Operational Decision Tree Avalanche Forecasting

Walter Rosenthal
Donald Bren School of Environmental Science and Management
University of California, Santa Barbara

Kelly Elder
Rocky Mountain Research Station
USDA Forest Service
Fort Collins, Colorado

Robert E. Davis
Cold Regions Research and Engineering Laboratory
US Army Engineer Research and Development Center
Hanover, New Hampshire

Abstract: Decision tree models of maximum avalanche size class run daily at Mammoth Mountain, California. A classification tree grown on an eight-year subset of all weather and avalanche records shows an absolute accuracy on avalanche control days of from about 60-70% in a given year; accepting overestimates increases this to 70-80%. Errors arise from the rarity of large events, exclusion of the smallest most frequent events, and tree sensitivity to small changes in key predictor variables. A complete 19-year data set yields a pair of decision trees forecasting both maximum size class and maximum crown size over the entire mountain. Tested against a twentieth year, the size class tree may be more accurate for extreme events but performed slightly worse overall than the original tree. Coupling the size class and crown trees identified both class 5 avalanches during the test year. A third set of trees, driven by hourly data from a remote instrument network, distributes maximum class and crown sizes over geographic sub-regions of the mountain. These are striking for both small size and low misclassification rate. If a major source of error is chaotic avalanche behavior, decision trees may prove most valuable for providing probability estimates from given sets of initial conditions.

Keywords: avalanche, decision trees, avalanche forecasting

1. Introduction

Decision tree models of avalanche activity can be valuable forecasting and training tools. They rank important relationships among avalanche variables and provide clear graphic display of the relative importance of variables on a given day. They provide probabilities of different results and can alert workers to unusual conditions.

Binary decision trees may be either classification or regression trees (Breiman *et al.*, 1984). Data are arrayed in a learning sample of the response variable (*e.g.*, size class) and predictor variables (new snow depth, 24-hour wind speed, etc.). They are split on the single predictor value that produces the two most homogenous response variable subsets. For example, more than 0.5m of new snow in 24 hours may produce mostly class 3-5 avalanches while those with less new snow may produce mostly class 1, 2 or no avalanches. Each of the two subsets is then reevaluated and recursively split until the final subsets shrink below a specified size or meet a criterion of homogeneity.

Davis *et al.* (1999) used decision trees to evaluate and rank storm, snow and weather factors influencing dry slab avalanches at Alta and Mammoth Mountain. That paper details decision tree methods, variables studied and rankings of the variables as predictors of maximum size class, sum of avalanche sizes, and whether a given day is an avalanche day. We use their variables in all three parts of this study.

Elder and Davis (2000) extended the previous work by developing classification trees for avalanche education and training at Alta and Mammoth Mountain. Those trees, too large to reproduce, graphically demonstrate logical and hierarchical relationships among the many variables affecting avalanche release.

**Corresponding author*: Walter Rosenthal, P.O. Box 7457, Mammoth Lakes, CA 93546; tel: 760-935-4464; email: walter@icess.ucsb.edu

We focus here on decision trees for avalanche forecasting, reporting on three studies at Mammoth Mountain. The first tests the tree from Elder and Davis (2000) against avalanche data from the last three seasons. The second develops classification trees for maximum size class and crown size from all data collected from the winters of 1983 through 2001, testing them on data from winter 2002. Finally, hourly data from a network of snow and weather sensors are used to forecast region-specific avalanche activity around the mountain. We describe methods and results in each section.

The combined studies address five questions:

1. Beyond education, are such models useful forecasting tools? Can they alert avalanche workers to unusual conditions that put them at risk?

2. Does increasing learning sample size improve the decision tree forecasts, particularly for larger, more dangerous events?

3. Size class is a subjective judgment. Can trees provide accurate quantitative estimates of physical features like maximum crown size?

4. Can spatially distributed hourly measurements drive region-specific forecasts of class and crown size?

5. Are there fundamental limits to accuracy that can be anticipated for these and other statistical models? How are such models best used?

2. A three-year retrospective study

We tested the Mammoth classification tree from Davis and Elder (2000) against all data from the winters 2000 through 2002. That tree, estimating maximum size class for a given day, was grown on an eight-year subset of all data acquired between 1989 and 1998.

The learning sample contained 890 cases, of which 160 were avalanche control days. Most of those had a maximum avalanche size of class 2 or 3, with only a handful of class 1 days, 24 class 4 days and seven class 5 days. Most class 1 days were excluded because of inconsistencies and uncertainty about record reliability. The largest size class estimated by the tree is class 4.

The tree ran daily on avalanche control mornings during the winters of 2001 and 2002. It has classified 132 avalanche control days over the last three years, as summarized in Table 1 and detailed in Figures 1-3.

Table 1. *Maximum class size tree accuracy for the last 3 seasons at Mammoth Mountain. N is the number of avalanche control days; hits are days the tree correctly forecast the maximum class size.*

Season	N	Hits	Hits + Overestimates
1999/00	46	27 (59%)	32 (70%)
2000/01	56	40 (71%)	45 (80%)
2001/02	31	15 (48%)	25 (81%)

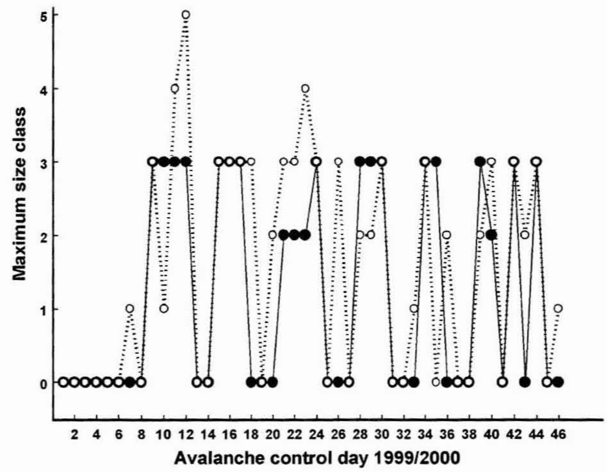

Figure 1. *Predicted and actual maximum avalanche size class, 1999/2000. (• = predicted size class; ○ = actual size class).*

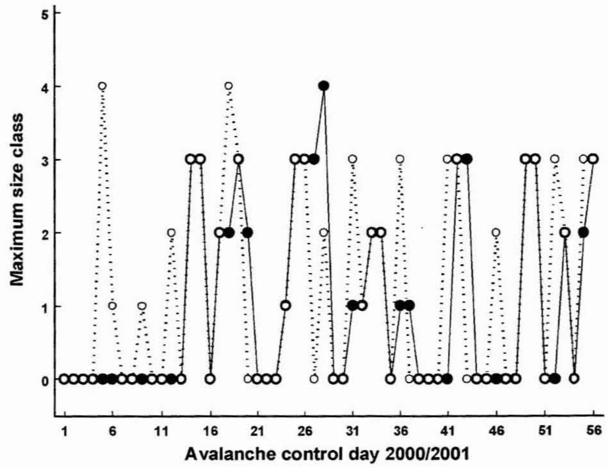

Figure 2. *Predicted and actual maximum avalanche size class, 2000/2001. (• = predicted size class; ○ = actual size class).*

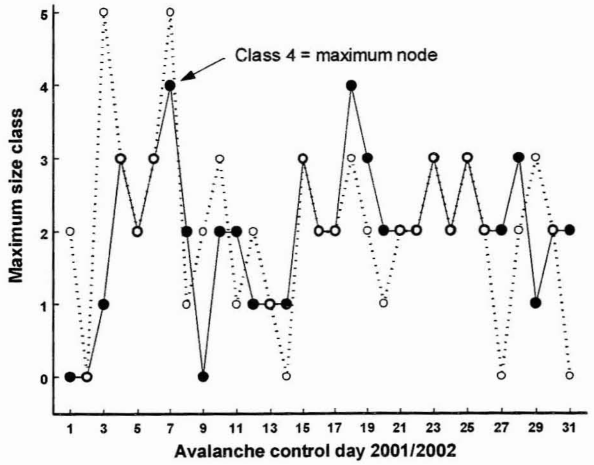

Fig. 3. *Predicted and actual maximum avalanche size class, 2001/2002. (● = predicted size class; ○ = actual size class).*

We accept overestimates because the trees should alert workers to maximum likely occurrences and the factors accounting for them. Class boundaries are fuzzy and overestimating anticipated avalanche size carries no penalty.

3. Class and crown size trees grown from a 19-year learning sample

Decision trees are insensitive to rare occurrences. These are sometimes isolated in small terminal nodes but often lumped with other events occurring under similar circumstances. The primary benefit anticipated from a larger learning sample is better classification of large avalanches.

Mammoth Mountain has recorded 742 avalanche control days during more than 3300 days of operation since 1982/83. We grew trees predicting maximum size class and maximum crown size from 696 control days during the first 19 years of the data set. 2001/2002 was withheld as a test year. Table 2 gives the size distribution for the largest avalanches on each control day in the sample.

Table 2. *Distribution of maximum avalanche size class for 696 control days at Mammoth Mountain, 1982/1983 through 2000/2001.*

Size:	None	1	2	3	4	5
Occurrences:	124	128	180	179	61	24

Data errors, identified graphically, were corrected from original records. Unlike Davis *et al.*, we included all class 1 events after correcting or excluding obvious mistakes. We did not shift avalanche data back into storm cycles to account for slides released by control work on clear days following storms (about 30% of avalanche control days would qualify). We used their variables plus several others indicative of clearing after a storm (*e.g.*, morning precipitation, morning cloud cover, previous control work done on the upper mountain). This approach followed from the exigencies of simple use during daily operations.

The maximum class size tree has 65 terminal nodes and a misclassification error rate of 0.35. The tree for maximum crown size has 76 terminal nodes and a misclassification error rate of 0.43. Predicted and actual maximum class and crown sizes for winter 2001/2002 are summarized in Table 3.

Table 3. *Maximum size class and crown size tree accuracies for winter 2001/2002 at Mammoth Mountain. N =31.*

Response	Hits	Hits + Overestimates
size class	13 (42%)	21 (68%)
crown size	4 (13%)	20 (65%)

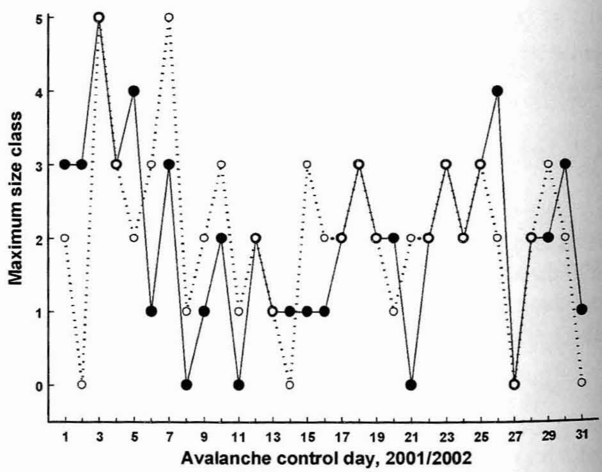

Figure 4. *Maximum size class during the 2001/2002 season, actual and predicted by the tree grown on a 19 year learning sample. (● = predicted size class; ○ = actual size class).*

The maximum class size tree (Figure 4) correctly predicts one of the two class 5 avalanches. However, it exhibits somewhat lower overall accuracy than the tree tested in Section 2. Davis *et al.* excluded most class 1 avalanche days while we excluded all non-control days, but *neither* tree did well with class 1 or 0 (i.e., none) last year. Each correctly identified only 2 out of 8 cases. Further, Table 2 shows our learning sample to be well balanced with respect to number of cases per avalanche class. The effect of shifting avalanche control days back into the storm cycles remains to be investigated.

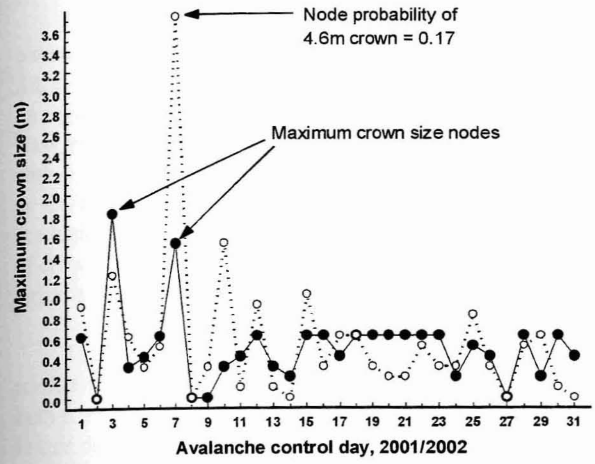

Figure 5. *Maximum crown size during the 2001/2002 season, actual and as predicted by the decision tree grown on a 19 year learning sample. (● = predicted size class; ○ = actual size class).*

Figure 5 shows predicted and actual maximum crown size for the 31 control days during 2001/2002. The fit between predicted and actual crown sizes is much better than suggested by few hits shown in Table 3. The relatively high misclassification rate is also misleading. The tree tracks crown size well throughout the season, with only two exceptions, which we consider first.

Day 10 is a clear miss, with probabilities of 0.05 and 0.24 for crowns of 0.9m and 0.6m respectively. Day 7, also underestimated by the size class tree (Figure 4), produced a class 5 avalanche with a 3.7m crown. However, day 7 falls into one of the two terminal nodes with the largest mean crown sizes. The tree predicts a 1.5m crown that day but estimates a 0.17 probability of a larger crown (4.6m, comparable to the actual one).

Figure 6 shows days 7 and 10 to be anomalies. Plotted are the errors between estimated and actual maximum crown size. The plotted mean (0.06m) and standard deviations (±0.29m) are for the set *without* those two days. Errors are then approximately normally distributed. Applying the t-test to those data, the 95% confidence interval for the mean is from -0.05 to 0.17m and the null hypothesis that the mean error equals 0 cannot be rejected.

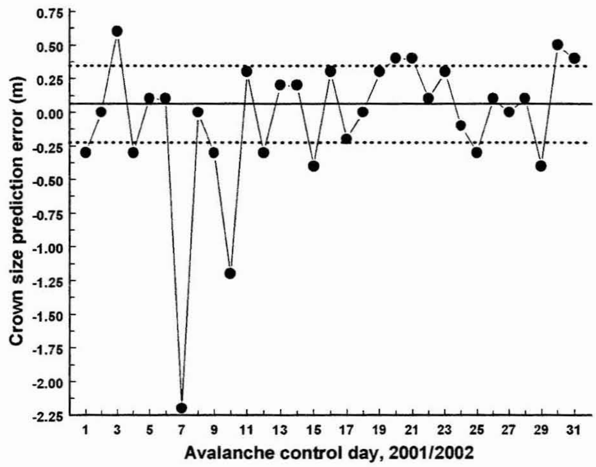

Figure 6. *Crown size errors. Mean (solid line) and standard deviation (dotted lines) are shown for the set without days 7 and 10.*

A strong case exists for joint use of crown size and size class trees. Class 5 avalanches occurred on two days. The size class tree predicted the first while the crown size tree issued a significant probability of the second. Had both trees been operational in late 2001 there would have been warning of class 5 avalanches on both days.

4. Regional trees driven by hourly remote instrument data

The trees described so far predict occurrences over the entire mountain, but different sets of starting zones receive different storm loading and respond differently to that loading. Since fall 1999, a network of remote instruments has logged hourly data from six sites around the mountain. Four sites (two on the summit ridge, two at mid-mountain) provide wind speed and direction, maximum gust, temperature and relative humidity near the major starting zones. A fifth site measures the same variables at the base of the mountain while the sixth logs snow depths, water equivalence, temperature and relative humidity.

Local learning samples for maximum size class and crown size on summit ridge and mid-mountain avalanche paths were prepared by averaging hourly

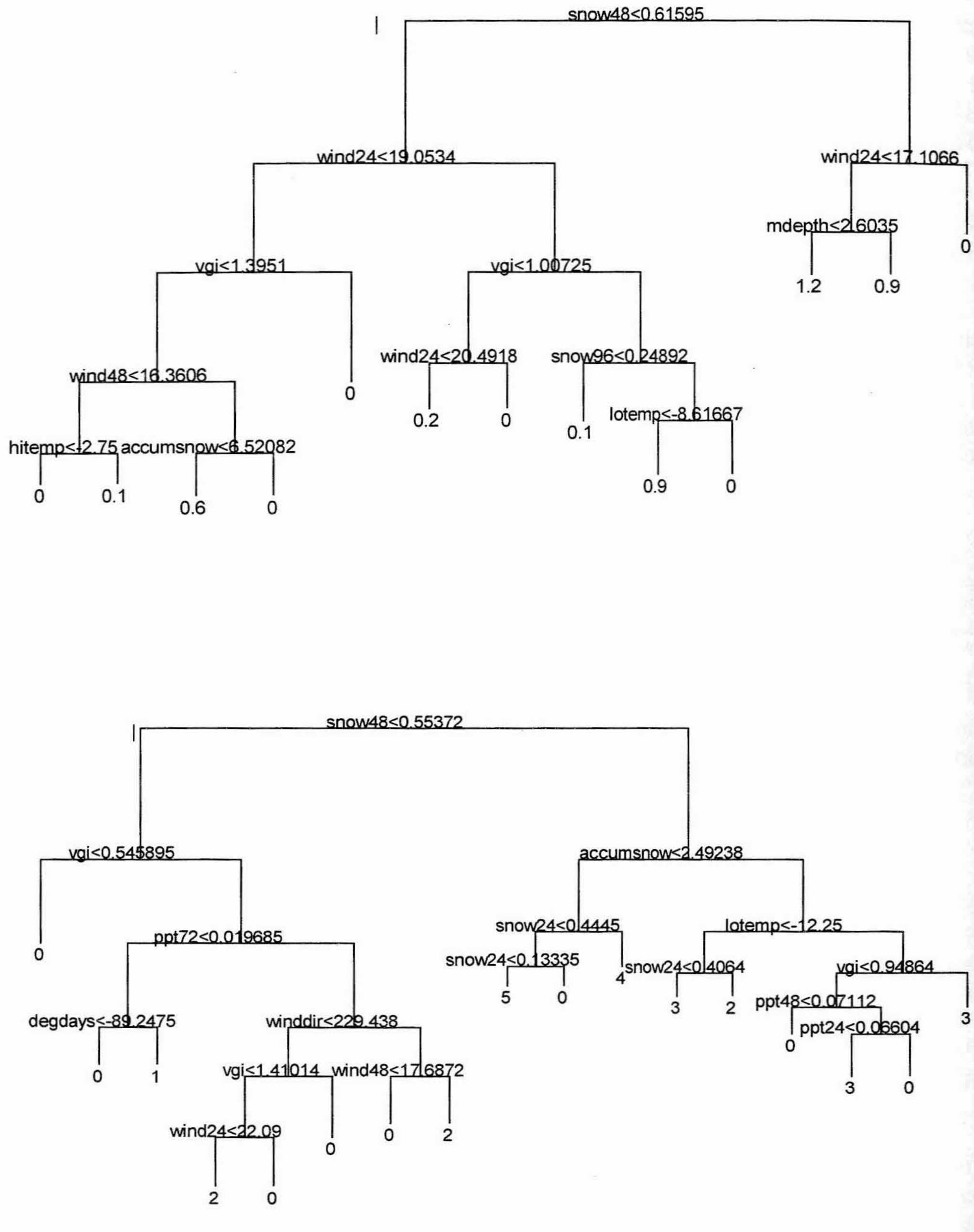

Figure 7. *Classification trees for maximum crown size (top) and size class (bottom) for 37 paths off the summit ridge of Mammoth Mountain. Trees are driven by local instrument data and snow plot data only. Splitting criteria are given for all nodes. Lesser values go to the left descendent node, greater values to the right. Step size is proportional to the reduction in node deviance at each split.* ***Key***: *snow48 = 48hr snowfall; mdepth = snow depth at master stake; wind24 = average 24hr wind speed; vgi = Vapor Gradient Index (Davis and Elder [2000]); ppt72 = 74hr water equivalence; accumsnow =season total new snow accumulation to date; degdays = degree days; lotemp/hitemp = 24hr low and high temperatures; winddir = wind direction (clockwise from north).*

local weather measurements into 24-, 48- and 72-hour bins for the last three years and coupling these with the corresponding precipitation totals. This produced local predictor variables analogous to those in the learning samples for the trees discussed earlier.

Trees were grown for both maximum class and crown size for the summit ridge and lower mountain avalanche paths. Trees for both regions were similar, though only the summit set is discussed here.

Figure 7 displays the full trees for maximum crown size and size class for 37 paths on the summit ridge. Most striking is their small size coupled with low misclassification error rates. The crown and size class trees have 13 and 17 terminal nodes respectively, with misclassification rates of 0.24 and 0.28. Allowing as few as two cases per terminal node and splits producing single-case nodes only increases the number of crown size terminal nodes by six and size class terminal nodes by ten.

The hierarchical splits in Figure 7 illustrate both predictive and explanatory aspects of decision trees. In both trees the first split producing the greatest reduction in tree deviance is on 48-hour snowfall. In both cases the critical value is about 0.6m of new snow, with higher values producing larger avalanches. Following the right main branch of the crown size tree, the next important factor is 24-hour average wind speed. Note that in both trees all critical values for 24-hour wind speed are about 20m/s. That sustained winds over 17m/s should produce no crown is at first counterintuitive, but such winds (depending on direction) either strip the mountain or produce hard slabs resistant to artillery, hand charges and ski cutting. Below the 17m/s threshold crowns are about a meter thick, but the larger ones tend to occur in thinner packs with less than 2.6m of snow at the master stake. Thinner snow packs are subject to larger temperature gradients, greater depth hoar development, and therefore larger releases.

Table 4 gives the distribution of maximum size classes for the summit ridge for the last three years. Note the relative under-representation of class 1 events when compared with Table 2.

Table 4. *Distribution of maximum avalanche size class for 37 paths on the summit ridge of Mammoth Mountain for 104 control days, 1999/2000 through 2001/2002.*

Size:	None	1	2	3	4	5
Occurrences:	57	5	22	17	2	1

Scarcity of small events is a likely cause of size class tree simplification. Similarly, crown sizes on the upper mountain are often much larger than those on the lower. The mean maximum crown size for paths off the summit ridge is 2.5 times that for mid-mountain. Average wind speeds, taken at the top and at mid-mountain, show similar ratios of about 2.5 across seasons, during storm periods and on control days. To the extent that regions are more homogeneous in their loading and avalanche characteristics than the mountain as a whole, tree size should shrink and accuracy should improve.

5. Discussion

Each split in a decision tree divides the data with a plane perpendicular to the axis of the splitting variable. A parallelepiped classifier emerges in which predictor variables in a given sub-volume of the data space produce (ideally) similar responses. To the extent that small differences in the initial conditions produce large differences in the response, the approach will fail.

Sensitive dependence on initial conditions is the classic definition of chaos. Chaotic behavior would account for the failure of the complete, much larger learning sample to produce decision trees of greater overall accuracy. It would also account for both the large tree size and relatively large error rates: adjacent and even identical points in the data space produce vastly different avalanche outcomes. For example, the largest crown recorded at Mammoth Mountain is 6.6m. It occurred on a day that produced only ten other avalanches, all class 1 and 2, with crowns from 0.05-0.3m. Inspecting the terminal nodes of the decision trees shows this is common.

Rosenthal and Elder (2002, this volume) present evidence that slab avalanching is a chaotic process. This accords with other findings of size distributions consistent with self-organized criticality (Birkeland and Landry, 2002).

If a major source of decision tree error is chaotic avalanche behavior, probability estimates of multiple avalanche measures may prove more important than the absolute accuracy of any one tree. Terminal nodes are often heterogeneous and their values are simply the means or modes (depending on tree type) of the responses within them. This suggests perturbing initial conditions to produce ensemble forecasts of measures like crown size, size class and path length.
New problems will arise such as how to employ such forecasts and how to resolve inevitable conflicts among them.

6. Conclusions

We return to the five questions posed in the introduction, drawing on the results of the three studies.

1. Decision tree models can be useful forecasting tools and are generally conservative in that they err on the side of overestimating the size of avalanches to which workers may be exposed.

2. Increasing the size of the learning sample does not improve overall tree accuracy but may improve estimates of extreme events. The use of multiple trees estimating different avalanche features like size class and class size improves estimates of extreme events.

3. Physical features such as maximum crown size can be well modeled and forecast with decision tree methods.

4. Hourly data from a remote instrument network can drive region-specific forecasts of class and crown size. The trees are small, with low misclassification rates compared with those for models applied to the entire mountain. This is partly because of the size of the learning sample and partly because of more homogeneous regional conditions.

5. Chaos may impose fundamental limits on the accuracy that can be anticipated from any statistical model. The most useful products of such models may be probability distributions of workers being exposed to large avalanches.

5. Acknowledgements

We thank the Mammoth Mountain Ski Patrol for weather and avalanche data and for their continuing support of this project. We also thank the Sierra Nevada Aquatic Research Laboratory of the University of California for their support. This work was partially funded by U.S. Army Project 4A762784AT42.

References

Birkeland, K.W. and C.C. Landry. *Power-laws and snow avalanches*. Geophysical Research Letters **29**(11), 10.1029/2001GL014623, p. 49-1, 2002.

Breiman, L., J.H. Friedman, R.A. Olshen, and C.J. Stone, *Classification and Regression Trees*, Wadsworth, Belmont, CA, 1984.

Davis, R.E., K. Elder, D. Howlett and E. Bouzaglou, *Relating storm and weather factors to dry slab activity at Alta, Utah and Mammoth Mountain, California, using classification and regression trees*. Cold Regions Science and Technology **30**, p. 79, 1999.

Davis, R.E. and K. Elder, *A between-storm indicator of avalanche activity*. ISSW 2000 Proceedings, p. 135.

Elder, K. and R. E. Davis, *Decision trees predicting avalanche response: Tools for training?*, ISSW 2000 Proceedings, p. 140.

Rosenthal, W and K. Elder, *Evidence of chaos in slab avalanching*. ISSW 2002 Proceedings (this volume).