



Modèles de risque pour prédire le travail des enfants

Un aperçu des différentes approches visant à identifier les enfants à risque de travail des enfants dans le cacao

Septembre 2021



International
COCOA
Initiative

Protéger les enfants et leur famille dans les communautés productrices de cacao

La fondation International Cocoa Initiative est une organisation associative à but non lucratif qui œuvre pour améliorer la vie des enfants et des adultes des communautés productrices de cacao. Nous sommes experts dans le domaine du travail des enfants et du travail forcé dans le secteur cacaoyer. Nous conseillons les gouvernements et les entreprises de manière à les guider dans leurs pratiques et à exercer une influence dans leurs processus de décision, et nous travaillons avec les ONG actives dans le secteur. Nous nous engageons à atteindre une production durable du cacao qui protège les droits des enfants et des adultes à l'échelle mondiale.

Soutenu par :



Cette étude a été conçue et développée par l'International Cocoa Initiative (ICI) et a été rédigée par Laurent Foubert. Des membres de l'ICI et de la *Deutsche Gesellschaft für Internationale Zusammenarbeit* (GIZ) y ont apporté leurs contributions.

Nous remercions les partenaires suivants, qui ont partagé des informations sur les projets de modélisation du risque de travail des enfants, et qui ont accepté de les inclure en tant qu'études de cas dans ce rapport : Cargill, ECOM, Nestlé et Tony's Chocolonely.

Image de couverture : session de sensibilisation au travail des enfants à Didoko, Côte d'Ivoire, © Nestlé, 2019.

www.cocoainitiative.org | info@cocoainitiative.org

Secrétariat ICI en Suisse

Chemin de Balxert 9,
1219 Châtelaine | Switzerland
+41 22 341 47 25

Bureau national ICI en Côte d'Ivoire

Il Plateaux, 7ème Tranche, Lot 3244, Ilot 264,
Abidjan-Cocody | Côte d'Ivoire
+225 27 22 52 70 97

Bureau national ICI au Ghana

No. 16, Djanie Ashie Street,
East-Legon | Accra | Ghana
+233 302 998 870

Résumé

Des progrès considérables ont été réalisés ces dernières années dans la création de modèles de risque capables de prédire le travail des enfants au niveau individuel ou des ménages. Ce rapport décrit différentes approches utilisées pour développer et tester les modèles de risque de travail des enfants dans le contexte des zones productrices de cacao en Afrique de l'Ouest. Ces modèles ont pour but d'améliorer la manière de cibler les interventions visant à prévenir et combattre le travail des enfants là où elles sont les plus nécessaires. Ils s'inscrivent dans l'effort plus large visant à mettre à l'échelle les mesures de protection des enfants contre le travail des enfants et à garantir le respect de leurs droits fondamentaux.

Qu'est-ce qu'un modèle de risque ?

Un modèle de risque est une approche statistique qui sert à **prédire un résultat** (p. ex. le travail des enfants) **pour une unité d'observation donnée** (p. ex. un enfant ou un ménage) **à partir d'un ensemble d'indicateurs**.

La première étape consiste à « calibrer » le modèle en appliquant des méthodes statistiques (p. ex. régression logistique, régression multiple, modèle de classes latentes, etc.) à un ensemble de données contenant des informations sur le résultat et sur les indicateurs pour une population similaire à la population cible. La deuxième étape consiste à alimenter le modèle avec des informations concernant les indicateurs observés au sein de la population cible, afin d'obtenir une prévision pour chaque unité. Si la prévision est un résultat binaire (p. ex. si l'enfant est astreint au travail des enfants), alors la prévision peut être interprétée comme « risque » (ou possibilité) que l'incident se concrétise.

Cette étude aborde les questions suivantes :

1. Quelles sont les caractéristiques des modèles de risque qui ont été développés à ce jour dans le secteur cacaoyer pour prédire le travail des enfants, et quels sont leurs résultats ?
2. Quelles leçons ont été tirées de ces expériences et quelles recommandations ont été formulées ?

Méthodologie

Ce rapport se base sur six projets déployés par l'ICI, ses membres et d'autres acteurs du secteur cacaoyer, qui utilisent des modèles de risque pour prédire le statut de travail des enfants ou des ménages en Côte d'Ivoire et au Ghana. Certains de ces projets ont été développés dans le cadre d'essais opérationnels, afin d'accroître la rentabilité et la possibilité de mise à l'échelle des activités de suivi et de remédiation du travail des enfants. D'autres ont été développés afin de générer des connaissances théoriques concernant le potentiel de modélisation du risque de travail des enfants. Les modèles utilisent des sources de données différentes, notamment les enquêtes nationales de

prévalence du travail des enfants¹, des registres d'agriculteurs tenus par les coopératives, et des données recueillies par les systèmes de suivi et de remédiation du travail des enfants (SSRTE). De plus, ils appliquent des méthodes statistiques différentes, y compris l'apprentissage automatique (Machine Learning) et des prédictions basées sur des régressions.

Suite à un examen de tous les projets, nous mettons en évidence les leçons principales tirées et proposons des recommandations pour développer des modèles de risque efficaces et pour déployer des approches basées sur le risque. Pour chacune des six approches examinées dans ce rapport, nous fournissons également des informations sur le contexte spécifique du projet, ses objectifs, la méthode utilisée pour prédire le travail des enfants et la performance des modèles.

Résultats

Les études de cas présentées ici montrent que la création d'un bon modèle de risque commence par des données solides, c'est-à-dire précises, fiables et à jour. Des outils de collecte de données et des recenseurs bien formés sont essentiels, aussi bien pour recueillir les données initiales nécessaires pour développer les modèles, et pour ensuite tester si les prédictions générées reflètent la situation sur le terrain. Sans données de haute qualité, même la « méthode statistique parfaite » ne suffira pas pour obtenir un modèle hautement efficace. Ceci signifie qu'avant qu'un modèle de risque soit développé et déployé à grande échelle, il pourrait d'abord être nécessaire de renforcer les capacités de collecte et de gestion des données, en particulier dans les cas où les données sont recueillies par des agriculteurs, des membres de la communauté ou des coopératives dont la collecte de données n'est pas le travail principal.

Le développement d'un modèle de risque de travail des enfants ne devrait pas se transformer en quête de « l'ensemble parfait de variables prédictives ». Un large panel de variables différentes peut être utilisé pour prédire efficacement le risque de travail des enfants, en fonction du contexte spécifique et de la disponibilité des données. Il est important de noter que les modèles de risque efficaces utilisent des facteurs qui n'ont *pas de lien causal* avec le travail des enfants, et qui *pourraient ne pas avoir de relation statistiquement significative* avec la prévalence du travail des enfants. Ceci dit, de nombreux modèles de risque intègrent certains facteurs qui sont considérés comme faisant partie des causes profondes du travail des enfants. Les facteurs habituellement utilisés dans les modèles de risque qui sont associés à un risque inférieur de travail des enfants comprennent l'accès à une main-d'œuvre adulte, l'eau potable et l'éducation de qualité.

Les études de cas présentées dans ce rapport prouvent que la modélisation du risque devrait adopter une approche basée sur l'enfant. Comme le montrent les exemples, la présence d'informations de base concernant l'enfant et son ménage augmente toujours la capacité d'un modèle à prédire avec précision le travail des enfants.

Au-delà de ces considérations, l'utilité des approches basées sur le risque pour un acteur donné dépendra de la stratégie, du contexte et des contraintes auxquels il fait face. Tandis que les études

¹ Voir par exemple: «Évaluation des progrès accomplis dans la réduction du travail des enfants dans les régions productrices de cacao de Côte d'Ivoire et du Ghana», Sadhu, S., Kysia, K., Onyango, L., Zinnes, C.F., Lord, S., Monnard, A. and Arellano, NORC at the University of Chicago, 2020. Cette étude est appelée «Étude NORC» dans la suite de ce rapport.

de cas présentées dans ce rapport prouvent que les modèles de risque peuvent être utilisés pour réduire le nombre de ménages visités dans le cadre d'un système de suivi et de remédiation du travail des enfants ou pour fournir en priorité un soutien préventif aux ménages à risque, il existe de nombreux autres cas d'utilisation – voire d'autres solutions plus appropriées. Pour chaque situation dans laquelle un modèle de risque de travail des enfants est utilisé, il est impératif de déterminer qui le modèle vise à inclure ou exclure, quelles stratégies sont en place pour les individus ou les ménages n'étant pas considérés comme à risque, à quelle fréquence l'évaluation du risque devrait être réalisée et comment l'efficacité des modèles employés peut continuellement être examinée et améliorée.



Étant donné qu'il n'existe pas d'approche universelle pour concevoir et utiliser des modèles de risque de travail des enfants, il est particulièrement important que toutes les parties prenantes continuent de partager les informations relatives à leurs approches, leurs résultats et leurs conclusions dans ce domaine, pour permettre aux autres de s'inspirer de leur travail, l'objectif général étant de mettre à l'échelle les activités de prévention et de lutte contre le travail des enfants.

Comment les modèles de risque sont-ils utilisés ?

Les modèles de risque peuvent être utilisés pour une multitude de buts. La liste suivante comprend des exemples tirés de projets existants et d'autres qui n'ont pas encore été testés :

- Pour identifier les ménages ayant le risque le plus élevé de travail des enfants (p. ex. parmi les membres d'une coopérative ou d'une communauté), afin qu'ils reçoivent en priorité des visites d'observation ou du soutien
- Pour identifier les coopératives ou les communautés présentant le risque le plus élevé de travail des enfants, de manière à ce que les interventions soient ciblées sur les zones les plus à risque

- Pour élargir le champ d'intervention de manière à inclure les ménages « à risque », ainsi que ceux où du travail des enfants a été identifié
- Pour renforcer les actions d'observation de routine du travail des enfants – par exemple, dans le contexte d'un SS RTE – afin d'augmenter les chances d'identifier des cas de travail des enfants (p. ex. en raison du caractère cyclique du travail des enfants et de la possible réallocation du travail entre les enfants d'un même ménage, en réponse à des chocs ou à un changement de circonstances)

Considérations pratiques lors du développement et de l'utilisation des modèles de risque de travail des enfants

- Pour développer et calibrer un modèle de risque, **au moins un ensemble de données existant est nécessaire**, qui comprenne le **phénomène étudié** (p. ex. si le ménage a recours ou non au travail des enfants) et **toutes les variables utilisées comme indicateurs** (p. ex. informations détaillées concernant les ménages et leurs membres), tirées d'un échantillon qui soit comparable à la population cible ou qui la recoupe.
- Pour utiliser un modèle afin de prédire le risque de travail des enfants, **les données relatives à la population cible doivent être complètes et précises**. Les approches basées sur le risque ne peuvent pas être utilisées pour évaluer le risque parmi des individus/ménages pour qui les données sont incomplètes ou imprécises.
- **Des capacités de gestion des données et d'analyse statistique sont nécessaires**, à l'interne ou à l'externe, afin de calibrer et de perfectionner un modèle de risque, de manière à ce qu'il puisse répondre aux besoins et aux contraintes opérationnelles de chaque acteur le déployant.

Leçons et recommandations

Leçons

- **Les modèles hautement performants peuvent réduire le nombre de ménages ciblés**, tout en **garantissant que la plupart des enfants contraints de travailler soient atteints** : si elle est appliquée dans le contexte de l'observation du travail des enfants au niveau des ménages – et indépendamment de la performance du modèle – une approche basée sur le risque peut induire une diminution de 50 % du nombre de visites d'observation, tout en identifiant plus de 95 % des enfants astreints au travail des enfants. Ceci aurait pour résultat une diminution importante des coûts d'observation, tout en continuant d'identifier la grande majorité des cas.
- **Il n'existe pas de modèle universel adapté à tous les cas** : les modèles de risque peuvent et doivent être adaptés au contexte dans lequel ils sont utilisés, aux données disponibles et aux contraintes opérationnelles.
- **Des données de qualité récoltées auprès de la population ciblée sont essentielles pour générer des prévisions utiles** : les erreurs de mesure et les données obsolètes diminuent la précision du modèle, quelle que soit la méthode statistique, tandis que les données manquantes excluent les individus ou les ménages de la couverture basée sur le risque. Pour

cette raison, renforcer les capacités de collecte et de gestion des données des parties prenantes locales pourrait être un prérequis avant d'utiliser une approche basée sur le risque.

- **Des informations sur les enfants individuels, comme leur âge et leur sexe, améliorent la qualité des prévisions** : incorporer l'âge, le sexe et le statut scolaire de l'enfant améliore systématiquement la précision des modèles visant à prédire le travail des enfants.
- **Le statut de travail des enfants doit être prédit au niveau du ménage, plutôt qu'au niveau de l'enfant** : cette approche permet d'identifier plus efficacement les cas de travail des enfants et facilite la planification des visites d'observation.
- **Une évaluation récurrente (annuelle) du risque de travail des enfants à l'aide d'un modèle prédictif est probablement plus efficace qu'une activité ponctuelle** : chaque modèle de risque implique un certain niveau d'erreur, ce qui signifie que certains enfants astreints au travail dangereux des enfants passeront entre les mailles du filet. Les modèles de risque hautement performants qui incluent l'âge de l'enfant parmi leurs indicateurs semblent avoir la capacité d'identifier, à la longue, les cas n'ayant pas été identifiés par le passé, grâce à plusieurs tours successifs d'évaluation. Ainsi, évaluer le risque pour une population ciblée sur une base annuelle à l'aide des données les plus récentes pourrait non seulement permettre aux enfants de 5 à 17 ans qui étaient précédemment absents des données d'être pris en compte, mais cela pourrait également mener à l'identification d'enfants qui sont passés entre les mailles du filet, puisque leur note de risque change à mesure qu'ils grandissent.
- **Le potentiel d'économies d'une approche basée sur le risque dépend de plusieurs facteurs** : ces facteurs comprennent la prévalence du travail des enfants, la performance du modèle et l'efficacité des mesures de soutien. Dans les zones où la prévalence de travail des enfants est plus élevée, les économies qu'un modèle de risque pourrait potentiellement engendrer sont plus faibles, mais elles pourraient augmenter avec le temps, grâce à un soutien correctement ciblé qui réduirait le taux de prévalence et augmenterait les économies potentielles.

Recommandations

Avant de considérer la mise en place d'une approche basée sur le risque, il convient :

- **De définir clairement le but du modèle de risque** – par exemple, une identification plus rapide des cas ou une diminution des visites d'observation – **et les contraintes opérationnelles, et de développer le modèle en conséquence**. Les modèles prédictifs comprendront toujours un degré d'incertitude ; même le meilleur modèle « manquera » une certaine part de cas. Cet aspect doit être clairement compris et des mesures d'atténuation doivent être mises en place pour gérer les conséquences opérationnelles et éthiques des choix faits.
- **D'évaluer la capacité technique et le temps disponible** pour développer un modèle. Un système de gestion des données efficace est nécessaire, ainsi que du personnel qualifié capable de le faire fonctionner et de mettre régulièrement à jour le modèle prédictif.
- **D'évaluer la disponibilité de données récentes, complètes et de haute qualité**. Si l'un de ces trois critères n'est pas rempli, ces données ne doivent pas être utilisées pour développer un modèle de risque. Au lieu de cela, les efforts doivent d'abord être concentrés sur la

collecte de données de qualité. S'il est prévu d'intégrer l'approche dans des structures locales (groupes d'agriculteurs, coopératives, etc.), leurs pratiques en matière de collecte et de gestion des données doivent être évaluées. Sans la capacité nécessaire pour récolter et maintenir des ensembles de données complets et précis, développer un modèle de risque n'aura aucune utilité.

Durant le développement d'un modèle de risque, il convient :

- **De se concentrer sur des indicateurs faciles à recueillir et faciles à évaluer**, et de s'assurer que tout a été mis en œuvre pour réduire les erreurs de mesure et les données manquantes.
- **De limiter le nombre d'indicateurs** utilisés par le modèle à environ 10–12², afin de faciliter la récolte de données et l'accès à des données de haute qualité.
- **D'utiliser des données fiables pour calibrer le modèle**, idéalement issues d'un grand échantillon dont la couverture recoupe autant que possible la zone géographique de la population ciblée.
- **D'incorporer les indicateurs de base relatifs aux enfants (p. ex. l'âge, le sexe et le statut scolaire)** au modèle de risque afin d'en augmenter la performance.
- **De configurer le modèle de manière à ce qu'il prédise le risque pour le ménage, plutôt que pour l'enfant**, car cela améliore la capacité d'un modèle à identifier les cas de travail des enfants et le rend plus pratique au niveau opérationnel.
- Lorsque cela est possible, **de choisir soigneusement une valeur de seuil** en fonction de sa distribution, du contexte local et des connaissances préalables (p. ex., le taux de prévalence national), ainsi que de la stratégie et des contraintes opérationnelles.

Lorsque le modèle est **en cours d'utilisation**, il convient :

- Lorsque cela est possible, **d'exécuter le modèle de risque chaque année de manière récurrente**, en utilisant des données actualisées, afin qu'il puisse, au fil du temps, identifier tous les enfants à risque.
- **De constamment évaluer la performance d'un modèle de risque** par rapport aux taux de prévalence connus et aux résultats des visites d'observation et d'ajuster le modèle si les taux d'identification sont plus faibles que prévu.

² Les modèles de risque hautement performants décrits dans ce rapport utilisent entre 10 et 12 indicateurs. Il n'existe pas de preuve qu'augmenter le nombre d'indicateurs améliore les modèles.