Australian Government
Forest and Wood Products Research and Development Corporation

MANUFACTURING & PRODUCTS

PROJECT NUMBER: PN05.2024

Defining statistical methods for monitoring structural properties of machine graded pine

SEPTEMBER 2006

This release can also be viewed on the FWPRDC website
www.fwprdc.org.au

FWPRDC
PO Box 69, World Trade Centre
Melbourne 8005, Victoria
T +61 3 9614 7544 F +61 3 9614 6822

Defining statistical methods for monitoring structural properties of machine graded pine

Prepared for the

Forest and Wood Products Research and Development Corporation

by

G N Boughton and D J Falck

The FWPRDC is jointly funded by the Australian forest and wood products industry and the Australian Government.

Australian Government

Forest and Wood Products Research and Development Corporation

Publication: Defining statistical methods for monitoring structural properties of machine graded pine

Project no: PN05.2024

Researchers:
G N Boughton and D J Falck
TimberED Services Pty Ltd

Final report received by the FWPRDC in September 2006

Forest and Wood Products Research and Development Corporation
PO Box 69, World Trade Centre, Victoria 8005
T +61 3 9614 7544 F +61 3 9614 6822 E info@fwprdc.org.au
W www.fwprdc.org.au

ISBN 978-1-920883-11-9

Executive Summary

This paper outlines generic tools that can be used in quality assurance programs to indicate an appropriate level of confidence in the data that is produced by product testing. These tools can be used by management to make technically informed decisions about programs used to monitor properties of production, and monitoring strategies that are based on random position testing of random specimens.

The outputs of this paper will enable producers to compare commercial implications of a range of options for daily testing and analysis of structural properties. An efficient and effective monitoring system will return reliable test data with known confidence. Once a particular monitoring strategy has been adopted, the outputs can be used to establish Test Comparison Values for use with the test results. Test Comparison Values can be monitored with feedback from the monitoring results.

This paper presents tables, equations and graphs that enable producers to determine Test Comparison Value that are used to assess test results for both MoE and strength. Where the test results exceed the Test Comparison Value, then there is the required confidence that the production in that period meets the required properties.

Producers can select from a range of confidence levels to suit their corporate risk management policies.

A number of different analysis methods have been presented:

Mean MoE

- Non-parametric
- Log-normal fit to the data
- From MSG data for the shift

5%ile MoE

- Non-parametric
- From Log-normal fit to the data

5%ile strength (bending or tension)

- Non-parametric
- From Log-normal fit to the data
- From Log-normal fit with known CoV
- From tail fit to Log-normal distribution
- From tail fit to 2 parameter Weibull distribution

The body of the report presents all of the background, methodology and technical justification for the outcomes. However, quality assurance managers can make use of the tables and equations to interpret test results on randomly sampled timber from batches, by using Appendices A and B without digesting the remainder of the report.

Table of contents

1. Background

Monitoring the strength and stiffness of structurally graded timber is a vital task in any quality assurance program. For the purpose of this study, the process of Monitoring Properties is taken to mean the taking of samples from production, the structural testing of the timber, and analysis to obtain an estimation of the properties of the entire production during that period. It can also include the collection of data from the machine stress grader to draw conclusions on the properties of the production run. Data can also be amalgamated between consecutive runs on the same product to estimate the longer-term properties of the product.

In order to make appropriate interpretations of the test results, it is necessary to have an understanding of the certainty of the test results. This project was conceived to deliver some tools that would enable producers to attach some statistically justifiable confidence to their data.

It was envisaged that the outputs of the project would enable producers to evaluate the various testing and analysis options open to them, and put in place statistically justified acceptance criteria for their test results.

The aim of monitoring of properties as used in this project is:

> to estimate whole of production properties from samples of production and use this information to control production parameters to deliver product with complying properties.

It follows that monitoring of properties is a process used to answer the question: *Do we have the required confidence that the properties of the product in that period at least meet the design value?*

In particular, the properties that are of importance to producers are the average MoE, 5%ile MoE and 5%ile strength properties of the production. It recognises that for constant grading parameters, some timber properties may change over time as the resource material changes, and that as a result some changes may need to be made to the production parameters. In quantifying the uncertainty in property estimations, the project enables producers to set acceptance levels for test results that will detect a drift in properties that may compromise the confidence they have in their product compliance. Having detected the shift in properties, the producer is then in a position to vary the production parameters so that the product remains within acceptable limits.

1.1 Sources of uncertainty in results

The aim of the testing and evaluation is to estimate the properties of the (untested) production during the period in which the samples were taken. This involves some sources of uncertainty – called error. In most cases, the error is hidden, and does not involve a "mistake". It is simply the term used to describe the difference between an estimate of a property and the actual property itself. The following is a list of the types of error that are likely to be part of an estimate of production properties.[1,2]

- Test result accuracy – machine inaccuracies, measurement errors
- Blunders – eg data entry errors

- Sampling error – difference between an estimate derived from a sample and the true value that would result if the whole production run was tested.
- Systematic error – over- or under-estimation by the analysis methods
- Interpretation errors – wrong decisions

Of all of these sources of error, this report only models the uncertainty caused by sampling and systematic errors.

1.2 Sampling and Systematic Errors

A test sample of timber includes only a small part of the total population of that particular grade and size. Intuitively, if a different group of pieces had been sampled, the results (average MoE or 5%ile strength) may have been slightly different. Statistics can be used to estimate just *how* different the result could have been. It is important to note that we can never determine the exact average MoE or 5%ile strength of the population. The best we can do is an estimate using the data from a tested sample of the production. Obviously the only way to know the exact values would be to test every piece of timber!

- Sampling errors are unavoidable where random samples are taken. The sample may not always accurately represent the character of the material produced. In many cases it does, but there is no way of knowing whether a small change in test result is due entirely to a change in the quality of the product or just the fact that the sample just happened to have better or poorer quality pieces than the production it was meant to represent. This is quite distinct from any conscious or sub-conscious biasing of the sample by people.
- In addition to the sampling errors, different analysis methods introduce small errors by generally over- or under-estimating the production property. (Over-estimation doesn't mean that every result will be high, but in the longer term, the method will more often estimate properties that are higher than the actual property.)

Sampling error is usually the most significant sources of discrepancy between test results of samples and the properties of production. This is the basis of some classical statistical definitions.

The population of material has a distribution of properties that can be represented by a mean and standard deviation. These values can be estimated from samples, but statistics recognises that the estimates will give different values.

Parameter	Population property	Estimate from sample
Mean	μ	$\overline{X}$
Standard deviation	σ	s

1.3 Confidence in Estimates

Wherever an estimation is made, we can attach a level of confidence to the estimate. The confidence level is a measure of how certain we are that a value will be within a given range. In the context of the monitoring of structural properties, this means:

The confidence level is a measure of our certainty that the production property is greater than the Design Value.

For example, if there is 80% confidence that the results show a mean MoE greater than the Design Value, then roughly 8 times out of 10, the assessment is right, but there is a 2 in 10 chance that it is actually less than the Design Value. A higher level of confidence – 95% gives roughly 19 out of 20 times that the assessment is right, but a 1 in 20 chance that it will be wrong.

Test Comparison Values are derived for a given level of confidence from the Design Values. The test results (estimate of the production property) can then be compared with them. If the test results are higher than the Test Comparison Values, then the producer can have the required level of confidence that the production property is greater than the Design Value. The Test Comparison Value is therefore always greater than the Design Value. Its calculation is detailed in Section 5.1 of this report. Higher levels of confidence lead to higher Test Comparison Values.

Some key points for compliance testing and assurance are:

- An estimate of a property greater than the Test Comparison Value means that the confidence in compliance of the production run is greater than the required confidence level.
- An estimate of a property less than the Design Value means that the confidence in the compliance of the production run is less than 50% (i.e. more likely to be non-compliant than compliant).

2. Objectives of this Project

To provide the industry with tools that can easily be used by management to make technically informed decisions about quality assurance strategies and programs.

To develop generic tools that can be used in quality assurance programs to indicate an appropriate level of confidence in the data that is produced by product testing.

2.1 Scope of the project

To present information that will enable producers to determine whether or not production batches meet design properties based on test data from random samples with random position testing and analysis using the techniques listed in Section 3.1.

Production batches can be defined as the quantity of graded material produced in the period over which a single sample of timber (consisting of a number of pieces drawn at random from production) was collected for an analysis. For example:

- Where an analysis is performed every shift, then the production batch is all of the timber produced in a single shift.
- Where an analysis is performed at the completion of every production run, then the production batch is all of the timber produced in a single production run.
- Where the data is accumulated over three shifts of one size, then the production batch is all of the timber produced in the three shifts.

Issues of Periodic Monitoring were not addressed in this study.

The report does not specifically address remedial action in the case of test data failing to meet the required confidence level. These decisions are a significant part of the mill QA strategy. However, an illustration of some possible strategies are provided in the examples in Section 6.

The study was restricted to interpretation of the data from a single analysis. It therefore focused on the results of a single batch.

3. Methodology

The methodology used in this study was as follows:

- Test Data provided by nine Australian mills was used to derive statistical distributions of the data that could be used to simulate mill production.
- Monte Carlo simulations[3] using known product properties (taken to be representative of Australian production as they were derived from the data provided by the mills) were used to replicate random sampling regimes. The sample results were used in various analysis techniques to estimate the properties of the production.
- The estimates of the properties of the production were used to derive confidence limits on the estimates.
- The confidence limits were used to derive the Test Comparison Values (TCV) that would give the required confidence that the data exceeded the Design Value. The TCVs were then plotted against sample size, and formulae were derived to represent the relationship.
- The results of the analysis were presented as tables or equations so that producers could easily establish the Test Comparison Values needed to give the required confidence that their product met the design properties.
- Examples were prepared to show the use of the confidence limits

The detail of the methodology is presented in Appendix C.

3.1 Analysis models

The analysis models selected for consideration in this report are all currently used or of interest within the Australian softwood industry. The following methods of analysis were addressed:

For establishing mean MoE:

- Non-parametric average of random sample
- Mean of a log-normal distribution through the random sample
- Mean of data taken from the machine stress-grader

For establishing 5%ile MoE:

- Non-parametric 5%ile of random sample
- 5%ile of a log-normal distribution through the random sample
- 5%ile of data taken from the machine stress-grader

For establishing 5%ile bending strength:

- Non-parametric 5%ile of random sample
- 5%ile of a log-normal distribution through the random sample
- 5%ile of a log-normal distribution through the random sample, constrained to a given CoV
- 5%ile of a log-normal distribution through the tail of a random sample
- 5%ile of a log-normal distribution through the tail of a random sample, constrained to a given CoV
- 5%ile of a 2 parameter Weibull distribution through the tail of a random sample

For establishing 5%ile tension strength:

- o Non-parametric 5%ile of random sample
- o 5%ile of a log-normal distribution through the random sample
- o 5%ile of a log-normal distribution through the random sample, constrained to a given CoV

The confidence levels of a CUSUM analysis were also addressed.

3.2 Methodology verification

Where possible the relationships were verified using classical statistical approaches[1,2] and by comparison with Australian and overseas standards.

Industry was consulted at a number of stages during the project:

- The industry was asked to provide test data to establish the nature of the distributions that were simulated in the study, and used to demonstrate the outcomes.
- In June 2005, representatives of a number of major Australian pine producers were visited and the preliminary findings from the project outlined. The findings were illustrated using their own data. Comments and feedback were sought during the visits.
- A draft of the report was circulated for comments prior to the preparation of the final report.

It is recognised that the Australian pine resource has varying characteristics depending on resource, species and grade, but the results were derived using generic parameters (such as CoV) and are presented for a range of values so that they are applicable to all resources.

The outcomes were trialled on a range of grades, sizes, resources and on two species (radiata pine and slash pine). Further details of the verification of the methodology are presented in C.4.3.

The draft report was also subjected to peer review prior to finalisation:

- The draft report was reviewed by an independent expert and the comments were incorporated into the final report.
- A paper based on the report was presented at the World Conference on Timber Engineering 2006 and was reviewed by a panel prior to acceptance for publication.

4. Methods of estimating production properties

The outputs from the project are outlined in Section 5 and presented in Appendices to the report.

This section briefly discusses the different methods of data analysis used in this report. Appendix C provides more detail on the analysis methods.

4.1 Production properties

The analysis methods referred to in Section 3.1 indicate a number of key properties of the production – mean MoE, 5%ile MoE and 5%ile strength[2]. The use of the outputs from this study also requires an estimate of the Coefficient of Variation (CoV) of the production properties and the number of specimens tested in the sample.

4.1.1 Mean

In monitoring properties of graded timber, an estimate of the mean MoE of the production batch is required.

The mean of some data is simply the arithmetic average of the data[1]. However, where we are interested in the mean of the entire production of a batch it is not generally possible to test all of the pieces in the batch, so an estimate of the mean must be made.

- This can be done by averaging the same property of a random sample from the batch.
- In some cases, a distribution fitted through the random sample test data can be used to estimate the mean of the batch.

4.1.2 5%ile

In monitoring properties of graded timber, an estimate of the 5%ile MoE and the 5%ile strength of the production batch may be required.

The 5%ile of some data is simply the data point that has 5% of the data below it and 95% of the data above it. Again, it is not possible to test all of the pieces in the batch, so an estimate of the 5%ile must be made.

- This can be done by ranking, and plotting the same property of a random sample from the batch and estimating a 5%ile from a cumulative frequency distribution.
- In some cases, a distribution (such as log-normal, or 2 parameter Weibull) can be fitted through the random sample test data and used to estimate the 5%ile of the batch.

4.1.3 Coefficient of Variation

In order to determine the Test Comparison Values using the tables or equations presented in this report, the CoV of the relevant production property must be estimated, and it is taken from a longer-term evaluation.

The definition of Coefficient of Variation (CoV) is the standard deviation divided by the mean and expressed as a percentage. For a production batch, neither of these values is known explicitly, so the Coefficient of Variation must be estimated from some test results on the production.

Typically, more than one hundred recent test results for the particular property would be used to provide a more reliable estimate of the CoV over the longer-term. (The CoV from a single test sample tends to be influenced by the sampling error for smaller sample sizes.) It is impossible to be definitive, but the intention is that the longer-term data represent a number of different production batches over a few months. Some alternatives for estimating longer-term CoV, depending on the sample size used, and the frequency of the production runs, are:

- At the commencement of a test program, the data from a previous Periodic Monitoring program, or data from similar products could be used as a starting point.
- The previous 200 to 400 random test data points for the size-grade-mill combination could be used. (Equivalent to a rolling data set of 200 to 400 points.)
- A rolling data set based on the test data from the previous 5 to 10 production runs.
- The test data for the size-grade-mill combination taken in the previous 2 to 3 months could be used.

An example of monitoring longer-term CoV of properties is given in Section 6.2.4.

Where proof testing is performed, the CoV used in evaluating Test Comparison Values is still the CoV of the whole of the production batch. This can be estimated from a distribution through the tail for which we have data. For 2 parameter Weibull tail fits, the CoV is usually less than the CoV estimated from other distributions, or the full data set.

- Where a 2 parameter Weibull distribution is used to find the 5%ile from a tail fit, the CoV estimated by the 2 parameter Weibull distribution tail fit is less than the long-term population CoV for the whole data. However, for this case only, a long-term estimate of the CoV from tail fitting should be used to find the Test Comparison Values.
- Where a log-normal distribution or non-parametric methods are used to find the 5%ile from a tail fit, the CoV estimated from a large number of test results for the product should be used to find the Test Comparison Values.

4.1.4 Number of specimens

The number of pieces in the random sample is also used to find the Test Comparison Values. In this study, the number of specimens is tied to the definition of a batch. The number of pieces (n) taken as a random sample during the time in which a single batch was produced:

- A single batch is the material produced during the period covered by a single analysis. (For example, if an analysis is performed after 3 shifts, then the entire production of those three shifts is a single batch. The analysis gives estimates of the properties of the batch based on the samples taken while the batch was in production.)
- Where proof testing is performed, it is the number of specimens taken (as distinct from the number of pieces broken.)

4.2 Methods of analysing test data

This section contains some brief information on the analysis methods that were used in the study for deriving properties from the test data. These methods are currently used in the industry to estimate production properties from random position tests on random samples of material. Further information on these methods is detailed in C.2. Their presence in the report does not imply that they are all recommended, it simply reflects the methods used in the industry.

The various methods are illustrated using some test data on bending test results from a single MGP grade. The same data (41 points) is used as the basis of Figure 1 through to Figure 5.

The 5%ile and means estimated from the analyses of the same data are as follows:

Table 1 – Estimations of bending strength using different analysis methods

Analysis method	Estimate of 5%ile	Estimate of mean
Non-parametric	19.0	37.5
Log-normal fit to all data	18.8	37.9
Log-normal fit to all data with CoV 40%	18.6	38.0
Log-normal fit to tail data	17.4	
Log-normal fit to tail data with CoV 40%	17.6	
2 parameter Weibull fit to tail data	17.7	

It can be seen that all of the estimates show a mean of around 38 MPa and a 5%ile of around 17 to 19, but the reality is that they are just estimates based on the particular samples selected, and we do not know the actual mean or 5%ile bending strength of the production. The example shows how the same data can give different estimates of the character of the material depending on the analysis method.

4.2.1 Non-parametric methods

This method uses the raw data from the test results without fitting any distributions or curves through it.[4,5]

- The estimation of the mean is found by finding the arithmetic average of all of the test points. All of the data is used to find this value.
- The estimation of the 5%ile is found by ranking the test results and linearly interpolating between the points either side of the 5th percentile. In effect, it is two test results only that define this estimate.

Figure 1 shows some test data for strength with the 5%ile value marked as a red line. The mean of this data is 37.5 MPa and is near the middle of the distribution. The 5%ile is 19.0 MPa and is found by interpolating between the 2nd and 3rd points in the cumulative frequency distribution.

It can be seen in Figure 1 that the data is not a smooth curve, and due to random sampling, there are ‘bumps’ in the distribution. While the average value is not significantly affected by these ‘bumps’, the 5%ile value can be significantly changed by localised irregularities in the data.

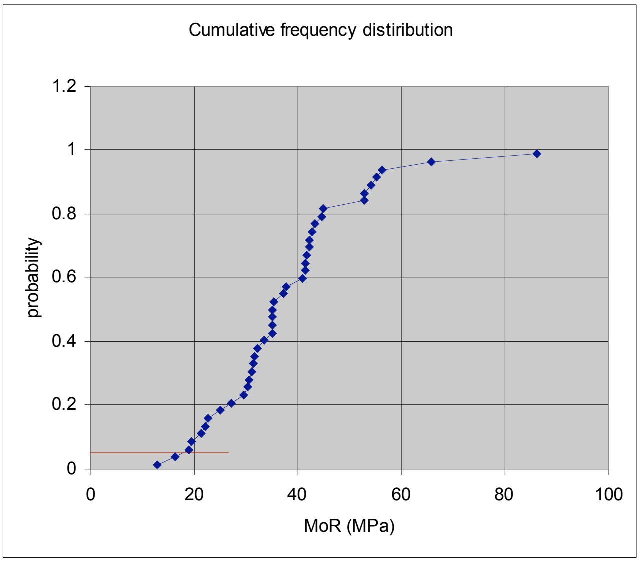


Figure 1 – Non-Parametric analysis of the raw test data.

4.2.2 Log-normal fit to the full data set

This method fits a log-normal distribution through all of the test data[1]. The shape of the curve fitted is given by the log-normal distribution, and its position along the horizontal axis by the average of the data. It is stretched to fit the range of the data.

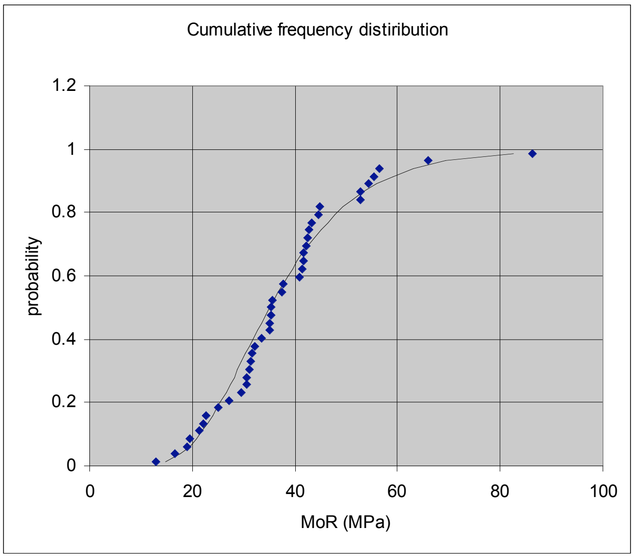


Figure 2 – Log-normal fit to the full raw test data.

The method involves fitting a normal distribution to the natural logs of the test data[6] using least squares fitting. The mean and the 5%ile are found from the fitted log-normal distribution rather than the raw data itself. The fit is illustrated in Figure 2. The black line through the data is the log-normal distribution fitted through the data.

The log-normal distribution through the data has a mean of 37.9 MPa and a 5%ile of 18.8 MPa. The mean is very close to the arithmetic average of the test data, but the

5%ile value was around 0.2 MPa lower than the non-parametric estimate of the 5%ile. Note that the discrepancy between the lognormal fit and the data is greatest at around the 25%ile. Had that irregularity occurred at a lower strength, then there may have been a very significant difference between the 5%ile estimated by non-parametric methods and that estimated using the log-normal fit through the data.

4.2.3 Log-normal fit to the data using an estimated population CoV.

The log-normal fit outlined in Section 4.2.2 uses the range of the test data to adjust the spread of the S-curve. However, particularly for small sample sizes, the spread of the data in the sample may not reflect the spread of the properties in the production it represents. The discrepancy is due to the random sampling.

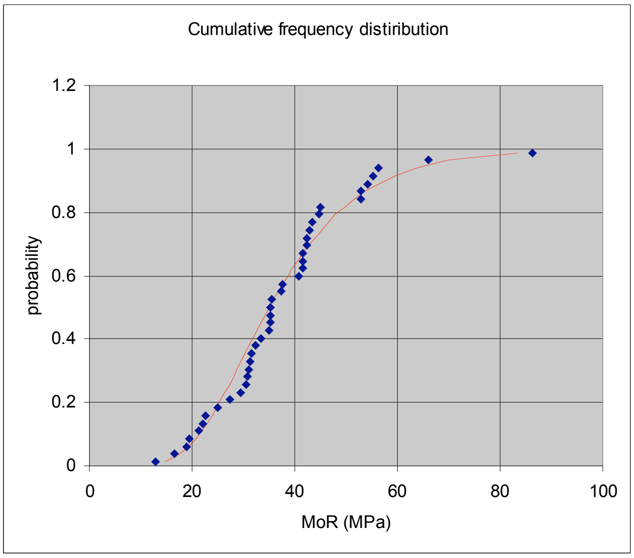


Figure 3 – Log-normal fit to the full raw test data with given CoV

However, the CoV of the production is the CoV of thousands of pieces of untested timber. This can best be represented by the CoV of a large number of tests of the same product. This is found from the longer-term characteristics of the product (derived from several months' data), and the longer-term CoV of the test data can be used to set the spread of the S-curve for the properties of the production.

Hence, the test data from a single batch is used to find the mean of the log-normal distribution, and this can be approximated to the mean of the production of the same product. The longer-term variation (given by the CoV of a large number of test data) is approximating the CoV of the production during the batch and the 5%ile strength is calculated from the mean of the production and the estimated CoV.

The sources of error here are the estimation of production mean – a function of sampling error, and the error in estimation of the CoV. The error in estimation of the CoV is a function of the variability of the CoV over time.

For example, the CoV of the data used as the basis of Figures 1 and 2 is 37.5%, and derived from this set of test data. However, over the long-term, the production of this grade in this mill seems to have a CoV of 40%. Hence this CoV can be used to model the test data with more reliability, as the CoV of the test sample is also affected by sampling error. Figure 3 shows the result.

The difference between Figure 2 and Figure 3 is fairly subtle. In Figure 3, the S-curve is slightly more stretched than in Figure 2, and it still seems to reflect the data over the full length of the curve. There is a slight difference in the average value and in the 5%ile value from this curve. The mean value returned by the distribution rose 0.5 MPa from the non-parametric mean to 38.0 MPa and the 5%ile value dropped 0.4 MPa to 18.6 MPa.

This method is less sensitive to points at the extremities of the data. These points are low probability points, and can have more influence on outcomes for small sample sizes than they would on larger samples.

The actual CoV of the production batch may vary slightly from time to time, and the true value can only be estimated some time after the event. Hence there may be a slight discrepancy between the CoV of the production estimated from long-term CoV and the actual production batch CoV at the time. This discrepancy introduces another source of error in the estimation of production properties and this has been incorporated in the following way. Two different variations in CoV have been used to estimate confidence limits:

- **Tight CoV** is long-term CoV varying by no more than 5% of the average longer-term value. In other words, the CoV of the batch could be expected to lie between 0.95 times the longer-term CoV and 1.05 times the longer-term CoV.
- **Loose CoV** is long-term CoV varying by no more than 10% of the average longer-term value. In other words, it could be expected to lie between 0.90 times the longer-term CoV and 1.10 times the longer-term CoV. This greater variability in CoV leads to more discrepancy in the estimated strength, so the analysis returns a higher Test Comparison Value to compensate.

If there is doubt about the variability of the CoV, the appropriate choice is “Loose CoV”, but if records of variation of longer-term CoV (averaged over a number of months) indicate the lower variability, then “Tight CoV” can be used. The definition and information on the estimation of CoV was given in Section 4.1.3.

4.2.4 Tail fitting

As it is only the 5%ile strength value that has significance in assessment of the strength of the production, there is a logic in using only values close to the 5%ile to fit a distribution that will best reflect the character of the data in that area – the lower tail of the distribution.[7]

Whereas Figures 2 and 3 fit a distribution using all of the data from the tests, Figure 4 shows fits that are only made to the lower tail of the distribution (probability = 15% or less). The points used to fit the distribution are shown filled and the others are shown open. Two different distributions have been fitted – a log-normal distribution (red) and a 2 parameter Weibull distribution (green).

For Figure 4, only the lower 15% of the data has been used to fit the curves, but in reality we know what the remainder of the test data is. It shows that the two fitted distributions follow the lower tail of the distribution really well, but when compared to the rest of the test data, they can be quite a long way from the higher strength data.

Sampling errors in the lower tail may mean that the fitted distribution may not represent the full distribution of production properties particularly accurately.

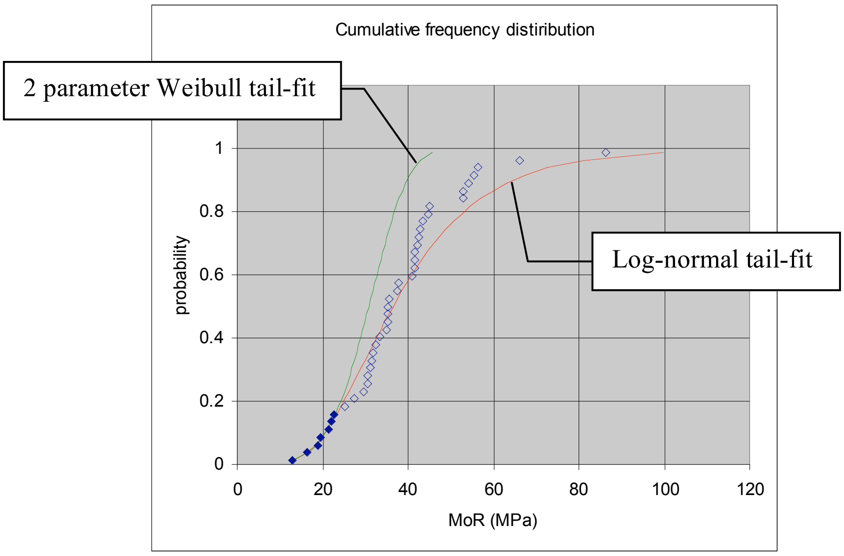


Figure 4 – Log-normal and 2 parameter Weibull fit to the tail test data

There are other variations to tail fitting that can also be implemented. The fitted distribution can be constrained to have a CoV that is expected for the production property. Figure 5 shows the result of this with the tail fit log-normal distribution with a CoV of 40% shown as a dotted red line and the tail fit 2 parameter Weibull distribution with an equivalent CoV (see Section C.4.5).

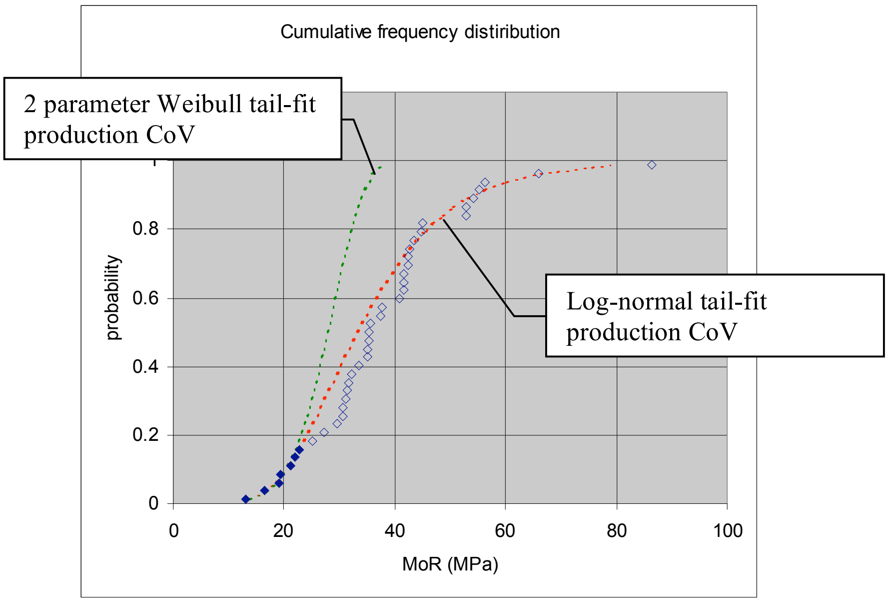


Figure 5 – Log-normal and 2 parameter Weibull fit to the tail with production CoV

The use of the estimated production CoV meant that the fitted distributions moved closer to the whole of the test data. However, the 2 parameter Weibull distribution fitted only through the tail data has a curve that is consistently to the left of the upper part of the data.

4.2.5 2 parameter Weibull tail fit to ISO 13910

ISO 13910[18] contains an informative appendix that gives a statistical method for fitting a 2 parameter Weibull distribution through the tail of test data, based on a minimum tail size of the larger of 15% of the sample size or 15 pieces. However in the analysis, the lower two points are discarded.

This method still returns a good fit for the tail of the distribution, but again is consistently to the left of the upper test values.

4.2.6 Goodness of data fit

Whichever distribution is used to fit the data, it is necessary to ensure that it matches the character of the data[8]. Normally this would be checked using an amalgamation of data from a number of batches to minimise the effect of sampling error.

Some analysis methods showed less sampling error than others. Different analysis methods also produced systematic over- or under-estimation of properties. This is discussed in Section C.4.1. The success of different analysis methods is summarised in Section 5.5.

5. Outputs

This section provides a description of the outputs of the project. The outputs of each analysis method are used to derive Test Comparison Value (*TCV*) for use in providing feedback from property monitoring programs.

5.1 Test Comparison Values

A Test Comparison Value (*TCV*) is the minimum value for a test result in order for there to be the designated confidence that the property from the production represented by the sample is above the Design Value (*DV*).[2,9]

- Where the test result is greater than the Test Comparison Value (*TCV*), then the producer has greater than the nominated confidence that the production meets Design Values.
- Where the test result is less than the Test Comparison Value (*TCV*), then the producer has less than the nominated confidence that the production meets Design Values.

The Test Comparison Values are found from the Design Value and multipliers given in this document.

Other points may also prove significant in the acceptance or rejection of production batches[6]:

- Where the test result is less than the Design Value, then the producer has less than 50% confidence that the production meets the Design Value.
- Where the test result is less than ($2DV - TCV$), then there is greater than the nominated confidence that the production DOES NOT meet the Design Value.

5.1.1 Test Comparison multiplier *M*

The Test Comparison Value (*TCV*) is simply found by multiplying the relevant Design Value (*DV*) by the Test Comparison multiplier (*M*).

$$TCV = DV \times M \qquad \textit{(eqn 5.1)}$$

Also

$$L = X\left(1 + A\frac{CoV}{\sqrt{n}}\right) \qquad \textit{(eqn 5.2)}$$

Where L = lower confidence limit in the evaluation of the property X, and A will be negative.

Eqn 5.2 is the basis of estimating confidence limits, and where X is a critical value (e.g. mean) of a distributed population, A is simply the probability points for the fitted distribution. (*Eqn 5.2*) has been used as the basis of the calibration of the method and the findings against classical statistical theory for mean values and some codified values for 5%iles. (Refer C.4.3 for more detail on the derivation.)

The definition of Test Comparison Value (*TCV*) is that where the test result (X) equals , the Test Comparison Value (*TCV*), then there is the required confidence that the production has the Design Value (DV). So in (*eqn 5.2*), by substituting the test value (X) with the Test Comparison Value (TCV) and setting the Design Value (DV) as the Lower Confidence Limit gives (*eqn 5.3*).

$$DV = TCV\left(1 + A\frac{CoV}{\sqrt{n}}\right) \qquad \textit{(eqn 5.3)}$$

Where A is a function of the test and analysis methods and the confidence level required (a negative value).
DV is the design value from AS1720.1
TCV is the Test Comparison Value as defined above.

This is the form of (*eqn 5.2*) that appears in AS/NZS4063 with the confidence level as 75% and the test result (X) = the Test Comparison Value (TCV).

(*Eqn 5.3*) can be rewritten for use in monitoring properties where the Design Value is known, and Test Comparison Values need to be set. This expression is given as (*eqn 5.4*).

$$TCV = DV\left(\frac{1}{1 + A\frac{CoV}{\sqrt{n}}}\right) \qquad \textit{(eqn 5.4)}$$

$$\text{giving } TCV = DV \times M \qquad \text{with} \qquad M = \left(\frac{1}{1 + A\frac{CoV}{\sqrt{n}}}\right) \qquad \textit{(eqn 5.5)}$$

For evaluating mean MoE from the average MSG values for the grade, a slightly different equation (eqn 5.6) is used.

$$\text{From MSG} \qquad TCV = DV \times M \qquad \text{with} \qquad M = \left(\frac{1}{B\left(1 + A\frac{CoV}{\sqrt{n}}\right)}\right) \qquad \textit{(eqn 5.6)}$$

Values of M are presented in Tables in Appendices A and B, and are used to calculate the Test Comparison Value (TCV) using (*eqn 5.1*).

A is presented as values that enable the calculation of M and hence TCV for any values of CoV or n using (*eqn 5.4*), (*eqn 5.5*) or (*eqn 5.6*).

5.2 Tables

The tables presented in Appendix A and B can be used to find M directly. Once M is known, a simple multiplication can be used to find the Test Comparison Value (TCV) using (*eqn 5.1*).

$$TCV = DV \times M \qquad \text{(copy } \textit{eqn 5.1)}$$

The tables all have a similar format:

- The heading of each table (across the top line of the table) indicates the method of analysis and the property being assessed.
- The confidence level for all of the information in the table is given in the second row of the table
- Each heading is shaded with a colour to highlight the Confidence level used for all of the data in the plot. (Blue 95%, Green 90%, Yellow 85%, Red 80%, White 75%)
- A number of different population CoVs are shown across each table (Population CoVs can be estimated from longer-term accumulation of test data as indicated in Section 4.1.3.) It is permissible to interpolate within a table between CoVs, however, the equations can be used to derive Test Comparison Values for CoVs not shown in the Tables.
- A number of different sample sizes are presented in each table. It is permissible to interpolate within a table between sample sizes, however, the equations can be used to derive Test Comparison Values for sample sizes not shown in the Tables.
- The numbers shown in the body of the table are values of *M* and can be used in (*eqn 5.1*) to find the Test Comparison Value directly.

The following tables have been extracted from Appendix A and B for illustration purposes only.

Non Parametric Mean MoE					
CL 95%	**CoV**				
No. Sample	**8%**	**10%**	**12%**	**15%**	**20%**
5	1.063	1.080	1.097	1.124	1.173
10	1.044	1.055	1.067	1.085	1.116
20	1.030	1.038	1.046	1.059	1.080
30	1.025	1.031	1.037	1.047	1.064
50	1.019	1.024	1.029	1.036	1.049
100	1.013	1.017	1.020	1.025	1.034
200	1.009	1.012	1.014	1.018	1.024

Log-normal 5%ile Strength (All data)								
CL 95%	**CoV (%)**							
No. Sample	**5%**	**10%**	**15%**	**20%**	**25%**	**30%**	**35%**	**40%**
5	1.063	1.135	1.217	1.312	1.423	1.554	1.713	1.907
10	1.044	1.092	1.144	1.202	1.266	1.337	1.417	1.507
20	1.031	1.063	1.098	1.135	1.175	1.217	1.263	1.312
30	1.025	1.051	1.079	1.108	1.138	1.170	1.205	1.241
50	1.019	1.039	1.060	1.081	1.104	1.127	1.152	1.177
100	1.013	1.027	1.042	1.056	1.071	1.087	1.103	1.119
200	1.009	1.019	1.029	1.039	1.049	1.060	1.070	1.081

5.2.1 MoE

Tables have been generated for the following:

- Non-parametric mean MoE *(Appendix A.1.1)*
- **Log-normal mean MoE** *(Appendix A.1.2)*
- **Mean MoE from MSG data** *(Appendix A.1.3)*
- Non-parametric 5%ile MoE *(Appendix A.2.1)*
- **Log-normal 5%ile MoE** *(Appendix A2.2)*

Analysis methods shown in bold give the lowest sample error (see Section 5.5).

5.2.1 5%ile bending strength and 5%ile tension strength

Tables have been generated for the following:

- Non-parametric 5%ile strength *(Appendix B.1)*
- **Log-normal 5%ile strength (All data)** *(Appendix B.2)*
- **Log-normal 5%ile strength from mean (Tight CoV)** *(Appendix B.3)*
- **Log-normal 5%ile strength from mean (Loose CoV)** *(Appendix B.4)*
- Log-normal 5%ile strength (Tail fit) *(Appendix B.5)*
- Log-normal 5%ile strength (Tail fit, Tight CoV) *(Appendix B.6)*
- Log-normal 5%ile strength (Tail fit, Loose CoV) *(Appendix B.7)*
- 2 parameter Weibull 5%ile strength (Tail fit) *(Appendix B.8)*
- **2 parameter Weibull 5%ile strength (Tail fit to ISO13910)** *(Appendix B.9)*

Analysis methods shown in bold give the lowest sample error (see Section 5.5).

Tables that include Tight CoV (where it is believed that the CoV at any time is within ± 5 % of the long-term estimate of Production CoV) or Loose CoV (where it is believed that the CoV at any time is within ± 10 % of the long-term estimate of Production CoV) have been provided for use with products of differing variability in CoV of the product.

5.3 Equations

(*Eqn 5.5*) can be used with the values of *A* presented in Appendix A and Appendix B to determine the Test Comparison multiplier *M* used to find the Test Comparison Value (*TCV*).

$$TCV = DV \times M \qquad \text{with} \qquad M = \left(\frac{1}{1 + A\frac{CoV}{\sqrt{n}}} \right) \qquad \text{(copy \textit{eqn 5.5})}$$

Equation
Log-normal 5%ile Strength (all data)

CL	Value of A
95%	-2.659

This is the general form of the equation presented and used throughout this document and has the following conventions:

- Each equation makes use of a constant A which is used for the given analysis method and confidence level.
- The heading of each equation indicates the method of analysis and the property being assessed.
- Each value of A is shaded with a colour to highlight the Confidence level used for all of the data in the plot. (Blue 95%, Green 90%, Yellow 85%, Red 80%, White 75%)

These equations can be used to evaluate M (and hence TCV) for any combination of population CoV and sample size (n).

5.4 Plots

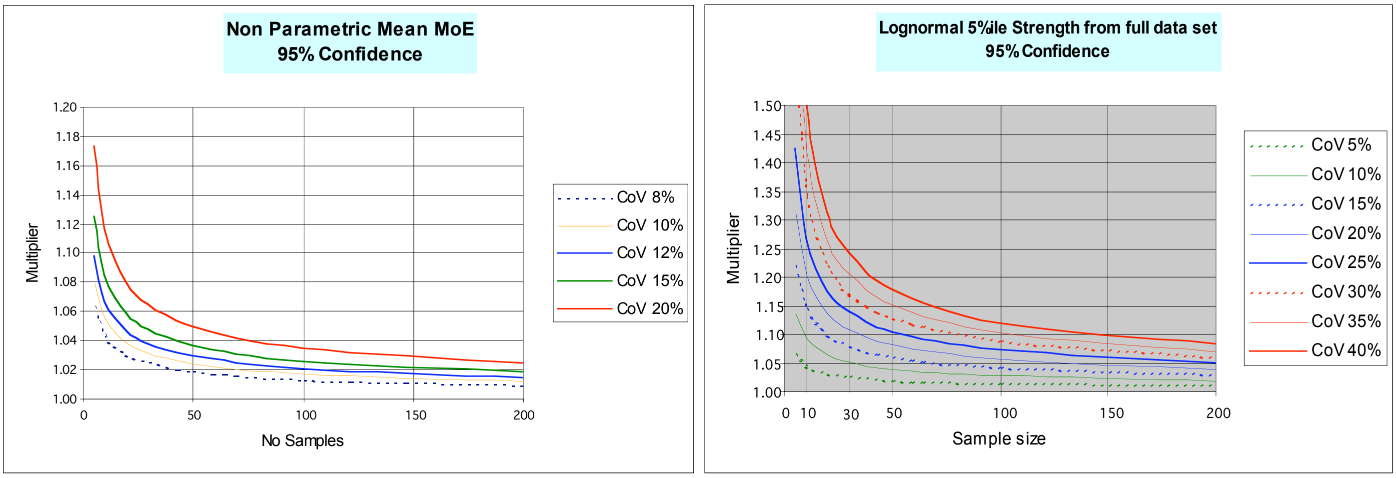
The graphs presented in Appendices A and B provide a visual representation of the values of M given in the Tables.
The graphs all have a similar format:

- The heading of each plot indicates the method of analysis and the property being assessed. The confidence level for all of the information on the plot is also given in the heading.
- Plots of M for MoE have a plain background, and plots of M for strength have a grey background.
- Each heading is shaded with a colour to highlight the Confidence level used for all of the data in the plot. (Blue 95%, Green 90%, Yellow 85%, Red 80%, White 75%)
- A number of different population CoVs are shown on each plot using different colour, strength or pattern lines to differentiate the curves. Higher CoVs are the uppermost curves. (Population CoVs can be estimated from longer-term accumulation of test data.)
- The horizontal axis is the number of specimens in each sample (sample size).
- The vertical axis shows the value of M for a given sample size, and population CoV.

All graphs show similar trends:

- The smaller the size of the sample, the larger the multiplier, and hence the larger the Test Comparison Value required for a given level of confidence. Sample sizes of 30 or more seem to have a lower slope. Increasing the sample size a little does not have a significant effect on M. However for sample sizes of less than 30, the plots have a high slope. Smaller sample sizes result in a significantly larger value of M.
- The larger the CoV of the product, the larger the multiplier, and hence the larger the Test Comparison Value required for a given level of confidence.
- The larger the level of Confidence required, the larger the multiplier, and hence the larger the Test Comparison Value required.

Figure 6 shows a graph of M for evaluation of MoE and for evaluation of strength.



(a) M for MoE from non-parametric average

(b) M for 5%ile strength from log-normal dist – all data

Figure 6 – Plots of M for different testing and analysis methods

While it is possible to determine the value of M from the graphs in Appendix A and Appendix B, it is more accurate to use either the equations to find M from a given value of A, sample size and population CoV, or by interpolation within the tables for M.

5.5 Comparison of Analysis methods and test strategies

A number of different test strategies and analysis methods have been outlined in Section 4.2. The outputs for each of these methods have been presented separately in Appendix A and Appendix B, but in this section, a summary of the results is presented in a single Table to enable a comparison to be made.

In (eqn 5.5) it is the value of A that indicates the estimation error. Smaller values of A imply smaller estimation errors. (This leads to a Test Comparison Value (TCV) that is closer to the Design Value, and will maximise recovery into the grade.)

Table 2 – Comparison of Analysis methods and test strategies

(Values of *A* from Appendix A for MoE in (*eqn 5.5*))

	95%	90%	85%	80%	75%
Non-parametric (Mean MoE)	-1.649	-1.290	-1.045	-0.854	-0.686
Log-normal (Mean MoE)	-1.657	-1.297	-1.052	-0.861	-0.693
MSG estimate of Mean MoE	-1.645	-1.282	-1.036	-0.842	-0.674
Non-parametric 5%ile MoE	-3.747	-3.110	-2.672	-2.325	-2.024
Log-normal 5%ile MoE	-2.682	-2.121	-1.737	-1.438	-1.178

(Values of *A* from Appendix B for 5%ile strength in (*eqn 5.5*))

	95%	90%	85%	80%	75%
Non-parametric	-3.698	-3.072	-2.651	-2.309	-2.021
Log-normal	-2.659	-2.104	-1.731	-1.431	-1.172
Log-normal from mean (Tight CoV)	-2.166	-1.806	-1.564	-1.372	-1.204
Log-normal from mean (Loose CoV)	-2.691	-2.331	-2.089	-1.896	-1.728
Log-normal from mean tail fit	-2.977	-2.385	-1.981	-1.659	-1.383
Log-normal from mean (Tail fit, tight CoV)	-2.423	-1.998	-1.713	-1.486	-1.297
Log-normal from mean (Tail fit, loose CoV)	-2.951	-2.524	-2.239	-2.011	-1.821
2 parameter Weibull tail fit	-6.295	-5.084	-4.286	-3.644	-3.083
2 parameter Weibull tail fit (ISO13910)	-3.106	-2.419	-1.949	-1.578	-1.260

It can be seen that for some analysis methods, there is a significant difference in values of *A*.

- The lowest values of *A* for finding the mean MoE were for the estimate using the average MSG of the production. A further advantage of this method is that the value of *n* used is very high. It gives consistently lower Test Comparison Value (*TCV*) than the other two methods.
- For estimating average MoE from samples there was effectively no difference between the non-parametric estimate and the estimate using the log-normal distribution.
- For estimating the 5%ile MoE, the log-normal distribution gives a lower value of *A* than the non-parametric estimate, for all confidence levels.
- In estimating the 5%ile strength, it is clear that the non-parametric estimate will lead to larger values of *A* and hence higher Test Comparison Values.
- In estimating the 5%ile using the 2 parameter Weibull tail fit, the smaller CoV of the tail of the 2 parameter Weibull distribution is used, but the Test Comparison Values will generally be higher than for fits through all of the data.
- A consistently low value of *A* was obtained for the estimation of the 5%ile strength using the log-normal fit from the mean with an estimated CoV that does not vary significantly over time (Tight CoV).
- Where a tail fit is to be used, the same analysis technique gave the best approximation.
- Where the CoV may vary in time more from the long-term population estimate of CoV (Loose CoV), then there is no advantage in using the production CoV. It is more effective to use the log-normal fit through the data.

6. Examples of use

This section illustrates the ways in which the outputs of this project can be used in monitoring timber properties and other aspects of structural timber production in Australia.

6.1 Selection of strategies for monitoring timber properties

In defining or refining a monitoring program, producers must make a variety of decisions. These decisions must reflect the corporate risk management strategy of the company, and also assist in the achievement of the aim for monitoring. In Section 1, this was stated as:

> *'to estimate whole of production properties from samples of production and use this information to control production parameters to deliver product with complying properties.'*

6.1.1 Example scenario

In this hypothetical example, a producer has decided to look at whether or not making changes to the property monitoring system will improve the profitability of the production. The current system is as follows:

- The producer produces MGP10 and MGP12, but does not separate an MGP15 product. Bending testing is used on each shift for both of the grades.
- On each shift, ten samples for each grade are drawn at even time intervals throughout the shift and subjected to a bending test. (A total of 20 tests are performed in each shift).
- The tests are conducted to a proof load of 1.25 times the 5%ile expected for the grade, and yield the average MoE for each piece and the MoR for each piece that fails at less than the proof load.
    - For MGP10, the expected 5%ile MoR is 20 MPa and the proof stress is therefore 25 MPa.
    - For MGP12, the expected 5%ile MoR is 32 MPa and the proof stress is 40 MPa.
- The company policy is to have 95% confidence in bending strength interpretations (roughly one wrong interpretation in 20 shifts), and 90% confidence in MoE interpretations (roughly one wrong interpretation in 10 shifts).
- The estimation of production MoE is from the arithmetic average of the 10 specimens, and the estimation of the production 5%ile bending strength is from the log-normal distribution through the tail of the strength data.
- The estimates of the CoV for the production of each grade are as follows:
    - Longer-term CoV of MoE for MGP10 = 15%
    - Longer-term CoV of MoE for MGP12 = 12%
    - Ratio of (min MSG/avg MSG) for MGP10 = 0.75
    - Ratio of (min MSG/avg MSG) for MGP12 = 0.79
    - Longer-term CoV of bending strength MGP10 = 40%
    - Longer-term CoV of bending strength MGP12 = 36%
    - Expected production of MGP10 = 10,000 per run
    - Expected production of MGP12 = 8,000 per run

The producer would like to see how much the Test Comparison Values could be reduced if there is:

- A change to test all specimens sampled to failure each shift,
- A change in the analysis method,
- An increase in the number of specimens taken per shift

A consequence of reduction in Test Comparison Values is that thresholds can be reduced and as a result, recoveries increase.

6.1.2 Current Test Comparison Values

Based on the current production methods and decisions, Test Comparison Values (*TCV*) can be calculated that will give the set confidence that each shift has properties that meet the design values (*DV*).

- If the test results for a product exceed the Test Comparison Value, then the producer has more than the set confidence in properties of the product.
- If the test results for a product are less than the Test Comparison Value, then the producer has less than the set confidence in the properties of the product.

MoE

For a non-parametric estimate of the mean MoE, and with 90% confidence, the Test Comparison Values can be set using the following equation:

$$TCV_E = DV_E \left(\frac{1}{1 - 1.290 \frac{CoV}{\sqrt{n}}} \right) \qquad \textit{(eqn 6.1)}$$

For MGP10, with CoV (on MoE) of 15%, and for a sample size of 10, the Test Comparison Value is based on the Design Value of 10,000 MPa. Hence from Appendix A1.1,

$$TCV_E = 10{,}000 \left(\frac{1}{1 - 1.290 \frac{0.15}{\sqrt{10}}} \right) = 10{,}000 * 1.065 = 10{,}650 \text{ MPa}$$

For MGP12, with CoV (on MoE) of 12%, and for a sample size of 10, the Test Comparison Value is based on the Design Value of 12,700 MPa. Hence

$$TCV_E = 12{,}000 \left(\frac{1}{1 - 1.290 \frac{0.12}{\sqrt{10}}} \right) = 10{,}000 * 1.051 = 13{,}350 \text{ MPa}$$

Bending strength

For an estimate of the 5%ile MoR, and with 95% confidence using a log-normal distribution fit to the tail of the data (and proof stress = ~1.25 times 5%ile strength), the Test Comparison Values can be set using the following equation with *A* from Appendix B.5:

$$TCV_R = DV_R \left(\frac{1}{1 - 2.977 \frac{CoV}{\sqrt{n}}} \right) \qquad \textit{(eqn 6.2)}$$

For MGP10, with CoV (on MoR) of 40%, and for a sample size of 10, the Test Comparison Value is based on the Design Value of 16 MPa. (Note that even though only one or two pieces may break, *n* is the number of specimens tested – not broken. It is still 10. Hence

$$TCV_R = 16 \left(\frac{1}{1 - 2.977 \frac{0.40}{\sqrt{10}}} \right) = 16 * 1.604 = 25.7 \text{ MPa}$$

For MGP12, with CoV (on MoR) of 36%, and for a sample size of 10, the Test Comparison Value is based on the Design Value of 28 MPa. Hence

$$TCV_R = 28 \left(\frac{1}{1 - 2.977 \frac{0.36}{\sqrt{10}}} \right) = 28 * 1.513 = 42.4 \text{ MPa}$$

In summary:
With the current monitoring system in place, in order to maintain the required confidence that each shift meets or exceeds the Design Values, then the test results should regularly exceed:

Test Comparison Value (current monitoring)	**MoE** (MPa)	**MoR** (MPa)
MGP10	10650	25.7
MGP12	13350	42.4

These Test Comparison Values are interpreted as follows:

- Where the test results after each shift are analysed and give estimates greater than the Test Comparison Values, then the confidence that the production meets or exceeds the design value is better than the required confidence.
- Where the test results after each shift are analysed and give estimates less than the Test Comparison Values, then the confidence that the production meets or exceeds the design value is lower than the required confidence.

For strength of both grades, the Test Comparison Values are both higher than the proof stress. This means that this type of testing cannot ever be successful in giving confidence that the production is meeting expectations.

6.1.3 Effect of testing all specimens

This will have no effect on MoE (as a result is already available for each specimen), but it will have an effect on MoR.

Bending strength
For an estimate of the 5%ile MoR, and with 95% confidence using a log-normal distribution fit to all of the data, the Test Comparison Values can be set using the following equation with *A* from Appendix B.2:

$$TCV_R = DV_E \left(\frac{1}{1 - 2.659 \frac{CoV}{\sqrt{n}}} \right) \qquad \textit{(eqn 6.3)}$$

For MGP10, with CoV (on MoR) of 40%, and for a sample size of 10, the Test Comparison Value is based on the Design Value of 16 MPa. Hence

$$TCV_R = 16 \left(\frac{1}{1 - 2.659 \frac{0.40}{\sqrt{10}}} \right) = 16 * 1.507 = 24.1 \text{ MPa}$$

For MGP12, with CoV (on MoR) of 36%, and for a sample size of 10, the Test Comparison Value is based on the Design Value of 28 MPa. Hence

$$TCV_R = 28 \left(\frac{1}{1 - 2.659 \frac{0.36}{\sqrt{10}}} \right) = 28 * 1.434 = 40.2 \text{ MPa}$$

In summary
This change will cause a small reduction in the strength TCV to give the required confidence that the shift had properties that exceeded the design values. There is no change to the MoE TCV.

Test Comparison Value (Break all pieces)	**MoE** (MPa)	**MoR** (MPa)	**Change (MoE)**	**Change (MoR)**
MGP10	10650	24.1	0.0%	-6.2%
MGP12	13350	40.2	0.0%	-5.2%

The reduction in MoR TCVs will make it easier to achieve the required confidence in strength and as a result there should be fewer occasions in which the monitoring indicates that there is a problem with strength.

The producer can then compare the improvement in recovery and success of the monitoring method as a result of the reduced Test Comparison Values with the cost of breaking every piece tested.

6.1.4 Effect of changing analysis methods

MoE
Already, the analysis method for MoE gives lowest possible Test Comparison Values and cannot be improved upon for systems in which the MoE is determined by sampling and testing. However, if the MoE is evaluated using the average value given by the Machine Stress Grader, then there can be an improvement in Test Comparison Value. Here *A* is taken from Appendix A.1.3

$$TCV_E = DV_E\left(\frac{1}{B\left(1-1.282\frac{CoV}{\sqrt{n}}\right)}\right) \quad \text{for 90\% CL} \qquad \textit{(eqn 6.4)}$$

With *B* and *CoV* found from the ratio of (minMSG/avgMSG) for each grade.

- For MGP10 with $\left(\frac{\min MSG}{avgMSG}\right) = 0.75$ and 10,000 pieces per run

$$B = 0.827 + \left(\frac{\min MSG}{avgMSG}\right)0.197 = 0.827 + 0.75\times 0.197 = 0.975$$

$$CoV = 0.377 - \left(\frac{\min MSG}{avgMSG}\right)0.334 = 0.377 - 0.75\times 0.334 = 0.127$$

$$TCV_E = 10{,}000\left(\frac{1}{0.975\left(1-1.282\frac{0.127}{\sqrt{10{,}000}}\right)}\right) = 10{,}270 \text{ average MSG}$$

- For MGP12 with $\left(\frac{\min MSG}{avgMSG}\right) = 0.79$ and 8,000 pieces per run

$$B = 0.827 + \left(\frac{\min MSG}{avgMSG}\right)0.197 = 0.827 + 0.79\times 0.197 = 0.983$$

$$CoV = 0.377 - \left(\frac{\min MSG}{avgMSG}\right)0.334 = 0.377 - 0.79\times 0.334 = 0.113$$

$$TCV_E = 12{,}700\left(\frac{1}{0.983\left(1-1.282\frac{0.113}{\sqrt{8{,}000}}\right)}\right) = 12{,}940 \text{ average MSG}$$

Bending strength

A small improvement can be made on the Test Comparison Values for strength by changing to an estimation of the 5%ile strength using a log-normal distribution with the CoV for the data analysis the value known to characterise the long-term production for each grade. Based on the current understanding of the CoV of production, and assuming that it is known to within 5% (Tight CoV), the following are the estimates of the CoV for strength. The method is valid providing the estimate of CoV for the analysis remains within the range defined by the following upper and lower bounds. If not, then the *A* values for Loose CoV should be used.

CoV for strength	estimate	Lower bound	Upper bound
MGP10	40.0%	38.0%	42.0%
MGP12	36.0%	34.2%	37.8%

For an estimate of the 5%ile MoR, and with 95% confidence using a log-normal distribution fit to all of the data and using the estimate of the population CoV, the Test Comparison Value can be set using the following equation with *A* from Appendix B.3:

$$TCV_R = DV_R\left(\frac{1}{1-2.166\dfrac{CoV}{\sqrt{n}}}\right) \qquad \textit{(eqn 6.5)}$$

For MGP10, with CoV (on MoR) of 40%, and for a sample size of 10, the Test Comparison Value is based on the Design Value of 16 MPa. Hence

$$TCV_R = 16\left(\frac{1}{1-2.166\dfrac{0.40}{\sqrt{10}}}\right) = 16 * 1.377 = 22.0 \text{ MPa}$$

For MGP12, with CoV (on MoR) of 36%, and for a sample size of 10, the Test Comparison Value is based on the Design Value of 28 MPa. Hence

$$TCV_R = 28\left(\frac{1}{1-2.166\dfrac{0.36}{\sqrt{10}}}\right) = 28 * 1.327 = 37.2 \text{ MPa}$$

In summary
This change (together with the testing of all specimens) will cause a more significant reduction in the strength Test Comparison Values to give the required confidence that the shift had properties that exceeded the design values. With the introduction of monitoring MoE using average MSG output, there is a small reduction in the MoE Test Comparison Values.

Test Comparison Value (MoR from ln & CoV) (MoE from avg MSG)	MoE (MPa)	MoR (MPa)	Change (MoE)	Change (MoR)
MGP10	10270	22.0	-3.5%	-14.4%
MGP12	12940	37.2	-3.2%	-12.3%

The use of the average MoE from the MSG to monitor product average MoE has much greater statistical significance than the use of samples because of the very large number of specimens used. Essentially sampling error is removed from the estimate but there is still a little uncertainty associated with the conversion of average MSG to average MoE. The resulting confidence will enable very fine control over average MoE from batch to batch.

The reduction in MoR Test Comparison Values of around 15% will make it easier to achieve the required confidence in strength and as a result there should be fewer occasions in which the monitoring indicates that there is a problem with strength. However, where the production appears to be strength limited, the high sampling error of strength compared with MoE will still cause production parameters (such as grade thresholds and visual over-ride limits) to be controlled by strength testing results.

6.1.5 Effect of increasing number of specimens taken each shift

If it is not possible to monitor the average MoE from the MSG and where the resource has very high strength, none of the changes discussed in 6.1.3 and 6.1.4 will enable the production thresholds to be reduced. However, by increasing the number of samples tested in each shift, there will be an improvement in the Test Comparison Values for both MoE and bending strength.

Where the production is strength limited, any further improvement in performance of strength testing will lead to an increase in grade recoveries.

Increasing the number of specimens tested in each shift increases production costs, but may enable recoveries to improve through reduced Test Comparison Values and hence thresholds. Two scenarios will be addressed:

- Doubling sample rate to 20 per shift
- Trebling sample rate to 30 per shift

MoE

In this case, we are assuming that MoE is still monitored by sampling and testing. The equation for determining the MoE Test Comparison Value is still given by *(eqn 6.1)*. This can be evaluated for the two scenarios.

	20 samples per shift	30 samples per shift
MGP10	$TCV_E = 10{,}000\left(\dfrac{1}{1-1.290\dfrac{0.15}{\sqrt{20}}}\right)$ $= 10{,}000 * 1.045 = 10{,}450$ MPa	$TCV_E = 10{,}000\left(\dfrac{1}{1-1.290\dfrac{0.15}{\sqrt{30}}}\right)$ $= 10{,}000 * 1.037 = 10{,}370$ MPa
MGP12	$TCV_E = 12{,}700\left(\dfrac{1}{1-1.290\dfrac{0.12}{\sqrt{20}}}\right)$ $= 12{,}700 * 1.036 = 13{,}160$ MPa	$TCV_E = 12{,}700\left(\dfrac{1}{1-1.290\dfrac{0.12}{\sqrt{30}}}\right)$ $= 10{,}000 * 1.029 = 13{,}070$ MPa

Bending strength

In this case, we are assuming that the bending strength is estimated by testing all specimens to destruction, and estimating the 5%ile by fitting a log-normal distribution with the CoV from long-term production data (as detailed in 6.1.4). The equation for determining the strength Test Comparison Value is still given by *(eqn 6.5)*. This can be evaluated for the two scenarios:

	20 samples per shift	30 samples per shift
MGP10	$TCV_R = 16\left(\frac{1}{1-2.166\frac{0.40}{\sqrt{20}}}\right)$ $= 16*1.240 = 19.8$ MPa	$TCV_R = 16\left(\frac{1}{1-2.166\frac{0.40}{\sqrt{30}}}\right)$ $= 16*1.188 = 19.0$ MPa
MGP12	$TCV_R = 28\left(\frac{1}{1-2.166\frac{0.36}{\sqrt{20}}}\right)$ $= 28*1.211 = 33.9$ MPa	$TCV_R = 28\left(\frac{1}{1-2.166\frac{0.36}{\sqrt{30}}}\right)$ $= 28*1.166 = 32.6$ MPa

In summary
Increasing the sampling rate in each shift (together with the testing of all specimens) will cause a significant reduction in the strength Test Comparison Values to give the required confidence that the shift had properties that exceeded the design values. There is also a reduction in the MoE Test Comparison Values.

Test Comparison Value (20 pieces)	**MoE** (MPa)	**MoR** (MPa)	**Change (MoE)**	**Change (MoR)**
MGP10	10450	19.8	-1.9%	-23.0%
MGP12	13160	33.9	-1.4%	-20.0%

Test Comparison Value (30 pieces)	**MoE** (MPa)	**MoR** (MPa)	**Change (MoE)**	**Change (MoR)**
MGP10	10370	19.0	-2.6%	-26.1%
MGP12	13070	32.6	-2.1%	-23.1%

The implication of this change is that fewer resources will be strength limited in production under this testing because of the significant reduction in strength Test Comparison Value. In addition, this mill will be able to lower thresholds by up to 300 MPa from the value required in the original monitoring scheme. This is because the MoE Test Comparison Value has reduced by that amount.

6.1.6 Summarising the options in this example

This example shows how the Test Comparison Value are affected by making changes to the monitoring system.

Changes to the monitoring system (sample size, test philosophy, analysis method) lead to

⇒ Changes to the Test Comparison Values. (These are the values that the test results have to exceed to give the required confidence in product properties.) Where the Test Comparison Values are close to the design value, there is better quality information on the product properties and the feedback between product properties and production values is better. This allows

⇒ Changes of grade thresholds or other parameters used in production. These in turn

⇒ Change the recovery in each grade.

Table 3 –Test Comparison Values for different Monitoring systems

Analysis method	Grade	MoE (MPa)	MoR (MPa)	Change (MoE)	Change (MoR)
Current System		Test Comparison Values		(Change/original)	
MoR – 10 pieces ln tail	MGP10	10650	25.7		
MoE – 10 pieces avg	MGP12	13350	42.4		
Change – break all pieces					
MoR – 10 pieces ln full	MGP10	10650	24.1	0.0%	-6.2%
MoE – 10 pieces avg	MGP12	13350	40.2	0.0%	-5.2%
Change – best analysis					
MoR – 10 pieces ln CoV	MGP10	10270	22.0	-3.5%	-14.4%
MoE – from avg MSG	MGP12	12940	37.2	-3.2%	-12.3%
Change – sample 20					
MoR – 20 pieces ln CoV	MGP10	10450	19.8	-1.9%	-23.0%
MoE – 20 pieces avg	MGP12	13160	33.9	-1.4%	-20.0%
Change – sample 30					
MoR – 30 pieces ln CoV	MGP10	10370	19.0	-2.6%	-26.1%
MoE – 30 pieces avg	MGP12	13070	32.6	-2.1%	-23.1%

It can be seen that changing the monitoring system by breaking all of the pieces made an improvement in the Test Comparison Values for strength, but did not affect the Test Comparison Values for MoE.

By incorporating both the breaking of all pieces in the sample, and a better analysis method, a significant reduction was made in the Test Comparison Values for both MoR and MoE.

Further improvements in the Test Comparison Values for strength could be achieved by further increasing the sample size.

6.2 Setting and using Test Comparison Values

6.2.1 Example scenario

In this different hypothetical example, a producer has decided on a Monitoring system and needs to set it up, use it and then monitor its operation. The system is as follows:

- The producer produces MGP10 and MGP12 and MGP15, but this example will simply follow the use of the system for MGP10 timber. Bending testing is used on each shift for all of the grades.
- On each shift, twenty samples of MGP10 are drawn at even time intervals throughout the shift and subjected to a bending test. (Smaller sample sizes are used for the other grades).
- The tests are conducted to failure and yield the average MoE and the MoR for each piece.
- The company policy is to have 80% confidence in both bending strength and MoE interpretations (roughly one wrong interpretation in 5 shifts – either indicated that it complies when it actually doesn't or indicating that it is non-compliant when it actually is).

- The estimation of production MoE for MGP10 is from the non-parametric mean of the 20 specimens, and the estimation of the production 5%ile bending strength is from the 5%ile of the log-normal distribution fitted through strength data.
- The estimates of the CoV for the production of MGP10 are as follows:
  - CoV MoE for MGP10 = 17%
  - CoV bending strength MGP10 = 35%
- The acceptance criteria adopted is as follows. (They are illustrations only. Each company must devise and adopt acceptance criteria that match their risk management strategies):
  - Where the test result is above the Test Comparison Value, accept.
  - Where the test value is below the Test Comparison Value but above the Design Value, accept and flag. If there are three such results in a row, then grade thresholds must be increased.
  - Where the test value is below the Design Value, then the grade thresholds must be increased immediately.
  - Where the MoE value is higher than 1.2 times the Design Value, and if the strength is well above the Test Comparison Value, the grade thresholds may be reduced.

The producer needs to:

- Establish Test Comparison Values for MoE and bending strength of MGP10,
- Monitor the production and make necessary adjustments to production parameters to ensure that adequate properties are maintained,
- Monitor the process to ensure that it remains valid.

6.2.2 Setting Test Comparison Values

The Test Comparison multipliers are found using the tables in Appendix A and B of this report.

MoE

The Table used to set the Test Comparison Value is chosen from the tables in Appendix A – Tables for Test Comparison Values for MoE.

- The test and evaluation method must correspond to the methods chosen for the monitoring system – in this case, testing random specimens and calculation of the mean using the arithmetic average of the MoEs of the specimen.
- The Confidence Level must be matched with the confidence level selected as appropriate for the risk management strategies of the company. The tables are colour coded to facilitate this choice. In this case, the confidence level for MoE is 80% and the tables will have red shading.
- The multiplier ***M*** is found using the column defined by the CoV for production MoE and the row for the number of specimens in the sample.

17% is part way between 15% and 20% - so can interpolate between the values

Non Parametric Mean MoE					
CL 80%	CoV				
No. Sample	**8%**	**10%**	**12%**	**15%**	**20%**
5	1.032	1.040	1.048	1.061	1.083
10	1.022	1.028	1.034	1.042	1.057
20	1.016	1.019	1.023	1.029	1.040
30	1.013	1.016	1.019	1.024	1.032
50	1.010	1.012	1.015	1.018	1.025
100	1.007	1.009	1.010	1.013	1.017
200	1.005	1.006	1.007	1.009	1.012

Linear interpolation between 1.029 and 1.040 gives $M = 1.033$, and with $DV_E = 10{,}000$ MPa, the Test Comparison Value is

$$TCV_E = DV_E * 1.033 = 10{,}000 * 1.033 = 10{,}330 \text{ MPa } (= 10.33 \text{ GPa})$$

Bending strength

The Table used to set the Test Comparison Value is chosen from the tables in Appendix B – Tables for Test Comparison Values for strength.

- The test and evaluation method must correspond to the methods chosen for the monitoring system – in this case, testing random specimens with each loaded to failure, and calculation of the 5%ile using a log normal fit to all of the data.
- The Confidence Level must be matched with the confidence level selected as appropriate for the risk management strategies of the company. The tables are colour coded to facilitate this choice. In this case, the confidence level for strength is 80% and the tables will have red shading.
- The multiplier ***M*** is found using the column defined by the CoV for production strength and the row for the number of specimens in the sample.

As the CoV is a tabulated value, there is no need to interpolate.

Log-normal 5%ile Strength (All data)								
CL 80%	CoV (%)							
No. Sample	**5%**	**10%**	**15%**	**20%**	**25%**	**30%**	**35%**	**40%**
5	1.033	1.068	1.106	1.147	1.190	1.238	1.289	1.344
10	1.023	1.047	1.073	1.099	1.128	1.157	1.188	1.221
20	1.016	1.033	1.050	1.068	1.087	1.106	1.126	1.147
30	1.013	1.027	1.041	1.055	1.070	1.085	1.101	1.117
50	1.010	1.021	1.031	1.042	1.053	1.065	1.076	1.088
100	1.007	1.015	1.022	1.029	1.037	1.045	1.053	1.061
200	1.005	1.010	1.015	1.021	1.026	1.031	1.037	1.042

The tabulated value gives:

$$TCV_R = DV_R * 1.126 = 16 * 1.126 = 18.02 \text{ MPa}$$

6.2.3 Using Test Comparison Values

The values of TCV_E and TCV_R calculated in Section 6.2.2 give the Test Comparison Values against which the compliance of production will be judged.

- Where the test results show values that hover around the Test Comparison Value, then the producer has a confidence that is close to the required level.
- Where there is a minor short-term excursion below the Test Comparison Value then for that shift, the producer does not have the expected confidence. The actual confidence is less. In this case, the producer has judged that this does not represent a significant problem for one or maybe even two shifts, as it may fall within the errors allowed for in the confidence level (in this case, one error in roughly five shifts).
- Where the test results are significantly below the Test Comparison Value, then the major excursion represents a significant loss of confidence (to less than 50% confidence in compliance). Where the test result is less than the Design Value, the production is more likely to be non-compliant than compliant. In terms of the Acceptance Criteria adopted by this producer in this example, remedial action will be required immediately.
- In this example, where the test results are well above the Test Comparison Values for both strength and MoE for some time, then the production settings (grade thresholds) may be able to be relaxed to improve recoveries.

Figure 7 presents an example of production outputs for MGP10 timber compared against the Test Comparison Values shown above.

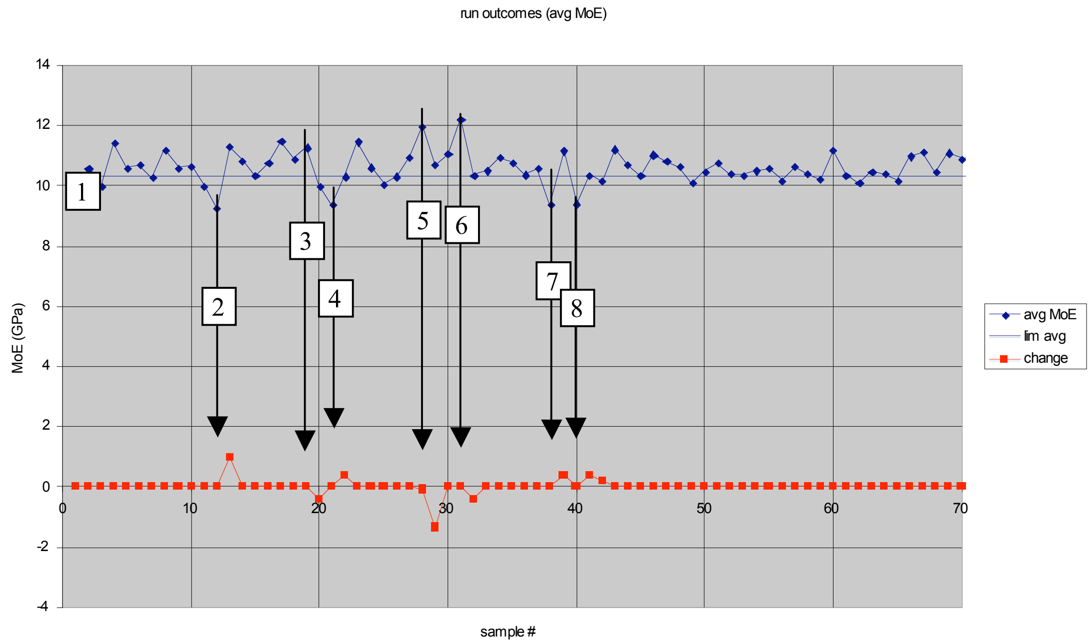
- The data on the plots represents the test results of 20 pieces analysed using the nominated method and plotted with one point representing each shift.
- The Test Comparison Values are plotted as a horizontal heavy line.
- In some cases, the results necessitated a change in production settings. These changes are indicated by the red line. (An upward change indicates an increase in threshold, and a downward change, a decrease in threshold.)

1 The data record starts off with relatively low values for both strength and MoE, but the next shift shows improvement with values that are in excess of the Test Comparison Values for both properties.

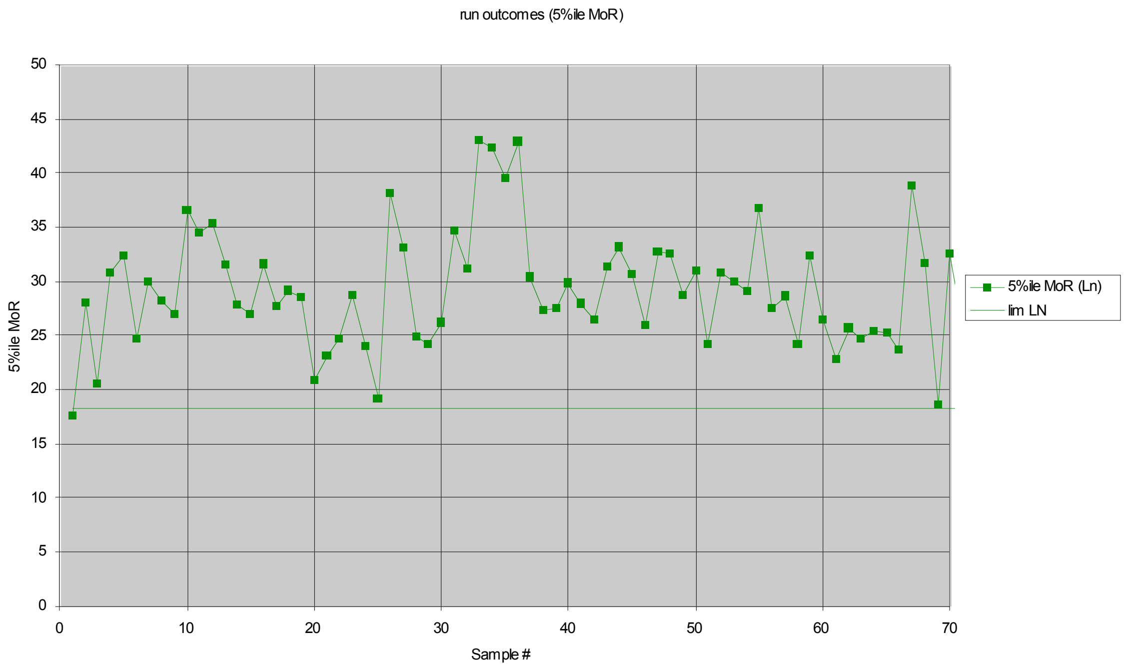
2 At 11 shifts, the average MoE appears to be a little less than the Test Comparison Value, but the next test result is even lower. As the test result is less than the Design Value (10 GPa), in accordance with the acceptance criteria adopted by this producer, grade thresholds would be raised in this situation as shown in the "change" line in Figure 7(a).

3 Shifts 17, 18 and 19 showed an average MoE that was well in excess of the Test Comparison Value. It is possible that the correction after shift 12 may have been too large. Even though the test value had not reached 1.2 times *E*, the grade thresholds were reduced a little in an effort to bring the results closer to the Test Comparison Value.

4 However on reducing the thresholds, the next two shift results were quite low. Shift 21 was on the Design Value, and shift 22 was well below it. The downward correction in grade thresholds was reversed just two shifts later. (Perhaps waiting until the results were greater than 1.2 times the Design *E* would have been better.)



(a) test results – average MoE (non-parametric mean)



(b) test results – 5%ilr MoR (log-normal fit, full data)

Figure 7 – Test results compared with Test Comparison Values

5 After 28 shifts, the test results had again become significantly higher than the Test Comparison Value with a test result of nearly 12 GPa for MGP10. The grade thresholds were again reduced, and the next test result was quite close to the Test Comparison Value.

6 The 31^{st} shift gave an average MoE test result of greater than 12 GPa so another reduction in grade thresholds could be justified. (In each of the cases in which a reduction in the grade threshold could be

achieved, a check was made that the strength data was also above the strength Test Comparison Value for the same shift.)

7 Production up to shift 38 was good with the average MoE test results generally at or a little above the Test Comparison Value. At shift 38, a low value was returned. As it was less than the design stiffness, an upward correction to grade thresholds was made. The next shift was above the Test Comparison Value.

8 Shift 40 was again less than the design value and the grade thresholds were again raised a little. This was the last correction necessary for the balance of this record.

Even though the grade threshold had changed a number of times through this example, the Test Comparison Value remained constant. It is tied only to the Design Values and sampling and analysis regimes, so does not vary with grade threshold changes or small changes in resource.

In this example, all of the changes to the grade threshold were necessary because of the MoE test results, rather than strength test results. The production appeared to have substantial strength, even though the MoE was sometimes border-line or too low. This product is an "MoE limited" product. It is the MoE test results that require the closest vigilance, and the strength data seemed to be averaging above 25 MPa. However, there were three test results that were very close to the strength Test Comparison Value, so strength still has to be monitored carefully each shift.

This example also illustrates that test results for strength have a larger variation than those for stiffness. For products that may be "strength limited" in production, the larger fluctuations of the 5%ile MoR results make control of the quality of product more difficult. Where production is "strength limited", the strength is typically not as responsive to changes in grade thresholds. So making adjustments to production parameters to rectify any strength shortfall is more difficult.

6.2.4 Monitoring Test Comparison Values

The Test Comparison Values were determined using values of CoV of the measured properties. It is advisable to check that the value of CoV used to determine the Test Comparison Values have not changed appreciably.

- Long-term stability of CoV can be verified by calculating the CoV of data accumulated during the previous 20 shifts for each of the properties measured. Because there are 20 pieces per shift, this represents the CoV of 400 pieces.
- For the strength data, the shift 5%ile was found using a log-normal fit to the shift data. In this case for consistency, the CoV was found from a log-normal fit through the 400 points (20 consecutive shifts worth of test data).
- For the MoE data, the mean MoE was found by the arithmetic average of the test specimens. The CoV was found from the standard deviation divided by the average for the 400 points.

The resulting plot is shown in Figure 8, and Table 4 summarises the Coefficients of Variation. In Figure 22, the solid horizontal lines are the values used in finding the Test Comparison Values.

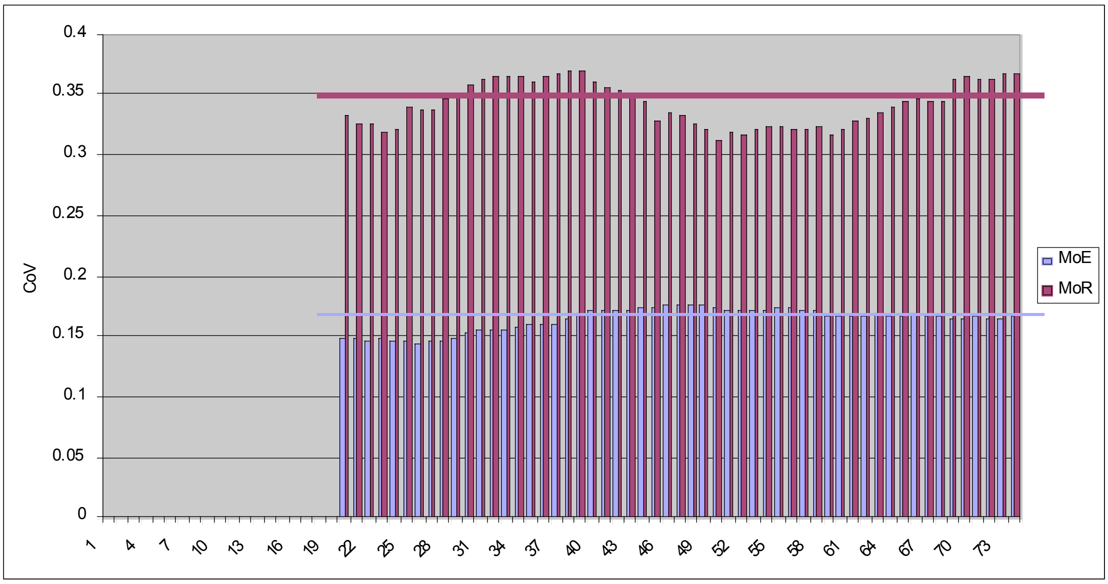


Figure 8 – CoV for test data for MoE and MoR
(400 point rolling data set used)

Table 4 – Coefficient of Variation in Monitoring data

	MoE	MoR
Assumed	**17.0%**	**35.0%**
avg	16.5%	34.3%
max	17.7%	37.1%
min	14.6%	31.4%
95%ile	17.6%	36.9%
5%ile	14.7%	31.9%
+/-	0.09	0.07

A small variation can be seen in the CoVs over the monitoring period. The average values of CoVs for both strength and MoE were a little less than the values used to find the Test Comparison Values. The values used were conservative, and therefore a change is not warranted. (Where the long-term CoV of production increases by more than a few percent over time, it is necessary to re-evaluate the Test Comparison Values using the new estimates of production CoV. Also if it is consistently below the value used for a few months, producers may want to re-evaluate Test Comparison Values with the new CoVs.)

For this example, there seems no reason to change the assumed CoV. The magnitude of the variation in the CoV through the sampling and testing period was around 7 to 9% with the change presented as a ratio of the initial value. (Had the MoR been estimated using a log-normal fit with the production CoV, this variation in CoV would be regarded as "loose CoV" as the variation is greater than +/- 5%.)

7. Conclusions

The results of this Project will give Australian timber producers a tool to assist in implementing a monitoring strategy. The outputs make it easy for producers to allow for sampling errors in estimating properties of production from test results.

The outputs of the project can be implemented by using the tables, equations and plots in Appendix A and B of this report to establish a monitoring program or refine an existing Program to meet the objectives and risk management strategies of each company.

The data used in testing the outcomes of this Project was drawn from softwood mills around Australia. The statistical methods described and the outputs have been validated with all types of Australian resource.

The project was based on only random sampling and testing, but examined a range of different analysis methods.

- For each analysis method, it was possible to determine whether the method under-estimated or over-estimated the required property, and find confidence limits on the estimation of the property.
- This led to the derivation of Test Comparison Values. If the test results exceeded the Test Comparison Values, then the producer has the nominated confidence that the production from which the samples were drawn exceeds the Design Values.
- A range of Confidence Levels, Coefficients of Variation of Data and sample sizes were used to derive expressions and relationships for the Test Comparison Values.
- No difference was found between the relationships used for bending strength and those used for tension strength.

Producers may use the outputs of the program to quantify the effects on their production of varying the following parameters in their monitoring program:

- Confidence level,
- Sample size,
- Using proof stress or testing all specimens to failure,
- Analysis method

The results of the project have been compared with classical statistical solutions where possible, and with equivalent parameters given in standards. No additional conservatism was incorporated in the estimates of mean MoE or 5%ile strength using any analysis methods.

The use of this data will enable production parameters to be monitored and set to ensure product properties with the maximum possible recovery.

8. Acknowledgements

We would like to acknowledge the assistance of the following:

- For provision of data to test the work, and comments on the draft report:
  - Geoff Stringer and Stephen Bolden, Hyne Ltd
  - Richard Schaffner and Greg Duff, Wespine Industries Ltd
  - Kim Harris and Angelo Guerrera, Carter, Holt, Harvey Ltd
  - Tony Haslett, Wayne Green, Des Pluckhan, Alan Seeby, and Alan Booby Weyerhaeuser Ltd
  - Eddie Brooks, Auspine Ltd

- For access to resources and comments on the draft report:
  - John Carson, consultant
  - Bob Leicester, consultant

- For provision of guidance in the project:
  - Technical Solid Wood Committee, A3P
  - Richard Woodman, Curtin University

- For funding of the work:
  - Forest and Wood Products Research and Development Corporation

9. References

1 R.E. Walpole, R.H. Meyers, S.L. Myers, Y. Keying, 2002 Probability and Statistics for Engineers and Scientists Seventh Edition, Prentice Hall, New Jersey.

2 Suddarth, S.K. 1994 "*Statistical fundamentals for wood Products*." Proceedings from Statistical Process Control Technologies: State of the Art for the Forest Products Industry, Forest Products Society, Madison WI pp3-15.

3 Palisade Corp 2005 *"@ RISK software"* Statistical tools incorporating Monte Carlo simulations, Palisade Corporation, CA

4 ASTM D 1990 2000 "*Standard Practice for Establishing Allowable Properties for Visually-Graded Dimension Lumber from In-Grade Tests of Full-Size Specimens*" ASTM West Conshohoken PA.

5 EN384 1995 "*Structural timber – Determination of characteristic values of mechanical properties and density*." European Committee for Standardization, Brussels, Belgium.

6 ISO 12491 1997 "*Statistical methods for quality control of building materials and components*", International Organization for Standardization, Switzerland.

7 ASTM D 5457 1998 "*Standard specification for computing the reference resistance of wood-based materials and structural connections for Load and Resistance Factor Design*" ASTM West Conshohoken PA.

8 Evans, J.W., Kretschmann, D.E., Herian, V.L., Green, D.W. 2001 "*Procedures for Developing Allowable Properties for a Single Species under ASTM D1990 and Computer Programs Useful for the Calculations.*" Forest Products Laboratory, Madison, WI.

9 Leicester, R.H. 1994 "*Statistical Control for Stress Graded Lumber*." Proceedings from Statistical Process Control Technologies: State of the Art for the Forest Products Industry, Forest Products Society, Madison WI pp63-70

10 AS/NZS 4063:1992 "*Timber – Stress-graded – In-grade strength and stiffness evaluation*." Standards Australia, NSW.

11 ASTM D 2915 2003 "*Standard Practice for Evaluating Allowable Properties for Grades of Structural Lumber.*" ASTM West Conshohoken PA.

12 ISO 2394 1998 "*General Principles on Reliability for Structures*", International Organization for Standardization, Switzerland.

13 EN519 1995 "*Structural timber – Grading – Requirements for machine graded timber and grading machines*." European Committee for Standardization, Brussels, Belgium.

14 Warren, W.G. 1977 "*The Potential for Application of CUSUM Charts to MSR Lumber. Part 1*". Fisheries and Environment Canada, Forestry Directorate, Western forest Products Laboratory, Vancouver, British Columbia.

15 Shelley, B.E. 1994 "*Lumber Quality Control with CUSUM – Issues and considerations.*" Proceedings from Statistical Process Control Technologies: State of the Art for the Forest Products Industry, Forest Products Society, Madison WI pp58-61

16 Durrans, S.R., Triche, M.H., Taylor, S.E. and Woeste, F.E. 1997 "Parameter and quantile estimation for the distributions of failure strength of structural lumber". Forest Products Journal Vol 47, no 4. pp80-88

17 EN14358 2003 "*Timber structures – Fasteners and wood-based products – Calculation of characteristic 5-percentile value and acceptance criteria for a sample*." European Committee for Standardization, Brussels, Belgium.

18 ISO13910 2005 *"Structural timber – characteristic values of strength-graded timber – sampling, full-sized testing and evaluation"*, International Organization for Standardization, Switzerland.