

Pré-atlas des longicornes de France métropolitaine :
Note méthodologique sur la validation des données
 J. Touroult, 15 mars 2019

Contexte

Dans cette étape de validation (et d'invalidation des erreurs manifestes), nous disposons :

- 1) de cartes par maille 10 X 10 km, permettant de visualiser la répartition des données sur un fond avec le relief et les départements
- 2) de tableaux reprenant l'ensemble des données.

L'ensemble des données servant au pré-atlas est disponible, c'est à partir de ces données que seront construites les cartes finales.

Le but est d'identifier :

- a) les données qu'il ne faut pas prendre pour la réalisation de ces cartes [validation au sens stricte]
- b) les données qu'il faut signaler car non représentative de la distribution, sans être fausses ou manifestement erronées (cas d'introduction accidentelle hors répartition). [choix de représentation de l'information dans le pré-atlas]

Délai : pour fin avril 2019 au plus tard.

a) La validation utilise les catégories prévues au SINP : colonne U « NIVEAU_VAL »

| Code validité SINP | Libellé | Utilisation / interprétation dans le cadre du pré-atlas des longicornes |
|--------------------|-------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 1 | Certain - très probable | A utiliser ponctuellement quand la donnée a fait l'objet d'une vérification auprès de l'observateur, quand elle vient d'une source vérifiable et qu'on a vérifié (exemple donnée signalée dans une publication avec illustration...), les données issues de révision récente... Exemple : une observation dans SPIPOLL dont on a été vérifié la photo, les données re-déterminées récemment de <i>Stenostola</i> ... |
| 2 | Probable | A utiliser par défaut quand la donnée est plausible au regard de la chorologie d'ensemble (aucun élément discordant). A utiliser largement dans le cas d'espèces communes et/ou faciles à déterminer, de données ayant déjà fait l'objet d'une validation (atlas/catalogue régional etc.). |
| 3 | Douteux | Donnée étrange (hors répartition et/ou à période décalée) mais non vérifiable et qui ne peut être totalement exclue. Ne concerne pas des données non représentative de la distribution mais qui sont exactes/possibles (introduction via le bois, implantation temporaire en plaine etc.) |
| 4 | Invalide | Données totalement aberrante (chorologie, écologie...) ou donnée vérifiée et erronée ou erreur déjà signalée (littérature...) Ne concerne pas des données non représentative de la distribution mais qui sont exactes/possibles (introduction via le bois, implantation temporaire en plaine etc.) |
| 5 | Non réalisable | A utiliser pour les espèces splittées récemment, quand on ne peut rien dire des données anciennes : concerne par exemple la majorité des <i>Leiopus nebulosus</i> , <i>Agapanthia cardui</i> , les deux <i>Stenostola</i> ... |
| | Non évalué | |

A retenir : les données de validité 3 (douteux) et 4 (invalide) ne seront pas utilisées du tout pour la réalisation des cartes du pré-atlas.

b) Autres données valides mais non représentatives de la distribution

Ce cas concerne les catégories de validation 1 (certain – très probable) ou 2 (probable) mais dont la donnée ne représente pas l'enveloppe écologique de distribution de l'espèce (espèce de montagne en plaine, ou

méditerranéenne dans le nord etc.). Il peut s'agir d'introduction temporaire (résineux de plaine), d'introduction avec le bois de chauffage etc.

Ces informations restent intéressantes pour le projet car pouvant témoigner d'une future extension de la distribution mais elles devront relever d'une sémiologie différente dans le pré-atlas.

Utiliser la colonne Z, avec une des valeurs suivantes (utiliser cette nomenclature)

- introduction ponctuelle, sans implantation durable
- implantation éventuelle à confirmer
- observation à confirmer.

On ne traitera pas dans cette rubrique les introductions durables qui sont considérées comme appartenant désormais à la distribution.

Procédure conseillée

Chaque co-auteur à un lot d'espèces à étudier, dont il doit aussi rédiger la notice. C'est sur ces espèces que chacun doit concentrer son effort mais il est possible d'examiner aussi les autres cartes pour détecter les données douteuses...

Pour cette validation nationale, intervenant dans beaucoup de cas après déjà une validation par le producteur des données (dans le cadre d'étude régionale, du programme de recherche etc.), l'effort de rechercher les données douteuses, invalides ou valides hors répartition doit être proportionné à la rareté de l'espèce. Il est par exemple vain de chercher à vérifier des données de *Rutpela maculata* ou de *Clytus arietis* mais utile de passer un peu de temps sur les *Necydalis* etc.

- 1) S'imprimer son lot de cartes (2 par pages) pour faire un aller retour simple entre la carte et les tableaux de données.
- 2) Pré-remplir la colonne U avec le statut 2 (probable) et la colonne W avec votre Nom Prénom et indiquer la date
- 3) Examiner chaque carte au regard de votre expertise sur le taxon
- 4) Pointer les éventuelles données « étranges », hors répartition, en marge, à une altitude inattendue etc.
- 5) Rechercher chacune de ces données « étranges » dans le tableau (INPN ou CardObs), en filtrant sur l'espèce et le département concerné,
- 6) A partir des informations disponibles (observateur, source, date d'observation, jeux de données), estimer le cas de figure, le cas échéant en recontactant l'observateur...
- 7) Reporter vos conclusions dans la colonne U (1 – certain, 3 – douteux, 4 – invalide, 5 – Non réalisable) ou dans la colonne Z pour signaler les données non représentatives de la distribution
- 8) Indiquer une justification aux invalidations dans la colonne Y, ou au statut 1 (certain) quand vous avez vérifié la source de la donnée.