I D E N T I F Y I N G U N S T A B L E S I T E S
O N L O G G I N G R O A D S

R. M. Rice
Principal Hydrologist

J. Lewis
Data Analyst

Pacific Southwest Forest and Range Experiment Station
Forest Service, U.S. Department of Agriculture
Arcata, California 95521 USA

INTRODUCTION Oxf.:116.5:(794)

Logging roads are an important source of forestry-related erosion. The amount of erosion on a forest road is determined by the interaction between how the road is constructed and maintained and the environment in which it is built. The roads in this study were constructed with large bulldozers, and most excavated material was sidecast. The roads studied were constructed primarily for trucks transporting logs. They averaged 11.5 years old. Available running surface averaged 5 m wide, ranging from 3 m to 9 m. About 68% of the roads were surfaced with crushed rock, 15% were unsurfaced, and 17% were paved with asphalt. The average road gradient was 4°, but some short segments were as steep as 9°.

This paper is the third in a series using the same data. The first paper (McCashion and Rice, 1983) reported on all of the erosion occurring on the study road segments and concluded that only about 60% of the erosion was caused by the roads and only about 24% could have been prevented by economically feasible engineering measures or minor changes in road alignment. The second paper (Rice and McCashion, 1985) attempted to use regression analysis to estimate road-related erosion. That attempt was unsuccessful, but a prediction method based on a linear discriminant function (Fisher, 1936) was developed which correctly identified as hazardous those road segments producing 82% of the road-related erosion. In this third paper, discriminant analysis is again used, this time contrasting conditions at the sites of large erosional features with a random sample of conditions on the study road segments.

STUDY AREA

The roads investigated in this study are in the Coast Range and Klamath Mountains of northwestern California. Study sites were from 10 to 60 km inland from the coast and extended from 40° to 41° 30' north latitude. For this study, the geologic parent materials of the area were grouped into six broad types: acid igneous. ultramafic, metamorphic. soft sedimentary, hard sedimentary, and the Franciscan Formation (APPENDIX). Over half of the sites were on either hard sedimentary rocks of the Cretaceous and Jurassic periods or on the Franciscan Formation, a highly sheared and complex unit consisting mainly of massive greywacke and interbedded shales. Slopes average 24°, and range from 0° to over 45°. Elevation ranges from sea level to 1920 m, averaging 843 m. Annual precipitation in the area averages 2000 mm, ranging between 1000 mm and 3000 mm. Snow makes up more than 30% of the precipitation only for the 25% of the study sites above 1200 m elevation.

DATA COLLECTION

Study sites were selected by using a stratified sampling design to ensure that the data spanned the ranges of important site variables and their interactions. Stratified sampling was also used to destroy naturally occurring associations between important site variables. The four stratification variables were road standard, terrain slope, geologic parent material, and 25-year/24-hour precipitation intensity.

Our stratification of the data seems to have been successful. The coefficient of determination of the regression of the common logarithm of terrain slope on the other stratification variables was 0.17. The coefficient of determination of the regression of the common logarithm of the 25-year/24-hour precipitation intensity on the other stratification variables was 0.18. The squared canonical correlations (Hotelling, 1936) of the road standard and geologic variables with the remaining variables were 0.13 and 0.28, respectively.

In previous investigations, each study site consisted of a 1.61-km road segment. The site conditions describing the road were the averages of those measured at stations at 0.32-km intervals along the road segment. Soil properties for the road were based on samples taken from a randomly chosen station. Within a study road segment, all erosional features appearing to have displaced more than 15 m^3 of soil were measured individually. Site conditions at each of these features were also recorded. The data for the present study consists of measurements taken from 107 erosional features (52 slides, 48 slumps, 7 gullies) and 108 control sites randomly chosen from the 0.32-km stations. We considered 21 variables describing the road and the site on which it was constructed (APPENDIX).

ANALYTICAL METHODS

First, the data were randomly divided in half. The first half was the developmental data, consisting of 54 erosion features and 54 control points. The second data set, including 53 erosion features and 54 control points, was used to verify equations developed with the first data set.

Two analyses were performed. In the first analysis, we investigated both site and road variables in order to gain insight into the conditions surrounding the road-related erosional features. This was our interpretive analysis. In the second analysis, we developed a prediction methodology which might be used in road design. Candidate variables for this analysis included only Road Standard and Site Variables. In this case we assumed that the designer might not know other details in the early stages of the road's design.

Both analyses used linear discriminant functions. In addition to the variables listed in the appendix, their logarithms (where appropriate) were also included as candidate variables. Some power and trigonometric functions were also tried. Logarithmic transformations were tried because other analyses (Furbish and Rice, 1983; Rice and Datzman. 1981; Rice and McCashion, 1985; Rice et al., 1985) had found logarithmic models more effective than models based on untransformed data. Other transformations were tried in the hope that they might improve the analyses. Both statistical and subjective criteria were used to choose the model to be verified with the test data set. We selected models for testing which had fewer variables than statistical tests indicated should be retained because similar analyses (Furbish and Rice, 1983; Rice et al., 1985) indicated a tendency to overfit developmental data sets.

RESULTS

Our first statistical test of the interpretive analysis was Mallows' Cp (Daniel and Wood, 1971) using an all possible subsets regression program (Norick and Sharpnack, 1977) as a surrogate for a discriminant analysis. The best equation based on that criterion included, in order of importance, log CUTHT, Geologic Parent Material, log FILLHT, Vegetation on Fill Surface, log 25YR24HR, Road Surface Material, and log COVER. The second test was a subjective comparison of the proportion of correct classifications by the discriminant functions using a jackknifed classification (Lachenbruch and Mickey, 1968). In a jackknifed classification, the classification of each observation is made using a function based on the remainder of the observations. A stepwise discriminant analysis computer program was used (Dixon and Jennrich, 1981). The classification accuracy improved substantially with each step until FILLHT had entered the equation. At that step, classification was 87.0% accurate. With the addition of the remaining variables, accuracy remained almost constant, increasing one percent then dropping to 86.1% in the last step.

We concluded that the best equation was one including only CUTHT, FILLHT, and Geologic Parent Material. Several factors led to that decision. That model had the best classification accuracy with the erosion sites, and the overall accuracy improved little with the addition of other variables. The coefficient of 25YR24HR indicated that erosional events were more frequent where rainfall was less intense, suggesting a spurious correlation and casting doubt on the significance of subsequent variables. Lastly, earlier analyses had tended to overfit developmental data sets. The cross-validation test suggests that our approach to variable retention led us to the least stable model. The accuracy of classification with the test data declined from 87% to 73.8%. The drop in accuracy was greater for this model than for models with more or fewer variables.-It appears that we should have been even more conservative in our criteria for retention of variables. The model based on CUTHT and Geologic Parent Material alone had the best accuracy with the test data set.

Our analysis of the data for a predictive model followed the same steps as the interpretive analysis. The choice of the best discriminant function, however, was easier. Mallows' Cp indicated that a model based only on SLOPE and Geologic Parent Material was best (Table 1). The jackknifed comparison yielded very similar indications. The slope and geology model had a jackknifed accuracy of 75.9%. The addition of CLAY to the model raised the accuracy 76.9% where it remained with the addition of PCTGT2MM. That accuracy declined with the addition of SAND, and it declined further to 73.1% with the inclusion of Road Standard. The performance of the slope and geology model was very consistent in the cross-validation test (Table 2). The overall accuracy was 75.8% and the errors were distributed similarly in the developmental data and test data.

Table 1. Linear discriminant function based on the analysis of site descriptors at 54 erosional features and 54 randomly chosen control sites on logging roads in northwestern California. (Note that the dummy variable FR has been decoded in this equation so that all geologic types take on a value of 0 or 1.)

Y =		(Canonical Variate)
	-2.35	(Constant)
	+0.04 SLOPE	
	-1.81 GR	Geologic Parent Material
	-0.28 FR	Geologic Parent Material
	-0.25 MM	Geologic Parent Material
	-0.12 UM	Geologic Parent Material
	+0.75 HS	Geologic Parent Material
	+1.72 SS	Geologic Parent Material

Table 2. Classification of erosion and stable sites by a linear discriminant function (see Table 1).

Actual Condition	Predicted Condition: Stable	Predicted Condition: Erosion	Correctly Classified (%)
A. Developmental Data			
Stable	39	15	72.2
Erosion	11	43	79.6
Total	50	58	75.9
B. Test Data			
Stable	39	15	72.2
Erosion	11	42	79.2
Total	50	57	75.7

DISCUSSION

In the introduction, we stated that erosion on a logging road was the result of the interaction between the road's construction and the environment in which it was built. The results of our interpretive analysis seem to support that statement. CUTHT was the variable most effective in distinguishing between erosion sites and control sites. CUTHT expresses the interaction between road width and slope. It increases as the slope becomes steeper or the road becomes wider. Geologic Parent Material is probably mainly indexing the inherent erodibility of the site. These two variables apparently provide about as much discriminating power as can be gleaned from the data which we were analyzing. This finding is the same as that of Furbish (1981) and Rice et al. (1985). In those studies, the equations giving the best classification of test data sets contained only two and three variables, respectively. It is tempting to think that the coefficients of variables entering the discriminant function in later steps are expressing real world relationships, but that may not be the case. Although a discriminant analysis may be fairly effective in separating stable sites from eroding sites, it does so by quantifing obvious simple relationships. Ours did not yield insights into subtle interactions between roads and their environments.

A chief advantage of a linear discriminant analysis is that it provides an objective estimate of the probability that an undeveloped site belongs to one of the groups from which the function was developed. Several assumptions must be met if such posterior probabilities are to be valid. In practical applications, some of the assumptions are usually violated. The two assumptions most frequently violated state that the variables are distributed multivariate normally and that the groups being classified have the same dispersion matrices. Lachenbruch (1975) and several other authors have shown that discriminant analysis is a rather robust technique which can tolerate some deviation from these assumptions. Since the sites in this analysis were selected with a regression analysis in mind, observations were chosen to span the range of the variables and to minimize the correlations among the more important variables. They should have been selected at random from the groups they represent to meet an assumption of discriminant analysis. We do not know what effect our violations of these assumptions may have on predictions using our results and therefore recommend caution. On the other hand, we believe that, in spite of its limitations, an equation such as ours is preferable to a procedure where performance cannot be evaluated objectively.

In order to compute posterior probabilities, it is necessary to know or estimate the prior probabilities of group membership. We used equal numbers of erosion sites and stable sites in our analyses. In nature, stable sites vastly outnumber erosion sites. The total length of logging roads in our study was 481 km. We estimate that the 107 erosional features we studied had a total length of about 2.9 km, making our estimate of the prior probability of an erosional feature to be 0.006. Since our discriminant function is composed of a continuous variable and a set of categorical variables, it is possible to graph the posterior probability as a family of curves depicting the effect of slope for each geologic type (Figure 1).

Two insights are evident from inspecting Figure 1. First, the probability of an erosional feature at any individual point is quite low for most conditions. Consequently, if these curves are to be of any use in evaluating erosion risk, the threshold of concern will have to be much lower than 50%.The highest posterior probability estimated for a site in our data was 35%, and the average for erosional features was only 5.25%. Second, although SLOPE has a great effect on the posterior probability, the effect of Geologic Parent Material is also profound. For example, at a slope of 66% (the average slope of the erosional features) the posterior probability of a granitic site is less than 0.1%, but it is 17% at a soft sedimentary site.

Another approach to evaluating erosion hazard is to rate road segments using the product of posterior probability and road length for segments which are relatively uniform with respect to the discriminating variables. For example, if the terrain slope along a road is approximately 70% across 500 m of hard sedimentary rock, then the posterior probability of failure is a constant 5% for the 500 m. The product, 25 m, represents the expected value for how much of the segment will include erosional features greater than 15 m^3 in volume if a road is built across the hillslope. Obviously this value should not be viewed as a prediction of failure size but as a measure of risk.

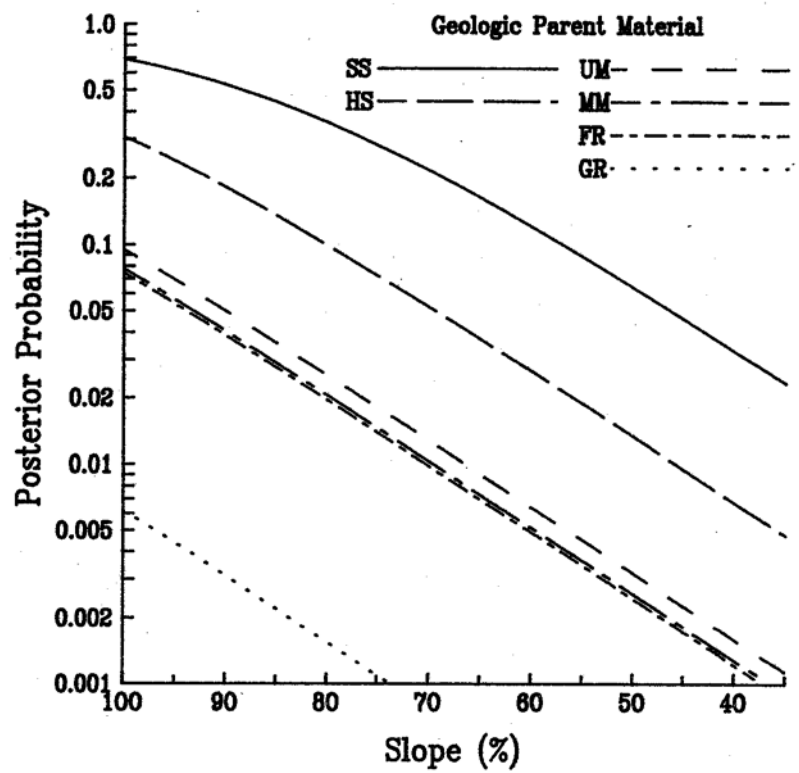

Figure 1. The probability that a site will produce a road-related erosional feature as predicted by slope and geologic parent material.

If longer roads need to be evaluated, the products can be summed for smaller segments. The sum-of-products method has the advantage of providing an intuitively meaningful hazard rating for road segments of whatever length the planner chooses as management units.

We recommend that our methodologies be only the first step in evaluating the erosion potential of a proposed road. The relationships in our equation are, after all, very simplistic. In practice, once high risk sites have been identified by their posterior probabilities, they should be evaluated by technical specialists in order to determine if the risk should be revised up or down because of the presence of risk factors other than slope and geology. In this fashion, the

objective risk appraisal of our methodology can be tempered by human judgment, hopefully benefitting from the strengths of both methods of estimating erosion risk.

SUMMARY

This paper is the third in a series concerning erosion from logging roads in northwestern California. It contrasts sites of large erosional features with randomly selected sites from 481 km of roads. The data were first divided so that relationships developed with one half could be tested on the second half. Linear discriminant functions were used to analyze the data. An analysis of both road and site variables found the best contrast to be based on cut height and geologic parent material. A prediction equation based only on site variables (slope and geology) had a 76% classification accuracy with both the developmental and test data sets. Major erosional features occupied only 0.6% of the length of roads studied. Therefore, if our function is to be useful, the threshold of concern will have to be quite low. In practice, high risk sites should be identified by their posterior probabilities and then be evaluated by technical specialists to determine if the risk should be revised up or down because of the presence of risk factors not included in our equation.

Key Words: forest roads, discriminant analysis, northwestern California, erosion.

ZUSAMMENFASSUNG

Diese Arbeit ist die dritte einer Serie über die Erosion durch forstwirtschaftliche Zufahrtswege im Norwesten Kaliforniens. Sie stellt gegenüber die Lagen von groβen erosionsartigen Bodenmerkmalen mit stichprobenartig ausgewählten Erosionsstellen auf der Strecke von 481 Sraβenkilometern. Die Werte warden zunächst durch swei geteilt damit die sich in der Hälfte entwiekelten Verhältnisse an der zweiten Hälfte geprüft werden konnten. Lineare Differentialfunktionen wurden zur Untersuchung der Data verwendet. Eine Analyse der Variablen der groβen Erosionsstellen bildeten die beste Gegenüberstellung basierend auf Schnittöhe and geologischem Quellenmaterial. Eine Vorhersagungsgleichung basierend ausschlieβlich auf Variablen der Erosionsstellen (Neigung and Geologi) hatte eine Einteilungsgenauigkeit von 76% mit Entwicklungswertegruppen und Testwertegruppen. Größere erosionsartige Bodenmerkmale befanden sich nur auf 0,6% der Gesamtlänge von den untersuchten Zufahrtswegen. Wenn daher unsere Funktion nützlich sein soll, sollten Veränderungen der Bodenstruktur mit empfindlicher Achtsamkeit verfolgt werden. In der Praxis sollten hochriskante Erosionsstellen bezüglieh ihrer Vorhersagefähigkeit kenntlich gemacht end anschlleβend von technischen Fachleuten ausgewertet werden, um zu bestimmen ob das Risiko höher oder niedriger eingestuft werden soll, aufgrund nicht in unserer Vorhersagungsgleichung enthaltenen Risikofaktoren.

LITERATURE CITED

Daniel, C., Wood, F. S., 1971: Fitting Equations to Data. Wiley-Interscience. New York. 342 p.

Dixon, W. J., Jennrich, R., 1981: Stepwise regression. Pages 251-263 in Dixon. W. J., ed., BMDP Statistical Software. Univ. of California Press, Berkeley.

Fisher, R. A., 1936: The use of multiple measurements in taxonomic problems. Ann. Eugenics 7:179-188.

Furbish, D. J., 1981: Debris slides related to logging of streamside hillslopes in northwestern California. M.S. thesis, Humboldt State Univ., Arcata, Calif. 65 p.

Furbish, D. J., Rice, R. M., 1983: Predicting landslides related to clearcut logging in northwestern California, U.S.A. Mt. Res. and Dev. 3(3):253-259.

Hotelling, H., 1936: Relations between two sets of variates. Biometrica 29:321-377.

Lachenbruch, P. A., Mickey, R. M., 1968: Estimation of error rates in discriminant analysis. Technometrics 10:1-11.

Lachenbruch, P. A., 1975: Discriminant Analysis. Hefner. New York, 128 p.

McCashion, J. D., Rice, R. M., 1983: Erosion on logging roads in northwestern California: how much is avoidable? J. For. 81(1):23-26.

Miller, J. F., Frederick, R. H., Tracey, R. J., 1973: Precipitation frequency atlas of the western United States. NOAA Atlas 2. Vol. 11, California. U.S. Dep. of Commerce, Washington, D.C. 71 p.

Norick. N., Sharpnack. D. 1977: WINNOW: an all possible subsets regression computer program. USDA Forest Service, Pacific Southwest Forest and Range Experiment Station, Berkeley, Calif. 26 p.

Rantz, S. E., 1968: Average annual precipitation and runoff in north coastal California. Hydrologic Investigations Atlas HA-298, USDI Geol Survey, Washington, D.C. 4 p.

Rice, R. M., Datzman, P. A., 1981: Erosion associated with cable and tractor logging in northwestern California. Proc International Symposium on Erosion and Sediment Transport in Pacific Rim Steeplands. IAHS Publ No. 132:362-375.

Rice, R. M., Pillsbury, N. H., and Schmidt, K. W., 1985: A risk analysis approach for using discriminant functions to manage logging-related landslides on granitic terrain. J. For. 31(3):772-784.

Rice, R. M., McCashion, J. D., 1985: Site conditions related to erosion on logging roads. Proc. International Symposium on Erosion, Debris Flow and Disaster Prevention. The Erosion Control Engineering Soc, Japan. pp. 69-74.

APPENDIX

Road Variables

Position on Slope

LOW Valley bottom and lower one-third of the slope
MID The middle one-third of the slope
UPPER The upper one-third of the slope
RIDGE The ridge (dummy variable designated by coding the above three variables as -1.)

Road Surface Material

LITE Road surfaced with 5 cm or less of gravel
HEAVY Road surfaced with more than 5 cm of gravel
PAVED Road surface sealed with a permanent hard material
NONE Road surface consists of existing natural material (dummy variable)

Road Standard

SEASON Roads used in dry season only, 3 to 4 m wide, surfaced with natural soil or thin rock, and with or without permanent drainage structures
AW2ND All-weather secondary roads suitable for year-round use, 3.5 to 5 m in drivable width, surfaced with gravel or crushed rock of moderate depth, and with drainage structures adequate for a 25-year storm
MAIN Permanent all-weather roads used as main collectors of traffic, about 9 m in drivable width, regularly maintained, heavily surfaced with gravel or pavement, and with drainage structures to accommodate a 25-year storm (dummy variable)

Vegetation on Cut Surface

CUTBARE Cut surface essentially devoid of vegetation
CUTWOOD Cut surface revegetated with woody vegetation
CUTGRASS Cut surface revegetated with herbaceous vegetation (dummy variable)

Vegetation on Fill Surface

FILLBARE Fill surface essentially devoid of vegetation
FILLWOOD Fill surface revegetated with woody vegetation
FILLGRASS Fill surface revegetated with herbaceous vegetation (dummy variable)

Amount Logged

UNCUT No logging adjacent to the road
CLEARCUT More than 60% of the timber removed
PARTCUT Some logging, but less than 60% of the timber removed (dummy variable)

GRADE The average road gradient (%)
CUTHT Road cut height (m)
FILLHT Height of road fills (m)
WIDTH Width of road operating surface (m)

Site Variables

SLOPE Average terrain slope (tangent)
ASPECT Coded 1-8 in rough approximation to microclimate (Rice and Datzman, 1981)
COVER Proportion of ground covered by the crowns of the dominant vegetation (%)
25YR24HR Maximum amount of precipitation in 24 hours resulting from a storm having a 25-year return period (cm) (Miller et al., 1973)
MAP Mean annual precipitation (cm) (Rantz, 1968)
SAND Proportion of soil composed of particles between 2 and 2×10^{-2} mm in diameter (%)
CLAY Proportion of soil composed of particles smaller than 2×10^{-3} mm in diameter (%)
SILT Proportion of soil composed of particles between 2×10^{-2} and 2×10^{-3} mm in diameter (%)
PCTGT2MM Proportion of soil composed of particles greater than 2 mm in diameter (%)

Shape of Slope

CONCAVE Average slope configuration normal to contours is concave
CONVEX Average slope configuration normal to contours is convex
UNIF Uniform slope (dummy variable)

Geologic Parent Material

GR Acid igneous rocks, mainly granitics
UM Ultramafic rocks
MM Metamorphic rocks, including metasedimentary rocks
SS Soft sedimentary rocks, mainly of Cenozoic era
HS Hard sedimentary rocks, mainly of Cretaceous and Jurassic periods
FR The Franciscan Formation, a highly sheared and complex unit consisting mainly of massive graywacke and interbedded shales (dummy variable)