

Avis CSRPN N° 2022-04

AVIS DU CONSEIL SCIENTIFIQUE RÉGIONAL DU PATRIMOINE NATUREL DE LA RÉUNION

Protocole de validation SINP des données de flore vasculaire.

RÉUNION PLÉNIÈRE DU 31 MARS 2022

PÉTITIONNAIRE : CONSERVATOIRE BOTANIQUE NATIONAL DE MASCARIN

Contexte et objet de la demande

Le Système d'information de l'inventaire du patrimoine naturel (SINP) vise à fluidifier la mise en partage des données naturalistes afin de faciliter les actions de conservation ou de prise en compte du patrimoine naturel dans l'aménagement du territoire. Impulsé par le ministère en charge de l'environnement au niveau national et les DEAL/DREAL au niveau régional, il repose sur la mise en place progressive d'organisation, d'outils et de référentiels, aux niveaux national et régional.

Le Muséum national d'histoire naturelle (MNHN) a produit plusieurs documents de cadrage en 2016 et 2017 qui caractérisent les types de contrôle à appliquer aux données d'occurrences de taxons : identification des doublons, contrôles de cohérence et de conformité et validation scientifique.

Déclinant localement ce cadre méthodologique national, le protocole régional de validation des données constitue un élément central du SINP. Il vise à formaliser, diffuser et mettre en œuvre des règles permettant d'attribuer un niveau de fiabilité à toute donnée d'observation naturaliste circulant dans le SINP. Ceci est d'autant plus important que les données alimentant le SINP peuvent être de sources très variées, dont celles issues des sciences participatives, ce qui le différencie d'autres dispositifs comme le Système d'information sur l'eau.

À La Réunion, ce protocole régional est le fruit d'un travail collaboratif entre les têtes de réseau du SINP, le Parc national et la DEAL. Après examen par le CSRPN, il a été publié en septembre 2018. Formalisant les règles de validation des observations naturalistes versées au SINP de La Réunion, ce protocole évolue régulièrement pour intégrer les nouvelles connaissances. Accessible à l'adresse https://www.borbonica.re/protocole_validation/, ce protocole définit les règles appliquées à chaque nouveau jeu de données intégrant le portail du SINP 974 « Borbonica » (www.borbonica.re). Ainsi, 92 % des données disponibles sur Borbonica sont évaluées « certaines » ou « probables » .

Référent scientifique régional du SINP, le CSRPN a vocation à émettre des avis sur différentes thématiques intéressant le SINP : protocoles de collecte ou de détermination de données, protocole de validation de données, référentiels d'espèces sensibles, ... Concernant la validation des données, le CSRPN a déjà rendu trois avis favorables avec remarques :

- avis n° 2018-07 du 31 août 2018 : protocole de validation des données (volet faune) pour le SINP,
- avis n° 2019-14 du 19 décembre 2019 : protocole de validation du SINP de La Réunion,
- avis n° 2021-04 du 21 mai 2021 : actualisation du protocole de validation des données concernant les reptiles, les amphibiens, les poissons et les macro-crustacés d'eau douce.

Le CSRPN est sollicité sur le volet « validation des données de la flore vasculaire », proposé par le conservatoire botanique national de Mascarin (CBNM), tête de réseau pour le pôle thématique « flore et habitats naturels ». Ce volet complétera ainsi le protocole examiné en juin mai 2021.

Remarques préalables

Selon un logigramme précis, le protocole prévoit un processus de validation « automatique » basé sur le logiciel Borbonica et les différents référentiels, complété par un processus de validation « manuelle » faisant appel à l'expertise de spécialistes.

Validation automatique

Les quatre premiers filtres de validation automatique apparaissent très pertinents, à la différence du cinquième :

- « taxon présent dans l'index des Trachéophytes de La Réunion ? » et « taxon présent dans Borbonica ? » : ces deux premiers filtres permettent d'alerter sur de nouvelles espèces saisies dans le SINP, et entraînent une vérification manuelle.

Les taxons sont intégrés dans le référentiel taxonomique (TaxRef) sur la base de publications scientifiques et d'échantillons présents dans l'herbier universitaire de La Réunion, référencé au niveau international dans l'index « herbariorum ». Par conséquent, si le taxon n'est pas représenté dans l'herbier, il n'est pas intégré dans TaxRef. En complément de cet herbier universitaire, l'herbier de travail du CBNM permet d'actualiser l'index de la flore de La Réunion.

Le protocole de validation des données devra prévoir un dispositif pour vérifier l'exactitude des noms scientifiques et de la synonymie. Si besoin les corrections d'erreurs pourraient à priori être effectuées manuellement par les experts validateurs.

L'harmonisation des noms scientifiques s'impose également (par exemple, le Benjoin est nommé *Terminalia bentzoë supsp bentzoë* ou *Terminalia bentzoë*). Par ailleurs, concernant les espèces indigènes, la base de données Borbonica doit privilégier l'intégration des données relatives aux spécimens sauvages en milieu naturel, et ne pas les amalgamer avec celles concernant des individus plantés en milieux urbains et péri-urbains.

- « taxon dans un de ses habitats d'occurrence » : ce troisième filtre apparaît fondamental et permet de corriger des erreurs flagrantes, et donc de les invalider. Néanmoins, dans l'annexe fournie, ce filtre ne concerne que les espèces difficiles à reconnaître et non pas, comme le propose l'ensemble du logigramme, toutes les espèces. Un nouveau tableau avec toutes les espèces permettrait donc de vérifier l'ensemble des espèces et des stations qui ne correspondent pas à leurs habitats d'occurrence.

- « taxon difficile à identifier ? » : également pertinent, ce quatrième filtre mérite toutefois de reconsidérer certaines espèces ciblées comme difficiles à reconnaître, alors qu'elles ne le sont pas.

- « taxon rare à l'échelle de La Réunion » : ce cinquième filtre, n'apparaît pas pertinent. En effet la rareté n'est pas un critère absolu de difficulté pour reconnaître une espèce, de nombreuses espèces très rares étant très faciles à identifier. Par ailleurs, si l'espèce n'est pas rare, la donnée restera forcément qualifiée de « probable » ou « très probable », ce qui est discutable. De plus, il n'y a pas de liste des espèces rares, et les critères permettant de définir une espèce comme rare ne sont pas précisés.

Une alternative pourrait être de proposer le filtre « observateur expérimenté ? », prévu par la validation manuelle, car un observateur expérimenté communique des données généralement fiables. Sur la base de critères à définir, le CBNM, avec l'appui de la DEAL, pourra dresser une liste de botanistes expérimentés, dont les données seraient jugées comme correctes, au moins sur les taxons non difficiles à identifier ; à noter que certains observateurs sont plus expérimentés sur certains groupes d'espèces que d'autres.

Sur la base de ce statut d'observateur expérimenté, les données pourront ainsi être qualifiées comme « certaines », même en l'absence de preuve. En effet, il n'est pas pertinent de demander à

des botanistes expérimentés de fournir systématiquement un échantillon d'herbier ou des photos. Cependant, il faudra veiller à ne pas créer un goulot d'étranglement engendré par un trop grand nombre de données à valider manuellement. Aussi, lors de la mise en place du protocole de validation de la flore dans le SINP, faudrait-il prévoir une phase de test pour calibrer le dispositif et éviter cet écueil.

Validation manuelle

Certains filtres de validation manuelle sont à reconsidérer.

- « preuve existante et suffisante -part d'herbier, échantillon- ? » : ce filtre est d'une part assez contraignant, car il suppose que l'observateur produise systématiquement un échantillon pour que le taxon soit qualifié de « certain » ; d'autre part, il ne peut s'appliquer aux espèces protégées sans demande de dérogation, de surcroît dans certains types d'espaces protégés. S'agissant d'observateurs expérimentés, la demande d'un échantillon ne devrait être prévue que pour des espèces difficiles à identifier, sur la base d'une liste d'espèces proposée par le CBNM ; ceci permettrait de qualifier la donnée « certaine » au lieu de « probable ».
Il n'est pas logique que la donnée soit qualifiée de « probable » en ce qui concerne des espèces dont le statut serait jugé « douteux » suite à l'application du filtre automatique « taxon dans un de ses habitats d'occurrence ». En effet, ces espèces peuvent être identifiées de manière certaine par les experts validateurs et sont aussi reconnaissables grâce à des photos de bonne qualité. L'expert pourra juger si le taxon peut ou non être reconnu grâce à une photo ou bien s'il faut demander dans ce cas un échantillon. À noter qu'il sera probablement difficile de valider des données historiques en l'absence de telles preuves.
- « photos de l'observation ? » : il convient de définir quels sont les cas dans lesquels une photo est exigée ; dans le cas d'un nouveau taxon, un échantillon sera à verser à l'herbier.
- « identification correcte ? » : il est nécessaire de fournir à terme un document d'accompagnement pour aider les observateurs à lever les difficultés de détermination.

Autres propositions d'amélioration

Valider l'expertise des producteurs de données fiables

À ce jour, il n'existe pas de dispositif de reconnaissance de l'expertise de producteurs de données fiables. Cette liste d'observateurs de référence mériterait d'être dressée pour faciliter les étapes de la validation. À l'instar du mécanisme de reconnaissance des acquis de l'expérience, ce dispositif permettrait de légitimer ces producteurs de données selon des critères de qualification à définir. Citons par exemple un nombre de données valides par observateur pour un nombre donné de taxons, ou un nombre de déterminations exactes de taxons difficiles, etc.

Critères d'identification des taxons

Un document d'accompagnement du protocole de validation semble indispensable pour décrire les principaux critères de détermination des taxons.

Dispositif de pré-validation au sein des outils propres aux observateurs

Les structures partenaires du SINP pourraient être encouragées à intégrer un processus de pré-validation des données, notamment dès lors de la saisie de leurs données dans leurs propres bases, pour éviter d'alourdir les étapes de contrôle.

Élargir et consolider la composition du comité d'experts validateurs

Il est vivement recommandé d'élargir le comité de validateurs à des experts botanistes d'autres structures que le CBNM, afin d'associer d'autres compétences sur certains groupes, de renforcer et de mutualiser les compétences de l'ensemble des experts. Des organisations collégiales de ce type sont mises en place dans le cadre d'outils participatifs (tel Faune France, ...) ou de comités divers d'homologation, de façon à répartir la charge de validation et effectuer les contrôles manuels rapidement. Le CBNM est invité à s'inspirer de ces démarches.

S'assurer du respect du standard minimum du protocole SINPLe protocole de validation devra bien souligner les champs minimaux à renseigner, tels que le taxon observé, l'observateur, la localisation et la date (cf. standards SINP accessibles via <https://standards-sinp.mnhn.fr/>)

Avis final du CSRPN

Appréciant la rigueur de cet exercice complexe, le CSRPN émet un avis favorable quant au protocole de validation SINP des données de la flore vasculaire, sous réserve de :

- modifier le protocole de validation et le logigramme, car si les quatre premiers filtres de validation automatique sont pertinents, le cinquième mérite d'être remplacé par un filtre « observateur expérimenté ou non »,
- produire une liste d'espèces difficiles à déterminer, incluant celles pour lesquelles un échantillon et/ou une photo sont à fournir,
- valider l'expertise des producteurs de données fiables, selon une méthode à définir,
- élargir et consolider la composition du comité d'experts validateurs, en sollicitant des experts botanistes de différentes structures,
- procéder à une phase de test du protocole de validation des données avec quelques jeux de données, afin de :
 - calibrer le protocole de validation en fonction de la quantité de données qui nécessiteront une validation manuelle,
 - valider la liste des espèces difficiles à identifier.

Par ailleurs, le CSRPN recommande qu'à terme le CBNM mette à disposition des observateurs un outil d'aide à l'identification des taxons, qui permettra de limiter les erreurs de détermination.

Fait à Saint-Denis, le 05 juillet 2022

Le Président du CSRPN



Patrick FROUIN