

Institut national de la statistique et de
la démographie (INSD)



**ADAPTATION DES METHODOLOGIES
STATISTIQUES AUX SITUATIONS
SECURITAIRES ET SANITAIRES
DIFFICILES :**

**STRATEGIES DE RESILIENCE OU RUPTURE AVEC
LES APPROCHES TRADITIONNELLES**

Novembre 2022

Ministère de l'économie, des
finances et de la prospective

Secrétariat général

Institut national de la statistique
et de la démographie

BURKINA FASO
Unité - Progrès -
Justice



ADAPTATION DES METHODOLOGIES STATISTIQUES AUX SITUATIONS SECURITAIRES ET SANITAIRES DIFFICILES :

STRATEGIES DE RESILIENCE OU RUPTURE AVEC LES APPROCHES TRADITIONNELLES

Novembre 2022

AVANT-PROPOS

Le Burkina Faso a adopté en 2021 son quatrième schéma directeur de la statistique (SDS 2021-2025) dont l'objectif général est de permettre au Système statistique national (SSN) de produire et de diffuser en temps opportun, des données statistiques correspondant aux besoins des utilisateurs. La bonne planification et l'exécution des opérations statistiques (enquêtes, recensements) concourent à la réalisation de cet objectif. Cependant, bien que les efforts déployés aient permis d'engranger des acquis majeurs, le contexte sécuritaire compromet de plus en plus la mise en œuvre des activités statistiques, avec son corollaire de reports voire de suspensions d'opérations de collecte dans certaines localités.

Face à ces difficultés, les équipes techniques sont souvent amenées à adopter les initiatives suivantes :

- exclure certaines zones à forts défis sécuritaires dès la phase d'échantillonnage ;
- remplacer des zones inaccessibles par d'autres plus accessibles ;
- Abandonner la collecte dans les zones inaccessibles bien qu'initialement incluses dans l'échantillon.

Ainsi, dans ce contexte, il devient difficile, pour les équipes techniques en charge de ces enquêtes, d'avoir un contrôle sur l'échantillon ; toute chose qui influe sur la qualité de l'estimation des indicateurs.

Au final, dans l'optique de garantir une production statistique régulière et résiliente, il est impératif que le système statistique national engage des réflexions sur des méthodologies adaptatives. Ces dernières impliquent une modernisation du processus de production et de diffusion de l'information statistique, ainsi que l'utilisation de nouvelles sources de données générées par les technologies de l'information et de la communication.

C'est dans ce cadre que l'Institut national de la statistique et de la démographie qui assure le secrétariat technique du conseil national de la statistique, entreprend à travers un processus participatif, l'élaboration d'un document d'adaptation des méthodologies statistiques aux situations sécuritaires et sanitaires difficiles.

Le présent document fait une analyse de la situation et trace les grandes orientations qui serviront de base à la mise en œuvre des futures opérations statistiques. Toutefois, l'Institut demeure attentif à toute suggestion contributive dans le sens d'améliorer la qualité de ce document.

Le Directeur général

Boureima OUEDRAOGO

Chevalier de l'Ordre du Mérite de l'Economie et des Finances

SOMMAIRE

Avant-propos	i
Sommaire	ii
Liste des acronymes et abréviations.....	iv
Liste des tableaux	v
RESUME EXECUTIF	vi
I. DIAGNOSTIC DE LA SITUATION	3
1.1. Les situations rencontrées et initiatives spontanées	3
1.1.1. Cas des recensements.....	3
1.1.2. Cas des enquêtes par sondage.....	4
1.2. Analyse des forces, faiblesses, opportunités et menaces des opérations statistiques	6
1.2.1. Analyse des forces, faiblesses, opportunités et menaces de la phase de conception....	6
1.2.2. Analyse des forces, faiblesses, opportunités et menaces de la phase terrain.....	7
1.2.3. Analyse des forces, faiblesses, opportunités et menaces de la phase de traitement des données	8
II. STRATEGIE DE RESILIENCE	10
2.1. Phase de conception	10
2.2. Phase de collecte	10
2.3. Phase de traitement.....	11
2.4. Analyse des forces, faiblesses, menaces et opportunités des stratégies de résilience	12
III. RUPTURE AVEC LES APPROCHES TRADITIONNELLES	14
3.1. Conception	14
3.2. Collecte	17
3.3. Traitement.....	20
IV. STRATEGIE DE MISE ŒUVRE.....	22
V. PLAN D' ACTIONS.....	26
CONCLUSION.....	34
RECOMMANDATIONS.....	34
BIBLIOGRAPHIE	35
ANNEXES	a
Annexe 1: Diagnostic du RGPH 2019	a
Annexe 2 : Diagnostic de l'EPC.....	b
Annexe 3: Diagnostic de l'ERI-ESI	b
Annexe 4 : Diagnostic de EAM.....	c
Annexe 5 : Diagnostic de EDSBF V	c
Annexe 6 : Diagnostic de EHCVM 1.....	d
Annexe 7 : Diagnostic de EHCVM-2.....	e
Annexe 8 : Diagnostic de l'enquête COVID	f
Annexe 9 : Diagnostic du RIC	f

Annexe 10 : Diagnostic des statistiques des prix	f
Annexe 11 : Forces et faiblesses des Enquêtes ménages	h
Annexe 12 : Opportunités et menaces.....	j
Annexe 13 : Forces faiblesse de l'enquête entreprise :.....	k
Annexe 14 : Opportunités et menaces de l'enquête entreprise	m
Annexe 15 : Forces et faiblesse de l'Enquête prix	n
Annexe 16 : Opportunités et menaces de l'enquête prix.....	o
Annexe 17 : Forces et faiblesses du Recensement	p
Annexe 18 : Opportunités et menaces du recensement.....	r
Annexe 19 : Analyse comparative des Procédures/stratégies.....	s

LISTE DES ACRONYMES ET ABREVIATIONS

ANPE : Agence Nationale pour la Promotion de l'Emploi

EPA : Enquête Permanente Agricole

EPC : Enquête Post-Censitaire

ERI-ESI : Enquête Régionale Intégrée sur l'Emploi et le Secteur Informel

EAM : Enquête sur l'Audience des Médias au Burkina Faso

EDS : Enquête Démographique et de Santé ;

EHCVM : Enquête Harmonisée sur les Conditions de Vie des Ménages ;

INSD : Institut National de la Statistique et de la Démographie

ONI : Office National d'Identification

ONEA : Office National de l'Eau et de l'Assainissement

SSN : Système Statistique National

SONABEL : Société Nationale Burkinabè d'Electricité

RGPH : Recensement Général de la Population et de l'Habitation

COVID : Coronas Virus Desease ;

ZD : Zone de Dénombrement

LISTE DES TABLEAUX

Tableau 1 : Synthèse des difficultés rencontrées et initiatives d'adaptation pour les recensements.....	3
Tableau 2 : Synthèse des difficultés rencontrées et initiatives d'adaptation pour les recensements.....	5
Tableau 3 : Analyse des forces, faiblesses, opportunités et menaces de la phase de conception	6
Tableau 4 : Analyse des forces, faiblesses, opportunités et menaces de la phase de collecte de données	7
Tableau 5 : Analyse des forces, faiblesses, opportunités et menaces de la phase de traitement des données.....	8
Tableau 6 : Conditions d'utilisation des méthodes résilientes de collecte de données	11
Tableau 7 : Forces et faiblesses des stratégies de résilience.....	12
Tableau 8 : Opportunités et menaces des stratégies de résilience	12
Tableau 9: Analyse des forces et faiblesses des méthodes	16
Tableau 10: Analyse des opportunités et menaces des méthodes	16
Tableau 11: Analyse des forces et faiblesses des méthodes de collecte.....	18
Tableau 12: Analyse des opportunités et menaces des méthodes	19
Tableau 13: Résumé des coûts du plan d'actions	26
Tableau 14: Plan d'actions budgétisé	27
Tableau 15 : Exemple de cartographie des risques en matière de collecte de données	w

RESUME EXECUTIF

Synthèse méthodologique

Dans le cadre de l'exécution de cette étude, l'approche participative a été utilisée. L'objectif de cette approche a été de permettre à l'ensemble des structures du système statistique national de faire un diagnostic des approches utilisées dans la mise en œuvre des opérations statistiques d'envergure nationale. A l'issue d'un atelier de réflexion organisé, un tour d'horizon de l'ensemble des opérations des cinq (05) dernières années a été fait.

Après avoir dressé le diagnostic des méthodologies initiales adoptées, un cadre d'échange a été ouvert pour faire une analyse des forces, faiblesses, opportunités et menaces de ces approches au regard du contexte sécuritaire au niveau national. Cet exercice a abouti à l'élaboration d'une stratégie de résilience, et de faire des propositions de solutions de rupture avec les approches traditionnelles. Enfin, une stratégie de mise en œuvre des approches méthodologiques innovantes est mise en place.

Du diagnostic des opérations statistiques d'envergure nationale

La plupart des enquêtes statistiques du Burkina Faso utilise un sondage à deux degrés, et la base de sondage du RGPH est utilisée au premier degré. Tout comme pour le déroulement des recensements, la collecte des données se fait par interviews directs dans les ménages échantillons, dans les marchés cibles ou dans les entreprises sélectionnées. Dans la mise en œuvre de ces opérations, plusieurs difficultés subsistent mais des initiatives d'adaptation permettent aux concepteurs de diminuer les sources de biais. Ces difficultés sont le plus souvent observées au niveau de la conception, de la collecte et du traitement de données.

Des difficultés de mise en œuvre des approches méthodologiques traditionnelles

A l'issue du diagnostic, plusieurs difficultés existent quant à l'utilisation des approches traditionnelles dans le contexte sécuritaire et sanitaire au niveau national. Ces difficultés et risques sont entre autres :

- l'instabilité de la base de sondage liée aux déplacements des populations
- les probabilités de tirage des unités primaires perturbées par les déplacements des populations ;
- le défaut de connexion internet par endroit
- la possibilité d'enlèvement du personnel de terrain
- le risque de perte de matériel
- le risque de propagation de maladies transmissibles (COVID-19 par exemple)
- la fermeture de certaines entreprises
- l'inaccessibilité de certaines localités

Même si l'on note des forces des méthodes traditionnelles de collecte des données, on constate que les contextes sécuritaire et sanitaire entraînent des menaces, surtout l'accessibilité des zones de collecte, engendrant ainsi des sources de biais qui doivent être corrigés.

Des stratégies de résilience face aux contextes sécuritaire et sanitaire difficiles

Au regard des difficultés, plusieurs aspects structurels doivent être revus pour tenir compte des contraintes sécuritaires. Il s'agit :

- de la prise en compte du volet sécuritaire dans la budgétisation des opérations statistiques ;
- de la modification des plans de sondage en tenant compte de quatre (04) domaines de tirages : urbain accessible, urbain non accessible, rural accessible, rural non accessible. L'allocation des grappes pourrait se faire de façon proportionnelle entre le milieu accessible et le milieu non accessible dans chacune des régions concernées ;

- au niveau de la phase de collecte, tenir compte de plusieurs aspects notamment : le recrutement d'agents locaux, le déplacement des ménages, la gestion des risques psychologiques ;
- au niveau du traitement de données, tenir compte de la restriction de la portée géographique des résultats, réaliser une post-stratification et/ou une repondération, utiliser la méthode d'imputation par les plus proches voisins.

De la rupture avec les approches traditionnelles

Les méthodes de sondage traditionnelles s'appuient sur les tailles de population résidente dans les zones enquêtées. Compte tenu du défi sécuritaire et au vu du déplacement de ces populations la mise à jour de la taille des ZD devient difficile voire impossible sur une partie du territoire. D'autres méthodes alternatives sont envisagées notamment :

- ❖ L'utilisation de la densité de téléphones mobiles par zone pour mettre à jour les tailles des ZD ;
- ❖ L'utilisation des zones de couverture des relais téléphoniques comme unité primaires
- ❖ L'utilisation des images satellitaires et photos aériennes pour dresser la base de sondage du second degré ;
- ❖ Au niveau de la collecte, envisager des approches comme les enquêtes par SMS, WhatsApp, en ligne, par e-mail ;
- ❖ Au niveau du traitement, envisager l'utilisation des données des téléphonies pour redresser les pondérations, mettre en place des modèles économétriques de prédiction des indicateurs, exploiter les données de l'Etat civil, satellitaire et données secondaires.

De la stratégie de mise en œuvre des approches innovantes

L'ossature de la stratégie de mise en œuvre comprend les axes stratégiques suivants :

- ❖ AXE 1 : Amélioration de la qualité de la production statistique en s'adaptant aux contextes sécuritaire et sanitaire difficiles
- ❖ AXE 2 : Modernisation et rationalisation du système de production statistique en tirant profit des avancées technologiques et des récents développements méthodologiques
- ❖ AXE 3 : Assurer l'autonomie et la pérennité de la production statistique en apportant des réponses aux difficultés transversales

Le montant global de la mise en œuvre du plan d'actions se chiffre à près de cinq milliards de franc CFA sur trois ans.

Recommandations

Pour une meilleure mise en œuvre des approches méthodologiques innovantes adaptées au contexte sécuritaire et sanitaire, les recommandations suivantes sont formulées :

- Faire un plaidoyer auprès de l'Etat et des partenaires techniques et financiers pour faciliter la mobilisation des ressources financières et humaines pour faciliter la mise en œuvre des stratégies ;
- Présenter les résultats de l'étude à l'ensemble des structures du système statistique national ;
- Mettre en place un comité de suivi des opérations statistiques utilisant ces approches innovantes ;
- Renforcer les capacités des acteurs du SSN.

INTRODUCTION

Le Burkina Faso a entrepris depuis 2003 l'élaboration de schémas directeur de la statistique pour assurer la disponibilité et la qualité des données statistiques afin de suivre les objectifs nationaux et internationaux de développement. La mise en œuvre de ces différents schémas a permis *(i) de mettre en place et de consolider le cadre législatif, institutionnel et organisationnel, (ii) de développer et de diffuser une importante production statistique, (iii) de renforcer les capacités humaines, matérielles et financières et de management du système statistique.*

Depuis le début de la crise sécuritaire en 2015, les acquis engrangés suite aux efforts de l'Etat et de ses partenaires au développement sont mis à rude épreuve, entraînant des reports et des annulations d'opérations de collecte de données.

Pour les enquêtes et recensements qui ont pu être réalisés, les mesures prises, compte tenu du contexte sécuritaire, sont susceptibles d'affecter la qualité des données.

En effet, la crise sécuritaire a conduit au déplacement de plus de 10% de la population burkinabè, et de ce fait, l'utilisation des informations de la cartographie pour les enquêtes devient problématique. La plupart des enquêtes par sondage réalisées par le système statistique national utilise un sondage à deux degrés, dont le premier degré est relatif au tirage des zones de dénombrement (ZD) avec une probabilité proportionnelle à la taille. Le déplacement des populations ayant affecté la taille de certaines ZD, l'utilisation des données de la cartographie ou du recensement pour concevoir d'autres enquêtes d'envergure nationale introduit un biais dans les probabilités d'inclusion réelles, donc dans les pondérations.

En plus du biais introduit par les données de la cartographie, les contraintes d'accès à certaines localités ont souvent conduit les concepteurs de ces enquêtes à :

- écarter certaines zones à fort déficit sécuritaire dès la phase d'échantillonnage ;
- remplacer les zones inaccessibles par d'autres zones plus accessibles ;
- ne pas enquêter les zones inaccessibles.

Ces pistes de solution sont, elles-mêmes, sources de biais dans les estimations des indicateurs calculés. Avec l'exclusion des zones à fort défi sécuritaire lors de la phase d'échantillonnage, on calcule des indicateurs valides pour les zones sécurisées, mais on impute à tort ces indicateurs à l'ensemble des zones dès lors que la taille de l'échantillon est suffisante.

De même, en remplaçant les zones inaccessibles par d'autres zones, il est nécessaire de recalculer les probabilités d'inclusion de ces zones afin d'en tenir compte dans l'extrapolation des résultats. Cet exercice étant très complexe, le personnel chargé du traitement des données est souvent amené à faire des simplifications en supposant que les probabilités d'inclusion initiales sont conservées.

Enfin, lorsque sur le terrain des zones sont inaccessibles pour des raisons sécuritaires, et que la collecte n'a pas lieu dans ces zones, il n'est plus possible d'imputer les résultats à l'ensemble du champ géographique de l'enquête, tant que des traitements particuliers ne sont pas réalisés pour compléter les données.

En sus de la situation sécuritaire difficile, le Burkina Faso, à l'image des autres pays du monde, a fait face à la pandémie de la COVID19, ce qui a mis à rude épreuve les opérations

statistiques en utilisant les approches traditionnelles. Et le pays n'est pas à l'abri d'autres situations sanitaires difficiles.

La poursuite de l'utilisation des approches traditionnelles de collecte, d'analyse et de diffusion des données ne sont plus à même de garantir la qualité des données du système statistique national. Il est donc indispensable d'adapter les méthodes en modernisant tout le processus de production et de diffusion de l'information statistique, ainsi que l'utilisation des nouvelles sources des données générées par les technologies de l'information et de la communication.

L'objectif général de cette étude est de renforcer les capacités du SSN à s'adapter aux contextes sécuritaire et sanitaire difficiles, et fournir une production statistique régulière résiliente et adaptée aux besoins.

De façon spécifique, il s'agit de :

- réaliser le diagnostic des méthodes statistiques utilisées dans la mise en œuvre des opérations statistiques d'envergure nationale;
- proposer des stratégies de résilience face aux menaces sécuritaires et sanitaires;
- trouver des solutions de rupture avec les approches traditionnelles;
- mettre en place une stratégie de mise en œuvre des approches innovantes ;
- élaborer un plan d'actions triennal pour opérationnaliser la stratégie.

Le présent document est structuré autour de cinq parties. La première partie dresse le diagnostic de la situation. La seconde aborde les solutions à envisager dans l'immédiat et le court terme pour s'adapter aux contextes sécuritaires et sanitaires difficiles. La troisième partie présente les solutions à moyen et long terme pour rompre avec les approches traditionnelles utilisées présentement par le SSN. La quatrième partie décrit la stratégie de mise en œuvre. Enfin, la dernière partie concerne le plan d'actions permettant de réussir la mise en œuvre des solutions proposées.

I. DIAGNOSTIC DE LA SITUATION

1.1. Les situations rencontrées et initiatives spontanées

1.1.1. Cas des recensements

De façon générale, le SSN réalise plusieurs recensements selon le phénomène étudié. On peut entre autres retenir :

- le recensement général de la population et de l'habitation (RGPH) ;
- le recensement général de l'agriculture (RGA) ;
- le recensement industriel et commercial (RIC) ;
- le recensement du cheptel ;
- l'inventaire national des ouvrages d'eau potable et assainissement (INO-AEP ; INO-AEUE) ;
- l'inventaire forestier national (IFN).

Le recensement, par sa nature, concerne en principe toutes les unités d'observation présentes sur toute l'étendue du territoire national. Toutefois, pour le RGA, la première phase a été couplée avec le RGPH, et les autres phases ont suivi un plan classique d'une enquête par sondage.

La collecte des données sur le terrain se fait par interviews directs dans les ménages, dans les entreprises, ou au niveau des infrastructures. Il arrive aussi que le questionnaire soit déposé auprès des entreprises pour être rempli et récupéré après par un agent de collecte pour le cas du RIC. Aussi, pour le RGA, certaines informations sont mesurées directement sur les parcelles agricoles. La collecte des données se fait de plus en plus avec des smartphones, même s'il arrive souvent que des questionnaires en support papier soient utilisées en cas de problèmes avec le smartphone.

Pour la phase de traitement des données, CSPro est utilisé pour certaines vérifications, et SPSS est utilisé pour l'analyse des données. Le classeur Excel et le langage SQL sont aussi utilisés pour le traitement et l'analyse, surtout pour l'INO-AEP et l'INO-AEUE. Particulièrement pour le RGPH, des redressements de population sont effectués.

Le Tableau 1 présente les différentes difficultés rencontrées lors des recensements et inventaires, ainsi que les initiatives spontanées mises au point lors de ces opérations.

Tableau 1 : Synthèse des difficultés rencontrées et initiatives d'adaptation pour les recensements

	Difficultés rencontrées	Initiatives d'adaptation
Conception	<ul style="list-style-type: none">- Cartographie partielle du territoire national liée au contexte sécuritaire- Dénombrement partiel de certaines ZD- Problème de découpage de certaines ZD suivant les normes- Problème d'estimation des tailles des ZD (surestimé ou sous-estimé)- Changement de la méthode de dénombrement des ZD (toutes les équipes n'ont pas utilisé les mêmes méthodes d'estimation de la population)- Omissions d'ilots dans la cartographie- Repère mal notifié sur la carte- Absence de cartographie des entreprises	<ul style="list-style-type: none">- Création de nouvelles ZD pour les ilots omis lors du dénombrement-

	Difficultés rencontrées	Initiatives d'adaptation
Collecte	<ul style="list-style-type: none"> - Inaccessibilité de certaines ZD (problème sécuritaire) - Pertes de smartphones - Problème de rattachement de certaines entités administratives à leurs ZD dans l'application de collecte - Problème de limite entre entité territoriale - Réticence et refus de certains ménages pour se faire recenser - Indisponibilité de la population en raison des activités agricoles - Non-respect du délai de collecte dans certaines zones - Difficulté de délimitation des ZD sur le terrain par les agents. - Réticence des entreprises à répondre aux questionnaires et à déclarer leurs informations (chiffres d'affaires, nombre d'employés, cotisation sociales, impôt) ; - Problème d'exhaustivité des entreprises recensées ; - délais de réaction parfois longs au niveau des entrepreneurs informels 	<ul style="list-style-type: none"> - Recrutement d'agents recenseurs locaux pour le recensement - Prendre l'estimation pour la localité concernant les ZD partiellement recensées - Redéploiement des agents dans certaines zones - La prolongation de la période de recensement pour boucler la collecte dans les zones en retard - La mise à contribution des autorités coutumières pour dissuader les réticentes et résoudre les différends au sujet des limites territoriales - Collecte sur papier et saisie de données sur smartphone
Traitement	<ul style="list-style-type: none"> - Absence d'un protocole formel de traitement de données - Rapport de traitement de données non suffisamment détaillé - Absence de base unifiée prenant en compte les traitements effectués dans le cadre de la production des indicateurs pour les analyses thématiques - Difficultés d'imputation des chiffres d'affaires des entreprises ; 	<ul style="list-style-type: none"> - Méthode d'estimation des effectifs dans les zones non recensées à partir des images satellitaires (éclairage nocturne, réseau routier, bâti, etc.) - Mobilisation d'un expert dans le cadre d'une assistance technique - Utilisation des images satellitaires

1.1.2. Cas des enquêtes par sondage

Le SSN réalise régulièrement des enquêtes par sondage auprès des populations et des entreprises. Parmi ces opérations, on peut citer :

- l'Enquête Post-Censitaire (EPC) ;
- l'Enquête Régionale Intégrée sur l'Emploi et le Secteur Informel (ERI-ESI) ;
- l'Enquête sur l'Audience des Médias au Burkina Faso (EAM) ;
- l'Enquête Démographique et de Santé (EDS) ;
- l'Enquête Harmonisée sur les Conditions de Vie des Ménages (EHCVM) ;
- l'Enquête COVID ;
- l'Enquête sur les prix ;
- les enquêtes auprès des entreprises (IPI, IPPI, ETC, ICA, etc.) ;
- l'Enquête Permanente Agricole (EPA) ;
- l'Enquête Nutritionnelle Nationale (ENN) ;
- l'Enquête Nationale sur la Consommation du Bois-Energie (ECBE)
- Etc.

La plupart de ces enquêtes utilise un sondage à deux degrés et la base de sondage du RGPH est utilisée au premier degré. Au premier degré, on tire des grappes (ZD ou villages) de façon proportionnelle à leur taille, puis on énumère sur le terrain les grappes sélectionnées avant de tirer les ménages à enquêter. Cependant, pour le cas particulier des enquêtes auprès des entreprises l'échantillonnage est fait par la méthode du seuil d'inclusion. Aussi,

l'échantillonnage des marchés de l'enquête sur les prix, se base sur la méthode du choix raisonné.

Tout comme pour le déroulement des recensements, la collecte des données sur le terrain se fait par interviews directs dans les ménages échantillons, dans les marchés cibles ou dans les entreprises sélectionnées. Aussi, le questionnaire sur les enquêtes auprès des entreprises est souvent déposé pour remplissage. Lors de la collecte des données de l'EPA, certaines informations sont mesurées directement sur les parcelles agricoles. Généralement, la collecte des données des enquêtes par sondage se fait avec des smartphones. Toutefois, il arrive souvent que des questionnaires en support papier soient utilisés en cas de problèmes avec le smartphone.

Le traitement des données des enquêtes par sondage se fait généralement avec des logiciels de traitement statistique tels que SPSS, STATA et R. Le formulaire de collecte mobile est très fréquemment fait avec CPro, mais d'autres plateformes sont parfois utilisées. Les traitements effectués sur les données portent sur :

- le listing d'erreur ;
- la codification ;
- les imputations ;
- la vérification de la cohérence ;
- la correction des valeurs aberrantes ;
- la vérification des doublons.

Le Tableau 2Tableau 1Tableau 1 présente les différentes difficultés rencontrées lors des enquêtes par sondage, ainsi que les initiatives spontanées mises au point lors de ces opérations.

Tableau 2 : Synthèse des difficultés rencontrées et initiatives d'adaptation pour les recensements

	Difficultés rencontrées	Initiatives d'adaptation
Conception	<ul style="list-style-type: none"> - Non prise en compte des communes à fort défis sécuritaire dans la base de sondage - Difficultés d'exploitation de certaines ZD (croquis inexploitable) 	<ul style="list-style-type: none"> - Réduction de la taille de l'échantillon, tirage de ménages - Tirage des communes, puis ZD puis ménages - une analyse de la structure de l'échantillon et des vérifications préliminaires ont été réalisées - Remplacement de ZD inaccessible par une ZD voisine accessible - Augmentation du nombre de ménages échantillonnés dans certaines grappes pour compenser les pertes de grappes - Révision de la taille des ménages - Exclusion de certaines ZD à fort défis sécuritaires de la base de sondage - Reconstitution de la base de sondage

	Difficultés rencontrées	Initiatives d'adaptation
Collecte	<ul style="list-style-type: none"> - Inaccessibilité de certaines ZD en milieu rural au moment de l'enquête due au contexte sécuritaire - La démission d'agents au cours de la collecte ; - Faible possibilité de remplacement d'agents de terrain (les démissionnaires) ; - Refus de certains membres de ménage de se soumettre à l'enquête ; - La non disponibilité de la connexion internet dans certaines localités ; - Dysfonctionnement de certains appareils (tablettes, GPS) ; - Retard dans le paiement des agents de terrain - L'inaccessibilité géographique de certaines localités (du fait des intempéries) - Dysfonctionnement de certains smartphones - Les difficultés de fonctionnement des contacts téléphoniques - La réticence des ménages - 	<ul style="list-style-type: none"> - Création de groupe WhatsApp des enquêteurs et contrôleurs pour échanger sur les difficultés, donner des instructions, suivre l'évolution de la collecte - Un groupe WhatsApp de l'équipe technique qui a permis de résoudre certaines difficultés - La mise à disposition de tablettes supplémentaires par équipe de collecte - La réallocation de certaines lignes budgétaires - L'abandon des grappes inaccessible après le premier remplacement - Le glissement du calendrier du fait de la COVID 19 -
Traitement	<ul style="list-style-type: none"> - Des modules entiers ne sont pas renseignés - Non budgétisation de certaines activités liées au traitement de données - Imputation des loyers fictifs et la valorisation des biens durables 	<ul style="list-style-type: none"> - Vérification et imputations des données collectées sur la base d'une liste d'incohérences harmonisées - Imputation modulaire pour certaines sections - Post stratification en utilisant les effectifs du RGPH pour corriger la sur-représentation du milieu urbain - Réduction de la portée des résultats (résultats valables sauf les milieux rural des provinces inaccessibles)

1.2. Analyse des forces, faiblesses, opportunités et menaces des opérations statistiques

Les opérations statistiques réalisées par le SSN sont soumises à plusieurs contraintes, dues aux contextes interne et externe du système.

1.2.1. Analyse des forces, faiblesses, opportunités et menaces de la phase de conception

L'analyse du contexte interne de la phase de conception des enquêtes fait ressortir des forces, mais également des faiblesses, dont la teneur est résumée dans le Pour la phase de conception, il existe également des opportunités que le système peut saisir pour s'améliorer, mais des menaces dont il faut prévoir un plan de gestion. Ces opportunités et ces menaces sont consignées dans le Tableau 3Tableau 3.

Tableau 3.

Pour la phase de conception, il existe également des opportunités que le système peut saisir pour s'améliorer, mais des menaces dont il faut prévoir un plan de gestion. Ces opportunités et ces menaces sont consignées dans le Tableau 3Tableau 3.

Tableau 3 : Analyse des forces, faiblesses, opportunités et menaces de la phase de conception

<p>Forces</p> <ul style="list-style-type: none"> - Permet l'extrapolation des résultats grâce à l'utilisation des sondages probabilistes - Possibilité d'une mise à jour régulière de la base de sondage - l'énumération des ZD sur le terrain juste avant l'enquête peut améliorer la précision des indicateurs (représentativité statistique) - Le sondage à 2 degrés diminue les coûts 	<p>Faiblesses</p> <ul style="list-style-type: none"> - Vétusté de la base de sondage - Peut exiger une mise à jour de la base de sondage (sectionnement des grandes ZD et assemblage de petites ZD) et engendrer des coûts élevés - Le sondage à 2 degrés déprécie la qualité des indicateurs - Mauvais contrôle de l'échantillon (si enquête en ligne)
<p>Opportunités</p> <ul style="list-style-type: none"> - Existence d'outils pour l'échantillonnage spatial à travers les images satellitaires - Existence de nouvelles sources de données (générées opérateurs de téléphonies, supermarchés, e-commerce) - Informatisation des procédures de l'administration (bases de données des usagers des services...) ou dans le secteur privé - Taux de possession et d'utilisation de téléphones élevé ; - Couverture élevée du territoire en réseau de téléphonie 	<p>Menaces</p> <ul style="list-style-type: none"> - Instabilité de la base de sondage liée aux déplacements des populations - Inaccessibilité d'une grande partie de la base de sondage - Probabilités de tirage des unités primaires perturbées par les déplacements des populations

1.2.2. Analyse des forces, faiblesses, opportunités et menaces de la phase terrain

La phase de collecte de données sur le terrain est la phase la plus impactée par la situation sécuritaire et sanitaire. En effet, l'analyse interne du système fait ressortir plusieurs faiblesses, même s'il existe quelques forces qu'il convient de mentionner. Le Tableau 4 présente les forces, faiblesses, opportunités et menaces de la phase de collecte des données sur le terrain.

Tableau 4 : Analyse des forces, faiblesses, opportunités et menaces de la phase de collecte de données

<p>Forces</p> <p>Interview direct</p> <ul style="list-style-type: none"> - l'enquêteur établit la confiance avec l'enquêté du fait de sa présence physique - Taux de réponse élevé - Plus rapide à mener - Evaluer la compréhension du répondant et son interprétation des questions - possibilité de relance des enquêteurs - Encadrement rapproché <p>Interview au téléphone</p> <ul style="list-style-type: none"> - Écoutes à distance - Economie de temps de déplacements - Possibilité de suivre et d'enregistrer les entretiens - Rapidité d'exécution, surtout pour les enquêtes complémentaires <p>Interview en ligne</p> <ul style="list-style-type: none"> - Questionnaire auto-administré 	<p>Faiblesses</p> <p>Interview direct</p> <ul style="list-style-type: none"> - Influencer intentionnellement ou non des enquêteurs - Les biais de désirabilités sociales - Couteux (financier) - Possibilité de tricheries de la part des agents enquêteurs - La façon de se présenter de l'enquêteur peut engendrer des refus de répondre (qualité du contact, facteur vestimentaire, hostilité envers un groupe social) - Durée de terrain assez longue du fait des déplacements - Influence intentionnelle ou non des enquêteurs - Les biais de désirabilités sociales <p>Interview au téléphone</p> <ul style="list-style-type: none"> - Possibilité de tricheries de la part des agents enquêteurs - La façon de se présenter de l'enquêteur peut engendrer des refus de répondre (qualité du contact, hostilité envers un groupe social) - Dysfonctionnement de certains numéros de téléphones - La méfiance des enquêtés à répondre aux questions - La durée de la communication - Pas de possibilité de supports visuels
--	---

<ul style="list-style-type: none"> - Pas de présence d'enquêteur, donc aucune influence de celui-ci sur l'enquêté - Possibilité d'administrer des questions sensibles - Taux de réponse élevé - Plus rapide à mener - Pas de contrainte de disponibilité des répondants 	<ul style="list-style-type: none"> - Moindre attention des enquêtés (la télé allumée, les enfants qui pleurent, ...) - L'influence de l'enquêteur est réduite par rapport au face à face, mais toujours présente - Taux d'abandon en cours d'interviews plus élevé - « Ras-le-bol » des appels non sollicités <p>Interview en ligne</p> <ul style="list-style-type: none"> - Biais de sélection des enquêtés (niveau d'instruction, manque d'équipements, défaut de connexion internet) - Manque d'assistance pour une meilleure compréhension de certaines questions - Non-réponse partielle plus fréquente, due à l'absence de relance de la part de l'enquêteur - Accès limité aux enquêtés - réticence de réponses aux questions - augmentation du taux de refus - réponses inexactes de la part des enquêtés
<p>Opportunités</p> <ul style="list-style-type: none"> - Développement des TIC (outils de collecte mobile, outils de communication modernes avec les réseaux sociaux, Mail, collecte en ligne, plateformes Web, etc.) - Gestion modernisée des entreprises - Augmentation rapide des supermarchés et des ventes en ligne 	<p>Menaces</p> <ul style="list-style-type: none"> - Couverture du réseau - Défaut de connexion internet par endroit - possibilité d'enlèvement du personnel de terrain - risque de perte - risque de propagation de maladies transmissibles (CORONA VIRUS par exemple) - Fermeture de certaines entreprises - Inaccessibilité de certaines localités

Bien que l'on note des forces des méthodes actuelles de collecte des données, on constate que les contextes sécuritaire et sanitaire entraînent des menaces, surtout l'accessibilité des zones de collecte. Par ailleurs, la façon actuelle de collecter les données possède plusieurs faiblesses qu'il conviendrait de corriger.

1.2.3. Analyse des forces, faiblesses, opportunités et menaces de la phase de traitement des données

La phase de traitement des données est impactée par les problèmes rencontrés lors de la phase d'échantillonnage et/ou de collecte des données. La plus grande faiblesse dans la phase de traitement des données est l'impossibilité de faire une inférence statistique à l'ensemble du pays, dès lors que l'échantillonnage n'a pas pris en compte toutes les ZD, ou que toutes les ZD-échantillons n'ont pas pu être collectées lors de la phase terrain.

Le Tableau 5 présente les forces, faiblesses, opportunités et menaces de la phase de traitement des données des enquêtes.

Tableau 5 : Analyse des forces, faiblesses, opportunités et menaces de la phase de traitement des données

<p>Forces</p> <ul style="list-style-type: none"> - Existence de programmes de traitement pour chaque enquête 	<p>Faiblesses</p> <ul style="list-style-type: none"> - Absence de protocole de traitement - Mauvaise organisation des programmes de traitement de données - Absence d'une procédure d'archivage des bases et des programmes d'apurement - Impossibilité de faire des extrapolations - Réduction de la qualité des estimations
--	---

	- Comparabilité des statistiques (impossibilité de regrouper les modalités d'une même variable)
Opportunités - Existence de nouveaux outils innovants pour l'organisation (stockage et sécurisation) et le traitement des données - Développement de la data-science ; - Disponibilité des mégadonnées (ONI, Téléphonie, Eta civil, Cadastre, ANPE, Examens et concours, Données bancaires, données des impôts, SONABEL, ONEA)	Menaces - Déphasage entre équipements existant et nouvelles technologies ; - Loi de protection des données individuelles - Non organisation des attributions en matière de data science

Pour le traitement des données, il existe d'énormes opportunités que le SSN doit saisir pour se projeter dans l'avenir, et surpasser le problème sécuritaire. Il s'agit de la disponibilité des mégadonnées provenant de l'ONI, des clients des téléphonies mobiles, de l'état civil, du cadastre, de l'ANPE, des examens et concours, des banques, des impôts, de la SONABEL, de l'ONEA, de CAMPUSFASO, ...qui peuvent permettre de récupérer certaines informations à des fins de complément ou de redressement. Cependant, il faut noter que le domaine de la data-science n'est pas suffisamment règlementé au Burkina Faso, et que la loi portant sur la protection des données à caractère personnel pourrait faire entrave à l'exploitation des mégadonnées.

Pour toutes les phases des opérations statistiques, il est nécessaire de consolider les forces constatées, de trouver des solutions aux faiblesses identifiées, de saisir les opportunités qui s'offrent, et de trouver des mesures d'atténuation des effets des menaces qui se posent.

L'ensemble des solutions peut être regroupé en 2 grands blocs : le bloc des solutions qui consolident les forces et corrigent les faiblesses constatées, et le bloc des solutions qui profitent des opportunités offertes et atténuent les effets des menaces identifiées.

II. STRATEGIE DE RESILIENCE

Face aux limites exposées et aux difficultés rencontrées par les praticiens des enquêtes du SSN, il est possible d'envisager des solutions dans l'immédiat pour garantir la qualité des données statistiques produites et permettre une prise de décision en toute confiance. Ces solutions sont présentées selon les différentes phases d'une opération d'enquête.

2.1. Phase de conception

La conception technique des opérations de collecte des enquêtes par sondages doit également tenir compte des contraintes sécuritaires. La plupart des enquêtes par sondage du système statistique national suit des plans de sondages similaires. Ces enquêtes sont des enquêtes stratifiées et à 2 degrés. Les strates sont les 13 régions et les tirages au premier degré se réalisent par milieu de résidence.

Dans chaque région ayant des localités inaccessibles, constituer 4 domaines de tirages : urbain accessible, urbain non accessible, rural accessible, rural non accessible.

La représentativité recherchée étant le niveau national, le milieu de résidence et les régions pour la plupart des enquêtes, il est nécessaire de distinguer les régions n'ayant pas de localités inaccessibles des régions ayant des ZD inaccessibles.

Pour les régions dont toutes les localités sont encore accessibles, deux domaines de tirages peuvent toujours être envisagés à savoir la zone urbaine et la zone rurale comme cela est fait d'habitude. L'allocation des ZD ou grappes entre les milieux de résidence doit être faites de sorte à assurer la représentativité des principaux indicateurs pour l'ensemble du milieu urbain et l'ensemble du milieu rural du Burkina Faso.

Les régions soumises à l'insécurité et dont certaines localités ne sont pas du tout accessibles, il est possible d'envisager quatre domaines d'échantillonnage : urbain accessible, urbain non accessible, rural accessible, rural non accessible. L'allocation des grappes pourrait se faire de façon proportionnelle entre la zone accessible et la zone non accessible dans chacune des régions concernées.

2.2. Phase de collecte

A court terme, la collecte de données se fait par la méthode habituelle d'interview direct face à face dans les strates urbain et rural accessibles. L'on veillera à détecter les Personnes déplacées Internes (PDI) se trouvant dans les ménages de même que leur provenance, à travers des questions supplémentaires.

Pour les PDI vivants dans les sites d'accueil, la collecte pourrait se faire par téléphone.

Si l'on dispose des numéros de téléphone, l'interview est réalisé par un enquêteur au niveau central ;

Pour les zones non accessibles on pourrait recruter sur place et former l'agent enquêteur en ligne pour collecter les données via son smartphone. Au cas où, l'on ne dispose pas d'agent remplissant les conditions pour être enquêteur, on pourrait recruter un facilitateur résident dans la zone et l'enquêteur le guidera pour le choix de la personne à enquêter et réaliser l'interview à travers lui par téléphone.

❖ **Recrutement d'agent local :**

Le recrutement d'agents enquêteurs et facilitateurs consistera à recruter par cooptation des agents locaux qui ont un niveau requis pour mener à bien l'opération. Ces agents doivent résider dans la localité ou avoir une bonne connaissance de la localité où doit se dérouler l'opération. De préférence, l'agent doit vivre régulièrement dans la localité et être disponible pour participer à l'enquête.

La démarche pour choisir ces personnes doit être la plus discrète possible. Pour ce faire, l'accompagnement des autorités administratives et/ou locales des zones en question peut s'avérer nécessaires pour la désignation des enquêteurs / facilitateurs de ces zones spécifiques. De même l'appui d'autres personnes ressources ayant des informations ou des relations particulières dans ces communes ou villages peut être déterminant.

Il faut mettre en place un protocole pour acter la particularité du recrutement dans les zones à défis sécuritaires. Ce qui permettra d'avoir une bonne lisibilité et une transparence dans le recrutement.

❖ **Déplacements des ménages**

Pour certaines opérations dont l'énumération a été faite, on pourrait faire déplacer les ménages tirés des zones à défi sécuritaire vers des zones plus sécurisées.

❖ **Gestion des risques psychologiques**

Un psychologue devrait être recruté pour gérer les cas où les agents sont exposés à des risques. Le psychologue, devrait faire une évaluation psychologique des agents après le travail de collecte.

Tableau 6 : Conditions d'utilisation des méthodes résilientes de collecte de données

Méthode de collecte	Conditions à remplir pour son utilisation
Entretien direct en face à face avec des enquêteurs non-résidents	- Zone totalement sous contrôle de l'administration et des forces de défense et de sécurité
Entretien direct en face à face avec des enquêteurs locaux	- Zone totalement sous contrôle de l'administration et des forces de défense et de sécurité - Zone sous menace terroriste - Existence de personnes ayant le profil requis dans la localité
Déplacement des ménages pour l'entretien	- Zone sous menace terroriste - Possibilité des résidents de sortir et d'entrer dans la zone - Pas de personne dans la zone répondant aux critères pour être enquêteur
Enquête par téléphone	- Existence de réseau de téléphonie mobile dans la zone - Possession des numéros de téléphone des enquêtés

2.3. Phase de traitement

À court terme, deux méthodes de traitement des données peuvent être utilisées :

❖ **Restriction de la portée géographique des résultats**

Cette méthode permet d'extrapoler les résultats dans les zones couvertes seulement. Elle exclut donc les zones inaccessibles et les indicateurs calculés ne sont valides que pour les zones couvertes.

❖ **Post stratification et Repondération**

Elle consiste en une redistribution des poids de sondage et permet d'extrapoler les résultats à l'échelle national ou régional. Elle introduit cependant des biais dans les résultats.

❖ **Imputation par les plus proches voisins**

Au regard de la situation sécuritaire, plusieurs ZD sont inaccessibles. Pour pallier ce problème, les données de ces ZD inaccessibles peuvent être estimées à partir des ZD voisines collectées. Il s'avère nécessaire de mener une étude qui permettra de faire une classification des ZD faisant ressortir les similitudes par indicateur.

2.4. Analyse des forces, faiblesses, menaces et opportunités des stratégies de résilience

Tableau 7 : Forces et faiblesses des stratégies de résilience

Phases	Activités	Forces	Faiblesses
Collecte	Recrutement d'agent local	<ul style="list-style-type: none"> - Maitrise de la zone (langue parlée, us et coutumes) ; - confiance des enquêtés - Continuer à collecter les données dans les zones inaccessibles ; - Pérennité d'avoir des agents enquêteurs dans les zones inaccessibles ; 	<ul style="list-style-type: none"> - Difficulté de collecter les informations sensibles ; - Impossibilité de vérifier l'exactitude des informations collectées
	Déplacements des ménages	<ul style="list-style-type: none"> Obtention des données dont on en a besoin moins de risque pour les agents enquêteurs Continuer à collecter les données dans les zones inaccessibles 	<ul style="list-style-type: none"> - Exposition des ménages à l'insécurité ; - Impossibilité d'avoir les coordonnées GPS ; - Impossibilité de vérifier l'exactitude des informations collectées ; - Coût élevé du déplacement ; - Désignation d'une personne locale afin de mobiliser les ménages pour leur déplacement
	Gestion des risques psychologiques	Prise en charge du personnel de collecte	Accroissement du coût du budget
Traitement	Restriction de la portée géographique des résultats	Indicateurs valides que pour les zones accessibles	Données non extrapolables
	Post stratification et Repondération	Extrapolation des données	Biais probable dans les résultats
	Imputation par les plus proches voisins	Complétude des résultats	<ul style="list-style-type: none"> - Absence d'informations sur les nouvelles caractéristiques de la ZD - Biais probable dans les résultats

Tableau 8 : Opportunités et menaces des stratégies de résilience

Phases	Activités	Opportunités	Menaces
Collecte	Recrutement d'agent local	<ul style="list-style-type: none"> - Existence d'agents utilisés par des ONG - Développer la confiance entre les agents locaux et la structure. 	<ul style="list-style-type: none"> - Favoritisme dans le recrutement ; - Ternissement de l'image de la structure - Manque de compétences requises ;

Phases	Activités	Opportunités	Menaces
		-Bonne image de la structure qui conduit l'opération.	
	Déplacements des ménages	Disponibilité des moyens de transport aérien	Exposition des ménages à l'insécurité
	Gestion des risques psychologiques	Disponibilité permanente pour l'évaluation des risques psychologiques	Néant
Traitement	Restriction de la portée géographique des résultats	Néant	Néant
	Post stratification et Repondération	Néant	Néant
	Imputation par les plus proches voisins	Néant	Néant

III. RUPTURE AVEC LES APPROCHES TRADITIONNELLES

Les méthodes de sondage traditionnelles en statistique s'appuient sur les tailles de population résidente dans les zones enquêtées. Compte tenu du défi sécuritaire et au vu des déplacements de ces populations, la mise à jour de la taille des ZD devient difficile voire impossible sur certaines parties du territoire. D'autres méthodes alternatives doivent être envisagées notamment l'utilisation de Big Data.

3.1. Conception

❖ **Utilisation de la densité de téléphones mobiles par zone pour mettre à jour les tailles des ZD**

L'utilisation des données de téléphonie mobile est une alternative crédible pour contourner l'inaccessibilité des zones lors de la mise à jour de la cartographie. En effet, avec ces données, il n'est plus nécessaire d'aller sur le terrain pour estimer le nombre de personnes dans une ZD.

Etant donné que les croquis des ZD sont numérisés et que les contours de ces zones sont disponibles, il s'agira de superposer ces contours avec les données des réseaux de téléphonie ou des géants du numérique (Google, facebook, etc.) et de les utiliser pour approcher le nombre de personnes dans la zone.

Il est évident que certains tests devront être effectués avant d'appliquer cette méthode. Il est recommandé de réaliser cette superposition pour les données réelles et les données de téléphonie mobile pour plusieurs périodes, d'appliquer des méthodes de « Machine Learning » pour déterminer des coefficients qui permettront d'obtenir la taille de la ZD à partir des caractéristiques de la zone et des données de téléphonie.

Particulièrement pour les zones à fort défis sécuritaire, cet exercice doit être fait pour les périodes où les populations n'avaient pas été déplacées. Pour obtenir les populations réelles de ces zones avant les déplacements, il suffit de considérer les données de mise à jour de la taille des ZD lors des enquêtes par sondage. En effet, la plupart des enquêtes par sondage menées dresse la base de sondage du second degré avant de tirer les ménages échantillons. Ces informations pourront être exploitées, en complément des données du RGPH 5.

❖ **Utilisation des zones de couverture des relais téléphoniques comme unité primaires**

Cette stratégie constitue une rupture totale avec les approches traditionnelles utilisées. En effet, elle suppose l'abandon du concept de ZD comme unité primaire, et sa substitution par une aire dont le centre est un relais téléphonique. L'hypothèse faite ici est que la plupart du temps, les réseaux de téléphonie placent des relais dans les zones où la demande est forte, donc peuplées. Le système statistique pourrait profiter de ces équipements.

Pour l'application de cette stratégie, il est nécessaire de réaliser des études pilotes et calibrer les informations sur les relais. En effet, plusieurs relais peuvent couvrir une même zone, et un travail préliminaire devrait permettre de disposer d'une base de sondage au premier degré sans doublons. Il s'agira de positionner tous les relais de téléphonie mobile et de faire un mapping tenant compte du fait qu'un individu peut disposer de plusieurs numéros de réseaux différents. En outre, au second degré, un recours au numéro d'identification personnel de l'ONI peut être mis à contribution pour résoudre le problème des doublons.

Les résultats de l'étape de mapping constitueront la base de sondage du premier degré des enquêtes auprès des ménages. Un échantillonnage spatial à probabilités égales ou inégales permettra d'obtenir les unités primaires, avec des limites géographiques de ces unités primaires.

❖ **Utilisation des images satellitaires et photographies aériennes pour dresser la base de sondage du second degré**

Disposer d'images satellitaires fiables et à jour permet de repenser la constitution des échantillons pour les enquêtes auprès des populations. En effet, les concessions sont repérables à partir des images satellitaires, et leur repérage peut se faire sans le déplacement d'un agent. Les images de bonne qualité peuvent être obtenues à partir de l'utilisation de drones civils ou militaires.

L'adoption de cette stratégie nécessite un réaménagement ou une combinaison des méthodes statistiques utilisées par le SSN. Pour ne pas avoir à repérer toutes les concessions du pays (mais souhaitable), il est utile de procéder par une méthode alternative pour disposer des unités primaires avant de dresser la base de sondage du second degré avec les images. L'obtention des unités primaires peut être faite en utilisant par exemple la première approche (Utiliser la densité de téléphones mobiles par zone pour mettre à jour les tailles des ZD).

En appliquant cette approche, la conception doit prendre en compte le fait que toutes les concessions repérées à partir des images satellites ne sont pas habitées par des ménages. Il existe des concessions vides, des concessions occupées par des services administratifs, et d'autres par des activités commerciales. Ainsi, la taille finale de l'échantillon (en ménages) n'est pas maîtrisée. Il sera donc judicieux d'adopter un sondage à 3 degrés dont :

- Au premier degré on tire les ZD avec probabilité proportionnelle à la taille, puis on dresse la liste des concessions de chaque ZD-échantillon à partir des images satellites ;
- Au second degré on tire les concessions avec probabilité égale. Le nombre de concessions tirées peut inclure des concessions de remplacement, au cas où des concessions ne sont pas habitées par des ménages.
- Au troisième degré, on tire un ménage dans la concession avec probabilité égale, au cas où la concession est occupée par plusieurs ménages.

Il est important de porter une attention particulière au calcul des poids de chaque ménage de l'échantillon final pour ne pas introduire des biais dans l'estimation des indicateurs.

L'application de ces méthodes de rupture avec les approches traditionnelles doit être obligatoirement précédée d'une phase d'expérimentation pour s'assurer de leur faisabilité technique et financière.

L'énumération des concessions par image satellitaire pouvant être fastidieux, les méthodes de Machine Learning ou Deep Learning peuvent être utilisées pour la reconnaissance automatique des concessions.

Par ailleurs, les coordonnées GPS prises lors du RGPH5 et les données de l'IGB peuvent être exploitées pour faciliter la mise à jour des ZD échantillonnées.

Au-delà de l'identification des concessions, l'utilisation des images satellitaires combinées aux images aériennes peut permettre :

- l'identification des parcelles agricoles ;
- l'estimation des superficies emblavées ;
- l'estimations des rendements agricoles ;

- l'estimation de la faune et de la flore
- l'identification des ouvrages de mobilisation des ressources en eau
- le suivi la déforestation et la dégradation des terres ;
- etc.

❖ **Utilisation des plans d'adressage en milieu pour dresser la base de sondage du second degré**

Cette stratégie constitue à utiliser les numéros des rues et des portes pour constituer une base de sondage soit au premier degré soit au second degré. Le système statistique national pourrait utiliser les données des mairies, de la SONABEL, de l'ONEA, de la POSTE et des CADASTRES pour l'identification des concessions. L'INSD pourrait en collaboration avec les mairies améliorer l'adressage des concessions.

Tableau 9: Analyse des forces et faiblesses des méthodes

Méthode	Force	Faiblesse
Utilisation de la densité de téléphones mobiles par zone pour mettre à jour les tailles des ZD	<ul style="list-style-type: none"> • Constitution d'une base de sondage rapide ; • Améliore la précision des indicateurs ; • Réduction des temps de collecte 	<ul style="list-style-type: none"> • Faible capacité à associer les individus aux numéros de téléphone ; • Biais de couverture ; • Non couverture de l'ensemble du territoire par les réseaux de téléphonies mobiles
Utilisation des zones de couverture des relais téléphoniques comme unité primaires	<ul style="list-style-type: none"> • Constitution d'une base de sondage rapide ; • Améliore la précision des indicateurs ; • Réduction des coûts ; • Réduction des temps de collecte 	<ul style="list-style-type: none"> • Biais de couverture ; • Non couverture de l'ensemble du territoire par les réseaux de téléphonies mobiles ; • Implémentation complexe
Utilisation des images satellitaires pour dresser la base de sondage du second degré	<ul style="list-style-type: none"> • Constitution d'une base de sondage rapide ; • Améliore la précision des indicateurs ; • Réduction des temps de collecte 	<ul style="list-style-type: none"> • Coût élevé ; • Complexité du calcul des probabilité d'inclusion ; • Accès limité aux images de hautes résolutions • La qualité des images peut dépendre des saisons
Utilisation des plans d'adressage en milieu urbain pour dresser la base de sondage du second degré	<ul style="list-style-type: none"> • Constitution d'une base de sondage rapide ; • Faciliter de repérage des concessions et des ménages • Réduction des coûts ; • Réduction des temps de collecte 	<ul style="list-style-type: none"> • Biais de couverture liée à la non mis à jour des adresses ; • La prise en compte des habitations spontanés ; • La non prise en compte du milieu rural ; • L'inexistence d'un fichier unique du cadastre

Tableau 10: Analyse des opportunités et menaces des méthodes

Méthode	Menace	Opportunité
Utilisation de la densité de téléphones mobiles par zone pour mettre à jour les tailles des ZD	<ul style="list-style-type: none"> • Attaque terroriste sur les installations téléphoniques ; 	<ul style="list-style-type: none"> • L'existence de plusieurs compagnies de téléphonie ; • Une forte couverture du territoire ;

Méthode	Menace	Opportunité
	<ul style="list-style-type: none"> Opinion nationale peu favorable à l'utilisation des données personnelles Manque de coopération des compagnies de téléphonies 	<ul style="list-style-type: none"> La modernisation des données de l'ONI ; La modernisation de l'état civil
Utilisation des zones de couverture des relais téléphoniques comme unité primaires	<ul style="list-style-type: none"> Attaque terroriste sur les installations téléphoniques ; Opinion nationale peu favorable à l'utilisation des données personnelles Manque de coopération des compagnies de téléphonies 	<ul style="list-style-type: none"> L'existence de plusieurs compagnies de téléphonie ; Une forte couverture du territoire ; La modernisation des données de l'ONI ; La modernisation de l'état civil
Utilisation des images satellitaires pour dresser la base de sondage du second degré	<ul style="list-style-type: none"> L'environnement juridiques ; 	<ul style="list-style-type: none"> Disponibilité de plusieurs fournisseurs d'images ; Lancement très prochainement du satellite Burkina-Sat1
Utilisation des plans d'adressage en milieu urbain pour dresser la base de sondage du second degré	<ul style="list-style-type: none"> Non coopération des structures détentrices de données 	<ul style="list-style-type: none"> Disponibilité des données des mairies, la SONABEL, l'ONEA, La POSTE ; CADASTRE

3.2. Collecte

Les défis sécuritaires et sanitaires auxquels fait face le Burkina Faso impose de réfléchir à des solutions à moyen et long terme qui vont permettre de poursuivre les opérations de collecte. La collecte des données se fait habituellement par la méthode d'interview direct auprès des sujets (cibles) sur le terrain. Cette approche devient difficile à appliquer sur toute l'étendue du territoire national compte tenu de la situation sécuritaire. A cet effet, il faut explorer d'autres méthodes de collecte en plus de celles déjà utiliser (interview direct face à face ou par téléphone) qui permettent de couvrir toutes les ZD échantillonnées. Cela nécessite de mettre en place un dispositif de collecte performant, sécurisé, opérationnel et adapté.

Les méthodes envisagées pour assurer la collecte dans ces zones sont : les enquêtes par SMS, par *Interactive Voice Response* (IVR), par WhatsApp, en ligne, par téléphone satellitaire et par mail.

❖ Enquête par SMS

La collecte par SMS consisterait à concevoir un questionnaire et à le subdiviser par lots de questions permettant aux enquêtés de répondre facilement. Ainsi, la contrainte de disponibilité de l'enquêté au moment où l'enquêteur veut administrer ses questionnaires est levée. Une relance automatique est nécessaire lorsque le délai de réponse s'épuise.

❖ Enquête par WhatsApp

Elle consiste à faire un appel vidéo ou vocal avec l'enquêté afin d'administrer le questionnaire. Le questionnaire peut être envoyé sous format vocal dans la langue de l'enquêté pour que celui-ci réponde. Dans ce cas des relances peuvent être faites pour améliorer le taux de réponses. Lorsque possible, le questionnaire peut être transmis sous format numérique, rempli par l'enquêté et retourné.

Pour appliquer la méthode de collecte par WhatsApp, une dotation en connexion internet au profit des répondants peut être envisagée.

❖ **Enquête en ligne (CAWI)**

Elle va consister à :

- mettre le questionnaire sur une plateforme de collecte en ligne
- créer un lien pour ces interviews en ligne
- communiquer ce lien (par sms, WhatsApp, email, messenger) au répondant
- lui donner un délai pour répondre ;
- le relancer en cas de retard.

❖ **Enquête par mail**

Elle va consister à :

- disposer des bonnes adresses mail des répondants ou des ménages sélectionnés ;
- s'assurer que le répondant sait lire et écrire en français
- s'assurer que le répondant dispose d'un téléphone Android ou IOS permettant de se connecter à internet ;
- envoyer le questionnaire par mail, le répondant renseigne les réponses et le retourne par mail.

❖ **Enquête par téléphone satellitaire**

Pour combler la couverture des réseaux de téléphonie ordinaires dans les zones à défis sécuritaire, le recours à la collecte des données par téléphone satellitaire peut être nécessaire. Cette méthode consiste à doter les correspondants statistiques en téléphone satellitaire pour la collecte de données administratives, la collecte des données sur les exploitations agricoles, les ouvrages d'eau, etc. Elle vient renforcer les méthodes classiques de collecte de donnée et faciliter l'accès aux zones inaccessibles.

❖ **Interactive Voice Response (IVR)**

L'IVR est une technologie qui permet aux utilisateurs du téléphone d'interagir avec un système téléphonique géré par ordinateur en utilisant la voix et les tonalités DTMF (Dual-tone multi-frequency signaling) saisies sur un clavier. Cette méthode, utilisée par les compagnies de téléphonie, peut être utilisée pour collecter de données auprès des sujets disposant d'un téléphone portable. Le questionnaire doit être court et ne comporter que des questions fermées.

La réussite de ces méthodes de collecte nécessite la mise en place d'une importante stratégie de sensibilisation afin que les populations adhèrent à ces modes de collecte innovants.

Tableau 11: Analyse des forces et faiblesses des méthodes de collecte

Méthode	force	faiblesse
Enquête par SMS	<ul style="list-style-type: none">• Possibilité d'encourager le répondant avec des incitations pour améliorer le taux de réponses• Le répondant peut répondre aux questionnaire selon sa disponibilité Possibilité d'intégrer des questions ouvertes• Réponses honnêtes aux questions sensibles	<ul style="list-style-type: none">• Besoin d'une forte pénétration de la téléphonie mobile• Nécessite un taux d'alphabétisation élevé ;

	<ul style="list-style-type: none"> • Possibilité de comparer les réponses à • taux réponses élevé par rapport aux méthodes par e-mail • Possibilité de traduire le questionnaire dans la langue accessible au répondant 	
Enquête par WhatsApp	<ul style="list-style-type: none"> • Possibilité de relance des enquêteurs • Pas de risques liés aux déplacements des agents enquêteurs ou des enquêtés ; • Rapidité de l'exécution • Faible coût • 	<ul style="list-style-type: none"> • Restriction de l'échantillon aux utilisateurs de téléphone • Impossibilité d'apprécier l'environnement du ménage par observation ; • Forte probabilité de nombreux cas de refus ; • Forte probabilité de nombreux cas de questionnaires partiellement renseignés ; • Interruptions dû à la qualité du réseau • Cas de suspensions d'entretiens du fait de circonstances diverses (entretiens longs et lassants, événements sociaux et contraintes de l'enquêté, etc.) •
Enquête en ligne (CAWI)	<ul style="list-style-type: none"> • Questionnaire auto-administré • Pas de présence d'enquêteur, donc aucune influence de celui-ci sur l'enquêté • Possibilité d'administrer des questions sensibles 	<ul style="list-style-type: none"> • Biais de sélection des enquêtés (niveau d'instruction, manque d'équipements, défaut de connexion internet) • Défaut de connexion internet par endroit • Manque d'assistance pour une meilleure compréhension de certaines questions • Mauvais contrôle de l'échantillon • Non-réponse partielle plus fréquente, due à l'absence de relance de la part de l'enquêteur
Enquête par mail	<ul style="list-style-type: none"> • Réduction des coûts de réalisation des enquêtes ; • Rapidité d'exécution, surtout pour les enquêtes complémentaires 	<ul style="list-style-type: none"> • Couverture du réseau internet • Dysfonctionnement des adresses mail ; • Faible réactivité de la part des entreprises (pas d'accusé de réception) ; • La méfiance des enquêtés à répondre aux questions
Enquête par téléphone satellitaire	<ul style="list-style-type: none"> • Couverture complète du réseau • Prise en compte des zones inaccessibles 	<ul style="list-style-type: none"> • Forte exposition des agents de collecte • Coût élevé
Interactive Voice Response (IVR)	<ul style="list-style-type: none"> • Pas de présence d'enquêteur, donc aucune influence de celui-ci sur l'enquêté • Enquêtes auprès des cibles non lettrées • Possibilité d'administrer des questions sensibles 	<ul style="list-style-type: none"> • Coût élevé • La méfiance/reticence des enquêtés à répondre aux questions

Tableau 12: Analyse des opportunités et menaces des méthodes

Méthode	Menace	Opportunité
---------	--------	-------------

Enquête par SMS	<ul style="list-style-type: none"> • Les attaques terroristes sur les installations téléphoniques 	<ul style="list-style-type: none"> • L'existence de plusieurs compagnies de téléphonie ; • Taux de pénétration du téléphone
Enquête par WhatsApp	<ul style="list-style-type: none"> • Les attaques terroristes sur les installations téléphoniques 	<ul style="list-style-type: none"> • Utilisation de plus en plus élevée des réseaux sociaux
Enquête en ligne (CAWI)	<ul style="list-style-type: none"> • L'insécurité de plus en plus croissante • Présence de cas de Covid-19 	
Enquête par mail	<ul style="list-style-type: none"> • L'insécurité de plus en plus croissante • Insuffisance de ressources financières propres à l'Etat pour les opérations de collecte 	<ul style="list-style-type: none"> • Développement des TIC (outils de collecte mobile, outils de communication modernes avec les réseaux sociaux, Mail, collecte en ligne, plateformes Web, etc.) • Gestion modernisée des entreprises
Enquête par téléphone satellitaire	<ul style="list-style-type: none"> • Contexte législatif peu favorable • 	<ul style="list-style-type: none"> •
Interactive Voice Response (IVR)	<ul style="list-style-type: none"> • Attaques terroristes sur les antennes téléphonique 	<ul style="list-style-type: none"> • Présence de plusieurs compagnies de téléphonies •

3.3. Traitement

Les difficultés de traitement causés par l'inaccessibilité des zones sont essentiellement le biais des pondérations et les non réponses partielles ou totales. Diverses solutions peuvent être envisagées dans le long terme pour y remédier.

❖ **Données des téléphonies et satellitaires pour redresser les pondérations**

Les opérateurs de téléphonie mobile disposent d'une importante base de données sur les usagers. La mise en place d'un système de géolocalisation de ces données permettra de corriger la taille des ZD et les pondérations biaisées par les déplacements de population. Pour ce faire, il est non seulement nécessaire d'acquérir l'expertise dans le domaine mais aussi de signer une convention avec les opérateurs de téléphonie pour disposer des données.

❖ **Mettre en place des modèles économétriques de prédiction des indicateurs**

A long terme, il est primordial de mettre en place des modèles économétriques pour prédire certains indicateurs. Ainsi, à l'aide des données historiques des différentes enquêtes, un modèle économétrique est construit permettant de prédire les valeurs des indicateurs notamment pour les zones inaccessibles.

❖ **Exploitation des données d'état civil**

L'enregistrement de l'état civil fournit une base de données très riche qui peut être exploitée pour le redressement des pondérations et l'estimation de certains indicateurs. L'exploitation statistique des déclarations d'État Civil permet également de recueillir des données démographiques utiles pour l'analyse et le suivi démographique. A long terme, il s'agira de mettre en place un dispositif de centralisation des statistiques d'état civil.

❖ **Exploitation des données satellitaires et de photographie aérienne**

Les données satellitaires permettent de répondre à de nombreux besoins en statistique. Elles fournissent par exemple des informations sur la météorologie, le suivi de parcelles d'agriculture et sur le développement local. A long terme, il s'agira d'acquérir l'expertise dans l'exploitation de ces informations.

❖ **Exploitation des données secondaires**

Certaines informations existent déjà dans l'administration (les données d'identification de l'Office National d'Identification, de la CENI, de l'ANPE, de CAMPUSFASO, les statistiques

scolaires, sanitaires et les déclarations statistiques et fiscales), dans un but différent de celui de l'étude menée et qui sont à disposition pour une seconde utilisation. En cas de difficultés d'accessibilité de certaines zones pour la collecte des données primaires, on pourrait se tourner vers les données secondaires. De plus à partir de ces données, des modèles de prédiction des indicateurs peuvent être développés à moyen et long terme en utilisant les méthodes du big data.

IV. STRATEGIE DE MISE ŒUVRE

La stratégie de mise en œuvre des solutions méthodologiques innovantes repose sur les constats faits sur les faiblesses et les menaces des opérations d'enquêtes et de recensements réalisés par le SSN. La vision est de disposer d'un système de collecte de données :

- efficace ;
- rationnel ;
- pérenne.

Pour y parvenir, la stratégie s'articule autour de trois (03) axes stratégiques qui se déclinent en objectifs stratégiques et en actions.

AXE 1 : Amélioration de la qualité de la production statistique en s'adaptant aux contextes sécuritaire et sanitaire difficiles

L'ambition de cet axe est d'accroître l'efficacité et l'efficience du système de collecte et de traitement, en explorant d'autres voies pour mobiliser le personnel de terrain, en renforçant les capacités des acteurs, et en prenant en compte le contexte sécuritaire lors de la budgétisation des enquêtes

Cinq objectifs stratégiques s'en dégagent :

- prendre en compte le volet sécuritaire dans la budgétisation des opérations de collecte ;
- faciliter la collecte des données de terrain en recrutant des agents de collecte et des guides locaux ;
- améliorer et renforcer les capacités de traitement et d'analyse des données en tenant compte des difficultés rencontrées lors de la conception et de la collecte ;
- assurer le suivi des activités ;
- améliorer la communication sur la démarche qualité, faire le plaidoyer.

Spécifiquement, par rapport à la prise en compte du volet sécuritaire dans la budgétisation des opérations de collecte de données statistiques, elle doit prendre en compte le contexte sécuritaire et sanitaire. De ce fait, elle doit intégrer les rubriques suivantes :

- les frais de déplacement des enquêtés vers des zones sécurisées pour être interviewé
- les frais d'hébergement des enquêtés dans les zones sécurisées pour être interviewé
- les frais de restauration des enquêtés déplacés dans les zones sécurisées pour être interviewé ;
- les frais de guides locaux ;
- les frais de déplacement des agents pour la formation ;
- les frais d'accompagnement des forces de défenses et de sécurités ;
- les primes de risques sécuritaires pour des agents déployés dans les zones à risques ;
- des frais de retrait en cas de paiements électroniques ;
- frais d'un psychologue ;
- frais d'un spécialiste des questions sécuritaires (qui sera chargé de définir les procédures de sécurité à chaque étape de l'enquête, développer ou élaborer une cartographie des risques, situation sécuritaire et la collecte des données dans les ZD, former tous les acteurs de l'enquête sur les risques sécuritaire)

- frais de communication téléphonique (pour appeler les ménages qui se trouvent dans les zones d'insécurité, ne pouvant pas se déplacer vers les zones sécurisées, mais joignable au téléphone)

En plus des problèmes d'inaccessibilité des zones pour mener les entretiens, le même problème se pose pour le paiement des agents qui y travaillent. Ceci a pour conséquences un retard dans le paiement des agents enquêteurs et un rallongement des délais dans le calendrier des activités.

Des procédures plus souples doivent être adoptées par l'administration financière afin de faciliter l'opérationnalisation des activités des enquêtes. Ces procédures doivent s'appuyer sur les nouveaux moyens de paiement électroniques prenant en compte les frais de retraits.

La nouvelle forme de budgétisation doit tenir compte des zones à défi sécuritaire et un assouplissement dans la procédure et le délai des décaissements doivent être adoptés.

Un cahier de charge clair des organes de coordination doit être définis avec des conditions plus stricts relatives aux documents finaux demandés à chaque organe au moment de la conception :

- la budgétisation doit être menée en collaboration avec toutes les directions ou services intervenant dans l'opération du début à la fin ;
- la budgétisation doit tenir compte des implications en terme budgétaire de la particularité des zones à défi sécuritaire ;
- disposer d'un chronogramme de décaissement précis pour éviter des décaissements tardifs ;
- budgétiser toutes les activités nécessaires notamment les activités de traitement et d'analyse ;
- prévoir une prise en charge de tous les organes jugés indispensables dans la mise en œuvre et le suivi : comité technique, comités centraux et déconcentrés de suivi.... ;
- définir un cahier de charge pour chaque organe ou comité et conditionner certains niveaux de paiement au respect strict de ces cahiers de charge ;
- les activités de contrôles et de supervision doivent être budgétisés et prévues dès les premières heures de collecte pour orienter les équipes de terrains afin d'éviter les erreurs de début de collectes ;
- introduire les requêtes dans les délais.

AXE 2 : Modernisation et rationalisation du système de production statistique en tirant profit des avancées technologiques et des récents développements méthodologiques

Cet axe devrait, à long terme, par l'exploitation des mégadonnées et l'utilisation de l'intelligence artificielle, répondre aux questions de collecte des données, de leur traitement et analyse ainsi que de leur diffusion.

Trois objectifs stratégiques en découlent :

- améliorer le cadre réglementaire régissant l'utilisation des nouvelles sources de données ;
- renforcer les capacités des agents en exploitation de mégadonnées et en traitement automatique de données ;

- élaborer et mettre en œuvre une stratégie de sécurisation des personnes et des biens lors des opérations statistiques

La mise en place des stratégies innovantes dans le contexte sécuritaire requiert un renforcement de capacité de l'ensemble des acteurs intervenant dans le processus. Ces acteurs sont entre autres, ceux du SSN, le personnel d'appui de mise en œuvre des opérations statistiques d'envergure, etc. Les thématiques qui peuvent être abordées sont :

- prise en compte du volet sécuritaire dans la budgétisation des opérations statistiques ;
- traitement et imputation des données manquantes sur des localités ou des individus ;
- nouvelles approches méthodologiques envisagées pour la collecte des données en situation d'insécurité ;
- méthodes de traitement de données issues de l'utilisation de méthodes innovantes ;
- méthodes de redressement des pondérations et post-stratification ;
- formation sur les nouvelles sources de données ;
- méthodes de Machine Learning et Deep Learning;
- anonymisation et confidentialité des données ;
- voyages d'études pour l'exploitation des données satellitaires ;
- voyages d'études pour l'utilisation des données téléphoniques dans la production statistique ;
- voyages d'études pour l'exploitation des données du e-commerce/super marché dans la production statistique ;
- élaboration et mise en œuvre d'une stratégie de sécurisation des personnes et des biens lors des opérations statistiques ;
- Programmation et pilotage des drones.

AXE 3 : Assurer l'autonomie et la pérennité de la production statistique en apportant des réponses aux difficultés transversales

Cet axe tente de répondre aux aspects transversaux du système de production statistique que sont les problèmes cruciaux d'accessibilité des données sur le terrain et la garanti du secret statistique. Il s'agit autant de la mise en place d'une stratégie de disposer d'un correspondant statistique à tous les niveaux en raison des coûts élevés des déplacements et de l'impossibilité d'accéder certaines zones, que de disposer des moyens pour garantir le secret statistique lors de l'utilisation des nouvelles sources de données.

Deux objectifs stratégiques sont fondamentalement poursuivis :

- Renforcer le processus de déconcentration / décentralisation du système de collecte de données.
- Renforcer l'infrastructure informatique/Electronique

La plus grande menace étant l'inaccessibilité des zones d'enquête, un travail de maillage du territoire national permettra de surpasser cette difficulté. Pour cela, il est urgent de penser à mettre en place un réseau de correspondants statistiques jusqu'au niveau le plus bas possible.

Pour la mise en place de ce réseau, le SSN peut se baser sur deux approches :

- ❖ La première approche consiste à exploiter le canal du Ministère en charge de la décentralisation : chaque région choisie un correspondant statistique qui répond directement de l'INSD. Chaque commune nomme un correspondant qui répond du correspondant régional. Chaque correspondant communal est chargé de mettre en place son réseau de correspondants dans chaque village de la commune. Chaque correspondant du village est chargé de tenir à jour un fichier contenant la liste des concessions du village, les coordonnées géographiques de chaque concession, le nom et prénom d'un répondant dans la concession, et les numéros de téléphone de ce répondant.
- ❖ La seconde approche consiste à s'appuyer sur les directions régionales de l'INSD. Chaque direction régionale met en place son réseau de correspondants dans chaque commune de son ressort territorial, et chaque correspondant communal est chargé de mettre en place son réseau comme dans le cas précédent, et ainsi de suite jusqu'au niveau de chaque concession.

En mettant en place ce dispositif, on contourne toutes les difficultés de conception et de collecte préalablement identifiées. En effet, le niveau central disposera de la base de sondage de toutes les concessions du pays avec leurs coordonnées géographiques et les numéros de téléphone.

Il ne sera plus nécessaire de réaliser un sondage à deux degrés, dont le but principal est de réduire les distances à parcourir, puisqu'il n'y aura plus de distance à parcourir. On peut alors réaliser un tirage aléatoire simple selon des strates bien précis à définir. Il va de soi également que le mode de collecte à adopter dans ce cas soit une enquête au téléphone et/ou en ligne et/ou WhatsApp et/ou par e-mail selon la préférence et les contraintes des enquêtés.

Pour ce qui est du renforcement de l'infrastructure informatique, Il s'agira d'acquérir des équipements informatiques de hautes performances (centre d'appel, data center, ordinateur, etc.) pour le stockage et le traitement des données volumineuses.

Le passage à l'utilisation de méthodes innovantes demande des infrastructures informatiques de qualité. En effet, pour des enquêtes en ligne, il est nécessaire de disposer d'un centre d'appel et des équipements de grande capacité de stockage des fichiers audios. Aussi, l'exploitation des mégadonnées requiert des équipements informatiques de dernière génération qu'il convient d'acquérir. En tout état de cause, une étude de faisabilité est nécessaire pour spécifier le type de matériel à acquérir et leurs caractéristiques techniques.

Les approches méthodologiques innovantes requièrent des ressources humaines de qualité et compétentes adaptées au contexte actuel. Les différents aspects intégrés dans les approches traditionnelles demande une augmentation des ressources financières pour mieux adapter les budgets.

Les différentes actions d'innovation doivent être conçues sous forme de projet et soumis à un plaidoyer auprès des bailleurs du SSN pour financement.

V. PLAN D' ACTIONS

Le plan d'actions élaboré pour opérationnaliser les propositions faites s'articule autour des axes et objectifs stratégiques déclinés dans la stratégie de mise en œuvre.

Le Tableau 13 donne une synthèse des coûts du plan d'actions sur trois ans.

Tableau 13: Résumé des coûts du plan d'actions

Axe stratégique	2023	2024	2025	Total
Axe 1 : Amélioration de la qualité de la production statistique en s'adaptant aux contextes sécuritaire et sanitaire difficiles	461 000 000	336 000 000	321 000 000	1 118 000 000
Axe 2 : Modernisation et rationalisation du système de production statistique en tirant profit des avancées technologiques et des récents développements méthodologiques	455 000 000	425 000 000	295 000 000	1 175 000 000
Axe 3 : Assurer l'autonomie et la pérennité de la production statistique en apportant des réponses aux difficultés transversales	950 000 000	750 000 000	855 000 000	2 555 000 000
TOTAL GENERAL	1 866 000 000	1 511 000 000	1 471 000 000	4 848 000 000

D'un montant global d'environ **cinq milliards FCFA**, ce plan pourrait examiner, lors de la mise en œuvre, l'opportunité de partager les coûts avec les autres utilisateurs des mégadonnées comme l'IGB, METEO BURKINA, etc.

Tableau 14: Plan d'actions budgétisé

Objectifs stratégiques	Actions	Responsable	Structures associées	Années			
				2023	2024	2025	Total
Axe 1 : Amélioration de la qualité de la production statistique en s'adaptant aux contextes sécuritaire et sanitaire difficiles				461 000 000	336 000 000	321 000 000	1 118 000 000
1-1 Prendre en compte le volet sécuritaire dans la budgétisation des opérations de collecte	Elaborer un canevas de budgétisation des opérations prenant en compte le contexte sécuritaire	DG	DAF	10 000 000			10 000 000
	Faire adopter le canevas par les instances compétentes	DG	DAF	10 000 000			10 000 000
	Vulgariser le canevas de budgétisation des opérations prenant en compte le volet sécuritaire	DG	DAF	30 000 000	15 000 000		45 000 000
1-2 Faciliter la collecte des données de terrain en recrutant des agents de collecte et des guides locaux	Elaborer un protocole de recrutement des agents locaux	DRH		10 000 000			10 000 000
	Identifier les zones où il y a nécessité de mobiliser des agents locaux	DSCVM		20 000 000			20 000 000
	Recruter les agents locaux pour la collecte des données de terrain	DRH	DSCVM/DD/DSSE	15 000 000	5 000 000	5 000 000	25 000 000
	Mobiliser un psychologue pour prendre en charge les risques psychologiques	DRH		10 000 000	10 000 000	10 000 000	30 000 000
1-3 Améliorer et renforcer les capacités de traitement et d'analyse des données en	Organiser des sessions de formation des producteurs sur la prise en compte du volet sécuritaire et sanitaire dans la budgétisation des opérations statistiques	DAF		30 000 000	30 000 000	30 000 000	90 000 000

Objectifs stratégiques	Actions	Responsable	Structures associées	Années			
				2023	2024	2025	Total
tenant compte des difficultés rencontrées lors de la conception et de la collecte	Former les producteurs sur les techniques de collecte et de traitement des données des enquêtes au téléphone	DCSFR	DSCVM/DD/DSSE	30 000 000	30 000 000	30 000 000	90 000 000
	Organiser des sessions de formation des producteurs sur les méthodes de traitement et d'imputation des données manquantes sur des localités ou des individus	DCSFR		30 000 000	30 000 000	30 000 000	90 000 000
	Organiser des sessions de formation des producteurs sur les nouvelles approches méthodologiques envisagées pour la collecte des données en situation d'insécurité	DCSFR		30 000 000	30 000 000	30 000 000	90 000 000
	Organiser des sessions de formation des producteurs sur les méthodes de traitement de données issues de l'utilisation de méthodes innovantes.	DCSFR		30 000 000	30 000 000	30 000 000	90 000 000
	Former le personnel au protocole de sécurité/Sanitaire	DCSFR		45 000 000	45 000 000	45 000 000	135 000 000
	Organiser des sessions de formation des producteurs sur les méthodes de redressement des pondérations et post-stratification ;	DCSFR		30 000 000	30 000 000	30 000 000	90 000 000
	1-4 Assurer le suivi des activités	Evaluer les procédures techniques de mise en œuvre lors de la réalisation des opérations statistiques (ayant utilisé des approches innovantes ou non)	DCSFR		30 000 000	30 000 000	30 000 000

Objectifs stratégiques	Actions	Responsable	Structures associées	Années			
				2023	2024	2025	Total
	Mettre en place un comité de suivi pour des contrôles réguliers des procédures (statisticiens, universitaires, chercheurs, etc.)	DG		10 000 000	10 000 000	10 000 000	30 000 000
1-5 Améliorer la communication sur la démarche qualité, faire le plaidoyer	Organiser des Sessions de Communication sur l'adaptation des méthodologies statistiques aux contextes sécuritaires et sanitaires à l'endroit des directions techniques de l'INSD, des structures du SSN	DCSFR		12 000 000	12 000 000	12 000 000	36 000 000
	Communiquer aux utilisateurs les méthodologies statistiques adaptées aux contextes sécuritaires et sanitaires	DG		20 000 000	15 000 000	15 000 000	50 000 000
	Organiser des formations sur la démarche qualité au profit des acteurs du SSN	DCSFR		14 000 000	14 000 000	14 000 000	42 000 000
	Elaborer un document de plaidoyer pour la mobilisation des ressources	DG		15 000 000			15 000 000
	Organiser une conférence avec les partenaires (PTF).	DG		30 000 000			30 000 000
	Axe 2 : Modernisation et rationalisation du système de production statistique en tirant profit des avancées technologiques et des récents développements méthodologiques				455 000 000	425 000 000	295 000 000
2-1 Améliorer le cadre réglementaire régissant	Etablir des protocoles de partage et d'utilisation des données avec les fournisseurs de mégadonnées	DCSFR		10 000 000	5 000 000		15 000 000

Objectifs stratégiques	Actions	Responsable	Structures associées	Années			
				2023	2024	2025	Total
l'utilisation des nouvelles sources de données	Rédiger les textes d'application de la nouvelle loi statistique, spécifiques aux nouvelles sources de données (associé les producteurs, fournisseurs de données, commission de l'informatiques et des libertés, etc.)	DCSFR		15 000 000	5 000 000		20 000 000
	Sensibiliser les acteurs (entreprises privées, multinationales, etc.) sur l'utilisations des nouvelles sources de données.	DG		5 000 000	5 000 000	5 000 000	15 000 000
	Etablir une convention d'utilisation des drones avec le ministère en charge de la sécurité	DG		20 000 000			20 000 000
2-2 Renforcer les capacités des agents en exploitation de mégadonnées et en traitement automatique de données	Organiser des sessions de formation des producteurs sur les nouvelles sources de données (mégadonnées)	DCSFR		30 000 000	30 000 000	30 000 000	90 000 000
	Organiser des sessions de formation des producteurs sur les méthodes de machine learning avec le logiciel R/python	DCSFR		30 000 000	30 000 000	30 000 000	90 000 000
	Organiser des sessions de formation des producteurs sur les méthodes de machine learning avec le logiciel SPSS/STATA	DCSFR		30 000 000	30 000 000	30 000 000	90 000 000
	Organiser des sessions de formation des producteurs sur l'anonymisation des données et la protection des données à caractère individuel	DCSFR	DIMIS	30 000 000	30 000 000	30 000 000	90 000 000

Objectifs stratégiques	Actions	Responsable	Structures associées	Années			
				2023	2024	2025	Total
	Organiser des voyages d'études pour l'utilisation des données téléphoniques dans la production statistique	DCSFR		40 000 000	40 000 000	40 000 000	120 000 000
	Organiser des voyages d'études pour l'exploitation des données du e-commerce / Super marché dans la production statistique	DCSFR		40 000 000	40 000 000	40 000 000	120 000 000
	Formation en Deep Learning	DCSFR		30 000 000	30 000 000	30 000 000	90 000 000
	Organiser des voyages d'études pour exploitation des données satellitaires	DCSFR		40 000 000	40 000 000	40 000 000	120 000 000
	Former des opérateurs sur le pilotage des drones civils	DIMIS		20 000 000	60 000 000		80 000 000
	Former les agents sur la programmation des drones	DIMIS		20 000 000	60 000 000		80 000 000
	Former des agents sur l'administration d'un data center	DIMIS		40 000 000			40 000 000
	2-3 Elaborer et mettre en œuvre une stratégie de sécurisation des personnes et des biens lors des opérations statistiques	Elaborer une stratégie de sécurité	DG		15 000 000		
Mobiliser un officier de sécurité		DG	DRH	10 000 000	10 000 000	10 000 000	30 000 000
Elaborer et mettre à jour une cartographie de risques d'insécurité		DG	DSCVM	30 000 000	10 000 000	10 000 000	50 000 000
Signer un protocole entre l'INSD et l'agence de renseignement		DG					-

Objectifs stratégiques	Actions	Responsable	Structures associées	Années			
				2023	2024	2025	Total
Axe 3 : Assurer l'autonomie et la pérennité de la production statistique en apportant des réponses aux difficultés transversales				950 000 000	750 000 000	855 000 000	2 555 000 000
3-1 Renforcer le processus de déconcentration / décentralisation du système de collecte de données	Tester la précision des approches innovantes par rapport aux approches traditionnelles dans des zones d'études de petites surfaces	DCSFR	DSCVM/Directions régionales de l'INSD	50 000 000	50 000 000	50 000 000	150 000 000
	Elaborer une stratégie pour la désignation des correspondants statistiques (tenir compte des sensibilités au niveau communal)	DCSFR	Directions régionales de l'INSD	20000000			20 000 000
	Mettre en place des correspondants statistiques au niveau décentralisé (régional, provincial, communal, village/secteur ou quartier)	DCSFR	Directions régionales de l'INSD	10 000 000	20 000 000	90 000 000	120 000 000
3-2 Renforcer l'infrastructure informatique/Electronique	Développer une plateforme informatique et d'une application de collecte mobile pour les besoins de collecte d'information de l'ensemble du SSN	DIMIS	DCSFR	200 000 000	100 000 000		300 000 000
	Acquérir des ordinateurs portables adaptés au traitement des données volumineuses (CPU, GPU, RAM, stockage SSD)	DIMIS	DAF	20 000 000	20 000 000	20 000 000	60 000 000
	Acheter des licences des logiciels et services informatiques et statistiques	DIMIS		100 000 000	100 000 000	100 000 000	300 000 000
	Acquérir des drones civils	DG	DAF/DIMIS		30 000 000	165 000 000	195 000 000
	Acquérir des téléphones satellitaires	DG	DAF/DIMIS		10 000 000	30 000 000	40 000 000

Objectifs stratégiques	Actions	Responsable	Structures associées	Années			
				2023	2024	2025	Total
	Acquérir des images satellitaires	DG	DAF/DIMIS	300 000 000	300 000 000	300 000 000	900 000 000
	Contribuer à la mise en service d'un satellite prenant en compte les besoins en information du SSN	DG	DAF/DIMIS				-
	Mettre en place un centre d'appel à l'INSD	DIMIS		150 000 000	20 000 000		170 000 000
	Mettre en place un data center	DIMIS					-
	Mettre en place un centre d'appel dans d'autre centre (MAAH, ISSP, etc.)	MIN/AGRIC		100 000 000	100 000 000	100 000 000	300 000 000
TOTAL GENERAL				1 866 000 000	1 511 000 000	1 471 000 000	4 848 000 000

CONCLUSION ET RECOMMANDATIONS

CONCLUSION

Le présent document de travail méthodologique relève des procédures d'adaptation ou innovantes qui permettront de mener à bien toutes les activités inhérentes aux opérations statistiques dans un contexte de crises sécuritaire et sanitaire difficiles. L'analyse des différentes approches méthodologique et de collecte initiales adoptées dans les opérations statistiques relève qu'elles introduisait un biais dans les indicateurs.

Ce rapport se veut un document de recherche permettant d'explorer de nouvelles approches méthodologiques adaptées au contexte sécuritaire et permet d'améliorer les approches traditionnelles d'échantillonnage, collecte et de traitement de données.

Ces approches méthodologiques se veulent innovantes dans la mesure où elle intègre les technologies de l'information et de la communication dans la sensibilisation, la collecte, le transfert et la diffusion des données et explore l'univers des mégadonnées et de l'intelligence artificielle.

Pour réussir la mise en œuvre de ces approches innovantes par le système statistique national, des stratégies ont été développées. Il s'agit de garantir l'efficacité du système de production statistique, son efficience, et sa pérennité.

RECOMMANDATIONS

Pour une meilleure mise en œuvre des approches méthodologiques innovantes adaptées au contexte sécuritaire et sanitaire, les recommandations suivantes ont été formulées :

- Faire un plaidoyer auprès des partenaires techniques et financier pour faciliter la mobilisation des ressources financières et humaines pour faciliter la mise en œuvre des stratégies ;
- Présenter les résultats de l'étude à l'ensemble des structures du système statistique national ;
- Faire la promotion de ce document pour permettre aux utilisateurs et aux producteurs de statistique de prendre connaissance des résultats de l'étude et de les appliquer de façon appropriée ;
- Mettre en place un comité de suivi des opérations statistiques utilisant ces approches innovantes ;
- Renforcer les capacités des acteurs du SSN sur l'utilisation de ces approches qui permettent d'améliorer la production statistique.
- Engager un officier en sécurité pour des conseils de sécurité et notamment assuré la sécurité des personnes en période d'insécurité

BIBLIOGRAPHIE

1. Rapport général RGPH 2019
2. Rapport final ERI-ESI
3. Rapport final EPC
4. Rapport final EHCVM
5. Rapport final EAM
6. Rapport final Enquête COVID
7. Rapport final EPA
8. Rapport final IHPI, IPPI

ANNEXES

Annexe 1: Diagnostic du RGPH 2019

Phase	Difficultés	Initiatives
Echantillonnage	<ul style="list-style-type: none"> - Cartographie partielle du territoire national liée au contexte sécuritaire (111 villages non couverts en milieu rural) - Inaccessibilités de certaines ZD au moment du recensement - Dénombrement partiel de certaines ZD - Problème de découpage de certaines ZD suivant les normes - Problème d'estimation des tailles des ZD (surestimé ou sous-estimé) - Changement de la méthode de dénombrement des ZD (toutes les équipes n'ont pas utilisé les mêmes méthodes d'estimation de la population) - Omissions d'ilots dans la cartographie - Repère mal notifié sur la carte 	<ul style="list-style-type: none"> - Création de nouvelles ZD pour les ilots omis lors du dénombrement
Collecte	<ul style="list-style-type: none"> - Non-respect du protocole de recrutement dans certaines localités ; - Insuffisance de résultats chez certains agents (faible niveau des agents de collecte, non-respect des échéances de paiement, etc.) - Non-respect du cahier de charges de certains contrôleurs et superviseurs - la taille élevée de quelques ZD, - Dysfonctionnement de certains comités de suivi du RGPH5 (retard de disponibilité des ressources, insuffisances des ressources allouées aux comités, non allocation de ressources aux comités villageois) - Ressources non allouées aux communes au suivi du recensement - Problème de rattachement de certaines entités administratives à leurs ZD dans l'application de collecte - Problème de limite entre entité territoriale - Réticence et refus de certains ménages pour se faire recenser - Indisponibilité de la population en raison des activités agricoles - Difficultés techniques liées à l'application et aux appareils de collecte - Erreurs d'affectation des ZD - Dysfonctionnement de la cellule de veille - Insuffisance du matériel de collecte (smartphones) - Inaccessibilité de certaines ZD (problème sécuritaire) - Dysfonctionnement des appareils de collecte, - Non-respect du délai de collecte dans certaines zones - Difficulté de délimitation des ZD sur le terrain par les agents. 	<ul style="list-style-type: none"> - Recrutement d'agents recenseurs locaux pour le recensement - Prendre l'estimation pour la localité concernant les ZD partiellement recensées - Redéploiement des agents dans certaines zones - La prolongation de la période de recensement pour boucler la collecte dans les zones en retard - La mise à contribution des autorités coutumières pour dissuader les réticentes et résoudre les différends au sujet des limites territoriales - Collecte sur papier et saisie de données sur smartphone
Traitement	<ul style="list-style-type: none"> - Non budgétisation de certaines activités liées au traitement de données - Absence d'un protocole formel de traitement de données - Rapport de traitement de données non suffisamment détaillé - Absence de base unifiée prenant en compte les traitements effectués dans le cadre de la production des indicateurs pour les analyses thématiques 	<ul style="list-style-type: none"> - Méthode d'estimation des effectifs dans les zones non recensées à partir des images satellitaires (éclairage nocturne, réseau routier, bâti, etc.) - Mobilisation d'un expert dans le cadre d'une assistance technique

Annexe 2 : Diagnostic de l'EPC

Phase	Difficultés	Initiatives
Echantillonnage	<ul style="list-style-type: none"> - Non prise en compte des communes à fort défis sécuritaire 	<ul style="list-style-type: none"> - Adaptation de la méthodologie au contextes sécuritaire, sanitaire et à la situation d'urgence (la réduction de la taille de l'échantillon, tirage de ménages dans le plan d'échantillonnage, appariement et collecte simultanée) - Tirage des communes, puis ZD puis ménages
Collecte	<ul style="list-style-type: none"> - Non-concordance des limites de certaines ZD à l'EPC avec celles du RGPH - Disparition des numérotations RGPH sur les murs due aux intempéries - Inaccessibilité de certains ZD pour des raisons d'insécurité - Modifications du questionnaire EPC, ce qui a rendu la programmation difficile (conception d'une application de collecte et d'appariement) - Dysfonctionnement de l'application de collecte sur certains smartphones - Impossibilité de rechercher des individus et des ménages hors de la commune concernée - Réticence voire refus de certains ménages à accueillir les agents enquêteurs 	<ul style="list-style-type: none"> - Appariement complémentaire au bureau - Prise en compte des limites identifiées à l'EPC comme limite de référence et appariement élargi au-delà de la ZD
Traitement	<ul style="list-style-type: none"> - Absence d'évaluation de contenu (exactitude des réponses) 	

Annexe 3: Diagnostic de l'ERI-ESI

Phase	Difficultés	Initiatives
Echantillonnage	<ul style="list-style-type: none"> - Quatorze (14) communes dont neuf (09) de la région du Sahel et cinq (05) de la région du Nord ont été exclues de la base de sondage pour des raisons de sécurité 	<ul style="list-style-type: none"> - une analyse de la structure de l'échantillon et des vérifications préliminaires ont été réalisées
Collecte	<ul style="list-style-type: none"> - Indisponibilité des répondants pour la phase 2 - Problème de saut dans les retours sur questions précédentes (volet emploi) - Les codes des activités inhérentes aux UPI - Problèmes de transfert des UPI - Couverture partielle de certaines ZD - Problème d'exhaustivité dans la réception des UPI transférées - Omission en début de collecte des UPI de commerce 	<ul style="list-style-type: none"> - Assistance technique au profit des superviseurs
Traitement	<ul style="list-style-type: none"> - Des modules entiers dans la partie informelle n'étaient pas renseignés (module PV, DC) dans certaines UPI - Non budgétisation de certaines activités liées au traitement de données 	<ul style="list-style-type: none"> - Recodification de certaines activités - Des programmes de tabulation standard ont été développés par AFRISTAT et transmis aux équipes pays pour la production des tableaux des rapports du volet emploi, secteur informel et gouvernance, paix et sécurité - Le plan de tabulation et les syntaxes élaborés par AFRISTAT ont été utilisés

Phase	Difficultés	Initiatives
		<ul style="list-style-type: none"> - vérification et imputations des données collectées sur la base d'une liste d'incohérences harmonisées - Imputation modulaire pour certaine section - L'utilisation des programmes de contrôle de cohérence a permis de détecter et de procéder à des imputations après

Annexe 4 : Diagnostic de EAM

Phase	Difficultés	Initiatives
Echantillonnage	<ul style="list-style-type: none"> - Inaccessibilité de certaines ZD en milieu rural au moment de l'enquête due au contexte sécuritaire - Difficultés d'exploitation de certaines ZD (croquis inexploitable) 	<ul style="list-style-type: none"> - Les ZD inaccessibles de 3 provinces n'ont pas été enquêtées (pas de remplacement) ; - Remplacement par la ZD voisine
Collecte	<ul style="list-style-type: none"> - Inaccessibilité de certaines ZD ; - Insuffisance dans le fonctionnement du Comité technique (pas de régularité dans la tenue des réunions) - Insuffisance des ressources de l'étude - Difficultés techniques liées aux appareils de collecte 	<ul style="list-style-type: none"> - Création de groupe WhatsApp des enquêteurs et contrôleurs pour échanger sur les difficultés, donner des instructions, suivre l'évolution de la collecte - Un groupe WhatsApp de l'équipe technique qui a permis de résoudre certaines difficultés - La mise à disposition de tablette supplémentaire - La réallocation de certaines lignes budgétaires
Traitement	<ul style="list-style-type: none"> - Insuffisance de ressources allouée au traitement de données 	<ul style="list-style-type: none"> - Post stratification en utilisant les effectifs du RGPH pour corriger la sur-représentation du milieu urbain - L'existence d'un plan d'analyse - Post-stratification pour l'extrapolation des résultats - Réduction de la portée des résultats (résultats valables sauf les milieux rural des 3 provinces Soum, Yagha, Oudalan

Annexe 5 : Diagnostic de EDSBF V

Phase	Méthodologie	Difficultés	Initiatives
Echantillonnage	<ul style="list-style-type: none"> - Échantillon aléatoire stratifié à 2 degrés - Au premier degré, 600 grappes ou zones de dénombrement ont été tirées et au second degré, un échantillon de 26 ménages sont sélectionné dans chacune des grappes tirées ; (15 600 ménages avec environ 17 800 femmes et 7 600 hommes) 	Grappes dans des zones à défis sécuritaires	<ul style="list-style-type: none"> - Remplacement de certaine grappe par un tirage complémentaire - Augmentation du nombre de ménage échantillonnés dans certaines grappes pour compenser les pertes de grappes - Révision de la taille des ménages

Phase	Méthodologie	Difficultés	Initiatives
	-Enumération des grappes sélectionnées		
Collecte	- Adaptation des questionnaires au contexte national - Interviews directs sur le terrain dans les ménages	- L'inaccessibilité de certaines grappes pour raison d'insécurité ; - La démission d'agents au cours de la collecte ; - Faible possibilité de remplacement d'agents de terrain (les démissionnaires) ; - Refus de certains membres de ménage de se soumettre à l'enquête ; - La non disponibilité de la connexion internet dans certaines localités ; - Dysfonctionnement de certains appareils (tablettes, GPS) ; - Retard dans le paiement des contrôleurs de la collecte.	- Abandon des grappes inaccessible après le premier remplacement (58 n'ont pas été couvertes) - Glissement du calendrier du fait de la COVID 19 -
Traitement	- Edition des données pendant la collecte et correction après retour-terrain - Imputation - Edition des données après la collecte, imputation - Traitement globale de la base faite par l'équipe de ICF	- Application d'édition non finalisé à temps - Non disponibilité de certains membres de l'équipe d'édition après la collecte	- Renforcement de l'équipe d'édition - Prolongation de la période d'édition

Annexe 6 : Diagnostic de EHCVM 1

Phase	Méthodologie	Difficultés	Initiatives
Echantillonnage	EHCVM1 - Au premier degré, 585 zones de dénombrement sont tirées avec des probabilités proportionnelles à la taille de la population issue du fichier des ZD du RGPH-2006 mise à jour - Au second degré, un échantillon de 12 ménages est tiré à probabilité égale et de façon systématique dans chacune des zones de dénombrement - la taille totale de l'échantillon a été estimée à 7020 ménages - L'échantillon de 7020 a été enquêté en deux vagues de 3510 ménages	- Grappes dans des zones à défis sécuritaires	- Exclusion de certaines ZD à fort défis sécuritaires de la base de sondage - Reconstitution de la base de sondage
Collecte	- Interviews directs sur le terrain dans les ménages	- L'inaccessibilité de certaines ZD pour raison d'insécurité ;	- Appels téléphoniques de certains ménages - Dotation supplémentaire en smartphone

Phase	Méthodologie	Difficultés	Initiatives
		<ul style="list-style-type: none"> - L'inaccessibilité géographique de certaines localités (du fait des intempéries) - Dysfonctionnement de certains smartphones - 	<ul style="list-style-type: none"> - Report de la collecte dans certaines zones à la seconde vague - Déplacement de certains ménages - Remplacement de certains ménages - Collecte sur papier et saisie de données sur smartphone
Traitement	<ul style="list-style-type: none"> - Apurement des données pour corriger les valeurs aberrantes - Imputation des valeurs manquantes - Correction des pondérations - Utilisation des loyers déclarés pour estimer les loyers fictifs 	<ul style="list-style-type: none"> - Imputation des loyers fictifs et la valorisation des biens durables 	<ul style="list-style-type: none"> - Existence d'un protocole de traitement de données

Annexe 7 : Diagnostic de EHCVM-2

Phase	Méthodologie	Difficultés	Initiatives
Echantillonnage	<ul style="list-style-type: none"> - Au premier degré, 600 zones de dénombrement non panel (transversale) sont tirées avec des probabilités proportionnelles à la taille de la population issue du fichier des ZD du RGPH-2006 mise à jour. Aussi, 585 ZD panels ont été dénombrées - 6 ménages panels par ZD panel et 12 ménages par ZD transversale - la taille totale de l'échantillon a été estimée à 10 705 ménages dont 7200 ménages non panels et 3505 ménages panels 	<ul style="list-style-type: none"> - Grappes dans des zones à défis sécuritaires - 	<ul style="list-style-type: none"> - Exclusion de certaines ZD à fort défis sécuritaires de la base de sondage - Reconstitution de la base de sondage
Collecte	<ul style="list-style-type: none"> - Interviews directs sur le terrain dans les ménages 	<ul style="list-style-type: none"> - L'inaccessibilité de certaines ZD pour raison d'insécurité ; - L'inaccessibilité géographique de certaines localités (du fait des intempéries) - Dysfonctionnement de certains smartphones - 	<ul style="list-style-type: none"> - Appels téléphoniques de certains ménages - Dotation supplémentaire en smartphone - Report de la collecte dans certaines zones à la seconde vague - Déplacement de certains ménages - Remplacement de certains ménages - Collecte sur papier et saisie de données sur smartphone
Traitement	<ul style="list-style-type: none"> - Apurement des données pour corriger les valeurs aberrantes - Imputation des valeurs manquantes - Correction des pondérations - Utilisation des loyers déclarés pour estimer les loyers fictifs 	<ul style="list-style-type: none"> - Imputation des loyers fictifs et la valorisation des biens durables 	<ul style="list-style-type: none"> - Existence d'un protocole de traitement de données

Annexe 8 : Diagnostic de l'enquête COVID

Phase	Méthodologie	Difficultés	Initiatives
Echantillonnage	- 2500 ménages ont été tiré sur la base des 7 010 ménages issus de la base de sondage de EHCVM 2018 - Parmi ces ménages, 2062 ayant des contacts ont pu être enquêtés		
Collecte	Appel téléphonique	- Les difficultés de fonctionnement des contacts téléphoniques - La réticence des ménages - Temps de collecte jugé long - Insuffisance de crédit de communication	- Dotation supplémentaire en crédit de communication - Motivation des ménages avec des transferts de crédits de communication
Traitement	- Apurement des données pour corriger les valeurs aberrantes - Imputation des valeurs manquantes -		

Annexe 9 : Diagnostic du RIC

Phase	Méthodologie	Difficultés	Initiatives
Echantillonnage	- Toutes les entreprises formelles et informelles -	- Absence de cartographie des entreprises ; -	- Directives pour l'identification des entreprises (voire manuel)
Collecte	- Face à face - Dépôt de questionnaires	- Réticence des entreprises à répondre aux questionnaires et à déclarer leurs informations (chiffres d'affaires, nombre d'employés, cotisation sociales, impôt) ; - Problème d'exhaustivité des entreprises recensées ; - contrainte financière - délais de réaction parfois longs au niveau des entrepreneurs informels	
Traitement	- Apurement des données pour corriger les valeurs aberrantes - Imputation des valeurs manquantes	- Difficultés d'imputation des chiffres d'affaires des entreprises ;	- Imputation par les plus proches voisins ;

Annexe 10 : Diagnostic des statistiques des prix

Phase	Pratique	Difficultés	Initiatives
Echantillonnage	- Choix raisonné de marchés	- Couverture insuffisante du territoire et des marchés locaux	- Elargissement de la couverture nationale a d'autres régions
Collecte	- Face à face - Mesure directe	- Insuffisance de ressources financières, humaines et logistiques ; - Imprécision de certaines descriptions dues au changement d'agent enquêteur; - Manque de certains outils de mesures	- A compléter
Traitement	- Saisie des données ; - Apurement des données pour corriger les valeurs aberrantes - Imputation des valeurs manquantes	- Forte variabilité des prix, surtout des produits alimentaires (c'est normal) - Variables complexes telles eau, électricité, loyer....	A compléter

Annexe 11 : Forces et faiblesses des Enquêtes ménages

Phase	Procédure/Stratégie	Avant les contextes		Avec les Contextes	
		Force	Faiblesse	Influence de la situation sécuritaire	Influence de la situation sanitaire
Echantillonnage	Procédure : Constitution de base de sondage Tirage aléatoire de ZD Enumération et tirage des ménages Choix des individus	<ul style="list-style-type: none"> - Permet l'extrapolation des résultats grâce à l'utilisation des sondages probabilistes - Possibilité d'une mise à jour régulière de la base de sondage - Améliore la précision des indicateurs (représentativité statistique) 	<ul style="list-style-type: none"> - Vétusté de la base de sondage - peut exiger une mise à jour de la base de sondage (sectionnement des grandes ZD et assemblage de petites ZD) et engendrer donc un coût 	<ul style="list-style-type: none"> - Instabilité de la base de sondage liée aux déplacements des populations - Inaccessibilité d'une grande partie de la base de sondage 	Néant
Collecte	Stratégie 1 : face à face	<ul style="list-style-type: none"> -Établir la confiance avec l'enquêté du fait de sa présence physique -Taux de réponse élevé -Plus rapide à mener -Évaluer la compréhension du répondant et son interprétation des questions -possibilité de relance des enquêteurs 	<ul style="list-style-type: none"> -Influencer intentionnellement ou non des enquêteurs -les biais de désirabilités sociales -couteux (financier) -Possibilité de tricheries de la part des agents enquêteurs -La façon de se présenter de l'enquêteur peut engendrer des refus de répondre (qualité du contact, facteur vestimentaire, hostilité envers un groupe social) -durée de terrain assez longue du fait des déplacements - 	<ul style="list-style-type: none"> - possibilité d'enlèvement du personnel de terrain - Accès limité aux enquêtés - réticence de réponses aux questions -augmentation du taux de refus -réponses inexactes de la part des enquêtés -risque de perte de vie 	<ul style="list-style-type: none"> - risque de propagation du virus - - Accès limité aux enquêtés - réticence de réponses aux questions -augmentation du taux de refus
	Stratégie 2 : appel au téléphone	<ul style="list-style-type: none"> - Encadrement rapproché - Écoutes à distance - Possibilité de suivre et d'enregistrer les entretiens - Rapidité d'exécution, surtout pour les enquêtes complémentaires 	<ul style="list-style-type: none"> - Couverture du réseau - Dysfonctionnement de certains numéros de téléphones - La méfiance des enquêtés à répondre aux questions - La durée de la communication - Pas de possibilité de supports visuels - Moindre attention des enquêtés (la télé allumée, les enfants qui pleurent, ...) 	-La réticence des enquêtés à participer	Néant

Phase	Procédure/Stratégie	Avant les contextes		Avec les Contextes	
		Force	Faiblesse	Influence de la situation sécuritaire	Influence de la situation sanitaire
			<ul style="list-style-type: none"> - L'influence de l'enquêteur est réduite par rapport au face à face, mais toujours présente - Comme en face à face, l'enquêteur peut interpréter les réponses de l'enquêté - Taux d'abandon en cours d'interviews plus élevé - « Ras-le-bol » des appels non sollicités 		
	Stratégie 3 : Enquête en ligne	<ul style="list-style-type: none"> - Questionnaire auto-administré - Pas de présence d'enquêteur, donc aucune influence de celui-ci sur l'enquêté - Possibilité d'administrer des questions sensibles 	<ul style="list-style-type: none"> - Biais de sélection des enquêtés (niveau d'instruction, manque d'équipements, défaut de connexion internet) - Défaut de connexion internet par endroit - Manque d'assistance pour une meilleure compréhension de certaines questions - Mauvais contrôle de l'échantillon - Non-réponse partielle plus fréquente, due à l'absence de relance de la part de l'enquêteur 	Néant	Néant
Traitement		Existence de programmes de traitement pour chaque enquête	<ul style="list-style-type: none"> - Absence de protocole de traitement - Mauvaise organisation des programmes de traitement de données - Absence d'une procédure d'archivage des bases et des programmes d'apurement 	<ul style="list-style-type: none"> - Impossibilité de faire des extrapolations - Réduction de la qualité des estimations - Comparabilité des statistiques (impossibilité de regrouper les modalités d'une même variable) 	Néant

Annexe 12 : Opportunités et menaces

Phase	Procédure/Stratégie	Avant les contextes	
		Opportunités	Menaces
Echantillonnage	Procédure : Constitution de base de sondage Tirage aléatoire de ZD Enumération et tirage des ménages Choix des individus	<ul style="list-style-type: none"> - existence d'outils pour l'échantillonnage spatial à travers les images satellitaires - existence de nouvelles sources de données (générées opérateurs de téléphonies, supermarchés, e-commerce) - Informatisation des procédures de l'administration (bases de données des usagers des services...) ou dans le secteur privé - Taux de possession et d'utilisation de téléphones élevé ; - Couverture élevée du territoire en réseau de téléphonie 	<ul style="list-style-type: none"> - L'insécurité de plus en plus croissante
Collecte	Stratégie 1 : face à face		<ul style="list-style-type: none"> - L'insécurité de plus en plus croissante - Présence de cas de Covid-19
	Stratégie 2 : appel au téléphone		
	Stratégie 3 : Enquête en ligne		
Traitement		<ul style="list-style-type: none"> - Existence de nouveaux outils innovants pour l'organisation (stockage et sécurisation) et le traitement des données 	

Annexe 13 : Forces faiblesse de l'enquête entreprise :

		Avant les contextes		Avec les Contextes	
	Procédure/Stratégie	Force	Faiblesse	Influence de la situation sécuritaire	Influence de la situation sanitaire
Echantillonnage	Procédure : - Sélection des 49 communes urbaines ; - Exclusion des marchés et des unités sans local fixe ; - Ratissage le long des rues et dans les zones d'activités économiques (entreprises formelles et informelles) ; - Utilisation de données de sources administratives (registres des entreprises, DSF...)	- Existence de textes précisant les communes urbaines ; - Disponibilité de données de source administrative ; - Disponibilité de liste des marchés	- Couverture partielle du territoire ; - Absence de cartographie des entreprises ; - Insuffisance de la démographie des entreprises ; - Absence d'une mise à jour régulière du répertoire statistique des entreprises.	- Fermeture de certaines entreprises - Inaccessibilité de certaines localités -	Néant
Collecte	Stratégie 1 : face à face	-Etablir la confiance avec l'enquêté du fait de sa présence physique -Taux de réponse élevé -Plus rapide à mener -Evaluer la compréhension du répondant et son interprétation des questions -possibilité de relance des enquêteurs	-Influence intentionnelle ou non de la part des enquêteurs -Risque de triches de la part des agents enquêteurs -La façon de se présenter de l'enquêteur peut engendrer des refus de répondre (qualité du contact, facteur vestimentaire, hostilité envers un groupe social) -durée de terrain assez longue du fait des passages multiples auprès de la même unité.	- Risque d'enlèvement du personnel de terrain - Accès limité aux enquêtés - réticence de réponses aux questions -augmentation du taux de refus -réponses inexactes de la part des enquêtés -Risque de perte de vie	- risque de propagation du virus - Accès limité aux enquêtés - réticence de réponses aux questions -augmentation du taux de refus
	Stratégie 2 : Transmission des questionnaires par mail	- Réduction des coûts de réalisation des enquêtes ; - Rapidité d'exécution, surtout pour les enquêtes complémentaires	- Couverture du réseau internet - Dysfonctionnement des adresses mail ; - Faible réactivité de la part des entreprises (pas d'accusé de réception) ;	- La réticence des enquêtés à participer	Néant

		Avant les contextes		Avec les Contextes	
	Procédure/Stratégie	Force	Faiblesse	Influence de la situation sécuritaire	Influence de la situation sanitaire
			- La méfiance des enquêtés à répondre aux questions		
Traitement		Disponibilité de la syntaxe de traitement dès la fin de l'enquête pilote	<ul style="list-style-type: none"> - Absence de protocole de traitement - Mauvaise organisation des programmes de traitement de données - Absence d'une procédure d'archivage des bases et des programmes d'apurement 	Néant	Néant

Annexe 14 : Opportunités et menaces de l'enquête entreprise

		Avant les contextes	
	Procédure/Stratégie	Opportunités	Menaces
Echantillonnage	<p>Procédure :</p> <ul style="list-style-type: none"> - Sélection des 49 communes urbaines ; - Exclusion des marchés et des unités sans local fixe ; - Ratissage le long des rues et dans les zones d'activités économiques (entreprises formelles et informelles) ; - Utilisation de données de sources administratives (registres des entreprises, DSF...) 	<ul style="list-style-type: none"> - L'utilisation des TIC par les entreprises 	<ul style="list-style-type: none"> - L'insécurité de plus en plus croissante
Collecte	<p>Stratégie 1 : face à face</p> <p>Stratégie 2 : Envoie des questionnaires par mail</p>	<ul style="list-style-type: none"> - Développement des TIC (outils de collecte mobile, outils de communication modernes avec les réseaux sociaux, Mail, collecte en ligne, plateformes Web, etc.) - Gestion modernisée des entreprises 	<ul style="list-style-type: none"> - L'insécurité de plus en plus croissante - Insuffisance de ressources financières propres à l'Etat pour les opérations de collecte
Traitement		<ul style="list-style-type: none"> - Existence de nouveaux outils innovants pour l'organisation (stockage et sécurisation) et le traitement des données - Disponibilité de données de source administrative 	

Annexe 15 : Forces et faiblesse de l'Enquête prix

		Avant les contextes		Avec les Contextes	
	Procédure/Stratégie	Force	Faiblesse	Influence de la situation sécuritaire	Influence de la situation sanitaire
Echantillonnage	Procédure : Pour mémoire : méthodologie IHPC	- Prix de base représentatifs de toutes les régions - -	Couverture partielle du panier Couverture géographique partielle (certains marchés non pris en compte) Système de pondérations vieillissant	- Inaccessibilité de certains marchés ;	- Inaccessibilité aux marchés
Collecte	Stratégie 1 : face à face et mesure directe	- Observation directe - Données exactes	- Coût élevé - Demande de prix sans achat	- Risque d'enlèvement du personnel de terrain - Accès limité aux marchés - augmentation du taux de refus pour les demandes de prix sans achat --Risque de perte de vie	- risque de propagation du virus - Accès limité aux marchés
Traitement		Disponibilité d'un logiciel de traitement des données	- Logiciel de traitement non actualisé	Néant	Néant

Annexe 16 : Opportunités et menaces de l'enquête prix

		Avant les contextes	
	Procédure/Stratégie	Opportunités	Menaces
Echantillonnage	Procédure : Pour mémoire : méthodologie IHPC	- Augmentation rapide des supermarchés et des ventes en ligne - Prix affiché	- L'insécurité de plus en plus croissante
Collecte	Stratégie 1 : face à face et mesure directe	- Augmentation rapide des supermarchés et des ventes en ligne	- L'insécurité de plus en plus croissante - Insuffisance de ressources financières propres à l'Etat pour les opérations de collecte

Annexe 17 : Forces et faiblesses du Recensement

		Avant les contextes		Avec les Contextes	
	Procédure/Stratégie	Force	Faiblesse	Influence de la situation sécuritaire	Influence de la situation sanitaire
Cartographie	<p>Procédure :</p> <ul style="list-style-type: none"> - l'élaboration de croquis (villages ou secteurs) en se basant sur les informations existantes - Estimation de la taille de la population des villages ou secteurs. - Découpage en ZD (800-1000 habitants en milieu rural et 1000-1200 habitants en milieu urbain) 	<ul style="list-style-type: none"> - Permet une répartition quasi-égalitaire du travail des enquêteurs - Donne la chance a tout individu de la population d'appartenir à l'échantillon lors des enquêtes 	<ul style="list-style-type: none"> - Possibilité d'oublier des ilots - Sous-estimation et sur estimation de la taille réelle de certaines ZD - Possibilité d'avoir des ZD à cheval entre plusieurs villages - Limites physiques non matérialisées des entités administratives 	<ul style="list-style-type: none"> - Inaccessibilité de certaines entités administratives - Surpeuplement ou dépeuplement de certaines ZD du fait des déplacements massifs de populations (PDIs) 	<ul style="list-style-type: none"> - Alourdissement du coût de la cartographie
Collecte	Stratégie : Face à face	<ul style="list-style-type: none"> - Etablir la confiance avec l'enquêté du fait de sa présence physique - Taux de réponse élevé - Plus rapide à mener - Evaluer la compréhension du répondant et son interprétation des questions - Possibilité de relance des enquêteurs - La façon de se présenter de l'enquêteur peut engendrer des réponses exactes 	<ul style="list-style-type: none"> - Influence intentionnelle ou non de la part des agents recenseurs - Les biais de désirabilités sociales - Couteux (financier) - Risque de triche de la part des agents recenseurs - Les mauvaises attitudes des agents recenseurs peuvent engendrer des refus de répondre (qualité du contact, facteur vestimentaire, hostilité envers un groupe social) - La lourdeur de l'opération - Risques de perturbations pouvant compromettre le recensement 	<ul style="list-style-type: none"> - Risques d'enlèvement du personnel de terrain - Accès limité aux enquêtés - Réticence de réponses aux questions - Augmentation du taux de refus - Réponses inexactes de la part des enquêtés - Risque d'accidents et pertes de vies humaines 	<ul style="list-style-type: none"> - Risque de propagation du virus - Accès limité aux enquêtés - Réticence de réponses aux questions - Augmentation du taux de refus

		Avant les contextes		Avec les Contextes	
	Procédure/Stratégie	Force	Faiblesse	Influence de la situation sécuritaire	Influence de la situation sanitaire
			- Non actualisation des compétences relativement aux nouveaux outils		
Traitement		- Existence de programmes de traitement lors du recensement ;	<ul style="list-style-type: none"> - Absence de protocole de traitement - Mauvaise organisation des programmes de traitement de données - Absence d'une procédure d'archivage des bases et des programmes d'apurement - Insuffisance ou absence de rapports de traitement - Traitements indépendants de parties de la base de données non capitalisés - Non actualisation des compétences Non actualisation des compétences 	<ul style="list-style-type: none"> - incomplétude de la base de données du recensement - Réduction de la qualité des statistiques du recensement 	- Néant

Annexe 18 : Opportunités et menaces du recensement

		Avant les contextes	
	Procédure/Stratégie	Opportunités	Menaces
Cartographie	<p>Procédure :</p> <ul style="list-style-type: none"> - L'élaboration de croquis (villages ou secteurs) en se basant sur les informations existantes - Estimation de la taille de la population des villages ou secteurs. - Découpage en ZD (800-1000 habitants en milieu rural et 1000-1200 habitants en milieu urbain) 	<ul style="list-style-type: none"> - Existence d'outils pour la digitalisation (SIG, images satellitaires, GPS, outils de collecte mobile, etc.) - Appuis techniques des partenaires 	<ul style="list-style-type: none"> - L'insécurité de plus en plus croissante
Collecte	<p>Stratégie 1 : face à face</p> <p>Stratégie 2 : appel au téléphone</p> <p>Stratégie 3 : Enquête en ligne</p>	<ul style="list-style-type: none"> - Développement des TIC (outils de collecte mobile, outils de communication modernes avec les réseaux sociaux, Mail, collecte en ligne, plateformes Web, etc.) - Niveau d'instruction des populations de plus en plus élevé (disponibilité de diplômés) ; - Ouverture d'une école de formation continue 	<ul style="list-style-type: none"> - L'insécurité de plus en plus croissante - Présence de cas de Covid-19 - Insuffisance de ressources financières propres à l'Etat pour les opérations de collecte
Traitement		<ul style="list-style-type: none"> - Existence de nouveaux outils innovants pour l'organisation (stockage et sécurisation) et le traitement des données - Ouverture d'une école de formation continue 	<ul style="list-style-type: none"> - Déphasage entre équipements existant et nouvelles technologies ;

Annexe 19 : Analyse comparative des Procédures/stratégies

		Avantage(s)	Inconvénient(s)	Coût (Aucun-Elevé-Moyen-Faible)
Echantillonnage	Procédure A : - Constitution de base de sondage - Tirage aléatoire de ZD - Enumération et tirage des ménages - Choix des individus	- Permet une généralisation (extrapolation) des résultats (représentativité statistique) - Maîtrise du niveau de précision des indicateurs - Maîtrise de la dispersion des ménages et des individus à enquêter ; - Existence de marges de manœuvre sur le coût de réalisation (possibilité d'ajuster la taille de l'échantillon en fonction des contraintes budgétaires)	- Limitation de l'aléa du tirage des ménages aux ZD sélectionnées (Faible dispersion des ménages) - Non recommandé pour certaines études (étude des événements rares) ; - Niveau de précision assujettie au coût de réalisation (taille de l'échantillon) ; - Non maîtrise a priori de l'accessibilité des unités échantillonnées	Aucun si base e sondage disponible Faible, si la base de sondage doit être mise à jour
	Procédure B : - Sélection des entités géographiques à l'exception de celles à fort défis sécuritaires - Tirage des ZD - Tirage des ménages après énumération -	Pour la strate accessible : - Généralisation (extrapolation) des résultats (représentativité statistique) - Maîtrise du niveau de précision des indicateurs - Maîtrise de la dispersion des ZD, des ménages et des individus à enquêter ; - Existence de marges de manœuvre sur le coût de réalisation - Réduction des risques liées à l'insécurité (attaques, enlèvements, perte en vies humaines)	- - Limitation de l'aléas du tirage des ZD aux entités géographiques accessibles - Limitation de l'aléas du tirage des ménages aux ZD sélectionnées (Faible dispersions des ménages) - Non recommandé pour certaines études (étude des événements rares) ; - Niveau de précision assujettie au coût de réalisation (taille de l'échantillon) - Ne permet pas toujours d'extrapoler les résultats à l'échelle national - Perte d'informations sur les zones écartées	Aucun si base e sondage disponible Faible, si la base de sondage doit être mise à jour

	<p>Procédure C :</p> <ul style="list-style-type: none"> - Constitution de strates : zones accessibles, zones inaccessibles (sites d'accueil des PDI) - Tirage des ZD et des sites - Tirage des ménages après énumération 	<ul style="list-style-type: none"> - Permet d'approcher le niveau de précision comme dans la procédure A - Permet de rattraper ce que l'on perd au niveau les zones difficiles en saisissant l'information au niveau des déplacés - Généralisation (extrapolation) des résultats (représentativité statistique) - Maîtrise de la dispersion des ZD, des ménages et des individus à enquêter ; - Existence de marges de manœuvre sur le coût de réalisation 	<ul style="list-style-type: none"> - Des biais systématiques dans la strate des zones inaccessibles - Non exhaustivité et non-conformité de la liste des PDI avec la liste de leur zone de provenance - Limitation de l'aléas du tirage des ménages aux ZD sélectionnées (Faible dispersions des ménages) ; - Biais de sélection dans le tirage des sites d'accueil des PDI ; - Non applicabilité pour certaines études (étude des événements rares) ; - Niveau de précision assujettie au coût de réalisation (taille de l'échantillon) 	<p>Aucun si base e sondage disponible</p> <p>Faible, si la base de sondage doit être mise à jour</p> <p>Faible, pour l'identification des sites d'accueil</p>
	<p>Procédure D :</p> <ul style="list-style-type: none"> - Constitution de la base de sondage à partir des nouvelles sources des données ou big data (opérateurs de téléphonies mobiles, données biométriques, supermarchés, images satellitaires...); - Tirage d'individus 	<ul style="list-style-type: none"> - Respect de l'aléa dans le tirage - Forte dispersion de l'échantillon réduisant le biais de sélection des US ; - Existence de marges de manœuvre sur le coût de réalisation ; 	<ul style="list-style-type: none"> - Faible représentativité statistique des résultats du fait par exemple de la couverture réseau - Biais de sélection des répondants (Seulement ceux qui disposent de données des sources considérées dans l'étude) ; - Données parfois limitées en termes d'information voulues ; - Accessibilité réduite aux big data (procédures d'accès complexes); - Nécessité préalable d'équipement en infrastructures spécifiques ; - Difficultés dans la constitution de la base de sondage (Incomplétude de la base de sondage) - Possibilité d'altérer la qualité de l base de sondage (possession de plusieurs contacts téléphoniques multiples pour un individu) 	<p>Coût relativement élevé</p>

	<p>Procédure E :</p> <ul style="list-style-type: none"> - Identification de points focaux dans les villages et secteurs - Tirage de points focaux 	<ul style="list-style-type: none"> - La proximité des points focaux peut faciliter l'accès à l'information - Permet une généralisation (extrapolation) des résultats - Maîtrise du niveau de précision des indicateurs - Maîtrise de la dispersion des US ; - Existence de marges de manœuvre sur le coût de réalisation 	<ul style="list-style-type: none"> - Insuffisance de niveau formation si l'enquête requiert un certain degré de compréhension - Faible représentativité (qualité) de l'information collectée auprès des points focaux - Risque de distorsion de l'information (idéaux vs faits sociaux); - Données parfois limitées en termes d'information voulues ; - Impossibilité de collecter des données à caractère personnel (des individus et des ménages) ; 	Faible coût
Collecte	<p>Stratégie 1 : face à face dans le ménage</p>	<ul style="list-style-type: none"> - Etablir la confiance avec l'enquêté du fait de sa présence physique - Taux de réponse élevé - possibilité de relance des enquêteurs - Possibilité d'apprécier la qualité des réponses (gestuelle, signes du visage, observation de l'environnement ...) 	<ul style="list-style-type: none"> - Influence des répondants par les enquêteurs 	Coût moyen
	<p>Stratégie 2 : face à face avec déplacement du chef de ménage</p>	<ul style="list-style-type: none"> - Etablir la confiance avec l'enquêté du fait de sa présence physique - Taux de réponse élevé - Possibilité de relance des enquêteurs - Possibilité d'apprécier la qualité des réponses ; - 	<ul style="list-style-type: none"> - Influence des répondants par les enquêteurs - Impossibilité d'apprécier l'environnement du ménage par observation ; - possibilité de contrainte au niveau de la taille de l'échantillon pour réduire les coûts liés aux déplacements - Nombreux refus probables ; - Risques liés au déplacement de l'enquêté (accidents, attaques, etc.) 	Coût élevé
	<p>Stratégie 3 : enquête par appel au téléphone/ whatsapp</p>	<ul style="list-style-type: none"> - Possibilité de relance des enquêteurs - Pas de risques liés aux déplacements des agents enquêteurs ou des enquêtés ; - Rapidité de l'exécution 	<ul style="list-style-type: none"> - Restriction de l'échantillon aux utilisateurs de téléphone - Impossibilité d'apprécier l'environnement du ménage par observation ; - Forte probabilité de nombreux cas de refus ; 	Coût faible

			<ul style="list-style-type: none"> - Forte probabilité de nombreux cas de questionnaires partiellement renseignés ; - Interruptions dû à la qualité du réseau - Cas de suspensions d'entretiens du fait de circonstances diverses (entretiens longs et lassants, évènements sociaux et contraintes de l'enquêté, etc.) 	
	Stratégie 4 : enquête en ligne	<ul style="list-style-type: none"> - Pas de risque liés aux déplacements des agents enquêteurs ou des enquêtés ; 	<ul style="list-style-type: none"> - Restriction de l'échantillon aux utilisateurs de l'internet ; - Problème de couverture réseau - Forte probabilité de non réponse et des cas de refus; - Difficultés de constitution de la base de sondage - Pas de possibilité d'extrapoler les résultats 	Coût faible
Traitement	Post stratification et Repondération (redistribution des poids)	-permet d'extrapoler les résultats à l'échelle national ou régional	- introduction des biais dans les résultats	Faible
	Restriction de la portée géographique des résultats	- permet d'extrapoler les résultats dans les zones couvertes	<ul style="list-style-type: none"> - exclusion de certaines zones des résultats - validité des indicateurs seulement pour les zones couvertes 	Faible
	Exploitation de méthodes d'estimation utilisant les big Data (opérateurs de téléphonies mobiles, données biométriques, supermarchés, images satellitaires)	- permet d'avoir les résultats à l'échelle des zones ou populations couvertes par les Big data	<ul style="list-style-type: none"> - inaccessibilité des big data - méconnaissance de la portée des big data par ses producteurs (opérateurs de téléphonie, etc.) - Qualité et éthique 	Elevé

Le Tableau 56 donne un aperçu sur la cartographie des risques en matière de collecte de données.

Tableau 15 : Exemple de cartographie des risques en matière de collecte de données

<i>REGIONS</i>	<i>PROVINCES</i>	<i>COMMUNES</i>	<i>ZD</i>	<i>METHODES DE COLLECTE</i>
Accessible	Accessible	Accessible	Accessible	Utilisation d'agents enquêteurs sur le terrain -Appel téléphonique -WhatsApp -e-mail -Etc.
Accessible	Accessible	moyennement accessible	moyennement accessible	Utilisation d'agents enquêteurs sur le terrain -Utilisation d'agents enquêteurs locaux résidents sur le terrain -Appel téléphonique -WhatsApp -e-mail,
Accessible	Accessible	Non accessible; réseau téléphonie mobile disponible	-non accessible -réseau téléphonie mobile disponible	- Utilisation d'agents enquêteurs locaux résidents sur le terrain -déplacer les ménages dans les zones accessibles si possible - Appel téléphonique - WhatsApp - e-mail

REGIONS	PROVINCES	COMMUNES	ZD	METHODES DE COLLECTE
Accessible	Accessible	Non accessible; réseau téléphonie mobile disponible	-non accessible -pas de réseau téléphonie mobile disponible	- Utilisation d'agents enquêteurs locaux résidents sur le terrain -déplacer les ménages dans les zones accessibles si possible
Accessible	moyennement accessible	Non accessible; réseau téléphonie mobile disponible	Non accessible ; réseau téléphonie mobile disponible	- Utilisation d'agents enquêteurs locaux résidents sur le terrain -deplacer les ménages dans les zones accesibles si possible - Appel téléphonique - WhatsApp - e-mail
Accessible	moyennement accessible	Non accessible; réseau téléphonie mobile disponible	Non accessible; pas de réseau téléphonie mobile disponible	- Utilisation d'agents enquêteurs locaux résidents sur le terrain si possible - déplacer les ménages dans les zones accessibles si possible
Accessible	Non accessible; réseau téléphonie mobile disponible	Non accessible; réseau téléphonie mobile disponible	Non accessible; réseau téléphonie mobile disponible	- Utilisation d'agents enquêteurs locaux résidents sur le terrain -déplacer les ménages dans les zones accessibles si possible - Appel téléphonique - WhatsApp - e-mail
moyennement accessible	Non accessible -réseau téléphonie mobile disponible	Non accessible; réseau téléphonie mobile disponible	Non accessible; réseau téléphonie mobile disponible	- Utilisation d'agents enquêteurs locaux résidents sur le terrain -déplacer les ménages dans les zones accessibles si possible - Appel téléphonique

REGIONS	PROVINCES	COMMUNES	ZD	METHODES DE COLLECTE
				<ul style="list-style-type: none"> - WhatsApp - e-mail
moyennement accessible	Non accessible; réseau téléphonique disponible	Non accessible; réseau téléphonique mobile disponible	Non accessible; pas réseau téléphonique mobile disponible	<ul style="list-style-type: none"> - Utilisation d'agents enquêteurs locaux résidents sur le terrain si possible - déplacer les ménages dans les zones accessibles si possible
Non accessible; réseau téléphonique disponible	Non accessible; réseau téléphonique disponible	Non accessible; réseau téléphonique mobile disponible	Non accessible; réseau téléphonique mobile disponible	<ul style="list-style-type: none"> - Utilisation d'agents enquêteurs locaux résidents sur le terrain -déplacer les ménages dans les zones accessibles si possible - Appel téléphonique - WhatsApp - e-mail
Non accessible; réseau téléphonique disponible	Non accessible; réseau téléphonique disponible	Non accessible; pas réseau téléphonique mobile disponible	Non accessible; pas réseau téléphonique mobile disponible	<ul style="list-style-type: none"> - Utilisation d'agents enquêteurs locaux résidents sur le terrain -déplacer les ménages dans les zones accessibles si possible

Accessible : pas de menace pour accéder à la localité et pour travailler

Moyennement accessible : zone sous contrôle des forces de défenses et de sécurité et accessible par la route mais avec prudence, une attaque déjà effectué dans la localité

Non accessible, réseau téléphonie mobile disponible : attaques récurrentes, présence probable de mines sur les voies, enlèvements, zone sous le contrôle terroristes, absence de l'administration publique ...

Non accessible, pas réseau téléphonie mobile disponible : attaques récurrentes, présence probable de mines sur les voies, enlèvements, zone sous le contrôle terroristes, absence de l'administration publique ...

Coordination administrative

OUEDRAOGO Boureima, - Directeur Général de INSD

BERE Bernard, Directeur Général Adjoint de l'INSD

KIENDREBEOGO Sandaogo Hamadou, Directeur de la Coordination Statistique, de la Formation et de la Recherche de l'INSD

Auteurs principaux

ILBOUDO Daouda Abdoul Aziz, Chef du SERDM/ DCSFR

ZONGO Beb-Zinda Gérard Wilfried Arnaud, SERDM/ DCSFR

NABIE Binouni, SERDM/ DCSFR

BAKO Babou Jean Pierre, SERDM/ DCSFR

Contributeurs

BELEMKOABGA Lucien, Chargé de contrôle interne/DG

OUEDRAOGO Souleymane, Chargé de contrôle interne/DG

OUEDRAOGO W Salif, Service de la coordination statistique/DCSFR

SOMDA Rogatien, Service de la formation, du perfectionnement et du recyclage/DCSFR

KABORE Ali, Direction des statistiques et des synthèses économiques

SARA Karim, Direction des statistiques et des synthèses économiques

GUIATIN Youssouf, Direction des statistiques et des synthèses économiques

NIGNAN Gafarou, Direction des statistiques sur les conditions de vie des ménages

DIALLO Oumarou, Direction des statistiques sur les conditions de vie des ménages

ZONGO Issa, Direction de la démographie

HEMA K Djouma, Direction de la démographie

OUEDRAOGO Rasmané, Direction de la démographie

KOANDA Harouna, Direction de la démographie

BOUDA Moussa, Direction de l'informatique et du management de l'information statistique

GANSORE Moctar, Direction régionale de l'INSD- Sahel

KINDO Seydou Mamadou, Direction régionale de l'INSD- Sahel

RAMDE Bakary, Direction régionale de l'INSD- Est

MANLI Boureima, Direction régionale de l'INSD- Est

OUEDRAOGO Issaka, Direction régionale de l'INSD- Hauts-Bassins

BENGALY Labasse, Direction régionale de l'INSD- Hauts-Bassins

OUEDRAOGO Issaka, Direction générale des études et des statistiques sectorielles /Agriculture

PARE Issouf, Direction générale des études et des statistiques sectorielles / Eau

BASSINGA Gaetan, Direction générale des études et des statistiques sectorielles / Environnement

KIEMTORE Ali, Direction générale des études et des statistiques sectorielles / Emploi

Novembre 2022

DCSDF/INSD

Institut national de la statistique et de la démographie (INSD) Avenue Pascal ZAGRE, Ouaga 2000

01 B.P : 374 Ouagadougou 01 – Burkina Faso

Tél : (00226) 25 37 62 04 - Fax : (00226) 25 37 62 26

Site internet : www.insd.bf - Email : insd@insd.bf ou insbf@yahoo.fr