



**INTÉGRATION DES
STATISTIQUES: DÉFIS ET
SOLUTIONS PROPOSÉES**

FÉVRIER 2024

2024 CONFÉRENCE SUR LES
**DONNÉES ET STATISTIQUES
DE L'ÉDUCATION**

1. INTRODUCTION

Le cadre de suivi des TCG implique un niveau d'ambition substantiel et nécessite un niveau d'innovation correspondant. Les appels à une révolution des données après 2015 reflétaient ce sentiment d'urgence pour de nouvelles idées. Une suggestion clé de 2014 était que « plus les données peuvent être combinées, plus elles sont utiles » (Groupe consultatif d'experts indépendants sur la révolution des données pour le développement durable, 2014).

Il existe plusieurs indicateurs de l'éducation qui s'appuient potentiellement sur plusieurs sources de données ou types de sources de données :

- Certains indicateurs peuvent être alimentés par plusieurs sources du même type. Ceux-ci pourraient – en raison de différences dans les méthodologies, les objectifs ou les circonstances au moment de la collecte des données – produire des résultats qui ne sont pas entièrement comparables sans une analyse plus approfondie. Citons par exemple les indicateurs de résultats d'apprentissage basés sur différentes évaluations et les indicateurs d'indices de parité de richesse basés sur différentes enquêtes (qui peuvent mesurer la richesse de différentes manières).
- Certains indicateurs peuvent être alimentés par différents types de sources (Figure 1). Un exemple est le taux de non-scolarité, qui peut s'appuyer à la fois sur des données administratives et d'enquête. Un autre exemple est celui du développement professionnel continu des enseignants, qui peut s'appuyer à la fois sur des données administratives et sur l'évaluation des apprentissages. Certains pays peuvent opter pour une seule source de données plutôt qu'une autre.

Le problème que les statisticiens de l'éducation sont donc de plus en plus appelés à résoudre est de savoir comment intégrer plusieurs sources de données ou types de sources de données dans l'estimation des indicateurs.

Figure 1. Répartition des indicateurs mondiaux et thématiques de l'TCG 4, par source de données potentielle



Même avant les TCG, la multiplicité des sources de données et des types de sources de données avait également mis d'autres secteurs au défi de produire des estimations d'indicateurs sélectionnés. Par exemple:

- La nécessité d'utiliser plusieurs enquêtes avec des méthodologies différentes (ainsi que de combler les lacunes dans les données) pour estimer les indicateurs de malnutrition, tels que l'émaciation et le retard de croissance, a conduit à la création du groupe inter-agences Joint Child Malnutrition Estimates en 2011 ([UNICEF et al. , 2023](#)).
- La nécessité d'utiliser plusieurs sources de données administratives et d'enquête pour estimer les indicateurs de santé a conduit à la création du Groupe interinstitutions des Nations Unies pour l'estimation de la mortalité infantile et à l'adoption d'un modèle pour générer des estimations annuelles de la mortalité des moins de 5 ans ([Alkema et New, 2014](#)) et la création du Groupe interinstitutions des Nations Unies sur l'estimation de la mortalité maternelle ([Alkema et al., 2016](#)).

Des mesures similaires ont été récemment prises dans le domaine de l'éducation. Mais les opportunités s'accompagnent de défis. Ce document décrit ces étapes, les défis émergents et un programme potentiel pour une intégration plus efficace des données.

2. PROGRÈS RÉALISÉS

Plusieurs sources de données et types de sources de données ont été récemment utilisées pour estimer deux indicateurs d'éducation : le taux d'achèvement (indicateur 4.1.2) et le taux de non-scolarité (indicateur 4.1.4).

2.1. UN MODÈLE DE TAUX D'ACHÈVEMENT

Le Groupe interinstitutions et d'experts sur les indicateurs des TCG a adopté le taux d'achèvement à trois niveaux d'enseignement (primaire, premier cycle du secondaire et deuxième cycle du secondaire) comme indicateur mondial des TCG 4.1.2, l'une des six propositions réussies sur plus de 200 faites au cours de l'enquête. Examen complet 2020 du cadre de suivi des TCG. L'indicateur 4.1.2 est défini comme le « pourcentage d'une cohorte d'enfants ou de jeunes âgés de 3 à 5 ans au-dessus de l'âge prévu pour la dernière année de chaque niveau d'enseignement qui ont terminé cette année ». Le taux d'achèvement est une mesure de « flux » de réussite, qui prend en compte les inscriptions tardives et les redoublements élevés dans de nombreux pays les plus pauvres, qui conduisent de nombreux enfants à atteindre la fin de chaque cycle éducatif plusieurs années après l'âge officiel d'obtention du diplôme.

La combinaison de plusieurs sources de données d'enquête peut résoudre certains problèmes, tels que les cycles d'enquête peu fréquents (tous les trois à cinq ans) et diverses erreurs d'échantillonnage et non-échantillonnage qui génèrent des informations contradictoires entre différentes enquêtes dans le même pays. Toutefois, les enquêtes présentent également des avantages par rapport aux données administratives, comme un meilleur enregistrement des informations sur l'âge et la couverture universelle des programmes éducatifs. Il est également possible d'utiliser des informations rétrospectives pour reconstituer les taux d'achèvement historiques des cohortes plus âgées et de ne pas se limiter aux informations sur la cohorte la plus récente.

Le modèle développé est un modèle hiérarchique bayésien inspiré de l'approche utilisée pour estimer les indicateurs de santé mais adapté au contexte éducatif ([Dharamshi et al., 2022](#)). Il estime une tendance sous-jacente des valeurs cibles et partage des informations sur la mise à

l'échelle des paramètres entre les pays. Le retard d'achèvement est explicitement modélisé en précisant l'ampleur du retard en fonction de l'âge. Les problèmes de déclaration erronée de l'âge sont également abordés. De tels ajustements permettent au modèle de consolider les données de l'enquête en une tendance sous-jacente lisse du taux d'achèvement à partir de laquelle les taux d'achèvement annuels réels estimés pour chaque pays peuvent être extraits. En répondant aux divers problèmes de qualité des données associés aux données d'enquête, ces estimations sont également moins sensibles aux enquêtes individuelles, à l'année au cours de laquelle elles ont été menées et au type d'enquête qui se trouve être la dernière disponible dans un pays donné.

Des estimations ponctuelles continuent d'être communiquées pour des combinaisons de pays individuels et d'années d'enquête dans la base de données de l'ISU. Mais l'ISU fournit également les estimations du modèle parallèlement à ces estimations ponctuelles, tandis que le Groupe de coopération technique (TCG) sur les indicateurs de l'ICG 4 a approuvé l'utilisation des estimations du modèle pour les agrégats régionaux et mondiaux pour la base de données des TCG.

Le taux d'achèvement est la contrepartie, basée sur une enquête, d'un indicateur basé sur des données administratives, le taux brut d'admission en dernière année d'école. Une éventuelle extension future du modèle pourrait consister à combiner des données d'enquête et des données administratives (comme dans le cas du modèle de taux de non-scolarité présenté ci-après).

2.2. UN MODÈLE DE TARIFICATION EXTRASCOLAIRE

Le taux de non-scolarisation est la « proportion d'enfants et de jeunes appartenant à la tranche d'âge officielle pour un niveau d'enseignement donné qui ne sont pas inscrits aux niveaux d'enseignement préscolaire, primaire, secondaire ou supérieur » ([UIS, 2021](#)). Il s'agissait de l'indicateur phare en 2000-2015 du programme d'éducation pour tous et des objectifs du Millénaire pour le développement. La nécessité d'une méthodologie combinant les sources de données pour estimer les taux de non-scolarisation a été reconnue il y a 20 ans, lorsqu'il a été

reconnu qu'« une sorte d'approche composite pourrait être nécessaire pour estimer des séries chronologiques et produire des estimations pour l'année la plus récente » ([UIS et UNICEF, 2005](#)).

En l'absence d'une telle « approche composite », le calcul des taux de non-scolarité s'est basé sur les registres d'inscription issus des recensements scolaires. Cependant, l'utilisation de données administratives présente trois défis dans les pays les plus pauvres où les taux de non-scolarisation sont élevés. Premièrement, les dossiers d'inscription sont souvent incomplets, inexacts ou totalement absents. Deuxièmement, les estimations doivent combiner les chiffres des inscriptions avec une mesure de la population, qui provient d'une source différente et souvent incohérente. La qualité des estimations de population par âge unique, nécessaires pour calculer avec précision les taux de non-scolarité, n'est souvent pas élevée, ce qui conduit à des sauts dans la série chronologique des indicateurs – et parfois à un plus grand nombre d'enfants inscrits que le nombre d'enfants de cette population tranche d'âge. Troisièmement, avec de faibles taux d'enregistrement des naissances, la capacité des écoles à enregistrer avec précision l'âge des élèves est limitée.

Ces dernières années, bon nombre de ces pays ont mené des enquêtes auprès des ménages qui, malgré leurs propres faiblesses, peuvent contribuer à combler certaines lacunes et à relever les défis liés à l'âge et à la population. Un modèle bayésien hiérarchique basé sur une cohorte a donc été développé pour estimer les taux de non-scolarisation à l'aide de plusieurs sources de données ([ISU et Rapport GEM, 2022](#)). Le modèle reflète la progression naturelle des élèves tout au long du cycle scolaire. Les données provenant de sources administratives et d'enquêtes sont rapprochées, reconnaissant les différences dans la nature et la génération de données de ces deux types, tout en partageant des informations sur les biais et la variance entre les pays.

Le modèle introduit de nouvelles idées clés. Premièrement, il utilise une approche de cohorte pour relier les taux de non-scolarité, à l'instar de la modélisation démographique des processus démographiques. Cependant, c'est en raison du manque de données fiables que les risques d'entrée tardive, d'abandon scolaire, de redoublement et d'autres transitions éducatives pertinentes ne sont pas estimés ; au lieu de cela, les changements nets d'une année à l'autre sont estimés de manière flexible pour lisser les courbes sous-jacentes des cohortes de taux de non-scolarité. Deuxièmement, le modèle accepte les cas où il y a plus d'enfants scolarisés que

d'enfants d'un âge particulier, mais limite en même temps les taux de non-scolarisation entre 0 et 100 % afin de permettre l'utilisation de ces informations. Troisièmement, le modèle déplace l'attention des taux de non-scolarisation par niveau d'éducation vers les taux de non-scolarisation par âge, puisque les élèves entrent et sortent de l'école à tout âge.

Les résultats de ce modèle ont été rapportés pour la première fois en septembre 2022, y compris pour de nombreux pays qui ne disposaient pas de données administratives sur les taux de non-scolarisation depuis de nombreuses années. Même si les estimations des données administratives restent les données nationales officiellement déclarées dans la base de données de l'ISU, comme pour le taux d'achèvement, l'ISU fournit également des estimations modèles pour chaque pays, tandis que les estimations modèles constituent la source privilégiée pour les agrégats régionaux et mondiaux. Une mise à jour de septembre 2023 estimait qu'il y aurait 250 millions d'enfants, d'adolescents et de jeunes en 2022.

Malgré l'utilisation efficace de toutes les informations disponibles, le modèle est également soumis à ses propres défis. La principale est que, le modèle étant animé par une approche de cohorte et caractérisé par un certain degré d'inertie, il ne peut pas facilement intégrer un impact majeur sur la fréquentation scolaire d'une année sur l'autre provoqué par les situations d'urgence. D'autres défis concernent la cohérence et sont abordés dans la section 4.

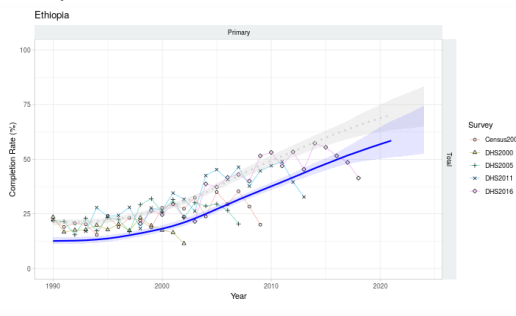
2.3. VISUALISATION DES ESTIMATIONS DU MODÈLE DE TAUX D'ACHÈVEMENT ET DE TAUX D'ABANDON SCOLAIRE

Les résultats des deux modèles sont disponibles sur le site Web VIEW, géré par le Rapport mondial de suivi sur l'éducation et l'Institut de statistique de l'UNESCO et mis à jour deux fois par an. Le site visualise les données d'entrée, invitant à un examen plus approfondi des problèmes statistiques résolus. Dans le cas du taux d'achèvement, la visualisation met en évidence le défi des inscriptions tardives. Un grand nombre d'enfants achèvent chaque niveau d'enseignement même après 3 à 5 ans au-dessus de l'âge prévu pour la dernière année. En fait, le modèle estime non seulement le taux d'achèvement officiel, mais également un taux d'achèvement « ultime », qui inclut ceux qui terminent leurs études avec jusqu'à 8 ans de retard (**Figure 2a**). Dans le cas

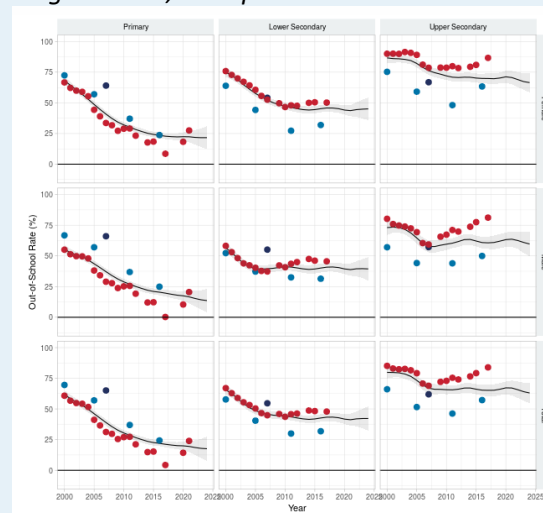
du taux de non-scolarisation, la visualisation met en évidence chaque source de données avec une couleur distincte, ce qui permet d'identifier la source des écarts (Figure 2b).

Figure 2. Visualisation des données d'entrée et des estimations du modèle pour deux indicateurs de l'éducation

a. Taux d'achèvement du primaire, Éthiopie



b. Taux de non-scolarisation, par tranche d'âge et sexe, Éthiopie



Source : <https://education-estimates.org/>.

3. DÉFIS ASSOCIÉS À L'ESTIMATION D'INDICATEURS À L'AIDE DE PLUSIEURS SOURCES DE DONNÉES

Les efforts récents visant à combiner plusieurs sources d'information pour estimer et communiquer les indicateurs de l' TCG 4 sont des mesures importantes prises par l'UNESCO pour s'acquitter de sa responsabilité d'informer la communauté internationale grâce à une utilisation efficace de données provenant de sources multiples. Néanmoins, les efforts continus visant à intégrer les données doivent également garantir que les meilleures pratiques soient suivies, que les besoins des pays soient satisfaits et que des estimations cohérentes soient calculées. Cette section décrit brièvement ces défis.

3.1. GARANTIR LES MEILLEURES PRATIQUES EN MATIÈRE DE REPORTING D'ESTIMATIONS BASÉES SUR PLUSIEURS SOURCES DE DONNÉES

Les données sur l'éducation sont rarement disponibles pour chaque population et chaque année, tandis que les données provenant de sources multiples et les différences dans les méthodologies de mesure peuvent donner lieu à des estimations incohérentes et non comparables. Malgré un engagement clair selon lequel les pays doivent diriger les rapports sur les TCG, le recours à des modèles peut être inévitable pour avoir une idée de la façon dont les indicateurs clés évoluent. Pourtant, les données et les méthodologies utilisées pour produire des estimations comportent souvent des caractéristiques ou des hypothèses qui affectent leur interprétation. Une interprétation précise et une utilisation responsable des estimations nécessitent de comprendre sur quoi reposent les données utilisées pour fonder les estimations (comment elles ont été identifiées, consultées et incluses ; quelle est leur qualité) et des méthodes complexes utilisées pour obtenir les estimations.

3.2. ASSURER LA PARTICIPATION ET L'APPROPRIATION DES PAYS DANS LA GÉNÉRATION DES ESTIMATIONS

Même si les meilleures pratiques sont suivies, d'autres protocoles sont nécessaires pour garantir que les pays qui, comme mentionné ci-dessus, sont censés diriger les rapports sur les TCG, puissent participer et s'engager dans les processus d'estimation. De tels protocoles sont nécessaires pour faciliter l'appropriation par les pays des estimations générées et pour aider à développer la capacité des statisticiens et des experts nationaux à contribuer aux améliorations du processus. Actuellement, il n'existe aucun mécanisme systématique permettant aux pays de demander des éclaircissements, de comprendre les méthodes qui sous-tendent les estimations et de contester les résultats qui contredisent leur propre compréhension de la situation réelle, ainsi que de contribuer de manière proactive aux sources de données et aux idées pour le développement des modèles.

3.3. ASSURER LA COHÉRENCE ENTRE LES ESTIMATIONS DU TAUX D'ABANDON SCOLAIRE ET DU TAUX D'ACHÈVEMENT

Les modèles d'achèvement et d'éducation extrascolaire visent à résoudre les incohérences entre et au sein de plusieurs sources de données et types de sources. Chacune de ces estimations est cohérente en interne : les estimations du taux d'achèvement pour une année donnée sont cohérentes avec les estimations passées et les estimations du taux de non-scolarisation à un âge donné sont cohérentes avec les taux de non-scolarisation d'une cohorte aux âges précédents. Mais il reste d'autres problèmes de cohérence à résoudre. Premièrement, les taux pour les femmes et les hommes sont actuellement estimés indépendamment les uns des autres et peuvent arbitrairement diverger. Deuxièmement, les taux d'achèvement et de non-scolarisation sont estimés indépendamment, même s'ils ne sont pas indépendants les uns des autres : chaque enfant qui termine l'école doit avoir passé un certain nombre d'années inscrit, ce qui impose une contrainte sur le nombre d'enfants qui peuvent avoir été non scolarisés école.

4. FUTUR AGENDA

Les défis décrits dans ce document laissent entrevoir des solutions potentielles qui peuvent s'appuyer sur les progrès réalisés jusqu'à présent. Ces solutions visent également à renforcer la communauté de pratique des statisticiens de l'éducation.

4.1. FORMALISER LES BONNES PRATIQUES EN MATIÈRE DE REPORTING DES ESTIMATIONS

Confrontée à des défis similaires liés à la disponibilité de sources multiples, la communauté internationale des statistiques de la santé, sous la direction de l'Organisation mondiale de la santé, a publié les Lignes directrices pour des rapports précis et transparents sur les estimations de la santé ([GATHER](#)) afin de définir les meilleures pratiques de déclaration pour les études qui estiment les indicateurs en utilisant plusieurs sources de données. Le processus a impliqué un groupe de travail qui a examiné les directives de déclaration existantes, généré une liste d'éléments de déclaration potentiels, reçu les commentaires d'une communauté plus large de chercheurs et d'utilisateurs d'estimations, et rédigé une liste de contrôle essentielle de 18 éléments qui devraient être déclarés chaque fois que des estimations sont publiées, afin de répondre aux besoins des décideurs et des chercheurs ([Stevens et al., 2016](#)).

La [liste de contrôle GATHER](#) des informations qui doivent être incluses dans les rapports d'estimations mondiales est organisée en quatre sections : (1) objectifs et financement, (2) sources de données, (3) analyse des données et (4) résultats et discussion. L'intention est que les informations sur les sources de données et les méthodes d'analyse, y compris les hypothèses et limites clés, soient présentées de manière accessible sans formation avancée en statistiques. Des explications sur la manière dont les nouvelles estimations se comparent aux estimations publiées précédemment et sur les raisons pour lesquelles elles diffèrent doivent être fournies.

Les lignes directrices visent également à promouvoir le libre accès aux données et au code source, dans la mesure où une documentation complète augmente la valeur de la recherche. Les données sous-jacentes aux estimations doivent être accessibles en ligne, ce qui peut nécessiter des ressources supplémentaires pour la documentation et l'archivage. Les exceptions incluent les cas où cela n'est pas possible, comme par exemple la propriété de tiers. Le partage du code source implique également un investissement de ressources, surtout si le code est entièrement documenté et disponible en ligne pour une utilisation standard. Au minimum, les segments clés du code doivent être partagés, mais les chercheurs ne doivent pas être tenus responsables de fournir une assistance aux utilisateurs.

Les auteurs doivent également fournir une mesure de l'incertitude associée aux estimations, comme les intervalles d'incertitude. Les estimations mondiales sont affectées par de multiples sources d'erreur, telles que les erreurs de mesure lors de la collecte des données, l'incapacité d'obtenir un échantillon véritablement aléatoire, les erreurs d'ajustement des données d'entrée pour tenir compte des sources de biais et l'utilisation d'un modèle pour calculer les estimations. Les utilisateurs de ces estimations doivent être informés de leur incertitude globale. Les lignes directrices en matière de déclaration sont conçues pour être suffisamment flexibles pour guider la déclaration des estimations, indépendamment de la disponibilité des données sous-jacentes et de la complexité des méthodes statistiques.

Il est proposé que :

- Le TCG lance un processus pour discuter des problèmes émergents liés à l'utilisation de modèles pour estimer les indicateurs de l'ITCG 4.

- [De GATHER](#) doivent être adaptées du secteur de la santé au secteur de l'éducation.

4.2. SOUTENIR LA PARTICIPATION ET L'APPROPRIATION DES PAYS DANS LES ESTIMATIONS

Le développement de modèles pour estimer les indicateurs de l'éducation a été initié par l'UNESCO. Cependant, maintenant que ces exemples sont disponibles, il est temps pour les pays d'examiner les résultats de manière systématique, de se familiariser avec la justification et les implications, d'identifier les erreurs et de demander des éclaircissements, de contribuer des idées aux domaines potentiels de développement du modèle et de fournir des informations supplémentaires et plus approfondies. -sources de données à jour. Familiariser les ministères de l'Éducation et la communauté des experts avec les indicateurs de l'TCG 4 basés sur des estimations en tant que nouveau moyen de suivre les progrès nécessite une communication approfondie.

Dans les statistiques sanitaires internationales, comme dans les estimations des indicateurs de mortalité, il existe des processus visant à renforcer les capacités nationales en matière de collecte de données, de compréhension des techniques d'estimation et d'interprétation des résultats. Des ateliers régionaux ont été organisés pour former des participants de différents pays aux techniques et méthodes de modélisation qui sous-tendent les estimations. Des experts ont été envoyés dans les pays pour dispenser une formation sur l'estimation de la mortalité infantile. Dans le cadre d'un processus d'examen des données, l'OMS et l'UNICEF, par l'intermédiaire de leurs bureaux extérieurs, consultent les gouvernements, qui fournissent des commentaires sur la plausibilité des estimations et la validité des données sous-jacentes. Ces efforts reposent sur le principe selon lequel l'estimation des indicateurs n'est pas simplement un exercice académique mais un élément fondamental d'une politique et d'une programmation efficaces.

Il est proposé que des mesures similaires soient également prises dans les statistiques de l'éducation, par lesquelles :

- L'Institut de statistique de l'UNESCO inclut des estimations modèles dans l'ordre du jour de son programme existant d'ateliers régionaux de développement des capacités pour familiariser les pays avec les méthodes.

- L'élaboration d'un inventaire des enquêtes (l'une des solutions proposées dans le document de conférence sur les enquêtes auprès des ménages) sert également à l'intégration des données, garantissant que les pays sont impliqués dans les entrées de données utilisées.

4.3. DÉVELOPPER UN MODÈLE COMMUN DE TAUX D'ABANDON SCOLAIRE ET D'ACHÈVEMENT

Le Rapport mondial de suivi sur l'éducation et l'Institut de statistique de l'UNESCO travaillent actuellement à l'élaboration d'un modèle informatiquement réalisable qui intègre et assure la cohérence des estimations des taux d'achèvement et de non-scolarité entre elles (et potentiellement avec les estimations d'autres indicateurs connexes), désagrégées par sexe.

Un modèle conjoint peut évaluer les écarts de manière systématique et aider à ajuster les taux d'achèvement et de non-scolarisation les uns par rapport aux autres : les estimations conjointes devraient être plus précises que les estimations indépendantes. À ce jour, les estimations du taux d'achèvement (qui s'appuient uniquement sur des sources d'enquête) ont été plus précises que les estimations du taux de non-scolarité (qui s'appuient sur des sources d'enquête et administratives). Cela implique que les taux de non-scolarisation sont plus susceptibles de changer considérablement dans certains pays dans le contexte d'un modèle commun.

Les estimations des taux d'achèvement et de non-scolarité sont liées les unes aux autres via la progression par âge et par niveau, c'est-à-dire le comportement d'entrée, de redoublement et d'abandon par âge et par niveau. En théorie, les estimations conjointes pourraient être améliorées en utilisant des données sur ces flux ; les estimations des taux d'achèvement et de non-scolarisation pourraient également servir à éclairer les estimations de progression. En pratique, aucune de ces deux ambitions n'est susceptible d'être réalisable. La couverture et la qualité des données sur la progression des classes d'âge au niveau des années d'âge ou des classes individuelles sont médiocres. Un modèle conjoint ne devrait donc s'appuyer que sur des hypothèses faibles concernant la progression, en utilisant occasionnellement des éléments d'information pertinents, par exemple sur les politiques de promotion automatique et de zéro redoublement. Il n'identifie ni n'exige pas non plus tous les éléments d'un modèle de progression complet.

Il est proposé de développer un modèle d'achèvement conjoint et extrascolaire, en tenant compte de ce qui suit :

- Tous les niveaux de scolarisation devraient être modélisés conjointement pour tenir compte correctement des effets d'une entrée tardive à l'école.
- Le modèle conjoint doit être suffisamment simple pour que la charge de calcul reste gérable.
- Le modèle devrait être structuré en termes d'implications en termes d'inscriptions du conditionnement sur les parts des différents groupes de finissants (ou non-achevés) et de ceux qui ne sont jamais entrés à l'école.
- Ces implications des tendances en matière d'inscription et d'achèvement peuvent être modélisées avec divers degrés de flexibilité (par exemple, un effet unique d'entrée tardive par rapport à différents taux d'entrée tardive de différents groupes de finissants ; niveaux constants d'entrée tardive ou évolution des tendances d'entrée tardive au fil du temps).
- Dans la mesure du possible, il conviendrait d'explorer la possibilité que ce modèle aide également à estimer d'autres indicateurs connexes (par exemple les taux de scolarisation).

4.4. DÉVELOPPER DES MODÈLES POUR ESTIMER D'AUTRES INDICATEURS QUI S'APPUIENT SUR PLUSIEURS SOURCES DE DONNÉES

Bien que les taux d'achèvement et de non-scolarisation aient été prioritaires pour l'élaboration du modèle, il ne s'agit là que de deux d'un ensemble plus large d'indicateurs qui pourraient bénéficier de l'utilisation systématique de multiples sources et types de données. Il est proposé que le TCG explore la possibilité de développer des modèles pour estimer d'autres indicateurs de l'TCG 4, qui peuvent s'appuyer sur :

- Plusieurs sources de données, notamment :
 - Pourcentage d'enfants en trop grand âge pour la classe (indicateur 4.1.5)
 - Taux de participation des jeunes et des adultes à l'éducation et à la formation formelles et non formelles au cours des 12 mois précédents, par sexe (indicateur 4.3.1)
- Plusieurs types de sources de données, notamment :
 - Taux de participation à l'apprentissage organisé (un an avant l'âge officiel d'entrée dans le primaire), par sexe (indicateur 4.2.2)

- Taux brut de scolarisation de la petite enfance dans (a) l'éducation préprimaire et (b) et le développement éducatif de la petite enfance (indicateur 4.2.4)
- Taux brut de scolarisation dans l'enseignement supérieur par sexe (indicateur 4.3.2).

Les taux d'alphabétisation des jeunes et des adultes (indicateur 4.6.2) entrent également dans cette dernière catégorie. Bien qu'un modèle existe déjà pour cet indicateur, un examen et une extension potentielle pourraient être justifiés.

RÉFÉRENCES

Alkema L. et coll. (2016). Global, regional, and national levels and trends in maternal mortality between 1990 and 2015, with scenario-based projections to 2030: a systematic analysis by the UN Maternal Mortality Estimation Inter-Agency Group. *Lancet*. 387 (10017), pages 462-474.

[https://www.thelancet.com/journals/lancet/article/PIIS01406736\(15\)00838-7/fulltext](https://www.thelancet.com/journals/lancet/article/PIIS01406736(15)00838-7/fulltext)

Alkema L. et New, JR (2014). Global estimation of child mortality using a Bayesian B-spline bias-reduction model. *Annals of Applied Statistics* 8 (4), pages 2122-2149.

<https://www.jstor.org/stable/24522377>

Dharamshi, A., Barakat, B., Alkema L. et Antoninis M. (2022) A Bayesian model for estimating Sustainable Development Goal indicator 4.1.2: School completion rates. *Journal of the Royal Statistical Society Series C: Applied Statistics*, 71 (5), pages 1822–1864.

<https://academic.oup.com/jrsssc/article/71/5/1822/7073267>

Stevens G. et coll. (2016) Guidelines for Accurate and Transparent Health Estimates Reporting: the GATHER statement. *Lancet* 388 (10062) E19-E23.

[https://www.thelancet.com/journals/lancet/article/PIIS0140-6736\(16\)30388-9/fulltext](https://www.thelancet.com/journals/lancet/article/PIIS0140-6736(16)30388-9/fulltext)

Rapport ISU et GEM (2022). *A Bayesian cohort model for estimating SDG indicator 4.1.4: Out-of-school rates*. Paris: UNESCO. https://www.unesco.org/gem-report/sites/default/files/medias/fichiers/2022/08/OOS_Proposal.pdf

UNICEF, OMS et Banque mondiale (2023). Levels and trends in child malnutrition: UNICEF / WHO / World Bank Group Joint Child Malnutrition Estimates: Key findings of the 2023 edition. New York: United Nations Children's Fund and World Health Organization.

<https://iris.who.int/bitstream/handle/10665/368038/9789240073791-eng.pdf?sequence=1>

UIS (2021). Metadata 4.1.4 Out-of-school rate (1 year before primary, primary education, lower secondary education, upper secondary education). Montreal: UNESCO Institute for Statistics. <https://tcg.uis.unesco.org/wp-content/uploads/sites/4/2021/09/Metadata-4.1.4.pdf>

UIS et UNICEF (2005). Children out of school: Measuring exclusion from primary education. Montreal: UNESCO Institute for Statistics.

https://uis.unesco.org/sites/default/files/documents/children-out-of-school-measuring-exclusion-from-primary-education-en_0.pdf

SITES INTERNET

VIEW. Visualizing Indicators of Education for the World. Global Education Monitoring Report and UNESCO Institute for Statistics. <https://education-estimates.org/>

WHO. Guidelines for Accurate and Transparent Health Estimates Reporting (GATHER). World Health Organization. <https://www.who.int/data/gather/statement/>