

BURKINA FASO

-----

MINISTERE DE L'AGRICULTURE  
ET DES RESSOURCES ANIMALES

-----

SECRETARIAT GENERAL

-----

DIRECTION DES STATISTIQUES  
AGRO-PASTORALES

METHODOLOGIE  
DE  
L'ENQUETE NATIONALE  
SUR LES  
STATISTIQUES AGRICOLES

VERSION  
DECEMBRE 93

## **SOMMAIRE**

*I°) INTRODUCTION*

*II°) ELABORATION DE LA BASE DE SONDAGE AU PREMIER DEGRE*

*III°) ELABORATION DE LA BASE DE SONDAGE AU DEUXIEME DEGRE*

*IV°) TAILLE DE L'ECHANTILLON*

*V°) TIRAGES DES ECHANTILLONS*

*VI°) FORMULES DE CALCUL DES ESTIMATEURS*

## **I °) INTRODUCTION**

*L'enquête nationale sur les statistiques agricoles (ENSA) est une enquête par sondage à deux degrés avec*

*stratification à chaque degré. Les villages administratifs et les secteurs des localités semi-urbaines sont les unités du premier degré. Au second degré on s'intéresse aux ménages agricoles à l'intérieur des villages administratifs ou des secteurs.*

*L'unité d'enquête choisie est le ménage agricole et non le ménage rural.*

*Le ménage agricole est un ménage où un ou plusieurs membres exploitent des parcelles pour le compte du ménage.*

*Le ménage est défini comme étant un groupe de personnes généralement unies par des liens de sang ou de mariage, logeant habituellement ensemble, produisant ou consommant ensemble, et dont l'autorité budgétaire relève au moins théoriquement d'une seule personne appelée chef de ménage.*

*L'enquête couvre l'ensemble du territoire national.*

*Elle se déroule en deux phases.*

*La première phase, d'une durée de trois mois de Mai à Juillet, permet de saisir par interview les informations sur les structures de l'agriculture. Il s'agit notamment de données sur la population agricole, le matériel agricole, le cheptel des ménages agricoles, l'utilisation de la production.*

*La deuxième phase, d'une durée de six mois d'Août à Janvier concerne un sous échantillon de l'échantillon de la première phase. Elle porte sur la collecte des statistiques liées à la production et au stockage, telles que les superficies par spéculation, les rendements par spéculation, les stocks paysans, les prévisions de récoltes, la main d'oeuvre etc...*

Le domaine d'enquête est la province. L'échantillonnage s'effectue donc au niveau de chacune des 30 provinces. Par voie de conséquence le niveau le plus fin de signification des résultats est donc la province.

## **II°) ELABORATION DE LA BASE DE**

### **SONDAGE AU PREMIER DEGRE**

#### **1°) Mise à jour de la liste exhaustive des villages et secteurs**

Dans chaque province, la base de sondage au premier degré est constituée par la liste exhaustive des villages administratifs et des secteurs des localités semi-urbaines, obtenue à partir du recensement général de la population de 1985. Les localités suivantes sont considérées comme urbaines et écartées de la base de sondage: Tenkodogo, Koudougou, Banfora, Fada N'Gourma, Bobo-Dioulasso, Ouagadougou, Dédougou, Yako, Gaoua, Kaya, Dori, Ouahigouya.

Toutefois une mise à jour de cette liste a été nécessaire. En effet beaucoup de quartiers ont été érigés en villages, certains villages ont été intégrés à des communes. Les tournées préalables dans les différentes localités ont été organisées pour obtenir auprès des préfets, responsables administratifs et responsables coutumiers les précisions sur les limites des villages ou secteurs telles qu'observées lors du recensement de 1985. Pour chaque village ou secteur, ces précisions ont été fournies aux enquêteurs.

#### **2°) Stratification des villages**

La méthode de classification automatique a été utilisée pour répartir les villages en 2 strates. Le critère de stratification est basée sur la dominance ou non des grandes

exploitations dans le village ou regard des variables niveau d'équipement, dimension moyenne des exploitations etc... . Aussi utilise-t-on les moyennes par exploitation de certaines caractéristiques comme variables d'identification du type de village.

Les variables ci-dessous ont été utilisées à cet effet. Elles ont été observées sur l'échantillon de l'enquête permanente agricole ( EPA ) et de l'enquête sur l'impact de la vulgarisation :

- Superficie cultivée en céréales,
- Superficie cultivée en culture de rente.
- Production en cultures céréalières,
- Production en cultures de rente.
- Quantité de céréales en stock en fin de campagne  
(Septembre-Octobre)
- Nombre de personnes,
- Nombre d'actifs,
- Nombre de d'abas,
- Nombre de charrues.
- Quantités de Burkina Phosphate utilisée,
- Quantité de semences sélectionnée utilisée,
  - Quantité de semence ordinaire utilisée.
- Nombre de bovins,
- Nombre d'ovins,
- Nombre de caprins.

Des données sur ces variables sont disponibles pour 213 villages à partir des enquêtes citées. Elles expriment des moyennes par ménages pour chaque village (six ménages environ par village).

Deux strates ont pu ainsi être identifiées. La 1ère comprend les villages où les ménages ont un grand potentiel de production , la deuxième ceux où les ménages ont un plus faible potentiel de production.

58,7 % des villages analysés sont ainsi de la strate 1. Du point de vue de la production, les ménages de ces villages cultivent en moyenne 3,59 ha et 0,80 ha de céréales et de culture de rente contre respectivement 4,75 et 0,78 ha de moyenne pour les villages de la strate 2.

Avec une superficie cultivée inférieure, les ménages dans la strate 1 ont en moyenne une production plus élevée en céréales (de 14%) et en produits de rente (de 115%).

Le stock moyen par ménages pour les villages de la strate 1 est presque le double de celui des ménages des villages de la strate 2.

Au plan effectif des personnes, on retient que la taille moyenne des ménages est de 8,40 personnes pour les villages de la strate 1 et de 12,08 pour ceux de la strate 2. La proportion d'actif est sensiblement la même dans les deux strates (45,6% contre 50,7%). Au niveau de l'équipement et des intrants agricoles, la strate 1 apparaît comme plus grande utilisatrice de moyens de production que la strate 2.

Les critères de stratification ont été transposés sur la liste intégrale du recensement générale de la population à l'aide d'une fonction discriminante.

La fonction discriminante retenue permet de classer correctement 72,35% des villages de l'échantillon de l'enquête agricole.

$$\begin{aligned} D = & - 8,8673 \times \text{nombre de ménage} \\ & - 3,0000 \times \text{population masculine} \\ & + 8,4929 \times \text{population féminine} \\ & - 2,1863 \times \text{population active} \\ & - 0,5476 \times \text{population agricole} \\ & - 300 \end{aligned}$$

\* Si  $D$  est positif l'élément est de la strate 1

\* Il est de la strate 2 dans le cas contraire

*La liste des villages a été stratifié sur cette base.*

### **III<sup>o</sup>) ELABORATION DE LA BASE DE SONDAGE AU DEUXIEME DEGRE**

#### **1<sup>o</sup>) Liste exhaustive des ménages des villages échantillons**

*Lors du premier passage de l'enquête, les villages échantillons sont recensés pour dégager une liste exhaustive des ménages par village avec pour chaque ménage les caractéristiques suivantes :*

- le caractère agricole ou non du ménage*
- la population du ménage*
- le nombre d'actifs agricoles du ménage*
- le nombre de charrues que possède le ménage.*

*Pour assurer l'exhaustivité, les ménages sont recensés par concessions lesquelles sont visitées et numérotées une à une en commençant par la concession du chef de village (ou responsable administratif) et en suivant un ordre donné.*

*La base de sondage au second degré est la liste exhaustive des ménages agricoles dressée par l'enquêteur (après élimination des ménages non agricoles).*

#### **2<sup>o</sup>) Stratification des ménages**

*La stratification des ménages a été faite suivant les*

mêmes techniques que pour les villages. La seule différence est qu'on utilise l'échantillon des ménages de l'EPA (978 ménages retenus) et non l'échantillon des villages.

Sur la base des 17 variables précédemment utilisées ces ménages sont classés en deux strates . La strate 1 qui représente 25,40 % du total des ménages agricoles est celle des grands ménages agricoles.

En effet les ménages de cette strate cultivent plus de superficie (2 fois plus) produisent plus de céréales (+158%) et de produits de rente (+213%) que ceux de la strate 2. Les ménages de la strate 1 ont des réserves en céréales beaucoup plus importantes (51,81 tonnes) contre (0,36 tonnes) pour ceux de la strate 2. La taille moyenne des ménages de la strate 1 est de 19,16 et celle de la strate 2 est de 8,47 (10,27 contre 4,14 pour la taille moyenne des actifs).

Les variables utilisées pour la détermination des fonctions discriminantes sont celles collectées facilement lors de l'établissement de la base de sondage au second degré.

Plusieurs variables notamment ceux ci-dessus cités (II°,2°) ont été utilisées pour établir des fonctions de discrimination. Selon que les provinces soit bien arrosées ou pas ces variables n'y ont pas le même pouvoir discriminant.

**Aussi le pays est divisé en deux zones :**

**Zone 1 : zone la plus arrosée.**

07 NAHOURI	18 BOULGOU
08 ZOUNDWEOGO	26 COMOË
11 SISSILI	27 KENEDOUGOU
12 MOUHOUN	28 HOUËT
13 KOSSI	29 PONI
15 GOURMA	30 BOUGOURIBA
17 TAPOA	



## **Zone 2 : zone la moins arrosée**

01 KADIOGO	16 GNAGNA
02 OUBRITENGA	19 KOURITENGA
03 GANZOURGOU	20 SENO
04 YATENGA	21 OUDALAN
05 PASSORE	22 SOUM
06 BAZEGA	23 BAM
09 BOULKIEMDE	24 SANMATENGA
10 SANGUIE	25 NAMENTENGA
14 SOUROU	

On dispose de données sur 480 ménages pour la zone 1 et 498 ménages pour la zone 2.

Le choix de la fonction tient compte du pourcentage de ménages correctement classés mais aussi de la simplicité des variables utilisées puisque cette stratification se fait sur le terrain par les contrôleurs.

### **Première zone**

La fonction discriminante  $D$  retenue combine les variables nombre de charrues possédées par le ménage et nombre d'actifs du ménage.

$$D = \begin{aligned} & 59 \times \text{nombre de charrues} \\ & + 30 \times \text{nombre d'actifs} \\ & - 172 \end{aligned}$$

Si la valeur de  $D$  est positive, le ménage est à classer dans la strate 1 sinon il appartient à la strate 2.

Cette fonction permet de classer correctement 84,13% des 480 ménages de la zone 1.

Pour cette zone une grille permettant de lire directement la strate du ménage à partir du nombre d'actifs et du nombre de charrues a été établie à l'intention des contrôleurs.

## **Deuxième zone**

A l'aide de la variable "nombre d'actifs agricoles", on classe correctement 85,77 % des ménages. Cette variable est retenue comme variable de discrimination :

- Si le nombre d'actifs est supérieur à quatre (4), le ménage est de la strate 1
- Si le nombre d'actifs est inférieure ou égale à quatre (4) , le ménage est de la strate 2.

## **IV°) TIRAGES DES ECHANTILLONS**

### **1°) Première phase**

#### **a°) Premier degré**

L'ensemble des villages de la province est divisé en deux strates.

Le tirage est à probabilités inégales, proportionnelles à la population de chaque village dans chaque strate. Le nombre de villages échantillons par strate est proportionnel à la population totale de la strate.

Le tirage est systématique et s'effectue après classement des villages par département et par ordre alphabétique.

NB Pour les unités primaires très grosses, le programme de tirage permet de les tirer plusieurs fois de façon à ce qu'ils aient un nombre plus élevé d'unités secondaires échantillons.

*Le nombre de ménages échantillons à retenir pour un village qui est tiré une fois étant de huit celui d'un village qui est tiré  $n$  fois est de  $n$  multiplié par huit.*

*Le nombre de fois qu'un village peut être tiré a été toutefois limité à quatre aussi pour une taille d'échantillon initial de 2 200, on se retrouve avec 2 077 villages distincts.*

### **b°) Deuxième degré**

*Tous les ménages agricoles recensés dans les villages échantillons sont classés en deux strates. Le tirage est aléatoire et à probabilité égale pour tous les ménages agricoles.*

*Le nombre de ménages à retenir par strates est proportionnel au nombre total de ménages agricoles de la strate.*

#### **Remarque**

*De fait on devrait donc avoir une probabilité égale de tirage quelle que soit la strate. Etant donné les arrondis sur le nombre de ménages à retenir par strate cette égalité n'est plus vérifiée. Les coefficients d'extrapolation sont donc calculés par strate à partir des nombres entiers des ménages.*

### **2°) Deuxième phase**

#### **a°) Premier degré**

*A partir de la liste des villages retenus pour la première phase, on effectue un tirage systématique à probabilités égales pour retenir un sous-échantillon des villages.*

#### **b°) Deuxième degré**

Tous les ménages échantillons des villages retenus pour la deuxième phase restent échantillons.

## **V°) TAILLE DE L'ECHANTILLON**

### **1°) Première phase**

Cette phase permettra de recenser des variables aussi variées et importantes les unes que les autres. L'échantillon est réparti entre les différentes provinces proportionnellement à leur population agricole. Le nombre d'enquêteurs initialement retenus est de 560. Dans certains CRPA d'anciens enquêteurs sont retenus comme enquêteurs pour la présente enquête (1 au Houët, 1 au Kéné Dougou, 1 à la Bougouriba, 3 à la Comoé, 1 au Poni). Aussi il faut noter que deux (2) villages échantillons du Boulgou ont disparu. Deux villages additionnels ont donc été tirés dans cette province.

Chaque enquêteur pouvant avoir à charge que quatre villages, le nombre de villages échantillons tirés est de 2270 (2077 villages distincts) soit un total de 18160 ménages échantillons.

#### **Répartition de l'échantillon village première phase** **par province**

PROVINCE	NOMBRE DE VILLAGES ECHANTILLONS	NOMBRE D'ENQUETEURS
BAM	52	13
BAZEGA	96	24
BOUGOURIBA	76	19
BOULGOU	122	30
BOULKIEMDE	100	25
COMOE	70	20
GANZOURGOU	64	16
GNAGNA	72	18
GOURMA	88	22
HOUET	116	29
KADIOGO	8	2
KENEDOUGOU	48	12
KOSSI	108	27

KOURITENGA	64	16
MOUHOUN	84	21
NAHOURI	32	8
NAMENTENGA	64	16
OUBRITENGA	96	24
OULDALAN	32	8
PASSORE	68	17
PONI	76	19
SANGUIE	68	17
SANMATENGA	108	27
SENO	68	17
SISSILI	80	20
SOUM	60	15
SOUROU	80	20
TAPOA	52	13
YATENGA	120	40
ZOUNDWEOGO	48	12

## 2°) Deuxième phase

Pour cette phase chaque enquêteur a en charge trois (3) villages échantillons. Deux des trois villages pris aléatoirement sont concernés par la pose des carrés de rendement.

La taille de l'échantillon par province est fixée de façon à permettre d'obtenir un coefficient de variation de 8% pour le total des superficies cultivées en céréales dans chaque province.

Les calculs partent des coefficients de variation calculés à partir des données de l'enquête permanente (EPA) de 1990-1991 dont le plan de sondage au niveau province est à deux degrés sans stratification.

Pour chacune des provinces la taille à retenir  $m'$  est déterminée à partir des relations suivantes.

Soit  $Y = 1/m \times \sum Y_i$  l'estimateur de la  
 $i$  superficie totale en  
céréales dans la province

Les superficies extrapolées par villages ( $Y_i$ ) étant des estimateurs sans biais de la superficie totale, on estime la variance de  $Y$  ( $v(Y)$ ) par  $S^2/m$ .  $S^2$  étant la quasi-variance de  $Y_i$  et  $m$  le nombre de villages échantillons.

$$v ( Y ) = S^2 / m$$

Pour un autre échantillon de taille  $m'$  on estime la variance de  $Y$   $v' ( Y )$  par  $S'^2 / m'$ .

On en déduit  $S^2 = v ( Y ) \times m = v' ( Y ) \times m'$

$S^2$  ne dépend que de l'univers.

$$\frac{v ( Y ) \times m}{Y^2} = \frac{v' ( Y ) \times m'}{Y'^2}$$

Soit  $CV$  le coefficient de variation de  $Y$  dans l'EAP 1990-91

$CV'$  le coefficient de variation de  $Y'$  souhaité pour  $Y$

$$CV^2 \times m = CV'^2 \times m'$$

C'est à dire 
$$m' = \frac{CV^2 \times m}{CV'^2}$$

On se donne  $CV' = 0,08$  quelle que soit la province

les tailles obtenues à partir de ces calculs ont subi des réajustements. On aboutit pour l'ensemble du pays à un total de 1097 villages tirés (soit 1001 villages distincts).

### **CRITERES DE REAJUSTEMENT**

- La taille minimale par province est de 30 villages, soit 240 ménages.
- La taille maximale est celle observée lors de la première phase.
- Malgré l'inexistence de données fiables sur la variabilité de la variable rendement, on tient compte de cette variable pour les réajustements en fonction de la connaissance qu'on a du terrain.



Répartition de l'échantillon village  
deuxième phase par province

PROVINCE	CV <sup>2</sup>	m	m'	NOMBRE ENQUETEURS
BAM	90,3	6	30	10
BAZEGA	240,3	12	45	15
BOUGOURIBA	49,0	9	30	10
BOULGOU	269,0	15	54	18
BOULKIEMDE	51,8	16	30	10
COMOE	153,8	24	42	14
GANZOURGOU	262,4	13	42	14
GNAGNA	114,5	12	30	10
GOURMA	306,3	12	48	16
HOUET	123,2	27	51	17
KADIOGO	10,9	3	8	4
KENEDOUGOU	108,2	12	30	10
KOSSI	278,9	20	60	20
KOURITENGA	151,3	6	30	10
MOUHOUN	100,0	12	30	10
NAHOURI	158,8	6	30	10
NAMENTENGA	118,8	8	30	10
OUBRITENGA	116,6	29	48	16
OULDALAN	441,0	8	30	10
PASSORE	46,2	9	30	10
PONI	256,0	9	36	12
SANGUIE	384,2	12	42	14
SANMATENGA	219,0	15	48	16
SENO	166,4	16	39	13
SISSILI	77,4	20	30	10
SOUM	201,6	16	42	14
SOUROU	51,8	12	30	10
TAPOA	62,4	12	30	10
YATENGA	68,9	21	42	14
ZOUNDWEOGO	894,0	6	30	10
Total			1097	367

$m'$  = Nombre de villages échantillons

$m$  = Nombre de villages échantillons EPA 90/91

$CV^2$  = Carré du coefficient de variation de la superficie de l'ensemble des cultures cérésières (EPA 90/91).

Remarque:

Il est à noter que les calculs se font avec l'hypothèse que l'on adopte le plan de sondage de l'EPA. La différence entre les deux plans de sondage se situent au niveau de la stratification aux deux degrés de sondage. L'on doit s'attendre à des coefficients de variation nettement inférieur à 8% par



effet de stratification.

## V°) FORMULES DE CALCUL DES RESULTATS

### 1°) Formules d'estimation et coefficients d'extrapolation pour le calcul des totaux

Ces coefficients d'extrapolation servent pour le calcul des totaux de toutes les variables observées au niveau de l'échantillon des ménages.

#### **Notation**

Soit :

##### **Au niveau provincial**

- $h$  indice de la strate du village ( $h=1$  ou  $h=2$ )
- $y$  variable observée au niveau des ménages échantillons
- $Ph$  population totale de la strate  $h$
- $Mh$  nombre total de villages de la strate  $h$
- $mh$  nombre de villages échantillons de la strate  $h$
- $i$  indice du village.

##### **Au niveau village**

- $k$  indice de la strate du ménage ( $k=1$  ou  $k=2$ )
- $Nhki$  nombre total de ménages de la strate  $n^{\circ}k$  du village  $n^{\circ}i$  de la strate  $n^{\circ}h$ .
- $nhki$  nombre total de ménages échantillons de la strate  $n^{\circ}k$  du village  $n^{\circ}i$  de la strate  $n^{\circ}h$ .
- $Phi$  population du village  $i$  dans la strate  $h$ .
- $j$  indice du ménage.

##### **Au niveau ménage**

- $Yhkij$  valeur de  $Y$  pour le ménage  $n^{\circ}j$  de la strate  $k$  appartenant au village  $n^{\circ}i$  de strate  $n^{\circ}h$ .

Soit :

- $Y1$  l'estimateur du total de  $y$  pour la strate  $h = 1$
- $Y2$  l'estimateur du total de  $y$  pour la strate  $h = 2$

- Y l'estimateur du total de y pour toute la province

$$\text{On a : } Y = Y_1 + Y_2$$

$$Y_h = \frac{1}{m_h} \times S_i \frac{P_h}{P_{hi}} \times S_k \frac{N_{hki}}{n_{hki}} S_j Y_{hkij}$$

On en déduit la formule suivante :

$$Y = S_h \frac{1}{m_h} \times S_i \frac{P_h}{P_{hi}} \times S_k \frac{N_{hki}}{n_{hki}} S_j Y_{hkij}$$

Ainsi, pour tout village i de la strate village h, nous avons deux coefficients d'extrapolation (Chi1 et Chi2)

- Pour k= 1 ,c'est à dire pour les ménages de la strate 1

$$Chi1 = \frac{1}{m_h} \times \frac{P_h}{P_{hi}} \frac{N_{h1i}}{n_{h1i}} \quad h = 1,2$$

- Pour k= 2 ,c'est à dire pour les ménages de la strate 2

$$Chi2 = \frac{1}{m_h} \times \frac{P_h}{P_{hi}} \frac{N_{h2i}}{n_{h2i}} \quad h = 1,2$$

## 2°) Formule d'estimation des moyennes.

### a°) Cas général

La moyenne des variables s'obtient en faisant le rapport de l'estimateur du total de la variable par l'estimateur du total des unités pour lesquelles la moyenne est calculée.

Par exemple pour calculer le nombre moyen de charrue par ménage agricole, on estimera le nombre total de charrues pour la province et le nombre total de ménages agricoles de la province. La moyenne est égale au nombre total de charrues divisé par le nombre total de ménages agricoles.

***b°) Cas particulier des rendements par spéculations***

*Au niveau de chaque province, on estimera le rendement de chacune des spéculations en faisant une moyenne simple des différents rendements obtenus sur les parcelles portant ces spéculations.*