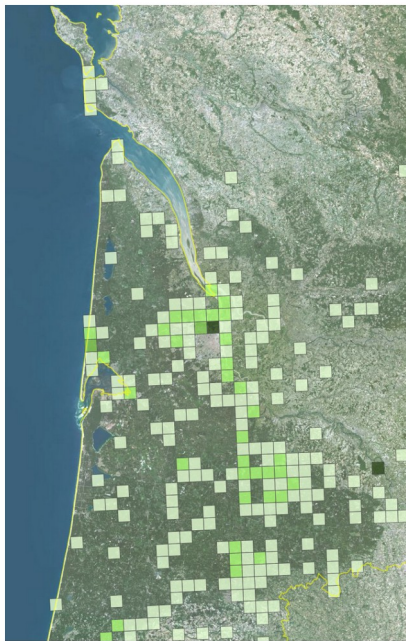




## Procédure de validation régionale des données d'occurrence de taxon de l’Observatoire FAUNA

Version 1.5



**SINP**  
Système d'Information  
sur la Nature et le Paysage

université  
de **BORDEAUX**



RÉGION  
**Nouvelle-  
Aquitaine**

**Coordination** : Luc CHEVALLIER

**Rédaction / Contributions** : Luc CHEVALLIER, Marie BARNEIX, Paul FROMAGE

**Référence** : FAUNA, 2021. Procédure de validation régionale des données d'occurrence de taxon de l'Observatoire FAUNA. Version 1.5. 16 p.

**Crédits photos** :

- Marie BARNEIX: Salamandre tachetée (*Salamandra salamandra*) ; Tabac d'Espagne (*Argynnis paphia*)
- Thomas RUYS : Bouquetin ibérique (*Capra pyrenaica*)

**Partenaires financiers** : Direction Régionale de l'Environnement, de l'Aménagement et du Logement Nouvelle-Aquitaine et Région Nouvelle-Aquitaine.

---

L'**Observatoire FAUNA** est un dispositif dédié à la coordination et à la valorisation des informations faunistiques. Il a pour vocation d'améliorer l'échange de données, l'accès aux connaissances et l'identification des enjeux liés à la préservation des espèces animales, à partir de l'expertise croisée de son réseau de contributeurs. En tant que centre d'appui technique pour les décideurs publics et les acteurs régionaux, sa finalité vise à disposer du socle fondamental de connaissances permettant d'orienter des actions de préservation de la faune sauvage et de ses habitats.



**En savoir plus** : <https://observatoire-fauna.fr/>

L'Observatoire FAUNA est adossé à l'Unité Mixte de Recherche "Biodiversité, Gènes et Communautés" (UMR BioGeCo) de l'Université de Bordeaux et de l'Institut National de Recherche Agronomique (INRA) et géré en partenariat avec l'ADERA.



## SOMMAIRE

<b>1. Périmètre et pré-requis.....</b>	<b>4</b>
<b>2. Validation automatique.....</b>	<b>5</b>
<b>2.1. Critères relatifs aux niveaux de validité existants.....</b>	<b>5</b>
2.1.1. La validation du producteur.....	5
2.1.2. La validation nationale.....	6
<b>2.2. Critères relatifs à l'écologie du taxon.....</b>	<b>6</b>
2.2.1. L'habitat.....	6
2.2.2. La période d'observation.....	7
2.2.3. La répartition géographique.....	8
<b>2.3. Critère relatif à l'identification du taxon.....</b>	<b>9</b>
<b>2.4. Attribution de la note finale de validité régionale automatique.....</b>	<b>10</b>
<b>2.5. Cas particulier des oiseaux soumis à homologation.....</b>	<b>12</b>
<b>3. Validation manuelle.....</b>	<b>13</b>
<b>4. Validation régionale.....</b>	<b>13</b>
<b>Bibliographie.....</b>	<b>14</b>
<b>Annexes.....</b>	<b>15</b>
<b>Annexe 1 : Complétudes des référentiels aquitains utilisés (à la date du rapport).....</b>	<b>15</b>
<b>Annexe 2 : Statuts de présence régionaux et départementaux du référentiel espèces de l'Observatoire FAUNA.....</b>	<b>16</b>

## 1. PÉRIMÈTRE ET PRÉ-REQUIS

Le Système d'Information de l'Observatoire FAUNA, vise à améliorer la collecte de données et la restitution d'informations sur les espèces animales et leurs habitats. Il s'appuie sur un large réseau de contributeurs et constitue l'une des bases thématiques supports de la plateforme régionale du Système d'Information sur la Nature et les Paysages (SINP).

Les observations, collectées selon des objectifs et protocoles variés puis mutualisées à travers le SI de l'observatoire, sont regroupés en jeux de données. Ces jeux de données ont vocation à nourrir les politiques publiques et les programmes de connaissance et de préservation du patrimoine naturel. Dans ce cadre, la fiabilité des données est un enjeu majeur de la gestion de ces informations. Elle conditionne à la fois la crédibilité de la démarche et la robustesse des analyses qui en résultent. Il est donc important que les différents utilisateurs soient en mesure de connaître le niveau de fiabilité de chaque donnée partagée dans le système.

Le terme "donnée" fait ici référence à l'observation d'une espèce animale. Également appelée "donnée d'occurrence de taxon" par le SINP, elle est décrite a minima par un taxon cité à une date et un lieu donnés par un ou plusieurs observateurs.

La procédure de validation régionale présentée dans ce rapport a pour objectif d'attribuer un niveau de fiabilité à chaque donnée partagée sur la faune. Elle vient à la suite des procédures de conformité et de cohérence, visant notamment le respect des règles fixées par les formats standards de données et métadonnées. Dans le cadre du SINP, cette validation régionale est complémentaire à la validation « producteur » et à la validation « nationale » (Robert et al., 2016).

La validation régionale s'appuie essentiellement sur des contrôles automatiques faisant intervenir des bases de connaissance existantes (référentiels, couches atlas, etc.). Cette validation automatique ne concerne que le périmètre minimal de la donnée à savoir : la validation du taxon, de la date et du lieu. Elle ne concerne également que les données de présence, la validation des données d'absence étant automatiquement considérée comme non applicable. Les données renseignées à la sous-espèce sont validées automatiquement au rang taxonomique supérieur : l'espèce.

*Remarque : Les bases de connaissances utilisées dans ce protocole permettent de disposer d'informations structurées au rang de l'espèce. Leur complétude peut varier en fonction des groupes taxonomiques et des territoires (voir annexe 1). Afin de pouvoir attribuer un niveau de fiabilité aux observations supra-spécifiques (genre, famille, etc.), ces éléments peuvent être agrégés à partir d'informations disponibles sur les taxons inférieurs. Par exemple, pour un genre donné, si au moins une espèce est considérée présente au niveau régional, le genre le sera également.*

Additionnellement à cette validation automatique, une validation dite « manuelle » peut être réalisée par des experts référents suite à l'analyse d'éléments complémentaires (preuve, méthode d'observation, etc). Elle peut aussi bien concerner le périmètre minimal de la donnée (taxon, date et lieu), que le périmètre élargi (dénombrement, stade, etc.). Cette validation « manuelle » prime sur la validation « automatique » et lorsqu'elle s'appuie sur cette dernière, elle est alors dite « combinée ».

Toutes les données partagées au sein du SINP régional et étant rattachées au référentiel taxonomique TAXREF ont vocation à être soumises au processus de validation régional, quel que soit leur statut (données d'origine publique ou privée).

Cette procédure s'appuie en partie sur les travaux du Guide méthodologique pour la conformité, la cohérence et la validation scientifique des données et des métadonnées du SINP (Robert et al., 2016).

## 2. VALIDATION AUTOMATIQUE

La validation automatique d'une donnée consiste à contrôler un ensemble de critères relatifs :

- aux niveaux de validité existants de la donnée ;
- à l'écologie du taxon observé ;
- à la difficulté de détermination du taxon observé.

Ces éléments permettront d'attribuer un premier niveau de validation régionale pour chaque donnée.

### 2.1. Critères relatifs aux niveaux de validité existants

Plusieurs niveaux de validation peuvent coexister pour une même donnée.

#### 2.1.1. La validation du producteur

La validation dite "producteur" s'appuie sur le niveau de validation identifié et communiqué à l'échelle de la donnée par le producteur. Ce critère a ainsi vocation à automatiser l'identification de données communiquées comme étant « douteuses » ou « invalides » par le producteur (figure 1).

*Remarque* : Ces procédures de validation, non standardisées, sont spécifiques à chaque producteur. Une correspondance avec le standard régional est donc réalisé avant l'intégration dans le SI de l'observatoire. L'utilisation directe de nomenclatures standardisées communes par les producteurs faciliterait ce traitement.

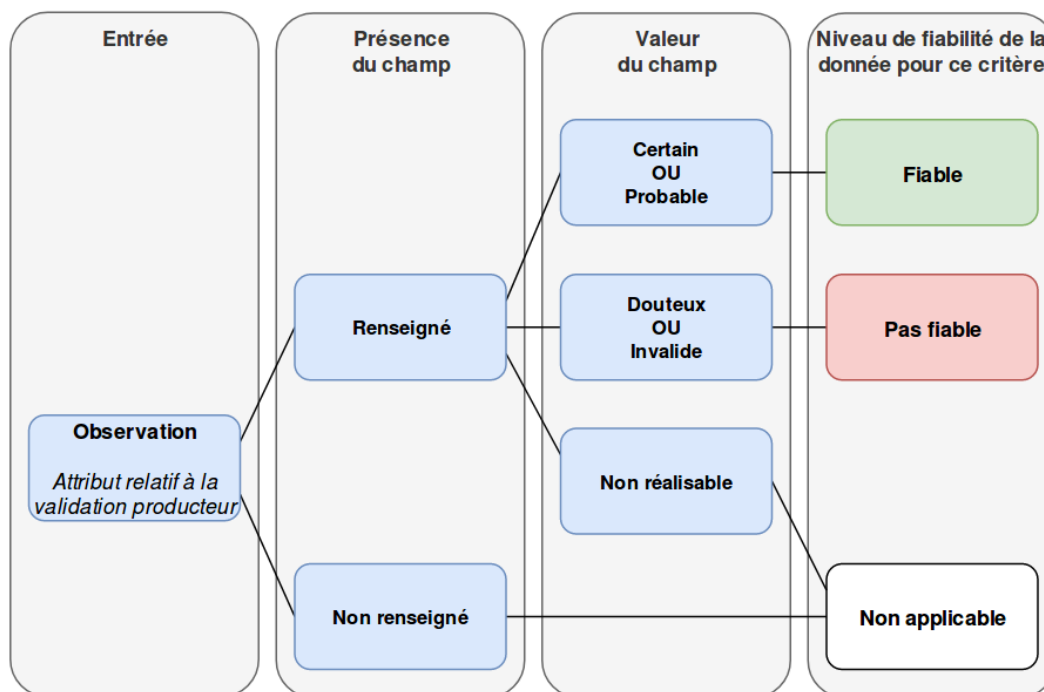


Figure 1 : Tests du critère "validation producteur"

## 2.1.2. La validation nationale

La validation dite « nationale » correspond à la validation coordonnée par l'UMS Patrinat, gestionnaire de la plateforme nationale des données partagées dans le cadre du SINP (Robert et al., 2017). Ce contrôle a vocation à automatiser l'identification au niveau régional de données communiquées « douteuses » ou « invalides » par le niveau national (figure 2).

*Remarque* : Le niveau de validation nationale n'étant actuellement pas communiqué, ce contrôle n'est pas effectué à ce stade.

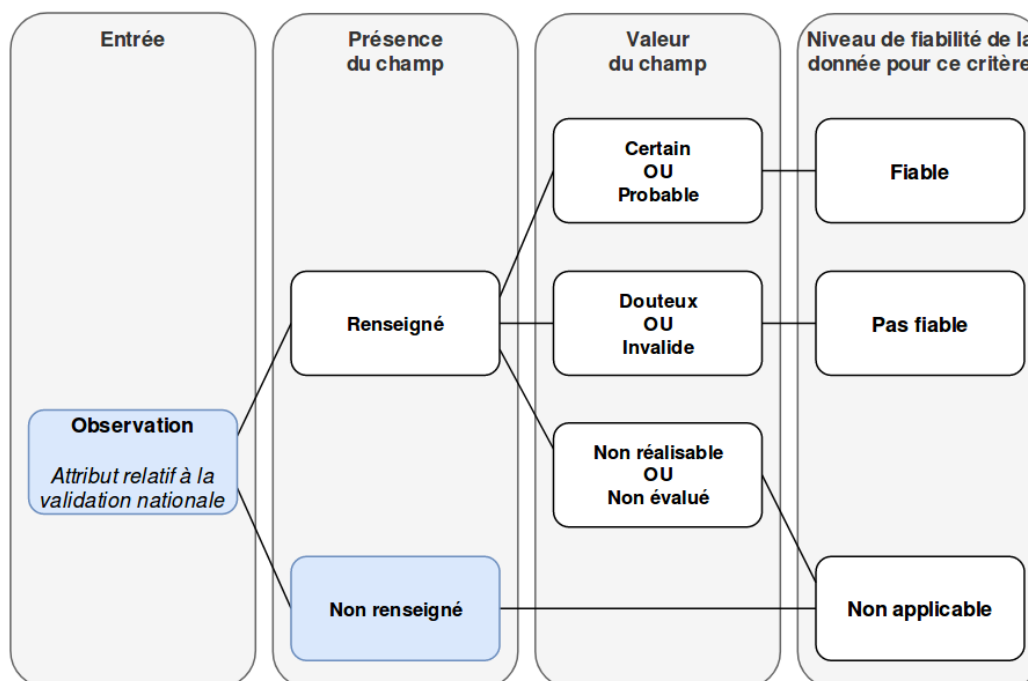


Figure 2 : Tests du critère "validation nationale". En blanc les tests non effectués actuellement.

## 2.2. Critères relatifs à l'écologie du taxon

Ces critères, établis à partir des bases de connaissance disponibles, permettent d'effectuer des contrôles sur la date et la localisation de l'observation, selon l'écologie connue du taxon observé.

### 2.2.1. L'habitat

Chaque espèce possède un ou des habitats de vie particuliers. Cependant, il n'existe actuellement pas de référentiel d'habitats d'espèces et de cartographie régionale précise de ces habitats naturels. Seule l'appartenance au domaine marin et continental est donc contrôlée par le biais de deux couches géographiques de référence (une pour le domaine marin et une pour le domaine continental) et du grand type d'habitat associé à chaque espèce dans le référentiel taxonomique TAXREF (figure 3).



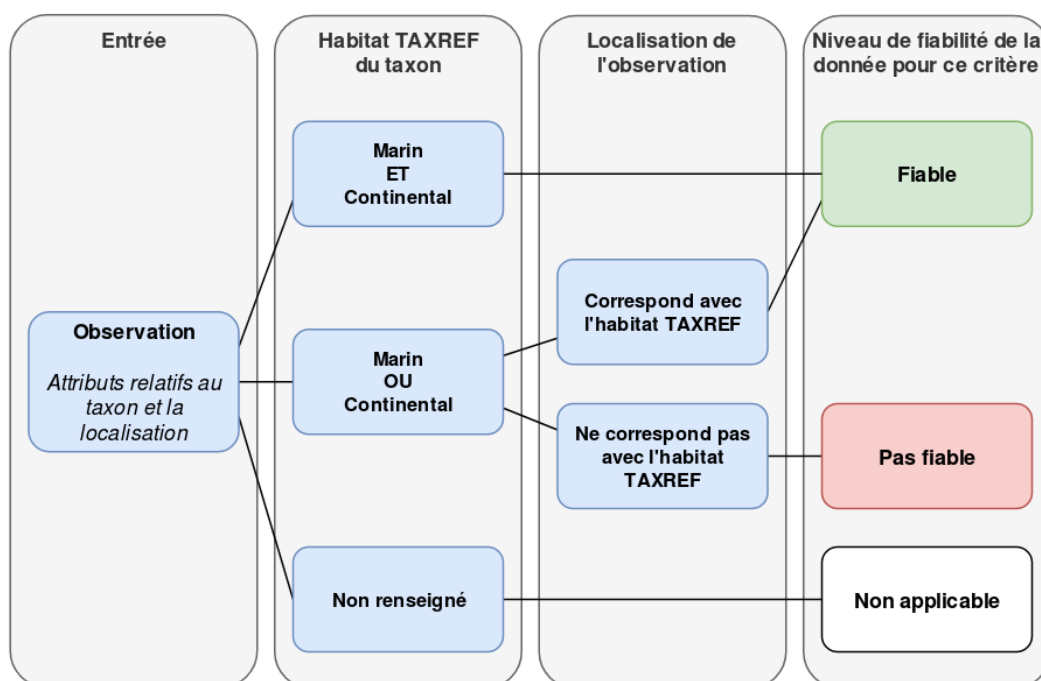


Figure 3 : Tests du critère “habitat”

### 2.2.2. La période d’observation

Une période d’observation correspond ici à une période favorable à l’observation du taxon considéré. Le critère mis en place consiste ainsi à vérifier la cohérence de la date d’observation avec ces périodes (figure 4). Ces périodes d’observation sont issues du référentiel des périodes d’observation de la faune, en cours d’établissement par l’observatoire FAUNA et ses partenaires (voir annexe 1).

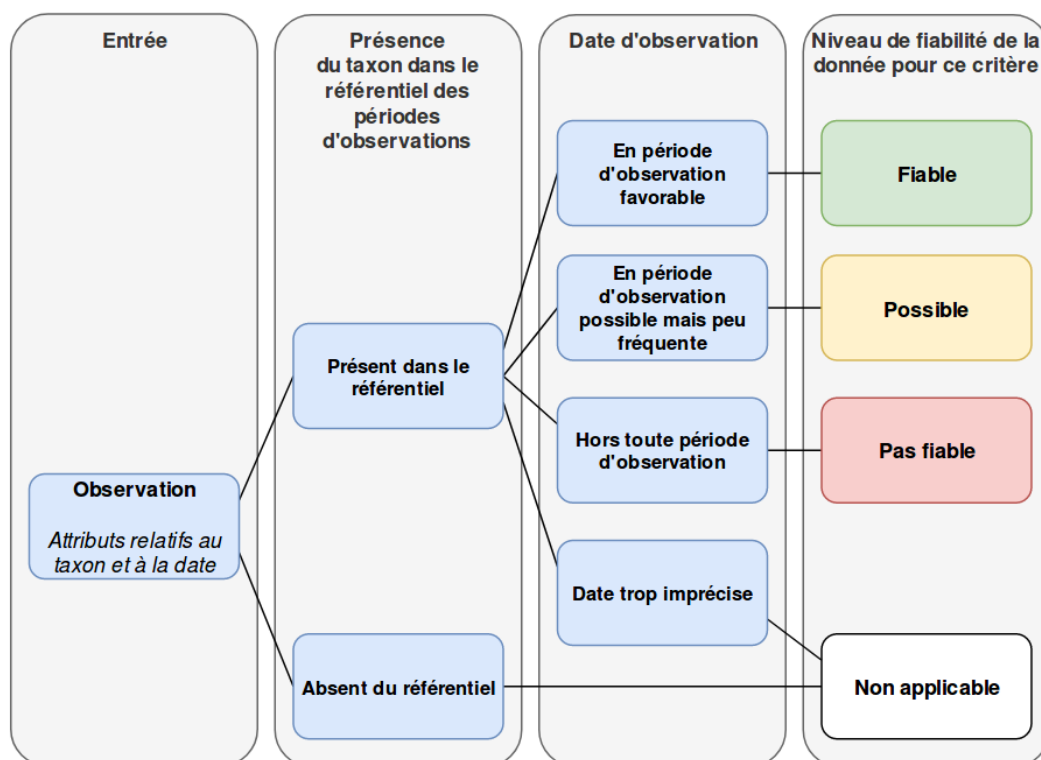


Figure 4 : Tests du critère “période d’observation”

*Remarque* : Ce référentiel est en cours d'établissement pour les groupes taxonomiques étudiés dans le cadre du collectif d'expertises coordonné par l'Observatoire FAUNA. Il recense à la fois des périodes durant lesquelles l'observation du taxon est favorable et des périodes d'observation possible mais peu fréquente. La valeur « non applicable » est attribuée par défaut à l'ensemble des espèces pour lesquelles ce référentiel n'est pas encore établi.

### 2.2.3. La répartition géographique

Le contrôle de la localisation du taxon s'appuie ici sur des bases de connaissances permettant de disposer d'informations de référence sur la distribution du taxon. Les différents tests de ce critère consistent ainsi à vérifier la correspondance entre la localisation du taxon et les distributions de référence (figure 5). Les bases de connaissances utilisées sont les suivantes :

- Les référentiels-espèces de l'Observatoire FAUNA (BARNEIX M. & PERRODIN J., 2021) : Il permettent de vérifier la cohérence de la localisation de l'observation avec les statuts de présence régionale et départementale définis pour l'espèce (les statuts sont explicités dans l'annexe 2).

*Remarque* : Ces statuts sont définis de manière exhaustive pour les principaux groupes taxonomiques (voir annexe 1). Certaines espèces, de groupes taxonomiques non traités à ce stade, peuvent tout de même disposer de ces statuts de présence consolidés. Les espèces faisant partie des groupes taxonomiques évalués de manière exhaustive mais n'ayant pas de statut sont absentes en région (correspond au « Considéré absent » de la figure 5).

- Les atlas régionaux : Ils permettent de vérifier la cohérence de la localisation de l'observation avec les mailles "atlas" du taxon (si disponibles et suffisamment récentes) (voir annexe 1).
- Les données du système qui ont été validées manuellement par des experts naturalistes : Elles permettent le contrôle, par un principe de congruence, de la localisation des observations lorsque les référentiels susmentionnés ne sont pas disponibles.

*Remarque* : Afin de procéder à ces contrôles, le rattachement de la localisation des observations à un département et à une maille 10x10km doit être effectué. Il arrive cependant qu'une observation intersecte plusieurs départements ou plusieurs mailles (du fait d'une localisation imprécise ou floutée géographiquement). Dans ce cas, la note du critère est calculée pour chaque entité intersectée et la valeur finale correspond à la valeur de fiabilité la plus élevée.



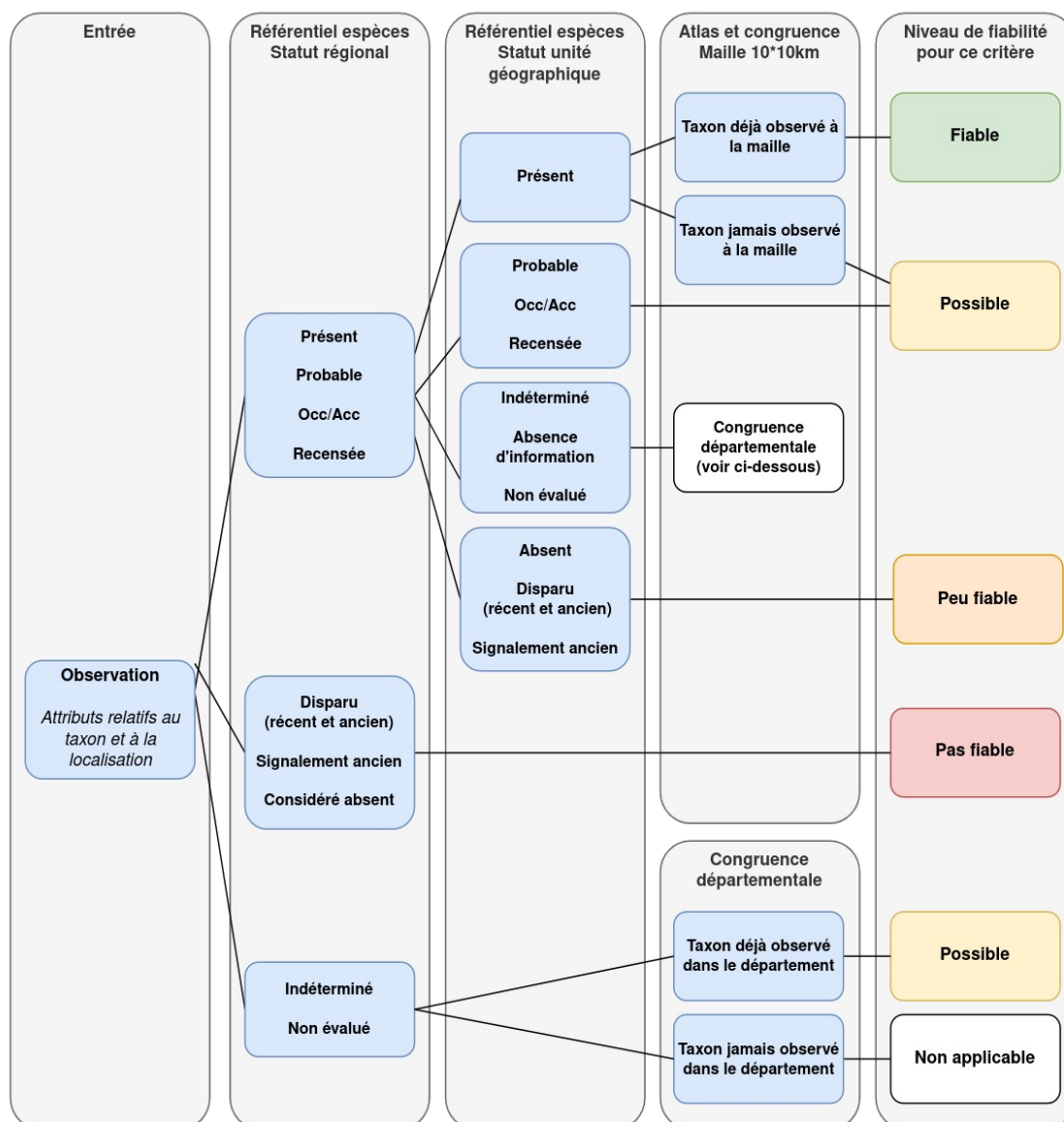


Figure 5 : Tests du critère "répartition géographique"

### 2.3. Critère relatif à l'identification du taxon

L'unique critère relatif à l'identification du taxon est le contrôle de la difficulté de détermination, qui peut être source d'erreur dans la donnée. Ce contrôle nécessite de disposer d'un référentiel des niveaux de difficulté de détermination des espèces présentes en Nouvelle-Aquitaine.

Afin de créer ce référentiel, quatre niveaux de difficulté ont été retenus :

- **Facile** : Espèce pouvant être identifiée aisément à vue, au chant ou via un indice de présence fiable. Le risque de confusion avec une autre espèce est très faible. Pas de difficulté particulière.
- **Intermédiaire** : Espèce pouvant être identifiée à vue, au chant ou via un indice de présence fiable mais nécessitant davantage d'attention/expérience. Le risque de confusion avec une autre espèce est possible mais n'empêche pas l'identification.
- **Difficile** : Espèce nécessitant une observation attentive avec capture. Le risque de confusion avec une autre espèce complique l'identification.
- **Très difficile** : Espèce nécessitant une capture et du matériel adapté (binoculaire, appareil de biométrie, etc) ou une analyse en laboratoire.

**Remarque** : Ce référentiel est en cours d'établissement pour les groupes taxonomiques étudiés dans le cadre du collectif d'expertises coordonné par l'Observatoire (voir annexe 1).

La valeur « non applicable » sera attribuée par défaut à l'ensemble des espèces pour lesquelles le niveau de difficulté n'est pas encore établi.

Chaque espèce dispose **a minima** d'un niveau de détermination global. Mais dans certains cas et pour certains taxons, une même espèce **peut** disposer de notes supplémentaires pour des sexes et/ou des stades de vie particuliers. Cependant, l'utilisation de ces niveaux supplémentaires est conditionnée par la complétion d'attributs facultatifs (sexe et stade) ainsi l'espèce dispose **toujours** d'un niveau global.

Ce critère est ainsi composé d'un unique test qui consiste à contrôler la présence du taxon dans le référentiel et le cas échéant de récupérer la note de difficulté de détermination correspondante (figure 6).

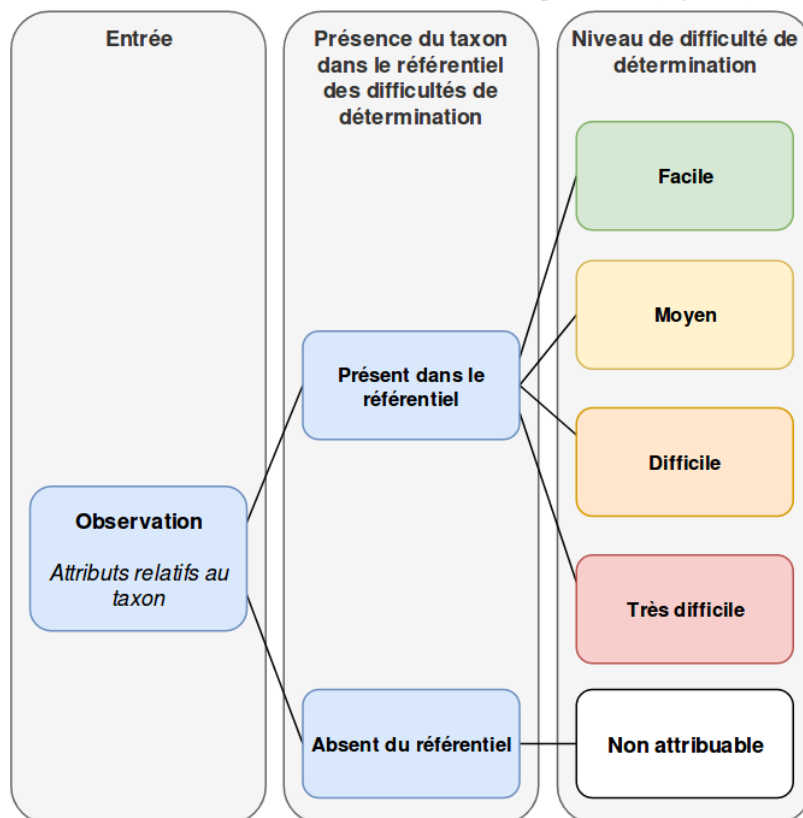


Figure 6 : Tests du critère “difficulté de détermination”

## 2.4. Attribution de la note finale de validité régionale automatique

Chaque observation est décrite par les résultats des critères présentés ci-avant. Elles sont assemblées selon la logique décrite à la figure 7 afin d'aboutir à la note finale du processus de validation.

**Remarque** : Le critère de détermination du taxon, contrairement aux autres critères, permet d'ajuster le niveau de validité, mais ne permet pas d'invalider directement une observation. En effet, l'observation d'une espèce difficile à déterminer n'implique pas forcément l'invalidité de cette dernière.

Au final, chaque observation se voit attribuer, suite au processus automatique, un niveau de validité qui détermine le degré de confiance que l'on peut lui accorder. Ce niveau de validation est accompagné de la date du jour de la validation puisqu'elle reflète l'état des connaissances à un moment donné. Par ailleurs, tous les résultats intermédiaires sont stockés en base de donnée afin de pouvoir être réutilisés.



Figure 7 : Regroupement final pour aboutir à la note de validation automatique

## 2.5. Cas particulier des oiseaux soumis à homologation

Afin de valider les observations d'espèces occasionnelles d'oiseaux, il existe au niveau national et ex-aquitain des comités d'homologation (respectivement, le Comité d'Homologation National et le Comité d'Homologation Aquitain). Toute observation d'oiseau considéré occasionnel sur le territoire doit ainsi être évaluée puis validée ou non par un de ces comités.

Chacune de ces entités fournit une liste d'espèces occasionnelles dont les observations sont soumises à homologation.

Pour la liste nationale (version du 9 mars 2017), seules les catégories CAF-A et -B sont soumises à homologation.  
Pour la liste aquitaine (version du 03 décembre 2019), certaines espèces ne sont soumises à homologation que sous certaines conditions.

Lors du processus de validation automatique des données, les données concernées sont identifiées (Figure 8) afin de pouvoir être évaluées par les comités. Cependant, tant qu'elles ne sont pas évaluées, ces données suivent le processus de validation de l'observatoire comme toutes les autres données.

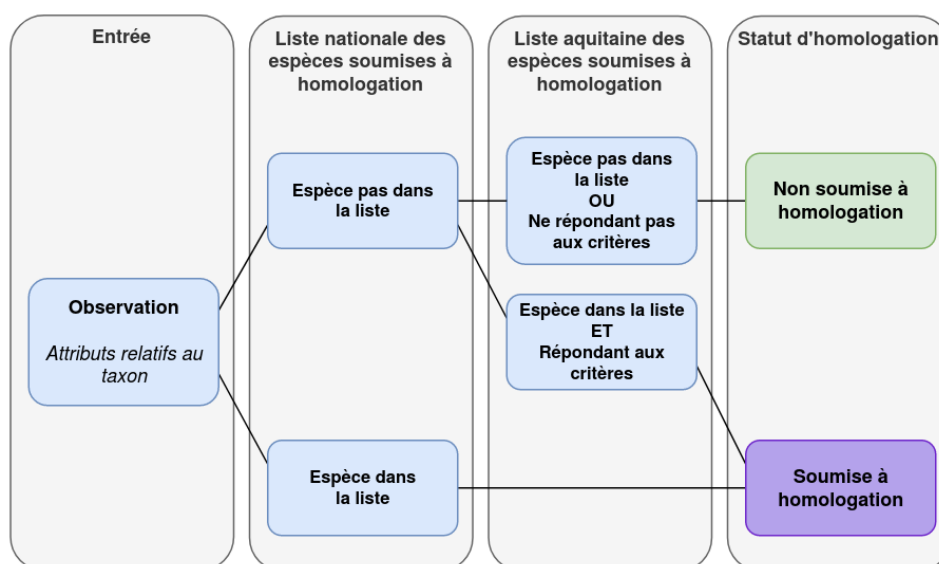


Figure 8 : Prise en compte des espèces soumises à homologation dans la validation automatique

### 3. VALIDATION MANUELLE

Suite au processus de validation automatique, une validation manuelle peut être réalisée par l'Observatoire ou par un collectif d'experts associés composé de membres spécialistes d'un ou plusieurs groupes taxonomiques.

La validation manuelle permet de confirmer ou modifier le niveau de validation attribué par le processus automatique. Au moyen d'une interface interactive sur la plateforme FAUNA, les validateurs peuvent s'appuyer sur :

- les résultats de chacun des critères du processus de validation automatique ;
- les bases de connaissances disponibles (notamment celles utilisées pour la validation automatique).

La validation manuelle permet également de valider ou d'invalider certaines données pour lesquelles le processus automatique n'est pas applicable (notamment en cas de manque de référentiels sur lesquels s'appuyer).

Les validateurs peuvent, s'ils jugent cela nécessaire, contacter le producteur de la donnée afin d'obtenir des informations complémentaires sur l'observation (des détails, des preuves, etc.).

Une validation manuelle peut dans certains cas être considérée comme obsolète et n'est donc plus prise en compte. C'est le cas lorsque :

- une nouvelle validation manuelle vient la supplanter ;
- la donnée a été modifiée après sa validation manuelle ;
- l'évolution des connaissances sur l'espèce entraîne la remise en question du niveau de validation (notamment lors de modifications des statuts de présence du référentiel espèce).

### 4. VALIDATION RÉGIONALE

La validation régionale correspond à la validation finale diffusée avec la donnée.

Cette validation régionale prend comme valeur le niveau de validation automatique lorsqu'aucune validation manuelle n'a été effectuée ou lorsque la validation manuelle est obsolète. Dans ce cas, la validation régionale est de type « automatique ».

Lorsqu'une validation manuelle a été effectuée et qu'elle n'est pas obsolète, elle prime sur la validation automatique : la validation régionale prend alors son niveau de fiabilité. Dans ce cas, la validation régionale est de type « combinée ».

Ce niveau final de validation régit la diffusion des données. Ainsi, les données ayant un niveau de validation régionale « Douteux » ou « Invalide » ne sont pas diffusées dans le SINP régional. Elles sont toutefois échangées avec la plateforme nationale INPN.

## BIBLIOGRAPHIE

- GARGOMINY, O., TERCERIE, S., RÉGNIER, C., RAMAGE, T., DUPONT, P., DASZKIEWICZ, P. & PONCET, L. 2018. TAXREF v12, référentiel taxonomique pour la France : méthodologie, mise en œuvre et diffusion. Muséum national d'Histoire naturelle, Paris. Rapport Patrinat 2018-117. 156 pp.
- BARNEIX M. & PERRODIN J. (coord), 2021. Méthodologie pour l'élaboration et la diffusion du référentiel-espèces sur la faune sauvage de Nouvelle-Aquitaine. Domaine continental et marin. Version 2. Observatoire de la Faune Sauvage de Nouvelle-Aquitaine. Pessac, 22 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., QUAINTEENNE 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.
- 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. 16p.



## ANNEXES

### Annexe 1 : Complétudes des référentiels aquitains utilisés (à la date du rapport)

	Référentiel	Atlas aquitains	Référentiel espèces de l'Observatoire (présence régionale et départementale)	Référentiel des périodes d'observation	Référentiel des difficultés de détermination
Groupe taxonomique	Oiseaux	Complété	Complété	Complété	Complété
	Mammifères hors chiroptères	Complété	Complété	Complété	Complété
	Chiroptères	Complété	Complété	Complété	Complété
	Amphibiens	Complété	Complété	Complété	Complété
	Reptiles	Complété	Complété	Complété	Complété
	Odonates	Complété	Complété	Complété	Complété
	Rhopalocères	Complété	Complété	Complété	Complété
	Zygènes	Complété			
	Écrevisses	Complété	Complété	Complété	Complété
	Orthoptères		Complété	Complété	Complété
	Poissons		Complété	Complété (Gironde)	Complété (Gironde)
	Mollusques		Complété	Complété	Complété
Araignées			En cours	En cours	En cours

## Annexe 2 : Statuts de présence régionaux et départementaux du référentiel espèces de l'Observatoire FAUNA

Statut	Définition
<b>Présence certaine (Pr)</b>	Dernière observation remontant à moins de 10 ans (inclus) et aucune preuve ou présomption de disparition depuis. L'espèce est considérée comme régulière sur le territoire (existence d'une population sédentaire et/ou migratrice régulière).
<b>Présence probable (Pb)</b>	Dernière observation sur la période récente remontant à plus de 10 ans et aucune preuve ou présomption de disparition. Il s'agit en général de taxons réguliers sous-prospectés.
<b>Présence occasionnelle (Occ)</b>	Une ou plusieurs observations sur la période récente mais dont la fréquence est estimée, par expertise, insuffisante pour considérer la présence d'une population régulière sur le territoire.
<b>Présence accidentelle (Acc)</b>	Une ou plusieurs observations sur la période récente mais dont le signalement est considéré, par expertise, comme exceptionnel sur le territoire. La différence avec le statut précédent est liée au caractère étonnant des données (en dehors de l'aire de répartition connue, évènement climatique particulier, introduction anthropique, etc.).
<b>Disparition récente (Dr)</b>	Plusieurs observations sur la période récente et disparition avérée depuis l'obtention de la dernière donnée (recherches poussées sans résultat, publication attestant la disparition).
<b>Disparition ancienne (Da)</b>	Dernière observation antérieure à la période récente et disparition avérée depuis l'obtention de la dernière observation (recherches poussées sans résultat, publication attestant la disparition).
<b>Signalement(s) ancien(s) (Sa)</b>	Dernière observation antérieure à la période récente et aucune preuve ou présomption de disparition. Ce statut s'applique indifféremment sur le caractère occasionnel ou accidentel.
<b>Recensée (R)</b>	Au moins un signalement sur l'ensemble de la période du référentiel. La date de dernière observation et le caractère régulier ou occasionnel/accidentel de la présence ne sont pas expertisés à ce stade. L'objectif de ce statut est de valoriser les premiers apports de connaissances (compilations de données bibliographiques non-numérisées par exemple), notamment pour les taxons moins étudiés.
<b>Indéterminée (Ind)</b>	Observation(s) remise(s) en cause ou en cours de vérification. Dans l'état actuel des connaissances, il n'est pas possible de se prononcer sur la présence ou l'absence du taxon. Ce statut s'applique aux espèces sujettes à une évolution taxonomique récente ou pour lesquelles le risque de confusion avec d'autres taxons est important. Il concerne également les espèces pour lesquelles le manque de connaissance ne permet pas de qualifier l'existence d'une population à l'endroit de l'observation (cas de coquilles pouvant être déplacés sur de longues distances par différents phénomènes par exemple).
<b>Non évaluée (NE)</b>	Espèce dont le statut de présence n'a pas été étudié ou consolidé.
<b>Absence certaine ou probable (Ab)</b>	Aucune observation recensée ou en cours de vérification et absence avérée de l'espèce, consolidée par expertise, au vu de sa chorologie notamment.
<b>Absence d'information (Ai)</b>	Aucune observation recensée ou en cours de vérification dans l'unité géographique considérée.