EVALUATION AND COMPARISON OF STATISTICAL AND CONVENTIONAL METHODS OF FORECASTING AVALANCHE HAZARD

By P. FÖHN, W. GOOD,

(Eidg. Institut für Schnee- und Lawinenforschung, 7260 Weissfluhjoch/Davos, Switzerland)

P. Bois and C. Obled

(Institut National Polytechnique de Grenoble, Laboratoires de Mécanique des Fluides, B.P. 53, 38041 Grenoble-Cedex, France)

ABSTRACT. Principal problems concerning the raw data and methodological limitations of statistical and conventional avalanche forecasting methods are summarized. The concepts of four statistical models based on multivariate data analysis, are outlined in a few words. In order to give an idea of the potential and quality of the different methods, test runs over two winters are discussed and a tentative score is established. Statistical models I and IV, together with the conventional forecast, attain a score of 70–80%, whereas statistical models II and III show a slightly poorer performance.

RÉSUMÉ. Evaluation et comparaison des méthodes statistiques et conventionelles pour la prévision des avalanches Les problèmes principaux quant aux données bruttes employées ainsi que les limites des différentes méthodes statistiques et de prévision conventionelle des avalanches sont exposés. Les idées de base aux quatre méthodes d'analyse des données sont esquissées brièvement. Pour donner une idée du pouvoir discriminatoire et de la qualité des modèles employés, les résultats de la prévision opérationelle pour deux hivers sont discutés. Les modèles statistiques I et IV ainsi que la méthode de prévision conventionelle réussissent dans 70 à 80 cas sur 100. Les modèles statistiques II et III cependant se sont avérés quelque peu inférieurs.

Zusammenfassung. Auswertung und Vergleich statistischer und konventioneller Methoden im Lawinenwarndienst. Die Probleme der Gewinnung von Ausgangsdaten und die verfahrenstechnischen Grenzen der statistischen und konventionellen Methoden im Lawinenwarndienst werden zusammenfassend dargestellt. Die Konzeption der vier statistischen Modelle — beruhend auf der Analyse vielparametriger Daten — wird in kurzen Worten dargetan. Zur Beurteilung der Leistungsfähigkeit und Qualität der verschiedenen Methoden werden Testreihen über zwei Winter analysiert und versuchsweise bewertet. Die statistischen Modelle I und IV ebenso wie die konventionelle Vorhersage erreichen den Zuverlässigkeitsgrad 70–80%, während die statistischen Modelle II und III etwas schlechter abschneiden.

I. INTRODUCTION

The problem of avalanche forecasting consists either in predicting the time and the size of a few single avalanche events (e.g. along a highway) or in foreseeing for the next 24 to 72 h the probable spectrum of the avalanche activity (number, size, type, aspect, altitude zone, slope inclination) in a whole mountain region. Whereas the first set of problems is rather suited for deterministic reasoning, the second contains many more random elements (various local snow conditions, local triggering) which directly call for a stochastic-statistical treatment.

Avalanche forecasting on a regional scale, which alone is treated here, may be compared with hail or thunderstorm prediction in meteorology. The tendency or probability of those local phenomena is still evaluated by empirical relations or multiple-regression techniques, although in meteorology physico-mathematical models are used extensively.

Different groups, namely Shcherbakov (1966), Bois and Obled (1973), Judson and Erickson (1973), Miller and Miller (1973), Bovis (1974), have reported their statistical models in recent years and it appeared that despite the different data sets and approaches, almost everywhere the same basic problems have been encountered.

With the aid of real-time test runs of four statistical avalanche forecast models, partly over a period of three years, and through a comparison with the conventional method, we will discuss these basic problems. The statistical model IV, the most recent one, is treated separately and rather in terms of methodology than results, because it has not yet been possible to use this model for day-to-day forecasts.

2. Avalanche observations and variable selection

Any avalanche prediction method, whether of conventional, statistical, or even physical nature, needs an appropriate avalanche variable in order to establish and to verify the method. As opposed to forecasting in meteorology or hydrology, where almost everyone can check if a given forecast has been verified by Nature, the observation of the phenomenon here poses a major problem:

(a) Avalanche activity is most intensive during snowstorm conditions, i.e. at a time when the majority of avalanches are triggered, we cannot observe them.

(b) Verification of avalanche activity in a whole area calls for a check of many slopes of different aspect, inclination, etc., a tremendous labour for any observation crew.

The avalanche observations, used in this study have been collected on Weissfluhjoch/Davos, Switzerland over a period of roughly twenty years and are affected by systematic errors, for the above mentioned reasons. Figure 1 confirms this. It displays the avalanche activity during the winter 1975/76 in three adjacent observation areas. The magnitude of avalanche activity is tentatively calculated by the area covered by avalanches per day; distinction is made between dry and wet snow avalanches and between those naturally and artificially released.

Avalanche activity seems to be most pronounced in the Parsenn area, which is most intensively observed, however the other areas sometimes yield avalanches when none were observed in the Parsenn area, whose records are used exclusively in this study.

In order to eliminate some of the observational errors, a discrete-threshold avalanche variable has been chosen in the present study, namely the "avalanche-day", a day on which at least one natural avalanche observed (cf. Bois and others, [1975]). In this way the magnitude of daily avalanche activity is irrelevant and avalanche data errors are restricted to those of timing and misinterpretation of avalanche types.

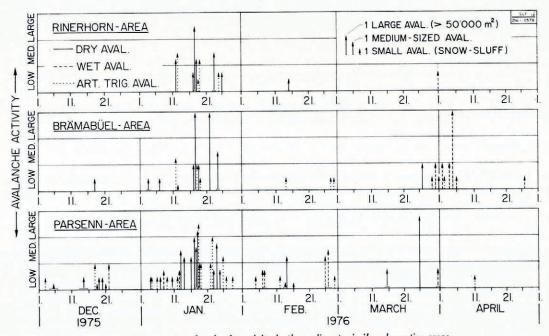


Fig. 1. Comparison of avalanche activity in three adjacent, similar observation areas.

The selection of the meteorological and snow-pack variables which were to be correlated with the avalanche activity, was mainly done on the basis of availability of long records. Table I displays the set of input variables and shows that most of the variables are of a meteorological nature, whereas the snow-pack is represented in a rather rudimentary manner. The reason for this is the well-known lack of quantitative interpretation of snow-cover and ramhardness profiles.

The input-variables have been measured at the index station Weissfluhjoch (2 540 m a.s.l.), situated at the level of the potential fracture zone of the majority of avalanches in the Parsenn area.

TABLE I. INPUT-VARIABLES FOR THE AVALANCHE FORECASTING MODELS I, II, IV

$\mathcal{N}o$.	Variables	Model I	Model II	Model IV
1	Total precipitation (last 24 h)		*	*
2	New fallen snow (last 24 h)	*	*	*
3	Water equivalent of new snow (last 24 h)		*	*
3 4 5 6 7 8	Maximum 3 h precipitation intensity (last 24 h)		*	*
5	Maximum wind-speed for day j	*	*	*
6	Global radiation for day i		*	*
7	Sunshine hours of day j	*	*	*
8	Cloudiness for day j	*	*	*
9	Maximum air temperature for day j	*	*	*
10	Minimum air temperature for day j	*	*	*
II	Total snow depth for day j	*	*	*
12	Penetration depth of cone penetrometer		*	*
13	Temperature of snow cover (10 cm below surface)		*	*
14	Number of avalanches on previous day		*	
14	Snow-drift conditions for day j			*

3. Forecasting methods

3.1. Conventional method

This method is based on empirically derived regularities in the relationship between weather, snow conditions, and avalanche activity. As opposed to the statistical approach, where index variables replace the processes involved, a combination of index-oriented thinking together with knowledge of the linked physical processes and personal experience enter into the forecasting procedure.

Advantages of this method are:

- (i) The synthesis of medium-scale meteorological analysis, local weather (snowfall, wind, etc.) and snow-pack examination covers the whole range of factors causing avalanches.
- (ii) The method is very flexible, even semi-quantitative information (e.g. failure or success of artificial triggering) can be integrated in the forecasting process.
- (iii) The output of the synthesis is given in verbal form, well suited for fast diffusion on radio and television.

Disadvantages are:

- (i) The reasoning about causes and effects is influenced subjectively.
- (ii) The memory of any forecaster is short compared with computer-oriented methods and often influenced by exceptional events.

3.2. Statistical model I

This model has been devised with the aid of seven commonly available input variables and the avalanche observations (cf. Table I) covering the time period of 1953/54-1969/70. The

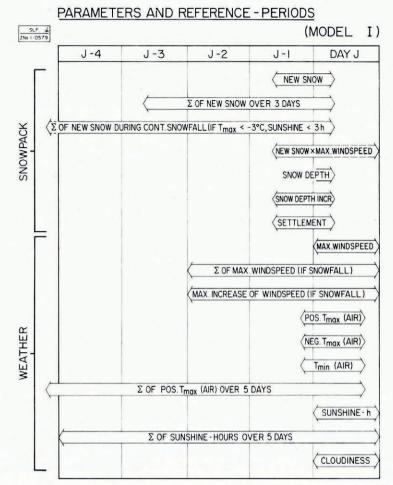


Fig. 2. List and time-reference of elaborated variables for model I. Forecast day is day j.

data were grouped month by month in order to keep the data sets more homogeneous, thus eliminating the seasonal trend. The two-step procedure consisted of:

- (1) the examination whether avalanche and non-avalanche days form one or several distinct groups,
- (2) finding discrimination criteria for the different groups.

The input variables (X_i) have been used to generate a set of elaborated variables (Y_p) which represents the various processes and their different time scales. These elaborated variables or parameters are displayed in Figure 2.

In order to render them independent, new variables formed as linear combinations of the old ones are introduced:

$$X_i \to \Upsilon_p \to \mathcal{Z}_{jp} = \sum_{p=1}^{17} a_{jp} \cdot y_{jp},$$

where y_{jp} denotes a value of the variable p on day j and a_{jp} are the coefficients. The selection criterion of the new axes may be stated as:

$$S_{p^2} = \frac{1}{n} \sum_{j=1}^{n} (z_{jp} - \bar{z}_p)^2 = \lambda_p = \text{maximum},$$

where λ_p is the eigenvalue of the pth eigenvector. Comparing the principal component loadings in terms of their variance contribution to the total data vector, the number of components can be reduced. In our case, 10 principal components were retained, carrying 92% of the total information.

Representing all data points in the new coordinate system and characterizing them as avalanche or non-avalanche days, we could see that the first group had to be subdivided into the groups: dry-avalanche days and wet-avalanche days. We had then to decide, if, by use of discriminant functions, any day characterized only by its input variables could be classified into one of the three groups. The problem is displayed in Figure 3. According to this figure we have to find two dividing surfaces in the reduced k-dimensional factor or component space (F_k) , which yield high selectivity and minimum allocation errors.

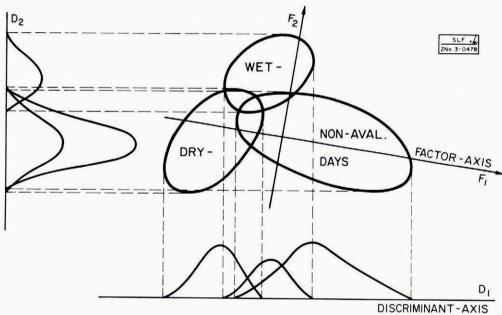


Fig. 3. Representation of dry-, wet-, non-avalanche days in a factor space and scheme of discrimination process.

By using three discriminant functions (dry-, wet-, non-avalanche days) any day may be classified operationally into one of the three groups:

$$f_{jl} = C_l + \sum_{k=1}^{10} C_{kl} \cdot y_{jk}$$
 $(l = 1, 2, 3),$

in which the C are discriminant coefficients and where the largest discriminant function value decides to which group a day j belongs.

The advantages of this model are:

- (i) This model is the only statistical model which can be implemented with commonly available input variables.
- (ii) Since with certain elaborated variables it covers a time range of 3 to 5 d (cf. Fig. 2), a reasonable memory effect is ensured.

The disadvantages are:

- Snow-pack conditions, most important for climax avalanches, are poorly represented.
- (ii) By using the whole input information (even though weighted by the principal components) the discrimination process is damped too much.

3.3. Statistical model II

On the basis of 14 input variables, this model generates by combination roughly 50 elaborated variables. As opposed to model I, the set of best discriminating variables is determined by objective step-wise discrimination programs BMD* and REDIS†.

The program BMD, starting with one variable, adds at each step a new variable and checks if the mean group vectors, e.g. the ones of the group of dry-avalanche days and of the group of non-avalanche days, are significantly different. Significance testing is based on the Wilks criterion:

$$\Lambda_k \equiv \frac{|W_k|}{|T_k|} = \prod_{i=1}^k \lambda_i = \text{minimum}.$$

 W_k and T_k describe the partial "within-group" covariance matrix and the partial total covariance respectively, while λ_k denote the eigenvalues of this quotient.

The backward elimination program REDIS does the same in the opposite direction: starting with all variables it eliminates step by step the least significant ones. Both selection methods finally keep 4 to 7 significant variables for every winter month, which are used for the discrimination process between avalanche and non-avalanche days.

The advantages of this model are:

- (i) The objective selection of significant "avalanche-prone" variables yields valuable information to any forecaster.
- (ii) The snow-pack conditions and the long memory of the avalanche phenomenon are better contained in certain variables.
- (iii) Because the discrimination between the groups is based on a few selected variables, the discrimination process is very distinct.

Disadvantages are:

- (i) Several input variables are not commonly available.
- (ii) Discrimination within the monthly samples—except for March—is made only between two groups (dry-avalanche days, non-avalanche days, etc.), although transition from dry to wet avalanches and vice versa may happen in any month several times.
- (iii) Sharp discrimination between avalanche and non-avalanche groups is valuable as long as every day of the adjustment sample is grouped faultlessly. In view of our avalanche observation problem we have serious doubts about this.

3.4. Statistical model III

The input variables are gained from the global meteorological situation over Western Europe (cf. Duband, 1970). 25 data points of height and temperature at the 700 mbar level and 11 points of surface temperature and pressures yield, by principal component analysis, 14 pressure and 6 thermal components. With the same stepwise procedure as described previously, the most significant ones are selected for the discrimination between

^{*} Program BMD 07M, BMD series, University of California, Berkeley, California, U.S.A. † Program REDIS, Bois, Obled, Laboratoire de Mécanique des Fluides, Grenoble, France.

avalanche and non-avalanche groups. On the supposition that a smaller seasonal trend will be seen in this type of data than in the previous models, December/January/February and March/April have been merged into two data sets.

Advantages of this model are:

- No local input variables other than an avalanche variable have to be fed into the model. Therefore it can be used in remote areas where detailed data sets are not available.
- (ii) The influence of the mesoscale meteorological situation on the avalanche activity can be tested.

Disadvantages are:

- (i) Global meteorological variables are often not available in due time.
- (ii) Process-oriented thinking, the ultimate goal, is rather hampered by such global models even though the forecast results may be reasonable.

4. Comparison of forecast results

In short, a method is as good as its results. During the winters 1973/74-1975/76 day-by-day forecast results have been obtained through the application of some of these models.

Figure 4 illustrates the comparison between the conventional method and statistical models I and II. Avalanche days are marked by large arrows, days with artificially triggered avalanches by dotted thin arrows. These arrows symbolize the approximate avalanche activity and yield the score for the methods. The verbal form of the conventional forecasts has been translated into avalanche "probabilities" for comparison. The general trend of the

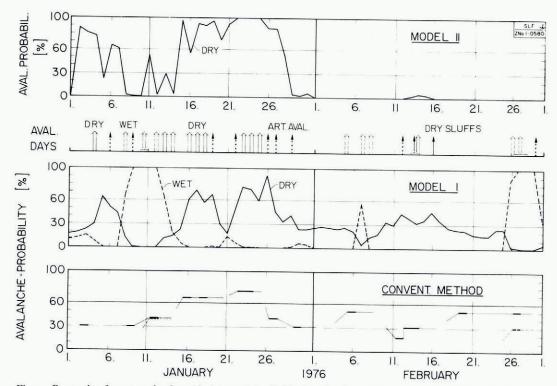


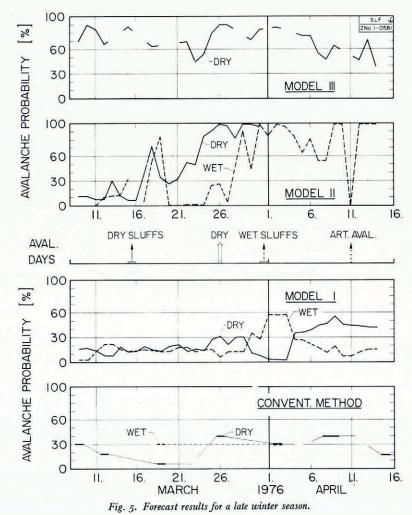
Fig. 4. Day-to-day forecast results for mid-winter period. Verification of a forecast is given by appropriate avalanche days.

avalanche activity in January has been recognized by all methods, however model II misses the wet-avalanche days because wet avalanches are not included in the model. In February all methods have difficulties in accounting for apparently safe periods mixed with short dry-and wet-avalanche cycles. In particular model II cannot recognize these short avalanche periods, because, as several simulation runs have shown, the memory variable (number of avalanches since beginning of winter) is far too large (cf. large number of avalanche days in January) to allow further avalanching in February.

Figure 5 compares four different methods for the late winter season. March is the only month for which model II discriminates between wet- and dry-avalanche days, whereas the

"global" model III yields only the combination of both.

The comparison between observed avalanche days and model results shows that during this rather safe period models II and III overestimate the avalanche probability to a large extent, whereas model I and the conventional method follow the general trend quite nicely, but seem to be too damped, in so far as the observed avalanche days should be rated with a probability between 60 and 100%.



Incorrect timing, i.e. announcing a small avalanche probability on a day when avalanches have been observed and jumping to large avalanche probabilities within the next two days, is observed with all models quite often, and may be related to the avalanche observation problem mentioned above.

5. Avalanche forecast by dynamic cluster analysis (model IV)

5.1. New approach

Up to now we have used the simplified approach that all avalanche days can be subdivided into two groups and that all days without avalanches belong to a homogeneous population (cf. Fig. 3). Due to the fact that specific avalanche types (slab, powder-snow, ground avalanches) occur very often during specific weather/snow conditions the thought was at hand, that an objective classification technique (dynamic cluster analysis) of the various avalanche and non-avalanche days might also be fruitful in avalanche forecasting (Bois and Obled, 1976).

5.2. Concept

Classification techniques for solving this problem are numerous as they are dependent on the initialization of the algorithm. We used one of the more elaborated methods, proposed by Diday (1971).

The main idea is to apply the method of clustering, e.g. to the avalanche days repetitively in two ways:

- 1. Determine the number of sub-populations or clusters required, by an iterative selection of the starting configuration.
- 2. Use the most probable value for the number of clusters k in a second run, with different starting configurations n times; n different partitions will result.

If the underlying structure is "strong", for instance because it is physically founded, then sets of individuals always appear grouped together, no matter which cluster they were initially in, thus yielding "strong patterns". Using this technique on our problem, we started with 120 individuals, i.e. avalanche days, asking for $3 \le k \le 7$ groups, defined by kernels of $7 \le m \le 10$ individuals and allowing for ten different initializations. The "strong patterns" (AK_l) contain between 10 and 17 individuals. These numbers are slightly reduced, especially in March/April, by eliminating the avalanche days of a sequence in order to get truly independent events.

5.3. Input data

Any day of the two winter groups, January/February and March/April, was defined by 14 raw variables (cf. Table I). By use of these, a set of 50 elaborated variables (\mathcal{T}_{jp}) has been compiled, which has been reduced by principal-component analysis to 20 orthogonalized variables (\mathcal{Z}_{jk}).

5.4. Constitution of avalanche groups

We used either linear or quadratic discrimination (SITYP) to classify all the avalanche days. Kernels AK_l were thus enlarged to groups AG_l . The discrimination in the vector space \mathcal{Z} was refined by a step-wise selection of variables Υ_p , providing a better physical understanding and characterizing the following groups:

January | February (cf. Fig. 6):

AG₁: Heavy and sudden snowstorm (up to 60 cm/d) with low temperatures.

AG2: Prolonged snowstorms with relatively warm air and high snow temperatures.

AG₃: Cold weather with strong winds and drift of recently deposited snow.

AG4: Fine weather, rather warm and marked settling of the snow.

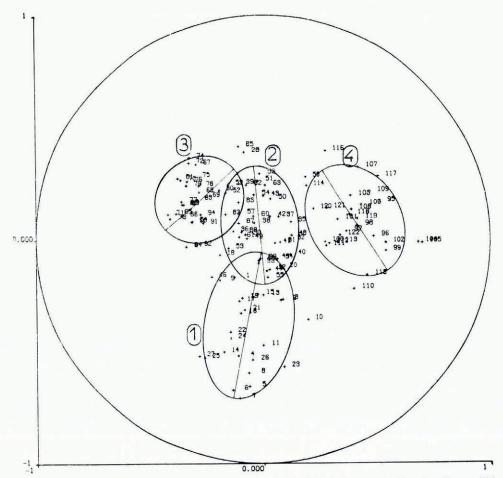


Fig. 6. Representation of typical avalanche days as avalanche groups (AG₁) in the factor space F₁ and F₂ January|February group).

March | April (cf. Fig. 7):

AG₁: Snowstorm with cold temperatures.

AG₂ and AG₃: Changing weather with fresh snow available. Group 2 responds to wind effects, while group 3 is sensitive to temperature changes and thermal effects.

AG₄: Prolonged fine and cold weather; importance of snow transport and temperature changes, includes information of snow-pack stratification.

AG₅: Warm weather with positive temperatures and possibly surface melting.

Intuitively, the consistency of our new classification with experience, I.C.S.I.-group conclusions (de Quervain and others, 1973) and possibilities for physical interpretation, backs up the opinion that this approach is sound.

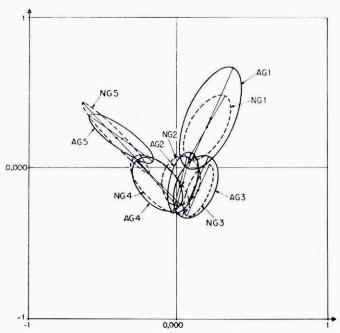


Fig. 7. Groups of avalanche (AG_i) and non-avalanche days (NG_i) in the factor space F_1 and F_2 (March/April group).

5.5. Models for avalanche occurrence

The variables defining the avalanche groups are essentially meteorological ones. We used the discrimination model (SITYP) to classify also the days without recorded avalanches. We find for type (1)—defined by the avalanche group AG_I of 25 individuals—75 other individuals, very similar but without avalanches, constituting a group NG_I . This type (1), defined by its snow and meteorological variables, occurred among the N days of the 12 year test period $NG_I + AG_I$ times. The fraction, $\frac{NG_I + AG_I}{N} = 0.13$, defines an a priori probability of this situation (1) occurring. By a similar reasoning, the a priori probability of an avalanche day occurring in type (1) is $\frac{AG_I}{AG_I + NG_I} = 0.25$.

Calibrating another discrimination model (OCCUR) between AG₁ and NG₁, we tried to answer why a given day of type (1) is to be found in group NG₁ rather than in group AG₁. We find that the variables discriminating best between AG₁ and NG₁ often differ from those separating the types of situation and are also essentially different for each model OCCUR {k}.

5.6. Operational models, results

The two-stage decision process consists roughly in deciding:

- 1. What type of situation, given the most important variables Y_j , is likely to occur today?
- 2. Assuming we are in a given type, what is the probability of avalanche occurrence (discriminating within the second set of variables).

Up to now, this model has only been tested in deferred operational forecast for the winters 1972/73 and 1973/74, which were out of the calibration samples. Results and a comparison with model I are given in Figure 8. As is to be expected, model IV in general has a sharp response.

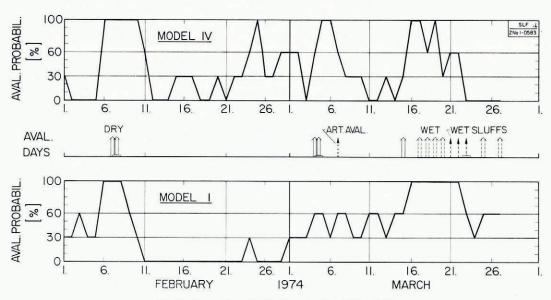


Fig. 8. Comparison of daily forecasts of model I and IV.

6. Conclusions

Statistical methods are a first, valuable step towards quantitativity and objectiveness in avalanche forecasting. The next breakthrough in avalanche forecasting, however, has to come through more accurate avalanche observations and physically more relevant snow-pack information.

Evaluating the various methods and comparing their results, we may conclude as follows:

- 1. The conventional method and the statistical models I and IV yield similar results. For model IV only preliminary results are as yet available. The mean score amounts to 70-80%, i.e. 70 to 80 d out of 100 are well classified.
- Models II and III (global model) show a slightly reduced score; in this case 60-70% of the days are well classified.
- 3. The statistical methods, especially models II and IV, using objective selection criteria have chosen partly the same significant "avalanche-prone" variables which have been used for many years by forecasters: New-snow depth (not the water-equivalent of new snow!), wind speed, air temperature, and radiation. However it is interesting to see that in model II the new-snow depth is replaced during mid-winter by the active snow layer (new-snow depth minus settlement) as the most significant variable.
- 4. Problems with the basic data set such as wrong timing, temporal correlation of avalanche occurrances, lack of good indices for structural instability of the snow-cover (an "avalanche-day" is a very rudimentary avalanche variable) are at present the most important problems.

Statistical methods show definite potential in backing up the memory and the decision process of forecasters; however, not being explanatory, they replace neither the qualitative reasoning of the conventional method not the quantitative—but still lacking—physically based approach of deterministic models.

REFERENCES

- 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.

 Bois, P., and Obled, C. 1976. Prévisions des avalanches par des méthodes statistiques. Comité Technique de la
- Société Hydrotechnique de France. Session des 18 et 19 Mars 1976, Paris.
- Bois, P., and others. [1975.] Multivariate data analysis as a tool for day-to-day avalanche forecast, [by] P. Bois, C. Obled and W. Good. [Union Géodésique et Géophysique Internationale. Association Internationale des Sciences Hydrologiques. Commission des Neiges et Glaces.] Symposium. Mécanique de la neige. Actes du colloque de Grindelwald, avril 1974, p. 391-403. (IAHS-AISH Publication No. 114.)
 Bovis, M. J. 1974. Development of methodology for evaluation and prediction of avalanche hazard in the San
- Juan mountain area of southwestern Colorado. Ch. 6. Statistical analysis. University of Colorado. Institute of Arctic and Alpine Research. Occasional Paper No. 13, p. 71-94.
- Diday, E. 1971. Une nouvelle méthode en classification automatique et reconnaissance des formes: la méthode
- des nuées dynamiques. Revue Statistique Appliquée, Vol. 19, No. 2, p. 19-33.

 Duband, D. 1970. Reconnaissance dynamique de la forme des situations météorologiques. Application à la prévision quantitative de précipitation. Grenoble, Electricité de France, Division Technique Générale. Production Hydraulique.
- 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.

 Miller, L., and Miller, D. 1973. The computer as an aid in avalanche hazard forecasting. Olympia, Washington, Evergreen State College. (Student Contract.)
- Quervain, M. R. de, and others. 1973. Avalanche classification. Proposal of the Working Group on Avalanche Classification of the International Commission on Snow and Ice, M. [R.] de Quervain, Chairman, L. de Crécy, E. R. LaChapelle, K. Losev, M. Shoda. Hydrological Sciences Bulletin, Vol. 18, No. 4, p. 391-402.
- Shcherbakov, M. P. 1966. Sposob prognoza lavinnoy opasnosti po intensivnosti snegopadov na Tyan'-Shane [Method of predicting avalanche danger from snowfall intensity in the Tyan'-Shan]. Trudy. Sredneaziatskiy Nauchno-Issledovatel'skiy Gidrometeorologicheskiy Institut, Tom 27, Vyp. 42, p. 130-41.

DISCUSSION

- J. W. Glen: I assume you record as avalanche days, not artificial-avalanche days, days when both naturally and artificially released avalanches occur?
- P. Föhn: Yes-during the perception stage of the study only natural avalanche-days were included, however at the verification stage we also give some attention to "artificial-avalanche" days, because they represent a certain instability of snow cover, even though this level is not exactly known ($\approx 30-50\%$ avalanche probability).