

DD. 76. 14

11626

ORGANISATION POUR LA MISE
EN VALEUR DU FLEUVE SENEgal

HAUT COMMISSARIAT

DIRECTION DU DEVELOPPEMENT
ET DE LA COORDINATION

CELLULE D'EVALUATION ET DE
PLANIFICATION CONTINUE

BANQUE DE DONNEES
DES AMENAGEMENTS HYDROAGRICOLEs
DU BASSIN DU FLEUVE SENEgal

MARS 1986

I INTRODUCTION

L'une des opérations les plus importantes, pour la planification continue du développement dans le bassin du fleuve Sénégal consiste en l'analyse et l'évaluation permanente de la culture irriguée ~~qui le secteur maritime du programme de l'OMVS devait entraîner le développement des autres secteurs (navigation, énergie, industrie, services etc...)~~

En effet, la rentabilité, en termes économiques et financiers, de l'important programme d'investissement prévu pour assurer la mise en valeur du bassin, dépend dans une très large mesure du succès de l'introduction à grande échelle, de la culture irriguée avec maîtrise complète de l'eau.

Afin de répondre à cette exigence du développement, l'OMVS a mis en place un système de suivi sur le terrain des performances des aménagements hydroagricoles du bassin.

Un des éléments essentiels de ce système repose sur l'enquête permanente "^{situation} Recensement des Aménagements Hydroagricoles du Bassin" avec les questionnaires "fiches signalétiques" et "bilans de campagne" conçus en 1980 par l'Etude Socio-Economique de l'OMVS.

Compte tenu de l'importance et de la diversité des données à collecter et à analyser, le recours à l'informatique pour le traitement est apparu nécessaire.

C'est ainsi qu'une banque de données sur les aménagements hydroagricoles du bassin a été créé et son exploitation se fait avec le logiciel "LEDA" de l'INSEE installé sur l'ordinateur IBM 4341 du Ministère des Finances du Sénégal, qui constitue un système complet extrêmement élaboré de dépouillement d'enquêtes.

Ce document a pour objet entre autres :

- de présenter les questionnaires "fiches signalétiques" et "bilans de campagne" qui constituent la source principale d'alimentation du système d'informations.

- de bien cerner la chaîne de traitement de l'information afin de situer les contraintes sur son déroulement.

- de présenter les informations contenues dans la banque de données.
- de présenter quelques traitements effectués sur la banque de données,
- de définir les nouvelles orientations nécessaires à la banque de données de manière à la rendre plus opérationnelle.

**II L'ENQUETE : "AMENAGEMENTS HYDRO-AGRICOLE
DU BASSIN DU FLEUVE SENEgal"**

Un des objectifs de la Cellule d'Evaluation et de Planification Continue consiste à préparer à la fin de chaque campagne agricole un bilan décrivant la situation des périmètres irrigués du bassin du fleuve Sénégal.

Cet objectif est poursuivi à deux niveaux :

1) Au niveau d'une campagne donnée

- faire le point de la situation des divers périmètres irrigués, à travers leurs caractéristiques, leurs performances ou contre-performances en considérant un certain nombre d'indicateurs agronomiques et socio-économiques tels que : les surfaces aménagées, le nombre d'exploitants, les rendements, les taux de sinistres des cultures, le coefficient d'intensité culturelle ...

- identifier des éléments favorables et/ou des facteurs limitants ayant conduit aux performances et/ou contre-performances enregistrées.

(*) Traditionnellement, l'année agricole en culture irriguée est divisée en 3 campagnes :

- une campagne d'hivernage
- une campagne de contre-saison froide
- une campagne de contre-saison chaude.

- analyser les résultats obtenus lors de la campagne et formuler des recommandations en direction des organismes et sociétés de développement ...

2) 2) Au niveau du système d'évaluation et de planification continue du développement hydroagricole du bassin

- mise à jour continue des fiches signalétiques des périmètres.

Cette opération bilan de campagne repose sur un vaste système d'informations qui s'articule autour de deux outils d'enquête :

- la fiche signalétique des périmètres : conçue comme un questionnaire, cette fiche se compose d'éléments caractéristiques de chaque périmètre. Elle est établie pour chaque périmètre et constitue ainsi sa "carte d'identité".

- le questionnaire bilan de campagne : il se compose principalement de variables caractéristiques de chaque campagne et constitue le "bulletin de santé du périmètre".

III CHAINE DE TRAITEMENT DE L'ENQUETE "RECENSEMENT DES AMENAGEMENTS HYDROAGRICOLES

Le chaîne de traitement se décompose en plusieurs phases qui sont toutes essentielles.

III.1 Phase de conception et de collecte de données

~~Les questionnaires bilan de campagne et fiche signalétique~~
des périmètres sont conçues par les experts de la Cellule en collaboration avec leurs collègues de l'OMVS et des sociétés nationales de développements (SAED, SONADER, OVSTM).

Pour chaque périmètre existant et pour tout nouveau périmètre, la fiche signalétique est présentée une seule fois sur le terrain.

Quant au questionnaire bilan de campagne, il circule campagne après campagne sur les périmètres et outre les données caractéristiques de la campagne, certaines des données qu'il recueille permettent de réactualiser les données correspondantes sur la fiche signalétique.

Dès la préparation d'une campagne, les questionnaires sont soit déposés auprès des encadreurs des sociétés nationales de développement pour leur remplissage, soit directement remplis par les contrôleurs d'enquête de la cellule lors de missions sur le terrain.

En fin de campagne, les questionnaires sont récupérés. Au retour des questionnaires, un certain nombre de contrôles et de vérifications nécessaires doivent être effectués. Ces contrôles simples consistent en des comptages, ou en une vérification de quelques variables essentielles et ont pour but de déclencher le cas échéant, un retour à l'enquêteur ou une demande de renseignements complémentaires à l'unité enquêtée.

Cette étape de contrôle est indispensable.

III.2 Phase de mise en forme et de saisie de l'information

Elle consiste d'abord au chiffrement (ou codification) et à la saisie de l'information.

Le chiffrement est l'opération manuelle qui assure la transformation des réponses aux questions en variables codées (en chiffres ou en lettres). Il se fait sur des feuilles indépendantes du questionnaires par un contrôleur d'enquêtes formé à cette technique.

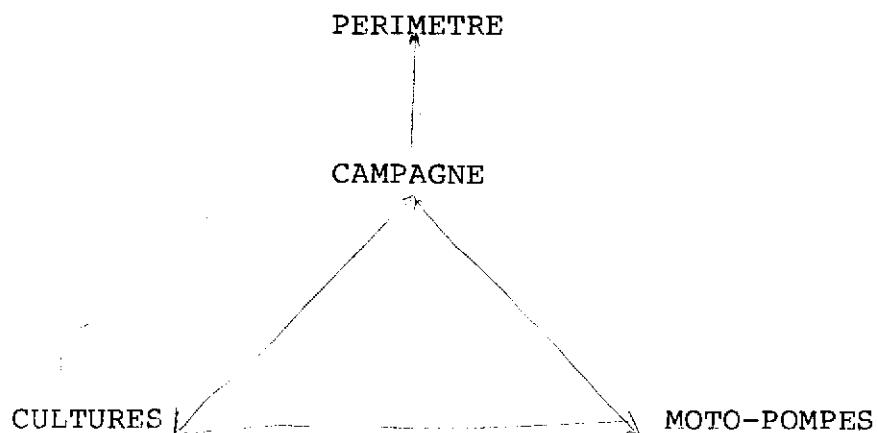
Les informations ainsi chiffrées, pour être exploitables sur ordinateur sont saisies sur un support disque. Les erreurs de saisie sont corrigées par l'opérateur vérificateur de saisie. Quant aux autres erreurs, il appartient au contrôleur d'enquêtes chargé du chiffrement de faire la correction.

III.3 Phase du traitement informatique de l'enquête

a) Banque de données des Aménagements Hydroagricoles du bassin : Organisations du fichier standard

L'ensemble des données ainsi collectées, codifiées et saisies constitue les fichiers de base de la banque de données. L'exploitation des fichiers se fait par le logiciel LEDA^(*) de l'INSEE qui est un générateur de tableaux statistiques.

Le fichier standard LEDA issu des fichiers de base de l'enquête est organisé en plusieurs niveaux reliés entre eux selon une structure d'arbre. On dit qu'il est hiérarchisé.



(*) voir annexe A

On peut trouver à un même niveau plusieurs unités statistiques (exemples : les cultures effectuées lors d'une campagne donnée et les moto-pompes utilisées lors de cette même campagne).

Le fichier standard LEDA reflète fidèlement cette arborescence. Il est caractérisé par :

- une structure logique

Dans le cas de l'enquête des aménagements hydroagricoles, les enregistrements PERIMETRE, CAMPAGNE, CULTURES, MOTO-POMPES, définissent les unités statistiques.

PERIMETRE est de niveau 1

CAMPAGNE est de niveau 2

CULTURES est de niveau 3

Moto-POMPES est de niveau 3

Le chemin parcouru pour accéder aux données d'une unité mineure à partir de l'unité première suit l'arborescence. Ce qui permet de reconstituer l'information d'origine.

Par exemple si l'on veut connaître les productions par spéculation d'un périmètre pour une campagne de donnée, il faudra descendre la branche PERIMETRE, CAMPAGNE, CULTURES.

Les enregistrements de niveau 1 PERIMETRE comprennent toutes les caractéristiques utiles du périmètre : type de périmètre, secteur d'encadrement, appartenance à une zone géographique, nom du village concerné, type de sol, etc...

Les enregistrements de niveau 2 CAMPAGNE portent sur les caractéristiques de la campagne : superficie aménagée, cultivée, récoltée,

Les enregistrements de niveau 3 CULTURES concernent les résultats d'exploitation du périmètre pour campagne : production, rendement, etc...

Les enregistrements de niveau 3 MOTO-POMPES concernent les moto-pompes utilisées par campagne sur un périmètre donné : type, marque, puissance etc...

- un dictionnaire qui, attaché au fichier, le décrit : LEDA peut ainsi reconnaître la structure et la composition de n'importe quel fichier standard.

L'édition du dictionnaire fournit à l'utilisateur :

- la description de la structure logique (nom et niveau de chaque unité statistique)
- les noms des variables disponibles
- les valeurs que peuvent prendre ces variables.

- Une structure physique du fichier standard est celle d'un fichier séquentiel d'enregistrements variables. Dans ce fichier chaque occurrence d'une unité statistique donne lieu à un enregistrement. L'occurrence de l'unité majeure précède dans le fichier les occurrences qui lui sont associées, des unités mineures.

b) Le contrôle de données du fichier standard

C'est une opération essentielle qui permet de tester la qualité de l'enquête de terrain, du chiffrement et de la saisie.

Le contrôle consiste à vérifier qu'il ne manque pas de maillons dans l'anchaînement obligatoire des enregistrements qu'impose la hiérarchisation du fichier.

Le contrôle formel consiste à s'assurer que tout enregistrement CULTURES ou MOTO-POMPES est bien rattaché à un enregistrement CAMPAGNE, et que tout enregistrement campagne est rattaché à un enregistrement PERIMETRE.

A l'inverse, le cas d'un enregistrement PERIMETRE non suivi d'enregistrement CAMPAGNE sera considéré comme normal (le périmètre peut ne pas être exploité pour une campagne donnée), mais afin de mesurer l'importance des superficies non exploitées par campagne et leurs causes, nous faisons suivre systématiquement pour tout périmètre existant, l'enregistrement CAMPAGNE avec la SNI^(*) et la cause de non exploitation. A signaler que le cas d'un enregistrement CAMPAGNE non suivi d'enregistrements CULTURES ou MOTO-POMPES est normal (cas d'un périmètre qui n'a pas exploité, ou qui a eu un sinistre total).

La méthode de correction d'erreurs que nous utilisons dans tous les cas, sauf si l'erreur est trivial est le retour au document de base.

A ce stade, le fichier standard est mis au propre sur le plan formel et débarrassé des erreurs individuelles flagrantes. Dès lors, la production de tableaux statistiques en série et leur édition sur imprimante sous forme de listing peuvent être envisagées pour l'analyse.

REMARQUES

- Pour les besoins des utilisateurs de la cellule et du Haut-Commissariat, des tableaux statistiques sont fournis sur simple demande adressée au responsable informatique de la cellule. Quant aux autres utilisateurs extérieurs à l'OMVS, les travaux ne se font qu'après approbation du chef de la cellule.

- En dehors des tableaux fournis à la demande des utilisateurs, la cellule produit périodiquement un ensemble de documents issus de l'exploitation destinés à nos collègues des sociétés nationales de développement et aux services nationaux intéressés.

(*) SNI = Surface nette irrigable

IV NATURE DES INFORMATIONS CONTENUES DANS LA BANQUE DE DONNEES

L'objet de ce chapitre est de présenter aux demandeurs de tableaux, les variables retenues pour la banque de données. Ces variables permettent de caractériser les entités périmètres irrigués, campagnes-agricoles et les résultats de campagne. Pour avoir plus de renseignements sur les valeurs des codes assignés aux variables, on pourra se reporter au document "Banque de données : description de fichier pour la gestion informatique" disponible à la Cellule.

IV. 1 Variables caractéristiques du périmètre irrigué

Elles peuvent être subdivisées en plusieurs groupes.

a) Variables permettant l'identification du périmètre

- nom du périmètre
- numéro de code du périmètre
- nom de code du pays
- organisme de tutelle
- existence d'une CUMA sur le périmètre
- type de périmètre
- date de démarrage
- type de génération du périmètre

b) Variables permettant la localisation du périmètre

- situation géographique
- nom de code du principal village concerné par l'aménagement
- nom de la principale communauté rurale
- nom de code du département ou de la région administrative
- numéro de code du secteur d'encadrement pour les périmètres encadrés
- origine du pompage
- distance de la ville la plus proche
- distance du centre d'expérimentation ou de recherche agricole le plus proche du périmètre
- nombre de villages exploitant le périmètre.
- distance moyenne des villages au périmètre

c) Variables liées à la nature de l'aménagement

- existence de levé topographique du périmètre
- ~~existence de l'analyse pédologique~~
- longueur du canal principal
- pourcentage de sol de type Hollaldé
- pourcentage de sol de type faux-hollaldé
- pourcentage de sol de type fondé
- pourcentage de sol de type autres sols
- mode de planage du périmètre au démarrage
- qualité du planage
- conditions pédologiques générales du périmètre
- causes de mauvaises conditions pédologiques
- qualité des aménagements

d) Variables liées aux opérations mécaniques sur le périmètre

- Principales tâches du pompiste (vidange, changement de filtres, graissage, réparation)
- Principales tâches du mécanicien (vidange, changement de filtres, graissage, réparation)
- Principales opérations mécanisées sur le périmètre aux démarriages (travail au sol, épandage d'engrais, semis, désherbage, fauchage, battage)

e) Variables liées à l'organisation sociale du périmètre

- pourcentage approximatif des wolofs parmi les attributaires de parcelles
- pourcentage approximatif de toucouleurs
- pourcentage approximatif de maures
- pourcentage approximatif de soninkés
- pourcentage approximatif de khassonkés
- pourcentage approximatif de bambaras
- pourcentage approximatif d'autres ethnies
- mode d'exploitation du périmètre au démarrage
- mode d'exploitation du périmètre actuellement

- recours au privé pour l'approvisionnement (en semences, en engrais, en produits phytosanitaires, en gasoil, en huile, en pièces détachées, en petits matériels divers, en gros matériel agricole, en moto-pompe)
- pénétration du privé au niveau de l'approvisionnement
- origine des semences
- recours au secteur privé (pour les opérations mécanisées pour le travail au sol, pour l'épandage des engrais, pour le semis, pour le désherbage, pour le fauchage, pour le battage)

d) Variables liées à l'évaluation des superficies sinistrées

- superficies sinistrées totalement (ou partiellement) par le riz rouge
- degré d'infestation sur les superficies sinistrées partiellement par le riz rouge
- méthodes de lutte utilisée contre le riz rouge
- superficies sinistrées totalement (ou partiellement) par le riz à rhizomes
- degré d'infestation sur les superficies sinistrées partiellement par le riz à rhizomes
- méthodes de lutte utilisée contre le riz à rhizomes
- superficies sinistrées totalement (ou partiellement) par d'autres adventices
- degré d'infestation sur les superficies sinistrées partiellement par d'autres adventices
- nature des adventices
- méthodes de lutte utilisées contre les adventices
- superficies sinistrées totalement (ou partiellement) par des insectes
- degré d'infestation sur les superficies sinistrées partiellement par des insectes
- nature des insectes
- méthodes de lutte utilisées contre les insectes
- superficies sinistrées totalement (ou partiellement) par les maladies
- nature des maladies

- méthodes de lutte utilisées contre les insectes
- superficies sinistrées totalement (ou partiellement) par des nématodes
- méthodes de lutte utilisée contre les nématodes
- superficies sinistrées totalement (ou partiellement) par d'autres types d'infestations
- degré de dégâts sur les superficies sinistrées partiellement par d'autres types d'infestations
- nature des infestations
- méthodes de lutte utilisée contre ces types d'infestation
- superficies sinistrées totalement (ou partiellement) par des oiseaux
- nature des oiseaux
- méthode de lutte contre les oiseaux
- superficies sinistrées totalement (ou partiellement) par des rongeurs
- degré de dégâts sur les superficies sinistrées partiellement par des rongeurs
- nature des rongeurs
- méthode de lutte utilisée contre les rongeurs
- superficies sinistrées totalement (ou partiellement) par d'autres déprédateurs
- degré de dégâts sur les superficies sinistrées partiellement par d'autres déprédateurs
- nature des déprédateurs
- méthodes de lutte utilisée contre les déprédateurs

e) Variables liées aux stockages et à la commercialisation de la production

- difficultés de stockage des récoltes
- difficultés de commercialisation
- difficultés d'évacuation
- usinage de la production
- utilisation des sous produits

f) Variables liées à la participation des attributaires
à l'entretien du périmètre

- entretien à l'affinage du planage
- entretien des diguettes
- entretien des drains
- entretien des canaux principaux
- entretien des canaux secondaires
- entretien des canaux tertiaires
- entretien des canaux arroseurs

g) Variables à l'appréciation de la campagne

- origine des mauvais rendements
- origine des bons rendements
- réhabilitation complète du périmètre

IV.3 Variables liées aux résultats économiques de la campagne

- production par spéculation et par variété
- rendement par spéculation et par variété
- superficie par spéculation et par variété

Il est à noter que, seules sont prises en compte les trois spéculations qui ont donné les meilleurs résultats ou qui sont les plus cultivées.

REMARQUE

Dans les fiches questionnaires "bilan de campagne" figurent les variables liées aux moto-pompes et aux compte d'exploitation des périmètres (intrants, prestations de services...).

Les données relatives aux comptes d'exploitation constituent un fichier qui est sur le point d'être réalisé, tandis que les données portant sur les moto-pompes et d'une manière générale sur le machinisme agricole feront l'objet d'une autre fichier.

Il est à noter que tous ces fichiers vont être inclus dans la banque de données des aménagements hydroagricoles.

V. TRAITEMENTS EFFECTUÉS SUR LA BANQUE DE DONNÉES

L'historique des campagnes étant reconstituée depuis 1975, toutes les données relatives aux campagnes jusqu'à la contre saison chaude 1985 sont introduites dans la banque de données et traitées.

Parmi les multitudes tableaux produits, on peut noter entre autres :

a) les tableaux faisant l'historique du développement agricole depuis la création des périmètres

Il s'agit principalement :

- d'un ensemble de tableaux qui fournissent par périmètre et pour chaque campagne depuis l'année de démarrage jusqu'à la campagne de contre saison chaude 1985, les données qui suivent :

- superficie nette irrigable
- superficie cultivée
- superficie récoltée
- taux de non exploitation et leurs causes
- taux de sinistre et leurs causes
- production et rendement par spéculation
- superficie cultivée, récoltée par spéculation

- d'un ensemble de tableaux faisant état de données agrégées au niveau de chaque campagne d'hivernage de 1975 à 1984 réparties :

- par pays et par type
- par pays, par secteur d'encadrement et par type de périmètre
- par pays, par organisme de tutelle et par type de périmètre
- par zone géographique et par type de périmètre
- pour l'ensemble du bassin et par type de périmètre.

Au niveau des données agrégées on peut citer :

- la superficie nette irrigable
 - la superficie cultivée
 - la superficie récoltée
 - la superficie non exploitée
 - la superficie sinistrée
 - le taux de sinistre
 - le taux de non exploitation
 - le taux de mise en oeuvre
 - le taux de productivité culturale
 - le nombre d'exploitants
 - la superficie cultivée par exploitant
 - la superficie récoltée par exploitant
 - la superficie sinistrée par exploitant
 - les coefficients d'intensité culturale
- b) les tableaux spécifiques devant servir à l'élaboration des rapports périodiques de campagne par la cellule
- c) les tableaux produits sur demande des utilisateurs pour leurs besoins propres.

V. PERSPECTIVES D'EVOLUTION DU BILAN DE CAMPAGNE ET DE LA BANQUE DE DONNEES

Le bilan de campagne permet de cerner les performances de la culture irriguée dans le bassin et de fournir des indications qualitatives et quantitatives devant guider l'analyse et l'évaluation qui doivent :

- déboucher sur les recommandations destinées aux sociétés nationales de développement et aux opérateurs intervenant dans le développement du bassin.
- proposer des mesures correctives et des solutions concrètes pour surmonter les contraintes et difficultés identifiées,
- définir les domaines méritant un approfondissement des connaissances (études ad hoc ponctuelles à mener)
- formuler des recommandations en direction de l'OMVS pour la planification du développement intégré du bassin
- formuler des recommandations précises pour un ajustement recherche/développement.

Cette opération bilan de campagne, bien que répondant dans ses grandes lignes aux objectifs premiers qui lui sont assignés comporte cependant certaines imperfections qui méritent d'être corrigées. En effet, vu la masse de données que renferment les questionnaires, il n'y a que très peu de variables utilisées comme base d'analyse dans nos rapports. Parmi les variables non utilisées, on peut citer entre autres :

- les variables ayant trait à l'évaluation des superficies sinistrées réparties par cause de sinistre qui à priori semblent difficiles à déterminer.
- les variables liées aux caractéristiques des périmètres
- les variables liées à la préparation de la campagne.

Dès lors, une discussion autour de ces questions mérite d'être engagée au niveau de la cellule afin de repréciser ou d'élaguer au besoin certaines variables.

Comme le bilan de campagne, la banque de données est un outil dynamique qui évolue au fil du temps et elle est sensible à toute perturbation ou à toute nouvelle approche intervenue au niveau des questionnaires. Aussi le travail d'actualisation et de mise à jour de la banque de données devient un processus permanent.

Dans les questionnaires bilans de campagne, les données ayant trait aux coûts de production ont été introduites depuis la campagne d'hivernage 1984. Ces données vont constituer le fichier "compte d'exploitation" de la banque de données qui est sur le point d'être finalisé. Un document concernant ce fichier sera produit ultérieurement.

Quant aux données relatives aux moto-pompes, deux solutions peuvent être envisagées :

- la première consiste à considérer les moto-pompes comme une entité propre. Dans ce cas, le problème est relativement simple. Il suffit de définir le contenu à donner à l'entité moto-pompes et de constituer le fichier correspondant.
- la deuxième revient à considérer les moto-pompes comme faisant partie d'un vaste système d'information sur le machinisme agricole.
- le principal avantage de cette solution réside dans le fait qu'elle permet de cerner tout un secteur comprenant plusieurs entités homogènes. Par contre, elle présente l'inconvénient d'être plus ou moins lourde, nécessitant une étude approfondie de toutes les entités devant faire partie du système et de déterminer tous les types de relations entre entités. La réalisation d'un tel projet, revient à créer une base de données. Parmi ces deux solutions, la cellule doit là aussi se prononcer.

VI. PERSPECTIVES DANS LE TRAITEMENT DE L'INFORMATION

a) Contraintes sur le déroulement de la chaîne de traitement de l'information

L'organisation générale du dépouillement consiste à définir le "chemin critique" c'est à dire à repréter les phases de la chaîne de traitement de l'information pour lesquelles un retard de réalisation entraînera automatiquement un retard du travail dans son ensemble. Ce qui revient à faire une étude critique des différentes phases afin de déterminer s'il y a lieu des contraintes dans l'exécution des tâches et au besoin d'y apporter des solutions adéquates.

De l'examen des différentes procédures de la chaîne de traitement, il ressort trois contraintes majeures :

- la lenteur de la collecte des questionnaires : c'est une étape difficile à maîtriser par la cellule qui est demandeur de services auprès des sociétés nationales de développement et des privés qui sont des détenteurs de données. A cet égard, il importe de trouver une méthode appropriée de collecte de questionnaires de concert avec nos différents collaborateurs.
- la lenteur du contrôle et de la codification des questionnaires : un contrôle plus assidu devra être exercé par les experts (l'agronome et l'informaticien), sur les contrôleurs d'enquête chargés d'effectuer ce travail.
- la lenteur de la production et de diffusion des tableaux : à ce niveau, une meilleure coordination entre les utilisateurs demandeurs de tableaux et l'informaticien doit être recherchée.

b) Matériel de traitement configuration de la station graphique de la cellule

- 1 ordinateur HP 9836 A comprenant :

- 1 microprocesseur Motorola 68000 (16/32 bits)
 - 1 mémoire centrale de 768 K octets
 - 2 lecteurs de disquettes 5 pouces 1/4 de 264 K de mémoire de masse
 - 1 écran graphique de 12 pouces 2 de diagonale présentant une résolution de 390 x 512 points
 - 1 interface HP - IB
-
- 1 unité de disquette 8 pouces HP 9895 A (1,2 M octets par disquette double face, double densité + possibilité de formation IBM 3740)
 - 1 imprimante graphique HP 2631 G vitesse 180 c/s
 - 1 traceur HP 7580 B (feuille de 622 mm x 1231 mm maximum)
 - 1 table à digitaliser + 1 contrôleur SUMMAGRAPHICS modèle ID (surface de travail active 914,4 x 1219,2 mm et résolution de 0,1 mm).

En outre, la cellule dispose d'un important lot de logiciels ^(*) orientés vers le traitement statistique, l'analyse numérique et l'analyse de données.

Remarque

Vu la masse importante de données à traiter, l'achat d'un disque dur d'une vingtaine de Mega octets ~~doit être~~ envisagé.
est

Configuration du Ministère des Finances

- 2 ordinateurs IBM 4341 (12000 K de mémoire centrale) sous MVS
- 2 x 32 consoles 3278 - CO2/3279 - CO2
- 24 unités disques 3375 - B1 (+ 400 M bytes chacune)

(*) voir annexe B

- 1 contrôleur de communication 3705
- 5 imprimantes 3203-5
- 2 lecteurs de disquettes 8 pouces 3540-002
(formatage 3740)
- 8 dérouleurs de bandes magnétiques.

Remarque

Tous les travaux relatifs à la banque de données des aménagements hydro-agricoles se font sur cette configuration - Avec l'acquisition du micro-ordinateur IBM PC/AT par la cellule, nous envisageons de transposer progressivement nos applications - Ce qui contribuera d'une certaine mesure à réduire le temps de traitement.

c) Fichiers cartographiques des données

La station cartographique de la cellule permet de produire des cartes à partir des fichiers cartographiques ⁽¹⁾

On distingue deux types de cartes :

- les cartes figuratives : ce sont des cartes de présentation de fonds uniquement (exp : inventaire des portefeuilles d'étude ...).
- les cartes thématiques : elles utilisent les fonds de plans de cartes figuratives et sur lesquelles sont superposées des diagrammes ou symboles correspondant à des indicateurs calculés dans la banque de données (exp: évolution des superficies par secteur d'encadrement de 1975 à 1983 ; évolution de la production de riz de 1980 à 1983 par secteur d'encadrement).

(1) un fichier cartographique renferme principalement les coordonnées de points (ville, barrage, ...) et les coordonnées des lignes (lit du fleuve, limites de délégations ou de secteurs d'encadrement).

Ce mode de représentation de données qui permet de visualiser simultanément sur carte l'effet de plusieurs indicateurs dans la banque de données est un instrument très utile qu'il convient de rationaliser l'utilisation pour l'évaluation et le suivi des aménagements hydro-agricoles.

d) Perspective dans l'utilisation de la technique de l'analyse de données

L'objectif de planification et de suivi de la situation du développement agricole dans le bassin du fleuve Sénégal assigné à la cellule exige des moyens fiables de collecte et de traitement de données.

La méthode classique d'analyse qui consistait à vouloir expliquer le comportement d'une population par une étude détaillée de variables prises individuellement n'est plus satisfaisante. Dans un environnement complexe comme le développement du bassin, les données collectées ne peuvent être analysées qu'au moyen de méthodes aptes à prendre en compte simultanément l'ensemble des variables, afin de dévoiler le jeu de leurs interactions - Ceci doit nous inciter à utiliser de plus en plus les méthodes d'analyse de données. En cela la cellule dispose d'une panoplie de logiciels qu'il convient de rentabiliser.

ANNEXE-A

SYSTEME LEDA

PRESENTATION DU LOGICIEL LEDA

1. Champ d'application

LEDA automatise la plus grande partie du dépouillement d'enquêtes, pratiquement de la lecture des cartes perforées à l'édition des tableaux.

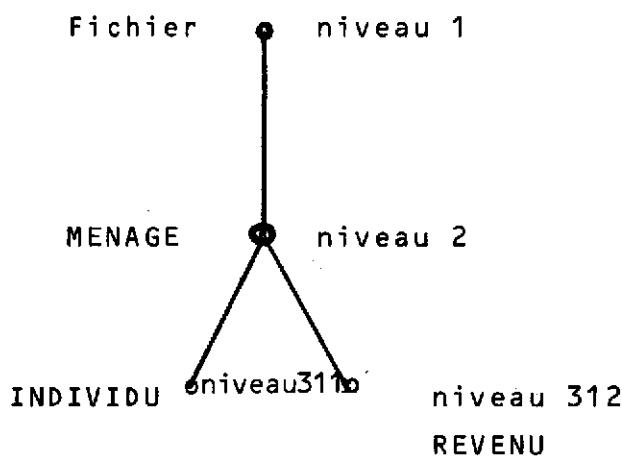
LEDA est composé d'un ensemble homogène de programmes utilisés dans de multiples combinaisons. Il se décompose en trois phases :

- la phase CASTOR réalise les contrôles de structures et de validité et crée le fichier standard,
- la phase POLUX opère sur ce fichier standard et crée un nouveau fichier standard. Une ou plusieurs phases POLUX permettent la détection des incohérences, redressement de l'échantillon, création de nouvelles variables,
- la phase HELENE, crée en une ou plusieurs fois les tableaux statistiques.

2. Le fichier standard

Le fichier d'une enquête est toujours organisé en plusieurs niveaux, reliés logiquement entre eux selon une structure d'arbre.

Exemple :



On peut trouver à un même niveau plusieurs unités statistiques (exemple : les individus qui constituent le ménage et les revenus du ménage).

Le fichier STANDARD LEDA reflète fidèlement cette arborescence.

Il est caractérisé par :

- une structure logique

On dit que l'unité statistique FICHIER est de niveau 1, l'unité statistique MENAGE est de niveau 2 et les unités statistiques INDIVI DU et REVENU de niveau 3.

- un dictionnaire qui, attaché au fichier, le décrit : LEDA peut ainsi reconnaître la structure et la composition de n'importe quel fichier standard.

- l'édition du dictionnaire fournit à l'utilisateur :

- la description de la structure logique (nom et niveau de chaque unité statistique)
- les noms des variables disponibles
- les valeurs que peuvent prendre ces variables.

Les variables sont classées en variables qualitatives (codes) ou quantitatives (quantités).

La structure physique du fichier standard est celle d'un fichier séquentiel d'enregistrements variables. Dans ce fichier chaque occurrence d'une unité statistique donne lieu à un enregistrement. L'occurrence de l'unité majeure précède dans le fichier les occurrences, qui lui sont associées, des unités mineures. On réalise ainsi une "mise en facteur" de l'information (gain de place, gain de temps de traitement).

A l'intérieur d'un enregistrement, les codes sont en mode caractère, les quantités en un mot binaire.

Tout enregistrement commence par l'indicatif de l'unité statistique qu'il représente.

Exemple de portion du fichier :

| | | | |
|-----------|-------------|----------------|----------------|
| 1 fichier | 21 Menage A | 311 individu 1 | 311 individu 2 |
| | | 312 revenu 1 | 312 revenu 2 |
| | | 312 revenu 3 | |
| | 21 Menage B | 311 individu 3 | 311 individu 4 |
| | | 312 revenu 4 | 312 revenu 5 |

ect...

.../

Les individus 1, 2 ainsi que les revenus 1, 2 et 3 sont reliés logiquement au ménage A.

Les individus 3, 4 ainsi que le revenu 5 sont reliés logiquement au ménage B etc...

Il n'y a pas de relation logique, entre le revenu 1 ou le revenu 2 ou le revenu 3 et les individus 1, 2 ni entre le revenu 4 et les individus 3, 4.

Si le fichier à exploiter est issu directement de la saisie, le fichier standard sera créé par la phase CASTOR.

Si le fichier à exploiter a une structure déjà connue et vérifiée, on pourra éviter cette phase à l'aide d'un module de lecture.

ANNEXE B

LOGICIELS DISPONIBLES

I. LOGICIELS DISPONIBLES SUR LE MATERIEL HELWETT PACKARD

I.1. CAT (cartographie analytique et thématique). Il permet la réalisation de cartes analytiques et thématiques au moyen d'une station de travail (voir configuration) en mode inter actif graphique. Il donne la possibilité de constituer :

- des fonds de plans à partir de documents existants, quelques que soient leurs échelles ;
- de reporter sur ces plans des situations existantes, des projets, des objectifs à atteindre ;
- d'habiller ces plans par un large éventail de symboles, de traits, et par de nombreuses possibilités de lettrages,
- de représenter sur ces plans des variables ou groupes de variables associés à une localisation spatiale, soit ponctuelle soit liée à un contour ;
- de composer des planchettes avec cartouches individualisés, légendes et plans à l'échelle fixée par l'utilisateur ;
- d'effectuer automatiquement les dessins en noir et blanc, ou en couleur ;
- de conserver toutes les données introduites, tant pour la constitution de plans que pour leur composition, leur habillage et leur représentation.

Par ses possibilités en matière de cartographie automatique, le CAT permet :

- une analyse de variables n'ayant pas la même distribution spatiale ;
- la mise en évidence de variables fortement corrélées
- la découverte de facteurs parfois inconnus, liés à la localisation ;
- la comparaison de cas particuliers, simultanément à une vue d'ensemble ;
- une compréhension plus rapide d'une situation ;
- un travail en groupe ou commission plus efficace.

I.2. Une bibliothèque de logiciels statistiques HP couvrant pratiquement tous les domaines tel que :

a) - la gestion des données statistiques avec toutes les possibilités de création de fichiers, de mise à jour, de concaténations de fichiers, de constitution de nouveaux fichiers à partir d'anciens...

b) - Les statistiques fondamentales (moyenne, écart type, variance, coefficients d'aplatissement et d'assymétrie...) matrices de corrélations, mesures de tendances statistique (maximum, minimum, moyenne, médiane, mode, quartiles, déciles...)

c) - L'analyse de régression

Les programmes de régression portent entre autres sur :

- la régression multiple linéaire
- la régression polynomiale
- la régression pas à pas
- la régression non linéaire
- les modèles standards de régression non linéaire du genre :

$$Y = A (X^T B) + C$$

$$Y = A \text{ Exp } (BX) + C$$

$$Y = A \text{ Exp } (BX) + C \text{ Exp } (DX) + E$$

$$Y = A \text{ Sin } (BX) + C \text{ Cos } (DX) + E$$

d) - La représentation graphique des données du genre :

- séries temporelles
- histogrammes
- distribution normale
- distribution de Weibull
- distribution des points sur un plan liant deux variables
- graphes semi-log, log-log
- graphes à 3 dimensions
- graphe d'Andrew
- etc...

e) - Statistiques générales

Ce chapitre permet d'estimer des paramètres d'une population (moyenne, variance...) à partir des statistiques d'échantillonnage correspondantes. Il est composé de 5 modules qui permettent de :

- voir si les observations d'un échantillon sont mutuellement indépendantes, si la moyenne d'un échantillon est significative, de comparer une donnée à une distribution normale, exponentielle, uniforme

- de comparer les moyennes de 2 échantillons
- de tester si les moyennes, médianes de 2 échantillons indépendantes sont égales et si les 2 populations ont la même distribution,
- de tester si la moyenne de plusieurs échantillons sont égales et s'il y a des différences significatives entre couples de moyennes
- d'étudier des séries de distribution continues (normales, Gamma, Beta, t de Student, , Laplace, Logistique, Weibull) et discrètes (Binominales, Poisson, hypergéométrique) et d'autres fonctions statistiques (coefficients binominales, Fonctions Gamma, Fonction Beta, N Factoriel).

f) - Analyse de données

Les différents facteurs qui ont une influence sur un phénomène sont traités tous ensemble et non plus seulement deux à deux.

Trois modèles sont présentés dans ce module :

- l'analyse factorielle des correspondances,
- l'analyse en composantes principales ;
- l'analyse de variance.

g) - Utilitaires de simulation de Monte Carlo

Des générateurs de nombres aléatoires sont fournis pour certains distributions (Beta, binomiale, χ^2 , exponentielle, Gamma, normale, Pareto, Poisson, e..) avec des tests pour déterminer leur caractère aléatoire (test de Kolmogorov - Smirnov, test χ^2) Ce modèle comprend aussi deux utilitaires élémentaires d'échantillonnage. Le premier permet de sélectionner un échantillon aléatoire simple sans remplacement ; le second arrage un tableau de nombres dans un ordre aléatoire.

h) - Analyse numérique

Ce module permet :

- la recherche de racines d'une fonction par :
 - . la méthode bisectron
 - . la méthode secante
 - . la méthode Muller
 - . la méthode Siljak

- de faire des intégrations :
- la règle Simpson
- La méthode de Filon
- La méthode adaptative Romberg
- de résoudre des équations différentielles
- de faire des interpolations
- etc...

1.3. PRESENTATIONS GRAPHIQUES HP

Permet de créer :

- des diagrammes en bâtons, en "camembert",
- des courbes,
- de combiner du texte avec des diagrammes. La sortie des dessins se fera soit sur imprimante en noir et blanc, soit sur table traçante avec des couleurs.

1.4. VISICALC

C'est un tableau qui intègre des fonctions logiques, mathématiques, de tri... Il est très utile pour des travaux répétitifs nécessitant beaucoup de calculs.

1. OBJECTIFS DU BILAN DE CAMPAGNE

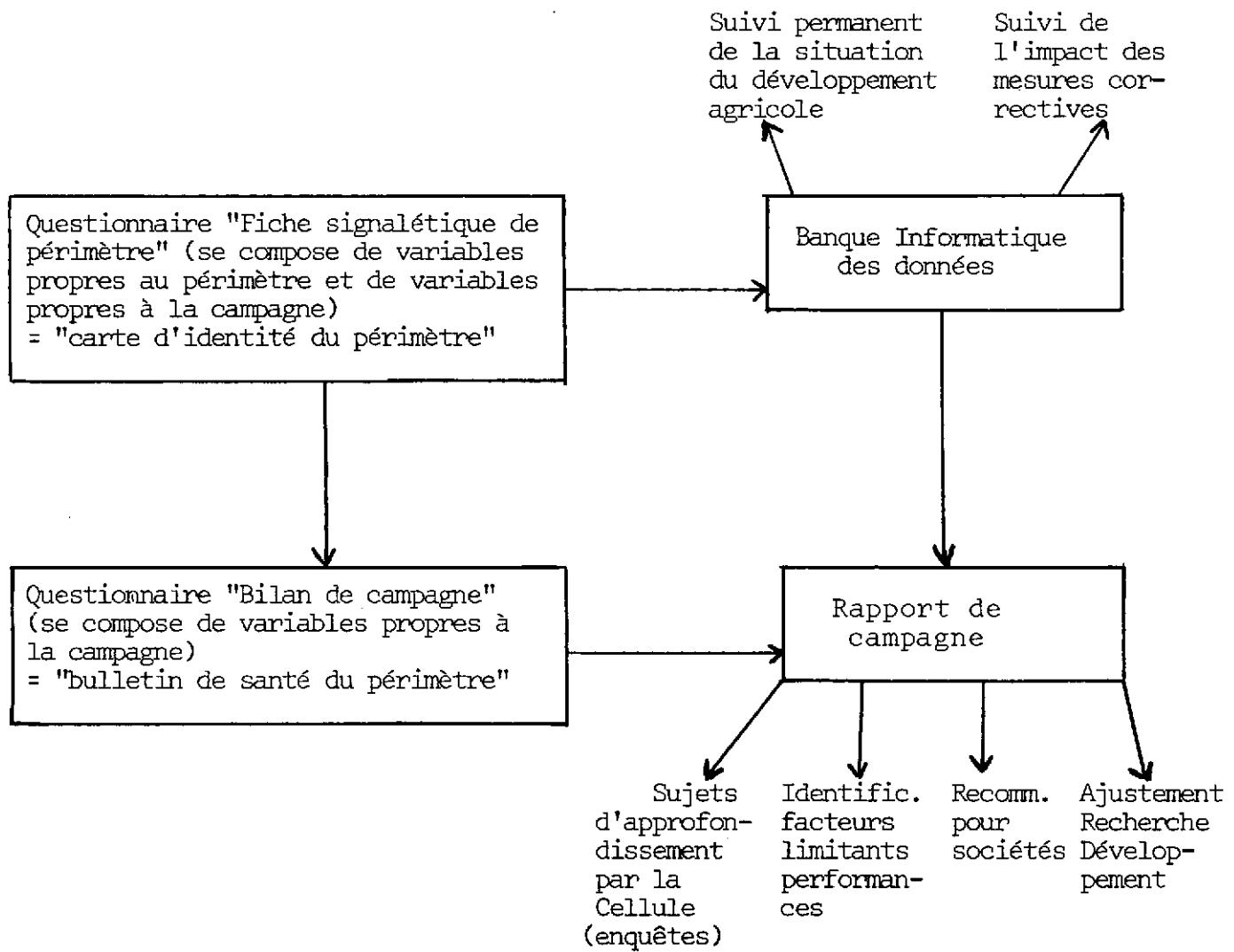
1.1. - Dans le cadre d'une campagne donnée

- faire le point de "l'état de santé" des divers pérимètres irrigués à travers leurs performances et un certain nombre d'indicateurs socio-économiques tels que rendements des cultures, taux de non-exploitation de superficies, taux de sinistres des cultures, coefficient d'intensité culturelle, niveau des charges et des revenus d'exploitation, taux de commercialisation, taux d'endettement.

1.2. - Dans le cadre plus large du système d'évaluation et de planification continue

- La mise à jour continue des fiches signalétiques des pérимètres et de la banque de données.
- La banque de données permettrait d'avoir à tout moment une vision précise de la situation du développement agricole du bassin du fleuve Sénégal, de suivre en vraie grandeur l'impact des mesures correctives adoptées, au niveau des politiques et des méthodes de développement aussi bien à l'échelle de l'ensemble du bassin d'un pays de l'OMVS, ou d'une zone géographique, qu'au niveau d'un type de pérимètre, voire de chaque aménagement hydro-agricole pris individuellement.

1.3. Résumé



2. PRESENTATION DES DIFFERENTS TABLEAUX

La Banque de données de la CEPC est aujourd'hui à jour, car l'on dispose des résultats de la contre saison chaude 1984. Par ailleurs, les dispositions sont prises pour le lancement du bilan de campagne hivernage 1984. Cependant, nous allons concentrer l'analyse dynamique du développement de la culture irriguée dans le bassin du fleuve Sénégal sur les résultats des campagnes d'hivernages entre 1975-1983. Cette analyse devra contribuer à l'enrichissement du diagnostic, base de toute planification qui se veut axée sur la définition de politiques économiques appropriées à la situation socio-économique du bassin du fleuve Sénégal, qu'on veut dynamiser pour :

- réduire le déficit céréalier actuel des différents pays de l'OMVS, et la vulnérabilité économique de la sous-région ;
- assurer un développement généralisé et accéléré de la région du fleuve ;
- assurer un effet optimal de ce développement sur la consolidation des éco-systèmes et le développement des régions avoisinantes par association de leurs productions avec celle du fleuve.

Afin de faciliter l'interprétation ultérieure des données, un certain nombre de valeur et de ratios sont présentés aux utilisateurs potentiels. Une partie de ces chiffres a déjà fait l'objet d'une interprétation soit dans le cadre de l'étude socio-économique pour la période (1975-1979), soit dans les différents bilans de campagne 1980-1981 et 1981-1982. Le bilan de campagne 1982-1983 est en cours de finalisation et le bilan de campagne 1983-1984 est engagé.

.../

~~Cependant, les bilans de campagne fournissent une analyse ponctuelle pour une année donnée, or le travail envisagé dans l'exploitation des chiffres ci-dessous présentés, devrait permettre de déceler les tendances actuelles du développement hydro-agricoles du bassin du fleuve Sénégal, tendances qu'ils faudraient peut être corrigé pour que les Etats de l'OMVS atteignent ou approchent les objectifs qu'ils se sont assignés.~~

Les tableaux qui sont au nombre de _____ se présentent de la manière suivante :

A) Pour les superficies

- 1) - superficie nette irrigable
- 2) - superficie cultivée
- 3) - superficie récoltée
- 4) - superficie sinistrée = différence entre la superficie cultivée et la superficie récoltée
- 5) - superficie non exploitée = différence entre la superficie nette irrigable et la superficie cultivée
- 6) - taux de sinistre = superficie sinistrée que divise la superficie cultivée
- 7) - taux de non exploitation = superficie non exploitée que divise la superficie nette irrigable
- 8) - taux de mise en valeur = superficie cultivée que divise la superficie nette irrigable
- 9) - la productivité culturelle = superficie récoltée que divise la superficie cultivée (ce ratio peut-être utilisé comme proxy des rendements)

B) Pour le nombre d'exploitant et les superficies

- 10) - superficie cultivée par nombre d'exploitants
- 11) - superficie récoltée par nombre d'exploitants
- 12) - superficie sinistrée par nombre d'exploitants
- 13) - nombre d'exploitant.

C) Pour les productions

- 14) - production de riz par pays et par type de périmètre
- 15) - production de riz par pays et par secteur d'encadrement
- 16) - production de riz par pays et par organisme de tutelle
- 17-18-19) - même tableau pour le maïs
- 20-21-22) - même tableau pour le sorgho
- 23-24-25) - même tableau pour le niébé.

Pour chaque spéculation, les tableaux indiquent :

- a) la production
- b) la superficie récoltée
- c) le rendement à l'hectare.

Ces différents paramètres sont présentés :

- 1) par pays et par type de périmètre (1)
- 2) par pays et secteur d'encadrement
- 3) par pays et organisme de tutelle
- 4) par zone géographique et par type de périmètre
(delta ; basse vallée ; moyenne vallée ; haute vallée ; haut-bassin.
- 5) pour l'ensemble du bassin du fleuve Sénégal et par type de périmètre.

(1) Les types de périmètres : * GP = grand périmètre
 * MP = moyen périmètre
 * PP = petit périmètre
 * CU = CUMA (Coopérative d'utilisation du matériel agricole
 * PV = périmètre villageois.

3. DIFFICULTES RENCONTREES - PERSPECTIVES

- Lenteur dans la collecte des données : très souvent, à cause du manque de cohérence entre les données fournies par les services du siège et les agents de terrain d'une même société de développement des recouplements et quelquefois des réunions de concertations sont nécessaires afin d'homogénéiser les données sur une base objective. Ce souci de rigueur dans la collecte des données est déterminant pour la qualité de l'information recueillie.
- Lenteur dans le traitement des données. Ce goulot d'étranglement sera facilement levé dès que l'OMVS disposera d'un ordinateur propre. Les tableaux statistiques et les résultats d'analyse pourront alors être rapidement fournis à la demande et les rapports de campagne seront disponible en temps opportun.
- Il serait opportun d'arriver à la situation où les résultats préliminaires d'un bilan de campagne soient disponible avant le lancement de la campagne suivante pour permettre aux S.N.D. de prendre les mesures correctives nécessaires notamment à une amélioration des rendements.

4. APERCU SUR LES CAMPAGNES AGRICOLES 1975-1983

Le nombre appréciable de périmètres à hauts rendements, répartis sur l'ensemble du bassin du fleuve Sénégal, confirme la viabilité de la généralisation de la culture irriguée en maîtrise de l'eau. En effet, les rendements ne sont pas mauvais lorsqu'une campagne est menée jusqu'au bout dans des conditions normales, ils peuvent être excellents et même exceptionnels. Il n'y a donc pas de problème majeur d'ordre technologique ou lié à l'acquisition de celle-ci par les paysans. Cependant, l'irrégularités des performances (due à des taux élevés de superficies non exploitées et sinistrées ainsi qu'à des écarts de rendements

~~quelquefois importants constitue une menace sérieuse sur la~~
viabilité économique du programme de mise en place des
périmètres irrigués et donc sur les possibilités réelles de
développement à grande échelle.

Les résultats obtenus pendant les différentes campagnes
met en évidence :

- des déficiences graves dans la conduite de l'eau. Ces déficiences sont imputables à une maîtrise imparfaite, soit interne, soit externe de l'eau au niveau des parcelles et proviennent de défauts d'aménagement (mauvais planage, mauvais calage de certains ouvrages, réseaux défectueux) ou des problèmes au niveau du pompage (retards de livraison des stations, mauvais entretien, panne fréquentes, pompes pas assez puissantes, tours d'eau mal organisés)
- des défaillances des prestations assurées par les sociétés nationales de développement. C'est ainsi que les façons culturelles en grande mécanisation, les problèmes d'approvisionnement en carburant, en lubrifiant et en engrais, le manque de pièces détachées pour la maintenance du matériel agricole et des stations de pompage ont constitué autant de problèmes qui ont marqué les différentes campagnes
- les risques d'extension des infestations : enherbement, riz sauvages, attaques d'insectes
- les difficultés rencontrées par la culture de contre-saison. Là où elle est possible, la culture de contre-saison se heurte souvent à des difficultés liées au chevauchement entre la fin de la campagne d'hivernage (quelquefois tardive) et le début de contre-saison. Dans ces conditions, les aléas du respect de calendrier cultural sont tels que la double culture n'est tentée que sur

~~une faible partie des superficies cultivables. Les~~
goulets d'étranglement se situent essentiellement au niveau de la conduite des façons culturales et de la récolte (battage, surtout), l'insuffisance et les problèmes de maintenance du matériel agricole étant, généralement, les facteurs limitants.

D'autres problèmes tout aussi importants méritent également une grande attention : coût élevé de l'eau d'irrigation qui grève lourdement les charges d'exploitation (riz), renouvellement des groupes motopompes par les communautés paysannes, endettement des paysans.

Une formation adéquate des exploitants ainsi qu'une dévolution progressive aux organisations paysannes de certaines fonctions assurées par les sociétés de développement devraient permettre de dominer en partie les difficultés évoquées.

ORGANISATION POUR LA MISE EN VALEUR
DU FLEUVE SENEgal (OMVS)

HAUT - COMMISSARIAT

CELLULE D'EVALUATION ET DE
PLANIFICATION CONTINUE

INTRODUCTION DE LA BANQUE DE DONNEES

ASSANE DIOP

NOVEMBRE 1984