

OBSERVATIONS DE TERRAIN ET TÉLÉDÉTECTION (BAIE DU ROI, SVALBARD NORD-OCCIDENTAL)

Dominique LAFFLY

Société, Environnement, Territoire, UMR 5603

Université de Pau et des Pays de l'Adour

Résumé

Les observations de terrain sont fondamentales en télédétection : selon la finalité de l'étude, la description des paysages observés sera plus ou moins détaillée. Une simple reconnaissance des principaux postes d'occupation du sol demande peu d'indicateurs et ne nécessite pas une localisation à la tâche près. En revanche, l'analyse du contenu des pixels en vue d'une modélisation statistique s'appuie sur une description thématique détaillée de points localisés très précisément. Dans ce cas les images utilisées doivent être correctement géoréférencées et la localisation sur le terrain se fait à l'aide d'un GPS différentiel. Reste le problème du choix des points à analyser ainsi que celui de l'observation de leur "contenu". Nous présentons dans cet article une méthode qui vise à répondre à ces deux questions. Notre approche est fondée sur la technique du sondage systématique non-aligné. Cette dernière permet de tenir compte à la fois de la luminance des points observés et de leur localisation pour définir le plan d'échantillonnage. Les exemples présentés sont issus d'une étude réalisée sur les sandurs de la presqu'île de Brøgger au Svalbard nord-occidental (images SPOT XSP).

Abstract

Field observation are very important for remote sensing application they represent the « field true ». Depending on the issue of the study landscape description will be more or less detailed. A simple observation of the most important land cover type doesn't need a lot of indicators and a very precise localisation at the scale of pixel. On the other and, pixel's inside modelisation for statistical monitoring needs a very detailed thematic description and a geographic localisation of each pixel. In taht case used images must be georeferenced with model including orbital parameters, DEM... and on the field it's necessary to uses a differential GPS to locate correctly each reference points. The most important problem concern in fact the selection of observable points and the way to realise their thematic description : wich points to use ? How to observe them ? We propose in this paper some reflection to answer at these questions. For that we a specific method of sampling the non-align systematic stratified sample in a double step approach. At a first scale to locate and to select field points - but we'll see the differences between theory and land condition -, at a second scale to define the way of inside's pixel oservation. To illustrate our discussion we propose a thematic study of water flow area on the sandur of King's Bay sub-island on North-West Svalbard.

Mots-Clés

Plan d'échantillonnage, ruissellement, télédétection, observations de terrain, Svalbard

Key-Words

Remote sensing, field observation, water flow, sample method, Svalbard

Depuis plusieurs années (tous les 2 ans depuis 1989) nous menons des missions de terrain au Svalbard (base CNRS J. Corbel, Ny-Ålesund) dans le but de cartographier les paysages à grande échelle à partir d'images satellitales. Les premiers travaux ont été consacrés à des vérifications de terrain afin de constituer une collection de données utilisées par la suite pour des classifications supervisées (Landsat TM). Au-delà de cette application classique de la télédétection notre objectif est de proposer aujourd'hui une réflexion méthodologique afin d'utiliser l'information radiométrique (SPOT XSP) pour cartographier – en termes de

probabilités – des composantes du paysage dont la résolution spatiale est inférieure à celle des pixels [1] [3] [4]. Les sandurs et strandflats de la presqu'île de Brøgger se prêtent bien à ce type de recherche par la relative simplicité des paysages qu'on y rencontre (relief peu contrasté, végétation restreinte, absence d'infrastructures humaines...).

Contrairement aux relevés de terrain effectués pour une classification supervisée, l'analyse détaillée du contenu des pixels demande davantage de précision. Dans le premier cas on se contente de vérifier la signification thématique d'ensembles relativement homogènes identifiés sur des compositions colorées. Les portions d'espace retenues sont relativement aisées à localiser (il s'agit de « taches » composées de plusieurs pixels) et leur description thématique a pour but essentiel de dégager des ensembles paysagers uniformes. Par exemple, au Svalbard, on identifiera la toundra, les zones humides, le minéral, la glace... Ces observations sont nécessaires afin de cartographier les principaux types d'occupation du sol. Pour chacun d'eux on peut ensuite envisager d'affiner les observations afin de décrire en détail le contenu thématique à l'échelle du pixel. Il ne s'agit plus d'identifier des « taches » homogènes ni d'estimer des types généraux de paysage. Chaque point retenu doit être localisé avec précision, sa description fait appel à une fiche détaillée: composition végétale, physionomie végétale, granulométrie, topographie... pour notre exemple. Un des principaux problèmes rencontrés sur le terrain¹ concerne alors la définition de plans d'échantillonnage fiables: quels points – et combien – observer sur le terrain? Quelle grille de lecture appliquer à chacun d'eux?

La pratique du terrain impose d'avoir recours à deux méthodes d'échantillonnage emboîtées. La première concerne la définition du nombre de points à observer ainsi que leur localisation. La seconde s'intéresse aux relevés effectués pour chaque point défini précédemment.

1. Échantillonnage de l'image: contraintes statistiques et contraintes de terrain

D'un certain point de vue, une image satellitale peut être considérée comme un recensement de l'espace géographique. Les individus sont les pixels, les variables les canaux spectraux. Une procédure probabiliste – sous contrainte spatiale – pourrait alors s'appliquer à une telle population afin d'en tirer un échantillon représentatif. Par exemple, un plan d'échantillonnage systématique stratifié non-aligné² permet de sélectionner des individus selon la distribution des variables en éliminant toute trace d'autocorrélation [8]. La figure 1 présente la localisation des points échantillonnés à partir de la scène SPOT XSP (la mer, les glaciers et la montagne sont masqués). Contrairement à un tirage aléatoire simple la répartition spatiale des échantillons est plus performante, on n'observe pas de zones de concentration des points et/ou de grands vides.

Fiable du point de vue statistique, cette procédure ne tient pas compte des contraintes de déplacement et de localisation sur le terrain. Il est pratiquement impossible de renseigner tous les points de la figure 1 dans un délai raisonnable. Cette méthode d'échantillonnage est appréciable pour des classifications automatiques, par exemple, où la signification des classes est interprétée a posteriori. Sur le terrain, afin de rendre possibles et plus faciles les observations, les points échantillonnés sont distribués le long de transects. Plus pratique, ce plan d'échantillonnage perd toute signification quant à sa représentativité statistique. Une solution intermédiaire consiste à respecter des quotas de classes de paysage issus d'une classification supervisée. Ainsi, les transects répondent aux contraintes de terrain³, les points observés y sont positionnés de manière à s'approcher de la masse relative des principaux ensembles paysagers.

Les 189 points présentés sur la figure 2 ont fait l'objet de relevés au cours de la mission de 1995. La classification supervisée utilisée afin de définir les quotas se référait à des observations datant de 1989 et 1991. Nous procédions par estimation visuelle des taux de recouvrement de la végétation et du minéral pour renseigner les fiches de description/terrain. Cette méthode empirique donne de bons résultats lorsqu'on évalue les principaux types de paysages. Elle n'est plus valable pour une description fine des milieux. Un second plan d'échantillonnage doit alors être mis en place.

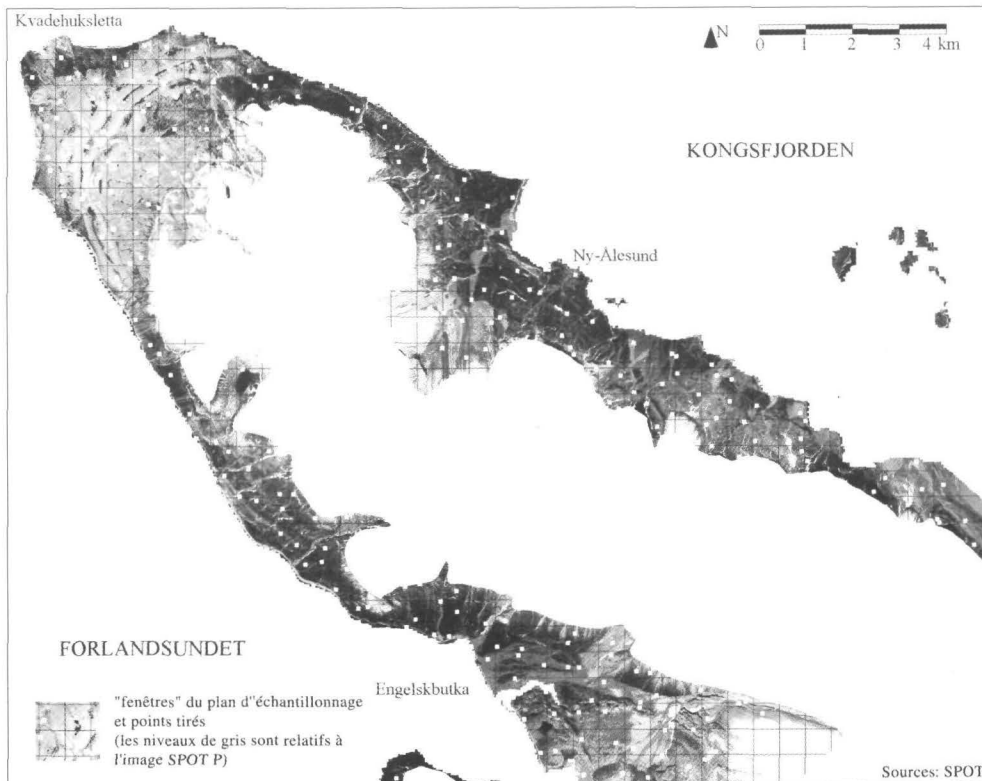


Figure 1 - Localisation des points échantillonnés selon un plan systématique non-aligné

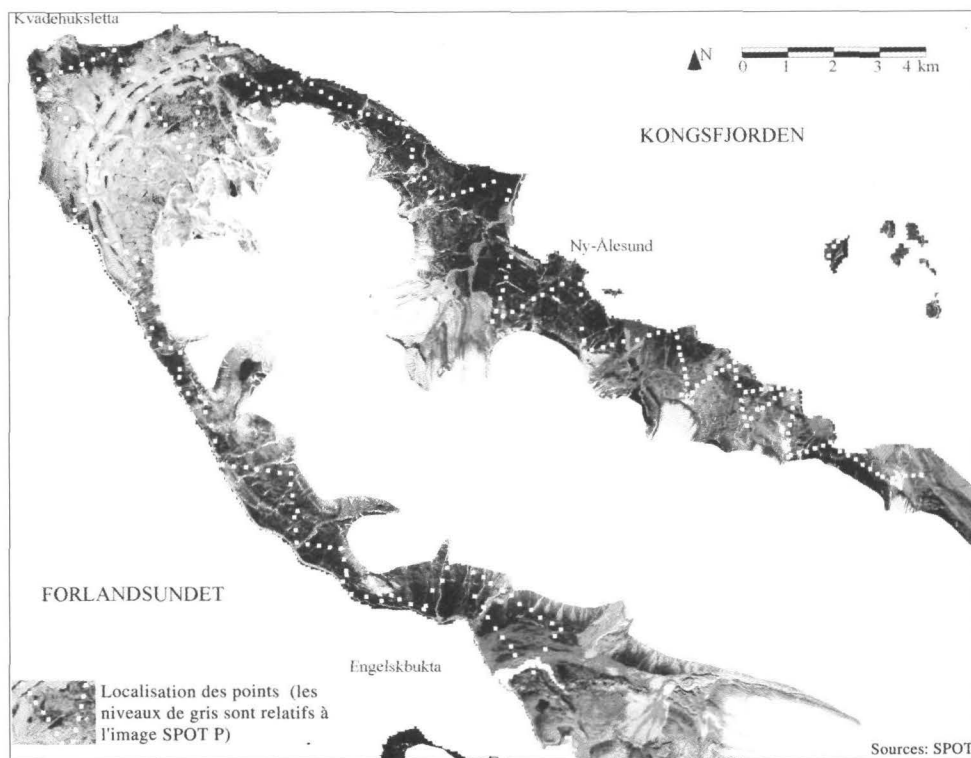


Figure 2 - Localisation des points échantillonnés selon un transect tenant compte des principaux ensembles paysagers

2. Échantillonnage des pixels : contraintes statistiques et technologiques

La méthode que nous présentons a pour but de définir une grille de lecture des pixels fondée sur une procédure statistique afin de réduire autant que faire se peut la part de l'aléatoire dans les mesures. La finalité de notre travail est d'utiliser les images satellitales – donc la radiométrie ou tout du moins la luminance apparente – pour cartographier en termes de probabilités des classes de paysages et les éléments qui les caractérisent. La typologie est issue d'une AFC et d'une CAH réalisées à partir d'observations fines de terrain (approche comparable à celle des phytosociologues, par exemple). C'est une démarche totalement différente de celle des classifications supervisées « classiques ». Pour simplifier, l'information radiométrique est utilisée afin de spatialiser (dans le sens de couvrir de manière continue) des données observées ponctuellement dont la résolution est inférieure à celle des pixels (il est également possible d'introduire d'autres données telles que l'altitude, l'orientation, la pente...).

À ce niveau de description on ne peut avoir recours à une estimation empirique du contenu des points échantillonnés. Ceux-ci doivent faire l'objet d'un véritable recensement exhaustif. De plus, et c'est une contrainte nouvelle pour la définition du plan d'échantillonnage, les pixels retenus doivent être aussi « purs » que possible. Nous entendons par « pur » un pixel pour lequel on observe à vue d'œil la même composition paysagère (par exemple, couverture continue de minéral ou de toundra rase...). Afin de réduire les risques d'erreur de localisation géographique et de chevauchement de milieux différents (fig. 3) nous élaborons une procédure de mesure de l'homogénéité locale (ou de l'hétérogénéité, comme on veut...) dont le principe consiste à dénombrer dans une fenêtre de 3 pixels de côté le nombre de points ayant la même valeur ou à peu près que celui situé au centre de la fenêtre. Plus le résultat est élevé, plus on a de chance de se situer dans une portion d'espace relativement « homogène » donc adaptée à nos observations. À titre expérimental, avant d'appliquer la méthode à l'ensemble de la presqu'île, nous procédons à des relevés aux alentours de la Base (fig. 4).

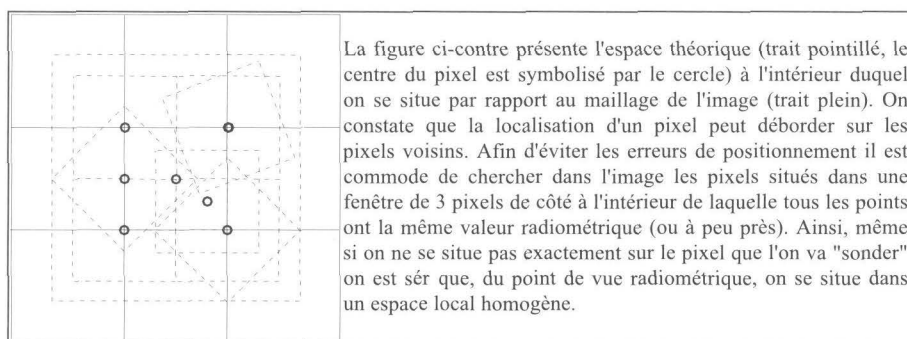
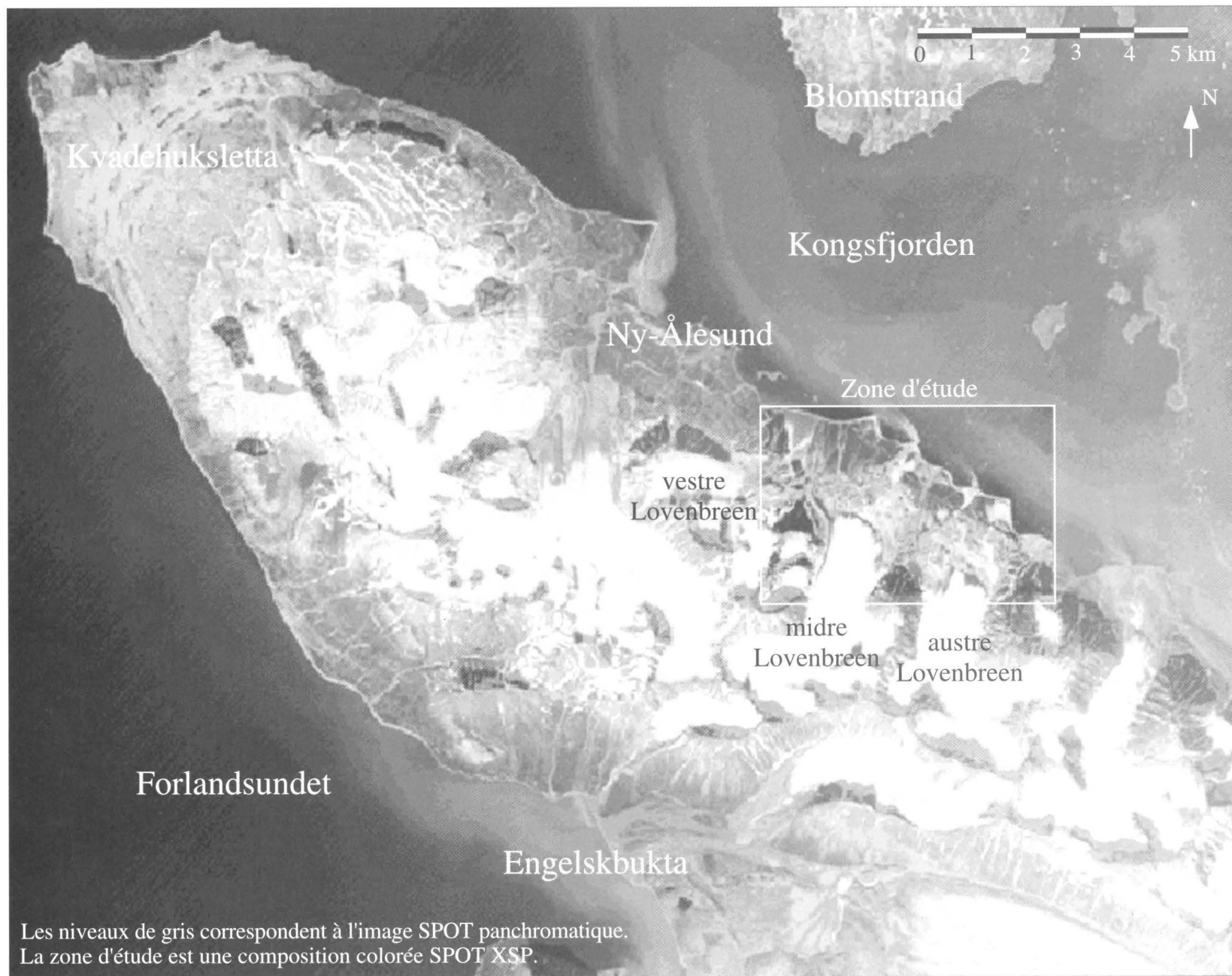


Figure 3 - Espace théorique de localisation d'un pixel

	XS brutes à 20 m			XS rééchantillonnées à 10 m					
	XS1	XS2	XS3	XS1_PPV	XS2_PPV	XS3_PPV	XS1_CC	XS2_CC	XS3_CC
0	3.4277	3.4277	3.4277	1.7213	1.7213	1.7213	1.7213	1.7213	1.7213
1	38.8674	33.4426	49.2027	0.6427	0.6259	0.6539	24.8994	22.4609	34.5585
2	24.7914	23.8674	27.2355	70.3987	66.4195	77.7925	23.1539	22.9154	29.0872
3	14.5827	13.8077	10.5067	16.6468	18.4203	13.6885	14.1487	15.2142	15.2627
4	8.6289	8.4575	3.7928	6.1047	5.7750	4.0071	12.2727	10.6595	7.5335
5	5.1788	5.7601	1.9076	3.1949	3.7127	1.5294	7.4106	6.2593	3.7668
6	2.6080	4.5007	1.3040	1.0153	2.5876	0.5551	7.2709	7.3677	2.8558
7	1.4083	2.7943	1.1997	0.1900	0.4005	0.0317	5.0410	5.2142	1.9877
8	0.4098	2.2057	0.6632	0.0745	0.2440	0.0112	2.2280	2.5671	0.9668
9	0.0969	1.7362	0.7601	0.0112	0.0931	0.0093	1.8536	5.6203	2.2597
Tot	100	100	100	100	100	100	100	100	100

Tableau 1 - Estimation des modifications introduites par la méthode de rééchantillonnage dans un test d'homogénéité locale (plus proche voisin (ppv) et convolution cubique (cc))



Les niveaux de gris correspondent à l'image SPOT panchromatique.
 La zone d'étude est une composition colorée SPOT XSP.

Figure 4 - Localisation de la zone d'observation autour de la Base

Rappelons que travaillant à partir d'images géoréférencées donc rééchantillonnées, la mesure d'homogénéité locale n'a que peu de rapport avec la luminance apparente des données brutes. Les valeurs « pseudo-radiométriques » que nous testons sont en fait très liées à la méthode de rééchantillonnage utilisée. Le tableau 1 ainsi que la figure 5 présentent les résultats d'un test que nous avons effectué afin d'avoir une idée précise de la perturbation du signal introduite par le rééchantillonnage. Les mesures d'homogénéité locale ont été effectuées sur une image SPOT XS brute de niveau 1A puis sur la même image rééchantillonnée à 10 m par la méthode du plus proche voisin et celle d'une convolution cubique. Il apparaît assez nettement des différences : la technique du plus proche voisin a tendance à minimiser la masse des extrêmes tout en exagérant la représentation des valeurs moyennes ; le calcul par convolution cubique opère un lissage de la distribution brute en assurant une sur-représentation des valeurs extrêmes. Concrètement, la méthode de convolution cubique – celle utilisée pour traiter les images du Svalbard⁴ – augmente sensiblement les surfaces « homogènes ». Malgré cela peu de pixels vérifient rigoureusement le critère d'homogénéité locale et ce même pour des milieux qui semblent « purs » à l'œil nu.

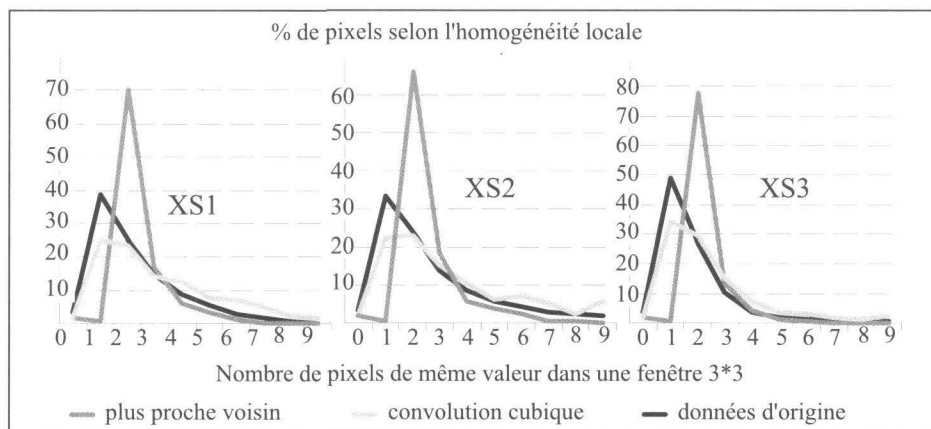


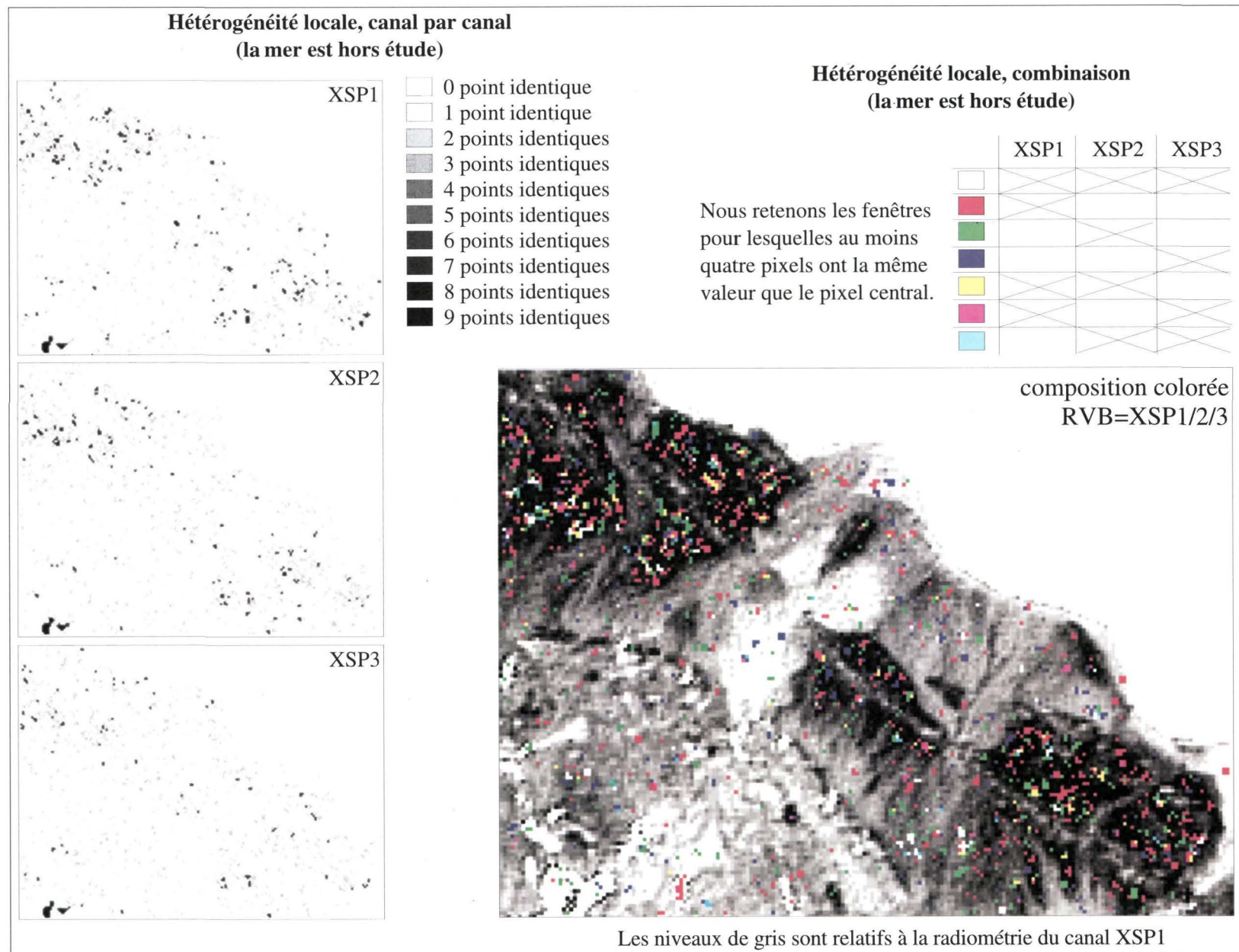
Figure 5 - Perturbation du signal selon la méthode de rééchantillonnage

Le tableau 2 présente les résultats obtenus sur une section de l'image autour de la base pour tous les milieux y compris la mer et les glaciers. Les taux sont encore plus faibles en ne retenant que les paysages de sandurs qui nous intéressent. La figure 6 est une composition colorée des résultats de ce test : les teintes primaires (rouge, vert, bleu) correspondent à des points qui vérifient les conditions sur un seul canal à la fois ; les teintes complémentaires (cyan, magenta, jaune) distinguent les pixels vérifiant les conditions sur deux canaux simultanément ; la couleur blanche précise les portions d'espace pour lesquelles les conditions sont vérifiées pour les trois canaux (la couleur noire correspondrait à l'inverse, nous avons préféré la remplacer par des niveaux de gris relatifs aux données XSP1 afin de faciliter l'interprétation de l'image).

	XSP1 %	XSP2 %	XSP3 %
Pas de points identiques	47,52	50,10	49,88
1 point identique	29,44	28,34	28,00
2 points identiques	12,86	12,23	11,75
3 points identiques	4,82	4,52	4,67
4 points identiques	1,99	1,61	2,20
5 points identiques	0,74	0,64	0,81
6 points identiques	0,23	0,15	0,27
7 points identiques	0,04	0,06	0,07
8 points identiques	0,03	0,03	0,03

Tableau 2 - Taux obtenus sur les critères d'homogénéité radiométrique locale

Figure 6 - Homogénéité locale des réponses spectrales autour de la Base



Les points que nous examinerons vérifieront pour les trois canaux simultanément le test d'homogénéité radiométrique locale avec au moins six pixels identiques parmi les neuf qui constituent la fenêtre testée. Ce sont ces derniers qu'il conviendrait de positionner sur les transects de cheminement milieu par milieu (cf. plus haut). Concrètement sur le terrain les pixels sont matérialisés par de la ficelle. Notre but étant de réaliser une description exhaustive des milieux, chaque tache délimitée est observée m² par m² à l'aide d'une grille de lecture phytosociologique⁵ (surface de 1 m² découpée en carrés de 10 cm de côté). Cette grille constitue l'unité, ou maille élémentaire de description. Nous réalisons ainsi un recensement des thèmes retenus. La distribution de fréquence des différentes informations est considérée par la suite comme un modèle empirique de probabilité utilisé afin de définir un plan d'échantillonnage fiable. On comprend aisément qu'il n'est pas possible de réaliser un recensement exhaustif de tous les points. Seuls quelques secteurs de paysages particuliers ayant un rapport avec le ruissellement seront finement observés pendant la mission de 1996. Au cours d'une prochaine mission les observations seront étendues à l'ensemble des paysages de la presqu'île.

Pour chaque milieu recensé nous disposons de 100 mailles élémentaires à l'intérieur desquelles les surfaces occupées par la végétation et le minéral ont été estimées à quelques cm² près. Nous appliquons un plan d'échantillonnage systématique stratifié non-aligné à cette population. Nous le rappelions plus haut, le sondage aléatoire simple, c'est-à-dire « le choix de *n* unités d'échantillonnage dans une population de *N* unités [...] tel que chacune des unités distinctes ait une chance égale d'être tirée » [6] est proscrit en raison de la nature spatiale des observations. La stratification des tirages élimine toute trace d'autocorrélation spatiale comme nous l'avions déjà signalé précédemment. La taille de l'échantillon est fixée selon un seuil de fluctuation autour des moyennes au fur et à mesure des tirages successifs (fig. 7). Les premiers donnent des moyennes très fluctuantes et des positions géographiques éparpillées dans le pixel. Assez rapidement la courbe a tendance à se stabiliser autour de la moyenne à mesure que le nombre de points augmente et que, par conséquent, le pixel se trouve mieux couvert spatialement.

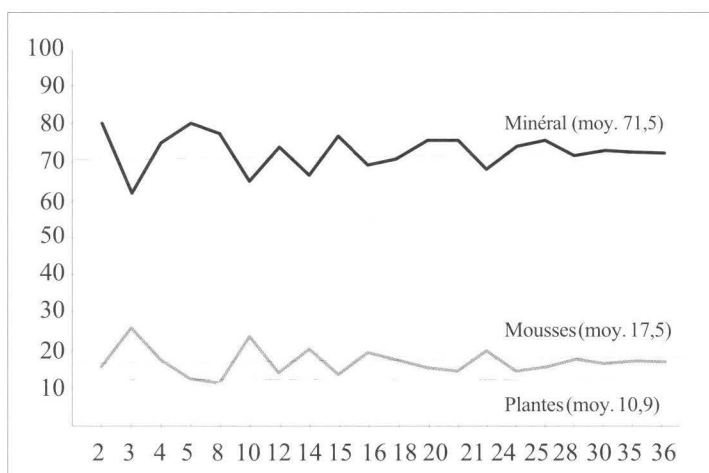


Figure 7 - Répartition des points échantillonnés au cours des tirages successifs : fluctuation autour de la moyenne (en % de la couverture d'un pixel)

Nous proposons à titre d'exemple d'application de ce qui vient d'être présenté, la cartographie d'un milieu de sandur obtenue selon des critères de distinction liés à deux dynamiques antagonistes de conquête végétale ou d'écoulement [11]. Les dix classes obtenues (fig. 8) correspondent à des profils différents et peuvent être regroupées en trois catégories : les milieux minéralisés représentent les secteurs les plus actifs des sandurs puisqu'ils correspondent aux principaux drains d'écoulement ; les milieux végétalisés s'opposent diamétralement aux premiers et correspondent à d'anciennes zones occupées par l'écoulement ; les milieux mixtes n'appartiennent pas aux deux secteurs précédents et illustrent l'antagonisme entre colonisation végétale et emprise de l'écoulement.

Le bilan végétal-minéral est en faveur du minéral puisque celui-ci domine ou conquiert 65 % de l'espace étudié. Le recul contemporain des glaciers alimente un écoulement concentré de plus en plus considérable, et de ce fait une quantité croissante d'eau doit circuler dans un nombre de moins en moins important de chenaux. Cette évolution répond à la logique de concentration de l'écoulement et de hiérarchisation du réseau hydrographique sur des marges englacées de milieux froids.

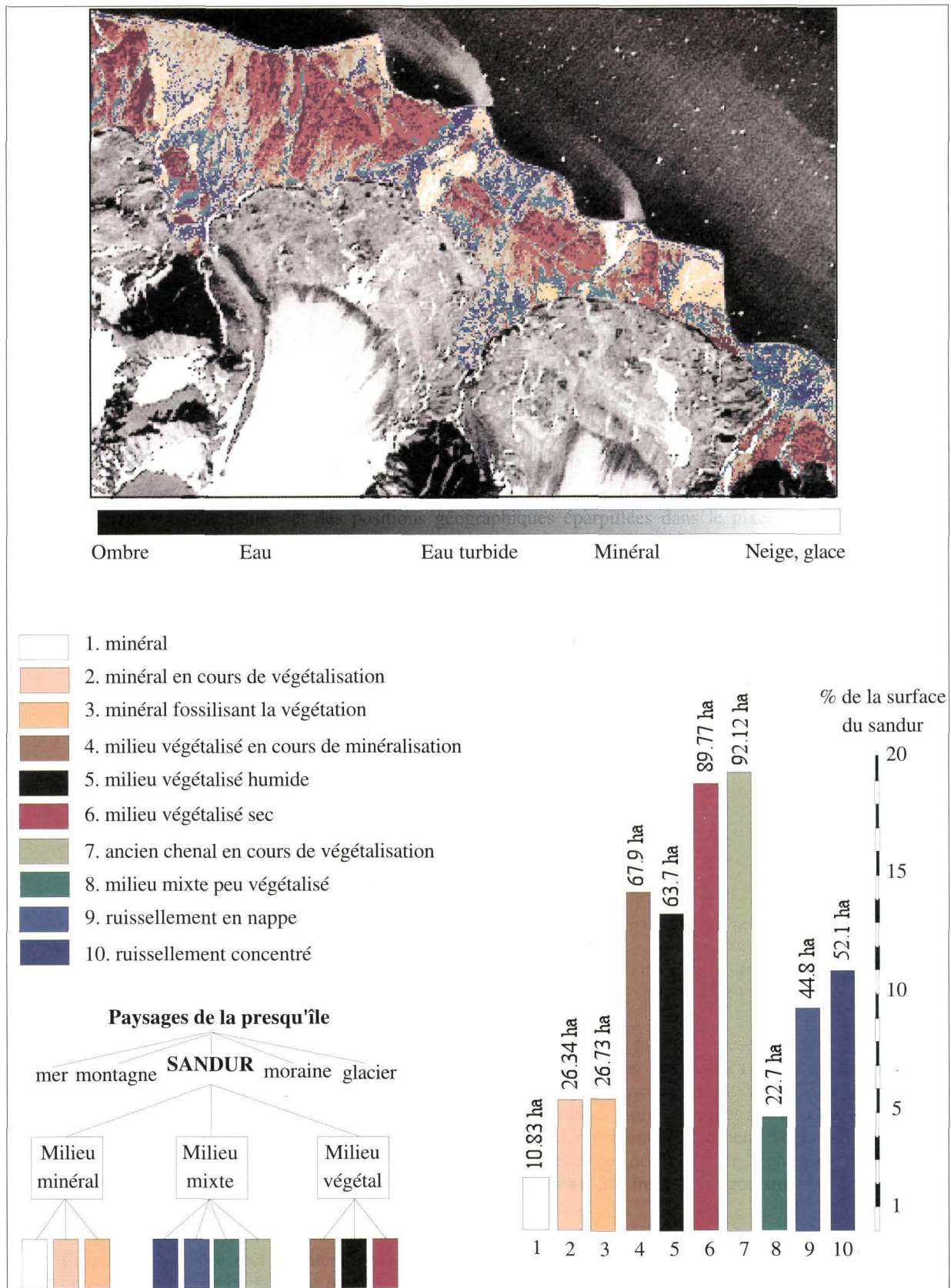


Figure 8 - Classification des milieux de sandur

Conclusion

Nous avons tenté d'apporter quelques éléments méthodologiques concernant les différents plans d'échantillonnage qu'il convient d'utiliser pour obtenir des échantillons représentatifs. Il en ressort que les contraintes de terrain et technologiques ne permettent pas de suivre exactement les résultats fondés sur les seules méthodes probabilistes. Cependant ces dernières fixent des critères vers lesquels il convient de tendre au cours de l'élaboration de l'échantillon. Ces procédures alourdissent considérablement les prétraitements et la pratique du terrain. Par conséquent, elles ne sont à envisager que dans un cadre précis d'analyse d'image tel que celui que nous avons présenté. Ce n'est certes pas le cas des traitements classiques en télédétection, par exemple, les classifications supervisées. En revanche, c'est une voie qui permet une assise statistique solide en vue d'une modélisation probabiliste des informations contenues dans les tachèles.

Références bibliographiques

- [1] BERTHOZ E., BROSSARD T., 1992 : « Les paysages du Svalbard vus du sol et vus du satellite: croisement des sources et enrichissements mutuels », *Photo-interprétation*, 3-4, pp. 113-122
- [2] BROSSARD T., DERRUELLE S., NIMIS P.L. et PETIT P., 1984 : « An interdisciplinary approach to vegetation mapping on lichen-dominated systems in high-arctic environment, Svalbard », *Phytocoenologia*, 12(4), pp. 433-453
- [3] BROSSARD T., JOLY D., 1994 : « Probability models, remote sensing and field observation: test for mapping some plant distributions in the Krossfjord area, Svalbard », *Polar Research*, 13, pp. 153-161
- [4] BROSSARD T., JOLY D., 1996 : « Using a GIS data base at a high resolution for geomorphological cartography of polar environment », 28^e congrès de l'UGI, La Haye, 4-10 août
- [5] BROSSARD T., DESSERTY G. et JOLY D., 1998 : « Le GPS comme source de données géographiques à grande échelle », *L'Espace Géographique*, pp. 23-30.
- [6] COCHRAN W.C., 1977 : « *Sampling techniques* », New-York, John Wiley & Sons
- [7] DYE R.H., 1975 : « *Restoration of Landsat image by discrete two-dimensional deconvolution* », Tenth International Symposium of Remote Sensing of Environment, Ann Arbor, Michigan
- [8] KEERSMAECKER M.L., 1987 : « Stratégie d'échantillonnage des données de terrain intégrées dans l'analyse des images satellitaires », *L'Espace Géographique*, n° 3, Paris, pp. 195-205
- [9] LAFFLY D., 1990 : « *Enrichissement des banques de données géographiques issues de la télédétection : combinaison et amélioration des images panchromatiques et multispectrales* », Association Internationale de Cartographie et Symposium international de cartographie thématique dérivée des images satellitaires, Paris, 2-4 octobre, pp. 20-25
- [10] LAFFLY D., 1995 : « *Evolutions et potentiels de l'espace comtois : recherche de méthodes par télédétection* », thèse, Université de Franche-Comté, 374 pages
- [11] MERCIER D., 1999 : « *Le ruissellement au Spitsberg. L'impact d'un processus azonal sur les paysages d'un milieu polaire* », thèse, Université de Clermont-Ferrand, dir. M.F. André, 532 pages

Notes

- 1- Nous n'évoquons pas ici le problème du géoréférencement des images et des points échantillons. Avec la précision demandée la meilleure solution sur le terrain consiste à se localiser à l'aide d'un GPS différentiel (Brossard & al, 96, 98). L'image SPOT XSP utilisée a été orthorectifiée et restaurée (Dye 1975, Laffly 1990, 95) par la société UNISFERE de Besançon.
- 2- Nous reviendrons par la suite plus en détail sur cette méthode.
- 3- Sans tenir compte des déplacements nécessaires afin de parcourir l'ensemble de la presqu'île (à pied lorsqu'on se trouve à proximité de la base, en bateau pneumatique pour les autres sites), ni du cumul des dénivelés, le cheminement le long des transects représente une distance de 110 km. Elle serait au moins dix fois plus importante avec l'autre méthode d'échantillonnage! La saison propice aux observations n'étant - dans le meilleur des cas - que de quelques semaines à cette latitude (début juillet à mi-août) on comprend aisément pourquoi la méthode par transect est inévitable.
- 4- Les images ont été restaurées c'est-à-dire que les valeurs utilisées pour la convolution cubique ne sont pas celles, discrètes, de luminance apparente du fichier brut mais un signal continu simulé à partir de la connaissance *a priori* de la fonction d'étalement du capteur (Point Spread Function).
- 5- La végétation étant très couvrante il est pratique d'avoir recours à une telle grille. En revanche, pour des milieux plus complexes (pelouses d'altitude, par exemple) elle n'est plus utilisable. Il est alors nécessaire de recourir à des échantillonnages linéaires de type Poissonet et Daget.