AMERICAN WOOD PROTECION ASSOCIATION

Observations on the Predictive Value of Short-term Stake Tests

Stan Lebow,
Bessie Woodward
Patricia Lebow
USDA, Forest Service, Forest Products Laboratory
Madison WI

ABSTRACT

This paper compares average ratings of test stakes after 3, 4, 5, and 7 years exposure to their subsequent ratings after 11 years. Average ratings from over 200 treatment groups exposed in plots in southern Mississippi were compared to average ratings of a reference preservative. The analysis revealed that even perfect ratings after three years were not a reliable indicator of good performance after 11 years. Only 53% of the treatment groups with an average rating of 10.0 after 3 years had ratings at least as great as the reference preservative after 11 years. This percentage increased to 72% for treatment groups with an average rating of 10.0 after 4 years and 84% for treatment groups with an average rating of 10.0 after 5 years, and 100% for treatment groups with an average rating of 10.0 after 7 years. The data indicate that 3 years of exposure is not sufficient to evaluate a preservative intended for use in ground contact in high hazard areas, and that even slight evidence of vulnerability after 5 years may be a predictor of inadequate future performance.

Keywords: Stake tests, ratings, test duration

INTRODUCTION

Although stake tests are commonly used to evaluate new wood preservative formulations, interpretation of stake test results continues to be problematic. It is recognized that parameters such as plot conditions (i.e., climate and soil properties), stake size, and duration of test can affect test results, but quantifying these effects when evaluating stake data remains difficult. In this paper we examine just one of these parameters, duration of test. Of particular interest is the length of the evaluation needed to obtain confidence in long-term durability or confidence in performance similar to established reference preservatives.

Prior to 2003, the AWPA standards did provide some guidance in interpreting results of stake tests, including duration of test, with the "Standard Procedure for the Calculation of the Performance Index of Preservatives in Stakes and Posts" (AWPA, 2002). However, this standard was mathematically complex and was eventually removed for lack of use. Standard E7, "Standard Method of Evaluating Wood Preservatives by Field Tests with Stakes," does not stipulate a minimum or necessary duration of exposure, leaving the decision to the discretion of the person conducting the test (AWPA, 2007). Appendix A of the AWPA standards states that a minimum of 3 years of stake data is expected in plots with high decay hazard (i.e., Hawaii or the Gulf Coast region) while longer (but unspecified) exposures are required in climates with lower decay hazard. It is left up to the discretion of the subcommittee evaluating the proposal to determine whether the length of the exposure is adequate, but 3 to 5 years of data has generally been considered to be sufficient. The subcommittee's decision on test duration is difficult because the failure pattern of a new preservative may not be known until after more extensive exposures. In this paper we discuss the relationship between stake ratings at 3-5 years and those obtained over a longer exposure period (11 years).

MATERIALS AND METHODS

Stake rating data from two of the U.S. Forest Service Forest Products Laboratory's test plots in southern Mississippi were analyzed. The location of these plots is considered to be within AWPA Deterioration Zone 5 (severe deterioration hazard, AWPA 2007). Although attack by native subterranean termites is severe at the site, colonization and attack by Formosan subterranean termites has not been reported. The two plots evaluated were installed in 1995 and 1996, and utilized 19 x 19 x 457 mm stakes with 20 replicates per treatment group. A total of 220 treatment groups are included in these two plots. They

include softwood and hardwood stakes treated with varying concentrations of both waterborne and oilborne preservatives. At each location the stakes were placed in rows with 305 mm spacing between stakes within each row and 915 mm between rows. The stakes were randomly assigned locations within the plot, and were buried vertically to half their length. An iron rod was used to create holes for stake placement. The stakes were inspected at 1, 2, 3, 4, 5, 7, 9 and 11 years after installation. At each inspection the stakes were scraped lightly to remove soil and given a visual rating for decay and/or termite attack according to the 10, 9, 8, 7, 6, 4, 0 scale as described in AWPA Standard E7 (prior to 2007).

A challenge in interpreting data from the 19 mm x 19 mm stakes is defining adequate or inadequate performance. Although it is expected that small stakes accelerate deterioration relative to commodity size material, the extent of acceleration is difficult to quantify. One way to define adequate performance is to compare durability to that of a reference preservative. In this report, we will use stakes treated with a conventional alkaline copper preservative to the ground contact retention (6.4 kg/m^3) as the comparative reference. The reference preservative stakes had an average rating of 9.10 in one of the plots and 9.35 in the other plot after 11 years. For the purposes of comparison, we used the lower value (9.10) as the measure performance of the reference preservative. Inadequate performance can also be subjective, but for the purposes of this paper inadequate performance was defined as an average group rating of 7.4 or below (average rating rounds to 7 or below). We also made the assumption that treatment groups with an average rating below 9.0 after 3 years would not be considered durable and confined the analysis to the 141 treatment groups with an average rating of 9.0 or above after 3 years.

RESULTS AND DISCUSSION

As shown in Figure 1, treatment groups with an average rating of 9 or above after 3 years of exposure exhibited a wide range of ratings after 11 years. Most treatment groups with average ratings below 9.6 after 3 years performed poorly, with average ratings falling below 7 after 11 years. Treatment groups with an average rating of 10.0 (all 20 stakes rated as perfect) after three years faired better, but only about half of these groups performed as well as the reference preservative (average rating of 9.10) after 11 years (Figure 2), and the average rating of one of the groups dropped to 3.8. Thus, perfect ratings or equivalent performance to a reference preservative after three years does not provide a high degree of confidence that a test system will be performing similarly to the reference preservative over the longer term. As expected, each increase in the duration of the test to 4, 5 or 7 years (Figure 2) improved the likelihood of performance similar to the reference preservative at 11 years. Over 70% of the treatment groups with an average rating of 10.0 after 4 years had ratings at least as great as the reference preservative after 11 years, and this percentage increased to 84% for treatment groups with an average rating of 10 after 5 years.

The collection of treatments resulted in a wide variety of mean stake evaluation patterns that are expected to be characteristic of future preservative stake evaluations. Using this collection as a database to sample from, 95% nonparametric bootstrap confidence intervals were constructed. One thousand bootstrap simulations, each of original size n=220, chosen with replacement, that repeated the above selection methods were run. Because of the selection methods, each simulation's sample is a random sample of the 141 treatment groups but with the simulation sample size a random variable. The confidence intervals generated using this methodology are shown in Figures 2 and 3.

It is notable that performance drops substantially for groups with ratings below 9.8 after 5 years (Figures 1, 3). Only 4% of treatment groups with average ratings between 9.8 and 10 after 5 years had average ratings drop below 7.4 after 11 years. However, over 40% of groups with average ratings between 9.6 and 9.8 after 5 years suffered declines to average ratings of 7.4 or below after 11 years (Figure 3). The percentage of poor performers increased to 80% for treatment groups with average ratings below 9.2 after 5 years. This indicates that even small differences in average ratings within the 9 - 10 range are important in predicting future performance. As shown in Figure 3, the greatest uncertainty in future performance is associated with groups that have average ratings in the range of 9.2 - 9.6 after 5 years. It is worthwhile noting that all of the preservatives evaluated in this plot are considered to be relatively resistant to leaching.

The interpretation of data from 19 x 19mm stakes is challenging because we do not fully understand the relationship between the performance of these small stakes relative to posts, poles or other commodity size members placed in ground contact. In this paper we chose 11 years as our indicator of "long-term" performance simply because that was the extent of data available. However, we do not know how specific

ratings in these plots after 11 years will compare to the performance of commodity size members in this or other locations. Until we better understand this relationship it will be difficult to determine a minimum testing period, or to establish acceptable minimum ratings at the end of a specified testing period.

CONCLUSIONS

This review of stake data indicates that when evaluating preservatives intended for use in ground contact in high hazard areas, a minimum of at least 5 years of exposure data is needed. It also indicates that the average rating of the test preservative should be at least as high as that of the reference preservative after 5 years. Even slight evidence of vulnerability after 5 years appears to be a strong indicator of poor future performance. It is recognized that these findings are likely to be specific to the plots evaluated and certainly are a function of the characteristics of the treatment groups. Still, it appears that caution is warranted in attempting to predict long term durability from just a few years of exposme data. In less severe decay hazard locations even longer exposures will be needed to provide confidence in long term durability and especially to predict long term performance in climates more severe than that of the test site.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the assistance of Apolonia Bocanegra in data entry and processing. The plots evaluated in this paper were established by Rod DeGroot and Doug Crawford, who have both subsequently retired.

REFERENCES

1. AWPA. 2002. Book of Standards. American Wood-Preservers' Association, Granbury, Texas.
2. AWPA. 2007. Book of Standards. American Wood-Preservers' Association, Birmingham, Alabama.

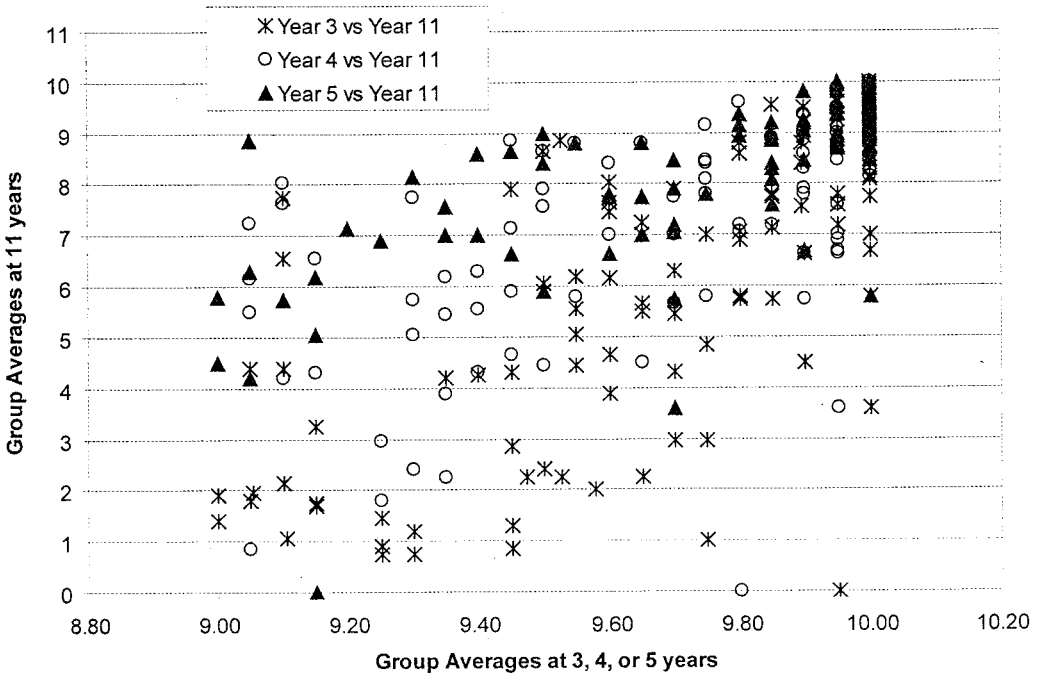

Figure 1. Relationship between a treatment group's average rating a 3, 4, or 5 years and its average rating after 11 years.

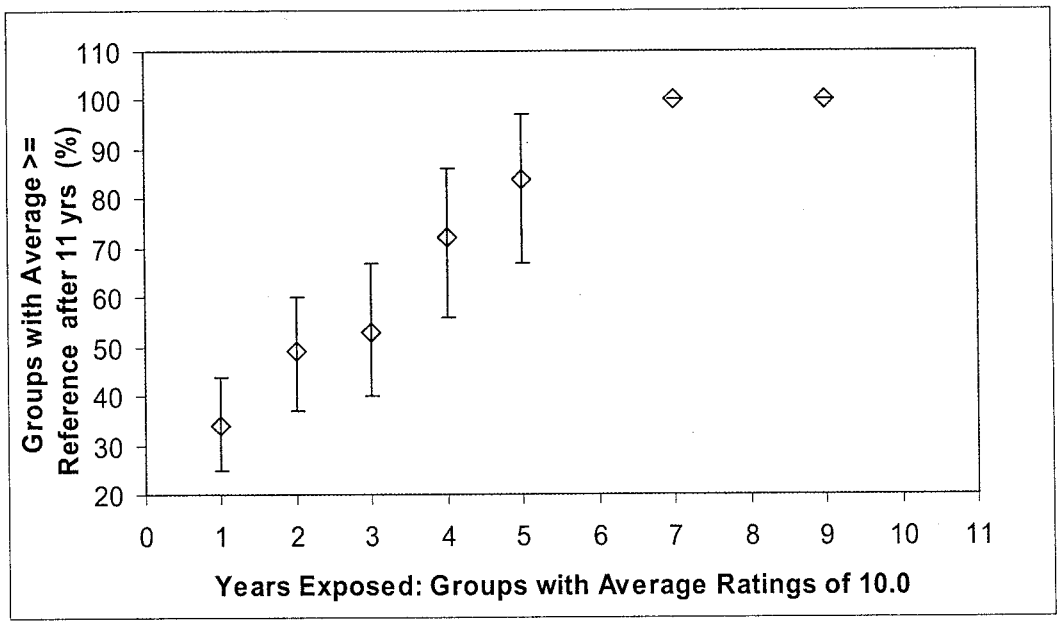

Figure 2. Percent of treatment groups with average ratings of 10.0 at a designated year that have a rating greater than or equal to the reference preservative after 11 years. Bars show 95% confidence intervals.

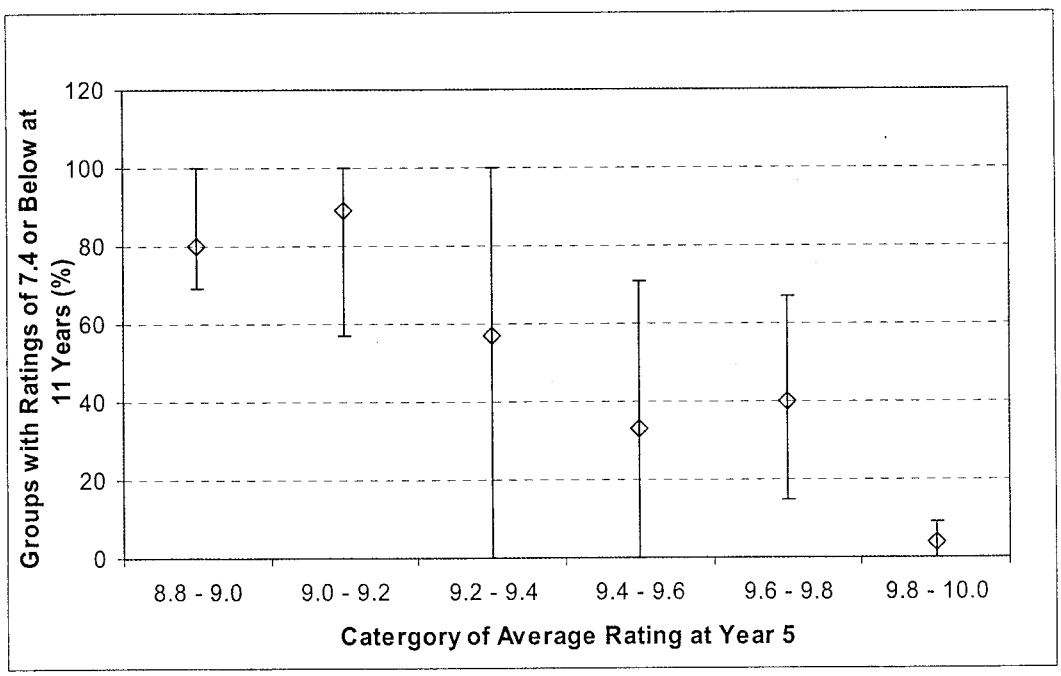

Figure 3. Relationship between average rating of treatment groups at 5 years and percent of treatment groups with average rating dropping below 7.4 after 11 years. Bars show 95% confidence intervals.

PROCEEDINGS

One Hundred Fourth Annual Meeting

of the

AMERICAN WOOD PROTECTION ASSOCIATION

Portland Marriott Downtown Waterfront
Portland, Oregon
May 18-20, 2008

VOLUME 104

AMERICAN WOOD PROTECTION ASSOCIATION
POBOX 361784 • BIRMINGHAM, ALABAMA 35236-1784 • USA

ISSN 0066-1198

Colin McCown, *Editor*
Beth Williams, *Assistant Editor*