

Décembre 2023

منظمة
الأغذية والزراعة
للأمم المتحدة联合国
粮食及
农业组织Food and Agriculture
Organization of the
United NationsOrganisation des
Nations Unies pour
l'alimentation et
l'agricultureПродовольственная и
сельскохозяйственная
организация
Объединенных НацийOrganización de las
Naciones Unidas para la
Agricultura y la
Alimentación

COMMISSION AFRICAINE DES STATISTIQUES AGRICOLES

Vingt-huitième Session

Johannesburg, Afrique du Sud: 4 – 8 décembre 2023

POINT 10 DE L'ORDRE: Nouveaux développements dans l'utilisation de sources de données alternatives pour les statistiques agricoles

Comblen les lacunes des données agricoles : Innovations dans les sources de données géospatiales et non conventionnelles

**Lorenzo De Simone, Bureau du Chef Statisticien
Christian Mongeau, Division des statistiques**

**ORGANISATION :
Organisation des Nations Unies pour l'Alimentation et l'Agriculture**

SOMMAIRE

Les statistiques officielles jouent un rôle crucial dans les processus décisionnels qui ont un impact sur les aspects économiques, démographiques, sociaux et environnementaux de la société. Elles reposent sur une collecte approfondie et précise des données. Toutefois, ce processus peut être considérablement entravé par des lacunes dans les données. Ces lacunes, qui dénotent l'absence de données ou des ensembles de données incomplets, peuvent sérieusement affecter la fiabilité, la représentativité et l'utilité de ces statistiques dans leur ensemble. Il est essentiel de combler ces lacunes pour que les statistiques sur lesquelles nous nous appuyons pour comprendre les différents aspects de la société soient aussi solides, représentatives et utiles que possible. Cette démarche permet non seulement d'améliorer la qualité des données, mais aussi de renforcer l'intégrité et l'efficacité des décisions prises sur la base de ces données.

Cette session est consacrée à la présentation des nouveaux développements dans l'utilisation de sources de données alternatives pour les statistiques agricoles dans deux domaines spécifiques. Tout d'abord, l'accent sera mis sur la science des données géospatiales en illustrant le projet EOSTAT et le projet de registre des agriculteurs financé par la Banque africaine de développement. Deuxièmement, nous partagerons des expériences pratiques dans l'exploitation de sources de données non conventionnelles pour obtenir des informations opportunes en présentant certaines des activités réalisées par le laboratoire de données de l'Organisation des Nations Unies pour l'Alimentation et l'Agriculture (qui sera également illustré dans ce document) et le laboratoire d'innovation de données de la Banque africaine de développement.

Lancé par le Bureau du Chef Statisticien et la Division des statistiques (ESS) de la FAO, le projet Observation de la Terre pour les statistiques (EOSTAT) vise à développer et à renforcer les capacités techniques des bureaux nationaux de statistiques (BNS) dans l'utilisation des données d'observation de la Terre (Big Data) et de l'intelligence artificielle comme source de données alternative et méthodes pour améliorer la qualité globale, l'exhaustivité, la granularité et l'actualité des statistiques agricoles, en mettant l'accent sur les surfaces cultivées et le rendement des cultures. Le présent rapport fait le

point sur le niveau des travaux au Sénégal, au Mali, au Rwanda, au Lesotho et au Zimbabwe et décrit les activités connexes à venir dans la région pour la prochaine période biennale. Les membres sont invités à exprimer leurs points de vue sur les progrès accomplis et les défis qui subsistent, et à fournir des conseils jugés nécessaires.

Le laboratoire de données de la FAO (Data Lab), créé fin 2019, vise à moderniser les statistiques de l'Organisation et à élargir la couverture des données en utilisant des méthodes et des technologies innovantes pour extraire des données à partir de sources non structurées. Son projet initial consistait à recueillir des données sous-nationales à partir de sites Web statistiques officiels afin de combler les lacunes et de valider l'ensemble de données nationales sur la production agricole. Ce projet sera détaillé dans la section suivante. Avec l'apparition de la pandémie de COVID-19 en 2020, le besoin d'informations opportunes pour soutenir la prise de décision est devenu critique, en particulier dans les situations d'urgence. La demande d'informations en temps réel et d'analyses automatisées provenant de sources non conventionnelles s'est donc accrue. Pour répondre à ces nouveaux besoins en matière de données, le Data Lab a développé et géré diverses bases de données, notamment des articles de journaux internationaux, des informations sur les prix quotidiens des denrées alimentaires et des données sur les pertes et gaspillages alimentaires tirées d'articles scientifiques et d'autres documents.

Ce document explore succinctement la question des lacunes en matière de données statistiques et examine comment les données géospatiales et les sources non conventionnelles peuvent contribuer à relever les défis posés par le manque de données.

1. INTRODUCTION

Les méthodologies statistiques traditionnelles ont toujours servi de base à la collecte de données et à l'analyse ultérieure. Ces approches, qui vont des enquêtes détaillées aux observations pratiques sur le terrain, ont permis d'obtenir des informations essentielles dans divers secteurs. Malgré leur valeur, les méthodes traditionnelles de collecte de statistiques officielles se heurtent fréquemment au problème des "lacunes dans les données", c'est-à-dire des domaines dans lesquels les informations sont soit manquantes, soit incomplètes.

Les lacunes dans les données peuvent provenir d'une multitude de sources, y compris, mais sans s'y limiter, les enquêtes à faible taux de réponse, les dossiers médicaux incomplets ou l'arrêt des méthodes de collecte de données en raison de circonstances imprévues telles qu'une pandémie mondiale. Ces lacunes peuvent se manifester au cours de la phase initiale de collecte des données ou se développer au fil du temps, lorsque les données deviennent obsolètes ou non pertinentes. La représentativité des données est particulièrement compromise lorsque certains sous-groupes de la population sont sous-représentés en raison de données manquantes, ce qui conduit à des conclusions biaisées et à une prise de décision politique potentiellement erronée. Des données incomplètes peuvent avoir un impact significatif sur les perspectives économiques individuelles et entraver la mobilité sociale, en particulier pour les groupes marginalisés, et peuvent également conduire à la formulation de politiques qui ne répondent pas de manière adéquate aux besoins de ces communautés, renforçant ainsi les préjugés systémiques et maintenant les disparités sociales.

La pandémie de COVID-19 a clairement mis en évidence la vulnérabilité des systèmes statistiques aux événements inattendus, les interruptions généralisées des activités de collecte de données dans le monde entier en étant un excellent exemple. La résilience et les progrès futurs des statistiques officielles dépendront de leur capacité à relever ces défis tout en maintenant l'objectivité et la fiabilité des données dans un environnement social et technologique en constante évolution.

Dans ce contexte, les méthodes de la science des données sont apparues comme un complément précieux aux statistiques traditionnelles, offrant de nouvelles perspectives et de nouveaux aperçus. La science des données utilise des techniques analytiques et des algorithmes avancés pour analyser des ensembles de données vastes et variés, y compris des sources de données non conventionnelles. Cette approche permet d'explorer des modèles et des prédictions cachés, offrant ainsi une compréhension plus dynamique et plus complète de questions complexes. En outre, la capacité de la science des données à traiter les données volumineuses et à appliquer des modèles d'apprentissage automatique a ouvert de nouvelles voies pour combler les lacunes en matière de données. Par exemple, la modélisation prédictive et l'analyse des sentiments peuvent fournir des informations sur l'opinion et le comportement du public, des domaines souvent ignorés par les méthodes traditionnelles.

Dans la suite de ce document, nous explorerons quelques méthodologies innovantes dans les domaines de la science des données géospatiales et des sources de données non conventionnelles afin de recueillir des informations opportunes.

2. PARTIE PRINCIPALE DU DOCUMENT

Big Data de l'observation de la Terre et le projet EOSTAT

Les Big Data de l'observation de la Terre constituent une source de données alternative précieuse qui peut soutenir le travail des Bureaux nationaux de statistique dans la production de statistiques agricoles et la communication d'informations sur les indicateurs des ODD, en réduisant l'ensemble du fardeau grâce à une efficacité accrue. Aujourd'hui, les conditions d'adoption des données d'observation de la Terre pour les statistiques agricoles sont très favorables : la disponibilité généreuse et sans précédent de données satellitaires gratuites et ouvertes à haute et moyenne résolution résultant de la démocratisation des données d'observation de la Terre, la quantité croissante d'expertise, de méthodes et d'outils pour les utiliser sont d'importants catalyseurs. Il est désormais possible d'accéder à des pétaoctets d'images satellite provenant de diverses sources de données (Landsat, Modis, Sentinel), ce qui permet de construire des séries temporelles denses pour détecter et discriminer les classes d'occupation des sols, la phénologie des cultures et les caractéristiques spectrales. Cela permet à son tour de cartographier avec précision l'occupation des sols et le type de culture, et d'obtenir des statistiques agricoles exactes. Le développement du stockage en nuage et de la technologie informatique, ainsi que l'essor de l'apprentissage automatique et de l'intelligence artificielle ont encore élargi l'horizon des options pour le déploiement d'infrastructures à faible coût et l'automatisation.

Malgré tout, l'utilisation réelle des données d'observation de la Terre à des fins opérationnelles dans les services nationaux de statistique reste très faible au niveau mondial, et en particulier dans les pays en développement, en raison d'une série d'obstacles techniques, financiers et administratifs. Le défi le plus important est certainement le manque de données in situ de qualité adéquate.

Les données in situ, ou plus simplement les données géoréférencées sur les cultures, sont nécessaires en grande quantité pour la formation, le test et la validation des classificateurs supervisés, qui sont ensuite utilisés pour "transformer" les images satellite en cartes catégorielles, notamment des cartes pour types de cultures.

Dans ce contexte, la FAO, par le biais du projet EOSTAT, aide les pays à renforcer leur capacité technique à utiliser les données d'observation de la Terre pour les statistiques agricoles officielles en travaillant ensemble sur les questions clés en collaboration avec les BNS, avec une stratégie articulée autour de trois actions stratégiques :

- L'action 1 se concentre sur l'examen des protocoles nationaux existants de collecte de données sur le terrain (enquêtes agricoles et recensements agricoles) et veille à ce qu'ils soient optimisés pour les applications d'OT, tant en ce qui concerne la répartition de l'échantillonnage entre les classes que la géographie, et à ce que les meilleures pratiques en matière de géoréférencement soient mises en œuvre.
- L'action 2 se concentre sur le développement et l'essai de méthodes de classification appliquées à des séries temporelles d'images satellite, qui sont plus économes en données que la forêt aléatoire, et sur le soutien à la recherche dans le domaine de l'apprentissage par transfert.
- L'action 3 articule la co-conception et la co-création du système national de surveillance des cultures pour en démontrer la faisabilité et ensuite soutenir leur opérationnalisation.

Depuis 2019, l'EOSTAT a été mise en œuvre dans 21 pays à travers le monde, renforçant les capacités techniques dans les domaines techniques clés suivants :

- Optimisation de la conception des enquêtes de terrain et intégration des meilleures pratiques en matière de géoréférencement au sein des enquêtes agricoles annuelles et du recensement agricole.
- Cartographie de l'occupation du sol
- Cartographie des types de cultures
- Modélisation du rendement des cultures
- Cartographie des parcelles de terrain
- Calcul des indicateurs SDG à l'aide des données EO

Dans le cadre de l'AFCAS 2023, la FAO fournit une mise à jour des travaux réalisés au Rwanda, au Lesotho, au Sénégal, au Mali et au Zimbabwe.

Cartographie de l'occupation du sol

Au Rwanda, en collaboration avec l'Institut national des statistiques du Rwanda (NISR), l'Office agricole du Rwanda (RAB), l'Autorité foncière nationale (NLA) et l'Agence spatiale rwandaise, un premier prototype national d'occupation du sol a été développé en utilisant une adaptation de la procédure automatisée développée par De Simone L. et al., en l'absence de données in situ, avec une précision globale de 75 %. Une enquête de terrain optimisée a été conçue pour être mise en œuvre au niveau national en décembre 2023/janvier 2024, afin d'élaborer la première carte nationale de référence de l'occupation du sol pour l'année de référence 2023, avec une précision attendue supérieure à 95 %. En 2024, en plus de la cartographie de l'occupation du sol, la cartographie des types de cultures et les activités de rendement des cultures seront également mises en œuvre. Le NISR collecte des données géoréférencées in situ dans le cadre de l'enquête agricole annuelle, en géoréférençant à la fois les limites des champs cultivés et le centroïde des parcelles, et constitue un modèle exceptionnel dans ce domaine.

Suivi des ODD et établissement de rapports

En collaboration avec le Bureau des statistiques (BOS) du Lesotho, la FAO a développé une solution nationale de suivi des ODD qui permet de calculer l'indicateur de l'ODD 15.4.2, l'indice de couvert végétal montagneux sur une base annuelle, de manière automatique à partir de données d'observation de la Terre. La solution est basée sur la méthodologie de l'Indice de couvert végétal montagneux développée par De Simone et al, 2021 au niveau mondial, et optimisée pour le Lesotho en utilisant les données nationales de couverture terrestre en conjonction avec le modèle numérique d'élévation global de la JAXA. Cette solution a permis au BOS de rendre compte des statistiques de l'Indice de couvert végétal montagneux désagrégées au niveau de la zone d'élévation de 2017 à 2022. Au fur et à mesure que la nouvelle ligne de base d'occupation du sol est mise à jour dans le système, la ligne de base de l'indicateur de l'indice de couvert végétal montagneux est également mise à jour avec un effet domino.

Cartographie des types de cultures, surfaces cultivées

Au Mali, EOSTAT est mis en œuvre en collaboration avec l'Institut d'Economie Rurale (IER) du Ministère du Développement Rural, la Cellule de Planification et de Statistique (CPS), et la Compagnie Malienne pour le Développement des Textiles (CMDT). Une campagne ad-hoc optimisée sur le terrain a été conçue et testée conjointement pour collecter des données in situ sur les cultures en 2023. La méthodologie comprend les étapes suivantes :

- Stratification basée sur l'intensité des cultures (0% - 30% ; 30% - 60% ; 60% - 100%) calculée à partir de la carte d'occupation du sol WorldCover de l'ESA.
- Sélection aléatoire de 300 segments (500m X 600m) dans les différentes zones d'intensité.
- Numérisation manuelle (à l'écran) de blocs de culture/parcelles homogènes à l'aide de l'imagerie Google Earth /Bing pour chaque segment.
- Intégration de la carte numérisée avec une application appelée MapMe, utilisée pour la navigation des équipes (conduite sur le lieu de chaque segment) ;
- Enquête basée sur ODK Collect, pour collecter des données de terrain (en répondant à un formulaire sur le type de culture et la zone de culture) ;
- L'application Qfield, pour confirmer les limites des blocs de culture/parcelles et les modifier en cas d'erreurs.

Le travail s'est concentré sur le coton, le sorgho, le maïs, l'arachide et le riz. La zone d'intérêt a été définie dans la ceinture cotonnière du Mali et plus précisément dans le district de Diola. Les données de l'OT ont été utilisées pour concevoir l'enquête agricole. En 2024, les données collectées sur le terrain seront utilisées pour produire une carte nationale des types de cultures et une carte des rendements des cultures. Les résultats seront comparés aux estimations de l'enquête agricole annuelle officielle.

Au Sénégal, l'EOSTAT a été lancé pour la première fois en 2019 en collaboration avec la Direction de l'Analyse, de la Prévision et des Statistiques Agricoles (DAPSA), l'Agence Nationale de la Statistique et de la Démographie (ANSD) et le Centre de Suivi Ecologique (CSE), Sénégal. Au cours de la première phase du projet, les données sur les cultures de l'Enquête Annuelle Agricole (AAS) pour l'année 2018 ont été utilisées pour produire le premier masque national des cultures et la première carte des types de cultures à l'échelle nationale pour les cultures majeures et mineures à l'aide d'une classification par forêt aléatoire. La précision des cartes était de 90 % et 76 % respectivement pour le masque des cultures et la carte des types de cultures. En 2022, une enquête expérimentale optimisée sur le terrain a été mise en œuvre conjointement avec la DAPSA afin de démontrer la valeur ajoutée de la numérisation des limites et des centroïdes des parcelles. Ces données in situ ont ensuite été utilisées pour produire un masque des cultures et une carte des types de cultures avec une précision globale de 91,8 % et 90,2 % respectivement. Enfin, la superficie des cultures a été estimée d'abord directement à partir des données in situ, puis à partir de la combinaison des données in situ et des données d'observation de la Terre, et donc à partir de la carte des types de cultures. L'exercice a montré que l'intégration de l'OT et des données in situ réduirait le coefficient de variation de l'estimation de la superficie de 50 % pour le mil et l'arachide. Cela prouve la valeur ajoutée des données d'observation de la Terre dans l'amélioration de la précision, en utilisant des données libres et ouvertes. L'obtention de résultats similaires par l'intensification de l'échantillonnage sur le terrain aurait entraîné des coûts plus élevés. Depuis 2023, l'AAS a officiellement introduit le géoréférencement des limites et des centroïdes des parcelles de culture.

Au Zimbabwe, l'EOSTAT est mise en œuvre en collaboration avec le ministère de l'agriculture, l'agence spatiale du Zimbabwe (ZINGZA) et les statistiques du Zimbabwe (Zim-Stat) dans le cadre du projet de production alimentaire d'urgence du Zimbabwe (ZEFPP) avec le soutien financier de la Banque africaine de développement (Facilité de production alimentaire d'urgence pour l'Afrique). En 2023, une enquête de terrain optimisée a été conçue et mise en œuvre, collectant 600 points de données circa in situ, et une première carte nationale du blé d'hiver a été élaborée en 2023 avec une précision globale de 98%.

Prévision du rendement des cultures

La prévision du rendement des cultures à l'aide de données d'observation de la Terre a été testée au Sénégal et au Cameroun, en utilisant respectivement une approche statistique et une approche basée sur les processus. Au Sénégal, des mesures de rendement de référence géolocalisées ont été collectées sur des centaines de parcelles de culture dans le paysage de Nioro. En fonction de la culture, la taille du carré de mesure variait entre 5 et 25 m². Un modèle de régression a été développé en utilisant les données d'observation de la Terre comme variable explicative du rendement. Les résultats ont montré une faible corrélation, mais l'échantillonnage des carrés de culture s'est avéré insuffisant et a probablement été la cause principale des mauvais résultats. Au Cameroun, un modèle basé sur les processus, l'approche systémique de la durabilité de l'utilisation des terres. Comme les données in situ n'étaient pas disponibles, le modèle de l'approche systémique de la durabilité de l'utilisation des terres a été calibré pour cinq produits principaux (maïs, riz, sorgho et manioc) en utilisant les statistiques de rendement de FAOSTAT. Le modèle a été utilisé pour produire des prévisions de rendement des cultures au niveau national, à une résolution de 10 mètres pour chaque produit et pour chaque année entre 2011 et 2019.

Innovation en matière de données pour combler les lacunes et obtenir des informations opportunes

En tant que biens immatériels essentiels, les données sont souvent confrontées à des problèmes de disponibilité dus à des facteurs tels que de faibles capacités statistiques, un financement insuffisant et une mauvaise culture de diffusion et d'utilisation des données. Pour combler ces lacunes, il est essentiel d'explorer les données web non structurées et de les combiner avec des méthodes innovantes de production d'informations. Les organisations nationales et internationales reconnaissent la nécessité de s'adapter à de nouvelles sources et méthodes de données à la lumière de la crise à laquelle sont confrontés les systèmes traditionnels de collecte de données. En réponse, l'Organisation des Nations Unies pour l'Alimentation et l'Agriculture (FAO) a créé le "Laboratoire de données pour l'innovation statistique" en 2019 (voir Fabi et al., 2022). Cette initiative vise à moderniser les processus statistiques de l'Organisation en améliorant l'actualité et la granularité de la collecte des données, en permettant une analyse automatisée et en facilitant la détection des signaux d'alerte précoce. Le Data Lab utilise des technologies avancées telles que le web scraping, le text mining (fouille de texte), l'analyse de données géo-spatiales et l'intelligence artificielle, ainsi que des sources de données non conventionnelles telles que les médias sociaux et les journaux en ligne.

Alors que certains domaines bénéficient d'une abondance de données (par exemple, les réseaux sociaux, les appareils intelligents), l'obtention d'informations dans d'autres domaines, comme les pertes de production agricole, peut s'avérer difficile et coûteuse. Le processus exige de concevoir des enquêtes, de compiler des informations et de normaliser les résultats, ce qui nécessite des efforts coordonnés, des ressources financières et des capacités humaines et techniques. Souvent, ces ressources font défaut dans certains pays, ce qui se traduit par des données peu fréquentes ou indisponibles. La situation est exacerbée par des facteurs tels que la diminution des capacités statistiques, la réduction des budgets et la faiblesse des cadres réglementaires.

La pandémie de COVID-19 en 2020 a encore souligné l'importance des sources de données non conventionnelles et des méthodes innovantes pour améliorer la couverture et fournir des informations rapides. La pandémie a accéléré la transformation vers ces nouvelles approches, soulignant leur urgence. Le dernier rapport sur les activités des Nations Unies en matière d'intelligence artificielle (UIT, 2023), qui présente 281 projets innovants menés par 40 entités, donne un bon aperçu de la manière dont les organisations internationales utilisent ces nouvelles méthodes et sources.

Les efforts du Data Lab de la FAO ont permis d'augmenter la couverture statistique de l'Organisation en comblant les lacunes en matière de données. Les sous-sections suivantes visent à détailler certaines des mesures prises par le laboratoire de données.

Obtenir des informations à partir d'articles de presse

Alors que la pandémie de COVID-19 se propageait à travers le monde en 2020, le besoin critique d'informations opportunes et exploitables est apparu très clairement, en particulier pour soutenir les processus décisionnels cruciaux dans une crise qui évolue rapidement. La pandémie a mis en évidence la demande pressante d'informations en temps réel ou quasi réel, avec un accent particulier sur les sources de données non traditionnelles et la nécessité d'une analyse automatisée de ces données. Pour répondre à ces nouvelles exigences en matière de données, le laboratoire de données pour l'innovation statistique s'est activement impliqué dans le développement et la gestion de diverses bases de données. Une initiative importante lancée en janvier 2020 a été la création d'une base de données complète qui comprend les tweets de plus de 500 comptes de journaux mondiaux. À la fin du mois d'avril 2023, elle contenait environ 28 millions de tweets et d'articles. Cette base de données couvre plusieurs langues, telles que l'anglais, l'arabe, l'espagnol, le français, l'italien, le portugais et le russe, et fusionne de manière unique le texte des tweets avec les articles d'actualité connexes lorsqu'ils sont tous deux disponibles.

Cette base de données est soumise à une série d'algorithmes avancés de traitement du langage naturel (NLP) afin d'enrichir les données d'informations précieuses. L'identification et la catégorisation des sujets pertinents constituent l'une des principales composantes de cette analyse. Les tweets et les articles sont classés dans des catégories pertinentes pour les objectifs de développement durable (ODD), notamment "chaîne de valeur alimentaire", "prix des denrées alimentaires", "changement climatique", "sécurité alimentaire", etc. Pour cette classification, nous utilisons un modèle de classification par apprentissage automatique, en particulier XGBoost, qui a été entraîné sur un ensemble de données préparé et validé par une expertise humaine sur différents sujets et dans différentes langues. La catégorisation de ces sujets est essentielle pour faciliter une analyse ciblée et générer des informations sur des sujets de préoccupation spécifiques. Cette approche permet de formuler des politiques éclairées et des réponses stratégiques aux crises.

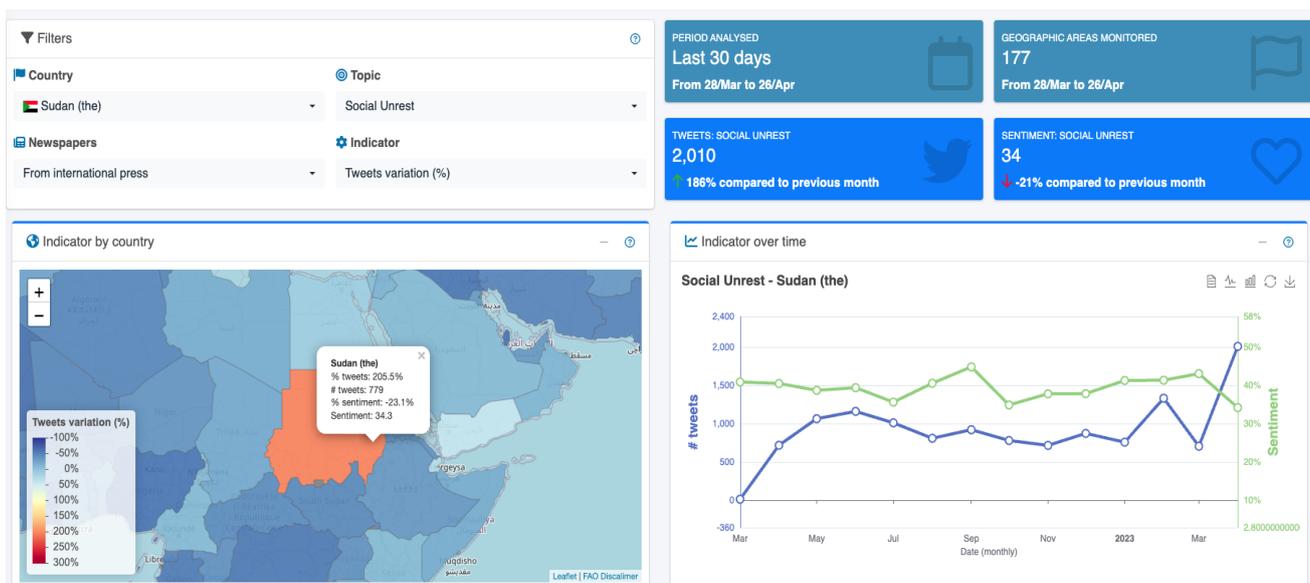
Outre la catégorisation des articles en divers sujets, nous appliquons également des méthodes pour extraire les mentions de pays dans les textes et calculer un indice de sentiment. L'identification des mentions de pays est particulièrement intéressante car elle révèle la façon dont les différents pays perçoivent ou discutent de divers sujets par rapport à d'autres nations. Cette analyse offre une vue unique des perspectives internationales, en montrant comment les nations commentent et comprennent les questions qui se posent au-delà de leurs frontières. En outre, l'analyse des sentiments joue un rôle important dans la compréhension du ton et de la nature de la couverture de divers sujets. En évaluant le sentiment, nous obtenons des informations sur la manière dont des questions spécifiques sont dépeintes et sur l'humeur dominante ou l'opinion publique dans différents pays concernant ces questions. L'analyse du sentiment est réalisée à l'aide d'une approche basée sur le dictionnaire, où les mots du texte sont comparés à un lexique prédéfini composé de mots positifs et négatifs. La formule appliquée est la suivante:

$$S = \frac{\#w^+}{\#w^+ + \#w^-} \times 100$$

où $\#w^+$ est le nombre de mots positifs et $\#w^-$ le nombre de mots négatifs trouvés dans le texte. L'indice va de 0 à 100, où 0 représente un sentiment très négatif, 100 un sentiment très positif et 50 un texte neutre.

Les informations sur les thèmes et les sentiments pour tous les pays du monde sont disponibles dans un tableau de bord convivial à l'adresse <https://www.fao.org/datalab/early-warnings/topics-explorer/>

Le graphique ci-dessous démontre clairement l'efficacité de notre outil d'analyse. Elle illustre la manière dont le sujet des "troubles sociaux" au Soudan a attiré l'attention de la presse internationale. Plus important encore, elle montre une baisse notable et constante du sentiment au fur et à mesure que les conflits dans le pays s'intensifient. Ce déclin du sentiment reflète effectivement la gravité et l'intensité croissantes de la situation au Soudan, telle qu'elle est rapportée par les médias du monde entier. Cette visualisation met en évidence non seulement la capacité de l'outil à suivre l'attention portée par les médias du monde entier à des événements spécifiques, mais aussi son aptitude à saisir l'évolution du ton émotionnel de la couverture, qui sert d'alerte précoce pour détecter d'éventuels problèmes. Cette analyse fournit une vue dynamique de la manière dont les perceptions internationales changent en réponse à des situations évolutives dans différents pays, offrant ainsi des informations précieuses sur la réaction mondiale à des événements importants.

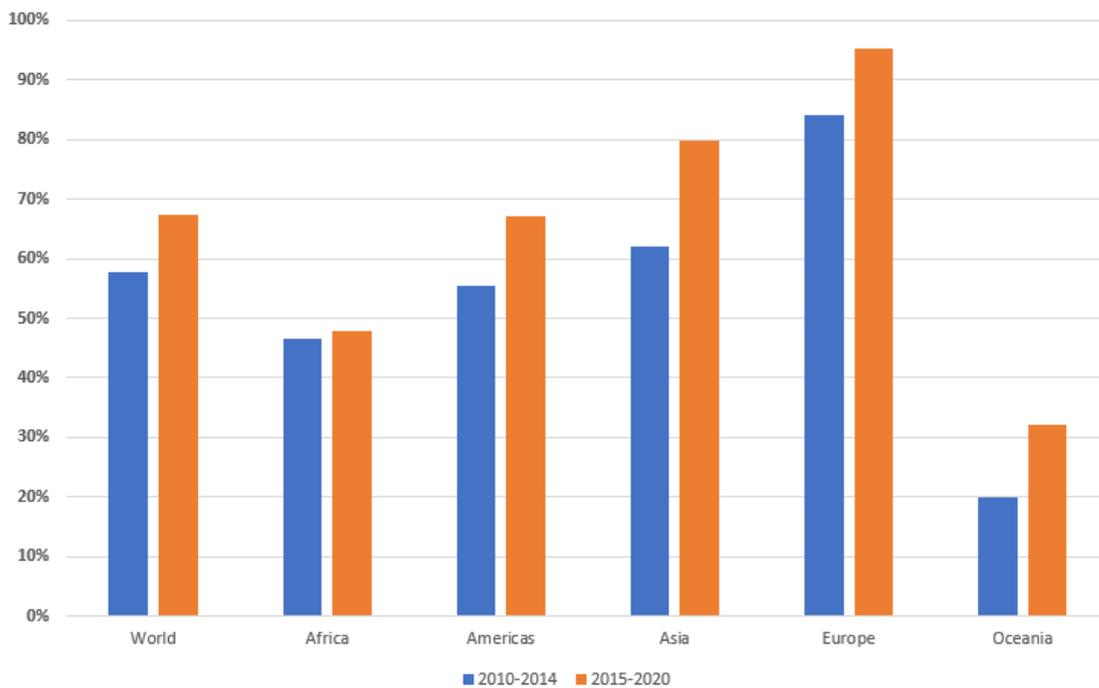


Comblent les lacunes dans les statistiques officielles de production et de superficie

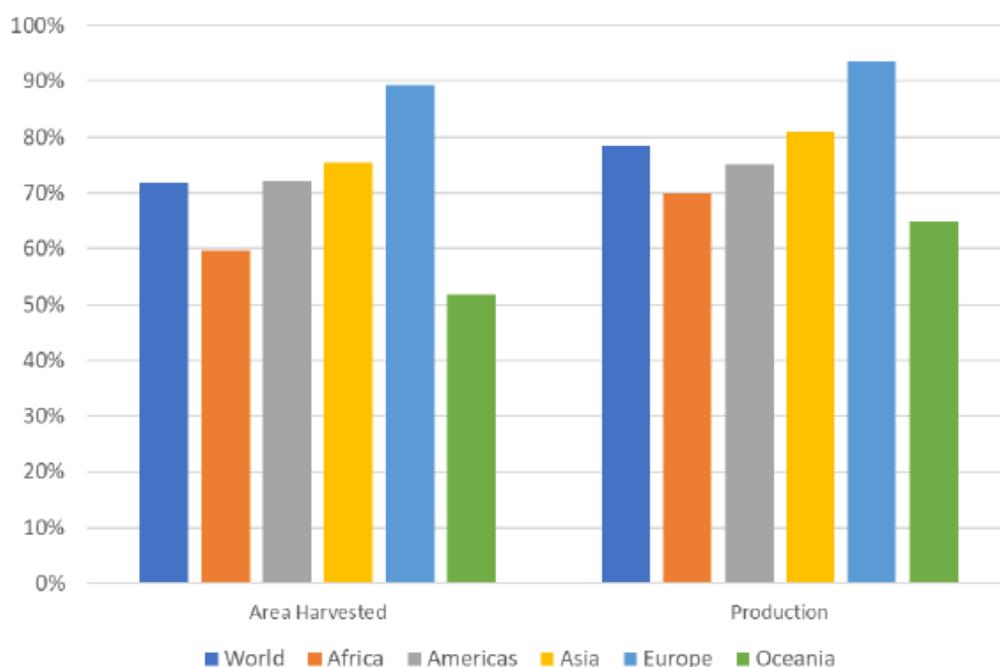
L'initiative Main dans la main de la FAO se concentre sur l'accélération de la transformation agricole et le développement rural durable pour lutter contre la pauvreté et la faim, en s'alignant sur les Objectifs de développement durable. Elle cible les pays aux capacités limitées ou confrontés à des crises, en utilisant la modélisation et l'analyse géospatiales avancées pour améliorer les revenus et réduire les vulnérabilités des populations rurales. L'initiative vise à améliorer les interventions politiques, les financements, les investissements et les réformes institutionnelles.

Le rôle du Data Lab dans la collecte de données agricoles, en particulier dans les pays où les données sont rares, est au cœur de l'Initiative Main dans la main. Pour ce faire, il utilise des techniques d'intelligence artificielle, en se concentrant sur les données sous-nationales pour combler les lacunes et valider les données existantes, ce qui permet d'améliorer la qualité des données et de soutenir les améliorations apportées par les gouvernements.

La figure ci-dessous montre les taux de réponse moyens aux Questionnaires sur la production agricole, c'est-à-dire le pourcentage de questionnaires reçus par rapport aux questionnaires envoyés. Bien que le taux de réponse aux Questionnaires sur la production agricole ait globalement augmenté au cours des dernières années, dans certaines régions, il est nettement inférieur à 50 % (Océanie) ou légèrement inférieur (Afrique).



Une vue comparable apparaît lorsque nous calculons un "indicateur d'exhaustivité" de l'APQ, qui est la proportion de points de données officiels par rapport à l'ensemble des combinaisons potentielles de pays/matières premières pour lesquelles au moins un point de données officiel a été fourni dans un APQ. Les résultats de ce calcul sont présentés dans la figure ci-dessous. Avec la figure précédente, il devient évident qu'il existe une grande marge de manœuvre pour améliorer l'étendue de notre couverture statistique.



Pour combler ces lacunes, le Data Lab utilise le web scraping pour collecter des données à partir de sources en ligne telles que les offices statistiques nationaux ou les ministères. Ce processus comprend l'extraction de données à partir de différents formats, la normalisation des noms de produits et d'administrations, et la résolution des barrières linguistiques.

Le processus de "web scraping" comporte des défis, tels que la gestion de divers formats de documents et la traduction de documents non anglophones. La normalisation des noms de produits et d'administrations utilise une approche de correspondance floue, avec une consultation d'experts pour les cas non concordants. En octobre 2023, le laboratoire de données a collecté un nombre important de points de données sur divers produits agricoles. Ces données sont utilisées pour recouper et compléter les données de production agricole de FAOSTAT, améliorant ainsi la qualité et l'exhaustivité de l'ensemble des données.

Les données collectées permettent non seulement d'améliorer l'ensemble des données existantes, mais jouent également un rôle crucial dans l'identification et la rectification des divergences dans les données rapportées, contribuant ainsi de manière significative à la précision et à la fiabilité globales des statistiques agricoles.

Pour donner une idée des résultats de la résolution des lacunes de données par le web scraping et le text mining comme illustré ci-dessus, nous rapportons dans le tableau suivant les résultats de l'exercice 2023 sur l'ensemble de données de la production agricole. Il indique le nombre de produits qui ont été récupérés sur le web et le nombre de produits qui ont été intégrés (ou acceptés) dans l'ensemble de données sur la production agricole. Il indique également le nombre réel de points de données acceptés pour chaque produit et le pourcentage de produits qui ont été intégrés sur le total des produits récupérés sur le web.

Pays	Produits mis au rebut	Accepté		
		Produits	Points de données	% produits
Afghanistan	20	0	0	0%
Bangladesh	85	4	15	5%
Bhutan	63	11	35	17%
Burkina Faso	13	0	0	0%
Burundi	13	5	5	38%
Cambodge	44	9	121	20%
RD Congo	15	6	10	40%
Dominique	39	29	200	74%

Equateur	32	6	141	19%
Gabon	20	9	39	45%
Gambie	15	0	0	0%
Haiti	14	2	3	14%
Indonésie	22	0	0	0%
Laos	12	0	0	0%
Lesotho	7	4	13	57%
Mali	6	0	0	0%
Mozambique	20	7	11	35%
Nepal	9	0	0	0%
Niger	17	2	13	12%
Pakistan	33	0	0	0%
Rwanda	65	6	12	9%
Sénégal	11	4	43	36%
Zambie	27	12	50	44%

Les données présentées dans le tableau précédent permettent d'obtenir un résultat significatif : en moyenne, 20 % des produits acquis par le web-scraping ont été intégrés avec succès dans l'ensemble de données sur la production agricole. Cette intégration met en évidence l'efficacité de l'utilisation des méthodologies de web-scraping et de text-mining dans l'amélioration de la disponibilité des données. Ces méthodes permettent non seulement d'élargir le champ d'application de l'ensemble de données, mais contribuent également à sa richesse et à sa précision.

En outre, l'intégration réussie d'une proportion substantielle de données extraites du web souligne le potentiel de ces techniques innovantes pour combler les lacunes en matière de données et mettre à jour les ensembles de données existants avec des informations plus actuelles et plus complètes. Le processus de récupération de données sur le web, en particulier, permet d'extraire un large éventail de données à partir de diverses sources en ligne, offrant ainsi une vue plus complète du secteur agricole. De même, l'exploration de texte facilite l'extraction et l'analyse d'informations pertinentes à partir de grands volumes de textes non structurés, garantissant que l'ensemble de données n'est pas seulement étendu mais aussi perspicace.

Extraction de données sur les pertes alimentaires via l'exploration de texte

La cible 12.3 de l'ODD vise à réduire de moitié les déchets alimentaires mondiaux par habitant d'ici 2030 et à réduire les pertes alimentaires dans les chaînes de production et d'approvisionnement. Pour suivre cette évolution, l'indice des pertes alimentaires et l'indice des déchets alimentaires ont été développés et gérés respectivement par la FAO et le Programme des Nations Unies pour l'environnement. L'indice des pertes alimentaires recense les produits agricoles perdus avant d'atteindre le stade de la vente au détail, souvent aux stades de production, de post-récolte et de distribution. En revanche, l'indice du gaspillage alimentaire se concentre sur le gaspillage alimentaire au niveau de la vente au détail et des consommateurs, ce qui implique que des aliments de bonne qualité sont jetés.

Cependant, les données sur les pertes et gaspillages alimentaires sont rares. En mai 2023, seuls 7 % des données sur les pertes alimentaires étaient officiellement communiquées, principalement par le biais de questionnaires adressés par la FAO aux pays. Bien qu'il existe une quantité importante de documents fournissant des estimations sur les pertes alimentaires, les statistiques officielles restent limitées. Pour combler cette lacune, la FAO a mis au point un processus semi-automatisé d'exploration de texte pour extraire des données sur les pertes alimentaires à partir de divers documents, notamment des articles scientifiques, des documents de travail et de la littérature grise. Ce processus comprend trois étapes principales : la collecte et le prétraitement automatisés des documents, un modèle statistique évaluant la pertinence des documents et l'extraction guidée des données. Le processus commence par la recherche de documents à l'aide de mots-clés liés à la perte (par exemple, "perte après récolte") dans les moteurs de recherche universitaires, suivie de l'extraction de métadonnées telles que les auteurs, les titres et les dates de publication. Une routine de résumé de texte et l'identification des mots-clés sont également effectuées.

La pertinence des documents téléchargés pour les chiffres de pertes réelles est déterminée à l'aide d'un classificateur d'apprentissage automatique, en particulier un modèle Forêts aléatoires. Les documents sont ensuite triés en fonction de leur pertinence pour l'extraction des données, c'est-à-dire que seuls les documents susceptibles de contenir des

pourcentages de pertes alimentaires selon le modèle sont conservés, tandis que les autres sont écartés. L'étape finale consiste en une validation manuelle par des analystes, qui rassemblent et valident les informations à l'aide d'un outil web qui met en évidence dans le texte des éléments clés tels que les noms de produits (par exemple, "maïs", "riz", etc.) ainsi que les noms de pays, les pourcentages, où ils sont placés dans un voisinage de mots liés à la perte ("perte", "pertes", "perdu", etc.). Il s'agit d'un processus "semi-automatisé" puisqu'il y a une étape finale de validation par un expert humain qui peut accepter ou rejeter le texte trouvé par les algorithmes.

En mai 2023, cette méthode a permis d'obtenir un ensemble de données comprenant près de 30 000 points de données couvrant 127 pays et 147 produits, en se concentrant principalement sur les données à partir de 2000. Cet ensemble complet de données est accessible au public sur le site web de la FAO à l'adresse suivante: <https://www.fao.org/platform-food-loss-waste/flw-data/>

L'ensemble de données FLW a deux objectifs principaux. Premièrement, il aide les chercheurs et les analystes à étudier les pertes et gaspillages alimentaires dans un domaine où les données sont rares. Deuxièmement, il est utilisé par la FAO pour estimer les données manquantes sur les pertes de produits primaires pour tous les pays. Pour imputer les pertes lorsque les données sont manquantes, la FAO a développé un modèle d'imputation qui prend en compte les données officielles et les données récupérées telles qu'elles ont été obtenues pour l'ensemble de données FLW. Il s'agit en particulier d'un modèle à effets aléatoires qui prend en compte divers facteurs tels que les prix de l'électricité et du pétrole, les conditions météorologiques et les investissements agricoles. Ce modèle permet de fournir des estimations annuelles des pertes pour les combinaisons pays-marchandises pour lesquelles les données ne sont pas disponibles. En agrégeant les données officielles et imputées, le niveau mondial des pertes alimentaires est estimé à environ 14 % en 2020.

3. CONCLUSIONS ET RECOMMANDATIONS

Les lacunes dans les données constituent un défi important pour l'intégrité des statistiques officielles et, par extension, pour l'efficacité de l'élaboration des politiques. La sous-représentation de certains sous-groupes de la population, le risque de conclusions biaisées et la possibilité de perpétuer les inégalités sociales figurent parmi les préoccupations majeures associées aux données incomplètes. Alors que le monde est confronté à des défis sans précédent tels que la pandémie de COVID-19, le rôle des statistiques officielles n'a jamais été aussi important. Il est impératif que les parties prenantes du système statistique plaident en faveur de changements qui renforcent la législation sur le partage des données, garantissent un financement souple et fiable et modernisent le système statistique afin de garantir sa résilience et sa pertinence face aux défis futurs.

Dans ce contexte, nous avons souligné comment les méthodes de la science des données contribuent de manière significative au progrès des connaissances dans les domaines de l'information géospatiale, améliorent les statistiques officielles et aident à acquérir des connaissances en temps opportun.

L'utilisation des données d'OT s'avère être un outil statistique pertinent, innovant et rentable qui, une fois affiné avec succès au sein d'un système opérationnel, peut aider les BNS des pays à améliorer leur capacité à produire des statistiques agricoles de haute qualité, hautement désagrégées et opportunes, qui sont généralement coûteuses et qui fournissent des estimations après la récolte, telles que les enquêtes agricoles traditionnelles.

Néanmoins, l'utilisation des données d'observation de la Terre reste problématique en raison du manque général de données de formation (données in situ) de qualité adéquate. Les enquêtes agricoles annuelles (EAA) et les recensements agricoles périodiques mis en œuvre dans les pays ne produisent généralement pas de données facilement compatibles avec les données d'observation de la Terre. En fait, ils n'incluent généralement pas le géoréférencement des centroïdes et des limites des parcelles. En outre, la répartition des échantillons entre les différentes classes de cultures n'est pas optimale pour la formation des algorithmes d'apprentissage automatique et d'apprentissage profond. L'échantillonnage est déséquilibré, avec plus d'échantillons pour les classes principales et peu d'échantillons pour les classes secondaires, ce qui entraîne des erreurs de commission et d'omission de la part de l'algorithme de classification. En outre, les échantillons de terrain sont rarement stratifiés par zones agroécologiques et ne représentent donc pas les différentes conditions de croissance des cultures dans le pays.

Dans ces conditions, des enquêtes ad hoc sont nécessaires pour démontrer l'utilisation des données d'observation de la Terre pour les statistiques agricoles, mais il ne s'agit pas d'une solution durable et elle ne devrait être adoptée qu'à court terme.

- Les enquêtes agricoles annuelles et les recensements agricoles périodiques représentent des mécanismes établis dans les pays pour la collecte de données sur le terrain et sont donc la meilleure source potentielle de données in situ durables. Dans ce contexte, la recommandation la plus urgente pour les BNS est de revoir les protocoles existants pour la collecte de données in situ dans les enquêtes agricoles annuelles et les recensements agricoles périodiques. Des ajustements devraient être apportés pour assurer une compatibilité totale avec les exigences de l'analyse cartographique

des cultures basée sur l'observation de la Terre. En particulier, les BNS doivent s'assurer que dans les enquêtes agricoles annuelles (EAA) et les recensements agricoles périodiques :

i) le centroïde des parcelles de culture et les limites des parcelles de culture sont géoréférencés.

ii) les coupes de cultures sont géoréférencées

iii) des GPS professionnels sont utilisés sur le terrain et que le géoréférencement ne repose pas sur le GPS interne du téléphone portable, dont la précision de positionnement est très faible les cultures sont toutes échantillonnées dans les différentes conditions de culture (par exemple, les zones agroécologiques ZAE) ;

v) les cultures mineures sont suffisamment échantillonnées, pas moins de 50 points de données par culture et par zone agroécologique ; les zones agroécologiques sont prises en compte pour la stratification.

- Les services nationaux de statistique qui géoréférencent déjà les données sur les cultures dans le cadre de enquêtes agricoles annuelles et qui utilisent déjà des données d'observation de la Terre pour produire des cartes de l'occupation des sols et des cartes des types de cultures pourraient tirer profit de l'expérimentation de différentes solutions analytiques, afin de maximiser la précision des cartes et de minimiser le biais dans l'estimation des superficies.
- En conséquence, il est recommandé:
 - pour les pays où la fréquence des nuages est élevée, utiliser les données radar (Sentinel-1) en remplacement ou en complément des données optiques (Sentinel-2);
 - expérimenter différentes combinaisons de bandes et d'indices de végétation, de sol et d'eau dérivés des données d'observation de la Terre pour mieux distinguer les différentes classes de cultures, en particulier dans les pays ayant des systèmes agricoles mixtes complexes ; envisager l'utilisation de classes de cultures mixtes pour cartographier les systèmes de cultures mixtes, ou passer à l'imagerie satellitaire commerciale à très haute résolution spatiale (de 5 mètres à moins d'un mètre de résolution) pour distinguer individuellement les types de cultures mixtes.
- Enfin, il est recommandé d'établir ou de renforcer la collaboration entre les OSN, les agences spatiales nationales compétentes et les centres de calcul à haute performance. La durabilité des solutions basées sur l'observation de la Terre dépend en fait également de l'infrastructure disponible qui peut permettre la création et l'analyse efficaces de séries temporelles denses d'images satellite. Une telle collaboration serait essentielle pour la mise en place d'un cube national de données d'images satellite. Cela permettra d'analyser des séries temporelles d'images d'observation de la Terre denses provenant de capteurs multiples, tant optiques que passifs : l'intensification des séries temporelles d'images permet de mieux décrire la phénologie de chaque classe de culture et, partant, de mieux les différencier. Les pays devraient également envisager la possibilité de créer leurs propres cubes de données nationaux, afin de mettre en place leurs systèmes nationaux de surveillance des cultures. Cela permettrait d'aborder une autre question clé, à savoir la confidentialité des données in situ, que les pays ne peuvent pas partager avec le public. L'utilisation de plateformes communes d'observation de la Terre telles que le Moteur Google Earth, AWS, etc. implique le téléchargement de données in situ dans ces systèmes, ce qui peut constituer un problème. L'établissement de cubes de données nationaux peut résoudre ce problème.

Prochaines étapes

Dans les années à venir, la FAO continuera à soutenir les BNS dans les pays en renforçant les capacités dans l'utilisation de sources alternatives de big data telles que les séries temporelles denses d'observation de la Terre pour produire des statistiques officielles agricoles, le rendement des cultures et les superficies améliorant la précision, l'actualité et la granularité ainsi que l'introduction des dimensions de catastrophe comme les inondations et la sécheresse ont des impacts dramatiques sur le rendement.

La FAO a prévu une série de projets pour 2024/2025 dans plusieurs pays d'Afrique. Au Rwanda, en Angola, au Kenya, au Mali, en Éthiopie, au Sénégal, en Ouganda et au Cameroun, la FAO soutiendra l'analyse des séries chronologiques d'images satellitaires, en appliquant des algorithmes d'apprentissage automatique ainsi que des algorithmes d'apprentissage profond et l'intégration avec des données de terrain provenant de SAA optimisés et/ou ajustés.

Cette entreprise de la FAO sera menée en collaboration et en alignement avec la portée plus large du travail du Comité conjoint d'experts de l'ONU sur l'agriculture et du Comité d'experts de l'ONU sur les Big Data. Ce comité conjoint a l'intention de soutenir la modernisation des systèmes statistiques nationaux à l'échelle mondiale en utilisant des sources alternatives de big data et l'innovation. Ce comité est composé des BNS, d'organisations internationales et d'universités. Dans le cadre de ce comité, une équipe de travail spécifique sur l'OT pour les statistiques agricoles a été créée et est coprésidée par la FAO, la Banque Mondiale et l'INEGI, avec un accent sur la télédétection et les big data d'OT.

Le programme de travail 2024-27 de ONU-CEAG/ONU-CEBD comprend le renforcement de domaines spécifiques afin de relever les défis qui ont été identifiés par le TT au cours de 2022/2023 selon le dernier rapport à la Commission statistique des Nations Unies.

Ces domaines sont les suivants :

- L'utilisation intégrée d'images satellite radar et optiques pour la cartographie des types de cultures et la surveillance des catastrophes.
- Enquêtes par OT et par drone
- L'amélioration de l'analyse des séries chronologiques par satellite et le développement de cubes de données d'images satellitaires
- Confidentialité des informations et partage des données in situ.
- Accès du BNS aux données, outils et informations générés par le TT.
- Amélioration et mise à jour de l'application de formation à l'OT:

Dans ce contexte, il convient de noter que le Hub régional de Big Data de l'ONU pour la Chine est passé en 2023 d'un Centre régional à un Centre mondial et a inauguré un laboratoire de télédétection. Le Centre mondial a nommé des experts internationaux de la FAO, de l'INEGI et de l'Institut national de recherche spatiale comme conseillers du laboratoire. Cela facilitera grandement la coopération internationale avec l'Académie chinoise des sciences et l'Institut de recherche spatiale et permettra l'enrichissement mutuel des projets grâce au partage de méthodes et de solutions innovantes et à l'accès aux images à très haute résolution spatiale et temporelle du satellite chinois GAOFEN.

En conclusion, le travail de la commission conjointe ONU-CEAG/CEBD sur l'observation de la Terre pour l'agriculture est rappelé ici pour sa pertinence en tant que forum idéal pour avancer collectivement sur les questions clés trouvées dans l'adoption des données d'observation de la Terre pour les statistiques officielles.

Dans ce document, nous avons présenté l'expérience de la FAO en matière de méthodes novatrices pour combler les lacunes des données et obtenir des informations. Le laboratoire de données de la FAO pour l'innovation statistique illustre l'impact significatif des méthodologies de la science des données et des sources de données non conventionnelles dans la transformation du domaine de l'analyse statistique et de la collecte de données. L'intégration par le laboratoire de technologies avancées telles que le web scraping, le text mining et l'analyse de données géo-spatiales met en évidence l'immense potentiel de la science des données pour améliorer la qualité et la portée des données. En puisant dans des données diverses et non structurées provenant de sources telles que les médias sociaux et les organes d'information en ligne, les articles scientifiques, les rapports et d'autres types de littérature, le laboratoire parvient à combler les lacunes des ensembles de données traditionnels et à fournir de nouvelles informations, élargissant ainsi l'éventail et enrichissant la profondeur et l'actualité de l'analyse des données.

En outre, l'utilisation de la science des données dans le traitement et l'analyse des données permet d'atteindre des niveaux d'efficacité et de précision sans précédent. Avec la capacité de naviguer dans de vastes volumes de données, de discerner des modèles et d'extraire rapidement des informations pertinentes, ces méthodologies s'avèrent inestimables dans un monde complexe et riche en données. Elles offrent un éclairage indispensable sur divers défis mondiaux, notamment la sécurité alimentaire, le changement climatique, le développement économique et la santé publique. La capacité d'analyser de vastes ensembles de données et d'en tirer des enseignements significatifs est essentielle pour s'attaquer à ces questions à multiples facettes dans notre monde interconnecté.

La confluence de la science des données avec des sources de données non conventionnelles révolutionne notre approche de la résolution des problèmes et de la prise de décision éclairée. En comblant les lacunes en matière de données et en fournissant des informations opportunes, ces pratiques innovantes favorisent une approche plus dynamique et plus éclairée dans différents domaines. À mesure que la technologie évolue et que l'abondance des données augmente, les possibilités de découvrir des informations latentes grâce à ces méthodes vont sans aucun doute se multiplier, marquant une ère passionnante pour les chercheurs, les organisations et les décideurs politiques.

L'évolution constante des méthodologies de la science des données, en particulier avec l'application progressive de l'intelligence artificielle (IA), est destinée à améliorer continuellement le traitement et l'interprétation des données. L'IA apporte une touche de transformation au domaine, en permettant des analyses plus complexes et l'exploitation efficace de vastes ensembles de données. À une époque où l'élaboration des politiques et le progrès sociétal reposent de plus en plus sur des informations fondées sur des données, l'amalgame de l'IA et des méthodes statistiques conventionnelles est vital. Cette intégration promet de surmonter les limites antérieures et d'ouvrir de nouvelles voies dans l'analyse des données, en veillant à ce que les statistiques officielles restent pertinentes, fiables et reflètent notre monde dynamique.

4. QUESTIONS ET INVITATIONS AUX PAYS MEMBRES DE L'AFCS

Les membres sont invités à :

- Prendre note du rapport et fournir un retour d'information s'ils le jugent nécessaire ;
- Exprimer leur point de vue sur la pertinence du programme EOSTAT pour la production de statistiques dans leurs pays respectifs ;
- Partager leur expérience et leurs défis concernant les données d'OT pour la cartographie de la couverture terrestre, le suivi et le rapport des indicateurs des ODD, y compris l'indice de couvert végétal montagneux, la cartographie des types de cultures, les estimations de la superficie et du rendement des cultures, et exprimer leurs besoins les plus urgents en matière de méthodologie et/ou de développement des capacités ;
- Prendre note des domaines de travail proposés par l'ONU-CEAG/CEBD pour 2024-27, partager des recommandations et des suggestions pour la finalisation de ce programme de travail et exprimer leur intérêt à devenir membres de l'équipe spéciale ;
- Partager et discuter les principaux défis rencontrés dans l'adoption de techniques de science des données pour l'analyse statistique, ainsi qu'explorer et identifier le type de soutien nécessaire pour surmonter efficacement ces défis ;
- Évaluer et partager les perspectives sur le rôle des sources de données non conventionnelles, telles que les médias sociaux et les nouvelles en ligne, dans l'augmentation des données statistiques traditionnelles. Cette discussion vise à comprendre si et comment ces sources peuvent contribuer au développement des systèmes statistiques ;
- Identifier et discuter des opportunités de partenariats public-privé dans le domaine de la collecte et de l'analyse des données. Il s'agit notamment de déterminer comment les sources de données non conventionnelles peuvent être intégrées et exploitées dans le cadre de ces partenariats ;
- Formuler une vision du rôle futur de l'intelligence artificielle (IA) dans l'amélioration des systèmes de données nationaux. Si une activité est en cours d'adoption, décrire les mesures prises pour concrétiser cette vision, en favorisant une discussion collaborative sur l'intégration stratégique de l'IA dans les pratiques en matière de données ;
- Explorer l'opportunité de collaborer avec le laboratoire Big Data de la FAO, en accord avec le financement obtenu pour un projet sur l'action anticipative dans les situations d'urgence et pour l'élaboration des politiques. Cette invitation est adressée aux pays qui cherchent à renforcer leurs capacités d'analyse dans les interventions d'urgence et la formulation des politiques grâce à des méthodologies avancées en matière de science des données et d'IA. Nous encourageons les membres à collaborer avec le Big Data Lab de la FAO pour améliorer leurs stratégies fondées sur les données et leurs processus de prise de décision.

5. RÉFÉRENCES

De Simone, L.; Navarro, D.; Gennari, P.; Pekkarinen, A.; de Lamo, J. Using Standardized Time Series Land Cover Maps to Monitor the SDG Indicator “Mountain Green Cover Index” and Assess Its Sensitivity to Vegetation Dynamics. *ISPRS Int. J. Geo-Inf.* **2021**, *10*, 427. <https://doi.org/10.3390/ijgi10070427>

De Simone, L.; Ouellette, W.; Gennari, P. Operational Use of EO Data for National Land Cover Official Statistics in Lesotho. *Remote Sens.* **2022**, *14*, 3294. <https://doi.org/10.3390/rs14143294>

Fabi, C.; Rosero Moncayo, J.; Mongeau Ospina, C.A., Silva e Silva, L.G. The FAO Data Lab on statistical innovation and the use of big data for the production of international statistics, *Statistical Journal of the IAOS*, **2022**, 38(3)

ITU (International Telecommunication Union), *United Nations Activities on Artificial Intelligence (AI) 2022*, **2023**. <https://www.itu.int/pub/S-GEN-UNACT-2022>