



Food and Agriculture
Organization of the
United Nations

COMMISSION AFRICAINE DES STATISTIQUES AGRICOLES

Vingt-septième Session

15 – 18 novembre 2021, Hôte virtuel – Dakar, Sénégal

POINT 5 DE L'ORDRE **DU JOUR**

LIGNES DIRECTRICES SUR LA DÉSAGRÉGATION DES INDICATEURS DES ODD À PARTIR DES DONNÉES D'ENQUÊTES

Clara Aida Khalil

Statisticienne

Bureau du Statisticien en Chef - FAO

Désagrégation des données et les ODD

- En adoptant l'Agenda 2030 relatif aux Objectifs de développement durable, les États membres se sont engagés à ne laisser personne derrière (LNOB) et d'aider premièrement les plus défavorisés : Ceci nécessite plus de **données désagrégées** que ce dont disposent actuellement la plupart des pays.
- Un **principe clé de la désagrégation des données** est au cœur du cadre de suivi des ODD :

« Les indicateurs des ODD devraient être ventilés, le cas échéant, par revenu, sexe, âge, race, appartenance ethnique, statut migratoire, handicap et lieu de résidence, ou d'autres caractéristiques conformément aux Principes fondamentaux de la statistique officielle ».

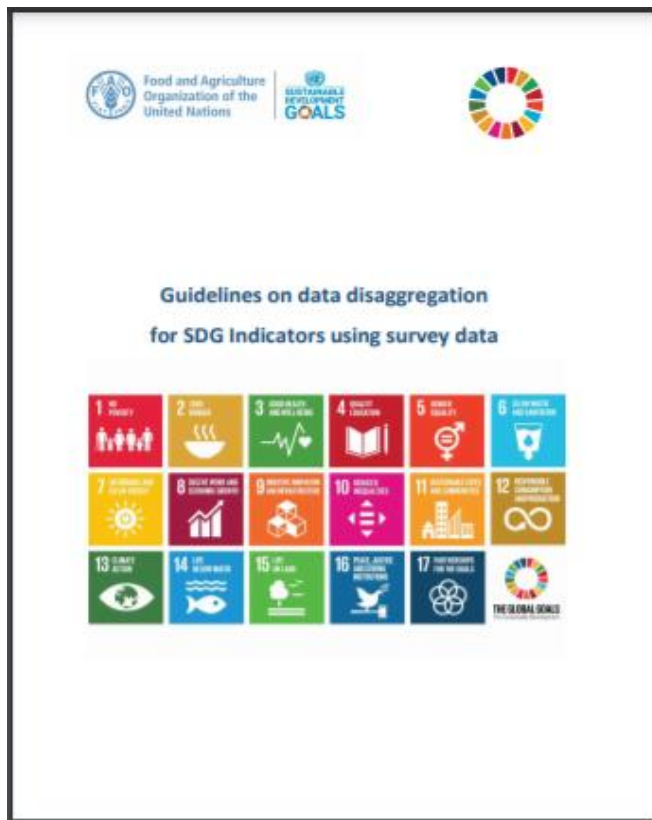


Désagrégation des données et les ODD

Le GIE-ODD a mis sur pied un groupe de travail sur la désagrégation des données pour renforcer les capacités nationales et élaborer les normes et outils statistiques nécessaires à la production de données désagrégées. Ceci a conduit à :

- L'élaboration d'un ensemble minimum de désagrégation et à la compilation des catégories et des dimensions de désagrégation des données actuellement en place ou prévues par les organes garants
- Une synthèse détaillée des normes et des catégories de désagrégation pour tous les indicateurs des ODD.
- Une compilation des politiques prioritaires par dimension de désagrégation.
- Une compilation des méthodes et outils de désagrégation des données.

Lignes directrices sur la désagrégation des indicateurs



En tant que membres du **GT sur la désagrégation** des données et **TA sur SAE**, la FAO a élaboré des « **Lignes directrices sur la désagrégation des indicateurs des ODD à partir des données d'enquête** »

Objectifs principaux de la publication des Lignes directrices :

- Proposer un guide méthodologique et pratique pour la production d'estimations désagrégées **directes et indirectes** relatives aux indicateurs des ODD
- Proposer des outils pour évaluer l'exactitude des estimations et élaborer des stratégies d'intégration des données, y compris les méthodes d'évaluation des **petits domaines** (SAE)

Lien pour accéder aux lignes directrices :

<http://www.fao.org/documents/card/en/c/cb3253en/>

Importance des Lignes directrices

- Environ 1/3 des indicateurs des ODD peut être calculés en utilisant les données d'enquête.
- 6 des 21 Indicateurs des ODD dont la FAO est l'organe garant peuvent dépendre des données des enquêtes agricoles et/ou sur les ménages.

PROBLÉMATIQUE : l'utilisation des techniques traditionnelle d'échantillonnage s'accompagne de contraintes liées à la production de données désagrégées et d'estimations fiables pour les sous-populations de petite taille. Les techniques innovantes pouvant permettre de résoudre certains de ces problèmes sont loin d'être adoptées par les bureaux nationaux de statistique.

Désagrégation des données à partir des enquêtes par sondage

Synthèse des Lignes directrices :

- Estimations directes relatives à un indicateur d'une sous-population donnée : à partir d'informations provenant d'échantillon de la sous-population elle-même.

Deux contraintes majeures :

- La taille de l'échantillon n'est souvent pas assez grande pour assurer des estimations fiables pour les petits domaines ;
 - Possibilité d'avoir des sous-domaines non échantillonnés.
- Ces contraintes peuvent être résolues :
 - Durant la phase de conception : en adoptant des plans d'échantillonnage qui garantissent un ensemble observé d'unités d'échantillonnage pour chaque sous-population pour laquelle doivent être générées des données désagrégées.
 - Durant la phase d'analyse : en produisant des estimations indirectes, tout en surmontant le nombre limité d'informations disponibles sur les « petits domaines » en s'appuyant sur d'autres sources de données.

Désagrégation des données lors de la phase de plan d'échantillonnage

Afin de produire des **estimations directes désagrégées**, les stratégies d'échantillonnage devraient inclure un nombre suffisant d'unités d'échantillonnage dans chaque domaine de désagrégation.

En outre, le fait d'avoir des unités d'échantillonnage dans chaque domaine de désagrégation améliore les **estimations indirectes** : réduction des biais et de la variance du modèle

- **Calcul direct** lorsque le nombre d'unités d'une sous-population peut être déterminé à partir du plan d'échantillonnage. Dans ces cas, le problème majeur constitue le choix du degré de suréchantillonnage à utiliser.
- **Plus complexe** lorsque le plan d'échantillonnage disponible ne permet pas de connaître d'avance les membres des sous-populations.

Désagrégation des données lors de la phase de plan d'échantillonnage (2)

Des approches intégrées et innovantes pour assurer une taille d'échantillonnage appropriée pour chaque domaine de désagrégation sont illustrées et expliquées (avec leurs avantages et inconvénients), notamment :

- Suréchantillonnage
- Stratification plus poussée
- Échantillonnage à plusieurs phases avec sélection des répondants
- Plans de stratification marginale
- Échantillonnage indirect

TRADITIONNELLE

INNOVANTE

Assure un échantillon suffisant pour chaque domaine de désagrégation sans augmenter la taille globale de l'échantillon

Des ensembles de logiciels sont proposés

Désagrégation des données lors de la phase d'analyse

- Les lignes directrices présentent des stratégies d'échantillonnage alternatives pour des estimations directes du domaine d'échantillonnage
- Les **estimateurs directs et indirects assistés par modèle** les plus courants sont présentés tout en introduisant leur contexte d'utilisabilité.
- **Les techniques d'estimation des petits domaines** (SAE) ont été illustrées en :
 - Présentant le flux des processus de leur application ;
 - Donnant un aperçu des principales approches au niveau des unités et au niveau des domaines ;
 - Indiquant les principales références sur le sujet ;
 - Proposant des outils pour évaluer la qualité des estimations des petits domaines.

Désagrégation des données lors de la phase d'analyse (2)

L'Estimateur par projection (Kim et Rao, 2012) a été présenté, expliqué et utilisé sur des données d'enquête réelle.

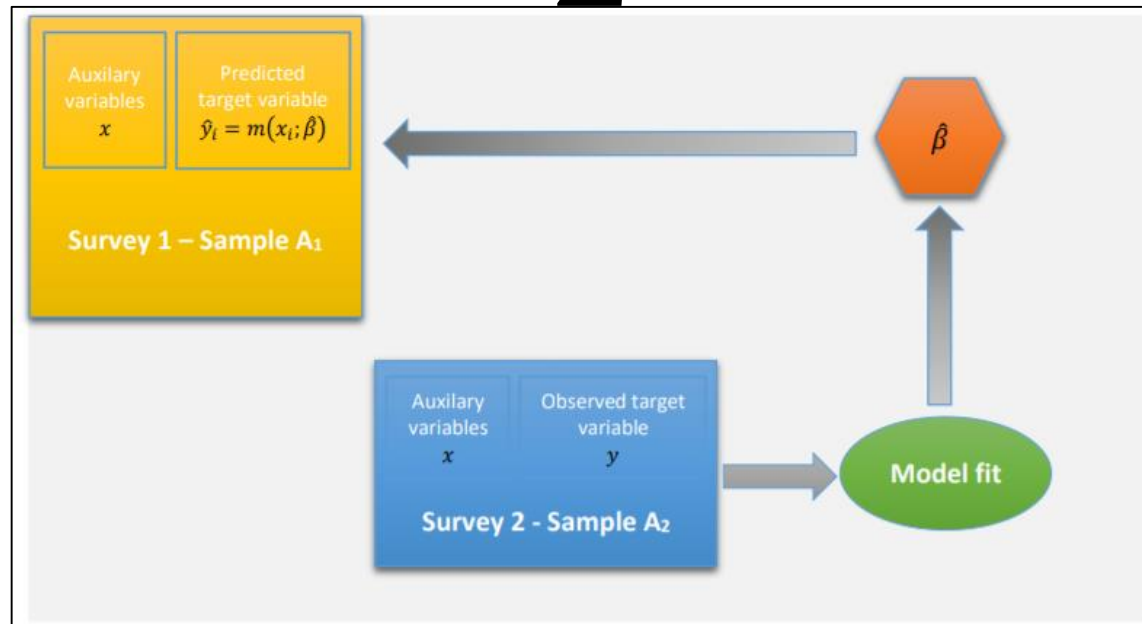
Il permet de produire des indicateurs désagrégés par **l'utilisation conjointe de deux enquêtes par sondage** :

- **La première enquête**, caractérisée par un échantillon A_1 de grande taille sans toutefois collecter des informations auxiliaires ou des variables d'usage général (exemple : variables socio-économiques).
- **La deuxième enquête** a un échantillon plus petit A_2 mais permet la collecte d'informations sur la variable cible y , ainsi que le même ensemble de variables auxiliaires disponibles dans A_1 .

Désagrégation des données lors de la phase d'analyse (3)

La variable totale y dans le domaine de désagrégation d peut être estimée par :

$$\hat{Y}_{PR,d} = \sum w_{i1} m(x_i; \hat{\beta}) y_{id}$$



Évaluation de l'exactitude des estimations

Les lignes directrices mettent l'accent sur l'importance d'estimer et de vulgariser les mesures d'exactitude pour :

- Permettre aux utilisateurs d'évaluer si une estimation peut être utilisée.
- Renforcer la confiance du public par rapport aux données et leur utilisation.

Des méthodes et des outils d'évaluation de l'exactitude des estimations directes sont proposés :

- Estimation de la variance de l'échantillon : à utiliser lorsque le calcul des indicateurs est basé sur les propriétés inférentielles de plans d'échantillonnage répétés.
- Variance du modèle : convient lorsque l'estimation repose **uniquement** sur des modèles utilisant des variables auxiliaires.
- Variance globale : convient lorsque les approches basées des modèles sont utilisées conjointement avec l'inférence suivant le plan

Un exemple pratique basé sur l'indicateur 2.1.2 des ODD

L'approche a été adoptée pour produire des estimations désagrégées relatives à l'indicateur 2.1.1 des ODD : la Prévalence d'une insécurité alimentaire modérée ou grave, évaluée selon l'échelle de mesure de l'insécurité alimentaire vécue (FIES). Objectif : Produire des estimations relatives à l'indicateur 2.1.2 par sexe, milieu rural/urbain, tranche d'âge, et par quintile de revenu.

Deux sources de données :

- La quatrième enquête intégrée sur les ménages du Malawi (IHS4) 2016-17
- Le module d'enquête FIES du Malawi collecté par le Gallup World Poll (GWP) - 2016

Les résultats sont présentés avec les mesures d'exactitude des estimations indirectes.

Un exemple pratique basé sur l'indicateur 2.1.2 des ODD(2)

La réalisation pratique de cette étude de cas peut se résumer en ces étapes :

- I. **Identification et recodage des variables auxiliaires.** L'application de *l'estimateur par projection* nécessite la disponibilité d'un même ensemble de variables auxiliaires dans les deux enquêtes à intégrer. Ces variables doivent également avoir la même structure et définition.
- II. **Définition de la fonction $m()$ et des paramètres de l'estimation de la projection.** La sélection de la forme fonctionnelle pour la fonction de lien $m()$ dépend essentiellement du type de variable y considérée (exemple : échelle, nominale, dichotomique).
- III. **Calcul de valeurs synthétiques.** En utilisant le paramètre de projection estimé, les valeurs synthétiques de la variable cible sont calculées dans l'ensemble de données de grande taille.
- IV. **Évaluation de l'exactitude des estimations.** Estimation de la variance, du coefficient de variation et des intervalles de confiance des estimations indirectes projetées.

Un exemple pratique basé sur l'indicateur 2.1.2 des ODD (3)

Quelques résultats :

		Insécurité alimentaire modérée ou grave			
		Prob.ms	CV (%)	CI_Plus faible	CI_Plus élevé
IHS4	Total	0.91	1.2	0.89	0.93
GWP		0.91	1.3	0.89	0.93
IHS4	Femme	0.91	1.4	0.88	0.93
GWP		0.90	1.5	0.89	0.94
IHS4	Home	0.91	1.9	0.87	0.94
GWP		0.91	2.0	0.87	0.94
IHS4	Rural	0.93	1.2	0.90	0.95
GWP		0.92	1.3	0.90	0.94
IHS4	Urbain	0.81	5.7	0.73	0.92
GWP		0.82	5.9	0.74	0.93
IHS4	15-24	0.91	2.0	0.87	0.94
GWP		0.89	2.1	0.85	0.93
IHS4	25-49	0.91	1.6	0.88	0.93
GWP		0.92	1.6	0.89	0.95
IHS4	50-64	0.87	3.6	0.82	0.94
GWP		0.90	3.5	0.84	0.96
IHS4	65+	0.97	1.6	0.94	1
GWP		0.98	1.7	0.95	1
IHS4	Inc_1	0.96	1.5	0.94	0.99
GWP		0.97	1.5	0.94	1
IHS4	Inc_2	0.96	1.5	0.93	0.99
GWP		0.96	1.6	0.93	0.99
IHS4	Inc_3	0.97	1.1	0.95	0.99
GWP		0.97	1.1	0.95	0.99
IHS4	Inc_4	0.89	3.6	0.82	0.95
GWP		0.88	3.7	0.82	0.94
IHS4	Inc_5	0.74	3.8	0.68	0.80
GWP		0.76	3.8	0.71	0.82

		Insécurité alimentaire grave			
		Prob.s	CV (%)	CI_plus faible	CI_plus élevé
IHS4	Total	0.73	2.4	0.67	0.75
GWP		0.71	2.8	0.67	0.75
IHS4	Femme	0.75	2.8	0.71	0.80
GWP		0.75	3.1	0.71	0.80
IHS4	Homme	0.70	3.6	0.65	0.75
GWP		0.67	4.2	0.61	0.73
IHS4	Rural	0.75	2.4	0.72	0.79
GWP		0.72	2.9	0.68	0.76
IHS4	Urbain	0.63	9.5	0.51	0.75
GWP		0.63	9.2	0.52	0.75
IHS4	15-24	0.72	3.8	0.66	0.77
GWP		0.67	4.5	0.61	0.73
IHS4	25-49	0.71	3.2	0.67	0.76
GWP		0.72	3.6	0.67	0.77
IHS4	50-64	0.74	6.9	0.64	0.84
GWP		0.75	7.1	0.65	0.86
IHS4	65+	0.86	5.4	0.75	0.96
GWP		0.87	5.8	0.78	0.98
IHS4	Inc_1	0.88	3.3	0.83	0.94
GWP		0.88	3.4	0.83	0.94
IHS4	Inc_2	0.82	3.8	0.75	0.88
GWP		0.81	4.2	0.75	0.88
IHS4	Inc_3	0.77	5.3	0.69	0.85
GWP		0.75	5.3	0.67	0.83
IHS4	Inc_4	0.68	6.0	0.60	0.76
GWP		0.64	6.5	0.56	0.72
IHS4	Inc_5	0.48	8.4	0.40	0.56
GWP		0.48	8.4	0.40	0.56



MERCI !

ClaraAida.Khalil@fao.org