



Protocole de validation des données du Système d'information sur la nature et les paysages de La Réunion (SINP 974)

-

Volet occurrences de taxons

Version 1.3

Résumé

Une multitude de jeux de données naturalistes est produite chaque année par de nombreux acteurs professionnels ou bénévoles. Ces données sont collectées selon des méthodes et protocoles différents et pour des objectifs variés : contribution aux diagnostics écologiques de zones d'études, appui aux actions de conservation ou de recherche, information des citoyens, publication d'indicateurs... Elles concernent des périmètres taxonomiques et géographiques divers, et les compétences et moyens des observateurs sont **hétérogènes**.

Dans le cadre du SINP (Système d'Information sur la Nature et les Paysages), quelle que soit leur utilisation, il est indispensable de **garantir la fiabilité des données** entrant dans le dispositif et de **diffuser le niveau de confiance** qui leur est associé. Cette exigence est un gage de crédibilité et de robustesse de l'ensemble de la démarche. Il s'agit aussi d'un **retour d'information utile au producteur de données** lui permettant de s'améliorer le cas échéant.

Le présent guide vise à formaliser dans un protocole les **règles de validation** des données d'observation naturaliste circulant dans le SINP, qu'elles concernent les domaines terrestre ou marin (zone économique exclusive). Le protocole est **rendu public** sur le portail internet du SINP 974 rendant ainsi transparentes les règles appliquées aux données fournies par les producteurs.

Il ne concerne que la **validation au niveau de la plate-forme régionale SINP** www.borbonica.re, d'autres validations pouvant avoir lieu en amont (validation par le producteur) ou en aval (validation par la plate-forme nationale SINP).

Le protocole de validation constitue ainsi la **déclinaison régionale du cadre méthodologique national** élaboré par le Muséum national d'histoire naturelle. Il a été **validé par le Conseil scientifique régional du patrimoine naturel** (CSRPN) de La Réunion.

Il conduit à formaliser des règles des différents volets de la validation, appliquées successivement par les administrateurs de Borbonica et les têtes de réseau :

- identification des **doublons** ;
- contrôle de la **conformité et de la cohérence** (validation sur la forme) ;
- **validation scientifique** (validation sur le fond) automatique (421 règles concernant 4873 taxons) et manuelle.

L'application de ces différents contrôles aboutit à l'**attribution d'un niveau de validation**, indice de la fiabilité associé à une donnée. Seules les données les plus fiables (certaines ou probables) sont accessibles au grand public, les professionnels, réputés experts, accédant à l'ensemble des données selon les conditions fixées par la charte régionale du SINP.

Dans la version actuelle, la validation scientifique ne concerne que les groupes taxonomiques suivants, les autres groupes ayant vocation à être traités lors de **prises à jour** ultérieures :

- | | |
|-------------------------------------|---|
| • tortues marines | • reptiles et amphibiens |
| • cétacés | • oiseaux |
| • chiroptères | • poissons et macro-crustacés d'eau douce |
| • mammifères terrestres non volants | • flore vasculaire |

Historique des versions du document

Version	Date	Commentaire
0.6	30/07/17	Version consolidée soumise à consultation des têtes de réseau du SINP 974
0.7	17/08/18	Version intégrant les retours des têtes de réseau du SINP 974, soumis à consultation du CSRPN
1.0	25/09/18	Version validée par le CSRPN (avec recommandations) ayant fait l'objet de retouches mineures. Les recommandations sont à apporter dans une version ultérieure.
1.1	15/04/19	Version intégrant différentes améliorations et compatible avec Taxref version 12
1.2	11/05/20	Version validée par le CSRPN du 19/12/2019 et intégrant d'autres modifications
1.3	21/09/22	Version complétée avec le volet sur la validation des données d'observation de flore vasculaire, validé par le CSRPN le 05/07/2022

Le détail des modifications apportées par chaque version est repris dans le document historique_version_v1.3 joint au protocole.

Rédacteurs

Rémi BOUILLY, Valentin Le Tellier (DEAL de La Réunion, coord.) sur la base des contributions de Pierre Valade (OCEA Consult'), Laetitia Faivre (OCEA Consult') et des têtes de réseau des pôles thématiques du SINP 974 : Nicolas Laurent (SEOR), Mickaël Sanchez (NOI), Gildas Monnier (GCOI), Sarah Caceres (ONCFS), Vanessa Estrade (Globice), Claire Jean (Kélonia CEDTM), Bertrand Mallet et Frédéric Picot du CBNM.

Le guide s'appuie grandement sur les principes nationaux définis dans le cadre du groupe de travail « validation » du SINP résumés dans le document suivant:

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., Quaintenne 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.¹

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.²

Jomier R., Robert S., *Extension pour occurrences de taxons version 1.2.1. Extension validation v1.0*. MNHN, Paris, 21 p.³

1 <https://sinp.naturefrance.fr/documentation-validation/>

2 http://spn.mnhn.fr/spn_rapports/archivage_rapports/2017/SPN%202017%20-%20%20-%20procedure_validation_scientifique_INPN_V1.pdf

3 <http://standards-sinp.mnhn.fr/extension-validation-v1-0-pour-occurrences-de-taxons-v1-2-1/>

Relecteurs

Têtes de réseau des pôles thématiques du SINP 974 : Nicolas Laurent (SEOR), Chloé Bernet (NOI), Gildas Monnier (GCOI), Sarah Caceres (OFB), Vanessa Estrade (Globice), Claire Jean (Kélonia CEDTM), Bertrand Mallet et Frédéric Picot du CBNM.

Auteurs volet « faune d'eau douce » : Pierre Valade (OCEA Consult'), Laetitia Faivre (OCEA Consult')

Équipe projet Borbonica : Jean-Yves Peseux (DEAL), Isabelle Bracco (DEAL), Jean-Cyrille Notter (PNRun).

Validation

Validation de la version 1.0.0 par le Conseil scientifique régional du patrimoine naturel, avec diverses recommandations (séance plénière du 31/08/2018, voir l'avis n°2018-07 sur <http://www.reunion.developpement-durable.gouv.fr/les-avis-du-csrpn-par-annee-r169.html>).

Validation de la version 1.2.0 par le Conseil scientifique régional du patrimoine naturel, avec diverses recommandations (séance plénière du 19/12/2020, voir l'avis n°2019-14 sur https://www.borbonica.re/actualites/20191219_csrpn/2019_14_protocole_de_validation_sinp.pdf).

Validation de l'ajout du volet Flore vasculaire (version 1.3.0) par le Conseil scientifique régional du patrimoine naturel, avec diverses recommandations (séance plénière du 31/03/2022, voir l'avis n°2022-04 sur https://www.borbonica.re/files/20220331_csrpn/2022_04_Avis_CSRPN_SINP_Protocole_Flore.pdf).

Diffusion

Diffusion au grand public sur le portail Borbonica du SINP 974 (https://www.borbonica.re/protocole_validation/).

Citation recommandée

DEAL (coord.), GCOI, Globice, Kélonia-CEDTM, NOI, OCEA Consult', OFB, SEOR, UMR PVBMT, 2020. Protocole de validation des données du Système d'information sur la nature et les paysages de La Réunion – volet occurrences de taxons (version 1.2). Rapport pour le SINP 974, 25 p. + 14 annexes.

1. La validation dans le cadre du SINP.....	8
1.1. Pourquoi un protocole de validation ?.....	8
1.2. Les différents volets de la validation.....	8
1.3. Mises à jour du protocole.....	10
2. Identification des doublons.....	11
2.1. Définition et périmètre.....	11
2.2. Doublons externes.....	11
2.2.1. Méthode mise en œuvre.....	11
2.2.2. Traitement des observations concernées.....	11
2.3. Doublons internes.....	12
2.3.1. Méthode mise en œuvre.....	12
2.3.2. Traitement des observations concernées.....	12
3. Contrôle de la conformité et de la cohérence.....	13
3.1. Définition et périmètre.....	13
3.2. Méthode mise en œuvre.....	13
3.3. Traitement des observations concernées.....	13
4. Validation scientifique.....	14
4.1. Définition et périmètre.....	14
4.2. Méthode mise en œuvre.....	14
4.3. Données de référence et autres données utilisées.....	16
4.3.1. Types de données pouvant être utilisées.....	16
4.3.2. Mises à jour.....	17
4.4. Organisation de la validation.....	18
4.4.1. Qui fait quoi.....	18
4.4.2. Validation d'un nom de taxon différent de celui indiqué par l'observateur.....	18
4.4.3. Validation des données d'observation négative.....	19
4.5. Déclinaison par pôle thématique.....	20
5. Échanges avec les producteurs.....	23
5.1. Échanges lors de la validation initiale de la donnée.....	23
5.2. Échanges au cours de la vie de la donnée.....	25
5.3. Bonnes pratiques pour l'amélioration de la qualité des données.....	25
Annexe 1 : Script SQL d'identification des doublons.....	26
Annexe 2 : Protocole de contrôle de la conformité et de la cohérence.....	27
Annexe 3 : Rapport d'import type.....	28
Annexe 4 : Validation scientifique des données de cétacés.....	33
Annexe 5 : Validation scientifique des données de tortues marines.....	35
Annexe 6 : Validation scientifique des données de chiroptères.....	39
Annexe 7 : Validation scientifique des données de mammifères terrestres non volants.....	43
Annexe 8 : Validation scientifique des données de reptiles et amphibiens.....	48
Annexe 9 : Validation scientifique des données d'oiseaux.....	54
Annexe 10 : Validation scientifique des données de poissons et macro-crustacés d'eau douce.....	68
Annexe 11 : Validation scientifique des données de flore vasculaire.....	75
Annexe 12 : Synthèse des tests réalisés lors de la validation automatique.....	79
Annexe 12 : Couches SIG de référence.....	80
Annexe 13 : Liste des statuts biogéographiques des taxons.....	81
Annexe 14 : Liste des sigles utilisés.....	82

Tableau 1: Les différents volets de la validation.....	9
Tableau 2: Répartition des différents contrôles de conformité et cohérence réalisés par type.....	13
Tableau 3: Les différents niveaux de validité de la donnée.....	15
Tableau 4: Thématiques pouvant générer des données de référence.....	16
Tableau 5: Liste des validateurs pour chaque tête de réseau.....	18
Tableau 6: Nombre de taxons auxquels une règle de validation est appliquée par groupe taxonomique.....	20
Tableau 7: Regroupement par catégorie des tests réalisés lors de la validation automatique :.....	21
Tableau 8: Répartition par groupe taxonomique des différents tests.....	22
Tableau 9: Types de mises à jour pouvant être apportées à une donnée.....	25
Tableau 10: Habitats préférentiels des <i>taxons</i> de cétacés fréquemment observés.....	33
Tableau 11: Logigramme de validation scientifique des données de cétacés.....	34
Tableau 12: Liste des <i>taxons</i> de tortues marines communes dans la région.....	37
Tableau 13: Liste des <i>taxons</i> de tortues marines communes à la côte.....	37
Tableau 14: Liste des <i>taxons</i> de tortues marines reproductrices.....	37
Tableau 15: Liste des habitats préférentiels des tortues marines (Habitats_preferentiels_TM_Run.ods).....	38
Tableau 16: Liste des taxons et types acoustiques de chiroptères présents à La Réunion.....	41
Tableau 17: Gamme d'altitudes connue de présence des gîtes et dortoirs de chiroptères.....	41
Tableau 18: Gammes d'altitudes connues de contact des chiroptères.....	41
Tableau 19: Périodes connues de présence de juvéniles de chiroptères.....	41
Tableau 20: Périodes connues de présence des taxons de chiroptères.....	42
Tableau 21: Liste des <i>taxons</i> de mammifères terrestres non volants naturalisés et des <i>taxons</i> non naturalisés mais déjà observés et facilement identifiables.....	46
Tableau 22: Liste des taxons de mammifères terrestres non volants pouvant être confondus.....	47
Tableau 23: Liste des taxons de reptiles et amphibiens indigènes et naturalisés de La Réunion.....	51
Tableau 24: Gammes d'altitudes connues d'occurrence de reptiles et amphibiens.....	52
Tableau 25: Habitats d'occurrence de reptiles et amphibiens connus.....	53
Tableau 26: Typologie des habitats utilisés pour la validation scientifique des données d'oiseaux.....	58
Tableau 27: Liste des espèces d'oiseaux communes nicheuses à La Réunion ou ne posant pas de problème d'identification et milieux de prédilection associés.....	59
Tableau 28: Liste des espèces migratrices communes d'oiseaux et.....	60
Tableau 29: Correspondance entre les niveaux de validité de Faune-Réunion et Borbonica.....	62
Tableau 30: Liste des taxons de crustacés d'eau douce présents ou ayant été présents à La Réunion.....	72
Tableau 31: Liste des taxons de poissons d'eau douce présents ou ayant été présents à La Réunion (source : Taxref v11.0 et dire d'expert).....	73

Illustration 1: Les différents types de contrôle des métadonnées et données.....	10
Illustration 2: Répartition globale des différents tests en fonction du niveau de validité automatique attribué.	22
Illustration 3: Échanges avec les producteurs lors de l'import initial des données.....	24
Illustration 4: Logigramme de validation scientifique automatique des données de tortues marines.....	36
Illustration 5: Logigramme de validation scientifique des données de chiroptères.....	40
Illustration 6: Logigramme de validation scientifique automatique des données de mammifères terrestres non volants.....	43
Illustration 7: Logigramme de validation scientifique manuelle des données de mammifères terrestres non volants.....	45
Illustration 8: Logigramme de validation scientifique automatique des données de reptiles et amphibiens.....	49
Illustration 9: Logigramme de validation scientifique manuelle des données de reptiles et amphibiens.....	50
Illustration 10: Logigramme de validation scientifique automatique des données d'oiseaux.....	55
Illustration 11: Logigramme de validation scientifique manuelle des données d'oiseaux.....	57
Illustration 12: Résumé des modalités d'intégration de la validation de Faune-Réunion dans Borbonica.....	63
Illustration 13 : Logigramme de validation scientifique automatique des données de poissons et macro-crustacés d'eau douce.....	69
Illustration 14: Logigramme de la validation scientifique automatique et manuelle des données flore du SINP 974.....	76

1. La validation dans le cadre du SINP

1.1. Pourquoi un protocole de validation ?

Une multitude de jeux de données naturalistes est produite chaque année par de nombreux acteurs professionnels ou bénévoles. Ces données sont collectées selon des méthodes et protocoles différents et pour des objectifs variés : contribution aux diagnostics écologiques de zones d'études, appui aux actions de conservation ou de recherche, information des citoyens, publication d'indicateurs... Elles concernent des périmètres taxonomiques et géographiques divers, et les compétences et moyens des observateurs sont **hétérogènes**.

Dans le cadre du SINP (Système d'Information sur la Nature et les Paysages), quelle que soit leur utilisation, il est indispensable de **garantir la fiabilité des données** entrant dans le dispositif et de diffuser le niveau de confiance qui leur est associé. Cette exigence est un gage de crédibilité et de robustesse de l'ensemble de la démarche.

Le présent document vise à **formaliser dans un protocole les règles de validation** des données élémentaires d'échange (**DEE**, données de base échangées dans le cadre du SINP) à La Réunion et des **métadonnées** associées (jeux de données et cadres d'acquisition), qu'elles concernent les domaines terrestre ou marin (zone économique exclusive). Le protocole est **rendu public** sur le portail internet du SINP 974 rendant ainsi transparentes les règles appliquées aux données fournies par les producteurs.

Il ne concerne que la **validation au niveau de la plate-forme régionale SINP Borbonica**⁴, d'autres validations pouvant avoir lieu en amont (validation par le producteur) ou en aval (validation par la plate-forme nationale SINP).

Le protocole régional s'appuie sur les publications nationales citées en introduction, qu'il décline localement. Depuis la publication de ces guides, la notion de « qualification » (ie « aptitude à répondre aux usages ») initialement envisagée a été abandonnée au niveau national, aussi le protocole ne parle que de « validation ».

Le protocole régional a été **validé par le Conseil scientifique régional du patrimoine naturel (CSRPN)** de La Réunion.

1.2. Les différents volets de la validation

La validation des données élémentaires d'échange se constitue de **4 volets** vérifiés à des moments et par des opérateurs différents (voir Tableau 1) selon les règles d'organisation définie dans les documents cadre du SINP :

- au niveau national : protocole national révisé le 02/10/2017 ⁵ ;
- au niveau régional : charte régionale validée le 29/06/2017 et révisée le 18/03/2021 ⁶.

4 www.borbonica.re

5 <http://www.naturefrance.fr/sinp/presentation-du-sinp/protocole-du-sinp>

6 https://www.borbonica.re/files/sinp/charte_sinp974_v1.1.pdf

Type	Volet	Description	Opérateur	Quand
Validation sur la forme	Doublons	Vérification que la même observation (caractérisée par un lieu, un taxon, un observateur et une date) n'a pas déjà été entrée dans le SINP au niveau national ou régional.	Administrateurs de Borbonica (PNRun et DEAL)	Avant l'import dans Borbonica*
	Conformité	Respect des règles fixées dans le cadre de la mise en œuvre des formats standards de données et de métadonnées, autant sur les aspects physiques que conceptuels (renseignement des champs obligatoires, format, utilisation des référentiels et des listes de valeurs/nomenclatures).		
	Cohérence	Respect de la logique combinatoire des informations transmises au sein des données, au sein des métadonnées et entre les données et les métadonnées.		
Validation sur le fond	Validation scientifique	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.	Administrateurs de Borbonica puis têtes de réseau SINP 974 **	Après l'import de Borbonica

Tableau 1: Les différents volets de la validation

* les différents contrôles sont réalisés sur les données standardisées importées dans la version de travail de Borbonica, avant import dans la version de production accessible aux adhérents et au grand public.

** sous la supervision globale du Muséum national d'histoire naturelle (MNHN) au niveau national et du CSRPN au niveau régional, qui sont les responsables scientifiques du SINP.

Ces différents contrôles concernent à la fois les métadonnées et les données comme indiqué dans l'Illustration 1 :

- les **métadonnées** des jeux de données et cadres d'acquisition (=dispositifs de collecte) sont concernées par l'identification des doublons et les contrôles de conformité et de cohérence ;
- les **données** sont concernées par l'identification des doublons, les contrôles de conformité et de cohérence et la validation scientifique.













Doublons	Conformité	Cohérence	Validation scientifique
  	  	  	

Illustration 1: Les différents types de contrôle des métadonnées et données

Les différents volets de la validation sont détaillés dans les parties suivantes.

1.3. Mises à jour du protocole

Le présent protocole sera **mis à jour** dès lors qu'il y aura modification de la procédure de validation, en tout ou partie (par exemple le traitement d'un nouveau groupe taxonomique). Une nouvelle version sera alors produite (Version X.x+1 pour les versions avec modifications mineures et X+1 pour les versions avec modifications majeures). L'historique des modifications apportées figure dans le document joint au protocole intitulé historique_version₁.2.odt.

Les **données de référence** sur lesquelles s'appuie le protocole sont **versionnées** pour assurer la traçabilité, comme les données cartographiques de l'Annexe 12 : Couches SIG de référence.

Enfin, les **données validées** portent la **version du protocole** faisant référence au moment de la mise en œuvre de l'étape de validation ainsi que la date de mise en œuvre de l'étape de validation.

2. Identification des doublons

2.1. Définition et périmètre

Cette étape consiste à vérifier que la donnée ou métadonnée (jeu de données et cadre d'acquisition) n'est **pas déjà présente dans Borbonica ou dans la plate-forme nationale (doublon externe)**. Lorsque la donnée ou métadonnée est déjà présente dans le système avec une date différente de mise à jour, il peut s'agir d'une donnée ou métadonnée venant mettre à jour la précédente ou d'un doublon généré par des flux croisés au sein du SINP.

Il est également vérifié si la donnée n'est pas dupliquée au sein du jeu de données fourni (**doublon interne**).

2.2. Doublons externes

2.2.1. Méthode mise en œuvre

Pour les données, le protocole mis en œuvre consiste tout d'abord à **filtrer en amont** les données à importer dans Borbonica en suivant ces règles :

- **ne pas importer de données comportant déjà un identifiant permanent SINP ne provenant pas de la plate-forme nationale.** En effet, comme ces identifiants ne peuvent être générés qu'au niveau de Borbonica ou de la plate-forme nationale, il est préférable d'importer les données directement depuis la plate-forme nationale afin d'éviter les doublons dans le dispositif ;
- vérifier **si la donnée n'a pas déjà été fournie par un autre acteur** : il s'agit de lister dans les jeux de données existants ceux concernant une structure associée au nouveau jeu de données et vérifier si les données ont déjà été intégrées à ce titre (par exemple données déjà fournies par un bureau d'études pour le compte d'un maître d'ouvrage) ;
- réaliser des **imports incrémentiels de données** pour un même producteur (ex : jdd_producteur_x_2016, jdd_producteur_x_2017) et sauf cas exceptionnel **ne pas réimporter intégralement un jeu de données déjà importé**. Seules les données modifiées depuis un import précédent seront réimportées. Ces données seront fournies séparément par le producteur de données et viendront écraser les données initialement importées sans sauvegarde particulière dans Borbonica. En revanche, les jeux de données brutes, le rapport et le script SQL d'import sont conservés sur la plate-forme interne de l'équipe projet Borbonica. Il en est de même pour les modifications qui seraient apportées par les administrateurs aux données déjà en base (modification du niveau de sensibilité, du rattachement taxonomique dans Taxref, etc.).

Dans un second temps, une fois les données importées dans la base de travail, il est vérifié qu'une **combinaison identique taxon/date** n'existe pas déjà par ailleurs dans les données en base. Cette vérification est précisée en regardant dans un second temps des informations complémentaires (**observateur, localisation, dénombrement, description des individus**). En effet, certains de ces champs peuvent ne pas être strictement égaux pour une même donnée du fait des décalages liés à des re-projections cartographiques d'une part, et de différences orthographiques d'autre part (initiales pour le prénom, nom orthographié avec ou sans accent, etc.). Le champ identifiant_origine peut également être analysé pour confirmer qu'il s'agit de la même donnée ou pas.

Une **fonction SQL** comportant comme paramètre le nom du ou des jeu(x) de données testé(s) est lancée pour lister les doublons probables qu'ils contiennent. Le script SQL de cette fonction figure en Annexe 1 : Script SQL d'identification des doublons.

2.2.2. Traitement des observations concernées

Les règles suivantes de gestion sont proposées :

- si la donnée ou métadonnée **a déjà été fournie au niveau régional**, la donnée la plus à jour et la plus complète est conservée. Cela signifie que la donnée déjà en base est modifiée avec les informations apportées par la nouvelle donnée le cas échéant, et que la nouvelle donnée est écartée lors de l'import du nouveau jeu de données pour ne pas être importée à nouveau. La donnée ancienne garde son statut de validation puisque les champs vérifiés lors de la validation (lieu, date, taxon) ne sont pas modifiés. Les choix appliqués à un jeu de données sont conservés dans le rapport d'import transmis au producteur (cf. 5 Échanges avec les producteurs) ;
- si la donnée ou métadonnée **a déjà été fournie à la plate-forme nationale**, les administrateurs de

Borbonica contactent la plate-forme nationale et le producteur pour clarifier l'organisation des flux de données du producteur concerné, afin de définir la marche à suivre, notamment une éventuelle mise à jour de la donnée à partir des nouvelles informations. L'import n'est pas réalisé tant que ce point n'est pas clarifié.

Ces règles peuvent également être appliquées **sur les observations déjà importées dans Borbonica**.

2.3. Doublons internes

2.3.1. *Méthode mise en œuvre*

Lors de chaque import, il est également vérifié si le jeu de données intégré comprend des doublons internes. Cette vérification est faite **de la même manière que pour les doublons externes**, avec la fonction automatique décrite plus haut.

2.3.2. *Traitement des observations concernées*

En cas de doublon avéré, les données sont écartées lors de l'import.

3. Contrôle de la conformité et de la cohérence

3.1. Définition et périmètre

Le contrôle de la **conformité** consiste à comparer les fichiers de données et métadonnées échangés avec les **standards d'échange** du SINP (données et métadonnées). Ces vérifications concernent les aspects physiques et conceptuels des standards, c'est-à-dire la présence des **champs obligatoires**, le **type**, les **règles** et le **format des fichiers échangés**, mais également la vérification des valeurs faisant appel à un **vocabulaire contrôlé** sous forme de listes de valeurs/nomenclatures ou de référentiels.

Le contrôle de la **cohérence** consiste à vérifier la **cohérence entre les différentes informations transmises** (intra ou inter fichier) pour les données et les métadonnées.

3.2. Méthode mise en œuvre

Pour les données, le protocole mis en œuvre s'appuie dans Borbonica sur la mise en place de **47 contrôles** de différents types (voir Tableau 2) :

- la mise en place de **contraintes de vérification ou d'autres règles** permettant de s'assurer du respect des listes de valeurs pour chaque champ ou de la cohérence entre les informations issues de plusieurs champs ou plusieurs tables. Une contrainte est une règle définie par l'administrateur de la base de donnée s'appliquant à un champ et bloquant de manière systématique toute valeur ne respectant pas cette règle ;
- la mise en place de **règles** permettant de s'assurer par exemple du caractère obligatoire de certains champs ou de leur format (date, entier positif...) ;
- le respect de certaines **bonnes pratiques lors de l'import** des données ;
- la vérification via une **fonction de contrôle SQL** d'un certain nombre de points une fois les données importées dans la base de travail.

Modalité de contrôle	Nombre de contrôles
contrainte de vérification	22
règle d'attribut	14
fonction de test	10
bonnes pratiques lors de l'import	1
Total	47

Tableau 2: Répartition des différents contrôles de conformité et cohérence réalisés par type

La liste détaillée des contrôles réalisés figure en Annexe 2 : Protocole de contrôle de la conformité et de la cohérence.

Pour les **métadonnées**, le protocole sera à affiner dans un second temps lorsque leur extraction depuis l'outil de gestion national sera opérationnelle depuis Borbonica.

3.3. Traitement des observations concernées

Lorsqu'un **champ obligatoire ou obligatoire conditionnel** ne répond pas aux exigences de conformité ou de cohérence, les données ou les métadonnées sont considérées comme non conformes et le producteur en est informé via le rapport d'import (cf. 5 Échanges avec les producteurs). Il doit alors corriger les informations non conformes pour que les données puissent être véhiculées dans le SINP.

Lorsqu'il s'agit d'un **champ facultatif**, les données et les métadonnées peuvent être intégrées à Borbonica, les attributs non conformes pouvant être déportés vers le champ commentaire ou un attribut additionnel destiné à stocker des champs non prévus par le standard national. Les champs incohérents sont mis à blanc ou avec la valeur par défaut requise.

4. Validation scientifique

4.1. Définition et périmètre

Cette étape vise à renseigner sur la **fiabilité**, dans le sens du **degré de confiance** que l'on peut accorder à la donnée. Elle fait intervenir des bases de connaissance ou de l'expertise directe. Aucun regroupement de données n'étant opéré dans Borbonica, la validation porte sur les données unitaires.

Les DEE circulant au sein du SINP portent **trois niveaux de validité**, résultat des validations scientifiques **producteur, régionale et nationale**. Les **différents niveaux de validité coexistent**, sont indépendants les uns des autres et ne se substituent pas entre eux. Ils sont stockés dans des champs différents de la base de données.

Les **différents niveaux de validité** pour chaque type de validité sont repris dans le Tableau 3. Le niveau de validité correspondant à un **état des connaissances à un moment donné**, il doit donc toujours être associé à une **date de validation**.

Le protocole national et la charte régionale du SINP⁷ prévoient que les données sont accessibles à tout adhérent et sont échangeables avec la plate-forme nationale **quel que soit leur niveau de validité**. En revanche, **seules les données validées scientifiquement sont diffusées au grand public** au niveau régional (valeurs « certain – très probable » et « probable » de la nomenclature des niveaux de validation).

Le protocole concerne l'**ensemble des taxons rencontrés à La Réunion (exotiques et indigènes)**., principalement de **rang taxonomique espèce ou sous-espèce**. Dans de rares cas, il peut concerner des rangs de niveau supérieur (genre, voire ordre comme pour les limicoles indéterminés par exemple).

4.2. Méthode mise en œuvre

La validation peut se concentrer sur la seule combinaison taxon/date/lieu (validation de **périmètre minimal**) ou bien concerner également d'autres champs (effectifs, statut biologique... on parle de validation de **périmètre élargi**). Dans un premier temps, la validation régionale du SINP 974 sera de **périmètre minimal**.

Elle est réalisée en deux phases :

- la validation dite **automatique** qui consiste en une validation faisant appel à des résultats d'expertises préalables (des référentiels, des bases de connaissances, etc.). Ce processus est réalisé de manière automatique (informatique) ;
- la validation dite **manuelle** qui consiste en une validation des informations faisant appel à une expertise directe (avis d'expert suite à l'analyse des informations transmises).

Dans les faits, la validation manuelle comme automatique consiste à **tester un certain nombre de règles** et à en déduire un niveau de validité pour la donnée, une règle se matérialisant sous la forme d'une **combinaison de ces éléments** :

1. **taxon** ou groupe de taxons ;
2. **champ(s)** testé(s) ;
3. **condition** à remplir pour ce(s) champ(s) ;
4. **niveau de validité** attribué lorsque la condition est remplie.

7 <https://carte.borbonica.re>

Valeur	Libellé	Processus automatique	Processus manuel
1	Certain - très probable	La donnée présente un haut niveau de vraisemblance (très majoritairement cohérente) selon le protocole automatique appliqué. Le résultat de la procédure correspond à la définition optimale de satisfaction de l'ensemble des critères du protocole 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 valeurs.	La donnée est exacte. Il n'y a pas de doute sur l'exactitude de l'observation ou de la validation. La validation a été réalisée notamment par un expert qui confirme la détermination du taxon par l'observateur et/ou du déterminateur.
2	Probable	La donnée est cohérente et plausible selon le protocole automatique appliqué 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.	La donnée présente un bon niveau de vraisemblance. Il n'y a, a priori, aucune discordance majeure mais il n'y a pas d'éléments complémentaires (notamment la présence d'une photo ou d'une source) permettant d'attribuer un niveau de validité plus élevé.
3	Douteux	La donnée concorde peu selon le protocole automatique appliqué. 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 (invalidé).	La donnée est peu vraisemblable. Il y a des éléments suffisants pour attester de sa validité mais elle est considérée comme douteuse.
4	Invalide	La donnée ne concorde pas selon le protocole automatique appliqué. 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.	La donnée a été infirmée (erreur de détermination, erreur de niveau de fiabilité). Elle est considérée comme invalide notamment au regard de l'aire de répartition connue et abiotiques de la niche écologique du taxon (détermination). Elle est considérée comme invalide.
5	Non réalisable	La donnée a été soumise à l'ensemble du processus de validation mais l'opérateur (humain ou machine) n'a pas pu la valider. Notamment: - état des connaissances du taxon insuffisantes - informations insuffisantes sur l'observation	
6	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 non finalisée). Le temps précis défini par la date de validation.	

Tableau 3: Les différents niveaux de validité de la donnée

4.3. Données de référence et autres données utilisées

4.3.1. Types de données pouvant être utilisées

Les validations, automatique et manuelle, s'appuient sur diverses **données de référence** qui peuvent concerner plusieurs des thématiques du tableau suivant :

Thématique	Commentaire
Congruence	La congruence comprend : - la validation par rapport à des couches de référence (distribution, phénologie, habitat...) validées par expertise ; - la validation par rapport à d'autres données avant intégration de celles-ci dans les couches de référence.
Rareté des données (en termes de fréquence d'observation)	Un niveau de rareté peut être calculé automatiquement en fonction des données présentes dans le système. Il peut être calculé relativement à un groupe taxonomique, éventuellement sectorisé par année, etc.
Difficulté de détermination	La difficulté de détermination tient en premier lieu au taxon, mais également à la zone géographique concernée, au protocole de collecte employé et à la méthode de détermination
Présence d'une preuve	Disponibilité de la preuve en elle-même ou existence d'une preuve
Formation des observateurs	Degré d'expertise lié à l'expérience ou aux formations dispensées aux observateurs dans le cadre d'un programme spécifique
Validation producteur	Présence d'une validation du producteur sur les données transmises. Paramètre à utiliser avec précaution pour garantir l'indépendance des validations.
Observateur et déterminateur distincts	Détermination déléguée ou reprise par une personne dédiée
Validation collective	Concept de double validation ou/ et de validation collective. Plusieurs experts analysent les données pour statuer ou mise en place d'une validation par une communauté d'acteurs

Tableau 4: Thématiques pouvant générer des données de référence

Les **exemples de référentiels** suivants peuvent apporter un certain nombre de réponses mais sont à expertiser et compléter au cas par cas par les têtes de réseau pour chaque groupe taxonomique :

- référentiel taxonomique Taxref (statuts biogéographiques) ;
- liste des taxons ayant fait l'objet de changements taxonomiques complexes et date de ces changements (exemple : scission d'une espèce en trois lors d'une révision publiée en 1998) ;
- liste des communes ayant fait l'objet de remaniements de type fusion ou scission ;
 - référentiels de difficulté de détermination ;
 - référentiels de rareté des données ou des taxons ;
- référentiels des protocoles d'acquisition de données, des méthodes de terrain et du matériel utilisé ;
- Couches de référence de répartition des taxons pouvant être déclinées selon différentes résolutions géographiques (commune, maille, département, etc.) et pouvant être constituées des données d'observations ou enrichies par une approche experte ou statistique d'estimation de la niche et de la distribution potentielle (approche utilisant elle-même des variables issues d'autres référentiels comme le climat et l'occupation du sol par exemple) ;
 - bases de connaissances sur les erreurs et confusions fréquentes ;
 - bases de connaissances sur la phénologie et sur les altitudes limites des taxons ;
- bases de connaissances sur les plantes-hôtes (ou autres types d'interactions) et couche de distribution de ces plantes-hôtes.

Il est également possible de **s'appuyer sur d'autres données du système**, par exemple la distance aux autres observations déjà connues ou le fait qu'un observateur a déjà eu des observations validées pour un taxon donné. Cependant, cette possibilité est à manipuler avec précaution, une première erreur de validation pouvant entraîner d'autres erreurs de manière exponentielle.

4.3.2. Mises à jour

Ces données étant évolutives, il est nécessaire que les référentiels et les bases de connaissance utilisés pour mettre en œuvre la validation soient **mis à jour et versionnés**. La fréquence de ces évolutions est variable selon les groupes taxonomiques concernés et sera adaptée avec les retours d'expérience issus des premières applications du protocole.

Dans le cas des mises à jour de référentiels, il est nécessaire de pouvoir **identifier les évolutions impactant la validation**, c'est-à-dire les cas pour lesquels la validation pourrait être réinitialisée automatiquement. Il pourra s'agir par exemple dans le cas de Taxref de l'analyse des cas des scissions de taxon complexes à la sortie d'une nouvelle version du référentiel. Chaque tête de réseau devra alors assurer une vérification des données des taxons concernés, en lien avec les administrateurs de Borbonica.

4.4. Organisation de la validation

4.4.1. Qui fait quoi

La validation scientifique manuelle et l'élaboration des règles de validation automatique en lien avec la plateforme régionale sont confiées aux **têtes de réseau du SINP 974**. Celles-ci remplissent un certain nombre de **critères rendus publics** au moment de la mise en place du pôle thématique qu'elles animent, notamment concernant les **compétences scientifiques** dont elles disposent ⁸.

Pôle thématique	Tête de réseau	Validateurs
Flore et habitats naturels	Conservatoire botanique national de Mascarin (CBNM)	Frédéric Picot
Bryophytes	Unité mixte de recherche « protection des végétaux et biologie des milieux tropicaux » (UMR PVBMT)	Claudine Ah-Peng
Insectes et arachnides	Unité mixte de recherche « protection des végétaux et biologie des milieux tropicaux » (UMR PVBMT)	Miharisoa Mirana Gauche Bernard Reynaud
Oiseaux	Société d'études ornithologiques de La Réunion (SEOR)	Nicolas Laurent
Reptiles et amphibiens	Nature Océan Indien (NOI)	Chloé Bernet Nicolas Huet
Tortues marines	Kelonia	Claire Jean Stéphane Ciccione Mayeul Dalleau
Chiroptères	Groupe chiroptères Océan Indien (GCOI)	Gildas Monnier Sarah Fourasté
Mammifères terrestres non volants	Office national de la chasse et de la faune sauvage (ONCFS)	Sarah Caceres
Cétacés	Globice	Violaine Dulau Vanessa Estrade
Poissons et macro-crustacés d'eau douce	<i>Pôle thématique non encore mis en place. Dans l'attente, seule une validation automatique est effectuée, sur la base du protocole de validation rédigé en 2018 par Ocea (Pierre Valade, Laëtitia Faivre)</i>	

Tableau 5: Liste des validateurs pour chaque tête de réseau

La procédure se déroule en deux temps :

- une **première validation automatique** des données déjà validées sur la forme, en s'appuyant sur des données de référence établies par les têtes de réseau. Cette première validation est assurée par les **administrateurs de Borbonica** (DEAL et PNRun) après import de données dans la base. Elle conduit à attribuer un niveau de validité automatique à tout ou partie des données ;
- une **seconde validation manuelle par les têtes de réseau** qui peuvent intervenir sur les données, quel que soit le niveau de validité attribué automatiquement à la donnée, y compris si la donnée n'a pas suivi la phase automatique. Cette validation n'est pas systématique, en particulier si la validation automatique a permis d'attribuer un niveau de fiabilité probable. Le **libellé des tests de validation automatique** appliqués aux données est associé à la donnée afin de faciliter le travail des validateurs.

4.4.2. Validation d'un nom de taxon différent de celui indiqué par l'observateur

En accord avec le producteur, la validation peut aboutir à **retenir un nouveau nom** pour le taxon observé. Ce nom est stocké dans un champ spécifique n'écrasant pas le nom cité à l'origine.

Par ailleurs, dans certains cas, il peut être utile de **valider les données au niveau du parent taxonomique**, par exemple lorsqu'il s'agit d'observations d'une espèce qui peut être facilement confondue avec une autre espèce du même genre (ex : *Rattus rattus* / *Rattus norvegicus* ; Faucon éléonore / concolore...). Cela peut

⁸ https://www.borbonica.re/tetes_de_reseau/

contribuer à **accroître la diffusion de données** dont le statut de validation est douteux à l'origine. Mais cela entraîne également une **perte de précision taxonomique** qui peut être préjudiciable pour la conservation si des raccourcis trop rapides sont faits par les utilisateurs de données : par exemple une donnée de *Phelsuma* sp. n'a pas du tout la même incidence qu'une donnée de *Phelsuma inexpectata*, même non validée.

Afin d'avoir une approche équilibrée et homogène entre les pôles, la validation d'observations au niveau du parent taxonomique est possible lorsqu'elle respecte le cadre suivant :

- **pour les taxons indigènes** : on reste au niveau du rang taxonomique initial (celui du nom cité par l'observateur), quitte à garder la donnée dans un statut de validation bas ;
- **pour les taxons exotiques**, on peut valider au niveau du parent taxonomique sous ces conditions :
 - limiter les cas où cela permet de passer d'un niveau de validité non visible grand public (douteux/invalidé/non réalisable/non évalué) à un niveau visible grand public (probable/certain) ;
 - limiter la perte de précision taxonomique : on peut remonter jusqu'au genre, mais pas au-delà (ex : pas au niveau de la famille ni de l'ordre) ;
 - dans tous les cas, conserver le nom cité initialement dans un champ dédié.

Qu'il s'agisse de retenir un nouveau nom de taxon ou de valider au niveau du parent taxonomique, les informations liées au taxon (ex : identifiant Taxref, statuts de protection) seront **alors mises à jour manuellement par les administrateurs** de Borbonica.

4.4.3. Validation des données d'observation négative

Les données d'observation négative sont **exclues de la phase de validation automatique**.

Lors de la phase de **validation manuelle**, une première étape consiste déjà à s'assurer que le producteur **ne s'est pas trompé** :

- vérifier que la valeur 0 du dénombrement était intentionnelle car certains logiciels de SIG (ex : Mapinfo) peuvent remplacer arbitrairement les valeurs NULL par des 0 ;
- comparer la valeur du dénombrement indiqué avec celle du champ de statut d'observation (statut_obs) qui peut prendre les valeurs « Présent » ou « Non observé ».

Si le producteur voulait effectivement signaler une observation d'absence, la validation peut se baser sur les questions suivantes : quel **protocole** a été suivi ? La **pression de prospection** a-t-elle été suffisante pour certifier l'absence du taxon le jour donné à l'endroit donné ?

4.5. Déclinaison par pôle thématique

Les règles générales décrites dans les parties précédentes ont été appliquées au sein des pôles thématiques du SINP 974. Les méthodes de validation scientifique automatique et manuelle en découlant sont **décrites pour chaque pôle de manière détaillée dans les annexes suivantes** :

- Annexe 4 : Validation scientifique des données de cétacés
- Annexe 5 : Validation scientifique des données de tortues marines
- Annexe 6 : Validation scientifique des données de chiroptères
- Annexe 7 : Validation scientifique des données de mammifères terrestres non volants
- Annexe 8 : Validation scientifique des données de reptiles et amphibiens
- Annexe 9 : Validation scientifique des données d'oiseaux
- Annexe 10 : Validation scientifique des données de poissons et macro-crustacés d'eau douce

A l'échelle de l'ensemble des pôles, **421 règles** concernant **4786 taxons** ont été identifiées (voir leur description détaillée dans l'Annexe 12: Synthèse des tests réalisés lors de la validation automatique et le nombre par groupe taxonomique dans le Tableau 6). Elles permettent de tester les observations par rapport à des **données de référence** (liste des taxons communs à La Réunion, habitat préférentiel, aire de répartition, altitude, profondeur, distance à la côte, phénologie...) mais aussi à l'existence de **preuves**. Ces règles sont **traduites en langage SQL** pour pouvoir être mises en œuvre dans Borbonica par les administrateurs de données. Leur regroupement par type figure au Tableau 7. **En cas de pluralité de règles pour un même taxon**, le statut final qualifiant la donnée est celui le plus pénalisant.

group2_inpn	nb_taxons
Amphibiens	217
Crustacés	72
Mammifères	680
Oiseaux	2438
Poissons	844
Reptiles	622
Total	4873

Tableau 6: Nombre de taxons auxquels une règle de validation est appliquée par groupe taxonomique

libelle	cetaces	chiropteres	mammiferes_terrestres	oiseaux	poissons_crustaces_eau_douce	reptiles_amphibiens	tortues_marines	Total
Observation acoustique avec preuve d'un type acoustique appartenant potentiellement à une espèce non décrite		2						2
Observation acoustique sans preuve d'un type acoustique appartenant potentiellement à une espèce non décrite		2						2
Observation avec preuve d'une espèce à une période où elle était présente sur l'île		3						3
Observation avec preuve d'une espèce commune en dehors de la période habituelle	1							1
Observation avec preuve d'une espèce commune en dehors de son habitat préférentiel	3							3
Observation avec preuve d'une espèce connue à La Réunion dont l'identification est complexe dans un secteur de présence avérée ou probable					47			47
Observation avec preuve d'une espèce connue à La Réunion en dehors d'un secteur de présence avérée ou probable					57			57
Observation avec preuve d'une espèce en dehors de la période connue de présence sur l'île		3						3
Observation avec preuve d'une espèce naturalisée ou d'une espèce non naturalisée déjà observée, facilement identifiable et ne pouvant être confondue avec une autre espèce			17					17
Observation avec preuve d'une espèce naturalisée pouvant être confondue avec une autre espèce			2					2
Observation avec preuve d'une espèce non commune	67						1	68
Observation avec preuve d'une espèce non connue à La Réunion et non décrite dans Taxref								
Observation avec preuve d'une espèce non connue à La Réunion mais décrite dans Taxref		202			842			1 044
Observation avec preuve d'une espèce non naturalisée jamais observée, ou déjà observée mais difficilement identifiable			335					335
Observation avec preuve dans les gammes d'altitudes connues		3						3
Observation avec preuve de juvénile en dehors des périodes connues de présence de juvénile		3						3
Observation avec preuve de juvénile pendant la période connue de présence de juvénile		3						3
Observation avec preuve en dehors des gammes d'altitudes connues		3						3
Observation d'un taxon éteint ou disparu de La Réunion						4		4
Observation d'un taxon indigène ou naturalisé dans sa gamme altitudinale d'occurrence connue et dans un habitat d'occurrence connu						2		2
Observation d'un taxon indigène ou naturalisé dans sa gamme altitudinale d'occurrence connue mais en dehors d'un habitat d'occurrence connu						2		2
Observation d'un taxon indigène ou naturalisé dans son aire de répartition connue						19		19
Observation d'un taxon indigène ou naturalisé en dehors de sa gamme altitudinale d'occurrence connue						3		3
Observation d'un taxon indigène ou naturalisé en dehors de son aire de répartition connue et de sa gamme altitudinale d'occurrence connue						21		21
Observation d'un taxon indigène ou naturalisé en dehors de son aire de répartition connue, dans sa gamme altitudinale d'occurrence connue mais en dehors d'un habitat d'occurrence connu						21		21
Observation d'un taxon indigène ou naturalisé, en dehors de son aire de répartition connue mais dans sa gamme altitudinale d'occurrence connue et dans un habitat d'occurrence connu						21		21
Observation d'un taxon non naturalisé						799		799
Observation d'une espèce commune dans la période habituelle	1							1
Observation d'une espèce commune dans son habitat préférentiel	3						2	5
Observation d'une espèce commune en dehors de son habitat préférentiel							2	2
Observation d'une espèce connue à La Réunion dont l'identification n'est pas complexe dans un secteur de présence avérée ou probable					10			10
Observation d'une espèce connue à La Réunion dont l'identification n'est pas complexe s'agissant d'un individu adulte, dans un secteur de présence avérée ou probable					23			23
Observation d'une espèce migratrice commune dans une période d'observation habituelle				14				14
Observation d'une espèce migratrice commune en dehors d'une période d'observation habituelle				14				14
Observation d'une espèce ne posant pas de problème d'identification dans un habitat préférentiel				9				9
Observation d'une espèce ne posant pas de problème d'identification en dehors d'un habitat préférentiel				9				9
Observation d'une espèce nicheuse commune dans un habitat préférentiel				19				19
Observation d'une espèce nicheuse commune en dehors d'un habitat préférentiel				19				19
Observation d'une espèce non commune				2 337			3	2 340
Observation d'une espèce non connue à La Réunion et non décrite dans Taxref								
Observation de gîte avec preuve dans les gammes d'altitudes connues		3						3
Observation de gîte avec preuve en dehors des gammes d'altitudes connues		3						3
Observation de gîte sans preuve dans les gammes d'altitudes connues		3						3
Observation de gîte sans preuve en dehors des gammes d'altitudes connues		3						3
Observation non acoustique d'un type acoustique appartenant potentiellement à une espèce non décrite		2						2
Observation sans photo d'une espèce non commune							1	1
Observation sans photo d'une espèce rare ou jamais observée dans la zone	67							67
Observation sans preuve d'une espèce à une période où elle était présente sur l'île		3						3
Observation sans preuve d'une espèce commune en dehors de la période habituelle	1							1
Observation sans preuve d'une espèce commune en dehors de son habitat préférentiel	3							3
Observation sans preuve d'une espèce connue à La Réunion dont l'identification est complexe dans un secteur de présence avérée ou probable					47			47
Observation sans preuve d'une espèce connue à La Réunion en dehors d'un secteur de présence avérée ou probable					57			57
Observation sans preuve d'une espèce en dehors de la période connue de présence sur l'île		3						3
Observation sans preuve d'une espèce naturalisée ou d'une espèce non naturalisée déjà observée, facilement identifiable et ne pouvant être confondue avec une autre espèce			17					17
Observation sans preuve d'une espèce naturalisée pouvant être confondue avec une autre espèce			2					2
Observation sans preuve d'une espèce non connue à La Réunion et non décrite dans Taxref								
Observation sans preuve d'une espèce non connue à La Réunion mais décrite dans Taxref		202			842			1 044
Observation sans preuve d'une espèce non naturalisée jamais observée, ou déjà observée mais difficilement identifiable			335					335
Observation sans preuve dans les gammes d'altitudes connues		3						3
Observation sans preuve de juvénile en dehors des périodes connues de présence de juvénile		3						3
Observation sans preuve de juvénile pendant la période connue de présence de juvénile		3						3
Observation sans preuve en dehors des gammes d'altitudes connues		3						3
Observation sur terre d'une espèce commune reproductrice hors habitat de ponte répertorié							1	1
Observation sur terre d'une espèce commune reproductrice sur un habitat de ponte répertorié							1	1
Observation sur terre d'une espèce commune non reproductrice							4	4
Total	146	458	708	2 421	1 925	892	15	6 565

Tableau 7: Regroupement par catégorie des tests réalisés lors de la validation automatique :
nombre de taxons concernés pour chaque test
(un même taxon pouvant être concerné par plusieurs textes, la dernière ligne de total comporte des doublons)

L'application de ces règles lors de la validation automatique permet d'attribuer un niveau de validité au sein de ces valeurs : probable, douteux, invalide, non réalisable, non évalué (voir Tableau 8 et Illustration 2). La validation manuelle permet de **préciser ce statut** éventuellement, notamment lorsqu'il y a des preuves ou que le niveau d'expertise de l'observateur est connu. Dans certains cas il est ainsi possible de qualifier les données de « certaines ».

Niveau de validité	Cétacés	Chiro.	Mamm. terrestres	Oiseaux	Poissons crustacés eau douce	Reptiles amphib.	Tortues marines	Total
Probable	4	18	2	42	31	18	2	117
Douteux	5	18	3		73	60	1	160
Invalide		9			42			51
Non réalisable	5	6	1		31			43
Non évalué		2		43			5	50
Total	14	53	6	85	177	78	8	421

Tableau 8: Répartition par groupe taxonomique des différents tests en fonction du niveau de validité automatique attribué

Les différents niveaux de validation attribués par la validation automatique

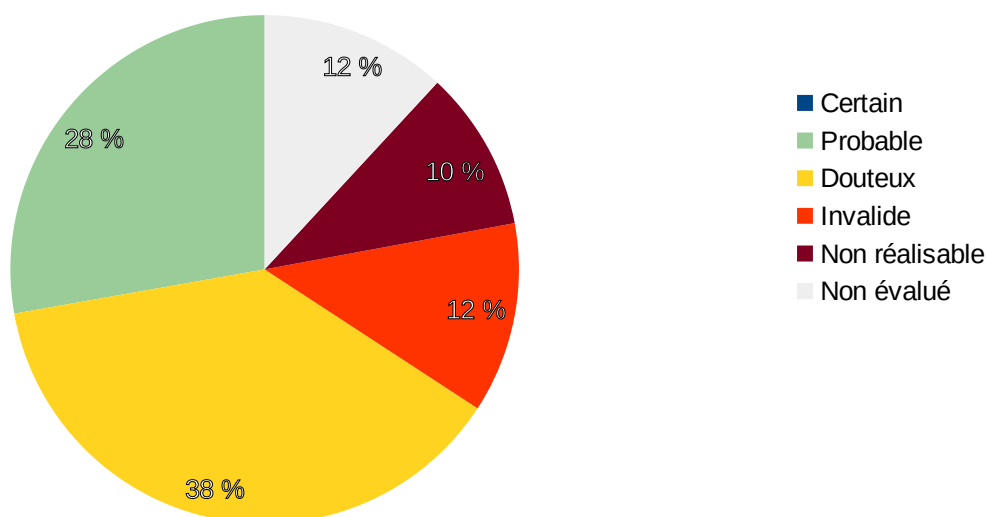


Illustration 2: Répartition globale des différents tests en fonction du niveau de validité automatique attribué

5. Échanges avec les producteurs

Les échanges avec les producteurs interviennent à différents stades de la vie d'une donnée. Ils sont essentiels pour permettre de **lever certains doutes** et assurer un juste **retour d'information** aux personnes et structures acceptant de partager leurs données dans le cadre du SINP.

5.1. Échanges lors de la validation initiale de la donnée

L'étape de validation initiale donne lieu à la transmission au producteur d'un **rapport d'import** produit par les administrateurs de Borbonica (cf. Annexe 3 : Rapport d'import type). Ce document porte sur la validation sur la forme, ainsi que la validation scientifique automatique.

Après accord du producteur sur les éventuels choix opérés pour importer son jeu de données, ce dernier est importé en base de production. Le producteur en est informé, ainsi que les **têtes de réseau qui peuvent alors valider manuellement** les données en lien avec le producteur.

En effet, les têtes de réseau **disposent du contact de l'observateur** lorsque celui est renseigné, a minima celui du producteur de données ayant transmis le jeu de données concerné. Ces informations peuvent être utilisées pour obtenir des éléments complémentaires nécessaires à la validation, comme des preuves. Ce contact est réalisé en **rappelant à l'observateur/producteur le contexte** dans le cadre duquel s'effectue la validation et en renvoyant vers le portail du SINP 974 pour plus d'informations sur la démarche⁹.

Chaque producteur se voit communiquer des codes d'accès à Borbonica pour y visualiser ses propres données et peut donc **connaître à tout moment le statut de validité attribué à chacune de ses données**.

En complément du rapport d'import, le producteur peut demander à tout moment un **rapport de validation scientifique** qui comporte le résultat de la validation manuelle et se présente sous la forme de statistiques sur les statuts de validité des données fournies par le producteur.

L'illustration 3 en page suivante schématise les **échanges avec le producteur** ayant lieu aux différents stades de la validation initiale d'une donnée. Ces modalités seront adaptées en fonction des premiers retours d'expérience.

9 www.naturefrance-fr/la-reunion

Logigramme : import de données des producteurs

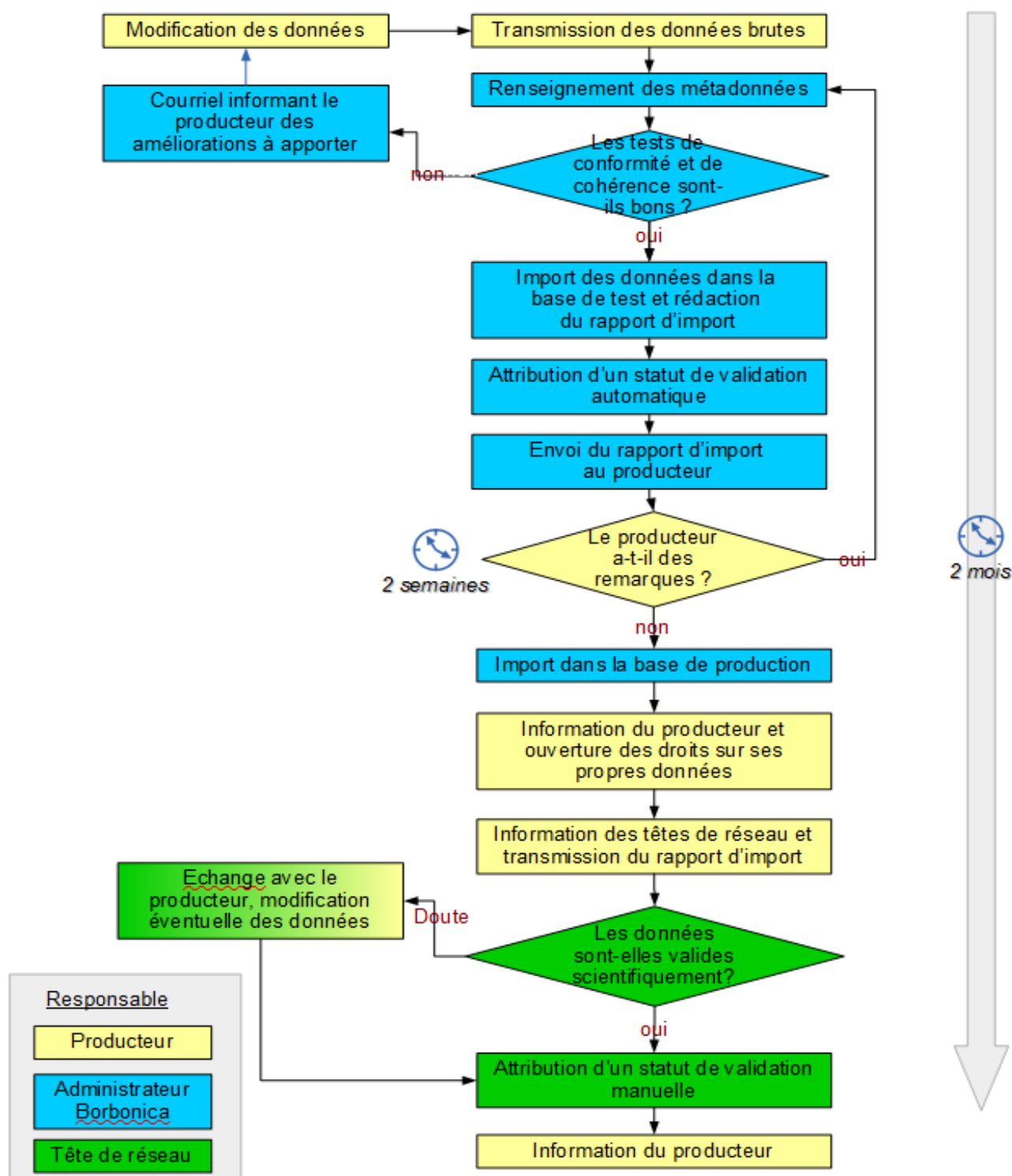


Illustration 3: Échanges avec les producteurs lors de l'import initial des données

5.2. Échanges au cours de la vie de la donnée

Plusieurs raisons peuvent conduire à modifier une donnée, que ce soit à l'initiative du producteur, des administrateurs de Borbonica ou des têtes de réseau.

Type de mise à jour	Exemple	Condition
Informations non présentes dans le jeu de données d'origine	Champs renseignés lors de l'import dans Borbonica : <ul style="list-style-type: none"> niveau de sensibilité niveau de validation scientifique rattachement à une entité géographique en cas de changement de référentiel (mailles 2 km et 10 km, commune, masse d'eau, espace naturel, ZNIEFF) rattachement à Taxref (cd_nom ou cd_ref) en cas de nouvelle version de Taxref ou de révision du rattachement ajout d'un champ au standard de données Borbonica calcul de l'altitude à partir d'un modèle numérique de terrain (précision 10 m de côté) 	Information du producteur par courriel collectif ou personnel selon les cas
Informations présentes dans le jeu de données d'origine	Modifications par exemple en cas d'erreur manifeste ou d'évolution des connaissances : <ul style="list-style-type: none"> nom du taxon etc. 	Accord du producteur Information du producteur sur les éventuelles conséquences sur le statut de validation

Tableau 9: Types de mises à jour pouvant être apportées à une donnée

Afin de garantir une traçabilité des opérations réalisées, les **modifications apportées aux données déjà importées sont historisées** dans Borbonica sous la forme :

- d'un tableau listant les modifications apportées, la date de modification et la liste des jeux de données impactés ;
- des scripts SQL de mise à jour.

Ces modifications seront qualifiées de **mineures ou majeures** selon leur étendue et nécessiteront selon les cas la mise à jour de la date de validation et/ou du niveau de validité.

5.3. Bonnes pratiques pour l'amélioration de la qualité des données

Un certain nombre de bonnes pratiques peuvent être adoptées par les producteurs lors de la collecte ou de la numérisation des données afin de les **aider dans la détermination** des taxons observés ou de **faciliter la validation scientifique**. En voici quelques-unes listées de manière non exhaustive :

- vérifier ses données** avant de les transmettre aux têtes de réseau. Si un observateur n'est pas sûr de la détermination taxonomique, il est recommandé d'indiquer « *sp.* » ou pour la flore « *confer* » (taxon nouveau) ou « *affinis* » (taxon proche de tel autre taxon).
- décrire précisément le lieu de l'observation ou l'habitat, ainsi que les individus observés ;
- transmettre les preuves associées aux observations quand nécessaire ;
- utiliser des guides d'identification (ex : celui du GEIR à cette adresse <http://www.especiesinvasives.re/especies-invasives/>) ;
- utiliser les coordonnées carrées et le système de projection UTM 40S ;
- Suivre les standards de données du SINP (voir le kit de saisie diffusé sur le portail Borbonica du SINP : https://www.borbonica.re/format_standard/) ;
- régler la date des appareils photo ;
- préciser les métadonnées associées aux outils (nom de l'outil, paramètres).

Annexe 1 : Script SQL d'identification des doublons

Voir script SQL d'identification des doublons en PJ (fonction_doublons_v2.0.sql).

Annexe 2 : Protocole de contrôle de la conformité et de la cohérence

Voir ces éléments en PJ :

- regles_coherence_conformite_sinp_v1.2.ods
- fonction_controle_coherence_conformite_v1.2.sql

Annexe 3 : Rapport d'import type

Jeu de données	
Auteur rapport	
Version rapport	
Relecture	
Emplacement	

1. Contexte

1.1 Rôle du producteur de données dans le SINP 974

Description du statut du producteur (adhérent à la charte régional, versements réguliers de données au SINP...).

Description du ou des jeux de données concernés par le rapport.

Le rapport dresse un bilan des différentes étapes réalisées au cours de son import dans Borbonica et précise en particulier les interprétations qui ont été faites pour convertir les lot de données dans le format de Borbonica.

1.2 Source des données utilisées

Description des différents fichiers transmis, nom du contact chez le producteur, date de transmission.

1.3 Les différentes étapes de l'import de données

L'import de données repose principalement sur la mise au format de Borbonica des données transmises dans le format natif du producteur. Pour mémoire, Borbonica respecte le standard de données occurrence de taxon mis en place dans le cadre de l'animation nationale du SINP (version 1.2.1, consultable sur un site internet dédié ou la structure et les listes de valeur spécifique sont diffusées¹⁰).

L'import repose sur différentes étapes, réalisées conformément aux règles préconisées par la groupe de travail national sur la validation des données du SINP ¹¹ :

1. contrôle des doublons
2. contrôle de la conformité
3. contrôle de la cohérence
4. validation automatique des données
5. calcul de la sensibilité des données

Il s'appuie sur les scripts SQL élaborés par l'équipe projet Borbonica (cf. annexe 2).

Les documents de référence utilisés pour la réalisation des contrôles et du calcul de la sensibilité sont les suivants :

- protocole de validation : version X.x.x ¹²
- référentiel de sensibilité : version X.x.x¹³.

10 http://standards-sinp.mnhn.fr/occurrences_de_taxons_v1-2-1/

11 Robert *et al.*, 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. téléchargeable ici : <https://inpn.mnhn.fr/docs-web/docs/download/247411>

12 https://www.borbonica.re/protocole_validation/

13 https://www.borbonica.re/donnees_sensibles/

2. Contrôle des doublons

Ce test vise à vérifier qu'aucun doublon n'est présent dans le jeu de données transmis. Il s'appuie sur les règles du protocole régional de validation.

3. Contrôle de la conformité

Le contrôle de conformité vise à vérifier le **respect des règles fixées dans le cadre de la mise en œuvre des formats standards** de données et de métadonnées autant sur les aspects physiques que conceptuels (renseignement des champs obligatoires, format, utilisation des référentiels et des listes de valeurs/nomenclatures).

Les adhérents au SINP sont invités à respecter le format standard diffusé sous la forme d'un kit de saisie (à mettre à jour) sur le portail du SINP¹⁴. La **comparaison entre le format fourni et le format cible**, réalisé dans cette partie, permet de statuer sur la conformité des données.

3.1 Métadonnées

Description des fichiers de métadonnées éventuellement fournis et de l'organisation pour le renseignement de l'application nationale de gestion de métadonnées du SINP (<https://inpn.mnhn.fr/mtd/>).

3.2 Structure de la base d'origine

La signification des champs des fichiers source est reprise en annexe 1. L'annexe décrit également la manière dont ils sont interprétés pour alimenter Borbonica.

3.3 Champs du standard non présents dans la base d'origine

Un certain nombre de champs non présents dans le jeu de données d'origine figurent dans Borbonica. Leur remplissage est explicité ci-dessous pour les principaux d'entre eux.

Liste des champs et justification des valeurs renseignées.

4. Contrôle de la cohérence des données

Ces tests visent à vérifier le respect de la logique combinatoire des informations transmises au sein des données, au sein des métadonnées et entre les données et les métadonnées (exemple : la date de début de l'observation est inférieure ou égale à la date de fin de l'observation, le cd_ref et le nom valide correspondent bien...).

5. Validation automatique des données

5.1 Validation par le producteur

Description de la validation éventuelle par le producteur et du niveau de validité attribué.

5.2 Validation régionale

Le protocole régional de validation des données du SINP identifie une liste de critères permettant d'attribuer automatiquement un niveau de validité à certaines données. Ces critères sont basés sur les connaissances existantes des taxons (aires de répartition, habitats préférentiels, phénologie...).

Indication du résultat de la validation automatique :

Niveau de validité	Libellé	Nombre de données	%
1	Certain	0	0%
2	Probable	589	31%
3	Douteux	24	1%
4	Invalide	0	0%
5	Non applicable	0	0%
6	Non évalué	1 259	67%
Total		1 872	100,0%

14 https://www.borbonica.re/files/format_standard/202204_kit_saisie_v7.zip

Une fois intégrées en base de production, les données feront alors l'objet d'une validation manuelle qui permettra de compléter ou préciser la validation automatique.

6. Calcul de la sensibilité des données

Le CSRPN a validé en 2017 un référentiel de sensibilité ¹⁵. Son application au jeu de données aboutit à ces niveaux de sensibilité (seuls les taxons dont certaines données sont sensibles sont listés) :

nom_cite	group2_inpn	sensi_niveau	sensi_libelle	nb_obs	nb_total_obs	%
Agonostomus telfairii	Poissons	m02	Floutage à la maille 2 x 2 km	1	1	100,0%
Amatitlania nigrofasciata	Poissons	m02	Floutage à la maille 2 x 2 km	3	3	100,0%
Ambassis ambassis	Poissons	m02	Floutage à la maille 2 x 2 km	2	2	100,0%
Anguilla marmorata	Poissons	m02	Floutage à la maille 2 x 2 km	3	3	100,0%
Trichogaster trichopterus	Poissons	m02	Floutage à la maille 2 x 2 km	3	3	100,0%
Valamugil sp	Poissons	m02	Floutage à la maille 2 x 2 km	1	1	100,0%
Varuna litterata	Crustacés	m02	Floutage à la maille 2 x 2 km	21	21	100,0%
Xiphophorus hellerii	Poissons	m02	Floutage à la maille 2 x 2 km	5	5	100,0%
Xiphophorus maculatus	Poissons	m02	Floutage à la maille 2 x 2 km	3	3	100,0%
Yirrkala tenuis	Poissons	m02	Floutage à la maille 2 x 2 km	3	3	100,0%
Total				350	350	100,0%

Pour mémoire, par mesure préventive, l'ensemble des données appartenant à des groupes pas encore traités par le référentiel sont considérées comme sensibles dans le référentiel de sensibilité faute de connaissances suffisantes pour le moment sur ce groupe.

7. Bilan de l'import

7.1 Bilan général

	Nombre de données	%
Données fournies	1 259	100,0%
Données ajoutées (dupliquées)	0	
Données écartées à la demande du producteur	0	0,0%
Doublons (non intégrés)	0	0,0%
Données non conformes et cohérentes (non intégrées)	10	0,8%
Données conformes et cohérentes (intégrées)	1 249	99,2%

15 https://www.borbonica.re/donnees_sensibles/

Les métadonnées renseignées pour ces 4 jeux de données sont reprises dans le tableau ci-dessous, ainsi que l'URL des fiches de métadonnées spécifiques créées dans l'outil de gestion national des métadonnées du SINP :

Signification	JDD1	JDD2	JDD3	JDD4
Identifiant unique auto-incrémenté de l'import.	17	18	16	19
Libellé de l'import	1er import en base de production	1er import en base de production	1er import en base de production	1er import en base de production
Identifiant du cadre d'acquisition	5473	5474	5493	5493
Identifiant du jeu de données	10727	10728	10747	10748
Libellé court du jeu de données	rnnesp_obsocc_201712	rnnesp_flore_201712	rnnesp_ipa_ika_2012-2014	rnnesp_peches_2012-2016
Description du jeu de données	Observations occasionnelles dans la RNN Etang de Saint-Paul avant le 31/12/2017	Inventaire de la Flore patrimoniale dans la RNN Etang de Saint-Paul avant le 31/12/2017	Suivi ornithologique par IPA et IKA réalisés dans la Réserve nationale de l'Étang de Saint-Paul en 2012 et 2014	Inventaire des poissons et crustacés d'eau douce capturés au filet maillant, nasse et pêche électrique dans la Réserve naturelle de l'Étang de Saint-Paul en 2012 et 2016
Identifiant permanent et unique de la fiche de métadonnées du jeu de données auquel appartient la donnée. Cet identifiant est attribué par la plateforme	da19d647-1603-4759-883e-64aee7d2d46e	d0225f42-093c-4dc2-a553-07dfbacb556e	fe04ca5e-ae4e-4353-aabd-0d42e131d967	5b9ff015-e1b6-4b05-97f8-cfe3de82cd22
Date de la réception des données (envoyée par le producteur)	2018-06-26	2018-06-26	2018-06-26	2018-06-26
Date à laquelle l'import a été effectué	2018-07-20	2018-07-20	2018-07-20	2018-07-20
Date minimale des observations importées pour le jdd	2013-05-23	2017-11-09	2011-12-14	2011-11-23
Date maximale des observations importées pour le jdd	2017-11-24	2017-11-09	2014-04-30	2016-07-05
Acteur référent (celui qui est responsable des données)				
Acteur ayant réalisé l'import				
Remarque générale sur l'import	Rapport d'import v 1.1	Rapport d'import v 1.1	Rapport d'import v 1.1	Rapport d'import v 1.1
Fiche de métadonnées INPN du cadre d'acquisition	https://inpn.mnhn.fr/mtd/cadre/edit/5473	https://inpn.mnhn.fr/mtd/cadre/edit/5474	https://inpn.mnhn.fr/mtd/cadre/edit/5493	https://inpn.mnhn.fr/mtd/cadre/edit/5493
Fiche de métadonnées INPN du jeu de données	https://inpn.mnhn.fr/mtd/cadre/jdd/edit/10727	https://inpn.mnhn.fr/mtd/cadre/jdd/edit/10728	https://inpn.mnhn.fr/mtd/cadre/jdd/edit/10747	https://inpn.mnhn.fr/mtd/cadre/jdd/edit/10748

7.2 Niveaux de diffusion pour le grand public

Le niveau de diffusion au grand public résulte de la **combinaison de plusieurs paramètres, hiérarchisés** ci-dessous (du plus important au moins important) :

- **statut de validation** de la donnée (champ validite_niveau) : seules les données dont la validité est 'Certaine' ou 'Probable' sont visibles ;
- **sensibilité** de la donnée (champ sensi_niveau) : voir partie 6 ;
- **statut public ou privé** de la donnée (champ ds_publique) : les données publiques sont diffusées sans floutage ;
- **souhait de diffusion du producteur** (champ diffusion_niveau_precision) : le producteur peut souhaiter une diffusion avec la précision géographique d'origine ou après floutage à la maille de 2 km.

L'application de ces règles de diffusion aux jeux de données importés conduit à ces niveaux, qui seront applicables lorsque les données auront été validées par les têtes de réseau régionales :

code_diffusion	libelle_diffusion	nb_obs	%
["d","m10"]	Diffusion grand public à la maille de 10 km	2	0,2%
["d","m10","m02"]	Diffusion grand public à la maille de 2 km	348	27,6%
["g","d","m10","m02","m01","e","c","z"]	Diffusion grand public avec géométrie précise	909	72,2%
Total		1259	100,0%

7.3 Préconisations pour les futurs versements de données

Préconisations au producteur pour améliorer la qualité de ses jeux de données.

Annexe 1 : Signification et liste des valeurs des champs de la base de données source

Voir tableau joint.

Annexe 2 : Scripts d'import dans Borbonica

Voir les scripts SQL en PJ.

Annexe 3 : Réconciliation taxonomique

Voir le fichier suivant en PJ : annexe_3_réconciliation_taxonomique.ods

Annexe 4 : Fichiers importés

Liste des fichiers joints

Annexe 4 : Validation scientifique des données de cétacés

1. Organisation mise en place par la tête de réseau

Les données de cétacés sont soumises successivement à une validation automatique puis à une validation manuelle selon le logigramme repris en page suivante au Tableau 11. Le processus se base sur des **données de référence** et l'existence de **photographies** accompagnant l'observation. La zone à photographier pour permettre la détermination dépend du taxon. La dorsale peut déjà permettre d'avoir une petite idée du taxon, mais le plus de photos possibles est préconisé.

La confirmation par l'**acoustique** n'est pas envisageable en milieu marin, car la propagation du son dans l'eau est rapide et le son peut porter loin. Il est donc possible d'entendre un taxon alors que visuellement un autre taxon est observé.

Enfin le **niveau d'expertise n'est pas pris en compte** dans la validation car le niveau de certitude dans l'identification d'un axon dépend non seulement de l'expertise de l'observateur mais également d'un axon. Par exemple, il est rare de se tromper sur l'identification d'une baleine à bosse, alors que le grand dauphin de l'Indo-Pacifique peut très facilement être confondu avec le grand dauphin commun.

2. Données de référence

La **liste des taxons rares ou jamais observés dans la zone** correspond à l'ensemble des espèces et sous-espèces des familles des *Balaenidae*, *Balaenopteridae*, *Delphinidae*, *Kogiidae*, *Physeteridae* et *Ziphiidae*, à l'exception des quatre taxons fréquents repris dans le Tableau 10. Cela concerne 53 taxons dans la version 11 de Taxref. Aucune distinction n'est réalisée entre les taxons rares ou jamais observés.

Les **habitats préférentiels et les périodes d'observation habituelles** pour les 4 taxons de cétacés fréquemment observés sont précisés dans le Tableau 10. Seuls les critères limitants, apparaissant en gris, sont utilisés pour tester les données. Ce tableau est également décliné sous la forme de 3 couches SIG dans l'Annexe 12 : Couches SIG de référence. Ce critère doit permettre par ailleurs de limiter les risques d'erreur entre certains taxons pouvant être confondus comme *Tursiops aduncus* et *Tursiops truncatus*.

cd_nom	Espèce	Habitat préférentiel			Période de présence
		Secteur	Profondeur	Distance à côte	
528706	<i>Tursiops aduncus</i>	Ouest – Nord – Sud	<60m	<1km	Toute l'année
60927	<i>Tursiops truncatus</i>	Ouest – Nord – Sud	>200m	>3,5km	Toute l'année
60916	<i>Stenella longirostris</i>	Ouest (hors Baie de Saint Paul)	<80m	<3,0km	Toute l'année
60867	<i>Megaptera novaeangliae</i>	Ouest – Nord – Sud	-	-	Juin à octobre

En gris : critère limitant

Tableau 10: Habitats préférentiels des taxons de cétacés fréquemment observés

Ces données sont issues des études réalisées par Globice, ayant abouti aux publications suivantes :

- Condet M. & Dulau-Drouot V. 2016. *Habitat selection of two island-associated dolphin species from the south-west Indian Ocean*. Continental Shelf Research, 125: 18-27.
- Globice 2015. *Rapport d'étude Gestion durable des cétacés de la Réunion*
- Dulau-Drouot V., Fayon J., Mouysset L., Boucaud V. (2012). *Occurrence and residency pattern of humpback whale in Reunion Island (France) during 2004-2010*. Journal of Cetacean Research and Management 12(2): 255-263
- Dulau-Drouot, V., Boucaud, V. & Rota, B. (2008) *Cetacean diversity off La Réunion Island (France)*. Journal of Marine Biology Association of UK, 88(6): 1263–1272.

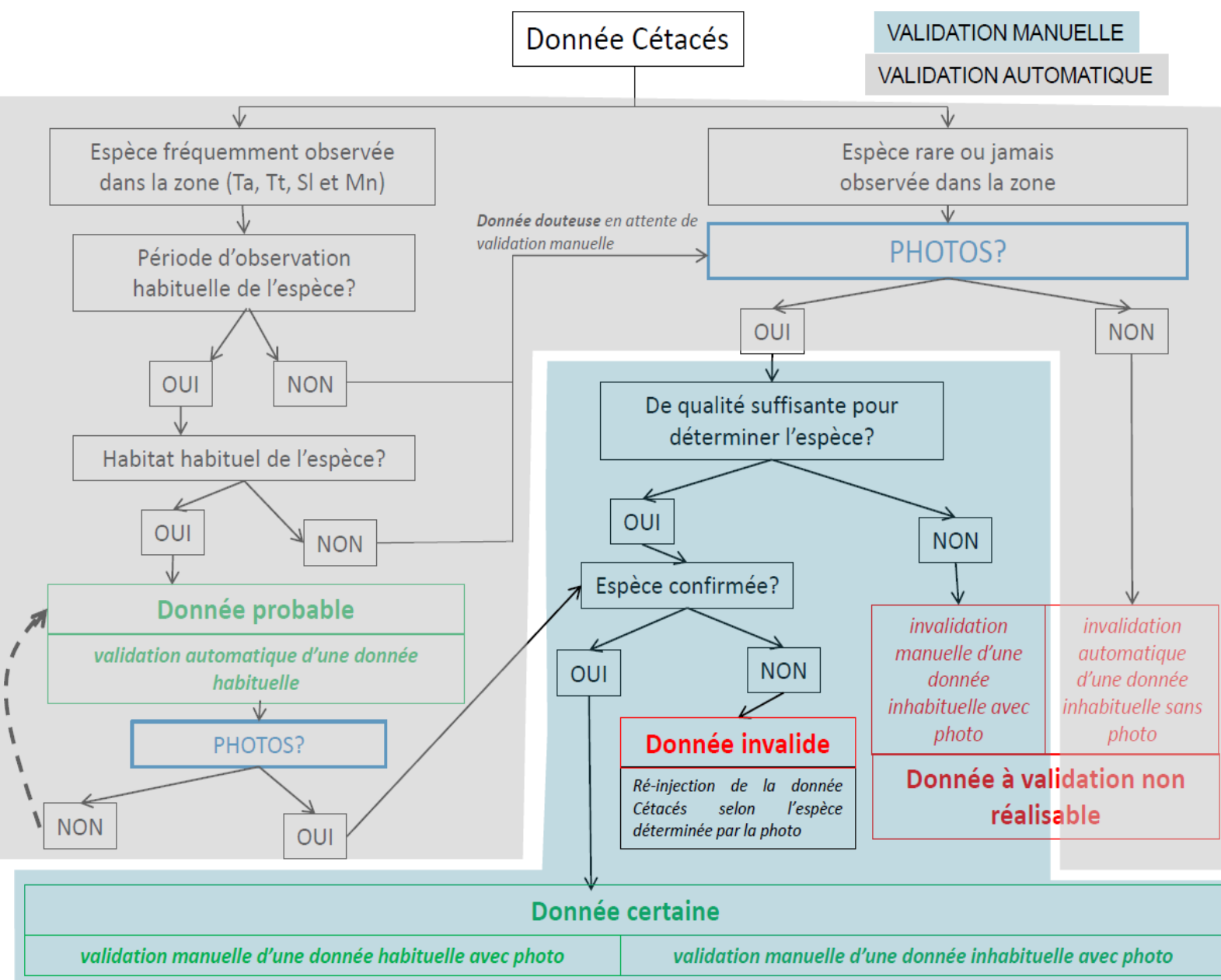


Tableau 11: Logigramme de validation scientifique des données de cétacés

Annexe 5 : Validation scientifique des données de tortues marines

1. Organisation mise en place par la tête de réseau

1.1 Procédure (actuelle) de validation des données par les producteurs

Les données collectées à La Réunion, ainsi que dans d'autres territoires français du sud-ouest de l'océan Indien, dans le cadre de **programmes scientifiques ayant des protocoles de collecte de données associés** sont bancarisées dans la **base de données TORSOOI** (TORTues marines du Sud Ouest de l'Océan Indien – www.torsooi.com) administrée par Kelonia, CEDTM et IFREMER.

Toutes les données saisies dans TORSOOI font l'objet d'un contrôle automatique à la saisie vérifiant le format (champs obligatoires, décimales, bornes min et max...), puis d'une validation manuelle après vérification de correspondance avec les fiches de terrain. Cette validation manuelle est réalisée une fois par an en début d'année n , et permet de vérifier et d'évaluer les données de l'année $n-1$. Elle consiste à extraire les données de l'année écoulée pour en contrôler ligne par ligne

1. la correspondance avec les données inscrites sur les fiches terrain utilisées par les observateurs formés (si disponibles),
2. la cohérence et la fiabilité par rapport aux connaissances disponibles sur les taxons dans la zone d'étude et à l'expertise des validateurs.

En cas de non correspondance, les données sont corrigées directement dans la base de données. Lors de ce processus, les données sont « qualifiées » une à une après vérification selon un critère de qualification (correspondant aux niveaux de validité précités). Les critères utilisés dans TORSOOI sont les suivants :

- **Non vérifiée** : Donnée saisie non vérifiée (valeur attribuée par défaut à la saisie).
- **Validée** : Donnée vérifiée à partir des fiches terrain papier ; correspondance totale entre les valeurs saisies et les relevés terrain. En l'absence de fiches terrain, il est considéré qu'il n'y a pas eu de relevé (cas particuliers des suivis en routine - journaliers).
- **Douteuse** : Donnée vérifiée d'après les fiches terrain, et concordante, mais une incertitude persiste sur l'un des paramètres du protocole de collecte (comme le zonage qui pourrait ne pas avoir été prospecté dans sa totalité par exemple), soit parce que l'information n'est pas disponible, soit parce qu'elle ne semble pas cohérente avec la valeur inscrite. Il peut aussi s'agir d'une donnée vérifiée sans fiche terrain (fiche perdue ou détruite par erreur après saisie dans la base de donnée) à partir des connaissances disponibles sur le taxon (l'aire de répartition connue, des paramètres biotiques et abiotiques de la niche écologique du taxon...), qui semble correcte, cependant, l'absence de fiche terrain ne permet pas de certifier la validité de cette donnée.
- **Invalide** : Donnée vérifiée d'après les fiches terrain, mais pas de correspondance trouvée entre le relevé terrain et la saisie, et la valeur inscrite sur le relevé terrain est aberrante. Cela peut aussi concerner un zonage aberrant, ou encore une erreur évidente pour un expert. Elle présente un trop bas niveau de fiabilité et 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, la preuve révèle une erreur de détermination).

En cas de doute, des échanges avec les producteurs de données permettent d'ajuster le niveau de qualification des données.

Les **données d'observation** (signalements transmis par le public, les associations ou les bureaux d'étude) sont contrôlées à la saisie **à partir des photos** jointes à l'information (si disponibles) ou après échanges avec le fournisseur de données. Sans photo, le taxon et l'identification ne peuvent être vérifiés, à moins de fournir un numéro de bague référencé (marques artificielles utilisées pour l'identification des tortues marines).

1.2 Procédure de validation automatique de la DEE après import

Le processus de **validation automatique** des données élémentaires d'échange, assuré par les administrateurs de Borbonica, suivra le schéma suivant :

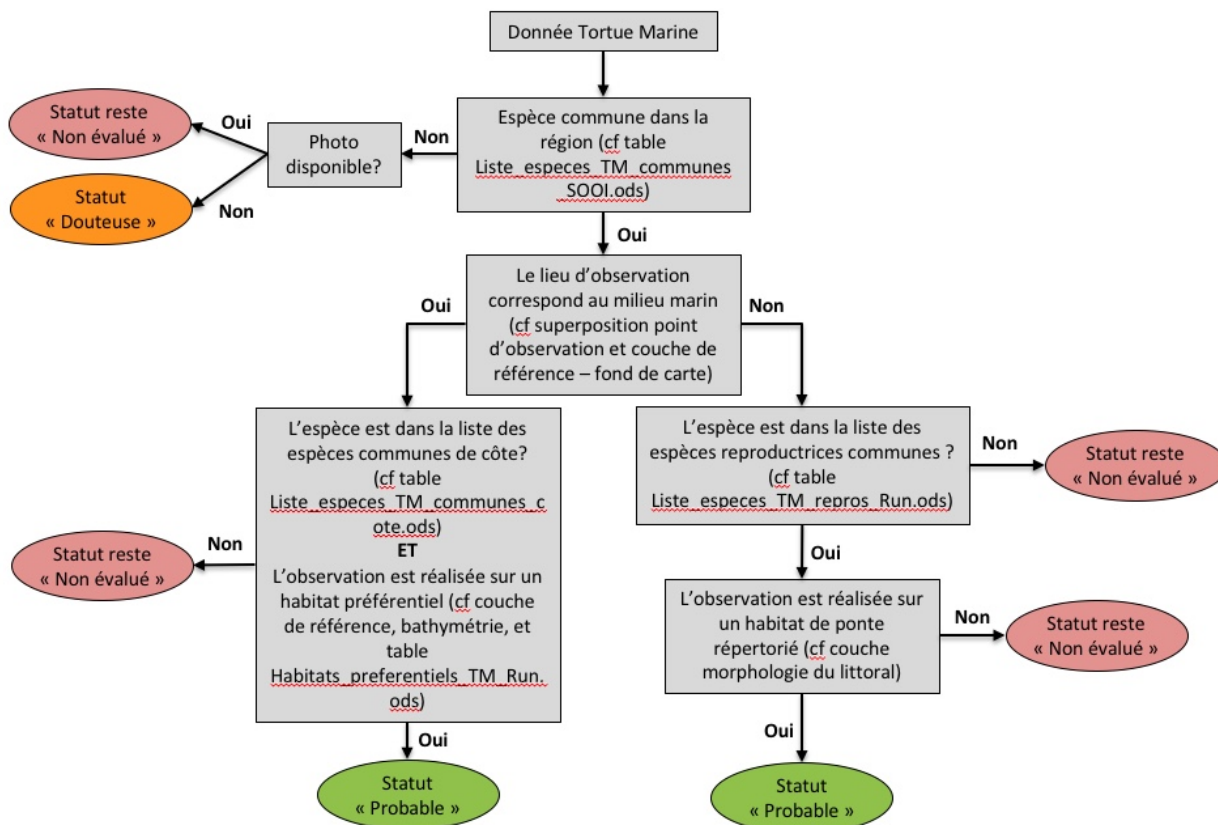


Illustration 4: Logigramme de validation scientifique automatique des données de tortues marines

Les données au statut « Douteuse » et « Non évalué » seront validées manuellement par les personnes identifiées par la tête de réseau selon la procédure de validation manuelle décrite ci-dessous.

1.3 Procédure de validation manuelle

La validation scientifique manuelle sera effectuée par les personnes identifiées pour la tête de réseau, à savoir Claire Jean, Stéphane Ciccione et Mayeul Dalleau, qui sont des experts « tortues marines » du sud ouest de l'océan Indien (cf fiche de critères de la tête de réseau). Elle sera effectuée **annuellement** après import des données dans le SINP sous un **délai de 2 mois**. Elle sera concentrée sur les **données au statut « Non évalué » après validation automatique**.

La validation des données scientifiques s'appuiera sur :

1. la procédure mise en place dans le cadre de TORSOOI (cf § 8.1.1), qui vérifie et qualifie les données bancarisées
2. les connaissances existantes (publications scientifiques, rapports d'expertise, bilans de projets...)
3. les données de référence disponibles sur la répartition des taxons à La Réunion et dans sa ZEE.

Elle concernera uniquement la **présence de taxons** de tortue marine et ainsi que **leur statut biogéographique**.

Les **données d'observation** (associations, bureaux d'études, public...) seront validées sur la base des connaissances des experts et des référentiels existants. L'expérience présumée de l'observateur sera également prise en compte dans ce processus de vérification et de validation. En cas de doute subsistant sur l'identification du taxon, d'autres informations pourront être utilisées si elles sont disponibles :

1. les **informations descriptives** associées à la donnée ;
2. des **photographies** ou des vidéos.

En fin de processus, les outils de visualisation des données cartographiques et d'analyse statistique de la plateforme Borbonica permettront d'identifier les éventuelles aberrations ou erreurs de qualification non détectées au cours du processus.

2. Donnée de référence

Les données de référence à utiliser pour la **validation automatique** sont les suivantes :

- Référentiel taxonomique national **Taxref** pour la faune, la flore et la fonge de France métropolitaine et d'outre-mer – Liste des taxons communs dans la région (issus de Taxref) (Liste_especes_TM_communes_SOOI.ods) :

cd_no m	nom_valide	nom_vern	id_court
77330	<i>Caretta caretta</i> (Linnaeus, 1758)	Tortue caouanne (La), Caouanne, Tortue carette	Cc
77338	<i>Chelonia mydas</i> (Linnaeus, 1758)	Tortue franche (La), Tortue verte	Cm
77347	<i>Eretmochelys imbricata</i> (Linnaeus, 1766)	Tortue caret (La), Tortue imbriquée, Tortue à écailles	Ei
77360	<i>Lepidochelys olivacea</i> (Eschscholtz, 1829)	Tortue olivâtre	Lo
77367	<i>Dermochelys coriacea</i> (Vandelli, 1761)	Tortue luth (La)	Dc

Tableau 12: Liste des taxons de tortues marines communes dans la région

- Taxons communs à la côte** à La Réunion (Liste_especes_TM_communes_cote.ods) :

cd_nom	nom_valide	nom_vern
77338	<i>Chelonia mydas</i> (Linnaeus, 1758)	Tortue franche (La), Tortue verte
77347	<i>Eretmochelys imbricata</i> (Linnaeus, 1766)	Tortue caret (La), Tortue imbriquée, Tortue à écailles

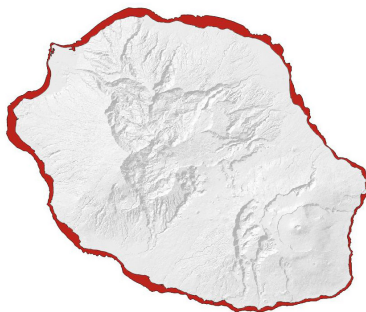
Tableau 13: Liste des taxons de tortues marines communes à la côte

- Liste des **taxons reproducteurs** connus à La Réunion (Liste_especes_TM_repro_Run.ods) :

cd_nom	nom_valide	nom_vern
77338	<i>Chelonia mydas</i> (Linnaeus, 1758)	Tortue franche (La), Tortue verte

Tableau 14: Liste des taxons de tortues marines reproductrices

- Distribution des taxons en mer à la côte** : [Zone d'habitat préférentiel \(zone d'alimentation côtière\) des tortues marines à La Réunion](#) – Atlas POLMAR_Sextant/ocean_indien (source Atlas POLMAR (KELONIA/DEAL/DOI))



- Répartition des taxons – **habitats préférentiels** connus à La Réunion :

cd_nom	nom_valide	Habitat_preferentiel	Profondeur_Run
77330	<i>Caretta caretta</i> (Linnaeus, 1758)	Pelagique	>200m
77338	<i>Chelonia mydas</i> (Linnaeus, 1758)	Cotier	<50m
77347	<i>Eretmochelys imbricata</i> (Linnaeus, 1766)	Cotier	<50m
77360	<i>Lepidochelys olivacea</i> (Eschscholtz, 1829)	Pelagique	>200m
77367	<i>Dermochelys coriacea</i> (Vandelli, 1761)	Pelagique	>200m

Tableau 15: Liste des habitats préférentiels des tortues marines (Habitats_preferentiels_TM_Run.ods)

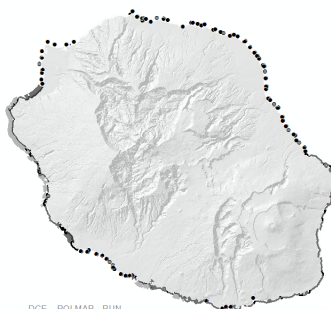
- **Limite terre-mer** : ajouter lien fiche MD

Source : couche des laisses de mer TRONCON_LAISSE.TAB issue de la BD Topo (IGN, avril 2014). Voir la couche SIG en Annexe 12 : Couches SIG de référence.

- **Habitats de ponte potentiels** : [Zones de nidification potentielles pour les tortues marines à La Réunion](#) – Atlas POLMAR_Sextant/ocean_indien (source Atlas POLMAR (KELONIA/DEAL/DOI)).

Voir la couche SIG dans l'Annexe 12 : Couches SIG de référence.

- **Géomorphologie des plages** (type et couleur du sable) : [Morphologie et évolution du littoral de la Réunion](#) (source : DIREN Réunion – BRGM, 2003) ET/OU [Morphologie du littoral "naturel" au niveau du jet de rive – île de La Réunion \(DCE-POLMAR-RUN\)](#) – Atlas POLMAR_Sextant/ocean_indien (source Atlas POLMAR (UBO/BRGM))



- **Bases de connaissances sur les erreurs et confusions fréquentes** :

Les deux espèces fréquentant les eaux côtières de La Réunion, tortue verte *Chelonia mydas* et tortue imbriquée *Eretmochelys imbricata*, sont souvent confondues car leurs couleurs sont proches, et elles disposent toutes 2 de 4 paires d'écailles costales (critère de détermination dans la clé d'identification des espèces de tortues marines).

Il peut aussi y avoir des confusions entre la tortue caouanne, *Caretta caretta*, et la tortue verte juvénile qui ont toutes deux une couleur brune orangée. La tortue verte adulte prend une couleur plus verdâtre et ponctuée de noir.

En cas de doute, seule une photographie permettra de valider la donnée et de lui attribuer la valeur « Certain – très probable ».

Annexe 6 : Validation scientifique des données de chiroptères

1. Organisation mise en place par la tête de réseau

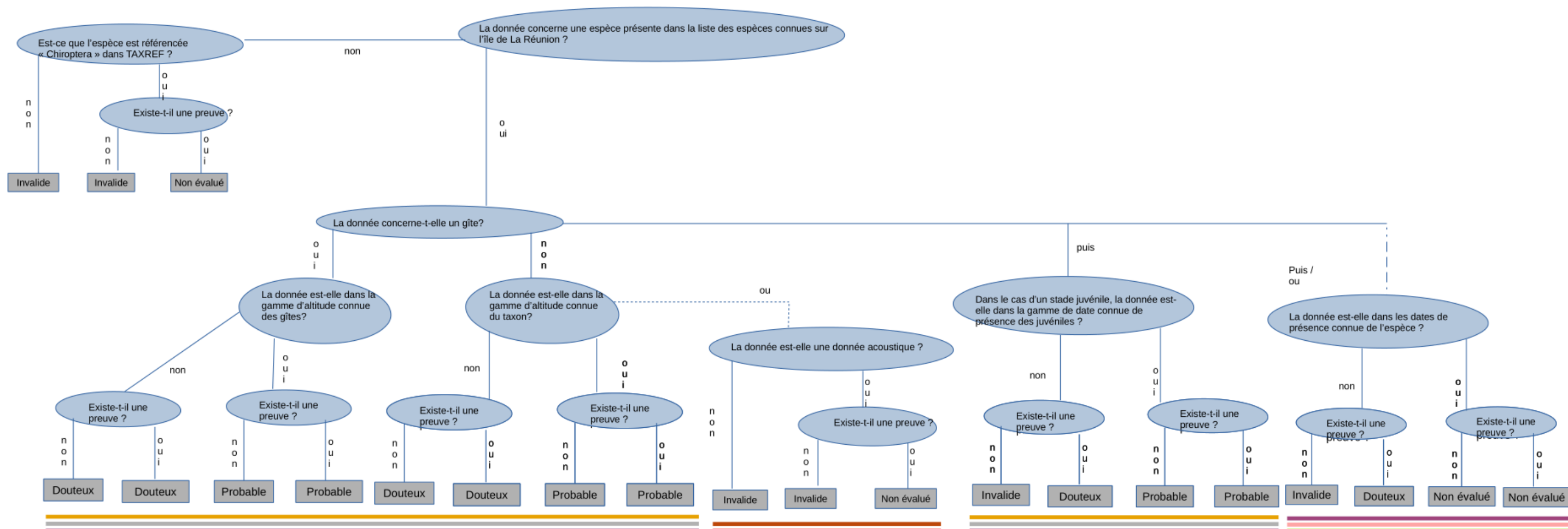
La **validation automatique** repose sur la réalisation successive de plusieurs tests (voir Illustration 5):

- *1_List_SP* : observation d'un taxon figurant dans la **liste des taxons présents à La Réunion et des types acoustiques** appartenant potentiellement à des taxons non décrits ;
- présence du taxon dans le référentiel taxonomique **Taxref** ;
- **type d'observation** (acoustique ou non) ;
- *2_Gamme altitude gîte* : observation de gîte réalisée dans une **gamme d'altitude connue** de présence des gîtes/dortoirs du taxon à La Réunion ;
- *3_Gamme altitude taxon* : observation réalisée dans une **gamme d'altitude de contact** du taxon connue à La Réunion ;
- *4_Gamme date Juv* : observation de juvénile réalisée à une **période connue de présence de juvénile** du taxon à La Réunion ;
- *5_Date présence-SP* : observation réalisée à une **période connue de présence du taxon** à La Réunion ;
- existence de **preuves**.

La validation s'effectuant en parallèle sur plusieurs champs (3 au maximum), le statut final qualifiant la donnée est celui **le plus pénalisant**.

La **validation manuelle** permet dans un second temps de préciser le statut de la donnée lorsqu'une **preuve** est fournie. Les données sont validées le plus rapidement possible à partir de la notification de versement des données au SINP 974 via le logiciel QGIS. Le **périmètre de validation** de chaque donnée peut si nécessaire être élargi au-delà des champs décrits plus haut au vu des valeurs des champs pouvant paraître hors-norme.

Logigramme de validation scientifique automatique des données de Chiroptères



Espèce concernée :

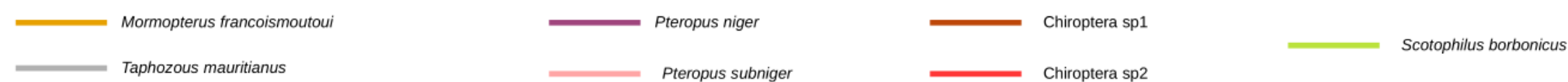


Illustration 5: Logigramme de validation scientifique des données de chiroptères

2. Données de référence

La validation automatique s'appuie sur les données de référence suivantes :

- **liste des taxons présents à La Réunion et des types acoustiques** appartenant potentiellement à des taxons non décrits :

cd_nom	lb_nom	reu *	codeEsp
418762	<i>Taphozous mauritanus</i>	P	TAPMAU
458690	<i>Mormopterus francoismoutoui</i>	E	MORFRA
418806	<i>Pteropus niger</i>	S	PTENIG
418807	<i>Pteropus subniger</i>	X	PTESUB
458692	<i>Scotophilus borbonicus</i>	Z	SCOBOR
-12 **	<i>Chiroptera sp1</i>	D ?	sp1
-13 **	<i>Chiroptera sp2</i>	D ?	sp2

* Statut biogéographique issu de Taxref (complété pour les deux types acoustiques)

** Identifiant issu de Borbonica dans l'attente d'une éventuelle intégration à Taxref

Tableau 16: Liste des taxons et types acoustiques de chiroptères présents à La Réunion

(voir Annexe 13 : Liste des statuts biogéographiques des taxons)

- **gammes d'altitudes de présence des gîtes/dortoirs** connues :

cd_nom	lb_nom	Alt_min_gîte (m)	Alt_max_gîte (m)
418762	<i>Taphozous mauritanus</i>	0	486
458690	<i>Mormopterus francoismoutoui</i>	0	2070
418806	<i>Pteropus niger</i>	0	334

Tableau 17: Gamme d'altitudes connue de présence des gîtes et dortoirs de chiroptères

- **gamme d'altitudes de contact** connues :

cd_nom	lb_nom	Alt_min_taxon (m)	Alt_max_taxon (m)
418762	<i>Taphozous mauritanus</i>	0	2150
458690	<i>Mormopterus francoismoutoui</i>	0	2320
418806	<i>Pteropus niger</i>	0	1100

Tableau 18: Gammes d'altitudes connues de contact des chiroptères

- **périodes connues de présence de juvéniles** :

cd_nom	lb_nom	date début juvénile	date fin juvénile
418762	<i>Taphozous mauritanus</i>	2 périodes : de début décembre à fin avril, puis de début juillet à fin septembre	
458690	<i>Mormopterus francoismoutoui</i>	début décembre	fin mai
418806	<i>Pteropus niger</i>	début octobre	fin mars

Tableau 19: Périodes connues de présence de juvéniles de chiroptères

- périodes connues de présence des taxons :

cd_nom	lb_nom	Présence taxon
418806	<i>Pteropus niger</i>	Avant 1801-01-01 et après 2001-01-01
418807	<i>Pteropus subniger</i>	Avant 1870-01-01
458692	<i>Scotophilus borbonicus</i>	Avant 1902-12-12
-12	<i>Chiroptera sp1</i>	Après 2009-03-18
-13	<i>Chiroptera sp2</i>	Après 2009-04-18

Tableau 20: Périodes connues de présence des taxons de chiroptères

La **fréquence de mise à jour des données de référence** dépendra de l'évolution des tests et des contrôles dans Borbonica ainsi que des potentielles évolutions taxonomiques. Le protocole complet présente les cas de réévaluations de certains critères. Il conviendra de statuer en fonction de la fréquence d'apparition de ces cas pour définir une fréquence de mise à jour.

Annexe 7 : Validation scientifique des données de mammifères terrestres non volants

1. Organisation mise en place par la tête de réseau

1.1 Validation automatique

La **validation automatique** repose sur la réalisation successive de plusieurs tests (voir Illustration 6) :

- taxons MTNV1 : présence sur la **liste des taxons naturalisés ou des taxons non naturalisés et facilement identifiables** ;
- taxons MTNV2 : **risque de confusion** avec d'autres taxons ;
- présence de **photos ou vidéos** disponibles.

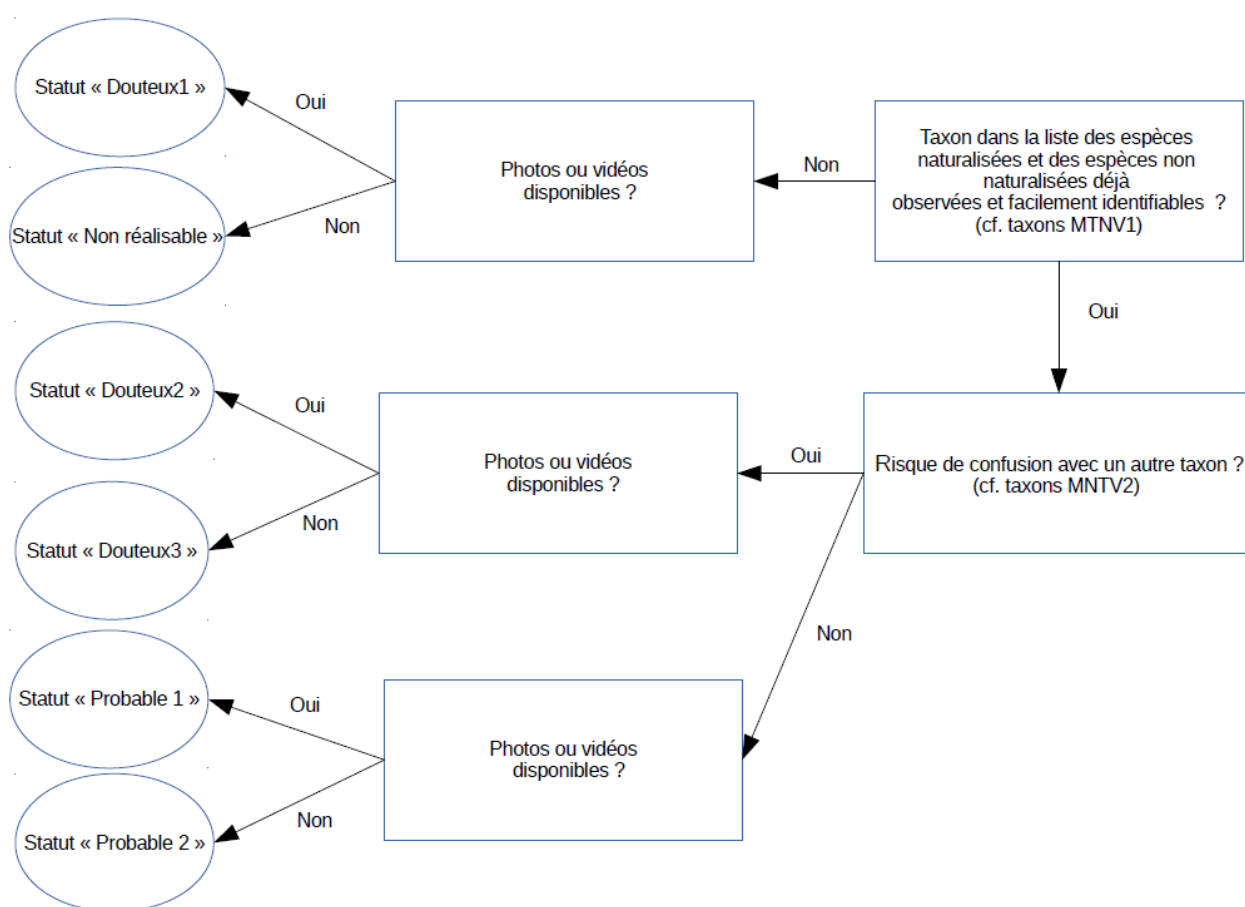
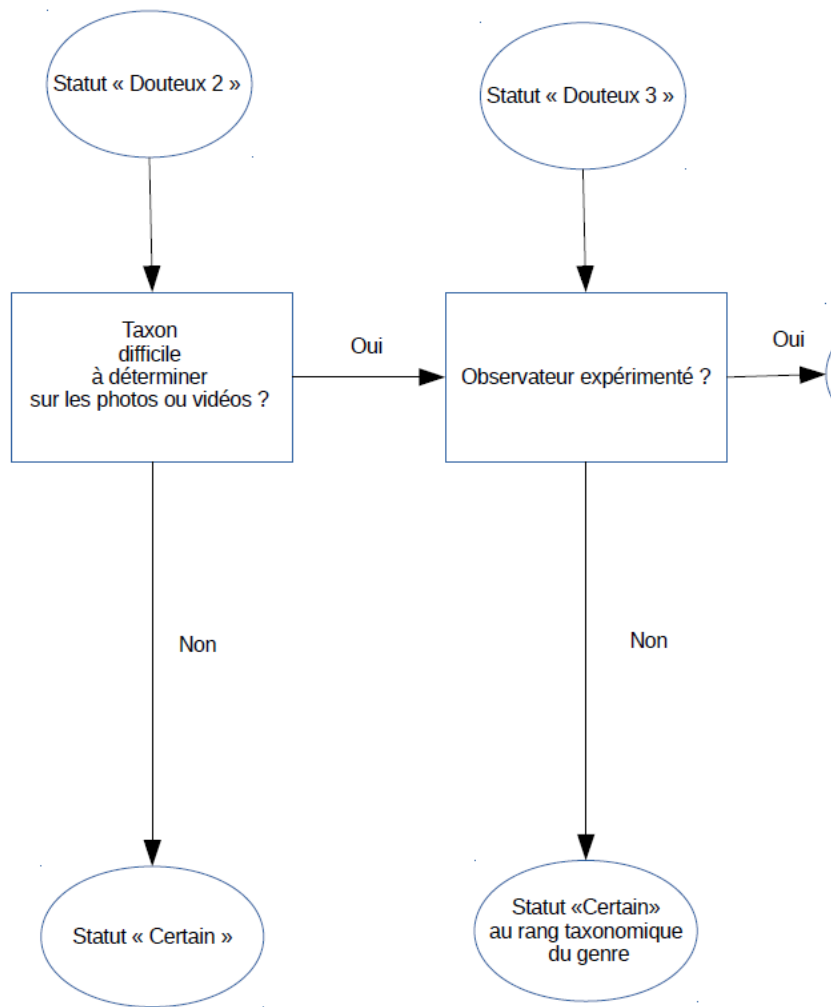
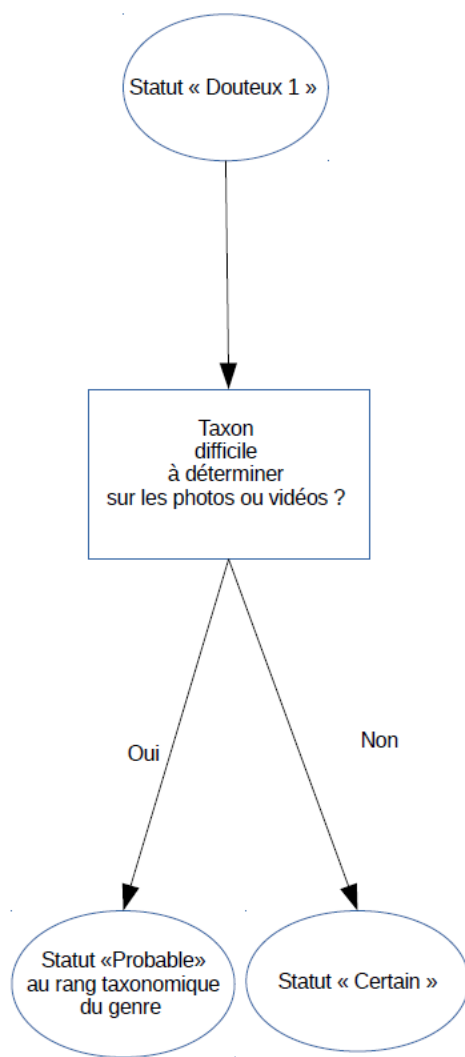


Illustration 6: Logigramme de validation scientifique automatique des données de mammifères terrestres non volants

1.2 Validation manuelle

La validation manuelle (voir Illustration 7) permet d'affiner les résultats en s'appuyant sur :

- l'**analyse de preuves**, en particulier leur qualité ;
- la **compétence** connue de l'observateur.



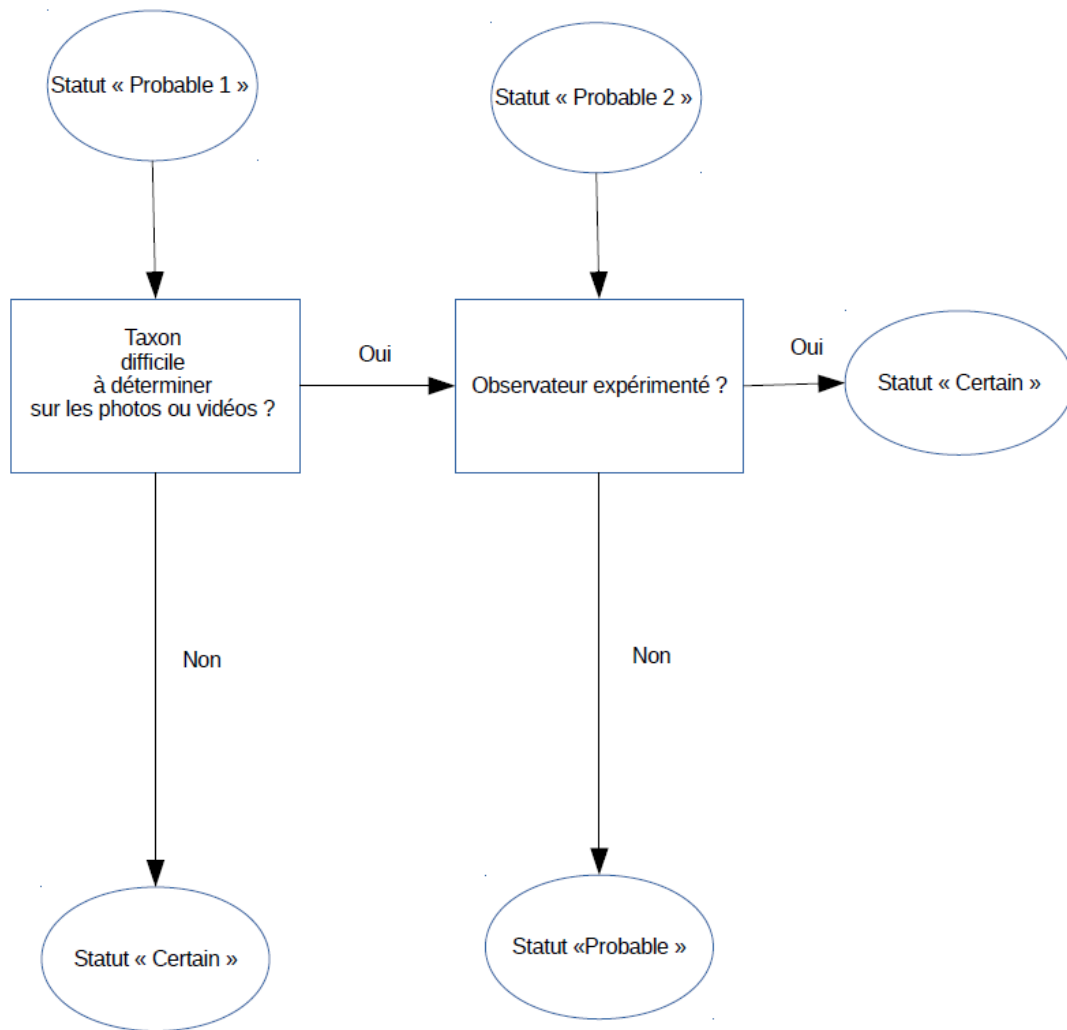


Illustration 7: Logigramme de validation scientifique manuelle des données de mammifères terrestres non volants

2. Données de référence

Les données de référence suivantes sont utilisées pour la validation automatique :

- **Liste des taxons naturalisés et des taxons non naturalisés mais déjà observés et facilement identifiables (MTNV 1) :**

famille	cd_ref	nom_valide	nom_vern	reu	occ_r un	Remarques
Tenrecidae	418758	<i>Tenrec ecaudatus</i> (Schreber, 1777)	Tenrec acaude, Tangué	I	I	
Canidae	162663	<i>Canis familiaris</i> Linnaeus, 1758	Chien	M	M	Echappé au moins une fois
Felidae	60595	<i>Felis catus</i> Linnaeus, 1758	Chat domestique, Chat haret	J	J	
Mustelidae	60731	<i>Mustela putorius</i> Linnaeus, 1758	Putois d'Europe, Furet	M	M	Echappé au moins une fois
Bovidae	199727	<i>Bos taurus</i> Linnaeus, 1758	Boeuf domestique, Vache domestique	I	J	
Bovidae	61097	<i>Capra hircus</i> Linnaeus, 1758	Chèvre domestique	I	J	
Bovidae	199754	<i>Ovis aries</i> Linnaeus, 1758	Mouton domestique		M	Captivité
Cervidae	459565	<i>Rusa timorensis</i> (de Blainville, 1822)	Rusa de Timor, Cerf de Java, Cerf rusa	J	J	
Cervidae	714551	<i>Rusa timorensis russa</i> (Müller & Schlegel, 1845)	Rusa de Timor, Cerf de Java, Cerf rusa	J	J	
Suidae	60981	<i>Sus scrofa</i> Linnaeus, 1758	Sanglier	M	M	Echappé au moins une fois
Leporidae	456605	<i>Lepus nigricollis</i> Cuvier, 1823	Lièvre indien, Lièvre à collier noir	I	I	
Equidae	199739	<i>Equus asinus</i> Linnaeus, 1758	Ane domestique		M	Captivité
Equidae	60976	<i>Equus caballus</i> Linnaeus, 1758	Cheval domestique	M	M	Echappé au moins une fois
Hominidae	815096	<i>Homo sapiens</i> Linnaeus, 1758	Homme moderne, Homme de Cro-Magnon	P		
Muridae	61568	<i>Mus musculus</i> Linnaeus, 1758	Souris grise, Souris domestique	J	J	
Muridae	197057	<i>Rattus Fischer</i> , 1803		J	J	
Muridae	61585	<i>Rattus norvegicus</i> (Berkenhout, 1769)		J	J	
Muridae	61587	<i>Rattus rattus</i> (Linnaeus, 1758)	Rat noir, Rat commun	J	J	
Soricidae	418760	<i>Suncus murinus</i> (Linnaeus, 1766)	Grande Pachyure, Musaraigne musquée, Pachyure musquée	J	J	

occ_run = statut biogéographique apprécié localement, différent de Taxref

Source des remarques : Soubeyran, 2008 ; Caceres & Mongin, 2011; Potin, 2013 ; GEIR 2018)

Tableau 21: Liste des taxons de mammifères terrestres non volants naturalisés et des taxons non naturalisés mais déjà observés et facilement identifiables

- **Liste des taxons non naturalisés déjà observés mais facilement identifiables ou jamais observés :**

Cette liste est construite en prenant l'ensemble des espèces et sous-espèces de mammifères connues dans Taxref, à l'exception de celles déjà présentes dans la liste précédente, de l'ordre des chiroptères (traités par

ailleurs) et des taxons vivant exclusivement en milieu marin (ordre des *Sirenia* et familles des *Balaenidae*, *Balaenopteridae*, *Delphinidae*, *Kogiidae*, *Monodontidae*, *Neobalaenidae*, *Odobenidae*, *Phocidae*, *Phocoenidae*, *Physeteridae* et *Ziphiidae*).

- **Liste des taxons pouvant être confondus (MTNV 2) :**

famille	cd_ref	nom_valide	nom_vern	reu	occ_run
Muridae	61585	<i>Rattus norvegicus</i> (Berkenhout, 1769)		J	J
Muridae	61587	<i>Rattus rattus</i> (Linnaeus, 1758)	Rat noir, Rat commun	J	J

occ_run = statut biogéographique apprécié localement, différent de *Taxref*

Tableau 22: Liste des taxons de mammifères terrestres non volants pouvant être confondus

Annexe 8 : Validation scientifique des données de reptiles et amphibiens

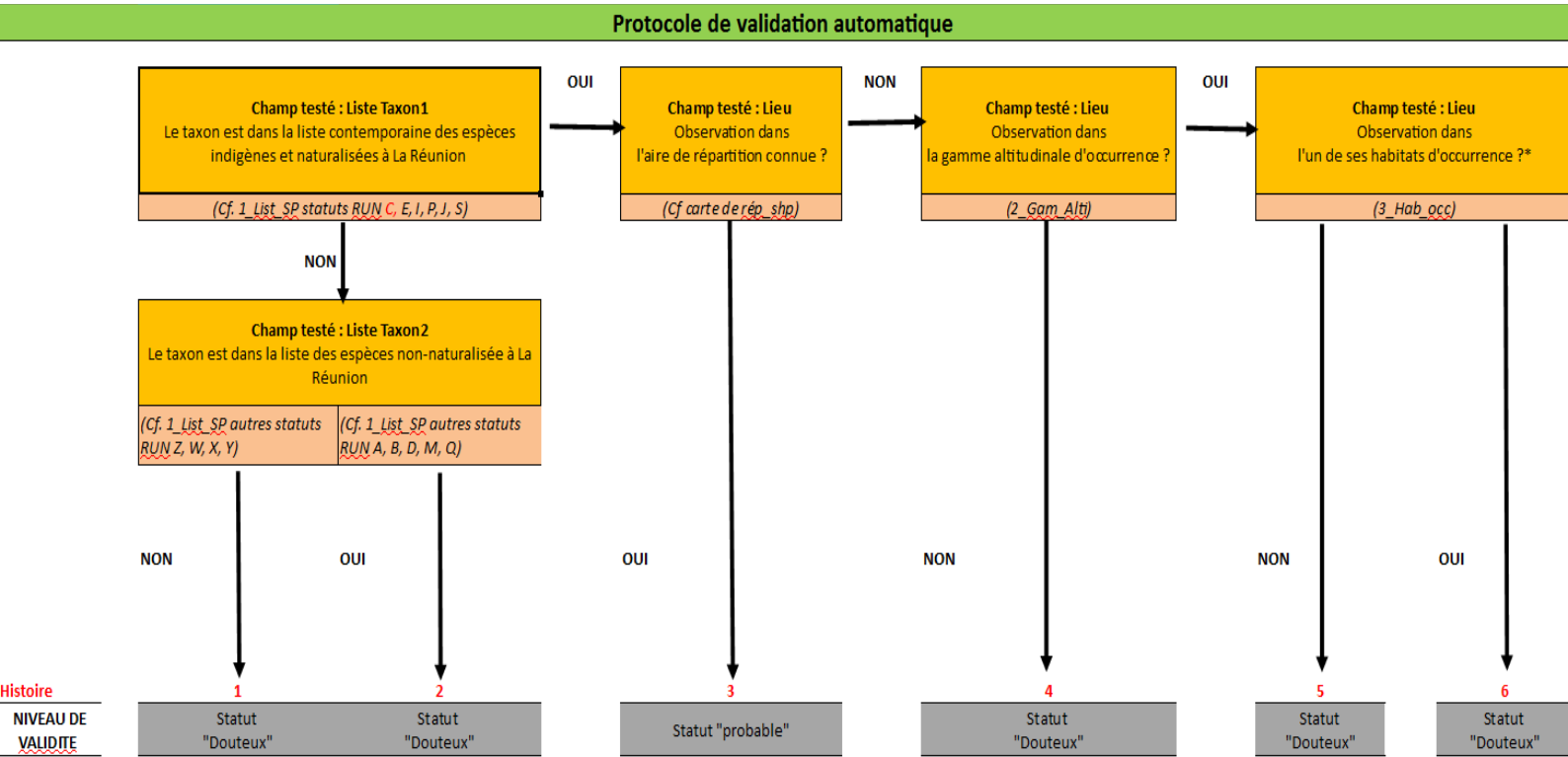
1. Organisation mise en place par la tête de réseau

La **validation automatique** repose sur la réalisation successive de plusieurs tests:

- **1_List_SP** : observation d'un taxon figurant dans une des listes suivantes : **liste des taxons indigènes et naturalisés** de La Réunion (statut P, E, I, J, S, C), celle des **taxons non naturalisés** (statuts A, B, D, M, Q) ou bien celle des **taxons disparus ou éteints** (statuts Z, W, X, Y). Le Tableau 23 liste les statuts des taxons ayant déjà observés à La Réunion ;
- observation au sein de l'**aire de répartition connue** du taxon ;
- **2_Gam_Alti** : observation réalisée dans la **gamme d'altitude d'occurrence** connue du taxon ;
- **3_Hab_occ** : observation réalisée dans un **habitat d'occurrence** connu du taxon.

La **validation manuelle** permet dans un second temps de préciser le statut de la donnée en s'appuyant sur l'analyse de **preuves** éventuelles (photos, vidéos ou échantillons) et le **niveau d'expérience** connu de l'observateur.

Ces différents tests de validation sont synthétisés dans l'Illustration 8 et l'Illustration 9 en pages suivantes.



Critères du protocole automatique : Liste Taxon, répartition, gamme altitudinale et habitat

* Le test est réalisé au niveau cartographique sur la couche "oiseaux_habitat" réalisée par Jean-Cyrille Notter dans le cadre du STOC (2012).

Les limites des polygones n'étant pas définies avec une précision fine, au regard de leur écologie, des espèces peuvent apparaître dans des polygones d'habitats semblant aberrants (ex : "*Phelsuma inexpectata*" dans "canne à sucre").

Illustration 8: Logigramme de validation scientifique automatique des données de reptiles et

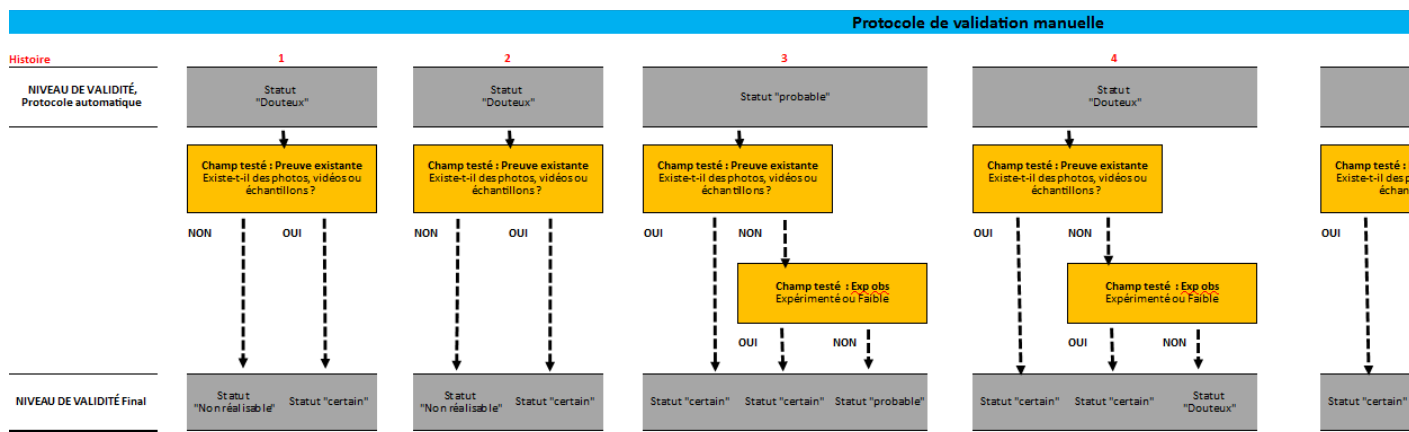


Illustration 9: Logigramme de validation scientifique manuelle des données de reptiles et

2. Données de référence

La validation automatique s'appuie sur les données de référence suivantes, qui devront être modifiées régulièrement en fonction de l'avancée des connaissances :

- **liste des taxons observés à La Réunion et statut biogéographique associé :**

ordre	famille	cd_nom	rang	lb_nom	reu *
Anura	Bufonidae	840666	ES	<i>Sclerophrys gutturalis</i>	I
Anura	Ptychadenidae	528674	ES	<i>Ptychadena mascareniensis</i>	I
Squamata	Agamidae	528683	ES	<i>Agama agama</i>	J
Squamata	Agamidae	528684	ES	<i>Calotes versicolor</i>	J
Squamata	Chamaeleonidae	528685	ES	<i>Furcifer pardalis</i>	I
Squamata	Colubridae	528688	ES	<i>Lycodon aulicus</i>	J
Squamata	Elapidae	838087	ES	<i>Hydrophis platurus</i>	B
Squamata	Elapidae	456383	ES	<i>Laticauda laticaudata</i>	B
Squamata	Gekkonidae	443978	ES	<i>Gehyra mutilata</i>	I
Squamata	Gekkonidae	418692	ES	<i>Hemidactylus frenatus</i>	J
Squamata	Gekkonidae	713653	ES	<i>Hemidactylus mercatorius</i>	I
Squamata	Gekkonidae	701839	ES	<i>Hemidactylus parvimaculatus</i>	I
Squamata	Gekkonidae	456375	ES	<i>Hemiphyllodactylus typus</i>	I
Squamata	Gekkonidae	528689	ES	<i>Nactus soniae</i>	Z
Squamata	Gekkonidae	528690	ES	<i>Phelsuma astriata</i>	I
Squamata	Gekkonidae	773857	SSES	<i>Phelsuma astriata semicarinata</i>	I
Squamata	Gekkonidae	432597	ES	<i>Phelsuma borbonica</i>	S
Squamata	Gekkonidae	528680	SSES	<i>Phelsuma borbonica borbonica</i>	E
Squamata	Gekkonidae	528681	SSES	<i>Phelsuma borbonica mater</i>	E
Squamata	Gekkonidae	528691	ES	<i>Phelsuma cepedianae</i>	B **
Squamata	Gekkonidae	528692	ES	<i>Phelsuma grandis</i>	J
Squamata	Gekkonidae	528679	ES	<i>Phelsuma inexpectata</i>	E
Squamata	Gekkonidae	418689	ES	<i>Phelsuma laticauda</i>	J
Squamata	Gekkonidae	528693	ES	<i>Phelsuma lineata</i>	I
Squamata	Iguanidae	350756	ES	<i>Iguana iguana</i>	M
Squamata	Scincidae	459101	ES	<i>Cryptoblepharus boutonii</i>	P
Squamata	Scincidae	528678	SSES	<i>Cryptoblepharus boutonii boutonii</i>	S
Squamata	Scincidae	714553	ES	<i>Gongylomorphus borbonicus</i>	Z
Squamata	Scincidae	528687	ES	<i>Leiolopisma ceciliae</i>	Z
Squamata	Typhlopidae	728244	ES	<i>Indotyphlops braminus</i>	I
Testudines	Chelydridae	199209	ES	<i>Chelydra serpentina</i>	M
Testudines	Emydidae	77424	ES	<i>Trachemys scripta</i>	M
Testudines	Emydidae	77425	SSES	<i>Trachemys scripta elegans</i>	M
Testudines	Geoemydidae	726070	ES	<i>Mauremys sinensis</i>	I
Testudines	Testudinidae	714556	ES	<i>Astrochelys radiata</i>	M
Testudines	Testudinidae	459614	ES	<i>Cylindraspis indica</i>	Z

Tableau 23: Liste des taxons de reptiles et amphibiens indigènes et naturalisés de La Réunion

* Statut biogéographique issu de Taxref - voir l'Annexe 13 : Liste des statuts biogéographiques des taxons

** Statut évalué localement, différent du statut attribué dans Taxref (I)

- **aire de répartition connues :** voir les couches SIG pour 18 taxons en Annexe 12 : Couches SIG de référence.
- **gamme d'altitudes d'occurrence connues :**

cd_nom	Espèce	Alt_min (m)	Alt_max (m)
528683	<i>Agama agama</i> (complexe)	0	500
840666	<i>Amietophrynus gutturalis</i>	0	1 900
528684	<i>Calotes versicolor</i>	0	2 900
528685	<i>Furcifer pardalis</i>	0	1 000
443978	<i>Gehyra mutilata</i>	0	1 250
418692	<i>Hemidactylus frenatus</i>	0	1 200
713653	<i>Hemidactylus mercatorius</i>	0	900
701839	<i>Hemidactylus parvimaculatus</i>	0	1 150
456375	<i>Hemiphyllodactylus typus</i>	0	900
728244	<i>Indotyphlops braminus</i>	0	100
528688	<i>Lycodon aulicus</i>	0	2 200
528690	<i>Phelsuma astriata</i>	0	100
773857	<i>Phelsuma astriata semicarinata</i>	0	100
432597	<i>Phelsuma borbonica</i>	0	2 800
528680	<i>Phelsuma borbonica borbonica</i>	0	2 800
528681	<i>Phelsuma borbonica mater</i>	0	2 800
528692	<i>Phelsuma grandis</i>	0	600
528679	<i>Phelsuma inexpectata</i>	0	250
418689	<i>Phelsuma laticauda</i>	0	600
528693	<i>Phelsuma lineata</i>	0	400
528674	<i>Ptychadena mascareniensis</i> (complexe)	0	900
459101	<i>Cryptoblepharus boutonii</i>	0	100
528678	<i>Cryptoblepharus boutonii boutonii</i>	0	100
726070	<i>Mauremys sinensis</i>	0	200

Tableau 24: Gammes d'altitudes connues d'occurrence de reptiles et amphibiens

- **Cartographie de l'occupation du sol :**

Voir la couche de cartographie de l'occupation du sol utilisée pour le STOC : oiseaux_habitat (Annexe 12 : Couches SIG de référence). Les types d'occupation du sol sont listés dans l'Annexe 9 : Validation scientifique des données d'oiseaux (Tableau 26).

• **Habitats d'occurrence connus :**

Espèce	cdnom	Habitat																							
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
<i>Agama agama</i> (complexe)	528683													X		X	X		X		X	X	X		X
<i>Calotes versicolor</i>	528684	X		X	X	X	X	X	X	X	X	X	X	X		X	X	X	X		X	X	X	X	X
<i>Cryptoblepharus boutonii</i>	459101												X				X								
<i>Cryptoblepharus boutonii boutonii</i>	528678												X				X								
<i>Furcifer pardalis</i>	528685						X			X	X	X	X	X		X		X	X		X	X	X		X
<i>Gehyra mutilata</i>	443978					X				X	X		X			X		X			X		X	X	X
<i>Hemidactylus frenatus</i>	418692	X				X		X		X			X	X		X	X	X	X	X	X	X	X	X	X
<i>Hemidactylus mercatorius</i>	713653																								X
<i>Hemidactylus parvimaculatus</i>	701839									X	X		X			X	X	X	X		X				X
<i>Hemiphyllodactylus typus</i>	456375						X			X	X	X	X			X		X			X	X	X		X
<i>Indotyphlops braminus</i>	728244									X						X		X	X						X
<i>Lycodon aulicus</i>	528688	X								X	X		X			X		X	X		X		X	X	X
<i>Phelsuma astriata</i>	528690																								X
<i>Phelsuma astriata semicarinata</i>	773857																								X
<i>Phelsuma borbonica</i>	432597	X	X	X	X	X	X	X	X	X	X	X	X	X		X	X	X			X	X	X	X	X
<i>Phelsuma borbonica borbonica</i>	528680	X	X	X	X	X	X	X	X	X	X	X	X	X		X	X	X			X	X	X	X	X
<i>Phelsuma borbonica mater</i>	528681	X	X	X	X	X	X	X	X	X	X	X	X	X		X	X	X			X	X	X	X	X
<i>Phelsuma grandis</i>	528692									X				X		X		X	X		X	X	X		X
<i>Phelsuma laticauda</i>	418689									X						X		X	X		X	X	X		X
<i>Phelsuma lineata</i>	528693															X							X		X
<i>Phelsuma inexpectata</i>	528679												X			X			X			X	X	X	X
<i>Ptychadena mascareniensis</i> (complexe)	528674					X	X			X	X		X	X		X	X	X	X		X	X	X	X	X
<i>Sclerophrys gutturalis</i>	840666		X	X	X	X	X	X	X	X	X	X	X	X		X		X	X		X	X	X	X	X

Tableau 25: Habitats d'occurrence de reptiles et amphibiens connus

Les limites des polygones n'étant pas définies avec une précision fine, au regard de leur écologie, des taxons peuvent apparaître dans des polygones d'habitats semblant aberrants (ex : "*Phelsuma inexpectata*" dans "canne à sucre" alors que sur le terrain le lézard était sur un vacoa à 5 m du champ de canne). Souvent il s'agit de points qui sont en limite de polygone. Il sera tenu compte de ce point lors de la validation manuelle.

Par ailleurs, lors de la validation manuelle, le fait que l'observation ait été réalisée sur des **structures artificielles** peut également être utilisé pour affiner le test, certains taxons en étant dépendants dans certains milieux.

Annexe 9 : Validation scientifique des données d'oiseaux

1. Organisation mise en place par la tête de réseau

1.1 Validation automatique

Les différentes étapes de la **validation automatique** sont indiquées dans l'illustration 10. Elle repose sur la succession des tests suivants :

- observation d'une taxon dans la liste des **espèces communes nicheuses ou ne posant pas de problème d'identification** ;
- observation d'un taxon dans la liste des **espèces communes migratrices** ;
- observation dans un **habitat préférentiel** (pour les nicheuses) ;
- observation **à une période habituelle** (pour les migratrices).

Elle est réservée aux taxons ne posant pas de problèmes de diagnose. Par exemple les taxons les plus **sujets à risques de confusion**, qui sont difficiles à distinguer, ne sont généralement pas intégrées dans la liste des taxons analysés, et doivent donc de facto faire l'objet d'une vérification manuelle. Le risque de confusion entre taxons dépend des périodes, du lieu, des conditions d'observation, et surtout du niveau de l'observateur. Mais s'il faut citer quelques grands classiques, les cas les plus fréquents sont listés ci-après, dans l'ordre de difficulté croissante à réaliser la détermination :

- Foudi de Madagascar femelle ou mâle internuptial / Moineau domestique / Veuve dominicaine femelle ou mâle internuptial / Travailleur à bec rouge femelle ou mâle internuptial ;
- Oiseau-lunettes gris / Oiseau-lunettes vert ;
- Astrild ondulé / Capucin damier ;
- Salangane des Mascareignes / Hironde de Bourbon ;
- Noddi brun / Noddi à bec grêle ;
- La plupart des limicoles (Bécasseaux, Chevaliers, Gravelots, Pluviers...), sachant qu'à La Réunion, ils sont en plumage internuptial, généralement terne et dépourvu des signes les plus caractéristiques du taxon telle qu'on la voit dans les livres et sur les photos. Il y a de plus de temps en temps des taxons inhabituels de limicoles qui arrivent à la Réunion, et dont l'identification est à la fois difficile et précieuse. Il faut donc rester très vigilant sur ce groupe de taxons ;
- Faucon concolore / Faucon d'Eleonore
- Sternes : plusieurs espèces différentes sont susceptibles d'être observées à La Réunion. Une même espèce peut présenter des différences de plumage assez importantes en fonction des individus et de la période.

Par ailleurs, si l'application des différents tests permet bien de qualifier de probable les données en accord avec les données de référence disponibles, **il n'est pas proposé de qualifier de douteuses** les observations ne concordant pas. En effet, les périodes de reproduction ou de présence pour les migrateurs ne sont pas très marquées (quand elles ne sont pas tout simplement inconnues) pour de nombreux taxons. Dans ces conditions, il paraît trop dangereux d'invalider une donnée qui serait hors du pic connu.

ETAPE 1 - TRAITEMENT AUTOMATIQUE APRES IMPORT

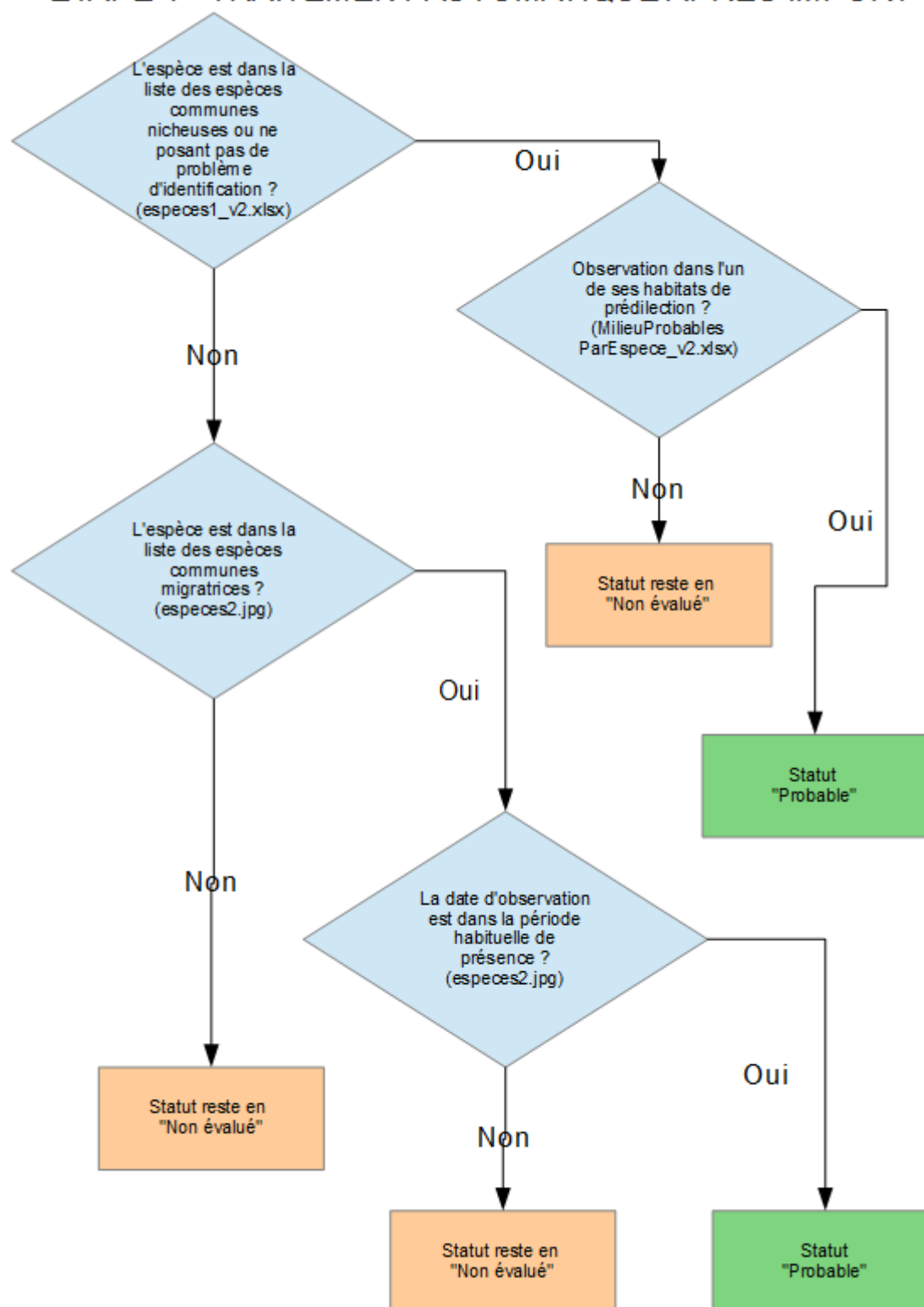


Illustration 10: Logigramme de validation scientifique automatique des données d'oiseaux

1.2 Validation manuelle

Les différentes étapes de la **validation manuelle** sont indiquées dans l'illustration 11. Cette validation s'appuie sur l'existence de **preuves** pour affiner la première analyse automatique ou traiter les taxons n'ayant pas été examinés lors de la première phase. Il peut s'agir :

- de **photographies**, si possible sous différents angles de vue ;
- d'un **commentaire** indiquant les parties de l'oiseau caractérisant ce taxon, l'habitat, le comportement de l'oiseau, les conditions dans lesquelles l'observation a été réalisée (jumelles, longue vue, météo, distance, durée, etc.) ;
- d'**enregistrements** de chants.

En fonction de la qualité et du niveau de détail de ces preuves, le niveau de validité pourra être qualifié de certain ou probable selon les cas :

- lorsque l'on dispose d'une photo de bonne qualité et qui comporte tous les éléments correspondant à la diagnose, le niveau certain pourra être attribué. Il en sera de même pour les descriptions détaillées des individus observés, en particulier pour les taxons très caractéristiques ou lorsque l'observateur est expérimenté ;
- dans la plupart des autres cas, le niveau probable sera attribué.

ÉTAPE 2 - TRAITEMENT MANUEL DES OBSERVATIONS RESTÉES EN STATUT "NON ÉVALUÉ"

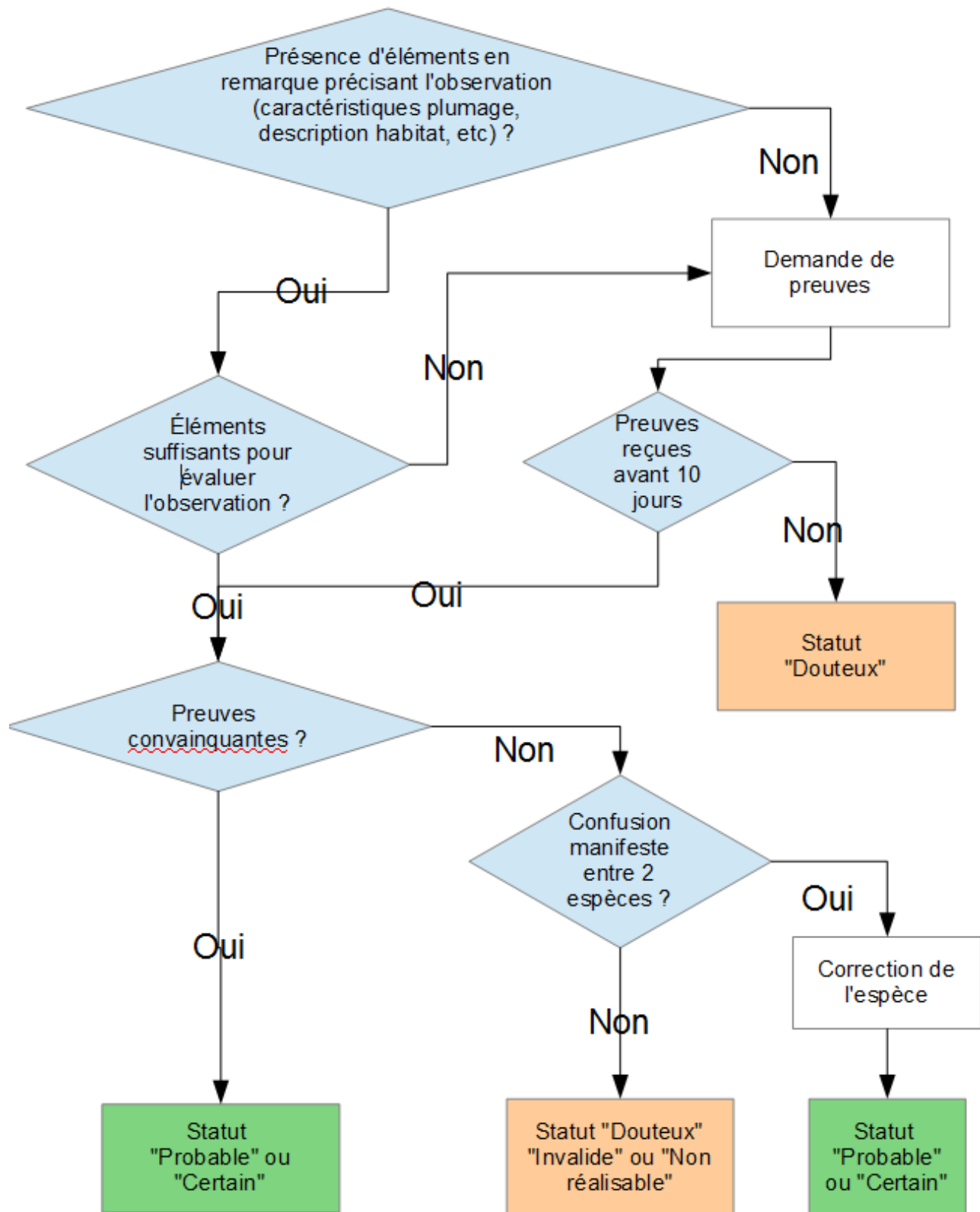


Illustration 11: Logigramme de validation scientifique manuelle des données d'oiseaux

2. Données de référence

La validation automatique s'appuie sur les données de référence suivante :

- **liste des espèces communes nicheuses ou ne posant pas de problèmes d'identification et milieux de prédilection connus** : voir Tableau 27 en page suivante.
- **cartographie des milieux propices à l'observation de *Phaton lepturus*** (paille-en-queue) : il s'agit du milieu marin et des zones terrestres situées à moins de 1000 m de la côte ou à 50 m de part et d'autre d'un cours d'eau permanent ou temporaire. Voir couche SIG en Annexe 12 : Couches SIG de référence
- **cartographie des habitats de prédilection** selon la cartographie utilisée pour le STOC : voir couche SIG en Annexe 12 : Couches SIG de référence

La couche est basée sur la typologie suivante :

id_milieu	code_milieu	descriptif
1	F	végétation d'altitude
2	C	Forêt à <i>Acacia heterophylla</i> (Tamarinaie)
3	C	Forêt tropicale de montagne au vent (et formations pionnières associées)
4	C	Forêt tropicale de montagne sous le vent (et formations pionnières associées)
5	O	Fourrés perhumides à <i>Pandanus</i>
6	B	Forêt tropicale de moyenne altitude au vent (et formations pionnières associées)
7	B	Forêt tropicale de moyenne altitude sous le vent (et formations pionnières associées)
8	G	Forêt de moyenne altitude des fonds de cirque sous le vent
9	G	Forêt semi-sèche
10	K	Forêt tropicale humide de basse altitude (et formations pionnières associées)
11	I	Coulées de lave récentes
12	R	Reliques de végétation littorale indigène et fourrés à <i>vacoas</i> littoraux
13	T	Zones humides
14	J	Végétation des remparts
15	E	Forêt secondaire
16	N	Savane
17	M	Sylviculture
18	S	Sol nu
19	E	Bouquet d'arbres
20	P	Zone en friche
21	H	Agriculture - Diversif
22	D	Canne
23	L	Elevage
24	A	Urbain et periurbain

Tableau 26: Typologie des habitats utilisés pour la validation scientifique des données d'oiseaux

id_faune_reu nion	cd_ref	nom_francais	nom_latin	famille	type_milieu																		
					A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
578	4552	Astrild ondulé	<i>Estrilda astrild</i>	Estrildidae				x	x			x				x		x		x			
1651	432754	Bulbul de la Réunion	<i>Hypsipetes borbonicus</i>	Pycnonotidae		x	x		x		x		x	x	x		x		x				
542	442300	Bulbul orphée	<i>Pycnonotus jocosus</i>	Pycnonotidae	x	x	x	x	x			x	x	x	x	x	x	x	x	x		x	x
1652	432596	Busard de Maillard (*)	<i>Circus maillardi</i>	Accipitridae	Tout milieu terrestre																		
189	2996	Caille des blés	<i>Coturnix coturnix</i>	Phasianidae				x				x		x		x	x			x			
1653	895901	Caille peinte	<i>Coturnix chinensis</i>	Phasianidae	x		x	x	x														
551	442042	Capucin damier	<i>Lonchura punctulata</i>	Estrildidae	x			x	x			x	x			x							
1305	199742	Coq bankiva	<i>Gallus gallus</i>	Phasianidae		x	x	x	x			x			x								
190	3003	Faisan de colchide	<i>Phasianus colchicus</i>	Phasianidae			x									x				x			
1658	418755	Foudi de Madagascar	<i>Foudia madagascariensis</i>	Ploceidae	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
1659	528812	Francolin gris	<i>Francolinus pondicerianus</i>	Phasianidae			x									x	x						
1664	441946	Géopélie zébrée	<i>Geopelia striata</i>	Columbidae	x			x	x			x							x		x		x
1668	442487	Hémipode de Madagascar	<i>Turnix nigricollis</i>	Turnicidae				x	x			x		x		x		x					x
1669	418748	Hirondelle de Bourbon	<i>Phedina borbonica</i>	Hirundinidae	x			x	x		x					x		x					
1001	601183	Martin triste	<i>Acridotheres tristis</i>	Sturidae	x			x	x			x				x		x		x		x	x
493	4525	Moineau domestique	<i>Passer domesticus</i>	Passeridae	x			x	x			x				x		x		x		x	x
1678	442536	Oiseau-lunettes gris	<i>Zosterops borbonicus</i>	Zosteropidae	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
1679	432587	Oiseau-lunettes vert	<i>Zosterops olivaceus</i>	Zosteropidae		x	x		x	x	x			x	x		x		x		x		
1681	528807	Perdrix de Madagascar	<i>Margaroperdix madagarensis</i>	Phasianidae	x			x		x		x				x							
12491	418705	Phaéton à brins blancs (*)	<i>Phaethon lepturus</i>	Phaethontidae	En mer ; Sur terre : à moins de 1000m de la côte ou à 50m de part et d'autre d'un cours d'eau permanent ou temporaire																		
308	3420	Pigeon domestique	<i>Columba livia</i>	Columbidae	x			x				x				x		x		x			x
1697	441613	Salangane des Mascareignes	<i>Aerodramus francicus</i>	Apodidae	x		x	x	x		x	x		x		x		x		x			x
1701	432586	Tarier de la Réunion	<i>Saxicola tectes</i>	Saxicolidae		x	x		x	x	x		x	x	x	x	x		x		x		
1702	442419	Terpsiphone de Bourbon	<i>Terpsiphone bourbonnensis bourbonnensis</i>	Monarchidae		x	x		x		x		x	x	x		x		x		x		
544	4548	Tisserin gendarme	<i>Ploceus cucullatus</i>	Ploceidae	x			x	x			x				x		x		x		x	
1704	459500	Tourterelle malgache	<i>Nesoenas picturata</i>	Columbidae	x	x		x	x			x		x	x	x	x	x		x		x	
1335	528816	Travailleur à bec rouge	<i>Quelea quelea</i>	Ploceidae	x				x							x				x			
553	418757	Veuve dominicaine	<i>Vidua macroura</i>	Viduidae	x																		

(*) Taxons ne posant pas de problèmes d'identification

Tableau 27: Liste des espèces d’oiseaux communes nicheuses à La Réunion ou ne posant pas de problème d’identification et milieux de prédilection associés

- liste des espèces communes migratrices et périodes de présence habituelles :

id_faune_reunion	cd_ref	nom_francais	nom_latin	famille	mois_presence
32	2489	Héron garde-boeufs	<i>Bubulcus ibis</i>	Ardeidae	1,5,6,7,8,9,10,11,12
1157	185992	limicole indéterminé	<i>Charadriiformes sp.</i>	Charadriidae	1,2,3,8,9,10,11,12
1036	3148	Gravelot de Leschenault	<i>Charadrius leschenaultii</i>	Charadriidae	9,10,11,12
1216	190643	Gravelot indéterminé	<i>Charadrius sp.</i>	Charadriidae	9,10,11,12
215	3165	Pluvier argenté	<i>Pluvialis squatarola</i>	Charadriidae	1,2,3,8,9,10,11,12
233	2616	Chevalier guignette	<i>Actitis hypoleucos</i>	Scolopacidae	1,2,3,8,9,10,11,12
236	3239	Tournepieuvre à collier	<i>Arenaria interpres</i>	Scolopacidae	9,10,11,12
1215	190184	Bécasseau indéterminé	<i>Calidris / Limicola / Tryngites sp.</i>	Scolopacidae	9,10,11,12
241	3195	Bécasseau sanderling	<i>Calidris alba</i>	Scolopacidae	9,10,11,12
250	2901	Bécasseau cocorli	<i>Calidris ferruginea</i>	Scolopacidae	9,10,11,12
222	2571	Courlis corlieu	<i>Numenius phaeopus</i>	Scolopacidae	1,2,3,8,9,10,11,12
1241	195290	Courlis indéterminé	<i>Numenius sp.</i>	Scolopacidae	1,2,3,8,9,10,11,12
230	2594	Chevalier aboyeur	<i>Tringa nebularia</i>	Scolopacidae	1,2,3,8,9,10,11,12
1262	198651	Chevalier indéterminé (Tringa)	<i>Tringa sp.</i>	Scolopacidae	1,2,3,8,9,10,11,12

Tableau 28: Liste des espèces migratrices communes d'oiseaux et périodes de présence habituelles à La Réunion

3. Cas particulier des données saisies sur l'outil Faune-Réunion

3.1 Fonctionnement de Faune-Réunion

Faune-Réunion est un **outil de sciences participatives** qui permet de collecter et partager les données saisies à La Réunion par des naturalistes débutants ou experts sur la faune vertébrée. L'outil se matérialise sous la forme d'une interface web accessible à l'adresse www.faune-reunion.fr et de l'application pour smartphone NaturaList. Administré par la SEOR, il compte de nombreux utilisateurs à La Réunion, comme dans toutes les régions françaises où il est développé et contribue activement au SINP 974 via des versements réguliers à Borbonica. Au 21/11/2019, il totalise 45 604 observations, en grande majorité relatives aux oiseaux (90%).

Plusieurs dispositifs de contrôle sont mis en place sur Faune-Réunion :

- **Date de présence** : des contrôles sur la **date de présence sur l'île des migrants** provoquent l'apparition automatique d'un message de demande de confirmation à la fin de la saisie. Ce message est informatif et non bloquant.
- **Observations accompagnées de photos** :
 - si la date de la photo ne correspond pas à celle de l'observation, une marque "Il y a un doute" est automatiquement positionnée, et les validateurs sont avertis par un message.
 - ces photos sont systématiquement vérifiées quelle que soit l'espèce ;
- **Espèces rares, très rares voire jamais observées** : A chaque espèce est associé un **statut de rareté** sur La Réunion (voir plus bas). Lorsqu'une observation d'espèce rare, très rare voire jamais observée est enregistrée :
 - un indicateur particulier mentionne « Donnée à soumettre au CHR » (Comité d'Homologation Régional) et elle est automatiquement positionnée en statut « Donnée probablement incorrecte contacter l'observateur » (on parle de « marque » dans vocabulaire de Faune-Réunion).

- l'observateur est systématiquement contacté par un validateur sauf si l'observation est accompagnée de photos permettant de la valider directement ;
- dans le cas des faucons (concolore / éleonore), s'agissant d'espèces peu faciles à différencier, les observations font en général l'objet d'une vérification approfondie, avec envoi d'un message à l'observateur si les éléments fournis initialement ne sont pas suffisants.
- **Espèces communes** : on peut considérer que les données d'espèces communes ne sont vérifiées que si elles sont accompagnées d'une photographie, ou qu'un élément inhabituel attire l'attention : altitude élevée pour une espèce côtière, loin de l'aire de répartition connue, nombre d'individus important pour une espèce non grégaire, etc.
- **Validation manuelle complémentaire** : les validateurs peuvent à tout moment positionner n'importe quelle observation dans l'une des marques ci-dessous. S'ils ne le font pas, c'est qu'ils considèrent la donnée comme valide. Concrètement :
 - les nouvelles observations (encore jamais visualisées par le validateur concerné) apparaissent dans une liste spécifique accessible à chaque validateur. Cette liste est indépendante des actions des autres validateurs. Comme il n'existe pas de validation explicite dans Faune-Réunion, les mêmes observations sont susceptibles d'être contrôlées, ou du moins visualisées, par 4 ou 5 personnes différentes.
 - depuis la page présentant le détail d'une observation, les validateurs peuvent directement positionner une marque, et écrire un message qui sera envoyé à l'observateur. Tous les validateurs reçoivent alors une copie de ce message, mentionnant le changement de statut.
 - en cas d'erreur sur l'espèce, il est demandé à l'observateur de la modifier, ce qu'il fait habituellement. Le validateur doit manuellement retirer la marque lorsque la donnée a été corrigée. Une observation peut aussi être purement et simplement supprimée s'il s'agit d'une erreur grossière ou d'une intention de nuire.
 - Les différentes marques disponibles pour les validateurs sont les suivantes :
 - La donnée est correcte, demander un complément d'information
 - Il y a un doute sur cette donnée, démarrer une discussion sans marquer la donnée
 - La donnée est probablement incorrecte, demander une modification
 - La donnée est incorrecte, marquer la donnée comme non retenue

5 validateurs valident activement les données sur Faune-Réunion :

- Martin Riethmuller
- Laurent Brillard
- Alexandre Boyer
- Jérôme Dubos
- Damien Chiron

Les vérifications portent essentiellement sur la **détermination de l'espèce**. Dans certain cas, la vérification porte aussi sur le sexe et l'âge (Busard de Maillard, essentiellement), ce qui va plus loin que la validation opérée dans Borbonica.

3.2 Intégration de la validation de Faune-Réunion dans Borbonica

Afin d'éviter des doubles validation, il est intéressant intégrer la validation de Faune-Réunion dans Borbonica. 3 cas de figure sont distingués pour cela, détaillés ci-dessous et synthétisés dans l'illustration 12.

Cas 1 : données marquées

Il s'agit des **données peu communes**, qui ont été vérifiées par un validateur qui a choisi de laisser la marque automatique (ou de la faire évoluer), ou bien qui n'ont pas encore été vérifiées. Le niveau de validité correspondant à la marque est alors indiqué dans Borbonica et la validation est considérée comme manuelle. La correspondance entre les niveaux de Faune-Réunion et Borbonica est indiquée dans le Tableau 29.

Icone pour l'utilisateur	Libellé sur la fiche de l'observation	Libellé dans l'export (champ admin_hidden)	Niveau de validité dans Borbonica
Point d'interrogation bleu	La donnée est probablement incorrecte, demander une modification	Probablement incorrecte	Douteux
Sinusoïde Jaune	La donnée est correcte demander un complément d'information	Incomplet	Douteux
Point d'exclamation rouge	La donnée est incorrecte	Refusée	Invalide

Tableau 29: Correspondance entre les niveaux de validité de Faune-Réunion et Borbonica

Cas 2 : Observations systématiquement vérifiées

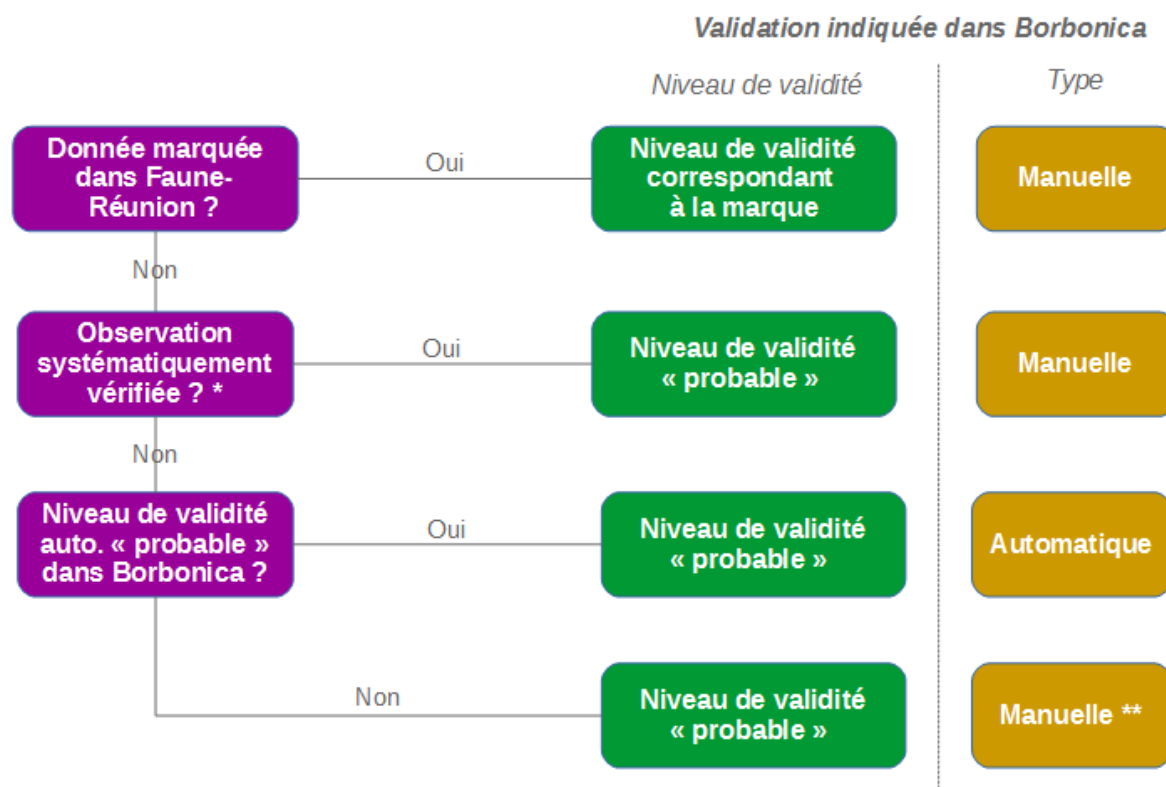
Il s'agit des **observations d'espèces rares, très rares, voire jamais observées, ou des observations accompagnées de photos**. Si les observations ne sont pas marquées, c'est qu'elles ont été examinées et validées par un expert. Le niveau de validité « probable » est alors retenu dans Borbonica et la validation est considérée comme manuelle.

Cas 3 : Observations non systématiquement vérifiées

Deux sous-cas sont identifiés :

- si le résultat de la validation automatique Borbonica aboutit au niveau « probable » : le résultat de la validation automatique est intégralement conservé (niveau, date, commentaires) ;
- si le résultat de la validation automatique Borbonica aboutit à un niveau autre que probable : un niveau « probable » est retenu en se basant sur la procédure Faune-Réunion, et la validation est considérée comme manuelle. Le commentaire de validation automatique est intégré à titre d'information dans le champ commentaire_validation. Cela permet d'enrichir la donnée et, si le validateur le souhaite, de faire quelques vérifications ciblées sur les taxons de son choix.

Dans tous les cas pour les validations retenues comme manuelles, sans éléments plus précis, le validateur indiqué dans Borbonica est « Équipe de validateurs oiseaux de Faune-Réunion » et la date de validation correspond par défaut à la date de transmission du jeu de données.



* Deux cas de figure :

- Espèce rare, très rare ou jamais observée
- Observation avec photo

** Conservation des commentaires de validation automatique pour mémoire

Illustration 12: Résumé des modalités d'intégration de la validation de Faune-Réunion dans Borbonica

3.3 Niveaux de rareté des espèces déclaré dans Faune-Réunion

Les espèces concernées par chaque niveau de rareté sont reprises ci-après :

- **Espèces rares :**

id_faune_re union	cd_nom	cd_ref	nom_francais	nom_latin	famille
	1005	1005	Pétrel de Bulwer	Bulweria bulwerii	Procellariidae
	3140	3140	Grand Gravelot	Charadrius hiaticula	Charadriidae
	3374	3374	Guifette leucoptère	Chlidonias leucopterus	Laridae
	190717	190717	Guifette indéterminée	Chlidonias sp.	Laridae
1657	441814	895898	Echenilleur de La Réunion	Coracina newtoni	Campephagidae
	199413	199413	Corbeau familier	Corvus splendens	Corvidae
1392	199343	199343	Faucon concolore	Falco concolor	Falconidae
175	2681	2681	Faucon d'Eléonore	Falco eleonora	Falconidae
1192	192519	192519	Faucon indéterminé	Falco sp.	Falconidae
1660	418704	418704	Frégate ariel	Fregata ariel	Fregatidae
1661	418703	418703	Frégate du Pacifique	Fregata minor	Fregatidae
1529	192624	192624	Frégate indéterminée	Fregata sp.	Fregatidae
	186040	186040	Océanite indéterminé	Hydrobatidae sp.	Hydrobatidae
1084	1039	1039	Océanite de Wilson	Oceanites oceanicus	Oceanitidae
	532691	532691	Sterne bridée	Onychoprion anaethetus	Laridae
12490	432685	432685	Phaéton à brins rouges	Phaethon rubricauda	Phaethontidae
	3258	3258	Labbe parasite	Stercorarius parasiticus	Stercorariidae
	418726	581178	Sterne huppée	Sterna bergii	Laridae
293	3343	3343	Sterne pierregarin	Sterna hirundo	Laridae
	2610	2610	Chevalier bargette	Xenus cinereus	Scolopacidae

- **Espèces très rares :**

id_faune_re union	cd_nom	cd_ref	nom_francais	nom_latin	famille
	4557	4557	Bengali rouge	Amandava amandava	Estrildidae
	1975	836222	Sarcelle d'été	Anas querquedula	Anatidae
	418710	418710	Crabier de Madagascar	Ardeola idae	Ardeidae
	199317	199317	Pétrel de Jouanin	Bulweria fallax	Procellariidae
	3231	3231	Bécasseau à queue pointue	Calidris acuminata	Scolopacidae
	3192	3192	Bécasseau maubèche	Calidris canutus	Scolopacidae
	3226	3226	Bécasseau tacheté	Calidris melanotos	Scolopacidae
243	3206	3206	Bécasseau minute	Calidris minuta	Scolopacidae
	459627	459627	Guifette moustac	Chlidonias hybrida	Laridae
	3371	3371	Guifette noire	Chlidonias niger	Laridae
	528809	528809	Coucou de Madagascar	Cuculus rochii	Cuculidae
	1002	1002	Damier du Cap	Daption capense	Procellariidae
	549574	442421	Albatros à cape blanche	Diomedea cauta	Diomedeidae
	186036	186036	Albatros indéterminé	Diomedeidae sp.	Diomedeidae
1656	418721	418721	Drome crabier	Dromas ardeola	Dromadidae
	528811	528811	Aigrette dimorphe	Egretta dimorpha	Ardeidae
	440194	440194	Rollier malgache	Eurystomus glaucurus	Coraciidae
	441952	441952	Glaréole de Madagascar	Glareola ocularis	Glareolidae
	192799	192799	Glaréole indéterminée	Glareola sp.	Glareolidae
1667	440187	440187	Gygis blanche	Gygis alba	Laridae
	442022	442022	Goéland dominicain	Larus dominicanus	Laridae
	193860	193860	Goéland indéterminé	Larus sp. (magna)	Laridae
226	2568	2568	Barge rousse	Limosa lapponica	Scolopacidae
	994	994	Pétrel géant	Macronectes giganteus	Procellariidae
	253474	253474	Pétrel géant subantarctique	Macronectes halli	Procellariidae
13120	199322	199322	Fou du Cap	Morus capensis	Sulidae
	3741	3741	Bergeronnette printanière	Motacilla flava	Motacillidae
	2576	2576	Courlis cendré	Numenius arquata	Scolopacidae
			Courlis cendré (N.a.orientalis)	Numenius arquata orientalis	
356	3803	3803	Loriot d'Europe	Oriolus oriolus	Oriolidae
	2415	2415	Océanite frégate	Pelagodroma marina	Oceanitidae
	2534	814245	Combattant varié	Philomachus pugnax	Scolopacidae
	442280	442280	Pétrel noir de Bourbon	Pseudobulweria atterima	Procellariidae
	3261	3261	Labbe à longue queue	Stercorarius longicaudus	Stercorariidae
297	3361	3359	Sterne voyageuse	Sterna bengalensis	Laridae
	3364	3364	Sterne de Dougall	Sterna dougallii	Laridae
	199376	528761	Sterne royale	Sterna maxima	Laridae
	418728	625683	Sterne de Saunders	Sterna saundersi	Laridae
1144	199320	199320	Fou masqué	Sula dactylatra	Sulidae
	440197	440197	Fou à pieds rouges	Sula sula	Sulidae
	186049	186049	Fou indéterminé	Sulidae sp.	Sulidae
	713021	713021	Albatros à bec jaune	Thalassarche carteri	Diomedeidae
	442424	442424	Albatros à sourcils noirs	Thalassarche melanophris	Diomedeidae
	2607	2607	Chevalier sylvain	Tringa glareola	Scolopacidae

- **Espèces jamais observées :**

id_faune_reunion	cd_nom	cd_ref	nom_francais	nom_latin	famille
	886117	886117	Puffin cendré	Calonectris borealis	Procellariidae
			Râle de Cuvier	Dryolimnas cuvieri	
	2679	2679	Faucon hobereau	Falco subbuteo	Falconidae
	3696	3696	Hirondelle rustique	Hirundo rustica	Hirundinidae
	2419	2419	Océanite tempête	Hydrobates pelagicus	Hydrobatidae
			Faisan argenté	Lophura nycthemera	
	418713	418713	Milan noir (parasitus)	Milvus migrans parasitus	Accipitridae
	442172	442172	Prion de la Désolation	Pachyptila desolata	Procellariidae
	2522	2522	Ibis falcinelle	Plegadis falcinellus	Threskiornithidae
	442268	442268	Pétrel gris	Procellaria cinerea	Procellariidae
	418712	418712	Canard à bosse	Sarkidiornis melanotos	Anatidae
			Espèce non listée	Species ignota	

Annexe 10 : Validation scientifique des données de poissons et macro-crustacés d'eau douce

1. Organisation mise en place

Les différentes étapes de la **validation automatique** sont indiquées dans l'illustration 13. Elle repose sur la succession des tests suivants :

- observation d'un taxon dans la **liste des taxons décrits par Taxref** ;
- observation d'un taxon dans la **liste des taxons signalés à La Réunion** ;
- observation d'un **taxon dans une zone de présence avérée ou probable** ;
- taxon dont l'**identification est complexe**, au moins à certains stades ;
- présence de **preuve disponible**.

L'utilisation de filtres additionnels, notamment sur la phénologie des taxons ou leurs habitats préférentiels, n'a pas été employée au vu des connaissances restreintes sur les espèces concernés. Le renseignement des données d'inventaire dans Borbonica permettra, à l'inverse, de recueillir un maximum d'informations afin d'alimenter les connaissances liées aux taxons et de mettre en place de nouveaux filtres à terme.

Un filtre additionnel relatif aux gammes altitudinales, pour lequel l'état des connaissances est davantage étayé, est intégré dans l'approche cartographique sur les zones de présence avérée ou probable.

A l'issue de la phase automatique, la **validation manuelle** pourra affiner le niveau de validité attribué via l'**examen de preuves** (photographies, description détaillée, séquençage moléculaire). En revanche, les observations ne disposant pas de preuves seront qualifiées de « non réalisables », l'opérateur ne pouvant statuer sur le niveau de fiabilité faute d'éléments suffisants.

A noter que pour le moment aucun pôle thématique n'étant en place sur la thématique de la faune d'eau douce, aucune donnée ne pourra faire l'objet d'une validation manuelle.

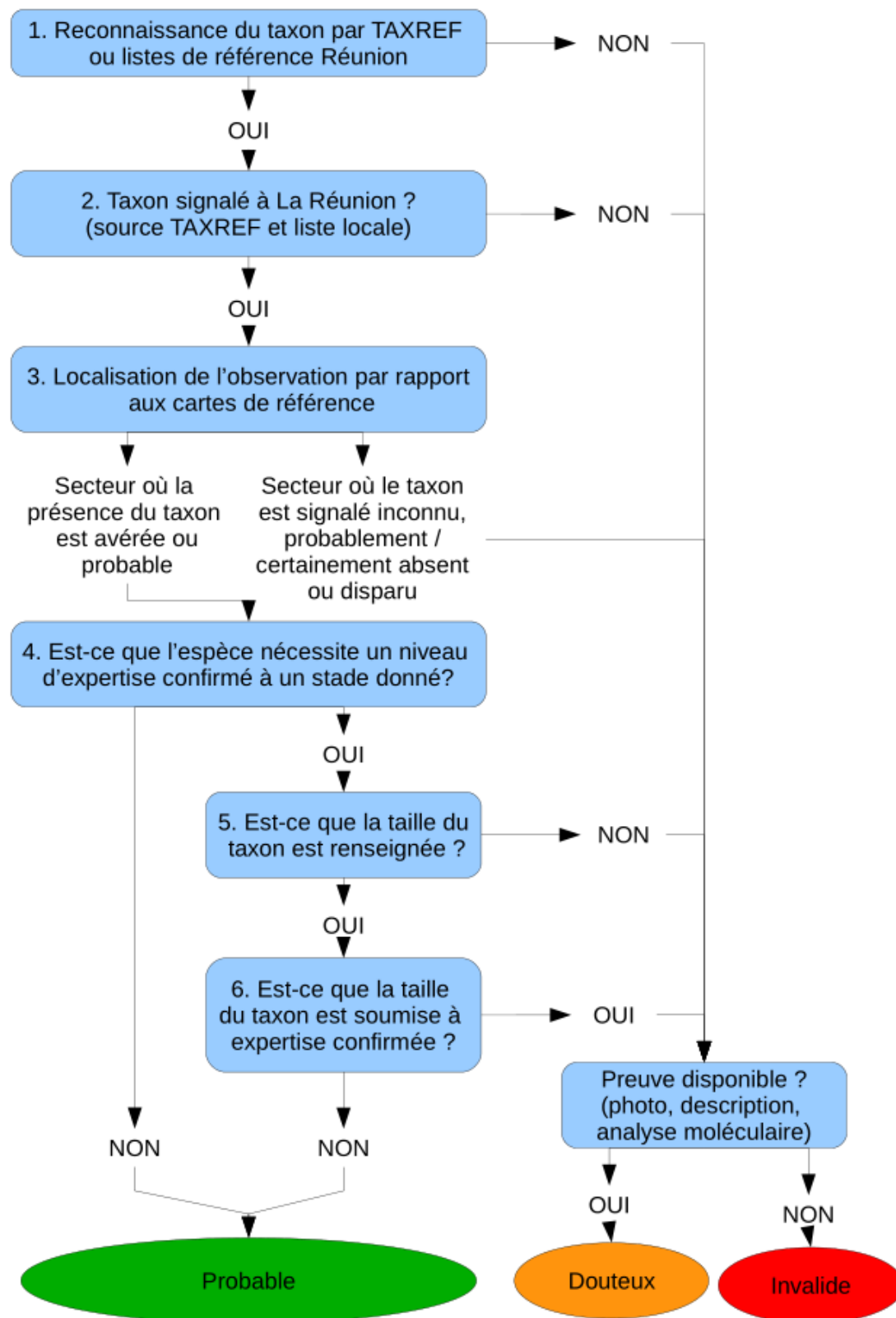


Illustration 13 : Logigramme de validation scientifique automatique des données de poissons et macro-crustacés d'eau douce

2. Données de référence

La validation automatique s'appuie sur les données de référence suivante :

- **liste des taxons de crustacés et de poissons présents ou ayant été présents à La Réunion et leur statut biogéographique** (source : Taxref v11.0 et dire d'expert) : voir Tableau 30 et Tableau 31 pages suivantes.

Cette liste a été constituée à partir de la liste des familles, espèces et sous-espèces répertoriées à La Réunion par Taxref, indigènes et exotiques (statuts biogéographiques P, B, E, S, C, I, J ou M), vivant au moins une partie de leur cycle en eau douce. Elle a été complétée à dire d'expert par des taxons indigènes référencés avec un nom valide erroné dans TAXREFv11 et par des taxons exotiques encore non identifiés dans TAXREF, bien qu'observés à La Réunion. Cela permet de devancer les évolutions de TAXREF sur la prise en compte de taxons nouvellement découverts par exemple. On aboutit in fine à une liste de **49 taxons de crustacés et 103 taxons de poissons**.

Ces ajouts se basent principalement sur un travail de **veille scientifique** permettant de disposer des dernières versions de l'état des connaissances sur les peuplements concernés, les données issues de Borbonica (jeux de données DEAL) compilant une partie des observations réalisées par les opérateurs œuvrant en eau douce à La Réunion, la base de données acquise par OCEA au fil de ses missions d'inventaires, ainsi que sur des échanges fréquents avec l'expertise locale (FDPPMA, Hydrô Réunion, Groupe Espèces Invasives de La Réunion, etc).

- **Cartes de référence sur la distribution des taxons**

Ces cartes ont été établies pour les principaux taxons (espèces ou familles : 11 crustacés et 30 poissons référencés dans le Tableau 30 et le Tableau 31 pages suivantes) à partir d'éléments de connaissance sur la distribution des taxons (linéaires de cours d'eau) croisés à un maillage du territoire (carrés de 250 m de côté). Pour les autres taxons, souvent rares et difficiles d'identification, ces cartes pourront être produites ultérieurement, sous réserve d'amélioration de la connaissance de leur taxonomie et de leur distribution. Pour chaque maille est ainsi renseigné le code de présence (Robert et al. 2017) :

- 1 : présence inconnue. Il s'agit de portions de cours d'eau où il n'existe pas suffisamment de données d'inventaire pour statuer sur la présence ou l'absence du taxon. Cette valeur a été donnée par défaut pour les taxons les plus communs à La Réunion ;
- 2 : Absence probable ou certaine. Il s'agit des portions de cours d'eau où un minimum de 4 inventaires ont été réalisés sans que le taxon ait été observé. Pour les taxons nouvellement introduits, comme la caridine *Neocaridina sp.*, l'ensemble des cours d'eau en dehors de la zone où elle a été observée a été noté « 2 ». De même, pour certains taxons à très fortes affinités marines, l'absence probable ou certaine a été renseignée pour les cours d'eau où elle n'est pas avérée ou supposée présente ;
- 3 : Disparition avérée. Un seul taxon a été renseigné avec ce code : l'écrevisse *M. hirtimanus* pour laquelle il existe des observations anciennes ;
- 4 : Présence probable. La présence probable a été renseignée à « dire d'expert » au cas par cas pour guider la validation automatique des données et pour pallier, dans la mesure du possible, aux manques de données d'inventaires,
- 5 : Présence avérée. Il s'agit de tronçons de cours d'eau où le taxon a été observé au cours de ces 15 dernières années.

Les principaux éléments utilisés pour décrire la distribution des taxons sont :

- Keith, P., Vigneux, E., & Bosc, P. (1999). *Atlas des poissons et des crustacés d'eau douce de la Réunion*. (P. Muséum national d'Histoire naturelle, Ed.) (Patrimoine).
- Keith, P., G. Marquet, P. Valade, P. Bosc, et E. Vigneux. 2006. *Atlas des poissons et des crustacés d'eau douce des Comores, Mascareignes et Seychelles*. Muséum national d'Histoire Naturelle, Paris,

Collection Patrimoines Naturels, 65.

- DEAL, en cours, *Etude en vue de la protection des espèces de poissons et de crustacés des eaux douces de La Réunion*.
- DEAL, *Comptes-rendus des inventaires par pêche électrique en rivière ou plan d'eau à La Réunion*.
- OCEA, non publié, *Compilation de données d'inventaires des poissons et des macro-crustacés d'eau douce de La Réunion*.

Ces cartes ne constituent pas des données de référence sur la distribution des taxons de poissons et de macro-crustacés, mais bien un premier bilan des connaissances sur la distribution des taxons, dans le cadre de la démarche de validation automatique des données taxonomiques. **Ces éléments cartographiques devront évoluer** à partir de la compilation des données d'observation dans Borbonica.

Dans cette première version les observations de reproduction (pontes, frai) n'ont pas été prises en compte. Complexes à décrire et encore méconnues, ces observations devront faire l'objet d'une validation manuelle.

famille	nom_valide	nom_vern	cd_nom	habitat	reu	carte	identification_complexe
Atyidae	Atyoida pilipes (Newport, 1847)	Crevette, saltarelle koros	680116	4*	P*		tous stades
Atyidae	Atyoida serrata (Spence Bate, 1888)	Crevette Bouledogue	418785	4*	P	Oui	<15mm
Atyidae	Caridina longirostris H. Milne Edwards, 1837	Chevaquine	418787	4*	P		<15mm
Atyidae	Caridina nilotica (Roux, 1833)	Chevaquine	535062	4*	P		<15mm
Atyidae	Caridina serratirostris De Man, 1892		432658	4*	P	Oui	<15mm
Atyidae	Caridina typus H. Milne Edwards, 1837	Chevaquine	418786	4*	P	Oui	<15mm
Atyidae	Neocaridina heteropoda	Crevette naine (var red cherry)	-67	2*	I*	Oui	tous stades
Palaemonidae	Macrobrachium australe (Guerin-Meneville, 1838)	Chevrette, grand bras	418790	4*	P	Oui	<25mm
Palaemonidae	Macrobrachium hirtimanus (Olivier, 1811)	Chevrette lecroc, Chevrette gros crocs, Ecrevisse, Bouquet malais	528800	4*	W	Oui	<25mm
Palaemonidae	Macrobrachium lar (Fabricius, 1798)	Camaron, chevrette, bouquet singe	418788	4*	P	Oui	<25mm
Palaemonidae	Macrobrachium lepidactyloides (de Man, 1892)	bouquet Malais	432662	4*	D		<25mm
Palaemonidae	Macrobrachium lepidactylus (Hilgendorf, 1879)	Ecrevisse	418789	4*	P	Oui	<25mm
Palaemonidae	Macrobrachium rosenbergii (de Man, 1879)	Chevrette d'elevage (La), Bouquet geant (Le), Chevrette (La), Crevette tropicale d'eau douce (La), Crevette geante d'eau douce (La)	538982	4*	Y*		<25mm
Palaemonidae	Palaemon concinnus (Dana, 1852)	Crevette, Chevrette	418791	4*	P	Oui	<25mm
Parastacidae	Cherax quadricarinatus Von Martens, 1868	Ecrevisse a pinces rouges	528798	4*	M		<25mm
Portunidae	Scylla serrata (Forsk., 1775)	Crabe de mangrove	432665	4*	P	Oui	<25mm
Potamidae	Potamon bouvieri Rathbun, 1904	Crabe d'eau douce	-95	2*	D*		tous stades
Sesarmidae	Parasarma plicatum (Latreille, 1803)		595657	4	P		tous stades
Varunidae	Ptychognathus hachijoensis Sakai, 1955		595346	4	I		tous stades
Varunidae	Varuna litterata (Fabricius, 1798)	Crabe d'eau douce	418793	4	P	Oui	<20mm

Habitat : codes issus de Taxref (2 = eau douce ; 4 = marin et eau douce)

reu : une * indique une valeur à dire d'expert, différente de Taxref v11.0

Tableau 30: Liste des taxons de crustacés d'eau douce présents ou ayant été présents à La Réunion

famille	nom_valide	nom_vern	cd_nom	habitat	reu	carte	identification_complexe
Ambassidae	Ambassis ambassis (Lacepède, 1802)	Ambache du large, Ambasse	418771	4	P		tous stades
Ambassidae	Ambassis gymnocephalus (Lacepède, 1802)	Ambache du large	423148	4	P		tous stades
Ambassidae	Ambassidae Klunzinger, 1870		418434	4	P	Oui	tous stades
Ambassidae	Ambassis natalensis (Gilchrist & Thompson, 1908)	Poisson d'argent	418772	4	D*		tous stades
Ambassidae	Ambassis urotaenia Bleeker, 1852	Ambache du large, Ambasse	528790	4	P		tous stades
Anguillidae	Anguilla bengalensis labiata (Peters, 1852)	Anguille marbree africaine	426003	4	P		tous stades
Anguillidae	Anguilla bicolor bicolor McClelland, 1844	Anguille bicolore	418766	4	P	Oui	non
Anguillidae	Anguilla marmorata Quoy & Gaimard, 1824	Anguille marbree	418765	4	P	Oui	<150mm
Anguillidae	Anguilla mossambica (Peters, 1852)	Anguille du Mozambique	418767	4	P	Oui	<150mm
Carangidae	Caranx sexfasciatus Quoy & Gaimard, 1825	Carangue vorace	419256	4	P		tous stades
Centrarchidae	Micropterus salmoides (Lacepede, 1802)	Achigan a grande bouche, Black-bass a grande bouche, Perche d'Amerique, Perche noire, Perche truite, Perche truitee	69346	2	Y*		tous stades
Chanidae	Chanos chanos (Forsskal, 1775)	Chano	423547	4	P	Oui	tous stades
Cichlidae	Amatitlania nigrofasciata (Gunther, 1867)		528788	2	I	Oui	non
Cichlidae	Oreochromis Gunther, 1889		200263	1	P	Oui	tous stades
Cichlidae	Oreochromis macrochir (Boulenger, 1912)		535102	2	I		tous stades
Cichlidae	Oreochromis mossambicus (Peters, 1852)	Tilapia du Mozambique	419269	4	I		tous stades
Cichlidae	Oreochromis niloticus (Linnaeus, 1758)	Tilapia du Nil, Tilapia, Lapia	200266	2	I		tous stades
Cichlidae	Parachromis managuensis (Gunther, 1867)	Managuense	528794	2	I	Oui	non
Cichlidae	Tilapia zillii (Gervais, 1848)		425883	4	I		tous stades
Cyprinidae	Carassius auratus (Linnaeus, 1758)	Carassin dore, Poisson rouge	67208	2	I	Oui	non
Cyprinidae	Ctenopharyngodon idella (Valenciennes, 1844)	Amour blanc (L'), Carpe amour, Carpe herbivore, Chinoise	67246	2	I		tous stades
Cyprinidae	Cyprinus carpio Linnaeus, 1758	Carpe commune, Carpat, Carpeau, Escarpo, Kerpaille	67058	2	I	Oui	non
Cyprinidae	Tanichthys albonubes Lin, 1932		698292	2	I	Oui	non
Eleotridae	Butis butis (Hamilton, 1822)	Cabot noir	418774	4	B		tous stades
Eleotridae	Eleotris acanthopoma Bleeker, 1853	Cabot noir	533446	4	P*	Oui	<40mm
Eleotridae	Eleotris fusca (Forster, 1801)	Cabot noir	418775	4	P	Oui	<40mm
Eleotridae	Eleotris klunzingerii Pfeffer, 1893	Cabot noir	562828	5	P	Oui	<40mm
Eleotridae	Eleotris mauritiana Bennett, 1832	Cabot noir	586405	4	P	Oui	<40mm
Eleotridae	Hypseleotris cyprinoides (Valenciennes, 1837)	Eleotris cyprin	418777	4	B*		tous stades
Eleotridae	Ophiocara porocephala (Valenciennes, 1837)		424994	4	D		tous stades
Esocidae	Esox lucius Linnaeus, 1758	Brochet	67606	2	Y*		non
Gobiidae	Awaous commersoni (Schneider, 1801)	Loche	418779	4	P	Oui	<30mm
Gobiidae	Cotylopus acutipinnis Guichenot, 1863	Cabot de cascade (adulte), Bichique, Cabot bouche ronde	418903	4	S	Oui	<30mm
Gobiidae	Glossogobius giuris (Hamilton, 1822)	Loche, Gobie	419271	4	P		tous stades
Gobiidae	Glossogobius kokius (Valenciennes, 1837)		424268	4	S		tous stades
Gobiidae	Sicyopterus fasciatus (Day, 1874)		439162	NA*	P*		tous stades
Gobiidae	Sicyopterus lagocephalus (Pallas, 1770)	Cabot bouche-ronde, Cabot a tete de lievre (adulte), Bichique (jeune)	418783	4	P	Oui	<40mm
Gobiidae	Stenogobius polyzona (Bleeker, 1867)	Gobie raye, Cabot raye	418784	4	P	Oui	<30mm
Kuhliidae	Kuhlia caudavittata (Lacepede, 1802)	Doule a queue rayee, Poisson pavillon des Mascareignes	424606	4	P		<50mm
Kuhliidae	Kuhlia rupestris (Lacepede, 1802)	Poisson plat, Doule de roche	418773	4	P	Oui	tous stades
Kuhliidae	Kuhlia sauvagii Regan, 1913	Poisson plat, Doule de roche	559719	4	P	Oui	tous stades
Locariidae	Pterygoplichthys pardalis (Castelnau, 1855)	Pleco léopard, Poisson-chat	-8	2*	I*	Oui	non
Lutjanidae	Lutjanus argentimaculatus (Forsskal, 1775)	Vivaneau des mangroves	424726	4	P		tous stades
Lutjanidae	Lutjanus fulvus (Forster, 1801)	Vivaneau queue noire	419260	4	P		tous stades
Megalopidae	Megalops cyprinoides (Broussonet, 1782)	Tarpon	424810	4	P	Oui	tous stades
Mugilidae	Agonostomus telfairii Bennett, 1832	Chitte	418900	4	P	Oui	<50mm
Mugilidae	Liza melinoptera (Valenciennes, 1836)	Mulet otomebora	424719	4	P	Oui	tous stades
Mugilidae	Liza vaigiensis (Quoy & Gaimard, 1825)	Mulet mopiro	424723	4	P	Oui	tous stades
Mugilidae	Moolgarda seheli (Forsskal, 1775)	Mulet a tache bleue, Mulet, Muge	560256	4	P	Oui	tous stades
Mugilidae	Mugil cephalus (Linnaeus, 1758)	Mulet a grosse tete, Muge, Mulet cabot	69772	4	D*	Oui	tous stades
Mugilidae	Valamugil buchanani (Bleeker, 1853)	Mulet, Muge	418769	4	P	Oui	tous stades
Mugilidae	Valamugil cunnesius (Valenciennes, 1836)	Mulet, Muge	425955	4	P	Oui	tous stades
Mugilidae	Valamugil robustus (Gunther, 1861)	Mulet, Muge	425957	4	P	Oui	tous stades
Muraenidae	Strophidon sathete (Hamilton, 1822)	Murene fil geante	425809	4	P		tous stades
Ophichthidae	Pisodonophis cancrivorus (Richardson, 1848)	Serpenton a longue nageoire	425199	4	P		tous stades
Ophichthidae	Yirkkala tenuis (Gunther, 1870)		418898	4	P		tous stades
Osphronemidae	Osphronemus goramy (Lacepède, 1801)	Gourami geant, Gourami	162673	4	Y		tous stades
Osphronemidae	Trichopodus trichopterus (Pallas, 1770)	Gourami bleu	784805	2	I		tous stades
Poeciliidae	Gambusia holbrooki Girard, 1859	Gambusie	68827	4	I		tous stades
Poeciliidae	Poecilia reticulata Peters, 1859	Guppy	68823	2	J	Oui	non
Poeciliidae	Xiphophorus hellerii Heckel, 1848	Mailleur, Xipho, Porte-epée	425982	2	J	Oui	non
Poeciliidae	Xiphophorus maculatus (Gunther, 1866)	Platy	425983	2	I		non
Salmonidae	Oncorhynchus mykiss (Walbaum, 1792)	Truite arc-en-ciel	67804	4	J	Oui	non
Syngnathidae	Microphis argulus (Peters, 1855)	Syngnathe	418768	4	P*	Oui	tous stades
Syngnathidae	Microphis millepunctatus (Kaup, 1856)	Syngnathe	560582	4	P	Oui	tous stades

Habitat : codes issus de Taxref (2 = eau douce ; 4 = marin et eau douce)

reu : une * indique une valeur à dire d'expert, différente de Taxref v11.0

Tableau 31: Liste des taxons de poissons d'eau douce présents ou ayant été présents à La Réunion (source : Taxref v11.0 et dire d'expert)

- **Liste des taxons dont l'identification est complexe** : voir Tableau 30 et Tableau 31 pages précédentes.

La taxonomie des taxons de poissons et de crustacés de La Réunion a significativement évolué ces dernières années, avec la redécouverte de taxons historiquement invalidés. Des progrès sont encore attendus sur la validation taxonomique de ces taxons avec l'utilisation des outils moléculaires. Enfin, des taxons restent définitivement complexes à identifier, comme par exemple les mulets. Pour ces taxons, un niveau d'investigation morphologique ou moléculaire doit être mis en œuvre pour que l'identification puisse être validée .

Aussi, il est proposé de poser un filtre sur la validation automatique des taxons, en retirant les taxons dont la détermination est complexe ou douteuse. Le filtre peut concerner un taxon tous stades confondus, ou bien pour des tailles correspondant aux stades post-larvaires ou juvéniles pouvant nécessiter un niveau d'expertise pour la détermination.

Pour ces taxons, si la taille indiquée par la donnée correspond à un stade d'identification complexe, ou bien si la taille n'est pas renseignée dans la donnée, une validation manuelle sera nécessaire, sous réserve de disposer de preuves (photographie, description détaillée, analyse moléculaire...).

Annexe 11: Validation scientifique des données de flore vasculaire

1. Organisation mise en place par la tête de réseau

1.1 Procédure de validation automatique

Les différentes étapes de la procédure de validation automatique sont indiquées dans la Figure 1. Cette procédure fait appel à trois niveaux d'analyse permettant de vérifier la cohérence de la donnée soumise à validation avec la connaissance déjà acquise sur le taxon observé :

- La cohérence territoriale permet de vérifier si le taxon a déjà été observé à La Réunion et est donc inclus dans le référentiel taxonomique local et dans Borbonica ;
- La cohérence écologique permet de vérifier si le taxon est bien dans sa gamme d'habitats. Dans le cas où la gamme d'habitat n'est pas renseignée pour le taxon, l'observation passe directement en validation manuelle ;
- La difficulté d'identification du taxon qui est évaluée d'une part par le niveau d'expertise de l'observateur et d'autre part par une liste de taxons dont l'identification est considérée comme difficile. Le niveau d'expertise de l'observateur permet de reconnaître les acquis de l'expérience des contributeurs et de faciliter le processus de validation lorsqu'un observateur a la capacité d'identifier certains taxons. Un observateur est considéré comme expert pour un taxon donné lorsqu'il a déjà réalisé 5 observations valides (niveau « probable » ou « certain ») sur ce taxon ;

Ainsi, à l'issue des deux premiers niveaux d'analyse, si l'observation est cohérente d'un point de vue territorial et écologique, alors l'observation peut être qualifiée de « certaine » si au moins un des observateurs est expert pour le taxon donné, sinon l'observation est qualifiée de « douteuse » si le taxon appartient à la liste des taxons difficiles à identifier ou « probable » s'il n'appartient pas à cette liste.

Dans les cas où l'observation est qualifiée de « douteuse » à l'issue de la validation automatique, c'est le processus de validation manuelle qui pourra éventuellement la requalifier en fonction des critères présentés au chapitre suivant.

1.2 Procédure de validation manuelle

Les différentes étapes de la validation manuelle sont indiquées dans la Figure 1. Cette procédure diffère selon l'étape à laquelle la donnée quitte la validation automatique :

- **Lorsque l'observation ne présente pas de cohérence territoriale**, c'est-à-dire que le taxon observé n'est pas inscrit dans le référentiel taxonomique ou qu'il n'est pas encore représenté dans le SINP régional, la validation manuelle consiste à rechercher **l'existence de preuves suffisantes pour la détermination du taxon**. Pour cela, l'observateur doit pouvoir fournir une part d'herbier ou un échantillon comprenant toutes les parties nécessaires à l'identification du taxon. Dans ce cas, si l'identification initiale est correcte, le niveau de validité devient « certain » et le taxon est alors intégré à l'index des trachéophytes de La Réunion ; si l'identification initiale n'est pas correcte, le niveau de validité devient « invalide ». Si l'observateur ne peut pas fournir de part d'herbier ou d'échantillon mais seulement des photos de l'observation, alors le niveau de validité devient « non réalisable » car le taxon ne peut pas être ajouté au référentiel taxonomique sans une part d'herbier permettant l'identification certaine du taxon ;
- **Lorsque l'observation ne présente pas de cohérence écologique**, c'est-à-dire que le taxon observé est hors des étages de végétation dans lesquels il a déjà été observé, la validation manuelle consiste dans un premier temps à rechercher **l'existence de preuves suffisantes pour la détermination du taxon** de la même manière que dans le cas précédent. Dans un second temps, en cas d'absence de preuves suffisantes pour l'identification du taxon, il s'agit de vérifier à la fois **la cohérence géographique à l'échelle de la maille kilométrique** (présence d'une observation valide du taxon dans la maille d'observation ou dans l'une des mailles adjacentes) et **le niveau d'expérience de l'observateur** (au moins 5 observations valides de l'observateur sur le même taxon). Si l'une des deux conditions est remplie, le niveau de validité de l'observation sera « probable ». Si aucune de ces conditions est remplie, le niveau restera « douteux » ;

- Enfin, lorsque l'observation est réalisée par un observateur non expert pour le taxon et que ce taxon est considéré comme difficile à identifier, la validation manuelle suit la même procédure que dans le cas précédent (pas de cohérence écologique).

L'ensemble de la validation manuelle est réalisé par le CBNM en tant que tête de réseau flore pour le SINP 974. Cependant, le CBNM se laisse la possibilité de consulter des experts botanistes issus d'autres structures notamment afin d'associer des compétences spécifiques sur certains groupes taxonomiques mais aussi afin de renforcer et de mutualiser les compétences de l'ensemble des experts. Dans un premier temps, cela pourra prendre la forme de consultations opportunistes, mais il est aussi possible d'envisager de réunir régulièrement (1 à 2 fois par an) un comité de validateurs afin d'examiner les observations les plus difficiles à traiter.

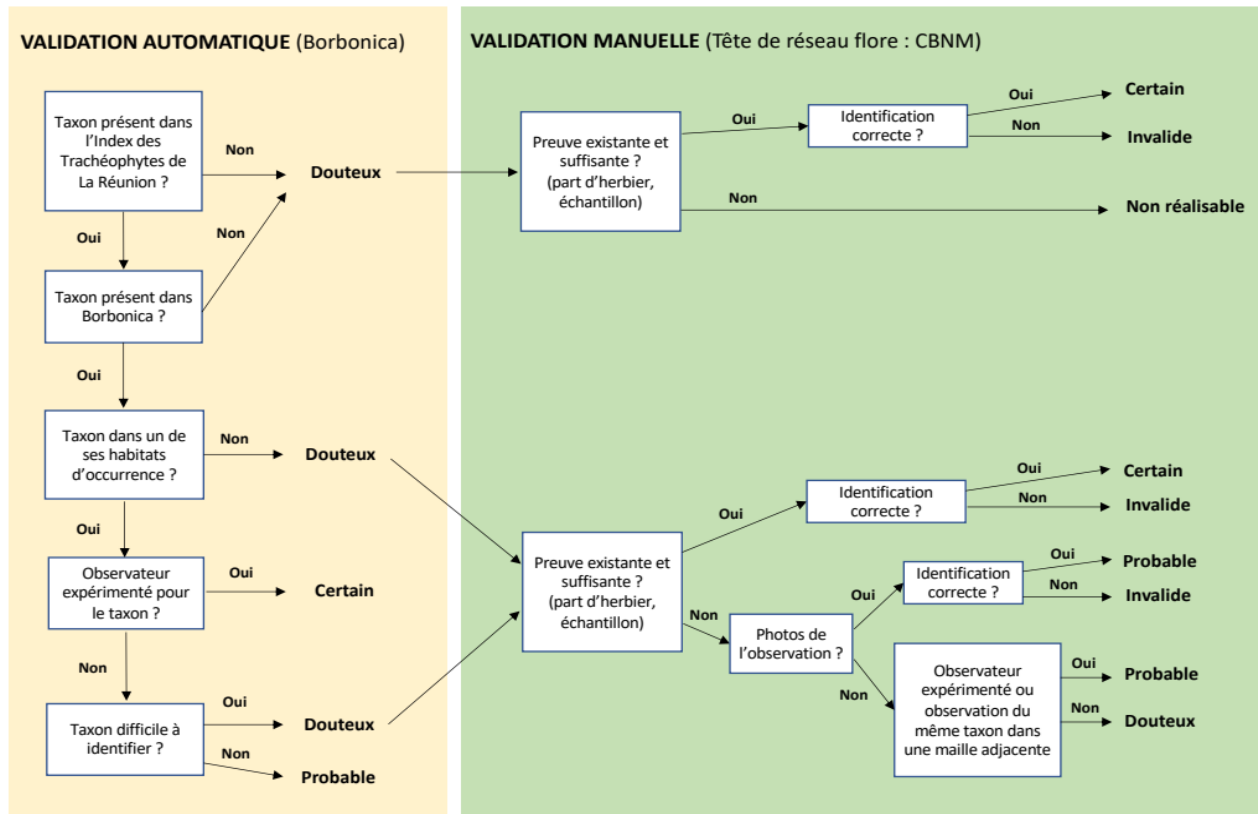


Illustration 14: Logigramme de la validation scientifique automatique et manuelle des données flore du SINP 974

2. Données de référence

La validation automatique s'appuie sur les documents de référence suivants :

- L'index commenté de la flore vasculaire (Trachéophytes) de La Réunion (Boullet 2020);
- Liste des grands étages de végétation connus pour chaque taxon à dire d'expert en se basant sur les observations déjà validées dans Mascarine ;
- Liste de taxons posant des difficultés d'identification (966 taxons).

Ces données de référence ont vocation à évoluer au fil du temps, en incrémentant les nouvelles observations valides. Cela permettra d'améliorer pour chaque taxon la connaissance de la répartition altitudinale et selon les grandes séries de végétation.

De même, le référentiel taxonomique évoluant régulièrement avec l'acquisition de connaissances et les changements nomenclaturels ou taxonomiques éventuels, un travail de revalidation manuelle pourra être nécessaire dans certains cas. En tant que tête de réseau SINP et producteur du référentiel taxonomique régional pour la flore, le CBNM sera chargé de réaliser ce travail de revalidation lorsque les évolutions taxonomiques le nécessiteront.

3. Aspects réglementaires

Le protocole de validation présenté dans ce document s'appuie dans certains cas sur l'existence de preuves suffisantes à l'identification du taxon. Ces preuves se basent sur le prélèvement d'échantillons sur des spécimens sauvages. Il est donc important de rappeler ici que tout prélèvement de tout ou partie d'un végétal nécessite d'une part l'accord du propriétaire du terrain sur lequel il se situe, et d'autre part d'une dérogation spécifique dans le cas d'un taxon protégé par l'arrêté ministériel du 27 octobre 2017 (238 taxons concernés).

Pour plus d'informations sur la réglementation, il est possible de se rapprocher de contacter de consulter le site internet de la DEAL de La Réunion à cette adresse :

<https://www.reunion.developpement-durable.gouv.fr/flore-protgee-de-la-reunion-a962.html>

4 Aspects réglementaires

Cette annexe constitue une première version du protocole de validation des données flore du SINP 974.

Dans un premier temps, la tête de réseau va procéder à une phase de test du protocole sur les données, qui vise à répondre aux objectifs suivants :

1. Établir une première liste d'observateurs expérimentés pour certains taxons ;
2. Obtenir la quantité de données qui nécessitent une validation manuelle ;
3. Estimer l'effort à fournir pour traiter l'ensemble de ces validations manuelles ;
4. Affiner la liste des taxons difficiles à identifier.

Par ailleurs, dans le cadre de ses missions d'intérêt général, la tête de réseau flore produira des outils d'aide à l'identification des taxons à destination de tous les observateurs afin de limiter au maximum les erreurs d'identification. Ainsi, des clés de détermination seront produites, en priorité pour les groupes taxonomiques connus pour être difficiles à identifier avec les outils existants (*Psiadia* spp., *Polyscias* spp., *Heterochaenia* spp., *Turraea* spp., etc.).

Bibliographie

- Boullet V. (auteur) [Boullet V., Picot F. et Mallet B. (coord.)] 2020. – *Index de la flore vasculaire (Trachéophytes) de la Réunion : statuts, menaces et protections*. - Version 2020.1 (mise à jour du 2 décembre 2020). Conservatoire botanique national de Mascarin, Saint-Leu (Réunion), format numérique Excel ;
- DEAL, 2017. Charte du système d'information sur la nature et les paysages de La Réunion (SINP 974) (version 1.0). Rapport pour le SINP 974, 16 p. + 3 annexes ;

- DEAL (coord.), GCOI, Globice, Kélonia-CEDTM, NOI, OCEA Consult', ONCFS, SEOR, UMR PVBMT, 2020. Protocole de validation des données du système d'information sur la nature et les paysages de La Réunion – volet occurrences de taxons (version 1.2). Rapport pour le SINP 974, 25 p. + 14 annexes. ;
- 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., Quaintenne 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.

Annexe

- Index commenté de la flore vasculaire (Trachéophytes) de La Réunion (Boullet 2020).
- Liste des taxons posant des difficultés d'identification (liste de 966 taxons).

Fichier Excel : « *Liste_taxons_difficiles_Index2020.xlsx* ».

Auteur : Christian Fontaine, à dire d'expert à partir de l'Index commenté de la Flore de la Réunion version 2018, actualisé à partir de l'Index commenté de la Flore de la Réunion version 2020.

- Liste des étages de végétation d'occurrence pour chaque taxon (2187 taxons renseignés).

Fichier Excel : « *Gamme_ecologique_taxons_20211216.xlsx* ».

Auteur : Marie Lacoste, à dire d'expert à partir de l'Index commenté de la Flore de la Réunion version 2020 et de l'ensemble des données de la base de données Mascarine du CBNM.

Annexe 12: Synthèse des tests réalisés lors de la validation automatique

Voir le fichier *criteres_validation_v2.0.ods* joint au présent protocole.

Annexe 12 : Couches SIG de référence

Voir le dossier contenant les couches SIG suivantes :

Zones d'occurrence de taxon connues :

- tortues_marines_4 (sensible*)
- tortues_marines_5 (sensible*)
- tortues_marines_6
- cetaces_4_habitat_tt.shp
- cetaces_5_habitat_ta.shp
- cetaces_6_habitat_s.shp
- reptiles_amphibiens_3_agaaga.shp
- reptiles_amphibiens_3_amigut.shp
- reptiles_amphibiens_3_calver.shp
- reptiles_amphibiens_3_furpar.shp
- reptiles_amphibiens_3_gehmut.shp
- reptiles_amphibiens_3_hemfre.shp
- reptiles_amphibiens_3_hemmer.shp
- reptiles_amphibiens_3_hempar.shp
- reptiles_amphibiens_3_hemtyp.shp
- reptiles_amphibiens_3_indbra.shp
- reptiles_amphibiens_3_lycaul.shp
- reptiles_amphibiens_3_pheast.shp
- reptiles_amphibiens_3_phebor.shp
- reptiles_amphibiens_3_phegra.shp
- reptiles_amphibiens_3_phelat.shp
- reptiles_amphibiens_3_phelin.shp
- reptiles_amphibiens_3_pheine.shp (sensible*)
- reptiles_amphibiens_3_ptymas.shp
- validation_crustace_67.shp
- validation_crustace_418785.shp
- validation_crustace_418786.shp (sensible*)
- validation_crustace_418788.shp (sensible*)
- validation_crustace_418789.shp
- validation_crustace_418790.shp
- validation_crustace_418791.shp
- validation_crustace_418793.shp
- validation_crustace_432658.shp (sensible*)
- validation_crustace_432665.shp
- validation_crustace_528800.shp (sensible*)
- validation_oiseaux_phaeton_lepturus.shp
- validation_poisson_423547.shp
- validation_poisson_8.shp
- validation_poisson_185828.shp
- validation_poisson_200263.shp
- validation_poisson_418434.shp
- validation_poisson_418765.shp (sensible*)
- validation_poisson_418766.shp (sensible*)
- validation_poisson_418767.shp (sensible*)
- validation_poisson_418768.shp
- validation_poisson_418773.shp (sensible*)
- validation_poisson_559719.shp (sensible*)
- validation_poisson_562828.shp
- validation_poisson_418775.shp
- validation_poisson_418779.shp (sensible*)
- validation_poisson_418783.shp (sensible*)
- validation_poisson_418784.shp
- validation_poisson_418900.shp (sensible*)
- validation_poisson_418903.shp (sensible*)
- validation_poisson_424810.shp
- validation_poisson_425982.shp
- validation_poisson_528788.shp
- validation_poisson_528794.shp
- validation_poisson_533446.shp
- validation_poisson_560582.shp
- validation_poisson_586405.shp (sensible*)
- validation_poisson_67058.shp
- validation_poisson_67208.shp
- validation_poisson_67804.shp
- validation_poisson_68823.shp
- validation_poisson_698292.shp

* les couches sensibles ne sont pas diffusées conformément au référentiel de données sensibles du SINP974

Habitats naturels :

- cartographie de l'occupation du sol utilisée pour le STOC : oiseaux_habitat.shp

Autres couches de référence :

- limiter terre/ mer : limite_terre_mer.shp (source : couche des laines de mer TRONCON_LAISSE.TAB issue de la BD Topo – IGN, avril 2014)

Annexe 13 : Liste des statuts biogéographiques des taxons

Liste et signification des statuts biogéographiques utilisés dans Taxref (version 11) :

code	libelle_court	libelle_long
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. Les taxons hivernant quelques mois de l'année entrent dans cette catégorie. Par indig
B	Accidentelle / Visiteuse	Espèce occasionnelle (migratrice de passage) ou observée de manière exceptionnelle (taxon accidentel dans la zone géographique considérée).
E	Endémique	Espèce naturellement restreinte à la zone géographique considérée.
S	Subendémique	Espèce naturellement restreinte à une zone un peu plus grande que la zone géographique considérée, mais dont les principales populations se situent dans la zone géographique considérée (une espèce endémique des Mascareignes est considérée subendémique à L
C	Cryptogène	Taxon dont l'aire d'origine est inconnue, et pour lequel on ne peut donc pas dire s'il est indigène ou introduit.
I	Introduite	Taxon dont la présence dans la zone géographique considérée est due à une intervention humaine (intentionnelle ou non), ou taxon qui est arrivé là sans intervention humaine à partir d'une zone dans laquelle il est introduit. Ce statut inclut les taxons ét
J	Introduite envahissante	Taxon introduit et naturalisé/établi 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 a
M	Domestique / Introduite non établie	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'introduc
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	Taxon qui n'est plus présent à l'état sauvage dans la zone géographique considérée mais qui n'est pas globalement éteint.
X	Éteinte	Espèce globalement éteinte (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 éteinte	Taxon endémique et aujourd'hui disparu, donc globalement éteint (X).

Annexe 14 : Liste des sigles utilisés

- CBNM : Conservatoire botanique national de Mascarin
- CSRPN : Conseil scientifique régional du patrimoine naturel
- DEAL : Direction de l'environnement, de l'aménagement et du logement
- DEE : Donnée élémentaire d'échange
- GCOI : Groupe Chiroptères Océan Indien
- JDD : jeu de données
- MNHN : Muséum national d'histoire naturelle
- NOI : Nature Océan Indien
- ONCFS : Office national de la chasse et de la faune sauvage
- SEOR : Société d'études ornithologiques de La Réunion
- SQL : Structured Query Language, langage utilisé dans les systèmes de gestion de bases de données
- UMR PVBMT : Unité mixte de recherche « Peuplements végétaux et bioagresseurs en milieu tropical »
- ZEE : Zone économique exclusive