

Évaluation des surfaces naturelles par télédétection et enquête de terrain en Mauritanie

Vincent GODARD

Institut de Développement Rural (IDR), Université de Ouagadougou, 03 BP 7021
Ouagadougou 03, Burkina Faso

Lorsque l'on a pour objectif de décrire les formations végétales spontanées d'un secteur pastoral sahélien de plusieurs centaines ou milliers de kilomètres carrés, voire d'en évaluer les ressources fourragères, une connaissance exhaustive de la zone d'étude est souvent recherchée. Or, il n'est pas concevable de parcourir et reconnaître l'intégralité de telles surfaces ; seul un échantillonnage puis une extrapolation de ses résultats à l'ensemble du secteur pastoral peuvent permettre une compréhension globale du milieu.

L'enquête de terrain, comme pour n'importe quel inventaire d'occupation du sol, est l'élément de base de la phase de collecte. La télédétection en est un complément d'information, elle permet d'améliorer les estimations de superficie des différents thèmes issus de la seule enquête de terrain. L'enquête décrite dans cet article a eu lieu en octobre-novembre 1987 et mars-avril 1988.

Pour que les responsables sahéliens puissent un jour adapter la charge animale aux ressources du milieu et ainsi éviter surpâturage ou sous exploitation de celles-ci, il est indispensable de fiabiliser la phase d'élaboration de la cartographie et des statistiques d'occupation du sol et des ressources végétales (spontanées ou non). Une fois cette étape franchie, il est possible de réaliser des documents fiables ayant trait à la production et d'élaborer des projets de développement.

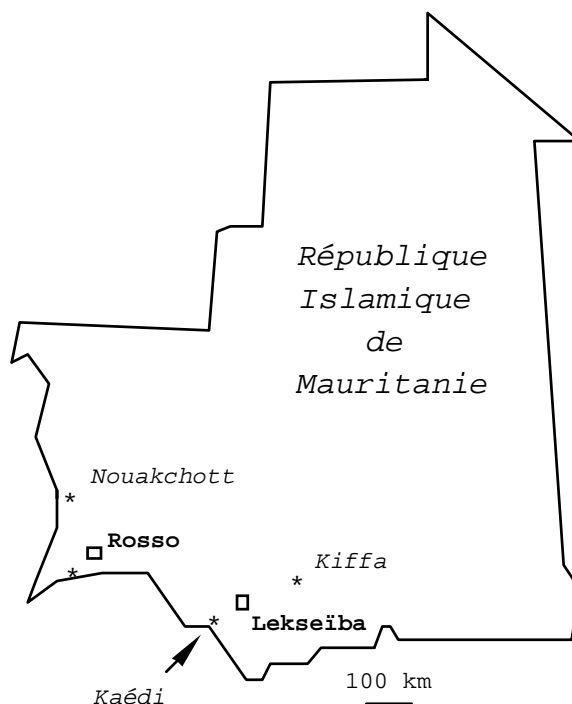
Pour atteindre cet objectif de représentativité et de fiabilité, nous nous sommes inspirés des méthodes pratiquées depuis quelques années en agriculture, que ce soit aux États Unis ou en Europe, pour l'évaluation de l'occupation agricole des terres (Houseman 1979, Bryant 1983, Fournier 1986, IAT 1989). Nous les avons adaptées aux inventaires de la végétation spontanée et au milieu tropical sec. Pour cela un certain nombre de problèmes devaient être résolus, à commencer par celui de la taille et de la répartition des unités d'enquête en fonction du temps d'inventaire envisagé.

Cette communication se propose donc après un descriptif de la méthodologie employée de présenter les résultats obtenus, résultats essentiellement statistiques dans la mesure où ce sont ceux pour lesquels l'amélioration est la plus facile à démontrer, puis, une discussion de ces résultats et des limites de la méthode sera présentée ; mais avant cela une description de l'aire d'étude est proposée.

1 Présentation de l'aire d'étude

Avec un peu plus d'un million de kilomètres carrés, la République Islamique de Mauritanie (RIM) est un des pays les plus vastes du continent africain, mais aussi un des plus désertiques. Les deux tiers de son territoire sont occupés par le désert du Sahara. Seule, sa frange sud échappe à ce domaine hyper-aride et permet à l'homme des activités telles que l'élevage et, plus aléatoirement, les cultures.

figure 1 - Localisation des zones d'enquête



La recherche méthodologique présentée ici s'est déroulée sur cette bande de territoire à l'intérieur de deux zones tests situées près des villes de Rosso dans le Trarza (zone pilote de Rosso) et de Kaédi dans le Gorgol (zone pilote de Lekseïba, cf. fig. 1).

Situé dans le domaine sahélien, notre champ d'expérimentation peut être caractérisé comme suit. Schématiquement, les précipitations sont comprises entre 100 et 5 à 600 mm par an, l'été est chaud et affecté d'une évaporation forte tandis que la période hivernale est fraîche ; les maxima de température interviennent aux périodes intermédiaires. Le domaine aérologique est également partagé durant l'année entre l'Harmattan en hiver et la mousson en été. Cependant, comme pour de nombreuses autres régions sahéliennes, depuis trente ans la dégradation des conditions climatiques est nettement perceptible sur les relevés pluviométriques.

Sur la période 1941-1970, la moyenne des précipitations a été de 295 mm à Rosso et de 373 mm à Kaédi. En revanche, sur la période 1971-1989, la moyenne

"trentenaire" est respectivement de 199 mm et de 245 mm, soit une baisse de plus de 30 p.100 entre ces deux périodes (Godard 1991 p.40).

De telles carences hydriques ne sont pas sans conséquences sur la végétation. Celle-ci est non seulement tributaire du déficit en volume des précipitations mais aussi de la raréfaction des jours de pluie. Notons enfin, que la saison des pluies concentre sur les mois de juillet, août et septembre 90 p.100 des précipitations (Godard 1991 p.28).

Dans la mesure où notre objectif est le suivi de la végétation spontanée, les dates des prises de vue et d'enquête sont tributaires des dernières pluies utiles. En effet, rappelons que nous souhaitons enquêter la végétation à son maximum de production et avant flétrissement. De ce fait, nous considérons que la dernière pluie utile est celle qui, supérieure ou égale à 20 mm (Boudet 1987 p.25), est suivie de deux décades sans pluies (Boudet 1981 p.9). Un modèle probabiliste de précipitation nous a permis d'estimer, pour les secteurs de Rosso et Kaédi, que les dates de prises de vue et d'enquête optimales se situaient un peu avant la mi-octobre (Godard 1991 p.51).

La végétation de ces secteurs nord-sahéliens est souvent caractérisé par une steppe "armée" aux peuplements arbustifs lâches et aux formations herbacées discontinues. Parfois des fourrés épineux plus denses se rencontrent dans les zones déclives aux conditions hydriques favorables, comme les lits d'oued. Les "arbres" et arbustes ont en général une taille comprise entre trois et six mètres avec une nette domination des légumineuses épineuses comme les acacias. La strate herbacée est dominée par des graminées pérennes des genres *Aristida* et *Cenchrus*.

En raison de la faiblesse des précipitations, les cultures pluviales sont absentes de la zone pilote de Rosso et subsistent à l'état de "traces" dans celle de Lekseïba. Seules, les cultures de décrues (Rosso et Lekseïba) ou les cultures irriguées (Rosso) ont une extension suffisante, dans ce secteur, pour être saisies par le plan d'échantillonnage. Toutefois, en raison des dates de prises de vue, seules les cultures irriguées étaient perceptibles sur l'image utilisée, car les zones de cultures de décrue étaient encore ennoyées à la même époque (octobre-novembre 1987). Elles n'ont pu être identifiées que lors de la mission de contrôle de saison sèche (mars-avril 1988).

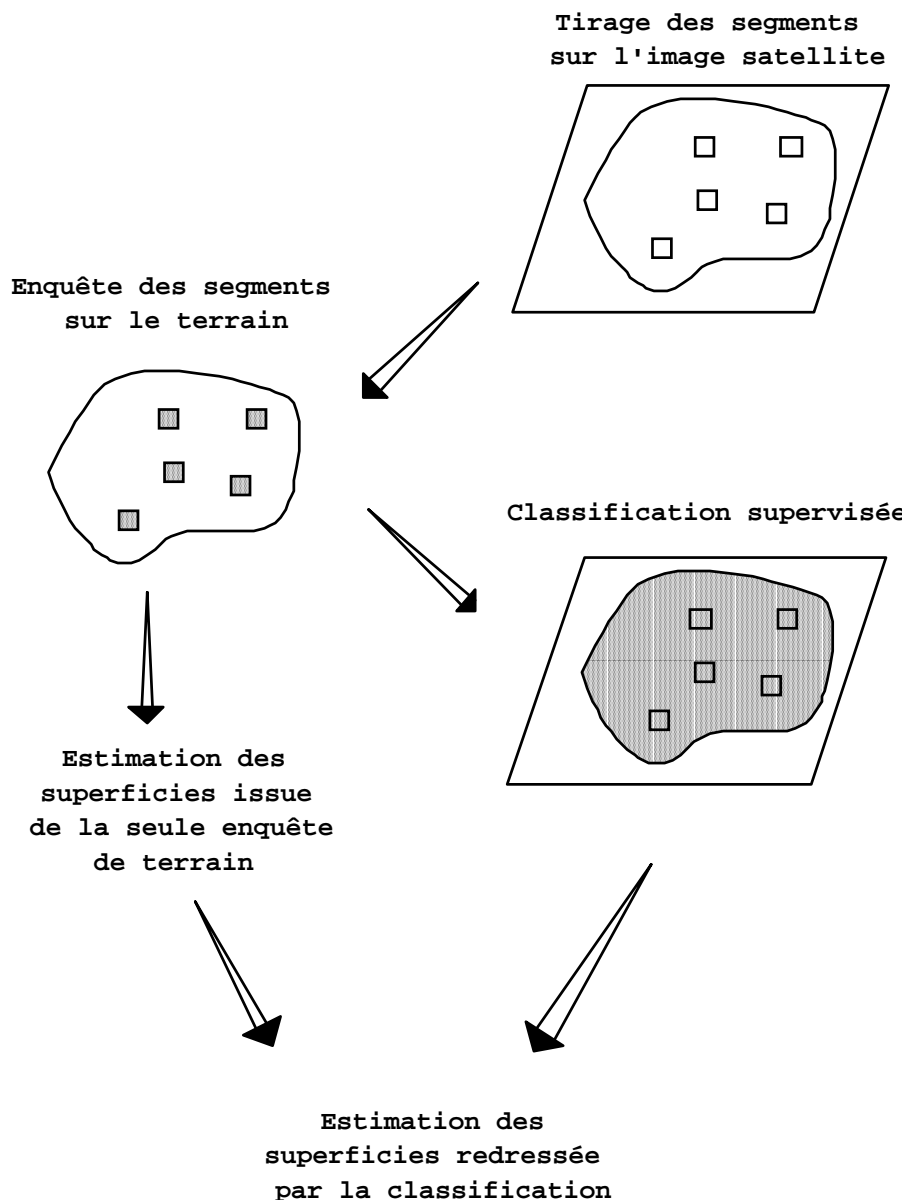
2 Méthode

2.1 Organisation générale

Dans la mesure où les missions de terrain sont toujours de courte durée, il est nécessaire de les préparer à l'avance. Ceci est d'autant plus indispensable que l'on souhaite recourir à un sondage pour guider la phase de collecte des données.

Lorsque l'on souhaite réaliser un sondage, il faut se préoccuper de ce qui peut constituer la base de sondage. En milieu agricole, dans les pays développés, il existe souvent des listes d'exploitants, ce n'est bien sûr pas le cas en milieu naturel. De plus en zone sahéenne, les seuls documents qui sont à jour sont en général les produits satellitaires, car les cartes et les photographies aériennes sont anciennes. Dans notre cas, elles datent bien souvent de la fin des années cinquante.

figure 2 - Déroulement de l'inventaire



L'imagerie satellitaire est donc le seul document qui peut nous servir de base de sondage. Comme nous souhaitons pouvoir calculer la précision de nos estimations de superficie, nous sommes obligés de recourir à un tirage aléatoire. Les unités d'enquête échantillonnées ne sont pas constituées de pixels indépendants mais de groupes de pixels jointifs appelés segments carrés. On parle donc de sondages aréolaires. La taille de ces segments est fonction de la complexité et de la diversité du paysage. Nous avons utilisé des segments de 36 ha sur les deux zones pilotes (cf. annexe).

Comme il est possible de le constater sur la figure 2, le tirage des segments et leur implantation s'effectuent sur l'image satellite. Celle-ci se présente en général sous la forme d'une composition colorée, mais il est possible de réaliser ces opérations sur un seul canal ou une image classée.

L'inventaire a lieu uniquement sur les segments dont la taille et l'emplacement ont été déterminés par le tirage sur l'image satellite. Il faut noter que seule l'enquête de terrain fournit les résultats statistiques d'occupation du sol. Enfin, si le tirage des segments est aléatoire et les résultats de classification cohérents avec l'enquête de terrain, alors il est possible d'améliorer les résultats issus du terrain par la télédétection.

L'enquêteur a à sa disposition sur le terrain des tirages à 1/50 000 de la composition colorée sur lesquels seront portés les segments, ainsi que des agrandissements à 1/10 000, de la composition colorée et de la classification automatique, de chacun des segments (Godard 1991 p.197 planche photo 6.1). Ces agrandissements sont complémentaires de la fiche de relevé, ils permettent de localiser précisément chaque poste d'occupation du sol inventorié. Notons toutefois que le tracé des limites, sur ces agrandissements, entre les différentes formations végétales spontanées n'est pas aussi aisé qu'en milieu agricole où l'emprunte du parcellaire est normalement beaucoup plus nette.

2.2 Réalisation de l'enquête de terrain

L'inventaire se fait sur les parcelles naturelles ou cultivées présentes à l'intérieur du segment. Celles-ci doivent présenter une certaine homogénéité et une certaine extension pour être localisables. Notons, que la localisation sur le terrain est un problème important en milieu naturel ; on aura donc tout intérêt à utiliser un système d'aide au positionnement comme les navigateurs par satellite de type GPS (*Global Positioning System*). Il est en partie possible d'apprécier l'homogénéité et l'étendue des parcelles avant d'être sur le terrain grâce à la consultation de préclassifications de l'image satellitaire. La préclassification est une classification automatique réalisée avant d'aller sur le terrain. Celle-ci est un bon descripteur préalable de l'imbrication thématique au niveau du segment.

La collecte de l'information sur la parcelle consiste en plusieurs descriptions synthétiques en relation avec la perception que peut en avoir le satellite à savoir, description topographique (pente, position, etc.), géopédologique (matériaux, texture et couleur de surface, humidité, etc.), sociologique (occupation humaine, fréquentation animale, etc.) et enfin végétale. Pour l'inventaire présenté, la description de la végétation est d'ordre physionomique avec une estimation du taux de couverture. Enfin, il peut être envisagé une quantification de la biomasse herbacée par coupe et pesée sur ces mêmes parcelles.

2.3 Dépouillement et traitements

L'intérêt du sondage aléatoire est de non seulement permettre l'estimation de superficie de chacun des thèmes mais surtout d'estimer la précision de cette estimation. Pour cela, il faut connaître l'occupation du sol par thème et par segment. En milieu agricole, le parcellaire est suffisamment net pour que l'on puisse directement mesurer la superficie de chaque parcelle, en tenant compte bien sûr des erreurs liées aux pixels de

bordure. Il n'en va pas de même en milieu naturel où le passage d'un thème à l'autre se fait en continu, sans rupture nette.

La superficie de chaque thème au niveau des segments enquêtés est déterminée à partir d'une grille de points superposée à un agrandissement du segment (cf. annexe) que l'enquêteur a complété sur le terrain. En l'absence de photographie aérienne, l'agrandissement provient d'un traitement de l'image satellite. C'est en général une composition colorée, mais parfois c'est le résultat d'une préclassification.

La proportion de chaque thème par segment est ensuite calculée (Godard 1991 pp.275-279). Elle sert d'une part à piloter les classifications supervisées (par les données collectées sur le terrain) et d'autre part, à fournir une première estimation des superficies. Comme nous le verrons plus loin, c'est l'enquête de terrain qui fournit les estimations de superficie alors que la télédétection n'intervient que comme complément d'information.

Les formulaires qui vont être présentés sont issus des méthodes d'inventaire que le SCEES a testées en milieu agricole tempéré (Fournier 1986) et peuvent être utilisés dans deux cas particuliers. Ils peuvent servir soit, à fournir très rapidement des estimations provisoires soit, parce qu'il n'y a pas de données satellitaires récentes pour redresser les estimations de superficie issues du seul terrain. Nous n'avons pas pour notre part testé les effets probables d'une dérive cartographique et statistique due à des classifications supervisées issues de données satellitaires trop anciennes.

- Estimation de la superficie du thème k à partir de la seule enquête de terrain :

$$S_k = S_{zp} * \frac{1}{m} * \sum_{i=1}^m P_{ki}$$

et

$$P_{ki} = \frac{n_{ki}}{N_i}$$

avec :

P_{ki} proportion d'occupation du sol du thème k dans le segment z_i ;

n_{ki} nombre de pixels du thème k tirés dans le segment z_i ;

N_i nombre total de pixels tirés dans le segment z_i ;

m nombre de segments inventoriés ;

S_{zp} superficie de la zone pilote ;

S_k superficie estimée du thème k sur la zone pilote.

- Variance de l'estimation de la superficie S_k :

$$\text{Var}(S_k) = S_{zp}^2 * \frac{1}{m} * \frac{1}{m-1} * \frac{M-m}{M} * \sum_{i=1}^m (P_{ki} - \bar{P}_k)^2$$

avec :

\bar{P}_k estimation de la proportion moyenne de la catégorie k ;

M nombre total de segments couvrant la zone pilote ;

$\frac{M-m}{M}$

très proche de 0,99 ou 0,98 dans la mesure où le taux de sondage est de l'ordre de 1 à 2 p.100.

Notons que comme l'indique Ph. Fournier (Fournier 1986 p.28) : "*Dans le cas d'un sondage aléatoire à un degré, avec des probabilités égales et sans remise, on utilise la formule de la variance d'une somme. Il n'est pas tenu compte du fait que le tirage des segments est plus ou moins stratifié, cet estimateur de la variance d'une somme est, en conséquence, calculé avec une surestimation certaine.*"

Muni de ces formulaires, il est maintenant possible de déterminer les estimations de superficie et la précision de ces estimations avant de réaliser la cartographie d'occupation du sol, si toutefois il existe des données satellitaires numériques satisfaisantes.

- Classification supervisée de l'image satellite

Les classifications supervisées utilisées pour réaliser la cartographie d'occupation du sol sont du type "maximum de vraisemblance". Elles ont deux fonctions. La première est de fournir un document visuel qui permet de localiser instantanément la répartition spatiale des différents thèmes, c'est bien sûr la carte. La deuxième est celle qui nous intéresse ici à savoir, la surface cartographiée pour chacun des thèmes. Ces surfaces obtenues par comptage sur la classification supervisée fournissent l'information exhaustive complémentaire qui est utilisée pour redresser les résultats obtenus à partir de la seule enquête de terrain.

Bien que nous ne développons pas l'aspect cartographie dans cet article, il nous faut toutefois ajouter que les classifications n'ont pas besoin d'être parfaites pour permettre un redressement satisfaisant des estimations de superficie. Nous verrons plus loin que l'amélioration de la précision des estimations est très nette alors que la précision globale des classifications n'est que de 86,7 p.100 à Rosso et de 92,3 p.100 à Lekseïba. La précision globale des classifications est obtenue en divisant la somme des pixels bien classés, lue sur la diagonale de la matrice de confusion, par la somme des pixels contrôlés.

- Estimation par la régression

Les formulaires que nous proposons maintenant sont issus en grande partie de *Théorie et pratique des sondages* (Desabie 1966 pp.215-219). A la différence des formulaires précédents, la mise en œuvre de ceux-ci est un peu plus longue, car en plus de la proportion d'occupation du sol estimée d'après le terrain, proportion d'ailleurs convertie ici en pixels, il faut avoir réalisé les classifications pour disposer des statistiques de chaque thème au niveau des segments et au niveau de la zone d'étude et cette fois-ci sur la carte.

Il n'est pas possible d'utiliser de manière fiable les seuls résultats de la classification supervisée comme statistiques d'occupation du sol. En effet, comme la classification n'est jamais parfaite, l'évaluation directe des superficies sur la carte est entachée d'un biais inconnu lié au choix des parcelles d'initialisation et à celui de la méthode de classification utilisée (Meyer-Roux 1981 pp.10-12). Par contre l'information fournie par le satellite est exhaustive : elle couvre l'ensemble de la zone d'étude.

Une régression linéaire est effectuée entre les données de terrain, thème par thème au niveau du segment, et leurs homologues obtenus par classification, toujours par segment. L'estimateur par la régression permet de tenir compte de la façon dont l'échantillon s'est comporté par rapport à l'ensemble de la zone relativement aux données satellitaires et de redresser l'estimation fournie par la seule enquête de terrain.

- Estimateur de superficie $\bar{y}_{reg k}$ au niveau du segment :

$$\bar{y}_{reg k} = \bar{y}_k + a \left(\frac{X_k}{M} - \bar{x}_k \right)$$

avec :

$\bar{y}_{reg k}$ superficie moyenne par segment de la catégorie k estimée par la régression ;

\bar{y}_k moyenne de la catégorie k, sur le terrain, au niveau des segments ;

a pente de la droite estimée comme suit :

$$a = \frac{\sum_{i=1}^m (x_{ki} - \bar{x}_k)(y_{ki} - \bar{y}_k)}{\sum_{i=1}^m (x_{ki} - \bar{x}_k)^2}$$

X_k valeur du thème k pour l'ensemble de la zone pilote sur la classification ;

m nombre de segments dans l'échantillon ;

- M nombre total de segments compris dans la zone pilote ;
- x_{ki} nombre de pixels comptés sur la classification pour le segment z_i ;
- \bar{x}_k moyenne de la catégorie k , sur la classification, au niveau des segments ;
- y_{ki} nombre de pixels comptés sur le relevé de terrain pour le segment z_i .

- Estimateur de superficie $Y_{reg\ k}$ au niveau de la zone pilote :

$$Y_{reg\ k} = M * \bar{y}_{reg\ k}$$

- Variance de l'estimateur $\bar{y}_{reg\ k}$ au niveau du segment :

Le calcul de la variance de l'estimateur $\bar{y}_{reg\ k}$ est destiné à déterminer la précision de cette estimation. Comme le fait remarquer J. Meyer-Roux (Meyer-Roux 1981 p.12) : "si X_k n'a pas de variance, $Y_{reg\ k}$ en a une car la droite des moindres carrés servant à le définir n'est qu'une estimation de la droite réelle inconnue.". En effet, X_k est mesuré sur la classification au niveau de la zone pilote, il n'a donc pas de variance.

Il est possible de déterminer la variance associée à l'estimation $\bar{y}_{reg\ k}$ de la façon suivante :

$$\text{Var}(\bar{y}_{reg\ k}) = \frac{(M - m)}{M} * \frac{1}{m} * \frac{1}{m - 2} * \sum_{i=1}^m e_{ki}^2$$

où
$$e_{ki}^2 = [y_{ki} - \bar{y}_k - a(x_{ki} - \bar{x}_k)]^2$$

Il est possible d'approcher cette variance par la formule suivante (Grobras 1987 p.139) :

$$\text{Var}(\bar{y}_{reg\ k}) = \frac{M - m}{M} * (1 - r^2) * \text{Var}(\bar{y}_k)$$

avec :

- r coefficient de corrélation entre les données de terrain et celles de la classification pour le thème k et par segment.

Nous n'avons eu recours à cette formule que pour effectuer le contrôle des résultats.

- *Variance de l'estimateur* $Y_{reg k}$ *au niveau de la zone pilote :*

$$\text{Var}(Y_{reg k}) = M^2 * \text{Var}(\bar{Y}_{reg k})$$

3 Résultats

Les premiers résultats que nous publions sont ceux obtenus à l'issue de la seule enquête de terrain, c'est-à-dire sans redressement par les données satellitaires numériques.

tableau 1 - Superficies estimées à partir de la seule enquête de terrain (zone pilote de Rosso)

Thèmes	Sigles	Superficies		Ecart types		CV	
		ha	%	ha	%		%
Surfaces en eau	Se	777	1,9	528	68,3		
Cultures irriguées	Ci	1 143	2,7	1 030	90,1		
Steppes arbustives hydromorphes	Sah	415	0,9	551	130,4		
St. herbeuses très claires des dépr. hydro.	Shy	2 827	6,8	1 445	51,1		
Sols nus sablo-limoneux	Sns	12 425	29,7	2 819	22,7		
Sols nus argileux	Sna	3 040	7,3	1 465	49,5		
Sables vifs	Sv	484	1,2	336	69,3		
St. herbeuses très claires des dunes	Shd	11 816	28,3	3 206	27,1		
St. herb. des ensablements et interdunes	Shi	7 613	18,2	2 265	29,7		
St. arbustives des ensablements	Sae	951	2,3	579	60,9		
Steppes "boisées" gommeraies	Sb	288	0,7	228	79,2		
Surface totale		41 779	100				

La façon la plus simple d'évaluer la qualité des résultats est d'en analyser les coefficients de variation (CV). Rappelons que le coefficient de variation s'obtient ici en divisant l'écart-type par la superficie, puis il est converti en pourcentage. Il permet d'apprécier et de comparer la précision relative de chaque thème.

Nous avons réparti les coefficients de variation en trois groupes. Il y a le groupe des thèmes dont le coefficient de variation est compris entre 10 p.100 et 30 p.100, par exemple pour Rosso, il s'agit des sols nus sablo-limoneux et à Lekseïba, des steppes herbeuses. On constate qu'ils occupent une superficie en général supérieure à 20 p.100 de la zone pilote.

Ensuite, il y a les thèmes dont le coefficient de variation est voisin de 50 p.100. Il s'agit par exemple à Rosso des sols nus argileux et à Lekseïba des prairies aquatiques. Leur superficie est inférieure à 10 p. 100

Enfin, certaines catégories d'occupation du sol ont un coefficient de variation qui peut atteindre 75 p.100. Leur superficie semble inférieure à 1 p.100.

tableau 2 - Superficies estimées à partir de la seule enquête de terrain (zone pilote de Lekseïba)

Thèmes	Sigles	Superficies		Ecart types		CV	
		ha	%	ha	%		
Surfaces en eau	Se	529	1,3	322	60,9		
Prairies aquatiques	Pa	927	2,2	499	53,8		
Steppes arbustives hydromorphes	Sah	359	0,9	296	82,4		
Steppes arbustives fermées	Saf	2 090	4,9	1 437	68,8		
Steppes arbustives	Sa	9 423	22,5	1 690	18,0		
Steppes herbueses	Sh	12 257	29,2	1 883	15,4		
Steppes arbustives denses dégradées	Sad	10 821	25,9	1 752	16,1		
Sols nus	Sn	3 929	9,4	1 060	27,0		
Affleurements rocheux	Af	1 568	3,7	1 057	67,4		
Surface totale		41 943	100				

On notera que la surface totale des deux zones pilotes n'est pas rigoureusement identique. Le fichier de Rosso était une matrice de 1022 pixels de côté, alors que celui de Lekseïba en comprenait 1024.

De l'analyse de ces résultats, il ressort que le plan de sondage mis en place n'est satisfaisant que pour les thèmes dont la superficie est importante, en général supérieure à 20 p.100 de la surface de la zone pilote. Pour ceux dont la superficie est inférieure à 5 p.100, les estimations sont trop imprécises pour bâtir une action d'aménagement ou de protection. Cependant, elles offrent un ordre de grandeur suffisamment précis pour fournir les premières estimations au sortir de l'inventaire. Cantonnées à cette fonction, elles en jouent pleinement le rôle. Le problème se pose tout autrement si, en l'absence de données satellitaires numériques récentes, voire contemporaines de l'enquête terrain, il faut se satisfaire de ces seules évaluations.

tableau 3 - Superficies estimées par régression (zone pilote de Rosso)

Thèmes	Sigles	Superficies		Ecart types		CV	
		ha	%	ha	%		
Surfaces en eau	Se	677	1,6	59	8,8		
Cultures irriguées	Ci	307	0,7	50	16,4		
Steppes arbustives hydromorphes	Sah	755	1,8	150	19,9		
St. herbueses très claires des dépr. hydro.	Shy	5 314	12,7	927	17,5		
Sols nus sablo-limoneux	Sns	10 363	24,8	505	4,9		
Sols nus argileux	Sna	1 703	4,1	285	16,7		
Sables vifs	Sv	622	1,5	52	8,3		
St. herbueses très claires des dunes	Shd	10 996	26,3	478	4,4		
St. herb. des ensablements et interdunes	Shi	7 952	19,0	1243	15,6		
St. arbustives des ensablements	Sae	1 141	2,7	360	31,5		
Steppes "boisées" gommeraias	Sb	239	0,6	36	15,2		
Total partiel		40 068	95,9				
Surfaces non affectées		1 711	4,1				
Surface totale		41 779	100				

On constatera qu'un peu moins de 5 p.100 des surfaces des zones pilotes ne sont pas affectées à l'issue des calculs d'estimation par l'estimateur par la régression. Une pratique courante veut que, pour une diffusion autre qu'à des fins de recherche, ce reliquat lié au mode de calcul soit ventilé entre toutes les catégories présentes.

Dès lors que l'on dispose de données satellitaires numériques suffisamment récentes pour redresser les estimations de la seule enquête de terrain, la précision des résultats est toute autre. Rappelons que ces données numériques servent non seulement

à redresser les estimations de terrain mais aussi à produire la cartographie d'occupation du sol et des ressources pastorales s'il y a lieu (De Wispelaere 1989).

Comme il est possible de le constater à la lecture de cette série de tableaux, la précision relative, au travers du coefficient de variation, s'est très nettement accrue. Alors que sans l'estimateur par la régression, il n'y avait aucun thème dont le coefficient de variation était inférieur à 15 p.100, ils sont maintenant quatre à Rosso (cf. tab. 3) et sept à Lekseïba (cf. tab. 4). De même, il ne reste plus qu'un seul thème dont le coefficient de variation est supérieur à 30 p.100, les steppes arbustives des ensablements à Rosso (cf. tab. 3).

tableau 4 - Superficies estimées par régression (zone pilote de Lekseïba)

Thèmes	Sigles	Superficies		Ecart types		CV
		ha	%	ha	%	
Surfaces en eau	Se	2 266	5,4	83	3,7	
Prairies aquatiques	Pa	2 199	5,2	269	12,2	
Steppes arbustives hydromorphes	Sah	1 113	2,7	141	12,7	
Steppes arbustives fermées	Saf	1 188	2,8	256	21,6	
Steppes arbustives	Sa	8 431	20,1	831	9,9	
Steppes herbeuses	Sh	9 126	21,8	946	10,4	
Steppes arbustives denses dégradées	Sad	11 630	27,7	961	8,3	
Sols nus	Sn	3 881	9,3	925	23,8	
Affleurements rocheux	Af	608	1,4	34	5,6	
Total partiel		40 442	96,4			
Surfaces non affectées		1 501	3,6			
Surface totale		41 943	100			

La précision assez moyenne de ce thème est en grande partie liée à une cartographie peu performante (Godard 1991 p.303). Comme celle-ci est utilisée pour redresser les estimations issues du terrain, le gain de précision est insuffisant. Le coefficient de variation passe de 60,9 p.100, sans l'estimateur par la régression (cf. tab. 1), à 31,5 p.100 avec l'utilisation de l'estimateur par la régression (cf. tab. 3).

Si l'on compare globalement l'amélioration de la précision liée à l'utilisation de l'estimateur par la régression, donc à l'apport des données satellitaires pour redresser les estimations de terrain, on constate que pour les deux zones pilotes les coefficients de variation ont beaucoup diminué. Si on calcule par zone pilote la moyenne des coefficients de variation, on constate qu'à Rosso ils passent de 61,5 p.100, sans l'estimateur par la régression, à 14,5 p.100 avec. De même à Lekseïba, ils passent de 45,5 p.100 à 12,0 p.100 dans les mêmes conditions. Dans les deux cas, la moyenne des coefficients de variation est inférieure à 15 p.100.

Quels sont les thèmes qui semblent le plus en avoir profité ? Comme on pouvait s'y attendre, ce sont logiquement les thèmes pour lesquels il y a une bonne corrélation entre le terrain et la classification et, phénomène particulièrement intéressant dans le cas présent, ce sont les thèmes de faible superficie, ceux que l'enquête de terrain du fait de son faible taux de sondage restituait les moins bien. A titre d'exemple, il faut bien sûr citer les steppes arbustives hydromorphes de Rosso qui passent ainsi d'un coefficient de variation de 130,4 p.100 (cf. tab. 1) à un coefficient de 19,9 p.100 pour une superficie de 1,8 p.100 (cf. tab. 4), superficie d'ailleurs largement sous estimée par le seul terrain.

L'apport de la télédétection a également permis de recadrer l'estimation des superficies des cultures irriguées. A l'issue de la seule enquête de terrain, les cultures irriguées sont créditées de 2,7 p.100 de la superficie de la zone pilote de Rosso, avec un coefficient de variation de 90 p.100 il est vrai (cf. tab. 1), alors qu'une fois redressée par les données satellite, cette superficie est ramenée à 0,7 p.100 (CV de 16,4 p.100, cf. tab. 3). La télédétection permet donc d'amortir les effets d'un plan d'échantillonnage peu adapté aux thèmes de faible extension, donc peu représentatifs du paysage mais dont l'importance est parfois capitale en terme de ressource.

Tout comme il est possible de contrôler l'efficacité d'un sondage stratifié par rapport à un sondage élémentaire, il est possible de contrôler l'efficacité des estimations par la régression par rapport à une estimation "élémentaire". Si l'on identifie l'estimation élémentaire à une estimation sans apport de la télédétection, nous considérons à la suite de J. Meyer-Roux (Meyer-Roux 1981 p.13) que l'efficacité Ef de la télédétection peut être mesurée par le rapport :

$$Ef = \frac{1}{1 - r^2}$$

où r est le coefficient de corrélation obtenu lors de la régression entre données de terrain et classification.

Une efficacité Ef égale à 3,4 correspond à un coefficient de corrélation de 0,84, exemple des steppes arbustives denses dégradées à Lekseïba (cf. tab. 6). Elle signifie que, sans image satellite et donc sans l'apport de l'estimateur par la régression, il faudrait 3,4 fois plus de segments pour faire l'inventaire de ce thème et obtenir la même précision, c'est-à-dire un coefficient de variation de 8,3 p.100.

Nous constatons donc que pour Rosso (cf. tab. 5), en fonction du coefficient de corrélation r, le gain de précision pour 20 segments par rapport au sondage élémentaire, va de 1,5 fois pour les steppes arbustives hydromorphes à 50 fois pour des thèmes comme les surfaces en eau ou les steppes "boisées" gommeraias. Il en va de même à Lekseïba.

tableau 5 - Efficacité de la télédétection (zone pilote de Rosso)

Thèmes	Sigles	Superficies %	r	Ef
Surfaces en eau	Se	1,6	0,99	50,3
Cultures irriguées	Ci	0,7	0,99	50,3
Steppes arbustives hydromorphes	Sah	1,8	0,58	1,5
St. herbeuses très claires des dépr. hydro.	Shy	12,7	0,78	2,6
Sols nus sablo-limoneux	Sns	24,8	0,98	25,3
Sols nus argileux	Sna	4,1	0,98	25,3
Sables vifs	Sv	1,5	0,99	50,3
St. herbeuses très claires des dunes	Shd	26,3	0,99	50,3
St. herb. des ensablements et interdunes	Shi	19,0	0,85	3,6
St. arbustives des ensablements	Sae	2,7	0,80	2,8
Steppes "boisées" gommeraias	Sb	0,6	0,99	50,3

A Rosso, la précision moyenne (calculée sans tenir compte de l'importance relative en superficie des différents thèmes étudiés) est 28 fois plus élevée qu'en l'absence de correction par la régression, elle est 11 fois plus élevée à Lekseïba pour

l'ensemble des catégories, en raison d'une classification un peu moins performante. Le coefficient de corrélation moyen (r) est de 0,86 contre 0,90 à Rosso. Autrement dit, sans l'apport des données satellitaires pour redresser les estimations, il aurait fallu 28 fois plus de segments à Rosso et 11 fois plus à Lekseïba pour obtenir la même précision.

On peut être surpris par le haut niveau d'efficacité de la méthode en regard de ce qui est obtenu en milieu agricole tempéré. A titre de comparaison, les résultats obtenus par R. Pastorelli et J.C. Porchier sur l'utilisation des terres dans le département de l'Indre (Pastorelli 1989 p.55) sont moins efficaces, leurs coefficients de corrélation sont en général plus faibles. Il faut peut-être en rechercher l'explication dans une nomenclature d'occupation du sol plus fine, voire peut-être trop fine par rapport aux possibilités de discrimination radiométrique liées à une date de prise de vue très tardive, aux dires des auteurs.

tableau 6 - Efficacité de la télédétection (zone pilote de Lekseïba)

Thèmes	Sigles	Superficies %	r	Ef
Surfaces en eau	Se	5,4	0,96	12,8
Prairies aquatiques	Pa	5,2	0,85	3,6
Steppes arbustives hydromorphes	Sah	2,7	0,88	4,4
Steppes arbustives fermées	Saf	2,8	0,97	16,1
Steppes arbustives	Sa	20,1	0,88	4,4
Steppes herbeuses	Sh	21,8	0,87	4,1
Steppes arbustives denses dégradées	Sad	27,7	0,84	3,4
Sols nus	Sn	9,3	0,52	1,4
Affleurements rocheux	Af	1,4	0,99	50,3

S'il est maintenant avéré que l'on ne peut plus faire l'économie des données satellitaires numériques dans le cadre d'un inventaire d'occupation du sol, il convient, au regard des avantages, d'évaluer les lacunes et imperfections de la méthode, ou du moins des exemples présentés ici.

Discussion

Redresser les estimations de superficies issues d'un échantillonnage aléatoire du terrain, à l'aide de données satellitaires numériques, a de nombreux avantages. Cela permet de limiter considérablement le nombre d'unités d'enquête à inventorier tout en conférant aux estimations une très grande précision.

En effet, lorsque l'on compare les précisions relatives des deux procédures d'estimation au moyen de leur coefficient de variation, le gain moyen de précision est très net. Les coefficients de variation passent en moyenne de 61 p.100 à 14 p.100 à Rosso et de 45 p.100 à 12 p.100 à Lekseïba dès lors qu'il est possible de corriger les estimations par la régression.

Si l'on parle en termes d'efficacité, il aurait fallu en moyenne 28 fois plus de segments à Rosso et 11 fois plus de segments à Lekseïba pour obtenir la même précision si l'on n'avait pas eu de données satellitaires numériques satisfaisantes.

Il faut constater que les 20 segments dont nous disposons par zone d'étude sont trop peu nombreux en cas d'absence des données satellitaires numériques. A quelques exceptions près, la précision est bien trop faible pour être satisfaisante. De plus, la faiblesse de l'échantillon entraîne d'autres problèmes dont il faut être conscient.

La télédétection ne peut révéler d'une part que ce qui est perceptible au travers de ses radiométries, et d'autre part que ce que l'enquête de terrain a permis d'identifier. Un petit échantillon a toutes les chances de ne révéler que les thèmes majoritaires dans le paysage à savoir, ceux dont la superficie est caractéristique du paysage. La petite taille de certains thèmes peut être compensée par une radiométrie très caractéristique qui fait prendre conscience au chercheur d'un "oubli" du plan de sondage. Cette remarque vaut pour des thèmes comme les cultures irriguées dont le comportement spectral est très tranché et facilite leur localisation, elle vaut beaucoup moins pour les thèmes qui, bien que singuliers, ne sont qu'une variation de ce qui les entoure. Nous en voulons pour preuve une vétiveraie "relique" (De Wispelaere 1989 p.16) non prise en compte par le plan de sondage et découverte par hasard sur le terrain au nord de la zone pilote de Rosso.

Oubli rédhibitoire pour les uns, rançon de la représentativité à moindre coût pour les autres, c'est l'objectif poursuivi qui doit trancher. Dans la mesure où, malgré un temps d'enquête très court, nous voulions être représentatif des paysages les plus caractéristiques de nos zones d'étude, à défaut d'être les plus originaux, cette méthode d'inventaire s'est révélée très satisfaisante.

Nous conseillons toutefois en mode opérationnel de ne pas hésiter à échantillonner 30 à 60 segments par zone pilote. Cela offre au moins la possibilité de se référer à la loi des grands nombres dans les calculs (!) et si le nombre des thèmes et leur répartition le permettent, d'obtenir des estimations relativement précises sur les thèmes majoritaires en l'absence de prises de vue satellitaires. Cela permet également un choix plus vaste pour l'initialisation et le contrôle de la cartographie.

Donc à coût d'enquête analogue, l'utilisation des classifications comme correcteur de l'échantillonnage de terrain a permis d'accroître la précision des résultats d'inventaire d'un facteur 28 à Rosso et 11 à Lekseiba.

Références bibliographiques

- BOUDET (G.) - 1981 - *Système de production d'élevage au Sénégal. Etude du couvert herbacé (2e campagne)*. Maisons-Alfort, IEMVT, 22 p. + annexes.
- BOUDET (G.), CARRIERE (M.), CHRISTY (Ph.), WEDDOUD (ould CHEIKH A.) & al. - 1987 - Pâturage et élevage au sud de la Mauritanie (Kaédi) : Étude intégrée sur les pâturages, leur conservation et leur restauration; le cheptel et les éleveurs. Maisons-Alfort, IEMVT, 282 p.
- BRYANT (C.R.), RUSSWURM (L.H.) - 1983 - *Les méthodes d'échantillonnage aréolaire selon les besoins et objectifs de la surveillance de l'utilisation des terres*. Environnement Canada, Direction Général des terres, 76 p. (doc. de travail n°24).
- DESABIE (J.) - 1966 - *Théorie et pratique des sondages*. Paris, Dunod, 480 p.

- GODARD (V.) - 1992 - Évaluation des surfaces naturelles par télédétection et enquête de terrain en Mauritanie. *STATECO*, n° 71, septembre 1992 : 27-58.
- DE WISPELAERE (G.), GODARD (V.), CARRIERE (M.) - 1989 - *Expérimentation méthodologique pour le suivi de l'environnement par télédétection dans le Trarza, le Gorgol et l'Assaba (Mauritanie)*. Maisons-Alfort, IEMVT, 181 p. + 4 cartes
- FOURNIER (Ph.) - 1986 - *Enquête sur l'utilisation du territoire effectué en 1985 par la méthode des segments*. Paris, SCEES, Série S, 63 p. (Méthodes et Applications Scientifiques n°13).
- FOURNIER (Ph.), GILG (J.P.) - 1985 - *Télédétection et observation terrain*. Toulouse, cours du GDTA, 22 p.
- GODARD (V.) - 1990 - Télédétection et "enquête terrain" dans le sud-ouest mauritanien : optimisation des enquêtes de terrain pour la cartographie de l'occupation du sol et le diagnostic de l'état de l'environnement. *Actes de la table ronde internationale : Apport de la télédétection spatiale à l'étude des paysages et des systèmes agraires tropicaux*. CEGET-CNRS, Talence 25-27 octobre 1989. Bordeaux-Talence, CNRS : 87-105.
- GODARD (V.) - 1991 - *Utilisation conjointe de la télédétection et de l'enquête de terrain lors des inventaires d'occupation du sol. Recherche appliquée au Sahel sud-mauritanien*. Maisons-Alfort, IEMVT. Thèse de doctorat de l'EHESS, 433 p.
- GROSBRAS (J.M.) - 1987 - *Méthodes statistiques des sondages*. Paris, Economica, 331 p. (Col. Économie et statistiques avancées).
- HOUSEMAN (E.) - 1979 - *L'échantillonnage d'aires dans l'agriculture*. Washington DC, Statistical Reporting Service, US Dept. of Agric. 88 p. (SRS n°20).
- IAT - 1989 - *Inventaires régionaux*. Flash Télé-Agri-News. Inst. des Applic. de la Téléd., Service Spécialisé Agriculture (CCRCCE), Ispra (Italie), 2/89 : 2-15.
- MEYER-ROUX (J.), SIGMAN (R.), CRAIG (M.), OZGA (M.) - 1981 - Cartographie et statistique en télédétection. - *Cahier de Statistique Agricole*. n°6/6 nov.-déc. 1981 : 9-14.
- PASTORELLI (R.) - 1986 - Statistique d'occupation du sol à partir d'images satellites. *Cahiers de Statistique Agricole*. n°1/6 janvier-février 1986 : 7-12.
- PASTORELLI (R.), PORCHIER (J.C.) - 1989 - Premiers résultats sur l'utilisation des terres à l'échelle départementale à partir des données du satellite SPOT. *Cahiers de Statistique Agricole*. n°3/6 mai-juin 1989 : 51-56.
- TOUZELET (M.), MEYER-ROUX (J.), JOANNARD (M.) - 1983 - Enquête sur l'utilisation du sol par la méthode des segments. Ardèche 1982. *Cahiers de Statistique Agricole*. n°2/6, mars-avril 1983 : 9-16.

Annexe : Le développé méthodologique

Cette annexe reprend et développe les particularités de la méthode de tirage, particularités liées aux conditions d'application rencontrées dans le Sahel mauritanien.

A1 Justificatif méthodologique

Nous avons donc eu recours à un tirage systématique qui, moyennant quelques précautions, peut être considéré comme aléatoire (Desabie 1966 p.98). De plus, pour éviter les erreurs liées à la présence de phénomènes périodiques dans le paysage, nous avons eu recours à un tirage systématique non-aligné, encore appelé tirage systématique et aléatoire (Fournier 1986 p.15). Ce tirage assure une bonne répartition de l'échantillon sur l'ensemble de la zone d'étude, mais le repérage des unités d'enquête sur le terrain est rendu complexe par leur dissémination.

En dehors de la stratification implicite liée au tirage systématique, stratification fondée sur la répartition spatiale de ce type de tirage, il n'a été procédé à aucune autre stratification. Il est à noter qu'une stratification basée sur des critères thématiques aurait probablement plus été une source de complication que d'amélioration des estimations en raison de la taille réduite des zones pilotes, 20 km de côté.

Pour faciliter la localisation des pixels échantillon sur le terrain et réussir avec une faible incertitude l'affectation "occupation du sol-radiométrie", il est recommandé de tirer des surfaces et non des points (Fournier 1985 pp. 7-8). Nous avons de ce fait réalisé un tirage en grappe. Cela revient à tirer aléatoirement au premier degré les coordonnées d'un pixel puis au deuxième degré, à enquêter la totalité d'une surface développée à partir de ce pixel. Cette surface, l'unité d'enquête, est un carré de taille fixe et connue appelé segment. Enfin, lorsque la grappe est constituée par une zone géographique, le segment par exemple, on parle de sondages aréolaires ou bien encore de sondages spatiaux.

Dès lors que l'on ne tire plus uniquement un point, il convient de déterminer la taille de l'unité d'enquête. Celle-ci va se définir en fonction du taux de sondage et du temps dont on dispose pour réaliser l'inventaire. En l'absence de référence en milieu naturel, nous avons retenu un taux de sondage courant en milieu agricole, à savoir compris entre 1 et 2 p.100 (Fournier 1986 p.13, IAT 1989 p.5). Notons toutefois, qu'en raison de la faible taille de nos zones tests, 20 km de côté, le nombre des unités d'enquête est assez faible, 20 segments par zone pilote. En ce qui concerne le temps d'enquête, il a été déterminé empiriquement sur les premiers essais (Godard 1990 p.94), puis affiné en fonction du type de prélèvements que l'opérateur doit effectuer.

Sur un segment de 36 ha (cf. *infra*), taille initialement retenue, soit 600 m de côté, il faut compter environ trois heures d'inventaire (Godard 1991 p.219). Ce temps d'enquête comprend une description de l'occupation du sol sur le segment à savoir, une analyse synthétique des états de surface (paramètres topographiques, géomorphologiques, édaphiques, etc.) et un descriptif physiognomique de l'ensemble des formations végétales rencontrées puis, une analyse du cortège floristique et une quantification de la phytomasse herbacée aérienne sur deux parcelles homogènes

situées à l'intérieur de chacun des segments. Pour une enquête limitée à l'inventaire d'occupation du sol, le temps d'investigation est sensiblement plus faible.

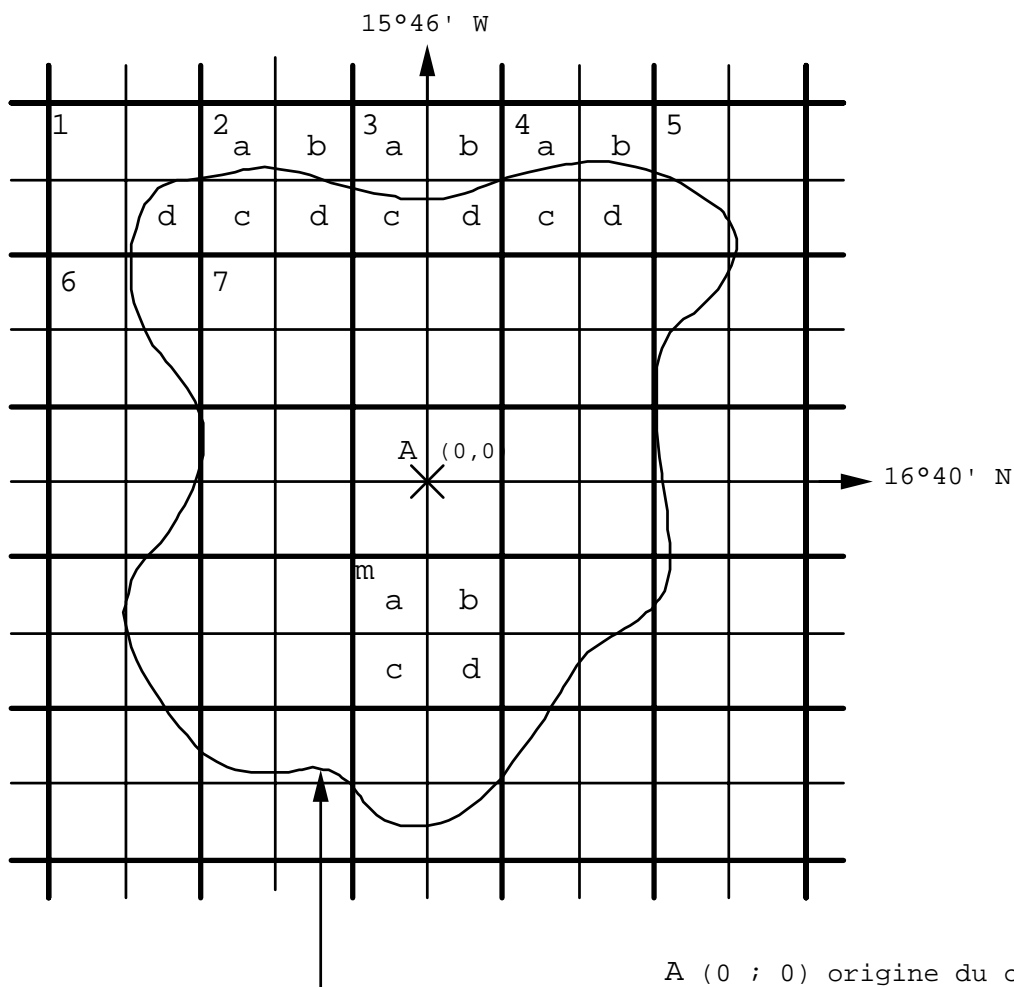
En effet, dans la mesure où toutes les parcelles ne portent pas de végétation, elles n'imposent pas systématiquement un temps d'inventaire de trois heures par segment. De même, en zone agricole le temps d'enquête sera inférieure car l'identification d'une culture est plus rapide que celle d'un groupement végétal. De ce fait, il est possible d'enquêter deux à trois segments par jour. A raison d'une dizaine de jours par secteur d'étude, et du fait de la mise au point de la méthode, il n'a été enquêté que 20 segments sur chacune des zones pilotes. Vingt segments de 36 ha, pour une zone pilote d'un peu plus de 40 000 ha, représentent un taux de sondage de 1,7 p.100.

Il aurait été possible d'enquêter un nombre un peu plus élevé de segments, d'une taille plus restreinte pour conserver le même taux de sondage, mais pour des raisons de calendrier d'enquête, nous avons dû démarrer les inventaires avant d'avoir pu déterminer la taille optimale du segment en fonction de la physionomie du paysage. Cette taille a été déterminée par la suite (cf. *infra*).

Le choix d'une unité d'enquête de 36 ha est directement inspiré des recherches menées en France par le Service Centrale de Enquêtes et Etudes Statistiques (SCEES) pour sélectionner, par paysage agricole, une taille optimale d'unité d'enquête. Celle-ci est fonction, de la taille moyenne du parcellaire et de la taille moyenne des exploitations agricoles (Fournier 1986 p.14) pour couvrir tous les types de cultures présents. De plus cette taille de segments doit tenir compte de la résolution du satellite utilisé, un multiple de 20 m dans le cas de SPOT et de 30 m dans le cas de TM, pour ne citer qu'eux. Il faut aussi tenir compte des possibilités de repérage et de localisation du segment sur le terrain.

figure 3 - Implantation des segments

La grille de sondage



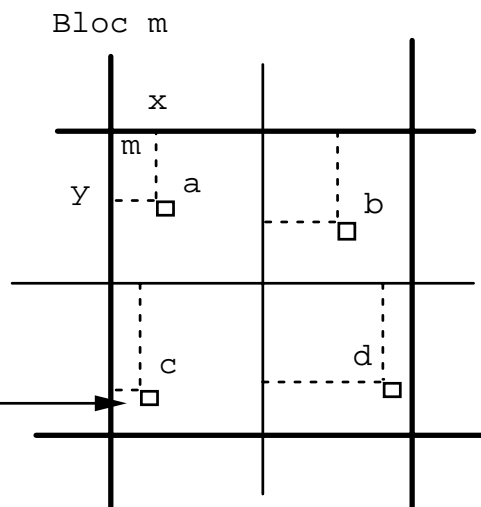
Limite de la zone d'étude

A (0 ; 0) origine du carroyag
ROSSO pour le tirage des segm

Implantation des segments

Un bloc est constitué de 4
carreaux élémentaires de 6
km de côté. La répartition
des segments est identique
d'un bloc à l'autre

Segment d'enquête mc



La plupart des expériences qui ont eu lieu en Europe, d'ailleurs exclusivement en milieu agricole, ont pris comme référence des segments carrés dont le côté est à quelques exceptions près un multiple de 100 m (système métrique oblige). C'est-à-dire par exemple, 500 m (Touzelet 1983 p.11) soit un segment de 25 ha, 600 m (Pastorelli 1986 p.8) soit un segment de 36 ha et 700 m (IAT 1989 p.3) soit un segment de 49 ha. En l'absence d'expérience en milieu naturel, tropical de surcroît, nous avons retenu la taille intermédiaire de 36 ha. De plus, son côté de 600 m est un multiple des pixels SPOT (30 * 20 m) et TM (20 * 30 m). Comme cela a été présenté au chapitre des résultats, un nombre un peu plus élevé de segments d'une taille légèrement inférieure n'aurait pas sensiblement modifié la précision des résultats. Cependant, la sélection des parcelles d'initialisation et de contrôle des classifications en aurait été grandement facilitée.

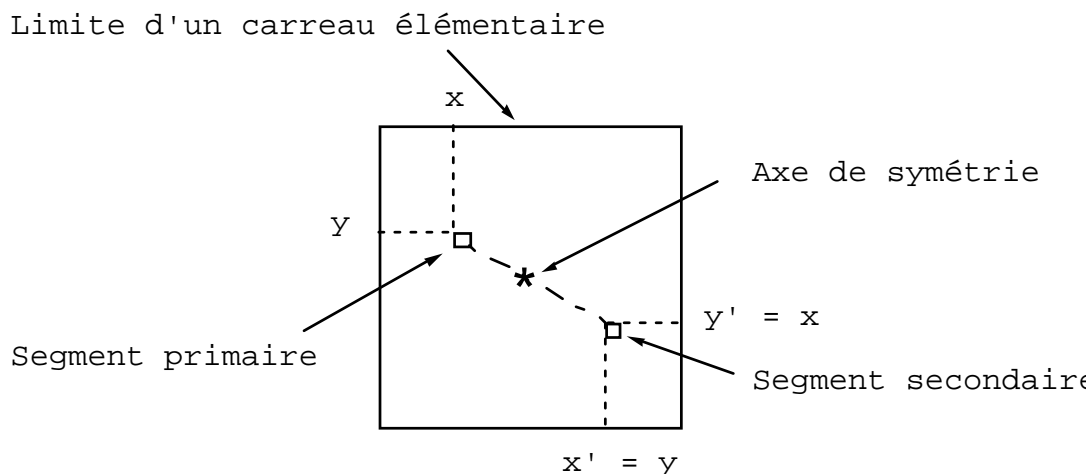
A.2 Tirage de l'échantillon

La mise en place du plan de sondage est assez simple une fois la taille du segment définie. Dans la mesure où le tirage est de type systématique non-aligné, il suffit de positionner aléatoirement une grille sur la zone d'étude. Ensuite, si l'on choisit un pas de grille valant dix fois la longueur d'un côté du segment, soit 6 000 m (10 * 600 m), alors tirer un segment par maille revient à sonder avec un taux de sondage de 1 p.100, deux segments par maille, 2 p.100, etc.

Pratiquement et pour gagner du temps, toutes les mailles de la grille sont groupées par quatre, ce groupe constitue un bloc. Pour chacune des mailles de ce bloc, sont tirées indépendamment les coordonnées x et y du segment qui lui est affecté (cf. fig. 3). Puis, ces quatre paires de coordonnées sont répétées à l'ensemble des blocs. On parle alors de la répétition a, b, c ou d de chacune des mailles origines.

En raison de la relative exiguïté des zones d'étude, et pour obtenir un minimum de 20 segments pour chacune d'entre-elles, nous avons dû réaliser un tirage à 2 p.100, car la méthode de tirage fait tomber tout ou partie de certains segments périphériques en dehors de la zone d'étude, segments qui sont dès-lors éliminés. Cela explique que le taux de sondage final soit en général un peu plus faible, 1,7 p.100 dans notre cas.

figure 4 - Renforcement de l'échantillon



Pour réaliser un tirage à 2 p.100, il suffit que chaque maille ou carreau élémentaire se voit attribuer un segment supplémentaire, dit segment secondaire, symétrique du premier qui lui a été affecté. On inverse par exemple les coordonnées x et y du segment primaire pour éviter un retraitage fastidieux (cf. fig. 4). Ce segment n'est retenu que s'il tombe en entier dans la zone d'étude, tout comme le segment primaire et s'il ne le chevauche pas. Il est donc tout à fait possible d'adapter ce mode de tirage aux besoins de l'inventaire en ajoutant ou supprimant ce segment supplémentaire sur une ou plusieurs mailles tirées au hasard, voire de supprimer une ou plusieurs mailles de la répétition pour alléger le taux de sondage sur de grandes zones d'étude.

Enfin, si une périodicité dans l'organisation du paysage apparaissait, ce qui constitue un risque de biais pour un tirage systématique, elle n'affecterait probablement pas les quatre répétitions. Il serait donc possible de supprimer la répétition affectée par cette périodicité et de conserver les trois autres pour poursuivre l'inventaire.

A.3 Recherche d'une taille optimale de segments

En milieu agricole, les dimensions du segment sont fonction de la taille moyenne de l'exploitation et du parcellaire agricole. Il n'est donc pas directement possible de recourir à ces deux paramètres pour adapter ou contrôler la taille du segment sur les formations spontanées. En revanche, il est envisageable de chercher à l'aide des données satellitaires s'il n'existe pas dans le paysage une "entité thématique", comparable au parcellaire agricole. Si cette entité est perceptible sur les images satellitaires, en déterminant sa taille moyenne, il doit être possible de l'utiliser pour améliorer l'adéquation de la taille des segments à la complexité du paysage. Rappelons que leur taille n'est pour le moment fonction que du taux de sondage et du temps d'enquête mais pas de la spécificité du paysage. Une telle entité est appelée par analogie, parcellaire naturel.

Pour déterminer la taille de ce parcellaire naturel, nous avons réalisé des classifications automatiques sur lesquelles différents descripteurs issus des techniques de la morphologie mathématique et de la géostatistique ont été testés. Cela revient à effectuer une analyse de formes sur des "entités radiométriques homogènes" supposées être représentatives "d'entités thématiques homogènes" à une échelle donnée. On n'a pas cherché à caractériser dans cette approche les thèmes mis en évidence par ce type de préclassification, mais uniquement la taille, tous thèmes confondus, du parcellaire.

Pour être comparables avec le paysage qui va être enquêté, les données satellitaires utilisées pour ce type d'analyse doivent avoir été enregistrées à une même saison que celle de la phase de terrain. En dehors de profonds remaniements dans le paysage, elles peuvent certainement être antérieures de quelques années à la mission. Pour notre part, il s'agit de données SPOT de niveau 1B enregistrées le 6 novembre 1987 pour Rosso (KJ 023-317) et le 22 octobre 1987 pour Lekseïba (KJ 028-318). Dans la mesure où, pour des problèmes de calendrier, nous n'avions pas pu mener l'analyse de la taille du parcellaire à son terme avant l'inventaire, une seule couverture satellitaire a été acquise, contemporaine de l'enquête de terrain, pour estimer les superficies. Cette couverture a également servi pour déterminer, mais a posteriori, la taille des segments. C'est un cas de figure improbable en phase opérationnelle.

A la suite de la classification automatique réalisée sur chacune des zones pilotes, une analyse de surfaces de l'ensemble des taches constituant le résultat des préclassifications a été pratiquée. Il en ressort qu'une fois éliminées les entités de moins de deux hectares, qui en raison de la méthode de calcul sont souvent des artefacts de traitements (Godard 1991 p.249), un segment de 36 ha comporte environ 9 parcelles (cf. tab. 7).

tableau 7 - Nombre de parcelles "radiométriques" par segment

Surface du segment en ha	Côté du segment en mètres	Nombre de parcelles	
		Rosso	Lekseïba
4	200	1,0	1,1
9	300	2,3	2,4
16	400	4,1	4,2
25	500	6,4	6,6
36	600	9,2	9,5
49	700	12,6	12,6
64	800	16,4	16,8

Le calcul est effectué pour des parcelles de plus de 2 ha. La taille médiane du parcellaire est de 3,9 ha à Rosso et de 3,8 à Lekseïba.

Ce tableau fournit un ordre de grandeur quant au nombre de parcelles que nous rencontrerons sur le terrain. Rappelons que celles-ci sont fonction de la relation thème-radiométrie, relation qui n'est pas toujours univoque et que de plus la taille et donc le nombre de parcelles dépendent du nombre de classes de la classification automatique (Godard 1991 p.231). Ce tableau ne doit donc être utilisé qu'en complément d'une détermination de la taille du parcellaire sur des contraintes de temps d'enquête et de taux de sondage.

Cette approche de la taille du parcellaire confirme a posteriori le choix d'un segment de 36 ha. Cependant, si nous avions à refaire cet inventaire, en fonction des

contraintes de temps d'enquête et de taux de sondage la taille des segments serait probablement réduite de 36 à 25 ha. Le nombre de parcelles à inventorier sur l'unité d'enquête serait plus réduit mais les segments seraient plus nombreux à être inventoriés pour faciliter l'initialisation et le contrôle des cartographies. Le résultat statistique n'en aurait pas fondamentalement été modifié, mais le coût d'accès est bien sûr dans ce cas plus élevé.

A ce niveau de l'analyse tous les paramètres nécessaires pour réaliser le sondage sont réunis. Une série de tirages d'échantillon va être effectuée sur la classification automatique précédemment décrite pour sélectionner l'échantillon le plus performant, c'est-à-dire celui qui est le plus représentatif de toutes les entités identifiées sur la préclassification.

La méthode la plus simple pour réitérer le tirage est de calculer de nouvelles paires de coordonnées au point A, origine de la grille de sondage (cf. fig. 3). Cela évite de recalculer les coordonnées de chaque segment dans la mesure où la grille et les segments se déplacent de manière solidaire.

Ces tirages successifs sur la préclassification ont pour but de contrôler que tous les thèmes, même ceux dotés d'une très faible superficie, sont bien pris en compte par plusieurs segments du plan de sondage. Le premier à remplir ces conditions est retenu pour l'inventaire. Si aucun n'était satisfaisant, il faut accroître le nombre des segments, avec ou sans réduction de leur taille suivant le taux de sondage envisagé. Cette étape franchie, l'inventaire proprement dit peut se dérouler.