

SYNTHÈSE
des entretiens avec un panel d'acteurs
travaillant à la production, au partage et
à la diffusion des données habitats

Stage - Alix Lafitte

Août 2020

UNITÉ MIXTE DE SERVICE PATRIMOINE NATUREL



Rédaction : Alix LAFITTE (UMS PatriNat)

Relecture : Solène Robert, Lucille Billon et Vincent Gaudillat (UMS PatriNat)

Experts mobilisés :

- BIOTOPE : Pierre AGOU, Cédric ELLEBOODE
- CBN Alpin : Jean-Michel GENIS, David PAULIN
- CBN de Bailleul : Emmanuel CATTEAU, Alexis DESSE
- CBN de Brest : Loïc DELASSUS, Sylvie MAGNANON
- CBN du Bassin Parisien : Silvère CAMPONOVO, Gaël CAUSSE
- CBN de Méditerranée : Virgile NOBLE
- CBN Sud-Atlantique : Aurélien BELAUD
- DREAL Centre : Damien PUJOL, Mathieu WILLMES
- DREAL Nouvelle-Aquitaine : Bertrand CHEVALIER
- DRIEE : Magali GIRARD
- GIP Loire : Sylvain CERISIER, Lise LEBAILLEUX, Kristell LEBOT
- IFREMER : Touria BAJJOUK, Noémie DELEYS, Emilie GAUTHIER
- IGN : Fabienne BENEST
- OFB : Annie BIROLLEAU, Elodie DAMIER, Gwenola DE ROTON, Guillaume FAUVEAU, Elodie GIACOMINI, Noémie MICHEZ, Mélanie ODION, Aurélien SCHMITT
- ONF : Francesco ROMANO
- Parc National des Ecrins : Camille MONCHICOURT
- UMS PatriNat (OFB-CNRS-MNHN) : Coline CHANET, Thibaut DE BETTIGNIES, Blandine DECHERF, Théo DUQUESNE, Vincent GAUDILLAT, Aurélie LACOEUILHE, Fanny LEPAREUR, Lise MACIEJEWSKI, Margaux MISTARZ, Judith PANIJEL, Rémy PONCET, Frédéric VEST

L'UMS Patrimoine naturel - PatriNat

Centre d'expertise et de données sur la nature



Depuis janvier 2017, l'Unité Mixte de Service 2006 Patrimoine d'expertise et de gestion des connaissances pour ses trois tutelle national d'Histoire naturelle (MNHN), l'Office français pour la biodiversité (OFB) et le Centre national de la recherche scientifique (CNRS).

Son objectif est de fournir une expertise fondée sur la collecte et l'analyse de données de la biodiversité et de la géodiversité présentes sur le territoire français, et sur la maîtrise et l'apport de nouvelles connaissances en écologie, sciences de l'évolution et anthropologie. Cette expertise, établie sur une approche scientifique, doit contribuer à faire émerger les questions et à proposer les réponses permettant d'améliorer les politiques publiques portant sur la biodiversité, la géodiversité et leurs relations avec les sociétés et les humains.

En savoir plus : patrinat.fr

Co-directeurs :

Laurent PONCET, directeur en charge du centre de données

Julien TOUROULT, directeur en charge des rapports et de la valorisation

Inventaire National du Patrimoine Naturel



Porté par l'UMS PatriNat, cet inventaire est l'aboutissement d'une démarche qui associe scientifiques, collectivités territoriales, naturalistes et associations de protection de la nature, en vue d'établir une synthèse sur le patrimoine naturel en France. Les données fournies par les partenaires sont organisées, gérées, validées et diffusées par le MNHN. Ce système est un dispositif clé du Système d'Information sur la Nature et les Paysages (SINP) et de l'Observatoire National de la Biodiversité (ONB).

Afin de gérer cette importante source d'informations, le Muséum a construit une base de données permettant d'unifier les données à l'aide de référentiels taxonomiques, géographiques et administratifs. Il est ainsi possible d'accéder à des listes d'espèces par commune, par espace protégé ou par maille de 10x10 km. Grâce à ces systèmes de référence, il est possible de produire des synthèses, quelle que soit la source d'information.

Ce système d'information permet de consolider des informations qui étaient jusqu'à présent dispersées. Il concerne la métropole et l'outre-mer, aussi bien sur la partie terrestre que marine. C'est une contribution majeure pour la connaissance naturaliste, l'expertise, la recherche en macroécologie et l'élaboration de stratégies de conservation efficaces du patrimoine naturel.

En savoir plus : inpn.mnhn.fr

SOMMAIRE

INTRODUCTION	4
ORGANISATION DE L'ENQUETE	4
I. La production, la gestion et la diffusion des données habitats dans les structures interrogées	5
I.1. Les données récoltées par les acteurs	5
I.2. L'état actuel de la bancarisation des données habitats	5
I.3. La diffusion des données mise en place par les structures	6
II. Le Standard d'Occurrence des Habitats du SINP (SOH) : connaissance générale du standard, difficultés rencontrées et propositions d'améliorations	9
II.1. La connaissance et l'utilisation du SOH parmi les sondés	9
II.2. L'avis général des utilisateurs du SOH	10
II.3. Les problèmes rencontrés lors de son utilisation	10
II.4. Les besoins et évolutions techniques envisagées au regard des problèmes soulevés	11
III. Le référentiel Habref : difficultés rencontrées et propositions d'améliorations	14
IV. Besoins, attentes et propositions en termes d'informations diffusées sur l'INPN, la plateforme nationale du SINP	15
IV.1. Les propositions pour accéder aux données souhaitées	15
IV.2. Le viewer cartographique et les requêtes associées	15
IV.3. Les autres besoins de diffusion exprimés par les sondés	17
V. Etat actuel de la validation des données et des besoins, attentes et propositions dans le cadre du SINP	18
V.1. L'état actuel de la validation des données habitats	18
V.2. La validation souhaitée dans le cadre du SINP	19
VI. Besoins, attentes et propositions en termes d'accompagnement sur la thématique habitat	20
VI.1. Les défauts d'accompagnement des acteurs	20
VI.2. Les besoins d'accompagnement exprimés par les acteurs autour de l'utilisation du SOH	21
VI.3. L'accompagnement des acteurs sur les sujets de la diffusion des données habitats	22
VII. Besoins, attentes et propositions en termes d'outils	22
CONCLUSION ET PERSPECTIVES	24
ANNEXES	26

INTRODUCTION

La gestion des données sur les habitats est un enjeu fort pour la préservation des milieux naturels en France. En janvier 2020, l'UMS PatriNat a conduit un séminaire nommé 'Typologies, données et cartographies : état des lieux des connaissances et perspectives sur les besoins opérationnels'. En vue de ce séminaire, un questionnaire a été diffusé auprès des acteurs opérant sur la thématique des habitats. Un des objectifs de ce questionnaire était d'identifier les besoins et les difficultés rencontrées par ces acteurs afin d'améliorer le partage des données sur les habitats dans le cadre du SINP.

Ce travail a été approfondi par un travail d'audit ciblé auprès d'un panel de personnes ressources sur les problématiques de partage et de diffusion des données habitats. Leurs retours ont ensuite été analysés. Au-delà du bilan dressé ici, ils nourriront les réflexions sur les actions à mener sur la thématique des données d'habitats, notamment pour décider des évolutions du standard d'échange des données d'occurrence des habitats du SINP (SOH), des perspectives de développement du volet habitats du SINP et de la diffusion des données habitats sur l'INPN.

Précaution de lecture :

Ce travail a été réalisé au cours d'un stage de 2 mois.

Il convient de noter, avant lecture, que cette synthèse a été rédigée de manière à respecter, le plus fidèlement possible, les retours des acteurs, quel que soit leur niveau de connaissance par rapport aux sujets abordés. Les commentaires sont des avis et propositions propres au contexte de travail des sondés et de leur structure.

ORGANISATION DE L'ENQUETE

Afin de mener à bien cet audit, une liste de personnes ressources sur les sujets du partage et de la diffusion des données habitats a été élaborée. Une liste interne à l'UMS PatriNat a d'abord été créée. Ensuite, les retours du questionnaire dispensé lors du séminaire habitats de janvier 2020 ont été analysés afin de dresser la liste des personnes ressources externes. Cette liste a été complétée par d'autres contacts connus des membres de l'UMS PatriNat et par des personnes ressources complémentaires identifiées lors des entretiens.

En parallèle, une trame d'enquête (Annexe.1.), inspirée par les retours des réseaux naturalistes lors du séminaire habitats de janvier 2020, a été réalisée. Les questions ont été laissées relativement ouvertes (entretiens semi-dirigés) pour ne pas orienter les réponses et laisser l'opportunité aux enquêtés d'aborder d'autres sujets en lien avec les données sur les habitats. Des modalités de réponses possibles ont néanmoins été imaginées pour pouvoir questionner plus amplement la personne interviewée.

Les entretiens se sont déroulés entre mi-juillet et mi-août. Les personnes ont été contactées par mail et ont été invitées à partager leurs disponibilités pour qu'un entretien physique, téléphonique ou par visioconférence puisse être convenu.

Au total 44 personnes ont été interviewées au cours d'une trentaine d'entretiens. 11 sondés appartenaient à l'UMS PatriNat et 1 était un ancien membre de l'équipe. Le reste des sondés provenaient de structures variées. 10 sondés appartenaient aux CBN (Alpin, du Bassin Parisien, Sud-Atlantique, de Bailleul, méditerranéen, et de Brest). 8 sondés faisaient partie de l'OFB et travaillaient dans le service ECUMM, le programme Life MARHA ou bien dans des Parcs Naturels Marin. Pour finir, 3 membres de l'IFREMER, 3 membres des DREAL (Centre et Nouvelle-Aquitaine), 1 membre de la DRIEE Ile-de-France, 3 membres du GIP Loire, 2

membres du bureau d'étude BIOTOPE, 1 membre de l'IGN, 1 membre de l'ONF et 1 membre du Parc National des Ecrins ont également participé aux entretiens.

En ce qui concerne leur domaine d'action, les acteurs travaillaient à 25 % sur le milieu marin. Les autres travaillaient en milieu terrestre. Par manque de contacts, aucun acteur ultramarin n'a pu être interrogé.

Les retours des acteurs ont ensuite été compilés et analysés pour pouvoir identifier les pistes d'amélioration pour le partage et la diffusion des données concernant les habitats.

I. La production, la gestion et la diffusion des données habitats dans les structures interrogées

I.1. Les données récoltées par les acteurs

Les données récoltées par les acteurs sont majoritairement sous formes de couches SIG (points, lignes ou polygones). Des relevés non spatialisés peuvent également être utilisés, notamment dans le cadre des relevés phytosociologiques.

Les données habitats des CBN et bureaux d'études consistent souvent en des relevés phytosociologiques et des données Natura 2000. Certains acteurs considèrent par ailleurs que les données phytosociologiques sont scientifiquement plus fiables que les données d'autres typologies. Les données peuvent ensuite être traduites dans d'autres typologies grâce aux correspondances.

Les données bancarisées peuvent provenir de programmes d'étude particuliers. Ainsi, les données habitats du CBN SA sont majoritairement issues d'un programme d'étude des sites Natura 2000, réalisé en commun avec la DREAL de Nouvelle-Aquitaine en 2014. Ce programme consistait en la standardisation et le regroupement des données habitats de ces sites, un travail fastidieux du fait des divers formats de données compilés. Le CBN Alpin possède également une majorité de données Natura 2000.

En milieu marin, la façon d'appréhender les habitats semble être un peu différente du milieu terrestre. Les protocoles d'acquisition, transformation et détermination des données habitats sont très précis et variés. Il faut entre autres noter que les données marines d'habitats sont déduites de l'observation de plusieurs stations. La donnée peut être issue de programmes internes ou de marchés publics. L'identification des habitats marins peut aussi être issue de caractéristiques physiques du milieu déduites d'imageries acoustiques et optiques (roche, sable, etc.). Ces données brutes peuvent être réutilisées pour déduire l'habitat selon une autre typologie choisie grâce à des méthodes précises (EUNIS, habitats benthiques de méditerranée et Atlantique et aussi Cahiers d'habitats). Ces exemples de particularités du milieu marin montrent que la comparaison des données pourrait être rendue difficile dans le cas où des métadonnées complètes n'étaient pas associées aux données. Elles en assurent, en effet, la qualité et la traçabilité.

I.2. L'état actuel de la bancarisation des données habitats

La plupart des structures possèdent des bases de données internes et des standards de données antérieurs au SOH.

En milieu marin, les données brutes sont très lourdes et hétérogènes et nécessitent la gestion dans plusieurs bases de données à l'IFREMER : QUADRIGE, SEXTANT, SURVAL.

Ces bases de données sont utilisées par de nombreux acteurs. En Méditerranée, la plateforme Medtrix d'un bureau d'étude indépendant est utilisée pour la gestion des données habitats.

Les typologies gérées dans les bases de données sont souvent : Prodrôme des végétations de France (PVF), EUNIS, CORINE biotopes, Habitats d'Intérêt Communautaire (HIC). Certains standards prennent en compte la liberté de pouvoir renseigner la typologie souhaitée (grâce notamment à un code typologie dans la table habitat). En milieu marin, les typologies Habitats marins benthiques d'Atlantique et Biocénoses benthiques de Méditerranée commencent à être utilisées.

Certains des standards des structures auditées semblent pouvoir être facilement mis en correspondance avec le SOH, notamment car le *cd_hab*, identifiant unique proposé par le référentiel Habref, semble être de plus en plus souvent utilisé et géré dans les bases de données. Ainsi, sur les aspects techniques, le partage des données dans le cadre du SINP semblerait être peu problématique.

Les standards internes peuvent parfois être multiples au sein d'une seule et même structure, ce qui peut rendre la gestion des données difficile. Certaines structures, comme l'ONF, ne possèdent pas encore de standards internes unifiés et pensent alors adapter le SOH à leur structure pour améliorer la qualité de leurs données et permettre de mieux les valoriser.

Afin de limiter les problèmes de standardisation, des outils de saisie en ligne sont développés comme Géonature¹, développé par le Parc National des Ecrins. Il est open source et possède un module occurrence des habitats, créé à partir des champs du SOH.

Sur la question de la gestion des mosaïques habitats dans les synthèses, les standards des CBN semblent être adaptés à leur gestion. En effet, dans le modèle de données, la table de l'unité cartographique est souvent différente de celle de l'habitat, ce qui permet de pouvoir entrer plusieurs habitats pour un même polygone et de pouvoir associer chaque habitat à un taux de recouvrement. Ensuite, les polygones peuvent également être associés à une liste d'habitats sans qu'il n'y ait de précisions sur le pourcentage de recouvrement de la zone d'étude (données ZNIEFF notamment).

Finalement, dans le cadre du programme CARTHAM, les mosaïques d'habitats sont gérées grâce à des champs supplémentaires issus d'un habitat principal. Chacun est associé à un pourcentage de recouvrement.

I.3. La diffusion des données mise en place par les structures

Les données habitats peuvent parfois être diffusées par les structures qui les produisent ou par celles qui les centralisent. La diffusion peut se faire à la demande d'un commanditaire. Des plateformes et des interfaces ont également pu être mises en place pour permettre la consultation des données habitats. Au regard des difficultés rencontrées, des choix ont parfois dû être faits pour pouvoir mettre en place ces outils.

➤ La mise à disposition des données

Tout d'abord, certains acteurs ne diffusent leurs données qu'au moment du rapportage européen des habitats Natura 2000. D'autres ne diffusent pas de données brutes par manque de moyens et de temps. Aussi, certaines données ne sont pas diffusables à cause de leur confidentialité ou parce qu'elles sont privées.

¹ <https://geonature.fr/>

Actuellement, les données sont souvent rendues accessibles après demande d'un commanditaire et extraction depuis la base de données. Ce service peut parfois être payant.

D'autres structures ont mis les données à disposition pour téléchargements via des sites web. Les données peuvent être accessibles via des webservices ou via des flux WMS ou WFS par exemple.

Chaque structure a pu adopter un ou plusieurs formats de diffusion. Les formats bruts des données peuvent alors être le .shp, le .csv, le .xlsx, le .acddb, le .pdf, le .xml. Par exemple, le programme Turboveg qui regroupe des relevés phytosociologiques permet leur diffusion en format .xml.

Des projets visant à diversifier les formats d'accès aux données peuvent être en cours dans certaines structures.

Les structures peuvent également diffuser leurs données grâce à des interfaces de consultation. Elles ne sont, pour la plupart, pas finalisées bien que les CBN semblent posséder des outils assez avancés. Les sites de l'OBVNA du CBN Sud-Atlantique, de SILENE du CBN méditerranéen diffusent et valorisent les données habitats, majoritairement issues de programme d'étude de sites Natura 2000.

L'OBVNA affiche les données habitat par mailles de 5 kilomètres et par communes sur un viewer cartographique. L'affichage par polygones est en projet. Il est possible de choisir la typologie d'affichage parmi la classification phytosociologique (PVF et référentiel régional du CBN), EUNIS, CORINE biotopes et les Cahiers d'habitats. Les habitats sont affichés selon la zone géographique choisie et la période de récolte. Le site de l'OBVNA est en évolution, notamment avec le projet de pouvoir visualiser les habitats par aires géographiques plus vastes qu'actuellement.

SILENE est actuellement en métamorphose. Les données sont affichées par polygones et par points dans les plus petites échelles (zoom France entière). Le choix a été fait de ne pas afficher les données issues de correspondances automatiques mais seulement celles déjà renseignées dans des typologies au sein de la base de données. Les données sont affichées et téléchargeables par projet, ainsi les données et métadonnées restent similaires.

La plateforme DIGITALE² du CBNBL permet d'afficher les couches d'habitats et leurs fiches descriptives associées.

Finalement, le CBN de Brest utilise CARMEN³ et un catalogue Geonetwork⁴ pour diffuser ses données Natura 2000 et ses relevés phytosociologiques.

Certaines interfaces de consultation permettent uniquement de montrer les habitats appartenant à une zone définie, par exemple aires protégées (ZNIEFF, zones Natura 2000). D'autres outils correspondent à un programme particulier et ne montrent que les données associées à ce programme (Inventaire forestier de l'ONF).

Certaines interfaces sont encore en projet. Le CBN Bassin parisien prévoit de créer un viewer cartographique avec des entrées par syntaxons PVF, par Habitats d'Intérêt Communautaire, par EUNIS et CORINE biotopes. Plus généralement, le projet s'inspirerait de l'OBVNA du CBN Sud-Atlantique avec un accès par mailles, communes et polygones. Il sera possible de choisir d'afficher les données selon une emprise géographique (zone Natura 2000 par exemple) et selon l'année de production de la donnée.

² <https://www.cbnbl.org/digitale2>

³ En Normandie : <http://carmen.application.developpement-durable.gouv.fr/8/nature.map>

⁴ <http://www.cbnbrest.fr/geonetwork/srv/fr/catalog.search#/home>

D'autres outils comme DEPOBIO⁵ (Dépôt légal des données brutes de biodiversité) permettent d'afficher les données au moment du téléversement des données sur la plateforme. Un viewer cartographique des données est en projet.

En ce qui concerne le milieu marin, l'outil CARTOMER⁶ de l'OFB a été abandonné au profit d'une plateforme OFB nationale en projet. L'outil donnait accès aux données brutes, à des flux, à une cartothèque, aux métadonnées, à un viewer cartographique qui permettait d'afficher les données par polygones et points.

Les plateformes SEXTANT⁷ et QUADRIGE⁸ de l'Ifremer permettent actuellement d'accéder aux données relatives aux habitats marins. Elles sont aussi utilisées par l'OFB en milieu marin. La première plateforme permet d'avoir accès à une cartothèque et à des données SIG (.shp). La seconde donne accès aux données ponctuelles d'habitats. Des webservices ont été mis en place entre les acteurs et QUADRIGE pour les données ponctuelles et attributaires. Ces outils permettent d'afficher des cartes exhaustives par zones géographiques, par mots clés ou par typologies. Enfin, l'Ifremer possède également la plateforme SURVAL⁹ qui permet de cartographier des points, des lignes et des polygones de surveillance puis de télécharger les données. Elle donne aussi accès aux métadonnées et aux données attributaires. La mise à jour des données est quotidienne.

Les données peuvent finalement être diffusées sous d'autres formats comme les cartogrammes, les rapports d'études, des rapports internes (Directive Habitats, Faune, Flore par exemple) ou bien des articles scientifiques.

➤ Les problèmes associés à cette diffusion

Un premier problème qui a été remonté est le manque de succès des plateformes de diffusion. Par exemple, l'utilisation de SILENE ne serait pas aussi importante que prévue.

Ensuite, sur la question des chevauchements d'habitats, les problématiques d'affichage peuvent être de natures différentes. En milieu marin, les chevauchements sont très fréquents. Il a alors parfois été décidé de ne montrer que l'habitat avec le plus d'enjeu. L'affichage de plusieurs typologies simultanément ou l'affichage de données provenant de jeux de données différents ont pu poser problème. Pour certains outils, la gestion des données par maille ou l'affichage d'un unique habitat dans une seule typologie résout ce problème. Sinon, une infobulle regroupant tous les habitats superposés permet d'avoir accès aux informations des différents polygones. Une autre difficulté lors de l'affichage est que les habitats peuvent être contigus mais incohérents. Il est donc nécessaire de bien revalider les données avant diffusion.

Les nombreuses méthodes utilisées par les acteurs pour la détermination des habitats et leur structuration rendent la diffusion plus complexe pour les structures rassemblant plusieurs sources de données. Elles doivent, en effet, souvent restandardiser les données pour pouvoir les diffuser.

Aussi, la délimitation spatiale du site d'étude et de l'objet cartographique ('station' dans le SOH) est subjective et dépend de l'observateur. Dans la méthodologie de l'observateur, le site d'étude peut être de la taille de l'objet cartographique ou bien être plus large.

⁵ <https://depot-legal-biodiversite.naturefrance.fr/>

⁶ <http://cartographie.afbiodiversite.fr/>

⁷ <https://sextant.ifremer.fr/>

⁸ <https://wwz.ifremer.fr/Recherche/Departements-scientifiques/Focus/Quadrige>

⁹ <https://wwz.ifremer.fr/surval/>

Pour finir, la diffusion des données habitats nécessite parfois de faire des choix sur la quantité d'information affichée. En effet, les cartes d'habitats peuvent être plus ou moins précises et être liées à plus ou moins de métadonnées. Ces informations additionnelles n'ont pas forcément pu être prises en compte pour la diffusion.

Pour résumer :

- ▶ Les données produites/gérées par les acteurs sondés sont majoritairement des **données géolocalisées** (couche SIG) et correspondent majoritairement à des **relevés phytosociologiques** et à des **données Natura 2000**.
- ▶ Les structures sont nombreuses à **posséder des standards d'échanges et des bases de données**. Elles évoquent cependant des **difficultés** à gérer et partager les données. En effet, la **standardisation des données transmises par les différents acteurs peut être approximative**, ce qui demande alors un travail conséquent pour normaliser les données.
- ▶ Les **données en milieu marin** semblent présenter des **particularités méthodologiques** qui rendent indispensable la **diffusion de métadonnées** associées, quand bien même leur diffusion est primordiale en tout milieu. Les données étant très lourdes, des bases de données différentes en fonction des objets ont dû être mises en place chez certains acteurs.
- ▶ Les **structures diffusent les données habitats** selon différentes modalités (plateformes web, sur demande, webservices...). Des **choix d'affichage** ont parfois dû être pris pour pouvoir diffuser les données au regard des différents problématiques. Certaines **contraintes de confidentialité ou de sensibilité** ne permettent pas toujours la diffusion des données. Le manque de moyens et de temps peut également constituer un frein notable à la diffusion.

II. Le Standard d'Occurrence des Habitats du SINP (SOH) : connaissance générale du standard, difficultés rencontrées et propositions d'améliorations

Certains sondés ont fait remonter leur avis sur le standard (SOH), les problèmes qu'ils ont rencontrés lors de son utilisation ainsi que des pistes d'améliorations pour que celui-ci puisse mieux répondre à leurs besoins.

II.1. La connaissance et l'utilisation du SOH parmi les sondés

La question de la connaissance du SOH a été posée aux sondés. Nous avons donc pu calculer les proportions de sondés connaissant et utilisant le SOH (Figure.1.). On observe que 15 % des sondés connaissent et utilisent le SOH. Une majorité (39 %) en a entendu parler mais ne le maîtrise pas. Un autre tiers ne le connaît pas du tout. Le reste des réponses n'a pas pu être traité dans cette partie par manque de précisions. Il faut nuancer la connaissance du SOH car presque 40 % des sondés le connaissant sont des personnes internes à l'UMS

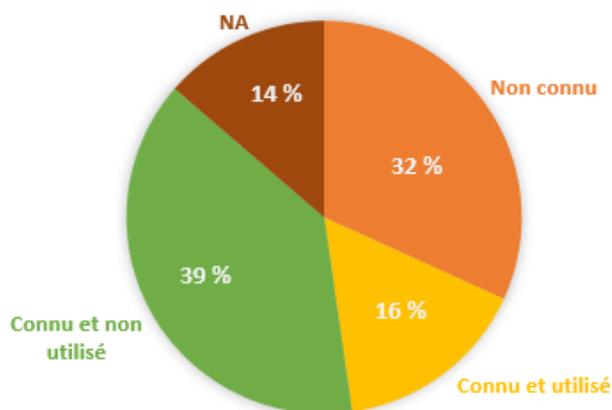


Figure.1. La connaissance et l'utilisation du SOH chez les acteurs

PatriNat. Aussi, de nombreux sondés ont avoué mal le connaître. Ces résultats semblent corroborer ceux de l'analyse du questionnaire Habitats du séminaire de Janvier 2020.

II.2. L'avis général des utilisateurs du SOH

L'intérêt de l'existence d'un tel standard est reconnu. Les acteurs semblent s'accorder sur le besoin d'avoir un langage commun. Il permet d'éviter que les données soient trop hétérogènes. Il permet aussi de pouvoir assurer une meilleure traçabilité des données. L'intégration et l'exploitation des données se trouvent facilitées par son utilisation.

Une minorité des acteurs interrogés trouve le SOH complet et facile à prendre en main. Toutes les possibilités du standard ne sont pas toujours utilisées. Les champs obligatoires permettent d'avoir un socle commun sans toutefois être trop contraignants.

Cependant, la majeure partie des acteurs jugent le standard peu pratique et éloigné de la réalité du terrain. Le standard a souvent été jugé trop complexe, ce qui rend la standardisation chronophage et rédhibitoire. La recherche d'un standard le plus complet possible pourrait être à l'origine de sa faible utilisation. Cela semble être à contre-courant de son but premier, favoriser le partage de données. En effet, certains sondés recherchent un standard le plus simple possible pour limiter les coûts d'intégration et accélérer la standardisation.

La complétude du guide semble effrayer les acteurs moins connaisseurs en gestion de bases de données. Le manque de concertation lors de sa création a également été soulevé. Certains acteurs ne voient pas d'objectifs clairs à son existence, ce qui les freinent dans son utilisation, notamment quand de nombreux standards locaux existent déjà.

Enfin, les acteurs ont insisté sur le besoin d'avoir un standard stable dans le temps pour éviter de devoir restandardiser les données trop fréquemment.

II.3. Les problèmes rencontrés lors de son utilisation

Le SOH a pu présenter certains problèmes d'utilisation. Ils concernent différents pans du standard comme la gestion des mosaïques et des typologies secondaires ou bien les tables, attributs et nomenclatures. Enfin, les spécificités du milieu marin ont semblé entraîner des problèmes lors de son utilisation.

➤ Les problèmes concernant l'intégration des mosaïques d'habitats

L'intégration des mosaïques d'habitats dans le standard peut s'avérer trop complexe pour les utilisateurs n'ayant de fortes compétences en base de données. Certains acteurs travaillent donc en utilisant les hauts niveaux hiérarchiques des typologies pour éviter les problèmes de gestion des mosaïques.

Il a également été remonté que le SOH n'est pas adapté aux imbrications d'habitats liées aux études phytosociologiques. Les habitats peuvent se révéler très petits et rendre la standardisation difficile.

➤ Les problèmes concernant la gestion des déterminations secondaires

La gestion via le SOH des correspondances et des déterminations secondaires semble floue pour certains acteurs. Il ne semble pas très aisé de rentrer plusieurs typologies dans le standard ce qui pousse les acteurs à ne rentrer qu'une seule typologie bien que l'étude en comprennent plusieurs. Il peut alors y avoir une perte d'informations et de qualité.

Le SOH pourrait également ne pas assez prendre en compte la phytosociologie. Il a donc été proposé de créer une extension phytosociologique.

➤ Les problèmes sur les tables, champs, et cardinalités du standard

En ce qui concerne le modèle du SOH, la table STATION semble être mal nommée. Elle pourrait être mal comprise par les acteurs qui ne savent pas quelles limites donner à la station et à la zone d'étude. La table CARHAB semble trop spécifique pour être intégrée dans un standard national.

Pour les champs, ils sont globalement vus comme trop complexes et devraient être simplifiés. Les notions auxquelles ils sont reliés ne sont pas forcément explicitées ce qui peut rendre leur compréhension difficile. Des champs peuvent parfois être inutiles, trop compliqués, notamment les conditions pour remplir un champ, ou scientifiquement faux de l'avis de certains sondés. Le champ *SensibilitéHab* devrait être supprimé car il n'y a pas de standard national sur la sensibilité des habitats ; il n'y a donc pas de contrôle possible sur les valeurs prises par cet attribut. Le champ *Surface* de la table STATION ne possède pas d'unité dans le guide méthodologique. Le champ *Typosolatlternative* devrait être retravaillé car il accepte des valeurs non prises en compte dans la nomenclature qui seront alors difficile à gérer dans la base de données. Il faudrait alors ajouter un champ *autre* dans la nomenclature. De même, certaines autres nomenclatures seraient incomplètes selon certains sondés.

L'attribution des UUID pose question pour certains acteurs. Ils ne savent pas qui doit les attribuer pour bien s'assurer de leur unicité.

Dans le cadre de DEPOBIO, des champs ont pu être supprimés car jugé trop compliqués à implémenter rapidement : il s'agit de la granulométrie et de la détermination secondaire. D'autres acteurs ont également modifié les champs du SOH pour que leurs libellés soient plus parlants du point de vue des utilisateurs.

➤ Les problèmes spécifiques au milieu marin

En milieu marin, le standard est jugé incomplet, notamment sur les référentiels de protocoles de collecte et de travail sur les données. Les métadonnées très importantes et complètes en milieu marin devraient pouvoir être renseignées dans le SOH pour éviter une perte de précision et d'informations.

II.4. Les besoins et évolutions techniques envisagées au regard des problèmes soulevés

Comme nous l'avons vu précédemment, le standard d'occurrence des habitats présente un certain nombre de problèmes qui, dorénavant identifiés, pourront être corrigés. Les sondés ont pu livrer leur avis sur les évolutions du SOH qui leur permettraient de mieux répondre à leurs besoins. Ces évolutions concernent la gestion des mosaïques d'habitats et des déterminations secondaires, l'implémentation de nouveaux champs, tables et nomenclatures. La création de SOH alternatifs a aussi été proposée par quelques acteurs.

➤ Les améliorations proposées pour pouvoir mieux intégrer les mosaïques d'habitats

Il a été remonté qu'un polygone ne devrait être relié qu'à un seul habitat dans le but de gagner en précision.

Selon certains sondés, les problèmes liés à la gestion des mosaïques d'habitats pourraient être résolus grâce à un modèle calqué sur le programme CARHAB. Il faudrait alors

que l'objet géographique soit géré au niveau du paysage. Cela permettrait d'avoir plusieurs échelles emboîtées d'habitats.

Il a également été proposé de créer un champ 'nom de la mosaïque d'habitats'. Ce nom de mosaïque pourrait alors correspondre à un niveau hiérarchique plus élevé dans la typologie choisie.

Pour gommer les problèmes de subjectivité de tailles du polygone entre les acteurs et méthodologies, il a été proposé de ne plus faire de polygones d'habitats mais des mailles. La comparabilité des données pourrait en être améliorée.

➤ Les besoins d'évolution de la gestion des typologies et correspondances dans le SOH

En ce qui concerne les problèmes de gestion des déterminations secondaires, une première solution a été avancée par plusieurs acteurs. Il s'agit de supprimer le concept de détermination secondaire du SOH et d'implémenter un code de typologie dans la table HABITAT. Ainsi, cette table comprendrait un code habitat (*cd_hab*), un nom cité, un code typologie (*cd_typo*), un code d'habitat d'intérêt communautaire et un recouvrement. Ce recouvrement devrait être inférieur ou égal à 100 % pour une typologie donnée. Cette modification répond au besoin de renseigner plusieurs typologies différentes en parallèle, pour être plus proche de ce qui se fait sur le terrain. De plus, ce *cd_typo* pourrait permettre de valider plus facilement les données.

Une autre idée d'évolution de la table de détermination secondaire a été suggérée. Il s'agit de supprimer la table de détermination secondaire et de créer une table TYPOLOGIE dans le standard. Elle permettrait de limiter le nombre de lignes occupées dans la base de données par un seul habitat (en fonction des différentes typologies). La table TYPOLOGIE permettrait également de pouvoir entrer une typologie locale dans le standard sans qu'elle ne soit dans le référentiel Habref grâce à un 'code typologie personnel' supplémentaire. Cette nouvelle table serait reliée à la table HABITAT grâce à une clé primaire *cd_typo* et à une cardinalité 1,n.

Outre ces deux propositions, d'autres commentaires ont été faits sur la gestion des correspondances et typologies dans le standard. La plupart des sondés pensent que les données provenant de correspondances sont moins fiables que les données provenant d'une typologie source. Il serait donc important de pouvoir renseigner le plus de typologies possibles dans le standard pour gagner en qualité de données et d'identifier la typologie primaire utilisée.

Le besoin de créer une hiérarchisation des typologies a ainsi été exprimé. Elle permettrait de connaître quelle est la typologie utilisée en premier sur le terrain.

Finalement, certains acteurs pensent que le standard devait intégrer directement la phytosociologie, la synphytosociologie et la gestion des habitats selon le programme CARHAB sans pour autant avoir des extensions comme pour CARHAB actuellement. Des acteurs venant des bureaux d'études et de CBN ont manifesté leur désir de voir le prodrome des végétations de France devenir la typologie principale du standard même s'ils reconnaissent que les typologies EUNIS et Habitats d'Intérêt Communautaire sont également fondamentales. D'autres acteurs pensent qu'EUNIS devrait être la typologie d'entrée du standard.

➤ Les autres changements évoqués sur les tables, champs et cardinalités

Les champs *nom_cité* et *cd_hab* devraient être rendus obligatoires pour certains sondés. Cela a notamment été précisé en milieu marin pour les typologies Habitats marins benthiques d'Atlantique et Biocénoses benthiques de Méditerranée. Une autre possibilité est que le renseignement obligatoire du *cd_hab* soit conditionné par la présence ou non de l'habitat dans

Habref. D'autres trouvent que le nom cité n'a pas autant sa place dans le SOH que dans le standard d'échange pour les taxons. Pour certains sondés, ne garder que le code habitat, sans le nom cité, permettrait de pouvoir limiter la lourdeur de saisie, faciliter la gestion des données et éviter les erreurs de saisie.

De nouveaux champs ont été proposés par les sondés. Il s'agit d'un champ définissant le statut diffusable ou non de la donnée, par exemple, dans le cas où la donnée n'est pas encore complète.

D'autres champs ont été proposés pour améliorer l'intégration des données Natura 2000 dans le standard. Le champ *Usage* pourrait être lié à une nomenclature et renseignerait l'usage qui est fait de l'habitat. Les champs *Gestion*, *Etat de conservation* et *Dynamique* de l'habitat ont aussi été proposés.

Finalement, il a été suggéré de rajouter un champ dans le SOH permettant de rentrer des fichiers image. Il permettrait de pouvoir, par exemple, partager un schéma explicatif du terrain dans lequel la donnée a été récoltée.

➤ Les demandes de création de SOH alternatifs

Quelques acteurs ont indiqué leur souhait de disposer de SOH alternatifs répondant mieux à leurs besoins.

Il a ainsi été proposé par certains sondés de créer un 'mini-SOH' destiné aux acteurs possédant moins de connaissances en gestion en base de données. Il aurait le minimum de champs requis pour avoir un langage commun sans pour autant rebuter les acteurs ayant un temps et des moyens limités pour la standardisation. Un nouveau format pourrait permettre d'intégrer des formulaires de saisie permettant de faciliter son utilisation et de limiter les erreurs de saisie.

Certains voudraient tout remettre à plat pour créer un SOH en concertation, plus proche des besoins du terrain. Il ne contiendrait que les champs ayant une utilité commune à tous les acteurs.

En ce qui concerne le milieu marin, la création d'un SOH spécifique au milieu marin a été proposée. Il pourrait permettre, selon certains sondés, d'avoir un référentiel complet des protocoles et méthodes associés au milieu marin. Ce référentiel serait basé sur les protocoles et méthodes du BODC.

En résumé :

- ▶ Le standard ne semble **pas bien connu** des sondés et **très peu affirme l'utiliser**.
- ▶ Le standard est jugé **trop complexe** bien que certains apprécient sa complétude.
- ▶ Les **principaux problèmes** remontés par les acteurs concernent les **déterminations secondaires**, certains champs sont jugés **inadaptés** et certaines **nomenclatures incomplètes**, notamment pour le milieu marin.
- ▶ Certains sondés ont alors fait part de leurs **propositions pour de futures évolutions du SOH** concernant les **déterminations secondaires**, les **mosaïques d'habitats**, ainsi que **certaines champs et nomenclatures**.

- ▶ Après correction du SOH, il a été proposé par les sondés de faire tester et valider le standard avant qu'il ne soit publié afin de voir s'il corrige bien les problèmes remontés et répond bien aux attentes des utilisateurs.
- ▶ Pour finir, il conviendrait de noter que les acteurs semblent globalement **méconnaître les outils du SINP**. En effet, des demandes concernant l'**intégration des métadonnées** et de la **phytosociologie** dans le SOH ont été formulées alors que **des standards dédiés sont disponibles**. Un travail d'animation autour de ces sujets serait probablement à organiser.

III. Le référentiel Habref : difficultés rencontrées et propositions d'améliorations

Certains problèmes rencontrés lors de la standardisation semblent parfois ne pas être liés au standard d'échange en tant que tel mais relatif au référentiel d'habitats Habref. Tout d'abord, la compréhension du référentiel peut parfois se révéler difficile. Le fait qu'Habref soit un référentiel de typologies et non une typologie peut prêter à confusion. Ensuite, les mises à jour régulières du référentiel semblent parfois freiner son utilisation.

Il a également été remonté que certains habitats pouvaient posséder les mêmes descriptions mais des codes habitats différents. Cela peut rendre le choix du code difficile et arbitraire. Il a donc été demandé de réaliser un travail de 'nettoyage' dans le référentiel.

Le manque de complétude du référentiel a également été indiqué. Certains sondés aimeraient que les typologies régionales soient prises en compte. En parallèle, il a été demandé d'homogénéiser les différentes typologies phytosociologiques locales, notamment en ce qui concerne les noms et codes attribués aux syntaxons. Les Habitats d'Intérêt Communautaire mériteraient également d'être complétés et harmonisés. Ensuite, la typologie EUNIS ne semble pas assez adaptée aux habitats de France. Les habitats en zones humides pourraient ne pas être assez intégrés dans le référentiel. Enfin, la symphytosociologie devrait être intégrée dans le référentiel.

En ce qui concerne les correspondances entre typologies, elles sont vues par certains comme indispensables pour mieux utiliser et valoriser les données habitats. Elles présentent néanmoins certaines limites et leur utilité est donc critiquée par certains sondés. Ces problèmes concernent d'abord le fait que les correspondances ne sont pas objectives (notamment PVF) et pas nécessairement très stables dans le temps. D'autres sondés reconnaissent que les données provenant de correspondances sont moins fiables. Il y a donc une nécessité de valider scientifiquement les typologies et les correspondances.

En résumé :

- ▶ Il semblerait que les acteurs souhaitent voir **le référentiel évoluer**.
- ▶ Certaines **typologies** devraient être **améliorées** (HIC, EUNIS, phytosociologie) et **d'autres ajoutées** (symphytosociologie et typologies locales).
- ▶ Les **correspondances** devraient également être **consolidées**.
- ▶ Néanmoins, les acteurs souhaitent également que le référentiel soit **stable dans le temps**.

IV. Besoins, attentes et propositions en termes d'informations diffusées sur l'INPN, la plateforme nationale du SINP

Les avis sur ce projet sont divers. Certains acteurs ont manifesté leur enthousiasme à l'idée de pouvoir accéder aux données habitats plus facilement. D'autres ont fait preuve de scepticisme. Ils souhaitent que l'intérêt de cette diffusion via l'INPN soit clarifié, notamment au regard des différentes plateformes déjà existantes. La qualité et la mise à jour des données qui y seront accessibles sont également au cœur des questionnements des acteurs. Les craintes se situent sur la redondance des données entre plateformes. Les données pourraient en effet être déjà accessibles sur les plateformes de partenaires et leur diffusion sur l'INPN pourrait être de moins bonne qualité.

IV.1. Les propositions pour accéder aux données souhaitées

Différents formats pour pouvoir accéder aux données habitats ont été proposés par les sondés. D'abord, un formulaire d'accès aux données sur le site de l'INPN a été avancé. Certains acteurs ont alors demandé à avoir accès aux différents utilisateurs des données pour connaître les différents usages faits à partir des données mises à disposition.

Ensuite, l'accès aux données brutes a été le plus souvent demandé. Le téléchargement direct, via un requêteur national, permettrait d'accéder aux données souhaitées (couches SIG d'un habitat ou d'une typologie donnée). Leurs formats pourraient être le .csv, le .xlsx, le .shp, le .accdb et le .xml. Les données devraient alors être les plus précises possibles pour être mieux valorisables/réutilisables. Elles se devraient d'être toujours au même format. Elles pourraient également toutes être liées à des typologies EUNIS et être associées à leur rapport d'étude.

Le téléchargement de données de synthèse a également été demandé par de nombreux acteurs (par exemple, par mailles).

Certains acteurs demandent que cet outil de téléchargement soit intuitif et simple d'utilisation pour ne pas se perdre dans les nombreuses requêtes envisageables. La possibilité de réaliser des études statistiques téléchargeables sur les habitats requêtés serait la bienvenue pour certains acteurs.

La mise en place de flux WMS et WFS pour avoir accès aux données SIG (mailles ou objets géographiques) a été réclamée par quelques acteurs. Ils permettraient de récupérer les données toujours à jour automatiquement tout en limitant les erreurs de saisie dans les documents .csv et .shp.

Une cartothèque/atlas pourrait être mise en place. Elle permettrait de diffuser des cartes de synthèse comme celles concernant les habitats dans les zones protégées (ZNIEFF, Natura 2000, PNR). On pourrait également y retrouver des cartes de répartition départementales, une demande très fréquente des acteurs. Ces cartes seraient plus exhaustives du point de vue surfacique et plus facilement utilisables pour les acteurs ne souhaitant pas manipuler les nombreuses requêtes d'un viewer ou d'un requêteur national.

IV.2. Le viewer cartographique et les requêtes associées

Il a été demandé aux sondés quels choix d'affichage ils proposeraient de retenir dans le cadre d'un viewer cartographique sur l'INPN. Cette proposition de nouvel outil a été plutôt bien reçue et a aussi fait l'objet de certaines recommandations.

Tout d'abord, l'outil se doit d'être performant pour pouvoir zoomer et dézoomer assez rapidement. Certains pensent que l'outil doit être le plus simple possible pour un usage du grand public.

L'organisation et les requêtes du viewer pourraient être calquées sur celui des espèces. Globalement, les sondés souhaitent avoir accès au plus de requêtes possibles si bien que toutes celles proposées lors des entretiens semblent avoir été retenues. Elles concernaient les choix d'affichage par habitats, typologies, niveaux hiérarchiques, correspondances, validation, dates, zones géographiques, jeux de données, observateur. Les choix sont, ainsi, détaillés ci-dessous.

➤ Affichage par maille ou polygones

La représentation par mailles ou par polygones sont toutes les deux envisagées pour la plupart des sondés bien que répondant à des besoins différents.

Les acteurs sont très nombreux à demander une représentation par maille car elle permet d'avoir une vision plus synthétique de la répartition des habitats. Elle serait par exemple très utile aux acteurs non experts en données habitats. Elle permettrait aussi de créer des cartes plus justes scientifiquement. En effet, les limites entre habitats, au niveau de leur détermination et de leur géographie sont très subjectives et dépendent du contexte de l'étude.

D'un autre côté, la représentation par polygones permet d'obtenir une plus grande précision des données. Cela pourrait s'avérer utile pour des travaux d'experts. La justesse scientifique d'une représentation d'habitats provenant d'auteurs ou d'études différentes est néanmoins questionnée par certains sondés. Si une telle représentation était choisie, il a été proposé d'avoir le choix du type d'objet affiché (points, lignes ou polygones).

➤ L'affichage par habitat et typologie et les correspondances

Le choix de l'affichage par habitat est unanime. Il devrait être basé sur ce qui se fait avec le viewer cartographique des espèces. Le polygone affiché pourrait alors donner toutes les typologies et correspondances auxquelles l'habitat peut être associé.

Le choix de pouvoir afficher les habitats provenant d'une ou plusieurs typologies semble intéresser les acteurs. Certains souhaiteraient également pouvoir afficher un ou plusieurs niveaux hiérarchiques dans une typologie souhaitée. Les niveaux hauts pourraient se révéler informatifs pour certains acteurs, notamment avant un travail de terrain. Ils pourraient également permettre de gommer l'imprécision de données et de calculer des surfaces d'habitats.

Les sondés ont montré leur besoin important en correspondances. Néanmoins, nombreux ont été ceux souhaitant que les données issues de correspondances automatiques, moins fiables, soient affichées de façon différente des données déterminées et validées dans une typologie. Une autre suggestion est de pouvoir sélectionner si l'on veut afficher ou non les données provenant d'une correspondance automatique, par exemple, dans le cadre d'une consultation plus experte.

➤ Les requêtes permettant d'appréhender la qualité des données

La question de la qualité de la donnée diffusée est très importante pour les sondés. Ainsi, ils sont nombreux à demander un affichage par jeux de données, par auteur ou par organisme producteurs de la donnée. Le cadre d'acquisition de la donnée a également été suggéré par certains acteurs. Ces requêtes pourraient permettre de s'assurer de la cohérence des données affichées et de leur fiabilité.

➤ Les requêtes par date

La requête par date a souvent été demandée avec notamment la possibilité de choisir une échelle de temps dans laquelle les données ont été récoltées. Des demandes de requêtes par seuil de date (avant 1990 par exemple) ou pour chaque rapportage européen ont également été soumises. L’affichage de plusieurs couches d’habitats provenant de plusieurs dates différentes a aussi été proposé.

De nombreux acteurs ont également exprimé le besoin de pouvoir visualiser l’évolution temporelle et spatiale des habitats. Cela pourrait se révéler particulièrement utile dans le contexte du changement climatique.

➤ Le choix de la zone géographique

La plupart des sondés ont exprimé le besoin de pouvoir croiser les données habitats avec des zones géographiques particulières. Il leur semblerait ainsi judicieux de pouvoir sélectionner des zones manuellement sur la carte et d’avoir le choix de n’afficher que les habitats de zones définies comme les communes ou les zones protégées par exemple.

Aussi, les acteurs marins souhaiteraient pouvoir choisir de n’afficher que les habitats des milieux marins. D’autres affichages par aires marines protégées, par façades, par types d’eaux (douces de transition...) ou par étages grâce à la bathymétrie ont également été proposés.

➤ Les autres besoins d’affichage exprimés par les sondés

D’autres besoins d’affichage ont été partagés par les sondés. Certains ont proposé la possibilité de pouvoir croiser les données habitats avec des données espèces ou environnementales comme le substrat en milieu marin. L’état de conservation de l’habitat pourrait également être affiché.

Aussi, une couche de fiabilité de la donnée pourrait être affichable sur le viewer cartographique, notamment pour le milieu marin en prenant appui sur le programme MESH.

IV.3. Les autres besoins de diffusion exprimés par les sondés

La diffusion des données habitats a fait l’objet d’autres recommandations de la part des acteurs.

Certains acteurs réclament que deux types de diffusions soient mises en place : une pour une utilisation grand public et une autre pour les experts.

Le besoin d’associer les métadonnées à la représentation a été soulevé. Elles permettraient de mieux comprendre les jeux de données et leur contexte. Il serait alors également plus simple de relier la donnée à son producteur et d’éviter toute confusion entre le producteur de la donnée et son diffuseur. Sur le même sujet, il a été demandé d’avoir la possibilité de connaître leur traçabilité. Ensuite, pour certains sondés, la confidentialité et la sensibilité des données pourraient constituer des freins à une éventuelle diffusion. Il conviendrait alors de créer une liste de sensibilité des habitats à la diffusion.

Une mise en valeur des contributeurs au SINP a été demandée. Elle permettrait une plus forte adhésion au programme de diffusion de leurs données.

Des acteurs en CBN souhaitent la création d’une couche de répartition validées par les CBN pour la phytosociologie et la typologie HIC. Elle pourrait être accessible sur le site de l’INPN.

En milieu marin, l'indice de confiance serait une information supplémentaire importante lors de la diffusion des données.

En résumé :

- ▶ La future diffusion des données habitats semble plutôt **bien accueillie** par les acteurs.
- ▶ Le **téléchargement des données brutes** est très largement demandé par les sondés.
- ▶ **D'autres formats de diffusion** ont également été proposés comme des **flux, cartes de synthèses...**
- ▶ La préférence pour un affichage par mailles ou par polygones semble dépendre du contexte d'utilisation des données.
- ▶ Des requêtes utiles sur le **viewer cartographique** ont été partagées par les sondés. Elles portent notamment sur le **choix de dates, zones géographiques, typologies et correspondances ou sur la qualité de la donnée.**
- ▶ L'importance de la **diffusion des métadonnées** a été soulevée.
- ▶ D'autres demandes plus spécifiques ont pu être formulées. Elles traitent du besoin d'une **diffusion experte** et d'une **plus tournée vers le grand public**, de la **valorisation des acteurs** et de spécificités de diffusion liées à certaines structures et domaines d'actions.

V. **Etat actuel de la validation des données et des besoins, attentes et propositions dans le cadre du SINP**

V.1. **L'état actuel de la validation des données habitats**

La validation des données diffusées est au cœur des préoccupations de nombreux acteurs. De nombreuses structures valident déjà les données habitats qu'elles récoltent, notamment les CBN. Cependant, cette phase de validation peut s'avérer couteuse, chronophage et demander un haut niveau d'expertise. Ainsi, au CBN de Bailleul, seules les données phytosociologiques peuvent parfois faire l'objet d'une validation avant d'être diffusées ou partagées.

De plus, la validation des données avant diffusion peut poser problème dans certaines structures, notamment si les typologies évoluent entre temps. Cela demande un travail conséquent de revalidation/reprise des données.

Le CBN de Bailleul a partagé sa méthode de validation. Il pratique trois vérifications, une chronologique, une géographique et une taxonomique. La vérification géographique permet de décider où se situe l'habitat quand un polygone est à cheval entre deux communes par exemple. La validation taxonomique permet de vérifier la plausibilité de la localisation spatiale d'un syntaxon grâce à des avis d'experts. Chaque habitat est validé dans une ou plusieurs typologies. Les données sont également vérifiées avant envoi au commanditaire dans le cas d'une demande d'extraction.

Finalement, lors de la diffusion des données habitats, les plateformes des CBN Sud-Atlantique, de Bailleul et méditerranéen montrent l'état de validation de la donnée. Ainsi, l'OBVNA indique si la donnée a été validée. SILENE ne peut afficher que les données validées ou douteuses. DIGITALE2 n'affiche que les données validées.

V.2. La validation souhaitée dans le cadre du SINP

La quasi-totalité des acteurs s'accordent sur le fait que les données doivent être validées avant une éventuelle diffusion au sein de l'INPN.

Une première validation technique de la donnée a d'abord été proposée. Elle pourrait s'effectuer grâce à un outil de validation des données qui, au moment du partage des données, vérifierait la bonne standardisation des données.

Ensuite, le choix du processus de la validation experte et scientifique des données a été soulevé par les CBN. Cette validation nécessite l'appui des réseaux de naturalistes et des producteurs de données.

Des algorithmes de vérification automatique pourraient être créés avec, par exemple, des indicateurs de qualité de la donnée. Ces algorithmes pourraient être réalisés dans le cadre d'une co-étude portée par des experts métiers et des informaticiens. Ces vérifications pourraient porter sur la typologie, les correspondances automatiques, la localisation, la présence de l'habitat dans une commune... Seules les données plus spécifiques et non prises en compte dans les algorithmes requerraient une validation manuelle par un expert.

En ce qui concerne la plausibilité de la localisation de l'habitat, la validation pourrait s'effectuer aux niveaux régional ou départemental. Il a aussi été imaginé de créer des mailles de synthèse validées par des experts (notamment en phytosociologie). Elles pourraient permettre de valider la localisation des habitats. Elles seraient mises à jour régulièrement, particulièrement si elles synthétisaient des habitats aux déterminations plus difficiles. Un autre moyen mis en avant pour vérifier la plausibilité d'un syntaxon en phytosociologie serait de comparer sa localisation avec sa présence ou non dans le référentiel phytosociologique local.

Pour finir, quelques recommandations ont été faites pour pouvoir diffuser les données. Ainsi, quelques sondés pensent tout de même que la diffusion de données non validées pourrait être envisageable. Il faudrait alors que le caractère validé ou non de la donnée soit bien précisé afin d'éviter toutes confusions. Il pourrait aussi être judicieux de séparer les données téléchargeables en deux dossiers différents : données validées et non validées.

Certains CBN ont aussi exprimé leur besoin d'avoir un espace ne diffusant que les données issues de CBN. Cela permettrait d'assurer la validation scientifique des données affichées et d'avoir des données issues de programmes et de dates similaires.

En résumé :

- ▶ La **validation des données** est un enjeu très important pour la majeure partie des sondés malgré son important **coût humain et financier**.
- ▶ La validation concerne la **chronologie**, la **taxonomie** et la **géographie** de la donnée.
- ▶ Les **plateformes de diffusion des CBN** informent sur l'**état de validation**
- ▶ Une **validation de la standardisation** des données **lors du partage** devrait être mise en place.
- ▶ Une **validation scientifique** des données pourrait être effectuée **par les réseaux naturalistes et producteurs**. Une **validation automatique** a également été proposée.
- ▶ Les **données diffusées** devraient être **séparées clairement selon leur état de validation**.

VI. Besoins, attentes et propositions en termes d'accompagnement sur la thématique habitat

De nombreux problèmes rencontrés par les acteurs ne semblent pas directement liés au standard mais à un manque d'accompagnement général sur les notions liées aux habitats naturels et l'utilisation du SOH. Un manque de dialogue entre les acteurs nationaux et locaux a été rapporté.

L'existence du gabarit semble être un bon début pour quelques acteurs mais globalement, l'accompagnement des acteurs semble encore trop faible. Il constitue même un des principaux problèmes lors de l'utilisation du SOH alors qu'un fort besoin a été exprimé.

Concernant la future diffusion sur l'INPN, un accompagnement des utilisateurs a également été demandé.

VI.1. Les défauts d'accompagnement des acteurs

➤ Les documents actuels du SOH ne sont pas jugés adaptés

Un problème majeur dans l'utilisation du SOH semble être le manque de documents clairs qui expliquerait comment on utilise le standard. Le guide actuel du SOH est considéré comme incompréhensible ou rebutant par la plupart des sondés. Si bien que ce document est peu connu et assez mal maîtrisé. Il est aussi parfois jugé flou car certains acteurs ne voient pas comment remplir certains champs du SOH.

Il est jugé inadapte pour une utilisation par des naturalistes peu connaisseurs en bases de données. Il est toutefois jugé adapté pour les géomaticiens et gestionnaires de données.

➤ Sur le partage des données

Les modalités de partage des données peuvent paraître floues pour certains sondés. Ainsi, des questions ont été posées concernant la nature des documents à envoyer, leur format, le lien entre les documents et le destinataire SINP des données.

Il conviendrait également que l'UMS PatriNat se charge de créer un référentiel des sites web partageant et diffusant leurs données sur les habitats. Certains sondés ont, en effet, fait part de leur confusion entre les nombreux sites de projets différents et d'aires géographiques variées. Ce référentiel pourrait également identifier où sont les données les plus récentes car les acteurs ont parfois du mal à trouver la dernière mise à jour de la donnée.

Plusieurs acteurs ont fait remonter leurs besoins de retours d'informations lorsqu'ils envoient les données à l'UMS PatriNat. Les retours pourraient concerner la bonne standardisation des données ou la plausibilité de la cartographie des habitats au vu des données validées. Certains acteurs souhaiteraient pouvoir comparer leurs données à d'autres sources pour pouvoir vérifier leur standardisation.

Il a donc été proposé par certains acteurs de réaliser une période de test durant laquelle les acteurs utilisent le SOH et partagent leurs données avec l'UMS PatriNat. Ils seraient ensuite conseillés sur les problèmes de standardisation des données.

Les modalités légales qui obligent les acteurs à transmettre leurs données peuvent sembler floues et nécessiteraient peut-être d'être mieux explicitées, notamment pour les partenariats entre l'UMS PatriNat et d'autres organismes publics (conventions, protocole...).

VI.2. Les besoins d'accompagnement exprimés par les acteurs autour de l'utilisation du SOH

Généralement, le fort besoin de vulgarisation et de publicité du SOH a été partagé par de nombreux sondés. Des nouveaux formats de vulgarisation pour expliquer son utilité et expliquer son fonctionnement ont été proposés. Ensuite, les acteurs ont été nombreux à demander des outils intuitifs pour les aider à intégrer et gérer leurs données d'occurrence des habitats.

Des nouveaux formats ont été réclamés pour mieux expliciter les liens entre Habref et le SOH ainsi que différentes définitions en lien avec les données habitats comme celles de données/métadonnées, d'habitats naturels ou d'habitats d'espèces, ou les notions d'occurrences de taxon ou d'habitat. Ces notions pourraient également être présentées sur l'INPN via des infobulles définissant ces termes ou bien via un glossaire.

Des documents aidant les acteurs tout au long de leur travail, lors de la collecte, de la standardisation, du partage, de la diffusion et de l'utilisation des outils de diffusion, ont été demandés.

Les acteurs sont très nombreux à demander des fiches méthodologiques sur l'utilisation du SOH. Elles pourraient faire 1 à 2 pages, être très visuelles et créer une forme de curiosité de la part des éventuels utilisateurs.

Il a également été proposé de créer de courtes vidéos introductives expliquant l'intérêt de l'utilisation d'Habref et du SOH, ainsi que leur fonctionnement. Elles pourraient également permettre d'expliquer plus globalement ce qu'est le SINP, l'INPN ou d'autres notions en lien avec le partage de données habitats.

Une FAQ pourrait être réalisée au regard des erreurs les plus courantes observées lors de l'utilisation du SOH. Des contacts de personnes ressources avec lesquelles les acteurs pourraient prendre contact ont aussi été demandés.

Des propositions pour la réalisation de séminaires permettant de présenter les enjeux et de former à l'utilisation des outils ont été faites. Ces formations présenteraient également les différents outils existants. Elles expliqueraient les notions liées à la gestion des données.

Des réunions de travail entre les acteurs ont pu être proposées. Elles permettraient de rassembler les retours des acteurs sur le partage des données et d'écouter leurs besoins sur l'évolution des outils. Elles pourraient également permettre de se concerter sur la création d'un nouveau SOH, demandé par certains acteurs.

Le guide méthodologique actuel devrait être conservé car il est jugé utile pour les géomaticiens. Cependant, il devrait être amélioré et son utilisation facilitée grâce à des listes de valeurs dynamiques par exemple. De nombreux exemples devraient être ajoutés pour montrer comment remplir les champs, notamment pour les cas complexes de détermination secondaire et de mosaïques d'habitats. Le guide pourrait également être séparé en plusieurs parties plus digests et le modèle de données devrait être plus schématique.

Pour finir, il est à noter que, si différents outils d'accompagnement autour de l'utilisation du SOH étaient choisis, il deviendrait nécessaire de bien expliciter quels documents utiliser en fonction des besoins. En effet, le site de téléchargement du standard est jugé austère par certains acteurs qui ne savent pas quel standard y choisir.

VI.3. L'accompagnement des acteurs sur les sujets de la diffusion des données habitats

Les acteurs ont été nombreux à partager que si plusieurs moyens de diffusion étaient choisis, il y aurait alors besoin d'accompagner les acteurs lors de l'utilisation des outils mis à disposition. Dans l'exemple du requêteur ou viewer cartographique, les utilisateurs pourraient se perdre dans les nombreuses requêtes possibles, notamment s'ils n'ont pas le temps de s'approprier les outils assez amplement.

En résumé :

- ▶ L'**accompagnement** des acteurs a été jugé **trop faible** par les sondés.
- ▶ Les **documents** actuels leur paraissent parfois **inadaptés** et un **flou** existe sur les **modalités de partage** des données habitats.
- ▶ Des **nouveaux formats d'accompagnement** sur les notions d'habitats naturels et sur l'utilisation du SOH ont été proposés. Il s'agit entre autres **de nouveaux guides, fiches, vidéos, FAQ et formations**.
- ▶ La **future diffusion** des données devra également **faire l'objet d'un accompagnement** pour faciliter l'utilisation des outils.
- ▶ D'autres remarques ont été partagées par quelques acteurs. Ainsi, l'accompagnement pourrait ne pas suffire pour avoir des données bien standardisées selon certains. Il faudrait avoir des **moyens supplémentaires** dans les structures pour que le personnel soit formé et ait le temps pour effectuer la standardisation.
- ▶ D'autres pensent que **l'automatisation du partage** permettrait de **limiter les problèmes de standardisation et d'accompagnement**.

VII. Besoins, attentes et propositions en termes d'outils

Certains sondés indiquent que la bonne standardisation des données par les différents acteurs est garante de la bonne diffusion des données. Il conviendrait donc de bien les accompagner grâce à des outils de saisie intuitifs pour éviter toute erreur de standardisation.

Ce besoin en outils de saisie a été formulé par un grande partie des sondés, notamment pour les structures ne possédant pas de gestionnaires de données. Certains ont proposé des formulaires de saisie SIG et tableur. Ils pourraient favoriser la qualité des données en facilitant l'entrée de données complexes par les utilisateurs.

Une plateforme en ligne pour saisir les données habitats a également été proposée. Ces outils pourraient permettre d'utiliser le SOH sans avoir à le connaître. Ils permettraient également aux prestataires de mieux comprendre et réaliser la standardisation requise des données. Cette plateforme pourrait intégrer les correspondances et les codes Habref (via webservices) afin de pouvoir relier automatiquement le code habitat et le nom de l'habitat lors de la saisie. Cet outil faciliterait ainsi la saisie, en limitant les erreurs, ainsi que la vérification, la validation et la diffusion des données. Aussi, certains acteurs semblent attendre l'arrivée du module habitats de Géonature pour pouvoir y saisir/importer leurs données habitats.

Un outil de saisie mobile a été proposé et pourrait permettre de générer les polygones d'habitats sur le terrain grâce à un GPS et de les intégrer directement dans la base de données.



En résumé :

- ▶ Différents outils de saisie des données d'habitats ont été proposés. Il peut s'agir de **formulaire de saisie Excel ou SIG** ou bien **d'outils de saisie en ligne**.
- ▶ Ils pourraient permettre de **faciliter la standardisation** des données et de **limiter les erreurs de saisie**.
- ▶ Néanmoins, si plusieurs outils étaient choisis, il conviendrait de pleinement **s'assurer de la complémentarité des différents outils** mis à disposition sans qu'ils ne soient redondants.

CONCLUSION ET PERSPECTIVES

Suite aux entretiens menés auprès de 44 personnes des réseaux naturalistes, les retours ont été analysés pour pouvoir conclure quant à l'état actuel du partage et de la diffusion des données habitats, aux problèmes et besoins exprimés sur le SOH, la diffusion et l'accompagnement des acteurs ainsi que sur la validation des données et l'utilisation d'Habref.

Tout d'abord, il est possible de retenir de cette enquête que les acteurs sondés produisent majoritairement des données au format SIG. Les structures utilisent globalement des standards d'échanges (internes à leur structure ou réseau pour la plupart) et disposent de bases de données. Certains gèrent des métadonnées (données sur les données), composant indispensable pour pouvoir mieux valoriser les données et s'assurer de leur qualité.

Les structures sont nombreuses à valider leurs données avant toute diffusion même s'ils indiquent que la validation peut s'avérer chronophage et nécessiter une grande expertise.

Certaines structures diffusent leurs données habitats à la demande ou via des plateformes web ou encore des webservices. L'affichage des habitats en ligne a pu leur poser quelques difficultés, notamment concernant le choix des éléments à afficher (complétude des informations notamment). Certaines structures ne diffusent pas leurs données, pour cause de manque de moyens, de confidentialité ou de sensibilité des données.

En ce qui concerne le standard SOH, il ne semble pas bien connu des sondés et peu utilisé. Si sa complétude a pu être appréciée, les sondés indiquent néanmoins une trop grande complexité et des imprécisions dans les terminologies utilisées ainsi que des lacunes au niveau des listes de valeurs (nomenclatures) notamment pour ce qui concerne la prise en compte des habitats en milieu marin. Les sondés ont également indiqués des difficultés à y gérer les déterminations secondaires. Il a alors été suggéré de modifier le SOH pour corriger ces problèmes.

Certains problèmes évoqués par les sondés vis-à-vis de la standardisation pourraient aussi avoir pour origine l'utilisation d'Habref. Sur ce sujet, certains sondés ont ainsi proposé que certaines typologies soient améliorées (EUNIS, phytosociologie...) et d'autres ajoutées (symphytosociologie, typologies locales...). Un travail de consolidation des correspondances et une stabilité dans le temps du référentiel ont été proposés.

Les acteurs semblent bien accueillir la diffusion des données habitats via l'INPN malgré quelques appréhensions, notamment vis-à-vis de la qualité et de la fraîcheur des données qui y seront diffusées mais également vis-à-vis de la redondance de diffusion avec d'autres plateformes. Les sondés ont proposé que soient mis en ligne deux grands types de données pour répondre aux différents contextes d'utilisation : des données brutes en premier lieu, mais également des données de synthèse (par maille notamment). Les sondés proposent que les données soient requêttables par habitats, dates, zones géographiques, typologies et correspondances ou encore par niveau de qualité/validité.

L'accompagnement actuel autour du partage et de la diffusion des données sur les habitats est jugé insuffisant et mériterait d'être renforcé. Ainsi, certaines personnes interrogées ont indiqué que les termes mériteraient d'être mieux définis (notion de donnée brute, d'habitat, etc.).

Actuellement, les documents d'accompagnement du SOH sont jugés inadaptés et les sondés ont proposé la création de nouveaux guides, fiches, vidéos, FAQ et la mise en place de formations pour les aider à développer ou renforcer le partage et la diffusion des données sur les habitats.

La mise en place d'outils de saisie permettrait de faciliter la standardisation des données tout en limitant le besoin de connaissance du standard. Ces outils pourraient prendre la forme de formulaires de saisie Excel ou SIG ou bien d'outils de saisie en ligne. La complémentarité des différents outils devrait par ailleurs être assurée pour éviter les confusions et les redondances.

Grâce à ce travail prospectif, des perspectives de nouvelles missions à amorcer peuvent être dressées. Le groupe de travail ayant créé le SOH pourrait être rouvert pour pouvoir le réviser et apporter les modifications nécessaires. Un ou plusieurs groupes de travail SINP dédiés aux habitats pourraient être créés et permettraient d'aborder les questions de flux, de sensibilité et de validation scientifique. Un outil de saisie habitat pourrait être construit dans le cadre du projet GINCO.

En ce qui concerne l'accompagnement des acteurs, des pages d'aides pourraient être développées sur l'INPN. Une animation pourrait être mise en place sur les sujets des données habitats. Ces sujets pourraient également être abordés lors des journées techniques du SINP.

ANNEXES

Annexe.1. La trame d'enquête suivie au cours des entretiens avec les réseaux naturalistes

- I. L'utilisation du Standard d'Occurrence des Habitats du SINP (SOH)
- a. Sur le terrain, quels types de données récoltez-vous ?
SIG, relevés...
 - b. Connaissez-vous le SOH ?
 - c. Utiliseriez-vous le SOH ?
Si non, pourquoi ?
 - d. Que pensez-vous du SOH en général ?
 - e. Quels problèmes avez-vous rencontré lors de votre utilisation du SOH ?
 - f. Quels sont vos principaux besoins concernant un standard d'échange, notamment le SOH ?
 - g. Auriez-vous des suggestions pour améliorer le SOH ?
Aspects techniques ? quels attributs manquants ? quelles cardinalités ? quelles nomenclatures (listes de valeurs) ?
 - h. Quelles sont les principales améliorations envisageables concernant l'accompagnement des acteurs sur le partage des données d'habitat pour vous ?
Fiches méthodologiques SOH, outils de saisie...
- Comment ?
- II. La diffusion des données d'occurrences des habitats
- a. Participez-vous à la diffusion/valorisation des données d'occurrence des habitats ?
Articles scientifiques, synthèses,...
 - b. Comment vos données habitats sont-elles rendues accessibles ?
site web, viewer cartographique, formulaire, téléchargement direct, webservice
- Sous quel format ?
.csv (du SOH ?), .xlsx, .shp, .accdb...
- c. Quels problèmes avez-vous rencontré lors de la diffusion de vos données ?
Aspects techniques (base de données, droits...)
 - d. Pour une diffusion nationale au sein de l'INPN, comment voudriez-vous que les données soient rendues accessibles ?
 - e. Avez-vous des recommandations pour la diffusion des données habitats ?
support de diffusion, paramètres importants...