

## AN INDEX OF REGIONAL SNOW-PACK STABILITY BASED ON NATURAL SLAB AVALANCHES

By ARTHUR JUDSON and RUDY M. KING

(Rocky Mountain Forest and Range Experiment Station, 240 West Prospect Street, Fort Collins, Colorado 80526, U.S.A.)

**ABSTRACT.** An index of regional snow-pack stability based on occurrences of natural slab avalanches was developed using a statistical distribution and a sequential testing procedure. The study interprets avalanche information on 185 paths in the Colorado Front Range. Results show general agreement with operational hazard estimates; test results have real-time evaluation potential.

**RÉSUMÉ.** Un indice de la stabilité régionale basée sur les déclenchements naturels d'avalanches. On a mis au point un indice régional de stabilité du manteau neigeux fondé sur les déclenchements naturels d'avalanches en utilisant leur distribution statistique et une procédure de contrôle séquentiel. L'étude interprète des informations avalanches sur 185 couloirs dans le Colorado Front Range. Les résultats

montrent une bonne concordance générale avec les estimations opérationnelles du danger; les résultats des essais ont une capacité d'évaluation en temps réel.

**ZUSAMMENFASSUNG.** Ein Index der regionalen Schneedeckenstabilität auf der Grundlage natürlicher Lawinenabgänge. Auf der Grundlage der Abgänge natürlicher Schneebrettlawinen wurde ein Index für die regionale Schneedeckenstabilität hergeleitet; dabei fand ein Verfahren der statistischen Verteilung und der sequentiellen Prüfung Verwendung. Die Studie stützt sich auf Lawineninformationen von 185 Bahnen in der Colorado Front Range. Die Ergebnisse zeigen im allgemeinen Übereinstimmung mit operationellen Gefahrschätzungen; die Versuchsergebnisse können in Echtzeit gewonnen werden.

### INTRODUCTION

Avalanches provide a qualitative clue to snow-pack stability. Fresh avalanches strongly affect snow safety decisions and stability evaluations, lend credibility to avalanche warnings, and serve as test criteria for the performance of predictive models. However, no quantitative guidelines for measuring regional snow-pack stability exist at present, and estimation of the mechanical stability of a regional snow-pack with shear frames is impractical. Since very little is known about the complex relationship between avalanches seen in a limited area and overall snow-pack stability, researchers have not found any generally acceptable means of using avalanches to measure model performance. Indices for such purposes have been restricted to the "yes-no" of Obled's "avalanche day" (Bois and Obled, 1973), to the summed mass of debris in the runout zone (Salway, 1979), or to simple combinations of number and size (Perla, 1970; Judson and Erickson, 1973; Judson and others, 1980). A better means of indexing regional snow-pack stability with avalanches will produce more accurate stability evaluations and avalanche prediction techniques. This paper presents a semiquantitative method of linking snow-pack stability with natural slab avalanches. The approach uses probability theory and a sequential probability-ratio testing procedure to evaluate existing daily snow-pack stability. Results are referenced to daily avalanche evaluations made by forecasters from the U.S. Forest Service Avalanche Warning Center in Colorado.

### OBSERVATIONS AND DATA

Avalanches were observed on 185 paths in Colorado's Front Range during the nine winters 1974-75 through 1982-83. These self-consistent, high-quality records were collected daily by experienced Forest Service avalanche personnel with help from the Colorado Department of Highways and the Urad-Henderson Mines. Good access was available by highway, over-snow vehicle, and skis. When poor visibility prevented an adequate view of avalanches in the 200 km<sup>2</sup> sample area, the record was reconstructed on the next

good-viewing day. Dates for prior events were estimated by comparing their snow surface features with those from avalanches whose occurrence time was known. The overall record had comparatively few errors, and those that did occur did so mainly during rare cases of extended poor visibility. Only natural slab avalanches that traveled more than 50 m slope distance were retained for analysis.

Most of the starting zones lie at or above timberline between 3500 and 3800 m above sea-level. Catchments are located on all aspects in multiform terrain. Winter climate is continental with mostly small but frequent snow-falls, frequent wind transport of snow, and cold temperatures (Judson, 1977). Resulting avalanches occur mainly in cold dry snow from November through March, but may involve dry, damp, or wet snow in April. Both wet and dry slab avalanches are included in this research.

### THEORY

Previous studies on the complex relationship between avalanches and snow-pack stability are limited. Recent studies by Judson (1983) and Judson and King (1983) indicate that the majority of starting zones produce avalanches independently of one another due to spatial variations in stability. The degree of independence changes from one winter to the next and is related to terrain features unique to each starting zone, avalanche frequency, snow conditions, prior avalanche patterns, and other factors. Over a period of several winters, it appears that the probability of an avalanche on a given day, is more or less unique for every path, but similar mean stability levels may be common to groups of "similar" release zones when the general snow properties are about the same on individual paths of the group. An indication of a possible structure is seen in Judson's (1983) study, where he found higher joint probabilities of avalanches between avalanche paths in separate groups than among individual paths. Groups in that study were based on proximity and terrain similarity. The authors suspected that a group structure based on avalanche frequency and snow conditions would give better results. For example, the probability of joint

avalanche release between groups with equivalent frequency will be higher than among groups with different frequencies (Judson, 1983). Probabilities of joint avalanche release should further improve if one compares frequency by snow conditions. Therefore, it seems reasonable to assume that the probability of an avalanche on paths with similar snow-pack properties and like frequencies is approximately equal, but varies with time. These two simplifying assumptions form the basis of the present study.

To avoid depleting the nine-year data base, days with like snow-pack properties were approximated in a general manner by dividing the November-April avalanche season into bimonthly segments. This allowed for the shallow and weak snow conditions of early winter, the midwinter transition period, and the deeper snow-packs of spring. Within each bimonthly period, the 185 paths of the sample were ranked by number of avalanches and assigned to classes of high, moderate, and low frequency.

Class distinctions were made arbitrarily because there were no obvious demarcation points in the frequency distribution. Paths in the first decile were assigned to the high-frequency class. The top 25% of the remaining paths were classed as moderate-frequency, and the rest of the paths became the low-frequency class. The choice of three frequency classes was subjective, based on our perception of the frequency structure. More or fewer classes could be used if dictated by patterns in the data.

The development of a regional snow-pack stability estimator is as follows: let  $X$ ,  $Y$ , and  $Z$  denote the number of avalanches in a sample of  $n_x$ ,  $n_y$ , and  $n_z$  paths, where  $X$ ,  $Y$ , and  $Z$  refer to the high-, moderate-, and low-frequency classes for each day in the bimonthly periods.  $X$  can take on values between 0 and  $n_x$ ,  $Y$  between 0 and  $n_y$ , and  $Z$  from 0 to  $n_z$ . If one assumes statistical independence among individual paths, the number of paths that fail in any frequency class should be a binomially distributed random variable. The joint probability of a realization of  $X$ ,  $Y$ , and  $Z$  is:

$$P(X = x, Y = y, Z = z) = \binom{n_x}{x} p_x^x (1 - p_x)^{n_x - x} \binom{n_y}{y} p_y^y (1 - p_y)^{n_y - y} \binom{n_z}{z} p_z^z (1 - p_z)^{n_z - z} \quad (1)$$

where  $p_x$ ,  $p_y$ , and  $p_z$  are the probabilities of an avalanche occurring on individual paths in each group on a given day. A relationship between  $p_x$ ,  $p_y$ , and  $p_z$  is then hypothesized such that for each day within a bimonthly period:

$$a_x p_x = a_y p_y = a_z p_z = p. \quad (2)$$

Since the majority of paths lie in the low-frequency class, further simplification by setting  $a_z \equiv 1$  produces

$$a_x p_x = a_y p_y = p_z \equiv p. \quad (3)$$

$a_x$  and  $a_y$  are estimated from the data. Significance of an avalanche is then an inverse function of frequency, which follows the reasoning of field personnel who assign less weight to failures on high-frequency paths and more weight to events on lower-frequency paths. This is done because the high-frequency starting zones often respond in small volume when the snow-pack on most other areas is generally stable. As will be discussed later, the data support setting  $a_z$  to one. Parameter  $p$  is the regional snow-pack stability index for the population of individual paths within the Colorado Front Range. The value of  $p$ , which varies between 0 and 1, reflects the degree of snow-pack stability in the region. Letting  $b_x = 1/a_x$  and  $b_y = 1/a_y$ ,

$$p_x = b_x p, \quad p_y = b_y p, \quad \text{and} \quad p_z \equiv p. \quad (4)$$

For a realization of  $X$ ,  $Y$ , and  $Z$ , the maximum likelihood estimator of  $p$  for the region is

$$\hat{p} = \frac{x + y + z}{b_x n_x + b_y n_y + n_z}. \quad (5)$$

Hence, when avalanches are visible,  $p$  can be estimated regardless of whether failures occur on high-, moderate-, or low-frequency paths. A more complete picture of snow-pack stability in the population could be given after  $p$  is estimated by obtaining  $\hat{p}_x$  and  $\hat{p}_y$  from Equation (4).

PARAMETER ESTIMATES AND GOODNESS OF FIT

Numerical maximum likelihood estimation (MLE) techniques were used to estimate  $b_x$ ,  $b_y$ , and  $b_z$  for each snow-pack category. The daily  $\log_e$  (ln) likelihood,

$$\begin{aligned} \ln L_{ij} &= \ln[P(X = x_{ij}, Y = y_{ij}, Z = z_{ij})] = \\ &= \ln \left\{ \left[ \binom{n_x}{x_{ij}} (b_x p_{ij})^{x_{ij}} (1 - b_x p_{ij})^{n_x - x_{ij}} \right] \right. \\ &\quad \left[ \binom{n_y}{y_{ij}} (b_y p_{ij})^{y_{ij}} (1 - b_y p_{ij})^{n_y - y_{ij}} \right] \times \\ &\quad \left. \left[ \binom{n_z}{z_{ij}} (b_z p_{ij})^{z_{ij}} (1 - b_z p_{ij})^{n_z - z_{ij}} \right] \right\}, \end{aligned}$$

was summed within each snow-pack category and across years:

$$\ln L = \sum_{i=1}^g \sum_{j=1}^{N_i} \ln L_{ij}, \quad (7)$$

where  $L_{ij}$ ,  $x_{ij}$ ,  $y_{ij}$ , and  $p_{ij}$  represent the likelihood function, the number of observed avalanches by frequency classes, and regional probability of avalanches for the  $j$ th day in the  $i$ th year for a particular season;  $N_i$  is the number of days with avalanches in the  $i$ th year.  $\ln L$  represents a composite likelihood for a given snow-pack category for the period of record. A quasi-Newton minimization method using subroutine ZXMIN (IMSL, [1982]) estimated  $b_x$  and  $b_y$ , given  $b_z=1$ , such that  $\ln L$  was maximized, considering  $p$  as estimated by  $\hat{p}$  conditioned on  $\hat{b}_x$ ,  $\hat{b}_y$ , and  $\hat{b}_z$ . Allowing  $b_z$  to vary produced no improvement in the fit, thus supporting the decision to hold  $b_z$  constant at one.

A simplified procedure for obtaining  $b_x$  and  $b_y$  given  $b_z = 1$  was tried for comparative purposes. The parameter  $b_x$  was obtained by normalizing the mean number of avalanche days for the high-frequency paths to the mean frequency (avalanche days) of the low-frequency class in the respective snow-pack categories.  $b_y$  was similarly obtained. This set of parameters is referred to as mean estimated parameters in later discussion. Estimated parameters for all snow-pack categories are listed in Table I.

Goodness of fit of the model to the data was assessed by forming daily  $\chi^2$  statistics for goodness of fit and accumulating across days within a snow-pack category and years. In notation similar to Equation (6), if  $x_{ij}$ ,  $y_{ij}$ , and  $z_{ij}$  are estimated as  $\hat{x}_{ij} = b_x \hat{p}_{ij} n_x$ ,  $\hat{y}_{ij} = b_y \hat{p}_{ij} n_y$ , and  $\hat{z}_{ij} = b_z \hat{p}_{ij} n_z$ , respectively, where  $\hat{p}_{ij}$  is the estimated stability parameter  $p$  for the  $i$ th- $j$ th day, then

$$\chi_{ij}^2 = \frac{(\hat{x}_{ij} - x_{ij})^2}{\hat{x}_{ij}} + \frac{(\hat{y}_{ij} - y_{ij})^2}{\hat{y}_{ij}} + \frac{(\hat{z}_{ij} - z_{ij})^2}{\hat{z}_{ij}}. \quad (8)$$

Goodness of fit with MLE and mean estimated parameters was evaluated both on a daily ( $\chi_{ij}^2$ ) and on a pooled basis

TABLE 1. PARAMETER ESTIMATES AND GOODNESS OF FIT FOR THE 1974-82 FRONT RANGE AVALANCHE DATA

| Paths/group <sup>1</sup>                | Estimated parameters | Number of days with avalanches | $\chi^2$            | Significance level  |       |
|---|----------------------|--------------------------------|---------------------|---------------------|-------|
| <i>November-December</i>                |                      |                                |                     |                     |       |
| $n_x = 13$<br>$n_y = 30$<br>$n_z = 90$  | MLE                  | $b_x = 5.89$                   | 126                 | 315.25              | 0.003 |
|   |                      | $b_y = 2.36$                   | 8                   | 69.05 <sup>2</sup>  |       |
|   |                      | $b_x = 1.00$                   | 118                 | 246.20 <sup>3</sup> | 0.264 |
|   |                      |                                |                     |                     |       |
|   | mean                 | $b_x = 7.60$                   |                     | 319.94              | 0.002 |
|   |                      | $b_y = 3.00$                   | 8                   | 73.35 <sup>2</sup>  |       |
| $b_z = 1.00$                            |                      | 118                            | 246.59 <sup>3</sup> | 0.258               |       |
| <i>January-February</i>                 |                      |                                |                     |                     |       |
| $n_x = 16$<br>$n_y = 36$<br>$n_z = 104$ | MLE                  | $b_x = 6.81$                   | 170                 | 449.05              | 0.001 |
|   |                      | $b_y = 2.92$                   | 16                  | 119.75 <sup>2</sup> |       |
|   |                      | $b_z = 1.00$                   | 154                 | 329.30 <sup>3</sup> |       |
|   | mean                 | $b_x = 7.80$                   |                     | 448.47              | 0.001 |
|   |                      | $b_y = 3.20$                   | 15                  | 122.91 <sup>2</sup> |       |
|   |                      | $b_z = 1.00$                   | 155                 | 325.56 <sup>3</sup> |       |
| <i>March-April</i>                      |                      |                                |                     |                     |       |
| $n_x = 14$<br>$n_y = 31$<br>$n_z = 92$  | MLE                  | $b_x = 4.42$                   | 178                 | 543.46              | 0.001 |
|   |                      | $b_y = 2.28$                   | 18                  | 180.50 <sup>2</sup> |       |
|   |                      | $b_z = 1.00$                   | 160                 | 362.95 <sup>3</sup> |       |
|   | mean                 | $b_x = 11.50$                  |                     | 540.62              | 0.001 |
|   |                      | $b_y = 3.80$                   | 13                  | 204.67 <sup>2</sup> |       |
|   |                      | $b_z = 1.00$                   | 165                 | 335.95 <sup>3</sup> |       |

<sup>1</sup>  $n_x$ ,  $n_y$ , and  $n_z$  refer to the number of avalanche paths in the high-, moderate-, and low-frequency groups, respectively.

<sup>2</sup> Pooled  $\chi^2$  for individual days having significant lack of fit ( $\alpha = 0.05$ ).

<sup>3</sup> Residual pooled  $\chi^2$ .

$$\left[ \sum_{i=1}^g \sum_{j=1}^{N_i} \chi_{ij}^2 \right]. \text{ Individual } \chi_{ij}^2 \text{ were evaluated as having}$$

two degrees of freedom and the pooled

$$\sum_{i=1}^g \sum_{j=1}^{N_i} \chi_{ij}^2 \text{ was given } \left[ 2 \sum_{i=1}^g N_i - 3 \right] \text{ degrees of freedom.}$$

Significance was assessed at  $\alpha = 0.05$ .

Considering all days in each bimonthly period, there is significant disagreement between model and data (Table I.) However, by partitioning the pooled  $\chi^2$  into two components, one for days exhibiting significant lack of fit on an individual basis and another for the remaining days, it appears that the serious disagreement is concentrated in only 10% of the days in each snow-pack category. For the remaining 90% of the days on which avalanches occurred, the model is an adequate representation.

For the days exhibiting significant lack of fit on an individual basis, two patterns were apparent: one was characterized by three or four avalanches from high-frequency paths with no avalanches in the other classes; the other involved days with many avalanches

on low-frequency paths with successively fewer than expected avalanches on paths with higher frequency. The former pattern occurred on 40% of the poor-fit days (4% of the total avalanche days), contained small soft-slab surface avalanches, and was strongly concentrated in the March and April data. Three-day precipitation totals preceding such avalanches varied from 7 to 40 mm, and all but one of the 24 h amounts were less than 25 mm. The occurrence of small- to moderate-sized storms on a generally stable spring snow-pack often produces scattered surface instability on the more active paths. Such episodes are not viewed with as much concern because they involve comparatively few avalanches on any particular day, and the avalanches are of small magnitude. The second pattern appeared on 30% of the poor-fit days (3% of the total avalanche days). Accompanying avalanches were generally bigger than with the first pattern, and these avalanches were concentrated in January and February. Three-day precipitation totals prior to these avalanches varied from 20 to 63 mm and daily totals often exceeded 25 mm. The deep new-snow layers accompanying significant storms may violate the structure of the model, but there were other large storms that produced avalanches from all frequency classes where the goodness of fit was acceptable. Missed observations associated with poor visibility produced by certain synoptic weather patterns

and the effects of prior avalanching on high-frequency paths both decrease the number of observed avalanches on active paths, but have little effect on the number of avalanches from the more numerous paths (Table I) in the moderate- and low-frequency classes. A better explanation and refinement of the model to account for such days will require further study. No clear pattern appeared on the remaining 3% of the avalanche days.

Goodness of fit associated with the mean estimated parameters was similar to that observed for the MLE parameters. This indicates that the mean estimated parameters, which are more readily obtainable, could be employed without appreciable loss of sensitivity. They could also be used at locations with a less extensive avalanche record than was used in this study.

SEQUENTIAL PROBABILITY RATIO TESTING

Statistical sequential analysis is any statistical procedure in which the number of observations needed to satisfy the sampling objective depends, as the sampling proceeds, on the results of prior observations. In particular, sequential probability ratio testing is a methodology useful for distinguishing between two or more statistical hypotheses. In brief, observations are taken in sequence until all but one hypothesis are rejected by the data. Wald (1947) summarized the methodology and a more recent treatment was given by Wetherill (1975).

Our use of the methodology is to provide a consistent mechanism for attaching physical meaning to the regional snow-pack stability parameter  $p$ . Using avalanche hazard categories similar to those employed by operational avalanche warning centers (Williams, 1980), we wish to distinguish among the following hypothesized conditions:

- 1) high stability  $H_1: p = p_1,$
- 2) moderate stability  $H_2: p = p_2,$
- 3) low stability  $H_3: p = p_3,$

where  $p_1 < p_2 < p_3$ . Competing sequential methods to evaluate this set of hypotheses are given by Sobel and Wald (1949) and Armitage (1950). Although Wetherill (1975) prefers the approach of Sobel and Wald, in this study Armitage's technique was used because it offers greater flexibility in specifying error probabilities of making the wrong decision. Evaluation of our data by both methods produced only minor differences in results. Rarely were different decisions indicated.

Details of the methodology can be found in Armitage (1950), but a brief description of the technique is given here. Let  $L_1, L_2,$  and  $L_3$  (here the  $\log_e$  transformation is not used) denote the value of Equation (6) evaluated for the data in hand on any given day, letting  $p = p_1, p_2,$  and  $p_3,$  respectively. The hypotheses  $H_1, H_2,$  and  $H_3$  are evaluated based on the ratios  $L_1/L_2, L_1/L_3,$  and  $L_2/L_3$ . If  $L_1/L_2$  is greater than one, there is reason to think  $H_1$  is more likely than  $H_2$  to be true. Conversely, if  $L_1/L_2$  is less than one,  $H_2$  is more likely. To reach a decision requires consistency among all three ratio values. For instance, to accept  $H_1$  requires both  $L_1/L_2$  and  $L_1/L_3$  to be sufficiently larger than one to satisfy test criteria determined by pre-established error probabilities. These error probabilities, denoted by  $w_{ij}$ , are the probability of accepting  $H_i$  when in fact  $H_j$  is true. They partially protect the user from making a wrong decision. The degree of protection is varied according to the consequences of making a particular error. If the data in hand produce ratio values that are inconsistent with each of the three hypotheses (or are insufficiently different from one), then the testing procedure cannot reach a decision, and more observations must be obtained. Sampling continues, with intermediate pauses to carry out a hypothesis testing exercise, until all hypotheses but one are rejected.

Probabilities of a natural slab avalanche in each of

the three hazard categories had to be specified in terms of what might be reasonably expected from a large sample of uncontrolled paths at undeveloped sites in the mountains. After some reflection on this problem, it was decided that the chances of release in the low through high avalanche hazard categories could be approximated by setting  $p_1, p_2,$  and  $p_3$  equal to one in a thousand, one in one hundred, and one in twenty, respectively. The choice of these values was mainly experiential and was made without direct reference to warning-center estimates.  $w_{ij}$  error probabilities were similarly selected, with the rationale that, where public risk is involved, slight overwarning is better than underwarning. To declare the snow-pack stable when it is unstable ( $w_{13}$ ) is a serious error that involves greater risk than the reverse situation of  $w_{31}$ . It was also felt that to err between low and moderate hazard ( $w_{12}$  or  $w_{21}$ ) involved less risk than an error between moderate and high hazard ( $w_{23}$  or  $w_{32}$ ). Guided by these general concepts, we chose the following  $w_{ij}$  error probabilities:

$$w_{12} = 0.10, \quad w_{21} = 0.20,$$

$$w_{13} = 0.0001, \quad w_{31} = 0.01,$$

$$w_{23} = 0.001, \quad w_{32} = 0.10.$$

Since optimum values for these probabilities are unknown, further refinement or adjustments are not planned until more is known about the relationship between avalanches and regional snow-pack stability.

Based on applying this methodology to the 1974-82 data and a limited simulation of sequential sampling from a hypothetical population containing 1000 avalanche paths with the proportion of paths in each frequency category similar to the observed data, some insight can be obtained into how the methodology functions in practice. First, it must be understood that the 1974-82 data were not recorded with the objective of applying a sequential testing procedure. Therefore, estimation of snow-pack stability and evaluation of the hazard hypotheses was based on the whole sample. It was not possible to add more samples when test results were inconclusive; conversely, if a decision was reached, it was not possible to estimate whether the decision could have been based on fewer observations, although entry of data into the sample could have been simulated. The testing procedure is more sensitive to observations in the high- and, then, moderate-frequency categories. If failures are observed in the higher frequency categories, smaller samples are required to reach a decision because proportionally larger increments are added to the likelihood functions and a higher hazard estimate occurs. Intuitively, we would rather assign more weight to observations in the low-frequency class because this group represents the bulk of the population and avalanches here may represent greater instability than those from more frequent paths. However, likelihood ratio testing of the frequency distribution given by Equation (1) dictates otherwise, and the result is a testing procedure which, if it errs, will generally do so by declaring a hazard higher than true, not lower.

As noted by Wetherill (1975) and Corneliusson and Ladd (1970), the sample size needed to reach a decision using a sequential testing procedure is highly variable and poorly estimated by Wald's (1947) theory. However, sequential sample sizes will generally be smaller than those needed for testing procedures with fixed sample size because early detection of avalanches in a sequential sampling exercise can produce a quick decision. Hypothesis testing of our sample of 185 paths resulted in a number of inconclusive decisions. This is consistent with our simulation results which indicated average sample numbers (ASN) of about 150 to make a decision when actual  $p = 0$ . As  $p$  was increased toward  $p = 0.01$ , ASN increased to about 225 with subsequent gradual decrease back toward 150 and below as  $p$  was further increased toward  $p = 0.05$ . Reaching a decision in the neighborhood of  $p = 0.01$  requires more extensive

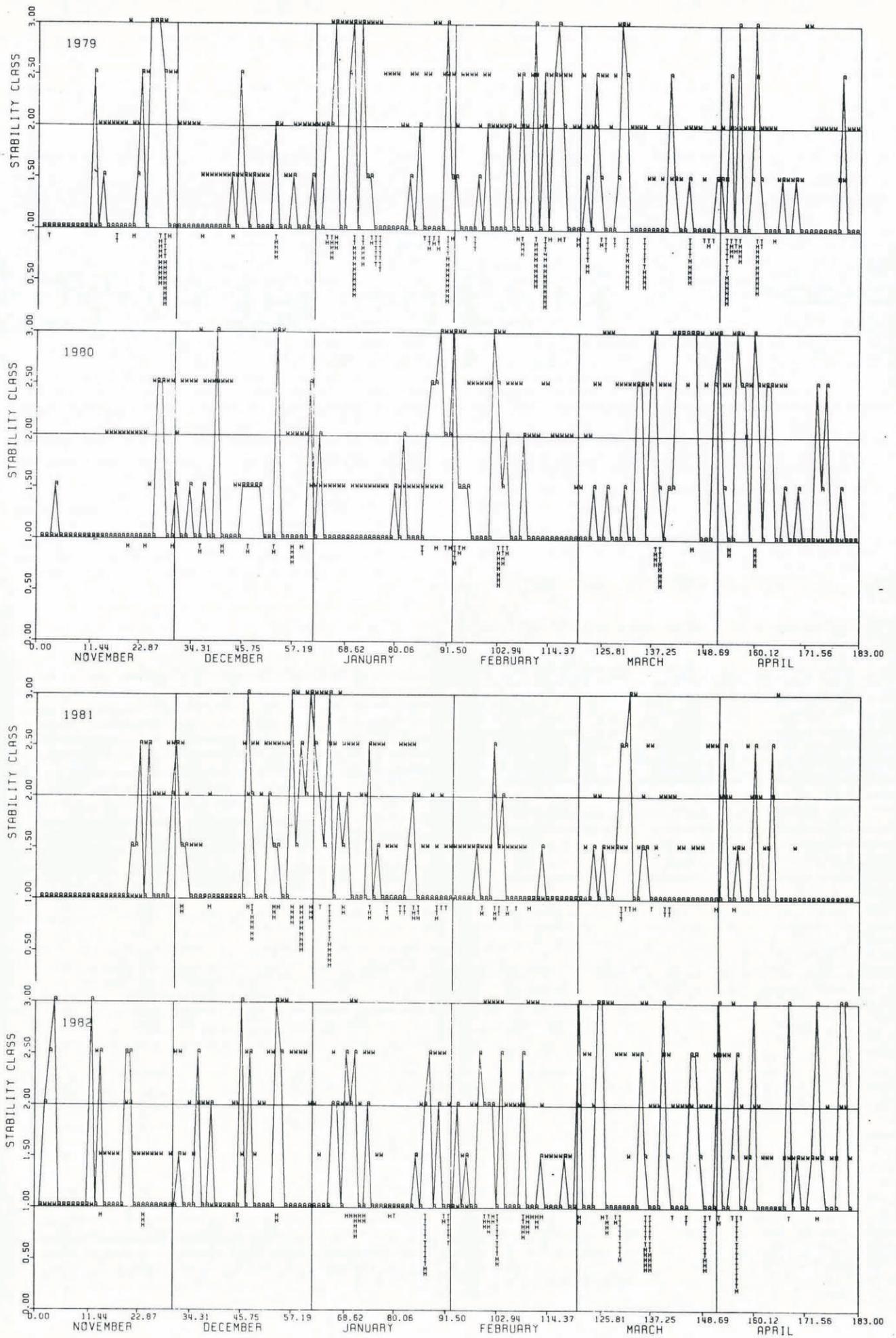


Fig. 1. Comparison of sequential test results with Warning Center hazard estimates and avalanche control information for winters 1979–82. A indicates the daily sequential test results, which are connected by a line. W refers to daily Warning Center stability estimates. T indicates a control attempt (try) without an avalanche, and H (hit) refers to a control attempt that caused an avalanche.

sampling because all three hypotheses are realistically possible, whereas for  $p$  near 0 or 0.05, it is really only necessary to distinguish between two hypotheses. Conceptually, ASN would continue to decrease with increasing  $p$ , but it is difficult to provide specific estimates because the possible mix of failures among the three frequency categories becomes increasingly variable for larger  $p$ . The above figures are only averages, and the variability was large, ranging in our limited simulation from 5 to about 600.

If the sample sizes needed for this testing procedure seem unreasonably large, it must be realized that even when snow-pack stability is low, avalanches are comparatively rare. In a population of 1000 paths in which the majority of paths fall in the low-frequency class, the difference between the hypothesized high and moderate snow-pack stability conditions is the appearance of avalanches on only an additional nine paths (1 versus 10 in a 1000), two of which would be expected to appear in a random sample of 200. The difference between moderate and low stability is more distinguishable (10 versus 50 avalanches in 1000 paths), and generally results in smaller expected sample sizes as  $p$  increases beyond 0.03. Mathematically, a decision could be made with only two paths if both were in the high-frequency group and both had failed. Realistic application of the method to a large area should require a minimum sample of about 100 paths, distributed among the frequency classes, in order to represent the population properly.

#### COMPARISON WITH OPERATIONAL FORECASTS

Likelihood ratio test results for the avalanche seasons of 1979-82 are plotted with daily estimates of avalanche hazard from the Colorado Avalanche Warning Center in Figure 1. For clarity, the likelihood ratio test results are referred to as "the stability index" in the remainder of this paper. The two systems are not directly comparable because the Warning Center estimates potential hazard from natural and triggered avalanches, and the plotted Warning Center estimates apply to some areas in the northern Colorado mountains that are not represented by our sample. Nonetheless, the daily Warning Center snow-pack stability evaluations provide a valuable independent reference for the model. Center forecasters appraise the potential for release on a four-point scale that corresponds to low, moderate, high, and extreme hazard, but the extreme category occurs so infrequently that only the first three hazard levels are displayed. The 1.50 and 2.50 scale values are used when the stability index was inconclusive, and for Warning Center hazard estimates between low and moderate or moderate and high. A hazard rating of three is the threshold value for issuance of public warnings in the U.S.

Natural avalanches fell on 30% of the winter days during the four test winters. Following the first avalanches of early winter, the stability-index produced hazard estimates somewhat lower than those from the Warning Center. A reversal in this trend appeared in spring when instability was confined to the surface layers of the snow-pack, a case in which avalanche activity from high-frequency paths dominated the procedure. Snow-pack stability indicated by natural avalanches is a rapidly fluctuating parameter with a marked lack of persistence. The stability index indicates that rapid stabilization from high to low hazard often occurs in one or two days, whereas slower stabilization rates and strong persistence are estimated by conventional methods. There were only three cases in the four winters when the stability index indicated high hazard persisting for two or more consecutive days compared to 21 such cases for Warning Center estimates. Similarly, the maximum number of consecutive days with a stability of 1.5 or more was 10, as indicated by the stability index compared to 85 by Warning Center personnel. These differences in trend, response, and persistence in snow-pack stability result from a

combination of real differences and a necessary safety margin used by Warning Center personnel.

#### SNOW-PACK STABILITY AND ARTIFICIAL RELEASE

The combined results of daily avalanche control by the highway department, a mining operation, and a ski area are plotted in the lower part of the seasonal displays in Figure 1. Each T indicates a "try" without release, and an H represents a hit, which is a "try" with release. The control attempts were limited to cold, dry snow conditions, which are common to the winter months, and are also prevalent on many days in spring. There is good overall agreement between snow-pack stability estimates and control results, but poor agreement obtained on several individual cases such as the one on 5 April, 1983, nullifies the relationship on a quantitative basis. The correlation coefficient between the stability index and shooting success, defined as a ratio of the number of hits to the number of "tries," was  $r = 0.58$ . Average shooting success, based on days with 10 or more attempts, was 35% on days with no natural avalanches versus 51% on days with snow-pack stability 1.5 or greater. This increased to 76% when the stability index indicated high hazard. There were fewer "tries" on stable days, which reflects the lower hazard perceived by control personnel. Some problems associated with using control results as an index of regional snow-pack stability include: the very small sample size of controlled avalanches, the tendency for control efforts to be concentrated in a few isolated pockets within the sample area, the fact that information from heavily controlled sites is mostly limited to the stability of surface layers, and the tendency for control teams to vary the number of shots in accordance with the reaction of the snow to the first few control attempts. Moreover, control efforts are not equally distributed across all frequency classes, but are concentrated in the high-frequency group. And since snow removal and traffic considerations generally take precedence over avalanche control along most highways, snow-pack stability information based on such control results is frequently unavailable when most needed. Further research is needed to determine the degree of residual instability in the snow-pack after the last natural avalanche occurs.

#### SUMMARY AND DISCUSSION

A framework has been proposed for summarizing and interpreting the information available from avalanche occurrence on individual avalanche paths within generalized snow-condition categories contained in the three bimonthly periods of early, middle, and late winter. Avalanches on paths having heterogeneous probabilities of release were pooled into a composite regional stability index denoted as parameter  $p$  of the statistical distribution, chosen to approximate the probability of failure in the population. The stability index offers more resolution than the avalanche day, and estimation of  $p$  is possible even when poor visibility or other factors preclude observation of an entire sample. An analysis of goodness of fit suggests that there is adequate agreement between observations and fitted distribution about 90% of the time across the three snow-condition categories. The disagreement is somewhat patterned, suggesting that improvement might result if a more sophisticated algorithm incorporating weather, snow-pack stratigraphy, antecedent failure history, and other factors was developed to partition the population into the X, Y, and Z categories on a continuing basis rather than assuming constancy within two-month periods. We suspect that the distinctions among X, Y, and Z groups become quite muddled under certain weather patterns and hope to explore this avenue of investigation. A testing procedure incorporating more than three frequency classes might offer further improvement.

A stability index (sequential testing procedure) was also proposed as a mechanism to interpret consistently

the magnitude of  $p$  as it applies to avalanche hazard estimation. Application of the testing methodology to our data produced reasonable results in general agreement with independently generated hazard estimates from an operational avalanche warning program and information from control efforts. Successful release of snow during control attempts appeared to be strongly related to the degree of hazard indicated by the stability index. Average sample sizes needed to reach a sequential testing decision appeared large, ranging from 150 to 225 paths, although a decision could be quickly reached after observation of only a few avalanches with low snow-pack stability (high avalanche hazard). These tests would provide real-time evaluation potential if a series of dispersed observers were available to convey the avalanche information rapidly to a central location.

We feel our initial efforts suggest a promising new approach to systematically relating snow-pack stability with avalanche occurrence, although further refinement of the present methodology is indicated. As already noted, a more sophisticated approach for assigning a particular path to the high-, medium-, or low-frequency classes based on snow texture and stress would improve the reliability of the methodology. Presently, there is some difficulty in specifying the potential population, especially in the low-frequency class. Previously unrecorded events which occur on new paths every year increase the size of the Z group and complicate the definition of sample size. An underestimate of  $n_z$  would cause overestimation of  $p$  and reduced sensitivity of the sequential testing procedure to information from the z group. Although the proposed methodology requires only a simple "yes-no" avalanche observation for each path, without any detailed attempt at classifying each avalanche event, a subjective value-judgment is needed for the size of avalanche that constitutes a hazard; the values of  $p_1$ ,  $p_2$ , and  $p_3$ ; and the  $w_{ij}$  error probabilities; these would depend on the purpose of the investigation. After these judgments have been made for a particular application, consistent non-subjective evaluation of observed failure frequency is possible.

#### ACKNOWLEDGEMENT

The authors extend sincere appreciation to Dr Charles Obled of the Institute de Mécanique, Grenoble, France, whose stimulating discussion and useful ideas helped guide our initial attack on this intricate subject. The authors also thank the referee whose helpful suggestions improved the clarity of the presentation.

#### REFERENCES

- Armitage, P. 1950. Sequential analysis with more than two alternative hypotheses, and its relation to discriminant function analysis. *Journal of the Royal Statistical Society*, Vol. 12, No. 1, p. 1327-44.
- Bois, P., and Obled, C. 1973. Vers un système opérationnel de prévision numérique des avalanches à partir de méthodes statistiques. *Hydrological Sciences Bulletin*, Vol. 18, No. 4, p. 419-29.
- Corneliusson, A.H., and Ladd, D.W. 1970. On sequential tests of the binomial distribution. *Technometrics*, Vol. 12, No. 3, p. 635-46.
- IMSL. [1982.] [International] M[athematical] and S[tatistical] L[ibrary]. Reference manual. Edition 9. Houston, Texas, IMSL.
- Judson, A. 1977. Climatological data from the Berthoud Pass area of Colorado. *U.S. Dept. of Agriculture. Forest Service. Research Paper RM-42.* [Rocky Mountain Forest and Range Experiment Station, Fort Collins, Colorado.]
- Judson, A. 1983. On the potential use of index paths for avalanche assessment. *Journal of Glaciology*, Vol. 29, No. 101, p. 178-84.
- Judson, A., and Erickson, B.J. 1973. Predicting avalanche intensity from weather data: a statistical analysis. *U.S. Dept. of Agriculture. Forest Service. Research Paper RM-112.* [Rocky Mountain Forest and Range Experiment Station, Fort Collins, Colorado.]
- Judson, A., and King, R.M. 1983. Spatial variations in snow stability inferred from artillery control. *Journal of Glaciology*, Vol. 29, No. 103, p. 508-11.
- Judson, A., and others. A process-oriented model for simulating avalanche danger, by A. Judson, C.F. Leaf, and G.E. Brink. *Journal of Glaciology*, Vol. 26, No. 94, p. 53-63.
- Perla, R.I. 1970. On contributory factors in avalanche hazard evaluation. *Canadian Geotechnical Journal*, Vol. 7, No. 4, p. 414-19.
- Salway, A.A. 1979. Time-series modelling of avalanche activity from meteorological data. *Journal of Glaciology*, Vol. 22, No. 88, p. 513-28.
- Sobel, M., and Wald, A. 1949. A sequential decision procedure for choosing one of three hypotheses concerning the unknown mean of a normal distribution. *Annals of Mathematical Statistics*, Vol. 20, No. 4, p. 502-22.
- Wald, A. 1947. *Sequential analysis*. New York, John Wiley and Sons, Inc.; London, Chapman and Hall Ltd.
- Wetherill, G.B. 1975. *Sequential methods in statistics. Second edition*. New York, John Wiley and Sons, Inc.; London, Chapman and Hall Ltd.
- Williams, K. 1980. Credibility of avalanche warnings. *Journal of Glaciology*, Vol. 26, No. 94, p. 93-96.

MS. received 13 March 1984 and in revised form 14 May 1984