

CIME: UN SYSTEME EXPERT DE CARTOGRAPHIE PAR TELEDETECTION EN REGION DE MONTAGNE

D.BLAMONT. U.P.R.299/CNRS; 1 PL.A.BRIAND, 92190 MEUDON

C.MERING. LIA/ORSTOM; 72-74 ROUTE D'AULNAY, 93140 BONDY

RESUME: Les milieux himalayens présentent des conditions difficiles à la cartographie par télédétection. Cette complexité rend nécessaire l'introduction d'autres variables que la radiométrie. De plus, la nécessité de la reproductibilité d'une méthode de télédétection justifie le recours à une démarche de type *Système Expert* qui dissocie la méthode de la spécificité des problèmes posés. Fondé sur la représentation simultanée de deux savoir-faire, le système CIME (cartographie intelligente en milieu montagnard) est destiné à piloter et à optimiser l'activation séquentielle de traitements numériques en vue de cartographies thématiques.

ABSTRACT: The Himalayan landscapes are so complex that their mapping using remote sensing requires other data, in addition to radiometry. An expert system has to be established in order to guarantee the reproductibility: CIME (Smart Cartography system in mountainous area), separating methodology and thematic, elaborates thematic maps by driving and optimizing the sequential activation of numerical treatments.

1.- Les éléments de complexité de la cartographie [Blamont & Méring 88]**1.1.- les facteurs naturels et anthropiques**

∞ **La multiplicité des unités de paysage:** Dans le centre du Népal, l'amplitude altitudinale d'un versant peut atteindre 5.000 m. Il en résulte un étagement des milieux allant du sub-tropical à l'alpin et une très grande variété des unités de paysage. Les diverses expositions et situations à l'intérieur du massif font que la nature des étagements varie d'un versant à l'autre.

∞ **La taille des unités de paysage:** les pentes oscillant autour de 40°, les étages de la végétation sont très étroits: la taille des unités de paysage en est réduite d'autant.

∞ **Les marges:** qu'elles soient naturelles ou d'origine anthropique, les limites entre les différentes unités de paysage sont le plus souvent floues et peuvent constituer des unités en soi.

∞ **L'exploitation des milieux:** l'exploitation des forêts est fonction de leur localisation et des pentes: ce facteur de diversité ajoute aux irrégularités naturelles.

1.2.- Les facteurs techniques: les effets de l'ensoleillement

∞ **Les voiles atmosphériques:** ils sont plus importants au dessus des versants à l'ombre

∞ **Les dynamiques de la radiométrie:** elles sont plus faibles sur les versants à l'ombre.

∞ **Les différences de valeur de la radiométrie:** puisque les différences de valeurs sont aussi des différences de nature, on n'opère aucune correction du signal en fonction de l'exposition. Grâce à un modèle de l'ensoleillement à l'heure de prise de vue, sont déterminés et traités séparément trois secteurs à l'intérieur desquels l'influence de l'exposition sur la radiométrie est considérée comme marginale par rapport aux différences de réflectance des unités de paysage: les secteurs à l'ombre; les secteurs à la lumière; les secteurs de lumière rasante. Soulignons cependant une coïncidence certaine entre l'influence instantanée de l'éclairement sur les différences radiométriques et son influence générale sur les différences des formes de la végétation: marquer ainsi celles-là lors de la procédure permet non seulement les traitements mais aussi de cartographier celles-ci. Néanmoins, à l'intérieur de chacun des secteurs, l'influence de la pente sur la radiométrie reste forte: elle sera prise en compte par l'introduction de mécanismes de correction du signal en fonction de la pente et de l'altitude.

1.3.- Les éléments des traitements

On procède à la partition d'une image satellitaire par une succession de traitements numériques portant sur les valeurs associées aux pixels en vue de l'identification, la caractérisation et la délimitation d'ensembles de ces pixels et leur assimilation à des taxons définis avant l'analyse de l'image mais après un travail de terrain qui a permis de dresser la liste d'un ensemble cohérent et considéré comme complet des formes d'occupation de l'espace (le paysage) et de repérer des parcelles-test dont une partie (les parcelles d'entraînement) est utilisée pour déterminer les lois de seuillage sur les variables que l'autre partie (les parcelles de contrôle) sert à valider.

1.3.1.- Variété des formes iconiques: les objets traités sont les pixels des parcelles: bien qu'ils soient considérés isolément dans le Système, les procédures de traitement statistique font référence aux parcelles dans leur cohésion. La différence de représentation adoptée pour les traitements numériques et le traitement symbolique fait pendant à une différence de statut des procédures en oeuvre dans ces deux phases. Par ailleurs, la notion de classe fait référence au figuré géographique d'une zone, par exemple "forêt dense sapin". Cependant, bien que la définition des pixels puisse être de l'ordre de la taille des objets géographiques étudiés, la connaissance de ces derniers n'a rien à voir avec celle des pixels, même s'ils semblent se correspondre.

1.3.2.- Nombre important de descripteurs: la complexité des paysages ne permet pas une classification à partir de la seule radiométrie: d'autres variables sont attachées aux pixels et sont soit calculées à partir de celle-ci (indice de végétation verte -établi sur MSS 5 & 7- ou de texture -établi sur MSS6- [Blamont & al. 84]), soit extrinsèques aux données satellitaires (l'altitude, la pente, l'ensoleillement).

1.3.3.- Des procédures de contrôle constant: plutôt que de tester toutes les chaînes (séquences d'utilisations successives de différents traitements portant sur les différentes variables), des procédures de contrôle permettent à chaque étape soit d'interrompre soit de poursuivre une chaîne de traitements.

2.- Le système [Mering & al 88]; [Mering & Blamont 89]

2.1.- Pourquoi simuler la démarche

L'utilisation simultanée de toutes les variables au sein d'un seul traitement ne donne pas de résultats satisfaisants: la pertinence d'une variable pouvant être masquée par la plus grande pertinence d'une autre, d'une part, une variable utilisée avec une certaine efficacité à un stade pouvant être réutilisée à un stade ultérieur plus fin, d'autre part. Il s'agit donc de déterminer l'ordre optimum d'utilisation des traitements et des variables en fonction d'un objectif spécifique. Dans le cas présent, cet ordre se trouve pouvoir être différent d'un secteur d'éclaircissement ou d'un thème à l'autre, d'une portion d'image ou d'une époque à l'autre. Une manipulation serait très coûteuse en temps qui ne dissocierait pas les procédures de mise en route et l'organisation en chaîne des traitements d'une part, de ces traitements et de la spécificité des problèmes posés d'autre part. Le **Système CIME**, grâce à cette dissociation, teste toutes les chaînes possibles.

2.2.- La nature des connaissances

- ∞ Les connaissances thématiques (des taxinomies mais aussi des problématiques établissant entre les objets des relations autres que d'emboîtement; par exemple: la chaîne dégradée se trouve sur les pentes les plus accessibles -par la proximité et la valeur de la pente)
- ∞ Les connaissances liées à l'expression iconique des objets (radiométrie, texture...mais aussi les relations topographiques des objets (proximité, éloignement, imbrication, adjacence...)).
- ∞ Les connaissances en analyse d'image: ne reposant pas sur une simulation de la vision, le système fait appel au résultat explicite de traitements numériques. Pour l'instant, seules les classifications dirigées sont utilisées et en particulier celles ressortissant d'une approche probabiliste (D.N.P.: *discrimination bayésienne non paramétrique* [Celeux 80]) qui produit un seuillage des variables descriptives permettant de partitionner l'ensemble des pixels de la scène.

2.3.- La représentation des connaissances

Bien que de nature variée, les connaissances ne sont représentées que sous deux formes:

2.3.1.- les connaissances factuelles: les connaissances sur les pixels de contrôle et sur les concepts thématiques généraux constituent la **base de faits**. Chaque pixel sera représenté comme un objet décrit par des attributs ayant une ou plusieurs valeurs numériques (radiométrie, indices, altitude, pente...) ou symboliques (la classe représentée par la zone-test à laquelle il appartient, la liste des classes (cl après) qui lui seront associées après une étape (phase d'analyse). Les connaissances factuelles non associées aux pixels (ex: l'attribut étape, attaché à un objet nommé contexte) sont également représentées comme des objets dans la base de faits. Les différents ensembles d'objet sont différenciés par leurs types.

2.3.2.- les connaissances procédurales: elles définissent, sous une forme propositionnelle, les moyens d'obtenir une nouvelle connaissance à partir d'un ensemble de faits déjà établis. Qu'il s'agisse des règles concernant la connaissance thématique ou algorithmique, elles sont données en vrac et insérées dans la **base de règles** sous une forme déclarative simple:

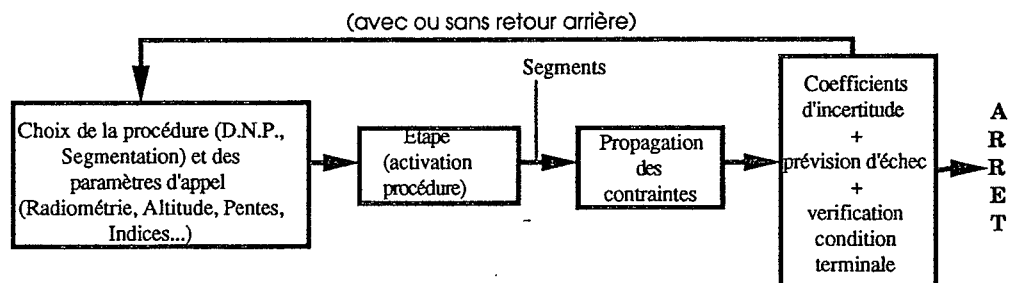
Si Conditions Alors Conclusions.

3.- Le fonctionnement du système

3.1.- Supervision du symbolique sur le numérique

Les pixels étudiés par le système appartiennent aux **parcelles-test**. Lors d'une procédure de seuillage (étape), le système étiquette chaque pixel en lui associant un ensemble de classes potentielles composé d'un certain nombre des classes *a priori*, taxons définis sur le terrain: l'étape est validée si cet ensemble comporte la classe *a priori* du pixel. Chaque segment est alors considéré comme une contrainte propagée sur les pixels des parcelles de contrôle: un seuillage préalable n'est en effet pas remis en cause et les classes associées aux segments sont définies par une formulation symbolique négative (ex.: non_sapin, non_chêne, etc...). Le but d'une étape est donc d'éliminer certaines des classes potentielles associées à chaque pixel lors des seuillages précédents. Les règles de seuillage ou sont déclaratives ou proviennent d'un traitement automatique (il faut alors que les parcelles-test soient représentatives de l'ensemble des situations possibles pour cette variable).

3.2.- Constitution d'une chaîne



Une chaîne se caractérise par l'ordre des traitements des variables quantitatives (par exemple ensoleillement, radiométrie, altitude, texture, pente... : l'objectif final est de trouver une chaîne qui classe correctement chaque pixel des parcelles de contrôle. Comme il peut exister plusieurs chaînes toutes imparfaites, la meilleure (celle contenant le plus grand nombre de pixels classés correctement ou avec la moins grande incertitude) sera appliquée à l'ensemble de l'image, mais le système conserve la trace de toutes les chaînes (Fig. 2), (voir page 40).

3.3.- Règles et coefficients de contrôle

Les règles activant les traitements, contrôlant les résultats puis entérinant les choix sont indépendantes de la thématique et peuvent être utilisées pour d'autres applications; les coefficients servant au contrôle des traitements et au choix des chaînes sont au nombre de deux: le coefficient d'incertitude C_i est calculé à partir des pixels de contrôle, à chaque étape, après propagation des contraintes sur l'ensemble de ces pixels. Il est construit pour évaluer le degré d'incertitude du classement des pixels à cette étape et sert à contrôler l'évolution du classement au sein d'une chaîne et à orienter vers une poursuite de la chaîne ou un abandon et un retour arrière; Le coefficient de vraisemblance C_v est calculé à partir des pixels de contrôle à la fin de chaque chaîne. Il est construit pour évaluer la vraisemblance du classement résultant de cette chaîne et permet de comparer deux chaînes, la meilleure chaîne étant celle dont le coefficient de vraisemblance est le plus élevé. Il est égal à la somme des pixels de contrôle exactement classés à l'issue d'une chaîne. Des fonctions du système permettent de calculer ces indices à l'aide des formules suivantes:

$$C_i = 1/N \{ \sum \{ \sum \sum \delta_j^k; k \in K(i); j=1,m \} i=1,N \}$$

$$C_v = 1/N \{ \sum \{ \sum \{ \sum \delta_j^{n(i)} * (1 - \delta_1^k); k \in K(i); l=1,m, l \neq j; \} j=1,m; \} i=1,N \}$$

où:

- δ_j^k (pour C_i) & δ_1^k (pour C_v) est le symbole de Kronecker (=1 si $l=k$; =0 si $l \neq k$)

- $K(i)$ la suite ordonnée des numéros des classes affectées a posteriori au pixel i ;

- $n(i)$ le numéro de la classe a priori du pixel i ;

- m le nombre total de classes a priori;

- N le nombre de pixels analysés;

Par ailleurs, à l'issue de chaque étape le système vérifie la pertinence de la connaissance du terrain en signalant tout pixel des parcelles de contrôle dont la classe a priori ne figure pas dans la liste des c_l après et en interrompant alors la chaîne. Cette remise en cause de l'information initiale (erreur d'échantillonnage ou mauvais choix des classes) est un apport très positif des techniques de l'Intelligence Artificielle. La distinction entre validation interne (la minimisation du risque de Bayes dans le cas de la discrimination non paramétrique du même nom) et externe (par reconnaissance des parcelles-test a posteriori) permet d'utiliser la stratégie d'ensemble pour des procédures de classification différentes pourvu que leurs résultats se traduisent par un seuillage.

CONCLUSION

Le système a été élaboré pour la cartographie des formes de l'occupation des sols dans le Centre du Népal à partir d'images Landsat. Les premières chaînes ont permis la production de différentes cartes en fonction des thèmes privilégiés sur chacun des secteurs d'éclairage (fig.1, p.40). Ces cartes représentent non seulement les types de végétation et les catégories de champs mais encore l'état de la végétation. Le système, qui accepte l'introduction d'autres variables et d'autres procédures de traitement que celles utilisées actuellement, doit être maintenant testé sur des images SPOT mais aussi pour d'autres thématiques et/ou d'autres régions.

BIBLIOGRAPHIE

- [Blamont & al. 84]: D. Blamont, C. Méring & J.F. Parrot. Numerical processes for identification of landscape units in mountainous areas. XV^e congress of ISPRS. Rio de Janeiro. 1984
- [Blamont & Méring 88] D. Blamont & Méring: CIME: An expert system for landuse mapping by Remote Sensing: the example of a mountainous area (central Himalaya). Asian Conference on Remote Sensing. Bangkok 1988
- [Celeux 80]: G. Celeux & Y. Lechevallier. Méthode de discrimination non paramétrique asymptotiquement efficace au sens de Bayes. Rapport de recherche n°52. INRIA; 1980.
- [Mering & al 88]: C. Méring, D. Blamont, J.G. Ganascia & J.F. Monjanel. CIME: Une Application Des Systèmes Experts à la Télédétection. Actes des 8^e journées internationales sur les Systèmes Experts et leurs applications. Vol 3, pp 427 & 448. Avignon, 1988,
- [Mering & Blamont 89] C.Mering & D.Blamont: Application des systèmes Experts à la Cartographie par télédétection. La Recherche et les satellites à haute résolution spatiale.Orléans. 1989.

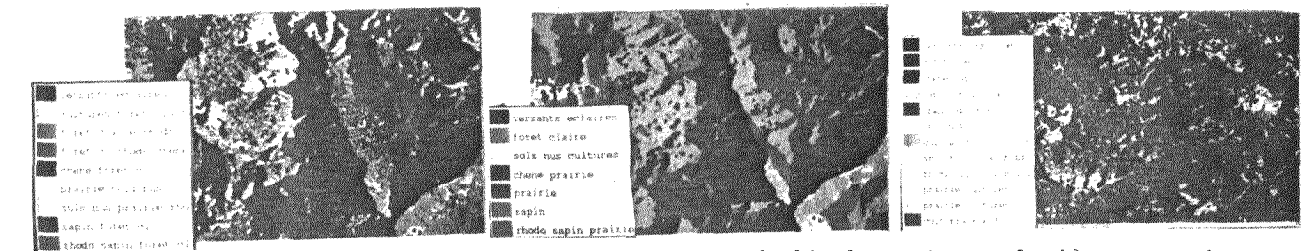
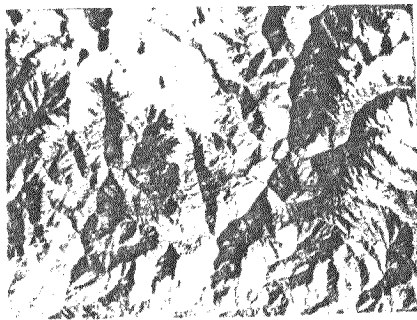


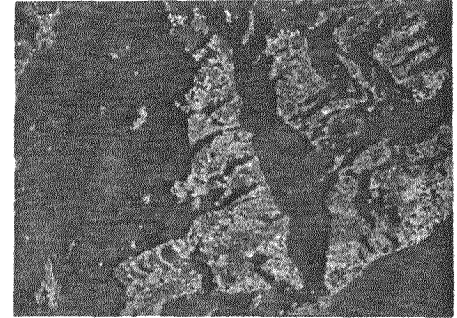
FIG. 1 : cartes intermédiaires obtenues sur versant à l'ombre et en lumière rasante



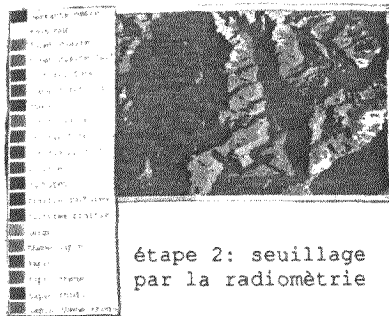
M.S.S.7 non masqué
Centre du Népal



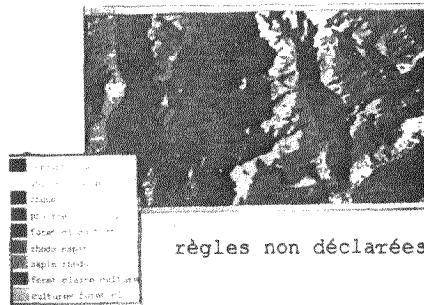
le modèle numérique de terrain
masqué par le modèle
d'ensoleillement



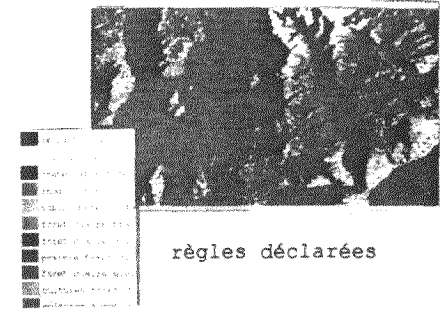
M.S.S.6 masqué
(étape 1)



étape 2: seuillage
par la radiométrie



règles non déclarées



règles déclarées

étape 3: seuillage par l'altitude

FIG. 2: CIME: une application des systèmes experts à la télédétection
une séquence d'une chaîne du traitement du secteur à la lumière
la chaîne est constituée des étapes suivantes:
masque par l'ensoleillement-DNP_radiométrie-seuillage par l'altitude

(fig. mentionnées page 64) - D. BLAMONT - C. MERING.



Figure 4 : Ratio $\frac{TM3}{TM5}$
FONTE DU COUVERT NEIGEUX
Alpes du Nord. Image LANDSAT TM
195-029 1/4 NW. 18 avril 1984

Fig. mentionnée p. 51
J.P. DEDIEU et al.

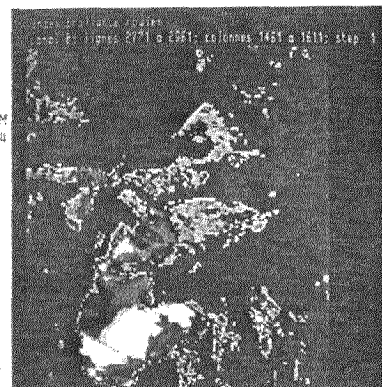


Figure 6
INDEX DE BRILLANCE
($IB = R^2 + PIR^2$)
BILAN DE MASSE DES GLACIERS
Glacier des Poules (Oisans)
Orientation NE et SE.
Image SPOT XS . KJ 051-259
5 octobre 1986.

Fig. mentionnée p. 51
J.P. DEDIEU et al.

Noir: vallées sans neige (Arc au nord,
Drac au sud-ouest, Durance au sud-est)
et nuages.
Gris: manteau neigeux peu transformé en
surface (Vanoise, Maurienne).
Blanc: manteau neigeux en métamorphose
(vallées internes, versants SE, Piémont
italien, sud-Oisans, Queyras, retenue
du Mont-Cenis).

Blanc: neige de chute récente
Bleu: zone d'accumulation (névé)
Jaune: zone d'ablation (glace vive)
Rouge et Magenta : dépôts morainiques
et rochers vifs.