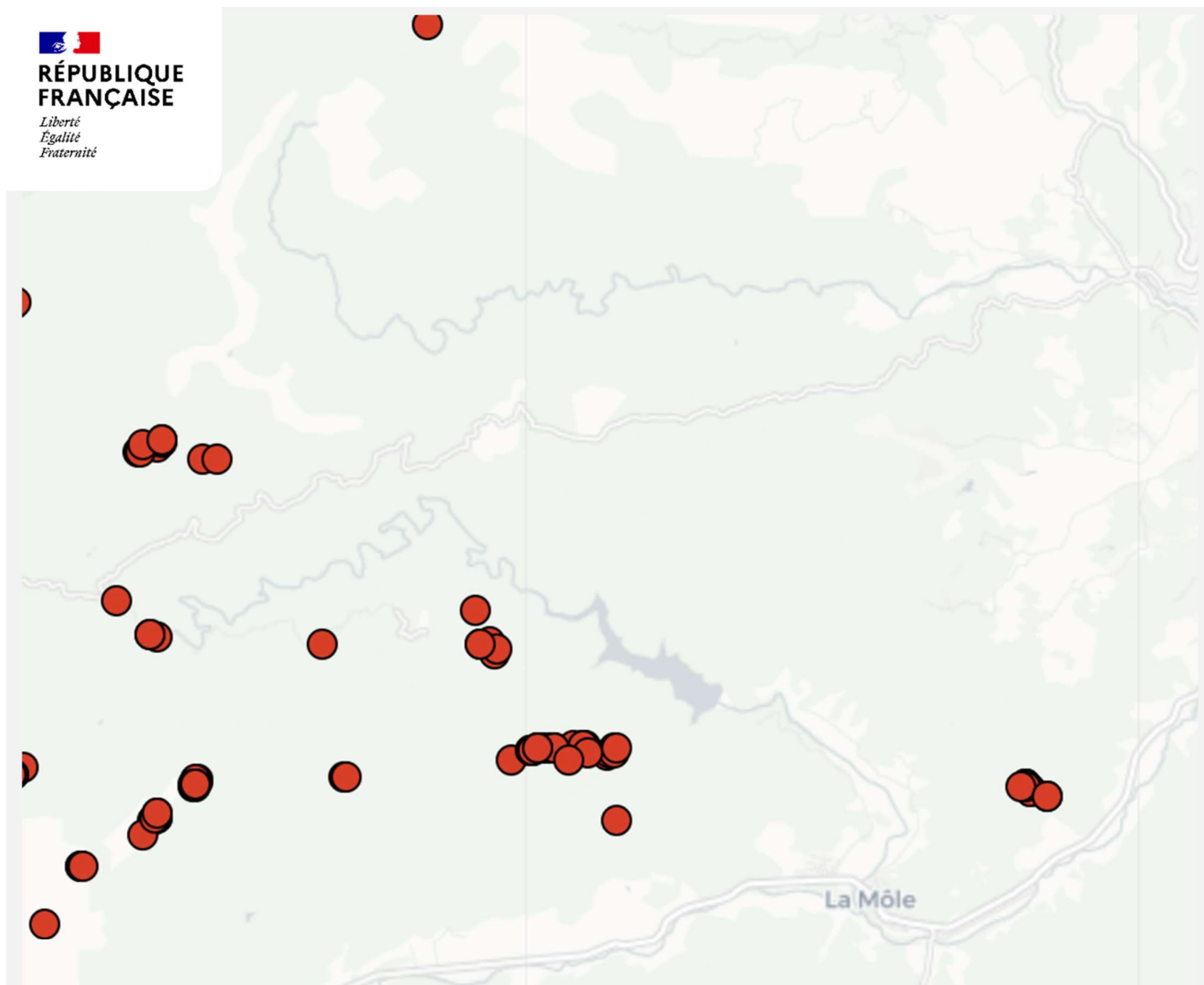




RÉPUBLIQUE
FRANÇAISE

*Liberté
Égalité
Fraternité*

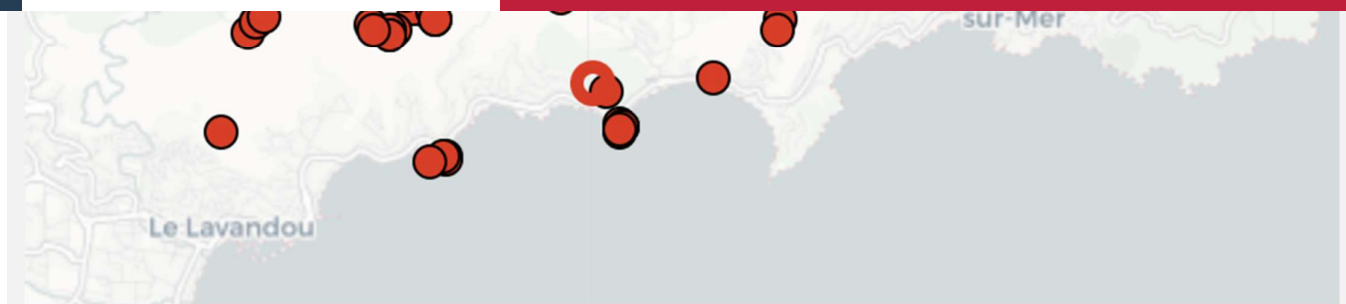


PROCÉDURES INTERNES



CONSERVATOIRES
BOTANIQUEs NATIONAUX
MÉDITERRANÉEN
ALPIN
CORSE

PROCEDURE D'INTEGRATION DES DONNEES D'OBSERVATIONS FLORISTIQUES DANS SIMETHIS



Contexte et application

Objet : Description de la procédure de contrôle, d'intégration et de validation scientifique des données d'observation floristiques que réalisent le Conservatoire botanique national méditerranéen, le Conservatoire botanique national alpin et le Conservatoire botanique national de Corse dans le cadre de leurs missions et de leur utilisation du système d'information SIMETHIS. La procédure vaut également pour le traitement des données dans le cadre des plateformes SINP pour les CBN méditerranéen et alpin.

Périmètre : Système d'information ; mission d'intérêt général « Gestion, diffusion et valorisation de données sur la flore, la fonge, les végétations et les habitats ».

Public cible : Personnel interne CBN ; gouvernance des SINP régionaux ; fournisseurs de données

Liaisons : -

Métadonnées et suivi

Référent de la procédure : Virgile NOBLE – Directeur scientifique du Conservatoire botanique national méditerranéen

Mots-clés :

Gestion de données, Import, Validation, Observations floristiques, SINP, Conformité.

Historique des modifications :

Version	Date	Auteur(s)	Description des modifications	Validé par
1	21/09/2021	Guilhem De Barros	Création initiale de la procédure.	Virgile NOBLE (CBNMed) Jean-Michel Genis (CBNA) Paula Spinosi (CBNC)
1.1	18/11/2025	Guilhem De Barros, Virgile Noble	Actualisation et remise en forme du contenu pour validation CSRPN PACA	Virgile NOBLE

⇒ Avis favorable du CSRPN PACA émit le 27/11/2025 (avis n°2025-30).

Sommaire

PREAMBULE.....	3
1. DEFINITIONS.....	3
2. PROCESSUS D'INTEGRATION.....	4
2.1. SOURCE DES DONNEES	4
2.2. RECEPTION DES DONNEES EXTERNES.....	5
2.3. CONTROLES DES DONNEES	5
2.4. INTEGRATION DES DONNEES DANS SIMETHIS.....	7
3. VALIDATION SCIENTIFIQUE DES OBSERVATIONS.....	11
3.1. PRE-VALIDATION SCIENTIFIQUE AUTOMATIQUE	11
3.2. VALIDATION SCIENTIFIQUE EXPERTE	13
4. Bilan de l'intégration et de la validation scientifique	14
5. SPECIFICITES LIEES AU SINP.....	15
5.1. GENERATION DES UUID.....	15
5.2. GENERATION DES METADONNEES	15
5.3. FLUX DES DONNEES VERS LES SINP REGIONAUX.....	16

PREAMBULE

Ce document décrit le processus d'intégration et de validation des observations floristiques dans SIMETHIS, le système d'information métier des Conservatoires botaniques nationaux alpin, méditerranéen et de Corse.

SIMETHIS comporte un module (FLORE) dédié à la gestion des données d'observations floristiques (flore vasculaire, algues, lichens, champignons) et à leur qualification. Ce module est utilisé à la fois pour les activités propres aux trois conservatoires et pour l'administration des données floristiques des plateformes régionales du SINP :

- **SILENE** : Provence-Alpes-Côte d'Azur
- **BiodivOcc** : Occitanie (ex-Languedoc-Roussillon)
- **Biodiv'AURA** : Auvergne-Rhône-Alpes (départements alpins)

Cet outil garantit la cohérence, la traçabilité et la qualité des données floristiques, dans le cadre des missions de conservation et de diffusion des connaissances.

1. DEFINITIONS

Intégration des données : processus permettant d'insérer les données dans la base réceptrice comprenant la vérification de la conformité et de la cohérence des données, le rattachement des données aux référentiels en vigueur et le contrôle des doublons. Une fois les données importées, elles peuvent passer par une étape de validation scientifique et de qualification et être diffusées dans les différents outils existants et faire l'objet d'analyses.

Validation des données (au sens du SINP) : la validation des données correspond au contrôle de la conformité et de la cohérence et à la validation scientifique des données.

Conformité (au sens du SINP) : la conformité désigne le respect des règles fixées par les standards de données et de métadonnées, autant sur les aspects physiques que conceptuels (renseignement des attributs obligatoires, format, type, utilisation des référentiels et des listes de valeurs). Par exemple, le champ « date » doit obligatoirement être renseigné et doit respecter la norme ISO 8601, comme spécifié dans le standard de données du SINP.

Cohérence (au sens du SINP) : la cohérence désigne le respect de la logique combinatoire des informations transmises (au sein des données, au sein des métadonnées, entre données et métadonnées). Par exemple, la date de début de l'observation doit être inférieure ou égale à la date de fin de l'observation : les champs « date début » et « date fin » sont cohérents entre eux.

Validation scientifique (au sens du SINP) : la validation scientifique est un processus d'expertise visant à renseigner sur la fiabilité de la donnée. On distingue : la validation scientifique dite automatique qui consiste en une validation faisant appel à des résultats d'expertise amont (des référentiels, des bases de connaissance, etc...) et la validation scientifique dite manuelle qui consiste en une validation des informations faisant appel à une expertise aval (avis d'expert basée sur l'analyse des informations transmises).

Qualification (au sens du SINP) : la qualification est l'aptitude à répondre aux usages. Elle est le résultat d'une sélection selon des critères inhérents à l'objectif visé, et est donc variable en fonction de cet objectif. Par exemple, une donnée qui ne comporte pas d'information sur les effectifs (ou autre évaluation quantitative de la population) d'une espèce végétale sera jugée utilisable pour réaliser un atlas (c'est l'information de présence de l'espèce qui est importante) mais non utilisable pour étudier la démographie d'une population.

2. PROCESSUS D'INTEGRATION

2.1. SOURCE DES DONNEES

Les données produites par les Conservatoires proviennent de plusieurs sources :

- **Saisie directe** par le personnel des CBN dans les interfaces web de **SIMETHIS** (observations inédites ou bibliographiques) ;
- **Outils nomades** utilisés directement sur le terrain ;
- **Importation de jeux de données internes.**

Les données **externes** reçues pour intégration proviennent principalement :

- Des **réseaux de botanistes correspondants** des CBN.
- De **partenariats bilatéraux.**
- De l'intégration des jeux de données dans le cadre du **SINP** pour les régions :
 - Provence-Alpes-Côte d'Azur,
 - Occitanie (partie ex-Languedoc),
 - Auvergne-Rhône-Alpes (départements alpins).

Ces observations sources sont principalement sous forme de fichiers tabulés (textes, tableurs) ou de données SIG, mais peuvent aussi être saisies directement depuis un ordinateur ou via des outils nomades. Une faible proportion est transmise sous forme de communications orales ou écrites (figure 1).

Les données saisies via les interfaces web de SIMETHIS ou les outils nomades associés sont immédiatement soumises aux contraintes de conformité, de cohérence et de structure. Elles suivent également les règles de validation (cf. § 3). Le présent document se concentre sur la description détaillée des étapes de traitement des données produites en externe, afin d'assurer leur intégration et leur qualification dans le système.



Figure 1 : Schéma général de réception des données

2.2. RECEPTION DES DONNEES EXTERNES

Les fichiers reçus sont stockés localement sur les serveurs, dans un dossier dédié aux données à intégrer. Un **premier contrôle** est effectué afin de vérifier la présence des informations minimales indispensables :

- **Taxon** observé
- **Date** de l'observation
- **Observateur** et son **organisme d'appartenance**
- **Localisation**

En cas d'informations manquantes, un retour est effectué auprès du fournisseur pour compléter les données.

À ce stade, un contrôle supplémentaire porte sur :

- La présence de l'**identifiant permanent** au format **UUID** (voir § 5.1).
- La présence des **métadonnées**, ou a minima des éléments nécessaires pour les générer (voir § 5.2).

Les éventuels **doutes concernant l'organisme fournisseur** ou l'**origine publique/privée** de la donnée sont également traités à ce moment.

2.3. CONTROLES DES DONNEES

Avant l'intégration dans **SIMETHIS**, chaque observation externe subit trois types de contrôles, conformément au processus de validation des données défini par le **SINP** :

- 1- **Contrôle de conformité et de cohérence** : vérification que les données respectent les formats attendus et que les informations sont logiquement cohérentes (ex. : date valide, taxon existant, localisation plausible).
- 2- **Détection des doublons** : identification des observations déjà présentes dans le système afin d'éviter les redondances.
- 3- **Structuration** : harmonisation des données pour garantir leur compatibilité avec le modèle d'intégration de **SIMETHIS**.

Une fois ces contrôles effectués, les informations sont **rattachées aux référentiels en vigueur** pour permettre leur intégration.

2.3.1. STRUCTURATION DES DONNEES

Lorsque les données sont fournies sous forme de tableur ou de couche SIG, un contrôle spécifique est réalisé sur leur structuration. Une bonne structuration facilite l'intégration et limite les modifications nécessaires, tandis que des données mal structurées exigent un travail important de normalisation.

Exemples de problèmes fréquents

- Dates non formatées correctement.
- Observateurs non individualisés (ex. : plusieurs noms dans une même cellule).

Traitement des données mal structurées

- Si le travail de normalisation est trop important, les données ne seront pas intégrées.
- Le fournisseur sera invité à transmettre des données mieux structurées.

Retour au fournisseur

Si nécessaire, les Conservatoires adressent un retour détaillé au fournisseur, précisant :

- Les points à améliorer.
- Les recommandations pour une structuration conforme.

2.3.2. CONTROLE DE CONFORMITE ET DE COHERENCE

Ce contrôle vise à vérifier la présence des **informations minimales requises** (standards SINP) :

- **Quoi** : Nom scientifique du taxon observé. Les noms vernaculaires doivent être évités, car ils peuvent correspondre à plusieurs noms scientifiques.
- **Quand** : Date de l'observation (au minimum l'année).
- **Qui** : Observateur(s) et, le cas échéant, organisme(s) d'appartenance.
- **Où** : Localisation de l'observation (exemple : coordonnées géographiques).

L'absence d'une de ces quatre informations rend l'intégration impossible.

Les informations complémentaires potentiellement fournies (facultatives) ne sont pas contrôlées (méthode de collecte (*comment*), type d'habitats, effectifs, etc.).

Deux vérifications de cohérence sont réalisées :

1- **Cohérence temporelle (Date)** : la date de l'observation doit être antérieure à la date de réception des données.

2- **Cohérence spatiale (Localisation)** :

- Si la commune est fournie, une analyse spatiale compare la commune citée et la commune calculée à partir des coordonnées.
 - Si elles diffèrent : contact du fournisseur pour définir l'action (correction de la commune, correction de la localisation ou rejet de l'observation)
 - Si elles coïncident : la localisation est qualifiée (voir § 2.4.4).
- Si la commune n'est pas fournie : elle est déduite automatiquement à partir des coordonnées géographiques par croisement géographique avec les référentiels de l'IGN.

2.3.3. RECHERCHE DE DOUBLONS

Ce contrôle vise à déterminer, **avant intégration**, si les observations à traiter sont **déjà présentes en base de données**. La vérification repose sur la comparaison des éléments suivants entre les observations entrantes et

celles déjà enregistrées :

- Taxon observé
- Date de l'observation
- Observateur(s)
- Localisation

Traitement des doublons

Les doublons identifiés par cette méthode **ne sont pas intégrés** en base de données.

Dans certains cas, ils peuvent **provoquer une mise à jour** de la donnée préexistante, en accord avec le **producteur**.

Cas particulier :

Si une même observation initiale est transmise à plusieurs reprises mais avec des éléments de localisation variables (ex. une fois avec une localisation précise, une fois avec une localisation floutée), elle ne pourra être considérée comme un doublon, sauf si d'autres informations évaluées par expertises permettent d'affirmer qu'il s'agit d'un doublon.

2.4. INTEGRATION DES DONNEES DANS SIMETHIS

2.4.1. DELAI D'INTEGRATION

Les délais d'intégration peuvent varier selon l'origine et la nature des données transmises.

Données fournies en fichier : hors problèmes majeurs ou jeux de données très volumineux, le délai d'intégration est fixé à six mois maximum entre la réception et l'intégration.

Observations nécessitant une saisie manuelle : Aucun délai garanti n'est applicable, en raison de la charge de travail variable et des priorités opérationnelles.

2.4.2. FORMATAGE ET RATTACHEMENT AUX REFERENTIELS

L'intégration des observations floristiques dans la base SIMETHIS implique trois opérations principales :

- Le formatage des données ;
- La standardisation des informations ;
- Le rattachement aux référentiels (taxonomiques, géographiques, observateurs, sources, etc.).

Chaque composante de l'observation est traitée comme suit :

Quoi (Taxon) : si le code taxon (cd_nom ou cd_ref) n'est pas fourni : rattachement au référentiel taxonomique TAXREF en vigueur via un outil automatique basé sur la comparaison des noms. Une post-expertise par un expert peut être nécessaire pour les rattachements les moins probables. Dans certains cas, il n'est pas possible de rattacher le taxon au référentiel taxonomique. Le rattachement est alors réalisé sur une valeur temporaire (Taxon à vérifier). Ces données nécessitent l'intervention d'un expert afin d'être rattachées correctement si cela est possible.

Quand (Date) : la date doit être au format AAAA-MM-JJ (année-mois-jour). Si ce n'est pas le cas, elle est convertie. La date d'observation est gérée par deux champs qui permettent de définir une période : date de début et date de fin. Ce système permet de gérer :

- Périodes (ex. : du 18/07/2020 au 27/08/2020),
- Dates incomplètes (ex. : 2019 → 2019-01-01 ; 2019-12-31),
- Dates approximatives (ex. : bibliographie).

Qui (Observateur) : rattachement aux référentiels des observateurs et des organismes.

Où (Localisation) : conversion des coordonnées dans l'un des deux systèmes de coordonnées géographiques acceptés par SIMETHIS (Lambert 93 ou WGS 84). Si la localisation est livrée dans un autre système, elle est convertie.

Comment (Méthode) : si l'information est présente, rattachement au référentiel des protocoles de collecte. Sinon elle est indiquée comme inconnue ou déduite des données par avis d'expert.

Autres champs d'information

En plus des cinq informations principales (Quoi, Quand, Qui, Où, Comment), l'intégration peut donner lieu au renseignement de champs complémentaires, tels que :

- Source de la donnée : organisme fournisseur.
- Type de donnée : bibliographique, herbier, terrain, etc.
- Altitude : obtenue à partir des données d'origine et/ou par croisement avec le modèle numérique de terrain.
- Autres informations utiles pour la qualification et la traçabilité.

2.4.3. CAS DES IMPORTATIONS VIA DES FLUX

Certaines structures partenaires disposent de la capacité de générer un flux de données à partir de leurs bases. La mise en place d'une actualisation des données via un flux présente l'avantage d'automatiser, en temps réel, l'intégration des informations. Elle permet également de gérer les mises à jour en identifiant les différences entre les valeurs existantes et celles du flux.

Les règles suivantes s'appliquent :

- Mise à jour des champs annexes : les champs annexes sont actualisés directement.
- Modification taxonomique : si le cd_ref fourni est différent à la fois du cd_ref d'origine et de celui présent dans SIMETHIS, cela est considéré comme une nouvelle détermination. Cette modification

nécessite donc une nouvelle validation.

2.4.4. QUALIFICATION DES DONNEES

La qualification des données vise à apporter une valeur ajoutée à l'observation floristique en lui associant des précisions supplémentaires sur certains champs. Cette étape facilite l'exploitation ultérieure des données, en offrant de nouvelles possibilités de tri et de filtrage.

Parmi les quatre informations obligatoires, seules les suivantes font l'objet d'une qualification :

- **Quoi (Taxon)** : validation scientifique de la détermination (confirmation ou correction).
- **Quand (Date)** : qualification de la date (ex. : date exacte, période, date approximative issue de bibliographie).

La qualification de la date repose sur le champ booléen **date_averé**, qui indique si la date fournie est fiable :

- **Vrai** : la date ou la période livrée correspond à la période réelle de terrain.
- **Faux** : la date ou la période livrée est déduite approximativement à partir d'autres informations (ex. : bibliographie, estimation).
- **Où (Localisation)** : attribution d'un niveau de précision (ex. : coordonnées exactes, commune, maille).

La localisation géographique d'une observation peut provenir de différentes sources, avec des niveaux de précision variables :

- **Données GPS précises** : précision de quelques centimètres à quelques mètres selon le type de GPS utilisé.
- **Pointage manuel** : sur carte topographique ou photographie aérienne.
- **Description manuscrite** : lieu-dit, commune, massif, etc.
- Chaque méthode influence le **niveau de précision attribué** à la localisation (ex. : coordonnées exactes, maille, commune). Le tableau 1 liste les principales valeurs retenues dans SIMETHIS.

A la localisation est associée la résolution en mètres qui correspond à la distance de précision de la donnée :

- Précision de la mesure du GPS ;
- Précision du pointage sur carte ou sur photographie aérienne ;
- Estimation du rayon du cercle dans lequel peut se trouver l'observation (localisation approximative) ;
- Pour les localisations surfaciques ou linéaires, rayon du cercle englobant le polygone ou la ligne.

Tableau 1 : Les différentes qualifications pour la localisation dans SIMETHIS.

Description	Libellé qualification
Localisation GPS, objet ponctuel. Objet linéaire ou surfacique pour lequel l'observation est en tous points de celui-ci.	Pointage précis(P)

Observation localisée au lieu-dit ou objet linéaire ou surfacique pour lequel l'observation n'est pas en tous points de l'objet.	Pointage approximatif (T)
Observation localisée au centroïde de la commune	Commune (C)

2.4.5. IDENTIFIANT UNIQUE

SIMETHIS gère **trois types d'identifiants uniques** pour assurer la traçabilité et l'interopérabilité des données :

- **Identifiant interne SIMETHIS** : généré automatiquement par le système pour chaque observation intégrée.
- **Identifiant de la base source** : fournit par le producteur. Il permet de retrouver la donnée dans la base d'origine si nécessaire.
- **Identifiant UUID (Universal Unique Identifier)**. Il peut être fournit par le producteur, si sa base permet de le générer ou créé lors de l'import dans SIMETHIS si non fournit. Dans ce cas, l'UUID est renvoyé au producteur pour intégration dans sa base ou archivage.

2.4.6. INTEGRATION

L'intégration des données dans **SIMETHIS** s'effectue via l'outil « **Import par lots** ». Après sélection du fichier et description des données, le système exécute automatiquement les étapes suivantes et **interrompt le processus en cas de non-conformité** :

- 1- **Chargement du fichier** sur le serveur SIMETHIS et **vérification des données** :
 - Présence des champs obligatoires,
 - Encodage correct,
 - Formatage conforme.
- 2- **Analyse de cohérence** des données par rapport aux **référentiels**.
- 3- **Intégration des données** : calcul automatique de l'**altitude** et de la **commune**.
- 4- **Pré-validation scientifique automatique** (voir section suivante).
- 5- **Mise à jour de l'historique des intégrations** et **création d'un fichier bilan** de la pré-validation scientifique.

3. VALIDATION SCIENTIFIQUE DES OBSERVATIONS

La validation scientifique est le mécanisme de contrôle permettant d'attribuer à chaque observation floristique un niveau de validité (tableau 2).

Elle peut être de deux natures :

- **Automatisée** : évaluation initiale de la probabilité de pertinence de l'observation.
- **Expertise** : intervention des experts botanistes référents si nécessaire.

Tableau 2 : Les différentes valeurs pour la qualification du niveau de validité des données dans SIMETHIS et leur correspondance avec les valeurs du SINP.

Code	Libellé SIMETHIS	Équivalent SINP
NULL	Non réalisée	Non évalué
0	En cours de validation	Non évalué
1	Certain	Certain - très probable
2	Probable	Probable
3	Douteux	Douteux
4	Invalide	Invalide
5	Non réalisable	Non réalisable

3.1. PRE-VALIDATION SCIENTIFIQUE AUTOMATIQUE

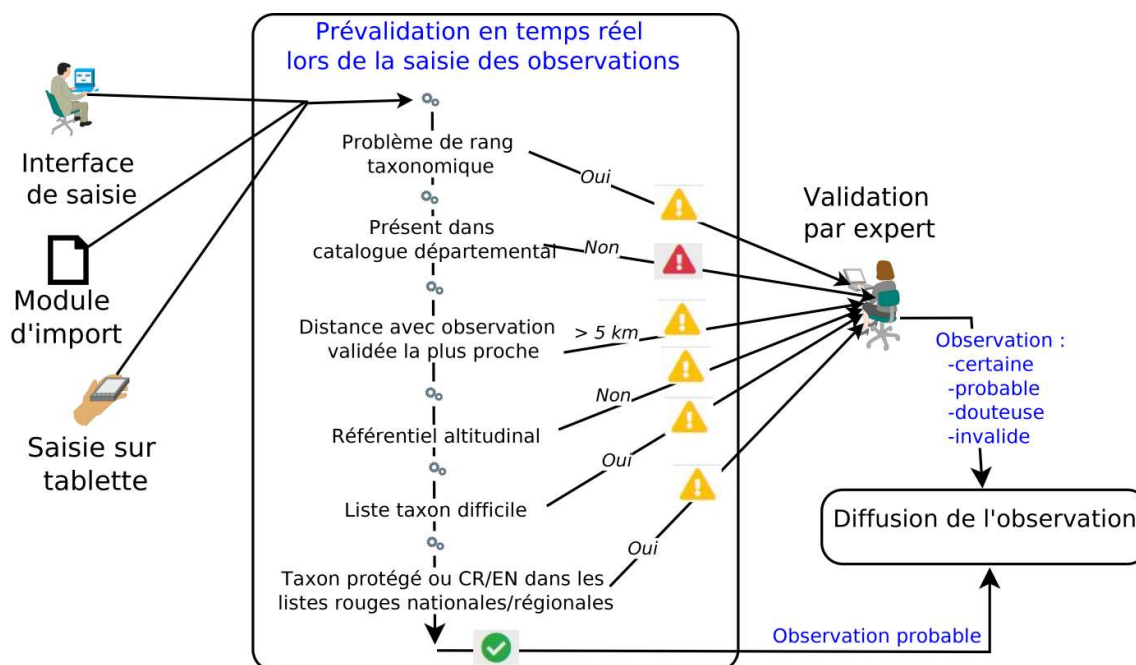
La validation scientifique automatique fait appel à six niveaux d'analyse (figure 2) :

- **Rang taxonomique** : rattachement à un rang spécifique ou infraspécifique.
- **Catalogue départemental** : comparaison avec le catalogue de référence listant la présence des taxons par département.
- **Cohérence territoriale** : vérification si le taxon a déjà été observé récemment (≥ 2000) à proximité du lieu de l'observation. La distance retenue est un rayon de 5 kilomètres pour les plantes vasculaires et 10 kilomètres pour les autres groupes taxonomiques.
- **Cohérence altitudinale** : vérification que l'altitude de l'observation est dans la gamme altitudinale connue pour le taxon (statistique sur les observations valides et précises géographiquement présentes dans SIMETHIS).
- **Détermination** : absence de la liste des taxons complexes à déterminer définie par les experts (listes départementales).
- **Enjeu réglementaire ou de conservation** : taxon non protégé et non fortement menacé (EN, VU dans les listes rouges régionales).

Les listes de taxons pour lesquelles toute nouvelle observation devra systématiquement être contrôlée de manière manuelle par un expert référent sont les suivantes :

- **Taxons protégés** : protection nationale, régionale et sa déclinaison départementale.
- **Taxons très menacés** : dont la cotation UICN régionale (si disponible) ou nationale est dans l'une des catégories UICN suivantes : EX, EW, RE, CR*, CR, EN
- **Taxons de détermination complexe** : listes départementales élaborées à dire d'expert contenant les taxons difficiles à déterminer, mal connus, souvent confondus, etc.

Figure 2 : Schéma général de description des étapes du processus de validation scientifique appliqué par les Conservatoires botaniques nationaux méditerranéen, alpin et de Corse.



Par suite de l'application de la procédure de pré-validation automatique, **deux situations se rencontrent** :

- **Pré-validation positive** : les observations répondant positivement à l'ensemble des critères sont automatiquement tagguées au niveau de validité "probable" et non spécialement considérée par les experts.
- **Pré-validation négative** : les observations répondant négativement à au moins l'un des critères sont orientées vers la validation experte.

3.2. VALIDATION SCIENTIFIQUE EXPERTE

3.2.1. ANALYSE DE L'EXPERT

À la suite de la pré-validation automatique, toutes les observations à contrôler sont analysées par un expert référent qui statuera du niveau de validité en fonction de :

- Ses connaissances ;
- Des éléments de contexte (expérience de l'observateur, période d'observation, conditions écologiques, etc.) ;
- Des données existantes ;
- Des éventuels échanges avec l'observateur.

L'expert dispose des éléments de synthèse des résultats de l'analyse automatique et d'un système d'indicateur (vert / jaune / orange / rouge) pour doser son effort de contrôle :

- Indicateur vert  : taxons répondant à tous les critères de l'analyse automatique mais protégés ou menacés.
- Indicateur jaune  : taxons répondants à tous les critères de la pré-validation automatique mais listés comme posant régulièrement des difficultés (listes des taxons difficiles).
- Indicateur orange  : taxons ne répondant pas à au moins un des critères de l'analyse automatique (hors enjeu réglementaire ou de conservation)
- Indicateur rouge  : taxons non connus ou non confirmés (≥ 2000) dans le département

L'expert pourra décider d'intervenir sur le rattachement taxonomique pour préciser un rang infraspécifique par exemple. Le nom du taxon d'origine, fourni par le producteur, est systématiquement conservé. L'intervention sur le rattachement taxonomique relance le processus de pré-validation automatique.

L'expert décide finalement d'attribuer à l'observation un des niveaux de validité suivant :

- **Certain** : si une preuve de l'observation (part d'herbier, photographies) a été vue par l'expert.
- **Probable** : si les éléments en sa possession lui rendent l'observation crédible.
- **Douteux** : si les éléments en sa possession lui rendent l'observation peu crédible sans toutefois pouvoir le vérifier dans des délais raisonnables).
- **Invalide** : si les éléments de vérification prouvent une erreur (cas de données anciennes publiées par exemple).
- **Non réalisable** : en cas d'absence totale d'éléments pouvant aider à une vérification (nom inconnu ou

non identifiés dans les référentiels taxonomiques en vigueur).

Les alertes rouges font l'objet d'un double contrôle (expert + référent validation). Leur validation entraîne l'actualisation automatique des référentiels départementaux (catalogues).

3.2.2. SUIVI DES MODIFICATIONS ET METADONNEES

Chaque modification réalisée sur le rattachement taxonomique et sur le niveau de validation est systématiquement archivée et associée à la donnée. Cela permet de disposer d'un suivi détaillé de l'historique des modifications du rattachement taxonomique et de documenter les évolutions de la validation scientifique de la donnée.

Figure 3 : Aperçu du suivi des modification du rattachement taxonomique et de la validation scientifique dans SIMETHIS.

Date	Modification	Par
19/03/2025	<i>Daucus carota</i> L., 1753 modifié en <i>Daucus carota</i> subsp. <i>carota</i> L., 1753	MICHAUD Henri
06/07/2023	Validation en probable	ANDRIEU Frédéric
06/07/2023	Envoi en validation	MORANDO Marine

4. Bilan de l'intégration et de la validation scientifique

Dans le cas d'intégration de jeux de données de structures tierces, après intégration et validation des observations transmises, un bilan d'intégration et un bilan de la validation peuvent être produits automatiquement et être transmis sur demande au fournisseur.

Le bilan d'intégration contient les informations générales de suivi suivantes :

- Date de réception ;
- Date d'intégration ;
- Nombre d'observations reçues ;
- Nombre d'observations intégrées ;
- Nombre d'observations par niveau de validation ou fichier bilan de la validation ;
- Problèmes rencontrés et si besoin fichier(s) des observations concernées.

Le bilan de la validation comprend les informations suivantes :

- Id_source : identifiant dans la base de données source
- Id_permanent : identifiant permanent de la donnée au format SINP

- Taxon source : nom du taxon fourni par le producteur
- Cdref_retenu : cd_ref utilisé après validation
- Taxon retenu : nom du taxon après validation
- Modif_rattach_taxon : le rattachement du taxon a-t-il été modifié
- Methode_qualification : qualification manuelle ou automatique
- Resultat_qualification : niveau de validité
- date_validation : Date à laquelle a été faite la validation
- nom_valisateur : Nom du valideur si qualification manuelle

5. SPECIFICITES LIEES AU SINP

5.1. GENERATION DES UUID

L'identifiant unique universel (UUID) de l'observation est une information obligatoire dans le standard occurrence taxon du SINP. Dans le cas où les données fournies ne contiennent pas d'UUID ceux-ci sont générés lors de l'intégration des observations floristiques dans SIMETHIS pour les observations et les relevés comme évoqué au § 2.4.5.

5.2. GENERATION DES METADONNEES

SIMETHIS dispose d'un module de gestion des métadonnées permettant de saisir les informations relatives aux jeux de données (JDD) et aux cadres d'acquisition (CA).

Les informations saisies sont compatibles avec le standard SINP Métadonnées et sont indiquées dans la figure 3.

Ces informations doivent être fournies en même temps que les données. Si elles sont incomplètes, les fournisseurs sont contactés pour complément. Dans le cas où le JDD ou le CA livrés possèdent des identifiants uniques universels (uuid) ils seront intégrés à la place des UUID générés automatiquement.

Figure 3 : Schéma général de description des étapes du processus de validation scientifique appliqué par les Conservatoires botaniques nationaux alpin et méditerranéen.

Nouvelle fiche "Cadres d'acquisition"	Nouvelle fiche "Jeux de données"
Libellé *	Cadre d'acquisition *
Description	Libellé *
Objectif	Libellé court *
Niveau territorial	(30 car. max)
Territoire visé	Description
Objectif	Domaine marin
Précision sur le territoire visé	Domaine terrestre
Contact principal	Type de données
ID du CA parent	Objectif
meta cadre	Territoire visé
Dates	Méthode de collecte
	Origine des données
	Statut de la source
	Contact principal
	Financier du jeu de données
	Producteur du jeu de données

☒ Enregistrer
 ☐ Annuler

5.3. FLUX DES DONNEES VERS LES SINP REGIONAUX

Au sein de SIMETHIS, des fonctions permettant de convertir les données au format des SINP régionaux peuvent être lancées à la demande. Elles permettent le transfert des données soit par la génération de fichiers texte (.csv) soit en écrivant directement dans une base de données par l'intermédiaire de FDW, un système de flux entre base de données.

Dans les 2 cas, la liste des données envoyées est archivée afin de permettre la gestion des insertions, mises à jour, suppressions lors des envois successifs.

Enfin, une table de SIMETHIS archive le nombre de données envoyé vers les SINP pour un suivi quantitatif dans le temps des données transmises (tableau 3).

Tableau 3 : liste des envois effectués vers les plateformes SINP régionales en 2025.

bd_cible	date_export	nb_insert	nb_update	nb_delete
SINP PACA Med	2025-03-07	21275	47750	170
SINP PACA Alp	2025-03-07	10890	32140	1402
SINP AURA	2025-03-13	266444	1276313	307482
SINP OCC	2025-04-01	49902	131120	428
SINP PACA Med	2025-09-09	143770	376597	3368
SINP PACA Alp	2025-09-17	48628	113585	478
SINP AURA	2025-09-18	536495	751727	6855
SINP OCC	2025-11-06	568086	1171848	2903

ANNEXE

Conseil scientifique régional du patrimoine naturel de la région Provence-Alpes-Côte d'Azur		
AVIS N° 2025 - 30		
Date : 27/11/2025	Objet : procédure d'intégration des données d'observation floristiques pour Simethis et le SINP régional Silene	Vote* : favorable

Le conservatoire botanique national méditerranéen (CBN Med) présente le contexte et l'application de la procédure d'intégration des données d'observation floristiques pour Simethis et le système d'information du patrimoine naturel (SINP) régional Silene. Il s'agit de décrire la procédure de contrôle, d'intégration et de validation scientifique des données d'observation floristiques que réalisent le Conservatoire botanique national Méditerranéen, le Conservatoire botanique national alpin et le Conservatoire botanique national de Corse dans le cadre de leurs missions et de leur utilisation du système d'information SIMETHIS. La procédure vaut également pour le traitement des données dans le cadre des plateformes SINP pour les CBN méditerranéen et alpin. Le CBN Med indique qu'environ 70 % des données passent en flux automatique et les 30 % restantes sont validées par des experts. Cela représente environ 50 000 données annuelles par conservatoire.

Le CSRPN s'interroge sur le traitement des données régionales qui sont transmises par le SINP national.

Le CBNMed indique que ces données passent les mêmes filtres que les données régionales, la plateforme nationale ayant son propre système de validation.

Avis 2025-30 :

Suite à ces échanges, le CSRPN émet un avis favorable à l'unanimité sur la procédure d'intégration des données d'observation floristiques pour Simethis et le SINP régional Silene.

Avis :	Favorable ☒	Favorable sous conditions ☐	Défavorable ☐
--------	---	--	--------------------------------------

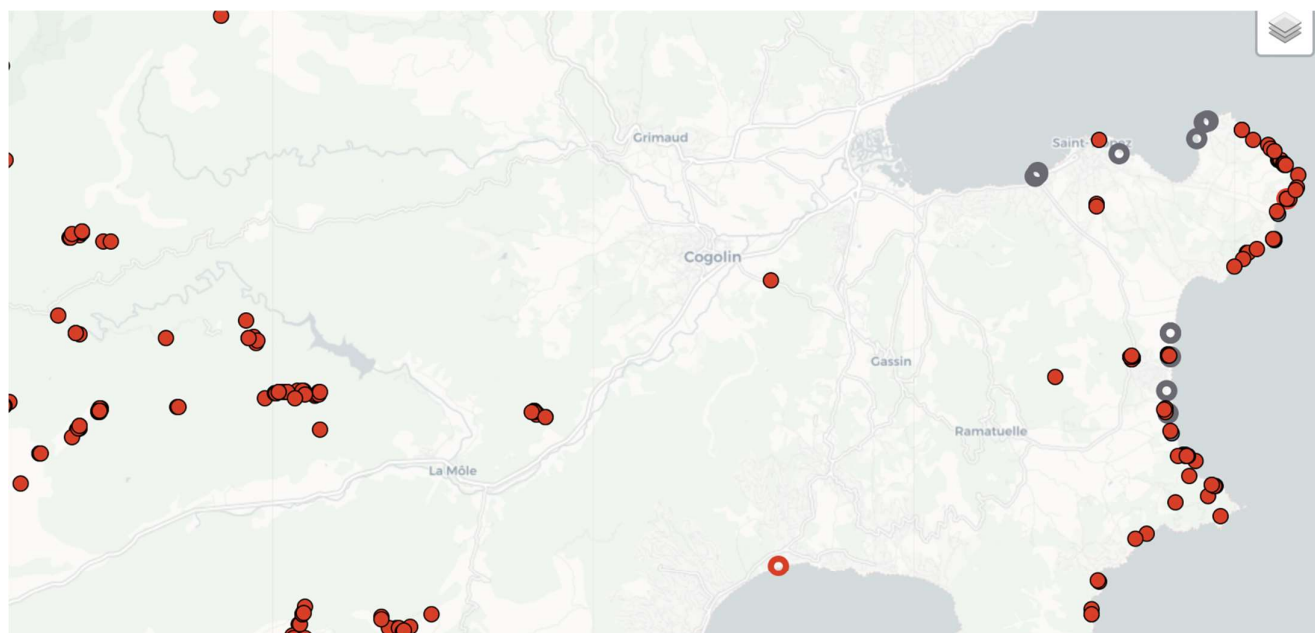
Le Président du Conseil Scientifique
Patrick Grillas



novembre 2025



**CONSERVATOIRES
BOTANIQUE NATIONAUX**
MÉDITERRANÉEN
ALPIN
CORSE



Contact

**Conservatoire botanique
national méditerranéen**
34 avenue Gambetta
83400 Hyères
04 94 16 61 40
contact@cbnmed.fr

**Conservatoire botanique
national alpin**
Domaine de Charance
05000 GAP
04 92 53 56 82
cbna@cbn-alpin.fr

**Conservatoire botanique
national de Corse**
Office de
l'Environnement de la
Corse
14, Avenue Jean Nicoli -
20250 CORTE
04 95 45 04 00
cbnc@oec.fr