



UNITED
NATIONS

EP

UNEP/MED WG.467/12



UNITED NATIONS
ENVIRONMENT PROGRAMME
MEDITERRANEAN ACTION PLAN

8 août 2019
Français
Original : anglais

7^{ème} réunion du Groupe de Coordination de l'Approche Écosystémique

Athènes, Grèce, 9 septembre 2019

Point 6 du l'ordre du jour : Système d'information pilote IMAP, Normes des données et Dictionnaires des données, problèmes relatifs à l'assurance qualité et premiers éléments de la politique de gestion des données IMAP

Système d'information pilote IMAP Assurance qualité et Contrôles qualité

Pour des raisons de coût et de protection de l'environnement, le tirage du présent document a été restreint. Il est aimablement demandé aux délégations d'apporter leur copie de ce document aux réunions et de s'abstenir de demander des copies supplémentaires.

PNUE / PAM
Athènes, 2019

Note du Secrétariat

Dans le cadre du Programme de travail et budget pour 2018-2019 de l'ONU Environnement/PAM (Décision IG.23/14), l'INFO/RAC dirige les travaux sur la mise en œuvre du point 1.5.1 du Programme de travail relatif au développement de « *la plate-forme «Info/IMAP» et de la plate-forme pour la mise en œuvre de l'IMAP pleinement opérationnelles et développées, connectées aux systèmes d'information des composantes du PAM et à d'autres plates-formes de connaissances régionales pertinentes, afin de faciliter l'accès aux connaissances des gestionnaires et des décideurs, des parties prenantes et du public en général.* »

CAR/INFO, en pleine conformité avec les besoins du système du PAM, et pour la mise en œuvre du programme intégré de surveillance et d'évaluation de la mer Méditerranée et des critères d'évaluation connexes (IMAP) va fournir une plate-forme Web pour la gestion des données et de l'information de tous les ensembles de données liés à la surveillance et à l'évaluation des écosystèmes marins relevant de son mandat.

La plate-forme devrait permettre:

- se conformer aux exigences en matière d'informations à la lumière des obligations énoncées dans la convention de Barcelone;
- évaluer, définir et surveiller les stratégies de prise de décision en atteignant l'objectif d'adéquation et d'utilisation des données.

Dans le cadre de ce projet, la création du système d'information pilote IMAP, y compris l'infrastructure de communication intégrée, revêt une importance particulière. Le système collectera et intégrera des données provenant de différentes sources et fournisseurs de données et fournira des informations à différents groupes d'utilisateurs cibles.

Ce document a pour objectif de décrire les contrôles qualité (QC) mis en œuvre pour le système d'information pilote IMAP et le processus d'assurance qualité (QA) qui garantissent l'application effective afin d'améliorer efficacement la qualité des données collectées et publiées par le système d'information pilote IMAP.

Table des matières

I.	Assurance Qualité et Contrôles Qualité officiels	1
II.	Développements futurs pour des Contrôles Qualité de Niveau Supérieur	4
III.	Conclusions.....	5

I. Assurance Qualité et Contrôles Qualité officiels

La qualité des données n'est pas un produit, mais est un processus. L'assurance qualité est le processus dont l'objectif est d'améliorer la qualité, et elle est basée sur la définition de ce qui est bon et de ce qui ne l'est pas.

La première étape du processus d'assurance qualité a été la définition des normes de données (DSs), des dictionnaires de données (DDs), et des contrôles officiels associés (QCs) pour les modules de surveillance associés aux 11 Indicateurs Communs IMAP sélectionnés. Une description détaillée des DSs et DDs est fournie dans le document *Standards de données (DSs) et Dictionnaires de données (DDs) pour l'ensemble de base d'indicateurs communs sélectionnés*.

Dans un DS, les informations sont agrégées au sein de différentes feuilles de calcul Excel et, pour chaque feuille de calcul, plusieurs champs de formats différents sont définis. Lorsqu'un champ doit être rempli en sélectionnant une valeur incluse dans une liste prédéfinie de valeurs acceptables, ces listes constituent le dictionnaire de données associé au DS. Les données sont conformes aux DSs et DDs si et seulement si chaque contrôle qualité officiel suivant est satisfaisant:

- a) **Format** - Chaque champ est conforme à son format, c'est-à-dire que sa valeur est du texte, des chiffres ou dates selon le format requis ;
- b) **Codage unique - Les codes** utilisés pour identifier chaque ligne des feuilles de calcul Excel sont uniques, c'est-à-dire qu'il n'y a pas plus d'une ligne avec le même code au sein des feuilles de calcul où ces lignes définissent les objets associés. Par exemple, au sein des feuilles de calcul où les stations sont définies, chaque ligne a un code unique qui identifie la station. En général, il peut y avoir des feuilles de calcul où l'information est liée aux stations et des codes, identifiant ces stations, qui ne sont pas uniques. Par exemple dans les feuilles de calcul, où les valeurs de concentration de différents paramètres correspondant à une station sont représentées, le code de cette station est répété pour chaque valeur de concentration;
- c) **Liaison cohérente** - Les codes utilisés pour relier les informations présentes dans différentes feuilles de calcul doivent être cohérents. Par exemple, dans une feuille de calcul, il y a le code de la station, et ce dernier doit être présent dans la feuille de calcul Excel où les stations sont définies.
- d) **Expression régulière** - Chaque champ est conforme à une expression régulière spécifique lorsque ces expressions régulières sont requises. Par exemple, le champs "pourcentage" doivent être remplis avec une valeur numérique ayant le format 0-100 ou le champ "temps" doit respecter l'expression HH:MM:SS ;
- a) e) **Valeurs acceptables** - Chaque champ pour lequel il existe une liste de valeurs acceptables est rempli avec une et une seule valeur de cette liste, c'est-à-dire qu'il est conforme au dictionnaire de données (DD) associé à la norme de données (DS);

Les Dictionnaires de données sont l'ensemble des listes de valeurs acceptables, et sont représentés de deux manières. Dans un premier cas, la liste a la valeur suivie de sa description après le signe "=". Par exemple, la liste des valeurs acceptables représentant la matrice associée à la surveillance chimique est la suivante:

W = eau
S = sédiments
B = Biota

Dans ce cas, les valeurs acceptables à utiliser dans le champ "Matrice" sont "W", "S" ou "B".

Dans un second cas, la liste est constituée par un tableau comportant plusieurs colonnes, telle la liste des contaminants:

ID_Contaminant	Étiquette	Numéro CAS	Matrice	Obligatoire
CAS_309-00-2	Aldrin	309-00-2	Les sédiments	Y
...

Dans ce cas, le champ qui nécessite une telle liste doit spécifier la colonne de la liste à laquelle il se réfère. Par exemple, le champ “*DeterminHazSubsName*” qui se réfère à la liste des contaminants, doit être rempli avec une des valeurs de la colonne “*Étiquette*” de la liste, tandis que le champ “*DeterminHazSubsID*” doit être rempli avec une des valeurs de la colonne “*ID_Contaminant*” de la liste. En outre, “*DeterminHazSubsName*” et “*DeterminHazSubsID*” doivent être remplis de façon cohérente, de sorte que si “*DeterminHazSubsName*” est rempli avec la valeur “*Aldrin*”, dans la même ligne de la feuille de calcul Excel, “*DeterminHazSubsID*” doit être rempli avec la valeur “*CAS_309-00-2*”.

Les ensembles de données conformes à tous les contrôles qualité officiels cités ci-dessus (cf de a) à e)) doivent être considérés comme **formellement conformes** ou de bonne qualité d'un point de vue officiel.

La deuxième étape du processus d'assurance qualité consiste à mettre en œuvre des contrôles qualité officiels associés à chaque DS et DD afin de vérifier que les ensembles de données soient formellement conformes. En particulier, pour chaque DS correspondant à un module de surveillance, la liste ci-dessous des contrôles qualité formels a été définie et mise en œuvre dans le système d'information pilote IMAP. Chaque ligne de la liste est représentée par:

- **QC_Code** : Code unique identifiant le contrôle qualité;
- **Objectif écologique**: Objectif écologique auquel correspond DS;
- **Indicateur commun** : Indicateur commun IMAP auquel DS correspond;
- **DS / Module** : Standards de données ou module de surveillance auxquels s'applique le contrôle qualité;
- **Feuille de calcul Excel** : Feuille de calcul de la DS à laquelle s'applique le contrôle qualité;
- **Champ** : Champ inclus dans la feuille de calcul auquel le contrôle de qualité s'applique ;
- **Description** : Description du contrôle qualité appliqué au champ.

Il est reporté ci-dessous un exemple de contrôle qualité réalisé pour le Standard de données P1 - Contaminants dans l'eau de mer.

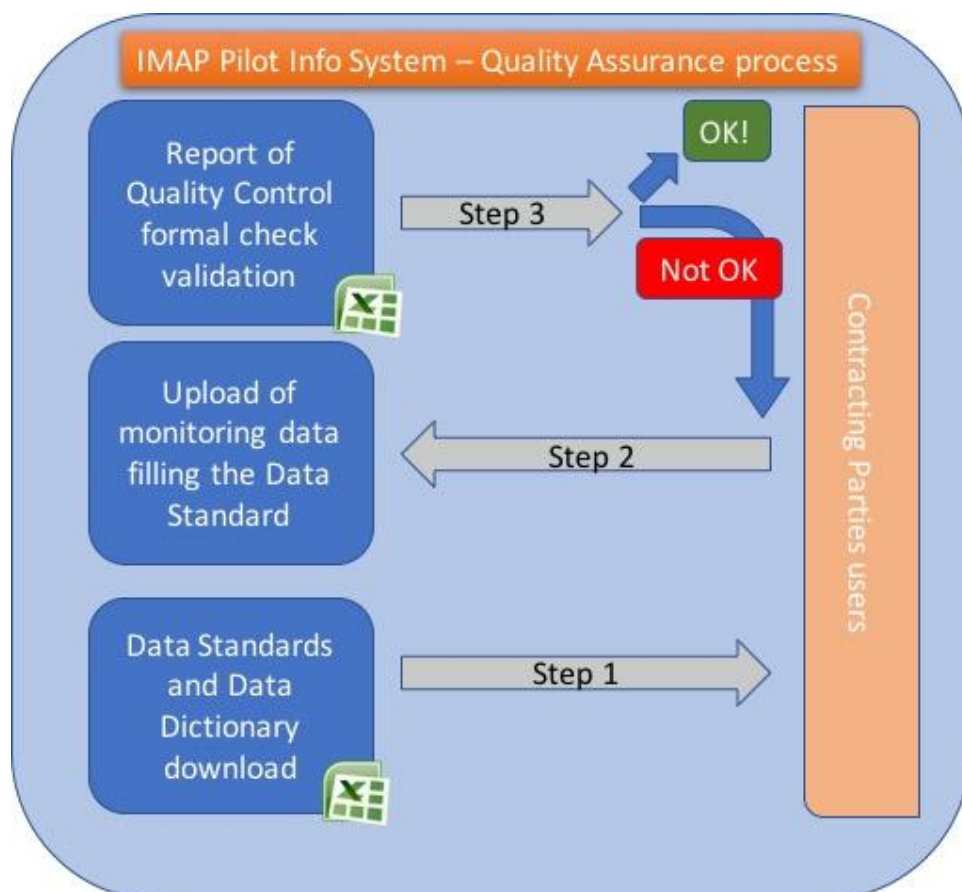
QC_Code	Objectif écologique	Indicateur commun	DS / module	Feuille de calcul Excel	Champ	Description
QC-E09-17-P1-00001	OE9 - Contaminants	Indicateur commun 17: Concentration des principaux contaminants nocifs mesurés dans la matrice correspondante (OE9, liée au biote, aux sédiments, à l'eau de mer)	P1 - Contaminants dans l'eau de mer	Stations	Code postal	La valeur correspond au code pays ISO 3166-1 alpha-2
QC-E09-17-P1-0000 2	OE9 - Contaminants	Indicateur commun 17: Concentration des principaux	P1 - Contaminants dans l'eau de mer	Stations	NationalStationID	Valeur unique dans la feuille de

		contaminants nocifs mesurés dans la matrice correspondante (OE9, liée au biote, aux sédiments, à l'eau de mer)				calcul
--	--	--	--	--	--	--------

Le processus de collecte et de contrôle qualité des ensembles de données est mis en œuvre pour chaque norme de données à travers la typique *communication à trois mains* :

1. **Étape 1** : l'utilisateur, en tant que Parties Contractantes, télécharge le standard de données correspondant au module de surveillance pour lequel il souhaite transférer des données de surveillance.
2. **Étape 2** : après avoir complété le standard de données avec les données de surveillance, l'utilisateur enregistre le fichier dans le système pour le flux de données qui correspond au standard de données utilisé.
3. **Étape 3** : Le système produit un rapport qui valide la vérification officielle du contrôle qualité (QC) avec les résultats du contrôle qualité officiel appliqué au fichier enregistré. Si le contrôle qualité est validé, le fichier est considéré comme " formellement conforme " (*OK!*), si c'est pas le cas (*Pas OK*), l'utilisateur doit corriger le fichier et l'enregistrer de nouveau dans le système afin de passer avec succès tous les contrôles qualité officiels

L'image suivante représente schématiquement le processus d'assurance qualité pour améliorer la qualité des ensembles de données transférés au système :



Le rapport de validation de la vérification officielle du contrôle qualité est produit sous forme de fichier Excel contenant les informations suivantes pour chaque ligne qui ne passe pas le contrôle qualité:

- **QC_Code** : Code unique qui identifie le contrôle qualité qui n'a pas été validé;
- **DS / Module** : Standards de données ou module de surveillance auxquels s'applique le contrôle qualité
- **Feuille de calcul Excel** : Feuille de calcul Excel du DS pour lequel le contrôle qualité n'a pas été validé
- **Champ** : Champ inclus dans la feuille de calcul Excel pour lequel le contrôle qualité n'a pas été validé
- **Ligne** : Numéro de la ligne de la feuille de calcul contenant le champ pour lequel le contrôle qualité n'a pas été validé.
- **Description** : Description du contrôle qualité non validé.

The following table contains an example with two rows of DS - P1 -Contaminants in seawater that are not compliant:

QC_Code	DS / module	Feuille de calcul Excel	Champ	Rangée	Description
QC-E09-17-P1-00001	P1 - Contaminants dans l'eau de mer	Stations	Code postal	34	La valeur correspond au code pays ISO 3166-1 alpha-2
QC-E09-17-P1-0000 2	P1 - Contaminants dans l'eau de mer	Stations	NationalStationID	45	Valeur unique dans la feuille de calcul

Dans l'exemple ci-dessus, l'utilisateur doit corriger ce qui a été inséré dans les lignes 34 et 45 pour les feuilles de calcul "*Stations*", respectivement dans les champs "*CountryCode*" et "*NationalStationID*" afin de rendre le fichier formellement conforme. Il est également clair que le codage unique de chaque contrôle qualité officiel et la production de rapports détaillés sur les résultats de leurs applications sont des outils indispensables du processus d'assurance qualité pour améliorer la qualité des données.

II. Développements futurs pour des Contrôles Qualité de Niveau Supérieur

Les ensembles de données officiellement conformes, c'est-à-dire qui ont passé les contrôles qualité confirmés du paragraphe précédent, sont stockés dans le système d'information pilote IMAP, un Système de Gestion de Base de Données Relationnelle (SGBDR), qui comprend également une extension SIG. Une fois que les données sont collectées dans ces infrastructures et qu'elles sont représentatives du bassin méditerranéen, il est possible de définir et d'appliquer des contrôles qualité de niveau supérieur concernant , par exemple, les problèmes suivants:

1. Contrôle des intervalles acceptables ou des valeurs maximales ou minimales des paramètres sur la base d'une analyse statistique des données de surveillance, de revues de la littérature scientifique ou de contraintes fixes dues à des caractéristiques physiques ou chimiques, comme par exemple l'intervalle de pH 0-14;
2. Emplacement géographique des points de surveillance à inclure dans la mer Méditerranée;

Mais un processus global d'assurance qualité pour la qualité des données inclut également d'autres contrôles qualité de niveau supérieur à différentes phases de la surveillance, telles que:

- Collecte d'échantillons
- Traitement des échantillons
- Déterminations analytiques

Les éléments ci-dessus sont directement liés aux procédures de laboratoire du contrôle et de la qualité des données avec un processus d'accréditation, une utilisation de matériaux de référence certifiés ou des protocoles de surveillance normalisés, et une participation du laboratoire aux tests de compétence.

Le système d'information pilote IMAP a été conçu pour effectuer, dans les futures phases, des contrôles qualité de plus haut niveau de type 1) et 2) et aussi pour recueillir des informations supplémentaires sur la qualité des données de laboratoire, sur les procédures de contrôle, et sur l'application des protocoles de surveillance. Ces informations seront incluses dans les DDs et DSs spécifiques et associées aux données de surveillance afin d'appliquer une catégorisation pour signaler des ensembles de données pour EO5 et EO9 sur l'exemple fourni dans le document UNEP / MED WG.473 / 9 – *Systèmes pour la qualité des bases de données et pour l' Assurance Qualité et le Contrôle qualité (QA/QC) de données sur la pollution* :

- **Catégorie A.** Laboratoires / Parties Contractantes déclarant la réussite de tests de compétence (z -score < 2) et / ou l'accréditation pour le produit chimique ou le paramètre analysé; les métadonnées ont été complétées et transmises à temps (délai maximum de 2 ans);
- **Catégorie B.** Laboratoires / Parties Contractantes déclarant des tests de compétence pour le produit chimique ou le paramètre analysé ($2 < z < 3$) et / ou l'accréditation; les métadonnées ont été complétées et transmises à temps (délai maximum de 2 ans);
- **Catégorie C.** Laboratoires / Parties Contractantes n'ayant pas participé aux tests de compétence lors des 2 dernières années; les métadonnées ont été complétées et transmises à temps. Elle pourrait également inclure la littérature scientifique présentant un rapport d'assurance qualité complet;
- **Catégorie D.** Laboratoires / Parties Contractantes n'ayant pas participé aux tests de compétence lors des 5 dernières années; les métadonnées ont été complétées mais n'ont pas été transmises à temps. Elle comprend également la littérature scientifique sans une particulière assurance qualité;
- **Catégorie E.** Laboratoires/ Parties Contractantes présentant des erreurs de déclaration importantes, mais les métadonnées pourraient être complétées et transmises à temps.

III. Conclusions

Le système d'information pilote IMAP effectue un processus d'assurance qualité qui comprend des contrôles qualité reconnus pour les 11 Indicateurs Communs IMAP sélectionnés. Il est également en mesure d'effectuer des contrôles qualité de niveau supérieur si des données de surveillance représentatives sont disponibles dans le système et si des informations détaillées, sur les procédures de contrôle et de qualité des données de laboratoire et sur l'application des protocoles de surveillance, sont fournies par les Parties Contractantes.