

Obtenir des données de qualité et bien organiser sa recherche

Atelier de formulation des projets MRV 2019-2021

14-18 octobre, Ouidah, Bénin



Belgium
partner in development

CEBioS



museum



Operational Directorate Natural Environment
OD Nature | OD Natuur | DO Nature

Prof. Dr. *Corneille* EWANGO EKOKINYA

Université de Kisangani

UNIKIS

Faculté de Gestion des Ressources Naturelles Renouvelables & Faculté des Sciences

Head of Delegation DR Congo GBIF

Forest Ecology, Biodiversity & Ecosystems Management Research Group

"Mettre le savoir au service de l'homme, la société et des forêts"

Inspiring people to care about the planet ... to impact and change way of life of millions

INTRODUCTION

- Rappel des objectifs de l'atelier

MÉTHODES DE COLLECTE DES DONNÉES

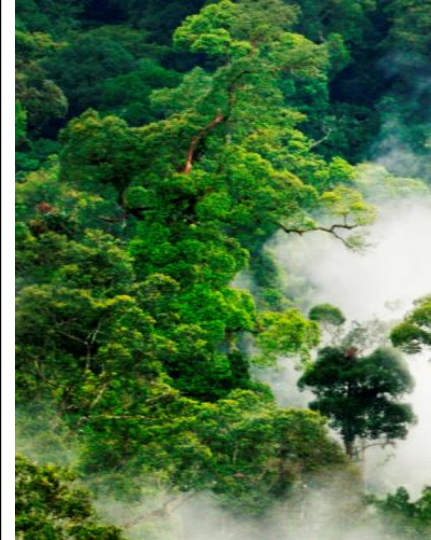
- Principes méthodologiques de la collecte de données
- Orientations générales pour collecter des données

COLLECTER LES DONNÉES

- Rassembler les données existantes
- Produire de nouvelles données
- Adapter les données pour les besoins de l'inventaire
- Importance de développer des Méthodologies Standardisées de collecte des données
- De la qualité et la gestion des données
- Nécessité de la gestion des données
- Data sharing

ORGANISER SA RECHERCHE

- 6 Étapes vers la réussite
- MRV & Interface Science-Politiques



Monitoring, Rapportage et Vérification (MRV)

Obtenir des données de qualité et bien organiser sa recherche

Comment une politique peut-elle être évaluée sans mesures? Les mesures de la biodiversité sont difficiles mais nécessaires. Cette difficulté tient essentiellement à la nature même de la biodiversité (diversité des espèces et des environnements, diversité des méthodes d'évaluation, diversité des systèmes de surveillance, etc.).

Les objectifs de l'atelier sont multiples

Connaître & Comprendre les MRV pour la biodiversité

- Présenter le concept MRV à de nouveaux partenaires scientifiques (institutions sœurs) ;
- Discuter des lignes directrices pour le format de l'appel à projets MRV à venir, sur base des expériences MRV passées ;
- présenter et discuter les propositions d'indicateurs et de méthodologies pour les projets à venir ;
- travailler en pratique sur certains aspects clés de ces projets : la qualité et la gestion des données, le(s) indicateur(s) proposé(s) et l'interface science-politique.

La collecte de données est le premier pas et une partie intégrante du développement et de l'actualisation d'un inventaire de biodiversité et sur le suivi des ressources naturelles renouvelables pour surmonter les défis auxquels le MRV devra répondre.

MÉTHODES DE COLLECTE DES DONNÉES

2.1 INTRODUCTION

- La collecte de **données** est une partie intégrante du développement et de l'actualisation d'un inventaire de biodiversité et de MRV sur les ressources naturelles.
- Des activités de collecte de données formalisées doivent être établies, adaptées aux circonstances locale, régionale et nationale des habitats et écosystèmes cibles, et périodiquement révisées dans le cadre de la mise en œuvre de *bonnes pratiques*.
- Dans la plupart des cas, la création de nouvelles sources de données sera **limitée par les ressources disponibles** et il sera nécessaire de donner des priorités à celles-ci, en prenant en compte *les résultats de l'analyse de catégories de source clés*.
- Des procédures de collecte de données sont nécessaires pour trouver et traiter les données existantes (c'est-à-dire des données collectées et stockées à des fins statistiques autres que l'inventaire), ainsi que pour produire de nouvelles données par le biais de sondages ou de campagnes de mesure.
- D'autres activités incluent le maintien des flux de données, l'amélioration des estimations, la création d'estimations pour de nouvelles catégories et/ou le remplacement de sources de données existantes lorsque celles actuellement utilisées ne sont plus disponibles.

L'expression « données » peut se définir comme des informations factuelles (par ex. des mesures ou des statistiques) utilisées comme base de raisonnement, de discussion ou de calcul.

La collecte de données est l'action d'obtenir et de compiler des informations provenant de différentes sources.

Principes méthodologiques de la collecte de données

□ Les principes méthodologiques de la collecte de données sur lesquels se fondent *les bonnes pratiques* sont les suivants :

- Se concentrer sur la collecte de données nécessaires à l'amélioration des estimations de catégories de source clés les plus importantes, avec le plus grand potentiel de changement ou avec la plus grande incertitude.
- Choisir des procédures de collecte de données qui améliorent itérativement la qualité de l'inventaire conformément aux objectifs de qualité des données.
- Mettre en place des activités de collecte de données (priorétisation des ressources, planification, mise en œuvre, documentation, etc.) qui mènent à une amélioration continue des ensembles de données utilisés dans l'inventaire.
- Collecter des données/informations à un niveau de détail approprié à la méthode utilisée.
- Examiner régulièrement les activités de collecte de données et les besoins méthodologiques, pour améliorer l'inventaire de manière progressive et efficace.
- Conclure des accords avec des fournisseurs de données pour assurer des flux d'information cohérents et continus.

Tout en collectant les données, les bonnes pratiques recommandent de reconnaître les besoins futurs liés à la collecte de données.

Orientations générales pour collecter des données

- Les méthodes de collecte des données cherchent à comprendre et à produire des données mesurées et abordent où trouver et comment utiliser les variables et leurs indicateurs ou facteurs par défaut.

- Elles couvrent :
 - le développement d'une stratégie de collecte de données visant à respecter les objectifs de qualité des données en termes d'opportunité, mais également de cohérence, d'exhaustivité, de comparabilité, d'exactitude et de transparence en utilisant les d'acquisition, d'utilisation et de vérification des données.

 - Des activités d'acquisition de données, notamment la création de nouvelles sources de données, traitant de données limitées et de confidentialité, et ayant recours à l'opinion d'experts.

 - La transformation de données brutes en une forme utile pour l'inventaire.

2.2 COLLECTER LES DONNÉES

□ Cette section propose des orientations générales pour la collecte de données existantes, la création de données nouvelles et l'adaptation des données à l'inventaire. Elles s'appliquent à la collecte de données sur les :

1) **Maintenir l'apport de données pour l'inventaire**

➤ *Les bonnes pratiques recommandent d'engager les fournisseurs de données dans le processus de compilation et d'amélioration de l'inventaire* en les impliquant dans des activités telles que :

- ✓ Proposer une estimation initiale pour la catégorie, montrer des incertitudes potentiellement importantes et inviter les fournisseurs de données potentiels à collaborer pour améliorer les estimations,
- ✓ Des ateliers scientifiques ou statistiques sur les entrées et les sorties de l'inventaire,
- ✓ Des contrats ou des accords spécifiques pour un apport régulier de données,
- ✓ Des actualisations informelles régulières/annuelles sur les méthodes qui utilisent leurs données,
- ✓ Établissement de termes de référence ou de mémorandums de compréhension pour le gouvernement et/ou les organisations scientifiques qui fournissent des données pour clarifier quelles données sont nécessaires à l'inventaire, comment elles sont sélectionnées et fournies à l'organisme/ chercheur(s) chargé de l'inventaire et quand.

Ces activités permettront d'assurer que les données les plus appropriées sont disponibles pour l'inventaire et que les données sont correctement comprises par l'organisme chargé de l'inventaire.

Elles permettront également d'établir des liens avec les organisations qui fournissent des données.

Le cas échéant, il peut être utile d'examiner si des accords légaux existants ou nouveaux peuvent garantir l'apport de données dans l'inventaire.

2.2 COLLECTER LES DONNÉES

2) Données restreintes et confidentialité

- ❑ Les fournisseurs de données peuvent limiter l'accès à l'information parce que celle-ci est confidentielle, non éditée ou non encore finalisée.
 - Il s'agit normalement d'un mécanisme visant à empêcher l'utilisation non appropriée de données et une exploitation commerciale non autorisée de celles-ci, ainsi qu'à sensibiliser à de possibles imperfections des données.
 - Cependant, il arrive parfois que l'organisation ne dispose simplement pas des ressources nécessaires pour collecter et contrôler les données. Il est conseillé, lorsque ceci est possible, de coopérer avec les fournisseurs de données pour trouver des solutions afin d'apaiser leurs préoccupations en :
 - ✓ Expliquant comment on envisage d'utiliser les données,
 - ✓ Se mettant d'accord par écrit sur la quantité de données qui sera rendue publique,
 - ✓ Identifiant comment l'utilisation des données dans les inventaires permettra d'augmenter l'exactitude,
 - ✓ Proposant une coopération pour établir un ensemble de données mutuellement acceptable,
 - ✓ Et/ou en accréditant et en reconnaissant les données fournies dans l'inventaire.

La protection de la confidentialité est un des principes fondamentaux d'une agence nationale de statistiques (ANS) voir des institutions de recherche scientifiques.

- Ceux-ci s'engagent à protéger les informations qui exposent manifestement les opérations, les biens, les attitudes ou toute autre caractéristique des répondants individuels.
- Si les répondants ne sont pas convaincus que les informations qu'ils fournissent aux ANS/Institutions scientifiques sont absolument confidentielles, la qualité des informations collectées peut en souffrir.
- Les données individuelles détaillées doivent dès lors être traitées et rassemblées de manière à en tirer les informations qui sont importantes pour l'utilisateur, sans dévoiler les données individuelles.

2.2 COLLECTER LES DONNÉES

3) Opinions d'experts

- ❑ L'opinion d'experts sur le choix méthodologique et le choix des données d'entrée à utiliser est finalement la base de tout développement d'inventaire et les spécialistes des différents secteurs peuvent jouer un rôle particulièrement utile pour pallier aux informations non disponibles, pour sélectionner des données dans toute une série de valeurs possibles ou émettre des opinions sur les degrés d'incertitude.
 - On peut trouver des experts qualifiés dans les gouvernements, les institutions de recherche scientifiques, les instituts techniques, l'industrie et les universités, etc.

- ❑ L'opinion d'experts peut permettre de choisir la méthodologie appropriée, la valeur de paramètre parmi la plage de valeurs fournie, les données sur les activités les plus appropriées, la manière la plus appropriée d'appliquer une méthodologie, ou d'établir la combinaison appropriée de technologies en usage.

- ❑ Il est nécessaire d'obtenir l'opinion d'experts, même si l'on applique des méthodes statistiques classiques aux ensembles de données, étant donné que l'on doit juger si les données sont un échantillon aléatoire représentatif et, dans l'affirmative, choisir les méthodes d'analyse des données.
 - Ceci nécessite une opinion en matière de méthodes et de statistiques. Une interprétation est particulièrement nécessaire pour les ensembles de données de petite taille, à biais élevé ou incomplets.

 - Dans tous les cas, l'objectif est d'être aussi représentatif que possible pour réduire les biais potentiels et augmenter l'exactitude.

Le terme *sollicitation de l'opinion d'experts* désigne les méthodes officielles pour obtenir des données en consultant des experts

MRV-Biodiversité : Valorisation des données existantes

- Quid du MRV? - **M**onitoring, **R**apportage et **V**érification (**MRV**)
= Un Mécanisme de Suivi et Evaluation + Rapportage
- Comment et pourquoi valoriser les données de biodiversité existantes dans:
 - Les archives dans les universités,
 - Les instituts de recherche,
 - Le site web CHM, GBIF).
 - Importance de ne pas toujours recommencer à collecter sur le terrain si des données existantes ne sont pas exploitées

Renforcement du mécanisme de suivi écologique et socio-économique et valorisation de données de référence

Les différents types de jeu de données

- ❑ Les quatre types de jeu de données connues sont simples au début, mais deviennent progressivement plus riches, mieux structurées et plus complexes.
- ❑ Les détenteurs de données sont encouragés à publier les données les plus riches possibles pour assurer leur utilisation dans le plus large éventail d'approches et questions scientifiques. Toutefois, tous les jeu de données ne possèdent pas les mêmes niveaux de précision.

1) Métadonnées seules - offre aux chercheurs un outil précieux pour découvrir et connaître la disponibilité des informations pas encore disponibles en ligne. Cela peut également aider à évaluer l'importance relative et la valeur des collections non numérisées et fixer des priorités pour de futures numérisations. De manière littérale, les métadonnées sont "les données sur les ... des informations additionnelles sur des aspects tels que la qualité des données

2) Listes taxonomiques - Les jeux de données peuvent également fournir un catalogue ou une liste de noms d'organismes, ou taxons. Même s'ils peuvent inclure des détails supplémentaires comme les nom communs et les citations des spécimens, ces listes taxonomiques ou référentiels classent généralement l'information par taxonomie, géographie et statut ou en combinant les trois.

3) Les données d'occurrence d'espèces - ont suffisamment de détails cohérents pour fournir des informations sur la localisation des organismes individuels dans le temps et l'espace - c'est-à-dire qu'ils offrent des preuves de la présence d'une espèce (ou d'un autre taxon) dans un endroit et une date donné.

- Les occurrences figurants dans ces jeux de données peuvent quelquefois ne fournir que des informations générales sur les localités, et parfois juste pour identifier le pays. Mais dans de nombreux cas, ces ensembles de données donnent de plus amples précisions sur les lieux et des coordonnées géographiques pour appuyer une analyse poussée et une cartographie de la distribution des espèces.

4) Données d'échantillonnage - Parfois, les jeux de données fournissent de plus amples détails, mettant en évidence non seulement l'enregistrement d'une espèce à un endroit et une date donnée, mais également la possibilité d'évaluer la composition des communautés de groupes taxonomiques plus larges ou même l'abondance des espèces en plusieurs moments et lieux. Ces ensembles de données proviennent généralement des protocoles standards développés pour mesurer et suivre la biodiversité comme les transects, les comptages d'oiseaux ou les prélèvements d'eau de mer ou d'eau douce. En indiquant, lors d'un échantillonnage, les méthodes, événements et l'abondance relative des espèces enregistrées, ces ensembles de données améliorent les comparaisons pouvant être faites avec des données collectées en utilisant les mêmes protocoles à différents endroits et moments - dans certains cas, cela permet aux chercheurs d'en déduire l'absence d'espèces particulières sur des sites spécifiques.

5) Données spatiales - Coordonnées : un code documentant une position sur la surface de la terre, exprimé suivant un SRS (spatial reference system). En pratique; souvent latitude/longitude. Géoréférencement : le procédé qui consiste à assigner une référence géographique à un enregistrement donné. Datum (système géodésique)

2.2.1 Rassembler les données existantes

❑ **Bien que la liste ci-dessous ne soit pas exhaustive, elle propose des pistes pour des sources possibles de données spécifiques à un pays :**

- Agences nationales de statistiques/centres de surveillance de la biodiversité et de recherches scientifiques
- Experts par secteur, organisations parties prenantes
- Autres experts nationaux
- Base de données sur la biodiversité
- Autres experts internationaux
- Organisations internationales publiant des statistiques comme, par ex., les CHM, GBIF, Herbaria & Musées).
- Bibliothèques de référence (bibliothèques nationales)
- Articles scientifiques et techniques dans des livres, journaux et rapports sur l'environnement/la biodiversité.
- Universités
- Recherche sur internet d'organisations et de spécialistes
- Rapports nationaux des inventaires des Parties à la Convention-cadre des Nations Unies sur biodiversité (CDB) et les changements climatiques

‡ **Quelquefois, selon la taille et la structure de l'échantillon original, des données brutes peuvent être rassemblées de manière à protéger la confidentialité tout en produisant des informations utiles pour les besoins de l'inventaire de biodiversité.**

‡ **Si, cependant, il est nécessaire de protéger la confidentialité, les ANS, ou l'organisme qui a collecté les données au départ, sont normalement les seuls qui peuvent effectuer ce traitement supplémentaire des données brutes.**

Comment y-parvenir? (1)

1) Filtrage des données disponibles

- **Les activités de collecte des données doivent commencer par un filtrage initial des sources de données disponibles.**
 - Il s'agira d'un processus itératif dans lequel s'accumulent les détails des données disponibles.
 - Ce filtrage peut être lent et nécessiter un questionnement jusqu'à ce qu'une opinion finale puisse être faite quant à l'utilité d'un ensemble de données pour un inventaire.
- **L'objectif pour lequel les données ont été collectées à l'origine peut être un indicateur important de la fiabilité de ces données.**
- **Les organismes de régulation et les organes statistiques officiels/ institutions de recherche sont responsables de la représentativité des échantillons et de l'exactitude des mesures et adoptent ainsi souvent des normes convenues.**
 - Parce qu'elles sont soumises à une procédure d'examen plus élaborée, les statistiques officielles ne sont souvent disponibles qu'après une longue période de temps mais les données préliminaires peuvent être disponibles plus tôt.
 - Ces données préliminaires peuvent être utilisées si leur validité est documentée et peut être vérifiée par rapport aux objectifs de qualité des données établis par le système de gestion de la qualité des données.

2) Affiner les besoins en données

- Une fois que l'organisme chargé de l'inventaire a sélectionné un ensemble de données, à moins que les données publiées ne puissent être utilisées sous leur forme originale, l'étape suivante consistera à développer une demande de spécification et de données plus officielle.
- Une définition claire des besoins en données garantira que les données fournies correspondent à ce qui est attendu. Les spécifications doivent comprendre des détails tels que :
 - ✓ Définition de l'ensemble de données (par ex. séries temporelles, détails sur les secteurs et sous-secteurs, portée (locale, régionale & nationale), besoins en données sur les incertitudes, facteurs, variables, indicateurs et/ou unités de données sur les activités),
 - ✓ Définition du format (par ex., feuilles d'analyse) et structure (par ex., quels différents tableaux sont nécessaires et quelle est leur structure) de l'ensemble de données,
 - ✓ Description de toute hypothèse concernant la portée (locale, régionale & nationale), les secteurs inclus, l'année représentative, le niveau de technologie/de gestion, et les facteurs, variables, indicateurs et/ou unités ou paramètres d'incertitude,
 - ✓ Identification des routines et des échelles de temps pour des activités de collecte de données (par ex. à quelle fréquence l'ensemble de données est-il actualisé et quels éléments sont actualisés),
 - ✓ Référence à la documentation et aux procédures d'acquisition et d'utilisation,
 - ✓ Nom de contact et organisation,
 - ✓ Date de disponibilité.

‡ Il peut être utile d'obtenir un engagement vis-à-vis de ces spécifications de la part de l'organisation qui fournit les données.

‡ Maintenir et actualiser ces spécifications régulièrement, en cas de changement de besoins en données, aide également à documenter les sources de données et offre des directives actualisées pour

Comment y-parvenir? (2)

3) Choisir entre données nationales et internationales publiées

- Dans la plupart des cas il est préférable d'utiliser des données nationales car celles-ci sont généralement mieux actualisées et donnent de meilleurs liens vers ceux qui ont fourni les données.
- La plupart des ensembles de données internationaux se basent sur des données nationales et, dans certains cas, les données émanant d'organes internationaux réputés peuvent être plus accessibles et mieux s'appliquer à l'inventaire.
- Dans certains cas, des groupes comme des associations professionnelles internationales ou des organes statistiques internationaux disposeront d'ensembles de données nationales spécifiques pour les industries ou d'autres secteurs économiques que n'auront pas les organisations nationales.
- Souvent les données internationales ont été soumises à un examen et une vérification supplémentaires et peuvent avoir été ajustées pour améliorer la cohérence, bien que ceci n'améliore pas nécessairement les estimations si les données ajustées sont recombinaisonnées avec des informations nationales.
- Les pays sont encouragés à développer et à améliorer les sources de données nationales pour éviter de dépendre des données internationales.
- Recouper les données nationales avec toutes données internationales disponibles permet d'évaluer l'exhaustivité et d'identifier des problèmes possibles avec un ensemble de données.

4) Données de substitution

- Il est préférable d'utiliser des données qui sont directement liées à l'objet quantifié plutôt que d'utiliser des données de substitution (c'est-à-dire des données alternatives qui ont une corrélation avec les données qu'elles remplacent).
- Cependant, dans certains cas, des données directement applicables peuvent ne pas être disponibles ou présenter des lacunes (par ex., si les enquêtes ou les programmes d'échantillonnage ne sont pas fréquents).
- Dans ces cas, des données de substitution peuvent permettre de combler les lacunes et produire une série temporelle cohérente ou une moyenne nationale.
- Les données de substitution doivent être liées physiquement et statistiquement aux émissions de l'ensemble des installations pour lesquelles les informations ne sont pas disponibles.
- Ces données alternatives doivent être choisies selon les informations et les circonstances spécifiques à un pays et une relation entre les données et les émissions (c'est-à-dire un facteur d'émission) doit être développée en utilisant des informations d'un sous-ensemble représentatif d'installations dont les émissions sont connues.
- L'utilisation de données de substitution pour obtenir une estimation initiale d'une émission ou d'une absorption peut aider à prioriser les ressources.
- Lors de la sélection et de l'utilisation des données de substitution pour évaluer les émissions ou les absorptions, les *bonnes pratiques recommandent aux pays de procéder aux étapes suivantes* :
 - i. Confirmer et documenter la relation physique entre les émissions/absorptions et les données de substitution sur les activités.
 - ii. Confirmer et documenter une corrélation significative sur un plan statistique entre les émissions/absorptions et les données de substitution sur les activités.
 - iii. En utilisant une analyse de régression, développer un facteur spécifique à un pays liant les émissions/absorptions aux données de substitution.

La plupart des ensembles de données internationaux se basent sur des données nationales et, dans certains cas, les données émanant d'organes internationaux réputés peuvent être plus accessibles et mieux s'appliquer à l'inventaire

2.2.2 Produire de nouvelles données

- ❑ Il peut être nécessaire de produire de nouvelles données si des facteurs d'émission, des données sur les activités ou d'autres paramètres d'estimation représentatifs n'existent pas ou ne peuvent pas être estimés à partir des sources existantes.
- ❑ Ces activités nécessitent souvent des ressources importantes et sont envisagées à juste titre lorsque la catégorie de source est *clé et qu'aucune autre option n'est possible*.
- ❑ *Pour optimiser* l'utilisation des ressources, il est recommandé, autant que possible, de produire les données requises dans le cadre d'une extension d'un programme existant plutôt que d'initier des programmes totalement nouveaux.

1) Produire des données par des mesures

- Des mesures doivent être utilisées, conformément aux normes standards de collecte des données.
- Des mesures peuvent également être utilisées pour directement quantifier la biodiversité, les ressources naturelles renouvelables ou pour calibrer et vérifier les modèles utilisés pour produire des données.
- Lorsque l'utilisation des données de mesure est envisagée, les **bonnes pratiques** recommandent de vérifier si elles couvrent un échantillon représentatif, c'est-à-dire typique d'une proportion raisonnable de la catégorie entière, et également si une méthode de mesure appropriée a été utilisée.
- Les meilleures méthodes de mesure sont celles qui ont été développées par des organisations de normalisation officielles et testées sur le terrain pour déterminer leurs caractéristiques opérationnelles.
 - L'utilisation de méthodes de mesure normalisées améliore la cohérence des données mesurées et fournit à l'organisme chargé de l'inventaire des informations supplémentaires sur la méthode telles que les niveaux d'incertitude statistique, les limites de détection plus faibles, la sensibilité et les limites supérieures de mesure, etc.

2) Relation entre les données et les modèles

- Bien que des modèles soient fréquemment utilisés pour évaluer des systèmes complexes et puissent être utilisés pour produire des données, les modèles sont un moyen de transformer les données et n'éliminent pas le besoin des données de les diriger.

Des résultats fiables et comparables peuvent être obtenus en utilisant un programme de mesure bien conçu avec des objectifs définis, des méthodes appropriées, des instructions claires au personnel chargé des mesures, des procédures de traitement des données et d'établissement des rapports bien définies et une documentation appropriée.

Produire des données par des mesures - - Organiser sa recherche

ÉLÉMENTS GÉNÉRIQUES D'UN PROGRAMME DE MESURE

Objectif de mesure	Énoncé clair du ou des paramètre(s) à déterminer, par ex. les évaluations de la biodiversité ou des ressources naturelles renouvelables dans un écosystème
Protocole de méthodologie	Description de la méthodologie de mesure à utiliser. Celle-ci doit inclure: <ul style="list-style-type: none">• Les composés à mesurer et toute autre condition de référence associée;• Les méthodes pour garantir que des échantillons représentatifs sont choisis qui reflètent la nature de la catégorie de source clé et l'objectif de la mesure;• L'identification de toute technique normalisée à utiliser;• L'équipement analytique requis et ses conditions opérationnelles;• Toutes conditions d'exactitude, de précision ou d'incertitude;• Besoins en capture de données à satisfaire;• Régimes d'acquisition & d'utilisation à suivre.
Plan de mesure avec des instructions claires au personnel chargé des mesures	Plan de mesure pour le personnel chargé des mesures incluant: <ul style="list-style-type: none">• Le nombre de points d'échantillonnage pour chaque paramètre à mesurer et comment ceux-ci doivent être sélectionnés;• Nombre de mesures individuelles à réaliser pour chaque point d'échantillonnage et ensemble de conditions;• Dates et périodes de mesure de la campagne de mesure;• Arrangements relatifs à l'établissement des rapports;• Informations supplémentaires liées à la source ou au procédé à collecter pour permettre le traitement des données ou l'interprétation des résultats;• Conditions (ou ensemble de conditions) de la source (ou pour une usine industrielle la capacité, la charge, le combustible ou l'intermédiaire) à respecter durant la mesure;• Personnel responsable des mesures, autres personnes impliquées et ressources à utiliser.
Procédures de traitement des données et d'établissement des rapports, et documentation	Procédures traitement des données, y compris; <ul style="list-style-type: none">• Procédures d'établissement des rapports qui rendront compte des mesures, de la description des objectifs de mesure et du plan de mesure;• Documentation nécessaire pour permettre, par des calculs, de tracer les données de base collectées et les conditions d'opération des procédés à partir des résultats.

Lors de mesures relatives à un écosystème, il convient de porter une attention particulière à la définition des conditions d'échantillonnage

2.2.3 Adapter les données pour les besoins de l'inventaire

- Lorsque l'on utilise des données existantes, que l'on réalise de nouvelles mesures ou que l'on combine les deux, il est important de garantir que le niveau de détail et la portée des données correspondent, y compris en ce qui concerne les secteurs/procédés/réduction, les lieux, le type de sol, les composés et les années comprises.

1) Lacunes dans les ensembles de données

- Les inventaires de biodiversité ou ressources naturelles renouvelables nécessitent des estimations cohérentes entre séries temporelles et entre catégories.
- Cependant, il existe des méthodes permettant de combler les lacunes si des données manquent pour un an ou plus ou si les données ne représentent pas l'année ou la portée (locale, régionale ou nationale) requise. Pour y remédier de lacunes dans les données ou d'incohérence, il convient de:
 - **Comblent les lacunes des données périodiques** : Des lacunes apparaîtront dans les séries temporelles si les données ne sont pas disponibles annuellement (par ex., les longues et coûteuses enquêtes sur les ressources naturelles, comme les inventaires sur les forêts nationales, sont compilés tous les cinq ou dix ans).
 - ✓ On peut devoir inférer des données de séries temporelles pour compiler une estimation annuelle complète pour les années entre les enquêtes, ainsi que pour les prévisions et les rétrodictions (par ex. lorsque des estimations sont nécessaires pour 1990-2004 et les données de sondage ne sont disponibles que pour 1995 et 2000).
 - **Révision des séries temporelles** : Afin de respecter les délais, les organisations statistiques peuvent utiliser la modélisation et les hypothèses pour compléter les années les plus récentes de leurs estimations.
 - ✓ Ces estimations sont ensuite affinées l'année suivante lorsque toutes les données ont été traitées.
 - ✓ Les données peuvent avoir été soumises à un examen ultérieur des données historiques pour corriger les erreurs ou actualiser de nouvelles méthodologies.
 - ✓ Il est important que l'organisme chargé de l'inventaire recherche ces changements dans les séries temporelles des données source et les intègre dans l'inventaire.

Adapter les données (2)

- **Incorporer des données améliorées** : Alors que la capacité à collecter des données s'améliore généralement avec le temps si bien que les chercheurs mettent en œuvre des méthodes de niveau supérieur, les données ne sont pas nécessairement acceptables pour les années précédentes avec des méthodes de niveau supérieur.
 - ✓ Par exemple, lorsque des programmes d'échantillonnage et de mesure directs sont introduits, il peut se produire des incohérences dans les séries temporelles car le nouveau programme ne peut pas mesurer les conditions passées.
 - ✓ Quelquefois cette question peut être réglée si les nouvelles données sont assez détaillées (par ex., si les facteurs, les variables et indicateurs faisant l'objet de mesures récentes peuvent être distingués de celles plus anciennes ne faisant pas l'objet de mesures) et les données sur les activités historiques peuvent être stratifiées en utilisant l'opinion d'experts ou des données de substitution.

- **Compenser pour des données détériorées** : Les techniques de raccordement, telles que sur la cohérence des séries temporelles, peuvent être utilisées pour gérer des ensembles de données qui se sont détériorées dans le temps.
 - ✓ La détérioration peut être due à des changements de priorité dans les institutions, une restructuration économique ou une diminution des ressources (par ex., certains pays avec des économies en transition ne collectent plus certains ensembles de données qui étaient disponibles au cours de l'année de référence, ou ces ensembles de données peuvent contenir différentes définitions, classifications et niveaux d'agrégations).

- **Couverture incomplète** : Lorsque les données ne sont pas entièrement représentatives de toute les portées, par ex., mesures prises pour 3 écosystèmes sur 10 ou données de sondage sur les activités agricoles pour 80 pour cent du pays, elles peuvent toujours être utilisées mais elles doivent être combinées avec d'autres données pour calculer une estimation nationale.
 - ✓ Dans ces cas, l'opinion d'experts ou la combinaison de ces données avec d'autres ensembles de données (données de substitution ou données exactes) peuvent être utilisés pour calculer un total national.
 - ✓ Dans certains cas, les données de sondage ou de recensement sont collectées dans un programme national évolutif qui échantillonne différentes provinces ou sous-secteurs annuellement avec un cycle répété qui établit un ensemble de données complet après quelques années.
 - En gardant à l'esprit cette cohérence des séries temporelles, il est recommandé que les hypothèses faites une année s'applique également les autres années, et que l'on demande aux fournisseurs de données de calculer des données annuelles représentatives avec une couverture complète.

Adapter les données (3)

2) Combiner des ensembles de données de manière numérique

- ❑ Un chercheur/organisme chargé de l'inventaire recevra parfois plusieurs ensembles de données potentiels pour la même estimation, par exemple, une série de mesures indépendantes sur la biodiversité.
 - Si les données font référence à la même quantité et ont été collectées de manière relativement uniforme, alors leur combinaison augmentera l'exactitude et la précision.
 - On peut également combiner des données en mettant en commun des données brutes et en réestimant la moyenne et les limites de confiance de 95 pour cent, ou en combinant les résumés de statistiques en utilisant la relation établie dans les calculs de statistique.
 - Il est également possible de combiner des mesures d'une seule quantité réalisée avec différentes méthodes qui produisent des résultats avec différentes distributions de probabilité sous-jacentes.
 - Cependant, les méthodes pour cela sont plus complexes, et dans la plupart des cas, il suffira probablement d'utiliser l'opinion d'experts pour décider si les résultats doivent être moyennés ou si des estimations plus fiables doivent être privilégiées par rapport à d'autres.
- ❑ Lorsque l'on utilise **des données qui ne sont pas homogènes** (par ex. à cause de la présence de méthodologie de collecte pour certains chercheurs mais non dans d'autres), **les estimations de l'inventaire doivent être stratifiées** (sous-divisées) **pour que toutes les strates soient homogènes** et que le total pour la catégorie source représente la somme des strates.
 - Les estimations d'incertitude peuvent alors être obtenues en traitant chaque strate comme une catégorie individuelle.

Adapter les données (4)

3) Données d'inventaires régionales (CHM, GBIF, Universités et centres de recherche)

- Dans certaines circonstances, les statistiques sur les activités et les ensembles de données sur la biodiversité régionale sont plus détaillées, mieux actualisées, plus exactes et/ou complètes qu'au niveau national.
- Dans ces cas, un inventaire compilé et ensuite regroupé au niveau régional peut donner un inventaire de meilleure qualité pour un pays qu'un inventaire compilé en utilisant des statistiques et des ensembles de données nationaux moyennés.
- Dans de tels cas, et afin de se conformer aux *bonnes pratiques*, les inventaires peuvent être *compilés* entièrement ou en partie au niveau régional pour autant que :
 - Chaque composé régional est compilé conformément aux *bonnes pratiques* y relatives, au choix des niveaux, à la cohérence des séries temporelles et à l'exhaustivité.
 - L'approche utilisée pour regrouper les inventaires régionaux et combler toute lacune au niveau national est transparente et conforme aux méthodes de *bonnes pratiques* présentées dans *les Lignes directrices*.
 - L'inventaire final du pays est conforme aux exigences de qualité des *bonnes pratiques* en matière **d'exhaustivité, de cohérence, de comparabilité, d'opportunité, d'exactitude et de transparence**.
 - ✓ En particulier les estimations de secteur calculées dans différentes régions, et ensuite regroupées dans l'inventaire final, doivent être autocohérentes.
 - ✓ Aucune observation ou occurrence ne doit être omise ou comptée deux fois dans l'inventaire final et les différentes parties de l'inventaire doivent utiliser les hypothèses et les données de manière cohérente pour autant que cela soit pratique et approprié.

Importance de développer des Méthodologies Standardisées de collecte des données

– Comment les développer ?

1) Les dimensions de la qualité des données

➤ **Certains des aspects de la qualité des données fréquemment observés sont :**

- **la validité** : les données mesurent ce qu'elles sont censées mesurer.
- **la fiabilité** : les données sont mesurées et collectées de manière cohérente en fonction de définitions et de méthodes normalisées ; la répétition des mesures produit les mêmes résultats.
- **l'exhaustivité** : tous les éléments de données sont inclus (conformément aux définitions et méthodes spécifiées).
- **la précision** : les données sont suffisamment détaillées.
- **l'intégrité** : les données sont protégées de tout biais ou manipulation délibérée pour des raisons politiques ou personnelles.
- **l'actualité** : les données sont à jour (actuelles) et les informations sont disponibles à temps.

2) Normes de qualité générales

Si la qualité peut être un concept ambigu, sa définition dans le cadre des données est depuis longtemps clairement établie. Dans sa [définition de la qualité en statistique](#), Eurostat présente un ensemble de sept dimensions qui définissaient initialement les données statistiques mais qui peuvent également s'appliquer à d'autres types de données :

Pertinence	Les données répondent-elles aux besoins actuels et potentiels des utilisateurs
Exactitude et fiabilité	Les données sont-elles exemptes d'erreurs. Dans le cas de statistiques, l'exactitude désigne la proximité entre la valeur estimée et la valeur réelle (inconnue) au sein de la population.
Actualité et ponctualité	Quel est le délai entre la publication des données et le phénomène qu'elles mesurent et la fréquence des mises à jour correspond-elle au calendrier prévu
Accessibilité et clarté	Les utilisateurs peuvent-ils accéder facilement aux données et les métadonnées permettent-elles de bien les comprendre
Continuité	Les données peuvent-elles aider à comparées les tendances dans le temps ou l'espace (écosystèmes)
Comparabilité	Les données peuvent-elles être comparées dans le temps ou géographiquement
Cohérence	Les données sont-elles conformes aux définitions et méthodologies reconnues

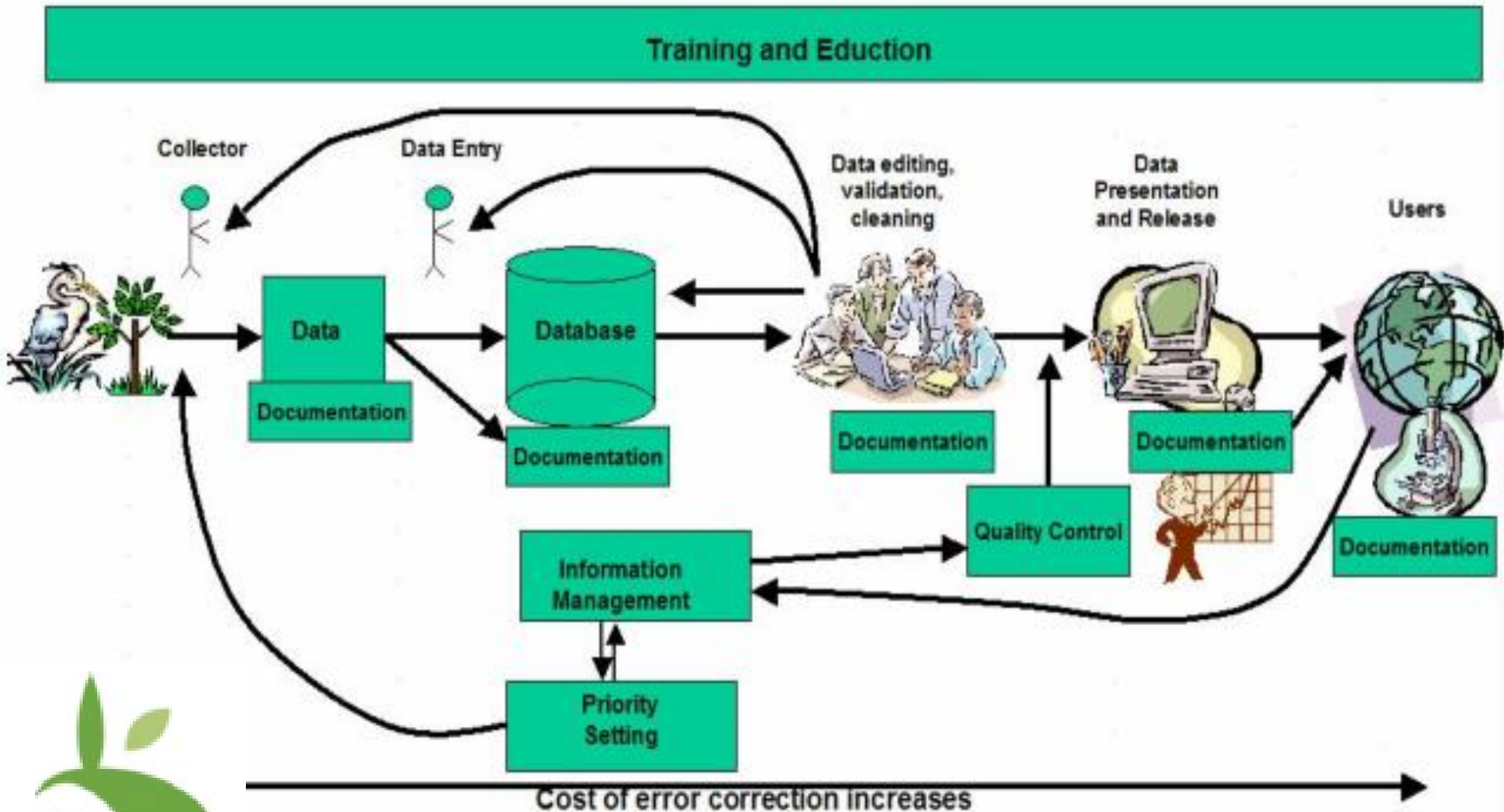
De la qualité et la gestion des données

□ Assurer le “Contrôle de qualité” - Metadonnées

- Expliquez la nécessité de suivre les collectes/interviews et de vérifier la qualité des données/interviews en temps réel.
- Discutez de la manière de traiter les erreurs des collecteurs/enquêteurs.
- Expliquez ce qu’il faut faire des questionnaires remplis, et comment régler les problèmes imprévus.
- Mettez en exergue le fait que le superviseur doit prendre des notes sur le terrain et expliquez ce qui doit être noté.
- Discutez du calendrier d’exécution de l’enquête et de la nécessité de garder le contact avec le coordinateur de l’enquête.

La qualité des données est un concept relatif qui dépend de l’usage qui est fait de ces données...

Chaîne des données et qualité



Principes du Nettoyage des Données

- **La structure générale du nettoyage de données (Maletic & Marcus 2000) est :**
 - Définir et déterminer les erreurs types ;
 - Chercher et identifier les occurrences d'erreurs ;
 - Corriger les erreurs ;
 - Documenter les cas d'erreurs et d'erreurs types ;
 - Modifier le procédé d'entrée de données afin de réduire les erreurs futures.



Référence suggérée:

Chapman, A. D. (2005). Principes et méthodes de nettoyage de données, version 1.0. Trad. Roosen, A. Copenhague: Global Biodiversity Information Facility, 90 pp. Disponible en ligne sur http://links.gbif.org/gbif_nettoyage_donnees_manual_fr_v1.pdf

Nécessité de la gestion des données

- La prise de décisions concernant l'élaboration des politiques, la planification et l'aménagement repose en grande partie sur une information traitée et non sur des données brutes.
- Les données doivent être interprétées avant de pouvoir être utilisées.
- Les données primaires brutes représentent souvent un volume très important de sorte qu'elles ne peuvent être utilisées efficacement que dans le cadre d'un système de gestion.
- Les fonctions d'un système de gestion d'une base de données sont les suivantes:
 - assurer la conformité des données avec les classifications types;
 - assurer la validité des données;
 - assurer l'intégrité des données et leur cohérence interne;
 - mettre en sécurité et préserver les données primaires;
 - permettre un accès facile aux données primaires;
 - traiter efficacement les données suivant les besoins;
 - permettre d'intégrer différentes séries de données, de manière à accroître leur utilité globale.
- Un principe fondamental consiste à conserver toutes les données telles qu'elles ont été collectées, sous leur forme primaire.
 - Cela réserve une certaine souplesse dans la manière dont les données peuvent être traitées (c'est-à-dire filtrées, agrégées, transformées) et garantit que tous les calculs sont reproduits à partir de données-sources incorporant toutes les révisions.

Étant donné l'investissement considérable que représente la collecte des données et les faibles coûts du stockage et du traitement, il n'y a guère de raison de ne pas conserver des données complètes sous leur forme primaire.

Data sharing

Le Partenariat améliore l'efficacité

- ❑ Notre mission est d'accroître l'accès à l'information et son utilisation pour la conservation de la biodiversité et le développement durable des projets MRV sur la biodiversité pour:
 - 1) collecter et améliorer des données,
 - 2) regrouper, synthétiser et publier des données,
 - 3) rendre les données plus largement accessibles aux utilisateurs finaux potentiels, et
 - 4) tirer parti des données de la biodiversité pour éclairer la conservation de la biodiversité.

- ❑ La stratégie des MRV est de relier les données à l'utilisation des connaissances dans les domaines de la conservation et du développement durable, où la demande d'informations peut soutenir les stratégies de gestion et prise de mesures de suivi de la biodiversité et de leurs écosystèmes.

Organiser sa recherche – 6 étapes vers la réussite

‡ Devoir posséder suffisamment d'information pour prendre une décision informée sur la meilleure méthode à suivre:

- La méthode de collecte et d'analyse des données dépend de nombreux facteurs, notamment le contexte, la question qui est examinée, l'objet de la collecte des données, la nature et la taille de l'organisme.
- Le plus important est de s'assurer que la collecte de données est effectuée d'une manière conforme aux normes et qu'elle respecte les règles sur l'accès à l'information et la protection de la qualité des données.
- Dans l'intérêt de l'efficacité et de l'efficience, il est recommandé que les données recueillies servent à éclaircir un problème ou cerner des possibilités.

‡ Afin de protéger la crédibilité et la fiabilité des données, l'information doit être collectée au moyen de techniques de collecte de données acceptées (standardisées).

- **Étape 1 : Déterminer les problèmes et/ou possibilités que présente la collecte de données** - - Il faut ensuite décider comment procéder. À cette fin, il pourrait être utile de se livrer à une évaluation interne et externe dans le but de comprendre ce qui se passe à l'intérieur et à l'extérieur de l'organisme à collecter les données.
- **Étape 2 : Sélectionner les problèmes et/ou possibilités et établir des objectifs**
 - L'étape 2 est axée sur le choix d'un problème prioritaire ou de possibilités de collecte de données et l'établissement d'objectifs.
 - L'organisme examine les problèmes et les possibilités cernés dans le cadre de l'évaluation interne et externe effectuée à l'étape 1 et choisit les problèmes et possibilités prioritaires à partir desquels fonder un projet de collecte de données. Exemples de questions qu'un organisme peut se poser pour classer les problèmes et possibilités par ordre de priorité aux fins de la collecte de données :
 - ✓ Y a-t-il une raison fondamentale ou une occasion de recueillir des données desquelles d'autres problèmes et/ou possibilités semblent découler?

Établissement d'objectifs - - Les objectifs définis pour chaque problème et/ou possibilité peuvent dépendre d'une hypothèse ou d'une prédiction qui peut être testée à l'aide de techniques de collecte de données et d'analyse.

6 étapes vers la réussite (2)

○ **Étape 3 : Planifier une approche et des méthodes**

- Les organismes/chercheurs prendront des décisions sur, entre autres, les personnes qui feront l'objet du sondage, la méthode de collecte des données, les sources des données qui seront utilisées et la durée du projet de collecte de données.
 - Ces décisions peuvent être prises en consultation avec un expert.
 - Les méthodes et approches découleront des objectifs énoncés à l'étape 2 et varieront considérablement selon certains facteurs, dont le contexte de l'organisme, la taille de l'organisme, ses ressources et la complexité des problèmes et/ou possibilités choisies.

○ **Étape 4 : Recueillir les données**

- Au moment de planifier la meilleure façon de recueillir des données, à l'étape 4, il est important de tenir compte des aspects pratiques et d'appliquer les meilleures méthodes de résolution des problèmes logistiques qui surgissent souvent à cette étape.

○ **Étape 5 : Analyser et interpréter les données**

- Cette étape consiste en l'analyse et l'interprétation des données recueillies.
- Que les données soient quantitatives ou qualitatives, l'analyse peut se révéler plus ou moins complexe selon les méthodes utilisées et la quantité de données compilées.
- Chaque organisme/chercheur devra déterminer s'il possède la capacité interne et les compétences nécessaires pour analyser et interpréter les données lui-même, ou s'il a besoin de recourir à l'aide d'une expertise externe.

○ **Étape 6 : Donner suite aux résultats (Rapportage)**

- Une fois qu'un organisme/chercheur a analysé et interprété les résultats des données recueillies, il peut décider d'y donner suite, de collecter d'autres données du même genre ou de modifier son approche.
- Des informations quantitatives et qualitatives peuvent former une base solide à l'élaboration d'un plan d'action efficace en vue d'atteindre des objectifs organisationnels stratégiques en matière de ressources humaines, pour la comparabilité, la continuité, d'équité et de diversité, élaborés à la suite du processus de collecte de données.

Développer des méthodologies de collecte de données standardisées permet la comparabilité et la continuité dans le temps et géographiquement pour détecter les

PRINCIPALES CONCLUSIONS

Développer les MRV contribuent à:

- ❑ **La Surveillance de la biodiversité, recherche, gestion de l'information et communication de l'information** - - Des informations de surveillance à long terme, standardisées, spatialement complètes et facilement accessibles, complétées par des recherches sur les écosystèmes, fournissent les conclusions les plus utiles pour des évaluations de l'état et des tendances pertinents pour les politiques. Le manque de ce type d'information dans de nombreux domaines a entravé le développement de cette évaluation.
- ❑ **Changements rapides et seuils** - - Une compréhension croissante des changements rapides et inattendus, des interactions et des seuils, en particulier en ce qui concerne le changement climatique, souligne la nécessité d'une politique qui réponde et s'adapte rapidement aux signaux de changement environnemental afin d'éviter des pertes de biodiversité majeures et irréversibles.



PRINCIPALES CONCLUSIONS (2)



Interface science-politique

□ Détecter le changement et agir

➤ Trois points de décision complémentaire pour la conservation de la biodiversité :

- 1. Lorsque les seuils sont franchis (il est peut-être trop tard)** - L'action est différée jusqu'à ce que l'évidence du changement soit claire. Par exemple, les espèces déclinent en dessous du nombre minimum viable; trop peu d'habitat est laissé pour supporter des groupes entiers d'espèces; extinctions. Les options d'action sont limitées et coûteuses. Les interventions sont drastiques et ont peu de chances de réussir. La récupération, si elle se produit, est lente. Les impacts socio-économiques sont inévitables.
- 2. Lorsque des changements d'écosystème sont détectés (il n'est pas trop tard)** - - Une action plus précoce, basée sur des preuves du changement en cours, offre davantage d'options pour atténuer les impacts. Les mesures sont prises lorsque les relations de cause à effet sont au moins partiellement comprises et lorsque les résultats de la recherche permettent de penser avec certitude que la biodiversité va probablement se dégrader si aucune mesure n'est prise. Conduit à une probabilité élevée d'inverser ou de stabiliser les impacts et de réduire les facteurs de stress avant qu'il ne soit trop tard.
- 3. Lorsque les premiers signaux indiquent qu'il peut y avoir un changement (prévention possible)** - - Les premiers signaux de changement sont une source d'information utilisée lors de l'élaboration d'options politiques, en vue d'une gestion proactive; et veiller à ce qu'une surveillance et des recherches appropriées soient en place. Les premiers signaux ne peuvent être détectés que dans quelques endroits ou seulement chez certains animaux ou plantes d'une population. Ces changements peuvent s'inscrire dans le cadre de fluctuations naturelles - ou peuvent indiquer des changements plus importants à venir. Une action rapide dès maintenant peut prévenir des problèmes à l'avenir, être moins onéreuse et ses conséquences moins graves.



Merci de votre



Belgium
partner in development

CEBioS



Laboratoire de Biomathématiques &
d'Estimations Forestières

attention

museum

Operational Directorate Natural Environment
OD Nature | OD Natuur | DO Nature

"Mettre le savoir au service de l'homme, la société et des forêts"

Inspiring people to care about the planet ... to impact and change way of life of millions