



Conservatoire botanique national du Bassin parisien

Une structure au cœur du développement durable

Connaître

Comprendre

Conserver

Communiquer

Notice pour la Validation et l'Interprétation des données à l'échelle du Bassin parisien

DOCUMENT DE TRAVAIL
Version Octobre 2017



Conservatoire **botanique** national du Bassin parisien

Une structure au cœur du développement durable

Connaître
Comprendre
Conserver
Communiquer

Notice pour la Validation et l'Interprétation des données à l'échelle du Bassin parisien

Document de travail

Version Octobre 2017

Ce document a été réalisé par le Conservatoire botanique national du Bassin parisien, délégation Centre-Val de Loire, sous la direction scientifique de

Frédéric Hendoux, directeur

Conservatoire botanique national du Bassin Parisien / Muséum national d'Histoire naturelle

61 rue Buffon CP 53, 75005 Paris Cedex 05

Tel. : 03.86.78.79.60 – Fax : 01 40 79 35 53

E-mail : cbnbp@mnhn.fr

Coordination : Jordane Cordier, Sébastien Filoche & Cyril Tabut

Sommaire

1 - Préalables à la validation.....	4
1-1 – Les champs dénommant une population dans Flora.....	4
1-2 – FLORE DE REFERENCE et REFERENTIEL OFFICIEL.....	4
1-3 – Garder l'historique de validation.....	4
1-4 – Le niveau taxonomique de COMPLETUDE.....	5
1-5 – Principes et types de validation.....	6
1-6 – Codage des types de validation.....	6
1-7 – Notion de Groupe et de taxon « cryptique » incluant une ambiguïté taxonomique.....	7
2. Propositions de règles de validation et d'interprétation de lots de données.....	8
2-1 – Circuit de transmission des demandes règles de validation et interprétation. .	8

1 - Préalables à la validation

Avec l'agrégation nationale des données SI-Flore, la sortie de flores concernant le TAG (Ile de France, Flora gallica, Flore med.), certaines validations sont à coordonner et homogénéiser au niveau du TAG pour plus de cohérence des jeux de données.

Pour la gestion de certaines données (série, ensemble complexe de taxons, évolution taxonomique historique,...), il est nécessaire de convenir de règles de renvois pour ceux qui ne peuvent pas être gérées automatiquement dans Taxref, ce dernier n'étant pas prévu à cet effet.

1-1 – Les champs dénommant une population dans Flora

Correspondance des dénominations des champs affichés sur le site Internet :

Nom des champs (formulaire relevé espèces du site internet du CBNBP)	Nom cité		Nom scientifique	Taxon de référence
Nom des champs (formulaire validation du site internet du CBNBP)	Taxon cité	numéro de taxon	Nom de référence	Nom valide
Nom des champs Tax_Ref V8.0		CD_NOM	NOM_COMPLET	NOM_VALIDE

Tab. : Hétérogénéité d'affichage et correspondance taxref des champs permettant de nommer une population dans Flora.

1-2 – FLORE DE REFERENCE et REFERENTIEL OFFICIEL

L'ouvrage de référence partagé pour le TAG pour la gestion des données est **Flora Gallica**. Des articles postérieurs à Flora Gallica, relayés au niveau national, en particulier dans les référentiels pourront également servir de support partagé à la validation.

Flora Gallica - Flore de France, Coordinateurs : Jean-Marc Tison & Bruno de Foucault. Éditions Biotope, 1216 pages. ISBN 9782366620122

Le référentiel utilisé par le CBNBP est TAXREF v7.0 mise en ligne le 1er décembre 2014 par l'INPN.

1-3 – Garder l'historique de validation

Le CBNBP conserve une trace de la donnée source intacte telle qu'elle a été fournie afin de permettre :

- Gestion d'un historique : traçabilité des corrections (validations) successives ;
- réversibilité des corrections (validations et interprétations) ;
- transmission extérieure (SINP, SI Flore) ;
- identification par le fournisseur des validations du CBNBP.
-

L'historique (nature, date...) des validations est conservé pour chaque donnée modifiée et est visible sur le site du CBNBP (avec login).

Taxon							
Yacon (C6) : Batellia renouillensis (L.) Parl.	(812410)						
A garder							
envoyé par Cordier le 26/01/2014							
Commentaire : D après le diagnostic figurant dans Flora gallica, l'ensemble des populations de Batellia renouillensis se rapporte à cette so-sap.							
Population							
Observation :							
Strate :	Abondance :						
Commentaire :							
Page de citation de l'espèce :							
Surface occupée en m2 :							
Nombre d'individus :	Unité d'échantillonnage :						
Comptage précis :	Dynamisme de la population :						
Indicateur de la station :	Phénologie :						
Besoin de matériel vivant	Risque de herbier						
Type :	Nombre part herbier :						
Echantillonnage :	Herbier de dépôt :						
Nombre d'individus échantillonnés : 0							
Lieu de dépôt :							
Historique des validations							
Date	Taxon (C6)	Description	Personne	Commentaire	Taxon de référence	Statut Yacon	Statut validation
29-04-14	Batellia renouillensis (L.) Parl.	Le Yacon (C6) est correct mais il n'est pas associé à la bonne espèce : il y a une ambiguïté taxonomique(M61)	Cordier	D après le diagnostic figurant dans Flora gallica et les commentaires de répartition de Flora gallica, l'ensemble des populations de Batellia renouillensis se rapporte à cette so-sap.	Batellia renouillensis subsp. nevadensis Molina-Alca, A.Gálvez, Pizarro & Bard-Klein, 1994(812413)	A garder	Efficacité
Dernière entrée	Batellia renouillensis (L.) Parl.				Batellia renouillensis (L.) Parl., 1854(815486)		

Fig. : Visualisation actuelle sur le site Internet de la validation au niveau de la population.

1-4 – Le niveau taxonomique de COMPLETUDE

Le niveau de précision taxonomique support de la complétude en termes d'espèces et de données a été défini au niveau de la **sous-espèce**.

Pour les espèces représentées par une seule sous-espèce, il sera préconisé de nommer ou saisir au niveau de l'espèce afin de ne pas apporter de fausses précisions de détermination. Les règles d'interprétation permettront de préciser la sous-espèce pour la consultation des données à la consultation (c.f. 1.5).

Pour les variétés et formes, les règles d'interprétation renvoyant vers ses niveaux ne seront pas mises en place (sauf évolution particulières du référentiel) et le renseignement de ces niveaux sur le terrain à la libre appréciation de l'observateur.

1-5 – Principes et types de validation

La **validation taxonomique** repose sur l'usage du Taxref. Elle vise :

- une **nouvelle traduction du [Nom cité¹]** conduisant à la modification du [CD_Nom] associé au **[Nom cité]**, par attribution d'un nouveau [CD_Nom] validé (le plus souvent correspondant à celui d'un nouveau Nom_Valide) déterminé par les experts du CBNBP ;
- des **corrections du [Nom cité]** et/ou [CD_Nom] liées à une relecture ou à la demande de modifications de son observation par un observateur.

Deux outils de validation opérationnels au CBNBP selon le volume de données concernées :

- Donnée ponctuelle ou plusieurs données ne pouvant faire l'objet d'une règle commune : utilisation du **module internet de validation manuelle** effectué par le référent flore ou responsable de chaque délégation ;
- Lots de données faisant l'objet d'une même action de validation taxonomique : proposition d'une **règle de validation** aux coordinateurs du Groupe de travail flore interne du CBNBP).

Il est également mis en place **des règles d'interprétation taxonomique** qui interviennent pour préciser un rang taxonomique infraspécifique particulier quand un seul taxon infraspécifique est reconnu pour le TAG ou dans une dition du TAG. Il permet d'apporter une précision supplémentaire à la donnée et d'améliorer les transmissions et affichages, plus particulièrement pour le SI Flore et le SINP. L'interprétation taxonomique est réversible et ne fera pas l'objet d'une conservation historique à l'échelle de la donnée mais pourra faire l'objet d'un historique au niveau de la règle.

L'interprétation peut être aussi employée temporairement comme phase de test et évoluée vers une validation.

Ces règles d'interprétation et de validation permanentes s'appliquent automatiquement aux données entrantes (ex pour les données entrantes de *Stipa pennata*, validation automatique vers *Stipa gallica* Celak.).

¹ Nom du taxon cité à l'origine par l'observateur

1-6 – Codage des types de validation

La validation conserve les principes établis par le document : Poncet L. & Vest F., 2005. Validation en ligne des données Flora. CBNBP, 7p.

Pour rappel, elle vise :

- la correction du taxon cité après relecture (**S2 ; M2**) ou interrogation de l'observateur (**S1 ; M1**) ;
- la saisie du taxon de référence auquel est associé le taxon cité : c'est la liaison qui est faite avec le référentiel. Le forçage doit être motivé (**MR**).
- Doute sur la dénomination ou présence de l'espèce (**D**).

1-7 – Notion de Groupe et de taxon « cryptique » incluant une ambiguïté taxonomique

Le groupe permet de gérer les données imprécises d'observations d'ensemble complexe de taxons « cryptiques ». Il permet à l'interrogation de voir :

- au niveau de Groupe : l'ensemble des données agglomérées de cet ensemble ;
- au niveau de chacun des taxons composant le groupe : l'ensemble des données **certaines et validées** de chaque taxon précis.

Le groupe est un niveau sans valeur taxonomique. Il est créé pour résoudre des contraintes de gestion de données non résolues avec l'emploi de Taxref (dont ce n'est l'usage initial). La création d'un groupe entraîne la création d'un identifiant CD_Nom unique propre au CBNBP.

Par conséquent, les données au niveau du Groupe ne peuvent être transmises et partagées (= **données absentes ou non valides**). Les données pour les taxons précis peuvent être transmises car rattachées à un CD_REF officiel de taxRef.

L'usage des groupes nécessite donc la validation impérative des données attribuées aux taxons précis qui le composent.

2. Propositions de règles de validation/ interprétation de lots de données et de prise de note associée

2-1 – Circuit de transmission des demandes règles de validation et interprétation

Les validations réalisées à l'aide de l'application internet sont effectuées par les référents flore et les responsables de délégation.

Pour les règles affectant un lot de données, la demande comportera les informations suivantes :

- Colonnes détaillant le lot de données à valider

CD_Nom	CD_REF (Taxref_7)	Taxon (Taxref 7)	type	code de validatio n	observateur	date début	Date fin
--------	----------------------	---------------------	------	---------------------------	-------------	---------------	-------------

- CD_Nom à modifier (Taxref 7) ;
- CD_REF avant validation (Taxref 7) ;
- Taxon = Nom_Valide (Taxref 7) ;
- Type **RE-Interprétation vers** ou **Validation vers** ;
- Code de validation (S1, S2, MR1, MR2, MR3, D1, D2, D3) ;
- Observateur ;
- Période d'observation.

- Colonnes précisant la nouvelle dénomination et sa justification

Remarques	Nom cité New	CD_Nom	CD_REF	Nom_Valide
-----------	-----------------	--------	--------	------------

- **Remarques** : commentaire motivant la demande avec précision de l'ouvrage de référence ou nom de l'expert. Nom de la personne proposant la règle ;
- **Nom cité New** : seulement pour les validations de types S1, S2, M1 & M2 ;
- **CD_Nom** : nouveau CD_Nom (Taxref 7) ;
- **CD_REF** : nouveau CD_REF en lien avec le Cd_Nom (ne pas forcer les renvois du TaxREF) ;

- **Nom_Valide** (Taxref 7) en lien avec le CD_REF.

Procédure d'établissement des règles :

- 1- Production/révision/annulation des règles :** les référents bota du CBNBP
- 2- Envoi aux coordinateurs GT Flore** pour validation et vérification de la cohérence de territoire, voire la proposition d'extension de la proposition à d'autres territoires du TAG ;
- 3- Transmission au SI pour mise en œuvre.**

2-2 – Règles générales de validation et d'interprétation

Il s'agit davantage de principes partagés. Des règles pourront être proposées sur ce principe et feront l'objet d'un accord par les correspondants régionaux flore.

Gestion des niveaux infraspécifiques

Cas 1 : Espèces présentant une seule sous-espèce sur le territoire national.

- **REGLE D'INTERPRETATION** au niveau infraspécifique (MR1) ;
- **PRISE DE NOTE** : au niveau de l'espèce, en raison le plus souvent de clés de détermination généralement indisponibles (diagnose absente de Flora Gallica).

Cas 2 : Espèces présentant une seule sous-espèce sur le territoire d'agrément, (ou au niveau régional) mais plusieurs taxons infraspécifiques au niveau national

- **Cas 2.1 : Niveau infra spécifique reconnu par Flora Gallica :**
 - o **REGLE D'INTERPRETATION** au niveau infraspécifique (MR1) ;
 - o **PRISE DE NOTE** : Les deux niveaux sont proposés (liste à cocher) afin de rendre compte de ce qui est réellement déterminé.
- **Cas 2.2 : Niveau infra spécifique non reconnu ou contesté par Flora Gallica :**
 - o **Pas de règle d'interprétation automatique** de l'ensemble des données espèces au niveau infraspécifique.
 - o **PRISE DE NOTE** : Les deux niveaux sont proposés (liste à cocher) afin de rendre compte de ce qui est réellement déterminé.

Cas 3 : Espèces présentant **plusieurs sous-espèces sur le territoire d'agrément** (ou au niveau régional) :

- **Cas 3.1 : Niveau infra spécifique reconnu par Flora Gallica :**
 - **REGLE D'INTERPRETATION ou VALIDATION :** Validation via l'application Internet ou des règles après consultation en interne pour améliorer le jeu de données actuel (ex : *Plantago major*...) par des interprétations ou validations par observateurs et/ou par dates (ex : *Plantago major*, renvoi à la sous-espèce major lorsque les observateurs utilisent *Plantago intermedia* pour l'autre taxon). (M1)
 - **PRISE DE NOTE :** encourager le renseignement systématique des niveaux infraspécifiques (proposition des sous-espèces communes dans la liste à cocher pour les espèces banales). Espèce figurant en italique pour signifier la présence d'autres infrataxons non présents dans la liste à cocher.
- **Cas 3.2 : Niveau infra spécifique non reconnu ou contesté par Flora Gallica :**
 - **Pas de règle d'interprétation** de l'ensemble des données espèces au niveau infraspécifique.
 - **PRISE DE NOTE :** seule l'espèce est citée. Signaler par une « * » l'existence d'autres infrataxons non présents dans la liste à cocher.

Cas 4 : Espèces présentant **plusieurs variétés ou formes sur le territoire d'agrément** :

- **Pas de validation/interprétation automatique** des données (sauf cas très particulier comme *Valerianella locusta* f. *locusta*). (MR1, M1)
- **PRISE DE NOTE :** seule l'espèce est citée.

Gestion du niveau espèce

Cas 5 : Complexes d'espèces (valides selon Taxref) présentant **un seul taxon** confirmé **sur le territoire national ou sur le territoire d'agrément** (*Stipa gallica/pennata*, *Veronica scheererii/prostrata*, ...).

- **REGLE DE VALIDATION** des données sous l'espèce valide selon Flora Gallica (MR1).
- **PRISE DE NOTE :** Seul le taxon valide est proposé dans la liste à cocher pour les plus banals. En cas de détermination précise, permettre l'identification de celle-ci par

le [Nom cité] : sur un BIG en indiquant obligatoirement *sensu stricto* à la suite du nom ou par l'emploi d'un BEE en précisant l'auteur de la détermination et l'ouvrage.

Cas 6 : Complexes d'espèces (valides selon Taxref) présentant **plusieurs taxons possibles sur le territoire d'agrément valides selon taxref** dont un seul est reconnu comme valide selon Flora Gallica (indiqués comme inclus ou à distinction non justifiée) (*C. stoebe/maculosa* ; *Ulmus minor/carpinifolia*) **hors séries, agrégats...**

- **REGLE DE VALIDATION** des données sous l'espèce valide selon Flora Gallica (MR1).
- **PRISE DE NOTE** : Permettre la prise de notes avec les deux dénominations. Seul le taxon valide est proposé dans la liste à cocher pour les plus banals. En cas de détermination précise, permettre l'identification de celle-ci par le [Nom cité] : sur un BIG en indiquant obligatoirement *sensu stricto* à la suite du nom ou l'emploi d'un BEE en précisant l'auteur de la détermination et l'ouvrage.

Cas 7 : Complexes d'espèces (valides selon Taxref) présentant **plusieurs taxons confirmés sur le Territoire d'agrément et valides selon Flora Gallica.**

- **Cas 7.1 : absence d'autres espèces dans le genre au niveau au TAG (ex *Elodea*, *Reynoutria*)**
 - o **VALIDATION** : Renvoi au genre des données jugées imprécises (Niveau supérieur) (MR1 ou M1 ou D) ;
 - o **PRISE DE NOTE** : pour la liste à cocher, identifier un nom générique (genre ou Groupe) et un renvoi à la saisie vers le CD_Nom du genre, suivi des taxons les plus banaux.
- **Cas 7.2 : présence d'autres taxons dans le genre au niveau du territoire d'agrément (*Arenaria serpyllifolia/leptoclados* ; *Carex viridula*, *Carex spicata*), inclus les séries (*Taraxacum*), les agrégats (*Centaurea decipiens*)....**
 - o **INTERPRETATION** : Création d'un groupe nécessaire (MR1 ou M1 ou D).
 - o **PRISE DE NOTE** : Si cet unique taxon appartient à un groupe, permettre l'identification par le [Nom cité]. Dans la liste à cocher le groupe et les taxons les plus banals. En bas de bordereau, les mentions déterminées avec précision suivies obligatoirement de *sensu stricto*. Recommander l'emploi du BEE pour les genres les plus complexes (*Rubus*, *Taraxacum*...)

Attention pour les taxons extrêmement communs: *Chenopodium album*, *Arenaria serpyllifolia*, *Stellaria media*, *Galium palustre*, *Ornithogalum divergens*, *Hypericum maculatum*, *Veronica hederifolia*, *Leucanthemum ircutianum* pour lesquels un groupe a été créé, lors de l'interrogation de la base de donnée, ou lors de la transmission des données, les cartes de ces taxons peuvent apparaître comme vides (ou presque) alors qu'ils sont considérés comme très communs dans les flores et atlas. Afin de limiter cet impact, il est important du moins en interne que la détermination précise de ces taxons soit maîtrisée quand elle est accessible. Une signalétique particulière devra être mise en place au niveau du site internet pour l'ensemble des taxons appartenant à un groupe afin de signaler les limites des cartes de répartition proposée et le meilleur niveau de consultation de la base de données.

Références citées :

Laurent Poncet & Frédéric VEST Vest, 2015. Validation en ligne des données Flora, version 4. CBNBP. 7p.

Flora Gallica - Flore de France, Coordinateurs : Jean-Marc Tison & Bruno de Foucault. Éditions Biotope, 1216 pages. ISBN 9782366620122