STATISTICAL ASSESSMENT OF REGIONAL AVALANCHE DANGER USING PARAMETERS FROM THE SNOW COVER MODEL "SNOWPACK"

Michael Schirmer *, Michael Lehning, Jürg Schweizer
WSL Institute for Snow and Avalanche Research SLF, Davos, Switzerland

ABSTRACT: Numerical avalanche prediction with statistical methods using meteorological input parameters has shown insufficient results, mainly since snow cover information was lacking. Snow cover data were not used because they were not readily available (manual observations). With the development and increasing use of snow cover models this deficiency can now be rectified and model output can be used as input for forecasting models. We used the output of the physically based snow cover model SNOWPACK combined with meteorological variables to investigate and establish a link to regional avalanche danger. Snow stratigraphy was simulated for the location of an automatic weather station near Davos (Switzerland) over nine winters. Only dry snow situations were considered. Statistical methods, including classification trees, artificial neural networks, support vector machines, hidden markov models and nearest neighbor methods, augmented with several feature selection algorithms, were trained using the forecasted regional avalanche danger (European avalanche danger scale) as target parameter. The best results were achieved with a nearest neighbor method which used the avalanche danger level of the previous day as additional input. This method was combined with a feature selection based on fisher's discriminant analysis and a genetic algorithm. A cross-validated hit rate of 73% was obtained for all days and 55% for days when the avalanche danger level either increased or decreased. This study suggests that SNOWPACK parameters are able to improve numerical avalanche forecasting.

KEYWORDS: Avalanche forecasting, numerical simulation, snow cover model, snow cover stability nearest neighbor, feature selection, genetic algorithm.

1. INTRODUCTION

Regional avalanche forecasting attempts to predict current and future snow stability relative to a given triggering level on the scale of a mountain range or a considerable fraction thereof (e.g. McClung and Schaerer, 2006). Forecasts are issued on a daily basis to warn the public about the level of avalanche danger. These public bulletins play a key role in the prevention of avalanche fatalities. Adequate avalanche warnings, combined with avalanche education and efficient rescue have probably prevented an increase of avalanche fatalities in parallel with the increased recreational use of avalanche terrain.

Reliable and consistent avalanche forecasts are therefore very much needed. To assess the avalanche danger level, most avalanche warning services still today rely on a combination of manual observations, automatic weather stations, weather forecasts (including model output) and snow profiles (Meister, 1995). For the locations of the automatic weather stations in the Swiss Alps, the amounts of new snow and drifting snow are additionally derived from the numerical snow cover model SNOWPACK (Lehning et al., 1999). Based on all these data, the forecaster uses experience, intuition and local knowledge of the mountain range to estimate and describe the avalanche danger in the public bulletin.

Over the past decades there have been many attempts to create an objective process of danger evaluation, which may also work as a support tool for the avalanche warning service. The French model chain SAFRAN-Crocus-MEPRA (SCM) provides an automated avalanche danger prediction for virtual slopes (Durand et al. 1999). Several studies have been used observations of avalanche activity as an indicator of the avalanche danger (e.g. Buser, 1983; Heierli, 2004; Pozdnoukhov, in press). As Schweizer et al. (2003) pointed out, the problem with this target parameter is that it does not distinguish between lower danger levels and that observations might be inconsistent, mainly due to limited visibility during times of avalanche activity. Schweizer and Föhn (1996) used a commercial decision-making

Corresponding author address: Michael Schirmer, WSL Institute for Snow and Avalanche Research SLF, Flüelastrasse 11, CH-7260 Davos Dorf, Switzerland; tel: +41 81 4170282;
fax: +41 81 4170110; email: schirmer@slf.ch

software to forecast the avalanche danger level. Input parameters included snowpack information. They trained the model using a verified danger level instead of the forecasted one. The verification was based on additional data and observations, and data not yet available at the time of the forecast. The cross-validated hit rate was 63%, but adding knowledge in the form of expert rules to the system, the performance improved to a hit rate of about 70%. However, their models did not run fully automatically but required manual input of the snow cover information. Therefore, Brabec and Meister (2001) used a nearest neighbor method with only meteorological parameters and snow information such as penetration depth or surface characteristics so that the model could be run for the whole area of the Swiss Alps. The accuracy was only about 52%. The absence of snow cover information was given as a reason for the insufficient results.

Model output from snow cover models such as SNOWPACK can provide the snow cover information with the required resolution in space and time. This study explores the question of whether the performance of data-based forecasting models can be improved with modeled snow cover data as additional input.

2. METHODS

2.1 Data

The forecast of the regional avalanche danger degree (Fig. 1), which is issued every day at 8:00 am for the region of Davos (Switzerland), was used as target parameter.

Modeled snowpack data were generated for the location of an automated weather station (Weissfluhjoch, 2540 m a.s.l.) in the test region over a time period of nine winters (1999 – 2007: 1229 days). We focused on dry snow situations with HS $\geq$ 75 cm. The snow cover model SNOWPACK was used to model settling and layering of the snowpack as well as its energy and mass balance (Bartelt and Lehning, 2002; Lehning et al., 2002a,b; Lehning and Fierz, 2008) This model requires meteorological parameters as input. A stability index (Schweizer et al. 2006) defined the potential weak layer interface in the modeled snowpack. Motivated by a study which evaluated observed snow stratigraphy in regard to stability (Schweizer and Jamieson, 2003), characteristics of the weak and adjacent layer and of the slab were considered. These model parameters were completed with measured and calculated meteorological and snow parameters (e.g. wind velocity, longwave emission or surface albedo).

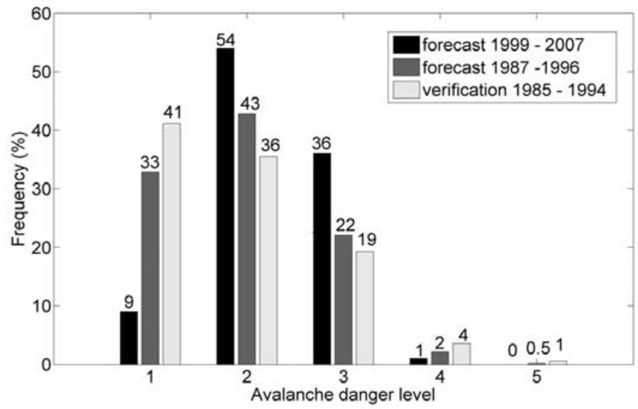

Fig. 1. Relative frequency of the avalanche danger levels for the region of Davos. Black columns for the forecasted levels during the time period considered in this study, dark grey for the study of Brabec and Meister (2001) and light grey for the verified levels used in the study of Schweizer and Föhn (1996).

2.2 Feature selection

Since the model delivers a huge number of parameters at high temporal resolution, a feature selection algorithm is useful for many reasons: First for data reduction, which increases the speed of the final model construction, second to achieve an overall improved performance of the final model and third to understand which parameters are important (Guyon and Elisseeff, 2006).

For many parameters, it may make sense to also consider their sum, mean, extreme values or gradient, for different time intervals. This multiplies the number of parameters. A simple univariate feature selection, a fisher's discriminant analysis (e.g. Bishop, 2006), was performed to determine for each parameter the two most important derived variables. Subsequently a rating of this feature set was carried out with fisher's discriminant analysis, classification trees (Breiman et al., 1998) or recursive feature elimination using support vector machines (Guyon et al. 2002). Only parameters were considered which were not pair wise linearly correlated (r^2 > 0.8). Additionally, for the nearest neighbor method a genetic algorithm was used as a search tool (Li et al., 2001). For all methods described parameters were scaled linearly to [-1,1].

2.3 <u>Statistical methods</u>

Several statistical methods were used in order to find an optimal link between input parameters and predicted regional avalanche danger. Three methods are described, which each show typical pros and cons. A simple classification tree with only one parameter, the 3d-sum of the new snow, performed very well and was therefore used as a benchmark for more complex models with more (especially modeled) input parameters (Breiman, 1998). Artificial neural networks (ANN) (e.g. Bishop, 2007) were trained for each winter. The results of the ANN-method discussed in the next section were gained with a recurrent network. The same nearest neighbor (KNN) approach as used by Brabec and Meister (2001) was applied for a direct comparison to previous work. In addition, this method was modified in two ways: the avalanche danger level of the previous day was used as an additional input to predict the current day. Second, the classification was modified. A decision boundary was used to determine if the danger level should jump to a new value on a particular day. Using 10 nearest neighbors, three of these days must show a higher (or lower) danger level. This decision boundary was obtained by optimizing the POD/SR pair (Heierli et al., 2004). Subsequently the danger level was determined with a majority vote between the neighbors showing higher (or lower) danger levels. For breaking a tie the nearest neighbor among them was used.

The quality of a method was assessed by a cross-validated hit rate (HR) for all danger levels (Low, Moderate, Considerable, High) combined with probability of detection (POD) and the success rate (SR) for each level separately (Wilks, 1995). Since both input and target parameter are auto-correlated, e.g. a weak layer might have an influence over a long time period, random cross-validation is not a useful method (Brabec and Meister, 2001). Therefore each winter was forecasted using a model and a feature selection based on the remaining eight winters. It seems useful to draw special attention to days on which the danger level changed: First, these days might be the most important to predict correctly. Second, an obviously trivial method which forecasts the next day using the danger level of the present day, showed good hit rates for all days because the danger level is characterized by a very high persistence (Table 1). Introducing hit rates (HR*) for only those days on which the danger level changed, can help to identify methods which tended to a persistent forecast. Similarly, also the hit rate for these days was considered on which the danger level increased or decreased.

Table 1: Relative transmission matrix from one avalanche danger level (rows) to next level (columns).

		to avalanche danger level (%)			
		1	2	3	4
from avalanche danger level (%)	1	74	24	2	-
	2	4	82	13	1
	3	-	20	78	2
	4	-	-	86	15

3. RESULTS

The various feature selection algorithms provided similar results. Important variables were the new snow amounts (*HN*, *HN*3d), the deformation stability index (Lehning et al., 2004), the modeled surface crust thickness on south slopes, the snow transport index (Lehning et al., 2006), the calculated surface albedo, the three-hour gradient of the outgoing longwave radiation, the wind velocity and the relative humidity. The deformation index turned out to be important despite the fact that recently a bug has been discovered in the description of the temperature dependence.

Table 2 summarizes the performance of three different methods. The classification tree with only one input parameter produced only the most frequently used danger levels Moderate and Considerable. The POD and SR values of the other danger levels were therefore 0%. The hit rate was 65% using all days (HR), 56% for days on which the danger level changed compared to the previous day and 59% for days on which the modeled output changed value (HR*). Using prior probabilities, classification trees are in general also able to handle unbalanced target parameters as existing in our case (Fig. 1), but improvements on the POD/SR pairs for the danger levels Low and High produced insufficient overall hit rates.

The ANN-method was able to predict all four danger levels, but predicted the intermediate danger levels (2 and 3) with much better accuracy (Table 2). The distribution between the danger levels improved compared to the simple classification tree and the hit rate increased slightly to 69% at the same time.

The nearest neighbor method, which used avalanche danger level of the previous day as additional input could predict all four danger levels with a more balanced accuracy. The hit rate even increased to 72%. Figure 2 shows a comparison between the KNN-method and the forecasted avalanche danger for a representative winter.

Table 2: Quality scores of three different methods described in the text. The first value in HR* is the hit rate for the days where the forecast changed value, the second where the model output changes. In the next columns are the POD (first value) and SR (second value) for each level. Entries in %.

	HR	HR*	POD/SR level 1	POD/SR level 2	POD/SR level 3	POD/SR level 4
tree	65	56	0	81	58	0
		59	0	64	66	0
ANN	69	58	33	72	72	27
		57	39	72	72	26
KNN	72	54	58	82	64	54
		57	79	73	73	64

Since the distribution of the forecasted avalanche danger levels (Fig. 1) is substantially different to previous studies (Schweizer and Föhn, 1996; Brabec and Meister, 2001), results are not directly comparable. Therefore the method used by Brabec and Meister (2001) with mainly meteorological data as input was applied to the same time period and a hit rate of 55% was achieved.

3. DISCUSSION AND CONCLUSIONS

The objective of this paper was to analyze whether modeled parameters of the snow cover model SNOWPACK improve the performance of forecasting models that use statistical methods, in comparison to models that are only based on measured weather and snow input data.

The simple classification tree which used only one measured (not modeled) parameter (*HN*3d) already distinguished well between the two most frequent danger levels of Moderate and Considerable. With measured meteorological input parameters as used by Brabec and Meister (2001) a more balanced model performance on all danger levels resulted in a strong decrease in the overall hit rate, independent of the statistical method used. With additional SNOWPACK parameters the ANN- and KNN-method could optimize both aspects. These findings suggest that for balanced performance between all danger levels *and* for a good overall accuracy, more complex models and/or more parameters, in particular modeled parameters were needed. Still, a hit rate of around 70% reveals a remarkable discrepancy between the operational avalanche danger forecast and a statistical model.

Since we have tested a wealth of methods with a huge range in complexity, we feel that the main problem is that is seems impossible to reproduce the human decision on the avalanche danger with the input used in our analysis. This missing link has already been recognized as a problem in the earlier study of Schweizer and Föhn (1996), though, they have used observed snow cover parameters. Furthermore a study by Schweizer et al. (2008) showed that a level of uncertainty exists in the detection of unstable slopes with Rutschblock tests and snow profiles (POD of 70%). Such snow profile interpretations are important arguments for the human prediction or verification of the regional avalanche danger level. Two possible conclusions remain: (i) additional information, which is not formalized at present enters the decision process such as the experience and intuition of the individual, (ii) the forecasted danger level is not a good target parameter since it might be erroneous due to incorrect data at the time of the forecast or due to variations in human perception (McClung and Schaerer, 2006).

An operational prediction of the avalanche danger level with a statistical model augmented with SNOWPACK parameters as additional input variables would include the following steps: First the present snow cover is simulated with measured data, second the development of the snow cover is predicted with forecasted meteorological data for the next day. This predicted snow cover provides the additional input parameters for the statistical methods presented. This would be a fully automated process which could be applied to the whole area of the Swiss Alps. In contrast to the French SCM chain (Durand et al., 1999), this system delivers not a stability index which can be compared to avalanche activity, but a direct forecast of the regional avalanche danger.

This study showed that the uncertainty in the prediction of the avalanche danger level needs to be quantified. For example, the shift in the danger level distribution to the intermediate danger levels that obviously occurred during the last couple years should be clarified.

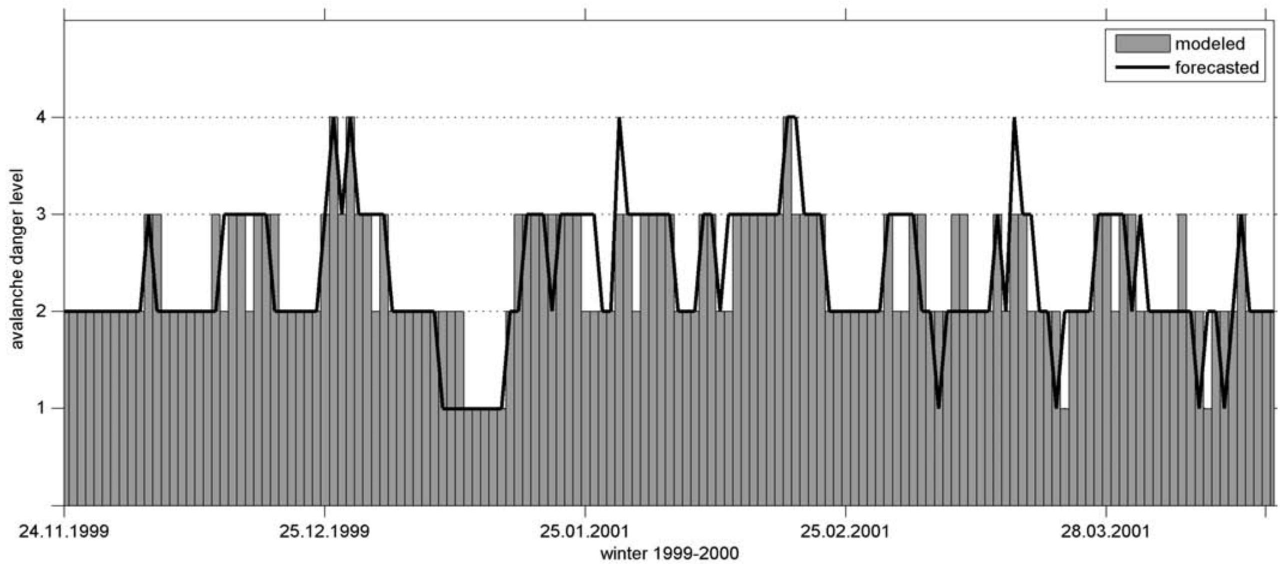

Fig. 2. The avalanche danger forecasted and modeled with the KNN-method for the winter 1999-2000.

AKNOWLEDGEMENTS

This work has partially been funded by the Swiss National Science Foundation and the Swiss Federal Office of the Environment. This study would not have been possible without the work of the avalanche warning service. We are very grateful to Tobias Sauter, Joachim Heierli, and Charles Fierz for their insightful commentaries on the ongoing work and Nick Dawes as a reviewer for valuable suggestions.

REFERENCES

Bartelt, P., Lehning, M., 2002. A physical snowpack model for the Swiss avalanche warning part I: numerical model. Cold Regions Science and Technology 35 (3), 123–145.

Bishop, C., 2006. Pattern Recognition and Machine Learning. Springer, New York, 738 pp.

Brabec, B., Meister, R., 2001. A nearest-neighbor model for regional avalanche foresting. Annals of Glaciology 32, 130–134.

Breiman, L., Friedman, J., Olshen, R., Stone, C., 1998. Classification and regression trees. CRC Press, Boca Raton, U.S.A, 368 pp.

Buser, O., 1983. Avalanche forecast with the method of nearest neighbours: an interactive approach. Cold Regions Science and Technology 8 (2), 155–163.

Durand, Y., Giraud, G., Brun, E., Mérindol, L., Martin, E., 1999. A computer based system simulating snowpack structures as a tool for regional avalanche forecasting. Journal of Glaciology 45 (151), 469–484.

Guyon, I., Elisseeff, A., 2006. An introduction to feature extraction. In: Guyon, I., Gunn, S., Nikravesh, M., Zadeh, L. (Eds.), Feature extraction: foundations and applications. Springer, Berlin, 1-28.

Guyon, I., Weston, J., Barnhill, S., 2002. Gene selection for cancer classification using support vector machines. Machine Learning 46, 389–422.

Heierli, J., Purves, R., Felber, A., Kowalski, J., 2004. Verification of nearest neighbours interpretations in avalanche forecasting. Annals of Glaciology 38, 84–88.

Lehning, M., Fierz, C., 2008. Assessment of snow transport in avalanche terrain. Cold Regions Science and Technology 51 (2-3), 240–252.

Lehning, M., Grünewald, T., Fierz, C., 2006. Assessment of mountain snow transport based on measured wind and simulated snow cover. In: Sterbenz, C. (Ed.), ISSW 2006 Proceedings. (International Snow Science Workshop, Telluride, CO USA, 1-6 October). Telluride, 815–819.

Lehning, M., Fierz, C., Brown, B., Jamieson, B., 2004. Modeling instability for the snow cover. Annals of Glaciology 38, 337–338.

Lehning, M., Bartelt, P., Brown, B., Fierz, C., 2002a. A physical snowpack model for the Swiss avalanche warning part III: meteorological forcing, thin layer formation and evaluation. Cold Regions Science and Technology 35 (3), 169–184.

Lehning, M., Bartelt, P., Brown, B., Fierz, C., Satyawali, P., 2002b. A physical snowpack model for the Swiss avalanche warning part II. snow microstructure. Cold Regions Science and Technology 35 (3), 147–167.

Lehning, M., Bartelt, P., Brown, B., Russi, T., Stöckli, U., Zimmerli, M., 1999. Snowpack model calculations for avalanche warning based upon a network of weather and snow stations. Cold Regions Science and Technology 30 (1-3), 145–157.

Li, L., Wineberg, C., Darden, T., Pedersen, L., 2001. Gene selection for sample classification based on gene expression: study of sensitivity to choice of parameters of the ga/KNN method. Bioinformatics 17, 1131–1142.

McClung, D., Schaerer, P., 2006. The Avalanche Handbook, 3rd Edition. Seattle, Washington, 342 pp.

Meister, R., 1995. Country-wide avalanche warning in Switzerland. In: ISSW 1994 Proceedings. (International Snow Science Workshop, Snowbird, Utah, USA, 30 October-3 November). Snowbird, 58–71.

Pozdnoukhov, A., Purves, R., Kanevski, M., in press. Applying machine learning methods to avalanche forecasting. Annals of Glaciology 49.

Schweizer, J., McCammon, I., Jamieson, J., 2008. Snowpack observations and fracture concepts for skier-triggering of dry-snow slab avalanches. Cold Regions Science and Technology 51 (2-3), 112–121.

Schweizer, J., Bellaire, S., Fierz, C., Lehning, M., Pielmeier, C., 2006. Evaluating and improving the stability predictions of the snow cover model SNOWPACK. Cold Regions Science and Technology 46 (1), 52–59.

Schweizer, J., Jamieson, B., 2003. Snowpack properties for snow profile interpretation. Cold Regions and Science Technology 37 (3), 233–241.

Schweizer, J., Kronholm, K., Wiesinger, T., 2003. Verification of regional snowpack stability and avalanche danger. Cold Regions Science and Technology 37 (3), 277–288.

Schweizer, J., Föhn, P., 1996. Avalanche forecasting - an expert system approach. Journal of Glaciology 42 (141), 318–332.

Wilks, D., 1995. Statistical methods in the atmospheric sciences. Academic Press, New York, 467 pp.