(y-\left(\beta_1+\beta_2 x\right)\right)^2-\tfrac{1}{2{\sigma_x}^2}\left(x-\mu_x\right)^2\right)dx \\
&\propto \int_{x_L}^{x_H} \exp\left(-\tfrac{1}{2\sigma^2}\left(y-\left(\beta_1+\beta_2 x\right)\right)^2-\tfrac{1}{2{\sigma_x}^2}\left(x-\mu_x\right)^2\right)dx
\end{aligned}$$

In order to find the probability density, we need to normalize the expression. We do this by adding a normalizing constant:

$$p\left(y\middle|x_L \le x \le x_H\right) = C\int_{x_L}^{x_H} \exp\left(-\tfrac{1}{2\sigma^2}\left(y-\left(\beta_1+\beta_2 x\right)\right)^2-\tfrac{1}{2{\sigma_x}^2}\left(x-\mu_x\right)^2\right)dx$$

In order to find the value of C, we integrate the expression over all possible values of y and set this equal to 1 (we know that a probability distribution should always equal to 1 when integrated over all possible values).:

$$\int_{-\infty}^{\infty} C \int_{x_L}^{x_H} \exp\left(-\tfrac{1}{2\sigma^2}\left(y-\left(\beta_1+\beta_2 x\right)\right)^2 - \tfrac{1}{2\sigma_x{}^2}\left(x-\mu_x\right)^2\right) dxdy = 1$$

$$\frac{1}{C} = \int_{-\infty}^{\infty} \int_{x_L}^{x_H} \exp\left(-\tfrac{1}{2\sigma^2}\left(y-\left(\beta_1+\beta_2 x\right)\right)^2 - \tfrac{1}{2\sigma_x{}^2}\left(x-\mu_x\right)^2\right) dxdy$$

We can see that, naturally, the normalizing constant also depends on the chosen range of IP.

Putting the expression for the unnormalized probability density and the normalizing constant together, we get the distribution of y, given range of x:

$$p\left(y \middle| x_L \le x \le x_H\right) = \frac{\int_{x_L}^{x_H} \exp\left(-\tfrac{1}{2\sigma^2}\left(y-\left(\beta_1+\beta_2 x\right)\right)^2 - \tfrac{1}{2\sigma_x{}^2}\left(x-\mu_x\right)^2\right) dx}{\int_{-\infty}^{\infty} \int_{x_L}^{x_H} \exp\left(-\tfrac{1}{2\sigma^2}\left(y-\left(\beta_1+\beta_2 x\right)\right)^2 - \tfrac{1}{2\sigma_x{}^2}\left(x-\mu_x\right)^2\right) dxdy} \tag{3.16}$$

This expression for the probability density of the natural logarithm of the material property, given the indicating property lies within our limits, assumes that all population parameters are known. In our case, they are not, and therefore we need to use t-distributions instead of the normal distributions (as expressed in (3.11) and (3.15)). Besides this difference, the way of thinking is the same.

The distribution of the indicating property would be t with location $\overline{x}$, scale $\sqrt{1+\tfrac{1}{n_x}}s_x$ and $n_x - 1$ degrees of freedom. The distribution of the material property would be t with center $\widehat{\beta}_1 + \widehat{\beta}_2 x$, scale $\sqrt{1+\begin{bmatrix}1 & x\end{bmatrix}\mathbf{V}_{\boldsymbol{\beta}}\begin{bmatrix}1 & x\end{bmatrix}^T}s$ and $n-2$ degrees of freedom:

$$p\left(\tilde{y} \middle| \tilde{x}_*\right) = \frac{\Gamma\left(\left(n-1\right)/2\right)\left(1+\dfrac{\left(\tilde{y}-\left(\widehat{\beta}_1+\widehat{\beta}_2\tilde{x}_*\right)\right)^2}{\left(n-2\right)\left(1+\begin{bmatrix}1 & \tilde{x}_*\end{bmatrix}\mathbf{V}_{\boldsymbol{\beta}}\begin{bmatrix}1 & \tilde{x}_*\end{bmatrix}^T\right)s^2}\right)^{-(n-1)/2}}{\Gamma\left(\left(n-2\right)/2\right)\sqrt{\left(n-2\right)\pi\left(1+\begin{bmatrix}1 & \tilde{x}_*\end{bmatrix}\mathbf{V}_{\boldsymbol{\beta}}\begin{bmatrix}1 & \tilde{x}_*\end{bmatrix}^T\right)}s}$$

$$p\left(\tilde{x}\right) = \frac{\Gamma\left(n_x/2\right)}{\Gamma\left(\left(n_x-1\right)/2\right)\sqrt{\left(n_x-1\right)\pi\left(1+\tfrac{1}{n_x}\right)}s_x}\left(1+\frac{\left(\tilde{x}-\overline{x}\right)^2}{\left(n_x-1\right)\left(1+\tfrac{1}{n_x}\right)s_x^2}\right)^{-n_x/2}$$

Using this we put it in the same expression as before (removing any constants not dependent on x or y):

$$p\left(y \middle| x_L \le x \le x_H\right) \propto \int_{x_L}^{x_H} p\left(y \middle| x\right) p\left(x\right) dx$$

$$\propto \int_{x_L}^{x_H} \frac{\left(1+\frac{\left(x-\overline{x}\right)^2}{\left(n_x-1\right)\left(1+\frac{1}{n_x}\right)s_x^2}\right)^{-n_x/2}}{\sqrt{\left(1+\left[1 \quad x\right]\mathbf{V}_{\boldsymbol{\beta}}\left[1 \quad x\right]^T\right)}} \left(1+\frac{\left(y-\left(\widehat{\beta}_1+\widehat{\beta}_2 x\right)\right)^2}{\left(n-2\right)\left(1+\left[1 \quad x\right]\mathbf{V}_{\boldsymbol{\beta}}\left[1 \quad x\right]^T\right)s^2}\right)^{-(n-1)/2} dx$$

where n is the sample size for our regression and n_x is the sample size for calculating the distribution of the indicating property. Normalizing the expression, we get:

$$p\left(y \middle| x_L \le x \le x_H\right) = \frac{\int_{x_L}^{x_H} p\left(y \middle| x\right) p\left(x\right) dx}{\int_{-\infty}^{\infty}\int_{x_L}^{x_H} p\left(y \middle| x\right) p\left(x\right) dxdy}$$

$$= C\int_{x_L}^{x_H} \frac{\left(1+\frac{\left(x-\overline{x}\right)^2}{\left(n_x-1\right)\left(1+\frac{1}{n_x}\right)s_x^2}\right)^{-n_x/2}}{\sqrt{\left(1+\left[1 \quad x\right]\mathbf{V}_{\boldsymbol{\beta}}\left[1 \quad x\right]^T\right)}} \left(1+\frac{\left(y-\left(\widehat{\beta}_1+\widehat{\beta}_2 x\right)\right)^2}{\left(n-2\right)\left(1+\left[1 \quad x\right]\mathbf{V}_{\boldsymbol{\beta}}\left[1 \quad x\right]^T\right)s^2}\right)^{-(n-1)/2} dx$$

$$\frac{1}{C} = \int_{-\infty}^{\infty}\int_{x_L}^{x_H} \frac{\left(1+\frac{\left(x-\overline{x}\right)^2}{\left(n_x-1\right)\left(1+\frac{1}{n_x}\right)s_x^2}\right)^{-n_x/2}}{\sqrt{\left(1+\left[1 \quad x\right]\mathbf{V}_{\boldsymbol{\beta}}\left[1 \quad x\right]^T\right)}} \left(1+\frac{\left(y-\left(\widehat{\beta}_1+\widehat{\beta}_2 x\right)\right)^2}{\left(n-2\right)\left(1+\left[1 \quad x\right]\mathbf{V}_{\boldsymbol{\beta}}\left[1 \quad x\right]^T\right)s^2}\right)^{-(n-1)/2} dxdy$$

$$p\left(y \middle| x_L \le x \le x_H\right) = \frac{\int_{x_L}^{x_H} \frac{\left(1+\frac{\left(x-\overline{x}\right)^2}{\left(n_x-1\right)\left(1+\frac{1}{n_x}\right)s_x^2}\right)^{-n_x/2}}{\sqrt{1+\left[1 \quad x\right]\mathbf{V}_{\boldsymbol{\beta}}\left[1 \quad x\right]^T}} \left(1+\frac{\left(y-\left(\widehat{\beta}_1+\widehat{\beta}_2 x\right)\right)^2}{\left(n-2\right)\left(1+\left[1 \quad x\right]\mathbf{V}_{\boldsymbol{\beta}}\left[1 \quad x\right]^T\right)s^2}\right)^{-(n-1)/2} dx}{\int_{-\infty}^{\infty}\int_{x_L}^{x_H} \frac{\left(1+\frac{\left(x-\overline{x}\right)^2}{\left(n_x-1\right)\left(1+\frac{1}{n_x}\right)s_x^2}\right)^{-n_x/2}}{\sqrt{1+\left[1 \quad x\right]\mathbf{V}_{\boldsymbol{\beta}}\left[1 \quad x\right]^T}} \left(1+\frac{\left(y-\left(\widehat{\beta}_1+\widehat{\beta}_2 x\right)\right)^2}{\left(n-2\right)\left(1+\left[1 \quad x\right]\mathbf{V}_{\boldsymbol{\beta}}\left[1 \quad x\right]^T\right)s^2}\right)^{-(n-1)/2} dxdy}$$

In order to make the expression more readable, we do the following:

$$p\left(y \middle| x_L \le x \le x_H\right) = \frac{\int_{x_L}^{x_H} \sigma_s^{-1}\left(1+\frac{\left(x-\overline{x}\right)^2}{\left(n_x-1\right)\sigma_{sx}^2}\right)^{-n_x/2}\left(1+\frac{\left(y-\mu_s\right)^2}{\left(n-2\right)\sigma_s^2}\right)^{-(n-1)/2} dx}{\int_{-\infty}^{\infty}\int_{x_L}^{x_H} \sigma_s^{-1}\left(1+\frac{\left(x-\overline{x}\right)^2}{\left(n_x-1\right)\sigma_{sx}^2}\right)^{-n_x/2}\left(1+\frac{\left(y-\mu_s\right)^2}{\left(n-2\right)\sigma_s^2}\right)^{-(n-1)/2} dxdy} \qquad (3.17)$$

Where we have the scale σ_{sx} for the distribution of the indicating property, and the scale, σ_s and location, μ_s from the regression:

$$\sigma_{sx} = \sqrt{1+\tfrac{1}{n_x}}s_x \tag{3.18}$$

$$\sigma_s = \sqrt{1+\begin{bmatrix}1 & x\end{bmatrix}\mathbf{V}_{\boldsymbol{\beta}}\begin{bmatrix}1 & x\end{bmatrix}^T}s \tag{3.19}$$

$$\mu_s = \hat{\beta}_1 + \hat{\beta}_2 x \tag{3.20}$$

Be aware that (3.19) and (3.20) are both functions of x, and as such must be included in the integration in (3.17).

This expression for the distribution of the natural logarithm of the material property, given a range of indicating property, is not solvable in closed form, but we will use this expression for numerical integrations going forward.

When performing numerical integration in MATLAB (using the built-in functions), we cannot use matrix operations, only scalar ones, and as such we need to reformulate the scale as follows (using Gauss-Jordan elimination):

$$\begin{aligned}
&\begin{bmatrix}1 & x\end{bmatrix}\mathbf{V}_{\boldsymbol{\beta}}\begin{bmatrix}1 & x\end{bmatrix}^T = \tilde{\mathbf{X}}\mathbf{V}_{\boldsymbol{\beta}}\tilde{\mathbf{X}}^{\mathbf{T}} = \tilde{\mathbf{X}}\mathbf{W}\\
&\mathbf{W} = \mathbf{V}_{\boldsymbol{\beta}}\tilde{\mathbf{X}}^{\mathbf{T}} = \left(\mathbf{X}^{\mathbf{T}}\mathbf{X}\right)^{-1}\tilde{\mathbf{X}}^{\mathbf{T}}\\
&\mathbf{X}^{\mathbf{T}}\mathbf{X}\mathbf{W} = \tilde{\mathbf{X}}^{\mathbf{T}}\\
&\begin{bmatrix}a & b\\ c & d\end{bmatrix}\begin{bmatrix}w_1\\ w_2\end{bmatrix} = \begin{bmatrix}1\\ x\end{bmatrix}\\
&\Rightarrow \mathbf{W} = \begin{bmatrix}w_1\\ w_2\end{bmatrix} = \begin{bmatrix}\tfrac{1}{a} - \tfrac{b}{a}\tfrac{ax-c}{ad-bc}\\ \tfrac{ax-c}{ad-bc}\end{bmatrix}\\
&\begin{bmatrix}1 & x\end{bmatrix}\mathbf{V}_{\boldsymbol{\beta}}\begin{bmatrix}1 & x\end{bmatrix}^T = \frac{1}{a} + \left(x - \frac{b}{a}\right)\frac{ax-c}{ad-bc}\\
&\sigma_s = s\sqrt{1+\frac{1}{a}+\left(x-\frac{b}{a}\right)\frac{ax-c}{ad-bc}}, \mathbf{X}^{\mathbf{T}}\mathbf{X} = \begin{bmatrix}a & b\\ c & d\end{bmatrix}
\end{aligned}$$

We can find the cumulative distribution by integrating from $-\infty$ to y:

$$\frac{\int_{-\infty}^{y}\int_{x_L}^{x_H}\sigma_s^{-1}\left(1+\frac{\left(x-\bar{x}\right)^2}{\left(n_x-1\right)\sigma_{sx}^2}\right)^{-n_x/2}\left(1+\frac{\left(t-\mu_s\right)^2}{\left(n-2\right)\sigma_s^2}\right)^{-(n-1)/2}dxdt}{\int_{-\infty}^{\infty}\int_{x_L}^{x_H}\sigma_s^{-1}\left(1+\frac{\left(x-\bar{x}\right)^2}{\left(n_x-1\right)\sigma_{sx}^2}\right)^{-n_x/2}\left(1+\frac{\left(y-\mu_s\right)^2}{\left(n-2\right)\sigma_s^2}\right)^{-(n-1)/2}dxdy} \tag{3.21}$$

In the case where we would know the population parameters of the distribution of the indicating property, but not the parameters $\boldsymbol{\beta}$ and σ, we could create a similar expression, using a combination of a t-distribution and a normal distribution. We will not use such an expression in this thesis, but it should be straightforward to make, using the same way of thinking as when we made (3.16) and (3.17).

3.3 Influence of Grading on Distribution of Material Property

Given the expression in (3.17), we will now describe how we can find the influence of the grading on the distribution of the material property. In chapter 5, we will show examples from the sample data given in Appendix C.

As mentioned earlier, we are interested in two effects, different settings of a given grading machine, and use of different grading machines.

3.3.1 The influence of grading machine settings

We are interested in seeing how different ranges of indicating property affects the distribution of the material property. This can be found by looking at (3.17). For a given grade, considering the strength, the value of y when the cumulative distribution is 0.05 (the 5 % fractile) is known. Using this fact, the upper limit of the range of the indicating property for a given lower limit can be found, or vice versa. As an example, for grade C40, it can be done in the following way:

1. We set $y=\ln\left(40\right)$ (we are still working with the natural logarithms)
2. We set $x_H=+\infty$
3. The cumulative distribution is calculated from (3.21) for a random value of x_L
4. We change the value of x_L until the cumulative distribution equals to 0.05
5. We now know the range of x : $\left[x_L \quad x_H\right]$

Other ranges of x can be found similarly. The lowest possible value for x_L would be found when $x_H=+\infty$, and the highest possible value would be found when $x_L=x_H$. This last case is not a practical situation, as it is no range, just a single value, but it is the extreme case, and can be interesting to observe in order to see the possible ranges. When $x_L=x_H$, we would not use (3.17), but the expression found from the regression, (3.11), and the single value of x would be found similarly as described above.

Using these different grading machine settings (different ranges of x), the different distributions of y can be found, by putting the values of x_L and x_H back in to (3.17) (or, in the case of a single value of x, (3.11)). Plotting these distributions in the same figure would let us be able to observe the influence of the machine settings on the distribution, for a given grade.

3.3.1.1 Simulation

It is also possible to do this using simulation. A similar method as described in 3.2.2.1 would then be used. In order to find the range of IP, the following procedure can be followed:

1. We draw σ_x^2 from (3.14) (or we use a known value)
2. We draw μ_x from (3.13) (or we use a known value)
3. We choose a value for either x_L or x_H
4. We set a random value of the other limit, not chosen in step 3
5. We draw $\tilde{x}$ from (3.12) and discard any values outside the range of our acceptance criteria
6. We draw σ^2 from (3.8) (or we use a known value)
7. We draw $\boldsymbol{\beta}$ from (3.4) (or we use known values)
8. We draw $\tilde{y}$ from (3.9)
9. The draws of $\tilde{y}$ are sorted and we see what value is the one 5 % from the bottom. If this value is close enough to the grade we seek, we have a valid range of x. If it is not close enough, we go back to step 4 and choose a new value of the limit. Rinse and repeat until a valid range is found

This method can be quite time consuming, as it requires a lot of iteration to find the limits. Once the limits have been found, the cumulative distribution can be plotted, using the method described in 3.2.2.1.

3.3.2 The influence of different grading machines

In order to find the effect of using different grading machines, we have to perform the regression for the different machines, yielding different values of $\boldsymbol{\beta}$ and σ. In addition, the distribution of the indicating property would also be different, yielding different values of $\bar{x}$ and σ_x.

The different grading machines would give different values for IP, from the method described above. This because of different regression parameters, but also because the different grading machines might use a different scale for their IP. Because of this, in order to compare different grading machines, it is necessary to choose machine settings that are relatively the same. The two extremes (when $x_H = +\infty$ and $x_L = x_H$) can be considered the end points for all the different machines, and then, for example, the values for x_L can be distributed similarly between the two extremes for the different machines. As an example, let us say that we want to use three different ranges of IP for the different machines. We then choose the two extremes, and then the last value of x_L could be set as the middle value between the lowest possible x_L and the highest possible x_L.

Once the appropriate ranges of IP has been set for the different machines, we can choose the corresponding ranges for the different machines, and plot the distributions calculated from

(3.17) (or (3.11)). Given these plots, we can observe how the distribution varies between the machines. For example, we choose $x_H = +\infty$ for all machines.

To sum up, when comparing machine settings, we keep grading machine and timber grade constant, while varying the machine setting, and when comparing machines, we keep (relative) machine setting and timber grade constant.

4 RELIABILITY OF STRUCTURES

Once the distribution of the material property has been found, and it has been observed how this distribution varies for different machine settings and different grading machines, this can be used to look at the effect on the reliability of structures. The material property of interest is the material resistance (strength), and this will be used going forward. An in-depth explanation of reliability of structures in general will not be given, but a brief explanation will be presented. The notation used in this chapter follows mainly that given in (Schneider, 1997).

The reliability of a structure can be expressed in terms of the probability of failure. Failure means that the stress induced by the load on the structure exceeds the resistance of the material. This can be expressed in the form of the limit state function, G, which is defined as the material resistance (strength), R, minus the stress from the load, S:

$$G = R - S \tag{4.1}$$

The probability of failure, P_f, is given as the probability that the limit state function is non-positive:

$$P_f = \Pr(G \le 0) = \Pr(R \le S)$$

Both the material resistance and the stress are expressed by probability distributions. Using these distributions, the probability of failure can be calculated in the following way:

$$P_f = \int_{-\infty}^{+\infty} F_R(y) f_S(y) dy \tag{4.2}$$

Or, equivalently:

$$P_f = 1 - \int_{-\infty}^{+\infty} F_S(y) f_R(y) dy \tag{4.3}$$

Where f_S and F_S are the probability density and cumulative distribution of the stress, and f_R and F_R the density and cumulative distribution of the material resistance.

The distribution of the material resistance can be found in chapter 3, with f_R given in (3.17) and F_R expressed in (3.21). In order to find the probability of failure we also need to have a probability distribution of the load.

4.1 LOAD ON STRUCTURE

For the distribution of the load on the structure, we choose to use an extreme value (Gumbel Max) distribution with mean, $\mu = 1$ and standard deviation, $\sigma = 0.2$. The Gumbel Max distribution has the following probability density function (using s as the stress):

$$f_S(s) = \alpha \exp\left(-\alpha(s-u) - \exp\left(-\alpha(s-u)\right)\right) \tag{4.4}$$

The cumulative distribution function is given as follows:

$$F_S(s) = \exp\left(-\exp\left(-\alpha(s-u)\right)\right) \tag{4.5}$$

To determine the values for α , and u , we look at the expressions for the mean and standard deviation, which are:

$$\mu = u + \frac{\gamma}{\alpha} \approx u + \frac{0.577216}{\alpha} \tag{4.6}$$

$$\sigma = \frac{\pi}{\alpha\sqrt{6}} \tag{4.7}$$

where γ is the Euler–Mascheroni constant. Rearranging the expressions, we get:

$$\alpha = \frac{\pi}{\sigma\sqrt{6}} \tag{4.8}$$

$$u = \mu - \frac{\gamma}{\alpha} = \mu - \frac{\gamma\sigma\sqrt{6}}{\pi} \tag{4.9}$$

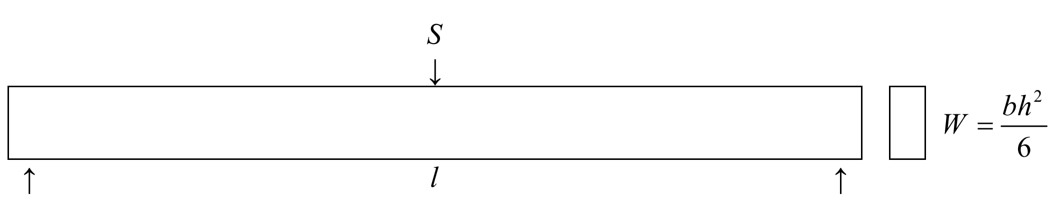
4.1.1 Relation between load and resistance

In order to find an appropriate relation between the distribution of the load and the distribution of the strength of our material, we look to our structural codes. In the structural codes, characteristic values for the load and resistance are used with safety factors and an ultimate limit is found by setting these equal, like this:

$$z \cdot s_k \cdot \gamma_S = {r_k}/{\gamma_R} \tag{4.10}$$

where s_k is the characteristic value of the load, γ_S is the safety factor for the load, r_k is the characteristic value of the resistance, γ_R is the safety factor for the material resistance, and z is a constant that depends on the structure.

To illustrate how one can find the value of z , we use a simple example of a beam with a point load in the middle of the span:



Using the bending moment as the measure we have the following for the load:

$$M_S = \frac{Sl}{4}$$

Similarly, we have the following expression for the moment resistance:

$$M_R = Wr$$

Setting the two moments equal to each other and using characteristic values and safety factors give us the following:

$$\frac{S_k \gamma_S l}{4} = \frac{W r_k}{\gamma_R}$$

Gathering all the factors that depend on the structure (geometry, dimensions) on one side, we get the following:

$$\frac{l}{4W} S_k \gamma_S = \frac{r_k}{\gamma_R}$$

Which is the same expression as stated in (4.10), with $z = l/4W$.

We want to find a value for z that, for our distributions of the load and resistance, satisfies equation (4.10):

$$z = \frac{r_k}{s_k \gamma_S \gamma_R} \tag{4.11}$$

The characteristic value of the resistance of the material is given as the 5 % fractile of the distribution, and we get this directly from our timber grade. The characteristic value of the load is given as the 98 % fractile of the distribution, and we can calculate it by setting the cumulative distribution equal to 0.98. We start by inversing the expression for the cumulative distribution:

$$\begin{aligned}
&F_S(s) = \exp\left(-\exp\left(-\alpha\left(s-u\right)\right)\right) \\
&-\exp\left(-\alpha\left(s-u\right)\right) = \ln\left(F_S\right) \\
&-\alpha\left(s-u\right) = \ln\left(-\ln\left(F_S\right)\right) \\
&s\left(F_S\right) = u - \frac{\ln\left(-\ln\left(F_S\right)\right)}{\alpha}
\end{aligned}$$

Incorporating (4.8) and (4.9), we get:

$$s\left(F_S\right) = \mu - \frac{\gamma\sigma\sqrt{6}}{\pi} - \frac{\sigma\sqrt{6}\ln\left(-\ln\left(F_S\right)\right)}{\pi}$$

$$s\left(F_S\right) = \mu - \frac{\sigma\sqrt{6}}{\pi}\left(\gamma + \ln\left(-\ln\left(F_S\right)\right)\right) \tag{4.12}$$

Setting $F_s = 0.98$ and $s(0.98) = s_k$, we get the expression for our characteristic value:

$$s_k = \mu - \frac{\sigma\sqrt{6}}{\pi}\left(\gamma + \ln\left(-\ln\left(0.98\right)\right)\right)$$

For our chosen distribution with mean $\mu = 1$ and standard deviation $\sigma = 0.2$ we then get the following characteristic value of the load:

$$s_k = 1 - \frac{0.2\sqrt{6}}{\pi}\left(\gamma + \ln\left(-\ln\left(0.98\right)\right)\right)$$
$$s_k \approx 1.518455$$

We now have the characteristic values, and we use the following safety factors:

$$\gamma_S = 1.5$$
$$\gamma_R = 1.3$$

For different timber grades, we can then calculate our z, using (4.11):

C24: $z \approx \dfrac{24}{1.518455 \cdot 1.5 \cdot 1.3} \approx 8.105404$

C30: $z \approx \dfrac{30}{1.518455 \cdot 1.5 \cdot 1.3} \approx 10.131755$

C40: $z \approx \dfrac{40}{1.518455 \cdot 1.5 \cdot 1.3} \approx 13.509006$

Now the distribution of the load is determined for three different grades (other values of z can be found similarly), and the limit state function can be written as:

$$G = R - zS \tag{4.13}$$

The distribution of zS also is a Gumbel Max, but with mean equal to $z\mu$. The coefficient of variation is as before, and the standard deviation will then, as a result, be $z\sigma$ (since z is always a positive number). We can see that this is correct by looking at (4.12):

$$s_z\left(F_S\right) = z \cdot s\left(F_S\right)$$
$$s_z\left(F_S\right) = z\left(\mu - \frac{\sigma\sqrt{6}}{\pi}\left(\gamma + \ln\left(-\ln\left(F_s\right)\right)\right)\right)$$
$$s_z\left(F_S\right) = z\mu - \frac{z\sigma\sqrt{6}}{\pi}\left(\gamma + \ln\left(-\ln\left(F_s\right)\right)\right)$$

We see that zS follows a Gumbel distribution with mean $z\mu$ and standard deviation $z\sigma$, as stated above.

4.2 LIMIT STATE

Using the values of z calculated above, we can now find the distributions of zS for the different grades, and we can use the limit state function given in (4.13):

$$G = R - zS$$

If we look at the expressions for the probability of failure, given in (4.2) and (4.3), we can see that it makes sense to use the last version, (4.3), since the distribution of the resistance is more easily expressed as the probability density (3.17), and the expression for the distribution of the load is simpler as the cumulative distribution, (4.5). We therefore choose to do this, and use the cumulative distribution of the load. The distribution of the load is the distribution of zS, as described above, with mean $z\mu$ and standard deviation $z\sigma$.

The distribution of the resistance and the distribution of the load must be expressed using the same unit of measure. Our distribution of the resistance is expressed as the natural logarithm of the strength, and so we also need to transform the distribution of the load using the natural logarithm. Using the transformation $y = \ln(s)$, we get:

$$F_S(s) = \exp\Big(-\exp\Big(-\alpha\big(\exp(\ln(s)) - u\big)\Big)\Big)$$
$$F_S^*(y) = \exp\Big(-\exp\Big(-\alpha\big(\exp(y) - u\big)\Big)\Big)$$

We can see that the cumulative distribution function is normalized correctly by checking the following:

1. When $y \to +\infty$, $\exp(y) \to +\infty$, and $\big(-\alpha(\exp(y) - u)\big) \to -\infty$
2. This means that $\exp\big(-\alpha(\exp(y) - u)\big) = \exp(-\infty) \to 0$
3. We then have $\lim_{y \to +\infty} F_S^*(y) = \exp(0) = 1$
4. Similarly we can show that $\lim_{y \to -\infty} F_S^*(y) = \exp(-\infty) = 0$

Which means that the cumulative distribution goes from zero to one, as it should. The cumulative distribution of the load, zS, which we will use going forward, can then be written as follows:

$$F_S(y) = \exp\left(-\exp\left(-\frac{\pi}{z\sigma\sqrt{6}}\left(\exp(y) - z\mu + \frac{\gamma z\sigma\sqrt{6}}{\pi}\right)\right)\right) \tag{4.14}$$

We choose to use subscript S in order to be consistent with (4.3), but be aware that we use the distribution of zS.

Where we have $\mu = 1$ and $\sigma = 0.2$, and z given for the grade we want to consider.

We have previously expressed the probability density of the resistance in (3.17):

$$f_R(y) = C\int_{x_L}^{x_H} \sigma_s^{-1}\left(1+\frac{(x-\overline{x})^2}{(n_x-1)\sigma_{sx}^2}\right)^{-n_x/2}\left(1+\frac{(y-\mu_s)^2}{(n-2)\sigma_s^2}\right)^{-(n-1)/2} dx$$

Where C is the normalizing constant.

In order to simplify the expressions, we choose to call the inner function of the integral $h(x,y)$. We can then have the following expression for the distribution:

$$f_R(y) = C\int_{x_L}^{x_H} h(x,y)\,dx$$

Using this we get the probability of failure expressed as follows:

$$\begin{aligned} P_f &= 1-\int_{-\infty}^{+\infty} F_S(y)\,f_R(y)\,dy \\ &= 1-\int_{-\infty}^{+\infty} F_S(y)\left(C\int_{x_L}^{x_H} h(x,y)\,dx\right)dy \end{aligned}$$

Since $F_S(y)$ does not depend on x, we can do the following:

$$P_f = 1-C\int_{-\infty}^{+\infty}\int_{x_L}^{x_H} h(x,y)\,F_S(y)\,dxdy \tag{4.15}$$

Where $h(x,y) = \sigma_s^{-1}\left(1+\frac{(x-\overline{x})^2}{(n_x-1)\sigma_{sx}^2}\right)^{-n_x/2}\left(1+\frac{(y-\mu_s)^2}{(n-2)\sigma_s^2}\right)^{-(n-1)/2}$ and $F_S(y)$ is given in (4.14).

We now have all we need to calculate the probability of failure of our general structure.

4.3 INFLUENCE OF GRADING ON RELIABILITY OF STRUCTURES

Ultimately, what we are interested in finding is how the grading of timber influences the reliability of the structure in where the timber is to be used. We now have all the tools we need to find and analyze this.

In order to find the probability of failure for the different grading machines and machine settings, we do the following:

1. We choose a timber grade we want to look at
2. Given the chosen timber grade, we can calculate the distribution of the load on the structure from (4.14)
3. We look at the different grading machines and settings as discussed in chapter 3.3, but instead of finding the distribution of the material resistance directly, we put the same parameters into (4.15) to calculate the probability of failure for the different combinations of grading machine and machine settings

In general, the probability of failure will be a discrete value for the given range of IP and grading machine, but if we want we can make a continuous graph for a given grading machine with the indicating property presented as the abscissa. To do this, it makes sense to make the lower limit of the range of the indicating property as the value along the horizontal axis (since it has a much smaller range than the upper limit, which can go to infinity). This can be done by expressing (4.15) as a function of x_L ($\ln(y)$ will be determined by the chosen timber grade and x_H is a function of x_L). An example of how to do this is presented in chapter 5.3.2

4.3.1 Simulation

The probability of failure can also be found by simulation. It can be performed in the following way:

1. Draw a value of the material resistance as expressed in chapter 3.2.2.1
2. Draw a value of the stress, using the distribution given in (4.14)
3. Calculate $G = R - S$ using the values drawn in steps 1 and 2
4. From all the draws, count how many of the G values that are negative and find the percentage of all the draws which resulted in a negative limit state function

This way of calculating the probability of failure can be quite time consuming, as it requires a lot of simulations for it to be accurate (since the probability of failure is very small). In addition, if the range of indicating properties is very small, there will be many draws of x that will be discarded, which again results in a longer time to compute. It is, however, a method which is very easy to understand, compared to the, perhaps a bit complicated, expression presented in (4.15).

5 EXAMPLE FROM DATA

Using the tools created in chapters 3 and 4, we will, in this chapter, present calculations and figures made from a data sample that contains material properties for a set of timber specimens, as well as corresponding indicating properties from a range of different grading machines. The data used can be found in Appendix C. All calculations and creation of figures were performed using MATLAB. The scripts and functions used can be found in Appendix B.

The effects of interest are, as mentioned in previous chapters, the use of different grading machines, and the use of different grading machine settings. The data consists of measurements of timber specimens from two different regions, and indicating properties from five different grading machines. The differences between regions are not of interest in this context, and as such only data from one region will be used. In terms of different grading machines, only the relative difference between grading machines, in terms of accuracy, is of interest. Therefore, three grading machines are chosen – Machine #1, the most accurate, Machine #3, the least accurate, and Machine #2, which is the grading machine with accuracy closest to the middle between the two other machines. The choice of grading machines was made after the regression, based on the estimate of the variance of the error, as expressed in (3.7) (low variance means high accuracy, and vice versa), but although the regression was performed for all five grading machines, only the regression calculations for the three chosen machines are presented below.

Calculations in MATLAB were performed using the transformation explained in chapter 3.1.1, but the figures presented below are of the data transformed back:

$$IP = \exp(x)$$
$$MP = \exp(y)$$

In our context, the material property is always the strength of the material, as mentioned earlier.

5.1 LINEAR REGRESSION

For each grading machine, the individual data points, (IP, MP), for the timber specimens are plotted in the same figure as the line of the estimate of the mean, $\mathbf{X}\hat{\boldsymbol{\beta}}$. This line is not completely straight in the figures presented, as the linear regression was performed on the transformed variables, x and y, while the figures are presented with regards to IP and MP. In addition, the cumulative distribution of both indicating property and material property is presented together with a plot of the lognormal distribution (this to see whether a lognormal distribution seems accurate). Naturally, the distribution of the material property is the same for all grading machines, as it is the same sample of timber specimens used.

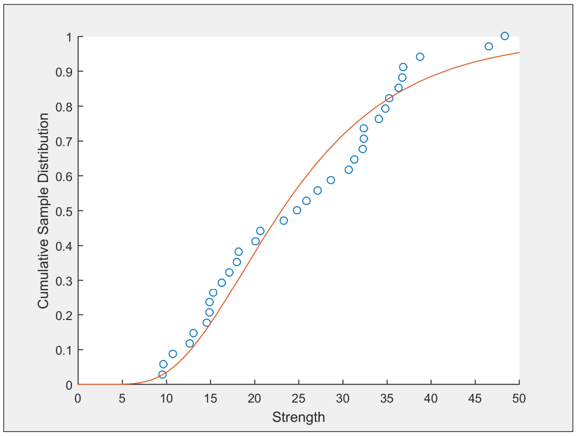


Figure 5.1: Cumulative distribution of the material property, strength

We observe from Figure 5.1 that the assumption of a lognormal distribution of the material property does not seem incorrect.

5.1.1 Grading machine #1

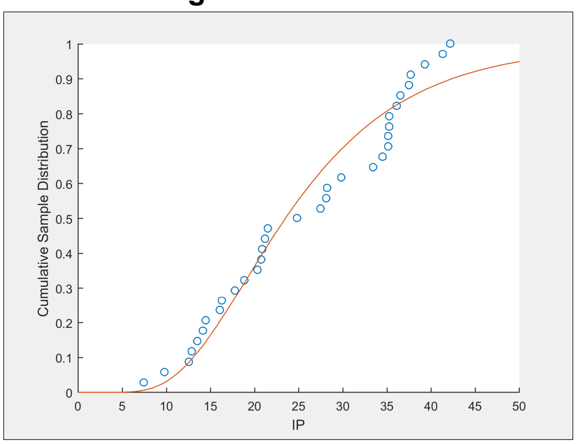


Figure 5.2: Cumulative distribution of the indicating property from Machine #1

Again, for the indicating property this time, it can be observed from Figure 5.2 that the assumption of a lognormal distribution does not seem to be incorrect.

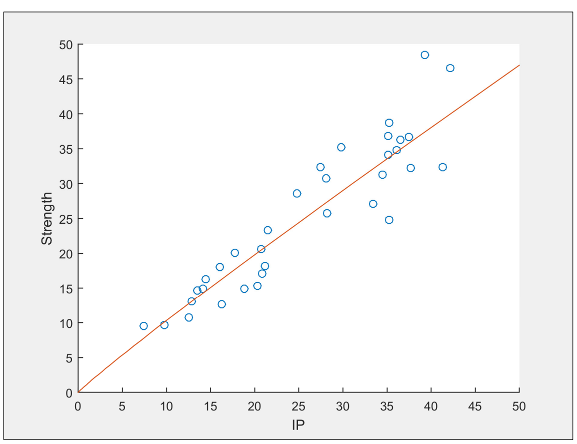


Figure 5.3: Data from Machine #1

We can see from Figure 5.3 that the values registered of the indicating property is relatively well concentrated around the estimated mean – the accuracy is relatively good.

5.1.2 Grading machine #2

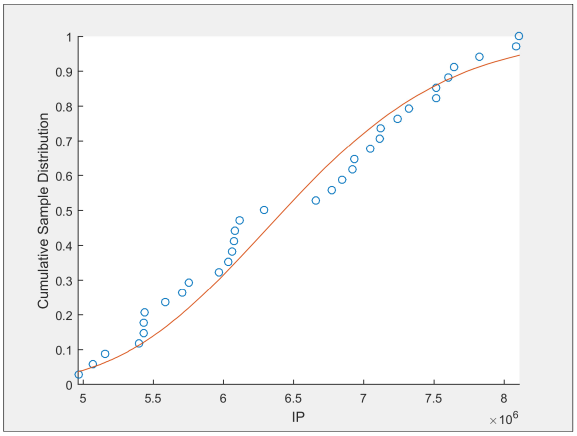


Figure 5.4: Cumulative distribution of the indicating property from Machine #2

A lognormal distribution, also in this case, appears to be a decent assumption.

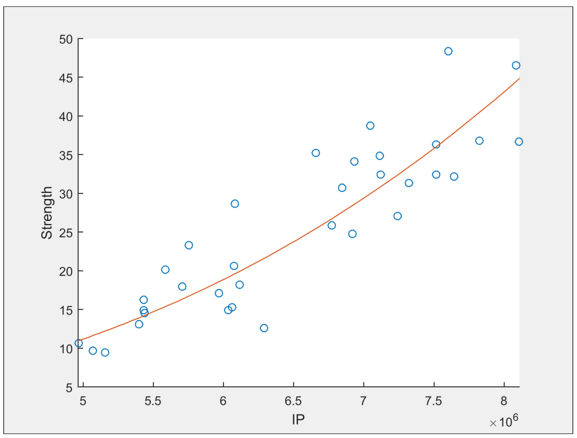


Figure 5.5: Data from Machine #2

By comparing Figure 5.5 to Figure 5.3, it can be observed that the errors are larger for Machine #2 than for Machine #1. The registered values of the indicating property is spread further away from the estimated mean.

5.1.3 Grading machine #3

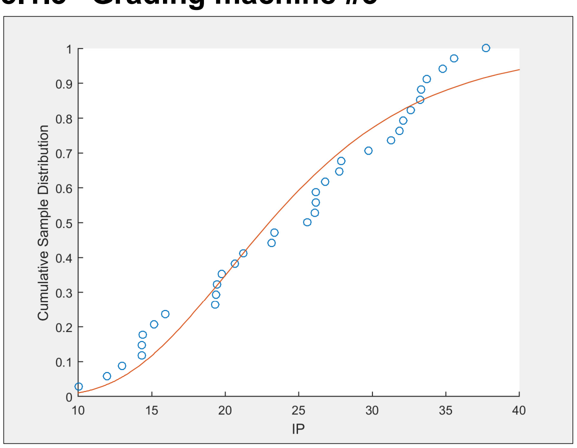


Figure 5.6: Cumulative distribution of the indicating property from Machine #3

As expected, the registered values of the indicating property from Machine #3 also follows a distribution that resembles a lognormal one.

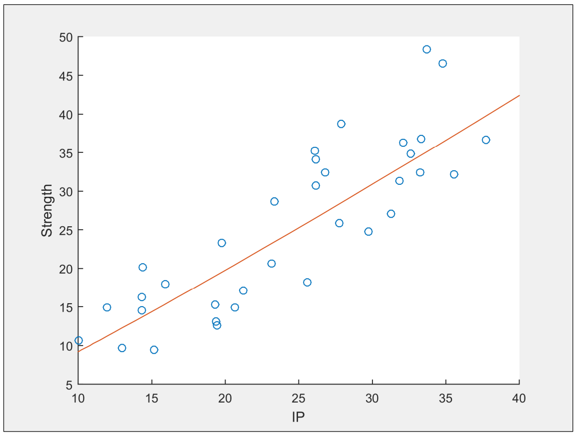


Figure 5.7: Data from Machine #3

Lastly, the data from Machine #3 can be seen to have an even larger spread than that of the two other grading machines. In other words, this grading machine is less accurate than Machine #1 and Machine #2. This is, naturally, as expected, since the machines were chosen this way, from most to least accurate: Machine #1, Machine #2 and Machine #3.

5.1.4 Comparison between machines

In Table 5.1 below, we show how the estimate of the variance of the errors, as calculated in (3.7), varies between the three different grading machines.

	Machine #1	Machine #2	Machine #3
s^2	0.0248	0.0415	0.0576
s	0.1574	0.2038	0.2400

Table 5.1: Estimates of the variance and standard deviation of the errors of our regression model

These variances may seem low, but they are calculated using the natural logarithm of our indicating and material property, therefore they are of the order of magnitude as they are. It is possible to transform the standard deviation back to the same unit of measure as the material property, but what is more useful is to observe the relative difference between them: we can see that Machine #1 is easily the most accurate and Machine #3 is the least accurate. This is the same conclusion as we made from observing the plots above.

5.2 TIMBER GRADING

From the data in Appendix C we want to look at how different ranges of indicating property affect the distribution of the material property for a given grade. We chose to look at four different ranges of indicating properties.

5.2.1 Ranges of IP

Before the distributions of the material property can be found, the settings for the different machines must be chosen. This means that different ranges of the indicating property must be defined. Our first range of indicating property is given by setting the upper limit to $+\infty$. Our last "range" of indicating property is given by setting $x_L = x_H$. This is not really a range, but just a single value of x, and is not something that is useful in actual grading, but we choose to include these two extreme ranges in order to show the limits in either direction: the largest range possible and the smallest "range" possible. In this last case we will not use (3.17), but (3.11) to calculate the distribution of the material property. Our second and third range of the indicating property are defined in the following way:

From our first range of indicating property (the largest range possible) we have:

$$x_L = x_{L1}$$

From our last "range" of indicating property (just a single value), we have:

$$x_L = x_{L4}$$

We name our lower limits for our remaining two ranges similarly, x_{L2} and x_{L3}.

We choose to set the distance between the lower limits like this:

- $x_{L2} - x_{L1} = a$
- $x_{L3} - x_{L2} = 2a$
- $x_{L4} - x_{L3} = 3a$

We can see that we have three equations and three unknowns, and this can easily be solved:

- $x_{L2} = \dfrac{x_{L4} + 5x_{L1}}{6}$
- $x_{L2} = \dfrac{x_{L4} + x_{L1}}{2}$

The reason we choose to distribute the lower limits in this manner, and not just distribute them evenly, is that the upper limit changes more for a given change of the lower limit, when the range is large, compared to when the range is smaller. Exactly how the ranges are defined is not important in itself; it is just presented here to show how they are distributed in the coming figures, and for reproducibility, should someone want to validate the figures and calculations.

5.2.1.1 Grade C24

We will now present some figures, showing the distribution of the material property for different ranges of the indicating property, using the most accurate grading machine from our data (we will show the effect of different grading machines later).

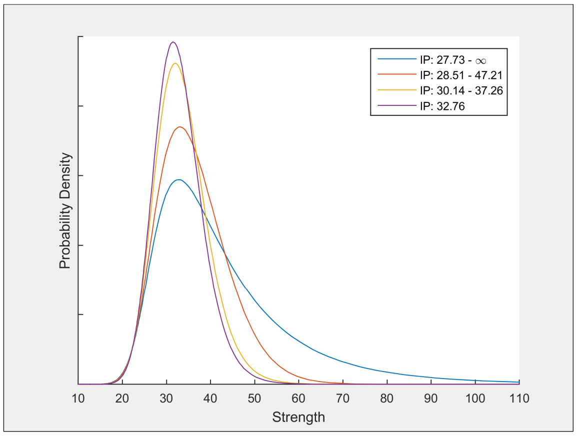


Figure 5.8: Distribution of the material property for different ranges of IP for grade C24

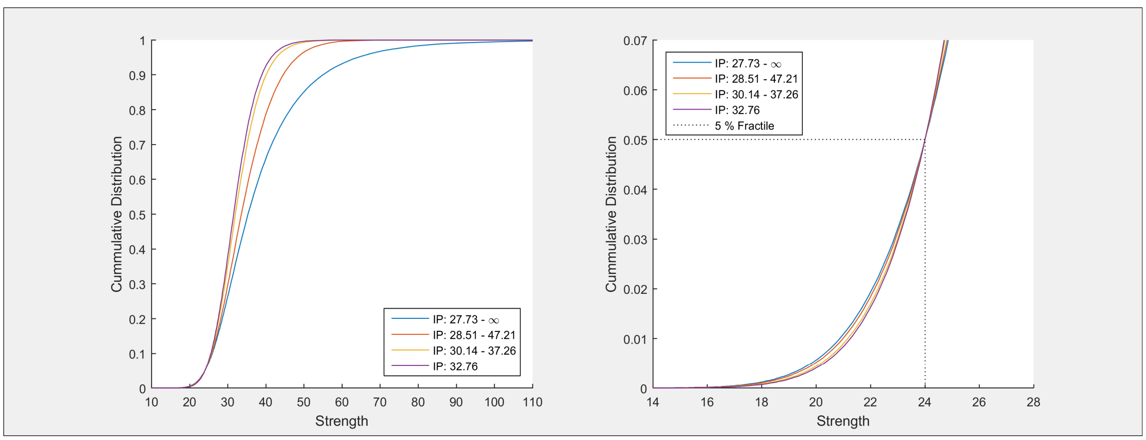


Figure 5.9: Cumulative distribution of the material property for different ranges of IP for grade C24

It can be observed that the right-hand tails of the different distributions vary quite significantly, but the left-hand tails have a relatively small variation. In Figure 5.9, it can be seen a zoomed-in view of the area around the 5 % fractile, and although the distributions follow a slightly different curve, they are still quite close to each other.

5.2.1.2 Grade C30

We now look at how the distributions are for timber grade C30:

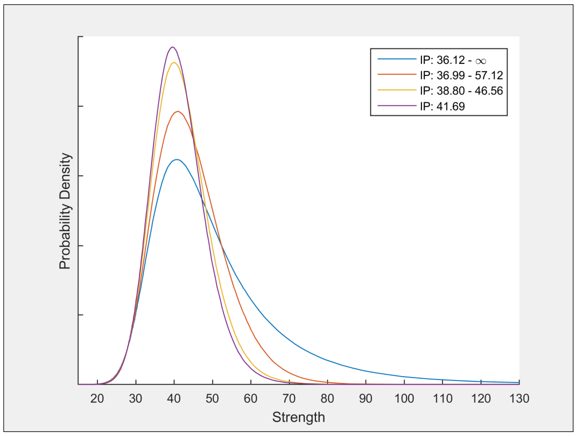


Figure 5.10: Distribution of MP for different ranges of IP for grade C30

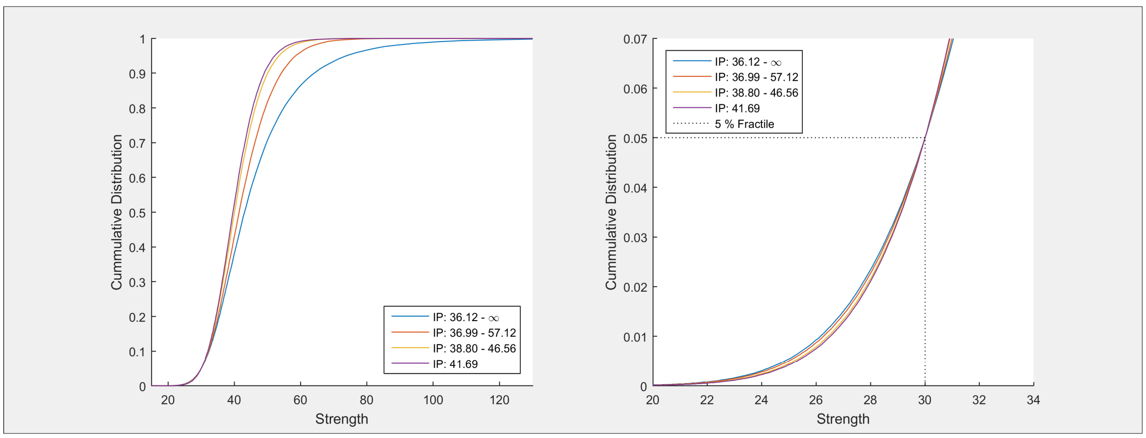


Figure 5.11: Cumulative distribution of MP for different ranges of IP for grade C30

We see that we have a similar situation for C30 as we do for C24.

5.2.1.3 Grade C40

Lastly, we also present the same figures for grade C40.

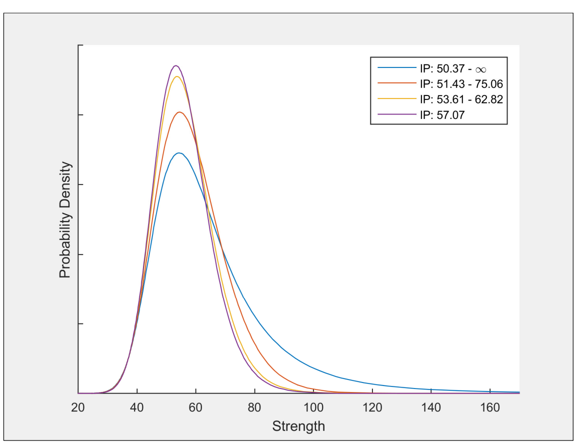


Figure 5.12: Distribution of MP for different ranges of IP for grade C40

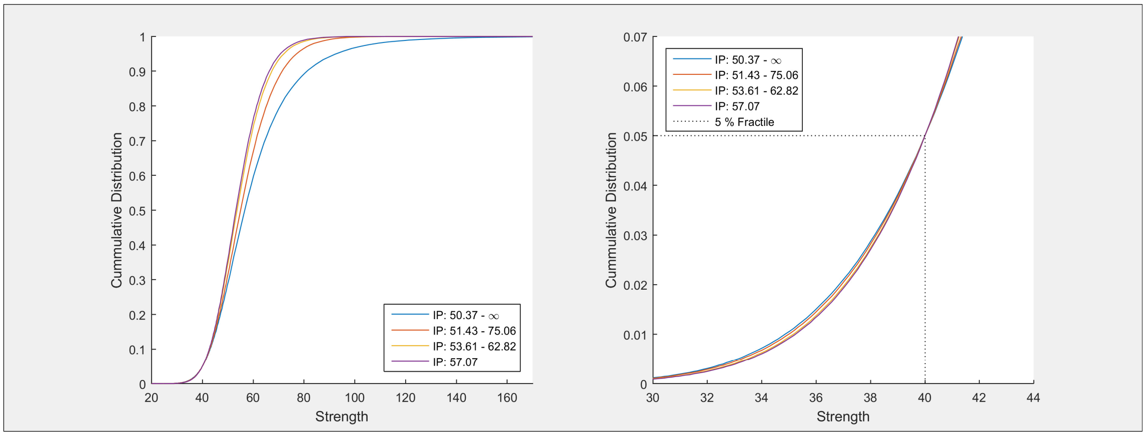


Figure 5.13: Cumulative distribution of MP for different ranges of IP for grade C40

Again, we observe that we have a similar situation, but the difference in the distribution of MP for the different ranges of IP is less than for the lower grades. This makes sense; if we look at Figure 3.3, we can see that as we are using values of IP on the far right side of the distribution, the weighting has a smaller impact – the curve is closer to horizontal, and as such the distribution of MP is closer to the t-distribution we found from our regression, in (3.11). Moving the upper limit towards infinity still results a relatively horizontal weighting. Changing this upper limit has a greater impact on the shape of the weighting function for the lower grades.

5.2.2 Different grading machines

In addition to the effect of the choice of range for the indicating property, we also want to look at the effect of the quality of the grading machine. We have three different grading machines, ranging from the most accurate, Machine #1, to the least accurate, Machine #3.

5.2.2.1 Grade C24

We choose to plot two different cases for the range of the indicating property: when the upper limit goes to infinity, and when we have a single value of IP, this way we can see the two extremes.

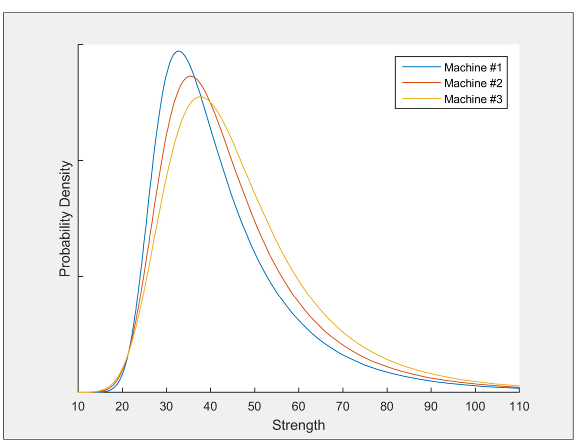


Figure 5.14: Distribution of MP for different grading machines for grade C24. $IP_H = +\infty$

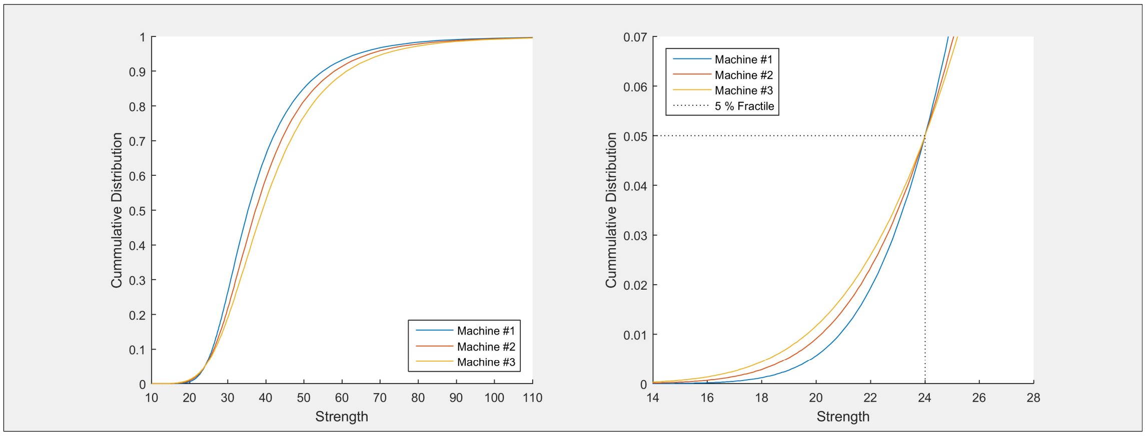


Figure 5.15: Cumulative distribution of MP for different grading machines for grade C24. $IP_H = +\infty$

We observe that the left-hand tail of the distributions have a larger difference than when comparing different settings for a single grading machine. This observation is of interest when considering the reliability, which we will come back to later.

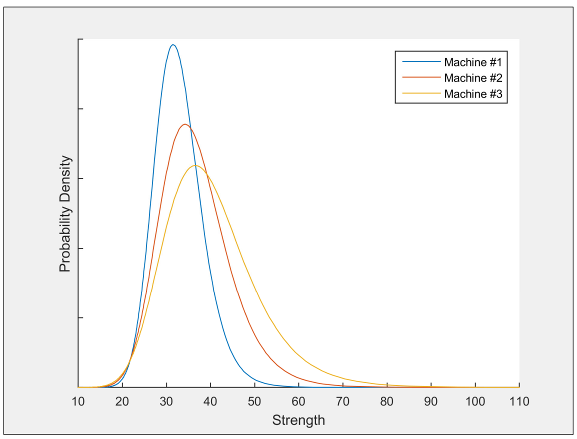


Figure 5.16: Distribution of MP for different grading machines for grade C24. $IP_L = IP_H$

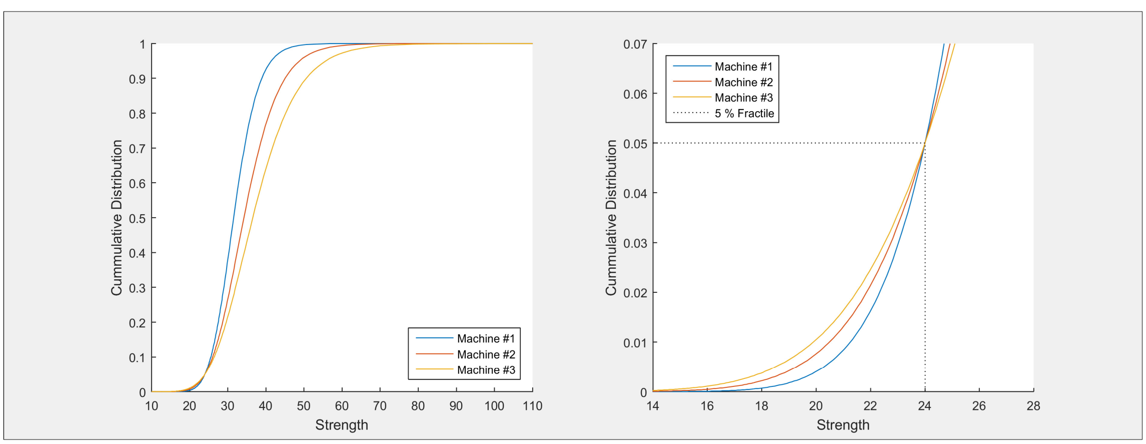


Figure 5.17: Cumulative distribution of MP for different grading machines for grade C24. $IP_L = IP_H$

Also when looking at the other extreme case, where the distribution follows a t-distribution, it can be observed that the left-hand tail has a larger variation than when comparing machine settings.

5.2.2.2 Grade C30

We perform the same calculations while considering timber grade C30.

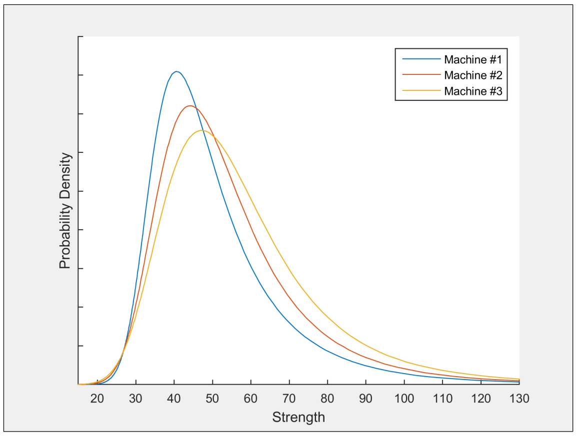


Figure 5.18: Distribution of MP for different grading machines for grade C30. $IP_H = +\infty$

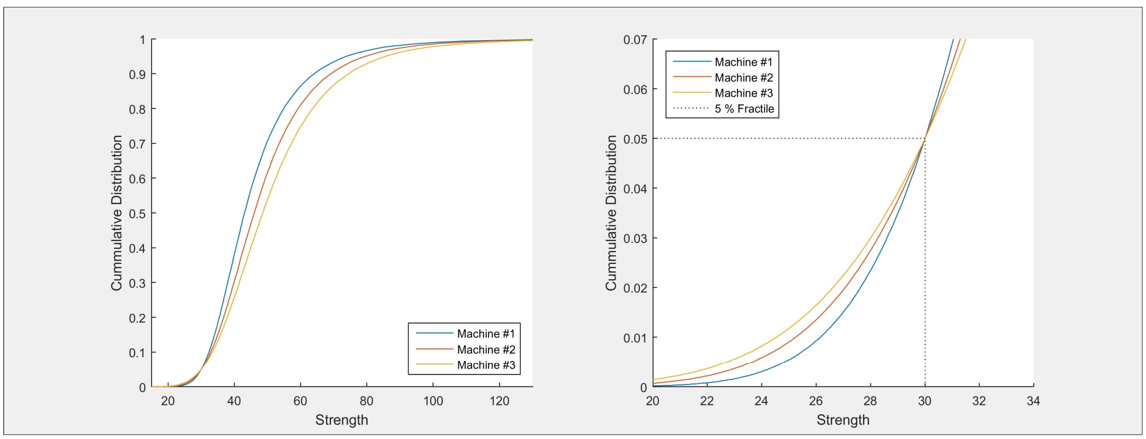


Figure 5.19: Cumulative distribution of MP for different grading machines for grade C30. $IP_H = +\infty$

The variation of the left-hand tail seems even larger for this grade, compared to C24.

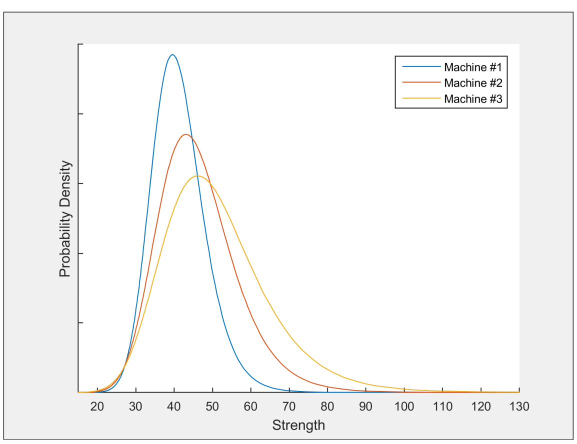


Figure 5.20: Distribution of MP for different grading machines for grade C30. $IP_L = IP_H$

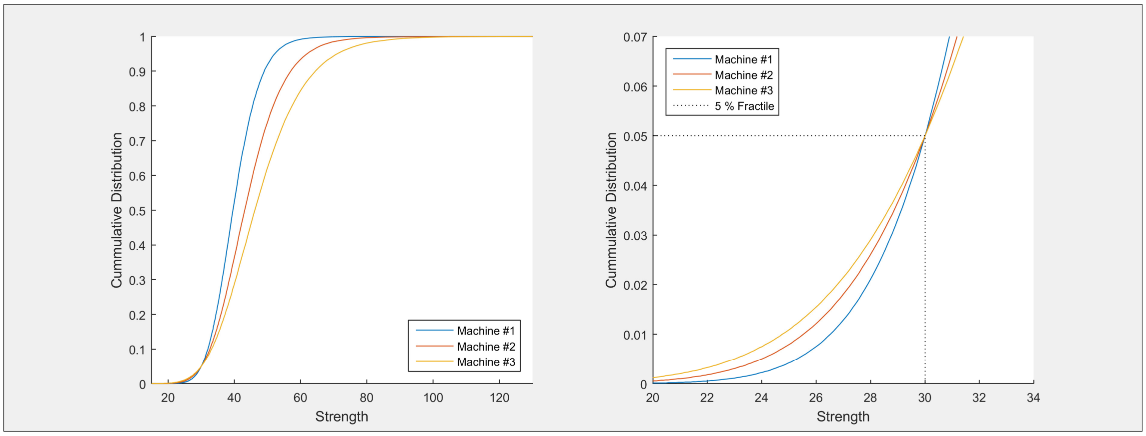


Figure 5.21: Cumulative distribution of MP for different grading machines for grade C30. $IP_L = IP_H$

Again, the left-hand tail has a similar situation for the extreme case with a single value of IP.

5.2.2.3 Grade C40

Lastly, we have a look at grade C40.

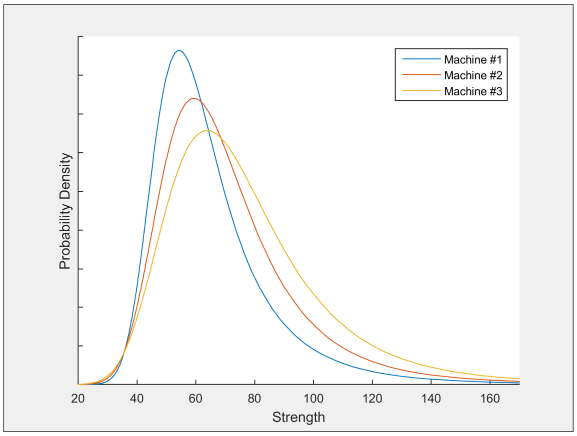


Figure 5.22: Distribution of MP for different grading machines for grade C40. $IP_H = +\infty$

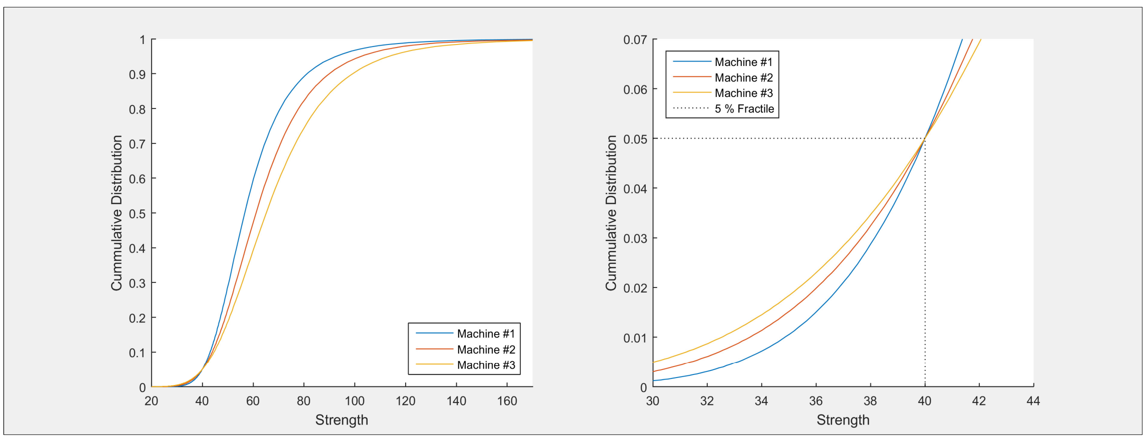


Figure 5.23: Cumulative distribution of MP for different grading machines for grade C40. $IP_H = +\infty$

The tendency we saw when comparing C24 and C30 seems to be confirmed here. It can be observed that the variation of the left-hand tail appears to be larger for higher grades.

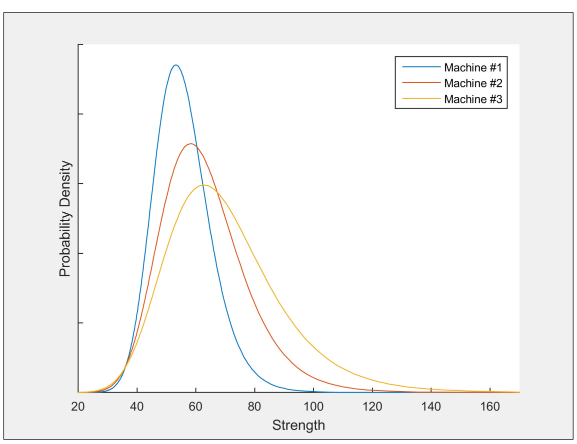


Figure 5.24: Distribution of MP for different grading machines for grade C40. $IP_L = IP_H$

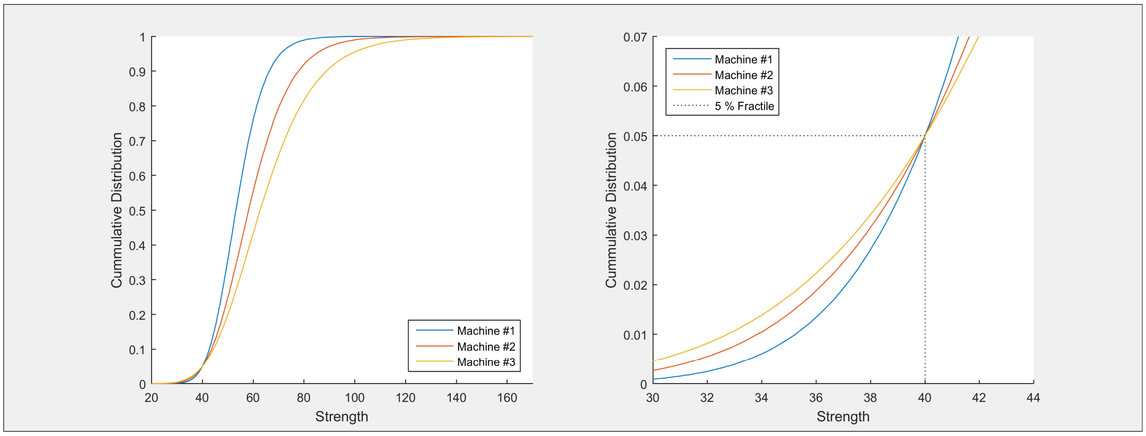


Figure 5.25: Cumulative distribution of MP for different grading machines for grade C40. $IP_L = IP_H$

As expected, the situation is the same, also for the other extreme machine setting.

5.3 Reliability

We have now come to the main topic of interest, the reliability of structures, and more specifically, the influence of the grading on this reliability.

5.3.1 Probability of failure for given machine settings

Firstly, we will present calculated values of the probability of failure, using (4.15), for the four different machine settings, as described in chapter 5.2.1. We will present the probabilities of failure for all three grading machines, and for three timber grades, C24, C30 and C40. In the next chapter, we will look at the probability of failure for any machine setting.

5.3.1.1 Grade C24

We start by looking at timber grade C24. The following table presents the calculated probabilities of failure, and in addition a relative comparison between the most and least accurate grading machine, as well as the two extreme cases of machine settings (upper limit to infinity (Setting #1), and upper limit equal lower limit (Setting#4)).

	Machine #1	Machine #2	Machine #3	#3/#1
Setting #1	$0.1869 \cdot 10^{-5}$	$0.5294 \cdot 10^{-5}$	$1.1900 \cdot 10^{-5}$	6.37
Setting #2	$0.1691 \cdot 10^{-5}$	$0.4846 \cdot 10^{-5}$	$1.1102 \cdot 10^{-5}$	6.57
Setting #3	$0.1445 \cdot 10^{-5}$	$0.4244 \cdot 10^{-5}$	$1.0068 \cdot 10^{-5}$	6.97
Setting #4	$0.1331 \cdot 10^{-5}$	$0.3976 \cdot 10^{-5}$	$0.9618 \cdot 10^{-5}$	7.23
#4/#1	0.712	0.751	0.808	

Table 5.2: Probability of failure, grade C24

From the data in Table 5.2, it can be observed that for timber grade C24, the least accurate grading machine, Machine #3, has a probability of failure that is around seven times the probability of failure for the most accurate machine, Machine #1 (between 6.37 times and 7.23 times). In other words, the probability of failure is around 600 % higher for Machine #3, than for Machine #1. It can also be observed that the probability of failure for the theoretical smallest range of indicating property (where the upper limit equals the lower limit) is around 75 % of the probability of failure with the largest range of IP (when the upper limit goes to infinity). In other words, the probability of failure can be reduced by up to approximately 25 %, by decreasing the range of IP (but a too small range of IP is not practical, and Setting #4 is never possible).

5.3.1.2 Grade C30

We look at a similar same table, but this time for timber grade C30.

	Machine #1	Machine #2	Machine #3	#3/#1
Setting #1	$0.1805 \cdot 10^{-5}$	$0.5370 \cdot 10^{-5}$	$1.2704 \cdot 10^{-5}$	7.04
Setting #2	$0.1669 \cdot 10^{-5}$	$0.5022 \cdot 10^{-5}$	$1.2079 \cdot 10^{-5}$	7.14
Setting #3	$0.1483 \cdot 10^{-5}$	$0.4567 \cdot 10^{-5}$	$1.1300 \cdot 10^{-5}$	7.62
Setting #4	$0.1400 \cdot 10^{-5}$	$0.4370 \cdot 10^{-5}$	$1.0973 \cdot 10^{-5}$	7.84
#4/#1	0.776	0.814	0.864	

Table 5.3: Probability of failure, grade C30

The data in Table 5.3 shows a similar tendency for timber grade C30, compared to the data for grade C24, as presented in Table 5.2. It can be seen, however, that the influence of the quality of the grading machine is larger – an increase in the probability of failure of around 650 % between the best and worst machine. At the same time, the influence of the machine settings is smaller – approximately, up to a 20 % reduction can theoretically be achieved by reducing the range of IP.

5.3.1.3 Grade C40

Lastly, we look at grade C40.

	Machine #1	Machine #2	Machine #3	#3/#1
Setting #1	$0.1880 \cdot 10^{-5}$	$0.6061 \cdot 10^{-5}$	$1.5377 \cdot 10^{-5}$	8.18
Setting #2	$0.1769 \cdot 10^{-5}$	$0.5767 \cdot 10^{-5}$	$1.4845 \cdot 10^{-5}$	8.39
Setting #3	$0.1620 \cdot 10^{-5}$	$0.5398 \cdot 10^{-5}$	$1.4213 \cdot 10^{-5}$	8.77
Setting #4	$0.1555 \cdot 10^{-5}$	$0.5244 \cdot 10^{-5}$	$1.3956 \cdot 10^{-5}$	8.97
#4/#1	0.827	0.865	0.908	

Table 5.4: Probability of failure, grade C40

The data for timber grade C40, presented in Table 5.4, shows again the same tendency – an increase in the influence of the accuracy of the grading machine, this time around a 750 % increase in the probability of failure when going from Machine #1 to Machine #3. The influence of the machine settings is even lower – a theoretical reduction of approximately 15 % by reducing the range of the indicating property.

5.3.2 Probability of failure for any machine setting

In order to have a graphical representation of the probability of failure, the following figures show the probability of failure as a function of the machine settings, for each of the three grading machines. In order to have the same scale for the machine setting for the different grading machines, the lower limit of the range of indicating properties was used. The minimum value of the lower limit is the value given when the upper limit goes to infinity, this referred to as "Max Range of IP" in the figures. The maximum value of the lower limit is the value when the upper limit equals the lower limit, in the figures this is referred to as "Single Value of IP". In addition to this, because the different grading machines have indicating properties of varying scale, the values were normalized to go from zero to one in the following way:

$$\frac{IP_L - IP_{L-\min}}{IP_{L-\max} - IP_{L-\min}}$$

It can easily be seen that when $IP_L = IP_{L-\min}$, the expression equals zero, and when $IP_L = IP_{L-\max}$, the expression is one.

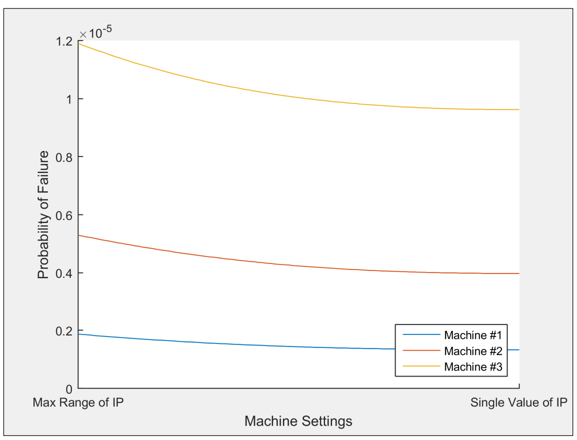


Figure 5.26: Probability of failure, C24

The graphs in the figure shows the same extreme values (for machine setting #1 and #4) as the ones given in Table 5.2, as expected. It is also easy to see that the accuracy of the grading machine has a much larger impact than that of the machine settings.

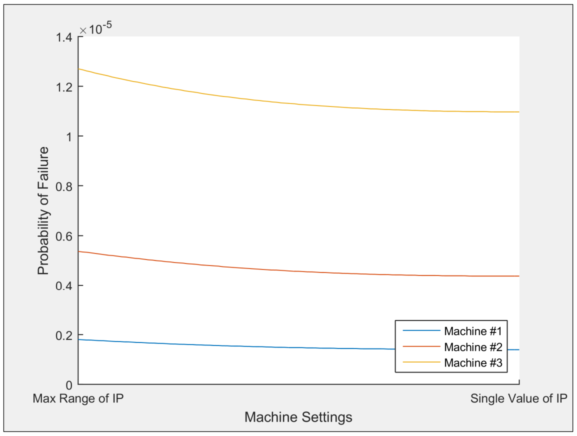


Figure 5.27: Probability of failure, C30

For timber grade C30, the figure shows the same as the values in the tables; the influence from the machine settings are somewhat smaller (curves are a bit "flatter", but also be aware that the vertical scale is slightly different), and the gap between the different machines is larger.

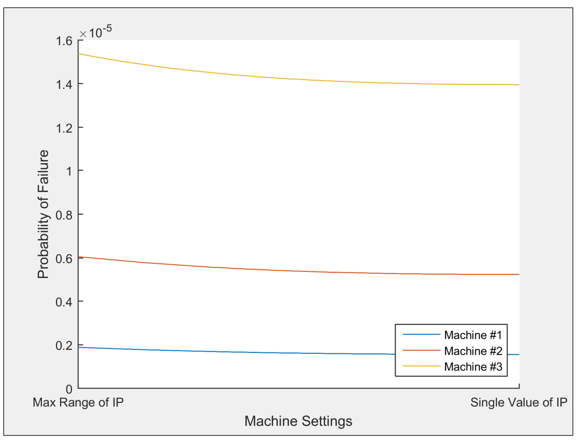


Figure 5.28: Probability of failure, C40

The last figure, for grade C40, again shows the same thing, a larger gap between the most and least accurate machine, and less influence from the machine settings.

6 CONCLUSIONS

From the calculations and figures presented in chapter 5, it can be seen that a change of machine settings, for any of the machines considered, can result in a decrease of the probability of failure of a structure by up to around 25 %. This result is based on the timber data given, and a different distribution of the indicating property might give a different result. In addition, for lower grades than C24 it is possible that the machine settings have a larger impact. Still, the result does indicate the order of magnitude that the grading machine settings have on the reliability of a structure where the graded timber is to be used.

In addition, the data presented in chapter 5 shows that the accuracy of the grading machine could have a significant impact on the reliability of a structure. It also shows that for higher grades, this impact seems to be even larger. Up to more than a 750 % increase in the probability of failure was observed from the analysis of the given data. The data given consists of five different grading machines (although we only looked at three of them, including the most and least accurate), and there might exist other machines with different accuracies, which could give, potentially, even larger differences (the range of grading machines available and/or in use is not a topic considered in this thesis).

To sum up, the most significant factor, of the factors considered, with regards to the influence of the grading on the reliability of a timber structure, appears, quite strongly, to be the accuracy of the grading machine used. The machine settings, in comparison, seem to have a relatively low influence on a structure's probability of failure, but it does still have some influence.

The methods and expressions presented in chapters 3 and 4 can be used to perform a similar analysis on other sets of data. This can be useful, for example, for analyzing samples from other wood species or if one has more accurate knowledge of the distribution of the indicating property for a given grading machine.

REFERENCES

Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., & Rubin, D. B. (2014). *Bayesian Data Analysis, Third Edition.*

Schneider, J. (1997). *Introduction to Safety and Reliability of Structures.*

Appendix A COMMON PROBABILITY DISTRIBUTIONS

In this section, we present common probability distributions used in this thesis. The notation is similar to that used in (Gelman, et al., 2014). In general, θ represents the variable that follows the given distribution.

A.1 UNIVARIATE NORMAL

A.1.1 Notation

$\theta \sim \mathrm{N}\left(\mu, \sigma^2\right)$

Location: μ
Scale: $\sigma > 0$

A.1.2 Density function

$$p(\theta) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{1}{2\sigma^2}(\theta-\mu)^2\right)$$

A.2 MULTIVARIATE NORMAL

A.2.1 Notation

$\boldsymbol{\theta} \sim \mathrm{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$

Location: $\boldsymbol{\mu} = \left(\mu_1, \cdots, \mu_d\right)$
Symmetric, positive definite, $d \times d$ variance matrix: $\boldsymbol{\Sigma}$

A.2.2 Density function

$$p(\boldsymbol{\theta}) = (2\pi)^{-d/2} |\boldsymbol{\Sigma}|^{-1/2} \times \exp\left(-\frac{1}{2}(\boldsymbol{\theta}-\boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\boldsymbol{\theta}-\boldsymbol{\mu})\right)$$

A.3 SCALED INVERSE-CHI-SQUARE

A.3.1 Notation

$\theta \sim \text{Inv-}\chi^2\left(\nu, s^2\right)$

Degrees of freedom: $\nu > 0$
Scale: $s > 0$

A.3.2 Density function

$$p(\theta) = \frac{\left(\nu/2\right)^{\nu/2}}{\Gamma\left(\nu/2\right)} s^{\nu} \theta^{-\left(\nu/2+1\right)} e^{-\nu s^2/2\theta}$$

A.4 UNIVARIATE T

A.4.1 Notation

$\theta \sim \mathrm{t}_{\nu}\left(\mu, \sigma^2\right)$

Degrees of freedom: $\nu > 0$
Location: μ
Scale: $\sigma > 0$

A.4.2 Density function

$$p(\theta) = \frac{\Gamma\left((\nu+1)/2\right)}{\Gamma\left(\nu/2\right)\sqrt{\nu\pi}\sigma}\left(1 + \frac{1}{\nu}\left(\frac{\theta-\mu}{\sigma}\right)^2\right)^{-(\nu+1)/2}$$

A.5 MULTIVARIATE T

A.5.1 Notation

$\boldsymbol{\theta} \sim \mathrm{t}_{\nu}\left(\boldsymbol{\mu}, \boldsymbol{\Sigma}\right)$

Degrees of freedom: $\nu > 0$
Location: $\boldsymbol{\mu} = \left(\mu_1, \cdots, \mu_d\right)$
Symmetric, positive definite, $d \times d$ scale matrix: $\boldsymbol{\Sigma}$

A.5.2 Density function

$$p(\boldsymbol{\theta}) = \frac{\Gamma\left((\nu+d)/2\right)}{\Gamma\left(\nu/2\right)(\nu\pi)^{d/2}}\left|\boldsymbol{\Sigma}\right|^{-1/2} \times \left(1 + \frac{1}{\nu}(\boldsymbol{\theta}-\boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta}-\boldsymbol{\mu})\right)^{-(\nu+d)/2}$$

A.6 GUMBEL MAX

A.6.1 Notation

$\theta \sim \mathrm{Gumbel}\left(\mu, \sigma^2\right)$

Mean: μ
Standard deviation: $\sigma > 0$

A.6.2 Density function

$$p(\theta) = \frac{\pi}{\sigma\sqrt{6}}\exp\left(-\alpha\left(\theta - \left(\mu - \frac{\gamma\sigma\sqrt{6}}{\pi}\right)\right) - \exp\left(-\frac{\pi}{\sigma\sqrt{6}}\left(\theta - \left(\mu - \frac{\gamma\sigma\sqrt{6}}{\pi}\right)\right)\right)\right)$$

Where γ is the Euler–Mascheroni constant.

Appendix B MATLAB SCRIPTS AND FUNCTIONS

In this section, we present the MATLAB scripts and functions used to perform calculations and create figures presented in this thesis. These scripts and functions are not meant to show how to perform numerical calculations optimally. There are probably much more optimal ways of doing it. They are merely presented here for reproducibility, and they could perhaps help illustrate a possible approach to how to solve the expressions presented in this paper. The version of MATLAB used is R2015a (8.5.0.197613).

B.1 SCRIPTS

B.1.1 data.m

The data.m script only imports the data from the Excel file in where the original data was stored, and saves this data to *.mat files:

```
%% Import Data
%
% This script imports the material and indicating properties from our
% Excel-file and saves this data to MAT-files.
%%
% We clear the workspace and import the data from the Excel file:

clear
rawdata = xlsread('data.xlsx','AL2:AW71');

%% Region 1 - Västergötland
%
% In this region there is one line that is excluded (due to "rupture in
% jaws"), and we need to remove this row when putting the data into our
% variables (because some cells are empty on this row - row 25)

str = [rawdata(1:24,1);rawdata(26:35,1)];
MOE = [rawdata(1:24,2);rawdata(26:35,2)];
dens = [rawdata(1:24,3);rawdata(26:35,3)];
IP_MOR_GE706 = [rawdata(1:24,4);rawdata(26:35,4)];
IP_MOE_GE706 = [rawdata(1:24,5);rawdata(26:35,5)];
IP_dens_GE706 = [rawdata(1:24,6);rawdata(26:35,6)];
IP_MOR_combiscan = [rawdata(1:24,7);rawdata(26:35,7)];
IP_MOR_escan = [rawdata(1:24,8);rawdata(26:35,8)];
IP_dens_escan = [rawdata(1:24,9);rawdata(26:35,9)];
IP_MOR_triomatic = [rawdata(1:24,10);rawdata(26:35,10)];
IP_dens_triomatic = [rawdata(1:24,11);rawdata(26:35,11)];
IP_Rosegrade = [rawdata(1:24,12);rawdata(26:35,12)];

% We save all the variables except rawdata to a file and then clear them
save data_1.mat str MOE dens IP_MOR_GE706 IP_MOE_GE706 IP_dens_GE706 ...
    IP_MOR_combiscan IP_MOR_escan IP_dens_escan IP_MOR_triomatic ...
    IP_dens_triomatic IP_Rosegrade
clearvars -except rawdata

%% Region 2 - Lappland
%
% From this region we import all rows, as there are no exclusions

str = rawdata(36:70,1);
```

```
MOE = rawdata(36:70,2);
dens = rawdata(36:70,3);
IP_MOR_GE706 = rawdata(36:70,4);
IP_MOE_GE706 = rawdata(36:70,5);
IP_dens_GE706 = rawdata(36:70,6);
IP_MOR_combiscan = rawdata(36:70,7);
IP_MOR_escan = rawdata(36:70,8);
IP_dens_escan = rawdata(36:70,9);
IP_MOR_triomatic = rawdata(36:70,10);
IP_dens_triomatic = rawdata(36:70,11);
IP_Rosegrade = rawdata(36:70,12);

% We save all the variables except rawdata to a file and then clear all
save data_2.mat str MOE dens IP_MOR_GE706 IP_MOE_GE706 IP_dens_GE706 ...
    IP_MOR_combiscan IP_MOR_escan IP_dens_escan IP_MOR_triomatic ...
    IP_dens_triomatic IP_Rosegrade
clear
```

B.1.2 regression.m

The regression.m script runs the regression on the imported data, and creates figures of the data and regression.

```
%% Linear Regression - Strength
%
% We look at the sample data given from a specific region and analyze the
% strength properties from this sample.

%% Importing data
%
% We start by importing the data which we will analyze by running a script
% which will import data from Excel to MAT-files (first checking if the
% MAT-files have already been created)

if exist('data1.mat','file') && exist('data2.mat','file')
    vars1 = whos('-file','data1.mat');
    vars2 = whos('-file','data2.mat');
    if all(ismember({'str' 'MOE' 'dens' 'IP_MOR_GE706' 'IP_MOE_GE706' ...
            'IP_dens_GE706' 'IP_MOR_combiscan' 'IP_MOR_escan' ...
            'IP_dens_escan' 'IP_MOR_triomatic' 'IP_dens_triomatic' ...
            'IP_Rosegrade'}, {vars1.name})) && ...
            all(ismember({'str' 'MOE' 'dens' 'IP_MOR_GE706' ...
            'IP_MOE_GE706' 'IP_dens_GE706' 'IP_MOR_combiscan' ...
            'IP_MOR_escan' 'IP_dens_escan' 'IP_MOR_triomatic' ...
            'IP_dens_triomatic' 'IP_Rosegrade'}, {vars2.name}))
    else
        data
    end
else
    data
end

clear

% We load the specific data we want below (in each loop)

%% Loops
%
% We will run the regression for our 2 different regions, and our 5
% different grading methods, and create the following for-loops:

% First for each of the two regions:
% 1 = Västergötland
% 2 = Lappland
for Region = 1:2

    % And next for each of the 5 grading machines:
    % 1 = GoldenEye 706 machine
    % 2 = Combiscan machine
    % 3 = E-scan machine
    % 4 = Triomatic machine
    % 5 = Rosegrade
    for Machine = 1:5

        %% Loading data
        %
        % We load the data we want to use:
```

```
datafile = ['data_' num2str(Region) '.mat'];
indicators = {'MOR_GE706','MOR_combiscan','MOR_escan', ...
    'MOR_triomatic','Rosegrade'};
indicator = indicators{Machine};
IPname = ['IP_' indicator];

load(datafile,'str',IPname)

% Now that we have our data, we put the data for the material
% property into the variable MP, and the data for the indicating
% property into IP:

MP = str;
IP = eval(IPname);
clear('str',IPname,'datafile','indicators','indicator','IPname')

n = size(MP,1); % Number of observations

%% Plot Data
%
% We plot the data in order to see the relation between the
% indicating property and the strength, and we also plot the
% cumulative sample distribution of IP:

% We define names of the figures:
machines = {'GoldenEye 706','Combiscan','E-scan', ...
    'Triomatic','Rosegrade'};
name1 = ['IP-Strength [Region ' num2str(Region) ' - ' ...
    machines{Machine} ']'];
name2 = ['Cumulative Sample Distribution - IP [Region ' ...
    num2str(Region) ' - ' machines{Machine} ']'];
name3 = ['Cumulative Sample Distribution - MP [Region ' ...
    num2str(Region) ']'];

% First we define our figures:
scrsz = get(groot,'ScreenSize');
fig(1) = figure(1);
fig(2) = figure(2);
fig(3) = figure(3);
set(fig(:),'Position',[(3*scrsz(3)-2*scrsz(4))/6 scrsz(4)/4 ...
    2*scrsz(4)/3 scrsz(4)/2],'Name',name1)
set(fig(2),'Name',name2)
set(fig(3),'Name',name3)

clear machines name1 name2 scrsz

% Next we define our plots:
figure(1) % Make active figure

scatter(IP,MP)
xlabel IP
ylabel Strength

figure(2) % Make active figure

scatter(sort(IP),1/n:1/n:1)
xlabel IP
ylabel('Cumulative Sample Distribution')
```

```
figure(3) % Make active figure

scatter(sort(MP),1/n:1/n:1)
xlabel Strength
ylabel('Cumulative Sample Distribution')

%% Distribution of observations
%
% We wish to look at how the observations of indicating properties
% and material properties are distributed - We assume a lognormal
% distribution, and transform our data before performing our
% regression

x = log(IP);
y = log(MP);

E_IP = mean(IP); % Sample mean of indicating property
E_x = mean(x);
var_IP = var(IP); % Sample variance of indicating property
var_x = var(x);
std_IP = std(IP); % Sample std deviation of indicating property
std_x = std(x);

E_MP = mean(MP); % Sample mean of material property
E_y = mean(y);
var_MP = var(MP); % Sample variance of material property
var_y = var(y);
std_MP = std(MP); % Sample std deviation of material property
std_y = std(y);

% Assuming a lognormal distribution of each population of the
% different properties, we choose to draw graphs using the
% estimates for mean and variance (to see if a lognormal
% distribution seems accurate)

min_IP = floor(min(IP)/10)*10; % Round lower limit down
max_IP = ceil(max(IP)/10)*10; % Round upper limit up

figure(2) % Make active figure

hold on
fplot(@(IP)normcdf(log(IP),E_x,std_x),[min_IP max_IP])
hold off
xlim([min_IP max_IP])

min_MP = floor(min(MP)/10)*10; % Round lower limit down
max_MP = ceil(max(MP)/10)*10; % Round upper limit up

figure(3) % Make active figure

hold on
fplot(@(MP)normcdf(log(MP),E_y,std_y),[min_MP max_MP])
hold off
xlim([min_MP max_MP])

clear min_MP max_MP % These values are no longer of interest

%% Finding the regression parameters
%
```

```
        % Next we calculate our regression parameters:

        % First we create our X-matrix:
        X = [ones(n,1) x];

        % Although, in our calculations, k is always equal to 2, we choose
        % to define it as follows:
        k = size(X,2);

        % We choose to define the transpose of X multiplied by X as a
        % separate variable:
        XTX = X.'*X;

        % We can now find an estimate for beta as follows:
        b = XTX\(X.'*y);

        % If we do not know the population variance, sigma, we can find an
        % estimate from the sample data:
        s2 = ((y-X*b).'*(y-X*b))/(n-k);

        %% Plot regression line
        %
        % We can now choose to plot our regression line into our figure:

        figure(1) % Make active figure

        % Since the regression line is linear before transforming back, we
        % choose to define the function after transforming back (to get a
        % smooth curve):
        PlotMP = @(IP)exp([1 log(IP)]*b);
        hold on
        fplot(PlotMP,[min_IP max_IP]);
        hold off
        xlim([min_IP max_IP])

        clear min_IP max_IP PlotMP

        % Lastly we save the figures and variables of intereset and clear
        % the others from the workspace
        figname = ['regression_region' num2str(Region) '_machine' ...
            num2str(Machine) '.fig'];
        varsname = ['regression_region' num2str(Region) '_machine' ...
            num2str(Machine) '.mat'];
        savefig(fig,figname)
        save(varsname,'b','E_IP','E_MP','E_x','E_y','IP','k','MP','n', ...
            's2','std_IP','std_MP','std_x','std_y','var_IP','var_MP', ...
            'var_x','var_y','x','X','XTX','y')
        clear figname varsname
    end
end
clear Region Machine fig
```

B.1.3 grading.m

The grading.m script is very slow to run, as it performs many numerical integrations. This script calculates ranges of the indicating property for the different machines, and creates plots for the distribution of the material property for the different machine settings, grading machines and timber grades.

```
%% Grading
%
% Given a linear regression of a set of indicating properties and
% corresponding material properties, we look at the effect of different
% ranges of indicating property and use of different grading machines on
% the distribution of the material property for different timber grades.
% We will use numerical integration to find our values (as opposed to
% simulation).
%
% THIS SCRIPT IS VERY SLOW - MAY TAKE OVER AN HOUR, DEPENDING ON THE
% TOLERANCE OF THE PLOTS!!!

%% Importing data
%
% We start by checking if the regression script has been run, and if not,
% run it.

if exist('regression_region1_machine1.mat','file') && ...
        exist('regression_region1_machine2.mat','file') && ...
        exist('regression_region1_machine3.mat','file') && ...
        exist('regression_region1_machine4.mat','file') && ...
        exist('regression_region2_machine1.mat','file') && ...
        exist('regression_region2_machine2.mat','file') && ...
        exist('regression_region2_machine3.mat','file') && ...
        exist('regression_region2_machine4.mat','file')
    vars1 = whos('-file','regression_region1_machine1.mat');
    vars2 = whos('-file','regression_region1_machine2.mat');
    vars3 = whos('-file','regression_region1_machine3.mat');
    vars4 = whos('-file','regression_region1_machine4.mat');
    vars5 = whos('-file','regression_region2_machine1.mat');
    vars6 = whos('-file','regression_region2_machine2.mat');
    vars7 = whos('-file','regression_region2_machine3.mat');
    vars8 = whos('-file','regression_region2_machine4.mat');
    if all(ismember({'MP','IP','n','x','y','E_IP','E_x','var_IP', ...
            'var_x','std_IP','std_x','E_MP','E_y','var_MP','var_y', ...
            'std_MP','std_y','X','k','XTX','b','s2'}, {vars1.name})) && ...
            all(ismember({'MP','IP','n','x','y','E_IP','E_x','var_IP', ...
            'var_x','std_IP','std_x','E_MP','E_y','var_MP','var_y', ...
            'std_MP','std_y','X','k','XTX','b','s2'}, {vars2.name})) && ...
            all(ismember({'MP','IP','n','x','y','E_IP','E_x','var_IP', ...
            'var_x','std_IP','std_x','E_MP','E_y','var_MP','var_y', ...
            'std_MP','std_y','X','k','XTX','b','s2'}, {vars3.name})) && ...
            all(ismember({'MP','IP','n','x','y','E_IP','E_x','var_IP', ...
            'var_x','std_IP','std_x','E_MP','E_y','var_MP','var_y', ...
            'std_MP','std_y','X','k','XTX','b','s2'}, {vars4.name})) && ...
            all(ismember({'MP','IP','n','x','y','E_IP','E_x','var_IP', ...
            'var_x','std_IP','std_x','E_MP','E_y','var_MP','var_y', ...
            'std_MP','std_y','X','k','XTX','b','s2'}, {vars5.name})) && ...
            all(ismember({'MP','IP','n','x','y','E_IP','E_x','var_IP', ...
            'var_x','std_IP','std_x','E_MP','E_y','var_MP','var_y', ...
            'std_MP','std_y','X','k','XTX','b','s2'}, {vars6.name})) && ...
            all(ismember({'MP','IP','n','x','y','E_IP','E_x','var_IP', ...
            'var_x','std_IP','std_x','E_MP','E_y','var_MP','var_y', ...
            'std_MP','std_y','X','k','XTX','b','s2'}, {vars7.name})) && ...
```

```
            all(ismember({'MP','IP','n','x','y','E_IP','E_x','var_IP', ...
            'var_x','std_IP','std_x','E_MP','E_y','var_MP','var_y', ...
            'std_MP','std_y','X','k','XTX','b','s2'}, {vars8.name}))
    else
        regression
    end
else
    regression
end

clear

%% Choice of data
%
% We are interested in seeing the effect of different grading machines and
% the effect of different ranges of indicating property, we therefore
% choose to run this script for one region (differences between regions are
% not of interest here), and we choose to look at 3 different machines:
% from the regression we choose the one with the smallest variance, the one
% with the largest variance and one in the middle. We are only interested
% in the relative difference between good and bad grading machines, so we
% will exclude the names of the machines here, and just call them, from
% best to worst: "Machine #1", "Machine #2" and Machine #3".

% We first choose one of our two regions:
% 1 = Västergötland
% 2 = Lappland
Region = 1;

% Next we choose which grading machines to look at:
variance = zeros(5,1);
for Machine = 1:5
    regfile = ['regression_region' num2str(Region) '_machine' ...
        num2str(Machine) '.mat'];
    load(regfile,'s2')
    variance(Machine) = s2;
end
clear Machine regfile s2

machine = zeros(3,1);

% We find which machine has the lowest variance:
[~,machine(1)] = min(variance);

% Next we find the machine with the highest variance:
[~,machine(3)] = max(variance);

% And lastly we find the machine with the value closest to the mean of the
% highest and lowest:
mid = (variance(machine(1))+variance(machine(3)))./2;
temp = abs(variance - mid);
[~,machine(2)] = min(temp);
clear mid temp variance

% We want to run the script for each of three grades:
% 1 = C24
% 2 = C30
% 3 = C40
% The values of the fractiles we are interested in for the different
% grades:
```

```
grades = {24 30 40}; % The 5 % fractiles

%% Input parameters
%
% We choose some input parameters, the percent fractile we are interested
% in (this is normally 5 % for strength), the number of different ranges of
% indicating property we want to look at, and the accuracy of the graphs we
% will make:

fract = 0.05; % The fractile we are interested in
numRange = 4; % Number of different ranges of IP for each grade
if numRange < 2
    error('numRange must be an integer of value 2 or greater')
end

% Since the function for our distribution is very slow to compute, we
% choose to set a custom tolerance for the plots (we can set this lower to
% get more precise plots - default value 2e-3):
tol = 2e-3;

%% Finding limits of indicating property
%
% Given the calculations from the regression we want to find our limits for
% the indicating property. We use our custom functions 'find_xL', 'find_xH'
% and 'find_x'.

limits = zeros(numRange,2,3,3); % Preallocation

for Machine = 1:3
    % We load the data we want to use:
    regfile = ['regression_region' num2str(Region) '_machine' ...
        num2str(machine(Machine)) '.mat'];

    load(regfile)

    clear regfile

    for Grade = 1:3
        grade = grades{Grade};

        % We start by finding the mimimum lower limit possible (when upper
        % limit goes to infinity):
        % Upper limit:
        limits(1,2,Grade,Machine) = Inf;
        % Lower limit:
        limits(1,1,Grade,Machine) = find_xL(b,s2,n,XTX,E_x,var_x,n, ...
            limits(1,2,Grade,Machine),fract,grade);

        % Next we find the theoretically maximum lower limit (when upper
        % limit equals lower limit). This is not a valid range of IP (as
        % xL=xH), but we calculate it to show the extremes in either
        % direction:
        % Upper limit = lower limit:
        limits(numRange,:,Grade,Machine) = find_x(b,s2,n,XTX,fract,grade);

        % Next we define our remaining lower limits (xL), and calculate the
        % corresponding upper limits (xH):
        if numRange > 2
            a = (limits(numRange,1,Grade,Machine)- ...
```

```
                limits(1,1,Grade,Machine))./sum(1:(numRange-1));
            for t = 2:(numRange-1)
                % Lower limit:
                limits(t,1,Grade,Machine) = ...
                    limits((t-1),1,Grade,Machine)+(t-1).*a;
                % Upper limit:
                limits(t,2,Grade,Machine) = find_xH(b,s2,n,XTX,E_x, ...
                    var_x,n,limits(t,1,Grade,Machine),fract,grade);
            end
            clear a
        end
    end
end

clear Machine Grade grade t

%% Range of plots
%
% We need to set the limits for our plots, and we will set one range when
% plotting the distributions of ln(MP), and one range when plotting the
% distributions of MP (MP = Material Property, in our case strength). We
% will make different ranges for the different grades. The widest spread is
% from IP-range with upper limit to infinity and from machine 3 (the least
% accurate).

% We need to load the data we will use:
regfile = ['regression_region' num2str(Region) '_machine' ...
    num2str(machine(3)) '.mat'];

load(regfile)

clear regfile

% Preallocation of array with ranges for  MP (for use in figures):
plotRange = zeros(3,2);
deltaL = 1e-4; % How far from 0 we want our lower value to be
deltaU = 5e-3; % How far from 1 we want our upper value to be
for Grade = 1:3
    plotRange(Grade,:) = plot_range(b,s2,n,XTX,E_x,var_x,n, ...
        limits(1,1,Grade,3),limits(1,2,Grade,3), ...
        deltaL,deltaU);
end
clear deltaL deltaU Grade

%% Plots
%
% Now that we have our limits, we can plot our distributions, using our
% custom function, 'distr'. We start by saving the different plots'
% abscissas and ordinates, and then we will combine the different plots
% into comparative figures.

plots = cell(2,2,numRange,3,3); % Preallocation

for Machine = 1:3
    % We load the data we want to use:
    regfile = ['regression_region' num2str(Region) '_machine' ...
        num2str(machine(Machine)) '.mat'];

    load(regfile)
```

```
    clear regfile

    for Grade = 1:3
        grade = grades{Grade};


        %% Range of IP
        %
        % Now we make plots for the different ranges of indicating property
        for t = 1:(numRange)
            % First we define our function handles for the probability
            % density and the cumulative distribution. These are different
            % when we have a single value of x (xL=xH).
            if t < numRange
                [PDF,CDF] = distr(b,s2,n,XTX,E_x,var_x,n, ...
                    limits(t,1,Grade,Machine),limits(t,2,Grade,Machine));
            else
                % For a single value of x, we have a t distribution as
                % follows:

                % We calculate the mean of y:
                mean_y = b(1)+b(2).*limits(numRange,1,Grade,Machine);

                % And we calculate the scale of the distribution:
                scale_y = sqrt(s2*(1+ ...
                    [1 limits(numRange,1,Grade,Machine)]* ...
                    (XTX\[1;limits(numRange,1,Grade,Machine)])));
                PDF = @(y)pdf('tLocationScale',y,mean_y,scale_y,n-2);
                CDF = @(y)cdf('tLocationScale',y,mean_y,scale_y,n-2);
                clear mean_y scale_y
            end

            % Name of plot:
            plotname = ['plot_region' num2str(Region) '_machine' ...
                num2str(Machine) '_C' num2str(grade) '_IPrange' ...
                num2str(t)];

            % Plot of the cumulative distribution of MP:
            [Xplot,Yplot] = fplot(CDF,log(plotRange(Grade,:)),tol);
            plots{1,1,t,Grade,Machine} = exp(Xplot);
            plots{1,2,t,Grade,Machine} = Yplot;

            % Plot of the probability density of MP:
            [Xplot,Yplot] = fplot(PDF,log(plotRange(Grade,:)),tol);
            plots{2,1,t,Grade,Machine} = exp(Xplot);
            plots{2,2,t,Grade,Machine} = Yplot;

            clear Xplot Yplot

        end
        clear PDF CDF t
    end
    clear Grade grade
end
clear Machine machine

%% Save
%
```

```
% We choose to save the plotdata and limits of indicating property:
save grading.mat limits plots

%% Figures
%
% Now that we have the plot data for each combination of range of IP, grade
% and machine, we make our figures of interest.

% We want to have 2 figures for each combination, and define the size of
% these first:
scrsz = get(groot,'ScreenSize');

fig(1) = figure(1);
set(fig(1),'Position',[(3*scrsz(3)-4*scrsz(4))/6 scrsz(4)/4 ...
            4*scrsz(4)/3 scrsz(4)/2])

fig(2) = figure(2);
set(fig(2),'Position',[(3*scrsz(3)-2*scrsz(4))/6 scrsz(4)/4 ...
            2*scrsz(4)/3 scrsz(4)/2])

clear scrsz

%% 1. Range of IP
%
% For a given machine and grade, we look at how the different range of IP
% effects the distribution of MP:

for Machine = 1:3
    for Grade = 1:3
        grade = grades{Grade};

        % We define the names on our legends.
        plotNames = cell(numRange,1);
        for t = 1:numRange
            if exp(limits(t,1,Grade,Machine)) >= 1000
                limL = sprintf('%.0f',exp(limits(t,1,Grade,Machine)));
            else
                limL = sprintf('%.2f',exp(limits(t,1,Grade,Machine)));
            end
            if limits(t,2,Grade,Machine) == Inf
                limU = '\infty';
            elseif exp(limits(t,2,Grade,Machine)) >= 1000
                limU = sprintf('%.0f',exp(limits(t,2,Grade,Machine)));
            else
                limU = sprintf('%.2f',exp(limits(t,2,Grade,Machine)));
            end
            if t < numRange
                plotNames(t) = cellstr(['IP: ',limL,' - ',limU]);
            else
                plotNames(t) = cellstr(['IP: ',limL]);
            end
        end
        clear limL limU t


        % We define names of the figures:
        name1 = ['Cumulative Distribution - C' num2str(grade) ...
            ' [Machine #' num2str(Machine) ']'];
        name2 = ['Probability Density - C' num2str(grade) ' [Machine #' ...
```

```
    num2str(Machine) ']'];

% And now we can make our plots:

% The first figure shows the cumulative distribution of MP:
fig(1) = figure(1);
clf
set(fig(1),'Name',name1)

subplot(1,2,1)
hold on
for t=1:numRange
    plot(plots{1,1,t,Grade,Machine},plots{1,2,t,Grade,Machine})
end
hold off
legend(plotNames,'Location','southeast')
xlim(plotRange(Grade,:))
ylim([0 1])
xlabel('Strength')
ylabel('Cummulative Distribution')

% This subplot is a zoomed in view of the plot above:
subplot(1,2,2)
hold on
for t=1:numRange
    plot(plots{1,1,t,Grade,Machine},plots{1,2,t,Grade,Machine})
end
% We also want to plot and indicator of where the wanted fractile
% for the grade is:
plot([grade-10 grade grade],[fract fract 0],':k')
hold off
legend([plotNames;'5 % Fractile'],'Location','northwest')
axis([grade-10 grade+4 0 0.07])
xlabel('Strength')
ylabel('Cummulative Distribution')

% The second figure shows the probability density of MP:
fig(2) = figure(2);
clf
set(fig(2),'Name',name2)
hold on
for t=1:numRange
    plot(plots{2,1,t,Grade,Machine},plots{2,2,t,Grade,Machine})
end
hold off
set(gca,'YTickLabel',[])
legend(plotNames,'Location','northeast')
xlim(plotRange(Grade,:))
xlabel('Strength')
ylabel('Probability Density')

clear name1 name2 plotNames

%% Save
%
% Lastly we save the figures:
figname = ['grading_machine' num2str(Machine) '_C' ...
    num2str(grade) '.fig'];
savefig(fig,figname)
clear figname
```

```
    end
end
clear Machine Grade grade t

%% 2. Different grading machines
%
% For a given grade and distribution of indicating property, we look at how
% different grading machines effect the distribution of the material
% property. The indicating properties are, naturally, different from
% machine to machine, but we use, relatively, a similar distribution
% between the two extreme cases (upper limit to infinity and upper limit
% equal to lower limit).

for Grade = 1:3
    grade = grades{Grade};
    for t = 1:numRange

        % We define the names on our legends.
        plotNames = cell(3,1);
        for Machine = 1:3
            plotNames(Machine) = cellstr(['Machine #' ...
                num2str(Machine)]);
        end
        clear Machine

        % We define names of the figures:
        name1 = ['Cumulative Distribution - C' num2str(grade) ...
            ' [IP range #' num2str(t) ']'];
        name2 = ['Probability Density - C' num2str(grade) ...
            ' [IP range #' num2str(t) ']'];

        % And now we can make our plots:

        % The first figure shows the cumulative distribution of MP:
        fig(1) = figure(1);
        clf
        set(fig(1),'Name',name1)

        subplot(1,2,1)
        hold on
        for Machine=1:3
            plot(plots{1,1,t,Grade,Machine},plots{1,2,t,Grade,Machine})
        end
        hold off
        legend(plotNames,'Location','southeast')
        xlim(plotRange(Grade,:))
        ylim([0 1])
        xlabel('Strength')
        ylabel('Cummulative Distribution')

        % This subplot is a zoomed in view of the plot above:
        subplot(1,2,2)
        hold on
        for Machine=1:3
            plot(plots{1,1,t,Grade,Machine},plots{1,2,t,Grade,Machine})
        end
        % We also want to plot and indicator of where the wanted fractile
        % for the grade is:
        plot([grade-10 grade grade],[fract fract 0],':k')
        hold off
```

```
        legend([plotNames;'5 % Fractile'],'Location','northwest')
        axis([grade-10 grade+4 0 0.07])
        xlabel('Strength')
        ylabel('Cummulative Distribution')

        % The second figure shows the probability density of MP:
        fig(2) = figure(2);
        clf
        set(fig(2),'Name',name2)
        hold on
        for Machine=1:3
            plot(plots{2,1,t,Grade,Machine},plots{2,2,t,Grade,Machine})
        end
        hold off
        set(gca,'YTickLabel',[])
        legend(plotNames,'Location','northeast')
        xlim(plotRange(Grade,:))
        xlabel('Strength')
        ylabel('Probability Density')

        clear name1 name2 plotNames

        %% Save
        %
        % Lastly we save the figures:
        figname = ['grading_IPrange' num2str(t) '_C' num2str(grade) ...
            '.fig'];
        savefig(fig,figname)
        clear figname

    end
end
clearvars -except limits plots
```

B.1.4 reliability.m

The reliability.m script calculates the probability of failure for a set of machine settings, but also creates graphs showing how the probability of failure varies for the different machine settings. The script is quite slow to run, but by commenting out the graph-part of the code, it runs a lot faster.

```
%% Structural Reliability
%
% We will, given our distribution of the material resistance, and a
% distribution of the load (stress), evaluate the reliability of the
% structure. We will do this by calculating the probability of failure.
%
% THIS SCRIPT IS VERY SLOW - MAY TAKE CLOSE TO AN HOUR, DEPENDING ON THE
% ACCURACY OF THE PLOTS!!!

%% Importing data
%
% We start by checking if the regression script has been run, and if not,
% run it.

if exist('regression_region1_machine1.mat','file') && ...
        exist('regression_region1_machine2.mat','file') && ...
        exist('regression_region1_machine3.mat','file') && ...
        exist('regression_region1_machine4.mat','file') && ...
        exist('regression_region2_machine1.mat','file') && ...
        exist('regression_region2_machine2.mat','file') && ...
        exist('regression_region2_machine3.mat','file') && ...
        exist('regression_region2_machine4.mat','file')
    vars1 = whos('-file','regression_region1_machine1.mat');
    vars2 = whos('-file','regression_region1_machine2.mat');
    vars3 = whos('-file','regression_region1_machine3.mat');
    vars4 = whos('-file','regression_region1_machine4.mat');
    vars5 = whos('-file','regression_region2_machine1.mat');
    vars6 = whos('-file','regression_region2_machine2.mat');
    vars7 = whos('-file','regression_region2_machine3.mat');
    vars8 = whos('-file','regression_region2_machine4.mat');
    if all(ismember({'MP','IP','n','x','y','E_IP','E_x','var_IP', ...
            'var_x','std_IP','std_x','E_MP','E_y','var_MP','var_y', ...
            'std_MP','std_y','X','k','XTX','b','s2'}, {vars1.name})) && ...
            all(ismember({'MP','IP','n','x','y','E_IP','E_x','var_IP', ...
            'var_x','std_IP','std_x','E_MP','E_y','var_MP','var_y', ...
            'std_MP','std_y','X','k','XTX','b','s2'}, {vars2.name})) && ...
            all(ismember({'MP','IP','n','x','y','E_IP','E_x','var_IP', ...
            'var_x','std_IP','std_x','E_MP','E_y','var_MP','var_y', ...
            'std_MP','std_y','X','k','XTX','b','s2'}, {vars3.name})) && ...
            all(ismember({'MP','IP','n','x','y','E_IP','E_x','var_IP', ...
            'var_x','std_IP','std_x','E_MP','E_y','var_MP','var_y', ...
            'std_MP','std_y','X','k','XTX','b','s2'}, {vars4.name})) && ...
            all(ismember({'MP','IP','n','x','y','E_IP','E_x','var_IP', ...
            'var_x','std_IP','std_x','E_MP','E_y','var_MP','var_y', ...
            'std_MP','std_y','X','k','XTX','b','s2'}, {vars5.name})) && ...
            all(ismember({'MP','IP','n','x','y','E_IP','E_x','var_IP', ...
            'var_x','std_IP','std_x','E_MP','E_y','var_MP','var_y', ...
            'std_MP','std_y','X','k','XTX','b','s2'}, {vars6.name})) && ...
            all(ismember({'MP','IP','n','x','y','E_IP','E_x','var_IP', ...
            'var_x','std_IP','std_x','E_MP','E_y','var_MP','var_y', ...
            'std_MP','std_y','X','k','XTX','b','s2'}, {vars7.name})) && ...
            all(ismember({'MP','IP','n','x','y','E_IP','E_x','var_IP', ...
            'var_x','std_IP','std_x','E_MP','E_y','var_MP','var_y', ...
            'std_MP','std_y','X','k','XTX','b','s2'}, {vars8.name}))
```

```
    else
        regression
    end
else
    regression
end

clear

%% Choice of data
%
% We are interested in seeing the effect of different grading machines and
% the effect of different ranges of indicating property, we therefore
% choose to run this script for one region (differences between regions are
% not of interest here), and we choose to look at 3 different machines:
% from the regression we choose the one with the smallest variance, the one
% with the largest variance and one in the middle. We are only interested
% in the relative difference between good and bad grading machines, so we
% will exclude the names of the machines here, and just call them, from
% best to worst: "Machine #1", "Machine #2" and Machine #3".

% We first choose one of our two regions:
% 1 = Västergötland
% 2 = Lappland
Region = 1;

% Next we choose which grading machines to look at:
variance = zeros(5,1);
for Machine = 1:5
    regfile = ['regression_region' num2str(Region) '_machine' ...
        num2str(Machine) '.mat'];
    load(regfile,'s2')
    variance(Machine) = s2;
end
clear Machine regfile s2

machine = zeros(3,1);

% We find which machine has the lowest variance:
[~,machine(1)] = min(variance);

% Next we find the machine with the highest variance:
[~,machine(3)] = max(variance);

% And lastly we find the machine with the value closest to the mean of the
% highest and lowest:
mid = (variance(machine(1))+variance(machine(3)))./2;
temp = abs(variance - mid);
[~,machine(2)] = min(temp);
clear mid temp variance

% We want to run the script for each of three grades:
% 1 = C24
% 2 = C30
% 3 = C40
% The values of the fractiles we are interested in for the different
% grades:
grades = {24 30 40}; % The 5 % fractiles
```

```
%% Input parameters
%
% We choose some input parameters, the percent fractile we are interested
% in (this is normally 5 % for strength), the number of different ranges of
% indicating property we want to look at, and the accuracy of the graphs we
% will make:

fract = 0.05; % The fractile we are interested in
numRange = 4; % Number of different ranges of IP for each grade
if numRange < 2
    error('numRange must be an integer of value 2 or greater')
end

% When creating the continous graph of probability of failure for a given
% machines, we choose how many points on the plot we want to calculate.
% Setting this value to 100 made this script run for approx 45 minutes.
% Since the line of the plot is fairly straight, it is not recommended to
% use any higher value than this (can reduce this and still get usable
% graphs).
plotPoints = 100;

%% Finding limits of indicating property
%
% Given the calculations from the regression we want to find our limits for
% the indicating property. We use our custom functions 'find_xL', 'find_xH'
% and 'find_x'.

% We first check if the grading script has already been run, if so, we just
% import the limits already calculated there:
if exist('grading.mat','file')
    vars = whos('-file','grading.mat');
    if all(ismember({'limits','plots'}, {vars.name}))
        load('grading.mat','limits')
    end
    clear vars
end

if exist('limits','var')
else
    limits = zeros(numRange,2,3,3); % Preallocation

    for Machine = 1:3
        % We load the data we want to use:
        regfile = ['regression_region' num2str(Region) '_machine' ...
            num2str(machine(Machine)) '.mat'];

        load(regfile)

        clear regfile

        for Grade = 1:3
            grade = grades{Grade};

            % We start by finding the mimimum lower limit possible (when
            % upper limit goes to infinity):
            % Upper limit:
            limits(1,2,Grade,Machine) = Inf;
            % Lower limit:
            limits(1,1,Grade,Machine) = find_xL(b,s2,n,XTX,E_x,var_x,n, ...
```

```
                    limits(1,2,Grade,Machine),fract,grade);

                % Next we find the theoretically maximum lower limit (when
                % upper limit equals lower limit). This is not a valid range of
                % IP (as xL=xH), but we calculate it to show the extremes in
                % either direction:
                % Upper limit = lower limit:
                limits(numRange,:,Grade,Machine) = find_x(b,s2,n,XTX,fract, ...
                    grade);

                % Next we define our remaining lower limits (xL), and calculate
                % the corresponding upper limits (xH):
                if numRange > 2
                    a = (limits(numRange,1,Grade,Machine)- ...
                        limits(1,1,Grade,Machine))./sum(1:(numRange-1));
                    for t = 2:(numRange-1)
                        % Lower limit:
                        limits(t,1,Grade,Machine) = ...
                            limits((t-1),1,Grade,Machine)+(t-1).*a;
                        % Upper limit:
                        limits(t,2,Grade,Machine) = find_xH(b,s2,n,XTX,E_x, ...
                            var_x,n,limits(t,1,Grade,Machine),fract,grade);
                    end
                    clear a
                end
            end
        end
    end

clear Machine Grade grade t

%% Distribution of load
%
% We have defined the distribution of our material resistance already (from
% our custom function 'distr.m'). We will now also define the distribution
% of the load.

% We choose to use an extreme value distribution for our load (Gumbel Max).
% We have the following parameters for the distribution:

% Mean value
E_S0 = 1;

% Coefficient of variation
cov_S = 0.2;

% Standard deviation
std_S0 = E_S0.*cov_S;

% The Gumbel Max distribution, expressed with regards to the mean and
% standard deviation, and with ln(strength) as the input variable, can be
% expressed with the following parameters:

a = pi()./(std_S0.*sqrt(6));
u = E_S0-double(eulergamma)./a;

%% Relation between distributions
%
% We want the distribution of the material resistance and the distribution
```

```
% of the stress (load) to follow a relation as indicated by the structural
% codes (the ultimate limit), and we do this by calculating the structural
% factor, z, for the different grades:

% Safety factors:
gamma_R = 1.3;
gamma_S = 1.5;

% Characteristic load:
s_k = u - log(-log(0.98))./a;

z = zeros(3,1);
for Grade = 1:3
    grade = grades{Grade};
    z(Grade) = grade./(s_k.*gamma_S.*gamma_R);
end
clear Grade grade gamma_R gamma_S s_k

%% Distribution of z*S
%
% We want to use the distribution of z*S when calculating the probability
% of failure, and this distribution is also a Gumbel distribution, but with
% an adjusted value for the mean and standard deviation:

% We express the distribution for the different grades:

F_S = cell(3,1);
for Grade = 1:3
    E_S = E_S0.*z(Grade); % New mean
    std_S = E_S.*cov_S; % New standard deviation
    a = pi()./(std_S.*sqrt(6));
    u = E_S-double(eulergamma)./a;
    F_S{Grade} = @(y)exp(-exp(-a.*(exp(y)-u)));
end

clear Grade E_S0 cov_S std_S0 E_S std_S a u

%% Probability of failure
%
% We can now calculate the probability of failure, using our custom
% function, 'failure':

Pf = zeros(numRange,1,3,3);
plots = cell(1,2,3,3);

for Machine = 1:3
    % We load the data we want to use:
    regfile = ['regression_region' num2str(Region) '_machine' ...
        num2str(machine(Machine)) '.mat'];

    load(regfile)

    clear regfile

    for Grade = 1:3
        for t = 1:numRange
            if t<numRange
                % The probability of failure is:
                Pf(t,1,Grade,Machine) = failure(b,s2,n,XTX,E_x,var_x,n, ...
```

```
                limits(t,1,Grade,Machine), ...
                limits(t,2,Grade,Machine), F_S{Grade});
        else
            % For a single value of x, we have a t distribution as
            % follows:

            % We calculate the mean of y:
            mean_y = b(1)+b(2).*limits(numRange,1,Grade,Machine);

            % And we calculate the scale of the distribution:
            scale_y = sqrt(s2*(1+ ...
                [1 limits(numRange,1,Grade,Machine)]* ...
                (XTX\[1;limits(numRange,1,Grade,Machine)])));
            f_R = @(y)pdf('tLocationScale',y,mean_y,scale_y,n-2);
            g = @(y)f_R(y).*F_S{Grade}(y);
            Pf(t,1,Grade,Machine) = 1-integral(g,-Inf,Inf);
            clear mean_y scale_y f_R g
        end
    end

    % In addition to calculating the probability of failure for given
    % values of xL and xH, we can make plots showing how the structural
    % reliability varies with the indicating property for the different
    % machines.

    % We can get a graph by choosing descrete points of xL, calculate
    % corresponding values for xH, and use this to make the plot.

    grade = grades{Grade};
    xL_min = limits(1,1,Grade,Machine);
    xL_max = limits(numRange,1,Grade,Machine);
    xL = (xL_min:((xL_max-xL_min)/(plotPoints-1)):xL_max).';
    xH = zeros(size(xL));
    Pf_plot = xH;

    for x_L = 1:plotPoints
        xH(x_L) = find_xH(b,s2,n,XTX,E_x,var_x,n,xL(x_L),fract,grade);
        Pf_plot(x_L) = failure(b,s2,n,XTX,E_x,var_x,n,xL(x_L), ...
            xH(x_L),F_S{Grade});
    end
    plots{1,1,Grade,Machine} = (exp(xL)-exp(xL_min))./(exp(xL_max)- ...
        exp(xL_min));
    plots{1,2,Grade,Machine} = Pf_plot;
    clear PlotPoints grade xL_min xL_max xL xH Pf_plot
  end
end

%% Plots
%
% We choose to plot the distribution of the probability of failure as a
% function of the lower limit of IP, comparing grading machines in the same
% figure. We create a separate figure for each grade.

% We define the names on our legends.
plotNames = cell(3,1);
for Machine = 1:3
    plotNames(Machine) = cellstr(['Machine #' num2str(Machine)]);
end
clear Machine
```

```
% We define our figure
scrsz = get(groot,'ScreenSize');
fig(1) = figure(1);
set(fig(1),'Position',[(3*scrsz(3)-2*scrsz(4))/6 scrsz(4)/4 ...
    2*scrsz(4)/3 scrsz(4)/2])

for Grade = 1:3
    grade = grades{Grade};
    % We name our figure:
    name = ['Probability of failure - C' num2str(grade)];
    fig(1) = figure(1);
    clf
    set(fig(1),'Name',name)

    % And we create our plots:
    hold on
    for Machine = 1:3
        plot(plots{1,1,Grade,Machine},plots{1,2,Grade,Machine})
    end
    legend(plotNames,'Location','southeast')
    xlabel('Machine Settings')
    ylabel('Probability of Failure')
    xlim([0 1])
    set(gca,'XTick',[0 1],'XTickLabel',{'Max Range of IP', ...
        'Single Value of IP'})

    %% Save
    %
    % Lastly we save the figures:
    figname = ['reliability_C' num2str(grade) '.fig'];
    savefig(fig,figname)
    clear figname
end

%% Save
%
% We save the data we have found and remove the other variables

clearvars -except F_S Pf plots
save reliability.mat
```

B.2 Functions

B.2.1 find_x.m

The find_x.m function calculates the value of x when $x_L = x_H = x$.

```
function x = find_x(E_beta,s2,n,XTX,fr,grade)
% The function returns the value of x (xL=xH)

% First we read each element of the XTX matrix in order to avoid matrix
% operations in the integration below
a = XTX(1,1);
b = XTX(1,2);
c = XTX(2,1);
d = XTX(2,2);

% We define our degrees of freedom (from the regression)
nu = n-2;

% Next we define the cumulative distribution of y as a function of x, and
% subtract the fractile we are interested in (If we want to find a 5 %
% fractile, we subtract 0.05), so that the function equals 0 at this point.
    function t = f(x)
        % We calculate the mean of y (logarithm of material property)
        E_y = E_beta(1)+E_beta(2).*x;
        % Next we calculate the scale of the distribution:
        scale_y = sqrt(s2.*(1+1./a+(x-b/a).*(a.*x-c)./(a.*d-b.*c)));
        % Now we can find the cumulative distribution of y, and subtract
        % the fractile:
        t = cdf('tLocationScale',log(grade),E_y,scale_y,nu)-fr;
    end

% Given the distribution, we can find the value of x which makes the
% function zero:
x = fzero(@f,(log(grade)-E_beta(1))./E_beta(2));
end
```

B.2.2 find_xH.m

The find_xH function calculates the value of x_H, given a value of x_L

```
function xH = find_xH(E_beta,s2,n,XTX,E_x,s2x,nx,xL,fr,grade)
% The function returns the upper limit, given the lower

% First we read each element of the XTX matrix in order to avoid matrix
% operations in the integration below
a = XTX(1,1);
b = XTX(1,2);
c = XTX(2,1);
d = XTX(2,2);

% We defines our degrees of freedom, from the regression, and for the
% distribution of the indicating property:
nuy = n-2; % Regression
nux = nx-1; % IP

% We calculate the square of the scale of the distribution of x:
sscale_x = (1+1./nx).*s2x;

% Now we define the inner function which we will integrate in order to find
% our cumulative distribution:
    function u = f(x,y)
        % We calculate the mean of y:
        E_y = E_beta(1)+E_beta(2).*x;
        % Next we calculate the square of the scale of the distribution:
        sscale_y = s2.*(1+1./a+(x-b/a).*(a.*x-c)./(a.*d-b.*c));
        % And finally we define our inner function:
        u = (1+(x-E_x).^2./(nux.*sscale_x)).^(-(nux+1)./2).*sscale_y.^ ...
            (-1/2).*(1+(y-E_y).^2./(nuy.*sscale_y)).^(-(nuy+1)./2);
    end

% Next we calculate the cumulative distribution of y as a function of xH,
% and subtract the fractile we are interested in (If we want to find a 5 %
% fractile, we subtract 0.05), so that the function equals 0 at this point.
    function t = g(xH)
        % We first need to find the reciprocal of the normalizing constant:
        function w = C
            w = integral2(@f,xL,xH,-Inf,Inf);
        end
        % Then we can calculate the distribution and subtract the fractile:
        t = integral2(@f,xL,xH,-Inf,log(grade))./C-fr;
    end

% Given the distribution, we can find the value of xH which makes the
% function zero:
xH = fzero(@g,(log(grade)-E_beta(1))./E_beta(2));
end
```

B.2.3 find_xL.m

The find_xL function calculates the value of x_L, given a value of x_H

```
function xL = find_xL(E_beta,s2,n,XTX,E_x,s2x,nx,xH,fr,grade)
% The function returns the upper limit, given the lower

% First we read each element of the XTX matrix in order to avoid matrix
% operations in the integration below
a = XTX(1,1);
b = XTX(1,2);
c = XTX(2,1);
d = XTX(2,2);

% We defines our degrees of freedom, from the regression, and for the
% distribution of the indicating property:
nuy = n-2; % Regression
nux = nx-1; % IP

% We calculate the square of the scale of the distribution of x:
sscale_x = (1+1./nx).*s2x;

% Now we define the inner function which we will integrate in order to find
% our cumulative distribution:
    function u = f(x,y)
        % We calculate the mean of y:
        E_y = E_beta(1)+E_beta(2).*x;
        % Next we calculate the square of the scale of the distribution:
        sscale_y = s2.*(1+1./a+(x-b/a).*(a.*x-c)./(a.*d-b.*c));
        % And finally we define our inner function:
        u = (1+(x-E_x).^2./(nux.*sscale_x)).^(-(nux+1)./2).*sscale_y.^ ...
            (-1/2).*(1+(y-E_y).^2./(nuy.*sscale_y)).^(-(nuy+1)./2);
    end

% Next we calculate the cumulative distribution of y as a function of xL,
% and subtract the fractile we are interested in (If we want to find a 5 %
% fractile, we subtract 0.05), so that the function equals 0 at this point.
    function t = g(xL)
        % We first need to find the reciprocal of the normalizing constant:
        function w = C
            w = integral2(@f,xL,xH,-Inf,Inf);
        end
        % Then we can calculate the distribution and subtract the fractile:
        t = integral2(@f,xL,xH,-Inf,log(grade))./C-fr;
    end

% Given the distribution, we can find the value of xL which makes the
% function zero:
xL = fzero(@g,(log(grade)-E_beta(1))./E_beta(2));
end
```

B.2.4 distr.m

The distr.m function returns the probability density function and the cumulative distribution function for the material property, given range of indicating property.

```
function [PDF,CDF] = distr(E_beta,s2,n,XTX,E_x,s2x,nx,xL,xH)
% This function returns the probability density function and the cumulative
% distribution function of the material property, given a range of the
% indicating property

PDF = @dens; % Function handle for the probability density
CDF = @cumul; % Function handle for the cumulative distribution

% We read each element of the XTX matrix in order to avoid matrix
% operations in the integration below
a = XTX(1,1);
b = XTX(1,2);
c = XTX(2,1);
d = XTX(2,2);

% We defines our degrees of freedom, from the regression, and for the
% distribution of the indicating property:
nuy = n-2; % Regression
nux = nx-1; % IP

% We calculate the square of the scale of the distribution of x:
sscale_x = (1+1./nx).*s2x;

% Now we define the inner function which we will integrate in order to find
% our distributions:
    function u = f(x,y)
        % We calculate the mean of y:
        E_y = E_beta(1)+E_beta(2).*x;
        % Next we calculate the square of the scale of the distribution:
        sscale_y = s2.*(1+1./a+(x-b/a).*(a.*x-c)./(a.*d-b.*c));
        % And finally we define our inner function:
        u = (1+(x-E_x).^2./(nux.*sscale_x)).^(-(nux+1)./2).*sscale_y.^ ...
            (-1/2).*(1+(y-E_y).^2./(nuy.*sscale_y)).^(-(nuy+1)./2);
    end

% Next we find the reciprocal of the normalizing constant:
    function w = C
        w = integral2(@f,xL,xH,-Inf,Inf);
    end

% And now we can integrate to find our density function:
    function v = dens(y)
        v = integral(@(x)f(x,y),xL,xH)./C;
    end

% and our cumulative distribution function:
    function t = cumul(y)
        t = integral2(@f,xL,xH,-Inf,y)./C;
    end
end
```

B.2.5 plot_range.m

The plot_range.m function calculates a range of material property to be calculated when making the plots.

```
function range_MP = plot_range(E_beta,s2,n,XTX,E_x,s2x,nx, ...
    xL,xH,deltaL,deltaU)
% This function returns a suitable range to plot our distribution, given a
% range of indicating property, and a delta value we want our cumulative
% function to be away from 0 and 1

% We read each element of the XTX matrix in order to avoid matrix
% operations in the integration below
a = XTX(1,1);
b = XTX(1,2);
c = XTX(2,1);
d = XTX(2,2);

% We defines our degrees of freedom, from the regression, and for the
% distribution of the indicating property:
nuy = n-2; % Regression
nux = nx-1; % IP

% We calculate the square of the scale of the distribution of x:
sscale_x = (1+1./nx).*s2x;

% Now we define the inner function which we will integrate in order to find
% our distributions:
    function u = f(x,y)
        % We calculate the mean of y:
        E_y = E_beta(1)+E_beta(2).*x;
        % Next we calculate the square of the scale of the distribution:
        sscale_y = s2.*(1+1./a+(x-b/a).*(a.*x-c)./(a.*d-b.*c));
        % And finally we define our inner function:
        u = (1+(x-E_x).^2./(nux.*sscale_x)).^(-(nux+1)./2).*sscale_y.^ ...
            (-1/2).*(1+(y-E_y).^2./(nuy.*sscale_y)).^(-(nuy+1)./2);
    end

% Next we find the reciprocal of the normalizing constant:
    function w = C
        w = integral2(@f,xL,xH,-Inf,Inf);
    end

% And we find our cumulative distribution function:
    function t = cumul(y)
        t = integral2(@f,xL,xH,-Inf,y)./C;
    end

% We now define two functions we use to find our limits (by setting them
% equal to zero):
    function v = lower(y)
        v = cumul(y)-deltaL;
    end
    function s = upper(y)
        s = cumul(y)-1+deltaU;
    end
range_MP = [floor(exp(fzero(@lower,0))./5).*5 ...
    ceil(exp(fzero(@upper,0))./5).*5];
end
```

B.2.6 failure.m

The failure.m function calculates the probability of failure.

```
function Pf = failure(E_beta,s2,n,XTX,E_x,s2x,nx,xL,xH,F_S)
% The function returns the probability of failure for a given range of x.

% First we read each element of the XTX matrix in order to avoid matrix
% operations in the integration below
a = XTX(1,1);
b = XTX(1,2);
c = XTX(2,1);
d = XTX(2,2);

% We defines our degrees of freedom, from the regression, and for the
% distribution of the indicating property:
nuy = n-2; % Regression
nux = nx-1; % IP

% We calculate the square of the scale of the distribution of x:
sscale_x = (1+1./nx).*s2x;

% Next we define the function from our distribution of the material
% property:
    function u = f(x,y)
        % We calculate the mean of y:
        E_y = E_beta(1)+E_beta(2).*x;
        % Next we calculate the square of the scale of the distribution:
        sscale_y = s2.*(1+1./a+(x-b/a).*(a.*x-c)./(a.*d-b.*c));
        % And finally we define our inner function:
        u = (1+(x-E_x).^2./(nux.*sscale_x)).^(-(nux+1)./2).*sscale_y.^ ...
            (-1/2).*(1+(y-E_y).^2./(nuy.*sscale_y)).^(-(nuy+1)./2);
    end

% Next we find the reciprocal of the normalizing constant:
    function w = C
        w = integral2(@f,xL,xH,-Inf,Inf);
    end

% And lastly we calculate the reliability of the structure:
g = @(x,y)f(x,y).*F_S(y);
Pf = 1-integral2(g,xL,xH,-Inf,Inf)./C;

end
```

Appendix C TIMBER DATA

Here is presented the data used to perform our calculations.

Gradewood ID-number	region_code	strength (150 mm incl. ki)	IP_MOR_GE706	IP_MOR_combiscan	IP_MOR_escan	IP_MOR_triomatic	IP_Rosegrade
216012	1	32.4	41.3	44.96683	42.93395	33.23348	7515840
216022	1	15.3	20.3	19.22404	19.3161	19.32841	6061567
216032	1	35.2	29.8	28.30006	26.27875	26.09348	6658325
216042	1	14.9	18.8	18.91271	14.65797	20.64039	6033634
216052	1	36.7	37.5	38.77105	36.19663	37.74203	8108252
216062	1	32.4	27.4	32.13787	28.93766	26.81351	7119214
216072	1	36.3	36.5	32.18785	31.70203	32.09888	7514819
216082	1	18.2	21.2	24.83597	26.13464	25.57706	6116840
216092	1	9.5	7.4	18.9101	13.97692	15.13009	5157186
216102	1	13.1	12.8	15.0881	16.66344	19.35136	5399252
216112	1	20.6	20.7	23.69823	22.20456	23.15007	6074059
216122	1	31.3	34.5	34.57905	33.78345	31.82603	7318400
216132	1	32.2	37.7	38.21084	36.5918	35.5279	7640377
216142	1	20.1	17.7	14.41349	15.19802	14.39964	5587993
216152	1	17.1	20.8	23.89394	23.06185	21.25357	5965571
216162	1	34.8	36.1	34.23926	34.32256	32.60841	7110718
216172	1	10.7	12.5	18.7365	12.47804	10.02928	4965357
216182	1	16.3	14.4	16.87079	13.09132	14.31505	5428229
216192	1	28.6	24.8	24.58251	24.0155	23.33336	6083686
216202	1	38.7	35.2	34.95924	35.45846	27.87585	7048585
216212	1	48.4	39.3	38.63916	37.06143	33.67408	7603326
216222	1	27.1	33.4	34.23937	33.75271	31.27417	7239466
216232	1	34.1	35.1	33.35844	32.19756	26.15563	6931609
216242	1	46.5	42.2	38.91471	36.06747	34.79463	8085256
216252	1	63.1	56.9	0	0	0	8940206
216262	1	14.6	13.5	15.05283	17.82512	14.32083	5434962
216272	1	12.6	16.3	21.29529	20.03409	19.43982	6288547
216282	1	36.8	35.1	37.59998	34.54115	33.28976	7823130
216292	1	14.9	14.1	14.36022	14.15219	11.94311	5430426
216302	1	9.7	9.8	13.7086	13.1065	12.94269	5072543
216312	1	24.8	35.2	35.0884	32.32302	29.71914	6917525
216322	1	23.3	21.5	26.62173	25.24488	19.78185	5752303

216332	1	18	16	20.49448	20.02645	15.9021	5706330
216342	1	30.7	28.1	30.17156	29.47452	26.14707	6846815
216352	1	25.8	28.2	23.89373	24.3172	27.7508	6770029
236012	2	20	15.5	14.95411	14.88039	11.38026	5113960
236022	2	14.9	10.1	14.79756	12.88746	12.93334	5064822
236032	2	29.6	27.4	30.59555	28.14708	28.09586	6679736
236042	2	24.5	27.3	24.78304	23.7453	22.97905	6579369
236052	2	21.4	15.7	17.24516	16.89156	15.0909	5365446
236062	2	25.5	25.5	25.76189	25.10198	22.51982	6361141
236072	2	34	27.8	28.39728	26.72549	24.99204	6496451
236082	2	29.8	26.8	22.42339	20.43882	23.71549	6322586
236092	2	16.2	16.7	17.57889	16.88639	19.87068	5715173
236102	2	33.9	25.2	25.72888	23.31152	24.39151	6428894
236112	2	27.8	21.1	23.79241	24.12462	19.14602	5488903
236122	2	26.1	24.9	23.0624	21.76518	23.94433	6361695
236132	2	34.3	32.2	34.79597	33.56081	29.38482	6850809
236142	2	23	24.2	22.81199	22.89673	21.15955	6066535
236152	2	26.1	17	17.52785	16.26016	15.58028	5489559
236162	2	24.9	24.3	23.25767	22.01458	21.85464	6161090
236172	2	19.7	20.9	24.37488	23.62634	16.27758	5730788
236182	2	8.8	4.6	7.198068	9.636651	6.648175	4223695
236192	2	15.5	18.1	19.57034	19.16547	19.97472	5659331
236202	2	25.3	23.7	22.92892	21.8336	15.81038	5567927
236212	2	13.7	14.2	14.00611	16.24645	11.39773	5046933
236222	2	26	24.3	22.06277	24.23895	22.98352	6228351
236232	2	24.2	22.9	22.0551	21.95341	23.48632	6368513
236242	2	27.3	21.4	20.26524	20.7869	24.39044	5864650
236252	2	22.1	14.5	17.49765	17.76282	20.80408	5485339
236262	2	24.1	18.2	18.0689	19.21476	18.45943	5485012
236272	2	16.4	12.4	18.34561	16.25041	13.32506	4980223
236282	2	34.8	26	27.06642	23.98848	23.09317	6165563
236292	2	31.3	22.2	16.88378	14.97225	20.93051	5810707
236302	2	26.9	23.8	18.5548	16.65842	18.26769	5788756
236312	2	33.5	28.2	27.03188	26.24042	26.55259	7001845
236322	2	36.2	27.3	28.57585	26.11068	15.92328	6373615
236332	2	25.8	20.2	21.10342	20.12929	16.36258	5878216
236342	2	19.5	26.6	19.80983	16.93289	22.2821	6396184
236352	2	27.6	26.7	25.91987	24.84277	23.59462	6563201