

Note technique

CONTRÔLE DES DONNÉES DE LA CAMPAGNE NATIONALE EXPLORATOIRE DE MESURE DES RÉSIDUS DE PESTICIDES DANS L'AIR AMBIANT (CNEP)

SYNTHÈSE

Les données de la campagne nationale exploratoire de mesure des résidus de pesticides dans l'air ambiant (CNEP) ont été intégrées dans la base nationale des données de qualité de l'air GEOD'AIR. A cette fin, le LCSQA a défini et mis en œuvre un ensemble de contrôles qualité aux différentes étapes du processus, de la saisie des données à leur mise en base, afin de garantir la cohérence et la fiabilité des informations bancarisées.

Au niveau régional, les données de mesure ont été compilées par chaque AASQA (association agréée de surveillance de la qualité de l'air) et centralisées en première étape dans PhytAtmo, base de données développée par les AASQA dans le cadre de leurs travaux antérieurs sur les pesticides, avant d'être transférées dans la base nationale GEOD'AIR.

Les données de la campagne extraites de la base nationale GEOD'AIR sont disponibles via le lien suivant :

<https://www.lcsqa.org/fr/rapport/resultats-de-la-campagne-nationale-exploratoire-de-mesure-des-residus-de-pesticides-dans>

1. CONTEXTE ET INTRODUCTION

La campagne nationale exploratoire de mesure des résidus de pesticides dans l'air ambiant [1] a conduit à l'intégration dans la base nationale de données GEOD'AIR d'environ 98000 données valides, associées à 50 sites de mesure et 75 substances (74 substances actives et 1 métabolite du glyphosate).

La présente note décrit les contrôles développés et mis en œuvre par le LCSQA aux différentes étapes de ce processus afin d'assurer la cohérence et la fiabilité des données mises en base.

Celles-ci se caractérisent par de nombreuses métadonnées propres à la problématique des pesticides. De ce fait, elles n'entrent pas dans le cadre usuel des données de surveillance réglementaire, dont la transmission vers le niveau national fait l'objet d'une supervision bien établie. Ces spécificités ont nécessité de construire des contrôles adaptés, en tenant compte des différentes sources d'erreur : saisie manuelle des données, multiplicité des métadonnées, opérations informatiques d'intégration et d'export de données.

2. CONTRÔLES QUALITÉ

2.1 Contrôles automatiques à la saisie

Un fichier Excel de saisie a été élaboré et consolidé par le LCSQA en collaboration avec les AASQA, afin de compiler des éléments de traçabilité relatifs au parcours de collecte et de validation des données.

Il comprenait des contrôles et des opérations automatiques destinés à prévenir les erreurs de saisie manuelle et à assister le travail des opérateurs :

- Listes de choix : type de préleveur, type de support de prélèvement, code qualité de la mesure, ...
- Calcul automatique des concentrations à partir des paramètres de masse et de volume et qualification de chaque mesure en fonction de la validation environnementale ;
- Mise en forme des données pour chargement dans la base PhytAtmo.

2.2 Contrôles après chargement dans PhytAtmo

Une première phase de contrôle de cohérence a été menée par le LCSQA sur les données intégrées dans PhytAtmo.

Ces vérifications ont notamment porté sur les points suivants :

- Conformité du remplissage des champs obligatoires (volume, concentration, date, ...) ;
- Conformité du format de certains champs : date, nom du site par exemple ;
- Vérification des codes qualité, des dates de mesure, des codes des substances ;
- Vérification de la cohérence entre les codes qualité et certaines informations (volume de prélèvement, conditions de transport par exemple) ;

- Vérification de l'existence de liens entre les différents objets constitutifs de la base : dates de prélèvement identiques, codes des sites, codes des polluants, ... ;
- Détection de doublons dans les données (enregistrements saisis deux fois) ;
- Saisie partielle des informations (périodes de prélèvement saisies sans données associées, statistiques météorologiques manquantes, ...)

En cas d'anomalie détectée, il a été demandé aux AASQA concernées de procéder aux corrections nécessaires et de réintégrer les données corrigées dans PhytAtmo. Cette étape s'est déroulée de décembre 2019 à mars 2020 et a été l'occasion de nombreuses itérations entre le LCSQA et les AASQA.

Les données ont alors été exportées de PhytAtmo pour être intégrées dans GEOD'AIR.

2.3 Intégration des données dans GEOD'AIR

En 2019, l'application GEOD'AIR a fait l'objet de développements spécifiques pour être en mesure d'accueillir les données de pesticides et leurs métadonnées qui diffèrent significativement (et logiquement) des métadonnées associées aux mesures de polluants réglementaires. Ces métadonnées couvrent la définition et les caractéristiques des points de prélèvement ainsi que des méthodes de prélèvement et d'analyse associées à chaque substance surveillée.

Cette dernière étape de bancarisation s'est déroulée sans difficulté particulière. Une dernière série de contrôles a permis de vérifier que la donnée n'avait pas été altérée lors de son intégration.

3. RÉFÉRENCES

[1] Résultats de la Campagne Nationale Exploratoire de mesure des résidus de Pesticides dans l'air ambiant (2018-2019). Rapport LCSQA 2020 (Marlière F., Salomon M. Létinois L.).