

APPORT DE LA TELEDETECTION DANS LE DOMAINE DES BANQUES DE DONNEES
D'OCCUPATION DU SOL - EXEMPLE DE LA MOSELLE FRANCAISE

SERGE PIEYNS ET MARC LOINTIER

Hydrologues - Service Hydrologique de l'ORSTOM

70-74, route d'Aulnay - 93140 BONDY

En 1978, par convention avec l'Agence Financière de Bassin RHIN-MEUSE (A.F.B.R.M.), le Service Hydrologique de l'Office de la Recherche Scientifique et Technique Outre-Mer (O.R.S.T.O.M.) mettait au point un modèle pluie-débit du bassin de la Moselle française, soit 11.600 km². Ce modèle, qui à l'heure actuelle a été calé sur à peu près la moitié du bassin, doit permettre à l'A.F.B.R.M. de reconstituer les débits moyens journaliers, en tout point du bassin, à partir des données pluviométriques. Il s'agit d'un modèle déterministe à réservoirs qui présente la particularité d'être à discrétisation spatiale, se référant en cela aux travaux de G. GIRARD de l'ORSTOM et de chercheurs de l'INRS Eau du Québec. Le bassin ayant été partagé en carreaux élémentaires de 25 km² de surface, la production, c'est-à-dire la répartition des précipitations entre les différents niveaux de réservoirs, s'effectue carreau par carreau, une fonction de transfert permettant ensuite de composer les coulements élémentaires en les transférant à l'exutoire du bassin versant concerné. La fonction de production appliquée à chaque carreau prend en compte, non seulement les données précipitations, mais également les données de température pour le calcul de l'ETP et de fonte de neige et, évidemment, les données physiographiques.

La banque de données physiographiques du bassin de la Moselle comporte huit fichiers rassemblant pour chaque carreau élémentaire un grand nombre d'informations sur, par exemple, l'altitude moyenne, maximale et minimale du carreau, les pentes et l'orientation des versants, la longueur et la largeur des cours d'eau principaux, la géologie, la pédologie, l'occupation du sol.

Dans un premier stade, le fichier d'occupation du sol a été constitué de manière tout à fait classique. On a déterminé, sur les cartes à l'échelle du 25000ème le pourcentage de la surface de chaque carreau élémentaire occupée

par les thèmes suivants : eau libre, marécage, feuillus, résineux, prairies, cultures, zones urbanisées, qui sont les plus largement représentés sur le bassin et jouent un rôle important dans les phénomènes de ruissellement, d'infiltration et d'évaporation. Cette opération, réalisée avec l'aide du Centre d'Etudes Géographiques de l'Université de Metz (C.E.G.U.M.), s'est avérée longue et fastidieuse, présentant en outre un certain nombre de difficultés inhérentes aux méthodes utilisées :

- Risques d'erreurs de planimétrage et de report des résultats sur les bordereaux de saisie mécanographique (90 cartes à planimétrer).
- Hétérogénéité de l'information, cartes dont la mise à jour s'étalait entre 1958 et 1978.
- Difficulté de ventiler au niveau du carreau les informations de l'Office National des Forêts (O.N.F.) et du Ministère de l'Agriculture établies sur une base communale.

Toutes ces raisons ont incité le Service Hydrologique de l'ORSTOM à répondre favorablement à la proposition de l'Opération Pilote Interministérielle de Télédétection (O.P.I.T.) qui, en 1979, nous demandait de mettre au point une méthodologie d'utilisation des données satellitaires pour la constitution d'un nouveau fichier d'occupation du sol.

1 - LA METHODOLOGIE

1.1 - Les données disponibles

Ce sont, d'une part, les informations fournies par des satellites, d'autre part, celles recueillies de manière plus classique, sur le terrain.

1.1.1 - Les données satellitaires

Nous avons utilisé les données transmises par les satellites LANDSAT 1 et 2, qui sont des satellites d'observation de la terre, à défilement, avec une orbite presque circulaire à 920 km d'altitude.

Tous les 18 jours, le satellite se retrouve pratiquement à la même heure sur la même orbite. Sur chaque orbite, on peut définir un certain nombre de scènes, chaque scène représentant une surface carrée de 185 km de côté et contenant 7,7 millions de points élémentaires ou PIXEL, élément qui mesure, au sol, 56 mètres sur 79 mètres. Le satellite numérise les données radiométriques restituées sur 128 niveaux pour chacun des 4 canaux du radiomètre embarqué et pour chaque pixel, chaque canal correspondant à une bande spectrale différente, les canaux 4 et 5 dans le visible, les canaux 6 et 7 dans le proche infrarouge. Ainsi, à chaque pixel sont associés 4 nombres, de 0 à 128, qui correspondent à 4 valeurs radiométriques, soit pour une scène, 31 millions de valeurs radiométriques.

Pour une vue donnée, qui est une scène à une date donnée, on disposera donc d'une bande magnétique comportant ces 31 millions de valeurs radiométriques et de 4 négatifs au millionième, 1 par canal, élaborés par transcription sur film, en 16 niveaux de gris des informations de la bande. Dans le cas de la Moselle, le bassin était entièrement couvert par les scènes 211-026 et 212-026 et nous avons pu disposer de 5 vues, entre le 28 août 1975 et le 13 août 1976.

1.1.2 - Les "données terrain"

L'utilisation des données satellitaires nécessite naturellement ce que l'on appelle généralement des "vérités terrain", pour choisir les zones test sur lesquelles on initialisera les traitements et pour faire une évaluation objective et précise des résultats de ces traitements.

Ces "données terrain" sont, d'une part, celles dont on disposait pour la constitution du fichier classique dans le premier stade de l'étude, c'est-à-dire les cartes à l'échelle du 25.000ème, les enquêtes du Ministère de l'Agriculture TER-UTI, les relevés des parcelles forestières de l'ONF,

d'autre part, les enquêtes et relevés que nous avons effectués sur les zones test, ainsi que les photos aériennes, mission 1972 et le fichier descriptif des étangs du plateau lorrain mis à notre disposition par l'A.F.B.R.M.

1.2 - Le traitement de l'information satellitaire

Les deux types de produit, mis à la disposition de l'utilisateur par le Service gestionnaire du satellite, bande magnétique et restitution au million, induisent deux types de traitement, qui sont pour nous complémentaires, le traitement photochimique et le traitement numérique.

1.2.1 - Le traitement photochimique

Ce type de traitement doit être considéré comme un moyen de classer les valeurs radiométriques et comme une étape avant le traitement numérique. Sans entrer dans le détail de toutes les manipulations que nous avons effectuées au cours de notre étude, nous citerons et donnerons quelques informations sur les étapes essentielles du traitement.

. Observation des vues au million de chaque canal

Cette première étape est tout à fait essentielle. Elle permet de juger de la qualité des vues disponibles, de faire un choix parmi les dates selon l'objectif fixé et notamment en utilisant le canal 4, de délimiter les zones de brume, qui n'apparaissent pas dans les canaux 6 et 7, proche infrarouge. Ce repérage des zones de brume a pour but d'éviter une initialisation du traitement numérique sur ces zones, pour lesquelles on constate en effet un décalage des radiances vers les fortes valeurs. D'autre part, la superposition d'une grille de 3240 colonnes et 2280 lignes, au pas de 100, permet de délimiter les zones d'étude dans les coordonnées LANDSAT.

. Agrandissement

Cet agrandissement des vues au million, pour les zones test, peut s'accompagner ou non d'une modification des paramètres photochimiques pour obtenir un étalement des valeurs de gris correspondant au thème sur lequel on travaille. Les échelles classiques d'agrandissement sont comprises entre le 500.000ème et le 100.000ème. L'observation de ces agrandissements au rétroprojecteur s'est avérée fort utile.

. Composition colorée

Cette technique fait appel au procédé de la thèse soustractive utilisée en photographie couleur. Pour chaque canal, partant d'un positif transparent, on utilise un film inversible diazoïque comportant l'un des trois pigments complémentaires. Trois pigments seulement pour 4 canaux LANDSAT, ce qui conduit à éliminer le canal 6 très proche du 7. En superposant les trois positifs, canal 4 (jaune), canal 5 (magenta) et canal 7 (cyan), on obtient un document coloré qui peut être examiné à la table lumineuse, ou mieux, au rétroprojecteur, ce qui permet des agrandissements importants autorisant un zonage des couleurs.

. Equidensité

Travaillant sur des agrandissements à l'échelle du 150.000ème des zones test, on a pu effectuer directement, avec un densitomètre digital, la comparaison de la radiométrie des différents thèmes recherchés, préalablement soigneusement repérés et délimités. La mesure est faite à l'intérieur d'un cercle de 1 mm de diamètre. En effectuant pour chaque thème une dizaine de lectures, on peut obtenir, pour chaque canal, un graphique du type de celui représenté en figure 2, pour le canal 7. Sur ce graphique on voit, par exemple, que le "thème eau", symbole E, se place toujours

vers les fortes densités, donc dans les faibles radiances et ce, quelque soit la date de la vue. Si l'on observe ce qui se passe pour les autres thèmes, on voit que les résineux, symbole R, ont eux une densité qui varie notablement selon la date de passage du satellite. L'examen du même graphique pour le canal 4 montrerait, au contraire, une très bonne stabilité de ce dernier thème vers les fortes densités. Ces observations montrent bien qu'en combinant les différents canaux, on va pouvoir assez facilement définir une position relative des radiométries des différents thèmes qui nous intéressent, les uns par rapport aux autres. Bien entendu, tous les documents utilisés pour cette étude densitométrique ont été obtenus sur le même type de film et dans des conditions d'exposition et de développement identiques.

A partir du moment où l'on possède de nombreuses valeurs de densité, on peut traiter les films par équidensité. Utilisant un film-trait, dont la réponse est parfaitement binaire, nous pouvons, par exemple, fixer le seuil de disparition du pigment de notre film à une valeur telle que seuls apparaîtront sur le film les thèmes dont les densités seront supérieures au seuil que nous aurons fixé. Si l'on combine plusieurs plages de densité et 4 pigments colorés : noir, jaune, cyan et magenta, on obtient une équidensité colorée, pratiquement équivalente à un traitement numérique sur écran couleur. Un autre avantage de l'étude densitométrique est de guider l'étude numérique et de faciliter le choix de l'ordre d'effacement des thèmes dont nous allons parler plus loin.

1.2.2 - Le traitement numérique

Si, comme nous l'avons déjà signalé, le traitement photochimique est une étape indispensable dans l'utilisation des données satellitaires, il n'en demeure pas moins qu'il dégrade l'information. Seul le traitement numérique permet de prendre en compte les 31 millions de valeurs radiométriques disponibles dans une vue donnée ! Pour traiter, de manière

performante au plan informatique cette masse de données, le Bureau de Télédétection de l'ORSTOM a mis au point une procédure, dite procédure LOTERIE, que nous avons utilisée pour cette étude de la Moselle. Cette méthode d'analyses multivariables a été développée pour s'adapter à la classification d'images numérisées, composées d'un très grand nombre de points ; elle repose sur deux notions principales que nous expliciterons après avoir donné quelques informations sur "l'atelier numérique" du Bureau de Télédétection de l'ORSTOM où nous avons réalisé nos traitements.

1.2.2.1 - "L'atelier numérique"

Cet "atelier numérique" se compose d'un ordinateur CII Honeywell Bull 643 de 256 K de mémoire, relié à l'époque de nos travaux à 13 périphériques et à une console de visualisation autonome PERICOLOR (figure 3). Grâce à cette console, une partie d'une vue LANDSAT, préalablement chargée sur disque, peut être visualisée, par transfert dans la mémoire du PERICOLOR de l'information qui représente 65.536 pixels, soit 0,8 % de la vue. En ne visualisant qu'un point sur 2, sur 3, etc..., on peut évidemment agrandir la zone d'étude. Le transfert réalisé, le PERICOLOR devient autonome et l'on peut réaliser divers traitements en utilisant le clavier de la console, par exemple affecter à chaque valeur radiométrique une couleur déterminée, faire varier la taille du pixel, dégrader l'image et bien d'autres opérations. Une console graphique autorise également la construction d'histogrammes qui, pour une vue donnée, permettent de visualiser dans quelle bande radiométrique se trouve l'information.

1.2.2.2 - La procédure LOTERIE

Elle s'appuie sur deux notions fondamentales, la notion de serpent et la notion de lot.

Notion de serpent (figures 4a à 4c)

Soit un pixel caractérisé par quatre valeurs radiométriques a, b, c, d . On appellera POLYNOMBRE du niveau 4, la suite des 4 valeurs entières notée : $B_1 = (a, b, c, d)$. Un tel polynombre est représenté à la figure 4a. Soit un second pixel représenté par le polynombre $B_2 = (e, f, g, h)$.

On appellera SERPENT, l'ensemble des polynombres dont les valeurs seront comprises entre les polynombres B_1 et B_2 (figure 4b). Un tel serpent est noté $S = \begin{pmatrix} e & f & g & h \\ a & b & c & d \end{pmatrix}$.

Au sein d'une même vue, on peut définir plusieurs serpents et l'on dira que 2 serpents sont disjoints s'ils n'ont aucun polynombre commun. Pour cela, il faut et il suffit que, sur un niveau, les 2 serpents n'aient aucun nombre en commun (figure 4c).

C'est cette propriété qui sera utilisée pour faire un classement des pixels de la vue.

Notion de lot

On appelle lot "L" l'ensemble des points d'une vue dont les polynombres appartiennent à un serpent "S". Si l'opérateur ayant préalablement défini le serpent d'un lot donné, par exemple le "lot eau", demande à la machine d'extraire les

points de la vue appartenant à ce serpent de définition " S_D ", il obtiendra un autre serpent, appelé serpent REEL " S_R " et il pourra visualiser sur le PERICOLOR l'image du lot et du serpent correspondant à l'eau.

La procédure LOTERIE

Si l'opérateur définit 3 bornes dans une vue à 2 niveaux, il obtiendra 9 lots et 9 serpents réels que l'on peut représenter sur un plan (figure 4d). La procédure LOTERIE permet donc de construire un ensemble de lots, définis par une ou plusieurs bornes, sur chaque canal LANDSAT. A l'inverse de nombreux algorithmes de traitement d'images, l'initialisation peut s'effectuer sur de vastes zones, limitant ainsi l'influence des phénomènes radiométriques particuliers et éliminant les hypothèses statistiques de répartition de l'information. Cette initialisation peut s'effectuer sur l'écran du PERICOLOR, l'opérateur remettant en cause les valeurs des bornes sur chaque canal lorsque l'image obtenue n'est pas conforme aux informations terrain. Lorsque le traitement apparaît correct, l'opérateur peut alors changer de zone dans la vue et vérifier la valeur de son initialisation et, de proche en proche, définir ainsi le "serpent du thème" recherché. Compte tenu du fait qu'en multipliant les bornes et en utilisant toute l'information radiométrique disponible, on atteint très vite un nombre de lots très importants, d'autres outils complètent la procédure LOTERIE permettant notamment de tester l'homogénéité et la stabilité d'un lot. De même, lorsque l'on teste le serpent d'un thème sur une autre zone de la vue que celle où il a été défini, on constate parfois que le résultat n'est pas toujours

satisfaisant. En effet, les phénomènes radiométriques qui prévalent pour la définition des bornes du serpent sur une zone donnée peuvent ne pas être exactement identiques sur l'ensemble de la vue. Le souci de trouver une démarche simple et rapide conduit donc à se poser la question suivante : sur les 8 valeurs du serpent, quelles sont les bornes réellement utiles dans la définition du lot du thème recherché ? Des essais au PERICOLOR permettent de répondre à cette question et de définir un nouvel opérateur, la PARENTE du thème, notée P, qui est en quelque sorte le serpent du thème simplifié, débarrassé des contraintes inutiles.

C'est ainsi par exemple que, sur la zone du Bois de RECHICOURT, on s'aperçoit qu'en ne gardant que 2 bornes, l'une dans le canal 5, l'autre dans le canal 7, on obtient la même image du lot forêt qu'en utilisant le serpent réel de ce lot.

$$S_R = \begin{pmatrix} 43 & 34 & 110 & 107 \\ 32 & 24 & 51 & 42 \end{pmatrix}$$

$$P = \begin{pmatrix} - & 34 & - & - \\ - & - & - & 42 \end{pmatrix}$$

Ce nouvel opérateur appliqué à l'ensemble de la vue permet d'obtenir de bien meilleurs résultats que l'utilisation du serpent réel, trop contraignant.

Nous avons donc utilisé la procédure LOTERIE et ses outils complémentaires, thème après thème, l'effacement sélectif d'un thème après son individualisation, simplifiant la recherche des bornes de thème suivant pour peu que l'on utilise les résultats de l'étude densitométrique qui nous a permis de classer les thèmes selon les valeurs relatives de leurs radiométries. Le graphique 5 résume la méthodologie que nous avons mise au point.

2 - LES RESULTATS

2.1 - Le produit obtenu

L'ensemble des traitements photochimiques et numériques que nous avons réalisés nous a finalement permis d'identifier, à l'intérieur des limites du bassin de la Moselle, les pixels appartenant à l'un des grands thèmes suivants : eau libre, feuillus, résineux, cultures, prairies, zones urbanisées. Un faible pourcentage de points élémentaires se retrouve "hors thème", soit parce que ces pixels n'appartiennent pas effectivement aux thèmes recherchés, soit parce que leur radiométrie, perturbée par des phénomènes particuliers, les place en dehors des limites des opérateurs utilisés.

Le produit final du traitement se présente sous deux formes :

- Une cartographie des thèmes. Nous avons réalisé une carte de l'ensemble du bassin à l'échelle du 250.000ème, une carte à l'échelle du 100.000ème du plateau lorrain et également une carte à l'échelle du 50.000ème de la région de SARREBOURG. Il s'agit d'une cartographie automatique réalisée à la table traçante à partir des résultats de la procédure LOTERIE, appliquée thème par thème.
- Une banque de données d'occupation du sol, obtenue en traitant les résultats des LOTERIES successives, à l'aide de deux programmes. Le premier qui nous a permis de positionner les carreaux élémentaires de notre modèle à discrétisation spatiale dans les coordonnées LANDSAT, le second qui, carreau par carreau, fait la statistique des pixels appartenant à un thème donné et nous fournit le pourcentage de la surface du carreau, occupé par les différents thèmes recherchés.

2.2 - Evaluation des résultats

Sans entrer dans le détail des enseignements fournis par cette étude, nous pouvons insister plus particulièrement sur le problème du

seuil de détection des principaux thèmes et du pourcentage d'erreur reconnu dans les produits finaux, les cartes et la banque d'occupation du sol. Le tableau 1 résume ces résultats à propos desquels il est possible de faire un certain nombre de commentaires.

On notera tout d'abord que les seuils de détection sont évidemment plus élevés dans le cas du traitement photochimique que dans celui du traitement numérique. La dégradation de l'information et la limite physiologique introduite par le pouvoir séparateur de l'oeil en sont responsables, cette première remarque justifiant si besoin en était l'intérêt du traitement numérique. On remarque en outre que les seuils de détection pour les zones urbanisées, les prairies et les cultures sont nettement plus élevés que pour l'eau libre, les feuillus et les résineux. Pour ce qui est de ces deux derniers thèmes, il faut voir là, principalement, la difficulté d'obtenir une "vérité-terrain" très précise à partir du moment où il n'est pas possible de la réaliser à l'époque du passage du satellite. Pour les zones urbanisées, il y a un réel problème, surtout pour les villages qui présentent très souvent des radiométries équivalentes à celles des cultures qui les entourent.

Le tableau 1 permet également de constater que si, pour les trois premiers thèmes les erreurs constatées sont assez faibles, de l'ordre de 2 à 3 %, ces erreurs atteignent 10 % pour les prairies et les cultures et 20 % en moyenne pour les zones urbanisées. On doit cependant tempérer ces résultats, tout au moins en ce qui concerne les prairies et les cultures, du fait même que la manière dont on a ventilé les résultats communaux à l'intérieur des carreaux élémentaires de notre modèle introduit très certainement déjà des erreurs. On constate en effet que le total des surfaces en prairies et cultures, calculé à partir des données de télédétection est équivalent à 2 ou 3 % près à celui qui ressort de ces relevés communaux, les différences constatées entre les deux fichiers, appelons les "classique" et "télédétection",

portent essentiellement sur la répartition des surfaces entre ces deux thèmes. En fait, on est pratiquement certain que l'erreur réelle commise pour ces deux thèmes est du même ordre que pour l'eau libre, les feuillus et les résineux, soit environ 2 à 3 %.

Pour les zones urbanisées, l'erreur commise varie beaucoup entre 10 et 30 % selon les dates utilisées. Il est possible, mais non certain, que l'on puisse grandement améliorer la qualité des résultats en jouant justement sur le diachronisme, en superposant plusieurs vues.

En conclusion, au terme de cette étude, nous avons tout lieu d'être satisfaits des résultats et nous pensons que l'utilisation des données satellitaires pour la constitution de banques de données d'occupation du sol est une méthode d'avenir. D'une part, la précision des résultats est tout à fait acceptable et en tout cas compatible avec l'usage que nous comptons en faire, d'autre part, cette technique lève certaines des difficultés que nous avons signalées dans notre introduction, hétérogénéité des documents, lourdeur de l'opération avec la multiplication d'opérations répétitives, risques d'erreur dans les planimétrages et de transfert de l'information sur support informatique.

L'utilisation des données satellitaires permet d'envisager une remise à jour périodique de la banque sans frais excessifs. Bien sûr, les opérateurs définis pour une vue donnée ne sont généralement pas transposables directement à une autre vue, mais la dérive radiométrique étant relativement faible, on n'aura dans la plupart des cas que de faibles ajustements à apporter aux bornes des différents serpens. Il reste que la résolution actuelle du LANDSAT n'est pas réellement adaptée au parcellaire français et que de nombreuses vues sont inutilisables, du fait des couvertures nuageuses fréquentes sous nos climats. En ce qui concerne la résolution au sol, le futur satellite SPOT devrait régler le problème qui reste malheureusement entier pour ce qui tient aux nuages.

TABEAU 1

THEMES	Traitement Photochimique	Traitement Numérique		
	Seuil de Détection (pixels)	Production de CARTES		FICHER erreur par rapport au fichier physiographique
		Seuil de détection	Erreur %	
Eau libre	5	3	5,2	3 %
Résineux	5	4	4	3,5 %
Feuillus	10	4	4,5	2 %
Zones urbanisées	30	30	11 à 30 suivant date	15 %
Prairies	16	8	5	10 %
Cultures	16	10	11	10 %

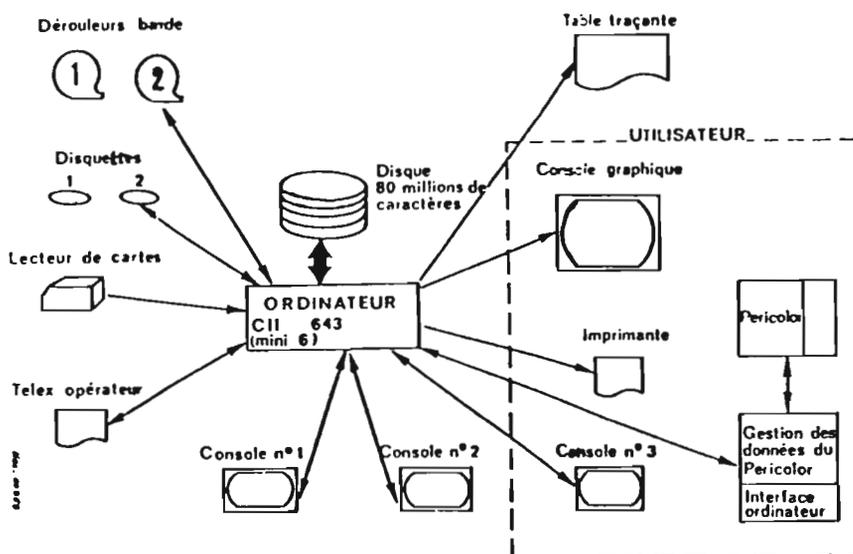


Figure 3 - L'Atelier Numérique du Bureau de Télécommunication de l'ORSTOM.

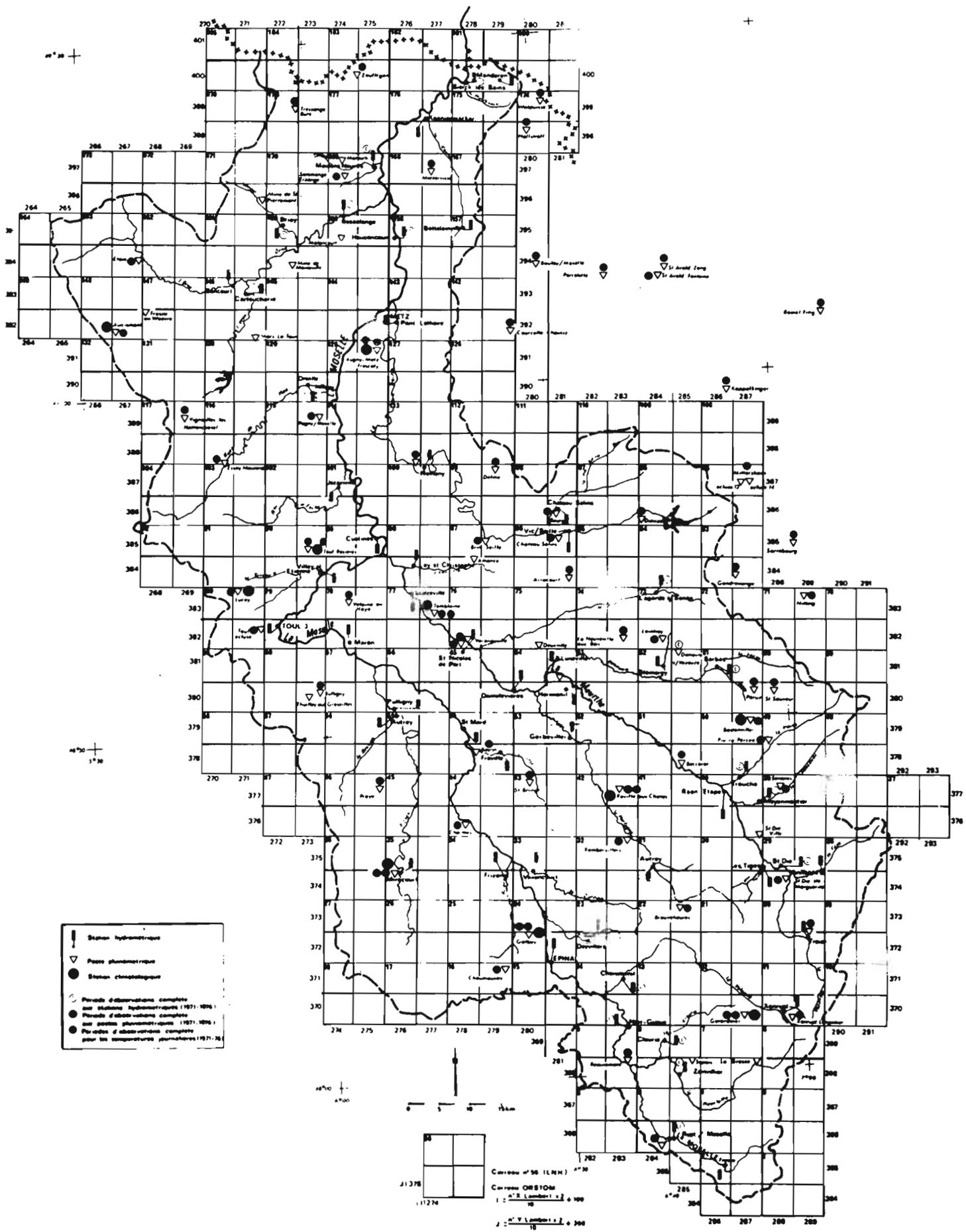


Fig. 1 - Modèle général de la Moselle. Équipement hydropluviométrique et climatologique.

CANAL 7

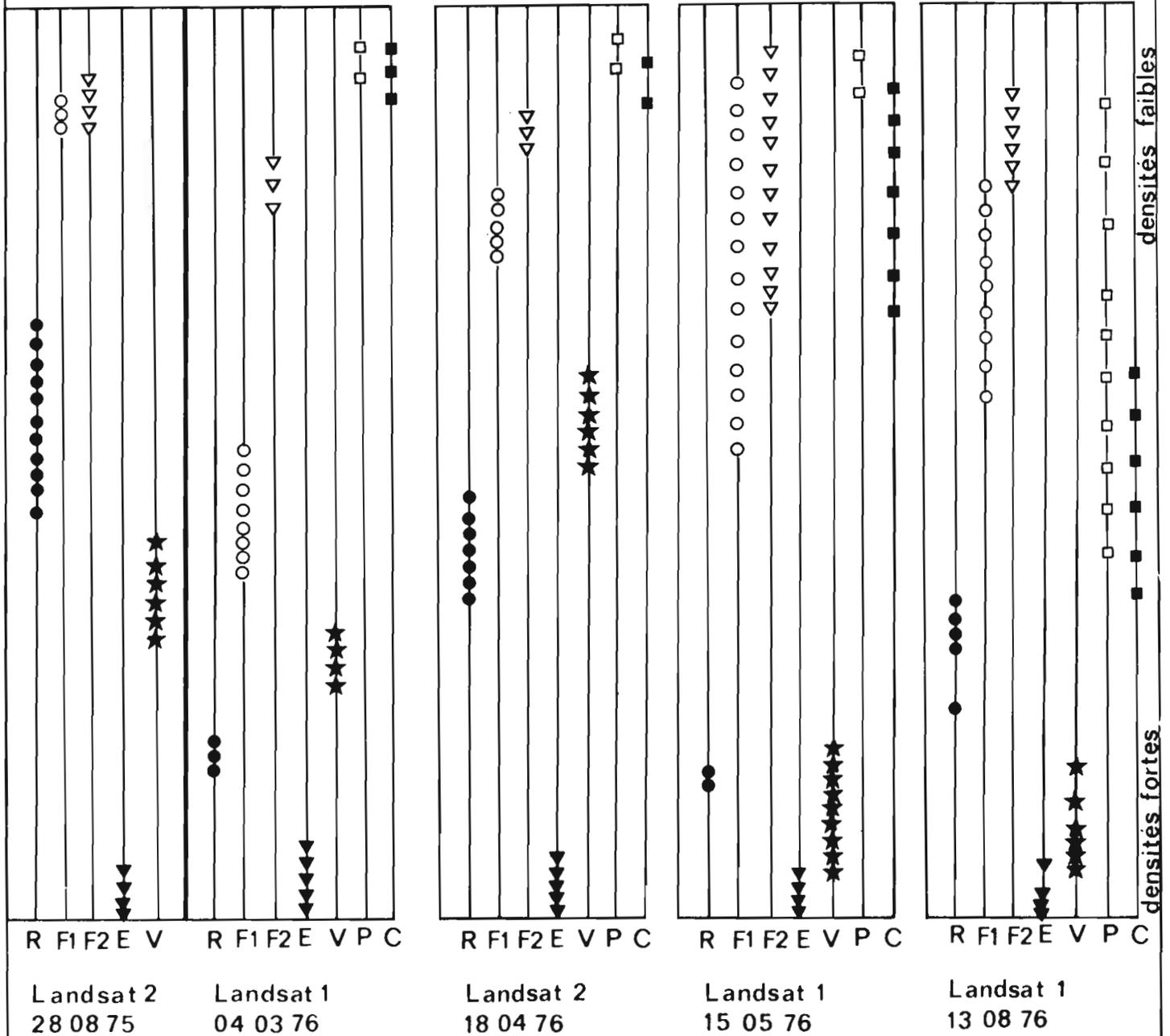


Figure 2 - Etude comparative des résultats densitométriques

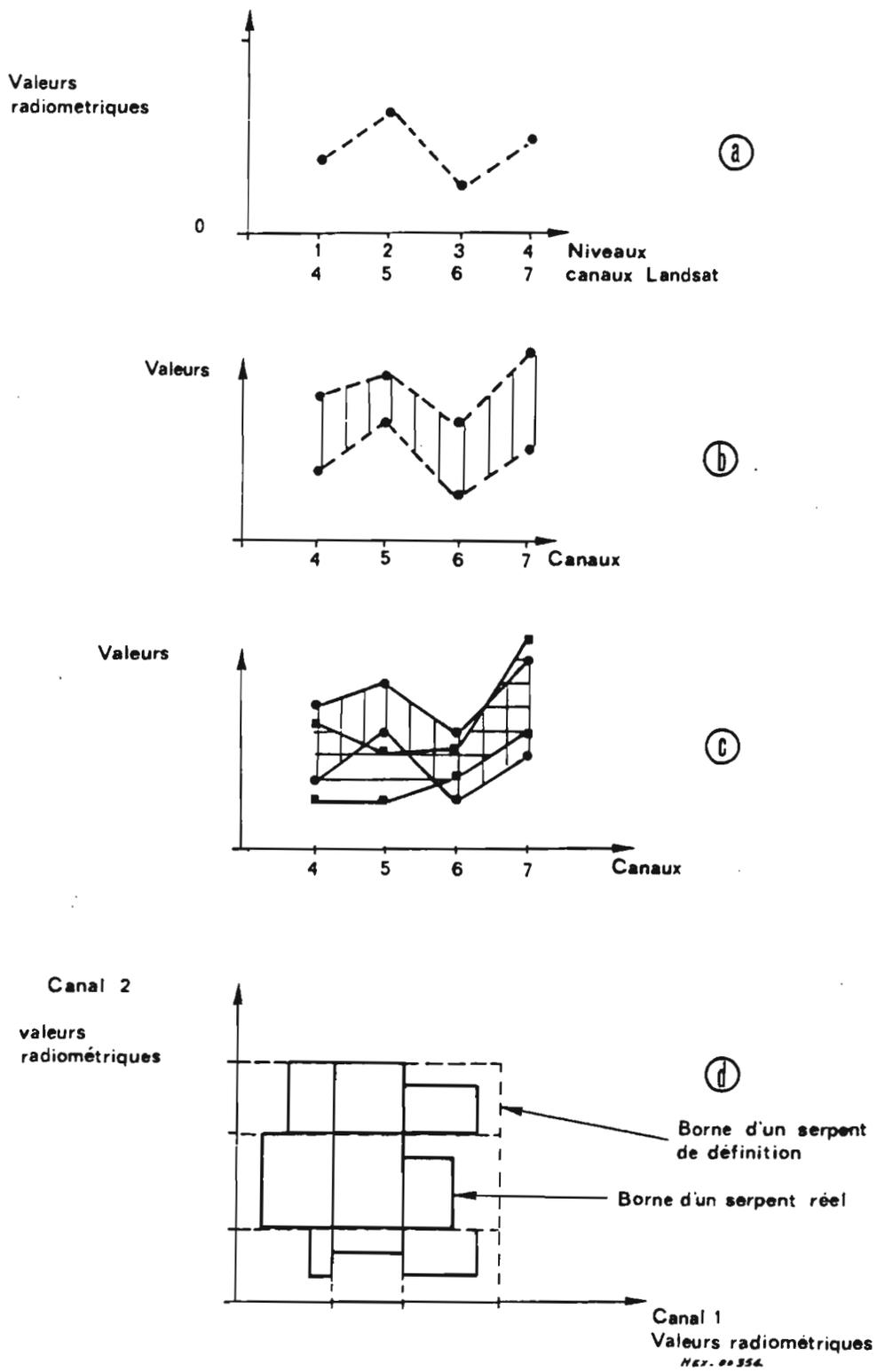


Figure 4 - Notion de serpent.

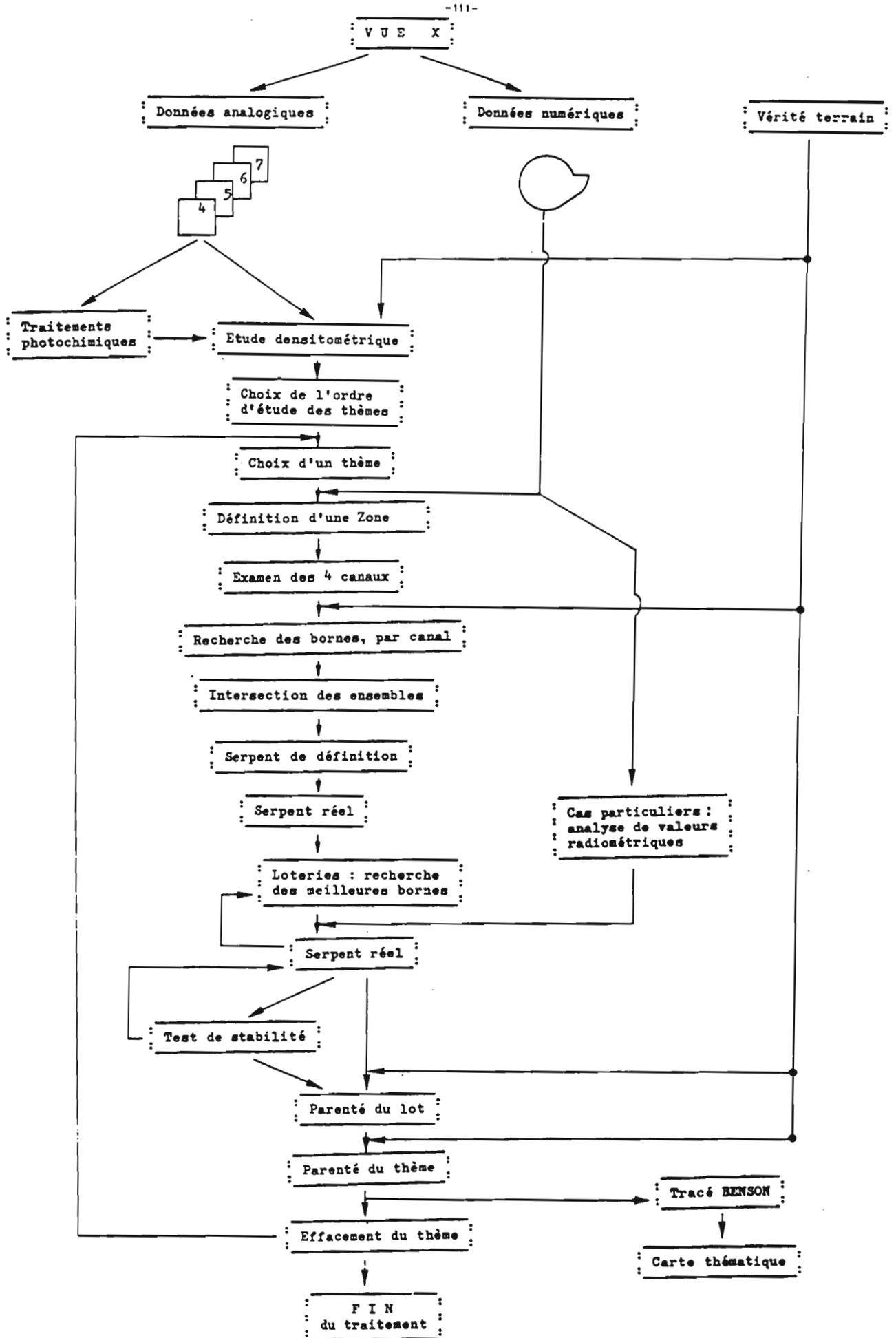


Figure - 5 - Schéma des étapes méthodologiques