

Estimating Erosion Risk on Forest Lands Using Improved Methods of Discriminant Analysis

JACK LEWIS AND RAYMOND M. RICE

Pacific Southwest Forest and Range Experiment Station, Forest Service, U.S. Department of Agriculture, Arcata, California

A population of 638 timber harvest areas in northwestern California was sampled for data related to the occurrence of critical amounts of erosion ($>153 \text{ m}^3$ within 0.81 ha). Separate analyses were done for forest roads and logged areas. Linear discriminant functions were computed in each analysis to contrast site conditions at critical plots with randomly selected controls. Bootstrapping was used extensively in the development and testing of the equations, in estimating prediction bias, and in placing confidence limits around parameters and posterior probabilities. The resulting three-variable equations had classification accuracy, corrected for prediction bias, of 77.7% for road plots and 69.2% for logged area plots. The use of linear discriminant functions facilitates the explicit consideration of erosion risk when planning land-disturbing activities.

INTRODUCTION

Excess erosion resulting from logging and forest roads has long been of concern to forest managers and the general public. Such concerns have prompted a number of studies in California. The findings of these studies [Dodge *et al.*, 1976; Rice and Datzman, 1981; McCashion and Rice, 1983; Peters and Litwin, 1983] were somewhat different, but they all concluded that most of the erosion occurring on timber-harvesting areas was from large mass wasting events found on a small fraction of the disturbed sites. Peters and Litwin [1983] proposed that because more than 85% of the erosion they measured was from sites yielding at least $100 \text{ yd}^3 \text{ ac}^{-1}$ ($189 \text{ m}^3 \text{ ha}^{-1}$), identifying those sites was the key to reducing erosion from logged areas and forest roads. Recommendations [Western Ecological Services Company, 1983] growing out of that study led to the definition of a critical site as any 2-acre (0.81-ha) square area enclosing more than 200 yd^3 (153 m^3) of erosional voids. This study develops a method for estimating the risk that logging or road construction will produce a critical site. Two data sets were analyzed: one from logged areas and one from forest roads (including landings).

DATA COLLECTION

Study Population

All of the sites used in the study were on private timberland in northwestern California (roughly, north of 37°N and west of 122°W). Field measurements were made between May 1985 and December 1986. The study population was composed of all timber harvest plans (THPs) on which logging was completed between November 1978 and October 1979 inclusive (638 THPs covering 22,922 ha). A THP in this context refers to a geographic area within which logging operations were performed under a written timber harvest plan as required by California forest practice regulations. Because of market conditions at the time, 1978-1979 was a peak period of logging activity and thus provided a large study population. These THPs were recent enough to see the

This paper is not subject to U.S. copyright. Published in 1990 by the American Geophysical Union.

Paper number 90WR00378.

logging disturbance and evaluate causes of failures, although brush had grown quite high in many areas. The 7-8 years which had elapsed after logging provided considerable time for loss of root strength following tree cutting [Ziemer, 1981; Kitamura and Namba, 1966; O'Loughlin, 1974] and for weaknesses in the planning or execution of the THP to be revealed by logging- and road-related mass wasting events.

Sample Selection

Access was requested from each landowner whose identity and address could be determined. The responses resulted in a sampling frame of 415 THPs (17,233 ha) to which landowners granted access. Most of the THPs to which we did not obtain access were small nonindustrial holdings.

In 1984, all of the THPs had been classified by forest practice inspectors as "critical" (thought to contain a critical site), "questionable," or "noncritical." "Critical" and "questionable" THPs are a somewhat higher proportion in THPs to which we were granted access than in the general population. Thus it seems unlikely that uncooperative landowners were refusing access because of erosion problems.

Critical sites. The sample of critical sites was based on a field reconnaissance of (1) all THPs which had been classified as "critical" or "questionable," and (2) a random sample of "noncritical" THPs. All critical sites found in logged areas became critical plots in the study. In the road analysis, however, we expected to find more critical sites than we could measure, so they were sampled. Sampling had to begin before the entire population was enumerated, so a selection probability was set at two-thirds (our best guess), and decisions to sample critical sites were based upon random numbers drawn when sites were encountered. The sampling plan called for obtaining at least 40 critical and 40 control plots for each analysis; we finished with 106 road critical plots as our selection probability turned out to be higher than needed.

Noncritical sites. Control plots were established at non-critical sites sampled at random from all THPs using a two-stage variable probability sampling method. Standard sampling with probability proportional to (THP) size [Raj, 1968] was not possible because we were still in the process of contacting landowners during the first field season. Consequently, we did not know the composition of the sampling

frame (i.e., to which THPs we would have access and how much noncritical area they contained). The sampling method we used was nearly equivalent to simple random sampling and allowed us to proceed with field work without violating the requirement that each noncritical site in the population have an equal probability of selection. The first stage in selecting control plots, known as SALT (sampling at list time [Norick, 1969; Thomas, 1985]), determined how many plots, if any, each THP was to be assigned. A second stage of simple random sampling was required to locate plots on logged areas within the THPs.

In SALT sampling, a list of random numbers is created in advance, based on the anticipated values of an auxiliary variable related to the variable of primary interest. As the auxiliary values become available, the random numbers determine which population units should be sampled.

Logged areas. For logged plots, the auxiliary variable was THP logged area from the THP work completion report (a surrogate for the area in noncritical sites, since only a very small proportion of any THP is occupied by critical sites). An upper bound (Y^*) exceeding the expected total number of permitted acres (landowner access granted) in the population was estimated in advance and a list of uniform random numbers between 0 and Y^* was then generated. When access to a THP was granted, the logged area of the THP was added to the cumulative total of logged area to which access had been granted. If the cumulative area interval defined by the addition of that THP contained n random numbers, then the THP was assigned n plots, which were located using random coordinates in a second stage of sampling.

Roads. For road plots, the auxiliary variable was length of haul roads used or reconstructed in the THPs. This value was determined from map measurements and by reading the THPs. A second stage of sampling was not required for road control plots. Instead, a set of rules for making an ordered traversal of a road network was used to associate numbers with road locations on the selected THPs. Then plot locations were determined by the random number positions within the intervals defined by selected THPs in the cumulative summation of road lengths.

Data Collection

The nominal sampling unit in the study was a 90-m square plot. Control plots were centered on a randomly selected point. Critical plots containing a single event of at least 153 m³ were centered on the cavity left by the erosion. If the critical plot was determined by two or three erosional features (none contained more than three), the plot's center was the midpoint of a line segment or centroid of a triangle defined by the centers of these features.

Rigid standards were established to ensure that each plot was a valid sample from its subpopulation (e.g., a road control plot). These standards related to the amount and timing of disturbance, the timing and causes of erosion, and topography in the plot. The following gives some of the more important standards:

1. Harvest plots were required to have visible evidence of yarding in or through the plot. If stumps were the only evidence, at least three trees must have been felled and yarded from the plot.
2. Failures that were part of erosion features predating the THP were considered critical sites if 153 m³ of additional void volume formed during or after the THP operations.

3. Erosion caused by a human disturbance other than logging or road building (e.g., mining, public roads, or building site development) could not qualify as a critical site.

4. Erosion from natural causes could qualify a site as critical, because naturally unstable areas may be aggravated by logging.

5. Areas relogged since 1979 were excluded from plots.

6. Terrain draining in a substantially different direction (defined by a ridge or first-order stream) than the plot center was excluded from plots.

As a result of deleting various portions of the plot due to such adjustments, a few plots containing less than 0.4 ha were disqualified from the study and 54% of the remaining plots had boundary adjustments. This reduced the average plot area from the nominal 0.81 ha to 0.72 ha.

In total, 172 variables were measured on each plot. Beside the field measurements, these included a number of variables determined from the THP report, county soil surveys, soil-vegetation surveys, and topographic, geologic, and precipitation maps. Only 44 of the 172 variables were intended for statistical analysis, and 12 of these were later dropped due to problems encountered in the field or in the analyses. The remaining 32 predictors are listed and defined in the notation section.

Field procedures. Each plot was investigated by two crews. The "classification crew" categorized THPs as non-critical or critical by locating critical sites. Classification crews were composed of two persons trained in forestry, geology, or hydrology. Crews worked independently throughout the study area, with membership rotating every two weeks to maintain consistency in methods and judgments. In addition to finding critical sites, the classification crews documented the reconnaissance, established plot boundaries, and measured and mapped site features such as roads, skid trails, landings, ridges, streams, and erosional features displacing 10 m³ or more of material. Debris flows and debris slides were responsible for the majority of the erosional features and volume (Table 1).

The second crew was an experienced "interdisciplinary team" composed of a forester, a geologist, and a soil scientist. They conducted an intensive analysis of each plot established by the classification crews and collected data on the topography, hydrology, geology, soils, vegetation, and management of the plot.

Point measurements at control plots (and at critical plots with only one erosional feature) were made at or near the plot center. Soil and geologic characterizations of critical sites were made at exposures representing the material that failed. However, when there was more than one feature of at least 10 m³ on the plot, a measurement feature was selected at random with probability proportional to void volume. This procedure provided characterizations of erosional events approximately in proportion to their contribution to total erosion.

Sample sizes. Since there were not enough critical sites in the interior forests to conduct separate analyses, interior plots were combined with plots on the coastal slopes. Stratified samples of control plots were chosen to represent both areas in proportion to their respective sizes (Table 2).

ANALYSIS

A three-step model selection procedure (Figure 1) was bootstrapped (bootstrapping is described below) before se-

lecting the final models. The steps in the bootstrapped procedure were (1) elimination of all but 12 sets of variables using automatic interaction detection (AID) [Sonquist *et al.*, 1973] (a "set" may be a single variable or a group of categorical dummy variables such as geologic type), (2) selection of the five best models of sizes one to five using all-possible-subsets regression with a dummy dependent variable, and (3) selection of the best discriminant model of each size based on their jackknifed classification accuracy (JCA). The data set was augmented with unrelated random variables (see the notation section) to help expose variables selected because of spurious correlations. Bootstrapping facilitated selection of an overall best model and provided estimates of variance and of the models' prediction bias. It also enabled us to place confidence limits about erosion risk estimates.

Automatic Interaction Detection

AID is an exploratory data analysis procedure developed at the University of Michigan's Institute for Social Research [Sonquist *et al.*, 1973]. AID produces a binary decision tree (essentially a dichotomous key) for predicting a continuous or dichotomous variable and is capable of uncovering very general kinds of predictor interactions. It is best used as a preliminary technique to other methods [Green, 1978]. We found it to be an excellent tool for screening out variables before performing all-possible-subsets regression, which is generally impractical to run with more than about 12 to 16 predictors.

The AID algorithm partitions a sample by locating a cutpoint on a selected predictor that best divides the sample into distinct groups (e.g., into critical and control plots). In the program we developed, each predictor is treated as an ordinal scale variable. Thus if a variable takes nine values in the sample, the cutpoint could be at any of eight locations. The process is repeated on each of the resulting groups in turn until some stopping criterion such as the total number of splits is exceeded. At each step, the split (group, predictor, and cutpoint) is chosen that maximizes the among-groups sum of squares of the dependent variable over the resulting more homogeneous groups.

AID analysis is a unique screening tool because each predictor is tested over numerous subgroups of the data as well as over the entire sample. In techniques such as

TABLE 2. Sample Sizes

| | Logged Areas | | Roads and Landings | |
|--------------------|--------------|---------|--------------------|---------|
| | Critical | Control | Critical | Control |
| Coastal plots | 46 | 34 | 96 | 39 |
| Interior plots | 5 | 15 | 10 | 15 |
| Total for analysis | 51 | 49 | 106 | 54 |

stepwise discriminant analysis that assume an additive model, a factor may be important for some subgroup of the sample, but if the factor is ineffective for the sample as a whole it will be bypassed. With AID, however, if a predictor can account for a substantial fraction of the variance over any of the various subgroups created by the partitioning process, that predictor will be retained.

The number of variables and sets to be considered at the all-possible-subsets regression phase was reduced to 12 after examining each variable's explanatory power over every subgroup created during the AID partitioning process. A variable's explanatory power over a subgroup was measured as the reduction in sum of squares of the dependent variable that could be attained by splitting on that variable. The criterion for keeping a variable was its maximum possible explanatory power over any subgroup. Sets of variables were evaluated according to the power of the best variable in the set.

All-Possible-Subsets Regression

Capitalizing on the fact that linear regression may be used to calculate discriminant functions [Flury and Riedwyl, 1982], all-possible-subsets regressions were computed with an adaptation of a program called WINNOW [Norick and Sharpnack, 1977]. WINNOW treats sets of categorical variables as inseparable and uses Mallows' C_p [Daniel and Wood, 1971] to select "best" models. The five "best" models of each size up to five sets (for a total of 25 models) were retained for further analysis.

Discriminant Analysis

Two-group discriminant analysis has been used effectively in several past studies of the slope stability classification problem [Rice and Pillsbury, 1982; Furbish and Rice, 1983; Rice *et al.*, 1985; Rice and Lewis, 1986]. The theory and application of this method are well known [Fisher, 1936]. Our methodology is unique, however, in the model selection process and in its use of bootstrapping.

Discriminant functions were computed for each of the 25 WINNOW models. Then, each function's JCA (jackknifed classification accuracy) was estimated using the procedure first presented by Lachenbruch and Mickey [1968]. In computing the JCA, individual observations are successively omitted from the calculation of a discriminant function. That function is then used to classify the omitted observation, continuing the procedure until all observations have been classified. The JCA of the model is the percentage of the observations classified correctly by these functions. The model of each size with the highest JCA was retained for evaluation as a best model.

The JCA typically increases with model size up to a point,

TABLE 1. Distribution of Erosion Among Failure Types

| Failure Type* | Critical Road Plots | | Critical Harvest Plots | |
|------------------------------|----------------------|----------------------|------------------------|----------------------|
| | Percentage by Number | Percentage by Volume | Percentage by Number | Percentage by Volume |
| Debris flow | 17.0 | 18.4 | 35.3 | 45.4 |
| Debris slide | 43.4 | 31.5 | 47.1 | 41.7 |
| Earthflow | 2.8 | 21.0 | 2.0 | 0.6 |
| Slump | 12.3 | 4.1 | 2.0 | 0.8 |
| Gully | 9.4 | 2.8 | 0.0 | 0.0 |
| Translational/ rotational | 6.6 | 18.2 | 3.9 | 7.2 |
| Rotational | 0.9 | 0.6 | 2.0 | 0.8 |
| Translational | 7.5 | 3.4 | 3.9 | 1.9 |
| Stream bank erosion | 0.0 | 0.0 | 3.9 | 1.5 |

*Categories adapted from Bedrossian [1983]. See Appendix.

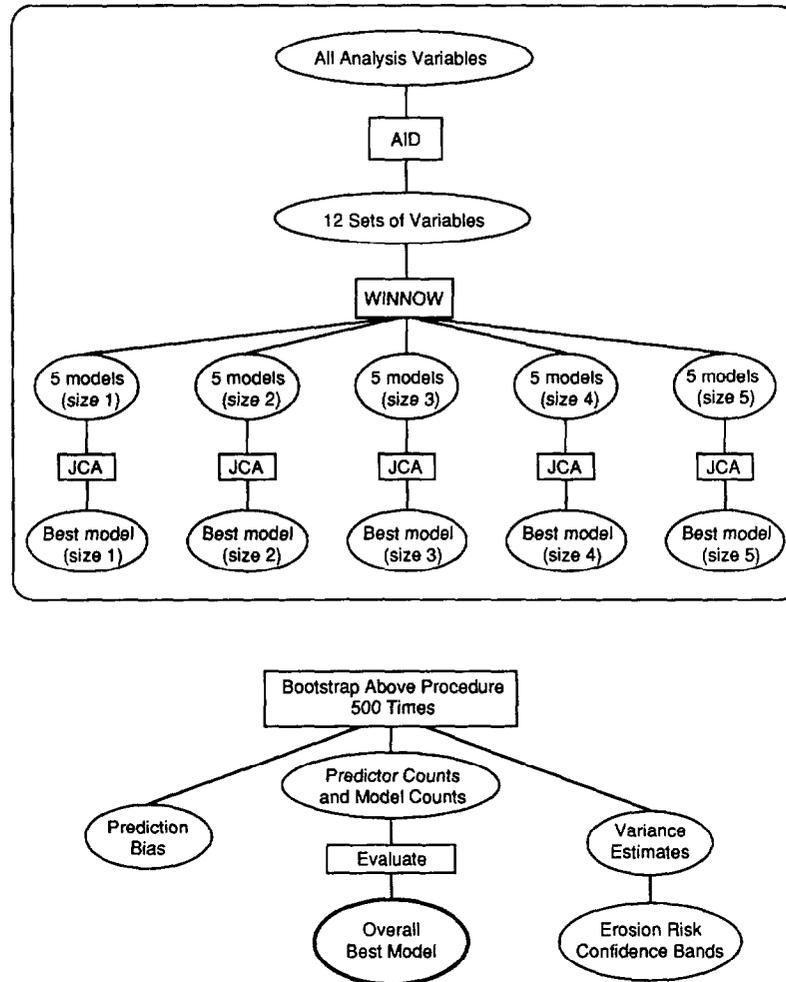


Fig. 1. Flow charts showing the discriminant model selection and evaluation procedure.

then tapers off or declines with the addition of more variables. Past experience [Furbish and Rice, 1983; Rice et al., 1985; Rice and Lewis, 1986] has shown that there is a tendency to overfit when adding variables which provide only marginal improvement to the model. Therefore choice of the best size model was necessarily subjective, depending on the JCAs and the stability of the models as demonstrated by bootstrapping (see next section). The selected model usually was the smallest model of several with similar JCAs.

Bootstrapping

Bootstrapping was used to determine the reliability of the various models. Would a slightly different sample have resulted in a different set of variables? If not, how precise are the coefficients? How reliable are estimates of the models' power to discriminate between stable and unstable sites? It is possible to explore such properties numerically through a technique called bootstrapping [Efron and Gong, 1983; Efron, 1982]. The bootstrap performs substantially better than the jackknife and cross validation for complicated prediction models [Gong, 1986] and is more versatile. The method mimics the process of selecting many samples from the population without measuring new data. Confidence intervals associated with a wide variety of parameter estimates from bootstrap samples usually closely match those

from real samples [Diaconis and Efron, 1983]. Bootstrapping also does not depend upon any distributional assumptions about the data.

Bootstrap samples are generated by drawing observations from the full sample with replacement. The bootstrap samples are the same size as the full sample but may lack some observations and contain multiple copies of others. To minimize simulation error, we used a balanced bootstrap [Davison et al., 1986], in which each observation is constrained to occur equally often in the aggregate of all bootstrap samples.

Estimators of interest (such as classification accuracy for stable and unstable sites) were calculated for each bootstrap sample, and their sample variance was calculated. Since the variable selection process was repeated for each bootstrap sample, the variance due to variable selection was accounted for as well as that due to sampling.

Stability of the model specification itself was examined by looking at the frequency with which each variable or suite of variables was chosen. Stable models include only variables which are consistently selected, so we used the bootstrap selection frequencies to help choose the final models.

After selection of each final model, another bootstrap was performed in which the process was constrained to the same variables for each bootstrap sample. This bootstrap pro-

TABLE 3. Classification Accuracy of Final Discriminant Models

| Site Type | Jackknifed Accuracy | Apparent Accuracy | Bootstrap | | Bootstrap Corrected Accuracy |
|---------------------|---------------------|-------------------|-----------|------|------------------------------|
| | | | s.d.* | Bias | |
| <i>Roads</i> | | | | | |
| Critical | 84.9 | 84.9 | 3.7 | 3.2 | 81.7 |
| Noncritical | 70.4 | 74.1 | 5.3 | 4.2 | 69.9 |
| All | 80.0 | 81.2 | 3.1 | 3.5 | 77.7 |
| <i>Logged Areas</i> | | | | | |
| Critical | 82.4 | 82.4 | 5.7 | 7.3 | 75.1 |
| Noncritical | 71.4 | 71.4 | 5.8 | 8.1 | 63.3 |
| All | 77.0 | 77.0 | 3.8 | 7.8 | 69.2 |

*Standard deviation.

duced estimates of sampling variance for the discriminant function coefficients.

Overall Best Models

Two similar three-variable functions were selected as the most effective discriminators of erosion risk. The discriminant function for forest roads was

$$DS = -0.0281 - 0.1142 \cdot \text{SLOPE} + 22.91 \cdot \text{H CURVE} + 1.0075 \cdot \text{HUE} \quad (1)$$

and for logged areas it was

$$DS = 5.032 - 0.1633 \cdot \text{SLOPE} + 20.69 \cdot \text{H CURVE} - 1.215 \cdot \text{WEAKROCK} \quad (2)$$

In these models, the discriminant score (DS) increases with slope stability. Both models seem stable and efficient. Tables 3 and 4 show the bootstrap statistics for the selected models. No other road model had a greater JCA than the 80.0% achieved with (1). Equation (2)'s JCA of 77.0% was exceeded in larger models, but we chose the three-variable function as a guard against over fitting the data, since only three variables consistently entered the bootstrap models.

Equation (1) occurred 264 times in 500 road bootstrap trials. The next most frequent three-variable model occurred only 28 times. Equation (2) occurred only 74 times, but its closest three-variable competitor occurred only 16 times and

the most frequent four-variable model occurred only 18 times. The individual variables in the two equations showed even greater robustness. The variables in the road function occurred 486, 363, and 381 times, respectively, in three-variable models. The next most frequent road variable appeared in only 47 three-variable models. Comparable figures for the logged area equation were 383, 181, 233, and 67. Similarly, the selection frequencies in four-variable models do not expose a fourth consistent variable for either roads or logged areas. In summary, the bootstrap seems to say that (1) and (2) are the best discriminant models that could be developed from our variables.

Bootstrap Estimate of Prediction Bias

Bootstrapping also provides several methods of estimating the prediction bias of the model [Efron, 1983]. We used the ordinary bootstrap estimator of bias. For that estimator, the difference is calculated between the classification accuracy of each bootstrap model on the full sample and on its bootstrap sample. The average of these differences is the ordinary bootstrap estimator of overoptimism (bias) for the classification accuracy.

Efron examined bootstrap estimates of bias in linear discriminant analysis, using small samples from simulated populations. He found that the ordinary bootstrap estimator of bias was usually an underestimate of the true bias, especially in highly overfitted situations, such as fitting a five-parameter model to samples of 14 observations. When fitting a two-parameter model to samples of size 20, downward bias was not exhibited by the bootstrap estimator.

Ours was a much more complicated analysis, in which we generally selected a three-parameter model from a sample containing 100 to 160 observations of 29 to 35 variables. We decided to continue Efron's investigations with a problem closer to our own. We simulated a multivariate normal population with the same means and covariance structure as the road data set. We performed 100 trials of sampling 50 observations of 29 predictors from each of two discriminant groups. In each trial, we selected a three-variable model using the three-step algorithm of Figure 1, bootstrapped it, and calculated its true prediction error rate (on the known population) and its apparent error rate (as measured by the sample). The mean difference between the 100 true and

TABLE 4. Coefficients of Final Road and Harvest Area Models

| Variable | Discriminant Coefficient | | | | Structural Coefficient† |
|---------------------|--------------------------|--------------|----------------|----------------|-------------------------|
| | Raw | Standardized | Bootstrap Mean | Bootstrap s.d. | |
| <i>Roads</i> | | | | | |
| CONSTANT | -0.0281 | | -0.0317 | 0.9636 | |
| SLOPE | -0.1141 | -1.125 | -0.1191 | 0.0242 | -0.660 |
| H CURVE | 22.9089 | 0.784 | 24.4221 | 6.7185 | 0.471 |
| HUE | 1.0075 | 0.842 | 1.0496 | 0.2644 | 0.624 |
| <i>Logged Areas</i> | | | | | |
| CONSTANT | 5.0317 | | 5.3054 | 1.0846 | |
| SLOPE | -0.1633 | -1.431 | -0.1725 | 0.0358 | -0.686 |
| H CURVE | 20.6886 | 0.728 | 22.2159 | 8.1749 | 0.369 |
| WEAKROCK | -1.2148 | -1.136 | -1.2968 | 0.3518 | -0.468 |

*Standard deviation.

†Simple correlation of predictor with the discriminant scores.

apparent error rates is the mean true bias, to which we compared the mean of the ordinary bootstrap bias estimates.

The mean ordinary bootstrap estimate of bias was 0.0562, compared to the mean true bias of 0.0584, and the mean squared error (MSE) of the estimator was 0.0023. Over 93% of the MSE was due to variation in the true bias or covariance between the true and estimated bias. The apparent downward bias of 0.0022 is inconsequential, contributing only 0.000005 to the MSE. The standard deviation of the bootstrap bias estimate was 0.0113 and only 35% of that was due to within-trial variation. (The mean of the standard errors calculated from individual bootstraps was 0.0039.) Thus for a given sample the standard error of the bootstrap bias estimate is about 7% of its mean. These simulation results give us confidence that for our data the ordinary bootstrap is a reliable method of estimating the bias which might result from using the same data to develop the models and estimate their classification accuracy.

The overall bootstrap-corrected accuracy of (1) is 77.7%, adjusted from an apparent accuracy of 81.2% (Table 3). The bootstrap corrected accuracy of (2) is 69.2%, adjusted from an apparent accuracy of 77.0%. Both of these corrected values seem high enough to make the functions useful tools for estimating the erosion risks associated with logging and road construction in the forests of northwestern California.

Posterior Probabilities

Posterior probabilities are the probabilities that new sites being classified by a discriminant function belong to a particular population (in our case, that they are critical sites). Before classifying a new site with the discriminant function, Bayesian classification procedures require us to consider the prior probability of a critical site. The prior probability is usually taken to be the relative frequency of such sites in the entire population. Prior probabilities are especially important in this problem, because the probability of a random site being critical is very low. In calculating classification accuracy up until now, we have considered that critical and control sites were equally likely in the population. The discriminant cutpoint was at 0, halfway between the mean discriminant scores of the two groups. Knowing that noncritical sites are much more common than critical sites, we should move the cutpoint closer to the smaller group in order to reduce the probability of misclassifying members of the larger group. Basing decisions on the posterior probabilities automatically makes such adjustments for the priors. The posterior probability (PP) of a critical site, given a particular discriminant score (DS), is

$$PP = [1 + ((1 - p_c)/p_c) \exp(DS)]^{-1} \quad (3)$$

where p_c is the prior probability of a critical site. (This formula applies only when DS is computed without normalization [Green, 1978, p. 176].) The cutpoint in many problems is chosen at a posterior probability of 0.50, because that choice minimizes the overall expected probability of misclassification. However, the cutpoint might be chosen at a different probability if the costs of misclassification or benefits of correct classification are not of equal magnitude for all outcomes. When the costs and benefits can be quantified, the cutpoint may be objectively determined. Alternative methodologies are described in the discussion and by Rice *et al.* [1985].

The prior probabilities of a critical site on a road or harvest area were determined by the proportion of area occupied by critical sites. For example, 21.5 km (an estimated population total of 239 sites \times 90 m) in critical road sites out of an estimated 1212 km in the population gives a prior probability of 0.0177 for roads. Applying a similar procedure to the area of logged plots gives a prior probability of 0.0050.

We used the bootstrap to calculate DS standard deviations for each observation in the data set from which the model was developed. For any observation, the DS was calculated once for each bootstrap model generated. The parameters in the model were free to change, so the variance is over many different model specifications. The discriminant scores from the final models, plus or minus one standard deviation, were projected onto the PP axis using the above equation and plotted at the DS computed from the full data set to show the expected variability of PP at a given DS (Figure 2). Clearly, the errors associated with risk estimates increase markedly as the estimated risk increases.

DISCUSSION

Through this investigation we have developed a methodology that can be used to better balance the interests in forest products and environmental protection. We have provided objective functions which can be used to compute the probability that logging or road construction on a given site in northwestern California will cause severe erosion. Before using such functions as guides, however, it is important to understand their capabilities and limitations.

Physical Basis

Within the limitations imposed by the form of the analyses, we are confident that the variables in the discriminant functions are valid surrogates for physical processes affecting the occurrence of critical sites. The variables in the models seem reasonable measures of risk for mass movements. Slope is crucial to the balance of forces promoting and resisting failure at a site. Horizontal curvature is very likely indexing two important aspects of slope failures: the accumulation of potentially unstable amounts of colluvium in swales [Dietrich and Dunne, 1978] and the convergence of subsurface water, leading to destabilizing pore water pressures [Sidel, 1986]. The third variable in both of the equations was some expression of the strength of the soil or rock at the site. The association of soil color with slope stability is probably related to soil drainage and age. Mottled subsurface g horizons, with matrix colors of low chroma and generally yellow hues, are commonly found in poorly drained soils. Such soils were identified with 88% of the debris flows occurring in Redwood National Park during 1981 and 1982 [LaHusen, 1984]. In contrast, well-drained soils, especially on stable older surfaces, often have a reddish hue due to the formation of unhydrated iron oxides. In summary, the variables in the equations are reflections of the parameters necessary for slope stability computations, and we consider it unlikely that the models are based on spurious correlations.

Climatic Effects

Granting that the variables are appropriate, what other extraneous influences should be considered? The weather

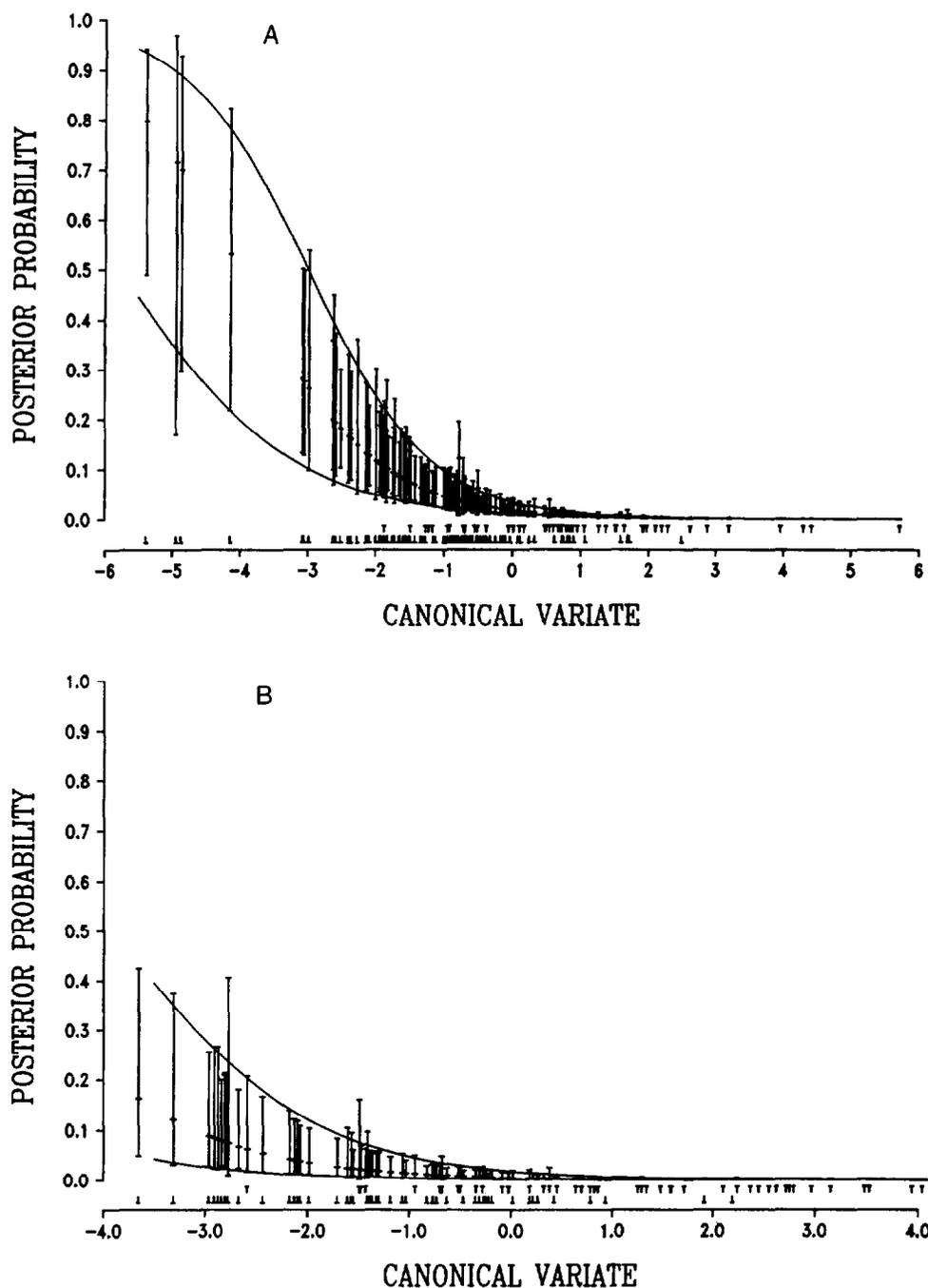


Fig. 2. Posterior probability of critical erosion: (a) road plots, in which the prior probability is 0.0177 and (b) logging area plots, in which the prior probability is 0.005. Vertical bars represent posterior probability plus or minus one bootstrapped standard deviation. Confidence bands were fitted with logistic regression. Beneath the confidence bands, T symbols represent the location of control plots on the discriminant axis and inverted T symbols indicate critical plots (erosion $>153 \text{ m}^3$ within 0.81 ha).

during the study period (i.e., after harvest or road construction and before the site visits) is probably the condition most capable of skewing the study results. If the plots had been subjected to exceptionally low environmental stresses, the models would probably have underestimated risks of critical erosion. If, on the other hand, there had been very severe weather during the study period, the models could have identified factors associated with relatively stable sites and

overestimated erosional hazards. It appears that neither of these conditions prevailed.

The estimated mean annual precipitation during the study period at study THPs was compared to estimated long-term averages. The study period mean was 11.5% greater than the long-term average (with 95% confidence limits of $\pm 2.1\%$).

Short period rainfall intensities are probably a better index of erosional stress than annual precipitation [Sidel, 1986].

One of our variables was the return period of the largest 24-hour storm occurring in the study period at the nearest precipitation station. Although the estimated return periods spanned a range from 0.5 years to 38 years, the median return period of the largest storm occurring on THPs during the study period was only 2.2 years. The true return periods, however, would be larger since we were forced by spotty rainfall records to use maximum calendar-day rainfall in place of maximum 24-hour precipitation.

To get an idea of what the true return periods might be, we collected data from weather stations in the study area which had records of both maximum 24-hour and daily precipitation. We gathered 83 data pairs from three of the wettest months during the study period (February 1980, January 1982, and February 1986). The mean ratio of maximum 24-hour to maximum daily precipitation for these data was 1.188, with a standard error of 0.021. Multiplying maximum post-harvest calendar-day rainfall by 1.188 increased its median by 2.5 cm and increased its median return period from 2.2 to 5.0 years. Thus most of the study plots appear to have experienced only slightly smaller stresses than might be expected over a "typical" 7- to 8-year period.

The return-period variable had very little predictive value in the analyses. In the road analysis bootstraps, its frequency of occurrence in the best models was 20th out of 29 variables and it occurred about one-fifth as often as one of the random variables. In the harvest area analyses, the return period variable was tied for last place and was well behind all of the random variables.

In summary, mean annual precipitation was slightly higher than usual, while maximum storm intensities were a little below normal; there was a poor correlation between maximum (24-hour) storm return period and the occurrence of critical sites. These results indicate that the study was not appreciably influenced by abnormal weather.

Sampling

Discriminant analysis requires that both groups be a simple random sample from their respective populations. Although the control sites very nearly constituted a simple random sample, the sample of critical sites was stratified by the forest practice inspectors' "critical," "questionable," and "noncritical" THP classifications. Consequently, we estimate that about 78% of the critical sites on "critical" and "questionable" THPs were sampled, but only about 45% of the critical sites on "noncritical" THPs. Critical sites from "noncritical" THPs were therefore underrepresented in the sample. If they are different from other critical sites, then the sample is biased.

To test this hypothesis statistically, we contrasted critical plots on "critical" and "noncritical" THPs. A T^2 test [Hotelling, 1931] based on 56 variables produced a probability of 0.047 that the multivariate group means were equal. Individual variables with the most significant differences were ANNLPPT, CABLE, TRACTOR, HUE, and MM. Since soil type and precipitation were used in the erosion hazard rating (EHR) system at the time, and logging method is probably correlated with EHR, it appears that the inspectors who made the preliminary classifications of the THPs generally classified THPs as "critical" when they contained areas meeting the criteria of the EHR in use at the time. That rating, however, has been found to be a poor predictor of erosion [Rice and Datzman, 1981].

We continued the investigation by conducting two additional analyses, one for roads and one for harvest areas. For these analyses, unbiased data sets were created by randomly discarding 30 critical road plots and 16 critical harvest plots from "critical" and "questionable" THPs. The discriminant analysis model selection procedure (Figure 1) was then run and bootstrapped for each of these new samples.

The same variables were selected as for the full samples, and the bootstrap strongly corroborated their selection. Moreover, the discriminant coefficients were all within one bootstrap standard deviation of their values in the full models. Chow's test [Chow, 1960] showed no significant differences between the models ($p > 0.50$ in both tests). Thus critical plot sampling bias appears to have had minimal effect on the results.

Acceptable Risk

A major difficulty to be overcome in the use of the predictive equations is establishing levels of "acceptable" risk marking the boundary between critical and stable sites. For any threshold used to trigger corrective actions, some percentage of critical and noncritical sites are misclassified (Table 5). The costs of such errors include erosional damage at undetected critical sites and the treatment costs or loss of timber revenues where corrective action is erroneously applied to stable sites. The choice of a threshold should be based on the economic, social and political ramifications of expected outcomes.

If the payoffs of various outcomes are quantified, then a simple method is available for computing a decision cutoff value on the discriminant axis. Although such evaluations may be subjective, we believe that use of such quantities will lead to improved decisions because of the discipline imposed by an explicit value system. For example, consider the following matrix in Table 6 displaying payoffs arising from interactions of predicted and actual outcomes at a logging site.

Mitigation effectiveness may be defined as

$$E = \frac{D - d}{V - v} - 1 \quad (4)$$

and let $C(n|c)$ be the cost of misclassifying a potential critical site, and $C(c|n)$ be the cost of misclassifying a noncritical site.

It is assumed that $V > v$, $D > d$, and that all quantities are positive. E will take only positive values for sensible mitigation measures whose cost ($V - v$) is less than the averted environmental cost ($D - d$). Also, if the type of mitigation is varied, d should decrease as v decreases. To maximize the expected payoff, the quantities in the matrix should be based on what is thought to be the most effective mitigation for the site(s) being considered. When potential environmental consequences are great and modifying the operation is unlikely to affect the probability of failure, the most effective mitigation is to avoid disturbing the site. For the no-action alternative, we assume both $v = 0$ and $d = 0$. (Note that if $D < V$, then $E < 0$ for the no-action alternative; but the no-action alternative would not be a sensible mitigation measure in such a situation.)

Given a particular site, the misclassification cost is simply the difference between the payoffs of correctly and incorrectly classifying the site. Thus

TABLE 5. Decision Table

| Discriminant Score* | Posterior Probability of Critical Site† | Percentage Classified Correctly | |
|---------------------------|---|---------------------------------|-------------|
| | | Critical | Noncritical |
| <i>Roads and Landings</i> | | | |
| -6.0 | 0.88 | 0 | 100 |
| -5.0 | 0.73 | 0 | 100 |
| -4.0 | 0.50 | 1 | 100 |
| -3.6 | 0.40 | 2 | 100 |
| -3.2 | 0.31 | 5 | 99 |
| -2.8 | 0.23 | 9 | 99 |
| -2.4 | 0.17 | 15 | 98 |
| -2.0 | 0.12 | 23 | 97 |
| -1.8 | 0.098 | 28 | 96 |
| -1.6 | 0.082 | 33 | 94 |
| -1.4 | 0.068 | 39 | 93 |
| -1.2 | 0.056 | 45 | 91 |
| -1.0 | 0.047 | 51 | 89 |
| -0.8 | 0.039 | 57 | 86 |
| -0.6 | 0.032 | 63 | 84 |
| -0.4 | 0.026 | 69 | 81 |
| -0.2 | 0.022 | 74 | 77 |
| 0.0 | 0.018 | 79 | 73 |
| 0.2 | 0.015 | 83 | 69 |
| 0.4 | 0.012 | 86 | 65 |
| 0.6 | 0.0098 | 89 | 60 |
| 0.8 | 0.0080 | 92 | 56 |
| 1.0 | 0.0066 | 94 | 51 |
| 1.2 | 0.0054 | 96 | 46 |
| 1.4 | 0.0044 | 97 | 41 |
| 1.6 | 0.0036 | 98 | 37 |
| 1.8 | 0.0030 | 98 | 32 |
| 2.0 | 0.0024 | 99 | 28 |
| 2.4 | 0.0016 | 100 | 21 |
| 2.8 | 0.0011 | 100 | 15 |
| 3.2 | 0.00073 | 100 | 10 |
| 3.6 | 0.00049 | 100 | 6 |
| 4.0 | 0.00033 | 100 | 4 |
| 5.0 | 0.00012 | 100 | 1 |
| 6.0 | 0.00004 | 100 | 0 |

| | | | |
|----------------------|---------|-----|-----|
| <i>Harvest Areas</i> | | | |
| -6.0 | 0.67 | 0 | 100 |
| -5.0 | 0.43 | 0 | 100 |
| -4.0 | 0.22 | 1 | 100 |
| -3.6 | 0.16 | 3 | 100 |
| -3.2 | 0.11 | 6 | 99 |
| -2.8 | 0.076 | 10 | 99 |
| -2.4 | 0.052 | 17 | 98 |
| -2.0 | 0.036 | 26 | 97 |
| -1.8 | 0.030 | 31 | 96 |
| -1.6 | 0.024 | 37 | 95 |
| -1.4 | 0.020 | 42 | 93 |
| -1.2 | 0.016 | 48 | 92 |
| -1.0 | 0.013 | 55 | 90 |
| -0.8 | 0.011 | 61 | 87 |
| -0.6 | 0.0091 | 66 | 85 |
| -0.4 | 0.0074 | 72 | 82 |
| -0.2 | 0.0061 | 77 | 79 |
| 0.0 | 0.0050 | 81 | 75 |
| 0.2 | 0.0041 | 85 | 71 |
| 0.4 | 0.0034 | 88 | 67 |
| 0.6 | 0.0027 | 91 | 63 |
| 0.8 | 0.0023 | 93 | 58 |
| 1.0 | 0.0018 | 95 | 54 |
| 1.2 | 0.0015 | 96 | 49 |
| 1.4 | 0.0012 | 97 | 44 |
| 1.6 | 0.0010 | 98 | 40 |
| 1.8 | 0.00083 | 99 | 35 |
| 2.0 | 0.00068 | 99 | 31 |
| 2.4 | 0.00046 | 100 | 23 |
| 2.8 | 0.00030 | 100 | 17 |
| 3.2 | 0.00020 | 100 | 11 |

TABLE 5. (continued)

| Discriminant Score* | Posterior Probability of Critical Site† | Percentage Classified Correctly | |
|----------------------------------|---|---------------------------------|-------------|
| | | Critical | Noncritical |
| <i>Harvest Areas (continued)</i> | | | |
| 3.6 | 0.00014 | 100 | 8 |
| 4.0 | 0.00009 | 100 | 5 |
| 5.0 | 0.00003 | 100 | 1 |
| 6.0 | 0.00001 | 100 | 0 |

*Based on the equation $DS = -0.02807 - 0.1142*SLOPE + 22.91*HCURVE + 1.0075*HUE$ for roads and landings and based on the equation $DS = 5.032 - 0.1633*SLOPE + 20.69*HCURVE - 1.215*WEAKROCK$ for harvest areas.

†Assumes the prior probability of a critical site = 0.0177 for roads and landings and 0.0050 for harvest areas.

$$C(n|c) = (v - d) - (V - D) = D - d - (V - v)$$

$$C(c|n) = V - v$$

According to Bayes' rule [Green, 1978], the discriminant cutpoint (t_c) which minimizes the expected cost is

$$\begin{aligned}
 t_c &= \ln \left(\frac{P_c}{1 - P_c} \frac{C(n|c)}{C(c|n)} \right) \\
 &= \ln \left(\frac{P_c}{1 - P_c} \frac{(D - d) - (V - v)}{V - v} \right) \\
 &= \ln \left(\frac{P_c E}{1 - P_c} \right) \tag{5}
 \end{aligned}$$

where P_c is, as before, the prior probability of a critical site. The cutpoint increases as the mitigation effectiveness increases, resulting in increased numbers of mitigated sites. This tends to happen when the potential reduction in environmental damage is great relative to the costs of mitigation. If E is near $(1 - P_c)/P_c$ (usually a large number), then t_c is close to 0 and approximately equal percentages of critical and noncritical sites are classified correctly (Table 5). As potential environmental consequences decrease relative to the costs of mitigation, t_c decreases until nearly all sites are classified noncritical. Because the prior probabilities of either road or logged area critical sites are so low, overall accuracy will always be approximately equal to stable site accuracy.

Although site specific applications of the discriminant

TABLE 6. Payoffs Arising From Interactions of Predicted and Actual Outcomes at a Logging Site

| Actual Outcome | Predicted Outcome | |
|----------------|-------------------|----------|
| | Noncritical | Critical |
| Noncritical | V | v |
| Critical | $V - D$ | $v - d$ |

V is value (profit) received from an unmitigated operation, v is value (profit) received from a mitigated operation, D is cost of environmental damage and repairs from an unmitigated operation, and d is cost of environmental damage and repairs from a mitigated operation.

functions would be ideal, a broader approach can be taken. If some general decision probability is chosen, the expected accuracy will still be known, but the extent to which benefits and risk are appropriately balanced at individual sites will be unknown. Since an optimal balance will not be achieved at most sites with the broader approach, it will reduce the potential benefits of erosion risk assessment.

Implementation

These models are strictly applicable only to the timberlands of northwestern California. For other areas, it would be wise to collect the data necessary to develop applicable regional predictive equations. The following implementation guidelines are recommended where discriminant models have been developed to identify potential critical sites when planning forest road construction and timber harvesting.

1. Determine the prior probabilities. These may be based upon the population used to construct the predictive model, or, if priors are available from subregions (such as coastal or interior forests), use these more specific probabilities.
2. Evaluate the economic, social, and political values of the four outcomes arising from identifying and failing to identify potential critical and noncritical sites. Determine what probabilities of misclassification are acceptable for potential critical sites and noncritical sites. Such evaluations are most valuable when done on a site-specific basis. If completing a payoff matrix, base profit estimates on the most cost-effective mitigation measures for the site. For a site with highly valued resources at risk, this may be the no-action alternative.
3. Choose a decision cutpoint compatible with the explicit value system arrived at in the previous step, using (4) and (5), tables such as Table 5, or the method described by Rice *et al.* [1985].
4. Evaluate the discriminant function using the appropriate predictive equation.
5. If the discriminant score exceeds the decision cutpoint, proceed with the operation using standard methods applied in the population where the model was developed.
6. If the discriminant score is less than the decision cutpoint, the operation should be restricted. Serious consideration should be given to avoiding any disturbance of the site, since similar sites used to develop the model failed when disturbed, and mitigation measures were undoubtedly applied to some of these. If the site must be disturbed, extraordinary precaution should be used.

CONCLUSIONS

The variables in the equations we have developed appear to be expressing three important site conditions related to erosion risk: the force of gravity promoting instability (slope steepness), the convergence of subsurface water (horizontal curvature and soil color), and strength of materials (soil color and parent rock strength). The road discriminant function has bias-corrected prediction accuracy of 69.9% for stable sites and 81.7% for critical sites. The bias-corrected accuracy of the harvest area function is 63.3% for stable sites and 75.1% for critical sites. These numbers, together with the plausibility of the variables in the equations, suggest these models can provide satisfactory estimates for the risk of large erosion events.

The equations are quite simple and their variables fairly easy to estimate in the field. Foresters routinely estimate slope, and horizontal curvature can be computed from two measurements of azimuth and distance. Some training in evaluating rock strength and soil color may be necessary. The calculation of the discriminant score or the posterior probability of a critical site can be done on a pocket calculator. Alternatively, tables have been developed to estimate approximate posterior probabilities.

None of the steps required for the utilization of the equations are particularly difficult, but they do require a new, more rigorous approach to erosion risk analysis. The difficult evaluations required to choose decision cutoff values can be viewed as a strength of the method. To a degree not normally achieved, the use of risk evaluations with discriminant models will expose the contending values in forestry decisions. It will make it possible to examine various balances between them by estimating the effects of alternative decisions.

APPENDIX: DEFINITIONS

Mass Movement Types (Adapted From Bedrossian [1983])

Rotational. A deep-seated landslide with a curved slide plane and a somewhat cohesive slide mass. Rotational slides are characterized by a steep head scarp above an intact slide mass that leads to an irregular, hummocky toe area. The slope steepness of the intact slide mass is reduced relative to the surrounding terrain, and the slope direction may be reversed, which leads to the development of sag ponds.

Translational. A deep-seated landslide with a linear slide plane and a somewhat cohesive slide mass. Translational slides are characterized by vertical cracks in the ground surface at the upslope margin of the slide and hummocky downslope terrain.

Translational/rotational. A combination of slide movements commonly involving rotational headward movement with downslope translational or earthflow transport.

Slump. This is a localized failure that is commonly rotational and associated with roads. It may occur on the fill slope or cut slope. It varies in size but the material usually does not move very far.

Earthflow. A slow to rapid movement of saturated soil and debris in a semiviscous, highly plastic state. Earthflows are characterized by irregular, hummocky terrain with grassland or meadow vegetation.

Debris slide. A relatively shallow-seated translational landslide that originates in unconsolidated rock, colluvium, or soil materials. Debris slides are characterized by a steep head scarp and jumbled toe deposits (when present).

Debris flow/torrent track. An erosion channel located on sideslopes or in streambeds that is formed by an extremely fast mass movement of water-laden soil, rock, and vegetative debris. Debris tracks are characterized by scoured channels that commonly originate at debris slide or fill failure sites.

Gully. An erosion channel resulting from the detachment and transport of soil and rock fragments by concentrated flows of surface runoff water. Gullies generally originate at a point of concentrated runoff water and are characterized by V-shaped or trapezoidal channel cross section shapes.

Undifferentiated. Small mass movement events, such as

cut-bank failures, without well-defined slide mechanisms or that do not fit any of the previously described failure types.

Parent Rock Strength

Plastic. Very low strength.

Friable. Specimen crumbles easily by rubbing with fingers.

Low. An unfractured specimen will crumble under light hammer blows.

Moderate. Specimen will break apart with moderate hammer blows.

High. Specimen will withstand a few heavy ringing hammer blows before breaking into large fragments.

Very high. Specimen will resist heavy ringing hammer blows and will yield with difficulty only dust and small flying fragments.

The test for rock strength is made on in-place specimens found at the nearest representative exposure of parent material. It is patterned after the strength element of the "Unified Rock Classification System" (URCS) [Williamson, 1984] that is concerned with rock as an engineering material. This system is directed toward evaluating weaker rocks whose strength differences might affect slope stability. That fact and the sampling problems associated with obtaining a truly representative sample led us to relax the URCS requirements concerning the type of hammer. We doubt that such refinements are warranted for this application. In fact, the flat end of a geologist's hammer was used in this study.

NOTATION

The following are categorical dummy variables (determined in the field).

Bedrock origin class

| | |
|----|---|
| FM | Franciscan Melange. |
| FC | Franciscan coherent (non-Melange Franciscan). |
| HS | hard sedimentary. |
| IE | igneous extrusive. |
| II | igneous intrusive. |
| MM | metamorphic. |
| SS | soft sedimentary. |

Forest type

| | |
|----|--------------------------------------|
| DF | pure Douglas fir. |
| HW | hardwoods and minor redwood/fir. |
| MC | mixed conifers. |
| PI | Ponderosa pine. |
| RW | redwood with or without Douglas fir. |

Logging method

| | |
|---------|-------------------------------|
| CABLE | cable yarded. |
| TRACTOR | tractor yarded. |
| NOYARD | not yarded (road plots only). |

Other field measurements

| | |
|----------|---|
| AMTLOG | estimated percentage of overstory crown cover removed. |
| ASPECT | azimuth in degrees; coded for analysis: N = 1, NE = 2, NW = 3, E = 4, W = 5, SE = 6, SW = 7, S = 8. |
| CURCOVER | current percent woody plant crown cover |

| | |
|----------|---|
| | (estimated overstory + understory); may exceed 100%. |
| DSURFH20 | minimum of STRDIST and WDIST. |
| HCURVE | horizontal curvature (m^{-1}) (concave is negative); this was computed as the inverse of the radius of a circle circumscribing a triangle determined by three points along a uniformly curved segment of the topographic contour line. |
| HUE | subsoil moist Munsell hue coded 1 to 5 from yellow to red (1 = 5Y; 2 = 2.5Y; 3 = 10YR; 4 = 7.5YR; 5 = 5YR.); see SUBSCLY for definition of subsoil. |
| PREOVER | adjacent (presumed preharvest) percent overstory crown cover (visual estimate). |
| PRETOTAL | PREOVER + PREUNDER (may exceed 100%). |
| PREUNDER | adjacent (presumed preharvest) percent understory crown cover (visual estimate). |
| ROCKSTR | parent rock strength (see definitions): 1 = plastic, 2 = friable, 3 = low, 4 = moderate, 5 = high, 6 = very high. |
| SLOPE | slope steepness in degrees; average of uphill and downhill slopes. |
| SLPBRK | Distance in meters upslope to nearest convex vertical slope break; defaulted if none or >91 m. |
| SOILAD | average soil depth (cm) to firm or unweathered bedrock; defaulted if >152 cm. |
| SOILCOMP | soil competency (numeric rating (0–100) based on wet consistence, dry strength, dilatancy, and toughness); SOILCOMP = $100 \times (WC + DS + DIL + TUF)/16$; where WC = wet consistence: very plastic = 0; plastic = 8; slightly plastic = 8; not plastic = 0. DS = dry strength: high = 2; med = 1; slight = 1; none = 0. DIL = dilatancy: none = 3; very slow = 2; slow = 1; quick = 0. TUF = toughness: high = 3; med = 1; slight = 1; none = 0. Wet consistence was determined according to "Soil Taxonomy" [USDA Soil Conservation Service, 1975, p. 176]. Dry strength, dilatancy, and toughness were determined according to the Unified Soils Classification System [Waterways Experiment Station, 1953]. All evaluations were based on subsoil (see SUBSCLY). |
| SOILRD | depth in centimeters to the first soil layer or condition substantially restricting the downward movement of water; defaulted if >152 cm. |
| STRDIST | slope distance (along downhill gradient) to nearest stream; measured from a topographic map if >91 m. |
| STRELEV | elevation change in meters to nearest stream; measured from topographic map if STRDIST >91 m. |
| SURFSND | percentage sand in <2-mm fraction of the surface soil horizon that lies between the soil surface and the subsoil (see SUBSCLY). |

- SUBSCLY** percentage clay in <2-mm fraction of subsoil (defined as the *B* horizon, if it occurs; otherwise it is the material that underlies the surface soil and which is differentiated from it by a change in texture, color, structure, or percent coarse fragments).
- VCURVE** vertical slope curvature (m^{-1}) (concave is negative); this was determined in the same manner as HCURVE except: (1) the points determining the triangle were taken along a downhill gradient, and (2) on road plots, curvature was averaged from measurements taken above the road cut and below the sidecast or fill.
- WDIST** slope distance in meters to nearest spring or pond; defaulted if elevation change >61 m.
- WEAKROCK** dichotomy for parent rock strength (see definitions): 1 = plastic, friable, or low; -1 = other.
- WELEV** elevation change in meters to nearest spring or pond; defaulted if elevation change >61 m.

Office measurements

- ANNLPT** mean annual precipitation, mm.
- LATDIS** lateral dissection (number of first-order drainages per 2400 m on a contour passing through the plot center).
- PPT1HR** five-year 1-hour precipitation, mm.
- PPT24HR** five-year 24-hour precipitation, mm.
- PPTRP** 24-hour return period (years) of maximum postharvest calendar-day precipitation (24-hour precipitation records were not generally available); this was determined from isopluvial maps for the nearest precipitation station with complete postharvest records.
- SLOPOS** slope position; ratio of ridge-plot slope distance to ridge-stream slope distance. Random variables included in analyses
- LOGNORM** lognormally distributed random variable.
- NORMAL** standard normal random variable.
- UNIFORM1** uniformly distributed random variable on (0, 1).
- UNIFORM2** uniformly distributed random variable on (0, 1).

Other notation

- AID** automatic interaction detection (a data-partitioning algorithm used for screening variables).
- DS** discriminant score.
- JCA** jackknifed classification accuracy.
- MSE** mean squared error.
- P_c prior probability of a critical site.
- PP** posterior probability (of a critical site).
- SALT** selection at list time (a variable probability sampling method).
- THP** timber harvesting plan, or the area covered therein.

- WINNOW** program for computing all-possible-subsets regression.

Acknowledgments. This study was conducted in cooperation with the California Department of Forestry and Fire Protection. The department provided the major funding, as well as initial planning, and supplied timber harvest plans. We also wish to thank the many individual landowners and industrial forestry organizations that granted us access to their lands and assisted our field crews in locating plots. The study would not have been possible without their cooperation.

REFERENCES

- Bedrossian, T. L., Watershed mapping in northern California. *Calif. Geol.*, 36(7), 140-147, 1983.
- Chow, G. C., Tests of equality of sets of coefficients in two linear regressions, *Econometrics*, 28(3), 591-605, 1960.
- Daniel, C., and F. S. Wood, *Fitting Equations to Data*, 342 pp., John Wiley, New York, 1971.
- Davison, A. C., D. V. Hinkley, and E. Schechtman, Efficient bootstrap simulation, *Biometrika*, 73(3), 555-566, 1986.
- Diaconis, P., and B. Efron, Computer-intensive methods in statistics, *Sci. Am.*, 249(5), 116-130, 1983.
- Dietrich, W. E., and T. Dunne, Sediment budget for a small catchment in mountainous terrane, *Z. Geomorphol.*, 29, suppl., 191-206, 1978.
- Dodge, M., L. T. Burcham, S. Goldhaber, B. McCully, and C. Springer, An investigation of soil characteristics and erosion rates on California forest lands, report, 105 pp., Calif. Div. of For., Sacramento, 1976.
- Efron, B., Estimating the error rate of a prediction rule: Improvements on cross-validation, *Tech. Rep. 78*, 46 pp., Stanford Univ. Div. of Biostatistics, Stanford, Calif., 1982.
- Efron, B., Estimating the error rate of a prediction rule: Improvement on cross-validation, *J. Am. Stat. Assoc.*, 78(382), 316-331, 1983.
- Efron, B., and G. Gong, A leisurely look at the bootstrap, the jackknife and cross-validation, *Am. Stat.*, 37(1), 36-48, 1983.
- Fisher, R. A., The use of multiple measurements in taxonomic problems, *Ann. Eugenics*, 7, 179-188, 1936.
- Flury, B. H., and H. Riedwyl, R^2 -tests, linear two-group discriminant analysis, and their computation by linear regression, *Tech. Rep. 82-42*, 36 pp., Purdue Univ. Dep. of Statistics, West Lafayette, Ind., 1982.
- Furbish, D. J., and R. M. Rice, Predicting landslides related to clearcut logging, northwestern California, U.S.A., *Mt. Res. Dev.*, 3, 253-259, 1983.
- Gong, G., Cross-validation, the jackknife, and the bootstrap: Excess error estimation in forward logistic regression, *J. Am. Stat. Assoc.*, 81(393), 108-113, 1986.
- Green, P. W., *Analyzing Multivariate Data*, 519 pp., Dryden, Hinsdale, Ill., 1978.
- Hotelling, G., The generalization of Students ratio, *Ann. Math. Stat.*, 2, 360-378, 1931.
- Kitamura, Y., and S. Namba, A field experiment on the uprooting resistance of tree roots (I), in *Proceedings of the 77th Meeting of the Japanese Forestry Society*, pp. 568-570, 1966. (English translation by J. M. Arata and R. R. Ziemer, USDA Forest Service, Arcata, California, 1976.)
- Lachenbruch, P. A., and M. R. Mickey, Estimation of error rates in discriminant analysis, *Technometrics*, 10(1), 1-11, 1968.
- LaHusen, R. B., Characteristics of management related debris flows, in *Symposium on Effects of Forest Land Use on Erosion and Slope Stability*, edited by C. L. O'Loughlin, and A. J. Pearce, pp. 139-146, Environment and Policy Institute, University of Hawaii, Honolulu, 1984.
- McCashion, J. D., and R. M. Rice, Erosion on logging roads in northwestern California: How much is avoidable?, *J. For.*, 81(1), 23-26, 1983.
- Norick, N., SALT (selection at list time) sampling: A new method and its application in forestry, M.A. thesis, 50 pp., Univ. of Calif., Berkeley, 1969.
- Norick, N., and D. Sharpnack, WINNOW: An all possible subsets

- regression computer program, report, 26 pp., USDA For. Serv., Pac. Southwest For. and Range Exp. Stn., Berkeley, Calif., 1977.
- O'Loughlin, C. L., The effects of timber removal on the stability of forest soils, *J. Hydrol.*, 13(2), 121-134, 1974.
- Peters, J. H., and Y. Litwin, Factors influencing soil erosion on timber harvested lands in California, report, 94 pp., West. Ecol. Serv. Co., Novato, Calif., 1983.
- Raj, D., *Sampling Theory*, pp. 47-53, McGraw-Hill, New York, 1968.
- Rice, R. M., and P. A. Datzman, Erosion associated with cable and tractor logging in northwestern California, *IAHS Publ. 132*, pp. 362-374, Int. Assoc. of Hydrol. Sci., Wallingford, England, 1981.
- Rice, R. M., and J. Lewis, Identifying unstable sites on logging roads, paper presented at 18th IUFRO World Congr., Int. Union of For. Res. Org., Ljubljana, Yugoslavia, July 21, 1986.
- Rice, R. M., and N. H. Pillsbury, Predicting landslides in clearcut patches, *IAHS Publ. 137*, pp. 303-311, Int. Assoc. of Hydrol. Sci., Wallingford, England, 1982.
- Rice, R. M., N. H. Pillsbury, and K. W. Schmidt, A risk analysis approach for using discriminant functions to manage logging-related landslides on granitic terrain, *For. Sci.*, 31(3), 772-784, 1985.
- Sidel, R. C., Groundwater accretion in unstable hillslopes of coastal Alaska, *IAHS Publ. 156*, pp. 335-343, Int. Assoc. of Hydrol. Sci., Wallingford, England, 1986.
- Sonquist, J. A., E. L. Baker, and J. N. Morgan, *Searching for Structure*, 236 pp., Institute for Social Research, University of Michigan, Ann Arbor, 1973.
- Thomas, R. B., Estimating total suspended sediment yield with probability sampling, *Water Resour. Res.*, 21(9), 1381-1388, 1985.
- USDA Soil Conservation Service, Soil taxonomy: A basic system of soil classification for making and interpreting soil surveys, *Agric. Handb. 436*, Washington, D. C., 1975.
- Waterways Experiment Station, Unified soil classification system, *U.S. Corps of Eng. Tech. Memo 3-323*, Vicksburg, Miss., 1953.
- Western Ecological Services Company, Study design for selection and characterization of critical erosion sites on timber-harvested lands and for evaluation of factors controlling erosion on forested lands, report, 115 pp., West. Ecol. Serv. Co., Novato, Calif., 1983.
- Williamson, D. A., Unified rock classification system, *Bull. Assoc. Eng. Geol.*, 21(3), 345-354, 1984.
- Ziemer, R., The role of vegetation in the stability of forested slopes, paper presented at 17th IUFRO World Congr., Int. Union of For. Res. Org., Kyoto, Japan, September 1981.

J. Lewis and R. M. Rice, Redwood Sciences Laboratory, Pacific Southwest Forest and Range Experiment Station, Forest Service, USDA, 1700 Bayview Dr., Arcata, CA 95521.

(Received June 21, 1989;
revised January 29, 1990;
accepted February 1, 1990.)

