



RAPPORT DE FAISABILITE
POC Nord
Programme CarHAB

CONSERVATOIRE BOTANIQUE
NATIONAL DE BAILLEUL

Octobre 2019





RAPPORT DE FAISABILITE

POC Nord

Programme CarHAB

**Chef de projet
et rédaction** Geoffroy VILLEJOUBERT
Emmanuel Catteau
Collaborateurs (EVS-Isthme, Irstea)

Prospections Geoffroy VILLEJOUBERT
Emmanuel Catteau
Audrey PACQUET

Relecture Benoit TOUSSAINT
Thierry CORNIER

Extraction des données Christophe MEILLIEZ
Florian WATRIN

Cartographie Geoffroy VILLEJOUBERT

**Direction et coordination
scientifiques** Thierry CORNIER
(Directeur général)

Photo couverture

G. VILLEJOURBERT

Référence à utiliser pour toute citation de l'étude

VILLEJOURBERT, G., CATTEAU E. & Collaborateurs, 2019. – Rapport de faisabilité Poc Nord – Programme CarHAB. Conservatoire botanique national de Bailleul. 57 p.

SOMMAIRE

Contexte de l'étude	7
Parc naturel régional Scarpe-Escaut.....	8
Parc naturel régional Avesnois	8
Méthodologie	11
Processus déployé pour la carte des physionomies.....	12
Processus déployé pour la carte des biotopes	13
Volume de données.....	23
Tests méthodologiques mis en œuvre	25
Nomenclatures cibles des cartographies	28
Typologie des biotopes (Tableau 10)	28
Typologie des physionomies (Tableau 11)	29
Description succincte des données en entrée	31
Répartition spatiale des données d'entraînement fournies	32
Cartographie des biotopes	34
Zones potentiellement humides (ZPH).....	35
Sensibilité des données radar à l'hydromorphie des prairies du PNRSE	37
Résultats	38
Méthode d'évaluation des résultats mise en place sur le POC.....	38
Évaluation des cartes des biotopes et ZPH produites	39
Évaluation de la carte prédictive des biotopes	41
Évaluation de la carte prédictive des physionomies	43
Évaluation de la sensibilité des radars à l'hydromorphologie des prairies	44
Évaluation de la carte prédictive des physionomies	44
Évaluation des charges et des coûts	48
Coûts liés à la production des données :.....	48
Présentation de cas d'usage des données produites dans l'étude	53
Méthode 1 : exploitation de la carte des zones potentiellement humides	54
Méthode 2 : exploitation du paramètre « gradient hydrique » de la carte des biotopes	55
Résultat.....	56
Points de vigilances identifiés pour le passage à l'échelle	56
Bibliographie.....	57

Figure 1 : Schéma simplifié du processus de production des modélisations cartographiques pour le POC Nord.....	12
Figure 2 : Schématisation des données d'entraînement utilisables pour la modélisation des biotopes	18
Figure 3 : Illustration de données "ponctuelles" fournies biaisant la modélisation (beaucoup de ces données correspondent à des végétations de chemin, de bord de route, de cours d'eau...)	21
Figure 4 : Résultat de la modélisation carte des biotopes VO issue des données "ponctuelles"	21
Figure 5 : Nombre de biotopes modélisables pour la première version du jeu de données d'entraînement (ponctuelles et linéaires, à gauche) et la seconde version (surfaciques, à droite)	25
Figure 6 : Données d'entraînement surfaciques produites par le CBNBL pour réaliser la modélisation prédictive des biotopes sur le PNRSE	32
Figure 7 : Données d'entraînement produites par le CBNBL pour réaliser la modélisation prédictive des physionomies sur le PNRSE	32
Figure 8 : données d'entraînement surfaciques produites par le CBNBL pour réaliser la modélisation prédictive des biotopes sur le PNRAV.....	33
Figure 9 : Données d'entraînement produites par le CBNBL pour réaliser la modélisation prédictive des physionomies sur le PNRAV	33
Figure 10 : Tableau récapitulatif des efforts méthodologiques réalisés sur le POC	34
Figure 11 : Représentation des données en Mode vecteur et Mode Raster	34
Figure 12 : Schéma de la méthode ZPH déclinée sur le PNRSE	36
Figure 13 : Zones potentiellement humides PNRSE	36
Figure 14 : Carte prédictive des niveaux d'humidité sur le PNRSE	37
Figure 15 : Échelles d'acidité et d'humidité	38
Figure 16 : Carte prédictive des biotopes VO	39
Figure 17 : Carte prédictive des biotopes VO du PNRSE, surreprésentation des biotopes aquatiques (32404, 33556) et mésophiles (25492)	40
Figure 18 : Carte prédictive des biotopes du PNRSE, V3.....	41
Figure 19 : Concordance du biotope, des niveaux d'humidité et d'acidité selon la modélisation en fonction de l'observation terrain sur le PNRAV	42
Figure 20 : Matrice de confusion de carte produite pour un niveau de nomenclature de niveau 4 pour (a) l'Avesnois et (b) Scarpe-Escaut. 2110 Minéral végétalisé indéterminé ; 2120 : Minéral non végétalisé indéterminé ; 3111 Pelouse permanente vivace ; 3210 : Herbacée haute indéterminée ; 3221 : Prairie fauchée ; 3222 Prairie pâturée; 4210 : Arbustif / Fourré / Fruticée mixte indéterminé; 5210 Forêt Semi-naturelle.....	45
Figure 21 : Matrice de confusion de carte produit pour un niveau de nomenclature de niveau 2 pour (a) l'Avesnois et (b) Scarpe-Escaut. 21 : Minéral ; 31 Pelouse ; 32 : Prairie et Herbacée Haute; 42 : Arbustif / Fourré / Fruticée ; 52 : Forêt mature	46
Figure 22 : Concordance entre la modélisation des physionomies et la donnée terrain	47
Figure 23 : Carte des niveaux d'humidité potentiels modélisés sur le PNRSE	55

Tableau 1 : Organismes et interlocuteurs impliqués dans le projet CarHAB pour le « POC Nord »	7
Tableau 2 : Caractéristiques environnementales communes aux territoires étudiés	9
Tableau 3 : Nombre de données d'entraînement fournies pour la modélisation des physionomies...	12
Tableau 4 : Informations sur les différentes versions des modélisations biotopes et zones potentiellement humides.....	13
Tableau 5 : Variables environnementale retenues pour le POC Nord	15
Tableau 6 : Valeurs de températures et de précipitations utilisées dans le calcul des paramètres environnementaux, l'exemple de Maubeuge	16
Tableau 7 : Volume de données considérées pour les données entraînement biotopes première version (< 1ha) et pour la seconde version (> 1ha).....	23
Tableau 8 : Volume de données retenues pour le premier jeu de données d'entraînement	24
Tableau 9 : Volume de données exploitées pour le second jeu de données d'entraînement	24
Tableau 10 : Biotopes modélisés pour la V1 (« ponctuelles » et linéaires) et la V2 (surfaciques) parmi les 36 biotopes retenus sur le territoire.....	28
Tableau 11 : Nomenclature des cellules paysagère de niveaux 4.....	30
Tableau 12 : Description des données d'entraînement	31
Tableau 13 : Explication de l'attribution de la valeur de concordance pour les biotopes et les physionomies	41
Tableau 14 : Nombre de données terrain analysées selon la physionomie modélisée. Les masques sont ceux de l'IGN et ne sont pas modélisés.....	47

Contexte de l'étude

Organismes et interlocuteurs impliqués dans l'étude du « POC Nord »

Les partenaires impliqués dans cette étude sont listés dans le Tableau 1 :

Tableau 1 : Organismes et interlocuteurs impliqués dans le projet CarHAB pour le « POC Nord »

Structure	Personnes et fonctions
Conservatoire botanique national de Bailleul (CBNBL)	<p>Chefs de projet local « POC NORD » Geoffroy VILLEJOUBERT, Emmanuel CATTEAU Stagiaire 2019 : Audrey PAQUET</p> <p>Fonctions Mise en œuvre des expérimentations, élaboration de notes méthodologiques, productions des données d'entraînement (biotopes et physionomies), contrôle des produits cartographiques réalisés (SIG et terrain), participation aux diverses réunions (Groupes de travail (GT1, GT2, GT3), Cotech, etc.)</p>
MNHN / UMS PatriNat – CNRS - AFB	<p>Chefs de projet national pour le « POC NORD » Joanie CATRIN, Rémy PONCET</p> <p>Fonctions Cadre scientifique des processus et méthodologies, mise au point de référentiels communs, méthodologie de modélisation, de vérification, participation aux diverses réunions (Groupes de travail (GT1, GT2, GT3), Cotech, etc.)</p>
Agence de l'eau Artois-Picardie	<p>Chefs de projet Programme de maintien de l'agriculture en zones humides (PMAZH) Estelle CHEVILLARD, Cécile GALLIAN</p> <p>Fonctions Encadrement du programme de maintien de l'agriculture en zones humides</p>
PNR Scarpe-Escout	<p>Encadrement projet PMAZH à l'échelle du PNR Aurore DLUGON, Christelle PARMENTIER, Gérald DUHAYON</p> <p>Fonctions Partenaires de l'étude, fourniture d'informations concernant les Parcs (cartographies, données) Interlocuteurs principaux pour le Parc</p>
PNR Avesnois	<p>Encadrement projet PMAZH à l'échelle du PNR Aurélien THURETTE, Fabien CHARLET, Guillaume DHUIEGE</p> <p>Fonctions Partenaires de l'étude, fourniture d'informations concernant les Parcs (cartographies, données) Interlocuteurs principaux pour le Parc</p>
IGN	<p>Encadrement de projet CarHAB2 Bénédicte MAISONNEUVE, Sophie REYNARD</p> <p>Fonctions Production des couches de segmentation, croisements biotopes physionomie, arbres de décision Participation aux diverses réunions (Groupes de travail (GT1, GT2, GT3), Cotech, etc.), pilotage du GT1</p>

Université Jean Monnet, SAINT-ETIENNE (Laboratoire EVS-Isthme UMR5600)	<p style="text-align: center;">Modélisateurs biotopes</p> Céline SACCA, Cyrille CONORD, Aurélien FOURNIER, Pierre-Olivier MAZAGOL, Bernard ETLICHER <p style="text-align: center;">Fonctions</p> Calcul des variables environnementales, modélisations prédictives des biotopes, modélisation prédictive des niveaux d'humidité, participation à l'élaboration de méthodologies : échantillonnage biotopes, méthodologie de vérifications terrains Participation aux diverses réunions (Groupes de travail (GT1, GT2, GT3), Cotech, etc.)
IRSTEA	<p style="text-align: center;">Modélisateurs physionomies</p> Samuel ALLEAUME, Sandra LUQUE, Mohammad EL-HAJJ, Marc LANG <p style="text-align: center;">Fonctions</p> Mise au point de la méthodologie d'échantillonnage de la physionomie et de la nomenclature (télétection et cellules paysagères) Test de télétection des zones humides par radar Participation aux diverses réunions (Groupes de travail (GT1, GT2, GT3), Cotech, etc.) Interlocuteur principal pour la modélisation des physionomies

Contexte biogéographique

Parc naturel régional Scarpe-Escaut

Le secteur concerné par l'étude des végétations est situé sur l'ensemble du Parc naturel Scarpe-Escaut et quelques communes voisines non rattachées au Parc (Hélesmes, Petite-Forêt, Tilloy-lez-Marchiennes, Saint-Aybert). Le site d'étude est constitué d'un ensemble de 59 communes, toutes n'ayant pas fait l'objet de prospections spécifiques.

La superficie totale occupée par l'ensemble de ces 59 communes est d'environ 500 km² soit 50 000 hectares. La moitié de cette surface est occupée par des zones urbaines ou artificialisées qui n'ont pas été prospectées.

Parc naturel régional Avesnois

Le secteur concerné par l'étude des végétations est situé sur le territoire élargi du programme « agriculture et zones humides » du 11^e programme d'intervention de l'AEAP, qui correspond au territoire des « 2 Helves » et aux 13 communes du programme initial « agriculture et zones humides ». Le site d'étude est constitué d'un ensemble de 29 communes, toutes n'ayant pas fait l'objet de prospections spécifiques. La superficie totale est d'environ 370 km². Un périmètre élargi a été envisagé avec le laboratoire EVS-Isthme afin de tester la possibilité d'extrapolation du modèle sur un secteur voisin (73 communes au total, pour une superficie d'environ 722 km²).

Les caractéristiques environnementales utilisées afin de définir le catalogue des biotopes à l'échelle du Nord et du Pas-de-Calais (hors littoral) sont précisées dans le Tableau 2.

Tableau 2 : Caractéristiques environnementales communes aux territoires étudiés

Dépt	Paramètres	Commentaires
59	Climat	Bioclimat collinéen, océanique, pas de sécheresse estivale, ombroclimat humide. Certains secteurs du PNR Avesnois présentent un climat à tonalité submontagnarde que nous n'avons pas cherché à modéliser ici.
	Géologie	Dans la plaine de la Scarpe, la géologie déjà peu diversifiée est encore « tamponnée » par des couches d'alluvions et de limons de plateau qui homogénéisent une grande partie du territoire à l'exception de buttes sableuses et des marges sud plus calcaires ; Dans l'Avesnois, la géologie est beaucoup plus contrastée, avec de nombreuses assises paléozoïques alternant selon un réseau dense de failles, plus ou moins couvertes de limons de plateau et d'alluvions.
	Relief	Très plat globalement. Le PNRSE est une plaine alluviale avec très peu de relief rendant l'utilisation des paramètres topographiques pour la modélisation peu pertinente. Le PNR Avesnois est légèrement vallonné avec une altitude s'échelonnant de 140 m environ à 272 m dans un profil alternant plateaux, versants et vallées.
	Acidité	Présence de séries acidiphiles, acidiclives, neutroclives et neutrophiles.
	Gradient hydrique	Mésoxérophile à aquatique : Mésoxérophile (EII. 4) (non réalisé sur le POC) Mésophile (EII. 5) Mésohygrophile (EII. 6) Hygrophile (EII. 7), nappe circulante Hygrophile (EII. 7), nappe stagnante Hydrophile (EII. 8), nappe circulante Hydrophile (EII. 8), nappe stagnante Aquat. & amphib. (EII. 10-12), nappe circulante Aquat. & amphib. (EII. 10-12), nappe stagnante

➔ Les processus appliqués sur les territoires de la Plaine de la Scarpe et de l'Escaut et de l'Avesnois ainsi que sur le territoire du Cher et du Loiret par le CBNBP seront applicables sur une bonne partie des plaines françaises.

Contexte de politique publique relatif à la protection ou la connaissance des habitats naturels et semi-naturels (projets en cours, attentes locales)

L'Agence de l'eau Artois-Picardie a développé un programme de maintien de l'agriculture en zones humides dans le cadre duquel il est prévu de cartographier 75 % des prairies humides de chacun des territoires-cibles, selon une typologie distinguant trois niveaux d'humidité : mésophile, mésohygrophile, hygrophile.

Le programme CarHAB de cartographie de la végétation de la France, tel qu'il a été expérimenté dans la plaine de la Scarpe et de l'Escaut en 2018, se révèle utile pour prédéfinir les prairies humides potentielles.

La proposition du Conservatoire botanique national de Bailleul est d'évaluer la pertinence de la cartographie CarHAB réalisée sur le territoire de la plaine de la Scarpe et de l'Escaut pour prédire la localisation de prairies humides. Les prédictions issues de la modélisation ont été étayées par une campagne de terrain sur les parcelles des 17 exploitations faisant l'objet d'un suivi technico-économique dans le cadre du PMAZH, pour lesquelles le diagnostic d'humidité doit être le plus fiable possible.

D'autre part, la cartographie prédictive selon la méthode CarHAB a été étendue sur un autre territoire participant au PMAZH afin d'en évaluer la pertinence dans un autre contexte écosystémique et d'étendre l'emprise de la carte. Le site retenu pour cette extension est dans le territoire du PNR Avesnois.

- ➔ L'attente principale de la part de l'AEAP est donc une carte prédictive des prairies humides du PNRSE, réalisée à partir de la carte des biotopes basée sur un indice en particulier qui est l'humidité. Cette carte correspond à carte déclinée de la carte des biotopes.
- ➔ Accompagnement des Parcs sur les possibilités d'utilisation des outils CarHAB dans le cadre de leur mission et notamment dans la Scarpe-Escaut avec la candidature des vallées de la Scarpe et de l'Escaut au label RAMSAR que portera le PNRSE.

Méthodologie

Afin de réaliser une méthodologie la plus pertinente possible et d'obtenir les cartes biotopes et physionomies les plus proches de la réalité, de nombreux échanges téléphoniques et mails ont été réalisés avec EVS-Isthme (biotopes) et l'Irstea (physionomies).

De nombreuses réunions ont également été réalisées avec les partenaires du programme dans l'objectif de réaliser une méthodologie similaire pour chacun des POCs.

Le processus déployé sur le POC suit le schéma général ci-dessous avec des étapes communes aux autres POC :

- Qualification en biotope de données d'occurrence de flore et de végétation et production de polygones d'entraînement biotopes (CBNBL) ;
- Qualification en physionomie de données d'occurrence de végétation et production de polygones d'entraînement physionomies (CBNBL) ;
- Modélisation des biotopes : classification supervisée par apprentissage machine (EVS-Isthme) ;
- Modélisation des physionomies : classification supervisée par apprentissage machine avec chaîne Iota 2 (IRSTEA, CESBIO) ;
- Échanges entre les modélisateurs et le CBNBL pour les révisions des modèles (seuillage, changement de données, etc.) → Obtention de nouvelles versions de cartographies ;
- Contrôles des biotopes et physionomies : bureau et terrain (CBNBL) ;
- Croisement de la carte prédictive des physionomies et des biotopes (incorporation dans la segmentation IGN) par l'IGN = carte HABITAT SOCLE

Une spécificité du POC PNR Scarpe-Escaut et PNR Avesnois est de tester la modélisation des niveaux d'humidité. Ce test nécessite le partenariat d'EVS-Isthme qui modélise une première carte des niveaux d'humidité et une évaluation de la qualité de cette carte produite. L'évaluation est faite à partir des données en possession du CBNBL (base de données DIGITALE), de la connaissance du territoire et de vérifications de terrain. Dans un objectif d'amélioration, des seuillages du paramètre écologique « humidité du sol » sont réalisés afin de correspondre au plus proche de la réalité terrain.

Cette carte des niveaux d'humidité constitue un exemple parmi les cartes dérivées possibles de ce projet CarHAB.

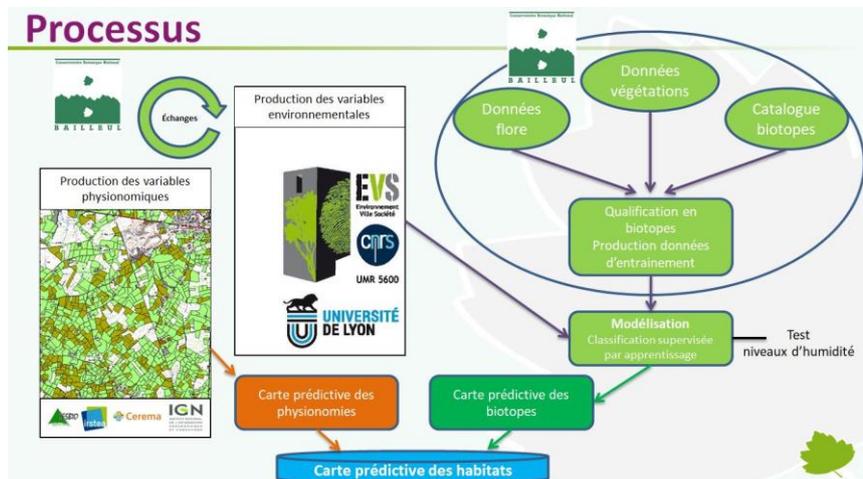


Figure 1 : Schéma simplifié du processus de production des modélisations cartographiques pour le POC Nord

Processus déployé pour la carte des physiologies

1. Définition des physiologies présentes sur le POC
2. Sélection par le CBNBL d'un maximum de polygones homogènes par physiologie en utilisant les images satellites (Sentinel2) et les orthophotos afin de constituer un jeu de données d'entraînement
3. Modélisation et croisements par les partenaires (Irstea, Cerema, Cesbio, IGN)

Tableau 3 : Nombre de données d'entraînement fournies pour la modélisation des physiologies

Identifiant national	Physiologie	PNRSE		PNRAv		Total
		Nombre polygones	Surface (ha)	Nombre polygones	Surface (ha)	
5210	Forestier pionnier à dominance de feuillus indéterminé	15	46,5	-	-	15
2120	Minéral non végétalisé indéterminé	9	7	-	-	9
2110	Minéral végétalisé indéterminé	22	21,4	9	5,9	31
3210	Herbacé haut indéterminé	18	13,3	24	16	42
3221	Prairie fauchée	32	44,8	42	61,5	74
3222	Prairie pâturée	23	41,4	35	78	58
4210	Arbustif / Fourré / Fruticée mixte indéterminé	2	4,6	8	36,7	10
3111	Pelouse permanente vivace	-	-	4	1,5	4
	Somme	121	179	122	199,6	243

Processus déployé pour la carte des biotopes

Concernant les modélisations, plusieurs versions ont été réalisées et sont indiquées dans le Tableau 4. La modélisation des zones potentiellement humides a principalement nécessité des calibrages de seuils des couches fournies par EVS-Isthme sur la base de remarques orales selon les connaissances du territoire.

Tableau 4 : Informations sur les différentes versions des modélisations biotopes et zones potentiellement humides

Productions	Date	Territoire concerné		Couches de données d'apprentissage en amont	Format reçu
V0 Carte des biotopes	30 avril 2019	PNRSE PNRAV	Une seule couche SIG Résolution 25 m	Données d'entraînement inférieures à 1ha fournies le 29 mars 2019 (« ponctuelles »)	web-sig/flux WMS
V1 Carte des biotopes	13 mai 2019	PNRSE PNRAV	Une seule couche SIG Résolution 10 m	Données d'entraînement inférieures à 1ha fournies le 29 mars 2019 (« ponctuelles »)	web-sig/flux WMS + Raster
V2 Carte des biotopes	04 juin 2019	PNRSE PNRAV	Deux couches (les PNR sont séparés) Résolution 10 m	Données d'entraînement inférieures à 1ha fournies le 29 mars 2019 (« ponctuelles »)	web-sig/flux WMS + Raster
V3 Carte des biotopes	15 juillet 2019	PNRSE PNRAV	Deux couches (les PNR sont séparés) Résolution 10 m	Données d'entraînement supérieures à 1ha fournies le 05 juillet 2019 (« surfaiques ») Les données ponctuelles sont abandonnées car induisent trop de biais	Raster et vecteur

V0 - Zones potentiellement humides (ZPH)	30 avril 2019	PNRSE	9 fichiers : .Zone Hydrophile/hygrophile isolée .Zone mésohygrophile isolée .Densité de canaux à 500 m sur PNRSE et reclassée .Densité de canaux et de mares à 250 m + 500 m et reclassées (pour la zone hydrophile/hygrophile isolée) .Densité de canaux à 250 m + 500 m et reclassée (pour la zone mésohygrophile isolée)		Raster
V1 - ZPH	20 mai 2019	PNRSE	.zhp_v1 : version "classique" .zhp_v2 : "classique" + densité de canaux et de mares .zhp_v3 : "classique" + géologie reclassée selon l'humidité .zhp_v4 : "classique" + géologie reclassée selon l'humidité + densité de canaux et de mares		Raster
V2 - ZPH	28 mai 2019	PNRSE	.nv_humidite_ok : couche réalisée à partir des données 2018 .ZPH_v4_rec : reclassement de ZPH_v4 de la V1 - ZPH .hydrophil_dc et hydrophil_dc : couches isolées à partir du reclassement ZPH_v4 de la V1 - ZPH et retravaillée avec densité de canaux.		Raster

V3 - ZPH	04 juin 2019	PNRSE	Une couche SIG suite aux diverses calibrations		Raster
V4 - ZPH	18 juillet 2019	PNRAv	Une couche SIG du reclassement de la zhp_v1 en fonction du deuxième vote pour la transition hydrophile → hygrophile		Raster des pixels concernés par le reclassement Vecteur de la couche SIG

✓ **Étape 1 = Définition des paramètres environnementaux dans la région**

Sélection des variables suivantes :

Tableau 5 : Variables environnementales retenues pour le POC Nord

Variable générale	Poste retenu	Code
Caractère littoral	Position intérieure	1.2
Étage thermique	Collinéen	2.A2
Ombroclimat	Humide	3.3
Continentalité	Océanique	4.2
Variante bioclimatique	Aucune	5.1
Humidité du sol	Mésoxérophile (Ell. 4)	6.2
	Mésophile (Ell. 5)	6.3
	Mésohygrophile (Ell. 6)	6.4
	Hygrophile (Ell. 7), nappe circulante	6.5
	Hygrophile (Ell. 7), nappe stagnante	6.6
	Hydrophile (Ell. 8), nappe circulante	6.7
	Hydrophile (Ell. 8), nappe stagnante	6.8
	Aquatiq. & amphib. (Ell. 10-12), nappe circulante	6.9 6.10
	Aquatiq. & amphib. (Ell. 10-12), nappe stagnante	
Acidité du sol	Acidiphile	7.1
	Acidicline	7.2
	Neutrocline	7.3
	Neutrophile	7.4

Pour les variables « Ombroclimat », « Étage thermique » et « Continentalité », les calculs ont été effectués selon les formules de Rivas-Martinez (Rivas-Martinez *et al.*, 2011)

Continentalité (Rivas-Martinez *et al.*, 2011 : 8)

Océanique (= Euocéanique de Rivas-Martinez) $14 < Ic < 17$

Ic (indice de continentalité) = T° moy du mois le plus chaud - T° moy du mois le plus froid

Étage thermique (Rivas-Martinez et al., 2011 : 17)

- It (thermicity index) = (Tmin du mois le plus froid + T°max du mois le plus froid + T° moy. annuelle)*10
- Tp (somme annuelle de T°positives) = [somme des T° mensuelles des n mois à moyenne >0°C]*10
 - ➔ Collinéen (supratempéré de Rivas-Martinez) : 120 < It < 190 ; 1100 < Tp < 1400
 - ➔ Planitiaire (mésotempéré de Rivas-Martinez) : 190 < It < 240 ; 1400 < Tp < 1700

Les étages thermiques peuvent être redécoupés par une variante bioclimatique

- **steppique** : pas dans l'humide ni dans l'océanique : plutôt à l'est de la France ;
- **subméditerranéenne** : si Io mensuel (P/T) au moins un mois en été < 2,8 ;
- **aucune** : sans variante.

Ombroclimat (Rivas-Martinez et al., 2011 : 18)

Humide = 9 > io annuel > 6

Pour récupérer l'information des stations météo, le site utilisé est <https://www.infoclimat.fr/climatologie/>.

Ce site compile les stations de particuliers et les données de météo France. Possibilité de choisir entre les normales 1970-2000 et 1980-2010 et de faire des copier-coller à partir des tableaux de chaque station. Nous avons utilisé les données 1980-2010.

Exemple avec la commune de Maubeuge dans l'est du département du Nord (le plus proche de l'avesnois)

- Étage thermique : It de 170 (et un Tp de 1270 (=10,6*12*10) ➔ étage Collinéen
- Variante bioclimatique ➔ aucune (pas de sécheresse estivale)

Tableau 6 : Valeurs de températures et de précipitations utilisées dans le calcul des paramètres environnementaux, l'exemple de Maubeuge

		Janv	Fév	Mar	Avr	Mai	Jui	Juil	Aou	Sept	Oct	Noc	Déc	Moy ann.
MAUBEUGE (59)	Tempé. maxi moy	5,8	6,9	10,6	14,2	18,7	21,3	23,6	23,7	19,8	15,5	9,5	6,6	14,7
Alt. 140 m	Tempé. moy	3,2	3,7	6,7	9,4	13,6	16,2	18,5	18,4	15,1	11,5	6,5	4,2	10,6
Topo : Vallée de la Sambre	Tempé. mini	0,6	0,6	2,8	4,6	8,6	11,1	13,4	13,1	10,3	7,6	3,6	1,7	6,5
Région naturelle :	Préci Cumul pluies moyen (mm)	81,1	64,7	80,8	54,7	70,9	79,1	71,2	73,1	61,8	79,5	78,8	85,1	880,8
Indice ombrothermique mensuel (P/T)	Iom	25,3	17,5	12,1	5,8	5,2	4,9	3,8	4,0	4,1	6,9	12,1	20,3	6,92

Indice de continentalité (T° moy mois le plus chaud - T° moy mois le plus froid) = 15,3 ➔ climat océanique

Variante Bioclim Rivas (mois Iom < 2,8) = 0

Balanced oceanic = aucune

Positive annual T°(Tp) = 1270

Thermicity Index (It) : 170

➔ Étage thermique : Collinéen (supratempéré de Rivas-Martinez)

Il est important de noter qu'il existe des zones « hybrides » dont l'un des deux indices appartient à la catégorie inférieure ou supérieure (planitiaire ou montagnard).

D'après la connaissance du territoire, ceci indique que toute l'amplitude du collinéen est couverte (de sa valeur la plus basse à sa valeur la plus haute) et que certaines zones sont à considérer comme ayant une tendance montagnarde par exemple.

✓ **Étape 2 = Filtre des paramètres environnement afin de déterminer la liste des biotopes possibles**

L'application de filtres sur les paramètres a permis de retenir 36 biotopes à l'échelle du Nord et du Pas-de-Calais (hors littoral).

Sur les **36 biotopes retenus**, 14 ne sont pas présents sur les POCs.

Au total, 22 biotopes modélisables sont définis.

En plus de ces 22 biotopes, il existe des séries que l'on appelle « séries dérivées » et qui se réaliseront dans plusieurs biotopes, sous l'effet de l'eutrophisation excessive des milieux.

D'autre part, certaines séries sont ponctuelles. Ces séries sont considérées comme ponctuelles car elles ne s'expriment que sur de petites surfaces (inférieures au seuil 2 500 m²). Elles ne sont donc pas représentables sur la carte des biotopes. Elles ne sont pas retenues dans le cadre du programme en raison du biais qu'elles provoquent dans la modélisation de ces POC. Le modèle surestime la présence de ces séries et indique les biotopes associés sur de trop grandes surfaces par rapport à la réalité.

L'objectif est donc bien de modéliser les biotopes les plus représentatifs du territoire. Pourtant, même pour ces biotopes, les données de biotopes de petites surfaces ne retranscrivent pas toujours suffisamment la réalité de terrain pour que le modèle soit juste en sortie (Figure 2), de même que les polygones de type « ligne », peu pertinents dans ce programme en raison d'une difficulté à les rattacher à des pixels homogènes. Ces deux types de données n'ont pas été pris en compte pour la modélisation finale.

Étape 3 = Extraction de données

À partir de la base de données du CBNBL (DIGITALE), une extraction des données de syntaxons et cellules présentes dans les deux PNR a été effectuée.

Extraction CBNBL (DIGITALE) concernant le site PNR Scarpe-Escaut et PNR Avesnois.

Observations de syntaxons présents dans le site
Polygones précis (IN, OC, Zo dans la terminologie du CBNBL), à l'exclusion des données rapportées à une commune, un lieu-dit, une maille...
Documents diffusables
Fiabilité : sans les données indiquées en « douteux », « erreurs » et « non traités »
Nature observation : Observée
Référentiel utilisé = 'BSS' (banque synsystématique)
Uniquement les phytocénoses ou microcénoses
Précision syntaxinomique : Seules les données de rangs synsystématiques inférieurs ou égaux à l'association végétale ont été retenues : association, groupement, sous-association, sociation, race, variante, sous-variante, forme, faciès

Transmissions de données par le PNRAv concernant le site PNR Avesnois.

Couche shape avec Excel associé contenant des noms de syntaxons tous rangs confondus
Seuls les rangs suivants ont été retenus : association, groupement, communauté basale, sous-association, sociation, race, variante, sous-variante, forme, faciès

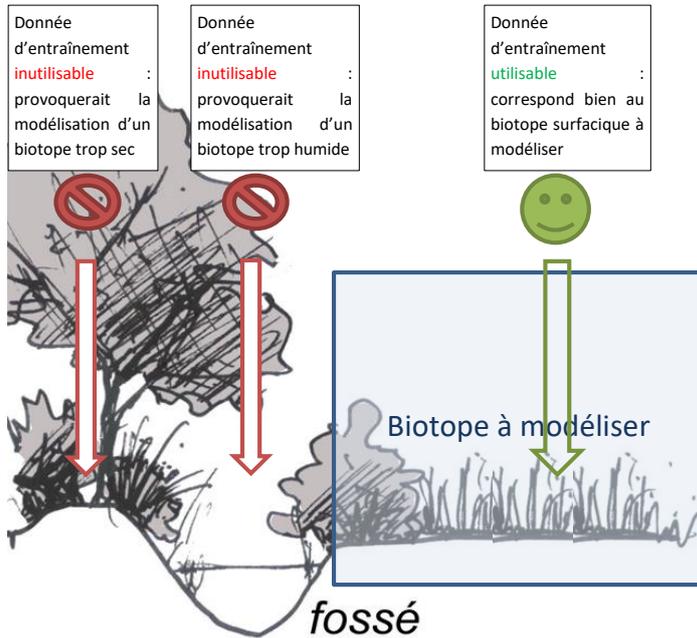


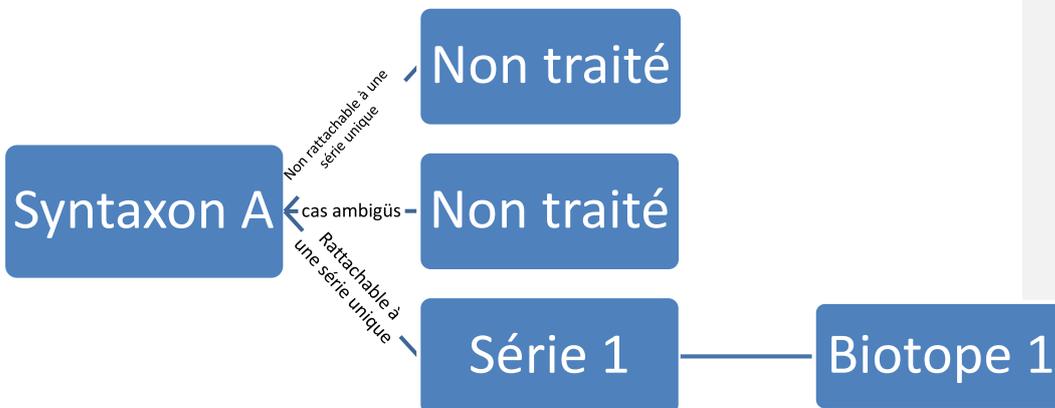
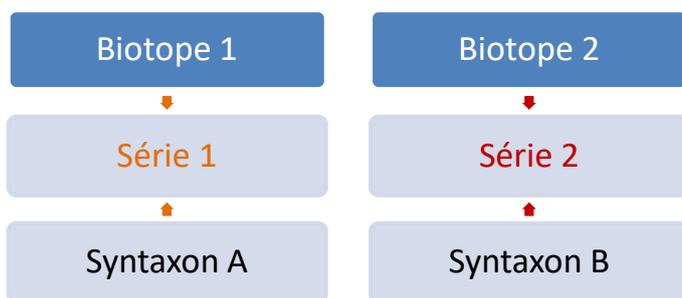
Figure 2 : Schématisation des données d'entraînement utilisables pour la modélisation des biotopes

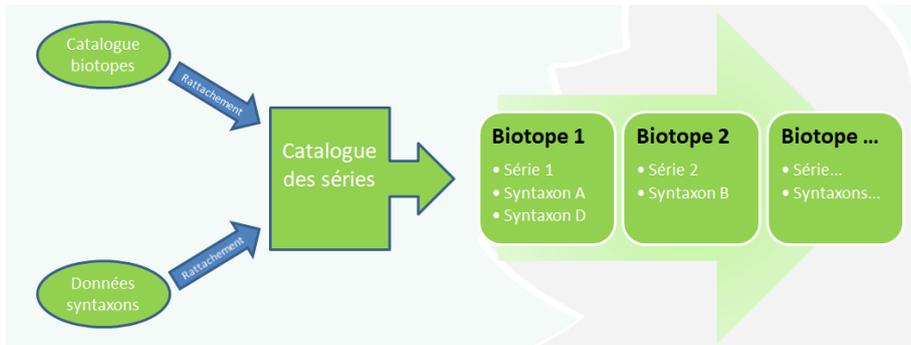
✓ **Étape 4 = Attribution séries / biotopes**

Pour chaque biotope défini en étape 2, une série a été associée.

En parallèle, les extractions ont permis d'obtenir la liste des syntaxons présents dans les deux PNR et pour chacun d'y associer une série quand cela était possible. L'objectif était d'avoir un syntaxon = une série. Les cas ambigus n'ont pas été traités pour la création de données d'entraînement, les cas avec plusieurs séries rattachées ont parfois fait l'objet d'une expertise quand le nombre de données d'entraînement pour un biotope donné était insuffisant.

Le lien entre les deux fichiers s'est donc fait via le champ « série »





✓ Étape 5 = Passage sous Qgis

Sous Qgis, les jointures ont été faites afin de remplir la table attributaire et les champs ont été complétés de la même façon que le protocole CBNBP.

Seuls les polygones suffisamment précis (max 3ha) ont été retenus, afin d'éviter que le polygone concerné ne couvrent plusieurs biotopes à modéliser.

Parfois, seulement 7 ou 8 données ont été fournies. Il a été estimé qu'il était plus judicieux de fournir moins de données mais de qualité pour certains biotopes (plus particulièrement dans le PNRAV) plutôt que de rajouter 3 mauvaises données afin d'arriver au nombre fixé de 10 polygones par biotope.

Un premier résultat de ce nombre de données faible est que le modèle n'arrive pas à apprendre suffisamment tout seul et qu'il est impératif d'augmenter la donnée dite « d'expert », donc créée par l'utilisateur en fonction des diverses informations en sa possession. Ces données sont plus informatives pour le modèle. Il est nécessaire d'ajouter des données afin de le calibrer au mieux ou bien de lui indiquer des poids d'importance différents selon la donnée. Cependant, nous avons fait le choix, dans un premier temps, de tester une production sans ces données expertes, afin d'évaluer une méthode applicable à moindres frais à l'échelle nationale.

Les accents ont été évités dans le jeu de données ainsi que pour les noms de fichiers afin de limiter les erreurs.

✓ Étape 6 = Constatations et ajustements

Les données trop précises (de petites surfaces et « non expertes ») ne retranscrivent pas suffisamment la réalité terrain pour que le modèle soit juste en sortie.

- La première version des données d'entraînement fournie sous forme « ponctuelle » ne reflétait pas le biotope général mais parfois un accident de terrain comme un fossé par exemple (Figure 3 : Illustration de données "ponctuelles" fournies biaisant la modélisation).

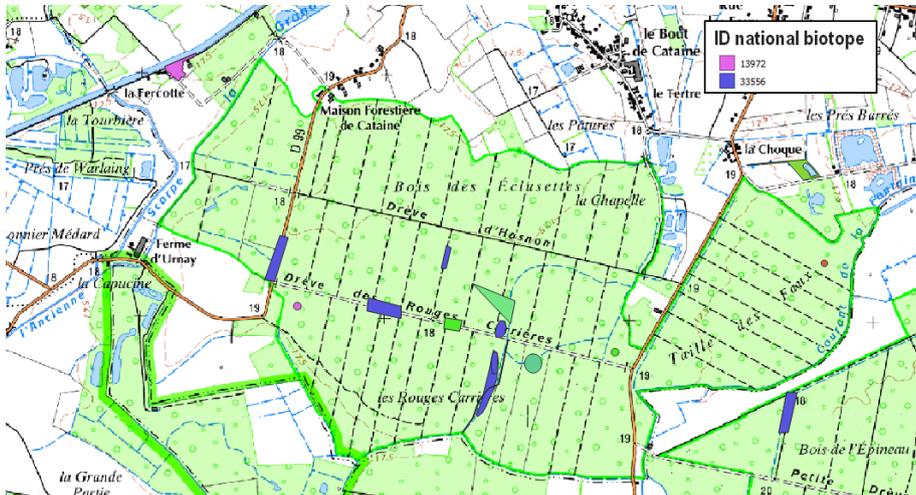


Figure 3 : Illustration de données "ponctuelles" fournies biaisant la modélisation (beaucoup de ces données correspondent à des végétations de chemin, de bord de route, de cours d'eau...)

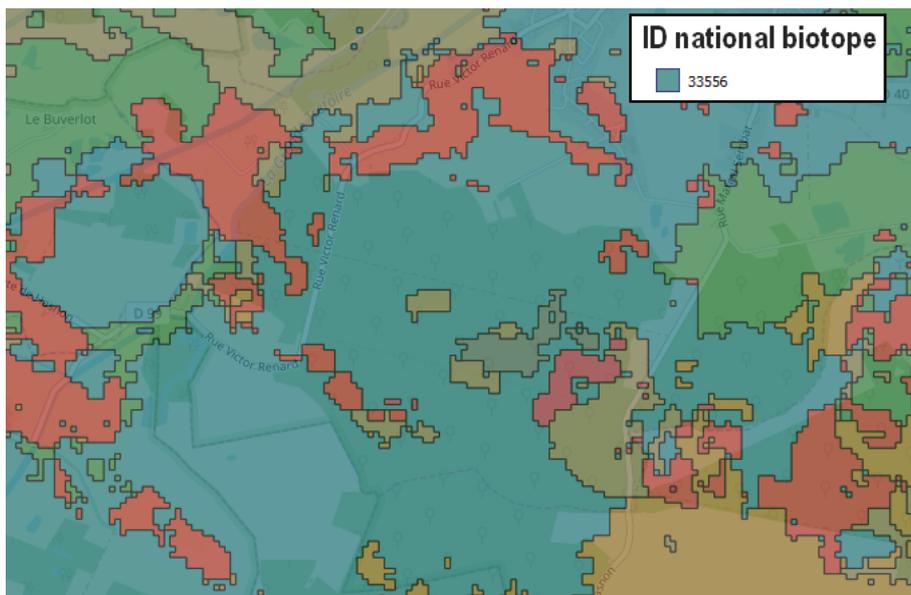


Figure 4 : Résultat de la modélisation carte des biotopes V0 issue des données "ponctuelles"

Le biotope 33556 est en réalité un biotope aquatique, sa modélisation ne devrait pas couvrir une surface aussi importante que celle apparaissant dans la modélisation des biotopes V0 (Figure 4). Le modèle a donc surestimé la présence de ce biotope localisé dans les fossés. Un autre biotope plus surfacique était attendu.

Le choix a été de ne pas retenir ces données « ponctuelles » et linéaires et de fournir des données surfaciques (> 1 ha). Afin de compléter le nombre de données fournies, quelques polygones inférieurs à 1 ha ont été fournis, cependant ces polygones ont tous été créés de manière « experte », assurant ainsi une donnée de qualité.

- Les polygones de type « linéaire » ont également été retirés des jeux de données d'entraînement car ils ont été jugés peu pertinents dans le cadre de la modélisation. Ces derniers présentaient une difficulté dans leur rattachement à des pixels homogènes, ce qui induisait donc des biais dans la modélisation. En effet, de par leur longueur, ces polygones impactent un grand nombre de pixels et induisent un poids important dans la modélisation. Les biotopes ponctuels et linéaires de plans d'eau et de cours d'eau pourront être extraits de couches spécifiques existantes par ailleurs (couches « mares », « réseau hydrographique »...).
- La première modélisation a été réalisée en une seule fois à l'échelle des deux Parcs et a entraîné des biais. Il a été décidé que l'utilisation des jeux de données d'entraînement serait séparée pour effectuer la modélisation sur chacun des parcs. Cette méthode est plus pertinente en raison de contextes écologiques différents entre les deux parcs et notamment en raison de la présence/absence de biotopes dans un PNR par rapport à l'autre.
- Les données « expertes » sont plus pertinentes pour l'entraînement du modèle. Il est nécessaire de favoriser ce type de données d'entraînement.

✓ **Étape 7 = Finalisation et transmission des données**

Les données suivantes ont été envoyées à EVS-Isthme

- « EVS biotopes POCNord.rar » avec les deux jeux de données QGIS (deux fichiers) ;
- Liste biotopes avec les biotopes utilisés contenant Numéro et nom (retrouvés dans la table QGIS) ;
- Nombre données Entraînement avec le nombre de polygones pour chaque biotope pour chaque POC (deux onglets).

Les couches fournies avaient pour table attributaire (convergence avec le protocole CBNBP) :

- Cooutm = numéro interne CBNBL
- ID_NAT = identifiant national du biotope (un biotope = un identifiant)
- L_BIOTOPE = nom du biotope (concaténation des 7 paramètres avec un underscore)
- ID_SAISIE = numéro unique d'identifiant du polygone
→ SE_xxx pour Scarpe escaut, Av_xxx pour Avesnois
- DATE = date de création du polygone
- CD_OPERATE = initiales de l'opérateur de création de la donnée (GV = Geoffroy Villejoubert)
- CBN_SYNTAX = la donnée d'entraînement s'appuie-t-elle sur l'information d'au moins une donnée syntaxon ? (OUI ou NON)
- CBN_CARTO = Le polygone créé existe-t-il dans la base CBNBL (=OUI) ou bien a-t-il été créé de toute pièce (=NON)
- CBN_REL_PH = le polygone créé utilise-t-il une information relevé phyto. N'ayant pas l'information pour les polygones, j'ai mis NON par défaut. Cette information n'est pas à prendre en tant que tel car il est certain que bon nombre de données syntaxons ont un relevé phytosociologique attaché. Cependant, cette information demandait un temps trop important et n'a donc pas été fournie afin de ne pas augmenter la charge de travail déjà conséquente.

- CBN_FLORE = le polygone créé utilise-t-il de la donnée flore (OUI ou NON)
- COMMENTAIR = indication sur la donnée de type « donnée expertisée » (créée de toute pièce). Si le champ est vide, c'est que le polygone existe dans la base de données DIGITALE.
- QUALITE = évaluation de qualité de la donnée fournie

Plusieurs niveaux :

Excellente (donnée créée et expertisée)

Très bonne (donnée CBN récente >2010)

Bonne (donnée CBN entre 2000 et 2010)

Moyenne (Donnée CBN <2000)

Et quelques ajustements pour certaines données sur la qualité (les données reçues par le PNRAV sont qualifiées en « Moyenne », les données qui avaient un niveau de confiance dans notre base ont été descendues d'un cran par rapport à leur position initiale).

Volume de données

→ Première version des jeux de données d'entraînement (données « ponctuelles » et linéaires)

→ Version non retenue

Le PNR Avesnois ayant réalisé des inventaires sur son territoire, les données ont été transmises par ce dernier avant l'intégration dans la base de données DIGITALE afin de pouvoir les traiter rapidement.

Au total ce sont 5938 données qui ont été analysées (extraction DIGITALE). Ces données sont triées afin de ne sélectionner que les plus pertinentes codant pour 18 biotopes dans le PNRSE et pour 13 biotopes dans le PNRAV. Le nombre de données analysées et retenues pour chaque PNR est détaillé dans le Tableau 7.

Tableau 7 : Volume de données considérées pour les données entraînement biotopes première version (< 1ha) et pour la seconde version (> 1ha)

	PNRSE données du CBN	PNRAV - données du CBN	PNRAV – données du PNR	Total données PNRAV	Total
Nombre total de données					
Nb données	4238	1278	422	1697	5938
Nb données < 1ha	2051	698	175	873	2924
Nb données > 1ha	2187	580	247	827	3014
Nombre total de données modélisables					
Nb données < 1ha modélisables	1068	386	12	398	1466
Nb données > 1ha modélisables	936	176	3	179	1115
Nombre total de données non modélisables					
Nb données < 1ha non modélisables	839	195	161	356	1195
Nb données > 1ha non modélisables	1067	352	244	596	1663

Nombre de données nécessitant une expertise

Nb données < 1ha dont plusieurs biotopes sont possibles	144	117	2	119	263
Nb données > 1ha dont plusieurs biotopes sont possibles	184	52	0	52	236

Nombre de données exploitées pour la première version du jeu de données d'entraînement

Nombre de données fournies dans le jeu de données d'entraînement	1096*			436	1532
Nombre de biotopes à modéliser	18			13	

*Les 1096 données fournies du jeu de données concernant le PNRSE ont été sélectionnées parmi les 1068 données modélisables et inférieures à 1ha ainsi que les 144 données inférieures à 1ha et nécessitant une expertise pour l'attribution d'un biotope.

Tableau 8 : Volume de données retenues pour le premier jeu de données d'entraînement

PNR	Nombre de biotopes	Nombre total de biotopes	Nombre de données	Surface totale d'apprentissage (ha)	Surface d'apprentissage (ha)
Scarpe-Escout	18	20	1532	272	160
Avesnois	13				112

→ Seconde version des jeux de données d'entraînement (données surfaciques) → Version retenue

Suite aux constatations de modélisation erronée, une seconde version des jeux de données d'entraînement a été fournie. Cette seconde version n'utilisait plus les données dites « ponctuelles » et linéaires mais uniquement les données surfaciques (supérieures à 1 ha). Les données inférieures à 1 ha retenues ne sont que des données « expertes » et donc de qualité excellente. Les données utilisées sont celles de la première extraction DIGITALE que l'on retrouve dans le Tableau 7. Les données ont été triées afin de ne garder que celles jugées comme pertinentes pour la modélisation codant pour 14 biotopes dans le PNRSE et pour 10 biotopes dans le PNRAv. Le nombre de biotopes, le nombre de données retenues et la surface d'apprentissage pour chaque PNR sont détaillés dans le Tableau 7.

Tableau 9 : Volume de données exploitées pour le second jeu de données d'entraînement

PNR	Nombre de biotopes	Nombre total de biotopes	Nombre de données		Surface d'apprentissage (ha)	
Scarpe-Escout	14	15	455	583	4269	5120
Avesnois	10		128		851	

→ Récapitulatif

La version retenue pour le jeu de données d'entraînement (seconde version) est celle utilisant les données surfaciques. Ces données ont été retenues et codent pour 15 biotopes. La seconde version ne prend pas en compte les données d'extraction « ponctuelles » et linéaires qui sont principalement des données liées à des habitats aquatiques, ainsi cinq biotopes sont abandonnés dans la seconde version (Figure 5).



Figure 5 : Nombre de biotopes modélisables pour la première version du jeu de données d'entraînement (ponctuelles et linéaires, à gauche) et la seconde version (surfaciques, à droite)

Tests méthodologiques mis en œuvre

Un premier objectif était de tester la pertinence de la modélisation des biotopes sur des territoires différents du point de vue écosystémique (PNR Scarpe-Escout et PNR Avesnois).

Un objectif décliné à partir du premier s'est dessiné : est-il possible de modéliser les biotopes en utilisant un jeu de données d'entraînement global ou est-il nécessaire de bien séparer les deux jeux de données en raison du contexte écosystémique différent ?

Dans le PNR Avesnois, un objectif secondaire a été testé et se rapprochant du test d'extrapolation du POC Loiret de la modélisation d'un territoire à partir de son voisin. Dans ce PNR, une zone d'étude principale était définie avec présence de données anciennes et une zone élargie où très peu de données étaient disponibles.

Un autre objectif était de tester la pertinence de la modélisation quant à l'identification des niveaux d'humidité. Celui-ci a principalement été testé et vérifié sur le PNRSE.

Afin de répondre à ces divers objectifs, plusieurs tests ont été mis en œuvre sur le POC Nord :

- Test de production de données d'entraînement biotopes
 - o Tester la création de données d'apprentissage « biotopes » en utilisant les différentes données disponibles dans la base de données DIGITALE du CBNBL (Flore, végétations, etc.).
 - Concluant, mais moins précis qu'une donnée experte

- Tester la création de données d'apprentissage « biotopes » en utilisant les anciennes données CarHAB1. Utilisation des données ponctuelles (< 1ha) non retenue en raison du biais d'extrapolation du modèle et utilisation de données surfaciques (> 1ha) retenue.
 - Concluant, mais moins précis qu'une donnée experte.
- Test de production de données d'entraînement physionomies
 - Tester la création de données d'apprentissage « physionomies » en utilisant les différentes données disponibles dans la base de données DIGITALE du CBNBL (végétations, données CarHAB1.).
 - Concluant pour le niveau 4 dans la mesure où l'information est précise (exemple de la distinction prairie fauchée/prairie pâturée).
- Tests de modélisation en utilisant un seul jeu de données unique.
 - Tester la modélisation d'une carte prédictive des biotopes à partir de données d'apprentissage définies sur deux territoires non contigus. Ce test a été réalisé en complément afin de voir si au niveau national il est possible d'utiliser des jeux de données issus de territoires aux contextes écosystémiques différents.

Le résultat indique que la modélisation de deux territoires aux caractéristiques environnementales différentes n'est pas pertinente. Les différentes versions de modélisation réalisées par la suite ont donc été réalisées de manière indépendante sur chacun des PNR.

 - Non concluant (différences écologiques trop importantes entre les deux PNR).
- Test d'extrapolation du modèle sur le PNR Avesnois (une zone avec données d'entraînement, une zone avec peu de données). Ceci se rapproche d'un des objectifs du POC Loiret
 - Tester la pertinence de l'extrapolation du modèle d'une zone avec peu de données d'entraînement à partir d'une zone voisine avec présence de données. Les deux zones sont contigües.
 - Non concluant, le jeu de données fourni en entrée n'était pas suffisamment conséquent pour fournir une modélisation juste.
- Test de vérification de la carte des biotopes et des ZPH
 - Contrôle qualité de la première version brute de la carte produite (V1) par SIG à dire d'expert et selon la base de données du CBNBL.
 - Non concluant, le jeu de données fourni était composé de données ponctuelles et linéaire induisant un biais dans la modélisation.
 - Contrôle qualité de la carte prédictive des niveaux d'humidité (zhp_v4 V2 – ZPH et ses versions améliorées) en collectant de la donnée *in-situ* notamment dans le cadre d'une mission réalisée sur le PNRSE (identification des niveaux d'humidité des prairies dans le cadre du Programme de maintien de l'agriculture en zones humides).
 - Concluant mais nécessitant des ajustements afin de se rapprocher au mieux de la réalité.

- Contrôle qualité de la dernière version de la carte des biotopes produite (V3) *in-situ* dans le PNR Avesnois (test d'une méthode terrain similaire à celle appliquée par le CBNBP, d'une durée de 6 jours réalisée parallèlement au contrôle des physionomies fin juillet).
 - Très concluant, cette dernière version ne nécessite que de légers ajustements.
- Test de vérification de la carte des physionomies
 - Contrôle qualité de la carte prédictive des physionomies : version de niveau 4 produite par l'Irstea. Contrôle de la carte prédictive des physionomies sur le PNRSE *in-situ* et SIG, contrôle sur le PNRAv *in-situ* (test d'une méthode terrain similaire à celle appliquée par le CBNBP, d'une durée de 6 jours réalisée parallèlement au contrôle des biotopes, fin juillet).
 - Moyennement concluant au niveau 4, plus performant dans l'Avesnois et peu satisfaisant dans la Plaine de l'Escaut. La modélisation niveau 4 est donc moyennement satisfaisante (distinction entre les prairies pas toujours évidente). Certaines données fournies sont à réviser afin de fournir un jeu de données plus proche de la réalité (problèmes concernant le minéral peu végétalisé notamment). Au niveau de restitution, niveau 2, la modélisation est satisfaisante.
- Test radar – télédétection des zones humides (Irstea)
 - Tester la sensibilité des images radar à l'hydromorphie des zones humides sur le PNRSE. Ce test a été réalisé par l'Irstea
 - Les résultats montrent une sensibilité et une séparabilité entre classes hydromorphiques des zones herbacées. Cependant, un effort plus important d'échantillonnage est nécessaire pour calibrer un modèle prédictif.
- Tester l'utilisation des séries dérivés afin d'augmenter les données d'entraînement
 - Test d'attribution de biotopes avec plusieurs niveaux de probabilités de réalisation (fort, moyen, faible, impossible) à partir de données de végétations codant pour plusieurs biotopes. Ce test a nécessité d'utiliser un ensemble de paramètres (données environnantes, carte des biotopes, carte des niveaux d'humidité) afin d'en déduire une potentialité du biotope à l'endroit concerné.
 - Non testé pour la modélisation, ceci n'a pas été nécessaire en raison d'un nombre suffisant de données d'entraînement.

Nomenclatures cibles des cartographies

Typologie des biotopes (Tableau 10)

La méthodologie générale a été réalisée afin de réaliser la liste des biotopes présents sur le territoire. 36 biotopes sont retenus pour le Nord et le Pas-de-Calais, hors littoral. Sur le POC Nord, 22 biotopes ont été jugés modélisables dans un premier temps. Cependant, en raison d'une petite couverture surfacique de deux biotopes, ceux-ci n'ont pas été considérés comme modélisables et seuls 20 ont été retenus (V1). Dans un second temps, les biotopes de type ponctuel, principalement aquatiques, n'ont pas été inclus dans le jeu de données d'entraînement pour la modélisation des biotopes. Ainsi, ce sont 15 biotopes qui ont été retenus pour la version finale (V2).

Tableau 10 : Biotopes modélisés pour la V1 (« ponctuelles » et linéaires) et la V2 (surfaciqes) parmi les 36 biotopes retenus sur le territoire.

ID NATIONAL	LIBELLE BIOTOPE SIMPLIFIE	
1300	Intérieur Collinéen Humide océanique Mésoxérophile Acidiphile	-
2452	Intérieur Collinéen Humide océanique Mésophile Acidiphile	V1 & V2
3604	Intérieur Collinéen Humide océanique Mésohygrophile Acidiphile	V1 & V2
4756	Intérieur Collinéen Humide océanique Hygrophile_circulant Acidiphile	-
5908	Intérieur Collinéen Humide océanique Hygrophile_stagnant Acidiphile	-
7060	Intérieur Collinéen Humide océanique Hydrophile_circulant Acidiphile	-
8212	Intérieur Collinéen Humide océanique Hydrophile_stagnant Acidiphile	V1 & V2
9364	Intérieur Collinéen Humide océanique Aquatique&Amphibie_circulant Acidiphile	-
10516	Intérieur Collinéen Humide océanique Aquatique&Amphibie_stagnant Acidiphile	V1
12820	Intérieur Collinéen Humide océanique Mésoxérophile Acidicline	-
13972	Intérieur Collinéen Humide océanique Mésophile Acidicline	V1 & V2
15124	Intérieur Collinéen Humide océanique Mésohygrophile Acidicline	V1 & V2
16276	Intérieur Collinéen Humide océanique Hygrophile_circulant Acidicline	V1 & V2
17428	Intérieur Collinéen Humide océanique Hygrophile_stagnant Acidicline	-
18580	Intérieur Collinéen Humide océanique Hydrophile_circulant Acidicline	V1 & V2
19732	Intérieur Collinéen Humide océanique Hydrophile_stagnant Acidicline	-
20884	Intérieur Collinéen Humide océanique Aquatique&Amphibie_circulant Acidicline	-
22036	Intérieur Collinéen Humide océanique Aquatique&Amphibie_stagnante Acidicline	V1
24340	Intérieur Collinéen Humide océanique Mésoxérophile Neutrocline	-
25492	Intérieur Collinéen Humide océanique Mésophile Neutrocline	V1 & V2
26644	Intérieur Collinéen Humide océanique Mésohygrophile Neutrocline	V1 & V2
27796	Intérieur Collinéen Humide océanique Hygrophile_circulant Neutrocline	V1 & V2
28948	Intérieur Collinéen Humide océanique Hygrophile_stagnant Neutrocline	V1 & V2
30100	Intérieur Collinéen Humide océanique Hydrophile_circulant Neutrocline	-
31252	Intérieur Collinéen Humide océanique Hydrophile_stagnant Neutrocline	V1 & V2
32404	Intérieur Collinéen Humide océanique Aquatique&Amphibie_circulant Neutrocline	V1
33556	Intérieur Collinéen Humide océanique Aquatique&Amphibie_stagnant Neutrocline	V1
35860	Intérieur Collinéen Humide océanique Mésoxérophile Neutrophile	-
37012	Intérieur Collinéen Humide océanique Mésophile Neutrophile	V1 & V2
38164	Intérieur Collinéen Humide océanique Mésohygrophile Neutrophile	V1 & V2
39316	Intérieur Collinéen Humide océanique Hygrophile_circulant Neutrophile	-
40468	Intérieur Collinéen Humide océanique Hygrophile_stagnant Neutrophile	-

41620	Intérieur Collinéen Humide océanique Hydrophile_circulant Neutrophile	-
42772	Intérieur Collinéen Humide océanique Hydrophile_stagnant Neutrophile	V1 & V2
43924	Intérieur Collinéen Humide océanique Aquatique&Amphibie_circulant Neutrophile	-
45076	Intérieur Collinéen Humide océanique Aquatique&Amphibie_stagnant Neutrophile	V1

Typologie des physionomies (

Tableau 11)

En raison d'un manque de données ou d'une expression peu marquée, il n'a pas été possible de générer systématiquement les données d'entraînement physionomies au niveau 4 :

- Difficulté de trouver certains polygones (PNRAv – « Forestier pionner (5210) », « Minéral non végétalisé (2120) ») ;
- Difficulté d'avoir un grand nombre de polygones pour des physionomies peu communes ;
- L'intitulé « pelouse bas marais/tourbière » n'est pas retenu. La distinction pelouse permanente annuelle et pelouse permanente vivace est retenue, bien que non ou peu présente sur le POC et difficilement distinguable de « Minéral non végétalisé : bancs de graviers/galets ». Cependant, la binarité vivace et bas-marais n'est pas correcte. La binarité possible serait pelouse permanente vivace sèche ou humide, mais c'est une entrée biotope qui n'a pas sa place dans la physionomie ;
- Il a été impossible de délimiter les « bancs de sables (2114 et 2124) » et « bancs de graviers (2115/2125) » au niveau des données d'entraînement. Le niveau 4 a donc été abandonné. La distinction « Minéral non végétalisé (2120)/Minéral végétalisé (2110) » a été conservée au niveau 3 mais elle nécessite une bonne connaissance de terrain (observations à plusieurs saisons au même endroit). Cependant sur le territoire étudié, la distinction avec le poste « Herbacé haut (3210) » a également été difficile. Le modèle a donc parfois surestimé certains de ces postes en raison d'une différenciation peu marquée ;
- La distinction « Surface en eau non végétalisée 6100/ Surface en eau végétalisée 6200 » n'a pas été conservée car le poste « surface en eau » est inclus dans un masque IGN.
- Le poste « Arbustif / Fourré / Fruticée mixte indéterminé (4210) » est à rapprocher de « Herbacé haute indéterminée (3210) » pour le POC Nord, du moins au niveau de la cellule paysagère qui sera une « cellule de recolonisation » ;
- Les postes « Verger (7111) », « Vigne (7211) » et « Culture annuelle (81000) » n'ont pas fait l'objet de fourniture de données d'entraînement car ces postes sont inclus dans les masques IGN ;
- Certains polygones de segmentation niveau 2 ont été parfois fusionnés ou redécoupés afin d'avoir une unité homogène et/ou le seuil minimal de surface ;
- Les deux PNR sont des zones avec assez peu de physionomies dans leur ensemble.

Tableau 11 : Nomenclature des cellules paysagère de niveaux 4

ID NIVEAU4	LIBELLE NIVEAU4	PHYSIONOMIE	PNRSE		PNRAv		Total
			Nombre polygones	Surface (ha)	Nombre polygones	Surface (ha)	
2110	Minéral indéterminé	végétalisé	22	21,4	9	5,9	31
2120	Minéral indéterminé	non végétalisé	9	7	-	-	9
3111	Pelouse permanente vivace		-	-	4	1,6	4
3210	Herbacé haute indéterminée		18	13,3	24	16,3	42
3221	Prairie fauchée		32	44,8	42	61,4	74
3222	Prairie pâturée		23	41,4	35	77,8	58
4210	Arbustif / Fourré / Fruticée mixte indéterminé		2	4,6	8	36,7	10
5210	Forestier pionnier à dominance de feuillus indéterminé		15	46,5	-	-	15
	Somme		121	179	122	199,9	243

Description succincte des données en entrée

Le tableau suivant synthétise les données utilisées en entrée

Tableau 12 : Description des données d'entraînement

<u>M1 – méthode de transcription de données (relevés) de production de données d'apprentissage biotope</u> <u>M2 – méthode expert de production de données d'apprentissage biotope</u> <u>M3 – méthode de production de données d'apprentissage physiologiques</u>					
TYPE	DESCRIPTION (nombre de données disponibles mais pas nécessairement utilisées au final)	PREPARATION DES DONNEES	DISPONIBILITE NATIONALE	MÉTHODE / PRODUIT OBTENU	PERTINENCE POUR LA MODELISATION
Relevé floristique (occurrence de taxon) CBN	Données non comptabilisées, utilisées à l'opportunité.	Aucune car usage comme donnée complémentaire.	Donnée la plus répandue dans le réseau des CBN.	M2 / V0-V1 et V2 biotopes	* (biotopes)
Occurrences surfaciques de syntaxons CBN	5938 polygones avec mention de 1 à n syntaxons.	Tri /conversion	Hétérogène en fonction des CBN	M1 et M2/ V0-V1 et V2 Biotopes M3/ Physiologie	*** (biotopes et physiologies)
Données de séries et de géoséries de végétation	246 points et environ 10 000 ha de séries.	conversion des séries, en biotopes Carhab2	Donnée très peu répandue et sur des territoires assez ciblés (PNR...)	V0-V1 et V2 des biotopes M3/ Physiologies	**** (biotopes et physiologies)

Répartition spatiale des données d'entrainement fournies

→ PNRSE



Figure 6 : Données d'entrainement surfaciques produites par le CBNBL pour réaliser la modélisation prédictive des biotopes sur le PNRSE

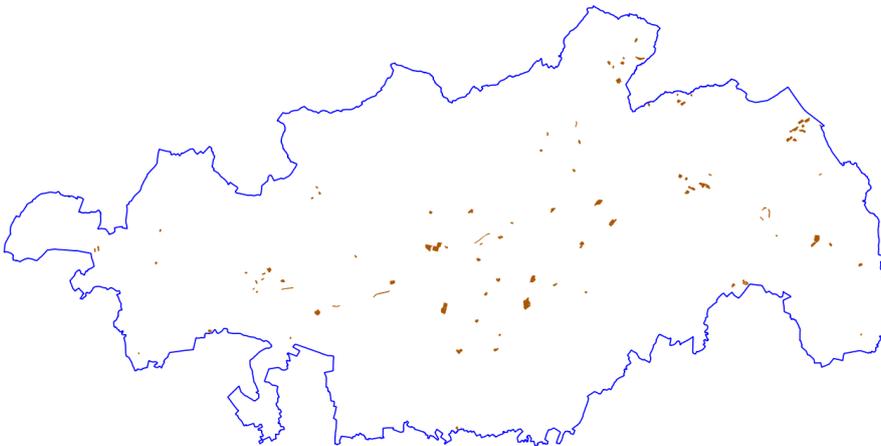


Figure 7 : Données d'entrainement produites par le CBNBL pour réaliser la modélisation prédictive des physiognomies sur le PNRSE

→ PNRAv

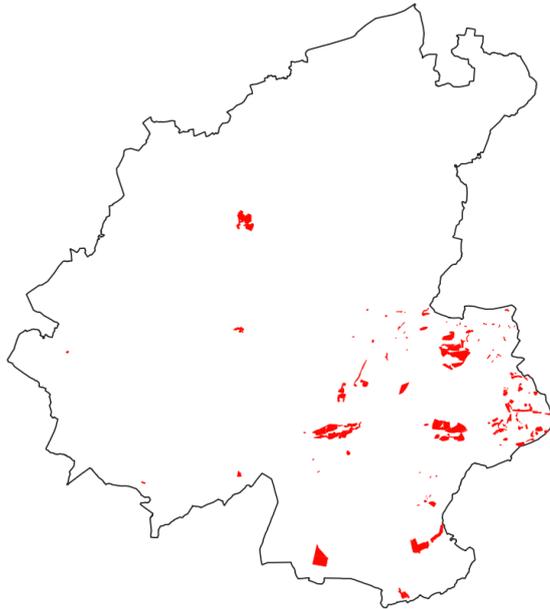


Figure 8 : données d'entraînement surfaciques produites par le CNBL pour réaliser la modélisation prédictive des biotopes sur le PNRAv

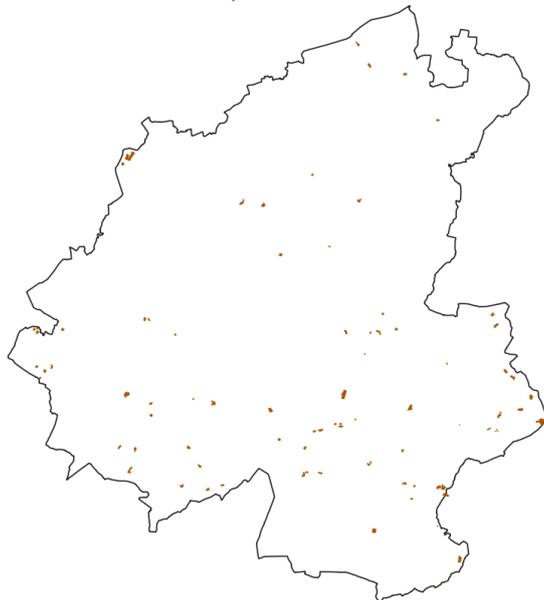


Figure 9 : Données d'entraînement produites par le CNBL pour réaliser la modélisation prédictive des physiognomies sur le PNRAv

Cartographie des biotopes

En matière de cartographie des biotopes, les principaux efforts ont porté (Figure 10) sur :

POC	Données d'apprentissage Biotopes B			Algorithme ~		Variables V		
	Effort de formatage des données	Adaptations Bruit, Incertitude Redondance	Création de protocole de validation terrain	Paramétrages	Versionnage	Modélisation spatiale des Biotopes	Aide à l'élagage des Biotopes	Elaboration de variables spécifiques
Loiret / Cher	X	X	XX	X	XXX	X		
Scarpe-Escaut / Avesnois	XX	XX	X	XX	XX	X		XXX
Dordogne	XX	XX		X	XX	X	XX	XXX
Vanoise	XXX	XXX		XXX	X	X	XX	XX

Figure 10 : Tableau récapitulatif des efforts méthodologiques réalisés sur le POC

→ Le formatage des données d'apprentissage qui a permis d'alimenter la réflexion sur la nécessité de standardiser les données d'entraînement (Bellenfant, 2019a). Il a ainsi été mis en évidence le bruit généré par les éléments linéaires inférieurs à la maille du pixel, lors de leur transformation au format raster (Figure 11). Pour exemple, la surface d'apprentissage passe de 2.72 km² à 11 km² une fois les données converties au format raster (pixel de 25 m), soit presque 5 fois plus. La différence correspond au bruit généré. Et bien que cette différence diminue à environ 6 km² si l'on change de résolution (pixel de 10 m), elle reste génératrice de biais pour le modèle d'apprentissage. Ce biais impacte de deux façons différentes le processus de modélisation :

- Le concept du biotope concerné est associé à des combinaisons de variables non pertinentes (bruit). Ceci diminue la séparabilité des biotopes pour le modèle, ce qui se traduit par une probabilité d'assignation plus faible d'une étiquette de biotope à un pixel,
- La séparabilité des biotopes basée sur l'association entre variables et données d'entraînement étant moindre, l'algorithme a tendance à sur-apprendre (overfit) afin de discriminer les données d'entraînement et le modèle en résultant intègre des relations spé cieuses qui nuiront à la qualité de l'extrapolation ;

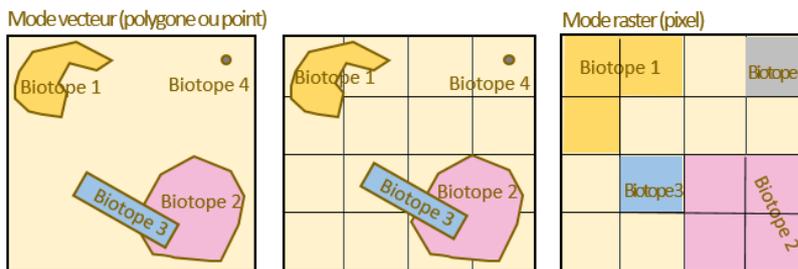


Figure 11 : Représentation des données en Mode vecteur et Mode Raster

- ➔ La gestion de la redondance au sein du jeu d'entraînement. Par redondance, on entend le fait qu'un même polygone soit étiqueté par au moins deux informations différentes de biotopes. Ceci peut résulter :
 - d'un artefact de numérisation des enveloppes de biotopes (chevauchements des polygones), ce qui admet une solution géomatique de détection et d'élimination des zones concernées,
 - de versions contradictoires d'expertise d'une zone, non identifiées comme telles lors de la fusion/consolidation des sources de données, ce qui est remédié en sensibilisant l'opérateur des bases de données gérées par le CBN,
 - ou enfin, de l'incertitude de l'expert botaniste (cas rencontré dans d'autres POC) pour assigner de manière non ambiguë un seul biotope à une donnée d'entraînement spatialisée. Le modélisateur n'étant pas en mesure de trancher (l'étiquetage de la donnée d'entraînement étant équiprobable), une stratégie de modélisation alternative adaptée à ce multi-étiquetage a commencé à être testée sur le POC Nord afin, non de lever des ambiguïtés dans les données d'apprentissage mais au contraire à tirer parti d'information incomplète. Par exemple, certains des points de validation de la modélisation en Avesnois n'ont pu être expertisés pour l'acidité du sol. Cette information manquante, l'assignation d'une étiquette de biotope est théoriquement impossible et disqualifie ainsi le point visité en tant que donnée d'entraînement. La stratégie de modélisation multi-étiquetage pourrait permettre d'utiliser l'information sur l'humidité du sol disponible malgré l'information d'acidité manquante (développements en cours) ;
- ➔ L'application du protocole de validation terrain mis en place sur le POC Loiret (Bellenfant, 2019b) pour comparaison et test de reproductibilité ;
- ➔ Plusieurs versions ont été livrées selon des résolutions différentes (25 m → 10 m), selon le traitement simultané ou individualisé des 2 PNR et selon le jeu d'apprentissage.
- ➔ Au niveau des variables, un important travail a été conduit sur la variable « zones potentiellement humides » afin de répondre aux attentes locales sur le PNR SE (voir partie « Contexte de l'étude »).

Zones potentiellement humides (ZPH)

Dans le cadre de la cartographie des biotopes, une variable « Zones Potentiellement Humides (ZPH) » destinée à alimenter l'apprentissage supervisé est produite à partir d'une méthode multicritère (Sacca, 2019). Les attentes en matière de zones humides sur le Parc naturel régional Scarpe-Escaut ont conduit à décliner la production de cette variable à l'échelle du parc. Cela s'est traduit par l'intégration de deux nouveaux critères : la géologie et la densité de canaux et de mares (Figure 12). En ce qui concerne la géologie, les substrats ont été classés de façon ordinaire à dire d'experts selon leurs rôles plus ou moins favorables à l'hydromorphie. À travers la densité de canaux et de mares, ce sont les secteurs les plus humides qui sont recherchés ; les canaux indiquant très souvent des secteurs drainés ; la présence de mares témoignant de l'imperméabilité du substrat.

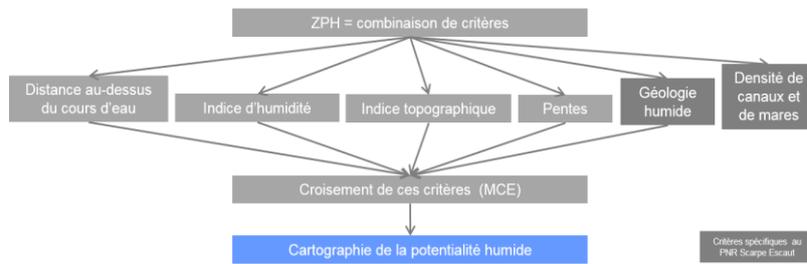


Figure 12 : Schéma de la méthode ZPH déclinée sur le PNRSE

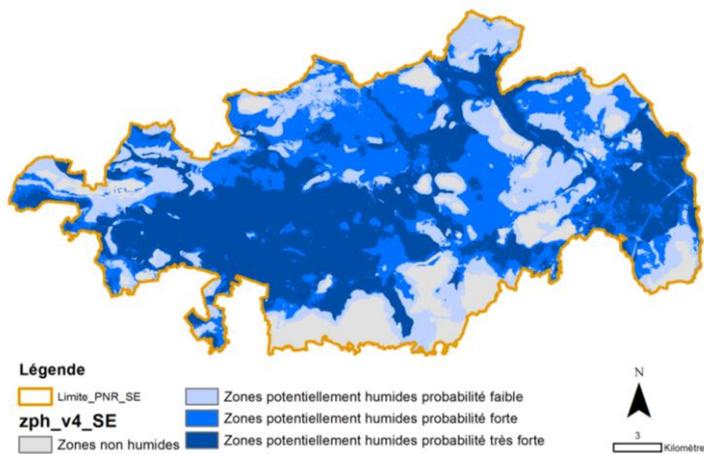


Figure 13 : Zones potentiellement humides PNRSE

Le résultat cartographique traduit pixel par pixel, la plus ou moins grande potentialité humide selon que la combinaison des critères est plus ou moins favorable (Figure 13). La discrétisation statistique des valeurs offre une première hiérarchisation du territoire selon la potentialité humide. La seconde étape a nécessité une étroite collaboration avec le Conservatoire Botanique National de Bailleul (CBNBL). Elle consiste à affiner les classes supérieures en les calibrant avec des données de végétation. Chaque classe est travaillée indépendamment et affinée si besoin. Pour exemple, la distinction entre l'hydrophile et l'hygrophile s'est faite conjointement sur les critères botaniques et densité de canaux et de mares. Certes plus longue, cette étape a néanmoins, largement amélioré la calibration résultats. Elle a ainsi permis la distinction des différents niveaux d'humidité (Figure 14).

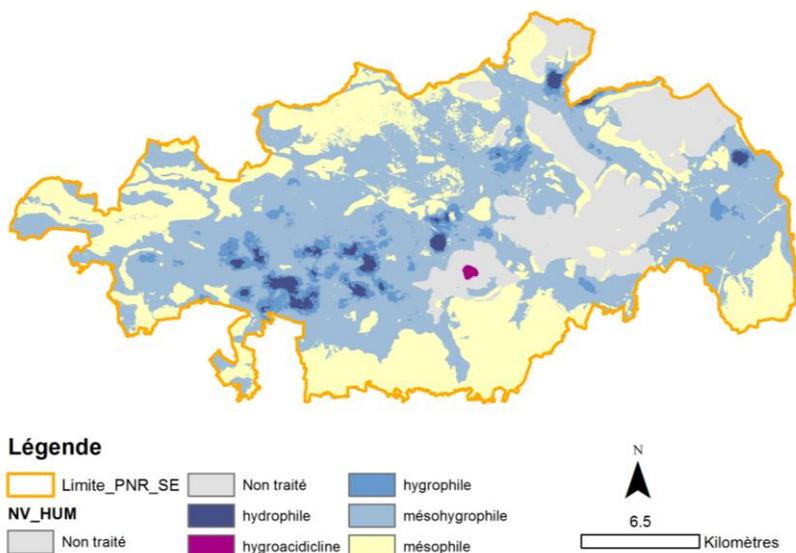


Figure 14 : Carte prédictive des niveaux d'humidité sur le PNRSE

Sensibilité des données radar à l'hydromorphie des prairies du PNRSE

Les images radars ont montré dans de nombreuses études un fort potentiel pour la cartographie des milieux humides. Le CBN de Bailleul a réalisé en 2018 des relevés de végétations ainsi que des cartes de niveaux d'humidité potentiels sur le parc régional Scarpe Escaut. Ces données d'échantillons peuvent contribuer à étudier la capacité des capteurs radars à caractériser les zones humides et plus particulièrement à capter l'hydromorphie des prairies humides. Grâce à sa forte répétitivité d'acquisition, Sentinel 1 permet d'accéder à la dynamique saisonnière des surfaces inondées et à la teneur en eau du sol. Un deuxième radar PALSAR sur ALOS est à tester puisqu'il fonctionne en bande L plus sensible à l'humidité du sol. Il s'agit donc de tester la détection d'une variable environnementale (hydromorphie) sur une physionomie particulière (prairies). Techniquement, il s'agit d'étudier le potentiel des données radar en bandes C et L pour la cartographie des coefficients d'humidité du sol définis selon la classification d'Ellenberg. Des mesures des coefficients d'humidité du sol associés à des unités spatiales (cellules paysagères) ont été utilisées. Les coefficients d'humidité se basent sur les indices d'Ellenberg (1988) qui attribue à chaque plante indicatrice une échelle d'humidité de 1 à 9. Étant donné que le signal radar en bandes C et L ne pénètre pas dans un couvert forestier bien développé, seules les cellules minérales peu végétalisées, herbacées et prairiales sont retenues pour l'analyse. Ces cellules ont des coefficients d'humidité du sol de 5, 6, 7 et 9. Le nombre des cellules est respectivement 106, 161, 21, et 2 pour les coefficients d'humidité du sol de 5, 6, 7 et 9.

Résultats

Méthode d'évaluation des résultats mise en place sur le POC

Les versions 0, 1 et 2 de la carte des biotopes ainsi que les premières versions des modélisations prédictives des physionomies ont fait l'objet d'avis d'expert au bureau.

Seules les versions 3 de la modélisation prédictive des biotopes et les versions 3 et 4 des zones potentiellement humides ont fait l'objet de vérifications sur le terrain à raison de six jours environs pour les biotopes et les physionomies. Sur les six jours, un peu moins de deux jours ont été consacrés aux vérifications pour la physionomie niveau 4 de l'Irstea.

La méthodologie de terrain pour effectuer le contrôle qualité de la carte prédictive des biotopes (version 3 du POC) est similaire à celle réalisée pour le POC Loiret (note méthodologique rédigée par le CBNBP).

L'évaluation des biotopes se fait selon deux paramètres (Acidité et humidité) en raison d'une homogénéité des autres paramètres sur ce POC. Ainsi, trois catégories d'évaluation sont définies :

- Bonne : l'information modélisée rattachée au paramètre évalué correspond à l'information observée ;
- Moyenne : l'information modélisée rattachée au paramètre évalué correspond à un niveau d'écart par rapport à l'information observée (Exemple : Mésophile – Mésohygrophile ; Acidiphile – Acidicline) ;
- Erronée : l'information modélisée rattachée au paramètre évalué ne correspond pas à l'information observée ou a deux niveaux d'écart par rapport à l'information observée (exemple : Mésophile – Hygrophile ; Acidiphile – Neutrocline).

Il est à noter qu'en raison de la faible amplitude de l'échelle d'acidité (4 niveaux, **Erreur ! Source du renvoi introuvable.**), la notation appliquée n'a pas la même incidence que celle de l'échelle d'humidité (6 niveaux, **Erreur ! Source du renvoi introuvable.**Figure 16**Erreur ! Source du renvoi introuvable.**). Une évaluation moyenne pour l'acidité correspond à un écart plus important entre la modélisation et le constat de terrain qu'une évaluation moyenne pour l'humidité. Cependant, elle a été appliquée afin de rester homogène par rapport à la notation de

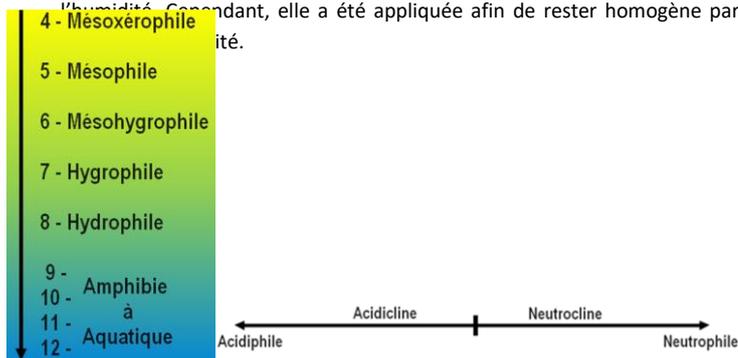


Figure 15 : Échelles d'acidité et d'humidité

Évaluation des cartes des biotopes et ZPH produites

Quelques tableaux de synthèse des résultats sont insérés dans les commentaires qui suivent.

Grille d'analyse de la modélisation des biotopes V0, V1 et V2	
→ Établie d'après analyses à dire d'expert (analyse de la carte au bureau/confrontation avec certains paramètres et données d'apprentissage produites et la connaissance du territoire) et fourniture d'un jeu de données ponctuelles	
Représentativité globale à petite échelle	Les données ponctuelles ont eu tendance à influencer le modèle dans le mauvais sens. Les biotopes associés étaient principalement des biotopes aquatiques et ne devraient pas représenter de si grandes zones surfaciques
Biotopes sous représentés	Absence de nombreux biotopes qui devraient être présents en raison d'une sur représentation des biotopes aquatiques
Biotopes sur représentés	Biotopes secs et thermophiles, biotopes aquatiques
Cartographie des zones potentiellement humides (ZPH) - V0	Incomplète et aux contours inexacts, nécessite un recalibrage
Tests supplémentaires et leviers envisagés	Relancer la modélisation avec les améliorations apportées aux paramètres suites aux échanges (paramètres calculés à 10 m au lieu de 25 m, données d'entraînement surfaciques au lieu de ponctuelles, modélisation basée sur zph_v4 de la V1-ZPH pour le PNRSE et la zph_v1 de la V1-ZPH pour le PNRAV, lancer la modélisation de façon indépendante pour chaque PNR).

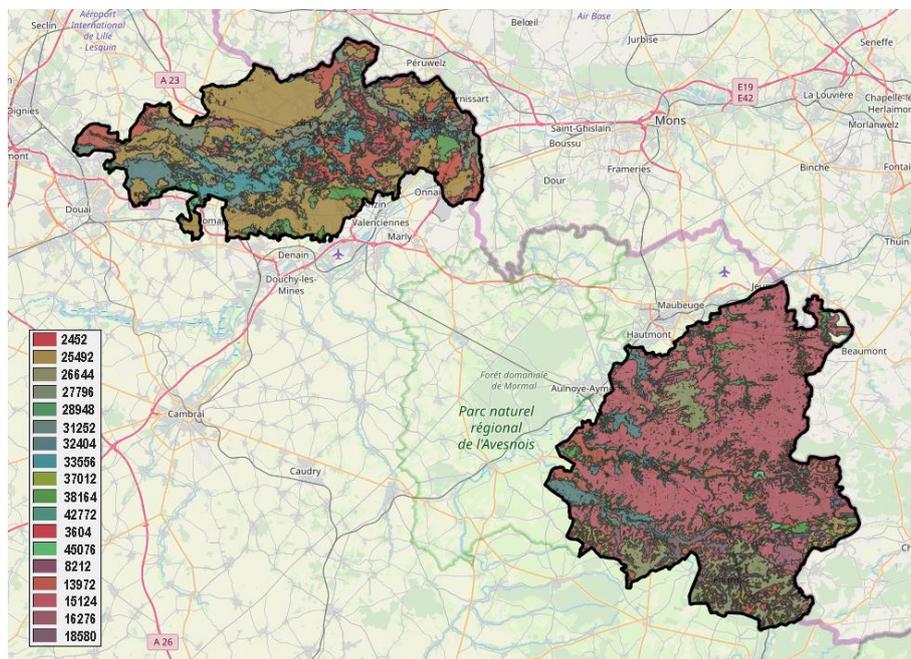


Figure 16 : Carte prédictive des biotopes V0

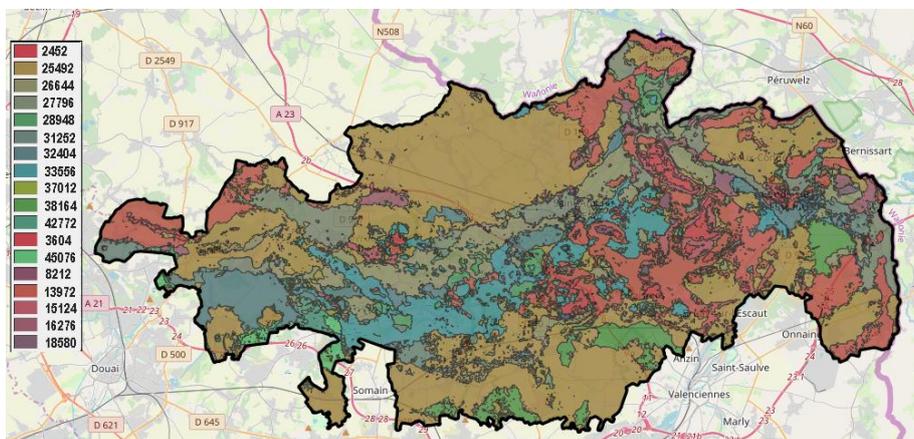


Figure 17 : Carte prédictive des biotopes V0 du PNRSE, surreprésentation des biotopes aquatiques (32404, 33556) et mésophiles (25492)

Grille d'analyse de la modélisation des biotopes V3

→ Établie d'après analyses à dire d'expert (analyse de la carte au bureau/confrontation avec certains paramètres et les données d'apprentissage produites et la connaissance du territoire) et fourniture d'un nouveau jeu de données surfaciées, mise à disposition du fond géologique par le BRGM

Représentativité globale à petite échelle	Les biotopes sont mieux représentés avec la géologie. Certains biotopes paraissent assez bien représentés mais sont parfois encore trop larges (13972, 26644, ...)
Biotopes sous représentés	Mésophiles et mésohygrophiles
Biotopes sur représentés	Biotopes hygrophiles et hydrophiles et prennent parfois le dessus sur les biotopes sous représentés Pour le PNRAv, les zones alluviales ont pris une trop grosse place dans la modélisation des biotopes et ont sur représenté les niveaux hygrophiles et hydrophiles. Les biotopes, concernant l'humidité, sont représentés avec un décalage d'un niveau le plus souvent
Biotopes non modélisables	Végétations aquatiques, non modélisées (abandon 10516, 22036, 32404, 33556 et 45076)
Cartographie des zones potentiellement humides (ZPH)	Le seuillage des niveaux a permis d'obtenir une version intermédiaire satisfaisante. Quelques ajustements sont à prévoir pour une carte plus affinée. Cette version se rapproche d'une carte prédictive des biotopes pertinente pour le PNRSE.
Tests supplémentaires envisagés	Relancer une modélisation avec ajouts de données d'apprentissage suite aux vérifications terrain (PNRSE = stage, PNRAv = vérification biotopes et physiologies) et un éventuel équilibrage (à la baisse) des biotopes sur représentés Affinage des seuils concernant les niveaux d'humidité Modéliser un paramètre séparément des autres (piste évoquée avec EVS-Isthme)

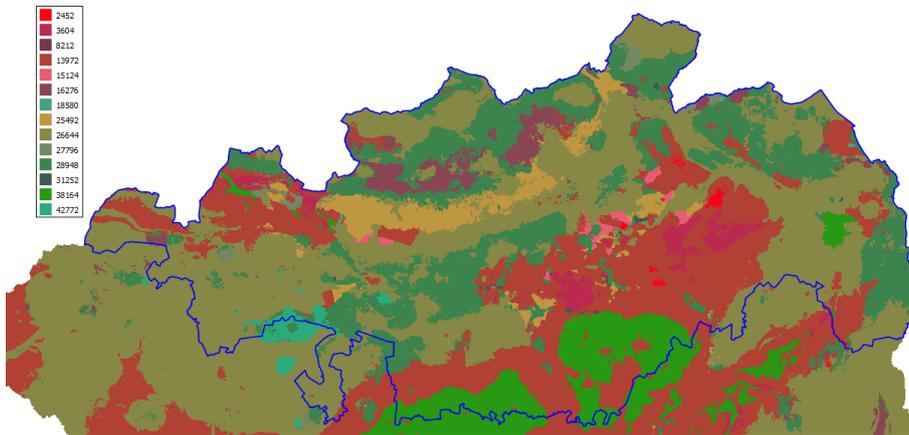


Figure 18 : Carte prédictive des biotopes du PNRSE, V3

Évaluation de la carte prédictive des biotopes

○ Analyse des données du PNRAV

109 données rattachées à un biotope ont été analysées sur le PNR Avesnois en fonction du paramètre acidité ou humidité et selon la typologie du biotope donnée. La concordance entre la modélisation et la réalité terrain nécessite des ajustements au niveau de la modélisation mais est globalement satisfaisante (Figure 19). En effet, la typologie des biotopes est faiblement bonne (16 %) en raison d'un des deux facteurs n'étant pas correct. Un seuillage de ces paramètres permettrait d'affiner ces résultats et de se rapprocher de la réalité. Les niveaux d'humidité modélisés présentent une estimation correcte avec 39 % de niveau bon et 40 % moyen (le niveau d'humidité est décalé d'un seul niveau). C'est le même constat pour l'acidité avec 39 % de niveau bon et 49 % de niveau moyen. Dans les deux cas, la modélisation donne assez peu d'information erronée, moins de 23 % d'erreur concernant l'humidité et moins de 11 % concernant l'acidité.

Les niveaux indiqués dans la Figure 19 ont été informés comme dans le Tableau 13.

Tableau 13 : Explication de l'attribution de la valeur de concordance pour les biotopes et les physionomies

Concordance d'un des deux paramètres	Concordance de l'autre paramètre		Résultat concordance biotope
Bonne	Bonne	→	Bonne
Bonne	Moyenne	→	Moyenne
Bonne	Erronée	→	Erronée
Moyenne	Moyenne	→	Erronée
Moyenne	Erronée	→	Erronée
Erronée	Erronée	→	Erronée

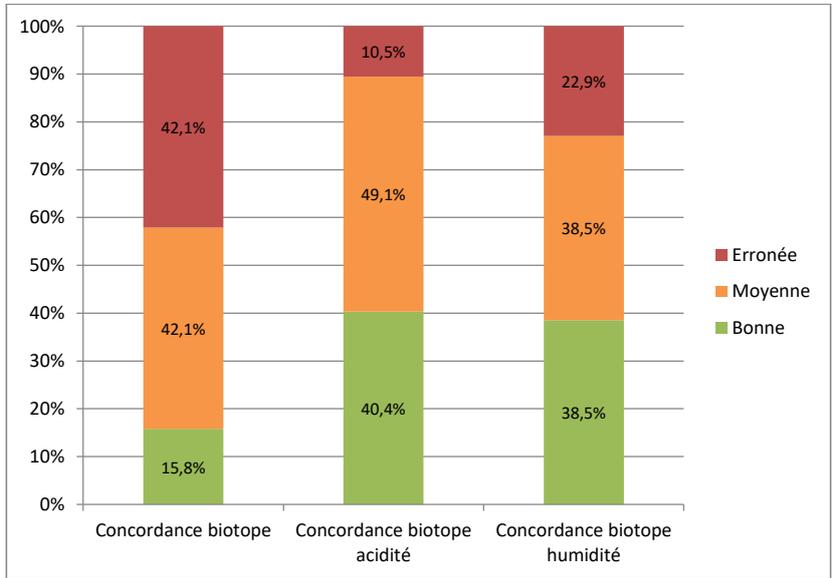


Figure 19 : Concordance du biotope, des niveaux d'humidité et d'acidité selon la modélisation en fonction de l'observation terrain sur le PNRAV

Évaluation de la carte prédictive des physionomies

Grille d'analyse des physionomies	
→ Établie d'après analyses à dire d'expert (analyse de la carte au bureau/confrontation avec les données d'apprentissage produites et connaissance du territoire) et vérifications <i>in-situ</i>	
Représentativité globale	Elle est moyenne du fait des masques (bâti, cultures) qui ne sont pas exacts mais aussi de la modélisation en elle-même (il faudrait entraîner le modèle avec tous les postes physionomiques y compris les cultures et les forêts) La distinction prairie fauchée et prairie pâturée n'est pas aisée (niveau 4). Au niveau de restitution (niveau 2), le code prairie est bon.
Physionomies sous représentées	Sous-représentation des « complexes de recolonisation », herbacée haute indéterminée (3210) : confusion avec des prairies à fauche tardive (3221) et complexes de recolonisation Aucune surreprésentation nette.
Physionomies sur représentés	2120 (Minéral non végétalisé) : confusion avec les cultures, zones bâties, chemin, prairies retournées 3110 (pelouses permanentes vivaces) : confusion avec les cultures, les prairies 4210 (Arbustif, fourré, fruticée mixte indéterminé) : confusion avec des postes forestiers (5210), coupes ou complexes de recolonisation
Physionomies non modélisables	Manque de données, difficultés de délimiter les échantillons : 2110 et 2120 (Minéral végétalisé et non végétalisé) 3112 (Pelouses annuelles) Déclinaisons niveau 4 du 211X (2111, 2112,...), 212X, 321X, 421X et 521X
Cartographie des zones potentiellement humides (ZPH)	N'est pas évalué dans le cadre des physionomies, paramètres exclusivement du ressort des biotopes
Tests supplémentaires Leviers pour pallier aux mauvais résultats pour une V2 de la physionomie	Appliquer les masques a posteriori de la modélisation Éliminer les polygones biaisant le modèle (2110 et 2120) Fusionner certains postes se rapprochant des complexes de recolonisation (fourrés, herbacée haute) Lancer la modélisation avec tous les postes physionomiques y compris les cultures et les forêts
Remarques	Intérêt des vérifications terrain pour mesurer précisément les atouts et limites d'utilisation de la carte des physionomies (bon pour les prairies de manière générale mais moins bon sur la distinction fauche et pâture, attention pour pelouses, minéral...)

Évaluation de la sensibilité des radars à l'hydromorphologie des prairies

L'utilisation du radar Sentinel 1 (Bande C) semble très pertinente pour capter l'humidité des végétations herbacées naturelles ouvertes et basses. Il est important de retenir que le signal radar ne pénètre pas les couverts ligneux importants (buissons ou arbres). Les données Sentinel 1 présentent l'avantage d'être disponibles gratuitement sur plusieurs plateformes associées à Copernicus. Une méthode de classification basée sur l'utilisation de la technique d'apprentissage automatique pourrait être envisagée mais reste à développer. Elle pourrait déboucher sur un algorithme permettant de cartographier les indicateurs d'humidité du sol. Bien que de telles approches existent sur la cartographie de l'humidité des zones de culture, cela nécessite une adaptation des algorithmes utilisés et une augmentation de l'effort d'échantillonnage pour la calibration et la validation de ces modèles.

Évaluation de la carte prédictive des physionomies

→ Matrice de confusion

Les matrices de confusion permettent d'évaluer la qualité d'une carte. Les lignes d'une matrice représentent les échantillons de références qui sont utilisés pour l'évaluation et les colonnes représentent les classes assignées par le modèle à ces échantillons.

Ainsi, si l'on observe une classe sur une ligne, on peut voir comment elle a été effectivement classée par le modèle. On peut donc détecter les erreurs d'omission : est-ce que tous les pixels de la classe ont été bien classés dans la bonne classe ?

Si l'on regarde une classe sur une colonne, on peut voir à quelle classe de la vérité terrain elle correspond. On peut donc détecter les erreurs de commission : est-ce que la classe affectée par l'algorithme correspond en réalité à cette classe ? Si non, dans quelles classes les erreurs sont-elles commises ?

Des indices résument les performances de classification par classe :

- ✓ le rappel résume les erreurs d'omission, une valeur de 1 indique que tous les pixels de références sont bien classés ;
- ✓ la précision résume les erreurs de commission, une valeur de 1 indique que tous les pixels auxquels on a assigné la classe appartiennent effectivement à la classe de référence ;
- ✓ le F-score est une moyenne harmonique des deux précédents indices.

L'Overall Accuracy (OA) représente le pourcentage de pixels bien classés et le KAPPA est une mesure de précision globale qui prend en compte l'accord aléatoire.

La moitié des échantillons de physionomie est utilisée pour l'entraînement et l'autre moitié est utilisée pour la validation.

Les résultats de la modélisation de la physionomie de la végétation au niveau 4 sont présentés dans la **Erreur ! Source du renvoi introuvable.** On peut tout d'abord noter que les résultats sont différents pour les deux zones. Alors que la carte de l'Avesnois (figure 20(a)) présente une bonne précision (OA de 0.84), la carte de Scarpe-Escaut (figure 20(b)) a une précision moins satisfaisante (OA = 0.60).

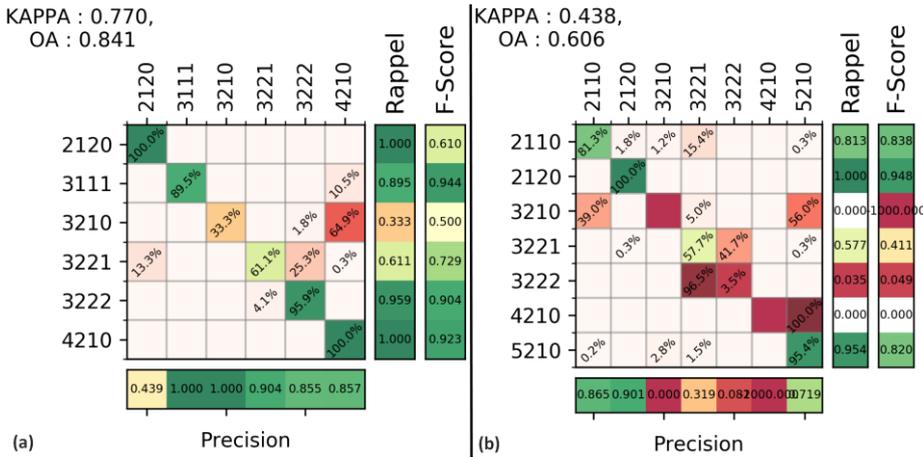


Figure 20 : Matrice de confusion de carte produite pour un niveau de nomenclature de niveau 4 pour (a) l'Avesnois et (b) Scarpe-Escaut. 2110 Minéral végétalisé indéterminé ; 2120 : Minéral non végétalisé indéterminé ; 3111 Pelouse permanente vivace ; 3210 : Herbacée haute indéterminée ; 3221 : Prairie fauchée ; 3222 Prairie pâturée ; 4210 : Arbustif / Fourré / Fruticée mixte indéterminé ; 5210 Forêt Semi-naturelle

Les principales confusions observées pour l'Avesnois concernent la classe 3210 (Herbacée haute) qui est classée par le modèle dans la classe 4210 (Arbustif / Fourré / Fruticée mixte). Pour la zone de Scarpe-Escaut, cette même classe est classée en 5210 (Forêt semi-naturelle) tandis que la classe 4210 (Arbustif / Fourré / Fruticée mixte) et elle-même classée en 5210. La classe Forêt semi-naturelle est donc certainement surreprésentée au détriment des deux autres classes. Il y a également des confusions importantes entre les prairies fauchées et pâturées (3221 et 3222). Si les résultats de la modélisation au niveau 4 sont seulement satisfaisants pour la zone de l'Avesnois, les résultats de la modélisation au niveau 2 de la nomenclature qui est le niveau de restitution attendu, sont meilleurs. Ces résultats sont présentés dans la Figure 21.

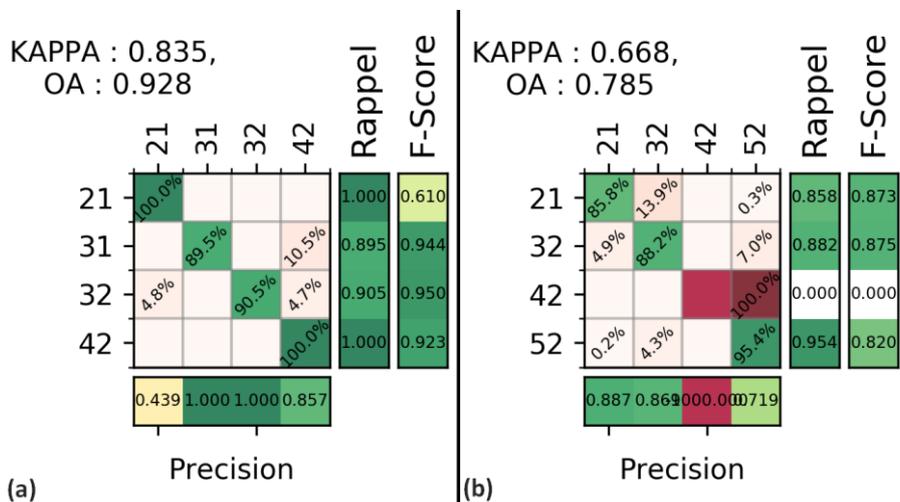


Figure 21 : Matrice de confusion de carte produit pour un niveau de nomenclature de niveau 2 pour (a) l'Avesnois et (b) Scarpe-Escaut. 21 : Minéral ; 31 Pelouse ; 32 : Prairie et Herbacée Haute; 42 : Arbustif / Fourré / Fruticée ; 52 : Forêt mature

On peut tout d'abord noter une très bonne précision globale (OA = 0.928) pour l'Avesnois et une bonne précision pour Scarpe-Escaut (OA = 0.785). La plus faible précision de la zone de Scarpe-Escaut est due au fait que la classe 42 (Arbustif / Fourré / Fruticée) est toujours confondue en 52 (Forêt mature) malgré le regroupement des classes. Cette confusion est probablement due à un déséquilibre du nombre de pixels pris en compte pour l'entraînement du modèle : le plus grand nombre de pixels de la classe 52 favorisant la prédiction de celle-ci.

Pour conclure, l'évaluation de la modélisation de la physionomie de la végétation montre que les résultats:

- ✓ sont variables d'une zone à l'autre ;
- ✓ sont bons pour le niveau 2 de restitution attendu ;
- ✓ sont bons pour l'Avesnois pour le niveau expérimental 4,
- ✓ sont décevants pour Scarpe-Escaut pour le niveau 4, avec notamment des confusions importantes entre prairies fauchées et pâturées, et les classes herbacée haute et arbustif / fourré / fruticée qui sont systématiquement classées en forêt semi-naturelle.

→ Vérification *in situ* dans le PNR Avesnois

La vérification *in situ* dans le PNRAV des physionomies indique une bonne concordance entre la donnée modélisée et la donnée observée (72 %). Cette vérification terrain a été réalisée sur 181 données de physionomies réparties comme indiqué dans le Tableau 14.

Tableau 14 : Nombre de données terrain analysées selon la physionomie modélisée. Les masques sont ceux de l'IGN et ne sont pas modélisés

Masque indéterminé	2
Masque culture	14
Masque forestier	36
Minéral non végétalisé	7
Pelouse	1
Prairie fauchée	34
Prairie pâturée	83
Arbuste/Fourré/Fruticée	4

Les résultats de la concordance entre la modélisation des physionomies et la donnée terrain sont présentés en Figure 22. Les postes ayant une concordance bonne sont ceux dont la donnée modélisée est identique à la donnée observée. Ceux qualifiés de « Concordance moyenne » correspondent aux données dont le niveau 4 est différent mais le niveau supérieur concordant (exemple d'une prairie de fauche modélisée et d'une prairie pâturée observée, le niveau 4 n'est pas le bon mais le niveau supérieur « Prairie » est bon). Enfin, la « concordance erronée » est indiquée quand les niveaux supérieurs sont différents (Pelouse modélisée et prairie observée, Masque forestier et prairie, etc.).

Comme indiqué en partie typologie des physionomies, les postes « Pelouse », « Arbuste/Fourré/Fruticée », « Minéral non végétalisé » sont ceux dont la modélisation est la moins pertinente. Certaines prairies ne sont pas indiquées en tant que tel en raison d'un masque occultant la modélisation.

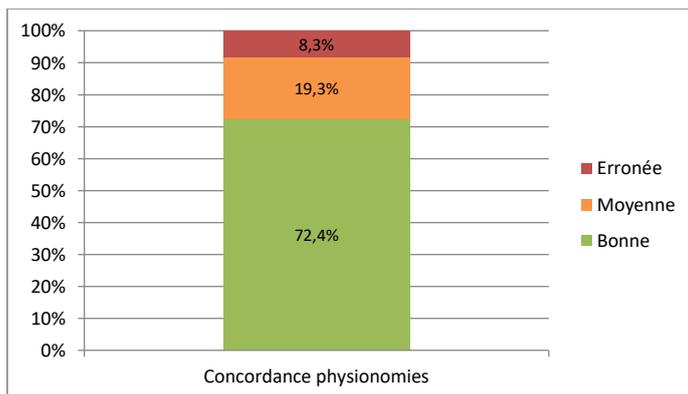


Figure 22 : Concordance entre la modélisation des physionomies et la donnée terrain

Remarque : si l'on se cantonne au niveau de restitution (niveau 2), la modélisation est bonne (mais les informations restituées moins précises).

Évaluation des charges et des coûts

Remarque sur le calcul de coût

Le calcul des coûts prend en compte l'intervention d'un chargé de missions en mode prestation (432 € euros/journée, soit un coût horaire de 54 €) et d'un coordinateur scientifique (688 €/journée, soit un coût horaire de 86 €). Signalons que ces coûts comprennent les fonctions supports du CBNBL (direction, informatique, traitement des données, comptabilité, etc.).

Coûts liés à la production des données :

- Définition des biotopes effectifs sur le territoire d'étude à partir du catalogue de biotopes

N.B. : Afin d'effectuer le tri des biotopes selon les variables exprimées sur le territoire, il est nécessaire d'avoir une connaissance du catalogue des biotopes. Cette constitution du catalogue n'est pas chiffrée dans les tableaux suivants mais serait à financer dans le cas du déploiement national

Phases	Temps passé (h)	Coûts
Tri des biotopes selon les variables exprimées sur le territoire	30	1 940 €
Croisement avec indices bioclimatiques (si concerné)	10	580 €
Total	40	2 520 €
Bilan		
Nombre de biotopes effectifs sur le territoire	36	
Estimation en mode opérationnel	Temps estimé (h)	Coût estimé
À l'échelle du Nord et du Pas-de-Calais (NPC)	280 (35 jours)*	19 600 €

*L'estimation en mode opérationnel est estimée de la façon suivante :

- Le POC Nord couvre 4 territoires phytogéographiques sur les 27 du Nord et du Pas-de-Calais (TOUSSAINT *et al.*, 2002), soit 1/7^{ème} du Nord et du Pas-de-Calais;
- 40 h ont été effectuées pour la définition des biotopes effectifs du POC Nord ;
- Sept fois plus de temps est à prévoir pour l'échelle NPC, soit 35 jours.

Cette estimation est à modérer. En effet, des économies d'échelles sont à prévoir comme par exemple de produire le travail bioclimatique à l'échelle française et ainsi ne pas reproduire plusieurs fois le travail à l'échelle locale. D'autre part, les territoires phytogéographiques présentent quelques hétérogénéités qui nécessiteront un travail spécifique. Compte tenu de ces deux tendances, l'estimation semble acceptable dans la perspective de dégager une tendance nationale.

De manière à obtenir une statistique extrapolable à l'échelle nationale, dans la mesure où le Nord et le Pas-de-Calais comptent 7 HER de niveau 2 (dont le littoral qu'on peut compter comme une HER), il semble possible de retenir le chiffre arrondi de 5 jours par HER.

○ Mise au format/production des échantillons de biotopes

Phases	Temps passé (h)	Coûts
Préparation des données existantes	30	1 864 €
Production des données d'entraînement	10 (dont 4 de données expertes)	540 €
Structuration des données d'entraînement	3	162 €
Total	43	2 566 €
Bilan		
Nombre de données analysées/mobilisées	>6 000	
Nombre de données d'apprentissage produites	1532 (1096 PNRSE+436 PNRAv)	
Temps moyen de préparation par données (min)	2 min et 30 sec : donnée préexistante *3 min et 30 sec : donnée experte	
Coût moyen de préparation par donnée	1,5 € pour la préparation + 1,4 € pour la production	
Estimation en mode opérationnel	Temps estimé (min)	Coût estimé
Temps et coût par biotope (25 préexistantes, 25 expertes)	262,5	308 €
Temps et coût pour une HER similaire (environ 20 biotopes,)	PNRSE (HER bien connue) : 1 300 min (2,7 j)	PNRSE (HER bien connue) : 1 517 €
	PNRAv (HER mal connue) : 3 950 min (8,2 j)	PNRAv (HER mal connue) : 4 608 €
	Moyenne pour une HER : 5,5j 75 % du travail est réalisé, soit en opérationnel, une moyenne par HER de : 7,3 j	Cout moyen pour une HER : 3 080 € Cout pour une HER en opérationnel : 4 088 €

*Le temps réalisé de 3 min et 30 sec pour la création d'une donnée experte ne prend pas en compte la préparation des données sources nécessaires à la production (sélection des espèces et des végétations guides). Le chiffrage actuel s'est principalement appuyé sur des données exploitables directement en raison d'un traitement préalable pour la préparation des données pré-existantes. Nous estimons donc le temps moyen de préparation par donnée experte à 8 minutes. Les calculs en mode opérationnel sont donc basés sur cette estimation.

- Mise au format/production des échantillons de physionomies

Phases	Temps passé (h)	Coûts
Préparation des données existantes	4	248 €
Production des données d'entraînement	15	970 €
Vérification de l'actualité de la donnée sur image		
Structuration des données d'entraînement	3,5	189 €
Total	22,5	1 407 €
Bilan		
Nombre de données existantes mobilisées	6500 données végétations, utilisation de la carte des physionomies CarHAB1	
Nombre de physionomies présentes sur le territoire	8, mais le nombre global serait à augmenter afin de prendre en compte la totalité des physionomies (dont les masques forêts, vergers, cultures, ...)	
Nombre de données d'apprentissage produites	243, soit environ 30 données par physionomie	
Temps moyen de préparation par données (min)	10	
Coût moyen de préparation par données	12 €	
Estimation en mode opérationnel	Temps estimé (min)	Coût estimé
Temps et coût par physionomie	300	360 €
Temps et coût pour une HER similaire (environ 10 physionomies)	3000 (6,3 j)	3 600 €

- Définition de l'arbre de décision pour le croisement biotope/physionomie

Phases	Temps passé (h)	Coûts
Tri des habitats potentiellement présents sur le territoire		
Définition de la nomenclature		
Définition des règles de décisions		
Total		
Bilan		
Nombre d'HIC présents sur le territoire		
Nombre d'autres habitats définis		
Temps moyen de définition par habitat (HIC et EUNIS) (min)		
Estimation en mode opérationnel	Temps estimé (h)	Coût estimé
Temps et coût pour un HER similaire		

Commenté [GV1]: NON réalisé pour l'instant

○ Echanges entre POC et modélisateurs

Phases	Temps passé (h)	Coûts
Echanges entre organismes (expertise technique)	Téléphoniques : EVS : 8h UMS : 1h Télédétecteurs : 3h Partenaires internes : 15h GT : 48h Echanges mails : 15h	5 250 €
Production nouvelles données d'entraînement/ règles de décisions	5	302 €
Structuration des données	1	54 €
Total	81	5 606 €
Bilan		
Nombre de versions produites	Biotopes : 4 Zones potentiellement humides : 5 Physio : 2	
Estimation en mode opérationnel	Temps estimé (h)	Coût estimé
Temps et coût pour une HER similaire	40	2 800 €

○ Coûts de validation des données produites (biotope, physio et habitat)

Méthode	Temps passé (h)	Coûts
Expertise	4 h (n'est pas pris en compte la mise en forme des données extraites)	280 €
Mobilisation des données existantes	3 h (n'est pas pris en compte la mise en forme des données extraites)	210 €
Contrôle terrain biotopes (V3)	84 h	5 432 €
Contrôle terrain physionomies	28 h	1 928 €
Total	119 h (soit 15j) pour les deux HER	8 400 €
Bilan		
Volume échantillonnage de validation	PNRSE : validation bureau (via les données d'un stage) PNRAv : 100 points biotopes et 131 points physionomies	
Estimation en mode opérationnel	Temps estimé (h)	Coût estimé
Temps et coût pour une HER similaire	Moyenne pour une HER : 60 h 75 % du travail est réalisé, soit en opérationnel, une moyenne par HER de : 80 h	5 600 €

- Coûts et délais liés aux réunions et échanges entre producteurs et chef de projet en local ainsi que pour la comitologie de suivi de projet (intervient tout au long du projet. A des fins d'estimation)

Coûts environnés liés à la formation et appropriation de la méthodologie

Phases	Temps passé (h)	Coûts
Information et partage des objectifs et de la méthode (GT1-1 à GT1-8, 2 réunions GT2 et 2 COTECH, réunions internes)	180	12 600 €
Nomenclature biotopes (dont 2 réunions GT3 biotopes, réunions téléphoniques, échanges mails).	70	4 900 €
Nomenclature physiologies	20	1 400 €
Échantillonnage physiologies (incl. réunion atelier)	20	1 400 €
Echantillonnage biotopes (incl. Rédaction note, réunion POC Planitiaire)	50	3 500 €
Notes de comptes rendus, notes informatives	30	2 100 €
Total	370	25 900 €

Coût réel interne

Phase	Total	COUT
Nombre de jours 2019	88	43 776€
Préparation + généralités (002)	18 (8j mis en 003)	8 032€ (-3 456 €)
Contrôle biotopes (003)	21,5 (+8)	12 104 € (+ 3 456 €)
Contrôle physio (008)	4	1 984 €
Biotopes - Données entrainement (004)	14,5	7 416 €
Rédaction (005)	2	1 120 €
Réunions (006)	15,5	8 104 €
Croisement physio+biotopes (007)	-	
Physio - Données entrainement (009)	3,5	1 640 €
Total	79	40 400 €
Ajout temps restant	+9j	+3 888 €
Ajout jours consommés 2018	+6j	+3 880 €
Ajout coût stage		+6 480 €
Ajout frais de déplacement mission		+1 200 €
Ajout frais déplacement Paris		+3 250
Ajout frais déplacement réunion région		+800 €
Total 100 % (environ)		59 900 €

- Coûts extrapolé à l'échelle du Nord et du Pas-de-Calais, de la région Hauts-de-France et de la France

NPC = 6 HER + Littoral

HdF = 11 HER + 1 Littoral = 12 HER

France = 13 régions

Rappel du coût journée (CBNBL) : 560 €

Phases	Temps passé (jours) - HER	NPC	HdF
Définition des biotopes effectifs (jours)	5	35	60
Mise au format/production échantillons biotopes (jours)	7,3	51,1	87,6
Mise au format/production échantillons physio (jours)	6,3	44,1	75,6
Echanges entre POC et modélisateurs (jours)	5	35	60
Coûts de validation des données produites (jours)	10	70	120
Total temps (jours)	33,6	235,2	403,2
Total coût	18 816 €	131 712 €	225 792 €

- Coûts supplémentaires pour la suite du programme (à faire remonter pour le rapport du GT 1 mais ne figurera pas dans le rapport du POC)
 - Formation des organismes publics et des utilisateurs locaux pour l'utilisation des données produites
 - Acquisition de données supplémentaires pour les territoires où les données sont trop peu nombreuses pour avoir une validité statistique ou ne permettent pas de construire des données d'entraînement ou de validation sur l'ensemble des biotopes
 - Rédaction de la notice des cartes, en lien avec la production des catalogues des séries.

Présentation de cas d'usage des données produites dans l'étude

Cas des zones humides sur le PNRSE

L'Agence de l'eau Artois-Picardie a développé un programme de maintien de l'agriculture en zones humides dans le cadre duquel il est prévu de cartographier 75 % des prairies humides de chacun des territoires-cibles, selon une typologie distinguant trois niveaux d'humidité : mésophile, mésohygrophile, hygrophile.

Le programme CarHAB de cartographie de la végétation de la France, tel qu'il a été expérimenté dans la plaine de la Scarpe et de l'Escaut (Catteau & Villejoubert., 2018), se révèle utile pour prédéfinir les prairies humides potentielles.

En 2019, la proposition a été d'évaluer la pertinence de la cartographie CarHAB réalisée sur le territoire de la plaine de la Scarpe et de l'Escaut pour prédire la localisation de prairies humides et de compléter cette cartographie par une campagne de terrain sur les parcelles dont le statut est douteux.

Cette méthodologie distingue 5 niveaux d'humidité :

- **Mésophile** (inondation très rare ou sols faiblement hydromorphes)
- **Mésohygrophile** (inondation de quelques semaines par an ou sols hydromorphes, saturés en eau en surface de quelques semaines par an ; Végétations des sols bien alimentés en eau, avec un engorgement temporaire de surface en début de saison de végétation)
- **Hygrophile** (inondation 1 à 3 mois/an ou sols hydromorphes, saturés en eau en surface de de 1 à 3 mois/an ; Végétations des sols présentant un engorgement à proximité de la surface durant une partie de la période de végétation.)
- **Hydrophile** (inondation supérieure à 3 mois/an ou sols hydromorphes, saturés en eau en surface de plus de 3 mois/an ; Végétations des sols présentant un engorgement à proximité de la surface durant l'essentiel de la saison de végétation)
- **Aquatique** (regroupe amphibie et aquatique ; des Végétations des sols saturés d'eau toute l'année et inondés pendant la majeure partie de l'année ou pendant toute l'année, sauf épisode climatique exceptionnel aux végétations de milieux aquatiques)

Cette information des niveaux d'humidité est accessible dans le programme CarHAB à partir de la cartographie des biotopes (un des paramètres écologiques est le gradient hydrique). La localisation des prairies est, quant à elle, renseignée dans la carte des physionomies.

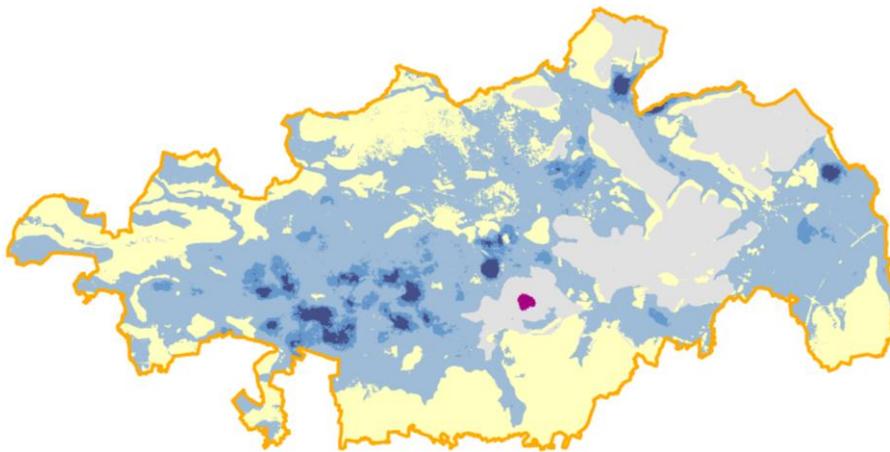
L'information sur le gradient hydrique est accessible de deux manières, soit en exploitant un travail spécifique réalisé par le laboratoire EVS-Isthme sur les zones potentiellement humides (ZPH), soit en exploitant le paramètre humidité dans la carte des biotopes.

Méthode 1 : exploitation de la carte des zones potentiellement humides

EVS-Isthme a obtenu le paramètre de gradient hydrique indépendamment des autres paramètres afin de produire une carte des zones potentiellement humides, aussi appelée carte des niveaux d'humidité (Figure 23).

La dernière version (V4) de la carte ZPH (04 juin 2019) correspond au reclassement de la V3 des ZPH qui s'appuie sur les retours et calibrations du CBNBL et tente d'optimiser pour chaque niveau d'humidité (« non traité » correspond aux zones dont l'indication du gradient hydrique n'a pas été fournie mais qui seront pris en compte sur le plan du gradient hydrique pour les prochaines versions).

Il est à noter que le niveau « aquatique » n'a pas été représenté car celui-ci fait l'objet d'un masque IGN. La couche finale sera un croisement du masque IGN et de la carte des ZPH.

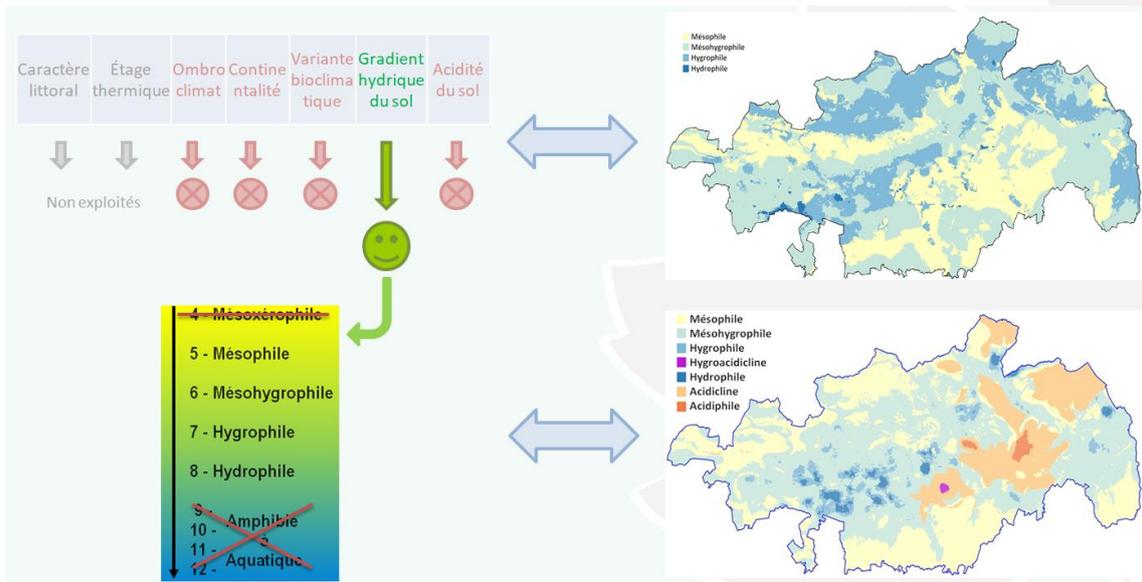


Légende



Figure 23 : Carte des niveaux d'humidité potentiels modélisés sur le PNRSE

Méthode 2 : exploitation du paramètre « gradient hydrique » de la carte des biotopes



Résultat

La carte des ZPH est très pertinente à l'échelle du PNRSE et sera à prendre en considération afin d'améliorer la carte des biotopes.

À terme, il est plus satisfaisant d'exploiter la carte des biotopes puisqu'elle bénéficiera des mises à jour et des progrès de la connaissance (nouvelles données d'entraînement) améliorant ainsi sa pertinence.

Points de vigilances identifiés pour le passage à l'échelle

- Biais lié aux biotopes ponctuels → ne pas fournir de données d'entraînement pour ces biotopes ;
- Les données de biotopes de petites surfaces ne retranscrivent pas toujours suffisamment la réalité de terrain pour que le modèle soit juste en sortie.
- Les polygones type « linéaire » sont peu pertinents dans le cadre de la modélisation de biotopes (difficultés de les rattacher à des pixels homogènes de par leur longueur) ;
- L'utilisation séparée des jeux de données d'entraînement est plus pertinente en raison de contextes écologiques entre les deux parcs et notamment en raison de présence/absence de biotope dans un PNR par rapport à l'autre ;
- Les données expertes sont beaucoup plus pertinentes pour l'entraînement du modèle. Il est nécessaire de favoriser ce type de données d'entraînement.

Suite à ces observations, le CBNBL a choisi de fournir un second jeu d'entraînement pour chaque PNR (PNRSE et PNRAV distincts) basé sur des données surfaciques (> 1ha) en ne tenant pas compte des polygones de type « linéaire ». Les biotopes de type « ponctuels » n'ont pas été fournis pour ces jeux de données car ils ont tendance à induire un biais dans la modélisation (présent sur de grandes surfaces). Ce sont 5 biotopes qui ne sont pas repris pour la modélisation.

Bibliographie

- Bellenfant, S., (2019a).** Note technique - Protocole d'échantillonnage des données de biotope pour la modélisation prédictive. Conservatoire botanique national du Bassin parisien.
- Bellenfant, S., (2019b).** Note technique - Contrôle de qualité terrain : méthodologie et résultats. Conservatoire botanique national du Bassin parisien.
- Catteau E. et Villejoubert G., (2018).** Inventaire et cartographie des végétations - Expérimentation de la méthodologie CarHAB dans le Parc naturel régional Scarpe-Escaut. – Rapport d'étude – CBNBL, Bailleul – 72 pages + Annexes pp.
- Ellenberg, H., (1988).** Vegetation ecology of Central Europe. Cambridge University Press.
- Rivas-Martínez S., Rivas Saéñz S. & Penas A., (2011).** Worldwide bioclimatic classification system. Global Geobotany, Vol. nº 1. December 2011. pp. 1-634 + 4 Maps.
- Sacca, C., (2019).** Note sur la variable « Zones Potentiellement Humides » produite dans le cadre du programme CarHab 2 et adaptée au PNR Scarpe-Escaut
- Toussaint B., Hendoux F. et Lambinon J., (2002).** Définition et cartographie des territoires phytogéographiques de la région Nord-Pas-de-Calais. – Revue Botanique, éd. Lejeunia / Société Botanique de Liège. 171(171) : 43 p.