CrossMark

Application of Hidden Markov Model for avalanche danger simulations for road sectors in North-West Himalaya

Jagdish Chandra Joshi[1] · Tankeshwar Kumar[2] · Sunita Srivastava[2] · Divya Sachdeva[1] · Ashwagosha Ganju[1]

Abstract Hidden Markov Model (HMM) has been developed for avalanche warning on 10 different road sectors in Pir-Panjal and Great Himalayan mountain ranges of North-West Himalaya. The model uses a data set of nine snow and meteorological variables—average air temperature, snow temperature index, snow drift index, snowfall in 24 h, snowfall in 48 h, snow water equivalent, snowfall intensity, standing snow and snowpack settlement collected during past 20 winters (1992–2012). The HMM is composed of four observations derived from the model input variables and four state variables. The state variables of the model are four levels of avalanche danger (No, Low, Medium and High). Single HMM has been developed to provide avalanche warning for both direct and delayed/wet avalanches with a lead time of two days. The HMM has been validated with (Case-1) and without (Case-2) incorporating delayed/wet avalanches using data collected during four winters (2012–2016) and compared with official Avalanche Warning Bulletin issued by Snow and Avalanche Study Establishment during these winters. The model has been validated through computation of accuracy measures such as percent correct (PC), bias, false alarm rate, probability of detection and Heidke Skill Score. The PC of the HMM for different stations for Case-1 varies from 80.1 to 98.6% for day-1 and 81.2 to 98.3% for day-2 and that for Case-2 from 82.2 to 98.6% for day-1 and 83.3 to 98.3% for day-2.

Keywords Avalanche warning · Hidden Markov Model · Snow temperature index

✉ Jagdish Chandra Joshi
joshjagdish@gmail.com

1   Snow and Avalanche Study Establishment, Sector 37-A, Chandigarh 160036, India

2   Department of Physics, Panjab University, Chandigarh 160014, India

🙲 Springer

1 Introduction

Avalanche forecasting over complex terrain of Himalaya has been a challenging task for snow and avalanche researchers in India. Ganju and Sharma (2000) discussed the complexities of avalanche forecasting over Himalaya. They classified Indian Himalaya into three climatic zones (Lower, Middle and Upper Himalayan climatic zones) based on precipitation regime, temperature and topographic elevation. Different regions in these climatic zones receive differential snowfall that result in a spatially variable snow pack making spatially variable pattern of avalanche occurrences. There are many road axes lying on these climatic zones, affected badly due to heavy snow fall and frequent avalanche activities during winter. As on today, avalanche forecasting for these road axes is done mainly with the help of numerical and conventional techniques. The numerical techniques used involve nearest neighbour and expert system approaches. The output of these numerical models in combination with avalanche forecaster's interpretation of snow pack stability is used to deliver operational avalanche forecast.

In the present study, the HMM has been developed for ten different road sectors in North-West Himalaya using combination of meteorological (Class III) and snowpack (Class II) data (McClung and Scherer 2006) of past 20 winters (1992–2012). For avalanche forecasting, Class I (snowpack stability tests) and Class II data (snowpack parameters) are more relevant than class III data (meteorological variables). In the present work, snow temperature index representing of snowpack layer hardness has been incorporated as class II data in the model input variables. As snowpack is highly spatially variable and climatic conditions of different regions vary, a separate HMM has been developed for each of the sectors. Unlike other operational models running at SASE, the HMM is independent of historical database for operational use of the model. Since most of the avalanche activities took place at the time of or immediate after snowfall, the HMM has mainly been developed for forecasting of direct action avalanches only. However, the model has been validated with and without incorporating delayed/wet avalanches. The model has also been compared with the official Avalanche Warning Bulletin (AWB) issued by SASE during winters from 2012 to 2016.

Avalanche forecasting models have been developed worldwide using various techniques such as nearest neighbour analysis, discriminant analysis, cluster analysis, classification and regression trees. Obled and Good (1980) tested different statistical methods to address the challenging problem of avalanche forecasting and compared their respective potential in operational forecast. They concluded that nearest neighbour and discriminant analysis appear more promising than cluster analysis but require further developments. Buser (1983, 1989) worked on the development of multivariate 'nearest neighbours' technique and used it for operational avalanche forecasting in the ski area of the Parsenn region for operational purpose. Following the suggestions of Obled and Good (1980) and Buser et al. (1987), McClung and Tweedy (1994) derived a numerical avalanche prediction scheme for avalanche forecasting on Kootenay Pass, British Columbia, Canada. They used parametric discriminant analysis (incorporating Bayesian statistics), cluster analysis and nearest neighbour analysis using the Mahalanobis distance as the distance metric. Boyne and Williams (1992) used classification and regression trees for avalanche classification for the region of Berthoud Pass, Colorado by taking number of avalanches per day as classification variable. Fromm (2009) used k-mean cluster analysis for grouping of weather conditions and discriminant analysis for classification of avalanche and non-avalanche days in the groups. Besides these attempts, many researchers worked on nearest neighbour method for short-term avalanche forecasting, e.g. Gassner et al. (2000), Brabec and Meister (2001),

Purves et al. (2003), McCollister et al. (2003) and Singh et al. (2005). Pozdnoukhov et al. (2011) used support vector machine for spatio-temporal forecasting of snow avalanches.

In recent years, researchers have worked on the prediction of wet avalanches. Helbig et al. (2015) predicted wet snow avalanche pattern over complex topography using combination of weather forecast data and terrain parameters. Bellaire et al. (2017) used physical-based snow cover model SNOWPACK and high-resolution NWP model for forecasting regional pattern of the onset of wet snow avalanches. In the beginning of winter season during snow storms mainly loose snow avalanches take place and depending on terrain it may convert to a gliding snow avalanche. Peitzsch et al. (2015) used classification tree analysis for distinguishing between glide and non-glide snow avalanches. In mid-winter due to accumulated layered snowpack, slab avalanche formation is a common feature. Marienthal et al. (2015) examined the usefulness of meteorological variables for predicting deep slab avalanche days. Vernay et al. (2015) introduced approach of ensemble avalanche forecasting to explicitly incorporate uncertainty to avalanche hazard forecasting.

Avalanche forecasting over Indian Himalaya was initiated by Agrawal and Ganju (1994) and Bhatnagar et al. (1994). The avalanche forecasting approach used by Agrawal and Ganju (1994) was based on the basic understanding of the physical processes affecting snow pack stability and use of avalanche forecaster's experience. Bhatnagar et al. (1994) used nearest neighbourhood criterion using snow and meteorological data. Naresh and Pant (1999) developed a knowledge-based system for forecasting snow avalanches of Chowkibal-Tangdhar Axis (J&K). Singh et al. (2005) coupled nearest neighbour technique and meso-scale model (MM5) output for prediction of avalanches 3 days in advance. Singh and Ganju (2008) developed a rule-based expert system for operational avalanche forecast over Himalaya. Joshi et al. (2010) used Avalanche Activity Index to categorize the levels of avalanche danger on different road axes over Himalayas. Joshi and Srivastava (2014) developed HMM for avalanche warning on Chowkibal-Tangdhar road axis (C-T road axis) in Indian Himalaya.

2 Study area and data

The study area falls in Lower and Middle Himalayan climatic zones of the Himalaya. The sectors in Lower Himalayan zone are characterized by warm temperatures, high precipitation, deep snowpack and short winter period of 3–4 months. The average height of mountains in this climatic zone varies from 2000 to 4000 m. The regions falling in Middle Himalayan zone are characterized by cold temperature, moderate precipitation, relatively shallow snow pack and longer winter period of 6 months. Winter climatology of different sectors and their geographic altitude of their representative meteorological station have been given in Table 1. Snow and avalanche characteristics of each of these regions are represented by meteorological data of a representative station of SASE in that region. The geographic location of different meteorological stations of SASE has been shown in Fig. 1.

The HMM has been developed using three types of data—snow and meteorological data, avalanche occurrence data and avalanche warning bulletins (AWB). The snow and meteorological data in each of these regions are recorded manually twice daily at 08:30 and 17:30 h at a representative meteorological station of Snow and Avalanche Study Establishment (SASE) since last more than four decades. Snow and meteorological data recorded during past 20 winters (1992–2012) at these meteorological stations have been used for development of the HMM. In the North-West Himalaya, winter persists from

Table 1 Climatology of different sectors in Indian Himalaya (1992–2016)

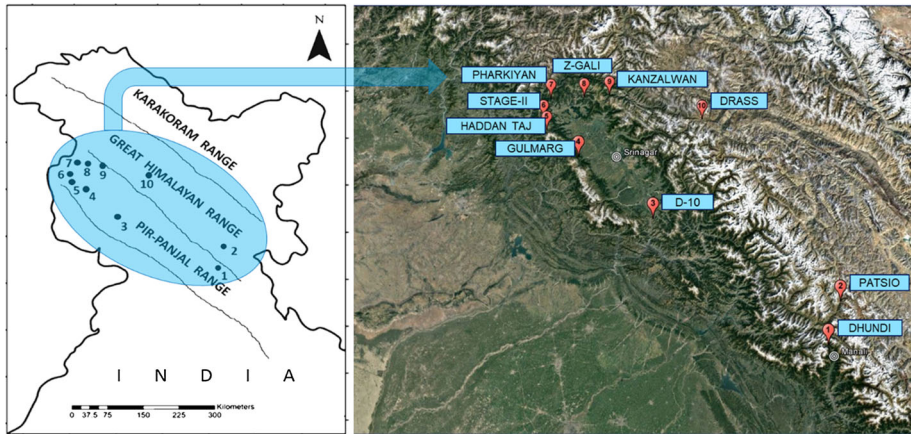| Sectors/ stations	Altitude (m)	Average lowest minimum temp (°C)	Average highest minimum temp (°C)	Average seasonal snowfall (cm)	Average lowest maximum temp (°C)	Average highest maximum temp (°C)	Average snowpack depth (cm)
Station-1	3050	− 10.8	7.2	1120	− 3.0	18.8	272
Station-2	3800	− 23.7	.1	445	− 8.5	12.7	184
Station-3	3200	− 1.6	7.7	613	− 6.5	17.9	235
Station-4	2800	− 15.4	4.6	702	− 3.4	17.8	181
Station-5	3080	− 17.2	5.6	924	− 5.8	14.7	261
Station-6	2650	− 12.8	7.7	950	− 3.2	19.4	229
Station-7	2960	− 13.2	6.1	953	− 5.2	16.3	260
Station-8	3192	− 16.0	3.0	956	− 6.8	13.0	253
Station-9	2440	− 18.3	− 4.0	922	− 3.0	19.6	196
Station-10	3250	− 28.3	2.7	354	− 9.8	16.3	130

Fig. 1 Meteorological stations of SASE in Pir-Panjal and Great Himalayan mountain ranges of North-West Himalaya (Google Earth Image)

November to April. However, the data collected during December, January, February and March have been used in the model because of less number of records of avalanche activities in November and April. The variables used in the development of model include average air temperature, Snow Temperature Index (STI), Snow Drift Index (SDI), snowfall

in 24 h, snowfall in 48 h, snow water equivalent, snowfall intensity, snow settlement and standing snow. The derived variable-STI attempts to summarize a physical effect that occurs in a snowpack over a period of time due to change in temperature (Kozak et al. 2002). The STI has been computed using degree-day method. For a day when maximum temperature exceeds a base temperature (-10 °C), it is summed for each day within the temperature index period. The base temperature of -10 °C chosen for the temperature index is based on the finding that sintering increases rapidly above -10 °C (Gubler 1982). The STI of a layer exposed for 'D' days is given by

$$\text{STI} = \sum\nolimits^{D}(T_{max} + 10\,^{\circ}\text{C}), \quad \text{when } T_{max} > -10$$

The model input variable-snow drift index (SDI) has been defined as the ratio of the friction velocity and threshold shear velocity of snow near snowpack surface. The friction velocity and threshold shear velocity have been defined by Campbell (1977) and Kind (1981), respectively. The SDI has been computed using the following relation:

$$\text{SDI} = u^*/u\text{th} = 0.226 \times u$$

where u^*, uth and u represent friction velocity, threshold shear stress and average wind speed respectively. When u^* exceeds uth, snow drift is initiated.

In addition to snow and meteorological data, avalanche occurrence and avalanche warning bulletins (AWB) of the same duration have also been taken in the analysis. The AWB data consist of the record of avalanche warning bulletins issued by SASE during winter for different regions of the Himalaya in the form of different levels of avalanche danger (No, Low, Medium and High). The qualitative interpretation of these danger levels has been given in Table 2.

Table 2 Scales of avalanche danger and their interpretation (SASE Publication)

Scales of avalanche danger	Interpretation
No	Generally safe conditions. Avalanches may not trigger
Low	Generally favourable condition. Triggering is generally possible only with high additional loads and on very few extreme slopes. Only sloughs possible and reach valley in small sizes. Valley movements are safe. Movements can be done on slopes with care (All safety measures shall be taken while crossing suspected avalanche paths)
Medium	Partly unfavourable condition. Triggering possible from the most avalanche prone slopes with low additional loads and may reach the valley in medium size. Avoid steep slopes. Routes should be selected with care. Valley movements with caution. Movements on slopes with extreme care (Rescue party shall stand by)
High	Unfavourable condition. Triggering possible from all avalanche prone slopes even with low additional loads and reach valley in large size. Suspend all movements. Airborne avalanche likely

3 Methodology

The HMM is an extension of Markov Chain model. The Markov chain model makes use of the property of Markov chain. According to the property of Markov chain, the probability of transition from one state to the next state depends only on the preceding state. The first-order Markov property (Rabiner 1989) is defined as follows:

$$P(q_t = j | q_{t-1} = i, q_{t-2} = k \ldots) = P(q_t = j | q_{t-1} = i)$$

where t denotes the time instance of a state, q_t is the state corresponding to time t, P is the probability of state transition and i, j, k are the state indexes.

In the HMM, there are two embedded stochastic processes in which the hidden process can only be observed through another stochastic process that produces the sequence of observations. The HMM is characterized by the number of observations, states and the model parameters. The model parameters include initial state probabilities, state transition probabilities and probabilities of observations in different states.

The flow chart of the methodology has been shown in Fig. 2. The HMM has been developed using four observations and four states variables. To define observations of the HMM, the meteorological variables have been categorized into different ranges and calculated index of avalanche (IA) for each range of all the variables. The IA of a range of a variable is an indicator of the probability of avalanche in that range. It is defined as the ratio of number of avalanche days in a range of a variable and total number of days in that range. The weather variables have been correlated with corresponding IA and the square of the correlation used as the weight of the variable. The weighted sum of the IA's of all the model input variables has been categorized into four categories representing four observations of the model. The categorization of observations is based on the database of past avalanche warnings. Four levels of avalanche danger—No, Low, Medium and High represent four states of the model.

The HMM parameters (Initial state probability, state transition probability and probability of observations in different states) have been computed using data of past 20 winters from 1992 to 2012 after the states and observations of the model have been defined. The probability of transition from one state to the other has been computed as the ratio of number of transitions from one state to the other and the total number of transitions taking
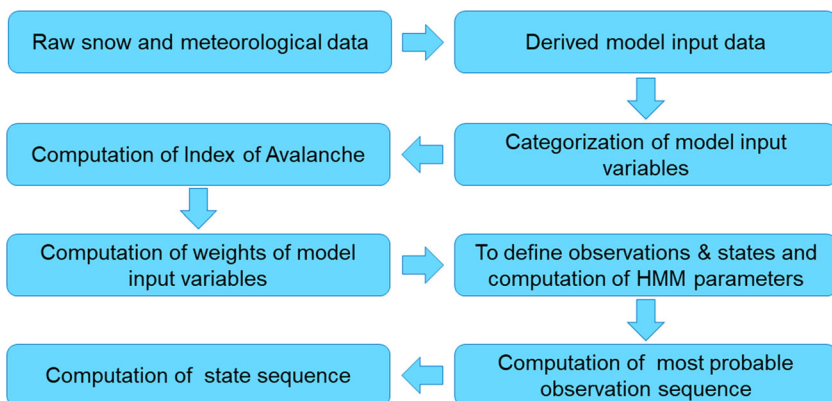

Fig. 2 Flow chart of HMM methodology for forecasting of avalanche danger sequence

place from that state. The initial state probability of a state has been defined as the ratio of the number of that states and total number of states. Similarly, the probability of an observation for a given state is defined as the ratio of the number of observations in that state and the total number of observations in that state. In the present study, a sequence of two observations has been used to predict the corresponding sequence of two states. The most probable observation and state sequences have been computed using Forward algorithm and Viterbi algorithms (Viterbi 1967), respectively.

4 Forward algorithm

Computation of most probable observation sequence is based on the computation of forward or backward variables. The direct computation of these variables is infeasible because of exponential growth of computation as a function of sequence length. For an observation sequence of length T and number of states N, it needs $(2T - 1) N^T$ multiplications and $N^T - 1$ additions. The forward algorithm on the other hand minimizes the computation cost to linear relative to T. It needs $N (N + 1) (T - 1) + N$ multiplications and $N (N - 1)$ $(T - 1)$ additions. In this algorithm, a forward variable is defined that accounts for the probability of partial observation sequence, $O_1, O_2, \ldots, O_t$ (up to time t), given state S_i at time t and model λ. The forward variable $\alpha_t (i)$ defined as:

$$\alpha_t(i) = P(O_1 O_2 \cdots O_t, \ q_t = S_i | \lambda)$$

The forward variable can be obtained inductively using initial state probability, π_i and probability of observation in a state, $b_i (O_1)$ as follows:

Step 1 Initialization

$$\alpha_i = \pi_i \times b_i(O_i), \ \ 1 \leq i \leq N$$

Step 2 Induction

$$\alpha_{t+1}(j) = \left[\sum_{i=1}^{N} \alpha_t(i) \times a_{ij} \right] \times b_j(O_{t+1}), \ \ 1 \leq t \leq T - 1, \ \ 1 \leq j \leq N$$

Step 3 Termination

$$P(O|\lambda) = \sum_{i=1}^{N} \alpha_T(i)$$

The initialization step represents forward probabilities as the joint probability of state S_i and initial observation O_1. The induction step as illustrated in Fig. 3 shows how state S_j can be reached at time $t + 1$ from N possible states, $S_i, 1 \leq i \leq N$, at time t. The product α_t $(i) \times a_{ij}$ in the induction step represents the probability of the joint event that $O_1, O_2, \ldots O_t$ are observed, and state S_j is reached at time $t + 1$ via state S_i at time t. This product when summed over all the states, S_i, at time t results in the probability of S_j at time $t + 1$ with all the accompanying previous observations. The termination step computes probability of

Fig. 3 Illustration of the computation of forward variable α_{t+1} (j)
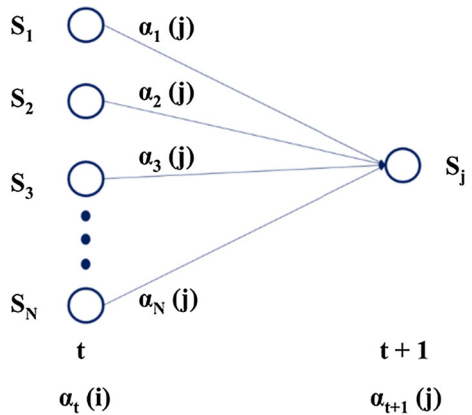
most probable observation sequence, $P(O|\lambda)$, as the sum of the terminal forward variable α_T (i).

5 Viterbi algorithm

The Viterbi algorithm (Viterbi 1967) is named after Andrew Viterbi, who proposed it in 1967 as a decoding algorithm for convolutional codes over noisy digital communication links. The Viterbi algorithm is a dynamic programming algorithm for finding the most likely sequence of hidden states called the Viterbi path that results in a sequence of observed events. To find the single best state sequence $S = \{S_1, S_2,... S_N\}$ for given observation sequence $O = \{O_1, O_2, ...O_T\}$, a probability function $\delta_t(i)$ is defined that gives highest probability along a single path, at time t, which accounts for the first t observations that end in state S_i.

$$\delta_t(i) = \max_{S_1 S_2...S_{t-1}} P(S_1 S_2...S_{t-1}, \quad S_t = i, O_1 O_2...O_t | \lambda)$$

Using induction $\delta_{t+1}(i)$ can be found as:

$$\delta_{t+1}(j) = \max_{1 \le i \le N} [\delta_t(i) \times a_{ij}] \times b_j(O_{t+1})$$

To retrieve the state sequence, it is necessary to keep track of the argument that maximizes $\delta_{t+1}(j)$, for each t and j. This is done by saving the argument in an array $\psi_t(j)$. The complete procedure to find out the best state sequence can be stated as follows:

Step 1 Initialization

$$\delta_1(i) = \pi_i \times b_i(O_1), \quad 1 \le i \le N$$
$$\Psi_1(i) = 0$$

Step 2 Induction

$$\delta_{t+1}(j) = \max_{1 \le i \le N}[\delta_{t-1}(i) \times a_{ij}] \times b_j(O_{t+1}), \quad 2 \le t \le T, \ 1 \le j \le N$$

$$\Psi_1(i) = \arg\max_{1 \le i \le N}[\delta_{t-1}(i) \times a_{ij}], \quad 2 \le t \le T, \ 1 \le j \le N$$

Step 3 Termination

$$P* = \max_{1 \le i \le N}[\delta_T(i)]$$

$$S_T* = \arg\max_{1 \le i \le N}[\delta_T(i)]$$

Step 4 State sequence backtracking

$$S_t* = \max_{1 \le i \le N}[\Psi_{t+1}(j) \times S_{t+1}], \quad t = T-1, \ T-2, \dots 1$$

The Viterbi algorithm is similar to the Forward algorithm. The main difference is that the Forward algorithm uses sum over previous states, whereas the Viterbi algorithm uses maximization.

In the present work, forecasting of avalanche danger is based on a sequence of two observations and corresponding states for forecasting with a lead time of 2 days. The complexity involved in the forecasting of observation sequence is of the order of N^2-T where N represents number of observations and T number of states. The computation involves $N (N + 1) (T - 1) + N$ multiplications and $N (N - 1) (T - 1)$ additions (Rabiner 1989). Thus, for increasing the lead time of forecast, length of the sequence of observations and states will be increased and consequently more computation will be involved.

6 Results and discussion

In Pir-Panjal and Great Himalayan mountain ranges, winter persists mainly for 6 months (November–April). During November and December months, snowcover just starts building up and due to shallow snowpack a very small number of avalanche activities take place. In January and February months, most of the direct action avalanches have been reported due to frequent snowfall and good snowcover build-up. In the months of March and April because of infrequent snowfall, rising temperature and snowpack settlement, direct as well as delayed/wet avalanches have been reported. The number of delayed/wet avalanches in the complete data set of all the sectors is limited. Therefore, the HMM has mainly been developed for direct action avalanches. However, the performance of the model during winters 2012–2014 has been discussed with and without inclusion of the delayed/wet avalanches.

The HMM has been developed for forecasting of avalanche danger on 10 different road sectors of Indian Himalaya with a lead time of 2 days. The weather variables used for the development of HMM have been categorized into different ranges and calculated index of avalanche for each range of all the variables. The plots of IA for all the model input variables for one representative station in both Pir-Panjal and Great Himalayan mountain ranges have been shown in Figs. 4 and 5, respectively.
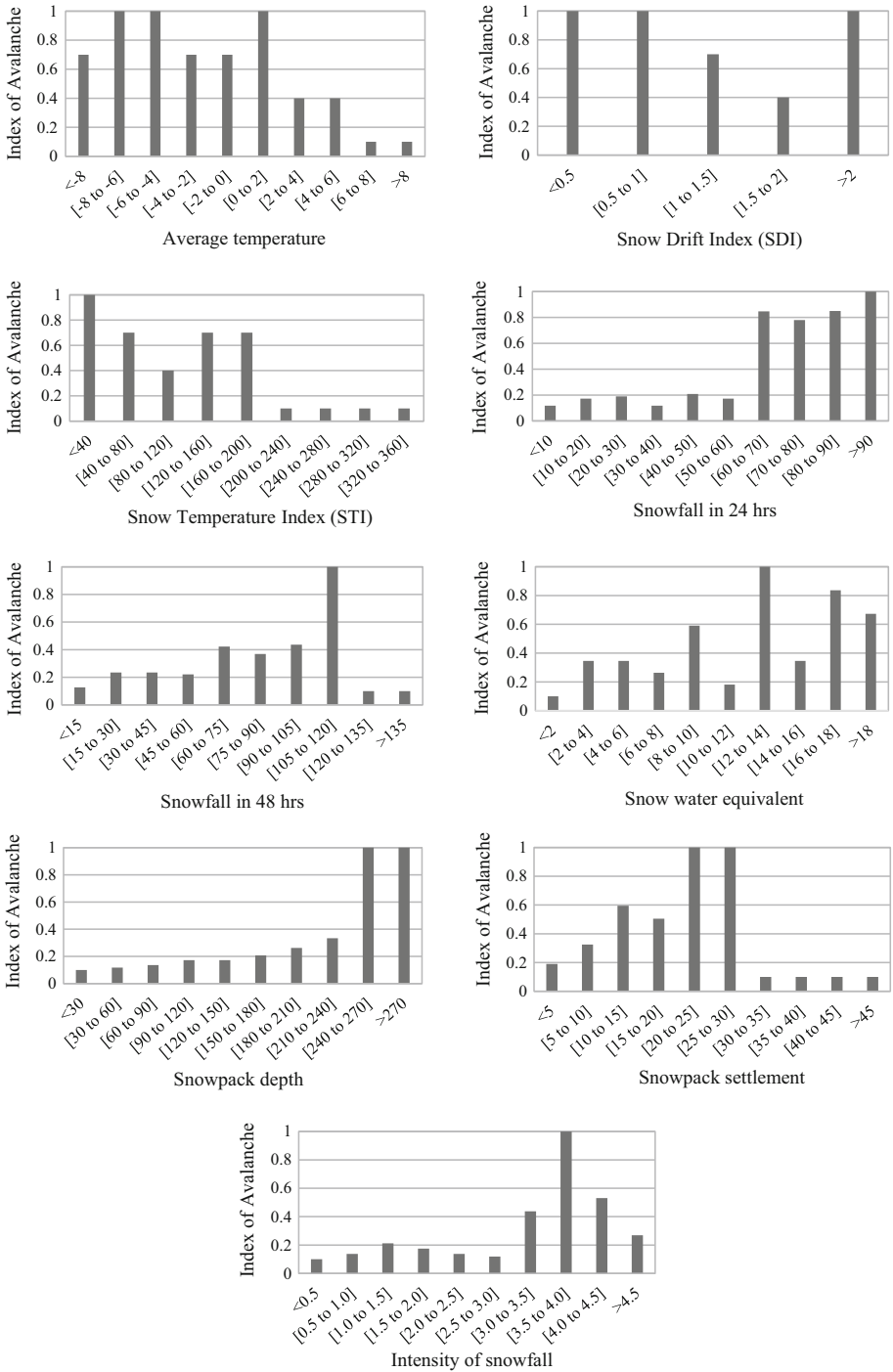
Fig. 4 Index of Avalanche for different categories of model input variables for station-2 in Great Himalayan range of Indian Himalaya
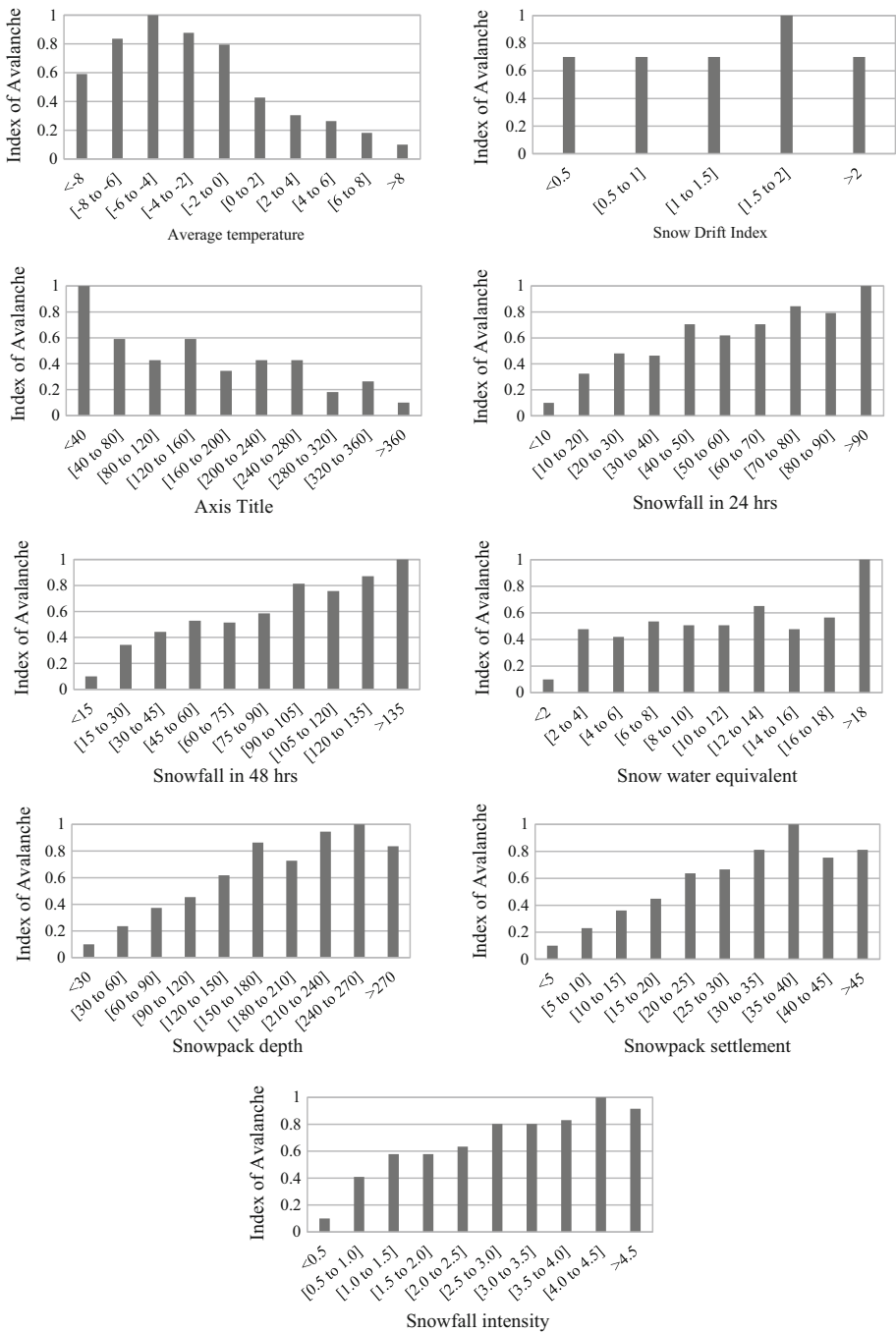
Fig. 5 Index of Avalanche for different categories of model input variables for station-6 in Pir-Panjal range of Indian Himalaya

The IA in both Pir-Panjal and Great Himalayan mountain ranges has been found higher for lower average air temperature. Since most of the avalanches in Indian Himalaya take place during or immediate after snowfall when temperatures remain sub-zero. Hence, the IA has been found higher for average temperature near zero or sub-zero. Moreover, for air temperatures below − 10 °C, the metamorphic process called sintering is slow and above this temperature it increases rapidly (Gubler 1982). Therefore, lower air temperatures lead to unstable snow pack and may result in an avalanche. For average air temperature above 0 °C, snowpack depletion takes place and relatively less avalanche activities take place resulting in smaller Index of Avalanche at higher air temperature.

Wind drifted snow plays an important role in avalanche formation. Snow drift is dependent on wind speed. Higher wind speed accelerates the process of snow drift. In Pir-Panjal mountain range for Station-6, avalanches triggered equally for smaller values of wind speed. For higher wind speeds, the IA first increases and then decreases. The avalanche activities during smaller wind speeds are mainly attributed to the overloading of snowpack due to snowfall. At higher wind speeds, snow is eroded from wind ward slopes and deposited on lee ward slopes, thus overloading the leeward slopes and triggering avalanches. For further higher wind speeds, either due to compaction of snowpack or due to already triggered avalanches, the IA has been found relatively smaller. In Great Himalayan mountain range for Station-2, the IA has been found high for lower as well as higher ranges of snow drift index and relatively low for intermediate ranges. For Station-2, high value of the IA for lower ranges of SDI is attributed both to the overloading of snowpack due to snowfall and drift deposition. The decrease in the value of IA for intermediate ranges of the SDI can be attributed to already triggered avalanches in previous snowfall. For higher ranges of the SDI, the slopes are re-loaded with snowfall and snow drift, thus triggering avalanches.

The STI represents duration of exposure of snowpack layer to solar radiation. It is an indicator of snow layer hardness. The IA has been found high for smaller STI and low for larger STI both in Pir-Panjal and Great Himalayan ranges. Smaller STI represents less exposure of the snow layer to solar radiation. Because of less exposer of the snow layer to solar radiation, the metamorphic changes are small and snowpack could not be stabilized and hence avalanches triggered. Larger exposure of the snow layers to solar radiation leads to larger number of melt-freeze cycles, thus hardening the snow layer and gaining strength to reduce avalanche activities.

Fresh snow fall can trigger avalanches in two ways. (1) It exerts stress on the snowpack layers beneath that may fail causing slab avalanches. (2) Fresh snow grains are not bonded well resulting in failure of fresh snow layer on a slope under the influence of gravity causing loose snow avalanches. All the variables related to fresh snowfall in both the mountain ranges (snowfall in 24 h, snowfall in 48 h and snow water equivalent) indicate that the IA increases with increase in the amount of these variables. For station-2 in Great Himalayan range, the smaller value of IA for higher range of snowfall in 48 h is attributed to the avalanches triggered during the previous snowfall and slopes are not loaded enough to re-initiate the avalanches.

The plot of IA and standing snow for both station-2 and 6 indicates that the IA increases with increase in snow depth. The standing snow represents stress generated by the snowpack on underlying layers. On a mountain slope, a threshold stress is required to trigger avalanche. As the standing snow increases, the stress on different layers increases leading to the failure of weakest snow layer and triggering avalanche. On some of the slopes after triggering of avalanches, further avalanches can occur only after sufficient accumulation of snow on that slope. Due to this, the IA decreased even for higher values of

standing snow on the slopes. Snow settlement is directly related to the amount of fresh snowfall and snowpack depth. Hence, the IA for snow settlement follows the similar pattern as that of snowfall and snowpack depth, i.e. the avalanche activities increase with increase in snow settlement.

The plot of IA and snowfall intensity indicates that the avalanche activities increase with snowfall intensity in both the mountain ranges. The rate of snowfall contributes towards strengthening or weakening of the snowpack. When the intensity of snowfall is low, snow pack stabilizes well, leading to stable snowpack and less avalanche activities. As the intensity increases, the snow pack does not get sufficient time in comparison to loading to stabilize and result in unstable snowpack and triggering of avalanches.

Forecasting of state sequence, i.e. avalanche danger depends on the most probable observation sequence corresponding to the initial observation. The most probable state sequence, i.e. day-1 and day-2 forecast for different stations in Pir-Panjal and Great Himalayan mountain ranges for different months during winter have been summarized in Table 3. The HMM forecast for day-1 and day-2 has been validated for four winters from 2012–2016 for all 10 road sectors in Indian Himalaya. Since avalanche database is dominated by non-occurrence cases of avalanche, the performance of avalanche fore-casting model cannot solely be assessed through computation of percent correct only. Hence, other accuracy measures such as bias, probability of detection, false alarm rate and Heidke skill score have also been analysed to assess performance of the HMM. The model has been validated separately for the cases with (Case-1) and without (Case-2) delayed/wet avalanches. The HMM forecast has also been compared with the official avalanche warning bulletin issued by SASE during the same winters (2012-16). The validation results for both the cases and AWB have been summarized in Table 4.

The AWB is issued by SASE with a lead time of 1 day only after analysis of different avalanche forecasting models running at SASE and snowpack stability assessment based on field data. The percent correct (PC) or overall accuracy of an avalanche forecasting model is a measure of correctness of the forecast for both occurrence and non-occurrence cases of avalanches. The PC of the HMM for different stations for Case-1 varies from 80.1 to 98.6% for day-1 and 81.2 to 98.3% for day-2 and that for Case-2 from 82.2 to 98.6% for day-1 and 83.3 to 98.3% for day-2. In the case of AWB, it varies from 73 to 89% for day-1. The PC of the HMM for both the cases for all the sectors has been found better than that of AWB. The HMM has shown better accuracy for Case-2 (excluding delayed/wet ava-lanches) than that for Case-1 (including delayed/wet avalanches) for all the road sectors in North-West Himalaya. Though the accuracy of HMM for both the cases has been found better than the AWB, the probability of detection (POD) of AWB has been found better than the HMM for all the stations. However, higher value of POD of the AWB is asso-ciated with highly biased forecast in favour of avalanche occurrence and higher false alarms. The POD tells about the capability of the model for correct prediction of avalanche occurrence cases. In avalanche forecasting, cost of type-II error (No avalanche danger warning when an avalanche actually occurred) is extremely high (life of human beings involved) as compared to type-I error (avalanche danger warning issued when actually avalanche did not occur) and, therefore, it is preferred to have type-II error as small as possible (Joshi and Srivastava 2014). The smaller value of type-II error of the model corresponds to over forecasting, i.e. higher bias. Therefore, higher bias with high POD and HSS can be considered as good forecast.

The Heidke Skill Score (HSS) greater than zero shows superiority of the model forecast over random forecast. The HSS of HMM has been found better than that of AWB for 5 stations where good avalanche feedback has been reported. For rest 5 stations where less

Table 3 State sequences corresponding to different initial observations in different months during winter in different sectors of North-West Himalaya

Stations	Months	Observation-1 (O_1)		Observation-2 (O_2)		Observation-3 (O_3)		Observation-4 (O_4)	
		Day-1	Day-2	Day-1	Day-2	Day-1	Day-2	Day-1	Day-2
Station-1	December	No	No	Low	No	Medium	Low	High	Medium
	Jan and Feb	No	No	Low	No	Medium	Medium	High	Medium
	March	No	No	Low	Low	Medium	Low	High	Medium
Station-2	December	No	No	Low	No	Medium	Low	Medium	Medium
	Jan and Feb	No	No	Low	No	Medium	Medium	High	Medium
	March	No	No	Low	No	Medium	Medium	Medium	Medium
Station-3	December	No	No	Low	No	Medium	Low	High	Medium
	Jan and Feb	No	No	Low	No	Medium	Low	High	High
	March	No	No	Low	Low	Medium	Low	High	Medium
Station-4	December	No	No	Low	No	Medium	Low	Medium	Low
	Jan and Feb	No	No	Low	No	Medium	Low	High	Medium
	March	No	No	No	No	Medium	Low	Medium	Medium
Station-5	December	No	No	Low	No	Medium	Low	Medium	Low
	Jan and Feb	No	No	Low	No	Medium	Low	High	Medium
	March	No	No	No	No	Medium	Low	Medium	Medium
Station-6	December	No	No	Low	No	Medium	Medium	Medium	Medium
	Jan and Feb	No	No	Low	Low	Medium	Low	High	Medium
	March	No	No	No	No	Medium	Low	High	Medium
Station-7	December	No	No	Low	Low	Low	Low	Medium	Medium
	Jan and Feb	No	No	Low	No	Medium	Low	High	High
	March	No	No	No	No	Medium	Medium	High	Medium
Station-8	December	No	No	Low	No	Medium	Low	Medium	Low
	Jan and Feb	No	No	Low	No	Medium	Low	High	Medium
	March	No	No	Low	Low	Medium	No	Medium	Medium
Station-9	December	No	No	No	No	Medium	No	High	Medium
	Jan and Feb	No	No	Low	No	Medium	Low	High	High
	March	No	No	No	No	Medium	Low	High	Medium
Station-10	December	No	No	No	No	Medium	No	High	Medium
	Jan and Feb	No	No	Low	No	Medium	Low	High	High
	March	No	No	No	No	Medium	Low	High	Medium

avalanche feedback has been received, the HSS and POD of AWB have been found better than the HMM due to over forecasting and associated higher bias. For some of the stations (Station-3 and 4) for day-2, the HSS of the model has shown negative values with high false alarm (1.0) implying that the model lacks forecasting skill. However, the higher bias of the HMM and AWB for these regions can be attributed to limited or delayed avalanche feedback from these regions. The bias occurred due to limited or delayed avalanche feedback can be reduced and the POD can be increased by employing suitable avalanche

Table 4 Forecast validation scores of HMM and official Avalanche Warning Bulletin (AWB) of SASE for different sectors in North-West Himalaya with (case-1) and without (case-2) delayed avalanches

	Percent Correct (PC)		Heidke Skill Score (HSS)		Probability of Detection (POD)		Bias		False Alarm Rate (FAR)	
	Day-1	Day-2	Day-1	Day-2	Day-1	Day-2	Day-1	Day-2	Day-1	Day-2
Station-1										
Case-1	97.5	98.1	.17	.35	.14	.29	.57	.57	.75	.5
Case-2	97.8	98.3	.19	.4	.18	.33	.67	.67	.75	.5
AWB	84.4	–	.133	–	.785	–	8.64	–	.909	–
Station-2										
Case-1	92	90.9	.26	.15	.42	.29	1.92	1.92	.77	.85
Case-2	92.4	91.3	.27	.16	.48	.32	2.16	2.16	.77	.85
AWB	86.9	–	.261	–	.75	–	3.9	–	.807	–
Station-3										
Case-1	98.1	97.8	.115	− .11	.142	0	1.3	1.3	.89	1
Case-2	98.2	98.1	.125	− .01	.2	0	2	1.8	.9	1
AWB	89	–	.095	–	.714	–	11.85	–	.94	–
Station-4										
Case-1	98.6	98.3	.16	− .08	.167	0	1	1	.833	1
Case-2	98.6	98.3	.16	− .08	.167	0	1	1	.833	1
AWB	75	–	.037	–	.833	–	31	–	.97	–
Station-5										
Case-1	86.3	85	.36	.296	.54	.47	1.5	1.5	.63	.685
Case-2	87	85.7	.375	.315	.58	.51	1.65	1.64	.644	.685
AWB	85.8	–	.43	–	.74	–	1.95	–	.62	–
Station-6										
Case-1	83.9	83.9	.256	.256	.547	.547	2.3	2.3	.76	.76
Case-2	84.8	84.7	.28	.27	.64	.62	2.7	2.7	.76	.76
AWB	81	–	.318	–	.85	–	3.3	–	.742	–
Station-7										
Case-1	84.6	83.7	.205	.163	.32	.28	1.22	1.22	.73	.77
Case-2	86.8	86	.263	.215	.44	.384	1.67	1.67	.73	.77
AWB	81.5	–	.357	–	.76	–	2.41	–	.684	–
Station-8										
Case-1	86.8	87.9	.22	.286	.29	.34	.97	.97	.701	.641
Case-2	87.5	88.6	.24	.307	.31	.38	1.06	1.06	.701	.641
AWB	78.9	–	.342	–	.855	–	2.92	–	.708	–
Station-9										
Case-1	80.1	81.2	.33	.37	.56	.60	1.52	1.52	.63	.605
Case-2	82.2	83.3	.38	.42	.68	.73	1.84	1.84	.63	.605
AWB	73	–	.307	–	.767	–	2.77	–	.685	–
Station-10										
Case-1	82.2	81.9	.017	.002	.016	.008	.04	.04	.6	.8
Case-2	86.3	86	.027	.007	.021	.01	.053	.053	.6	.8
AWB	74.3	–	.375	–	.8	–	2.05	–	.61	–

feedback mechanism in avalanche prone regions of North-West Himalaya and the same can be done in future for improvement in avalanche forecasting.

7 Conclusion

The HMM has been developed for avalanche forecasting in 10 different regions over North-West Himalaya with a lead time of 2 days. It has been developed using data of past 20 winters (1992–2012) and validated for four winters (2012–2014). The model performance has been validated for the cases with and without delayed/wet avalanches and compared with official Avalanche Warning Bulletin issued by SASE for North-West Himalaya. The accuracy of HMM for both the cases for all the sectors has been found better than that of the AWB. The HMM has shown better accuracy for Case-2 (excluding delayed/wet avalanches) than that for Case-1 (including delayed/wet avalanches) for all the road sectors in North-West Himalaya. The avalanche forecasting accuracy measures are affected either by delayed/wet avalanches or inaccurate avalanche feedback. Therefore, to improve avalanche forecasting in future separate studies should be carried out for forecasting of direct and delayed/wet avalanches. Moreover, the avalanche feedback mechanism should be improved for accurate avalanche feedback. The model performance can further be improved by including more snow pack parameters into the model. The lead time of avalanche forecast can be improved by coupling HMM with appropriate weather prediction model. Currently, the avalanche forecast for larger region in the HMM is based on the data of a single representative station in that region. For high-resolution/site-specific avalanche forecast, numerical weather prediction model output at a high-spatial resolution can be coupled with the HMM in future.

References

Agrawal KC, and Ganju A (1994) A semi-quantitative approach to avalanche forecasting. In: Proceedings of snow symposium, pp 430–439
Barbec B, Meister R (2001) A nearest neighbor model for regional avalanche forecasting. Ann Glaciol 32:130–134
Bellairea S, Herwijnen AV, Mitterer C, Schweizer J (2017) On forecasting wet-snow avalanche activity using simulated snow cover data. Cold Reg Sci Technol 144:28–38
Bhatnagar V, Goyal RK, Chaudhury S, Gupta HM, Mondal D (1994) On avalanche prediction using neighbourhood criterion. In: Proceedings of snow symposium, pp 414–416
Boyne HS, Williams K (1992) Analyses of avalanche prediction from meteorological data at Berthoud Pass, Colorado. In: Proceedings of the international snow science workshop, Breckenridge USA, pp 229–235
Buser O (1983) Avalanche forecast with the method of nearest neighbors: an interactive approach. Cold Reg Sci Technol 8:135–163
Buser O (1989) Two years' experience of operational avalanche forecasting using the nearest neighbor method. Ann Glaciol 13:31–34
Buser O, Butler M, Good W (1987) Avalanche forecast by the nearest method. IAHS Publication No. 162, pp 557–569
Campbell GS (1977) An introduction to environmental biophysics. Springer, New York

Fromm R (2009) Estimating the forecasting success of artificially triggering of avalanches with the combination of cluster and discriminant analysis. In: Proceedings of the international snow science workshop, Davos, Switzerland, pp 366–370

Gassner M, Etter H, Birkelend K (2000) An improved avalanche forecasting program based on the nearest neighbor method. In: Proceedings of international snow science workshop, Big Sky, Montana, USA, pp 52–59

Gubler H (1982) Strength of bonds between ice grains after short contact times. J Glaciol 28(100):457–473

Helbig N, Van Herwijnen A, Jonas T (2015) Forecasting of wet-snow avalanche probability in mountainous terrain. Cold Reg Sci Technol 120:219–226

Joshi JC, Srivastava Sunita (2014) A Hidden Markov Model for avalanche forecasting on Chowkibal-Tangdhar road axis in Indian Himalayas. J Earth Syst Sci 123(8):1771–1779

Joshi JC, Ganju A, Sharma V (2010) A new approach to avalanche warning over Indian Western Himalaya. Curr Sci 91(11):69–72

Kind (1981) Snow drifting. In: Gray DM, Male DH (eds) Handbook of Snow. Pergamon, Tarrytown, NY, pp 338–359

Kozak MC, Elder K, Birkeland K, Chapman P (2002) Predicting snow layer hardness with meteorological factors. In: Proceedings Snow Science Workshop. Penticton, British Columbia

Marienthal A, Hendrikx J, Birkeland K, Irvine KM (2015) Meteorological variables to aid forecasting deep slab avalanches on persistent weak layers. Cold Reg Sci Technol 120:227–236

Mc Clung DM, Tweedy J (1994) Numerical avalanche prediction: Kootenay Pass, British Columbia, Canada. J Glaciol 40(135):350–358

Mc Collister C, Birklend K, Hansen K, Aspinall R, Comy R (2003) Exploring multi-scale spatial pattern in historical avalanche data, Jackson Hole ski area, Wyoming. Cold Reg Sci Technol 37(3):299–313

McClung DM, Scherer P (2006) The avalanche handbook, 3rd edn. The Mountaineers, Seattle

Naresh P, Pant LM (1999) Knowledge based system for forecasting snow avalanches of Chowkibal-Tangdhar Axis. J&K Def Sci J 49(5):381–391

Obled C, Good W (1980) Recent developments of avalanche forecasting by discriminant techniques: a methodological review and some applications to the Parsenn area (Davos, Switzerland). J Glaciol 25(92):315–345

Peitzsch EH, Hendrikx J, Fagre DB (2015) Terrain parameters of glide snow avalanches and a simple spatial glide snow avalanche model. Cold Reg Sci Technol 120:237–250

Pozdnoukhov A, Matasci G, Kanevski M, Purves RS (2011) Spatio-temporal avalanche forecasting with support vector machine. Nat Hazards Earth Syst Sci 11:367–382

Purves RS, Morrison KW, Moss G, Wright DSB (2003) Nearest neighbors for avalanche forecasting in Scotland- development, verification and optimization of a model. Cold Reg Sci Technol 37(3):343–355

Rabiner LR (1989) A tutorial on Hidden Markov Model and selected applications in speech recognition. Proc IEEE 73(2):257–286

Sharma SS, Ganju A (2000) Complexities of avalanche forecasting in western Himalayas: an overview. Cold Reg Sci Technol 31:95–102

Singh D, Ganju A (2008) Expert system for prediction of avalanche. Curr Sci 94:1076–1081

Singh A, Srinivasan K, Ganju A (2005) Avalanche Forecast using numerical weather prediction in Indian Himalaya. Cold Reg Sci Technol 43:83–92

Vernay M, Lafaysse M, Mérindol L, Giraud G, Morin S (2015) Ensemble forecasting of snowpack conditions and avalanche hazard. Cold Reg Sci Technol 120:251–262

Viterbi A (1967) Error bounds for convolutional codes and an asymptotically optimum decoding algorithm. IEEE Trans Inf Theory 13(2):260–269