Cornice – development of a nearest neighbours model applied in backcountry avalanche forecasting in Scotland

[1&*]Ross Purves, [2&]Keith Morrison, [3]Graham Moss and [3]Blyth Wright

[1]Department of Geography, University of Zurich
[2]Geowise, Edinburgh, UK
[3]sportScotland Avalanche Information Service
[&]Formerly at the Department of Geography, University of Edinburgh, UK

Abstract: This paper describes the development and refinement of a nearest neighbours avalanche forecasting model in Scotland. Avalanche forecasting in Scotland is carried out in five areas, primarily for mountaineers who are either winter climbing or mountaineering in 'backcountry' locations where no avalanche control is carried out. Forecasting is primarily conventional, but computer-models are also used – NXD was introduced in the early nineties and based on our experiences with NXD the model described in this paper, Cornice, was developed and introduced.

Cornice has been developed in close consultation with avalanche forecasters and aims to provide a tool which integrates well with conventional avalanche forecasting by encouraging hypothesis testing. In order to facilitate this approach, model development must focus not only on the event we are forecasting, but mechanisms for outputting useful information in a range of ways to the forecaster.

In this paper we describe two novel elements of Cornice: the development of automated, objective methods for weighting of parameters through the use of a genetic algorithm and the use of graphical output from forecasts allowing the forecaster to visualise relationships between neighbours in terms of clustering in meteorological/snowpack variables and the identification of spatial clusters of avalanche occurrence in location, altitude and aspect.

We present results from the operational use of the model over the past three winters for the Ben Nevis area and examine its integration into a conventional avalanche forecasting framework.

Keywords: Avalanche forecasting; avalanches; computer models.

1. Introduction

This paper describes the development and refinement of a nearest neighbours avalanche forecasting model (Cornice) now being used operationally in Scotland. We describe how a model has been developed in consultation with forecasters carrying out primarily conventional forecasting, with the aim of providing a tool which enhances the decision making process by encouraging hypothesis testing.

1.1 Avalanche forecasting in Scotland

Avalanche forecasting in Scotland is carried out in five areas on a daily basis throughout the winter season (generally December to April). The forecasting is aimed at those participating in 'backcountry' activities – which in the Scottish context consist primarily of winter mountaineering, with small numbers of ski tourers (Anderson, 1994). The areas for which forecasts are provided are heavily used, with it being not uncommon for tens of participants to be undertaking winter climbs in a single corrie. The avalanche forecasts produced consist of a daily report of the current snowpack conditions, contributing weather and avalanche hazard and a forecast for the development of avalanche hazard over the next twenty-four hours.

The Scottish snowpack is shaped by a maritime climate at a relatively northerly latitude (~57°N) with mountains of comparatively low elevation (~1000m). Weather systems from the Atlantic track rapidly over Scotland and air temperatures and wind speeds and directions can fluctuate very rapidly. Barton (1987) reported data from an Automatic Weather Station on Cairn Gorm (1245m) including mean annual wind speeds of $13ms^{-1}$.

The forecast is produced by observers who make daily measurements of snowpack conditions, avalanche activity and general meteorological parameters within their forecast area. They receive a tailored meteorological forecast from the Meteorological Office and also have access to a nearest neighbours avalanche forecasting model (formerly an adapted version of the Swiss NXD and

* *Corresponding author address:* Ross Purves, Department of Geography, University of Zurich – Irchel, Winterthurerstrasse, 190, CH-8057, Zurich, Switzerland. Tel: +41 (0)1 635 6531; fax: +41 (0)1 635 6848; email: rsp@geo.unizh.ch

now Cornice) which they may utilize in producing the final avalanche forecast. Avalanche forecasts are prominently displayed at local access points to the hills and disseminated through a range of media sources including the internet (www.sais.gov.uk). Forecasts describe the areas of highest hazard (using the European Avalanche Hazard Scale) and the weather and snowpack conditions associated with the hazard. No avalanche control work is carried out, although forecasters may test-trigger small, safe slopes when they are in the field.

1.2 Avalanche forecasting and the application of computer models in avalanche forecasting

LaChappelle (1980) characterized conventional avalanche forecasting as an empirical process relying on the knowledge, experience and intuition of the forecaster. He noted how a range of deterministic and non-deterministic processes were used in arriving at a forecast and commented on the use of diverse and often redundant information sources. As described by LaChappelle conventional avalanche forecasting is essentially a hypothesis testing approach, with the forecaster collecting a wide range of information in order to test and refine an initial hypothesis.

Research has been carried out for many years on the use of computer models in avalanche forecasting both at regional and local scales. One approach which has been shown to be successful is the use of nearest neighbours models as originally developed by Buser (1983) and subsequently reported on and developed by a number of authors (Buser, 1989; Kristensen and Larson, 1994; Bolognesi, 1994; McClung and Tweedy, 1994; Gassner *et al.*, 2000; Brabec *et al.*, 2001). Nearest neighbours models work on the simple principle of finding the most similar days in the past to the forecast conditions and allowing the forecaster to compare the resulting events. Either some predetermined rule is then used to determine the likelihood of avalanche occurrence, or the forecaster may simply use the model as an *aide memoir*. Nearest neighbours models are attractive since they are intuitive in their operation and their data requirements are relatively modest, in general being homogeneous snowpack, weather and avalanche data over a period of some years. Such data are typically collected by backcountry avalanche forecasters in the normal course of their work.

2. Developing Cornice

Cornice was developed with the aim of providing a tool for use in conventional avalanche forecasting. With this in mind a number of issues were considered including:

- provision of a more user-friendly interface than the (then) DOS-based NXD;
- implementation of batch testing and automated weighting and scaling techniques;
- provision of a range of outputs which encourage the use of the model to test hypotheses.

Amongst other issues considering important in the development of Cornice was the development of an 'open' data format which would facilitate importing and exporting of data between software packages. Here we provide a brief description of the technical implementation of Cornice, before describing the latter two points in more detail.

2.1 Implementation of Cornice

Cornice is implemented in Java™. This has the advantage that code is in principle platform independent and in practice can be run under a wide range of operating systems. Data are stored using a bespoke XML file format with an associated DTD (Document Type Definition) which allows easy import and export of data and the use of a wide range of free software tools in viewing and processing the data (Ray, 2001).

As with all implementations of nearest neighbours models, each day in the dataset is represented by meteorological and snowpack variables considered to be representative of the event we are trying to forecast (in this case avalanches) and information about the nature and occurrence of any events. Furthermore, each variable is assigned a scale, which normalizes its value with respect to all other variables. Finally, variables may also be assigned a weight to indicate their relative importance in carrying out a forecast – for instance we may consider snow drifting to be much more important than air temperature and thus assign it a higher weight. Cornice, unlike NXD2000 (Gassner *et al.*, 2000) does not use elaborate variables which attempt to "introduce physical knowledge about the underlying phenomena". The calculation of scales and weights in Cornice is described in Section 2.2.

To perform a forecast the Euclidean distance of the forecast day is measured from every other day in the dataset as follows:

$$\begin{aligned} d_{ij}^{\ 2} &= (s_1 w_1 x_{1i} - s_1 w_1 x_{1j})^2 + \\ &(s_1 pd' w_1 x_{1i} - s_1 pd' w_1\ x_{1j})^2 + \\ &\ldots \\ &+ (s_n w_n x_{ni} - s_n w_n x_{nj})^2 + \\ &(s_n pd' w_n x_{ni} - s_n pd' w_n x_{nj})^2 + \\ &\ldots \end{aligned} \tag{1}$$

where

s_n is the scaling factor of variable n;
w_n is the weight of variable n;
x_{ni} is the value of variable n on day i; and
$pd'w_{in}$ is the weight applied to variable n on days previous to the day in question*.

---

* Since avalanches are the result of weather and snowpack history nearest neighbours models generally consider 3-5 days in identifying similar days, in other words they actually identify *similar patterns* of days.

The whole dataset is then sorted according to the calculated distances and the ten nearest neighbours are made available to the forecaster for use in their analysis of the forecast results.

2.2 Automated batch testing, scaling and weighting

2.2.1 Batch testing and scaling

One of the major drawbacks of NXD was that the computation of a 'batch forecast' returning the results of forecasts for every day in a dataset was not possible except through a tedious manual process. Both NXD2000 (Gassner *et. al.*, 2000) and Cornice address this problem through the provision of automated batch testing. In the case of Cornice batch tests return the number of correctly forecast avalanche days, the number of correctly forecast non-avalanche days and the number of no-visibility days. The unweighted average of the former two values is also returned for use in optimisation. The average is unweighted since even if a very small number of days in a dataset have avalanches it is important that these are correctly identified. Scaling the data is performed automatically in order to give each variable approximately the same range.

2.2.2 Weighting

Weights are used in nearest neighbours models to allow certain variables to have more influence in calculating the distance between days. In the original versions of NXD weighting was considered an 'expert' task and performed by the local avalanche expert. In NXD2000 (Gassner *et al*, 2000) weighting is still described as being carried out by trial and error, though correlation methods are also used on the source data with external software. (Felber, A., *peers. comm.*) However, although a local avalanche expert may well know which variables are important in predicting avalanches, it is much less likely that they will correctly predict weights for all days in a dataset in an n-dimensional space. Indeed, our experience in Scotland tended to suggest that such an approach could lead to the mode becoming *over tuned* to recent winters, or events which were considered to be particularly serious omissions. Over tuning implies that the model performs well on a small dataset, but when presented with conditions outside the tuning parameters performs much less well. We preferred to attempt a solution whereby we optimized the model over the entire dataset.

Since we were attempting to solve an optimization problem, namely which set of weights will allow the nearest neighbours model to accurately forecast the maximum number of avalanche and non-avalanche days we decided to implement an algorithm to automate this procedure. After a review of different methodologies it was decided to utilize genetic algorithms which are "adaptive methods which may be used to solve search and optimization problems" (Beasley, 1993a). Genetic algorithms essentially work by simulating evolution, where the fittest individuals survive and pass their characteristics to their offspring. In the case of a nearest neighbours model, fitness is characterised by the ability of a set of weights to forecast avalanche and non-avalanche days, and an individual's characteristics are determined by the weight ascribed to each variable. Figure 1 shows schematically the operation of a genetic algorithm for a population of four individuals, each with three weights assigned.

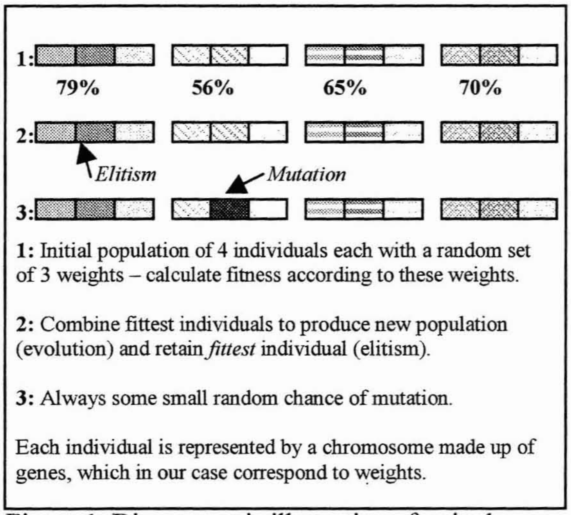

Figure 1: Diagrammatic illustration of a single iteration of a genetic algorithm

In a real dataset a population may consist of tens or hundreds of individuals, each with many genes (weights) and the algorithm may iterate either for a fixed number of iterations or until some pre-defined convergence criteria (such as a maximum fitness level) is reached. A good review of many of the issues relating to genetic algorithms and their implementation is in (Beasley, 1993a, Beasley, 1993b).

2.3 Model outputs

Model outputs were developed in conjunction with forecasters to provide useful information which would facilitate hypothesis testing. Three main modes of output have been implemented: textual output; symbolic output and map output. Textual outputs simply display the ten nearest neighbours and associated data with another window containing avalanche reports for the day in question. This mode is similar to what NXD already provided. Symbolic output displays variables graphically, allowing forecasters to easily identify clusters of data points of similar values (Figure 2). This mode of visualization aids identification of variables which might be important in sensitivity testing – for instance if the forecaster identifies that all of the nearest neighbours have similar wind directions they can perform tests to evaluate the effect of changing the forecast wind direction by small amounts.

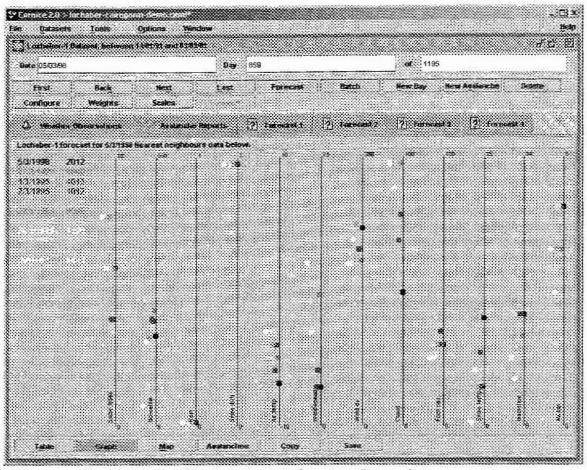

Figure 2: Symbolic output from Cornice

The final mode of output displays avalanche data for the ten nearest neighbours on a geo-referenced map (Figure 3). This allows the forecaster to examine the geographic clustering or otherwise of avalanches on nearest neighbour days – did all the avalanches occur on similar slopes or aspects, or are they widely dispersed through a range of slopes and aspects? Did an avalanche occur in an unexpected location on a previous day, and does this suggest unusual or unexpected areas of avalanche hazard?

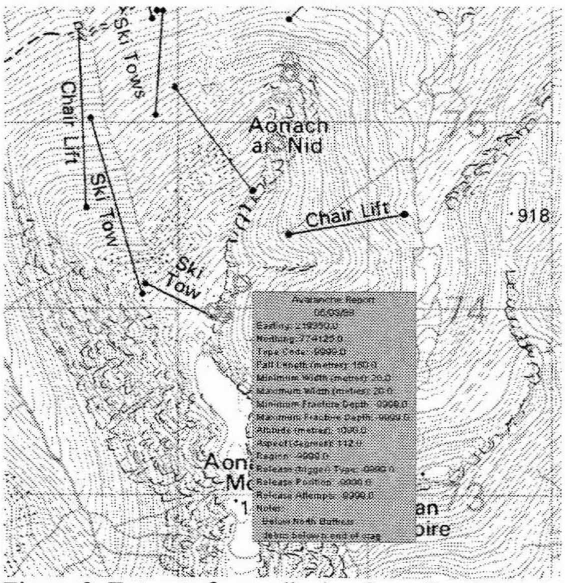

Figure 3: Extract of map display from Cornice showing avalanches and avalanche report display (Mapping data reproduced from Ordnance Survey maps, Crown Copyright)

3. Configuration and model results for Lochaber

We report on configuring the model for one of the five avalanche forecasting areas in Scotland, Lochaber. The Lochaber area includes Ben Nevis, Scotland's highest mountain (1344 m). It is a very popular winter climbing venue and has been scene of many avalanche incidents and accidents over the years. The variables used in Cornice for the Lochaber area are shown in Table 1.

Variable name	Description
Snow index	An index of the precipitation as snow on a day (measured in the field by the observer)
No settle snow index	The total sum of the snow index over the winter
Rain at 900m	A common occurrence in Scotland! (Yes/no)
Snow drifting	Drifting occurring during the observation period (Yes/no)
Air temperature	Midday air temperature measured at an automatic weather station (AWS)
Wind speed	24 hour vector mean wind speed from the AWS
Wind direction	24 hour vector mean wind speed from the AWS
Cloud cover	Cloud cover as a percentage
Foot penetration	Measured in centimetres at the pit site
Snow temperature	Snow temperature at a depth of 10cm at the pit site
Insolation	Theoretical insolation for day in the season

Table 1: Variables used in the Lochaber area in Cornice

It should be noted that a number of indices are used. In particular the snow index is a qualitative index of the precipitation as new snow in a given day. This was developed in response to the forecasters' dissatisfaction with measuring new snow depth at a fixed site, which due to the almost daily drifting experienced in Scotland they felt to be unrepresentative of avalanche hazard. The dataset consists of over one thousand days, starting in the winter of 1991 with around four months of data for every subsequent winter.

Table 2 shows the effect of scaling and weighting the data for Lochaber in comparison with simply using the raw values. Results are presented in three ways: as the number of correct/ incorrect forecasts for avalanche days and non-avalanche days; as the mean number of nearest neighbours with avalanche events for each class (this facilitates comparison with Gassner *et al.*, 2000) and as an overall percentage (this final figure is the one used by the genetic algorithm in calculating fitness).

	Raw	Scaled	Scaled and weighted
Avalanche Days	123 right 113 wrong (2.73)	166 right 70 wrong (3.97)	181 right 55 wrong (4.07)
Non-avalanche days	526 right 143 wrong (1.43)	551 right 118 wrong (1.07)	576 right 93 wrong (1.00)
Overall fitness	65%	76%	81%

Table 2: Performance of figures for Cornice

The performance figures shown here are very similar to those presented for NXD2000, demonstrating that genetic algorithms permit one to reach very similar solutions to expert weighting. Further analysis has been undertaken for three recent winters with differing avalanche characteristics (Figure 4) to evaluate whether model performance is consistent across a number of winters. These analyses can be used to infer whether the model is becoming over tuned to a particular winter or set of meteorological and snowpack conditions.

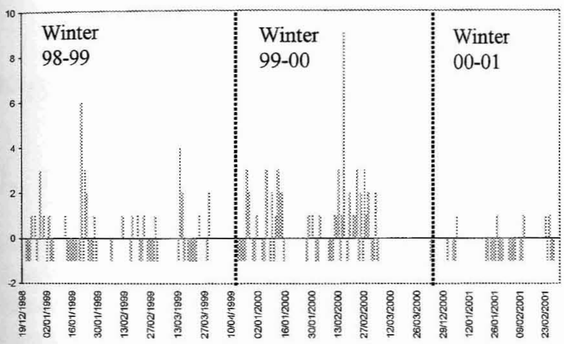

Figure 4: Avalanche activity for winters 1998-99, 1999-2000 and 2000-01 in Lochaber (Values of -1 indicate no visibility days*, avalanche days have number of avalanches as recorded in Cornice)

Table 3 shows the results of these winter to winter performance comparisons. It is clear that the model is performing in a similar fashion in winters with many avalanches (1999-2000) and very few avalanches (2000-2001).

	Winter 1998-99	Winter 1999-2000	Winter 2000-01
Avalanche Days	14 right 5 wrong (3.05)	20 right 6 wrong (3.92)	4 right 1 wrong (3.40)
Non-avalanche days	49 right 11 wrong (0.7)	44 right 7 wrong (0.67)	40 right 5 wrong (0.42)
Overall (%)	77%	81%	84%

Table 3: Winter to winter performance of figures for Cornice

* No-visibility days occur when drifting snow or other adverse conditions prevent forecasters from observing avalanche occurrence and are a very common feature in the dataset – no event is ascribed to such days.

4. Using the model in operational avalanche forecasting

It is clear that Cornice's performance is similar to that of NXD2000 and, although we would argue that the use of genetic algorithms allows objective weighting to be carried out in n-dimensional space, that differing approaches are converging towards very similar solutions. We would caution against tuning the model to individual avalanche days which may be omitted and argue that such an approach may instil unrealistic expectations of model performance.

Here, we seek to show how Cornice (or other models) can be best utilized in operational avalanche forecasting as a part of the conventional avalanche forecasting process. To illustrate this we have prepared a cartoon of the backcountry avalanche forecasting process, which shares elements of the analysis performed by McClung and Schaerer (1993) and LaChappelle (1980). We show a timeline through the forecaster's day, starting with gathering data to form initial hypotheses about the prevailing conditions and their evolution. These hypotheses shape the route taken by the forecaster in the field to make observations which may cause a re-evaluation of the initial hypotheses (and perhaps the gathering of more field data). Finally, these hypotheses are integrated with available forecast data for the following day - in the Scottish case mainly of meteorological parameters - to prepare the forecast.

Model runs appear at two points in the cartoon, in the initial planning phases and in the preparation of the forecast. They are given no more weight than any other element of the information gathering process and have interdependencies with the diverse information sources.

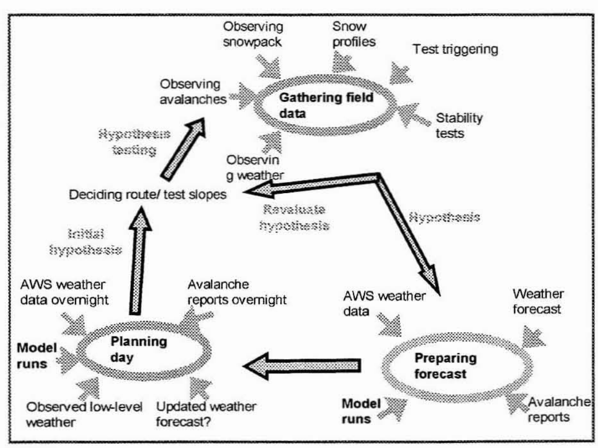

Figure 5: Information sources in conventional backcountry avalanche forecasting

The critical point here is that our model of the backcountry forecasting process is one primarily based on hypothesis testing. Thus, the role of the model is not to provide the observer with the avalanche hazard for the following day or to identify the probability of avalanches. Rather, it is another part of the information gathering and hypothesis testing process which allows the forecaster access to a large database of information on avalanche occurrence in similar conditions in the past. This information can then supplement the empirical experience referred to by LaChappelle (1980) and provide a further means for developing hypotheses about the prevailing conditions and resulting avalanche hazard.

6. Conclusions

In this paper we have introduced Cornice, a nearest neighbours model used in backcountry avalanche forecasting in Scotland. Cornice utilises optimisation techniques, namely genetic algorithms to carry out weighting. Model results over the data set are shown to be comparable with those from NXD2000 and we further illustrate consistent winter to winter performance. Elaborate variables are not used, but this does not appear to adversely affect the model's performance.

The development of Cornice was carried out in consultation with avalanche forecasters in Scotland, and we illustrate how different visualisations of the results may facilitate the model's use in hypothesis testing. Such approaches are likely to compliment the processes of conventional avalanche forecasting as described in LaChappelle's seminal paper (1980). We illustrate how Cornice fits into the information gathering process of the avalanche forecaster preparing an avalanche forecast as an additional tool, rather than as a replacement for the forecaster's knowledge and experience.

7. Acknowledgements

This work has benefited from many useful discussions with Othmar Buser and others at the Swiss Federal Institute for Snow and Avalanches. sportScotland funded experiments with Cornice. The forecasters of the SAIS have provided much useful feedback in the operational implementation of Cornice.

8. References

Anderson, C.M. 1994. *The Scottish Mountain Rescue Study 1964-1993*. HMSO, Edinburgh.

Barton, J. 1987. Weather observations on Cairn Gorm summit 1979-86. *Meteorol. Mag.*, 116, 346-353.

Beasley, D., Bull, D.R. and Martin, R.R. 1993. An overview of genetic algorithms: Part 1, Fundamentals. *University Computing*, 15(2) 58-69,

Beasley, D., Bull, D.R. and Martin, R.R. 1993.An overview of genetic algorithms: Part 2, Research Topics. *University Computing*, 15(4) 170-181, 1993

Bolognesi, R. 1994. Local avalanche forecasting in Switzerland: strategy and tools. In *Proceedings of the International Snow Science Workshop*, 463-472, Snowbird, Utah, USA.

Brabec, B., Meister, R., Stöckli, U., Stoffel, A. and Stucki, T. 2001. RAIFoS: Regional Avalanche Information and Forecasting System. *Cold Reg. Sci. Technol.*, 33, 303-311.

Buser, O. 1983. Avalanche forecast with the method of nearest neighbours: An interactive approach. *Cold Reg. Sci. Technol.*, 8, 155-163.

Buser, O. 1989. Two years experience of operational avalanche forecasting using the nearest neighbours method. *Annals of Glaciology*, 13, 31-34.

Gassner, M., Birkeland, K., Etter, H.J. and Leonard, T. 2000. NXD2000: An improved avalanche forecasting program based upon the nearest neighbour method. In *Proceedings of the ISSW*, 52-59, Big Sky, Montana, USA.

Kristensen, J. and Larson, C. 1994. An avalanche forecasting method based on a modified nearest neighbour method. In *Proceedings of the ISSW*, 22-30, Snowbird, Utah, USA.

LaChappelle, E. 1980. The fundamental processes of conventional avalanche forecasting. *J. Glac.*, 26(94), 75-84.

McClung, D.M. and Tweedy, J. 1994. Numerical avalanche prediction: Kootenay Pass, British Columbia. *J. Glac.*, 40(135), 350-358.

McClung, D. and Schearer, P. 1993. *The Avalanche Handbook*. Cordee.

Ray, E.T. 2001. *Learning XML*. O'Reilly, Cambridge.