

REPUBLIQUE TOGOLAISE
RECENSEMENT NATIONAL DE L'AGRICULTURE
2012

PLAN DE SONDAGE

par
Julien Amegandjin, consultant

Sommaire

1. Introduction	2
2. Bases de sondage	2
3. Domaines d'étude et strates	3
4. Taille de l'échantillon des ménages agricoles	3
5. Répartition des échantillons des unités primaires et des ménages agricoles	4
6. Modes de tirage	4
7. Mise en œuvre du tirage des unités primaires	5
8. Mise en œuvre du tirage des ménages agricoles	7
9. Probabilités et poids de sondage des unités de sondage	11
10. Mise en œuvre du calcul des poids de sondage	13
11. Estimation d'un total	16
12. Estimation d'un effectif, d'une moyenne ou d'une proportion	16
13. Mise en œuvre des estimations	17

Le présent plan de sondage porte sur :

- les objectifs du recensement ;
- les bases de sondage ;
- les domaines d'étude et les strates ;
- la taille de l'échantillon des ménages agricoles ;
- la répartition des échantillons ;
- les méthodes de tirage ;
- la mise en œuvre des tirages ;
- les probabilités d'inclusion et les poids de sondage des unités de sondage ;
- les méthodes d'estimation.

1. Introduction

Le recensement national de l'agriculture qui va être réalisé en 2012 vient 16 années après l'édition précédente réalisée en 1996. Il a pour objectif de contribuer à la réduction de la pauvreté et de l'insécurité alimentaire grâce à des stratégies et politiques plus efficaces définies et mises en œuvre sur la base d'informations statistiques structurelles fiables et pertinentes. Il a également pour objectif d'asseoir un système d'information permanent capable de stimuler la croissance du secteur agricole et de dynamiser les programmes de développement.

Les données du recensement national de l'agriculture de 2012 serviront tous les acteurs, pouvoirs publics, organisations de producteurs et de productrices, société civile, secteur privé et partenaires au développement dans la formulation, le suivi et l'évaluation des politiques et stratégies en faveur du monde rural.

La dernière édition du recensement national de l'agriculture ayant eu lieu en 1996, la nouvelle édition de cette opération de collecte de données de grande envergure était très attendue, tant les données agricoles de structure fiables sur le monde rural togolais faisaient cruellement défaut.

Le recensement national de l'agriculture de 2012 a été précédé d'un pré recensement de l'agriculture réalisé en 2011 dans le but de constituer pour, le recensement de l'agriculture, la base de sondage des ZD agricoles devant servir d'unités primaires ainsi que la base de sondage des ménages agricoles ou unités secondaires.

Le recensement national de l'agriculture de 2012 est réalisé par sondage aléatoire au moyen d'un gros échantillon de ménages agricoles.

2. Bases de sondage

Le plan de sondage du recensement repose sur un sondage aréolaire à deux degrés avec stratification au premier degré. Les unités statistiques du premier degré ou unités primaires sont les zones de dénombrement (ZD) agricoles définies lors du pré recensement agricole de 2011. Les unités statistiques du deuxième degré ou unités secondaires sont les ménages agricoles.

Les opérations de collecte du pré recensement ont utilisé les résultats des travaux cartographiques du dernier Recensement général de la population et de l'habitat (RGPH). Elles ont consisté à dénombrer les ménages agricoles dans les ZD rurales telles qu'elles sont définies lors des travaux cartographiques du RGPH. Elles ont également consisté à déterminer

la localisation géographique de chacun des ménages agricoles dénombrés ainsi que l'effectif des ménages agricoles pour chaque ZD agricole.

Les travaux du Pré Recensement Agricole de 2011 ont permis d'établir la liste des ZD agricoles du pays pour chaque région, chaque préfecture et chaque canton. Cette liste des ZD agricoles constitue la base de sondage du premier degré pour le recensement. De plus, pour chaque ZD agricole, la base de sondage comporte l'effectif des ménages agricoles, une variable qui servira d'information auxiliaire dans le tirage de l'échantillon du premier degré.

En outre, le Pré Recensement Agricole a permis d'établir pour chaque ZD agricole du pays, la liste des ménages agricoles ainsi que la localisation et l'identité de chacun d'eux. Cette deuxième catégorie de données définit la base de sondage du deuxième degré du recensement. Plus précisément, la base de sondage du 2^e degré est constituée par la liste des ménages agricoles des ZD agricoles tirées au 1^{er} degré.

3. Domaines d'étude et strates

Un domaine d'étude est une partie de l'univers sondé pour laquelle sont recherchés des résultats significatifs, c'est-à-dire des estimations séparées et d'une précision suffisante. Dans le cas du recensement national de l'agriculture de 2012, chacune des 5 régions qui comportent un milieu rural est considéré comme un domaine d'étude. La commune de Lomé qui ne compte pas de milieu rural ne fait pas partie du champ du recensement malgré l'existence d'activités de production horticole (légumes et fleurs) en son sein.

Il est fait appel à une stratification aréolaire reposant sur le découpage des régions en préfectures. Chaque préfecture est considérée comme une strate et le plan de sondage porte sur un total de 35 strates.

Le tableau 1 présente les domaines d'étude et les strates ainsi que la structure de la base de sondage des ZD et des échantillons des 2 degrés.

4. Taille de l'échantillon des ménages agricoles

La Direction du RNA de 2012 a retenu très tôt d'avoir pour les échantillons des deux degrés de sondage, des tailles sensiblement semblables à celles du dernier recensement national agricole de 1996. Ce dernier avait porté sur un échantillon de 1800 ZD agricoles et un échantillon de 7200 ménages agricoles.

Toutefois, pour l'édition de 2012, il a été retenu de réduire quelque peu les déplacements en réduisant à 1500 la taille de l'échantillon des unités primaires tirées et en augmentant le nombre de ménages agricoles à observer par ZD agricole de 4 à 6. Cela a donc conduit au choix d'un échantillon de 9000 ménages agricoles pour l'ensemble des cinq régions du pays.

Les contraintes budgétaires semblent avoir constitué l'unique élément considéré dans la décision au sujet du choix des tailles des échantillons. Aucune considération de précision ne semble avoir été prise en compte. Une exigence de précision n'aurait d'ailleurs pas été aisée à exprimer, car les résultats du RNA de 1996 ne comportent pas de publication d'erreurs de sondage pas plus que ceux des enquêtes agricoles annuelles.

Cette situation implique la recommandation que les résultats du nouveau recensement comprennent la précision des estimations pour les principales variables d'étude.

5. Répartition des échantillons des unités primaires et des ménages agricoles

5.1 Répartition de l'échantillon des unités primaires entre les domaines d'étude

Au total, 1500 ZD agricoles seront donc tirées mais pas de façon uniforme entre les 5 domaines d'étude. On estime qu'il faut un échantillon d'au moins 250 ménages agricoles dans un domaine d'étude pour obtenir des estimations d'une bonne précision. Une répartition proportionnelle de l'échantillon des ZD agricoles entre les 5 régions ne remplirait pas cette condition. Il s'ensuit qu'un sur échantillonnage au profit de la région Centrale, de la région de la Kara et de la région des Savanes est nécessaire et cette opération se fait en sous échantillonnant la région des Plateaux. Le tableau 1 présente la répartition de l'échantillon des ZD agricoles selon le domaine d'étude.

5.2 Répartition des échantillons des unités primaires entre les strates d'un domaine d'étude

Pour chaque domaine d'étude, il est fait appel à un sondage stratifié représentatif au premier degré. Autrement dit, le sous échantillon des unités primaires ou ZD agricoles de chacun des domaines d'étude est réparti entre les strates proportionnellement à la taille des strates. Il n'a pas été possible de faire appel à l'une des répartitions optimales classiques faute de disposer de données sur la dispersion des variables au niveau des strates.

La répartition des sous échantillons des unités primaires selon la strate au sein des domaines d'étude est présentée dans le tableau 1.

5.3 Répartition de l'échantillon des ménages entre les strates d'un domaine d'étude

Au deuxième degré, 6 ménages agricoles sont prélevés dans chacune des unités primaires ou ZD agricoles échantillonnées, ce qui conduit à un échantillon global de 9000 ménages agricoles.

La répartition de l'échantillon des 9000 ménages entre les strates ainsi que le taux de sondage global sont présentés dans le tableau 1. L'on notera que ce taux est bien évidemment uniforme au sein de chaque domaine d'étude du fait que l'on a utilisé un sondage stratifié représentatif.

Le taux global de sondage, celui des deux degrés réunis, est de 0,018 et correspond à l'observation de 18 ménages agricoles en moyenne sur 1000.

6. Modes de tirage

Les tirages des échantillons sont mis en œuvre de façon indépendante d'une strate à l'autre. Les unités primaires ou ZD agricoles sont tirées selon le mode de tirage systématique avec probabilités proportionnelles aux tailles des unités, un cas particulier de tirage avec probabilités inégales et sans remise. La probabilité de sélection d'une ZD à chaque tirage est choisie proportionnelle à la taille de la ZD, la taille étant définie ici par l'effectif des ménages agricoles de la ZD agricole.

Pour les tirages du deuxième degré, c'est-à-dire les tirages des unités secondaires (ou ménages agricoles), il est fait appel à un sondage aléatoire simple sans remise. Les ménages agricoles sont donc tirés avec probabilités égales et sans remise.

Un nombre constant de ménages agricoles est sélectionné dans chaque ZD agricole échantillonnée, soit 6 ménages agricoles.

7. Mise en œuvre du tirage des unités primaires

Le tirage des unités primaires est réalisé de façon indépendante d'une strate à l'autre. Comme cela est indiqué plus haut, les unités primaires ou ZD agricoles sont tirées selon le mode de tirage systématique avec probabilités proportionnelles aux tailles des unités (tirage SPPT).

Le tirage des unités primaires a été mis en œuvre à l'aide du logiciel TIRAGE 2.1, un logiciel destiné à la réalisation de tirages aléatoires. En organisant les tirages, il a fallu vérifier au départ pour chacune des 35 strates si elle comporte ou non des ZD agricoles atypiques. Pour ce faire, les probabilités d'inclusion d'ordre 1 ont été calculées pour chaque ZD de chaque strate. Elles satisfont toutes la condition

$$0 < \pi_r < 1 \quad \forall r = 1, 2, \dots, M$$

pour les M unités primaires de chaque strate, à l'exception de 14 ZD présentées dans le tableau 2 ci-après.

Selon ce tableau, neuf strates possèdent une ou deux ZD atypiques, c'est-à-dire des ZD de probabilités d'inclusion d'ordre 1 supérieures à 1. Les 14 ZD concernées seront tirées d'office avec une probabilité égale à 1.

Le tableau 3 présente pour les 9 strates ou préfectures concernées l'effectif des ZD tirées d'office ainsi que l'effectif des ZD restant à tirer.

Le tirage se fait en utilisant le 1^{er} module du 3^e groupe de modules du menu "Tirages dans un sondage à 2 degrés". Ce module est dénommé "Tirage SPPT de l'échantillon du 1^{er} degré". Le tirage est réalisé strate par strate et les résultats du tirage pour toutes les strates figurent dans la table dénommée "TableSPPT_1^{er} Degré" de Data.mdb.

A l'issue des tirages pour les 35 strates, la liste de 1486 ZD tirées au moyen de TIRAGE 2.1 se trouve dans la table "TableSPPT_1^{er} Degré" de Data.mdb. Cette liste est complétée par l'ajout des 14 ZD atypiques du tableau 2.

De la table "TableSPPT_1^{er} Degré", on peut déduire une liste plus documentée des 1500 ZD tirées en y incluant les principales informations de la base de sondage concernant ces ZD. Chacune des ZD ou unités primaires de cette liste est repérée par un numéro d'ordre allant de 1 à 1500 et par le champ ID_ZD qui est obtenue par concaténation des codes de la région, de la préfecture, du canton et de la ZD agricole. Le tableau 4 présente un extrait de la liste documentée des ZD composant l'échantillon du 1^{er} degré du sondage.

Tableau 1 : Structure de la base de sondage des ZD agricoles et des échantillons des deux degrés

Nom de la région	Nom de la préfecture	Effectif des ZD	%	Effectif des ménages agricoles	%	Répartition de l'échantillon des 1500 ZD		Proportion des ZD à tirer par préfecture	Effectif des ZD à tirer par préfecture	Effectif corrigé	Effectif des MA à tirer par préfecture	Taux global de sondage
						proportionnelle	corrigée					
Maritime	AVE	117		14 584				0,14	44,4	44	264	0,018
	BAS-MONO	105		12 585				0,12	38,3	38	228	0,018
	LACS	161		13 025				0,12	39,6	40	240	0,018
	VO	221		21 702				0,20	66,0	66	396	0,018
	YOTO	219		18 981				0,18	57,8	58	348	0,018
	ZIO	300		25 914				0,24	78,9	79	474	0,018
	Sous Total	1 123	0,26	106 791	0,22	325	325	1,00	325,0	325	1 950	0,018
Plateaux	AGOU	102		13 578				0,08	33,0	33	198	0,015
	AKEBOU	73		8 704				0,05	21,1	21	126	0,014
	AMOU	110		16 387				0,10	39,8	40	240	0,015
	ANIE	53		9 118				0,06	22,1	22	132	0,014
	DANYI	54		7 129				0,04	17,3	17	102	0,014
	EST-MONO	140		16 271				0,10	39,5	40	240	0,015
	HAHO	266		27 947				0,17	67,8	68	408	0,015
	KLOTO	74		7 796				0,05	18,9	19	114	0,015
	KPELE-AKATA	87		10 126				0,06	24,6	25	150	0,015
	MOYEN-MONO	102		11 302				0,07	27,4	27	162	0,014
	OGOUE	180		24 451				0,15	59,3	59	354	0,014
	WAWA	89		12 013				0,07	29,2	29	174	0,014
	Sous Total	1 330	0,30	164 822	0,33	502	400	1,00	400,0	400	2 400	0,015
Centrale	BLITTA	153		15 965				0,27	61,3	61	366	0,023
	PLAINE DE MO	37		4 235				0,07	16,3	16	96	0,023
	SOTOUBOUA	103		12 598				0,22	48,4	49	294	0,023
	TCHAMBA	125		14 877				0,25	57,2	57	342	0,023
	TCHAOUDJO	104		10 893				0,19	41,8	42	252	0,023
	Sous Total	522	0,12	58 568	0,12	178	225	1,00	225,0	225	1 350	0,023
Kara	ASSOLI	39		1 696				0,02	5,8	6	36	0,021
	BASSAR	119		13 312				0,16	45,4	45	270	0,020
	BINAH	72		9 896				0,12	33,7	34	204	0,021
	DANKPEN	111		12 899				0,16	44,0	44	264	0,020
	DOUFELGOU	93		11 780				0,15	40,1	40	240	0,020
	KERAN	91		11 122				0,14	37,9	38	228	0,020
	KOZAH	149		19 987				0,25	68,1	68	408	0,020
	Sous Total	674	0,15	80 692	0,16	246	275	1,00	275,0	275	1 650	0,020
Savanes	CINKASSE	49		5 205				0,06	17,5	17	102	0,020
	KPENDJAL	154		16 677				0,20	56,1	56	336	0,020
	OTI	180		20 845				0,25	70,1	70	420	0,020
	TANDJOARE	126		15 023				0,18	50,5	51	306	0,020
	TONE	225		24 006				0,29	80,7	81	486	0,020
	Sous Total	734	0,17	81 756	0,17	249	275	1,00	275,0	275	1 650	0,020
Total		4 383	1,00	492 629	1,00	1 500	1 500	-	1 500	1 500	9 000	0,018

Tableau 2 : Liste des ZD de probabilités d'inclusion supérieures à 1

Num	ID_ZD	Région	Préfecture	Canton	Nb_Magr
1	20303701	PLATEAUX	AMOU	EKPEGNON	487
2	20304101	PLATEAUX	AMOU	EVOU	567
3	20802101	PLATEAUX	KLOTO	ATCHAVE	605
4	30101406	CENTRALE	BLITTA	BLITTA-GARE	284
5	30402408	CENTRALE	TCHAMBA	GOUBI	289
6	30402511	CENTRALE	TCHAMBA	KABOLI	325
7	40500101	KARA	DOUFELGOU	AGBANDE-YAKA	557
8	40501903	KARA	DOUFELGOU	DEFALE	357
9	40701502	KARA	KOZAH	BOHOU	393
10	40704101	KARA	KOZAH	KPINZINDE	434
11	50203102	SAVANES	KPENDJAL	MANDOURI	333
12	50203105	SAVANES	KPENDJAL	MANDOURI	366
13	50300210	SAVANES	OTI	BARKOISSI	456
14	50506806	SAVANES	TONE	TOAGA	357

Tableau 3 : Effectif par strate des ZD tirées d'office et des ZD restant à tirer

Région	Préfecture	Effectif des ZD tirées d'office	ID des ZD tirées d'office		Effectif des ZD restant à tirer
PLATEAUX	AMOU	2	20303701	20304101	38
PLATEAUX	KLOTO	1	20802101		18
CENTRALE	BLITTA	1	30101406		60
CENTRALE	TCHAMBA	2	30402408	30402511	55
KARA	DOUFELGOU	2	40500101	40501903	38
KARA	KOZAH	2	40701502	40704101	66
SAVANES	KPENDJAL	2	50203102	50203105	54
SAVANES	OTI	1	50300210		69
SAVANES	TONE	1	50506806		80

8. Mise en œuvre du tirage des ménages agricoles

L'information dont on a besoin pour le tirage de l'échantillon des ménages est l'effectif N des ménages agricoles pour chacune des 1500 ZD agricoles échantillonnées. Les travaux du Pré Recensement Agricole de 2011 ayant été achevés juste avant le tirage des échantillons des deux degrés, une opération de redénombrement des ménages agricoles dans les ZD agricoles tirées au premier degré n'a pas été nécessaire. L'effectif N des ménages agricoles de chaque ZD agricole figure dans la base de sondage des ZD agricoles établie par le Pré Recensement Agricole.

Si 147 désigne l'effectif N des ménages agricoles de la ZD agricole enquêtée n° 130, par exemple, la base de sondage des ménages agricoles pour cette ZD agricole est la liste des 147 numéros d'ordre qui identifient les 147 ménages agricoles de la ZD.

Le plan de sondage du recensement national agricole prévoit de sélectionner 6 ménages de chaque ZD agricole échantillonnée. Le tirage des 6 ménages agricoles se fait en tirant avec probabilités égales et sans remise, 6 des 147 numéros d'ordre de ménage de la ZD agricole.

Au lieu de procéder aux tirages séparément par ZD agricole, on gagne beaucoup de temps à les effectuer en une seule fois pour l'ensemble des ZD agricoles échantillonnées de toutes les strates. Pour ce faire, on utilisera le 3^e groupe de modules du menu "Tirages dans un sondage à 2 degrés" du logiciel TIRAGE 2.1. Grâce au module "Construction de la table-source du 2^e tirage" de ce 3^e groupe de modules, les valeurs des effectifs N des ménages agricoles des ZD agricoles seront prises en compte ainsi que le nombre constant de ménages agricoles à tirer dans chaque ZD.

Plus précisément, le tableau 5 suivant présente un extrait du fichier Access qui est la source du tirage en une seule fois des 9000 ménages agricoles (ZD de n° 1 à 10 et de 1491 à 1500).

Le tableau 6 montre un extrait (ménages n° 1 à 25) du fichier Excel des 9000 ménages agricoles tirés.

Tableau 4 : Extrait du fichier donnant la liste des ZD tirées

Num_ZD	ID_ZD	Région	Code de la préf.	Nom de la préf.	Rang du tirage	Code du canton	Nom du canton	ZD	Nombre de Menages Agr
1	10101102	MARITIME	1	AVE	1	11	AKEPE	2	140
2	10101104	MARITIME	1	AVE	2	11	AKEPE	4	336
3	10101105	MARITIME	1	AVE	3	11	AKEPE	5	168
4	10101721	MARITIME	1	AVE	4	17	ASSAHOUN	21	176
5	10101724	MARITIME	1	AVE	5	17	ASSAHOUN	24	128
6	10101727	MARITIME	1	AVE	6	17	ASSAHOUN	27	212
7	10101728	MARITIME	1	AVE	7	17	ASSAHOUN	28	189
8	10101801	MARITIME	1	AVE	8	18	ASSAHOUN+ANDOPEME	1	165
9	10101804	MARITIME	1	AVE	9	18	ASSAHOUN+ANDOPEME	4	100
10	10101807	MARITIME	1	AVE	10	18	ASSAHOUN+ANDOPEME	7	57
11	10101811	MARITIME	1	AVE	11	18	ASSAHOUN+ANDOPEME	11	223
12	10101813	MARITIME	1	AVE	12	18	ASSAHOUN+ANDOPEME	13	174
13	10101815	MARITIME	1	AVE	13	18	ASSAHOUN+ANDOPEME	15	172
14	10101817	MARITIME	1	AVE	14	18	ASSAHOUN+ANDOPEME	17	101
15	10101819	MARITIME	1	AVE	15	18	ASSAHOUN+ANDOPEME	19	112
16	10101832	MARITIME	1	AVE	16	18	ASSAHOUN+ANDOPEME	32	96
17	10101835	MARITIME	1	AVE	17	18	ASSAHOUN+ANDOPEME	35	120
18	10101838	MARITIME	1	AVE	18	18	ASSAHOUN+ANDOPEME	38	154
19	10102001	MARITIME	1	AVE	19	20	BADJA	1	48
20	10102004	MARITIME	1	AVE	20	20	BADJA	4	31
21	10102009	MARITIME	1	AVE	21	20	BADJA	9	59
22	10102012	MARITIME	1	AVE	22	20	BADJA	12	277
23	10102015	MARITIME	1	AVE	23	20	BADJA	15	94
24	10102018	MARITIME	1	AVE	24	20	BADJA	18	76
25	10102603	MARITIME	1	AVE	25	26	DZOLO	3	181
26	10102605	MARITIME	1	AVE	26	26	DZOLO	5	184
27	10102607	MARITIME	1	AVE	27	26	DZOLO	7	82
28	10102609	MARITIME	1	AVE	28	26	DZOLO	9	130
29	10102612	MARITIME	1	AVE	29	26	DZOLO	12	190
30	10103902	MARITIME	1	AVE	30	39	KEVE	2	120
31	10103906	MARITIME	1	AVE	31	39	KEVE	6	133

Tableau 5 : Extrait de la table source des tirages du 2^e degré

Num_ZD	Identité_UP	Taille_UP	NombreUS_àTirer
1	10101102	140	6
2	10101104	336	6
3	10101105	168	6
4	10101721	176	6
5	10101724	128	6
6	10101727	212	6
7	10101728	189	6
8	10101801	165	6
9	10101804	100	6
10	10101807	57	6

1 491	50506202	60	6
1 492	50506205	117	6
1 493	50506208	104	6
1 494	50506210	108	6
1 495	50506213	111	6
1 496	50506215	151	6
1 497	50506801	121	6
1 498	50506803	79	6
1 499	50506806	357	6
1 500	50506809	132	6

Tableau 6 : Extrait du fichier des 9000 ménages agricoles tirés

Num_Men	ID_ZD	Rang_Tirage	Nombre_Tiré	ID_MEN	Région	Préfecture	Canton	MEN_AGR
1	10101102	2	32	10101102032	MARITIME	AVE	AKEPE	140
2	10101102	5	74	10101102074	MARITIME	AVE	AKEPE	140
3	10101102	3	86	10101102086	MARITIME	AVE	AKEPE	140
4	10101102	6	103	10101102103	MARITIME	AVE	AKEPE	140
5	10101102	1	122	10101102122	MARITIME	AVE	AKEPE	140
6	10101102	4	128	10101102128	MARITIME	AVE	AKEPE	140
7	10101104	1	105	10101104105	MARITIME	AVE	AKEPE	331
8	10101104	5	116	10101104116	MARITIME	AVE	AKEPE	331
9	10101104	6	130	10101104130	MARITIME	AVE	AKEPE	331
10	10101104	4	221	10101104221	MARITIME	AVE	AKEPE	331
11	10101104	2	291	10101104291	MARITIME	AVE	AKEPE	331
12	10101104	3	326	10101104326	MARITIME	AVE	AKEPE	331
13	10101105	1	14	10101105014	MARITIME	AVE	AKEPE	168
14	10101105	6	27	10101105027	MARITIME	AVE	AKEPE	168
15	10101105	3	53	10101105053	MARITIME	AVE	AKEPE	168
16	10101105	4	64	10101105064	MARITIME	AVE	AKEPE	168
17	10101105	2	114	10101105114	MARITIME	AVE	AKEPE	168
18	10101105	5	138	10101105138	MARITIME	AVE	AKEPE	168
19	10101721	2	38	10101721038	MARITIME	AVE	ASSAHOUN	176
20	10101721	3	89	10101721089	MARITIME	AVE	ASSAHOUN	176
21	10101721	6	112	10101721112	MARITIME	AVE	ASSAHOUN	176
22	10101721	5	142	10101721142	MARITIME	AVE	ASSAHOUN	176
23	10101721	4	148	10101721148	MARITIME	AVE	ASSAHOUN	176
24	10101721	1	161	10101721161	MARITIME	AVE	ASSAHOUN	176
25	10101724	4	21	10101724021	MARITIME	AVE	ASSAHOUN	128

9. Probabilités et poids de sondage des unités de sondage

Les notations présentées ci-après servent à établir les formules de définition des probabilités d'inclusion et des poids de sondage des unités de sondage.

- h désigne la strate dans un domaine d'étude ;
- m_h est l'effectif des UP (unités primaires) ou ZD agricoles tirées dans la strate h ;
- la strate h est composée de M_h unités primaires (UP) d'étiquettes 1, 2, ..., M_h ;
- l'unité primaire ou ZD agricole i de la strate h sera notée UP_{hi} ;
- N_{hi} désigne la taille de l'unité primaire UP_{hi} ;
- la taille N_{hi} est pour la base de sondage utilisée, l'effectif des ménages agricoles de l'unité primaire UP_{hi} ;
- N_h désigne la somme des tailles N_{hi} des unités primaires UP_{hi} et est défini par la relation

$$N_h = \sum_{i=1}^{M_h} N_{hi} \quad (1)$$

- n est le nombre constant de ménages agricoles à tirer au 2^e degré de sondage dans l'unité primaire UP_{hi} de la strate h .

Au premier degré, m_h unités primaires ou ZD agricoles sont tirées de la strate h selon le mode de tirage systématique avec probabilités proportionnelles aux tailles des unités.

Au 2^e degré, un nombre constant n de ménages agricoles est tiré par sondage aléatoire simple sans remise de chaque UP échantillonnée dans la strate h .

- P_{hi} désigne la probabilité d'inclusion de l'unité primaire UP_{hi} dans l'échantillon du 1^{er} degré ;
- $P_{j.hi}$ désigne la probabilité de sélection du ménage agricole j de l'unité primaire UP_{hi} de la strate h ;
- P_{hij} désigne la probabilité d'inclusion du ménage agricole j de l'unité primaire UP_{hi} de la strate h dans l'échantillon de l'enquête.

Le tirage systématique de l'unité primaire UP_{hi} avec probabilités proportionnelles à la taille de l'unité fait que la probabilité d'inclusion P_{hi} est proportionnelle à la taille relative N_{hi}/N_h de cette unité primaire, ce qui s'exprime par la double relation suivante :

$$P_{hi} = k \frac{N_{hi}}{N_h} \quad \forall i = 1, 2, \dots, M_h \quad (2)$$

On montre que pour tout plan de sondage de taille fixe m_h (taille de l'échantillon issu de la strate h), on a

$$\sum_{i=1}^{M_h} P_{hi} = m_h \quad (3)$$

On en déduit

$$m_h = \sum_{i=1}^{M_h} P_{hi} = k \sum_{i=1}^{M_h} \frac{N_{hi}}{N_h} = k$$

D'où

$$P_{hi} = m_h \frac{N_{hi}}{N_h} \quad (4)$$

soit encore

$$P_{hi} = m_h \frac{N_{hi}}{N_h} = m_h \frac{N_{hi}}{\sum_{i=1}^{M_h} N_{hi}} \quad (5)$$

On montre que les probabilités $P_{j.hi}$ et P_{hij} ont pour expression respectivement

$$P_{j.hi} = \frac{n}{N_{hi}} \quad (6)$$

$$\text{et} \quad P_{hij} = P_{hi} P_{j.hi} \quad (7)$$

D'où finalement

$$P_{hij} = m_h \frac{N_{hi}}{N_h} \frac{n}{N_{hi}} = \frac{m_h n}{N_h} \quad (8)$$

On en déduit le poids de sondage W_{hij} du ménage agricole j de l'UP i de la strate h qui est l'inverse de la probabilité d'inclusion P_{hij} , soit

$$W_{hij} = \frac{N_h}{m_h n} \quad (9)$$

De la formule (8) on observe que la probabilité d'inclusion P_{hij} ne dépend que de la strate h et aucunement du ménage agricole j ni de l'unité primaire i . Cela signifie que tous les ménages agricoles de la strate h indépendamment de l'unité primaire d'appartenance i ont la même probabilité de faire partie de l'échantillon sélectionné dans la strate h .

Il en résulte que tous les ménages agricoles d'une même strate h ont également le même poids de sondage W_{hij} comme l'indique la formule (9). Nous sommes dans le cas d'un échantillon dit auto pondéré.

Dans le cas particulier d'un échantillon auto pondéré, ainsi que le montre la formule (8), la probabilité d'inclusion d'un ménage agricole de la strate h est identique au taux global de sondage de la strate, c'est-à-dire le rapport du nombre des ménages tirés nm_h dans la strate au nombre total de ménages N_h de la strate.

Avec un échantillon auto pondéré, "l'enquête se dépouille comme un recensement", ce qui signifie tout simplement que la moyenne d'échantillon d'une variable d'étude estime sans biais la moyenne de la population sondée.

10. Mise en œuvre du calcul des poids de sondage

La mise en œuvre du calcul des poids de sondage est relativement simple avec ce plan de sondage du fait qu'il donne lieu à un échantillon auto pondéré. En effet, tous les ménages agricoles d'une même strate possèdent le même poids de sondage quelle que soit la ZD agricole d'appartenance. Seules les ZD agricoles atypiques tirées d'office possèdent des poids de sondage différents du poids de sondage caractérisant la strate.

La mise en œuvre du calcul des poids de sondage conduit à l'élaboration de deux fichiers, l'un présentant le poids de sondage pour chacune des 35 strates (ou préfectures) et l'autre, les poids de sondage pour les 14 ZD atypiques qui ont été tirées d'office.

Le fichier Excel 1 dénommé "Poids de sondage par strate" présente les poids de sondage pour les différentes strates du recensement. Le tableau 7 donne une image de ce fichier.

Le fichier Excel 2 dénommé "Poids de sondage des ZD tirées d'office" présente les poids de sondage des 14 ZD agricoles tirées d'office. Le tableau 8 en donne une image.

Un autre fichier Excel sera élaboré au moment du traitement des données pour présenter les poids de sondage redressés pour cause de non réponses totales.

Tableau 7 : Probabilités d'inclusion et poids de sondage initiaux par strate

Région	Strate ou préfecture	Effectif de ZD agricoles tirées dans la strate h	Effectif des ménages agricoles de la strate h en 2011	Nombre de ménages agricoles tirés dans la ZD i de la strate h	Probabilité d'inclusion du ménage j de la ZD i de la strate h	Poids de sondage initial du ménage j de la ZD i de la strate h
	h	m _h	N _h	n	P _{hij} = m _h n/N _h	W _{hij} = 1/P _{hij}
Maritime	AVE	44	14 584	6	0,01810	55,242
	BAS-MONO	38	12 585	6	0,01812	55,197
	LACS	40	13 025	6	0,01843	54,271
	VO	66	21 702	6	0,01825	54,803
	YOTO	58	18 981	6	0,01833	54,543
	ZIO	79	25 914	6	0,01829	54,671
Plateaux	AGOU	33	13 578	6	0,01458	68,576
	AKEBOU	21	8 704	6	0,01448	69,079
	AMOU	38	16 387	6	0,01391	71,873
	ANIE	22	9 118	6	0,01448	69,076
	DANYI	17	7 129	6	0,01431	69,892
	EST-MONO	40	16 271	6	0,01475	67,796
	HAHO	68	27 947	6	0,01460	68,498
	KLOTO	18	7 796	6	0,01385	72,185
	KPELE-AKATA	25	10 126	6	0,01481	67,507
	MOYEN-MONO	27	11 302	6	0,01433	69,765
	OGOU	59	24 451	6	0,01448	69,071
	WAWA	29	12 013	6	0,01448	69,040
Centrale	BLITTA	60	15 965	6	0,02255	44,347
	PLAINE DE MO	16	4 235	6	0,02267	44,115
	SOTOUBOUA	49	12 598	6	0,02334	42,850
	TCHAMBA	55	14 877	6	0,02218	45,082
	TCHAOUDJO	42	10 893	6	0,02313	43,226
Kara	ASSOLI	6	1 696	6	0,02123	47,111
	BASSAR	45	13 312	6	0,02028	49,304
	BINAH	34	9 896	6	0,02061	48,510
	DANKPEN	44	12 899	6	0,02047	48,860
	DOUFELGOU	38	11 780	6	0,01935	51,667
	KERAN	38	11 122	6	0,02050	48,781
	KOZAH	66	19 987	6	0,01981	50,472
Savanes	CINKASSE	17	5 205	6	0,01960	51,029
	KPENDJAL	54	16 677	6	0,01943	51,472
	OTI	69	20 845	6	0,01986	50,350
	TANDJOARE	51	15 023	6	0,02037	49,095
	TONE	80	24 006	6	0,02000	50,013

Tableau 8 : Poids de sondage des ménages des ZD atypiques tirées d'office

		ZD agricole	Strate ou préfecture	Probabilité d'inclusion de la ZD i	Effectif des ménages agricoles de la ZD i en 2011	Nombre de ménages agricoles tirés dans la ZD i	Probabilité de sélection du ménage j dans la ZD i	Probabilité d'inclusion du ménage j de la ZD i	Poids de sondage initial du ménage j de la ZD i
Numéro	Région	ID_ZD	h	P_{hi}	N_{hi}	n	$P_{j,hi} = n/N_{hi}$	$P_{hij} = P_{hi} * P_{j,hi}$	$W_{hij} = 1/P_{hij}$
1	Plateaux	20303701	AMOU	1,00	487	6	0,0123	0,0123	81,167
2		20304101	AMOU	1,00	567	6	0,0106	0,0106	94,500
3		20802101	KLOTO	1,00	605	6	0,0099	0,0099	100,833
4	Centrale	30101406	BLITTA	1,00	284	6	0,0211	0,0211	47,333
5		30402408	TCHAMBA	1,00	289	6	0,0208	0,0208	48,167
6		30402511	TCHAMBA	1,00	325	6	0,0185	0,0185	54,167
7	Kara	40500101	DOUFELGOU	1,00	557	6	0,0108	0,0108	92,833
8		40501903	DOUFELGOU	1,00	357	6	0,0168	0,0168	59,500
9		40701502	KOZAH	1,00	393	6	0,0153	0,0153	65,500
10		40704101	KOZAH	1,00	434	6	0,0138	0,0138	72,333
11	Savanes	50203102	KPENDJAL	1,00	333	6	0,0180	0,0180	55,500
12		50203105	KPENDJAL	1,00	366	6	0,0164	0,0164	61,000
13		50300210	OTI	1,00	456	6	0,0132	0,0132	76,000
14		50506806	TONE	1,00	357	6	0,0168	0,0168	59,500

11. Estimateur d'un total

Si y_{ij} est la valeur observée pour la variable d'étude y chez le ménage agricole j de l'unité primaire UP_i dans une strate donnée, alors on estime le total Y de la variable y par la grandeur $\hat{Y}_{HT}$ d'expression

$$\hat{Y}_{HT} = \sum_{i=1}^m \sum_{j=1}^n W_{ij} y_{ij} \quad (10)$$

où W_{ij} est le poids de sondage du ménage j de l'unité primaire UP_i .

La grandeur $\hat{Y}_{HT}$ appelée estimateur de Horvitz-Thompson. On montre que $\hat{Y}_{HT}$ est un estimateur¹ linéaire sans biais du total Y pour tout plan de sondage.

On déduit de la formule (9) la double relation suivante

$$\hat{Y}_{\pi} = \sum_{i=1}^m \sum_{j=1}^n W_{ij} y_{ij} = \frac{N}{mn} \sum_{i=1}^m \sum_{j=1}^n y_{ij} \quad (11)$$

où N et m sont respectivement l'effectif des ménages agricoles de la strate étudiée et le nombre de ZD agricoles tirées dans cette strate, et n le nombre constant de ménages agricoles tirées par ZD agricole.

Si la variable d'étude y est la taille du ménage agricole j de la ZD i , le total Y de cette variable est l'effectif total de la population de la strate sondée. On estime ce total grâce à la formule (11), où :

- y_{ij} est la taille du ménage agricole j de la ZD i ;
- N est l'effectif total des ménages agricoles de la strate sondée ;
- m est la taille de l'échantillon des ZD agricoles tirées dans la strate ;
- n est le nombre des ménages agricoles tirés dans chaque ZD échantillonnée.

12. Estimation d'un effectif, d'une moyenne ou d'une proportion

12. 1 Estimation d'un effectif

L'estimation d'un effectif n'est autre que l'estimation du total d'une variable indicatrice. Cette estimation est donc basée sur la formule (11).

Soit y la variable indicatrice relative à l'utilisation de tracteur pour les labours. La valeur y_{ij} de la variable indicatrice pour le ménage agricole j de la ZD i est :

- 1 si ce ménage agricole utilise le tracteur pour les labours ;
- 0 sinon.

Le total Y de cette variable est l'effectif des ménages agricoles de la strate étudiée qui utilisent le tracteur. Ce total est estimé par l'estimateur de Horvitz-Thompson.

¹ Ce résultat a été établi en 1952 par Horvitz et Thompson

12. 2 Estimation d'une moyenne

De même, de l'estimateur du total, on déduit l'estimateur de la moyenne d'une variable d'étude pour un domaine d'étude ou région. S'il s'agit d'une moyenne par ménage agricole, on divise le total estimé pour la variable d'étude par l'effectif total des ménages agricoles du domaine d'étude, une donnée qui figure dans la base de sondage fournie par le pré recensement de l'agriculture.

Cas de l'estimation d'une moyenne de l'ensemble du pays

L'estimation d'une moyenne de l'ensemble du pays doit se faire en tenant compte de ce que la répartition des échantillons selon les régions n'est pas uniforme. Trois régions ont été sur échantillonnées au détriment d'une des autres régions. En conséquence, l'estimation d'une moyenne au niveau national doit être une moyenne pondérée des estimations de région, les coefficients de pondération devant obéir à la structure des régions dans la base de sondage et non à celle de l'échantillon.

Le tableau 9 suivant présente les tailles relatives des régions (ou domaines d'étude) selon la base de sondage et selon l'échantillon des ménages agricoles.

Tableau 9 : Tailles relatives des régions

Région	Dans la base de sondage	Dans l'échantillon
Maritime	0,217	0,217
Plateaux	0,335	0,267
Centrale	0,119	0,150
Kara	0,164	0,183
Savanes	0,166	0,183
Ensemble	1,000	1,000

12. 3 Estimation d'une proportion

L'estimation d'une proportion n'est autre que l'estimation de la moyenne d'une variable indicatrice. L'estimation de cette moyenne particulière se fait donc cela est indiqué plus haut avec le traitement spécial à donner à l'estimation d'une proportion au niveau national.

13. Mise en œuvre des estimations

L'estimation d'un total au moyen de l'estimateur de Horvitz-Thompson se fait en construisant un tableau de 9000 lignes (une ligne par ménage agricole tiré) et de deux colonnes, l'une réservée au poids de sondage W_{ij} du ménage agricole et l'autre à la valeur de la variable d'étude y_{ij} pour le ménage j correspondant. Le produit scalaire des données des deux colonnes fournit l'estimation du total recherché.

En cas de non réponses totales, les poids de sondage initiaux doivent être redressés et ce sont les poids redressés qui seront pris en compte dans la colonne 1. Dans cette situation, seuls les ménages agricoles enquêtés figureront dans le tableau de calcul défini plus haut.

Les calculs peuvent être réalisés avec Excel, mais ils sont plus aisés et plus rapides avec un logiciel de calcul plus élaboré comme SPSS.