



## Annexe 5

# Méthodologie de validation des données

Les données transmises au Pôle Invertébrés seront soumises à un procédé de validation.

Celui-ci inclut :

- **Une méthode automatique**, destinée uniquement à traiter les données les plus redondantes et classiques, représentant des volumes souvent importants,
- **Une méthode manuelle**, destinée à traiter l'ensemble des cas susceptibles d'être litigieux, inhabituels ou portant sur des taxons jusqu'à lors trop peu connus.

Les règles permettant la validation automatique seront définies sur la base de jeux de données de référence, sur lesquels la fiabilité est jugée maximale (données d'experts, données analysées, revalidées, réexploitées etc). Les données ne répondant pas à ces filtres automatiques seront soumises à des experts chargés d'appliquer une validation manuelle.

Les critères de la validation automatique seront recalculés périodiquement, en intégrant les connaissances de l'ensemble des données validées précédemment par le mécanisme automatique ainsi que par les validateurs dans le cadre de la procédure manuelle.

La validation automatique ne peut en aucun cas écraser une validation manuelle. En revanche, un validateur peut décider, s'il le juge nécessaire, de corriger une validation automatique. Plusieurs statuts de validation manuelle peuvent se succéder en cas de nécessité.

Pour ce premier processus de validation automatique, les critères d'analyses resteront simples (taxon, localisation, période et altitude) et devront uniquement exclure du processus manuel les données redondantes et proches. Cette méthode simple pourra évoluer, à terme, vers un processus plus perfectionné appuyé sur une base de connaissances plus robuste.

Les adhérents seront informés périodiquement des données estimées douteuses ou invalides par les experts assurant la validation, afin de pouvoir apporter des précisions ou procéder à des corrections.