Prévision des avalanches par des méthodes statistiques

Aspects méthodologiques et opérationnels (Application à la région de Davos, Suisse)

PAR

Ph. Bois

ET

Ch. Obled

Institut de Mécanique de Grenoble et Institut National Polytechnique de Grenoble

Introduction

La pratique croissante des sports d'hiver et les accidents survenus au cours des cinq dernières années ont attiré l'attention du public, et des autorités responsables, sur la nécessité d'une prévision du danger d'avalanche, notion qui est d'ailleurs assez mal définie.

A l'échelle locale, cette prévision s'appuie essentiellement sur les connaissances d'observateurs, difficilement remplaçables, dont l'expérience ne semble pas transmissible à d'autres personnes ou transposable en d'autres lieux. A une échelle spatiale plus grande (une région des Alpes par exemple), il semble difficile de synthétiser l'information et d'extraire des données locales les valeurs nécessaires à une prévision régionale ; cela conduit à une prévision généralement trop floue, tant dans l'espace que dans le temps, pour permettre la mise en œuvre de mesures de protection ou de prévention.

Or, si les pratiques de haute montagne (ski de randonnées et alpinisme) s'effectuent sous la seule responsabilité du touriste, il n'en est pas de même pour la fréquentation des domaines skiables desservis par des remontées mécaniques, ou pour les voies publiques (routes et voies ferrées).

Afin que le problème soit bien défini, nous allons considérer un domaine skiable assez représentatif (100 km² dont 160 km de piste et une dizaine de remontées mécaniques), comportant des observations simultanées de diverses variables nivométéorologiques et de l'activité avalancheuse sur une période relativement longue (10 à 15 ans). Notre but sera de déterminer le *potentiel de prévision* contenu dans ces variables, et de définir une méthodologie qui serait, elle, transposable à d'autres zones. Elle pourrait fournir une base objective (car éventuellement reproduc-

tible) de prévision, et être ensuite affinée par le prévisionniste local, comme c'est le cas dans d'autres domaines (prévision de précipitations, par exemple).

Pour des raisons matérielles, notre choix s'est porté sur la zone de la Parsenn, située autour de Davos (Suisse), à proximité immédiate de l'Institut Fédéral Suisse pour l'Etude de la Neige et des Avalanches, qui nous a fait bénéficier de son inappréciable collaboration. (cf. carte - fig. 1).

1. Rétrospective des travaux antérieurs

Les premières tentatives (Ch. Obled, 1970) consistèrent à ajuster des lois de probabilité conditionnelle aux fréquences d'occurrence des avalanches, en fonction d'une ou deux variables. Cet ajustement se faisait non à partir de fréquences cumulées, méthode habituelle en hydrologie, mais à partir des fréquences observées pour chaque classe de la variable : $f_i = r_i/x_i$: nombre de journées avalanches sur nombre de journées totales pour lesquelles la variable appartient à la classe x_i . Ceci nous conduisit à utiliser la technique des probits, décrite par D.J. Finney (1964). La méthode avait été testée sur les grandes avalanches de couloirs de la Haute Maurienne et de la Haute Tarentaise, pour l'hiver 1969-1970. Toutefois, la méthode ne prenait en compte que les journées situées pendant ou immédiatement après une séquence de précipitations.

Malgré le petit nombre de variables prises en compte, les résultats furent assez prometteurs ; la figure 2 présente l'ajustement réalisé du risque d'avalanche en fonction des précipitations cumulées. D'autre part, l'application de cette méthode à l'hiver 1969-70 (cf. fig. 3) montre qu'elle s'applique bien aux avalanches dues à de fortes

1/ Secteur étudié

2/

3/

précipitations mais qu'elle présente des lacunes pour les autres types d'avalanches.

Ce travail avait suscité l'intérêt de l'Institut Fédéral Suisse, qui mettait à notre disposition les données de Davos. Une première étude (Bois et Obled, 1972) avait conduit à l'utilisation de l'analyse discriminante pour calculer, sur une base journalière, la probabilité d'occurrence ou de non occurrence d'avalanche, en fonction d'une vingtaine de variables nivométéorologiques. Cette étude fut reprise (Bois et Obled, 1974) en augmentant le nombre de variables explicatives et surtout en considérant que les journées avalancheuses ne constituaient pas un groupe homogène mais se décomposaient en journées à avalanches de neige sèche et journées à avalanches de neige humide. Des algorithmes de sélection permirent de ne retenir que les variables permettant de séparer, avec le plus de succès, les journées selon leur groupe (avalanches de neige sèche, avalanches de neige humide, journées non avalancheuses). Les variations saisonnières des variables explicatives nous obligeaient à considérer chaque mois séparément. Les modèles décisionnels obtenus avaient été ajustés sur 1960-70, testés sur 1971 et 1972 et utilisés opérationnellement en 1974 et 1975. On en trouvera des exemples (fig. 4 et 5). Toutefois, rappelons que, contrairement aux premiers essais, nous ne nous limitons plus seulement aux grands phénomènes (avalanches de couloirs) mais à toutes les manifestations de la zone pour tous les types de conditions rencontrées. Malheureusement la faiblesse des échantillons utilisés pour l'ajuste-

4/

5/

ment compromettait la robustesse des modèles quand on les appliquait à d'autres périodes (cf. sur les fig. 4 et 5, les divergences entre 2 modèles pourtant ajustés sur les mêmes données).

Conclusions des essais antérieurs

Les connaissances actuelles sur la physique des mécanismes de déclenchement montrent que la dichotomie : Avalanches/Non avalanche, est insuffisante et qu'il faut distinguer plusieurs types de phénomènes avalancheux.

Ceci provient de l'existence de *seuils*, et du fait que l'on mélange sous la même manifestation extérieure, l'avalanche, divers phénomènes. Par exemple, il est évident qu'une variation de température de 3°C entre -21° et -18° n'a pas le même effet qu'entre -3,5° et -0,5° (non linéarité) et ne correspond plus du tout au même processus si on passe de -1° à +2°C (Seuil de 0°C). Une première distinction entre avalanches de "neige sèche" et avalanches de "neige humide" apporte déjà beaucoup mais ce codage est entaché d'assez fréquentes incertitudes d'observations.

D'autre part, il ne nous semble plus justifié d'opposer, à diverses populations de journées avalancheuses, une seule population de journées dites normales, qui constituerait une population homogène. En conséquence, nous avons repris complètement notre méthodologie pour tenter de

nous rapprocher encore plus de la réalité physique et pour proposer une démarche qui n'est, peut être, que la version formalisée et quantifiée de celle qu'emploie, empiriquement, le prévisionniste expérimenté.

2. Données brutes et variables élaborées

a) Nous donnons dans le tableau (1) la liste des données brutes, c'est-à-dire correspondant à des mesures, effectuées sur une base journalière, et disponibles sur les hivers 1960-61 à 1973-74 inclus.

Variables brutes mesurées	Tableau 1
Variables "brutes"	
1) Valeur en eau de la neige fraîche recueillie à 08 H à la planche à neige, en mm	
2) Hauteur de neige fraîche recueillie à 08 H en cm	
3) Intensité moyenne de la chute entre le jour j à 08 H et j + 1 à 08 H quand elle dépasse 1 mm/h	
4) Vent maximum (ou max. des 3 observations) du jour j en m/s	
5) Rayonnement solaire global incident en cal/cm ² /jour	
6) Température le jour j à 7 H 30	
7) Température le jour j à 13 H 30	
8) Nombre d'heures d'insolation du jour j	
9) Nébulosité moyenne du jour j en 1/10	
10) Hauteur de neige au sol le jour j vers 8 H en cm	
11) Enfoncement de la sonde dans la couche superficielle en cm	
12) Nombre de cas de chasse-neige observés au cours des observations du jour j	
13) Quantité d'eau recueillie au pluviomètre le jour j à 08 H	
14) Température à 10 cm sous la surface de la neige le matin vers 7 H 30	

De ces données brutes l'on a déduit un nombre important de variables explicatives ou présumées telles. L'étude des phénomènes et les avis parfois divergents des experts nous ont incités à être le plus exhaustifs possible, ce qui nous a conduit à ce nombre de 50 variables, parfois redondantes. Nous donnons de ces variables une description simplifiée dans l'Annexe I, en conseillant aux lecteurs intéressés de se reporter aux rapports correspondants.

Nous pouvons constater que la situation proprement météorologique est représentée par de nombreuses variables ; les effets thermiques sont bien pris en compte. Par contre, les effets mécaniques de la situation météorologique sont plus difficiles à quantifier, notamment en ce qui concerne le transport de la neige par le vent.

Quant à la description de l'état du manteau neigeux, elle est, faute de données systématiques, assez simpliste et il ne faut considérer les variables qui tentent de la représenter que comme des index de stabilité ou d'instabilité.

b) Les données d'avalanche comportent, pour chaque avalanche observée dans la région envisagée, un code des-

criptif et la date présumée et/ou réellement observée du déclenchement. On comprend, vu les conditions d'observations (particulièrement par mauvais temps) que de nombreuses incertitudes demeurent sur les dates (0 à + 2 jours). C'est la raison pour laquelle nous avons choisi de travailler avec des variables journalières, alors que certaines données étaient disponibles sur un pas de temps plus fin. C'est aussi la raison pour laquelle nous avons choisi de caractériser l'activité avalancheuse par les journées où *au moins* une avalanche avait été répertoriée.

c) Enfin, il est exclu de travailler sur l'ensemble de la période hivernale (Novembre à Juin) considérée comme une seule période homogène. En effet, de nombreuses variables sont affectées de variations saisonnières dont il faut tenir compte : une température minimale de - 9° est plutôt élevée en Janvier mais plutôt basse en Avril, une durée d'insolation de 8 h caractérise un beau temps clair en Décembre ou Janvier, mais une journée partiellement nuageuse en Mars.

Ces exemples, et beaucoup d'autres, nous avaient conduit (Bois et Obled, 1974) à effectuer un traitement mois par mois (par exemple, en mettant bout à bout les 10 mois de Janvier de 1961 à 1970). Malheureusement, le nombre de journées avalancheuses au sein d'une telle série de 300 jours n'excède guère 15 à 20 %, soit, pour les mois les mieux fournis 50 à 60 journées avalancheuses en 10 ans. Il était difficile alors de scinder les phénomènes avalancheux en plus de 2 groupes (par exemple : "neige sèche" et "neige humide"), sans tomber dans des échantillons à effectifs faibles.

Pour ces raisons, nous avons rassemblé les données en bimestres, regroupant par exemple le mois de :

Janvier-Février (JAN-FE)	sur 12 ans	711 journées dont 122 avalanches
Mars-Avril (MAR-AV)	sur 12 ans	732 journées dont 123 avalanches

Dans une optique de *recherche*, nous négligerons Novembre-Décembre, faute d'une définition précise du "début de l'hiver", et Mai-Juin, pour des raisons équivalentes sur la "fin de l'hiver".

Ces 12 années couvrent la période 1961-72, les hivers 1972-73 et 1973-74 étant mis de côté pour servir de tests aux modèles ajustés sur 1961-72.

3. Premiers traitements et visualisation

Avant d'en arriver à la méthodologie finalement retenue, un grand nombre d'essais ont été réalisés, essentiellement en vue de visualiser les phénomènes et leur liaison avec les variables proposées. Rappelons que nous disposons d'un ensemble de données X_{oj} , constitué de N observations de 50 variables :

Soit pour le jour j $\rightarrow X_{jl} \quad (l = 1 \dots 50)$,

le jour j étant avalancheux ou non.

3.1. Visualisation des variables

Il s'agissait d'abord de voir s'il existait des liaisons entre certaines variables et les phénomènes avalancheux. Une

première technique consiste à comparer la distribution d'une variable X_l au sein d'un échantillon de journées sans avalanches et sa distribution au sein d'échantillons de journées avalancheuses (cf. fig. 6). L'écart entre les distributions peut être mesuré par un test de X^2 qui donne une idée de la liaison entre le phénomène avalancheux et la variable.

Une autre technique, beaucoup plus sophistiquée mais qui n'est en fait que la version multidimensionnelle de la précédente, consiste à appliquer l'analyse des correspondances (cf. Lebart et Fenelon, 1971) à l'ensemble des variables X_l , plus une, X_{l+1} , qui représentera le phénomène avalanche, en nombre d'occurrences observées le jour j .

En fait, l'analyse des correspondances s'applique théoriquement à des tableaux de contingence et pose quelques problèmes quand il s'agit de tableaux de mesures de variables continues. Pour tourner cet inconvénient, Nakache

(1972) a proposé de coder les données en classes, c'est-à-dire que :

$$X_{jl} \rightarrow x_{jlk} \quad k = 1, \dots, K$$

et

$$x_{jlk} = \begin{cases} 1 & \text{si } X_{jl} \in \text{classe } C_k \text{ de la variable } l \text{ le jour } j \\ 0 & \text{sinon} \end{cases}$$

On a donc un tableau de $K.L$ variables ne contenant que des 0 et des 1. L'analyse factorielle des correspondances fournit alors une visualisation, en général dans les 2 premiers axes factoriels, qui ne fait plus correspondre à la variable un seul point, mais K points.

Dans notre cas nous nous sommes limités à $K = 3$ classes.

Ex : Variable température max

$$C_1 = [< -10^\circ] \quad C_2 = [-10, -4^\circ\text{C}] \quad C_3 = [> -4^\circ\text{C}]$$

①
Température minimum du jour j
Minimum temperatur of day j

②
Cumul des degrés jours positifs
sur les 5 derniers jours.
*Total of positive degree-days
of the last five days*

③
Hauteur de neige fraîche disponible
Available fresh snow depth

- ④
Groupe 1	{	Avalanches de neige "sèche"
		<i>Dry snow avalanches</i>
" 2	{	Sans avalanche
		<i>No avalanche</i>
" 3	{	Avalanches de neige "humide"
		<i>Wet snow avalanches</i>

6/ Exemples d'histogrammes

Variable précipitation

$C_1 = [0 \text{ mm}]$ $C_2 = [.5, 15 \text{ mm}]$ $C_3 = [> 15 \text{ mm}]$

Variable avalanche

$C_1 = [0]$ $C_2 = [1 \text{ ou } 2]$ $C_3 = [\geq 3]$

Pour des raisons théoriques il est souhaitable que les classes soient les plus équiprobables possible, ce qui n'est pas réalisable pour certaines variables.

Dans le plan factoriel $F_1 - F_2$, on a alors pour chaque variable $X_1, X_2, \dots, X_L$ etc., une trajectoire de 3 points, et la proximité des trajectoires de 2 variables indique la liaison statistique entre les variables.

(ex : les 2 variables sont fortes, moyennes, ou faibles en même temps sur l'ensemble des N observations étudiées).

On donne un exemple sur les mois de Janvier et Mars (fig. 7 et 8). Sur ce dernier, on voit que la décomposition des avalanches en 2 groupes ("neige sèche" et "neige humide") apporte beaucoup et que, par exemple, la variable "avalanche de neige sèche" a une trajectoire voisine de celle des précipitations, tandis que l'"avalanche de neige humide" est très proche des températures positives cumulées.

Le fait de mettre en évidence la liaison, classe par classe, entre 2 variables, se rapproche d'une corrélation (de rang) et a fait dénommer cette technique régression par "analyse en correspondance" (Cazes, 1974).

3.2. Visualisation des observations

Nous cherchons ici à visualiser des points qui correspondent à chaque journée j , caractérisée par 50 variables, donc situées dans $\mathcal{R}^{50}$. Ceci étant impossible, l'analyse

factorielle cherche quel est le plan qui, dans $\mathcal{R}^2$ donc, donne de ce nuage de points l'image la plus ressemblante. La technique utilisée, dite méthode des composantes principales (Cooley et Lohnes, 1971, Lebart et Fenelon, 1971) consiste à chercher la combinaison linéaire (resp. la direction de $\mathcal{R}^{50}$).

$$Z_1 = u_{11} \cdot X_1 + u_{12} \cdot X_2 + \dots + u_{1L} \cdot X_L$$

qui minimise la variance de Z_1 (resp. sur laquelle la somme des carrés des projections des distances des N points 2 à 2 est maximum).

On cherche ensuite Z_2 , non corrélée à Z_1 (ou encore "orthogonale" à Z_1) etc. Dans le cas où l'on utilise des variables centrées réduites, on démontre que ces nouveaux axes-variables sont les vecteurs propres, associés aux valeurs propres décroissantes $\lambda_1, \lambda_2, \dots$ de la matrice de corrélation R des variables sur les N observations. L'information transportée par chacune de ces variables est égale au rapport de la valeur propre λ_i , liée à Z_i , sur la trace de la matrice de corrélation. On considère en général que les variables associées à des valeurs propres < 1 ne sont plus significatives.

Dans notre cas, tant sur les fichiers JAN-FE que MAR-AV, on constate qu'il y a environ 15 à 20 variables indépendantes parmi les 50 proposées. Pour les besoins de la représentation, on ne considère que les 4 ou 5 premières dont la répartition est énoncée par le tableau suivant (Tab. 2).

On donne également les variables initiales qui interviennent le plus, (ou qui sont corrélées avec) ces nouvelles variables Z encore appelées "composantes principales" (Tab. 2)

Nous allons maintenant en donner quelques exemples. Rappelons simplement que, dans ce traitement, la variable avalanche n'est utilisée que de façon exogène (c'est-à-dire

7/ Analyse des correspondances

8/ Analyse des correspondances

qu'elle n'est pas intervenue dans le traitement) : on a donc représenté chaque journée par un point, puis repéré le point différemment selon qu'il y avait eu ou non avalanche.

Tableau 2			
Janvier-Février		Mars-Avril	
N° de la variable	Carré du coefficient de corrélation multiple avec les axes $F_1 - F_2$	N° de la variable	Carré du coefficient de corrélation multiple avec les axes $F_1 - F_2$
1	• 58	1	• 55
2	• 61	2	• 60
5	• 63	5	• 46
9	• 47	12	• 72
12	• 42	14	• 56
14	• 43	15	• 53
16	• 44	16	• 69
18	• 63	18	• 62
19	• 55	20	• 55
20	• 56	21	• 55
22	• 62	22	• 54
24	• 51	35	• 50
31	• 44	40	• 58
38	• 59	41	• 57
42	• 44	42	• 61
$\lambda_1 = 9.5$ $\lambda_2 = 5.5$ } 30 %		$\lambda_1 = 11.5$ $\lambda_2 = 4.8$ } 33 %	

Les figures 9 à 12 montrent que dans le plan des deux premiers axes, les journées avalancheuses se séparent assez bien selon le type d'avalanches ; d'autre part, alors que les journées non avalancheuses sont groupées vers le centre (fig. 10 et 12), les journées avalancheuses sont réparties différemment (fig. 9 et 11). Il n'y a donc pas une discrimination franche, mais plutôt des différences de densité.

On a aussi visualisé les journées avalancheuses en les distinguant selon les valeurs des précipitations (nulles ou supérieures à 30 mm) : cf figure 13.

On a aussi visualisé les journées (fig. 13) en les distinguant selon les valeurs des précipitations reçues (beaucoup, peu ou pas).

3.3. Représentation temporelle – Trajectoires des journées avalancheuses

Nous avons voulu voir si le phénomène avalanche était lié à une "accumulation" sur certaines variables (mauvais temps prolongé, temps chaud prolongé) ou plutôt à une "évolution" rapide, à une variation brutale des variables. Pour cela, nous avons représenté les journées avalancheuses, et les 3 jours qui les précédaient par une ligne brisée dans le plan $F_1 - F_2$ (fig. 14).

Il est assez surprenant de constater que ces "trajectoires" ne sont pas quelconques mais présentent une allure assez typique, physiquement interprétable. Le phénomène n'est donc pas seulement fonction d'une situation (position du point j) mais aussi d'une évolution (position des points $j-1, j-2 \dots$).

Pour tâcher de mettre ceci en évidence de façon plus quantifiée, nous avons utilisé le schéma A suivant :

1) diviser le plan $F_1 - F_2$ en 8 secteurs et une zone centrale incertaine

9/ Jours avalancheux (Janvier-Février)

10/ Jours non avalancheux (Janvier-Février)

11/ Jours avalancheux (Mars-Avril)

12/ Jours non avalancheux (Mars-Avril)

13/ Jours avalancheux (Mars-Avril)

14/

2) construire les matrices de passage d'un secteur à l'autre, d'une part pour les points non avalancheux, et d'autre part pour les points avalancheux, en ajoutant ± 1 dans la case (r, s) si le jour j était dans le quadrant s sachant que $j - 2$ était dans le quadrant r .

La part du phénomène expliquée par la situation est appréciée par la valeur des termes diagonaux, tandis que la part d'évolution peut se mesurer à la valeur des termes non diagonaux.

On constate et on peut le vérifier à l'aide d'un test assez complexe (cf. Gordon Hilton - 1972) que ces matrices de passage diffèrent sensiblement pour les cas avalancheux et non avalancheux (Tab. 3).

Si l'on se place, à titre d'exemple, dans les axes $F_1 - F_2$ de mars-avril qui peuvent s'interpréter à l'aide du tableau 3, on constate :

- au vu des fréquences marginales le jour j , que les journées avalanches se concentrent dans les secteurs 3, 4 et 7, 8 alors que les journées normales sont mieux réparties entre les différents secteurs. En particulier le secteur 9 s'avère peu avalancheux.

- que les journées avalanches du secteur :

- 3 proviennent de 2 et surtout 3
- 4 proviennent de 4 et surtout 3
- 5 proviennent de 4 et surtout 3.

Ce secteur 3 étant celui des précipitations récentes, cela veut dire qu'un certain nombre d'avalanches nécessite d'abord des précipitations, sur lesquelles un autre effet (refroidissement, vent, réchauffement) s'exercera.

- que les journées avalanches des secteurs 7 et 8 se caractérisent par un phénomène d'"accumulation" puisque la plupart étaient déjà en 7 ou 8 l'avant-veille de l'avalanche.

- que, tant pour les journées avalanches que pour les autres, il y a plus de points au-dessus et à droite de la diagonale qu'en dessous. Ceci met en évidence une caractéristique climatologique intéressante. Les jours successifs ont tendance à suivre une trajectoire dominante dans le

sens trigonométrique et qui peut se résumer, pour mars-avril, à :

Celle-ci est encore plus nette en janvier-février.

3.4. Conclusions préliminaires

Nous n'avons malheureusement pas eu les moyens d'exploiter cet aspect "dynamique" du phénomène avalancheux, considéré comme le résultat d'une évolution ou d'une séquence de situations relativement précise.

Si nous nous limitons à l'aspect "statique", nous constatons que :

1) les situations avalanches, prises dans leur ensemble, ne se distinguent guère des situations normales en ce sens que, pour une situation avalancheuse donnée, on peut trouver des situations voisines non avalanches ;

2) par contre, la densité de probabilité du phénomène avalancheux varie sensiblement selon les régions du plan

Secteur auquel appartient la journée avalancheuse j										Tableau 3	
Avant-veille d'une journée avalancheuse		2	3	4	5	6	7	8	9	Total	%
	2	1	2	0	0	0	1	1	0	5	5.5
	3	1	8	9	4	2	1	0	0	25	27.5
	4	0	1	3	2	3	1	0	0	10	11.0
	5	0	3	2	0	0	1	0	0	6	6.6
	6	1	3	0	0	2	4	1	0	11	12.1
	7	2	0	0	0	0	10	6	1	19	20.9
	8	0	0	0	0	0	1	12	0	13	14.3
	9	0	0	0	0	0	0	2	0	2	2.2
Total		5	17	14	6	7	19	22	1	91	
%		5.5	18.7	15.4	6.6	7.7	20.9	24.2	1.1		100.0

Secteur auquel appartient la journée quelconque j											
Avant-veille d'une journée quelconque		2	3	4	5	6	7	8	9	Total	%
	2	6	7	3	1	0	4	7	6	34	7.6
	3	6	25	24	8	12	8	2	1	86	19.1
	4	1	9	17	25	7	2	1	0	62	13.8
	5	1	9	8	8	12	4	1	0	43	9.6
	6	5	6	2	2	11	9	1	0	36	8.0
	7	6	5	0	0	3	29	18	8	69	15.3
	8	12	4	0	0	1	7	39	20	83	18.4
	9	7	9	0	0	0	5	8	8	37	8.2
Total		44	74	54	44	46	68	77	43	450	
%		9.8	16.4	12.0	9.8	10.2	15.1	17.1	9.6		100.0

Note — On a éliminé les points du domaine (1) pour lesquels l'appartenance aux quadrants n'est pas significative.

$F_1 - F_2$, ce qui incite à penser qu'il y a plusieurs phénomènes différents, d'où l'idée de procéder à une classification des avalanches.

4. Classification par types de temps avalancheux

Le problème de la classification des avalanches a été abordé, depuis déjà longtemps, par des physiciens soucieux de décrire toutes les formes et variantes possibles du phénomène, et avec des idées préalables sur les mécanismes de déclenchement (accumulation, transport par le vent, réchauffement...). Cette démarche présente pour notre but plusieurs inconvénients :

1) Le souci de tout décrire conduit à créer des classes ou types d'avalanches qui ne sont que peu, ou pas du tout, observés en un lieu donné

2) Les mécanismes présumés sont parfois difficilement vérifiables, et de plus l'expérience n'est pas reproductible

3) L'information, c'est-à-dire les données dont nous disposons, peut être insuffisante pour distinguer entre 2 types pourtant bien définis par ailleurs.

C'est pourquoi nous allons adopter une démarche pratiquement inverse. Dans un but de prévision, il est apparu difficile de considérer comme un tout homogène l'ensemble des phénomènes avalancheux qui se sont manifestés dans la zone particulière que nous étudions. D'autre part, nous ne disposons pour les décrire que d'un ensemble limité de données, insuffisant pour rendre compte de l'ensemble des phénomènes, particulièrement en ce qui concerne les caractéristiques de résistance du manteau.

Nous allons donc demander, aux seules données dont nous disposons, de mettre en évidence, si elle existe, une certaine classification au sein des journées avalancheuses observées. Cette démarche est statistique en cela que seuls les groupements importants apparaîtront — il est exclu de considérer comme une classe un cas isolé, même fort intéressant — et que cette classification sera propre à la région observée et aux données utilisées.

Parmi les classifications "déterministes" proposées antérieurement, 2 voies se dégagent. L'une est dite "génétique" : elle classe les phénomènes en fonction de leurs causes connues ou présumées ; l'autre, dite "morphologique" regroupe les manifestations selon leur aspect extérieur observé a posteriori (avalanche de fond ou de surface, départ ponctuel ou fracturé, etc.). C'est évidemment de la première que nous chercherons à nous rapprocher.

4.1. Recherche des formes fortes

Il est inutile de présenter ici les tâtonnements qui nous ont conduit à choisir la technique d'agrégation dite "méthode des nuées dynamiques" (Diday 1974).

a) L'ensemble à classer est constitué par l'ensemble des journées *avalancheuses* du bimestre considéré, caractérisées par un certain nombre de variables.

Pour des raisons exposées en 3-2, nous n'avons pas pris les X_l ($l = 1$ à 50) mais les variables orthogonales Z_l ($l = 1 \dots 17$).

Toutefois, on a vu que la variance de ces variables était :

$$\text{var}(Z_l) = \lambda_l \text{ valeur propre associée à l'axe } l$$

On risquait donc de donner aussi à l'information Z_m un poids λ_m différent de celui de Z_l , et dans un rapport parfois important (1 à 15). Toutefois un doute demeure quant à la justification de ces pondérations. Elles peuvent provenir soit de la simple redondance de certaines variables (encore appelé effets de taille) et doivent alors être éliminées : par exemple, en travaillant sur les variables z_l , centrées et réduites :

$$z_l = \frac{Z_l}{\sqrt{\lambda_l}} \quad (l = 1 \dots 17)$$

Mais ces pondérations peuvent avoir un sens : par exemple, celui de dire que Z_1 associée à $\lambda_1 = 15$ est une information bien représentée dans notre ensemble alors que Z_{17} associée à $\lambda_{17} = 0.8$ l'est beaucoup moins.

Les deux points de vue étant également défendables, nous avons effectué les traitements dans les 2 cas et gardé seulement les résultats communs aux deux traitements.

b) Algorithme des "nuées dynamiques"

Nous en donnerons un organigramme simplifié qui met en évidence les étapes principales du calcul.

1) Tirage au hasard de m étalons par classe

2) (i) Affectation des journées restantes (soit $N - km$) aux "noyaux" provenant du tirage précédent (1 ou $i - 1$).

3) (i) Test sur la stabilité de la solution obtenue par rapport à l'affectation $2^{(i-1)}$ précédente. Si oui on va en 5, sinon :

4) (i) On tire au sein de chacun des k groupes formés, k nouveaux noyaux de m individus et on retourne en $2^{(i+1)}$

5) On compare le résultat obtenu aux classements précédents. Les ensembles de points qui restent groupés ensemble dans un groupe donné constituent une *forme forte*.

6) On retourne en 1.

On constatera que cette méthode requiert divers paramètres :

a) le nombre d'initialisations aléatoires NI souhaitées. Nous nous sommes limités à 10, mais il faut bien voir que le nombre de formes fortes augmente, tandis que leur contenu s'amenuise, quand ce nombre NI augmente.

b) le nombre $m = NE$ d'étalons dans chaque noyau, et

c) le nombre $k = KG$ de noyaux, donc de groupes demandés. Nous avons en général utilisé 7 ou 10 étalons pour définir le noyau. Quant au nombre de groupes, on s'aperçoit que cette contrainte n'est pas très forte car si on en demande trop, certains sont vides. On a ainsi conclu à l'existence de

- { 3 ou 4 groupes pour JAN-FE
- { 4 ou 5 groupes pour MAR-AV

En fait, ce qui nous intéressait le plus n'était pas tant les groupes formés que les formes fortes obtenues, soit :

JAN-FE	<i>FF1</i>	<i>FF2</i>	<i>FF3</i>	<i>FF4</i>	
nb d'individus	15	15	13	17	
MAR-AV	<i>FF1</i>	<i>FF2</i>	<i>FF3</i>	<i>FF4</i>	<i>FF5</i>
nb d'individus	15(12)	10(6)	14(8)	15(10)	16(9)

Dans ces formes fortes apparaissent souvent des individus correspondant à des journées successives très dépendantes, c'est pourquoi nous avons indiqué entre parenthèses, le nombre d'épisodes réellement indépendants.

4.2. Constitution des groupes

En ne considérant toujours que les journées avalanches, nous avons utilisé ces formes fortes comme échantillons de départ pour les analyses discriminantes suivantes :

- Analyse discriminante *linéaire* :

groupes de départ : FF_i

individus à classer : tous ceux ne figurant pas dans les FF_i

variables utilisées : Z_i (1 à 20)

On constituait alors des groupes :

$$FF_i + \text{individus affectés à } FF_i = AG_i$$

qui étaient repris par :

- Analyse discriminante *linéaire* :

groupes de départ : AG_i

variables utilisées : X_i ($i = 1, 50$)

recherche des individus mal affectés. Détermination pas à pas des variables qui discriminent le mieux

et :

- Analyse discriminante *quadratique* :

L'analyse précédente repose sur l'hypothèse que les variables X_i ont même matrice de covariance au sein de chaque groupe AG_i . La visualisation de ces groupes dans les deux premiers axes discriminants montre qu'il n'est rien. Il est alors préférable (Cacoullos T., 1973) d'en tenir compte en calculant la matrice de covariance de chaque groupe pris séparément. La technique utilisée (Romedor, 1973), qui effectue une sélection pas à pas, fournit un autre modèle et une autre détection d'individus mal affectés. Ceux-ci étant relativement rares, on arrive à 4 groupes pour les avalanches de JAN-FEV (resp. 5 pour MAR-AV).

La figure 15 représente la projection des journées avalanches de janvier et février dans le plan des deux pre-

15/ Avalanches de Janvier-Février, 4 groupes

miers axes discriminants ; on a représenté sur ces figures les ellipses d'inertie de chaque groupe (ces ellipses contiendraient 80 % des points si la distribution des points était multinormale au sein de chaque groupe). La figure 16 concerne les avalanches de mars et avril.

Notons qu'il aurait été souhaitable d'utiliser d'emblée une méthode de "nuées dynamiques" quadratique mais nous ne disposons malheureusement pas du programme correspondant.

4.3. Comparaison des groupes formés

a) Aspects théoriques dans le cas de plusieurs groupes

Il est tentant d'essayer d'interpréter les groupes formés vis à vis des variables sélectionnées. Toutefois, cette interprétation est très difficile dans le cas de plusieurs groupes. En effet, on démontre (Lebart et Fenelon, 1973) que l'analyse discriminante est une forme d'analyse canonique non centrée entre les variables "explicatives" X_i et un ensemble de variables booléennes qui représentent l'appartenance aux groupes, soit $YG_1 \dots YG_K$ avec,

$$YG_k = \begin{cases} 0 & \text{si l'individu } i \notin \text{groupe } G_k \\ 1 & \text{si l'individu } i \in \text{groupe } G_k \end{cases}$$

On construit alors les 2 variables canoniques :

$$ZY_k \text{ combinaison linéaire des } YG_k$$

$$ZX_i \text{ combinaison linéaire des } X_i$$

qui sont le mieux corrélées, puis les 2 suivantes, respectivement orthogonales aux premières, etc. On peut donc interpréter, comme en composantes principales, le premier axe discriminant ZX_1 , puis le second ZX_2 , etc. On représente d'ailleurs les individus dans ces deux axes, qui maximisent la liaison entre les variables d'appartenance aux groupes et les variables explicatives, donc la discrimination.

En d'autres termes, l'axe u_1 , correspondant à la variable ZX_1 , combinaison linéaire des variables X_j , est tel que le rapport de la variance due aux différences entre classes sur la variance totale le long de cet axe est maximal. Cet axe u_1 est d'ailleurs le 1er vecteur propre de la matrice $W^{-1} \cdot B$ (avec les notations usuelles) associé à la valeur propre μ_1 .

On appelle
$$\lambda_1 = \frac{\mu_1}{1 + \mu_1}$$

le pouvoir discriminant de cet axe (compris entre 0 et 1) et on peut de façon analogue à l'analyse en composantes principales, analyser le poids des variables sur cet axe,

16/ Avalanches de Mars-Avril, 5 groupes

par exemple en donnant le coefficient de corrélation *totale* entre l'axe et chacune des variables. Malheureusement cela ne nous renseigne toujours pas sur les effets de corrélation *partielle* dus aux liaisons entre variables.

On fera de même pour les axes suivants sachant que pour K groupes on a, au plus, $K - 1$ axes significatifs.

Toutefois, si l'on a utilisé 10 variables pour séparer 5 groupes, il est difficile aussi d'interpréter le rôle *global* de la seule variable X_j .

b) Cas de deux groupes

Dans ce cas, au lieu de se ramener à une analyse canonique à deux variables d'appartenance YG_1 et YG_2 , et un seul axe discriminant, on montre qu'on peut se ramener à une corrélation multiple entre les X_i et une seule variable Y , qui vaut $+\sqrt{\frac{n_2}{n_1}}$ si l'individu $i \in G_1$ et $-\sqrt{\frac{n_1}{n_2}}$ si $i \in G_2$. Dans ce cas le vecteur des coefficients de régression se déduit des coefficients des deux fonctions discriminantes.

Si le coefficient de la 1ère variable introduite X_i

est C_{11} sur la 1ère fonction

C_{12} sur la 2ème fonction

et si $C_{11} > C_{12}$ alors l'augmentation de X_i entraîne l'affectation à G_1 et inversement. On en conclut que ce groupe G_1 se caractérise par les valeurs élevées de X_i et inversement pour G_2 .

Par contre, l'interprétation de la 2ème variable X_k , analogue si le coefficient de corrélation r_{ik} est faible, se complique quand les deux variables sont corrélées, par exemple

$$r_{ik} > 0.3$$

car il faut alors interpréter des corrélations *partielles*.

c) Applications au bimestre JAN-FE

Nombre de groupes : 4

Valeurs propres : 5.10 2.17 1.51

Pouvoir discriminant : .836 .684 .602

L'axe 1 sépare essentiellement les groupes ①, ②, ③) et ④ avec les variables :

1, 14, 24, 31 et 42

qui représentent les précipitations récentes et les basses températures et jouent un rôle positif sur cet axe, tandis que les variables :

13, 18, 25, 32 et 45

représentent les remontées de températures, l'ensoleillement, le tassement de la neige fraîche et jouent un rôle négatif.

L'axe 2 sépare surtout (①, ②, ④) et ③. Les variables importantes sont :

14, 25 et 42

représentant les basses températures à 13 H, le nombre de jours sans précipitations depuis la dernière séquence de précipitation, et les basses températures dans la neige.

Cela nous permet donc de définir les groupes :

AG ① : Très mauvais temps. Beaucoup de précipitations et basses températures, fortes chutes de neige, assez brutales (30 à 60 cm/jour)

AG ② : Mauvais temps prolongé avec chutes de neige faibles mais répétées, une température de l'air moyenne pour le bimestre ou se réchauffant, et une neige relativement chaude

AG ③ : Beau temps froid (tant pour l'air que pour la neige) plutôt prolongé. Influence très nette du transport par le vent et du chasse neige sur la neige fraîche provenant des précipitations antérieures

AG ④ : Assez beau temps, plutôt chaud pour le bimestre avec un ensoleillement et des températures provoquant le tassement de la neige fraîche récente.

On en conclut que le groupe ① est d'origine mécanique (forte surcharge) et le groupe ④ d'origine thermique (réchauffement après précipitations). Le groupe ③ est plutôt mécanique (transport par le vent de la neige froide) et le groupe ② assez indéfinissable.

Une autre remarque consiste à dire que les avalanches ont lieu, soit pendant les séquences de précipitations qui peuvent être brutales ① ou prolongées ②, soit peu après, au retour d'un temps variable à beau qui peut être chaud et peu durable ④ ou au contraire froid et stable ③.

Remarque : Une rapide analyse des températures à 13 H pour les 2 jours qui suivent une séquence, par rapport à la température à 13 H du dernier jour de la séquence, montre que l'on a des probabilités comparables pour le refroidissement (40 %) et le réchauffement (35 %), un peu plus faible par contre pour des températures voisines ou inchangées (25 %).

d) Cas du bimestre MAR-AV

Nombre de groupes : 5

Valeurs propres : 4.85 2.77 1.31 0.86

Pouvoir discriminant : .829 .735 .568 .463

L'axe 1 sépare les groupes :

①, ②, ③, ④, ⑤

à l'aide des variables :

1, 2, 13

dans le sens négatif représentant les précipitations récentes, les basses températures et les refroidissements, tandis que les variables

4, 12, 15, 19

représentant la densité superficielle, les températures à 7 H et 13 H et la radiation incidente, jouent dans le sens positif.

L'axe 2 sépare les groupes (①, ⑤) de (②, ③) et ④. Les variables qui jouent dans le sens négatif sont

1, 6, 15 et 19

qui représentent les précipitations, le vent, les températures max. élevées et s'opposent au chasse-neige. Son interprétation est rendue délicate du fait de la liaison entre certaines des variables qui y participent.

On a donc plutôt essayé de comparer les groupes 2 à 2, ou en les regroupant : les 2 groupes les plus caractéristiques sont ① et ⑤ mais on peut être tenté de grouper

②, ③ avec ④ ou ① avec ②, ③. Ceci nous mène aux conclusions suivantes :

AG① : Mauvais temps avec fortes précipitations, basses températures et chasse-neige

AG② et AG③ : Ces 2 groupes sont assez voisins et : caractérisés par un temps mauvais ou variable, de la neige fraîche disponible au sol, et un vent prolongé avec chasse-neige. Ce type de temps suit des séquences de précipitations (appartenant à AG①) n'ayant pas donné de fortes intensités journalières. Le groupe ② se caractérise par des effets de chasse neige et de transport de la neige fraîche tandis que le groupe ③ est sensible à des effets de réchauffement de l'air, ensoleillement, refroidissement de la neige. Ces 2 groupes correspondent à un type de temps différent de ceux de Janvier-Février, en quelque sorte plus printanier, correspondant à un temps prolongé avec alternance de faibles chutes de neige et d'éclaircies.

AG④ : Beau temps plutôt froid avec fort ensoleillement mais une neige froide, sensible au chasse neige, aux fluctuations de température et aux paramètres structurels (nombre de couches dans le manteau, nombre de journées avalanches dans l'hiver)

AG⑤ : Temps chaud (plus au sens des températures, tant minimum que maximum, et des échanges convectifs que du rayonnement solaire). Ici le seuil de 0°C intervient de façon toute particulière et on peut certainement envisager une fusion, au moins superficielle, du manteau.

Les conclusions ressemblent fort à celles obtenues lors de l'étude du bimestre JAN-FE : Le groupe ① est d'origine mécanique (forte surcharge) de même que le groupe ③ (transport de la neige fraîche). Le groupe ② est sensible aux remontées de températures accompagnant un ciel variable avec éclaircies. En effet, le retour du beau temps, après une séquence de précipitations, s'accompagne beaucoup plus souvent d'un réchauffement qu'en Janvier-Février vu la saison relativement avancée et le rôle du rayonnement. Cela n'affecte pas forcément la neige qui peut rester froide en cas de beau temps prolongé, caractéristique du groupe ④. Par contre, un temps chaud, sur une neige déjà vieille, avec des effets convectifs notables (du type "foehn" ou vent du sud) provoque les avalanches humides du groupe ⑤ que l'on ne trouvait pas en Janvier-Février.

e) Comparaison des types de temps avalancheux de Janvier-Février et de Mars-Avril

On a effectué l'analyse factorielle discriminante des 4 groupes de Janvier-Février et des 5 groupes de Mars-Avril, de deux manières différentes :

- en valeurs réelles pour les X_l
- en valeurs désaisonnalisées $X_l - \bar{X}_l$, où $\bar{X}_l$ est la valeur moyenne de la variable l pour le bimestre considéré.

On constate que les centres de gravité des groupes dans un traitement comme dans l'autre, se répartissent selon un schéma "en équerre" (dû aux distributions particulières, tronquées, de certaines variables).

Seul le groupe ⑤ de Mars-Avril n'a pas d'équivalent en Janvier-Février, ce qui s'explique aisément. La stabilité de ces classifications, effectuées séparément sur chacun des bimestres, est à notre avis, surprenante. Si l'on y ajoute le sens physique, facilement interprétable, de chacun des groupes formés, il devient peu probable qu'elles soient le seul fait du hasard.

On constate aussi que le découpage effectué antérieurement, et qui distinguait "avalanches de neige sèche" et "avalanches de neige humide" avait pour seul avantage de mettre en évidence le groupe ⑤ de Mars-Avril mais confondait en un seul les groupes ①, ②, ③ et ④ ce qui explique certaines des faiblesses de ce système.

4.4. Conclusions.

Si l'on se reporte à la classification *génétique* proposée par le groupe de travail de la Commission Internationale des Neiges et Glaces, on constate d'abord que la plupart des causes avancées sont incluses, avec plus ou moins de finesse et de précision, dans les variables que nous utilisons (cf. annexe I).

D'autre part, on retrouve, au niveau des groupes formés, une famille où dominent les effets de "masse". Celle-ci se sépare ensuite entre un groupe où les effets proviennent des précipitations seules, et un autre où interviennent leur transport et leur remaniement par le vent. Dans l'autre famille, ce sont les effets "thermiques" ou "énergétiques" qui dominent. Ils se répartissent entre des effets à action rapide (cycles thermiques, redoux ou refroidissements brutaux) ou à action lente et prolongée.

La ressemblance frappante entre nos résultats et les propositions de l'ICSI est un indice encourageant de cohérence, tant pour les uns que pour les autres, puisque les résultats intuitivement avancés sont retrouvés par l'application objective et mécanique d'un algorithme de classification à des séries de données observées.

5. Recherche d'un modèle "occurrence/non occurrence" au sein de chaque type de temps préalablement défini

5.1. Classification des journées non avalancheuses

On a vu que les journées avalancheuses pouvaient se répartir en 4 ou 5 groupes assez bien séparés, correspondant à des *types de temps* différents. Toutefois ces types de temps sont d'origine *purement nivométéoro-*

logique, et le fait qu'une avalanche au moins se soit déclenchée ce jour là n'est pas entré directement dans leur définition. On peut donc chercher, parmi les journées non avalancheuses laissées de côté jusqu'à présent, si certaines ne ressemblent pas aux journées avalancheuses du type ① ou ②, etc.

Pour cela nous avons appliqué deux méthodes (linéaire et quadratique) sur les données suivantes :

Groupe de départ : les AG① précédemment obtenus

Variables utilisées : les X_i ($i = 1, 50$)

Individus à classer : toutes les journées non avalancheuses

On peut constater que certaines journées sont mal affectées, sensibles au modèle utilisé, etc, mais 80 % au moins se rattachent sans ambiguïté à l'un des groupes précédents.

On peut donc considérer non plus les journées avalancheuses du groupe ① mais celles qui ont été affectées au type ① bien que non avalancheuses. Si l'on ne considère que les plus représentatives, on constituera un échantillon NG① de journées du type ①, comparables à AG① pour ce qui est du type de temps, mais sans avalanche.

5.2. Mise au point du modèle "occurrences/non occurrence"

Il est tentant alors de demander à une analyse discriminante quelles variables ont décidé que la journée serait avalancheuse ou non. Pour chaque type ① successivement, nous avons opposé les groupes AG① et NG①. On constate que les variables séparant AG① et NG① sont de deux types :

— soit elles figurent dans la définition du groupe ① (par exemple mauvais temps avec précipitations) et c'est la valeur particulièrement forte de cette variable qui décidera des occurrences d'avalanches

— soit, au contraire, ce sont des variables complètement extérieures à la définition du type de temps qui expliquent les occurrences : par exemple le transport par le vent, dans les types de temps où beaucoup de neige fraîche est disponible au sol, ou au contraire, le passé du manteau quand les types de temps correspondent à des effets essentiellement thermiques.

Toutefois, nous serons très prudents encore dans l'interprétation. Si, dans le paragraphe 4.3 nous avons des problèmes d'interprétation de corrélations partielles, ici ils se compliquent du fait que nous nous trouvons au sein d'un sous-ensemble très particulier de la population, le type de temps ①.

Les figures 17 et 18 présentent la projection dans le plan discriminant des ellipses d'inertie des groupes avalancheux et non avalancheux.

Signalons aussi que, si les modèles d'affectation aux types de temps donnent peu d'ambiguïtés ou d'incertitude d'affectation, compte tenu des variables disponibles, il n'en est pas de même pour les modèles Avalanche/Non avalanche. Ceci semble indiquer que notre potentiel d'explication en ce domaine est plus faible : certaines variables explicatives n'y figurent pas, comme ces index de structure et de résistance du manteau dont nous ne disposons pas, ou sous une forme simpliste.

17/ Janvier-Février

18/ Mars-Avril

5.3. Interprétation des variables retenues

On se trouve désormais dans le cas de 2 groupes (cf. 4.3-b)

a) Bimestre Janvier-Février

Groupe ① : journées avalancheuses AG① : 28, non avalancheuses NG 1 : 5.

Le faible échantillon sans avalanche (quand on est dans ce groupe, la probabilité a priori d'avalanche est de 80 % au moins...) rend les résultats incertains. Toutefois, sachant donc que l'on a beaucoup de précipitations, l'activité avalancheuse s'accroît avec la quantité de neige fraîche

disponible, le vent du jour, et la fragilité du manteau caractérisée par le nombre de strates antérieures.

Groupe ② : $AG② = 34$ $NG② = 73$

Le modèle $AG②/NG②$ plafonne pour 7 variables, au delà desquelles la discrimination ne s'améliore guère. On sait que ce groupe suit ou termine une séquence de précipitations modérées.

Les variables les plus significatives sont les quantités de précipitations disponibles (neige fraîche au sol) et le chasse-neige, qui accroissent le risque. Par contre, le tassement de la couche de neige fraîche le réduit.

Groupe ③ : $AG③ = 26$ $NG③ = 37$

Il s'agit de beau temps froid suivant des précipitations, donc avec une couche de neige récente, froide peu stabilisée.

Les variables les plus significatives sont le vent et le chasse-neige du jour, qui jouent dans le sens du risque croissant. A cela s'ajoute le durcissement de la couche superficielle, détectée par la sonde de battage et peut être lié à la formation de plaques.

Groupe ④ : $AG④ = 34$ $NG④ = 92$

Ce groupe, relativement peu avalancheux, est caractérisé par des effets thermiques qui stabilisent la neige mais, s'ils se prolongent ou s'ils se produisent trop vite, donnent quelques avalanches. Ces avalanches viennent d'ailleurs un peu "en chaîne" et sont sensibles aux effets thermiques (ciel couvert, températures relativement élevées).

b) Bimestre Mars-Avril

Groupe ① : $AG① = 27$ $NG① = 23$

C'est un type de temps à précipitations moyennes à fortes, mais moins avalancheux qu'en Janvier-Février à cause des températures plus élevées qui rendent la neige poudreuse plus stable.

On constate d'ailleurs que si le risque augmente avec la neige fraîche disponible et le chasse-neige aux effets surtout mécaniques, il est accru par les basses températures de l'air et de la neige, et les refroidissements. Les hausses de températures ayant, dans un premier temps, un effet stabilisateur.

Groupe ② : $AG② = 27$ $NG② = 68$

Il s'agit d'un temps médiocre suivant une séquence de précipitations.

L'activité avalancheuse est sensible aux amplitudes de température et surtout au chasse-neige prolongé agissant sur la neige fraîche disponible.

Groupe ③ : $AG③ = 23$ $NG③ = 30$

Ce groupe se caractérise par un temps assez beau à variable, plutôt froid, sensible à une remontée des minimums mais surtout à la structure du manteau et à sa stratification.

Rappelons que ces 2 groupes (② et ③) sont voisins

et assez mal définis, sinon qu'ils caractérisent le temps intermédiaire entre séquence de précipitations et beau temps établi.

Groupe ④ : $AG④ = 28$ $NG④ = 27$

Il s'agit de beau temps, froid pour la saison, surtout sensible au chasse-neige depuis la fin des précipitations et à la structure du manteau (indices de "fragilité" tels que nombre de strates, nombre de situations avalancheuses antérieures, etc.)

Groupe ⑤ : $AG⑤ = 18$ $NG⑤ = 17$

Type de temps chaud et éventuellement perturbé (variable) d'autant plus stable que la neige est assez fortement tassée et sensible à la nébulosité et aux températures élevées.

6. Utilisation — Modèles d'explication — Potentiel de prévision

6.1. Conception et mise au point d'un modèle - Potentiel de prévision

On peut considérer que notre démarche a été soutenue par les hypothèses physiques suivantes :

- Le phénomène "avalanche" est un effet lié à des causes de nature fort différentes.
- La situation nivométéorologique, telle qu'elle est accessible à travers les variables dont nous disposons, contribue à une mise en condition du manteau, plus ou moins favorable au déclenchement, mais de façon différente selon les influences dominantes. On en est arrivé à définir, au vu des données et pour la région considérée, 4 ou 5 types de temps avalancheux.
- Au sein de ces types de temps, certains phénomènes, extérieurs ou non à la définition du type de temps, décident ou non du déclenchement.

a) Ces considérations nous amènent à proposer un modèle de prédiction à deux niveaux :

Etant donnée une journée j caractérisée par X_{j1} ($1 = 1,50$)

1) A quel groupe ou type de temps appartient-elle ?

$\Rightarrow \Pr(j \in G_i / \{X_i\})$

2) Au sein de ce type de temps, quelle est la probabilité que j soit avalancheuse ?

$\Rightarrow \Pr(1 \text{ avalanche au moins le jour } j / j \in G_i \cap \{X_i\})$

En fait, le calcul se complique légèrement du fait que l'on tient compte des probabilités *a priori* de chaque type de temps, c'est-à-dire que même en ignorant les $\{X_i\}$,

$$\Pr(j \in G_i) \neq \Pr(j \in G_k) \text{ si } j \neq k$$

De même, on a une estimation des probabilités *a priori* d'avalanches au sein des types de temps (cf alinéa 5.3). A ce niveau, pour être rigoureux et dans une optique pu-

rement bayésienne (cf. T.W. Anderson, 1958), on devrait faire intervenir les coûts d'erreurs. Or en ne tenant pas compte des probabilités a priori :

$\Pr(\text{Avalanche}/G_i) \neq \Pr(\text{Non avalanche}/G_i)$, on introduit indirectement un coût qui, s'il est de 1 pour la fausse alerte est supérieur à 1 (de 3 à 8) pour l'absence d'alerte en cas de danger caractérisé. Cette facilité de calcul nous a semblé suffisante.

b) Vérification des modèles

Ces modèles ont été appliqués aux hivers de la période 1960-1972, période qui, rappelons-le, a servi à l'ajustement des modèles ; le but de cette opération est de vérifier que le modèle est cohérent et reproduit correctement les phénomènes sur lesquels il a été calé.

Un autre intérêt consiste à voir évoluer les réponses du modèle et à acquérir une habitude de leurs réactions dans des conditions variées ; il y a là un apprentissage important. Nous donnons des exemples de résultats pour l'année 1972 (fig. 19).

6.2. Test du modèle sur les hivers 1972-1973 et 1973-1974

Pour ces deux hivers, n'ayant pas servi à l'ajustement du modèle, nous avons effectué le calcul des probabilités d'appartenance de chaque journée j aux différents groupes : cette probabilité est en général très "pointue", en ce sens qu'un modèle donné, linéaire ou quadratique, affecte en général une journée à un groupe et un seul. En cas de divergence entre les modèles, nous avons donné une légère préférence au modèle quadratique qui s'était avéré plus efficace en vérification.

Une fois le groupe choisi, nous avons calculé la probabilité, au sein de ce groupe, d'avoir au moins une avalanche, probabilité que nous avons ramenée à 4 niveaux :

0	Pas de danger
30 %	Danger faible ou apparaissant
60 %	Danger moyen
100 %	Danger important

Ce système conversationnel nous a permis de simuler le comportement d'un prévisionniste qui disposerait des 50 variables élaborées définies à partir des 15 variables brutes. Signalons toutefois que dans la simulation que nous avons ainsi réalisée, nous ne disposons pas d'un élément dont disposerait un prévisionniste : la vision matérialisée du temps qu'il fait et la connaissance intuitive de l'état du manteau neigeux grâce à la vue panoramique sur la région concernée ; dans les résultats présentés, nous nous en sommes tenus aux résultats du calcul (fig. 20, 21).

Cette nouvelle méthodologie aboutit à des modèles qui sont d'abord "explicatifs". Ajustés sur un certain nombre d'années et appliqués à ces mêmes données, ils délivrent un "signal" sous forme d'évolution de la probabilité d'avalanche en fonction du temps qui suit de façon très satisfaisante l'activité avalancheuse observée.

Cette méthodologie nous semble supérieure à celles

21/

que nous avons présentées précédemment, tant par son interprétation accessible au physicien, que par ses performances en prédiction et en prévision simulée.

Toutefois, trois améliorations nous semblent indispensables :

1) Utiliser, tant au niveau des algorithmes de classification qu'à celui de l'analyse discriminante des méthodes *quadratiques*, car les divers types de temps ont des dispersions très différentes sur certaines variables ; cependant le nombre de paramètres à estimer étant supérieur à celui des méthodes linéaires, elles requièrent des échantillons fournis.

2) Augmenter la robustesse des modèles — Cela suppose l'élimination maximale des erreurs d'observation d'avalanches ; par exemple, en Janvier 1973, il y a eu plusieurs avalanches le 28 et une autre signalée "entre

le 26 et le 28" : elle est très probablement liée à l'activité du 28 mais si on la place le 26 et que l'on utilise la journée du 26 dans un échantillon d'ajustement, on introduit certainement un biais.

Comme ces problèmes ne touchent que 10 à 20 % des cas avalancheux, on aura quand même intérêt à augmenter le fichier d'ajustement pour améliorer la robustesse de l'ajustement, en éliminant au besoin les cas trop litigieux.

3) Au niveau des variables explicatives, on pourrait éliminer des variables trop sophistiquées, quitte à réduire un peu l'efficacité de l'ajustement, de façon à avoir un modèle plus stable en test ou en prévision et plus facile à utiliser opérationnellement.

Il ne s'agit cependant que de mises au point, certes longues et coûteuses en temps de calcul, mais qui ne remettent pas en cause la méthodologie avancée.

Conclusions générales

Les études que nous avons menées nous ont convaincus qu'il existe, au sein d'une information nivométéorologique qu'il est possible de rassembler, un certain potentiel de prévision pour le phénomène avalancheux. La méthodologie que nous proposons se compare favorablement, même utilisée seule, avec les résultats d'un prévisionniste entraîné. Un dialogue homme-machine, réalisant l'association de cette prévision numérique avec l'expérience et le sens du terrain d'un prévisionniste entraîné devrait améliorer encore ses performances.

Cependant, le système présenté ne concerne, à l'heure actuelle que la seule région de Davos ; mais le service "Avalanche" (Lawinendienst) de l'Institut Fédéral de Davos dispose, à notre avis, des archives et des compétences nécessaires pour l'étendre à l'échelle de la Suisse, si les moyens matériels lui en sont donnés. Ailleurs, par contre, l'avenir de ces méthodes dépendra essentiellement de la volonté que l'on mettra à les utiliser et surtout à collecter avec soin les séries de données qu'elles nécessitent.

Remerciements

Nous tenons à remercier tout particulièrement M. de Quervain, directeur de l'Institut Fédéral Suisse pour l'Etude de la Neige et des Avalanches, et M. Good, son adjoint, pour l'intérêt qu'ils ont manifesté à nos travaux et l'aide matérielle et financière qu'ils nous ont apportée. M. Beck, observateur, et M. Foehn, météorologue à l'Institut nous ont beaucoup aidé dans la préparation et la critique des données.

Sur le plan méthodologique, MM. Guillot et Duband de la Division Technique Générale de l'E.D.F. et M. Cazes de l'I.S.U.P., nous furent d'un grand secours.

Et nous ne manquerons pas, pour terminer, de remercier M. de Crecy, président de la Commission de Glaciologie de la S.H.F., pour l'intérêt avec lequel il a suivi toutes nos études et les encouragements qu'il nous a prodigués.

Bibliographie

- OBLÉD (Ch.). — Vers une prévision numérique des avalanches. *Rapport D.G.R.S.T.*, Laboratoire de Mécanique des Fluides. Grenoble 1970.
- FINNEY (D.J.). — *Probit Analysis*. Cambridge University Press, 1964.
- BOIS (Ph.) et OBLÉD (Ch.). — Vers un système opérationnel de prévision des avalanches par des méthodes statistiques. *Bulletin AIHS*, Vol. 18 n° 4, Dec. 1973.
- BOIS (Ph.), OBLÉD (Ch.) et GOOD (W.). — Multivariate Analysis as a Tool for day by day avalanche forecast — *Symposium on Snow Mechanics* — Grindelwald 1974. pub. AIHS.
- CAZES (P.). — Régression et corrélation — *Cahiers du B.U.R.O.*, 1975.
- COOLEY (W.) et LOHNES (P.). — *Multivariate data Analysis*, John Wiley, 1971.
- LEBART (L.) et FENELON (J.P.). — *Statistique et Informatique Appliquées* — Dunod, 2ème ed. 1973.
- HILTON (G.). — An Algorithm for detecting differences between transition probability Matrices. *Applied statistics*, 1972.
- ROMEDER. — *Méthodes et Programmes d'analyse discriminante*, Dunod, 1973.
- CACOULOS (T.). — *Discriminant analysis and application*. Academic Press, 1972.
- QUERVAIN de (M.), CRECY de (L.) et al. — "Avalanche Classification". Proposals of the working group of the Internal Commission on Snow and Ice — *AIHS* — Bulletin Vol. 18 n° 4, Décembre 1973.
- NAKACHE (J.P.). — Influence du codage des données en Analyse factorielle des correspondances. *Revue de Statistique Appliquée*, 1973, vol. 21, n° 2.
- DIDAY (E.) et al. — Diverses publications IRIA sur "la méthode des Nuées Dynamiques".
- ANDERSON (T.W.). — *An introduction to multivariate analysis* John Wiley 1958.

Annexe I

Aperçu sur les variables "explicatives" utilisées

Variables

- 1 Précipitation du jour en mm d'eau
- 2 Précipitation du jour en cm de neige fraîche
- 3 Intensité de la chute du jour
- 4 Estimation de la densité de la couche superficielle
- 5 Cumul des précipitations sur les 3 derniers jours
- 6 Vent du jour
- 7 Différence entre la variable 1 et la quantité recueillie au pluviomètre
Index de transport par le vent en %
- 8 Chasse-neige observé
- 9 Cumul du vent sur les 3 derniers jours
- 10 Variable 7 mais en valeur absolue
- 11 Accroissement du vent pendant ou peu après les précipitations
- 12 Température de l'air à 13 h
- 13 Ecart de températures à 13 h entre les jours j et j - 1
- 14 Température à 13 h si inférieure à 3°C sinon 0
- 15 Température à 13 h si supérieure à 3°C sinon 0
- 16 Température de l'air à 7 h
- 17 Ecart des températures à 7 h entre les jours j et j - 1
- 18 Heures d'insolation du jour
- 19 Rayonnement solaire incident
- 20 Nébulosité moyenne
- 21 Rayonnement solaire absorbé
- 22 Pourcentage du rayonnement potentiel reçu
- 23 Nombre de séquences de précipitations depuis le début de l'enneigement
- 24 Cumul des précipitations de la dernière séquence
- 25 Nombre de jours sans précipitations depuis la fin de la dernière séquence
- 26 Nombre de jours sans précipitations ayant précédé la dernière séquence de précipitations
- 27 Maximum, depuis le début de l'hiver, de la variable 26
- 28 Nombre de cas de chasse neige depuis le début de la dernière séquence de précipitations

- 29 Nombre de jours avec chasse-neige depuis le début de la dernière séquence
- 30 "Storm Index". Précipitations cumulées pondérées par le vent
- 31 Hauteur de neige fraîche "efficace", calculée à partir de la hauteur au sol précédant les dernières précipitations
- 32 Tassement depuis la dernière hauteur maximale atteinte
- 33 Tassement relatif par rapport à la seule neige fraîche
- 34 Tassement entre les jours j et $j - 1$
- 35 Enfoncement de la sonde de battage
- 36 Variation de 35 entre j et $j - 1$ si croissant
- 37 Variation de 35 entre j et $j - 1$ si décroissant
- 38 Cumul sur les 5 derniers jours des heures d'ensoleillement
- 39 Cumul sur les 5 derniers jours du rayonnement incident
- 40 Cumul sur les 5 derniers jours des degrés jours positifs
- 41 Cumul sur les 5 derniers jours du rayonnement absorbé
- 42 Température dans la neige à 10 cm sous la surface vers 8 h
- 43 Variation de 42 entre j et $j - 2$
- 44 Variation de 42 entre j et $j - 1$
- 45 Variation relative de 42 entre j et $j - 2$
- 46 Cumul sur les 7 derniers jours de la température de la neige
- 47 Variation de la température de l'air à 7 h, entre j et $j - 2$, pondérée par la proximité à 0°C
- 48 Nombre d'avalanches observées les jours j et $j - 2$
- 49 Nombre de journées avalancheuses depuis le début de l'hiver
- 50 Nombre de journées avalancheuses sur nombre de séquences depuis le début de l'hiver.

Discussion

Président : M.L. de CRECY

M. le Président remercie M. OBLED de son exposé qui traite d'un problème difficile mais de grande actualité. Il ouvre ensuite la discussion.

M. PLAS (E.E.R.M. - C.E.N.) intervient en ces termes :

Le premier modèle de ce genre a été élaboré à partir de données d'archives françaises, avant 1970. Nous y avons participé avec M. OBLED (il n'est pas mentionné dans la bibliographie). Il était bien sûr moins élaboré que celui qui est présenté, ici, par MM. OBLED et BOIS, mais il a servi de base initiale pour les travaux qui ont suivi ; c'est précisément ce premier travail qui a suscité l'intérêt de l'Institut Suisse de Davos.

Mais pour parfaire ce modèle de prévision numérique, il nous paraît très intéressant, voire très important, de le compléter en prenant en compte des données de mesures - même si elles sont peut-être encore imparfaites - sur l'état du manteau neigeux en place. La "base de décision numérique" obtenue en serait certainement bien plus précise pour le prévisionniste.

Les premières études, répond M. OBLED, partaient d'une base assez différente et tentait d'estimer, à partir d'une seule variable - par exemple l'épaisseur de neige fraîche au sol - la probabilité d'occurrence d'une avalanche.

Dans notre modèle, nous avons tenté de formaliser et de quantifier la démarche, plus ou moins implicite, du "prévisionniste" ; à cet effet, dans une première phase, nous essayons, à partir de quelques paramètres reliés à la situation météorologique, de cerner des types de "situations avalancheuses", c'est-à-dire propices au déclenchement de certains genres d'avalanches ; dans un deuxième temps, à partir d'autres variables, nous cherchons à déterminer s'il y aura une avalanche ou non au sein de ce type.

Pour l'ensemble des paramètres que vous avez utilisés, remarque M. DELSOL, vous avez disposé d'un fichier assez détaillé mais qui vous a paru, semble-t-il, encore insuffisant. Avez-vous essayé de "tester" votre modèle sur un cas où les données étaient notablement plus incomplètes que celles disponibles à Davos ?

En effet, répond M. OBLED, nous n'avons pas été très satisfaits des observations d'avalanches. Malgré le soin méticuleux avec lequel sont effectuées les observations en Suisse, certains jours classés sans avalanches ont pu être sujets à des déclenchements, que le mauvais

temps n'a pas permis de repérer avant plusieurs jours. D'où une incertitude sur les dates de déclenchement.

D'autre part, le fichier des données nivométéorologiques n'est pas parfaitement adapté à notre problème. Nous disposons d'assez nombreuses mesures météorologiques dont certaines sont "redondantes" mais avons très peu de mesures concernant les paramètres caractérisant l'état du manteau neigeux à une époque déterminée.

Nous avons cherché à éliminer certaines variables initialement sélectionnées mais qui, à l'usage, nous ont paru peu sûres ; nous avons constaté une certaine baisse des performances du modèle mais nous n'avons pas fait d'étude systématique sur ce point.

Un dialogue s'engage entre M. GUILLOT et M. OBLED pour rechercher parmi les imperfections du modèle celles qui pourraient être palliées par l'entrée de nouvelles variables. Dans cet esprit, M. LEFEBVRE (SOGREAH) formule les suggestions suivantes :

Est-ce que de la même manière que la représentation des différentes variables sur des axes principaux simplifiés fait apparaître des types de situations météorologiques - ce qui constitue une confirmation de la réalité physique des données du modèle - vous auriez pu effectuer l'analyse en caractérisant le résultat d'une façon plus précise en ce qui concerne le type d'avalanche à craindre (poudreuse, plaque, neige mouillée)... ?

Est-ce que, malgré l'intérêt que cela pourrait représenter - dans la mesure où le modèle serait ainsi plus proche de la réalité physique - la complexité du traitement supplémentaire en rend inefficace l'application, où est-ce seulement au niveau des données de comparaison, déjà difficile à obtenir pour : "Avalanche ou Non-Avalanche" que cette distinction semble inapplicable ?

On pourrait parfois pour améliorer le fichier d'avalanche, répond M. OBLED, disposer d'un certain nombre de paramètres topographiques (longueur, largeur, épaisseur de l'avalanche) ou physiques (type d'avalanche, température de la neige), mais pour l'instant, en raison du faible nombre des données disponibles, on ne peut prétendre estimer que la probabilité d'occurrence ou de non-occurrence de l'avalanche (au sein d'un type de temps). Encore faut-il remarquer que le nombre d'avalanches observées est lié au nombre d'observations effectuées, ce qui introduit un biais supplémentaire dans tous les calculs statistiques.