

Cote du document:	EC 2020/111/W.P.6
Ordre du jour:	8
Date:	24 septembre 2020
Distribution:	Publique
Original:	Anglais

F



Investir dans les populations rurales

Examen des produits d'autoévaluation du IFAD

Note à l'intention des membres du Comité de l'évaluation

Responsables:

Questions techniques:

Thomas Eriksson

Directeur

Division des politiques et des résultats opérationnels

téléphone: +39 06 5459 2425

courriel: t.eriksson@ifad.org

Sara Savastano

Directrice

Division recherche et évaluation de l'impact

téléphone: +39 06 5459 2155

courriel: s.savastano@ifad.org

Transmission des documents:

Deirdre Mc Grenra

Cheffe

Gouvernance institutionnelle

et relations avec les États

membres

téléphone: +39 06 5459 2374

courriel: gb@ifad.org

Comité de l'évaluation — Cent onzième session
Rome, 22 octobre 2020

Pour: **Examen**

Table des matières

Sigles et acronymes	ii
I. Introduction et rôle de l'autoévaluation	1
II. Évolution du système d'autoévaluation au FIDA	2
III. Le système d'autoévaluation et ses produits au FIDA	3
A. Le système d'autoévaluation	3
B. Les produits actuels d'autoévaluation du FIDA et les changements proposés	4
C. Nouveaux produits proposés	10
D. Amélioration du système	10
IV. Des produits à la pratique: promouvoir l'apprentissage grâce à l'évaluation	11
A. Mobiliser les savoirs pour devenir une organisation en apprentissage permanent	11
B. Collaborer avec IOE pour renforcer l'apprentissage	12
C. Exploiter la technologie pour tirer des enseignements de l'évaluation	13
V. La marche à suivre et les prochaines étapes	13
 Appendices	
I. International financial institutions comparators	
II. Self-evaluation supporting systems and processes	

Sigles et acronymes

COSOP	Programme d'options stratégiques pour le pays
ESPP	Évaluation de la stratégie et du programme de pays
EvPP	Évaluation de la performance des projets
FIDA11	Onzième reconstitution des ressources du FIDA
IFI	Institution financière internationale
IOE	Bureau indépendant de l'évaluation du FIDA
ODD	Objectif de développement durable
ONS	Opération non souveraine
PRIME	Programme relatif au suivi-évaluation rural
PRISMA	Rapport annuel du Président sur la mise en œuvre des recommandations issues de l'évaluation et sur les suites données par la direction
RAP	Rapport d'achèvement de projet
RARI	Rapport annuel sur les résultats et l'impact des opérations du FIDA
RIDE	Rapport sur l'efficacité du FIDA en matière de développement
SGRO	Système de gestion des résultats opérationnels
VRAP	Validation des rapports d'achèvement de projet

I. Introduction et rôle de l'autoévaluation

1. **Le système d'autoévaluation du FIDA s'inscrit dans l'architecture de l'évaluation à l'échelle du FIDA.** Le système d'autoévaluation permet à la direction i) de s'acquitter de ses responsabilités en rendant compte au Conseil d'administration de sa performance et de ses résultats, et ii) de fournir au Bureau indépendant de l'évaluation (IOE) les éléments nécessaires pour qu'il puisse valider les autoévaluations du FIDA. En plus de fournir un cadre solide de mesure des résultats en matière de reddition de comptes, le système d'autoévaluation constitue également la base de la gestion des performances opérationnelles ainsi que de l'apprentissage et du partage des savoirs.
2. **Le présent document répond à une recommandation de l'examen externe indépendant par les pairs** autorisé par le Conseil d'administration. Il examine l'état des produits d'autoévaluation existants et propose des améliorations avant d'aborder les nouveaux produits d'évaluation pour refléter les résultats de l'évolution du modèle opérationnel du FIDA. Le document examine les moyens par lesquels le FIDA peut faire de l'apprentissage l'un des principaux moteurs de son action; il examine également comment il peut être disposé à adopter l'innovation au moyen de mesures d'incitation, pour une prise de risque éclairée et dans un souci de franchise, et comment il peut se préparer à adopter une culture de gestion adaptative pour s'adapter rapidement à l'évolution des circonstances, notamment à la pandémie de COVID-19.
3. **Cadre relatif à l'efficacité en matière de développement.** L'utilisation de données probantes comme base des décisions sur la conception et l'exécution des projets, programmes et stratégies est au cœur du Cadre stratégique du FIDA 2016-2025. C'est pourquoi le Cadre relatif à l'efficacité en matière de développement a été élaboré en 2016. Il comprend cinq ensembles d'activités fournissant une approche globale et cohérente de la mesure des résultats, et notamment de l'impact, étayés par une théorie du changement. L'objectif principal du Cadre relatif à l'efficacité en matière de développement est de veiller à ce que la prise de décisions fondées sur des données probantes soit intégrée aux stratégies de pays et aux cycles des projets.
4. **La fonction d'évaluation.** La fonction d'évaluation du FIDA est un complément essentiel du Cadre relatif à l'efficacité en matière de développement pour ce qui est d'évaluer la pertinence et la concrétisation du Cadre stratégique. L'évaluation a trois objectifs principaux: i) promouvoir la responsabilité en documentant les résultats; ii) encourager la gestion de la performance, en utilisant des décisions fondées sur des données probantes; iii) renforcer les mécanismes d'apprentissage et de retour d'information.
5. À l'instar d'autres institutions financières internationales (IFI), le Conseil d'administration, IOE et la direction ont chacun un rôle essentiel, mais spécifique, à jouer dans l'évaluation. La surveillance et la responsabilité institutionnelles reviennent au Conseil. La direction a pour responsabilité d'ensemble de conduire des autoévaluations, afin de mesurer et d'évaluer systématiquement la performance, les résultats et les impacts des stratégies et des projets, et d'en rendre compte au Conseil. L'évaluation indépendante, et notamment les validations réalisées par l'IOE, fournit un examen critique des données probantes, des résultats et des évaluations de l'autoévaluation. Elle est indépendante des responsables de la conception et de l'exécution des activités. Ensemble, ces éléments se renforcent mutuellement, l'évaluation promouvant une meilleure prise de décisions fondées sur des données probantes, appuyant une responsabilisation renforcée, améliorant l'apprentissage et favorisant une culture plus fortement axée sur les résultats, ce qui constitue le cœur du Cadre relatif à l'efficacité en matière de développement.

6. Le rapport de l'examen externe par les pairs de 2019 a constaté que "la structure et le fonctionnement d'ensemble du système d'évaluation au FIDA [...] étaient généralement comparables à ceux d'institutions financières internationales (IFI) du même type". Il a également mis en évidence les domaines dans lesquels le système d'évaluation à l'échelle du FIDA doit s'adapter pour suivre le rythme de l'évolution des pratiques d'évaluation dans d'autres IFI (voir appendice I) et à l'ONU. Une constatation clé, partagée par d'autres, est que l'apprentissage et le retour d'information sur l'évaluation n'ont pas été pleinement réalisés. Le rapport a recommandé une évolution stratégique pour faire de l'optimisation des résultats par l'apprentissage un objectif fondamental de l'évaluation au FIDA.
7. En ce qui concerne la fonction d'autoévaluation de la direction, l'examen externe par les pairs a constaté que les éléments de base étaient bien en place à la suite de la Onzième reconstitution des ressources du FIDA (FIDA11), mais qu'il restait encore une importante marge d'amélioration.

II. Évolution du système d'autoévaluation au FIDA

8. Depuis l'approbation de la Politique de l'évaluation, en 2011, le système d'autoévaluation du FIDA a mûri, et la direction a systématiquement amélioré la mesure de l'impact de ses opérations. L'architecture de l'autoévaluation au FIDA se distingue de celle de la plupart des autres IFI en cela qu'elle est intégrée au Cadre relatif à l'efficacité en matière de développement, lui-même trouvant sa source dans le Cadre stratégique.
9. Le Cadre relatif à l'efficacité en matière de développement a pour objectif de garantir que: i) les stratégies de pays déterminent les projets et les activités hors prêts qui abordent les principaux problèmes de développement sous-jacents; ii) les projets individuels sont conçus dès le départ avec une théorie du changement soigneusement formulée indiquant comment l'on prévoit que les intrants et les activités auront un impact; iii) les projets peuvent être évalués pour permettre des évaluations ultérieures de leur impact; iv) les systèmes de suivi-évaluation et de collecte de données sont suffisamment planifiés lors de la conception et appliqués en même temps que le projet pour s'assurer que la collecte des informations est effectuée de manière adéquate. Le Cadre relatif à l'efficacité en matière de développement et la matrice d'efficacité en matière de développement qui l'accompagne servent également d'outils pour évaluer la qualité du projet au point de départ.
10. Depuis 2016, date de la mise en place du Cadre relatif à l'efficacité en matière de développement, le FIDA a réalisé d'importants progrès dans la mise en place des éléments fondamentaux de la prise de décisions fondée sur des données probantes, et notamment:
 - a) l'intégration d'une focalisation sur les résultats tout au long des cycles des programmes de pays et des projets, y compris dans la théorie du changement, les cadres logiques et les systèmes de suivi-évaluation;
 - b) l'amélioration de la capacité de suivi-évaluation dans le secteur rural, par exemple le Programme relatif au suivi-évaluation rural (PRIME), une initiative de formation et de certification;
 - c) la capacité d'établir un lien entre le suivi-évaluation d'un projet, y compris les indicateurs de base, et la communication des résultats institutionnels harmonisés sur le Cadre stratégique et les objectifs de développement durable (ODD);
 - d) l'établissement d'un Système de gestion des résultats opérationnels (SGRO) pour saisir les données sur les produits et les effets directs sur la base des indicateurs de base convenus;

- e) l'utilisation de données probantes en temps réel dans la gestion du portefeuille à l'échelon régional et national et au niveau des projets, en employant des tableaux de bord de gestion.
- 11. Pris ensemble, ces éléments de base ont permis au système d'autoévaluation du FIDA de devenir la pierre angulaire de l'architecture des résultats de l'organisation et le fondement de toute la fonction d'évaluation au FIDA (appendice II).

III. Le système d'autoévaluation et ses produits au FIDA

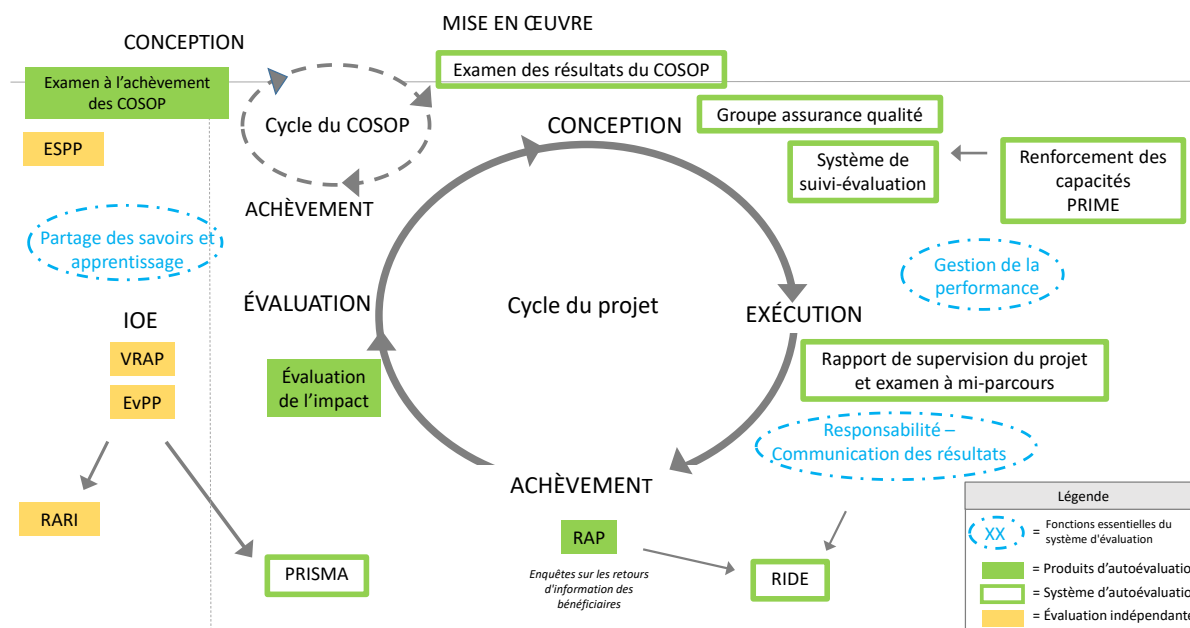
A. Le système d'autoévaluation

- 12. La définition du Comité d'aide au développement (CAD) de l'Organisation de coopération et de développement économiques (OCDE) est la suivante: "Un système d'autoévaluation est une combinaison de processus et de produits pour le suivi-évaluation systématique, empirique et transparent d'un projet, d'un programme ou d'une stratégie, en cours ou achevé, réalisé par le département des opérations responsable de cette activité."
- 13. Le système d'autoévaluation du FIDA est fondé sur des évaluations et des examens conduits en interne par la direction pour mesurer l'impact institutionnel du FIDA dans le système global de suivi-évaluation.
- 14. La figure 1 ci-dessous montre comment le système d'autoévaluation s'insère dans l'architecture globale d'évaluation. Elle montre également comment les produits d'autoévaluation font partie intégrante des cycles des programmes d'options stratégiques pour le pays (COSOP) et des projets du FIDA. Les systèmes de suivi-évaluation sous-jacents, les conceptions de projet basées sur le Cadre relatif à l'efficacité en matière de développement, l'assurance qualité et le suivi permanent des progrès à l'aide du SGRO sont également fondamentaux. Le mécanisme d'autoévaluation permet au FIDA de gérer la performance, de rendre compte des résultats et d'encourager l'apprentissage et le partage des savoirs.
- 15. La direction dispose de deux catégories de produits d'autoévaluation:
 - i) les rapports d'achèvement: les rapports d'achèvement des projets (RAP) et les examens à l'achèvement des COSOP;
 - ii) les évaluations de l'impact: elles portent sur un échantillon de 15% du portefeuille, comme l'a demandé le Conseil d'administration. Les différentes études sont agrégées suivant une méthode rigoureuse pour produire le rapport global d'évaluation de l'impact à la fin de chaque période de reconstitution.
- 16. Comme le montre la figure 1, d'autres examens sont effectués dans le cadre du système d'autoévaluation. Ce ne sont pas des produits d'autoévaluation formels, mais des outils de gestion et/ou des rapports de synthèse.
- 17. Les conclusions des autoévaluations constituent la base des rapports sur les résultats adressés par la direction au Conseil d'administration. Ces études comprennent le Rapport annuel sur l'efficacité du FIDA en matière de développement (RIDE) et le rapport d'évaluation de l'impact, qui détaille l'impact institutionnel des opérations en vue des reconstitutions. La direction rend également compte au Conseil d'administration de la mise en œuvre des recommandations d'IOE dans le Rapport annuel du Président sur la mise en œuvre des recommandations issues de l'évaluation et sur les suites données par la direction (PRISMA).
- 18. Le système d'évaluation indépendante d'IOE et ses produits s'appuient sur les données et les produits générés par le système d'autoévaluation. IOE fournit une gamme de produits, dont quatre correspondent directement à des produits d'autoévaluation: les validations des rapports d'achèvement de projet (VRAP), les évaluations de la performance des projets (EvPP), les évaluations de stratégie et de

programme de pays (ESPP), et les évaluations de l'impact. IOE s'appuie également sur les données collectées par le système d'autoévaluation pour produire son Rapport annuel sur les résultats et l'impact des opérations du FIDA (RARI).

Figure 1

Architecture de l'évaluation au FIDA



19. Bien que le système d'autoévaluation comporte de nombreux éléments et s'appuie sur les rapports de supervision sous-jacents et les systèmes de suivi-évaluation au niveau des projets, trois produits de base sont examinés plus en détail ci-dessous : les RAP, les rapports d'achèvement des COSOP et les évaluations de l'impact. La raison en est que ces produits correspondent plus directement aux produits d'évaluation et fournissent des données probantes sur les projets achevés et les stratégies de pays. Des synthèses d'autoévaluations comme le RIDE et le PRISMA sont également examinées.

B. Les produits actuels d'autoévaluation du FIDA et les changements proposés

Rapports d'achèvement de projet

20. Les RAP doivent être envisagés, de préférence, dans le contexte de la totalité du cycle du projet à partir de la phase de la conception (figure 1). L'examen de la qualité au point de départ a pour objet de s'assurer que les rapports de conception de projet décrivent les résultats attendus et le suivi-évaluation à utiliser pour recueillir les données sur les résultats. Les progrès sont suivis et évalués pendant l'exécution et consignés dans les rapports de supervision de projet et les examens à mi-parcours. Le suivi de l'information permet d'éventuelles corrections de trajectoire en cours d'exécution.
21. Il est essentiel, pour l'intégrité et l'efficacité du système d'évaluation à l'échelle du FIDA, que la qualité des RAP soit élevée. La direction donne une importance accrue aux RAP de qualité en distinguant, chaque année, le meilleur d'entre eux et en attribuant une précieuse récompense monétaire à une mission de supervision, à utiliser pour les activités de suivi-évaluation. Chaque année, entre 25 et 30 RAP sont soumis au FIDA par les emprunteurs, comme stipulé dans les Conditions générales du FIDA applicables au financement du développement agricole. Bien que la responsabilité en incombe aux gouvernements, les équipes du FIDA sont étroitement associées à la préparation des RAP et fournissent un renforcement des capacités et un appui technique. C'est aussi chaque gouvernement qui prend en

charge le coût de la préparation du RAP, généralement inclus dans le coût du projet, avec une contribution du FIDA d'environ 20 000 à 25 000 USD pour ses activités d'appui à l'achèvement. Les RAP doivent normalement être soumis six mois après la date d'achèvement du projet.

22. Le processus de préparation du RAP peut commencer dès lors que 85% du financement du projet sont décaissés. Le succès d'un RAP impliquant une analyse approfondie des données sur l'impact et les résultats, des enquêtes dédiées peuvent être nécessaires pour compléter les informations provenant du suivi-évaluation du projet et collecter des données essentielles. Le Cadre des retours d'information opérationnels des parties prenantes, approuvé par le Conseil d'administration du FIDA en décembre 2019, prévoit que les retours d'information se poursuivent tout au long du cycle du projet. Dans le pays, les retours d'information des parties prenantes sur la performance du projet, obtenus dans le cadre de larges consultations, sont recueillis à l'achèvement. La direction a l'intention de renforcer, durant FIDA12, la robustesse de la boucle des retours d'information des parties prenantes en introduisant un nouvel indicateur de base spécifique.
23. Le FIDA suit la méthode d'évaluation du CAD de l'OCDE pour apprécier la performance des projets à l'achèvement, en utilisant des critères communs (pertinence, efficacité, efficience et durabilité). Le FIDA applique en outre d'autres critères, conformément à ses stratégies, et notamment: l'impact sur la pauvreté rurale, la gestion de l'environnement et des ressources naturelles, l'adaptation aux changements climatiques et l'égalité femmes-hommes. Ces évaluations basées sur les notations fournissent une indication solide de la performance pendant la durée de vie d'un projet. L'approche est harmonisée avec celle d'IOE.
24. Les évaluations des notations contenues dans les RAP sont complétées par des résultats, effets directs et produits quantifiables, fondés sur le Cadre relatif à l'efficacité en matière de développement et reflétés dans le SGRO. De plus en plus, le FIDA utilise des indicateurs de base agrégés, mis en correspondance avec les objectifs stratégiques du FIDA et alignés sur les ODD. La mesure directe des effets est une importante rupture méthodologique par rapport aux jugements relatifs à la performance, souvent formulés sans preuves adéquates à l'appui. Le cadre des indicateurs de base de 2017 a efficacement intégré la gestion des résultats dans les systèmes de suivi-évaluation des projets, tandis qu'en 2020 de nouvelles directives ont établi une méthode obligatoire pour assurer la cohérence entre les projets. En outre, le FIDA vient de mettre au point une méthode de réalisation d'enquêtes sur les effets directs, qui s'appliquera à l'ensemble du portefeuille. Ensemble, ces outils augmenteront encore la solidité des effets directs communiqués à compter de FIDA11.
25. La ponctualité de la soumission des RAP par les gouvernements s'est améliorée. À l'heure actuelle, le retard moyen est de 2,7 mois par rapport au délai de six mois prévu pour la soumission des RAP, alors qu'il était de trois ans en 2016. Un processus d'examen rigoureux est en place et, si les gouvernements donnent leur approbation, sera mis en ligne sur le site web externe du FIDA, auquel ont accès les représentants au Conseil d'administration. Le FIDA s'est engagé à publier 90% des RAP d'ici à la fin de FIDA11; à ce jour, 74% ont déjà été affichés.
26. Dans ses validations de tous les RAP (VRAP), IOE fournit une notation globale des RAP basée sur la franchise, les enseignements, la portée et la qualité. À cet égard, le FIDA est aligné sur d'autres IFI, qui valident 100% des rapports d'achèvement ou un échantillon statistiquement représentatif de ceux-ci. Depuis 2014, environ 90% des RAP sont jugés satisfaisants, comme le montre le RARI 2020. L'écart d'ensemble s'est rétréci au fil du temps et se situe à un tiers de point de pourcentage.

Propositions de modifications concernant les RAP

27. **Le FIDA est un cas atypique par rapport aux autres IFI** car, dans d'autres institutions, ce sont les gouvernements qui sont responsables de la préparation des RAP. En revanche, la participation active du FIDA au processus complique la définition des rôles et des responsabilités et, parfois, sape la franchise concernant la performance globale du projet. Il arrive aussi, dans certains cas, que les notations et les évaluations respectives des gouvernements et du FIDA présentent des différences. En outre, des conflits d'intérêts peuvent survenir lorsque le FIDA examine les notes attribuées par le gouvernement à la performance du FIDA. La responsabilité du gouvernement dans la préparation des RAP peut avoir un impact sur la ponctualité et la divulgation complète des rapports, surtout s'il y a des désaccords sur le contenu. Pour résoudre ces problèmes, d'autres IFI ont assumé la responsabilité de la préparation des RAP – conformément à la définition de l'autoévaluation – tout en prévoyant une section pour les rapports d'achèvement du gouvernement et/ou ses observations.
28. **Assumer la responsabilité de la préparation des RAP.** Pour améliorer la qualité, la franchise, la ponctualité et la transparence des RAP, la direction propose d'adopter la pratique suivie par d'autres IFI et de transférer officiellement au FIDA la responsabilité de la préparation des RAP. Les gouvernements devraient encore fournir des rapports d'achèvement, à condition qu'ils soient conformes aux spécifications énoncées dans les Conditions générales du FIDA (ils pourraient toutefois être plus courts et se concentrer sur les questions présentant un intérêt particulier pour le gouvernement). Les coûts des RAP étant à l'heure actuelle largement à la charge des gouvernements, il conviendrait aussi d'allouer des budgets adéquats.
29. **Rationaliser les notations des RAP.** La direction propose de poursuivre la rationalisation des notations des RAP. Comme dans les autres IFI, les RAP se concentreraient sur l'autoévaluation et la notation: i) de la réalisation des effets en matière de développement; ii) de la performance du FIDA (y compris la qualité de la conception et de l'appui à la supervision); iii) de la performance de l'emprunteur/du bénéficiaire. Ces changements seraient reflétés dans de nouvelles directives relatives aux RAP.
30. Si cette approche est adoptée, la direction collaborera avec IOE afin de garantir une bonne harmonisation avec sa gamme de produits et ses méthodes de notation en évolution. Il conviendrait également de déterminer quelle serait la place la plus appropriée pour les notations des RAP dans le Cadre de gestion des résultats. Néanmoins, les RAP continueront de collecter des données et des éléments probants sur les effets en matière de développement et la gestion du projet; ces domaines ne feront simplement pas l'objet de notations subjectives.
31. **Actualiser les directives et les systèmes.** Il conviendra, une fois achevés le transfert de la responsabilité des RAP au FIDA et l'examen conjoint des notations, d'actualiser les directives et les systèmes des RAP. Ces directives devront aussi refléter l'évolution de la culture de l'évaluation et des données factuelles au sein du FIDA, ainsi que les améliorations des systèmes de suivi-évaluation dans l'ensemble de l'institution.

Examen des COSOP à l'achèvement

32. En 2018, les stratégies de pays sont devenues obligatoires pour tous les États membres recevant des allocations au titre du financement de base du FIDA, soit sous la forme de COSOP, soit de notes de stratégie de pays préparés par les divisions régionales, avec des contributions du Département de la stratégie et des savoirs et d'autres unités du FIDA. Les directives révisées relatives aux stratégies de pays exigent aussi la réalisation d'examens à l'achèvement des COSOP, qui constituent une autoévaluation des objectifs stratégiques d'un COSOP et de la performance du FIDA à l'aune de ces objectifs. Les examens à l'achèvement des

COSOP sont préparés dans les six mois suivant la fin d'un cycle de COSOP. Comme convenu avec IOE, ils suivent une méthode standard pour évaluer les programmes de pays, y compris les activités de projets et les activités hors projets. Les résultats sont mesurés par rapport aux indicateurs définis dans le cadre de résultats des COSOP.

33. Tous les examens à l'achèvement des COSOP comprennent des notations de l'efficacité et de la pertinence globales des COSOP ainsi que des activités hors prêts. Les examens à l'achèvement des COSOP sont autorisés par le directeur régional et approuvés par le Vice-Président adjoint du Département de la gestion des programmes et soumis au Conseil d'administration en même temps qu'un nouveau COSOP. Ils tirent de l'exécution des COSOP des enseignements pratiques et en temps réel qui peuvent éclairer la conception d'un nouveau COSOP. Ils sont rendus publics et contribuent à la base de savoirs du FIDA. Ils seront agrégés au niveau de l'institution et seront consignés dans le RIDE à la fin de FIDA11.
34. À la différence des RAP, les examens à l'achèvement des COSOP ne sont pas validés par IOE. Le Bureau réalise par contre des ESPP dans certains pays, qui couvrent généralement une période de 10 ans, s'étendant sur plus d'un cycle de COSOP. Si elles sont disponibles, les ESPP et leurs recommandations peuvent apporter des éclairages utiles aux nouveaux COSOP. Souvent, toutefois, le calendrier des examens à l'achèvement des COSOP et des ESPP n'est pas aligné. L'accord conclusif de l'ESPP est signé par le gouvernement client et la direction, alors que les COSOP eux-mêmes ne le sont pas. Le FIDA est la seule IFI où l'accord conclusif, s'il est disponible, est joint en appendice à un COSOP.

Propositions de modifications concernant les COSOP/examens à l'achèvement des COSOP

35. **Modifications relatives aux COSOP.** Lors de la préparation des nouveaux COSOP, les retours d'information et les enseignements tirés des précédentes interventions dans le pays seront renforcés, notamment les conclusions des évaluations de l'impact. À cet égard, la direction estime qu'une meilleure harmonisation du calendrier des examens à l'achèvement des COSOP, des ESPP d'IOE et des nouveaux COSOP renforcera les boucles d'apprentissage et de retour d'information. La direction propose que l'on fasse coïncider, si possible, les discussions dans le pays à l'achèvement d'un COSOP avec le gouvernement et d'autres parties prenantes, et l'atelier d'IOE sur l'ESPP.
36. En ce qui concerne l'évolution de la gamme de produits d'IOE, la direction envisage avec intérêt une discussion sur l'opportunité et la pertinence de l'accord conclusif dans le cadre d'une ESPP, et sur le point de savoir si un produit similaire à une VRAP pour les COSOP pourrait être approprié, en utilisant et en validant les données disponibles dans l'examen d'achèvement du COSOP.
37. La direction prévoit également de déployer en 2021 un module SGRO plus complet pour le cycle de vie du COSOP, ce qui facilitera un meilleur suivi du calendrier et de l'exécution des produits du COSOP, y compris leur synchronisation avec les rapports d'IOE. Cet outil permettra également à la direction de mieux évaluer la performance des programmes de pays en temps réel et, en particulier, de suivre les contributions à l'élaboration des politiques au niveau des pays.

Évaluations de l'impact

38. En appliquant les recommandations du Cadre relatif à l'efficacité en matière de développement, le FIDA adopte une approche rigoureuse pour la production de données factuelles de haute qualité sur l'impact imputable à ses investissements. Les évaluations de l'impact au niveau des projets, menées par la Division recherche et évaluation de l'impact, remplissent une double fonction: tirer des enseignements des projets pour éclairer les nouvelles conceptions de projets et les RAP, et fournir

une estimation de l'impact global des indicateurs institutionnels énoncés dans le Cadre stratégique.

39. En conséquence, le Conseil d'administration a demandé au FIDA de réaliser des évaluations de l'impact sur un groupe choisi de projets représentant 15% du portefeuille du Fonds. Les évaluations de l'impact sont conçues en étroite collaboration avec les parties prenantes locales et les homologues gouvernementaux pour garantir une compréhension approfondie du projet et des enseignements pertinents. Il s'agit là d'une approche systématique et rigoureuse unique qui permet au FIDA de tirer effectivement des enseignements et de communiquer ses résultats, et ainsi de contribuer à la réalisation des ODD.
40. Comme l'ont suggéré IOE et le Comité de l'évaluation, la méthode des évaluations de l'impact a été validée au moyen d'une analyse de sensibilité. L'évaluation de l'impact de FIDA11 a été imaginée en s'inspirant des enseignements tirés de FIDA10: elle a été conçue pour produire 24 études, soit 21% des projets en phase de clôture, un échantillon suffisamment large pour tenir compte de l'attrition potentielle et, plus récemment, des projets abandonnés en raison des restrictions dues à la COVID-19 et à d'autres facteurs. Pour FIDA12, enfin, il est prévu que la direction et IOE travailleront ensemble pour mettre en place le cadre de référence et l'approche générale pour la conduite de l'évaluation de l'impact.
41. Le FIDA est la seule IFI capable de rendre compte de l'impact global de ses opérations durant un cycle de reconstitution; les autres présentent leurs rapports projet par projet. La méthode d'estimation de l'efficacité globale en matière de développement comporte deux étapes. On effectue, premièrement, une agrégation au moyen d'une méta-analyse systématique des estimations de l'évaluation de l'impact au niveau de chaque projet pour calculer les impacts institutionnels globaux. On procède ensuite à une projection pour extrapoler les impacts au reste du portefeuille et estimer le nombre de personnes bénéficiant de chaque objectif stratégique (à savoir augmentation de la production, accès ou participation au marché, et résilience) et atteignant l'objectif global du FIDA – une mobilité économique accrue. Dans le cadre de FIDA11, la direction élargira l'analyse d'impact des thèmes à transversaliser du FIDA.
42. Les évaluations de l'impact ne sont pas validées par IOE. Au lieu de cela, IOE examine les rapports de synthèse préparés par la direction sur les évaluations de l'impact réalisées pour chaque reconstitution. Durant FIDA10, IOE a noté que les projets ayant fait l'objet d'une évaluation de l'impact avaient généralement de meilleurs RAP. Cela n'est pas surprenant et s'explique par le fait que ces RAP intégraient des données probantes provenant des évaluations de l'impact, ce qui se traduisait par des rapports de meilleure qualité et, par conséquent, des notes plus élevées. Pour faciliter l'intégration des conclusions des évaluations de l'impact dans le RAP, les projets choisis pour faire l'objet d'une évaluation de l'impact ont bénéficié de six mois supplémentaires pour la préparation des rapports.

Modifications proposées ou futures concernant les évaluations de l'impact

43. L'approche des évaluations de l'impact met principalement l'accent sur les évaluations de l'impact ex post en raison de la nécessité de rendre compte des résultats et de fournir une estimation des impacts institutionnels pour chaque reconstitution. Pour accroître les enseignements tirés des évaluations de l'impact, la direction a l'intention d'adopter des mesures visant à combler le fossé entre les évaluations de l'impact et d'autres produits d'autoévaluation dans le cadre de FIDA12. La direction améliorera les possibilités d'apprentissage des évaluations de l'impact ex post, qui sont effectuées à une date proche de l'achèvement du projet, grâce à une meilleure intégration de leurs conclusions dans d'autres parties du cycle du projet. Les enseignements tirés des évaluations de l'impact feront l'objet d'une étude systématique lors d'examens choisis de manière stratégique de la conception et de l'exécution de projets, en particulier les examens à mi-parcours.

Enfin, de futures évaluations de l'impact pourraient être effectuées pour évaluer l'impact d'interventions innovantes (par exemple, prêts axés sur les résultats, stratégie de collaboration avec le secteur privé, Mécanisme de relance en faveur des populations rurales pauvres).

44. Pour renforcer les capacités, la direction parachève actuellement les directives et les matériels d'apprentissage en ligne relatifs aux évaluations de l'impact à l'intention des parties prenantes, ainsi que les plans en vue d'une diffusion plus systématique et stratégique des conclusions des évaluations de l'impact.
45. Enfin, pour accroître la transparence, la direction favorisera les données ouvertes en mettant à la disposition du public des ensembles de données recueillies durant les évaluations de l'impact. Il conviendra d'élaborer au préalable une politique en matière de gouvernance des données, afin de garantir des normes éthiques et de protection de la vie privée concernant les données relatives aux bénéficiaires du FIDA.

Rapport sur l'efficacité du FIDA en matière de développement

46. Le RIDE est le principal instrument utilisé par le FIDA pour rendre compte chaque année de sa performance institutionnelle et opérationnelle en regard du Cadre de gestion des résultats convenu. Il n'est de ce fait pas conçu comme un produit d'autoévaluation, mais pour offrir une vue d'ensemble actuelle et complète de la performance du Fonds en s'appuyant sur une large gamme de données émanant de sources aussi bien quantitatives que qualitatives. Comme indiqué dans l'examen externe par les pairs, on peut s'attendre à ce que les progrès accomplis par la direction en matière de gestion des résultats et de production de rapports sur la performance se reflètent dans le RIDE. Ainsi, à mesure que se renforcent les produits d'autoévaluation, le RIDE lui-même devient un outil plus solide. Dans le RIDE 2019, la direction a également inclus des données provenant du portefeuille en cours pour présenter un compte rendu actualisé et global.
47. **Rationaliser le RIDE.** Sous sa forme actuelle, le RIDE rend compte d'un très grand nombre d'engagements et d'activités du FIDA. La direction a l'intention de rationaliser le RIDE pour se concentrer sur l'efficacité et les résultats en matière de développement. Il est proposé que les engagements (par exemple les progrès en matière de transversalisation) fassent l'objet d'un document distinct à soumettre au Conseil d'administration.

Rapport du Président sur la mise en œuvre des recommandations issues de l'évaluation et sur les suites données par la direction (PRISMA)

48. En 2015, la Politique de l'évaluation prévoyait que, afin d'assurer des retours d'information appropriés, le Président du FIDA devrait veiller à ce que les recommandations formulées par IOE dans ses évaluations soient adoptées aux niveaux opérationnel et stratégique. Le PRISMA annuel, préparé par la direction et transmis au Conseil d'administration, se compose de deux parties: le volume I contient un résumé et regroupe les mesures de suivi relatives aux recommandations récurrentes; le volume II comprend des précisions sur les suites données à chacune des recommandations. IOE prépare des observations écrites sur le PRISMA, qui sont également communiquées au Comité de l'évaluation et au Conseil. Dans sa forme actuelle, le PRISMA est davantage axé sur la responsabilité que sur l'apprentissage, ce qui ajoute encore au déséquilibre entre ces deux fonctions d'évaluation.
49. **Choisir la juste taille pour le PRISMA.** Le PRISMA est un instrument clé pour le suivi de la responsabilité au niveau de l'institution. D'autres IFI étudient actuellement les moyens par lesquels les recommandations issues d'évaluations indépendantes, la réponse de la direction et l'état d'avancement de leur mise en œuvre sont suivis et davantage axés sur l'apprentissage, sans toutefois perdre de vue la fonction de responsabilité. La direction estime que le volume I pourrait avoir un impact beaucoup plus fort et stratégique s'il suivait et rendait compte de l'état

d'avancement d'un nombre plus restreint de recommandations clés d'IOE. Le volume II de l'actuel PRISMA pourrait être supprimé, ou publié entièrement en ligne, la responsabilité du suivi des recommandations approuvées d'IOE restant à la charge des divisions opérationnelles compétentes, et faisant l'objet d'un suivi en temps réel.

C. Nouveaux produits proposés

Autoévaluation des activités dans le cadre des opérations non souveraines

50. L'approbation de la Stratégie de collaboration avec le secteur privé et d'un cadre pour les opérations non souveraines (ONS) permettra au FIDA de prendre appui sur le secteur privé dans le cadre de son programme de travail. Les procédures régissant les ONS de la conception à l'achèvement sont fondées sur des opérations souveraines. Toutefois, les opérations du secteur privé présentent des caractéristiques particulières, notamment le besoin de confidentialité, là où un RAP standard n'est pas réalisable. Étant donné qu'il sera important pour le FIDA de tirer des enseignements des ONS au cours de la période de FIDA11 puis de FIDA12, des produits d'autoévaluation modifiés, notamment des évaluations de l'impact pour les ONS, devront être élaborés. Ces produits devront s'inspirer de l'expérience d'autres IFI (notamment la Société financière internationale, la Banque européenne pour la reconstruction et le développement, la Banque interaméricaine de développement et la Banque asiatique de développement), mais devraient être adaptés au mandat spécifique du FIDA.

Examens de groupes de projets

51. L'autoévaluation des projets s'est jusqu'ici concentrée sur de simples opérations d'investissement souveraines. Pour favoriser l'apprentissage et aborder des questions d'intérêt stratégique, le FIDA conduira un essai pilote de regroupement d'examens d'achèvement de projets et d'évaluations de l'impact autour de groupes de projets similaires (par exemple, thèmes à transversaliser, types de projets). Pour saisir des enseignements en temps réel, ces examens pourraient également couvrir des opérations en cours. Cela offrirait des synergies avec les nouvelles évaluations, par IOE, de groupes de projets (voir section IV).

Examen de la durabilité après l'achèvement

52. La durabilité des projets du FIDA après leur achèvement est un facteur clé pour garantir un impact plus profond en matière de développement. Selon les autoévaluations du FIDA, les évaluations indépendantes d'IOE et l'examen consacré en 2018 au FIDA par le Réseau d'évaluation de la performance des organisations multilatérales (MOPAN), la durabilité est un domaine dans lequel la performance du Fonds est insuffisante. Afin de mieux en comprendre les raisons, le FIDA étudiera avec IOE la possibilité d'entreprendre des évaluations ex post quelque trois à cinq ans après l'achèvement du projet. Ces évaluations porteraient essentiellement sur la durabilité et chercheraient à déterminer si les données factuelles observées au moment de la clôture du projet sont toujours présentes.

D. Amélioration du système

Suivi-évaluation

53. Des produits d'autoévaluation de qualité élevée sont essentiels pour une évaluation efficace à l'échelle du FIDA. Les systèmes de suivi-évaluation et les informations qu'ils génèrent sous-tendent tous ces produits. Afin d'améliorer la gestion axée sur les résultats et l'établissement de rapports, des efforts plus systématiques sont nécessaires pour aider les gouvernements à renforcer et à utiliser leurs ressources de suivi-évaluation, et à optimiser la manière dont elles se rapportent à celles du FIDA. Donnant suite à un engagement pris au titre de FIDA12, la direction prépare un plan d'action cohérent et détaillé pour renforcer davantage la capacité de suivi-évaluation au niveau des projets, tant en interne que pour les clients. Ce plan s'appuiera sur l'initiative de renforcement des capacités PRIME et le déploiement du Programme de gestion axée sur les résultats pour la transformation du monde rural

(RESOLVE). Un programme d'apprentissage en ligne sur les méthodes de l'évaluation de l'impact est également en cours d'élaboration pour renforcer les capacités d'évaluation de l'impact au niveau des pays.

Système de gestion des résultats opérationnels

54. Comme indiqué ci-dessus, le SGRO est une plateforme de gestion des résultats bien développée et fonctionnelle. Afin de compléter la suite d'applications, un module complet couvrant tout le cycle du COSOP sera élaboré et déployé d'ici à la fin de la période de FIDA11; des modèles supplémentaires sont également envisagés pour couvrir les ONS et les opérations financées par des fonds supplémentaires (par exemple les partenaires de financement climatique). Pour faciliter l'apprentissage et le partage des savoirs, une fonction de marquage est également en cours de mise au point (voir section IV).

IV. Des produits à la pratique: promouvoir l'apprentissage grâce à l'évaluation

A. Mobiliser les savoirs pour devenir une organisation en apprentissage permanent

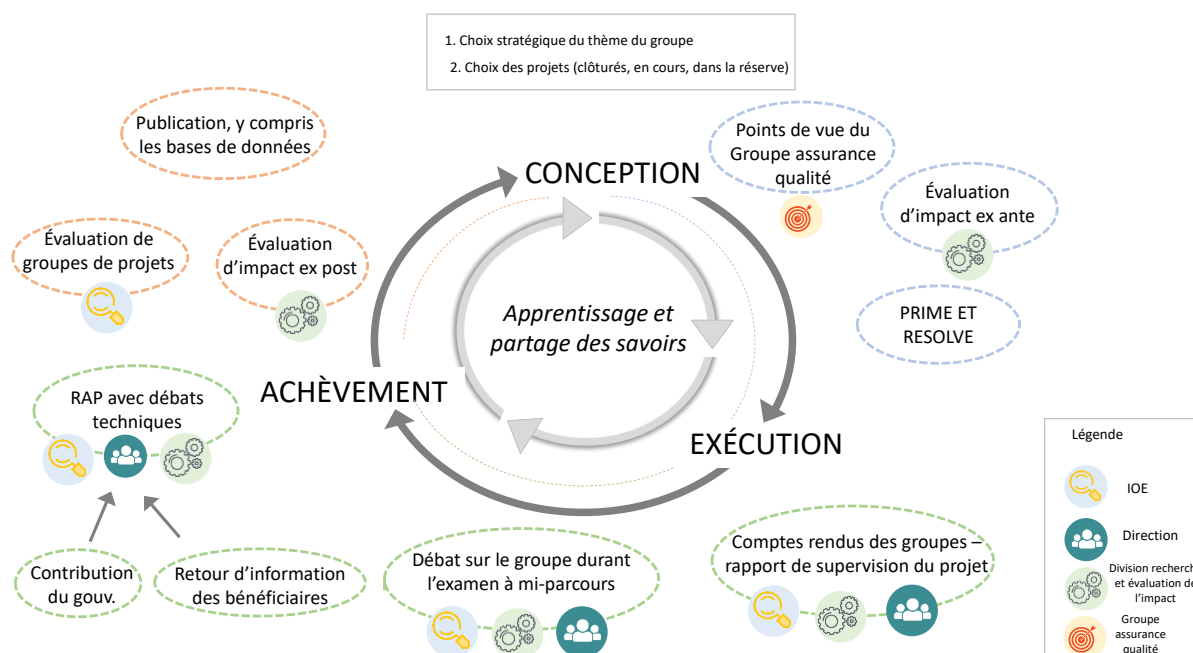
55. Le système d'autoévaluation du FIDA permet à la direction de rendre des comptes au moyen de rapports systématiques sur les résultats et grâce à une gestion efficace de la performance. Le FIDA peut toutefois améliorer encore ses résultats et son impact en tirant davantage parti de sa vaste réserve de savoirs et d'apprentissage, provenant notamment des autoévaluations et des évaluations indépendantes.
56. Bien que les systèmes soient en place, les changements de comportement et de culture nécessaires pour que le FIDA devienne une organisation en apprentissage permanent sont à la traîne. Divers facteurs contribuent à expliquer cet état de choses, et notamment l'inadéquation des mesures d'incitation. À l'instar d'autres IFI, les produits d'autoévaluation obligatoires sont souvent perçus par le personnel opérationnel comme un fardeau. L'accent est mis de manière excessive sur les notations et les processus de validation, ce qui peut constituer un frein à la franchise comme à l'apprentissage. Il conviendrait de rechercher, dans l'attribution des notes, un équilibre délicat entre le réalisme (pour réduire l'écart potentiel avec les notes d'IOE) et l'obtention de résultats positifs (même si l'échec peut également être porteur d'enseignements fructueux). En outre, les équipes sont également soumises à des pressions, l'urgence étant de fournir des volumes de prêts et d'atteindre les objectifs de performance dans la réalisation, ce qui laisse peu de temps pour la découverte et le transfert approfondis des savoirs.
57. Dans la perspective de FIDA12, le modèle opérationnel actuel donne le ton et l'orientation pour l'exploitation de l'apprentissage par le FIDA. Il est important de noter que la direction s'est engagée à actualiser le cadre général que constitue le Cadre relatif à l'efficacité en matière de développement, compte tenu en particulier des importants changements introduits pendant FIDA11. Le Cadre actualisé comprendra des mesures visant à renforcer l'architecture d'autoévaluation, en particulier les initiatives qui améliorent l'apprentissage et le partage des savoirs.
58. Des changements seront également nécessaires sur le plan du comportement et de la culture organisationnelle. Ces changements – accompagnés d'une prise de risque en connaissance de cause, d'incitations à tirer des enseignements des échecs, et d'une culture de gestion adaptative pour procéder à des ajustements de cap – favoriseront davantage l'innovation et un impact plus large. L'autoévaluation jouera un rôle essentiel.
59. **Être le champion de l'apprentissage.** Les signaux en provenance de la direction réorientent déjà la culture institutionnelle dans le sens d'un encouragement à l'innovation et à l'apprentissage, avec en premier lieu la création de l'Unité de

changement, réalisation et innovation. La direction s'est également engagée à accroître l'apprentissage pour une plus grande efficacité en matière de développement, comme énoncé dans la Stratégie en matière de gestion des savoirs et le Plan d'action pour la gestion des savoirs, afin d'encourager une prise de décisions davantage fondées sur des données probantes. La stratégie de gestion des savoirs met le FIDA en mesure d'agir de manière plus proactive pour intégrer efficacement la génération des savoirs à leur documentation et leur utilisation.

B. Collaborer avec IOE pour renforcer l'apprentissage

60. **Créer un cercle d'apprentissage vertueux.** Le projet de politique de l'évaluation note que l'autoévaluation et l'évaluation indépendante se renforcent et se complètent mutuellement et, grâce aux produits, processus et systèmes, aboutissent à de meilleurs résultats aux niveaux opérationnel et organisationnel. Tirer parti de l'alignement stratégique de toutes les composantes du système d'évaluation peut générer un cercle d'apprentissage vertueux pour le FIDA et pour ses clients. Ainsi, les évaluations de groupes de projets proposées par IOE pourraient aider à générer une boucle de retour d'information concernant l'apprentissage et les savoirs durant le cycle du projet (figure 3). Cette boucle commence avec le choix, par IOE et la direction, d'un thème de groupe, avec des projets à déterminer conjointement, y compris des projets en cours et des projets clôturés. Les groupes pourraient même comprendre des projets figurant dans la réserve – en portant une attention particulière au renforcement des systèmes de suivi-évaluation – et, éventuellement, prévoir même une évaluation d'impact ex ante. Les comptes rendus des groupes après les visites de supervision des projets en cours, y compris potentiellement les examens à mi-parcours, pourraient fournir des données probantes en temps réel (mise à jour, par exemple, du cadre de résultats) et des possibilités de gestion adaptative. Les évaluations de l'impact ex post pourraient également être regroupées.
61. La direction et IOE savent tous deux qu'il existe plusieurs domaines dans lesquels la collaboration sur les processus peut renforcer les synergies entre les évaluations indépendantes et les autoévaluations, et faciliter l'apprentissage. On peut citer, parmi ces domaines: i) la conception de produits d'évaluation (par exemple, des EvPP afin d'intégrer des projets plus anciens pour étudier la durabilité; ii) de nouvelles évaluations sous-régionales; iii) une planification améliorée de l'évaluation (par exemple, une meilleure harmonisation des échéanciers entre les COSOP, les examens à l'achèvement des COSOP et les ESPP); iv) les possibilités d'évaluations en temps quasi réel (par exemple, la sélection de projets dans les évaluations de groupes de projets); v) des discussions techniques ex ante sans mention des notations (par exemple, les examens des RAP sur les enseignements tirés); vi) la poursuite de l'harmonisation des approches à mesure que les produits évoluent (par exemple, responsabilité des RAP).

Figure 2

Évaluation de groupes de projets – un cercle d'apprentissage vertueux**C. Exploiter la technologie pour tirer des enseignements de l'évaluation**

62. **Marquage des enseignements tirés.** Les RAP, les examens à l'achèvement des COSOP et les évaluations de l'impact contiennent un trésor de savoirs, intégrés ou tacites, mais il est difficile de les extraire. Bien que des segments consacrés aux "enseignements tirés" soient prévus tout au long des étapes de la conception, de l'exécution et de l'achèvement, la stratégie en matière de gestion des savoirs a noté que "documenter les enseignements solides n'est pas encore systématique, en particulier au niveau des pays, et les enseignements et les informations disponibles par l'intermédiaire des systèmes du FIDA ne sont pas toujours cohérents ou tenus à jour". L'utilisation des technologies de l'information et des communications (TIC) peut contribuer à systématiser et à extraire ces enseignements en vue de concevoir des interventions de développement plus performantes. Des travaux sont actuellement en cours pour améliorer le SGRO afin de permettre le marquage des enseignements tirés et de les relier à des outils de recherche intelligents.
63. **Créer un gisement de résultats des évaluations et les diffuser systématiquement.** Au-delà des savoirs tacites, il est souvent difficile de trouver des rapports d'autoévaluation, spécialement dans le contexte d'une décentralisation accrue. Là aussi, tirer parti de l'application de systèmes TIC de pointe peut contribuer à faciliter l'apprentissage et le partage des savoirs, et à accroître la transparence. Disposer d'un portail complet, à la fois en interne et sur le site web externe du FIDA, permettrait un accès plus aisé à tous les produits d'autoévaluation (c'est-à-dire les RAP, les examens à l'achèvement des COSOP et les évaluations de l'impact). Enfin, pour faciliter encore davantage le cycle d'apprentissage vertueux décrit ci-dessus, des produits indépendants d'IOE pourraient être intégrés aux autoévaluations connexes, avec des possibilités de diffusion plus systématique.

V. La marche à suivre et les prochaines étapes

64. Une solide fonction d'évaluation est essentielle à la crédibilité, la transparence et l'apprentissage d'une institution et, plus important encore, à son efficacité en matière de développement. Afin de consolider cette approche à l'échelle de l'institution, la direction se réjouit à l'idée d'une collaboration avec IOE en vue de la révision de la Politique de l'évaluation et du Manuel de l'évaluation qui s'y rattache,

et de la recherche de synergies entre leurs ensembles de produits respectifs. La politique précisera les rôles du Conseil d'administration, de la direction et d'IOE dans la fonction d'évaluation, et définira les principes fondamentaux de l'évaluation. Reflétant cette approche au niveau de l'ensemble du FIDA, le Manuel de l'évaluation fournira des directives consolidées, tant pour les évaluations indépendantes que les autoévaluations. À cet effet, des consultations devront être entreprises avec toutes les parties prenantes.

65. Plus précisément, la direction prévoit que trois domaines seront adéquatement intégrés dans la Politique de l'évaluation et/ou dans le Manuel de l'évaluation. Il s'agira, premièrement, de la formulation des moyens de renforcer les complémentarités entre les autoévaluations et les évaluations indépendantes afin de souligner la dimension d'apprentissage et d'accroître l'efficacité en matière de développement.
66. Il faudra, deuxièmement, affiner et harmoniser le système et les méthodes de notation du FIDA, à mesure de l'évolution des ensembles de produits d'évaluation d'IOE et de la direction. Ainsi, dans la vision du Fonds qui est celle de FIDA12, c'est-à-dire une institution innovante, en apprentissage permanent et agile, les notations devront être en mesure de rendre compte des avantages d'une gestion adaptative.
67. Troisièmement, les améliorations au système d'autoévaluation décrites dans le présent document devront aussi être reflétées, à commencer par l'actualisation prévue d'un Cadre relatif à l'efficacité en matière de développement tourné vers l'avenir privilégiant les savoirs et l'apprentissage.

International Financial Institutions Comparators

1. IFAD shares many characteristics of self-evaluation systems with other comparable international financial institutions (IFIs). IFAD is also not alone in the present challenges of other IFIs' self-evaluation functions. This appendix reviews the key characteristics, products and challenges of the evaluation systems of five IFIs: the World Bank Group (WBG); the Asian Development Bank (ADB); the African Development Bank (AfDB); the Inter-American Development Bank (IDB); and, the European Bank for Reconstruction and Development (EBRD). Key practices of IFIs in the following aspects are summarized: rating approaches; management reporting using evaluation data; validation by independent evaluation functions; and, reporting on management accountability for implementing independent evaluation recommendations.

A. Self-evaluation systems and key challenges

2. Of the five IFIs considered here, four have had recent assessments of self-evaluation systems, beginning with the World Bank Group's "Behind the Mirror" (2016), followed by African Development Bank and the EBRD (both in 2019) and the Asian Development Bank (2020).¹ These assessments have shown that IFI self-evaluation systems have many common characteristics.
3. **Functions.** Across all IFIs, including IFAD, self-evaluation systems support three functions: results reporting and accountability; performance management; and learning. Information produced by self-evaluation systems is regularly used for performance management, including early detection of poorly performing projects or country strategies needing course correction. Aggregated, this information also provides the basis for high-level corporate results reporting to Boards.
4. **Key challenges.** Across IFIs, a common feature of self-evaluation systems is that their primary focus has been on performance management and accountability using results reporting. The third function, that of learning and knowledge sharing, has received less attention. Hence, a common theme among IFI self-evaluation systems is that the potential use of self-evaluations for learning and knowledge sharing is not being maximized. As summarized by one IFI: "Learning has taken the backseat to accountability...[with] little effort to extract and synthesize evidence and lessons or to inform operations."
5. To help explain the imbalance between accountability and results reporting versus learning, the assessment of self-evaluation systems point to disincentives for writing and using self-evaluations. These disincentives include: excessive focus on ratings and the potential for disconnect with later independent validations; attention to meeting volume delivery targets; and, the low perceived value of the knowledge created and/or difficulty in easily retrieving it. Finally, despite having the systems in place and various initiatives to strengthen M&E systems and capacity, weaknesses in underlying data generated by such systems, undermine the credibility of the self-evaluation system.

¹ *Independent Evaluation Group*. "Behind the Mirror. A Report on the Self-Evaluation Systems of the World Bank Group". Washington: World Bank, 2016. *Asian Development Bank*. "2020 Annual Evaluation Review", Special Topic on Project Self-Evaluation System, ADB Independent Evaluation Department, March 2020. *EBRD*. "Independent external evaluation of EBRD's evaluation system", C. Kirk, EBRD consultant, July 2019. *African Development Bank*. "Evaluation of the African Development Bank's self-evaluation systems and processes", IDEV Corporate Evaluations, December 2019.

B. Self-evaluation products

6. All IFIs share a common mix of self-evaluation products using the OECD DAC and ECG definition of “the formal, written assessment of a project, program, or policy by an entity engaged in that activity”. Self-evaluation products are typically characterized as mandatory (i.e. project completion reports and country strategy completion reports) and voluntary (e.g. impact evaluations/assessments). The self-evaluation systems of the IFIs cover most lending products and country strategies; few evaluate non-lending analytical and advisory activities. The underlying M&E systems and data, and their capture in implementation status and mid-term review reports, support the self-evaluation systems but are not considered formal self-evaluation products. Compliance with requirements of reporting using these systems is generally strong.
7. **Projects.** It is commonplace now for all IFI projects to have project objectives and results frameworks. Only the World Bank and IFAD require a theory of change at the project approval stage. The IDB and IFAD are the only two IFIs to base project quality assurance and results measurement on a development effectiveness framework, which also enables later evaluability. Baselines for results frameworks may be included at design but they are usually completed during the first year of project implementation. During project implementation, all IFIs have a standardized online project supervision systems, regular supervision missions, and a midterm review.
8. Project completion reports are required by all IFIs. For all IFIs, except IFAD, the preparation of the PCR is the responsibility of operational management, although the PCR is usually done by junior staff or consultants. Governments’ contribution to the completion process varies. For the World Bank, the government or borrower normally provides comments. Occasionally, the government prepares a separate completion report that is included in an annex to the PCR. For IDB, governments prepare inputs to the PCRs, while for the AfDB, governments are expected to be consulted for the PCR preparation. IFAD is an outlier in the sense that the PCR is still formally the responsibility of government, prepared with IFAD support. For all IFIs, PCRs are sent to the Board of Directors. PCRs are normally publicly disclosed for sovereign projects. IFAD is again an outlier because the PCR is a government document and hence government (completion and) approval is required.
9. **Country strategies.** All IFIs, with the exception of EBRD, have self-evaluations at completion of country strategy periods. Except for EBRD, there are independent evaluation of country strategies, but these like IFAD’s, cover period longer than a single country strategy period.
10. **Impact evaluations** are done by operational staff, frequently in research divisions. The longest established impact evaluation program is that of the World Bank. Over the last 15 years, the Development Impact Evaluation Initiative (DIME) has expanded and deepened its impact evaluation work. Most DIME impact evaluations are financed by trust funds with project selection done by operational teams and influenced by the interests of trust fund donors. Because of this, impact evaluations are not suitable for corporate accountability purposes. And, unlike IFAD, they are not used for aggregated reporting. The WB assessment recommended embedding more impact evaluations in projects which would help for measuring results, and in doing so, would help to strengthen M&E systems. Most other IFIs (except EBRD) also conduct impact evaluations.
11. **Private Sector Operations.** Self-evaluation completion reports are done for private sector investments and are usually prepared at early operating maturity rather than at closing. As they are considered a self-evaluation, they are done by operational management and not the investee. IDB Invest and the AfDB self-

evaluate all private sector projects, but IFC, ADB and EBRD select a sample. All IFIs use ratings with the criteria usually following the standard OECD or ECG criteria. At the IFC and AfDB, self-evaluation completion reports are not shared with the Board of Directors (due to confidentiality), while IDB Invest provides a heavily redacted PCR to its Board.

C. Rating practices

12. The use of ratings in self and independent evaluation of projects helps to provide structure and consistency. International standards generally regard the use of ratings as good practice. The Evaluation Cooperation Group's² Big Book on Evaluation Good Practice Standards (2012) sets out standard definitions and elements across standard dimensions (i.e., outcome assessed against relevance, efficacy, efficiency and sustainability; and performance of the institution and the recipient).
13. The advantages of ratings are that they: enable aggregation of overall performance, can help to compare projects, including over time, and can identify areas that did not work well and those that performed well. Ratings. The harmonization between self-evaluation and independent evaluation is important to ensure consistency in evaluation scope, criteria, and rating scales.
14. However, disadvantages have also arisen as noted in the several assessments of self-evaluation systems. Here, the key question is whether the focus of attention on ratings has come at the expense of the learning function of evaluation. As noted by one IFI: "Ratings can also become a straitjacket, particularly in self-evaluation, transforming what should be an objective and reflective process into a litigious and defensive one. A predictable effort to present the project in the most favourable light possible will come to the detriment of real accountability and learning."³ In addition to the focus on potential disagreement of ratings between self and independent evaluations, the so-called "disconnect", there is also the potential that ratings are seen as threatening or reflective of staff performance.
15. Many IFIs use a six point rating scale (i.e. HS, S, MS, MU, U, HU), while the AfDB and the ADB use a four point scale. There are pros and cons of each: a six point scale helps to provide more nuance and mitigates against a stark satisfactory or unsatisfactory rating. A six point scale may be too broad but can still have a tendency to rate performance as moderately satisfactory which can be aggregated as being "above the bar". To assign overall ratings based on sub-criteria ratings, most IFIs use an averaging methodology or (because this generates decimals) the ratings range system, while the WBG and IFAD use a judgment-based approach.
16. Finally, the adoption by most IFIs of the standard rating dimensions put forward by the ECG may facilitate comparisons between performance of these institutions. However, there are questions of whether the standardized definitions accurately reflect dimensions of importance to the institution and are as meaningful to staff (and therefore have more potential to increase learning) as compared to more tailored rating regimes. For example, the EBRD has replaced what were EBRD-specific ratings with standard ECG one. While IFAD has used many of the standard rating dimensions, it has retained a rural poverty outcome rating as well as ratings of key elements of its strategic framework (i.e. gender, innovation, scaling up, environmental and resource management, and climate change).
17. **World Bank.** Implementation Completion Reports (ICRs) are required for all investment projects. The WB's Independent Evaluation Group (IEG) conducts ICR

² The ECG consists of: AfDB, ADB, EBRD, IDB, WBG, IFAD, EIB, IMF, and IsDB. There are three observers: OECD-DAC, UN, and GEF.

³ "Project Self-evaluation in EBRD", EvD Discussion Paper for Board of Directors, January 2020.

Reviews (ICRRs) on all ICRs in order to independently validate the findings using desk-review. IEG uses the same rating scales as the ICR. There are three main ratings done by both Management and IEG: outcome (with three dimensions of relevance, efficacy and efficiency); Bank performance (with IEG rating constituent elements of quality at entry and quality of supervision); and, quality of M&E. In addition, IEG provides an overall quality of ICR rating. There is a six point scale: highly satisfactory, satisfactory, moderately satisfactory, moderately unsatisfactory, unsatisfactory, and highly unsatisfactory.

18. IEG carries out Project Performance Assessment Report (PPARs) on 20-25 per cent of the projects that close, which involves a field visit. There are four main ratings in the PPAR with an additional rating on monitoring and evaluation (M&E). Outcome: defined as the extent to which the operation's major relevant objectives were achieved, or are expected to be achieved. It has three dimensions: relevance, efficacy, and efficiency. Risk to development outcome: defined as the risk, at the time of evaluation, that development outcomes (or expected outcomes) will not be maintained (or realized). Bank performance: defined as the extent to which services provided by the World Bank ensured quality at entry of the operation and supported effective implementation through appropriate supervision. The rating has two dimensions: quality at entry and quality of supervision. Borrower performance: defined as the extent to which the borrower (including the government and implementing agency or agencies) ensured quality of preparation and implementation, and complied with covenants and agreements, toward the achievement of development outcomes. The rating has two dimensions: government performance and implementing agency(ies) performance. Monitoring & Evaluation (M&E) quality is rated on three dimensions: design; implementation; and utilization. The WB mostly uses a six point scale (highly satisfactory, satisfactory, moderately satisfactory, moderately unsatisfactory, unsatisfactory, and highly unsatisfactory). Possible ratings for risk to development outcome: high, significant, moderate, negligible to low, and not evaluable.
19. **IDB** also uses standard definitions for project evaluation which are harmonized between self-evaluation project completion reports and independent evaluation's validation. Performance ratings measures project performance against the project development objectives that were formulated at approval. The ratings are given on the four core criteria (relevance, efficacy, efficiency, sustainability) and an overall project outcome rating is derived from them. For non-sovereign operations, additional dimensions – additionality, work quality, and investment outcome – are also evaluated and scored. IDB is the only major IFI that does not assess and rate Bank or borrower performance as part of its project completion assessments. Project outcomes are rated on a six-point scale ranging from Highly Successful to Highly Unsuccessful.
20. **AfDB** uses a four-point scale to rate performance in self-evaluation PCRs and independent validations. Four key criteria of relevance, efficiency, effectiveness and sustainability are rated using possible options of: Highly satisfactory (4), Satisfactory (3), Unsatisfactory (2) and Highly unsatisfactory (1). The overall score of the project is the average of the scores of the four dimensions. Because this process generates decimal points, ranges are then used to attribute the final rating as an integer number within the current 4-point scale. They also rate AfDB's and the borrower's performance as well as gender and environmental-social sustainability but these are not included in the computation of the overall project performance score.
21. **ADB** requires that all public sector projects have a self-evaluation PCR. Evaluation of projects by IED validates approximately 80% of PCRs and reported in project completion report validations (PVRs) using desk reviews. In addition, field-based

project performance evaluation reports (PPERs) are done for 10%–20% of projects. PCRs assesses core criteria of relevance, effectiveness, efficiency, and sustainability that contribute to the overall project performance assessment and the project's success rating. Non-core criteria of development impacts (i.e., ADB, borrower and cofinancier performances) are also rated. ADB uses a four point scale. The assessments and ratings standards used in self-evaluations and independent evaluations by Independent Evaluation Department (IED) are consistent.

22. **EBRD.** In 2016, EBRD replaced EBRD-specific ratings with internationally comparable OECD-DAC ratings.
23. **IFAD.** PCRs include ratings on project performance on standard dimensions of relevance, effectiveness, efficiency and sustainability. IFAD and government's performance are also assessed and rated. In addition, IFAD provides a rating for overall rural poverty impact as well as: gender equality and women's empowerment; innovation, scaling up, environment and natural resource management, adaptation to climate change; targeting and outreach; and access to markets. An overall project achievement rating is also provided. All ratings use a six point scale (highly satisfactory, satisfactory, moderately satisfactory, moderately unsatisfactory, unsatisfactory, and highly unsatisfactory). Final aggregate scores reflect the best, informed judgement of the project completion review team and are not averages of the sub-criteria. IOE conducts desk-based validation reviews of all PCRs (PCRVs) and selects a rates the same criteria using the same six point scale. In addition, IOE rates the quality of the PCR on dimensions of candour, lessons, quality and scope. IOE conducts more in-depth project performance evaluations (PPEs) on about 10 completed projects each year using field visits.

D. Use of self-evaluation data (reporting)

24. One of the key functions of self-evaluation systems is accountability and the reporting of results to executive boards and, more broadly, stakeholders. All IFIs, including IFAD, use this information and report using corporate results scorecards as well as management dashboards which are populated with information collected from supervision reports to manage performance of ongoing operations and completion reports on results. Only the IDB and IFAD, however, produce a management development effectiveness report for regular annual reporting on results. This includes both corporate performance targets and implementation progress from the active portfolio as well as recently closed projects.

E. Validation practices of independent evaluation offices

25. The range and diversity of types of independent evaluation report varies across the different IFIs. However, almost all IFIs' independent evaluation products mirror the two major self-evaluation products for project and country strategy completion reports. Validations of PCRs by independent evaluation offices are now common practice and are done to ensure consistency and credibility of self-evaluations. PCRs are usually sent to the independent evaluation offices within 6 months of project closing. The World Bank, ADB and IDB validate 100 per cent of available PCRs for sovereign operations. The ECG good practice guidance indicates that if the sample is less than 100 per cent of PCRs, then a statistically representative sample should be selected. The AfDB validates a representative sample of PCRs up to a maximum of 65 a year. EBRD conducts 10-15 validations reports each year.
26. For private sector operations, the ADB and AfDB validate all completion reports for non-sovereign operations which are prepared at the time of maturity. The independent evaluation group of the World Bank validates a random sample of 40% of IFC's mature investments

27. These independent offices also conduct further in-depth evaluations of projects. In recent years, there has been a shift from many single operation evaluations to those focussing on higher level products such as clusters of project evaluations (e.g. by AfDB and EBRD) to further their enhance relevance and value.
28. All IFIs except EBRD conduct country strategy evaluations usually covering a period of 10 years.
29. Voluntary or demand driven impact assessments or evaluations are not validated by independent evaluators.

F. Management accountability reports

30. In 2016, the Evaluation Cooperation Group (ECG) constituted a working group to review "Evaluation Recommendations, Management Responses and Feedback Loops". All IFIs have a formal requirement for management to respond to all recommendations put forward by the evaluation function. Their report found that across IFIs there is a systematic process of following up and reporting on the implementation of accepted independent evaluation recommendations. The process covers: (i) a compilation of recommendations to be implemented by management; (ii) monitoring and assessment of their implementation; and, finally (iii) reporting on the implementation. Some IFIs (e.g. ADB, IDB and WBG) require an additional step, namely, the formulation of action plans for the implementation of recommendations.
31. Institutions differ in how roles and responsibilities are defined for different phases of the process. For monitoring, it is usually Management which self-assesses the status of implementation. However, in some cases (ADB, IDB, WBG), the evaluation functions validate this self-assessment; for IFAD, the evaluation function prepares a response to Management's report. The preparation of action plans usually is Management's responsibility. For reporting on the implementation status of recommendations, either Management (IFAD) or the evaluation function (ADB, IDB) may be prepare the reports. The content can either be stand-alone (e.g. IFAD's PRISMA) or can be integrated into broader reports. Technology helps compile recommendations and monitor the implementation of follow up actions in many IFIs (ADB, IDB, WBG, EBRD). IFAD relies on standardised excel files.
32. Discussions among the ECG members are taking place on how to best strengthen feedback loops for capturing lessons learned and to extract extra value added of higher-level strategic issues. Options to do this include thematic or summative evaluation that pick up recurring recommendations on a given theme.

G. Individual IFI Self-Evaluation Highlights

33. **World Bank Group.** In 2016, the IEG conducted a major review of the self-evaluations systems of the entire Group (i.e. IBRD/IDA, MIGA and IFC). Coming out of this, a landmark document "Evaluation Principles" (2019) setting out a wider set of 'evaluation principles' has recently been issued by the World Bank. This was issued jointly by Management at the World Bank Group and the Bank's Independent Evaluation Group (IEG). It is recognized by others as 'state of the art' good practice. *Utility, credibility and independence* as key principles and uses the OECD/DAC evaluation principles as core evaluation criteria. Self-evaluation forms the nexus of oversight, monitoring and research.
34. Self-evaluation in the World Bank (IDA/IBRD) includes mandatory self-evaluation of Implementation Completion Reports (ICRs) for projects at closing and Country Partnership Framework Completion and Learning Reviews (CLRs). Approximately 300 ICRs are done each year at an average cost of \$40-50,000 while for CLRs, the

average is about \$45,000. Voluntary self-evaluations include impact evaluations overseen by the Development Impact Measurement Initiative (DIME). Of the impact evaluations conducted by the WB, 62% are not related to a parent project, while 32% of investments projects completed during 2011-2015 had impact evaluations embedded. All evaluation products of the World Bank (i.e. not including IFC or MIGA) are presumed to be publicly disclosed.

35. IEG validates and rates all ICRs on outcomes against relevance, effectiveness and efficiency criteria as well as Bank and borrower performance, project M&E and the quality of the self-evaluation. The ICRs and their validation feed into WB corporate scorecards and IDA results measurement systems. It does not validate impact evaluations. IEG purposively selects a sub-set of projects (about 15%) and conducts Project Performance Assessment Reports. The Management Action Record tracks follow-up to all IEG recommendations and is updated annually. IEG's evaluative notes are circulated to the Board.
36. **African Development Bank.** The AfDB's self-evaluation system was found by its independent evaluation department in 2019 to be largely relevant and aligned with comparators' best practices. While the existing guidance tools and procedures are clear, the main weakness is the low level of compliance, lack of candour and overly optimistic performance ratings. This has resulted in a lack of credibility of the system and hence of its value added. Overall, the evaluation found that the three main functions of a self-evaluation system (i.e. performance management, accountability and learning) need strengthening.
37. Country strategy (CSP) and regional programs are all self-evaluated at mid-term and completion. A CSP-CR is conducted within 6 months of the closing date of the CSP and its findings incorporated into the next country strategy. The CSP-CR does not use ratings but a satisfactory or unsatisfactory designation; these are sent to the Board for information. CSPs are not validated by the independent evaluation group (IDEV) but they do conduct independent country program evaluations (CSPEs) that cover one or more CSP periods.
38. **Asian Development Bank.** In March 2020, ADB's Annual Evaluation Review contained a special theme chapter on the robustness and credibility of its project self-evaluation system. Overall, the architecture of ADB's project self-evaluation system was found to be robust, credible, and useful for accountability and learning. Like other IFIs, there were common weaknesses especially data reliability/management and capturing and using tacit knowledge embedded in self-evaluations throughout the project cycle
39. **Inter-American Development Bank.** The IDB's results architecture, like IFAD's, is anchored in a development effectiveness framework. Similar to IFAD's RIDE, the IDB produces an annual development effectiveness report. The Development Effectiveness Overview (DEO) presents impact of closing projects and lessons learned as well IDB Group corporate performance indicators. Among the IFIs, the IDB is a leading edge in developing a unified framework to evaluate sovereign and non-sovereign operations.
40. **EBRD** conducted an independent external evaluation of their Evaluation System in 2019. While the major evaluations produced by the EBRD's Evaluation Department are largely relevant, credible and useful, the review found that the self-evaluation system needs to be substantially reformed and rebuilt. It recommended the development a medium-term plan for strengthening the self-evaluation system and integrating it with the EBRD's results architecture. It also found that valuation quality is constrained by the quality of self-evaluation, the evaluability of operations and the limitations of EBRD's results monitoring systems. Although EBRD's Evaluation Policy gives responsibility to management for ensuring an effective self-evaluation system, in practice it is the independent evaluation department (EvD) that runs the system and reports on results.

Self-evaluation supporting systems and processes

A. Project Design and essential underpinnings

1. Although the project design phase and its reports are neither a monitoring nor an evaluation product, they are a core building block of the entire evaluation system with contributions from across the house. The DEF and the accompanying development effectiveness matrix (DEM) form the basis of a project's theory of change. The ensuing results framework, which includes output and outcome indicators, helps to improve the later evaluability of the project. Ideally, all indicators in the results frameworks contain baselines against which to monitor progress and evaluate results at completion. The IFAD11 RMF has committed that by the end of IFAD11, 70 per cent of ongoing projects will have baselines by the end of the first year of implementation. Key among the indicators are the core indicators (CI) which are mapped to the Strategic Framework and the SDGs. To help improve the quality and consistency of CIs, Management issued guidelines in 2020 which include a mandatory methodology to collect timely and reliable data at the project baseline, midterm and completion stages. Overall results frameworks are reviewed systematically during design phase review meetings, including by the Quality Assurance Group (QAG). Equally important, the project implementation manuals prepared during the design phase detail the monitoring and evaluation (M&E) systems that are needed to track progress against indicators.
2. Weak borrower capacity regarding M&E is a significant barrier to having a robust self-evaluation system. IFAD is recognized as a leader in the development and delivery of capacity building for clients in the area of M&E. The *Program for Rural M&E* (PRiME) is a global M&E and impact assessment training and certification programme which helps to improve in-country capacities for measurement and tracking at the project management unit level. In PRiME's first phase, 156 PMU staff from 147 IFAD projects in 79 countries were trained, with another 100 participants scheduled for phase two. In recent years, other IFIs have sent participants to these courses and IFAD operational staff jointly attend with project collaborators. *AVANTI* conducts assessments to identify gaps in national institutional M&E systems for the rural sector and develops action-plans to fill these gaps. The diagnostics allow IFAD government counterparts to implement targeted improvements to their management-for-results in the rural sector. A recent program, *Results based management for rural transformation* (RESOLVE), seeks to improve IFAD project performance by strengthening the implementation capacities of project management units managing IFAD-financed projects through tailored training and follow-up support in results-based and adaptive management.

B. Project Supervision Reports and Mid-Term Reviews

3. Project supervision reports (PSRs) are a critical element of IFAD's evaluation system. Project delivery teams are required to conduct a supervision mission for all projects to produce PSRs at least once a year. Problem projects are required to have at least two supervision missions a year. PSRs and the data from the underlying systems are an important means for reporting and managing the performance of ongoing projects. They help with the timely identification of potential problem projects and facilitate adaptive management, including course corrections, during implementation. In addition to capturing data on results frameworks, PSRs also include assessments by teams of key elements of project implementation, including the performance of project management and M&E systems.
4. Mid-term reviews (MTRs) are also mandatory and are a key decision point in the project life cycle providing opportunities for adaptive management. In fact, timely MTRs with project restructuring and other actions can help bring problem projects

back into good performance. A revised policy on restructuring was approved in 2018 to facilitate early corrective action; restructuring is not permitted in the last year of project implementation.

5. Management has been focusing on strengthening the quality, consistency and timeliness of supervision reports. Efforts are underway to ensure that PSR ratings are consistent with the narrative and evidenced-based. Teams are also giving more attention to updating log frames in the ORMS accurately and in a timely fashion following supervision missions. Regular portfolio review meetings track the timeliness and frequency of PSRs.
6. The performance of M&E systems are also rated during project supervision missions. Evidence in other IFIs has shown that good project M&E can boost performance, while poor project monitoring systems hinder project management. Management is actively working to strengthen the capacities of both project monitoring units as well as IFAD staff through such initiatives as PRiME and RESOLVE.

C. COSOPs

7. COSOPs are living strategies with an average duration of six years. COSOP Results Reviews (CRRs) are done at the midterm and are a self-assessment of progress towards results, lessons learned, risk factors encountered (e.g. exogenous shocks) and changes in country demand and priorities. CRRs may entail adjustments to the strategy if there are significant changes in the country context.

D. ORMS

8. The well-developed Operational Results Management System (ORMS) is critical part of IFAD's highly integrated technology platform, including financial management, procurement and a client portal. ORMS facilitates the tracking of results, project performance and aggregate reporting by capturing real-time data from decentralized settings, including the core indicators, that is then used for evidence-based decision-making by country teams. It is also used for performance management at regional and corporate levels, including with the use of dashboards tracking such things as disbursement targets, risk indices and project cycle timelines. Data and results across the project portfolio are openly shared on IFAD's corporate and external websites. Enabled by technology, this immediate level of transparency in IFAD's project portfolio is helping to promote accountability, drive quality, and increase impact. The platform continues to evolve, building in features such as multilingual capability, a risk framework and a stronger focus on the management of problem projects. Promoting enhanced learning from IFAD's portfolio with the use of innovative tagging features and use of artificial intelligence is also under development.