TIME-SERIES MODELLING OF AVALANCHE ACTIVITY FROM METEOROLOGICAL DATA

By A. A. SALWAY

(Division of Building Research, National Research Council of Canada, B.C. Regional Station, 3904 West 4th Avenue, Vancouver, British Columbia V6R 1P5, Canada)

ABSTRACT. Avalanche-hazard evaluation by field analysts is largely achieved along causal intuitive lines, depending for its success upon the experience of the analyst in his particular area. Several attempts have been made in the past to quantify such procedures by means of statistical models based upon meteorological measurements. Modified forms of a multivariate technique known as linear discriminant analysis have been tried with only partial success. Intercorrelated variables and autocorrelated data, omission of time-lagged terms, insufficient variation in the dependent variable, and sampling difficulties may have combined to weaken the discriminant approach. These problems and the nature of the phenomenon suggest that a stochastic transfer-function time-series approach may be a useful alternative method.

A numerical weighting scheme has been devised for the representation of avalanche activity for the Rogers Pass area of British Columbia in terms of terminus, size, and moisture-content codes for each event. From various types of correlation analysis performed on data for the period 1965–73, models were developed using the "best" weighting scheme for avalanche activity representation and the most promising meteorological variables, as indicated by the results of the correlation analysis.

These relatively simple models demonstrate a good fit to the actual data, in both a descriptive and a simulated-forecasting situation.

RÉSUMÉ. *Modélisation de séries chronologiques de l'activité des avalanches à partir de données météorologiques.* L'estimation du danger d'avalanche par des observateurs sur le terrain est en grande partie basée sur des déductions intuitives, dont le succès dépend de l'expérience personnelle de l'observateur dans son secteur particulier. On a fait plusieurs tentatives dans le passé pour quantifier ces processus par des modèles statistiques basés sur des données météorologiques. Des formes dérivées de la technique d'analyses à variables multiples, connues sous le nom d'analyses linéaires discriminantes, ont été essayées avec un succès seulement partiel. Des variables liées entre elles et des données auto-corrélées, l'omission de termes de décalage dans le temps, l'insuffisance des variations des valeurs de la variable dépendante et les difficultés d'échantillonnage se sont combinées pour affaiblir l'efficacité de l'approche discriminante. Ces problèmes et la nature du phénomène font penser que l'on peut recourir à une approche d'une fonction stochastique de transfert de séries chronologiques.

Un schéma numérique de pondération a été imaginé pour la représentation de l'activité des avalanches dans le secteur de Rogers Pass en Colombie Britannique par des codes donnant pour chaque événement le point extrême d'arrivée, l'étendue et le taux d'humidité. Différents types d'analyses de corrélation appliqués aux données de la période 1965–73 ont été mis au point, utilisant le schéma "optimum" de pondération pour représenter l'activité des avalanches et les variables météorologiques les plus significatives selon les résultats des analyses de corrélation.

Les courbes de prévision construites d'après ces modèles relativement simples montrent une bonne concordance avec les données réelles, aussi bien pour la description que pour la prévision d'une situation.

ZUSAMMENFASSUNG. *Zeitreihenmodelle der Lawinentätigkeit aus meteorologischen Daten.* Die Auswertung von Lawinenabgängen durch Feldbeobachtung folgt meist intuitiven Kausalketten; ihr Erfolg hängt von der Erfahrung des Beobachters in seinem speziellen Gebiet ab. In der Vergangenheit wurden mehrere Versuche zur Quantifizierung solcher Verfahren mit Hilfe von statistischen Modellen auf der Grundlage meteorologischer Messungen unternommen. Modifizierte Formen einer mehrparametrigen Technik, der sog. linearen Unterscheidungsanalyse, brachten nur teilweisen Erfolg. Dies liegt an der Korrelation der Variablen und der Autokorrelation der Daten, an der Vernachlässigung von zeitverzögeten Termen, an der ungenügenden Variation der abhängigen Variablen und an den Schwierigkeiten der Datengewinnung. Diese Probleme und die Natur des Phänomens legen es nahe, Zeitreihen auf der Basis von stochastischen Übertragungsfunktionen zu verwenden.

Für die Erfassung der Lawinentätigkeit im Gebiet des Rogers-Passes, British-Columbia, wurde ein numerisches Wichtungsschema entwickelt, das Ende, Grösse und Feuchtigkeitsgehalt jedes Abganges berücksichtigt. Anhand von Daten für die Periode 1965–73 wurden Modelle mit verschiedenen Typen der Korrelationsanalyse entwickelt, wobei auf Grund der Ergebnisse der Korrelationsanalysen das "beste" Wichtungsschema für die Wiedergabe der Lawinentätigkeit und die meistversprechenden meteorologischen Variablen herangezogen wurden.

Vorhersageprofile, abgeleitet aus diesen relativ einfachen Modellen, zeigen gute Übereinstimmung mit den tatsächlichen Daten, sowohl in einer beschreibenden wie in einer simulierten vorhersagenden Situation.

NOMENCLATURE

Y index of avalanche activity, dependent variable
$\hat{Y}$ predicted value of Y
S snow-pack depth
N new-snow depth
W precipitation
T_1 maximum air temperature
T_2 minimum air temperature
V wind speed
V_1 north component of wind speed
V_2 west component of wind speed
V_3 south component of wind speed
V_4 east component of wind speed
H relative humidity
X independent variable
$x = (X - \bar{X})$ deviation from mean of X
$y = (Y - \bar{Y})$ deviation from mean of Y
r_{xy} sample correlation coefficient between x and y
M number of observations in the sample
t time
x_t input time series
y_t output time series
$k = 1, 2, 3, \ldots, K$ time lag
$r_{xx}(k), r_{yy}(k)$ sample autocorrelation functions
$r_{xy}(k)$ sample cross-correlation function
$c_{xx}(0), c_{yy}(0)$ sample variances
$c_{xx}(k), c_{yy}(k), c_{xy}(k)$ sample covariances
ϕ autoregressive coefficient
θ moving average coefficient
p number of autoregressive coefficients
q number of moving average coefficients
B backward-shift operator, e.g. $Bx_t = x_{t-1}$
∇ difference operator, e.g. $\nabla x_t = x_t - x_{t-1}$
α_t *white-noise* series
β_t transformed output series after *prewhitening* input series
ν transfer-function coefficient for primary input series
n_t stochastic noise series
ϵ_t transformed stochastic noise series after *prewhitening* input series
u transfer-function coefficient for secondary input series
$\hat{\nu}, \hat{u}$ transfer-function coefficient estimates
a_t residual error

INTRODUCTION

At an elevation of approximately 4 350 ft (1 330 m), Rogers Pass provides an important east–west route through the Selkirk Mountains of British Columbia via the Trans-Canada Highway and the Canadian Pacific Railway. It is one of the most active avalanche areas in western Canada. A combination of steep-sided mountains, a characteristic of the Selkirk Range, and heavy winter snowfall has produced more than ninety major avalanche sites that affect a 50 km length of highway from the east gate of Glacier National Park to just beyond

the western boundary. The greatest concentration of sites exists between two narrow defiles formed by Mounts Tupper and MacDonald, just east of the Pass, and Mounts Fidelity and Fortitude in the western section. Slope aspects are predominantly north and south. Some of the sites are on open slopes but most are mountain gullies with well-defined scree fans as depicted in Figure 1.
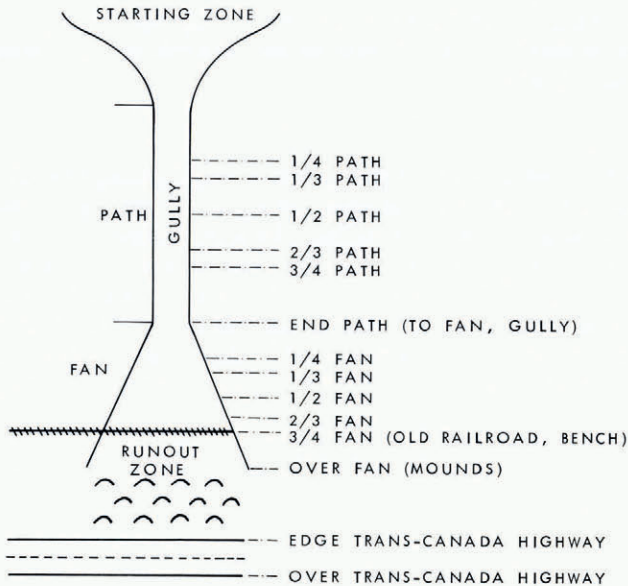

Fig. 1. Typical Rogers Pass avalanche site.

The Snow Research and Avalanche Warning Section (SRAWS) of Parks Canada conducts an ongoing programme of avalanche-hazard evaluation and control for the area. Operational objectives are to maintain an optimum balance between minimum highway closure time and the safety of the public and Parks personnel. This balance can only be achieved by the accurate evaluation of avalanche hazard backed by prompt action in the form of highway closure where necessary and appropriate use of artillery to cause avalanching under controlled conditions.

Avalanche-hazard evaluation relies essentially on three factors: the stability of the layers in the upper section of the snow-pack, the stability of the lower layers, and the availability of snow for avalanching at each site. Avalanches are often triggered by new snow in the starting zones, whereas both upper and lower layers further down the path may provide snow for the main mass of the avalanche. This hazard evaluation is made by the field analyst, often intuitively, depending for success upon his experience in the area.

It is not always possible, however, to identify periods of instability intuitively. This paper is concerned with the problem of statistical estimation and prediction of avalanche activity using meteorological data and analytical techniques. Mathematical models have been developed that describe the phenomenon in terms of the statistical behaviour of previous avalanche and meteorological patterns. It is hoped that such models will become a useful tool enabling the field analyst to evaluate the hazard situation more accurately and to identify and predict periods of instability with greater certainty.

LINEAR DISCRIMINANT ANALYTICAL PROCEDURE

Various attempts have been made, notably by Judson and Erickson (1973), Bois and others ([1975]), Föhn and others (1977), and Bovis (1974, 1976, 1977), to produce forecasting models for avalanche occurrences using modified forms of a multivariate statistical technique known as linear discriminant analysis. This procedure involves the assignment of cases into two or more groups, using linear discriminant functions. These functions consist of linear combinations of independent variables multiplied by appropriate coefficients (least-squares estimates) obtained by maximizing the ratio of the between groups variances to the within-groups variances (Rao, 1952).

Lagged variables

Current avalanche activity is affected by previous avalanching, precipitation, temperature, wind, and other weather variables. Models produced by the application of discriminant procedures rely on current weather variables, although Judson and Erickson (1973) introduced certain arbitrary decay terms to overcome this deficiency. Bovis (1974, 1976, 1977), Bois and others ([1975]), and Föhn and others (1977) used terms summed or averaged over several days prior to an avalanche day. Perla and Judson (1973) investigated the possibility of introducing fading memory terms, without arbitrary factors, into the discriminant analysis procedures. Discriminant analysis, however, does not readily lend itself to time-series applications.

A stochastic transfer-function time-series approach uses lagged values of variables such as precipitation, temperature, and wind. The coefficients of these terms are least-squares best estimates determined from actual data.

Intercorrelated variables and autocorrelated data

Strong interrelations of independent variables, a normal feature of weather data, are not handled well by conventional regression and discriminant methods. Furthermore, the meteorological time series are usually quite strongly autocorrelated, that is, adjacent observations in time are not independent. Variable selection procedures such as forwards selection, backwards elimination or stepwise selection tend to break down. Models are produced that do not necessarily contain the "best" set of independent variables (Draper and Smith, 1966, p. 163–95).

Time-series analysis procedures operate on dependent observations "where the nature of this dependence is of interest in itself" (Box and Jenkins, 1970, p. vii).

Variation of the dependent variable

Assigning all avalanche days to one class, whether the level of activity is high or low, may not be desirable. If avalanche activity can be treated as an ordinary dependent variable, allowing it to take on values corresponding to various levels of activity, it may more accurately reflect the changing meteorological conditions that give rise to the phenomenon.

Bovis (1974, 1976, 1977), has tried to overcome this drawback in the discriminant approach by stratifying avalanche events on the basis of magnitude for both dry and wet seasons. He recognizes that "The conditions which lead to an avalanche day having only one event are likely to be quite different from those which lead to days which involve a major avalanche 'cycle'" (Bovis, 1977, p. 94). Four magnitude classes of avalanche activity are recognized for an area. Stratification of events in this way, however, results in a reduction of the sample sizes. If the groups are too small, the discriminant analysis procedure breaks down. Overfitting may also occur when the number of independent variables is too large relative to the sample sizes. Assigning a range of values to avalanche activity eliminates the need for stratification.

Data imbalance between avalanche days and non-avalanche days

There is an imbalance between avalanche and non-avalanche days. Strong differences exist between the group variances, particularly for precipitation terms, and there are usually far more non-avalanche days. This can result in discriminant functions biased in the direction of the non-avalanche group, so that a greater proportion of avalanche days are misclassified than are non-avalanche days. To overcome this difficulty, Judson and Erickson (1973) use a weighted average of the discriminant means for each group as their discriminant index. This causes the probabilities of misclassification for avalanche and non-avalanche days to be approximately equal, but does little to improve the overall classification scheme.

Bois and others ([1975]) and Bovis (1974, 1976, 1977) equalized their discriminant groups by selecting a random sample of non-avalanche days equal in number to the avalanche days. The autocorrelation functions of meteorological and avalanche series may contain information about the processes involved and should probably not be eliminated by random sampling. Serial correlations can be exploited by use of time-series procedures.

AVALANCHE ACTIVITY AS A DEPENDENT VARIABLE

Field observations of avalanche occurrences

Naturally occurring avalanches are generally recorded twice daily at Rogers Pass, often after the event and according to a prescribed format. The site name, date, and, if possible, the time of occurrence are noted, along with the observer's estimate of size, terminus, and moisture content, as indicated by the designations in Table I and Figure 1. Continuous monitoring of such an intermittent phenomenon is often difficult owing to limitations on the availability of personnel, high hazard, and poor visibility, particularly during periods of intense activity when observations are most needed. These problems, combined with the necessarily subjective nature of the measurements, set a limit on the overall accuracy of the data and ultimately determine the level of random noise in the prediction models. Artificially released avalanches are noted during artillery application and can therefore be timed reliably when visibility is good. Artillery control is performed whenever the field analyst considers that the probability of avalanching is high. Size, terminus, and moisture content are also recorded whenever possible.

TABLE I. AVALANCHE ACTIVITY INDEX WEIGHTING SCHEMES

	(I)	(II)	(III)	(IV)
Terminus				
$\frac{1}{4}$ or $\frac{1}{3}$ path	1	1	1	1
$\frac{1}{2}$ path	2	2	2	1
$\frac{2}{3}$ or $\frac{3}{4}$ path	3	3	3	1
End path, to fan, or gully	4	4	4	1
$\frac{1}{4}$ fan	5	5	5	1
$\frac{1}{3}$ fan	6	6	6	1
$\frac{1}{2}$ fan	7	7	7	1
$\frac{2}{3}$ fan	8	8	8	1
$\frac{3}{4}$ fan, old railroad, or bench	9	9	9	1
Over fan, or mounds	10	10	10	1
Edge Trans-Canada Highway	11	11	11	1
Over Trans-Canada Highway	12	12	12	1
Size				
Small	1	1	1	0
Medium	6	2	1	1
Large	12	3	1	1
Moisture content				
Dry	1	1	1	1
Damp	3	1	1	1
Wet	6	1	1	1

For both artificially produced and naturally occurring avalanches, size is estimated in relation to the particular site, either from the visual appearance of the site and size of the deposit (for those that occur naturally, after the event) or from a visual impression of mass and energy if the avalanche is actually observed, as is often the case with those that are brought down artificially. Size is designated by the field observer as small, medium, or large.

The terminus classification gives an indication of the farthest point reached by the avalanche but does not include any information on the actual distance travelled. A low cloud base frequently obscures the starting zone, thereby preventing the point of origin or fracture line from being recorded. As individual sites consistently avalanche from the same rupture area, often at the base of cliffs, the terminus can be regarded as an index of distance travelled.

Moisture content is estimated by the observer from the appearance of the snow in the avalanche deposit or by gently squeezing a sample. If no ball forms, it is classed as dry; if one does form, it is damp; and if the surface becomes glassy, it is wet.

Avalanche activity index

Definition and physical interpretation. Prior to the application of statistical time-series procedures it was necessary to devise a suitable index of avalanche activity that could be used as a dependent variable. An index of mass movement (Salway, unpublished) was therefore defined as the product of terminus, size, and moisture content for a particular event after appropriate numerical codes had been assigned to these designations, as shown in Table I. For example, using the codes shown in the first column (weighting scheme (I)), a damp avalanche of medium size running $\frac{1}{4}$ fan would have an index given by:

$$\text{Avalanche activity index} = (3 \times 6 \times 5)$$
$$= 90$$

This index may be a good measure of avalanche activity. It includes not only size, and therefore an indication of the amount of snow picked up from lower zones after the initial movement, but also a measure of the energy associated with the avalanche in terms of the farthest point reached. Moisture content of the resulting deposit also is directly related to the type of the avalanche.

Computation. The avalanche activity index, when computed for individual events, can be summed for all sites or for a group of sites on a daily basis, resulting in values of avalanche activity that vary smoothly and continuously, lending themselves to the successful application of multiple regression or time-series techniques. LaChapelle (1970, p. 106) recognized that the greatest potential of the statistical approach lies in this direction when he stated that it "is most useful when dealing with hazard probabilities over large areas, where individual avalanches fall effectively at random, but the patterns of their occurrence in time are related to snow and weather". Effects of random errors caused by individual site peculiarities, the subjective nature of the data, and occasional unobserved avalanches are minimized. Not only does the index contain the three basic characteristics of the avalanches measured at Rogers Pass but the relative contribution of each measurement can be altered by choosing a different set of weights.

METEOROLOGICAL VARIABLES

Meteorological data from the Rogers Pass observatory for the period 1965–73 were transcribed from field books and stored on computer tape. They consist of daily observations of snow-pack depth, new-snow depth, precipitation, maximum and minimum air temperature, average wind speed and direction, and relative humidity. Table II lists these primary meteorological variables together with assigned symbols for future reference. Such primary factors, which are known to have considerable influence on avalanche activity, and certain secondary factors formed from the primary terms were included in the analyses.

TABLE II. METEOROLOGICAL VARIABLES

Algebraic symbols	Description	Units
S	Snow-pack depth	cm
N	New-snow depth	cm
W	Precipitation (water equivalent of new snow and/or rainfall)	mm
T_1	Maximum air temperature	°F
T_2	Minimum air temperature	°F
V_1	North wind component	km/h
V_2	West wind component	km/h
V_3	South wind component	km/h
V_4	East wind component	km/h
V	Unresolved wind speed	km/h
H	Relative humidity	%

CORRELATION ANALYSES

Several correlation-analysis procedures were performed in order to identify appropriate starting, finishing, and transition dates for avalanche activity periods, the "best" avalanche-activity weighting schemes and most suitable meteorological variables to be used in the subsequent development of the time-series models (Salway, unpublished). Simple linear correlation coefficients were used to provide a good initial indication of the potential strength of models based on these variables. They are defined by the familiar equation

$$r_{xy} = \frac{\sum\limits_{1}^{M} (X - \bar{X})(Y - \bar{Y})}{\left(\sum\limits_{1}^{M} (X - \bar{X})^2 \sum\limits_{1}^{M} (Y - \bar{Y})^2 \right)^{\frac{1}{2}}},$$

where Y and X are dependent and independent variables, respectively, and M is the number of observations in the sample.

Data-selection computer program

In order to allow complete flexibility in the application of the various types of analysis, a data-selection computer program was written. It takes the meteorological information for one or more ranges of dates, computes the appropriate avalanche activity indices for each day, and writes the entire record, including the index, into a file ready for the analysis programs. Apart from allowing freedom in the choice of dates, the program permits any combination of avalanche sites (or ranges of sites), avalanche types (artificial or natural), sizes, and moisture contents to be specified as criteria for including any one avalanche as part of the activity index.

Identification of dates

Starting and finishing dates were defined as those for the first and last avalanches of each of the eight avalanche seasons, 1965–73. Bois and others ([1975]) and Föhn and others (1977) grouped their data on a monthly basis, "for example, all Januarys in the 10 year sampling period are taken as the total population for that month" (Bois and others, [1975], p. 394). This procedure was adopted to eliminate the seasonal trend but assumes that similar conditions occur during the same month each year on a regular basis. Some winters may be more advanced than others, however, on a particular date each year. Bovis (1974, 1976, 1977) partitioned avalanche seasons into dry and wet avalanche periods but recognized that transition dates were hard to identify (Bovis, 1977, p. 95). Dry avalanches are not confined to the first, nor wet avalanches to the second, part of the winter. In this study, winters were

divided into a first period, during which avalanching is primarily precipitation dependent, and a second period during which it is primarily temperature dependent. Extreme temperature dependence is manifested in the spring avalanche cycles. Identification of optimum transition dates was achieved by visual inspection of the data patterns, combined with an incremental correlation-analysis procedure designed for this purpose. Correlation coefficients are computed between avalanche activity and certain meteorological variables, such as snowfall, after each sequential data record has been read and added to the work file. Thus, values of the correlation coefficient can be monitored as analysis of data for a winter progresses and the best "cut-off" date is selected for the earlier period. It was observed that for most winters such correlation coefficient values rise to a peak near the end of January or beginning of February, then drop off sharply as snow-fall becomes less significant than temperature. Periods identified as primarily precipitation dependent for the eight seasons were selected for subsequent analysis and use as the basis for development of the time-series models described in this paper. The primarily temperature-dependent periods will require a separate study.

"Best" avalanche-activity weighting schemes

The three best avalanche-activity weighting schemes were found to be schemes (III), (II), and (IV), Table I, in that order, indicating that terminus is a better measure than size and that exclusion of small avalanches would not reduce the correlations significantly. Varying the moisture content codes did not result in any further improvements. It should be noted that in scheme (III) terminus alone determines the value of the index, and that scheme (IV) implies equal weighting for all classifications, omitting small avalanches and is therefore simply a frequency count of medium and large avalanches.

"Best" meteorological variables

A composite meteorological variable, SWH (the product of snow-pack depth, precipitation, and relative humidity), was found to be the variable most highly correlated with avalanche activity. The relative significance of all the other meteorological variables is shown in Table III, in which correlation coefficient values, expressed as percentages, have been computed using Rogers Pass meteorological data on a daily basis for the first periods of the eight winters from 1965 to 1973, and the three "best" avalanche-activity weighting schemes. Several possible reasons can be given for the importance of precipitation, snow-pack depth, and relative humidity in the SWH term.

Precipitation, the "best" of the simple meteorological variables, is better correlated with avalanche activity than is new-snow depth alone, probably because it is a more direct measure of slope loading or the magnitude of shear stress imposed by the loading. The rate of increase of shear stress in relation to the rate at which snow gains strength through sintering and compaction undoubtedly determines the short-term potential for avalanching.

Tests of relative significance may be performed assuming approximate normality. Correlation coefficient values for SW compared with W alone (Table III) indicate a highly significant improvement at the 0.999 level, as described by Hotelling's t-test (Freeze, 1964, p. 108). This could not be due merely to the establishment of the minimum snow-pack depth required for avalanches to start (about 70 cm for the Rogers Pass area). This depth had already been established on or near the starting dates used in the analyses. The importance of depth of snow-pack as a major factor modifying precipitation is undoubtedly the result of the greater participation of the deepening snow-pack in avalanching.

Correlation coefficient values for SWH also demonstrate a statistically significant improvement over those for SW, although the effect does not appear to be great.

Wind is undoubtedly important in redistributing snow and loading lee slopes, but its effect on large-scale avalanching, as opposed to sloughing, is probably somewhat non-linear,

TABLE III. CORRELATION COEFFICIENTS: ROGERS PASS METEOROLOGICAL DATA 1965–73 DAILY OBSERVATIONS, FIRST PART OF WINTER, ALL SITES, ARTIFICIAL AND NATURAL AVALANCHES

Meteorological variables	Avalanche activity weighting schemes		
	(III) %	(II) %	(IV) %
S	37	38	38
N	67	65	63
W	71	68	66
T_1	22	20	18
T_2	21	19	18
V_1	-2	-2	-2
V_2	17	16	14
V_3	28	27	26
V_4	-12	-11	-11
H	15	14	13
$(S_{t-1}+N_t-S_t)$*	12	12	11
V^2	32	31	30
V	25	25	24
NV	58	57	55
T_2V	30	29	28
NT_2	57	53	52
$(T_{2,\,t}-T_{2,\,t-1})$†	22	19	18
T_2/S	-6	-6	-6
W/N‡	17	16	14
WT_2/N	21	19	18
WV/N	27	26	25
WH/N	18	17	15
NH	67	64	63
T_2H	21	19	18
VH	28	27	26
NVT_2H	52	50	49
WV	61	60	58
WT_2	59	56	55
WH	71	68	66
WVT_2H	52	51	50
SW	77	78	76
SWH	78	79	77
$SWHT_2$	75	74	74
$SWHT_2V$	62	63	63

739 daily records.

Correlation coefficient values $\geqslant 10$ are significant at the 99% level.

* Settlement term (S_{t-1} is the previous value and S_t the current value of S).

† Change in minimum air temperature ($T_{2,\,t-1}$ is the previous value and $T_{2,\,t}$ the current value of T_2).

‡ Density of new snow.

diminishing with increasing wind speed above a critical level (Seligman, 1936, p. 159–85; Mellor, 1968, p. 151; U.S. Dept. of Agriculture, 1961, p. 35; Perla and Martinelli, 1976, p. 35). This would explain the low linear correlation coefficients between wind terms and avalanche activity obtained in this study (Table III). A separate study will be required to evaluate the complex influence of wind on avalanching at Rogers Pass.

TIME-SERIES MODEL DEVELOPMENT

A versatile system was designed and programmed for the development of optimum transfer-function stochastic noise models. The main instrument in these procedures is a computer program written specifically for this study and incorporating the features required to perform the following steps.

(1) Sample autocorrelation functions, $r_{xx}(k)$ and $r_{yy}(k)$ for the input series x and the output series y, are computed up to 20 lags ($K = 20$) using the following equations,

$$r_{xx}(k) = c_{xx}(k)/c_{xx}(0),$$
$$r_{yy}(k) = c_{yy}(k)/c_{yy}(0),$$

where

$$c_{xx}(k) = \frac{1}{M} \sum_{t=1}^{M-k} x_t x_{t+k}, \qquad k = 0, 1, 2, ..., K$$

and

$$c_{yy}(k) = \frac{1}{M} \sum_{t=1}^{M-k} y_t y_{t+k}, \qquad k = 0, 1, 2, ..., K.$$

Sample cross-correlation functions $r_{xy}(k)$ are also computed from

$$r_{xy}(k) = c_{xy}(k)/(c_{xx}(0) \, c_{yy}(0))^{\frac{1}{2}},$$

where

$$c_{xy}(k) = \frac{1}{M} \sum_{t=1}^{M-k} x_t y_{t+k}, \qquad k = 0, 1, 2, ..., K,$$

M is the number of observations in the sequence, $c_{xx}(0)$ and $c_{yy}(0)$ are sample variances, and $c_{xx}(k)$, $c_{yy}(k)$, and $c_{xy}(k)$ are sample co-variances. These correlation functions are examined for "stationarity". A "stationary" process is said to be strictly stationary if its properties are unaffected by a change of time origin.

Deviations from the series means were used throughout the model development for computational convenience, i.e.

$$x_t = X_t - \bar{X}, \qquad y_t = Y_t - \bar{Y}.$$

If the stochastic time series exhibits "non-stationary" behaviour, usually indicated by a slow and linear tapering of the auto- or cross-correlation functions, it may be necessary to apply some degree of differencing to the data. "We assume that a degree of differencing d necessary to induce stationarity has been achieved when the estimated auto- and cross-correlations $r_{xx}(k)$, $r_{yy}(k)$, and $r_{xy}(k)$ of $x_t = \nabla^d X_t$, and $y_t = \nabla^d Y_t$ damp out quickly. In practice, d is usually 0, 1 or 2" (Box and Jenkins, 1970, p. 378). For avalanche activity and for most of the composite meteorological series the correlation functions decayed to insignificance rapidly (at or before the third lag). Differencing was therefore unnecessary.

(2) Partial autocorrelation functions (defined as the last autoregression coefficient obtained after successively fitting increasing orders of autoregressive process to the data) are computed for the input series up to twenty lags. Suitable autoregressive moving average (ARMA) (p, q) models are identified from the autocorrelation and partial autocorrelation functions and the parameters estimated using a least-squares approach. These models are of the form,

$$\phi_x(B) \, x_t = \theta_x(B) \, \alpha_t,$$

where B is the backward-shift operator defined by

$$Bx_t = x_{t-1},$$

and

$$\phi_x(B) \, x_t = (1 - \phi_1 B - \phi_2 B^2 - ... - \phi_p B^p) \, x_t$$
$$= x_t - \phi_1 x_{t-1} - \phi_2 x_{t-2} - ... - \phi_p x_{t-p}.$$

Similarly,

$$\theta_x(B)\ \alpha_t = \alpha_t - \theta_1\alpha_{t-1} - \theta_2\alpha_{t-2} - \ldots - \theta_q\alpha_{t-q},$$

$\phi_1, \phi_2, \ldots, \phi_p$, are autoregressive coefficients, x_t is the input series, $\theta_1, \theta_2, \ldots, \theta_q$, are moving-average coefficients, and α_t is a white-noise process (Box and Jenkins, 1970, p. 9–10).

(3) The ARMA model describing the first (primary) input series is used to transform it into an approximate white-noise process (random and uncorrelated), and the same transformation is applied to the output series. That is,

$$\alpha_t = \phi_x(B)\ \theta_x^{-1}(B)\ x_t,$$
$$\beta_t = \phi_x(B)\ \theta_x^{-1}(B)\ y_t.$$

Pre-whitening of the input series produces more efficient estimates of the transfer function (Box and Jenkins, 1970, p. 379).

(4) Assuming the model

$$y_t = \nu(B)\ x_t + n_t,$$

the transfer function

$$\nu(B) = (\nu_0 + \nu_1 B + \nu_2 B^2 + \ldots)$$

of the process described by

$$\beta_t = \nu(B)\ \alpha_t + \epsilon_t,$$

where ϵ_t is the transformed stochastic noise series

$$\epsilon_t = \phi_x(B)\ \theta_x^{-1}(B)\ n_t,$$

may then be identified and least-squares estimates of the parameters obtained using an all-combination approach up to five lags. The model, based on a single input series, may now be written,

$$y_t = \hat{\nu}(B)\ x_t + n_t,$$

where $\hat{\nu}(B)$ is the estimated deterministic transfer function and n_t is the stochastic noise. x_t and n_t are assumed to be independent.

(5) If a secondary input series is contemplated, it may be selected by calculating partial correlation coefficients between $(y_t - \hat{\nu}(B)\ x_{1,\ t})$ and the residuals of $x_{2,\ t}$ after regressing $x_{2,\ t}$ on all of $x_{1,\ t}$ contained in $\hat{\nu}(B)\ x_{1,\ t}$. $x_{1,\ t}$ is the primary input series and $x_{2,\ t}$ is each of the proposed secondary input series. The highest partial correlation coefficient indicates the best choice for the secondary series. Steps (3) and (4) are then repeated, using $x_{2,\ t}$ as the new input series and $(y_t - \hat{\nu}(B)\ x_{1,\ t})$ as the new output series. The model may now be written,

$$y_t - \hat{\nu}(B)\ x_{1,\ t} = \hat{u}(B)\ x_{2,\ t} + n_t,$$

where $\hat{u}(B)$ is the estimated transfer function for the secondary input series. The process may be repeated for several secondary series.

(6) Finally, the complete transfer-function component is subtracted from the output series to obtain the stochastic-noise component, that is,

$$n_t = y_t - \hat{\nu}(B)\ x_{1,\ t},$$

for one input series, or

$$n_t = y_t - \hat{\nu}(B)\ x_{1,\ t} - \hat{u}(B)\ x_{2,\ t},$$

for two input series, etc.

(7) Autocorrelation and partial autocorrelation functions for the stochastic-noise process are computed up to twenty lags and a suitable ARMA model describing the process identified and estimated,

$$n_t = \phi_n^{-1}(B)\ \theta_n(B)\ a_t.$$

(8) The transfer-function and stochastic-noise components are then combined,

$$y_t = \hat{v}(B)\, x_t + \phi_n^{-1}(B)\,\theta_n(B)\, a_t,$$

and efficient least-squares estimates of the parameters obtained. Any insignificant terms are eliminated and the final, complete model re-estimated.

(9) In this last step, tests of model adequacy are performed as described by Box and Jenkins (1970, p. 392–95). Among other things, insignificant autocorrelation in the residuals is confirmed.

Several sets of models have been developed for Rogers Pass using the foregoing procedures (Salway, unpublished). Presented here is an optimum time-series model based on the "best" weighting scheme ((III), Table I) for the total period 1965–73. Daily observations were used and all types of avalanche; slab and loose, dry, damp, and wet, were included. Artillery releases were combined with natural avalanches because these account for a significant portion, up to 40% per year, of the total avalanching at Rogers Pass. Artificial occurrences probably result from snow-pack instabilities in the same way as do natural occurrences.

The complete model obtained, after performing the previously outlined steps (1) to (9), was

$$\hat{Y}_t = 0.095 Y_{t-1} + 0.060 Y_{t-2} + 0.000\,685 SWH_t - 0.000\,094 SWH_{t-1} + \\ + 0.131 WT_{2,\,t} - 0.035 WT_{2,\,t-2}.$$

Y_{t-1} and Y_{t-2} are the first and second lags, respectively, of avalanche activity Y_t. SWH_t is the primary input series (SWH_{t-1}, the first lag) and $WT_{2,\,t}$ is the secondary input series ($WT_{2,\,t-2}$, the second lag). The model has an R-value (multiple-correlation coefficient) of 0.81 and a standard error of estimate of 61.3 compared to a standard deviation for the avalanche data of 103.3.

Figure 2 shows predicted values obtained with this model, plotted along with actual values of the avalanche activity index for the entire period 1965–73. Good agreement is demonstrated. The model can be used to predict avalanche activity for any winter that resembles the class of eight, defined by the period 1965–73.

Comparison with discriminant analysis

It is difficult to make a direct comparison of time-series predictions with results obtained using a discriminant-analysis approach. In the first case, avalanche activity is treated as a quasi-continuous variable whereas in the second case a probabilistic assignment of avalanche and non-avalanche days is made. Some idea of the relative power of the two methods is however desirable in view of work done by others using discriminant methods. Simulated discriminant results can be obtained from the time-series curves (Fig. 2) by choosing an appropriate level of avalanche activity for discrimination and defining successes as those predictions which fall on the same side of a level as actual values. Any day on which at least one avalanche of medium size occurs on any one avalanche site could be selected as the criterion for discrimination. Examination of the 1965–73 data set indicates that this criterion corresponds to an average avalanche-activity level of 30 for weighting scheme (III) (Table I). Two or three small avalanches usually occur with a medium avalanche on a particular day. Applying this criterion to the actual and predicted time-series curves for 1965–73 (Fig. 2) provides a measure of the performance of the time-series model in a simulated discriminant situation. 197 out of 237 avalanche days were correctly classified and 397 out of 524 non-avalanche days were correct. The misclassification percentages were therefore 17% for avalanche days, 24% for non-avalanche days and 22% overall. If an avalanche activity level of 100 is chosen, instead of 30, as the discriminant criterion, 94 out of 128 avalanche days and 595 out of 633 non-avalanche days are correctly assigned. These figures correspond to misclassification percentages of 27% for avalanche days, 6% for non-avalanche days, and 9%
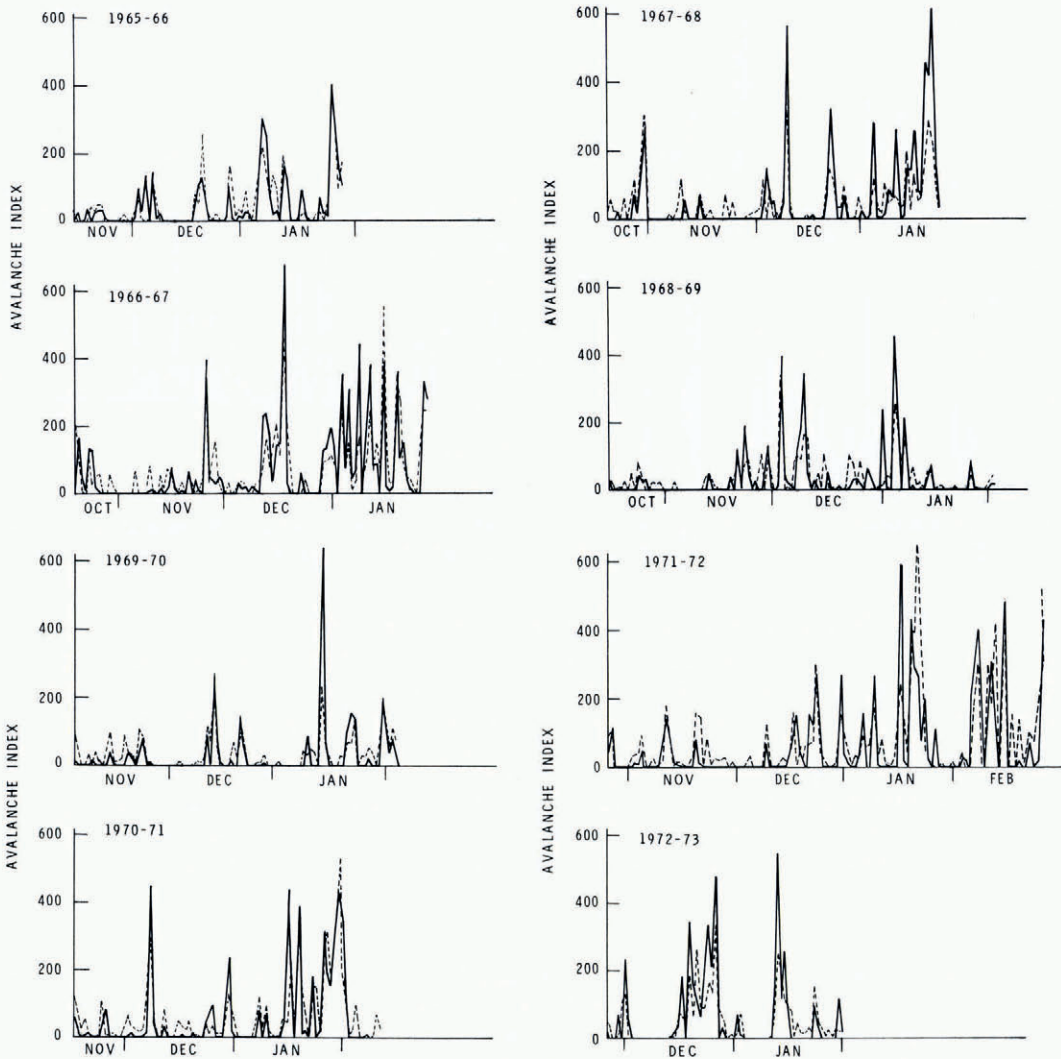
Fig. 2. *Avalanche activity prediction profiles for 1965–73* (——— *actual values,* – – – – *predicted values*) *based on the single model*

$$\hat{Y}_t = 0.095Y_{t-1} + 0.060Y_{t-2} + 0.000\,685SWH_t - 0.000\,094SWH_{t-1} + 0.131WT_{2,\,t} - 0.035WT_{2,\,t-2}.$$

overall. Bovis (1977) stratified avalanche activity into four levels for discrimination but Bois and others ([1975]) used only one. Any number of stratifications, or criteria for discrimination, may be applied to Figure 2. The misclassification percentages obtained can be favourably compared with those of Bois and others (1975) and Bovis (1977).

Simulated forecasts

Simulated forecasts were obtained by applying the model developed from 1965–73 data to three recent winters 1973–76. Results are shown in Figure 3 where it is clear that simulated forecasts closely match actual values of the avalanche index. The *R*-value has dropped only
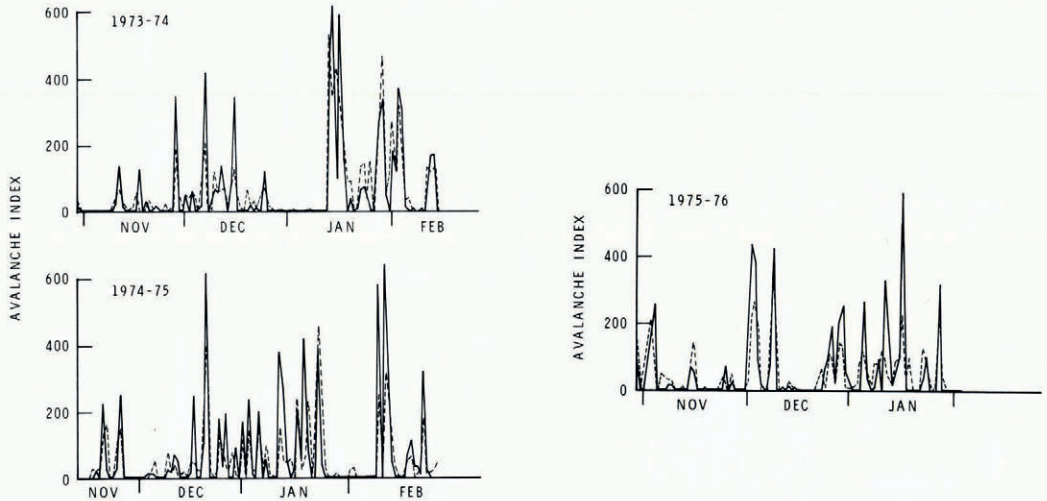
Fig. 3. *Simulated avalanche forecast for 1973–76* (——— *actual values,* – – – – *forecasted values*) *based on the single model*
$\hat{Y}_t = 0.095 Y_{t-1} + 0.060 Y_{t-2} + 0.000\,685 SWH_t - 0.000\,094 SWH_{t-1} + 0.131 WT_{2,\,t} - 0.035 WT_{2,\,t-2}.$

slightly to 0.80 compared with 0.81 for 1965–73. In a simulated discriminant situation as defined previously, using an avalanche activity level of 30 for discrimination, these 1973–76 forecasts yield the following results. 91 out of 102 avalanche days and 166 out of 212 non-avalanche days are correctly assigned. Misclassification percentages are therefore 11% for avalanche days, 22% for non-avalanche days, and 18% overall. Using an activity level of 100 for discrimination, 43 out of 56 avalanche days and 238 out of 258 non-avalanche days are correct giving misclassification percentages of 23% for avalanche days, 8% for non-avalanche days, and 11% overall. This analysis indicates that the model works well when applied to winters not included in the adjustment sample from which the model was developed.

SUMMARY AND CONCLUSIONS

It has been shown that a stochastic transfer-function time-series approach may be a useful alternative to discriminant analysis procedures for forecasting avalanches.

Accordingly, a numerical scheme was devised for representing avalanche activity in the Rogers Pass area of British Columbia in terms of size, terminus, and moisture content codes for each avalanche event. Avalanche activity indices have been computed for the entire area on a daily basis and treated as the dependent variable in various types of correlation analysis designed to establish optimum weights for size, terminus, and moisture content codes, appropriate starting, finishing, and transition dates for avalanche activity periods, and the most suitable meteorological variables to be used in the subsequent development of time-series models. A composite term, SWH (the product of snow-pack depth, precipitation, and relative humidity), was selected as the "best" meteorological variable since it had the highest linear correlation with avalanche activity for the period 1965–73. Highly temperature-dependent early and late-spring avalanching was excluded. After a summary of the theory of time-series model development, a model was presented involving SWH_t as the primary input series, and $WT_{2,\,t}$ (the product of precipitation and minimum air temperature) as a secondary input series. Lagged values of previous avalanche activity Y_{t-1} and Y_{t-2} also appear in the model. The importance of such terms has long been recognized by field analysts who make operational hazard evaluations and forecasts at Rogers Pass. "The results of the avalanche patrol are used

to verify the avalanche-hazard evaluation and to help with the avalanche hazard forecast" (Schleiss and Schleiss, 1970, p. 121). Perla and Martinelli (1976, p. 95) list "current and past avalanches" as the second basic input for stability evaluation and state that "The most precise stability evaluation for a slope must be based on knowledge of recent avalanche activity on nearby slopes". Predicted values obtained with this model are in excellent agreement with actual values of the avalanche activity index for the period 1965–73. The model performed well when applied to three winters 1973–76 in a simulated forecasting situation. A simulated discriminant analysis yielded low misclassification percentages for both the adjustment sample 1965–73 and the simulated forecasting sample 1973–76.

Such models can be improved, not only as further data become available but also when more frequent measurements are made, particularly during storm periods. Avalanche events should be recorded as precisely as possible, for such records are undoubtedly the most important limiting factor in the development of accurate models.

REFERENCES

Bois, P., *and others*. [1975.] Multivariate data analysis as a tool for day-by-day avalanche forecast, [by] P. Bois, C. Obled, and W. Good. [*Union Géodésique et Géophysique Internationale. Association Internationale des Sciences Hydrologiques. Commission des Neiges et Glaces.*] Symposium. *Mécanique de la neige. Actes du colloque de Grindelwald, avril 1974*, p. 391–403. (IAHS–AISH Publication No. 114.)

Bovis, M. J. 1974. Statistical analysis. (*In* Armstrong, R. L., *and others.* Development of methodology for evaluation and prediction of avalanche hazard in the San Juan mountain area of southwestern Colorado, [by] R. L. Armstrong, E. R. LaChapelle, M. J. Bovis, J. D. Ives. *University of Colorado. Institute of Arctic and Alpine Research. Occasional Paper* No. 13, p. 71–94.)

Bovis, M. J. 1976. Statistical analysis. (*In* Armstrong, R. L., *and* Ives, J. D., *ed.* Avalanche release and snow characteristics, San Juan Mountains, Colorado. *University of Colorado. Institute of Arctic and Alpine Research. Occasional Paper* No. 19, p. 83–130.)

Bovis, M. J. 1977. Statistical forecasting of snow avalanches, San Juan Mountains, southern Colorado, U.S.A. *Journal of Glaciology*, Vol. 18, No. 78, p. 87–99.

Box, G. E. P., *and* Jenkins, G. M. 1970. *Time series analysis, forecasting and control.* San Francisco, Holden-Day.

Draper, N. R., *and* Smith, H. 1966. *Applied regression analysis.* New York, John Wiley and Sons, Inc.

Föhn, P., *and others.* 1977. Evaluation and comparison of statistical and conventional methods of forecasting avalanche hazard, by P. Föhn, W. Good, P. Bois, and C. Obled. *Journal of Glaciology*, Vol. 19, No. 81, p. 375–87.

Freese, F. 1964. Linear regression methods for forest research. *U.S. Dept. of Agriculture. Forest Service. Research Paper* FPL 17.

Judson, A., *and* Erickson, B. J. 1973. Predicting avalanche intensity from weather data: a statistical analysis. *U.S. Dept. of Agriculture. Forest Service. Research Paper* RM-112.

LaChapelle, E. R. 1970. Principles of avalanche forecasting. *Canada. National Research Council. Associate Committee on Geotechnical Research. Technical Memorandum* No. 98, p. 106–13.

Mellor, M. 1968. Avalanches. U.S. Cold Regions Research and Engineering Laboratory. *Cold regions science and engineering.* Hanover, N.H., Pt. III, Sect. A3d.

Perla, R. I., *and* Judson, A. 1973. Study plan fading memory analysis of avalanche contributory factors. *U.S. Dept. of Agriculture. Forest Service. Line Project* No. FS-RM 1601.

Perla, R. I., *and* Martinelli, M., *jr.* 1976. Avalanche handbook. *U.S. Dept. of Agriculture. Forest Service. Agriculture Handbook* 489.

Rao, C. R. 1952. *Advanced statistical methods in biometric research*. New York, John Wiley and Sons, Inc.
Salway, A. A. Unpublished. Statistical estimation and prediction of avalanche activity from meteorological data. [Ph.D. thesis, University of British Columbia, 1976.]
Schleiss, V. G., *and* Schleiss, W. E. 1970. Avalanche hazard evaluation and forecast, Rogers Pass, Glacier National Park. *Canada. National Research Council. Associate Committee on Geotechnical Research. Technical Memorandum* No. 98, p. 115–22.
Seligman, G. 1936. *Snow structure and ski fields; being an account of snow and ice in forms met with in nature and a study on avalanches and snowcraft*. London, Macmillan and Co. Ltd.
U.S. Dept. of Agriculture. 1961. Snow avalanches: a handbook of forecasting and control measures. *U.S. Dept. of Agriculture. Forest Service. Agriculture Handbook* 194.