Nordic Hydrology 10, 1979, 251-260

Pattern Recognition Analysis of Snowdrifts

W. J. Rawls and T. J. Jackson
USDA-SEA-AR, Hydrology Laboratory, Beltsville, Md, USA

Terrain features and prevailing wind patterns play a dominant role in how snow is accumulated. Research was conducted to determine the location of snowdrifts on a 0.41 km^2 subbasin of the Reynolds Creek Experimental Watershed located in the Owyhee Mountains in southwestern Idaho. Snow depths at maximum accumulation were determined on a 7.6 m grid using aerial photography. The terrain variables used in the analysis were ground slope, ground aspect, elevation, distance from the top of the ridge, in the direction of the prevailing wind and the difference in elevation between the top of the ridge and the point. A pattern recognition technique called cluster analysis was used to determine the best terrain variables for locating snowdrifts. Our analysis indicated that slope and aspect were the best variables. Discriminant functions using these variables and the prevailing wind were developed to predict the location of snowdrifts. The functions were tested on a 16.7 km^2 area.

Introduction

Falling snow is strongly influenced by wind interacting with vegetation and local topography so that snow is deposited in irregular spatial patterns. In some circumstances much of the snowfall is concentrated within a small area of the watershed. As the snowmelt season progresses, each snowdrift functions as a small reservoir, supplying water over a prolonged period of time. Determining the size, shape, and location of these snowdrifts is essential to the developing and management of this water resource (Cox 1975).

Much research has been conducted relating snow accumulation and distribution to climatic, vegetative, and topographic factors. Gray (1978), Meiman (1970), Packer (1962), Kuz'min (1960), Anderson (1968), and Dickison (1978) summarized the relationships developed between snowcover depth and elevation, slope, aspect, and other factors in forested regions with rolling topography. Adams and Rogerson (1968) and Adams (1976) studied the snowcover character in different vegetational zones. McKay (1963) illustrates the importance of terrain and wind in establishing snow cover patterns on the prairies. Steppuhn and Dyck (1974) suggested that the areal variation of snowcover within units having similar landscape features was similar. Similar frequency distributions of depth, density, and water equivalent which may be found within similar areal units are cited as a statistical symptom of this phenomenon. Broader scale studies (McKay 1972) demonstrated differences in certain snowcover properties that may be expected in different climate and vegetative regions.

Thus relationships describing snow accumulation and distribution may be developed for hydrological purposes using climatic, vegetative, and topographic factors. However, Adams (1976) cautioned researchers about extrapolating these relationships to other regions or even to the same region during extreme climatic years. The purpose of this study was to develop techniques for determining the location of snowdrifts based on climatic and topographic features for mountainous rangeland, an area that has not been studied in the literature.

Study Area

The study area is located in the Reynolds Creek Experimental Watershed in the Owyhee Mountains in southwestern Idaho, about 80.5 km southwest of Boise, Idaho (Fig. 1). Reynolds Creek Experimental Watershed, operated by the USDA-SEA-AR Northwest Watershed Research Center, ranges in elevation from 1,402 to 2,195 m and is covered with sagebrush rangeland, except for scattered stands of douglas fir and aspen and mountain meadows. The topography has north- and northwest-trending ridges with gently inclined windward slopes and steep north- and east-facing lee slopes. Annual precipitation ranges from about 25 cm at the lower elevations to 152 cm at the higher elevations. Most precipitation at the higher elevations is in the form of snow with a predominately southwestern wind (Johnson and McAuthor 1973). This area represents large areas of southern Idaho, Oregon and Washington.

The Reynolds Mountain East Watershed located in the upper part of the Reynolds Creek Experimental Watershed (Fig. 1) was used for developing functions to describe the location of drifts. This 41 km^2 watershed ranges in elevation from 2,012 to 2,134 m and has the vegetation as the larger experimental

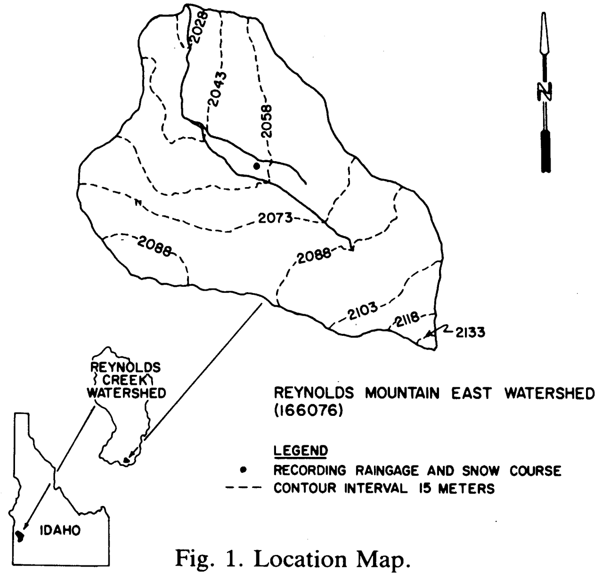

Fig. 1. Location Map.

watershed. The vegetation and topographic characteristics of the watershed are summarized in Table 1. The data base included on a 7.6 m square grid system (6,776 ponts) the elevation, ground slope, ground aspect, the distance from the top of the ridge in the direction of the prevailing wind, the elevation difference from the top of the western ridge and the snow depth for the March 7, 1966; March 17 and May 15, 1967; March 12, and April 30, and May 23, 1969 dates. The snow depths were determined using a photogrammetric technique (Cooper 1965) with a standard error of ± 15.2 cm. According to the snow course located in the Northeast part of the watershed (Fig. 1), the March dates represent maximum accumulation for 1966, 1967 and 1969 where the other dates represent depletion. Table 2 shows that compared with the 16 year snow course record the three maximum depths of accumulation represented a good range of maximum depths of snow accumulations.

An independent data set was used for testing the prediction equations developed from the Reynolds Mountain East Watershed data. These data were from the upper half of the Tollgate Watershed located in Reynolds Creek Experimental Watershed for which the Reynolds Mountain East Watershed is a part. This 16.7 km^2 area ranges in elevation from 1,524 to 2,194 m and has vegetation resembling that found in Reynolds Mountain East Watershed (Table 1). Except for snow depths, the same data as described for Reynolds Mountain East Watershed was found on a 305 m grid (414 points) on May 23, 1969. Since snow depth data were not available for this date we determined whether or not the grid points had snow and used this information to determine where snow drifts were located.

Table 1 – Percentage of Watershed Area in Vegetation, Slope and Aspect Classes.

Watershed	Vegetation Cover Classes (%)				
	0-25	26-50	50-75	76-100	
Tollgate	25	15	15	45	
Reynolds Mt. East	16	18	6	60	
	Slope Classes (%)				
	0-10	11-20	21-30	31-40	41-50
Tollgate	26	54	15	4	1
Reynolds Mt. East	13	48	21	16	2
	Aspect Classes (0)				
	0-90	91-180	181-270	271-360	
Tollgate	36	21	8	35	
Reynolds Mt. East	32	4	11	53	

Table 2 – Snow Course (176007) Characteristics

	Snow Course (176007) Depth (centimeters)
Year	
1966 Maximum	127
1967 Maximum	150
1969 Maximum	188
Maximum (1975) 16 Year Maximum[1]	231
Minimum (1977) 16 Year Minimum[1]	74
16 Year Average[1]	165
1 Standard Deviation about 16 Year Mean[1]	41

[1] Period of record 1961 to 1978

Methods

Pattern recognition involves detecting of patterns or categories within data sets and developing of efficient classifiers for assigning of observations to known categories. It is a collection of mathematical and statistical procedures that vary both in complexity and types of applications. A good overview of the concepts of pattern recognition is presented in Swain (1972). A more detailed description is presented in Fukunaga (1972), Duda and Hart (1973), and Young and Calvert (1974).

This investigation utilized a pattern recognition technique called cluster analysis to identify snow drift categories based on snow depths (Anderberg 1973). Cluster analysis utilizes various mathematical and statistical analyses to identify groupings within a data set. Groups will exhibit a high degree of association and will be distinct from other groups. To perform this task a distance function-similarity measure was used. We utilized Ball and Hall's (1965) clustering algorithim with some minor modifications.

Once we identified the categories within the data set using cluster analysis, the next step was to design a classifier that would categorize all of the observations into these categories. Discriminant analysis was used to develop the classifier in this investigation. Discriminant analysis utilizes a calibration data set to develop a function that separates groups based on observed variables. The mathematics of discriminant analysis are described in Duda and Hart (1973). In this investigation, a package routine described in Barr et. al (1976) was utilized. Classification is based upon a generalized distance function that utilizes the centroid vector and covariance matrix of each group.

Results

As an initial step, correlation and factor analysis were performed on the data. This analysis indicated that the snow depths at the three maximum accumulation dates were highly correlated (0.68, 0.82, and 0.86) and were grouped in the same factor indicating they were representing the same information. Of the topographic variables only elevation, distance from the ridge, and difference between the ridge elevation and point were highly correlated (0.72, 0.76, and 0.55). Ground slope and ground aspect were not significantly intercorrelated or correlated with any of the topographic variables.

To determine if there was a significant spatial grouping of snow depths, cluster analysis was performed on the March data for 1966, 1967, and 1969 and on the May 1969 data separately and in combination. Cluster analysis produced two significant groupings of snow depths which correspond to observed drift and no-drift areas. We felt that the groupings produced by the combination of the March dates (maximum accumulation) were best because they would not be time dependent. The spatial distribution of these groups is illustrated in Fig. 2. When we compared these with the spatial grouping produced for the May 23, 1969 data cluster analysis when only snow drifts remained (Fig. 3) we observed that the spatial patterns were quite similar. The mean snow depths for the three dates ranged from 0.79 to 1.19 m for the no-drift cluster and from 1.65 to 3.57 m for the drift cluster. The mean snow depths of the clusters were significantly separated by between one and three standard deviations. The means for the May 23, 1969 data gave similar results.

In attempting to explain the drift and no-drift cluster pattern given in Fig. 3, we performed cluster analyses separately and in combination on the following variables - elevation, ground slope, ground aspect, the distance from the top of the ridge in the direction of the prevailing wind, the difference in elevation between the top of the southwestern ridge and the point. These analyses indicated that a combination of ground slope and ground aspect were the best variables for explaining the drift pattern shown in Fig. 2. The mean slopes for the drifts and no-drift clusters were 20.6 and 11.7% respectively and were separated by one standard deviation. The mean aspects for the drift and no-drift cluster patterns were 125.2 and 211.1^0 respectively; however, they were not separated by one standard deviation. With this information we developed a discriminant function, using ground slope and aspect and the classifications of snow depth determined from the cluster analysis for the three March dates. The accuracy of the discriminant function was very poor. When we plotted the classification against slope and aspect, we observed that the discriminant function accuracy degraded in the aspect range between 210 and 240^0. Thus we felt that to develop an accurate relationship, a variable that would explain the above problem needed to be added. Since the prevailing wind was from the southwest (225^0) (Johnson and McAuthor 1973), we felt that dividing the discriminant function at 225^0 would account for the specific meteorological conditions and allow the functions to be more universally adaptable. Also, dividing the function would eliminate the problem that occurs in developing discriminant functions with this type of data because 0 and 360^0 aspect are the same. Discriminant functions were developed with the break at 210, 225, and 240^0 and with no significant change in accuracy. Therefore, the general prevailing wind pattern, 225^0 was chosen as the break. The discriminant function is illustrated graphically in Fig. 4.

The discriminant function predicted 80% of the no-drift points and 83% of the drift points for an overall accuracy of 81% for the calibration data set. Discriminant functions were also developed for the May 23, 1969 data and these are graphically illustrated in Fig. 4. This discriminant function predicted 83% of the no-drift points and 92% of the drifts points for an overall accuracy of 84%. In Fig. 4 the three date and one date functions are compared. The one date function has a narrower drift range for the 0 to 225^0 aspect range than does the data for the

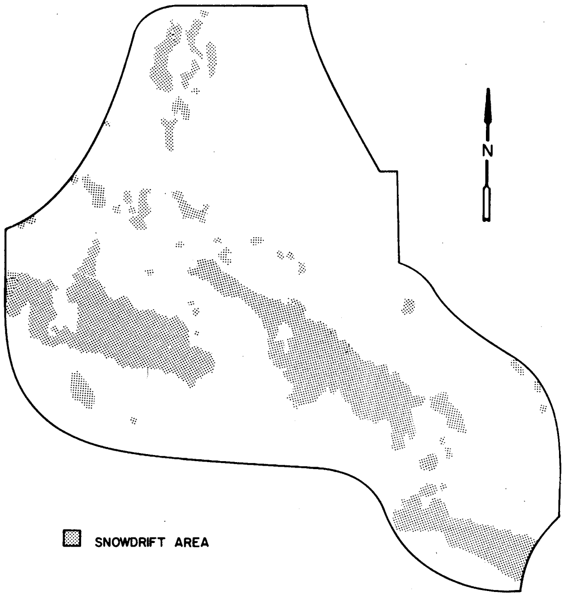

Fig. 2. Snowdrift classification derived from cluster analysis of Mar. 12, 1969, Mar. 7, 1967, and Mar. 17, 1966, data.

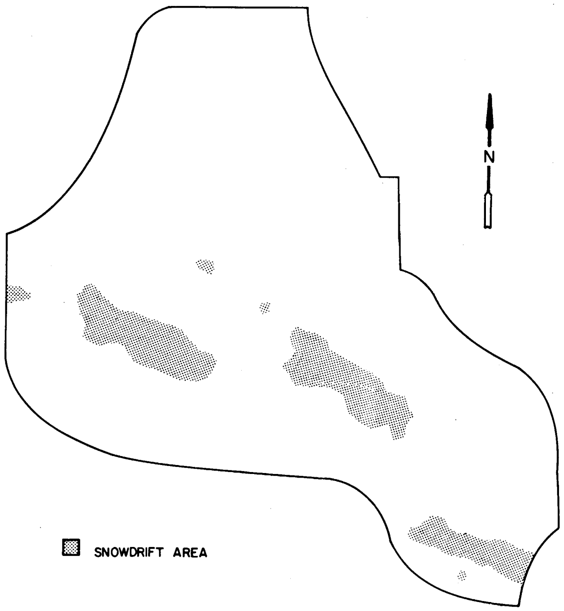

Fig. 3. Snowdrift classification derived from cluster analysis of May 23, 1969, data.

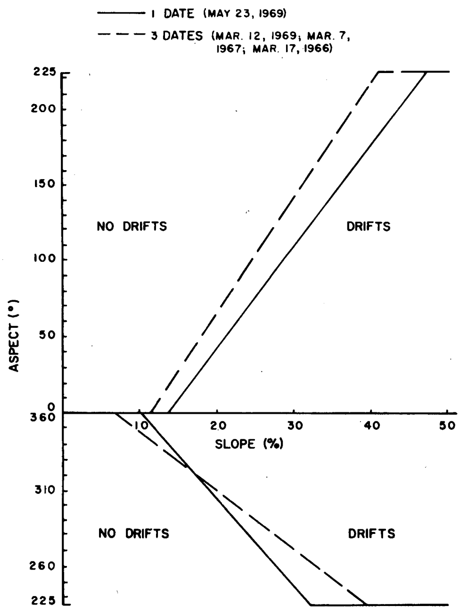

Fig. 4. Snowdrift driscriminant function regions.

three dates; however, in the 225 to 360^0 aspect region the lower slopes have a narrower range and the higher slopes have a wider range. The three date and one date discriminant functions were very similar, however, because of the variety of conditions included in the three date function we felt that this was the best discriminant function for discriminating between drift and no-drift areas based on snow depth.

The three date and one date discriminant functions were tested on the data base for the upper part of the Tollgate Watershed. This watershed had the same prevailing wind pattern as the Reynolds Mountain East Watershed. The three date discriminant function (Fig. 4) predicted 64% of the no-drift points and 65% of the drift points with an overall accuracy of 64%. The one date discriminant function (Fig. 4) predicted 72% of the no-drifts points and 49% of the drift points with an overall accuracy of 62%. The accuracy for the three date and one date discriminant functions indicated that the range between the functions (Fig. 4) is very sensitive and this area needs further clarification. Discriminant functions developed from the test data did not produce any increase in accuracy; thus further confirming the validity of the three date discriminant functions. The lack of prediction accuracy obtained for the test data can be attributed to scale of the data (305 m grid) and the time of year the data were obtained (late May). Also, the differences in relief, and slope and aspect distributions (Table 1) for Reynolds Mountain East Watershed and Tollgate Watershed could have contributed to the poor prediction accuracy obtained on the test data set. Considering the problems associated with the test data set, our results adequately validated the derived function for locating snow drifts. Further work is needed to expand the functions for a variety of topographic and metrological conditions so that the derived functions can be universally applied.

Conclusions

Terrain features and prevailing wind patterns play a dominant role in how snow is accumulated. Research was conducted to determine the location of snowdrifts based on climatic and topographic features for mountainous rangeland. The terrain variables used in our analysis were ground slope, ground aspect, elevation, distance from the top of the ridge in the direction of the prevailing wind, and the difference in elevation between the ridge and the point. A pattern recognition technique called cluster analysis was used to classify drift and no-drift areas and to determine the best terrain variables for locating snowdrifts. Our analysis indicated that slope and aspect were the best variables. Discriminant functions using these variables and the prevailing wind were developed to predict the location of snowdrifts. The functions were tested on a 16.7 km^2 area which included the test area and satisfactorily predicted the location of snowdrifts. Such relationships for predicting the location of snow drifts will enable better management of the vast amounts of water stored in snow drifts which is needed in making management decisions for irrigation, stock water, municipal water supplies hydro electric power generation, recreation, navigation and pollution control.

Acknowledgment

The authors gratefully acknowledge the assistance of Dave Robertson, Lee Perkins, John Zuzel and Lloyd Cox of the USDA-SEA-AR, Northwest Watershed Research Center for collecting and assembling the data used in this study. Also, we wish to thank Albert Rango, Eric Anderson, and John Zuzel for their critical reviews and helpful comments.

References

Adams, W. P. (1976) Areal differentiation of snowcover in East-Central Ontario. *Water Resour. res. 12(6),* 1226-1234.

Adams, W. P., and Rogerson, R. J. (1968) Snowfall and snowcover at Knob Lake, Central Labrador-Ungava. Proc. Eastern Snow Conf. 25, pp. 110-139.

Anderberg, M. R. (1973) *Cluster Analysis for Applications,* Academic Press, Inc., NY.

Anderson, H.W. (1968) Snow accumulation as related to meteorological, topographic, and forest variables in Central Sierra Nevada, California, Int. Assoc. Sci. Hydrol. Publ. 78, pp. 215-224.

Ball, G.H., and Hall, D. J. (1965) IOSDATA, A Novel Method of Data Analysis and Pattern Classification, Technical Report, Stanford Research Institute, Menlo Park, CA.

Barr, A. J. et al (1976) A User's Guide to SAS 76, SAS Institute Inc., Raleigh, NC.

Cooper, C. F. (1965) Snowcover measurement, *Photogrammetric Engineering, Vol. 31,* pp. 611-619.

Cox, L. M., Rawls, W. J. and Zuzel, J. F. (1975) Snow: Nature's Reservoir, Water Resour. Bull., 11(5), 1009-1013.

DeCoursey, D. G. (1973) Application of Discriminant Analysis in Design Review, *Water Resour. Res., 9(1),* 93-102.

Dickison, R. B. B., and Daugharty D. A., (1978) A square grid system for modeling snow cover in small watersheds, Proc. of Workshop/Meeting on Modeling of Snow Cover Runoff, Hanover, NH.

Duda, R. O., and Hart, P. E. (1973) *Pattern Classification and Scene Analysis,* John Wiley and Sons, NY.

Fukunaga, K. (1972) *Introduction to Statistical Pattern Recognition,* Academic Press, NY.

Gray, D. M. et al (1978) Snow Accumulation and Distribution, Proceedings, Workshop-/Meeting on Modeling of Snow Cover Runoff, Hanover, NH.

Kuz'min, P. P. (1960) Snowcover and snow reserves. Gidrometeorologicheskoe, Ixdatelsko, Leningrad. (translation), U.S. National Science Foundation, Washington, DC.

Johnson, C. L., and McAuthor, R. (1973) Winter Storm and Flood Analysis, Northwest Interior. Proc. ASCE Hydraulic Div. Spec. Conf., Boseman, MT pp 359-369.

McKay, G. A. (1963) Relationships between snow survey and climatological measurements. Int. Assoc. Sci. Hydrology, IUGG Assembly. Publ. No. 63 Berkeley, CA.

McKay, G. A. (1972) The mapping of snowfall and snowcover, Proc. Eastern Snow Conf. 29, pp. 98-110.

Meiman, J. R. (1970) Snow accumulation related to elevation, aspect and forest canopy, Proc. Workshop Seminar on Snow Hydrology, Queen's Printer of Canada, pp. 35-47.

Packer, P. E. (1962) Elevation, aspect, and forest cover effects on maximum snowpack water content in a western white pine forest. *For. Sci., 8(3),* 225-235.

Steppuhn, H., and Dyck, G. E. (1974) Estimating true basin snowcover, in Advanced Concepts and Techniques in the Study of Snow and Ice Resources, National Academy of Sciences, U.S. Int. Hydrological Decade, Washington DC. pp. 314-328.

Swain, P. H. (1972) Pattern Recognition: A basis for Remote Sensing Data Analysis, LARS Information Note 111572, Laboratory for Applications of Remote Sensing, Purdue University, West Lafayette, IN.

Young, T. Y., and Calvert, T. W. (1974) *Classification, Estimation and Pattern Recognition,* Elsevier, Inc., NY.

Received: 1 May, 1979

Address: Hydrology Laboratory, PPHI,
Rm. 139, Bldg. 007, BARC-West,
Beltsville, Maryland 20705,
U.S.A.