# AVALANCHE FORECAST: EXPERIENCE USING NEAREST NEIGHBOURS

O. Buser[1) ] W. Good[2)]

Abstract. NXDAYS, the non-parametric avalanche forecast model, is discussed in the context of historical development of multivariate, multinormal discriminant type models. Three years of experience permit the reviewing of its performance. To distinguish avalanche days from non-avalanche days, the critical level is 0.3, i.e. if 3 or more out of 10 nearest neighbours are avalanche days, the actual day to forecast will also be an avalanche day. Problems with data acquisition, data- and parameter sets, as well as forecast models are mentioned.

## INTRODUCTION

The experienced avalanche forecaster does an excellent job by merging observation, knowledge, intuition, scientific reasoning, and learning ability into a personal decision scheme. Most forecasters perform best in a given area or at a specific site. The analyst trying to formalize the forecaster's working scheme and his background knowledge is confronted with a problem several layers deep. Not only are the personal reasonings very often hidden and inaccessible but the accumulated data as well are inhomogeneous, incomplete and biased. The avalanche phenomenon itself is manifold and its modelling very exacting. Another aspect of the problem is the parameter set measured in the past and for extended periods. Previously the choice more often depended on feasibility criteria than on physical reasoning so the meteorological observation have prevailed and the snowcover itself has only been scantily represented. Jaccard (1965) suggested a kind of multivariate data analysis, to get hold of the recorded data. Obled presented his ideas on numerical avalanche forecast during the 1971 spring meeting of SHF (Obled, 1971). In 1973 Judson and Erickson introduced the "storm index" into their numerical analysis in order to predict avalanche intensity (Judson, Erickson, 1973). Thanks to the personal efforts of Bois and Obled to become familiar with the data from Weissfluhjoch, and to their experience with data analysis techniques, an interesting period of collaboration between Grenoble and Davos started in 1971 (Bois, Obled, 1972, 1973). A number of more and more complex models of the discriminant type were developed to a point where it became obvious that the way they were being treated, no further data could be squeezed out (Föhn et al., 1977). The idea of introducing non-parametric methods occurred together with the necessity of activating the still unused data from avalanche observations (besides Yes/No information of the "avalanche day") (Obled, Good, 1980). In order to see the difference between statistical (parametric) methods and the techniques of data analysis without multinormality assumptions (non-parametric methods), we give a short description of both.

[1)] Dr. Othmar Buser, [2)] Dr. Walter Good, research scientists at Swiss Federal Institute for Snow and Avalanche Research, Weissfluhjoch/Davos, Switzerland

## MODELS OF THE DISCRIMINANT TYPE

Table 1 lists the observed or "raw" variables. They have been collected on a day by day basis for several decennaries. Statistical methods rely heavily on the multinormality assumption. Parameters and populations have to be represented by bell shaped (Gaussian) two- or multi-dimensional density curves. In fact this is true in a few cases only and the errors introduced are therefore considerable. This situation as well as the physical reasoning one would like to introduce into the models call for transformations and combinations of the raw variables (Sums of new snow, degree-day, relative settlement of the snowpack,..). We named the resulting parameters elaborate or explanatory variables. An advantage of multivariate distributions is to avoid empty regions in parameter space, a problem not to be neglected with point methods like NXDAYS.

### Predictor space and avalanche day

Fortunately for the skier and the mountain dweller, avalanches do not occur too frequently. In the Parsenn Area (10 * 10 $km^2$) near Davos, an avalanche was observed in less than 1 day out of five. For model development, however, this is a handicap and the expected stratification (altitude of release zone, aspect of the slopes, and type of avalanches) (Jaccard, 1965; Bovis, 1977) has to be omitted, otherwise some avalanche classes are void. The predictor chosen was the "avalanche day", the day with at least one observed (and recorded) avalanche in the test area. It was not possible either to use the intensity of the phenomenon (number of avalanches) because of inhomogeneity in observation. It is just not feasible to observe an area 100 km$^2$ wide and under all weather (visibility) conditions with the same accuracy and resolution every day of the winter. This statement is very crucial to NXDAYS and will be dealt with in 'description of NXDAYS'.

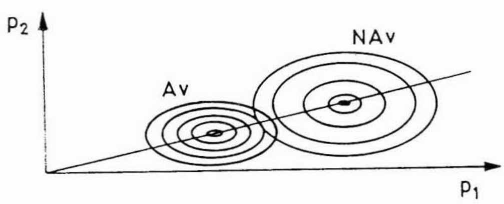



Figure 1.--Idealized avalanche- and non-avalanche populations. Equiprobability curves, principal componant axes 1 and 2 (shifted) and discriminant axis. From Buser et al.(1984)

The common part of all discriminant models is the use of an appropriate vectorspace where all days are represented by their state vectors. The "state" of the day under consideration is given by a selected set of raw and elaborate variables. In the best of cases, avalanche days and non-avalanche days would group in two distinct populations (fig. 1) and the probability of a given day belonging to either could be evaluated unambiguously.This ideal situation does not exist, however. The increasing complexity of the models tries to cope with the real situation.In addition we have to be aware that a really avalanche sensitive variable does not exist in the dataset. The only one that could create the requested dichotomy is the avalanche variable itself (IVAL), a binary parameter that is not available in real time, however (fig. 2).

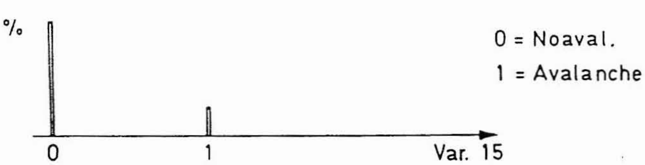



Figure 2.--Boolean variable IVAL. From Buser et al. (1984) This variable divides the whole population into avalanche- (Av) and non-avalanche days (NAv).

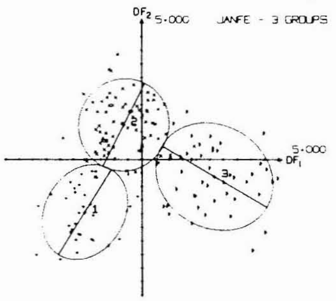
Probabilistic interpretation of the avalanche day

The probability of having an avalanche day is rather difficult to interpret; does a probability of 95 % mean a catastrophic situation or just near certainty for wet sluffs to occur? If the avalanche days by themselves could not be stratified, the attempt to do so was made using frequently occurring weather patterns. In the most complex discriminant model, and for the period of January-February, a selection of one out of three (four for March-April) weathertypes for the given day is made in a first appropriate parameterspace (see fig. 3.a). In the optimal coordinate frame for the chosen weathertype, a simple discrimination avalanche/non-avalanche is performed (fig. 3.b). The interpretation is now easier: "Heavy daily snowfall, strong winds, low air temperature" for instance or "overcast sky, high and rising air temperature" (Obled, Good, 1980), trigger different appreciation and decision patterns in the mind of the forecaster helping to interpret the above mentioned 95 % for an avalanche day. The question, however, where the avalanches occur remains unanswered. All discriminant models have to work with the lost identity of avalanche days.

Discriminant techniques are one method among others of dealing with complex data. LaChapelle in 1977 and Buser et al. in 1984 reviewed the different possible approaches and the various techniques to use for objective avalanche predictions (LaChapelle, 1974; Buser et al., 1984).

NON-PARAMETRIC MODELS

A different approach works with the discrete points for all days in a p (parameter)-dimensional space. In contrast to the above methods, the coordinate axes are single parameters and not linear combinations thereof. The coordinate frame is thus not orthogonal. For parametric - as well as non-parametric methods it is important to have "reasonable" variables. We have to compare fresh snow in mm with total snow height in cm. The latter at least is subjected to strong seasonal variations. Wind speed in 1/10 m/s and duration of sunshine (1/10 h) may in fact be anticorrelated, the comparison, however, is difficult. It is good practice to use dimensionless and reasonably normalized parameters. One possibility is to subtract mean values (or gliding means) from the variables and to divide the difference by the mean ($(x-\bar{x})/\bar{x}$). The statistical correlation between parameters, like amount of fresh snow (HNF) and penetration depth of the rammsonde (TIEF), give an emphasis on unsettled new snow. The correlation between direction (DIR) and speed (VENT) of prevailing winds on one hand and precipitation (PREC) on the other accentuate again the ponderation of fresh snow.
If an adequate weight vector does account for these correlations, ponderated distances between pairs of points (states of the given day) can be



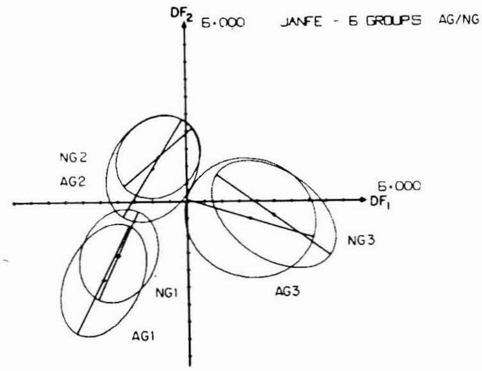




Figure 3.--Two level discriminant model SITYP for January-February.

a.: The first discriminant level splits all data into 3 clusters representing 3 weathertypes.

b.: The second discriminant level partitions each cluster into avalanche groups (AG) and non-avalanche groups (NG). From Obled, Good, 1980.

evaluated according to the rules of Euclidian geometry. The more similar the character of a day is to its "nearest neighbours", the shorter the distances are among them. We named this non-parametric procedure the method of "N nearest neighbours" (Obled, Good, 1980) and Buser, who implemented it gave it the acronym "NXDAYS" (Buser, 1983).

Description of NXDAYS

The basic ideas of the NXDAYS model are to rely on the prior probabilities of analogous, real avalanche situations and to use the information of the individual avalanche days and observed avalanches from the past. As probabilities without additional information are difficult to interpret, the second aspect proves to be more important to the forecaster.

Following the description of Buser (1983), NXDAYS asks the forecaster to input the variables representing the snow- and weather situation of the last 24 h (table 1). The distances between the state vector of today and the state vectors of all winter days for a period of 20 years are computed and ordered. NXDAYS responds with the ten nearest neighbours. The ratio of n avalanche days over N nearest days (n/N) yields the a priori probability of the actual day to be an avalanche day (fig. 5). For a good forecast method we would expect the ratio n/N to be either 0 or 1. Because we deal with the problem in a strictly punctual way in space and time and have to work with a rather poor set of raw variables, the "grey zone" between zero and one cannot be avoided. In order to get more insight into the structure of the collected data (parameters and predictor) NXDAYS offers several possibilities: - If we feel ourselves to be in a critical domain of the predictor space where information is either rare or dubious, we ask for a larger number (N= 30) of neighbours - If we do not fully agree with the selected days, we may change the weights of the ponderation vector. This computer experiment enables us to see the importance of the single parameters and the influence they have on a given situation, enhancing the "feeling" for the behaviour of the model in given situations. How does the avalanche probability change by increasing or decreasing the amount of new snow, by changing the direction and/or the speed of wind? - Is the situation really unique and do we agree with the proposed solution? This is not a vain question: especially if the dataset of previous situations extends over one or two years only, the model output might be wrong, no neighbours being in the vicinity. After the real

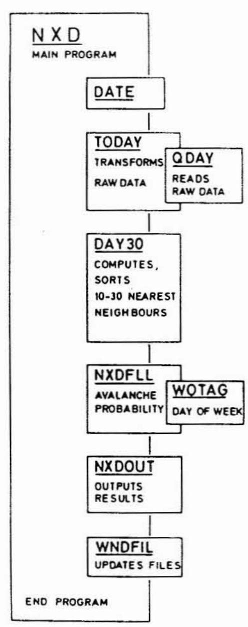



Figure 4.--Logical flow-chart of forecast model NXDAYS

time forecast has led to a satisfactory result, the files are updated with the verified parameters of the previous day.

A way to become familiar with this straight forward method is to look at the computer flowchart (fig. 4).

Experiences with NXDAYS

NXDAYS has been in operation for three winters in the test area of Weissfluhjoch. Different operators have worked with it. Critical comments and proposals have led to modifications and, it is to be hoped, to improvements.

We discuss examples of the winter 1983/84 (Regli, 1984). The forecaster was untrained both in avalanche forecast and electronic data processing. He had, however, a good scientific background. After a short training period, he worked independently. A second person, a young mountain guide, was hired to make extensive surveys of the test area. Because of previous experience, where the lack of data for selected days was felt to be crucial, his task was to ascertain the presence or the absence of avalanches. After the analysis, it is easy to say that this action was a kind of boomerang. According to the data collected, the Parsenn Area had nearly twice the number of avalanches compared to previous winters (4, instead of 2 days out of 10, see e.g. fig. 9). This drastically demonstrates the problems occurring if the data set (the recorded days in the past) are not homogeneous (extent of test area, frequency and density of field observations). This has to be kept in mind when discussing the winter 1983/84 and comparing it to previous winters.

As an illustration figure 5 gives an idea of the sequence of avalanche days from the beginning of February. It shows additional avalanche days in the second half of the month, from observations in areas adjacent to the original test site.

The general behaviour of model NXDAYS is represented in figure 6. Two probability values for each day to be an avalanche day are given: First computed from observed (certified) avalanche days (days with at least one avalanche in the test area), second from observed days and days where an observation was not possible but the occurence of avalanches probable (storm, fog, New-Year's Eve!).

The long term behaviour of the model is represented in figures 7 and 8. All the avalanche days are summed up and normalized to one. Figure 7 emphasizes the importance of a realistic discriminant level to separate avalanche days from non-avalanche days. The fine, unbroken line fits the avalanche days best. If the model finds 3 or more avalanche days in the neighbourhood of the actual day, the day to forecast will also be an avalanche day. Figures 6, 7 and 8 point to two different behaviours of the model, winter days versus spring days (transition around the $10^{th}$ of February). The model is rather insensitive to the probable avalanche days if the right discriminant level ($\geqslant 0.3$) is selected (see fig. 8). Figure 9 shows another effect of the "overobserved" winter. The sequences of the avalanche periods are too long. For the 7/8, 11, 19, and 28/29 January, NXDAYS proposes avalanche free days, interrupting the avalanche period.

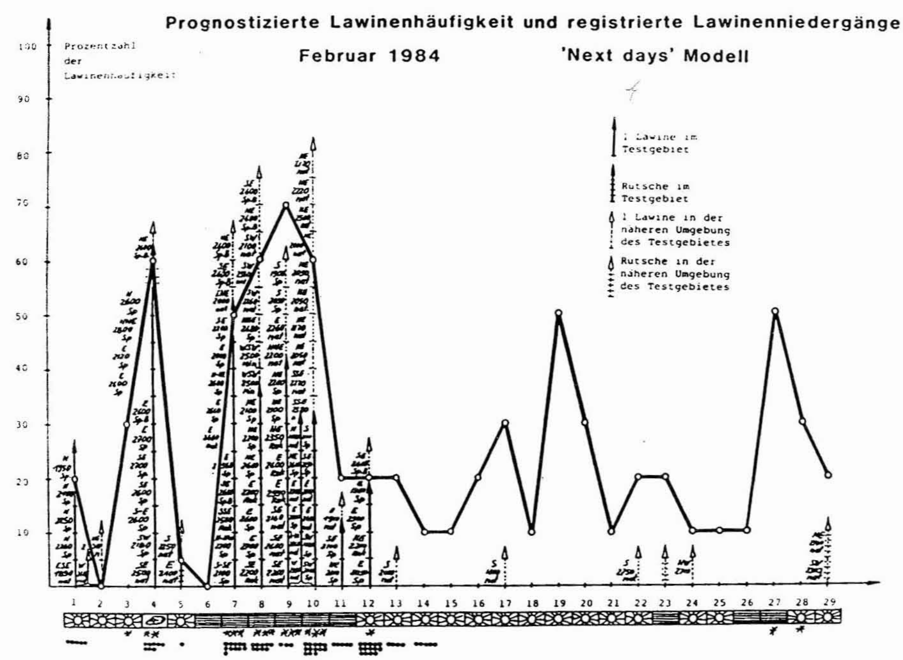



Figure 5.--February 1983/84: Observed avalanches and forecast probabilities (n/N) of avalanche days (Regli, 1984)

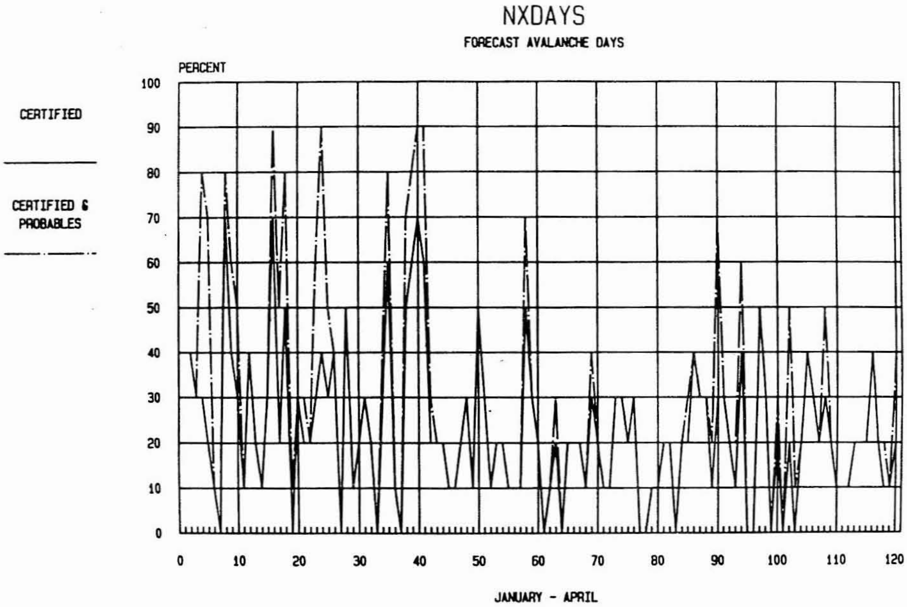



Figure 6.--NXDAYS 1983/84: Forecast avalanche days. Certified: From observed avalanches only. Certified & Probables: From observed and suspected (but not verified, because of storm, fog...) avalanches.

Improvements of NXDAYS

Inspite of the realistic behaviour of NXDAYS and of its potential not only to forecast an avalanche day but to furnish additional information as to where and how avalanches are likely to occur, the inherent weaknesses of the model itself are still important: As a complement to the raw variables, parameters have to be introduced in order to extend the memory of the model. This is especially critical for $H_S$, the amount of new snow for the last 24 h. It is important to know whether 30 cm of new snow will build up on a settled snowcover or whether it will be on top of another 30 cm from yesterday. Other parameters as discussed in (Obled, Good, 1980) will have to be analysed carefully: The settlement of the snowcover, the formation of surface hoar during extended periods of cold and clear weather, the temperature development in the snowpack, etc. This may either be done by introducing elaborate variables into the state vector or by computing the additional parameters backward, starting with the actual day. The first procedure, analogous to the discriminant models, would increase the time to compute distances considerably. The second carries out the evaluation in a subspace only. In a similar way the relationship between sequential days may be found.

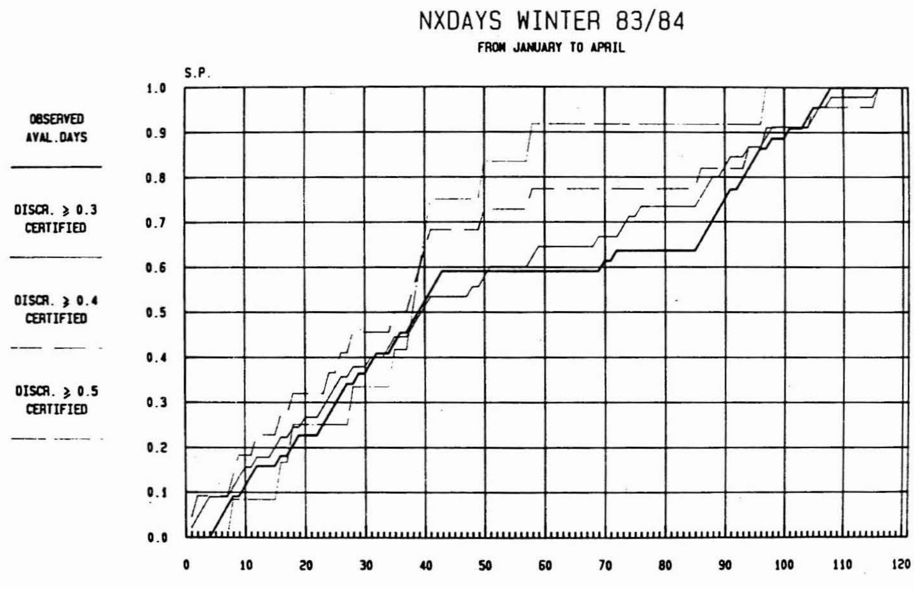



Figure 7.--NXDAYS 1983/84: Summed probabilities showing critical levels of probabilities for discriminating Av/NAv -days.

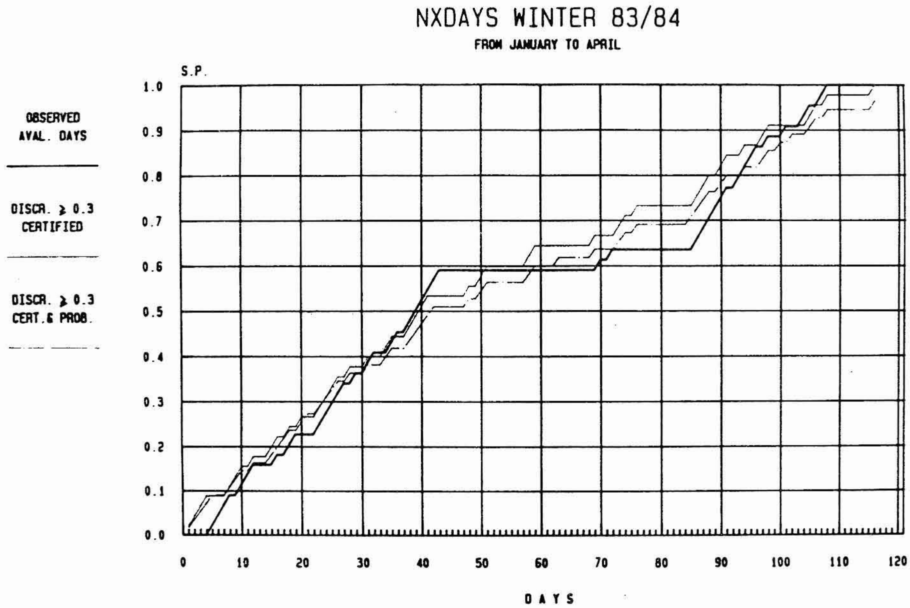



Figure 8.--NXDAYS 1983/84: Summed probabilities for the best level of discrimination (≥ = 0.3) and both interpretations of the data set (certified, cert. & prob.).

Limits to NXDAYS

Many problems left are not specific to NXDAYS. The 20 % prior probability for an avalanche day in the data set (winters 1960-1980) and the 40 % probability verified in a winter with exceptionally "good" observation emphasizes the importance of homogeneity. Observation, data collection, data handling, model- use and interpretation have to be done in a consistent way. The forecaster has to become familiar with all those aspects of his job.

CONCLUSION

NXDAYS is a clear and straightforward method to assess avalanche situations. In some basic aspects, the model behaves similarly to the

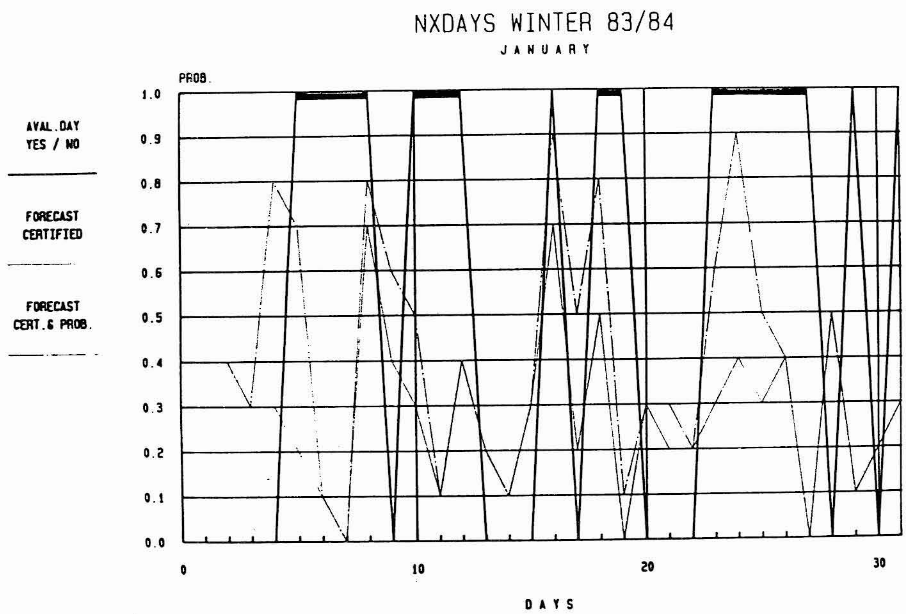



Figure 9.--NXDAYS January 1984: Probabilities for observed and forecast avalanche days. Note out-of phase periods 6/7, 11, 19, 28/29 January.

forecaster's intuition. The prior probability and its critical level ($\geq 0.3$) help to avoid drastic errors. Perhaps the most interesting feature for the forecaster, however, is the possibility to compare the actual situation to analogous ones in the past. This amounts to sharpening and extending the forecaster's memory. He can interactively check the behaviour and response of the model and his reasoning in real time or afterwards in less time-critical situations.

## ACKNOWLEDGEMENTS

The help of Beat Regli and Jachen Planta, and the work, discussions, and criticism of the forecasters from the avalanche group of the Institute is gratefully acknowledged. The ski patrol team of Parsenndienst helped to collect avalanche data. Monika Bütler was an efficient help in the preparation of this report. Evagreth Fopp did the careful typing of the camera ready copy.

## REFERENCES

Bois, Ph., and Ch. Obled. 1972. Analyse des données nivoclimatologiques en vue d'une prévision des avalanches. Institut National Polytechnique de Grenoble, Laboratoires de Mécanique des Fluides. Premier Rapport.

Bois, Ph., and Ch. Obled. 1973. Vers un systême opérationnel de prévision numérique des avalanches à partir de méthodes statistiques. Hydrol. Sc. Bull. XVIII, 4, 12, p. 419-429.

Bovis, M. 1977. Statistical Forecasting of Snow Avalanches, San Juan Mountains, Southern Colorado, USA. J. Glaciol. 18, 78, p. 87-99.

Buser, O. 1983. Avalanche Forecast with the Method of Nearest Neighbours: An Interactive Approach. Cold Regions Science and Technology, 8, p. 155-163.

Buser, O., and P. Föhn, and W. Good, and H. Gubler, and B. Salm. 1984. Different Methods for the Assessment of Avalanche Danger. Cold Regions Science and Technology. (in press).

Föhn, P., and W. Good, and Ph. Bois, and Ch. Obled. 1977. Evaluation and Comparison of Statistical and Conventional Methods of Forecasting Avalanche Hazard. J. Glaciol., 19, 81, p. 375-387.

Jaccard, C. 1965. Statistische Analyse der Lawinenereignisse und Lawinenvorhersage. Eidg. Institut für Schnee- und Lawinenforschung, Weissfluhjoch/Davos. Int. Bericht 450.

Judson, A., and B. Erickson. 1973. Predicting Avalanche Intensity from Weather Data: A Statistical Analysis. USDA Forest Service Research Paper RM-112.

LaChapelle, E. 1974. Snow Avalanches: A Review of Current Research and Applications. J. Glaciol. 19, 81, p. 313-324 (51 references).

Obled, Ch. 1971. Vers une prévision numérique des avalanches. Réunion de la section glaciologique de la Société Hydrotechnique de France (SHF) du 3-4 Mars 1971.

Obled, Ch., and W. Good. 1980. Recent Developments of Avalanche Forecasting by Discriminant Analysis Techniques: A Methodological Review and some Applications to the Parsenn Area (Davos, Switzerland). J. Glaciol. 25, 92, p. 315-345 (32 references).

Regli, B. 1984. Bewertung des Modells "Next days" für die Zwecke der lokalen Lawinenwarnung. Eidg. Institut für Schnee- und Lawinenforschung, Weissfluhjoch/Davos. Int. Bericht 620.

| | | |
|---|---|---|
| 1 | INTSI | Max intensity of precipitation (0.1 mm/h) |
| 2 | PREC | Precipitation (0.1 mm) |
| 3 | HNF | Fresh snow (mm) |
| 4 | PEG | Total snow height (cm) |
| 5 | TIEF | Penetration depth (cm) |
| 6 | T10 | Neg. snow temperature 10 cm below snowface (0.1 C) |
| 7 | CHAS | Snowdrift (index) |
| 8 | T | Temperature at 13.00 (0.1 C) |
| 9 | VENT | Max. windspeed (0.1 m/s) |
| 10 | DIR | Direction of max. wind (10-degrees north=36) |
| 11 | NEB | Cloudiness(mean of three observations) (%) |
| 12 | SOL | Sunshine (0.1 h) |
| 13 | RAD | Solar radiation (number of counts of integrator) |
| 15 | IVAL | Avalanche code (0=none, 88=uncertain, probably not recorded) |

Table 1. --List of raw variables.
13 variables are observed or measured and recorded once a day. The avalanche variable is updated from observer reports as soon as the data are available.