



SINP

Système d'Information
sur la Nature et le Paysage



Procédure nationale de validation scientifique des données élémentaires d'échange du SINP pour les occurrences de taxons

Version 1

-

Janvier 2017

Documents de référence :

Gargominy, O., Terceirie, S., Régnier, C., Ramage, T., Schoelinck, C., Dupont, P., Vandel, E., Daszkiewicz, P. & Poncet, L. 2015. *TAXREF v9.0, référentiel taxonomique pour la France : méthodologie, mise en œuvre et diffusion*. Muséum national d'Histoire naturelle, Paris. Rapport SPN 2015 – 64. 126 pp.

Haffner P., Touroult J., Da Costa H., Poncet L., 2012. *Atlas de la biodiversité départementale et des secteurs marins : Rapport méthodologique – version 1.1*. Service du patrimoine naturel, Muséum national d'histoire naturelle, Paris, 9 p.

Robert S., Barneix M., Body G., Castanet J., Caze G., Cellier P., Desse A., de Mazières J., Fromage P., Gourvil J., Jomier R., Juste A., Landry P., Lebeau Y., Lecoq M.E., Lescure J., Marage D., Meyer D., Pamerlon S., Papacotsia A., Poncet L., Quintenne G., Saltré A. & Touroult J. 2016. *Guide méthodologique pour la conformité, la cohérence et la validation scientifique des données et des métadonnées du SINP – Volet 1 : occurrences de taxons, Version 1*. Rapport pour le SINP, rapport MNHN-SPN 2016-77, 63 p.

Citation proposée :

Robert S., Dupont P., de Mazières J., Poncet L., Touroult J., 2017. *Procédure nationale de validation scientifique des données élémentaires d'échange du SINP pour les occurrences de taxons. Version 1*. Service du patrimoine naturel, Muséum national d'histoire naturelle, Paris. Rapport SPN 2017 - 2. 16 p.

Historique des versions :

Version	Date	Commentaire	Rédaction	Contribution et relecture
0	05/09/2016	Document préliminaire	Solène Robert	Julien Touroult Laurent Poncet
0.1	14/09/2016	Document soumis à relecture interne SPN élargie	Solène Robert	Pascal Dupont
0.2	14/10/2016	Document soumis à relecture interne SPN élargie	Solène Robert	Julien Touroult Jeanne de Mazières
0.3	21/12/2016	Modifications de terminologie et relatives aux évolutions prévues pour l'ABDSM	Solène Robert	Julien Touroult Pascal Dupont Jeanne de Mazières
1	16/01/2017	Version finalisée	Solène Robert	

Préambule :

La validation scientifique repose sur des processus d'expertises visant à renseigner sur la fiabilité, c'est-à-dire le degré de confiance que l'on peut accorder à la donnée. Ces processus font intervenir des bases de connaissance et/ou de l'expertise directe.

On distingue :

- La validation scientifique dite automatique : validation faisant appel à des résultats d'expertises préalables (des référentiels, des bases de connaissances, etc.). Ce processus est réalisé de manière automatique (informatique) dans le cadre du SINP.

- La validation scientifique dite manuelle (nonobstant le fait qu'une interface informatique puisse être utilisée) : validation des informations faisant appel à une expertise directe (avis d'expert suite à l'analyse des informations transmises). Ce processus est réalisé par un ou des experts dans le cadre du SINP.

Portée :

Ce document porte sur la validation scientifique des données d'occurrences de taxon transmises à l'INPN au titre de son rôle de plateforme nationale du système d'information sur la nature et les paysages (SINP). Il détaille les contrôles appliqués sur les données et traite de la traduction des résultats des différents contrôles en termes de niveau de validité et des conséquences pour la diffusion sur l'INPN, dans le cadre du SINP et dans le cadre de sa mission d'élaboration de couches de référence sur la répartition des taxons. Les données concernées par ces traitements sont les données élémentaires d'échange du SINP (DEE).

A noter :

Des procédures spécifiques de validation scientifique pourront être mises en place dans le cadre des inventaires nationaux. Ce document pourra être élargi dans ses futures versions pour porter les règles génériques définies dans le cadre de ces inventaires.

L'efficacité des contrôles est directement corrélée à la robustesse et à la complétude des référentiels qu'ils utilisent. L'analyse des résultats des contrôles par les experts devrait mettre en exergue les éventuelles lacunes de ces référentiels en ciblant sur les éléments à améliorer ou à actualiser.

Table des matières

A. Processus de validation automatique	5
1. Utilisation de TAXREF pour la validation scientifique	5
• Contrôle de reconnaissance par TAXREF (R)	5
• Contrôle du statut biogéographique (SB).....	6
• Contrôle de l'habitat (H).....	9
2. Utilisation d'ABDSM pour la validation scientifique	10
B. Attribution du niveau de validité à partir des contrôles automatiques.....	13
C. Processus de validation manuelle	16
D. Diffusion INPN et échanges SINP.....	16

A. Processus de validation automatique

La validation automatique fait appel à des couches de référence, des modèles probabilistes, des bases de connaissance ou à des référentiels pré-existants ou spécialement dédiés à l'opération de validation.

1. Utilisation de TAXREF pour la validation scientifique

TAXREF est le référentiel taxonomique national pour la faune, la flore et la fonge de France métropolitaine et d'outre-mer, élaboré et diffusé par le Muséum national d'Histoire naturelle (MNHN) dans le cadre de la mise en œuvre du Système d'Information sur la Nature et les Paysages (SINP). Ce référentiel est unique pour la France ; il a pour but de lister et d'organiser les noms scientifiques de l'ensemble des êtres vivants recensés sur le territoire national. Le référentiel TAXREF est consultable et téléchargeable sur le site internet de l'Inventaire National du Patrimoine Naturel¹.

A noter : les données élémentaires d'échange du SINP doivent comporter le code unique (cd_nom) des taxons selon le référentiel TAXREF dès lors que le taxon y est listé. Cela implique que les taxons cités par les producteurs des données et transmis aux plateformes régionales ou à la plateforme thématique occurrences de taxons aient été préalablement « réconciliés » avec ce référentiel. C'est-à-dire d'associer un cd_nom (identifiant unique du référentiel) au nom de taxon.

La réconciliation est l'étape qui consiste à associer un nom de taxon cité à une ligne dans un référentiel. Cette étape n'est pas décrite dans le présent document mais pourra faire l'objet d'une note spécifique décrivant la procédure de réconciliation utilisée par la plateforme thématique occurrences de taxon.

• Contrôle de reconnaissance par TAXREF (R)

Portée du contrôle R : toutes les données d'occurrence de taxons transmises sont concernées par le contrôle.

Description du contrôle R : le MNHN vérifie l'existence du cd_nom dans TAXREF :

- version en vigueur
- versions antérieures
- version en cours de mise à jour

La version en vigueur est la dernière version publiée du référentiel. Il s'agit de la version utilisée par l'INPN pour la diffusion.

Les versions antérieures sont les anciennes versions publiées du référentiel. Ces versions ne sont plus utilisées par l'INPN pour la diffusion mais y sont néanmoins bancarisées.

La version en cours de mise à jour est la version de travail de l'équipe Taxref qui préfigure la sortie de la version suivante du référentiel. Elle est alimentée au fur et à mesure de l'année par les experts en charge du référentiel.

Règles pour le contrôle R :

¹ <https://inpn.mnhn.fr/telechargement/referentielEspece/referentielTaxo>

Si le cd_nom existe dans la version de TAXREF en vigueur ou dans la version en cours de mise à jour, le test R est conclusif et positif. Résultat = **OK**

Si le cd_nom n'est pas présent dans TAXREF version en vigueur ou en cours mais existe dans l'une des précédentes versions, le test R est conclusif et négatif. Résultat = **KO**

Si le cd_nom n'est pas présent dans TAXREF, versions en vigueur, en cours, ou passées, alors il s'agit d'une erreur de conformité. Le fournisseur a utilisé une valeur qui n'existe pas dans TAXREF et doit corriger la donnée. Le test est inopérant. Résultat = **X**

• Contrôle du statut biogéographique (SB)

Les statuts biogéographiques définis dans TAXREF se veulent suffisamment généraux pour être transversaux à la faune, la flore et la fonge. Ils couvrent trois notions différentes : notion de présence (présence/absence), d'origine (indigénat ou introduction) et de géographie d'aire d'occupation (endémisme).

L'attribution des statuts biogéographiques aux taxons repose sur des publications scientifiques, et éventuellement sur les données d'experts dans le cadre de programmes nationaux (inventaires, listes rouges, etc.).

Tableau 1 : Définitions des statuts biogéographiques utilisés dans TAXREF

STATUT	DESCRIPTION	DEFINITION
P	Présent (indigène ou indéterminé)	Taxon présent au sens large dans la zone géographique considérée, c'est-à-dire taxon indigène ou taxon dont on ne sait pas s'il appartient à l'une des autres catégories. Le défaut de connaissance profite donc à l'indigénat. Par indigène on entend : taxon qui est issu de la zone géographique considérée et qui s'y est naturellement développé sans contribution humaine, ou taxon qui est arrivé là sans intervention humaine (intentionnelle ou non) à partir d'une zone dans laquelle il est indigène. (NB : exclut les hybrides dont l'un des parents au moins est introduit dans la zone considérée) <u>Sont regroupés sous ce statut tous les taxons catégorisés « natif » ou « autochtone ».</u> Les taxons hivernant quelques mois de l'année entrent dans cette catégorie.
B	Occasionnel	Taxon occasionnel, non nicheur, accidentel dans la zone géographique considérée (par exemple migrateur de passage).
E	Endémique	Taxon naturellement restreint à la zone géographique considérée.
S	Subendémique	Taxon naturellement restreint à une zone un peu plus grande que la zone géographique considérée mais dont les principales populations se situent dans la zone géographique considérée. Pour l'Outre-mer, on applique ce statut à l'endémisme régional : - pour la Guyane française = endémique du plateau des Guyanes, - pour les Antilles françaises = endémique des Petites Antilles, - pour Mayotte = endémique des Comores, - pour la Réunion = endémique des Mascareignes, - pour les TAAF = endémique de la South Indian Ocean Province.
C	Cryptogène	Taxon dont l'aire d'origine est inconnue et dont on ne peut donc pas dire s'il est indigène ou introduit.

STATUT	DESCRIPTION	DEFINITION
I	Introduit	<p>Taxon introduit (établi ou possiblement établi) dans la zone géographique considérée.</p> <p>Par introduit on entend : taxon dont la présence dans la zone géographique considérée est due à une intervention humaine, intentionnelle ou non, ou taxon qui est arrivé dans la zone sans intervention humaine mais à partir d'une zone dans laquelle il est introduit.</p> <p>Par établi (terme pour la faune, = naturalisé pour la flore) on entend : taxon introduit qui forme des populations viables (se reproduisant) et durables qui se maintiennent dans le milieu naturel sans besoin d'intervention humaine.</p> <p><u>Sont regroupés sous ce statut tous les taxons catégorisés</u> « non-indigène », « exotique », « exogène », « allogène », « allochtone », « non-natif », « naturalisé » (en anglais : <i>alien</i>) dans une publication scientifique.</p>
J	Introduit envahissant	<p>Taxon introduit dans la zone géographique considérée, qui produit des descendants fertiles souvent en grand nombre, et qui a le potentiel pour s'étendre de façon exponentielle sur une grande aire, augmentant ainsi rapidement son aire de répartition. Cela induit souvent des conséquences écologiques, économiques ou sanitaires négatives (IUCN, 2000).</p> <p><u>Sont regroupés sous ce statut tous les taxons catégorisés</u> « introduite envahissante », « exotique envahissant » ou « invasif » (<i>invasive</i> en anglais) dans une publication scientifique.</p>
M	Introduit non établi (dont domestique)	<p>Taxon introduit qui peut occasionnellement se reproduire en dehors de son aire de culture ou de captivité, mais qui ne peut se maintenir à l'état sauvage car ne pouvant former de populations viables sans intervention humaine, et qui dépend donc d'introductions répétées pour se maintenir dans la nature.</p> <p><u>Sont regroupés sous ce statut tous les taxons catégorisés</u> « introduit occasionnel », « subspontané », « échappé de culture ou de captivité » (en anglais : <i>casual alien</i>).</p> <p>Ce statut inclut les taxons strictement domestiques (faune) ou uniquement cultivés (flore).</p>
D	Douteux	Taxon dont la présence dans la zone géographique considérée n'est pas avérée (en attente de confirmation).
A	Absent	Taxon non présent dans la zone géographique considérée.
W	Disparu	<p>Taxon qui n'est plus présent à l'état sauvage dans la zone géographique considérée mais qui n'est pas globalement éteint.</p> <p>Rq : en cas de doute sur la présence ancienne ou non du taxon à l'état sauvage, utiliser le statut absent (A).</p>
X	Éteint	Taxon globalement éteint (= ayant totalement disparu de la surface du globe terrestre).
Y	Introduit éteint / disparu	Taxon introduit par le passé mais aujourd'hui disparu de la zone géographique considérée (W) ou éteint (X).
Z	Endémique éteint	Taxon endémique et aujourd'hui disparu, donc globalement éteint (X).
Q	Mentionné par erreur	Taxon mentionné par erreur comme présent sur le territoire considéré.

Ces notions sont contingentes à chacun des territoires administratifs de la France, à savoir :

- la métropole (continentale et Corse),
- la Guyane,
- la Martinique,

- la Guadeloupe,
- Saint-Martin,
- Saint-Barthélemy,
- Saint-Pierre-et-Miquelon,
- Mayotte,
- la Réunion,
- les îles Éparses (Tromelin, Les Glorieuses, Juan de Nova, Bassas de India, Europa),
- les Terres Australes et Antarctiques Françaises (îles Kerguelen, Crozet, Saint-Paul et Amsterdam, Terre-Adélie - hors îles Éparses),
- la Nouvelle-Calédonie,
- Wallis-et-Futuna,
- la Polynésie Française,
- l'île de Clipperton.

Ces territoires sont à considérer dans leur intégralité. Ainsi, une espèce qui est indigène d'au moins un site du territoire considéré est à prendre en compte comme indigène du territoire entier, même si elle s'avère introduite dans d'autres sites de ce territoire.

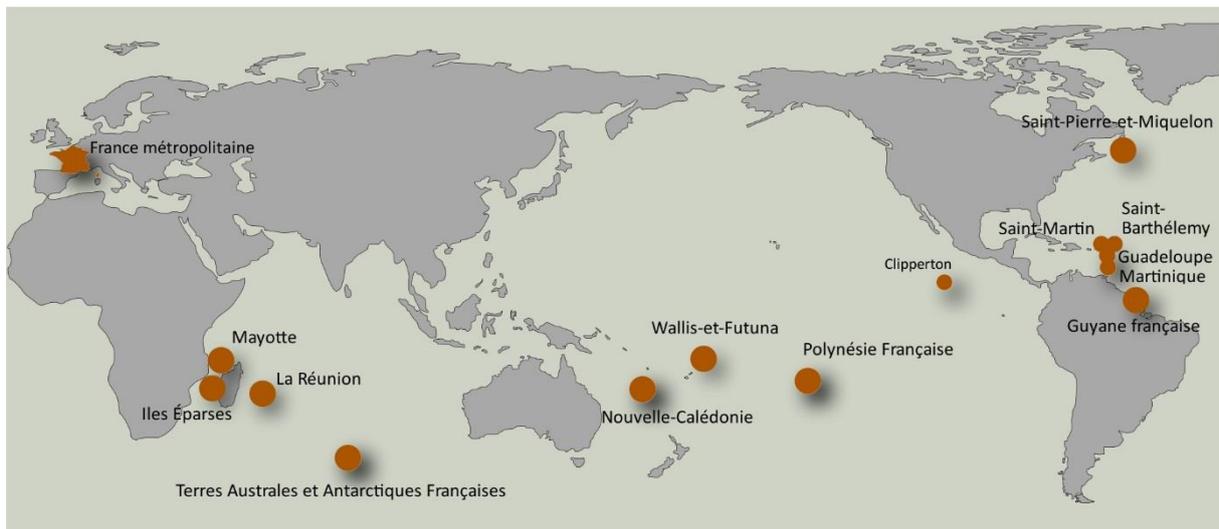


Figure 2 : Localisation des territoires français couverts par TAXREF

Portée du contrôle SB: le contrôle s'applique uniquement aux données d'observation (et pas de « Non Observation ») rattachées à TAXREF par un cd_nom (selon la version en vigueur ou la version en cours - soit R=OK) et qui sont de rang taxonomique égal ou inférieur au rang spécifique (espèce, sous-espèce, variété, sous-variété, forme, etc.).

Description du contrôle SB : le contrôle consiste à vérifier que la donnée d'observation est cohérente par rapport aux informations relatives au statut biogéographique contenues dans le référentiel en cours de mise à jour.

Règles du contrôle SB :

Pour le territoire concerné par la donnée, si le statut biogéographique est P, B, E, S, C, I, J ou M, alors le contrôle SB est conclusif et positif. Résultat = **OK**.

Pour le territoire concerné, si le statut biogéographique est A ou Q ou n'est pas stipulé (null), alors le contrôle SB est conclusif et négatif. Résultat = **KO1**.

Pour le territoire concerné, si le statut biogéographique est W, X, Y ou Z et que la donnée est récente (date_sup de moins de 10 ans), alors le contrôle SB est conclusif et négatif. Résultat = **KO2**.

Pour le territoire concerné, si le statut biogéographique est D, alors le contrôle SB est conclusif et négatif Résultat = **KO3**.

Pour le territoire concerné, si le statut biogéographique est W, X, Y ou Z et que la donnée est ancienne (date_sup de plus de 10 ans), alors le contrôle n'est pas conclusif. Résultat = **X**.

A noter : les données non concernées par le contrôle (voir portée du contrôle) prennent la valeur SB=null.

• **Contrôle de l'habitat (H)**

Portée du contrôle H : le contrôle s'applique uniquement aux données d'observation (et pas de « Non Observation ») rattachées à TAXREF par un cd_nom (selon la version en vigueur ou la version en cours - soit R=OK) et qui sont de rang taxonomique égal ou inférieur au rang spécifique (espèce, sous-espèce, variété, sous-variété, forme, etc.).

Les données situées sur des mailles 10x10 Km à la fois continentale et marine ne sont pas concernées par ce contrôle puisque, dans ce cas, toutes les valeurs d'habitat y sont acceptables.

Tableau 2 : Valeurs de la table TAXREF_HABITATS.

HABITAT	DESCRIPTION	REMARQUES
1	Marin	Espèces vivant uniquement en milieu marin
2	Eau douce	Espèces vivant uniquement en milieu d'eau douce
3	Terrestre	Espèces vivant uniquement en milieu terrestre
4	Marin & Eau douce	Espèces effectuant une partie de leur cycle de vie en eau douce et l'autre partie en mer (espèces diadromes, amphidromes, anadromes ou catadromes).
5	Marin & Terrestre	Cas des pinnipèdes, des tortues et des oiseaux marins (par exemple).
6	Eau saumâtre	Espèces vivant exclusivement en eau saumâtre.
7	Continental (terrestre et/ou eau douce)	Espèces continentales (non marines) dont on ne sait pas si elles sont terrestres et/ou d'eau douce (taxons provenant de Fauna Europaea).
8	Continental (terrestre et eau douce)	Espèces terrestres effectuant une partie de leur cycle en eau douce (odonates par exemple), ou fortement liées au milieu aquatique (loutre par exemple).

Description du contrôle H :

Le contrôle consiste à vérifier que l'habitat associé au taxon est cohérent avec la situation géographique de la maille 10x10Km où est située l'observation.

Le contrôle se base sur le maillage national 10x10 Km (Lambert 93). Chaque maille porte deux attributs :

- **Continental :** cet attribut est renseigné à true lorsque la maille couvre le domaine continental et à false lorsqu'elle ne couvre pas le domaine continental.
- **Marine :** cet attribut est renseigné à true lorsque la maille couvre le domaine marin et à false lorsqu'elle ne couvre pas le domaine marin.

A noter : une maille peut ainsi être à la fois continentale et marine.

Considérant la difficulté d'évaluer exactement la démarcation terre/mer, une imprécision (buffer) est appliquée par rapport à la limite de la BD Topo de 1 000 m (soit une imprécision de 2 000 m de largeur totale).

Règles pour le contrôle H :

Pour une maille strictement continentale, si l'habitat du taxon est 2, 3, 4, 5, 6, 7 ou 8, alors le contrôle H1 est conclusif et positif. Résultat = **OK**.

Pour une maille strictement continentale, si l'habitat du taxon est 1 alors le contrôle H1 est conclusif et négatif. Résultat = **KO**.

Pour une maille strictement marine, si l'habitat du taxon est 1, 4, 5 ou 6 alors le contrôle H1 est conclusif et positif. Résultat = **OK**.

Pour une maille strictement marine, si l'habitat du taxon est 2, 3, 7 ou 8, alors le contrôle H1 est conclusif et négatif. Résultat = **KO**.

Pour les mailles mixtes, le test n'est pas conclusif. Résultat = **X**.

A noter : les données non concernées par le contrôle (voir portée du contrôle) prennent la valeur H=null.

2. Utilisation d'ABDSM pour la validation scientifique

L'Atlas de la Biodiversité Départementale et des Secteurs Marins est un programme du MNHN dont le but est la réalisation de cartes nationales de répartition géographique d'espèces (et sous-espèces) sous la forme d'une présence/absence actuelle expertisée par département. Par extension, le milieu marin a été découpé en secteurs marins qui sont des unités calées en partie sur les départements terrestres.

Les cartes ABDSM permettent de contrôler la cohérence globale des données de résolution plus fine transmises pour intégration dans l'INPN. Elles permettent par exemple de détecter des observations situées en dehors de leur aire de répartition actuelle connue.

La démarche ABDSM donne un rôle clé à l'expertise. L'expert qui réalise la carte utilise à la fois les données disponibles, la littérature, sa connaissance biogéographique etc. Cet atlas génère un statut de présence/absence par taxon et par département ou secteur marin selon la liste des valeurs suivantes :

P1. « Pas d'information ». Dans l'état actuel de la connaissance, on ne peut pas se prononcer sur la présence ou l'absence.

P2. « Absence probable ou certaine ». Ce point recouvre l'absence (probable ou certaine), par nature plus difficile à démontrer que la présence. Ce statut est également attribué à un département dans lequel la présence de l'espèce est uniquement occasionnelle.

P3. « Disparition avérée ». Cas particulier : absence liée à une disparition avérée depuis moins d'un demi-siècle (si disparition ancienne, utiliser alors le code P2).

P4. « Présence probable ». Ce statut se base sur un ou plusieurs des critères suivants : une recherche de l'espèce incomplète mais présence de milieux favorables ; une écologie de l'espèce compatible avec l'hypothèse de sa présence ; une dernière observation fiable remontant à plus de 10 ans (20 ans pour les taxons peu étudiés)* par rapport à la date de référence, aucune recherche spécifique récente et aucune présomption de disparition depuis cette date [vertébrés, plantes et invertébrés bien étudiés (rhopalocères, orthoptères, odonates...)].

P5. « Présence certaine ». Correspond à un signalement sur la base d'au moins une observation avérée dans une période de 10 ans (20 ans pour les invertébrés peu connus)* précédant l'année de référence et aucune preuve ou présomption de disparition depuis l'obtention de la dernière donnée.

Un niveau de validation est affecté à chaque carte :

- **Initiale** : la carte n'est pas encore soumise au processus d'établissement de la carte de référence.
- **En cours** : la carte est soumise au processus d'établissement de la carte de référence. Au moins un département ou un secteur marin est associé à un statut de présence et un statut de reproduction et d'indigénat par un rédacteur.
- **A valider** : la carte est soumise au processus d'établissement de la carte de référence. Le ou les rédacteurs ont associé l'ensemble des départements et des secteurs marins à un statut de présence et un statut de reproduction et d'indigénat. La carte est soumise à un validateur.
- **Validée** : carte pour laquelle l'ensemble des statuts de présence et d'indigénat pour chaque département et secteur marin a été validée par un validateur. La carte est qualifiée de référence.

Portée du contrôle ABDSM : ne sont actuellement contrôlables que les données d'observation (et non de Non Observation) situées en France métropolitaine et pour lesquelles les taxons sont de rang spécifique ou infra-spécifique et rattachés à TAXREF par un cd_nom (selon la version en vigueur ou la version en cours - soit R=OK).

Les données pour lesquelles le statut biologique est noté « erratique », « halte migratoire » ou « passage en vol » ne sont pas concernées par le contrôle.

Tableau 3 : Statuts biologiques selon le standard DEE (au 14/06/2016)

Code	Libellé	Définition
0	Inconnu	Inconnu : Le statut biologique de l'individu n'est pas connu.
1	Non renseigné	Non renseigné : Le statut biologique de l'individu n'a pas été renseigné.
2	Non Déterminé	Non déterminé : Le statut biologique de l'individu n'a pas pu être déterminé.
3	Reproduction	Reproduction : Le sujet d'observation en est au stade de reproduction (nicheur, grvide, carpophore, floraison, fructification...)

* La fraîcheur des observations considérées est susceptible d'évoluer pour être adaptée selon les groupes taxonomiques.

4	Hibernation	Hibernation : L'hibernation est un état d'hypothermie régulée, durant plusieurs jours ou semaines qui permet aux animaux de conserver leur énergie pendant l'hiver.
5	Estivation	Estivation : L'estivation est un phénomène analogue à celui de l'hibernation, au cours duquel les animaux tombent en léthargie. L'estivation se produit durant les périodes les plus chaudes et les plus sèches de l'été.
6	Halte migratoire	Halte migratoire : Indique que l'individu procède à une halte au cours de sa migration, et a été découvert sur sa zone de halte.
7	Swarming	Swarming : Indique que l'individu a un comportement de swarming : il se regroupe avec d'autres individus de taille similaire, sur une zone spécifique, ou en mouvement.
8	Chasse / alimentation	Chasse / alimentation : Indique que l'individu est sur une zone qui lui permet de chasser ou de s'alimenter.
9	Pas de reproduction	Pas de reproduction : Indique que l'individu n'a pas un comportement reproducteur. Chez les végétaux : absence de fleurs, de fruits...
10	Passage en vol	Passage en vol : Indique que l'individu est de passage et en vol.
11	Erratique	Erratique : Individu d'une ou de populations d'un taxon qui ne se trouve, actuellement, que de manière occasionnelle dans les limites d'une région. Il a été retenu comme seuil, une absence de 80% d'un laps de temps donné (année, saisons...).
12	Sédentaire	Sédentaire : Individu demeurant à un seul emplacement, ou restant toute l'année dans sa région d'origine, même s'il effectue des déplacements locaux.

Description du contrôle ABDSM :

Ce contrôle consiste à utiliser les couches de référence départementales disponibles afin d'identifier rapidement les données incohérentes (taxon largement en dehors de son aire de répartition par exemple).

Seules les cartes ABDSM « validées » ou « à valider » sont utilisées pour ce contrôle.

A noter : les cartes ABDSM sont produites sur le nom de référence (cd_ref de TAXREF). Les noms utilisés dans les données à contrôler (ensemble des synonymes) sont donc préalablement rattachés au nom de référence (CD_REF) selon la version de TAXREF en vigueur.

Règles pour le contrôle ABDSM :

A1 : Si une donnée récente (moins de 20 ans) se situe dans un département à statut P2 ou P3 alors le contrôle est conclusif et négatif. Résultat= **A1**.

A2 : Si une donnée ancienne se situe dans un département P2 ou P3 alors le test n'est pas conclusif. Résultat = **A2**.

A3 : Quelle que soit la date d'observation de la donnée, si la donnée se situe dans un département à P4 ou P5, alors le contrôle est conclusif et positif. Résultat = **A3**.

A4 : Quelle que soit la date d'observation, si la donnée se situe dans un département P1, alors le contrôle n'est pas conclusif. Résultat= **A4**.

Règles spécifiques pour les données qui concernent des taxons migrants ou à forte dispersion depuis leur zone de reproduction (oiseaux et odonates considérés pour l'instant) (écrase le statut précédent le cas échéant) :

B1 : Si le statut biologique de la donnée prend la valeur « Reproduction », que la donnée se situe dans un département à statut P2 ou P3 et que la donnée est récente (moins de 20 ans) alors le contrôle est conclusif et négatif. Résultat = **A2**.

Autre cas particulier (écrase le statut précédent le cas échéant):

C1 : Si une donnée relative à un taxon dont le statut de présence Taxref est I, J ou M (= introduit) se situe dans un département à statut P2 ou P3, alors le contrôle n'est pas conclusif. Résultat=**C1**.

*A noter : les données non concernées par le contrôle (voir portée du contrôle) prennent la valeur **ABDSM=null**.*

B. Attribution du niveau de validité à partir des contrôles automatiques

Les contrôles définis ci-avant donnent lieu à une traduction en termes de niveau de validité. Les niveaux de validité utilisés sont ceux du SINP :

Valeurs	Libellés
1	Certain - très probable
2	Probable
3	Douteux
4	Invalide
5	Non réalisable
null	Non évalué

Libellés	Définition
Certain - très probable	La donnée présente un haut niveau de vraisemblance (très majoritairement cohérente) selon la procédure automatique appliquée. Le résultat de la procédure correspond à la définition optimale de satisfaction de l'ensemble des critères de la procédure automatique, par exemple, lorsque la localité correspond à la distribution déjà connue et que les autres paramètres écologiques (date de visibilité, altitude, etc.) sont dans la gamme habituelle de valeur.
Probable	La donnée est cohérente et plausible selon la procédure automatique appliquée mais ne satisfait pas complètement (intégralement) l'ensemble des critères automatiques appliqués. La donnée présente une forte probabilité d'être juste. Elle ne présente aucune discordance majeure sur les critères jugés les plus importants mais elle satisfait seulement à un niveau intermédiaire, ou un ou plusieurs des critères automatiques appliqués.
Douteux	La donnée concorde peu selon la procédure automatique appliquée. La donnée est peu cohérente ou incongrue. Elle ne satisfait pas ou peu un ou plusieurs des critères automatiques appliqués. Elle ne présente cependant pas de discordance majeure sur les critères jugés les plus importants qui permettraient d'attribuer le plus faible niveau de validité (invalide).
Invalide	La donnée ne concorde pas selon la procédure automatique appliquée. Elle présente au moins une discordance majeure sur un des critères jugés les plus importants ou la majorité des critères déterminants sont discordants. Elle est considérée comme trop improbable (aberrante notamment au regard de l'aire de répartition connue, des paramètres biotiques et abiotiques de la niche écologique du taxon). Elle est considérée comme invalide.
Non réalisable	La donnée a été soumise à l'ensemble du processus de validation mais l'opérateur (humain ou machine) n'a pas pu statuer sur le niveau de fiabilité. Notamment: -Etat des connaissances du taxon insuffisantes -Informations insuffisantes sur l'observation
Non évalué	Niveau initial ou temporaire. La donnée n'a pas été soumise à l'opération de validation ou l'opération n'est pas encore terminée (validation en cours). Elle n'est donc pas évaluée à un temps précis défini par la date de validation.

Affectation du niveau de validité

Les contrôles sont ordonnés (Cf. figure 3) et les contrôles R, SB et H sont des contrôles « invalidants », c'est-à-dire pour lesquels l'obtention d'un résultat négatif entraîne l'arrêt de la chaîne de contrôle.

Si le contrôle R de reconnaissance TAXREF a donné un résultat conclusif négatif (KO), alors les contrôles de statut biogéographique SB, habitat H et ABDSTM ne sont pas appliqués. Les données prennent un niveau de validité = « Invalide ».

Si le contrôle SB donne un résultat conclusif négatif (KO1 ou KO2) ou que le contrôle H donne un résultat conclusif négatif (KO), alors le contrôle ABDSTM n'est pas appliqué. Les données prennent un niveau de validité = « Invalide ».

Si le contrôle SB donne un résultat conclusif négatif (KO3), alors le contrôle ABDSM n'est pas appliqué. Les données prennent un niveau de validité « Douteux ».

Si le contrôle H donne un résultat conclusif négatif (KO), alors le contrôle ABDSM n'est pas appliqué. Les données prennent un niveau de validité « Invalide ».

Les données ayant passé avec succès les tests R, H et SB sont soumises au contrôle ABDSM.

Si ABDSM n'est pas disponible ou que les données sont hors périmètre du contrôle (voir portée du contrôle), alors le niveau de validation automatique est jugé « **non réalisable** ».

Si ABDSM est disponible et que la donnée entre dans le périmètre du contrôle :

- Si le test est conclusif positif : le niveau de validité est « **Probable** »
- Si le test est conclusif négatif : le niveau de validité est « **Douteux** »
- Si le test n'est pas conclusif : le niveau de validité est « **Non réalisable** »

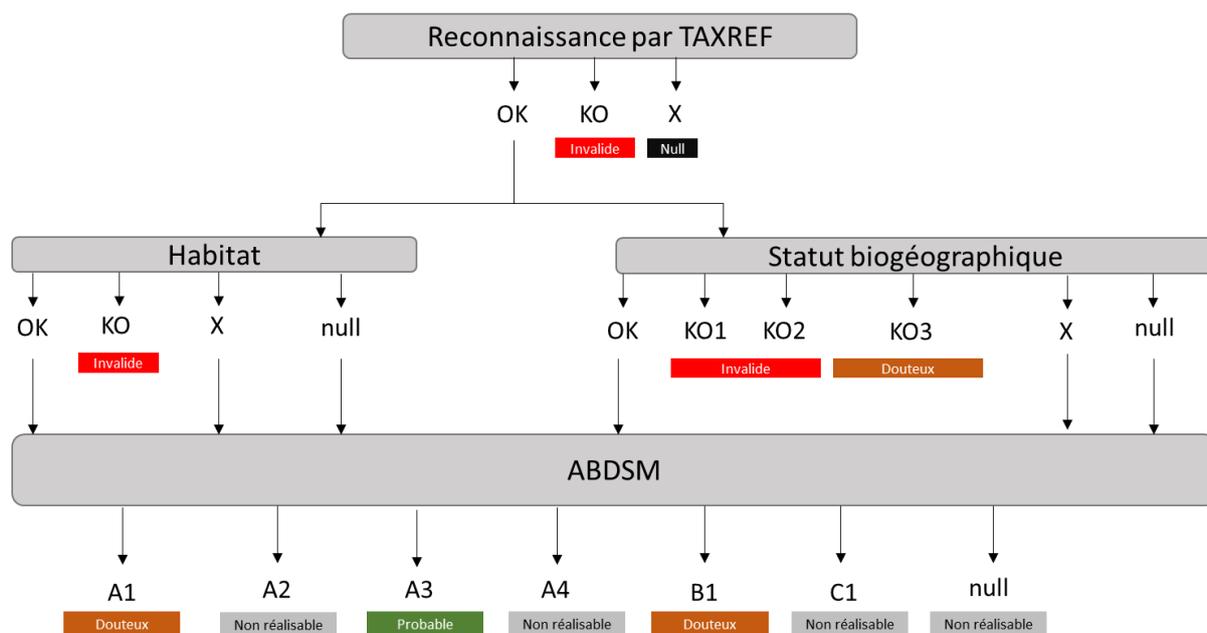


Figure 3 : Organisation des contrôles, résultats et attribution des niveaux de validation automatique

C. Processus de validation manuelle

La validation manuelle fait intervenir des experts mandatés par la plateforme en charge de la validation. Ces experts peuvent évaluer l'ensemble des données, quel que soit le niveau de validité attribué automatiquement à la donnée, y compris si la donnée n'a pas suivi la phase de validation automatique. Ils doivent avoir accès à l'ensemble des informations portées au niveau des données et au niveau des métadonnées nécessaires à assurer leur expertise.

Ils ont la possibilité de recontacter l'observateur, le déterminateur ou l'organisme producteur de manière à demander des éléments complémentaires à leur analyse.

La validation manuelle fait foi par rapport à la validation automatique.

La validation manuelle peut intervenir avant ou après la mise en diffusion des données.

Si une procédure plus fine de validation scientifique est mise en place dans le cadre des inventaires nationaux (automatique ou manuelle), cette validation fait également foi par rapport à la validation automatique présentée ci-avant.

D. Diffusion INPN et échanges SINP

Nous distinguons actuellement la diffusion des données dans l'INPN au titre de sa mission d'élaboration de couches de référence (i) de l'échange des données dans le cadre du SINP au titre de son rôle de plateforme nationale (ii).

Dans le premier cas (i), les données douteuses ou invalides sont écartées pour la diffusion.

Dans le second cas (ii), toutes les données, quel que soit leur niveau de validité peuvent être échangées entre les plateformes du SINP.

Le niveau de validité se traduit ainsi de la manière suivante :

Niveau de validité	Echange SINP	Diffusion INPN (cartes de référence)
Certain - très probable	True	True
Probable	True	True
Douteux	True	False
Invalide	True	False
Non réalisable	True	True
Non évalué	True	True