Is limit states design ready for the new millennium?
Gromala, David S.1, Rosowsky, David V.2, Ellingwood, Bruce R.3, and Murphy, Joseph F.4

ABSTRACT
When the U.S. National Bureau of Standards published NBS Special Publication 577 in 1980, it represented a quantum leap forward in extending the principles of reliability analysis to structural engineering. NBS SP577 provided the basis for the Load and Resistance Factor Design (LRFD) load factors in ANSI Standard A58.1 (now ASCE Standard 7). Similarly, when reliability-based limit states design (LSD) procedures for engineered wood construction were developed in the late 1980’s, it represented another quantum leap -- this time in extending the same principles to wood structures. During the past decade, probability-based limit states design has been relatively stable.

A potential conflict arises between the 1980s vision, in which reliability analysis would be used exclusively to develop LSD/LRFD codes and standards, and the evolving reality in which reliability analysis is becoming available to individual designers – potentially for project-specific or product-specific design. The authors discuss the limitations and potential pitfalls of reliability analysis when used in routine structural design – most notably the improper definition of the limit state and incomplete or erroneous accounting for all sources of uncertainty. While standards development committees have examined both of these in great detail, there is little guidance for individual engineers in this area.

The new millennium is an opportune time to reflect on differences between the 1980s vision of LSD/LRFD and the early 21st century realities of its implementation. To begin the discussion that might lead toward resolution of these issues, the authors propose a list of considerations which, if adopted as standard procedures for application of reliability analyses, could bridge a few of the gaps between theory and application of structural reliability in the 21st century.

INTRODUCTION
Early application of reliability analysis to structural engineering was centered in consensus-based standards development committees that included the leading experts in the field of structural reliability. These and other experts worked throughout the 1980s to distill the principles of theoretical reliability analysis into sets of acceptable procedures that could yield structural performance measures suitable for codification. In the United States, NBS SP577 (Ellingwood, 1980) first developed the load factors that have become the cornerstone of LRFD in ASCE Standard 7 (ASCE, 1998). The 1980s vision was that reliability analysis would be used almost exclusively to assist developers of LSD/LRFD structural codes and engineering -- not as a project-specific or product-specific design tool. This conceptual bias is reinforced in other international standards documents. The Canadian Standards Association (CSA, 1981) publication S408 discusses both the partial factor method and direct probabilistic methods. However, it states that: “The partial factor method…is recommended as a common basis for the checking of civil engineering structures in Canada.” The International Standards Organization (ISO, 1998) ISO 2394 similarly discusses both methods. It also states that: “The partial factors format is the format which is intended to be used for design calculations in normal cases.”

In spite of these directives by standards writers, the limited scope of structural reliability analysis in the 1980s vision may prove to be inaccurate as we enter the new millennium. One author (Ellingwood, 1994) notes the introduction of structural reliability as a topic in engineering education and the publication of several textbooks. He describes “an

1 Senior Engineering Specialist, Weyerhaeuser Company, Tacoma, WA, USA
2 Associate Professor of Civil Engineering, Clemson University, Clemson, SC, USA
3 Professor of Civil Engineering, The Johns Hopkins University, Baltimore, MD, USA
4 Research Engineer, USDA Forest Products Laboratory, Madison, WI, USA
explosion in systems reliability analysis and advanced stochastic computational methods for complex systems.” While such developments will certainly serve to further advance code provisions, the broad availability of these tools is leading to an increasing belief that structural reliability is suitable for project-specific design.

The potential pitfalls of reliability analysis fall into two major categories – improper definition of the limit state and incomplete or erroneous accounting for all sources of uncertainty. Consensus committees in their deliberations that have led to the adoption of the current generation of LSD and LRFD standards and specifications have examined both of these in great detail, and provide some control of such errors. The following section examines some of the pitfalls that may be faced by an engineer who performs a simplified reliability analysis on a “full structure” basis for a multi-story apartment building without the benefit of code-defined system-failure limit states or prescription of the sources of variability to be considered.

**CONCEPTUAL EXAMPLE: MULTI-STORY APARTMENT BUILDING**

This brief conceptual example illustrates just a few of the questions that might arise if a team of engineers were given the freedom to design a building with no directive other than a target reliability index. It is intended to provide a glimpse of just a few of the issues that must be addressed before reliability analysis can move to the project-specific phase. What would happen if the design specification for a multi-story apartment building included the building size and location, but regarding code requirements simply said “Design shall achieve a reliability index of 3.0”? In the absence of other information, the conscientious design engineers might begin digging for background information. Floor load statistics might be available in the literature. Snow data and wind data might come from the weather bureau. Seismic data might come from an old engineering textbook. Where would they find structural strength properties? The framing lumber data is published in some industry test reports. The engineered wood product data could be pulled from a promotional article by the manufacturer.

Once past the basic data collection, the engineers must decide how to assemble it. Should they model the composite behavior of the sheathed framing systems, or should they just bump the capacities by a “system factor” from the literature? They can combine the floor load statistics and the floor system strength statistics in the reliability analysis to get a floor component reliability against the limit state of floor collapse. If it's higher than the target of 3.0, it meets the specification. OK so far. Several of the floors have multi-span joists in them. Should they re-run the reliability analysis for multiple cases (i.e., adjacent, alternate span loading) like they did in the old days? That's a tough call.

Moving to multiple-level load collection, the problem gets tougher. What loads are collected into the wall columns on the first floor? What does it mean to combine load statistics from each floor level? Should they check the arbitrary-point-in-time load on three levels combined with the actual load statistics on the fourth level? Under this type of check, is the target reliability of 3.0 for the main support columns adequate?

The point of this example is to remind ourselves that the engineers need a sense of reference points in any analysis to be able to apply good engineering judgments to their design. When the code provides nominal values for loads, the engineers have a fairly good sense of the relationship between the nominal load and “typical” loads on the structure. For floor loads, the engineer knows that 40 psf is rare, but achievable in a residence and that 100 psf in a corridor is similarly rare, but achievable. This intuition permits the engineers to examine various load cases relative to both code provisions and judgment. It provides a level of confidence that code-permitted area reductions (or multiple-level reductions) in loads will not compromise the safety of the structure. Increases for “system” effects for today’s engineers are generally modest – again matching their judgment that the system may perform better, but that system increases shouldn’t squeeze all safety buffers out of the design. Until these (and a host of other) issues are resolved, structural reliability analysis cannot be consistently or safely applied to project- or product-specific design.

**STANDARDIZATION NEEDS**

The following five items are identified as being key standardization needs that must precede the evolution of probabilistic analysis from a code development tool toward a general purpose design tool.

1. **Standardized load distributions**
   The type of distribution and parameters chosen to represent each load case will have a significant effect on the computed reliability index. For example, early examples of reliability analysis for hot-rolled steel were based on closed-form
lognormal load, lognormal resistance reliability index calculations. Many analyses in the U.S. subsequent to publication of NBS SP577 used the load statistics listed therein as a basis for reliability calculations.

**Table 1. Load Distributions from References**

<table>
<thead>
<tr>
<th>Load</th>
<th>Reference (basis of nominal)</th>
<th>$R(\text{bar}) / R(n)$</th>
<th>COV</th>
<th>Distribution Type</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>D</td>
<td>SP 577</td>
<td>1.05</td>
<td>0.10</td>
<td>Normal</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Foschi, et al</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L</td>
<td>SP 577</td>
<td>See footnote 2</td>
<td>0.25</td>
<td>Type I</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Foschi, et al</td>
<td>0.812</td>
<td>0.272</td>
<td>Residential</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.925</td>
<td>0.236</td>
<td>Offices</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.480</td>
<td>0.133</td>
<td>Commercial</td>
<td></td>
</tr>
<tr>
<td>S</td>
<td>SP 577</td>
<td>0.82</td>
<td>0.26</td>
<td>Type II</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Foschi, et al</td>
<td>1.23</td>
<td></td>
<td>Type I</td>
<td>Saskatoon</td>
</tr>
<tr>
<td></td>
<td>(high/low)</td>
<td>1.19</td>
<td></td>
<td></td>
<td>Halifax</td>
</tr>
<tr>
<td>W</td>
<td>SP 577</td>
<td>0.78</td>
<td>0.37</td>
<td>Type I</td>
<td></td>
</tr>
<tr>
<td>E</td>
<td>SP 577</td>
<td>See footnote 4</td>
<td>(2.3)</td>
<td>Type II</td>
<td></td>
</tr>
</tbody>
</table>

1 Ratios differ between SP 577 and Foschi, et al due to differing data sources and 50-yr vs. 30-yr statistics.

2 Varies, based on area reduction equations, see SP 577

3 For expanded discussion of wind load information see Ellingwood, 1999.

4 Varies, site dependent, see SP 577

Similar analyses for wood products in Canada have attempted to duplicate the multiple load case scenarios examined by Foschi, et al. See Table 1 for examples of some assumed load distributions in several references. The authors recommend that standardized load distributions be referenced in future versions of load standards such as ASCE Standard 7 and CSA S408. Note: This is not intended to preclude research examination of these variables. In fact, continuing study to refine estimates of the true distribution of loads is encouraged. As a starting point for discussion of standardization needs in this area, consider the following:

- Gravity live load cases can generally be described in terms of Extreme Value Type 1 distributions. Even though early analysis of individual data sets led to the choice of other “best-fit” distribution types, these conclusions (based on relatively small data sets) can change as additional data are accumulated. The lack of standardization of distribution types is a complication that only serves to confuse the interpretation of reliability analysis results.

- Standardization of distribution parameters for gravity load cases is recommended. For snow loads, a means of smoothing these parameters based on regional snow characteristics would be superior to the use of discrete site-specific distribution parameters. This is the process used in NBS SP577.

- Reference periods for gravity load cases are generally assumed to be 50 years in the U.S. and 30 years in Canada. While it would be nice to eliminate this inconsistency for better cross-border interpretation of reliability analysis results, it would appear that this difference will be retained.

- Lateral load cases are generally be described in terms of Extreme Value Type I or Type II distributions. As with gravity load cases, standardization of distribution parameters and smoothing based on regional load magnitudes is recommended (where possible).
• Reference periods for lateral load cases are in a state of flux in load standards – evolving from the 50 year (30 year in Canada) reference period to near-maximum (i.e., once-in-a-lifetime or longer) periods. Standardization in this area across both wind and seismic load cases is recommended.

2. Standardized Resistance Distributions
As with loads, the type of distribution and parameters chosen to represent the resistance distribution has a significant effect on the computed reliability index. Once again, early reliability analyses for hot-rolled steel assumed a lognormal resistance distribution. Typically, the distribution form chosen is based on goodness-of-fit tests. Unfortunately, these tests are not reliable for small data sets. Many analyses in the U.S. subsequent to publication of ASTM D5457 (ASTM, 1993) use the two-parameter Weibull distribution, taking advantage of its tail-fit provisions for larger data sets. Procedures in Canada for wood products are generally consistent with this approach.

Once the distribution type is standardized, procedures for defining the distribution parameters must also be standardized. As a starting point for discussion of possible standardized procedures, consider the following:

• Virtually all wood products used in structural applications in North America are manufactured and graded in conformance to some type of product standard. The product standards generally define the statistical basis by which design values are derived. Products are periodically evaluated to ensure that they meet or exceed the target properties.

• The product standards effectively “peg” a point on the statistical distribution for a given property. The “peg” point is generally the 5th percentile or the mean. Once this point is defined, a complete family of two parameter Weibull distributions can be defined as a function of the coefficient of variation.

• Reliability analyses are then conducted using the parameters generated by the aforementioned procedures.

• Some wood products are subjected daily property monitoring while others are not. There is no current method for differentiating between the two. It is recommended that a study be initiated to assess the temporal variations in properties for various product types. These temporal variations can then be analyzed to compute a factor to adjust design values. The analysis would be similar to the analysis underlying the time effect factor (Ellingwood and Rosowsky, 1991).

• If a product is to be used solely within a defined set of applications, the reliability analysis should focus on the load distribution and configuration limits specific to that application.

• If a product is to be used in a broad set of applications (typical for wood products), the reliability analysis will yield different answers for different configurations. In the U.S., it was determined that a reliability analysis referenced to the snow load case at a nominal load ratio of three (S/D=3) yielded a reliability index that was roughly in the middle of a range of common design cases. In Canada, Foschi, et al conducted the reliability analysis over a range of “standard” load cases and chose a representative reliability index from those results.

• As we develop standardized load distributions for these analyses, our standards must somehow accommodate the fact that definitions of “best available information” for determining standardized distributions will evolve over time (see Ellingwood, 1999).

3. Standardized Reliability Analysis Procedures
While not a particularly good topic for a consensus-standard, it would be wise to provide designers with guidance on proper use of reliability analysis procedures.

• For standardization purposes, the analysis should include load effects from standard load distributions and resistance from standardized resistance distributions.

• Unless the structural application is clearly defined, the analysis of structural configuration should result in maximum permissible stresses. In other words, if a joist can span 20.3 feet prior to reaching the limit state (defined by the code-
prescribed checking equation), the analysis should be conducted at this span. Example: Some studies in the literature have computed the reliability index of the flexural limit state for joists at simple spans that are limited by a deflection check. While this is valuable information, the results of these studies can be misinterpreted by a designer or code official.

- The effects of superfluous variables should be omitted from the analysis. Superfluous variables include those with minor impacts on the analysis and those that will not generally be reliably quantified. An example of a minor variable is the addition of a distribution of member sizes with small variations from the norm. Examples of variables that will not generally be quantifiable are moisture content, temperature, and their impacts on strength properties. While these effects can be quantified in a broad sense (and are included in design on that basis), their inclusion as variables in the reliability analysis is inappropriate. Guidance on this issue should come from committees of experts.

- “Benchmark” limit states for consistent comparison of results from one analysis to another must be identified. For example, how does one compare the limit state of yield of a metal fastener in a wood-to-wood connection versus its maximum load carrying capacity? Should this judgment follow the logic used when assessing yield versus fracture in steel members?

4. Standardized Target Reliability Indices
For the reliability analyst, this is the “bottom line.” In code and standard development, the nearly universal philosophy is to match the target against current or historical practice. While this approach works well for simple limit states and individual structural members, it has significant limitations for other cases. For example, the logic of selecting target safety levels commensurate with consequences of failure would lead to increasing target reliability indices when moving from individual member design to assembly or system design to full structure design. Examine the following questions:

- Who can decide when a structure is safe enough? What criteria will be used to compare the computed $\beta$’s from various analyses? Who will determine whether certain parameters are “relevant variables” for inclusion in the analysis?

- Who will decide whether houses should have lower target $\beta$’s than auditoriums (since many more people can congregate in an auditorium) or whether houses should have higher target $\beta$’s (since so many more people spend so many more hours in their houses than in auditoriums)?

- Who will decide appropriate target $\beta$’s for systems analyzed under many various ways and compared to many different limit states? In Figure 1 (adapted from Rosowsky and Ellingwood, 1991), definitions of limit state and system size are critical to meaningful discussion of target reliability.

5. Standardized Design Implementation Procedures
Much of the early effort in LSD and LRFD has focused on proper definition of load factors and resistance factors for relatively elementary limit states. However, designers of real buildings face many issues that are not well defined in these new design formats. Coverage of the following topics is somewhat sketchy in most modern LSD and LRFD specifications -- but each decision impacts the reliability of the structure.

- The designer of a multi-story apartment building needs standardized and approved methods for combining loads from multiple sources, or for assessing analysis results that are based on single member analysis versus system analysis.

- The designer needs a complete “paper trail” that proves compliance with all legal requirements and with the state-of-the-art to properly protect public safety, and also to protect himself/herself against liability concerns.

Interpretation of the end result of a reliability analysis as “the answer” is an ever-present temptation. However, examination of a specific example is illuminating. The development of ASTM D5457 (ASTM, 1993) required the establishment of several “ground rules” for the analysis. Many design cases were examined. Many variables were considered for inclusion in the analysis. If distributions were assumed to be lognormal (like steel), $\beta$’s could be computed that were similar to those of steel. If other (more technically defensible) assumptions were made $\beta$’s were somewhat lower. Some argued that the large beneficial effects of system behavior should be included in the analyses – which would raise the computed $\beta$’s. Others argued that the detrimental effects of temporal variations in material properties should be included – which would lower the computed $\beta$’s, most dramatically for those products not subject to continuous property
reassessment (quality assurance) requirements. In the end, the committee acknowledged that one could include or exclude any number of variables – and chose to adopt the stance that, on the average, current and historical practice was likely to represent a good balance between safety and economy. This stance has likely erred on the conservative side (a good decision for the introduction of a new design format). As we enter into the second and third iterations of LSD and LRFD codes and standards, we challenge the committees to critically examine these calibration decisions to find areas where designs could be made somewhat more efficient.

**Figure 1. Conceptual System Reliability Estimates**

<table>
<thead>
<tr>
<th>Two Adjacent Members</th>
<th>Any Two Members</th>
<th>Single Member</th>
<th>Any One Member</th>
</tr>
</thead>
</table>

**SUMMARY AND RECOMMENDATIONS**

Direct application of reliability analysis to project- or product-specific design was not envisioned when initial versions of LSD and LRFD codes, standards, and specifications were being written. As we enter the new millennium, expanded teaching of structural reliability and widespread availability of reliability analysis software may be leading this field into new and uncharted territory. Before civil engineers or code authorities can properly evaluate the results of reliability analyses, they must fully understand the limitations and potential pitfalls of these methods when used in routine structural design.

The authors propose several specific steps toward the adoption of standardized reliability analysis techniques and underlying assumptions. Unless and until these issues are resolved, the authors continue to discourage the use of structural reliability analysis outside the bounds of either research or code/standard development. Once resolved, we look forward, albeit cautiously, to new and exciting applications of structural reliability in the 21st century.

**REFERENCES**


APPENDIX A. “STANDARDIZATION” USING PROCEDURES FROM A SINGLE SOURCE

The application of reliability analysis to structural wood products in Canada provides an interesting case study. In essence, the methods developed by Foschi, Folz and Yao (1989) have been adopted as de facto standards by the Technical Committee of CSA O86.1 in their draft “Standard Practice Relating Specified Strengths to Characteristic Structural Properties.” The advantages of this decision are that it formally accepts the procedures in place for sawn lumber and that it provides reasonable consistency as new products are subjected to the same procedures. The disadvantages of this decision are that some of the underlying analysis assumptions (and their design impacts) are unknown and that the procedures have not undergone the scrutiny that often accompanies consensus-based standards. This appendix reviews several steps in the reliability analysis of flexural sawn lumber members from Foschi, et al and discusses how consensus-standardization will improve future implementation.

Step 1. Select a calculation method.
- The first order reliability method (FORM) was chosen by the authors.
- **Issues:** There are no standardized procedures for this calculation method nor for the variables to be considered in the analysis.
- **Recommendation:** Develop a guide for assessment of various types of reliability analysis techniques. Develop a list of potential variables for consideration and a standard methodology for determining whether to include or exclude them from the analysis.

Step 2. Select relevant load cases for examination.
- Reliability results are provided for snow loads from a variety of locations across Canada (Vancouver, Halifax, Arvida, Ottawa, and Saskatoon), for various types of floor loads (residential, office, commercial), and for dead load. Each load type is modeled based on data from specific literature references chosen by the authors, converted to 30-year lifetime distributions using methods developed by the authors. There are no standardized procedures for selecting load cases for analysis nor for distribution types nor for distribution parameters for various structural load conditions.
- **Issues:** First, the translation of load survey data from various sources into 30-yr or 50-yr structural lifetime load distributions involves a host of complex and interrelated technical judgments. Second, structural wood products are used
in a variety of applications – thus, separate consideration of their computed reliability under various load cases (while important from an informational perspective) does little to assist in design judgments relative to assignment of appropriate material property values.

- **Recommendation**: Standards committees should recommend standardized load distributions to be used for structural reliability analysis. When the purpose of the analysis is to generate estimates of structural reliability for a given product that is used in a range of applications, a single standardized load distribution should be used. If the application is more specific, analysis under a specific load type may be appropriate. If a standards committee chooses to recommend analysis under several load distributions, they must also recommend very specific techniques for evaluating “compliance.”

**Step 3. Select relevant material property information.**

- Reliability results are computed for several lumber grades, sizes and species. Material property data are taken directly from the Canadian in-grade lumber testing program. There are no standardized procedures for selecting distribution types for analysis.

- **Issues**: As shown in the reference document, different test cells provide different estimates of reliability – even when the analysis is conducted using the “stabilizing” assumptions of low-tail data fits and consistent distributional assumptions.

- **Recommendation**: Material property information for reliability analyses should conform to a strict set of rules to eliminate spurious estimates of reliability that are simply due to manipulation of the distributional assumptions. The use of idealized distributions, as proposed by the CSA O86.1 draft practice and by ASTM D5457, eliminate “noise” from the analysis caused by random differences between data sets. If committees choose to use actual data rather than idealized distributions, aggressive grouping of like-data sets to achieve numerical stability is recommended.

**Step 4. Generate reliability analysis results.**

- Each load case (from step 1) is combined with each material data cell (from step 2) to generate a $\beta-\phi$ curve (reliability index versus resistance factor). The analysis illustrates a wide range of computed reliability for a given material depending on the load case under consideration.

- **Issues**: The presentation of an array of reliability indices, while highly useful for informational purposes, is difficult to evaluate.

- **Recommendation**: As discussed previously, the analysis should ideally be conducted under a single standardized load distribution. If multiple load distributions are used, the committee must provide guidance on how to interpret the results.

**Step 5. Interpret results: Select target reliability index and a resistance factor for design.**

- The analysis provides a range of reliability indices (at a constant resistance factor) or, conversely, a range of resistance factors (at a constant reliability index). The authors chose to characterize the results in terms of the average $\beta$ at a specific $\phi$. The range of $\beta$'s for each case is reported. There is no standardized method for determining whether the lowest $\beta$'s are safe enough or whether the highest $\beta$'s are overly conservative.

- **Issues**: The establishment of target reliability indices is a key component of the transition to Limit States Design. All facets of this decision must be examined and debated by affected parties prior to adoption.

- **Recommendations**: The use of the average $\beta$ of several products over several load cases is highly subjective. The preferred approach proposed previously is to adopt the “unified analysis” approach, wherein a single, idealized material characterization is analyzed under a single standardized load characterization. The resulting $\beta$ can be easily compared to $\beta$’s computed in the same manner for other products. An interim approach would be to replace the proposed average target $\beta$ with the minimum computed $\beta$ – which would be more consistent with the concept of setting minimum, rather than average, standards.