Quality assessment for structural lumber from mountain pine beetle-attacked timber: Data analysis

Phil Trott[1] and Conroy Lum[2]

Mountain Pine Beetle Working Paper 2009-02

[1]Canadian Wood Council

99 Bank St. Suite 400

Ottawa, ON, K1P 6B9

[2]FPInnovations – Forintek Division

2665 East Mall

Vancouver BC V6T 1W5

MPBI Project # 8.63B

Abstract

Canadian lumber producers enjoy access to the North American and overseas markets because of recognition afforded to the lumber quality and conformity assessment system. While the processing of post-mountain pine beetle (MPB) wood under this system is not new, the volume of such wood is unprecedented, as is the duration under which commercial operations must continue to process it. In response to these concerns, including the possibility that new issues may emerge from producing lumber from MPB wood, it is essential to monitor and demonstrate that the lumber quality system continues to meet its obligations, and to follow-up, if deficiencies are detected, with appropriate adjustment for the period over which the MPB-killed stands will be processed.

The Canadian Wood Council is the agency designated by the National Lumber Grades Authority (NLGA) to maintain and provide access to the Canadian In-Grade Lumber database. The proposed work will provide the baseline for the development of protocols for quality assessment through an in-depth analysis of the Canadian In-Grade Lumber databases and an evaluation of data collected from a trial sample targetting MPB-relevant defects.

Keywords: mountain pine beetle, lumber quality, lumber grading systems, shake

Résumé

Les producteurs de bois d'œuvre canadiens bénéficient d'un accès aux marchés nord-américains et étrangers grâce à la reconnaissance de la qualité de ce bois et du système d'évaluation de la conformité. Bien que, dans le cadre de ce système, la transformation de bois infesté par le dendroctone du pin ponderosa (DPP) ne soit pas nouvelle, le volume de ce bois et surtout le temps durant lequel les exploitations commerciales situées dans les régions touchées par le DPP doivent poursuivre cette transformation sont sans précédent. En réponse à ces préoccupations, y compris la possibilité que de nouveaux problèmes surviennent en ce qui a trait à la production du bois d'œuvre à partir de bois infesté par le DPP, il est essentiel de surveiller la qualité du bois d'œuvre et de démontrer que le système de contrôle de la qualité continue de respecter les obligations qui y sont rattachées. Il faut par ailleurs assurer le suivi, dans le cas où des lacunes seraient détectées, en procédant aux ajustements appropriés pour la période durant laquelle les peuplements tués par l'infestation de DPP seront transformés.

Le Conseil canadien du bois est l'organisme désigné par la Commission nationale de classification des sciages (NLGA) pour maintenir et offrir l'accès à la base de données sur le bois d'œuvre canadien par catégorie. Le travail proposé formera la base pour l'élaboration de protocoles d'évaluation de la qualité fondés sur une analyse en profondeur des bases de données sur le bois d'œuvre canadien par catégorie et sur une évaluation de données recueillies à partir d'un échantillon d'essai ayant pour cible des défauts liés au DPP.

Mots clés : dendroctone du pin ponderosa, qualité du bois d'œuvre, systèmes de classification du bois d'œuvre, fentes

Table of Contents

List of Tables

List of Figures

1   Introduction

The design values for major commercial species of lumber, such as SPF (spruce-pine-fir), are established from large-scale in-grade lumber testing programs. Due to the manner in which this methodology has been promoted, it is also seen as a possible way to provide information to assess whether there are quality issues associated with lumber from post-MPB wood. Unfortunately, in-grade lumber test programs only work well when sufficient production exists to enable a representative test sample to be randomly collected over the population of interest. In the case of concerns with specific characteristics of lumber from MPB wood, generating such a reliable sample may not be possible.

1.1   Overview

This data analysis study is intended to supplement Forintek's MPBI Project #8.63, *Draft Protocol for Evaluating Potential Strength Reducing Characteristics in Lumber after a Catastrophic Event: Pilot Application and Next Steps* (Lum 2009*)*, and will provide a perspective on defects that may be relevant to future study of MPB-affected lumber defects and their observed frequency and severity.

A review of industry lumber databases developed from earlier lumber property studies has been completed to determine what information on MPB-relevant characteristics is available and how useful it may be in evaluating post-MPB lumber. A relationship between the severity of characteristics and their associated mechanical properties, subject to the data's limitations, was investigated to support whether adjustments to the system may be warranted as recommended by quality assessment protocols. The final portion of this data analysis study reviewed data from a trial survey of No. 2 and better grade SPF lumber sampled from two mills in B.C.

1.2   Objectives

The general objective of this study was to supplement through data analysis the development of a quality assessment protocol for post-MPB wood (Lum 2009) that is consistent with standard North American procedures for testing and evaluating commercial structural lumber.

The specific data analysis objectives were to determine the usefulness and limitations of existing data collected in past pan-Canadian cross sectional studies and how they may be used to assist in developing regional longitudinal programs.

The second objective was to extract existing data on MPB-relevant characteristics from industry databases relevant to their frequency, and severity relative to grade rules. The techniques developed in this analysis are expected to form part of the evaluation framework of the quality assessment protocol.

The third objective was to develop relationships between characteristics and mechanical properties. If one was established, some recommendations could be made on how to adjust the grading rule, if warranted.

The fourth objective was to select a trial "MPB marker" characteristic based on data review to assess how much information could be extracted with a single MPB-relevant characteristic. Ideally, this sample MPB marker would be derived from useable frequency and severity information in past databases and be practical for visual assessments in the field.

The final objective was to support through analysis the development of an evaluation framework and review the efficacy of its procedures by analyzing trial survey data collected in two mills. The results of this evaluation should be considered preliminary and it is expected that further work will be required because of the statistically complex nature of classifying a large number of pieces with a certain defect, and the overlapping nature of the grading system.

2   Data Usefulness and Limitations

2.1   Data Limitations - Materials and Methods

A survey of available lumber data containing visually determined characteristics that may be relevant to MPB included the 1983 in-grade database and the NLGA 1999 National monitoring database.

2.1.1   *Materials – Lumber Databases Analyzed*

The in-grade study was a joint effort between lumber industries in Canada and the United States to test thousands of pieces of lumber to destruction to determine their in service characteristics. The objectives was to assign new design values based on *in-grade* testing which is believed to result in more accurate design values for lumber. It was agreed that this in-grade testing program should simulate, as closely as possible, the structural end use conditions to which the lumber would be subjected to. In Canada this cross sectional in-grade program sampled two grades, and three sizes from the entire growth range of three species groups including SPF.

The 1999 NLGA monitoring study was also an in-grade style testing program whose objectives were to gather information on a key size and grade of SPF and move toward a standard framework for longitudinal monitoring studies. The sampling and testing was very similar to the original in-grade program but the focus was restricted to a single size grade and species group - SPF 2x4 No. 2.

The 2007 pilot study undertaken under the MPBI protocol (Lum 2009) sampled No. 2 grade and better 2x4 SPF from two B.C. mills. The purpose of the study was to try to sample lumber so that each lot or group of 10 sequential production pieces contained at least one MPB-relevant defect. Extra pieces without an MPB-relevant defect were

collected to replace those with an MPB-relevant defect to understand the impact of potentially restricting some defects.

2.1.2 Methods – Lumber Data Review

A search for all types of characteristics was undertaken in both the 1983 and 1999 lumber studies; the characteristic needed to be visual to be useable in the protocol's field studies. Three types of visual characteristics were found to have been recorded for each piece of lumber. This common information on each piece tested included the grade controlling characteristic (GCC), the maximum strength reducing defect (MSRD), and the cause of failure (FC). Although the original intent in collecting these data was to demonstrate grade representativeness, the data were recorded in a format that would lend itself to analysis related to the development of the MPB protocol.

Ideally, the MPB protocol would be based on the visually determined characteristic with the least limitations with respect to data and be practically observable in the field. The GCC was defined as the single characteristic related to strength or visual quality that was responsible for the piece being assigned its visual grade. The MSRD was defined as the characteristics judged to have the greatest impact on strength for the entire piece whereas the FC was defined as the characteristics judged to be responsible for the onset of failure but it could only originate from a shorter section of the piece that was destructively evaluated because of testing setup. An added advantage of using MSRD was that it occasionally was also the cause of failure in the destructive test and therefore could serve as an effective predictor without destructive testing, which is a key component of a grading system.

2.2 Data Limitations - Results and Discussion

When considering the requirements for what type of visually recorded data was most useful to the MPB protocol, MSRD was considered to be the most appropriate. The reasons for this were that it was based on the concept of strength and is determined from the full length of the piece. It was also the only visually recorded data that underwent round robin evaluations in grading competitions. After a review of the available data sources and the objectives under which those data were collected, the maximum strength reducing defect (MSRD) was believed to be most useful for the development of a protocol but several important limitations described below needed to be considered.

2.2.1 Grading Limitations

It is important to note that all three visually recorded properties are somewhat subjective assessments. The reasons being that they were assigned based on visual judgement of an individual grader who compared and rated the expected impact on lumber strength between several competing types of defects, for example, between a knot of a certain size versus a shake of certain length. In both examples, the defects extended into the cross section of the piece and were only visible to the grader on the surface, requiring the grader to visualize the internal structure of each significant defect and record the defect having the most significant reduction on strength of the piece.

When recording defects for the in-grade and 1999 databases, graders only coded single defects with no mechanism to assess whether nearby defects could jointly affect mechanical properties. Therefore, these types of databases could not help assess whether it may be worthwhile to consider interaction of defects.

2.2.2 *Defect Coding Issues*

Before attempting analysis of frequency or severity it became apparent that the use of different recording schemes in the two data sources would make it difficult to group and simply compare the defects and develop meaningful conclusions. This became especially apparent when, for example, shake was considered as demonstrated in Table 2. The table was developed to demonstrate differences in recording schemes for severity of shake. If one considered the frequency of shake in the in-grade data for a No. 2 sample, three codes (11, 12, 13) were used to fully bracket No. 2 shake from the upper to lower spectrum of the No. 2 shake versus the use of only two codes (05, 03) for monitoring data which covered only the middle of the No. 2 shake. The Forintek coding scheme could account for fewer occurrences of shake in the same No.2 sample than the ASTM D4761 coding scheme, potentially introducing a bias in the frequency of shake in monitoring data.

Table 1. Various coding schemes for shake in relation to the national grade rule

Visual Grade	National Grade Rule description	Rating within grade	ASTM D4761 X1.1 description (used to code in-grade data)	Forintek Failure Coding description (used to code monitoring data)
Select Structural	Surface shake only less than 2 ft long	Upper	Light not through (01)	Not through < 2 ft long (01)
	End splits less than piece width	Middle	Medium not through (02) Others not through (03)	Not through > 2 ft long (02)
		Lower		
No. 1	Surface shake only less than 2 ft long	Upper		
	End splits less than piece width	Middle		
		Lower		
No. 2	Through shake less than 2 ft long well separated	Upper	Light through (11)	
	Single shakes not through less than greater of ¼ length or 3 ft	Middle	Medium through (12)	Shake breaks <2/3 the edge (05) Through shake <2ft average (03)
	End splits less than 1½ piece width	Lower	Others through (13)	
No. 3	Scattered through shake less than 1/3 length	Upper		
	Any surface shake End splits less than 1/6 piece length	Middle		Shake breaks >2/3 the edge (06) Through >2 ft average (04)
		Lower		

A comparison of both recording schemes in the appendix revealed that other MPB-relevant defects such as unsound wood were not recorded in the same way in both data sets. In the ASTM D4761 X1.1 coding scheme, unsound wood and peck were jointly assigned code 33 whereas the Forintek Failure coding scheme used 31 with three sub categories.

Both of these coding systems were developed to document lumber strength reducing characteristics but placed most of the emphasis on accuracy in coding the more commonly prevalent characteristics such as knots and slope of grain. In earlier mill visits a survey of production, grading practices, and review of lumber bundles underscored the need to develop a more comprehensive coding scheme for defects that may be associated with MPB-affected lumber.

As in-grade databases are not set up to provide necessary information to refine lumber grading rules because of increased occurrence of defects, a more detailed description of the selected MPB-relevant characteristic for each grade was considered with three categories of severity for each grade. It was believed that baselines of frequency and severity derived from MSRD data would somehow be developed and applied to future field surveys. To do this many factors needed to be taken into account including assumptions of equivalency between existing data sets, field input from experienced graders, consideration of data biases, and a method to relate past data to future data collected from surveys as described in the MPB assessment protocol.

Instead, an iterative approach is now believed to be more appropriate and the more detailed coding scheme would be applied only at the more regionally focussed stages of the protocol instead of doing so on a global scale, which would be impractical. Also, certain quality control rules should be determined on a case by case basis instead of being integral components of national programs. In this area it is expected that additional resources will need to be directed towards better documenting the characteristics found in lumber from a stand subject to a catastrophic event.

3 Frequency and Severity Baselines

3.1 Frequency and Severity Baselines - Materials and Methods

A review of data collected in both the in-grade and 1999 datasets revealed two groups of MPB-relevant defects that had useable data. Those groups were defined as a lengthwise separation of wood group (shakes, checks and splits) and a decay group (unsound wood, honeycomb, white speck and brown rot, heart stain). Although three types of visual assessments were recorded for each piece, the target frequencies and severities reported in this report were based on the MSRD (maximum strength reducing defects).

The frequency of shake and decay was determined in both databases for the SPF population. Determining severity was more complex because coding scheme equivalencies needed to be considered. Because only three levels of severity were recorded for shake in the No. 2 grade, only qualitative conclusions could be made about the data. Severity of the decay group was not recorded in much of the data so no

conclusions could be made. Species, specimen size and grade effects were also investigated as well as the predictability of the cause of failure from MSRD data.

3.2   Frequency and Severity Baselines - Results and Discussion

The frequency of shake as the MSRD in the No. 2 grade SPF populations ranged from 5% to 10% depending on coding assumptions and the database considered. An investigation of size effects revealed an increased occurrence of shake in wider widths possibly due to the increased susceptibility to shrinkage stresses. The frequency of the defects in the decay group ranged from 0 to 3% in the SPF population.

The incidence of MPB-relevant defects being simultaneously recorded as the MSRD and the cause of failure was tallied to understand how suitable field observations were in predicting destructive testing results. The number of times MSRD ended up also being the cause of failure ranged from 10% to 80% depending on the dataset. No testing mode effects were apparent for MSRD since their assessment was not related to the testing mode.

As mentioned earlier, developing a baseline for severity of shake for No. 2 grade was of little use because the two data sets available used different coding schemes that overlapped in the middle of the grade as shown in Table 1. When severity data from both studies was plotted, the majority of the data fell into the middle severity bin for No. 2 grade. The 1999 data set recorded the type of decay instead of recording decay severity; therefore, no severity baselines for decay were developed.

4   Property Relationships

4.1   Property Relationships - Materials and Methods

The original in-grade and 1999 databases were analyzed to compare the shake severity data to mechanical properties using both the ASTM and Forintek coding schemes which each recorded six types of shake. Table 1 was developed with input from senior graders to rank the severity of the shake type between databases. The MPBI survey data (Lum 2009) was also reviewed and the more detailed shake metrics were compared to strength.

The difficulty in relating severity codes between databases and the limited number of pieces containing shake made it challenging to develop meaningful models between shake and strength or stiffness, and impossible for decay because of missing data.

### 4.2   Property Relationships – Results and Discussion

Preliminary analysis of mechanical property data and characteristic data had revealed that it would be difficult to uncover meaningful relationships based on the groupings used to classify severity data because of the shortcomings that existed in theoretical relationships. It is important to note that the grading system was founded on theoretical relationships between defects and their impact on clear wood and the real effect on full-sized lumber specimens is not accurate. These relationships tended to focus on the impact on strength due to knots and slope of grain because they were more prevalent and believed to have the most significant impact.

A review of the shake severity data of No. 2 grade in-grade dataset revealed that no pieces contained the lowest shake code, leaving only two tiers of severity in the grade to develop a relationship between strength or stiffness. Because of this limitation, and the fact that the 1999 databases had only one tier of severity in the No. 2 grade, and that the sample size was limited, no meaningful trend was investigated. It was concluded that any relationships developed from such limited data would be of little significance.

However, a review of shake in the in-grade database for the three grades that the shake classifications spanned (select structural, No. 1 and No. 2) was performed to augment the sample size and confirm if at least a qualitative relationship existed across multiple grades. A review of the shake severity and strength data found that the more severe the shake was in terms of length and sides, the lower the resulting strength property. This qualitative observation supported what is currently found in the NLGA Grade Rule.

The MPBI survey data also confirmed that no relationship between shake severity, as coded by the Forintek coding system, and strength existed. A search for a relationship using the more detailed coding metrics also did not reveal any meaningful trends as displayed in Figure 1 below.

Figure 1. Modulus of Rupture versus shake attributes
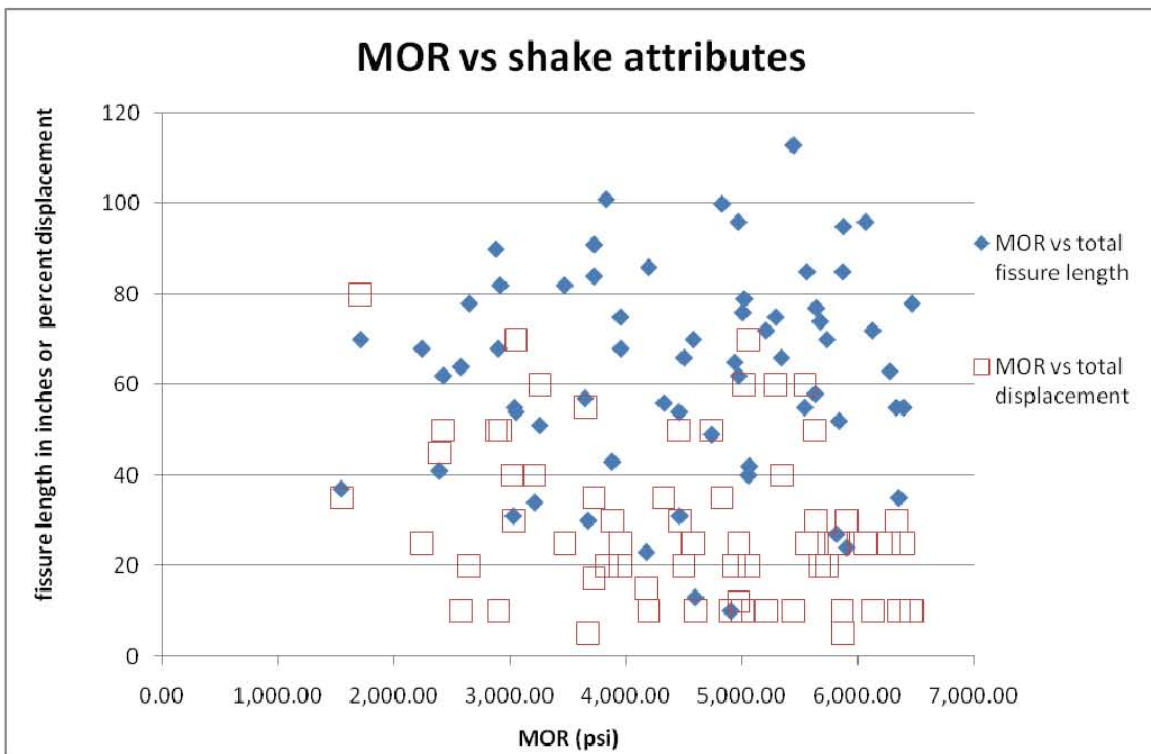
Because of the lack of a clear relationship between shake and strength, an analysis that sought shake outliers in lots of 10 pieces was considered to be more useful in the context of the MPBI protocol (Lum 2009). This approach is described in more detail in section 6 of this report.

5 Trial Defect Selection

5.1 Trial Defect Selection -Materials and Methods

A database review was performed to seek out which MPB-relevant defect contained meaningful information in support of the MPBI protocol (Lum 2009). Factors that were considered in this trial defect selection included: data quality, the level of detail in the severity classification, data quantity to support meaningful analysis, practical application in the field, potential relationship to mechanical properties, and impact on lot properties.

5.2 Trial Defect Selection – Results and Discussion

Based on past data, the two candidates for an MPB-relevant trial defect were shake and decay. As mentioned earlier, the frequency of decay was shown to be low and in some data sets no severity information had been recorded. Therefore, shake as a trial defect for the MPBI (Lum 2009) protocol was determined preferable when considering data available. Although the existing recording schemes were coarse compared to the precision used to record knot sizes or slope of grain, the potential to develop a more detailed coding scheme existed.

With the assistance of experienced graders, a trial recording scheme was developed to record more information about shake such as fissure length and number of faces affected. The details of this recording scheme were reported in the MPBI protocol report (Lum 2009).

6 Evaluation of the Framework

6.1 Evaluation of the Framework – Materials and Methods

Under the MPBI protocol (Lum 2009), a trial sample of No. 2 and better grade 2x4 SPF lumber was selected from two mills in B.C. in the areas affected by the mountain pine beetle. The purpose of the sample was to assess whether sequential pieces from production could be selected in such a way as to ensure at least one piece in each lot of 10 had shake, a replacement of pieces with shake could be simulated, and meaningful analysis could be performed. The MPBI sample (Lum 2009) collected 29 lots of 10 pieces at two mills and supplemented all the pieces with shake with extras containing no shake.

A preliminary review of the MPBI trial study data (Lum 2009) was performed to comment on whether a purposive sample would provide some guidance in refining grading practices that was not available from existing databases. A cursory review of data from this sample was performed to understand what benefits this type of analysis may yield in terms of empirical grade rule adjustments.

The resulting data was analyzed and lot statistics were plotted to reveal the impact of replacing some or all of the pieces containing shake with spares that didn't contain shake. Three cases were compared: the base case with at least one piece of shake per lot; the augmented case where no pieces contained shake; and the partially augmented case where only the most severe cases of shake were removed and replaced. Although the methodology suggested is a trial and may need further statistical review, the qualitative observations resulting from the analysis were reported.

6.2   Evaluation of the Framework – Results and Discussion

One way to consider variability was to consider a sample that was not purely from a random sequential production stream, by virtue of its artificially imposed frequency of pieces with shake, and replace the intentionally selected pieces containing shake with others without shake. This type of sample allowed us to develop an understanding of how changing the frequency of pieces with shake in a lot could affect the minimum modulus of rupture in each lot and to a lesser extent lot means.

The total data comprised of 58 lots of 10 pieces sampled from two mills. Lot minimums were the significant point of interest because they were believed to be more related to variability. A comparison between the base case and the augmented case revealed that replacement of each piece containing shake with another that excluded shake increased minimum modulus of rupture values for 15 of 58 lots. Table 2 below summarizes changes to lot minimums, means and maximums when pieces with shake were removed and partially removed.

Table 2. Comparison of lot statistics

	Augmented case compared to base case	Partially augmented case compared to base case
Lot minimum increased	15 out of 58	9 out of 58
Lot minimum decreased	3 out of 58	1 out of 58
Lot mean increased	24 out of 58	15 out of 58
Lot mean decreased	10 out of 58	3 out of 58

The partially augmented case removed two of the shake types that are classified as shake breaks and results were provided in the third column of Table 2; these demonstrated a similar but less pronounced trend of increasing lot minimums.

Figure 2 below shows the maximum, minimum and mean modulus of rupture values for some of the lots in the trial study. The grey bars represented the range of modulus of ruptures values in a lot and blue diamonds the lot means. The superimposed red lines demonstrate the modulus of rupture values in the same lots with pieces containing shake replaced with pieces excluding shake. The Figure 2 plot shows more increases than decreases in lot minimums and means for a wide range of strength levels seen in the sample.

Figure 2. Modulus of Rupture by Lot
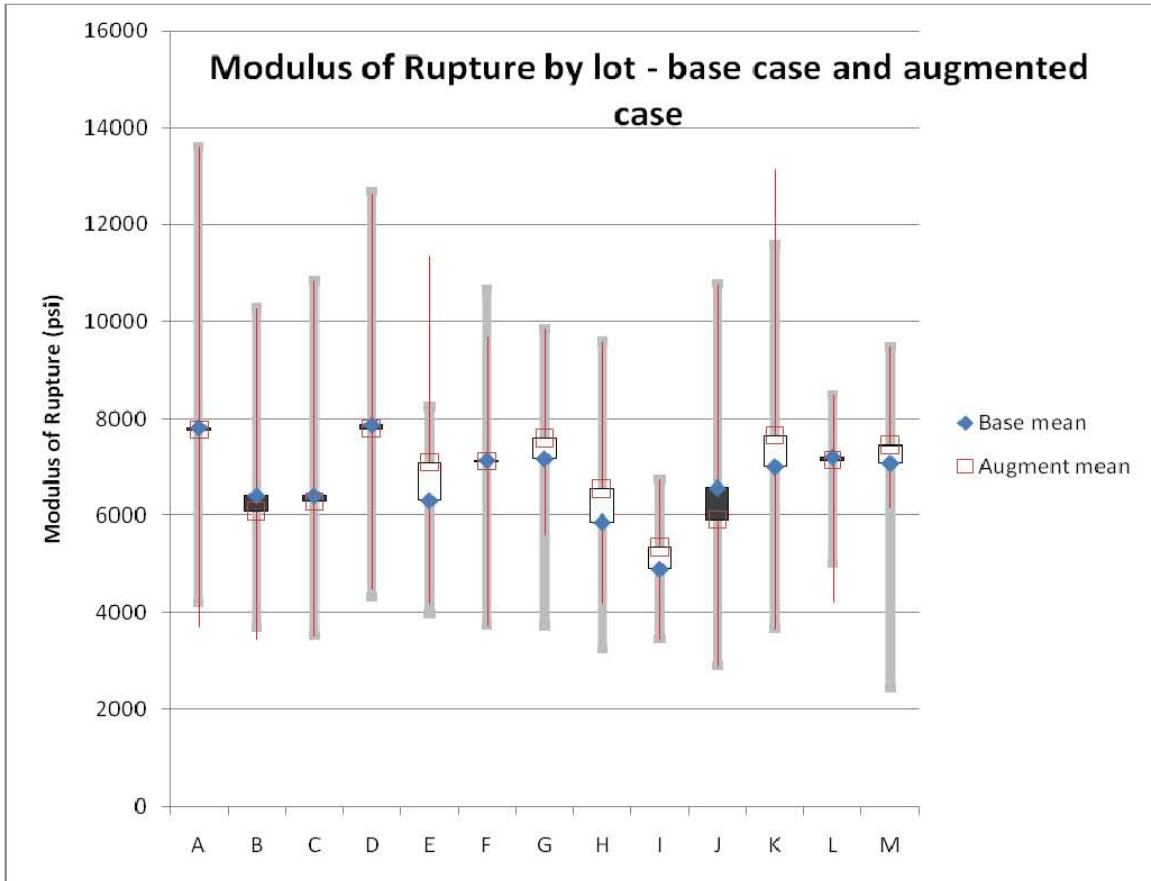
The MPBI trial sample (Lum 2009) was taken to assess the value of a different approach which in this case empirically augmented or replaced certain defects at the lot level to assess the impact. The exercise revealed some changes to lot properties which now needs to be analyzed to determine its significance and the value of this type of approach and if other similar approaches may be useful. Next steps should include a more thorough review of these observations by both statistical experts and the lumber industry.

7 Conclusions

Within the limitations of existing lumber databases and recent MPBI trial survey data (Lum 2009), shake was considered an adequate trial defect to begin to quantify lumber quality issues associated with the mountain pine beetle infestation.

The comprehensive review of existing lumber databases provided some intial frequency information of MPB-relevant defects from global population lumber studies. Although frequencies of defects observed in global samples will be different than those from regional samples, global baseline statistics should be part of the intial steps of an iterative monitoring process of lumber quality subjected to catastrophic events.

Analysis of lumber data did not uncover any meaningful baselines for severity of an MPB-relevant defect, nor did it discover any new relationships with mechanical properties. However, in the MPBI trial survey of two mills (Lum 2009), a simulated replacement of lumber pieces containing shake with those free of shake, shed some light on the practicality of this type of approach and provided some initial observations which needs further statistical and lumber industry review.

As more is learned about classifying lumber quality, it is expected that lumber quality assesment protocol will be structured in iterative stages and will need to be integrated with ongoing monitoring programs at both regional and global levels.

8 Acknowledgements

8.1 Collaboration with partners

The project hosted a small workshop with industry consultants in March 2007. Review and feedback was sought on the draft MPB assessment protocol (Lum 2009) and how data from this study could be used in support of characteristic surveys recommended in the protocol. Some feedback was obtained on the development of alternative models between characteristic severity and strength.

The project sponsored visits to two mills to review MPB defects as they were occurring in production. Options and practical considerations for the development of more detailed coding schemes were discussed with graders and statisticians.

9   Literature Cited

Lum, C. 2009. Proposed Protocol for Evaluating Potential Strength- Reducing Characteristics in Lumber after a Catastrophic Event: Pilot Application and Next Steps. MPBI Project #8.63. Natural Resources Canada, Canadian Forest Service, Pacific Forestry Centre, Victoria BC. Mountain Pine Beetle Working Paper 2009-01.

ASTM, D4761-05 - Standard Test Method for Mechanical Properties of Lumber and Wood-Based Structural Material, 2007.

NLGA, Standard Grading Rules for Canadian Lumber 2003, March 2007.

FP Innovations- Forintek Divisions, Forintek coding scheme, 2002.

Contacts:
Phil Trott
Canadian Wood Council
99 Bank St. Suite 400
Ottawa, ON, K1P 6B9
ptrott@cwc.ca