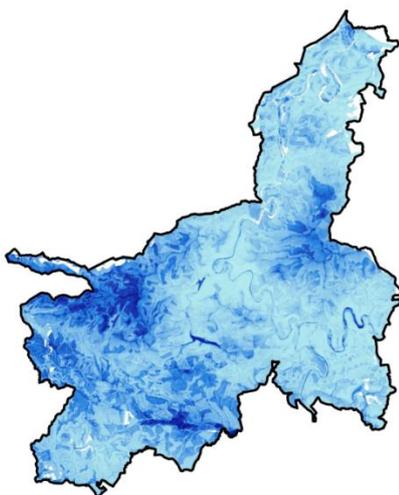


# Localisation des zones humides forestières potentielles dans le nord des Ardennes

Territoire du PNR/Bassin de la Meuse



## Auteurs :

AgroParisTech ENGREF – Promotion Gestion des Milieux Naturels 2014-2015

Raphaël BEC, Julie CRABOT, Amandine DIOT, Camille GAZAY, Claire GELY, Delphine JAYMOND

### **Avertissement**

Le travail présenté dans ce rapport a été réalisé par des étudiants dans un temps limité, et avec avant tout un objectif pédagogique. Malgré le soin qui y a été apporté, il peut comporter des erreurs dans les résultats, ou manquer de précision, par manque de temps. Un travail plus approfondi serait nécessaire avant une application directe et des résultats définitifs. Nous invitons donc le lecteur à garder un regard critique.

### **Remerciements**

Nous adressons tous nos remerciements à nos commanditaires pour leur disponibilité et l'ensemble des informations apportées, notamment Marie Lemoine (Agence de l'eau Rhin-Meuse), Jérémy Garin, Philippe Lecomte, Vincent Spanneut, Alain Lescouet (ONF), Virginie Graitson Schmitt, Violaine Laucoin et Julien Crugnola (PNRA), Céline Cordani, Xavier Caron et Nathalie Mear-Caubel (EPAMA).

Merci également à André Vincent, maire de Hargnies, pour nous avoir accueilli le jour de la présentation, ainsi qu'à Alain Bernard, maire de Monthermé, pour son investissement et son intérêt quant à notre projet.

Nous sommes également reconnaissants envers Lucien Dapvriil et Michel Colcy pour nous avoir accompagnés sur le marais des Hauts-Buttés, ainsi qu'envers Bernard Delhaye pour ses nombreux documents. Plus généralement, nous remercions toutes les personnes nous ayant accordé de leur temps au cours des entretiens.

Enfin, un grand merci à nos encadrants : Philippe Durand, Bernard Jabiol, Marieke Blondet, Jean-Claude Gégout et Vincent Perez, pour leur présence et leur soutien tout au long du projet.

# Sommaire

I.	Études sur les zones humides sur le territoire de PNRA-bassin de la Meuse.....	4
1.1	Introduction générale.....	4
1.1.1	Contexte.....	4
1.1.2	Présentation du périmètre d'étude.....	4
1.1.3	Présentation des acteurs.....	5
a.	Le Parc naturel régional des Ardennes.....	5
b.	L'Établissement public d'aménagement de la Meuse et de ses affluents.....	5
c.	L'agence de l'eau Rhin-Meuse.....	5
1.2	Aspects climatiques, géologiques et écologiques de la zone d'étude.....	6
1.3	Contexte socio-économique de la zone d'étude.....	7
1.4	Définition des zones humides et législation.....	8
1.5	Présentation des 3 études et de leurs objectifs.....	9
1.5.1	Localisation des zones humides forestières potentielles ;.....	9
1.5.2	Mise en place d'un outil d'aide à la hiérarchisation des zones humides ;.....	9
1.5.3	Élaboration d'une stratégie de restauration du marais des Hauts Buttés.....	9
II.	Objectifs de l'étude.....	10
III.	Matériels et méthode.....	10
3.1	Données à disposition.....	10
3.2	Méthode de régression logistique.....	11
3.3	Sélection des variables et construction du jeu de données.....	13
3.3.1	Variable à expliquer.....	13
3.3.2	Variables potentiellement explicatives.....	13
3.3.3	Sélection des variables explicatives par le modèle.....	15
3.4	Méthodologie dans l'utilisation des données flore et stations.....	17
3.4.1	Objectifs.....	17
3.4.2	Méthode.....	17
a.	Flore.....	17
b.	Stations.....	20
c.	Flore de zone humide.....	22
3.5	Méthodologie de validation par une phase de terrain.....	22
IV.	Résultats.....	23
4.1	Cartographie des zones humides potentielles.....	23
4.1.1	Validation du modèle par comparaison avec la présence réelle de zone humide sur les placettes de calibration.....	24
4.2	Résultats – Phase de validation (terrain).....	25
4.2.1	Carte des probabilités de présence de zone humide.....	25
4.2.2	Cartographie des placettes terrain selon leur classe de zone humide.....	26
4.2.3	Carte des probabilités de présence d'une zone humide et leurs liens avec l'engorgement.....	29
4.3	Comparaison avec les données existantes.....	30
4.3.1	Comparaison avec les données flore et stations.....	30
a.	Flore.....	30
b.	Stations.....	31
4.3.2	Comparaison avec les autres cartes et modèles.....	31
V.	Discussion, limites et perspectives.....	34
5.1	Facteurs d'imprécision dans la modélisation.....	34
5.1.1	Flore.....	35
5.1.2	Stations.....	36
5.2	Limites dans l'utilisation des résultats.....	36
5.3	Extrapolation du modèle.....	37
VI.	Conclusion.....	37

# I. Études sur les zones humides sur le territoire de PNRA - bassin de la Meuse

## 1.1 Introduction générale

### 1.1.1 Contexte

Face à la régression des zones humides au cours du siècle dernier, les outils pour leur protection et leur restauration se sont multipliés : convention RAMSAR, loi sur l'eau... Les différents organismes impliqués dans la gestion de ces milieux ont également pris conscience de leur importance écologique. Ainsi, dans le nord des Ardennes, l'ONF, le Parc naturel régional des Ardennes (PNRA) et l'Établissement public d'aménagement de la Meuse et de ses affluents (EPAMA) ont placé la préservation de ces zones au centre de leurs politiques environnementales. En effet, en raison de conditions climatiques actuelles et passées particulières et d'un contexte géologique et pédologique spécifique, le département des Ardennes présente de nombreuses zones humides, notamment en forêt. Certaines d'entre elles, à fort intérêt patrimonial, sont déjà protégées et nécessitent d'être davantage étudiées afin d'être correctement gérées ; d'autres doivent encore être localisées avant toute étude.

### 1.1.2 Présentation du périmètre d'étude

Le PNRA se situe au nord du département des Ardennes et couvre une superficie de 1200 km<sup>2</sup> avec un point culminant de 504m d'altitude. Étant à cheval sur les vallées de la Meuse et de la

Semoy, sur le plateau forestier ardennais et sur une partie de la Thiérache ardennaise, ce territoire est soumis à une gestion concertée entre les différents acteurs. Ainsi, notre étude porte plus particulièrement sur les 900km<sup>2</sup> de territoire commun au PNRA et au bassin versant de la Meuse, gérés en partenariat avec l'EPAMA, en charge de l'aménagement de la Meuse.

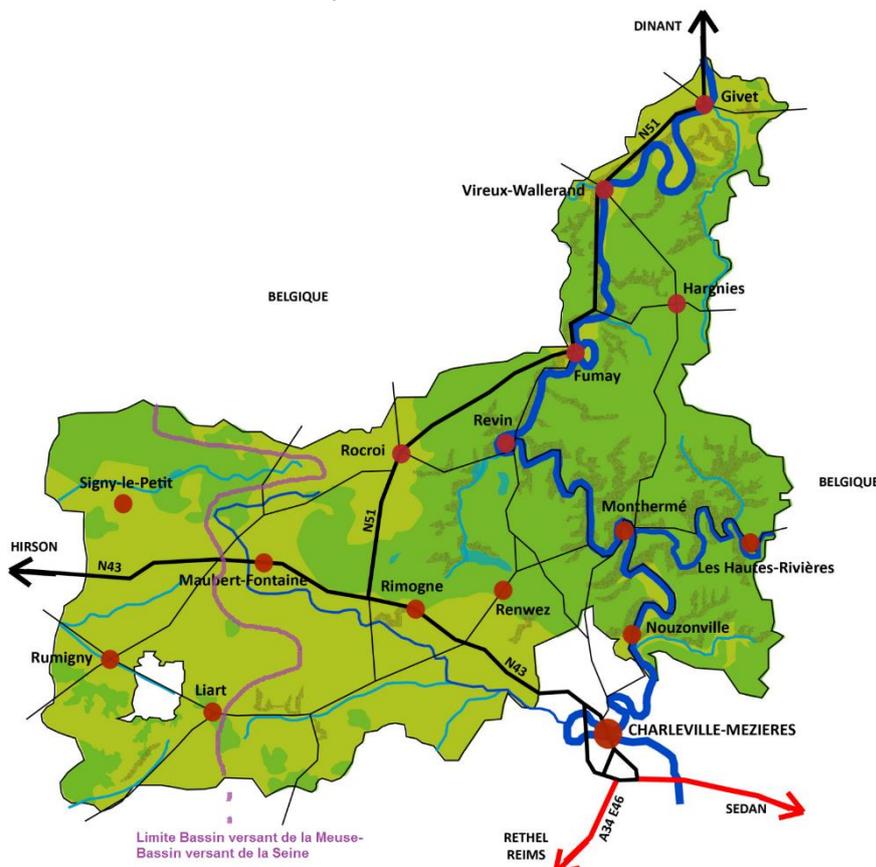


Figure 1 : Localisation de la zone d'étude (Source : PNRA)

### 1.1.3 Présentation des acteurs

#### a. Le Parc naturel régional des Ardennes

Labellisé en 2011, le Parc est une structure encore jeune dont la charte a été rédigée en 2012, celle-ci étant valable pour une période de 12 ans. Ce document fixe les objectifs à atteindre ainsi que les mesures et actions permettant d'y parvenir. Le Parc assume alors son rôle :

- Répondre aux enjeux de préservation et de valorisation du patrimoine naturel ;
- Favoriser un développement durable
- Promouvoir l'éducation au territoire.

Elle permet également d'assurer la cohérence et la coordination des actions menées sur les 91 communes adhérentes par les diverses collectivités publiques.

#### b. L'Établissement public d'aménagement de la Meuse et de ses affluents

Face aux inondations catastrophiques de décembre 1993 et de janvier 1995 les élus lorrains et champardennais, en concertation avec le Préfet de Lorraine, Préfet coordonnateur de Bassin, ont décidé de mettre en place l'Établissement Public d'Aménagement de la Meuse et de ses Affluents, l'EPAMA.

Il regroupe les régions Champagne-Ardenne et Lorraine, les départements des Ardennes, de la Haute-Marne, de la Meuse et des Vosges ainsi que de nombreuses communes ou groupements de communes riveraines de la Meuse.

Sa principale mission est la réalisation d'une étude complète de modélisation des écoulements en crue de la Meuse afin de mettre en place des mesures de protection. Cet objectif est aujourd'hui complété par l'étude des zones humides du secteur.

#### c. L'agence de l'eau Rhin-Meuse

L'AERM est un établissement public du ministère chargé du développement durable. Sa zone de compétence couvre 3 régions et 8 départements (Bas-Rhin, Haut-Rhin, Moselle, Meurthe-et-Moselle dans leur totalité ; Ardennes, Haute-Marne, Meuse, Vosges en partie). Dans le cadre de son dixième programme (2013 – 2018), l'agence de l'eau participe à la mise en œuvre des politiques nationales relatives à l'eau, qui s'inscrivent elles-mêmes dans les objectifs de l'Union européenne.

Elle contribue à réduire les pollutions de l'eau de toutes origines et à protéger les ressources en eau et les milieux aquatiques :

- grâce à des soutiens financiers à des investissements,
- par la construction et le développement d'outils de planification,

- par la production et l'exploitation de données sur l'eau pour la connaissance, la gestion et l'évaluation.

Constatant le mauvais état physique de nombreuses zones humides, la protection ou la restauration de ces dernières constitue une part importante du programme mis en œuvre.

## 1.2 Aspects climatiques, géologiques et écologiques de la zone d'étude

D'un point de vue climatique, les Ardennes sont à l'**interface entre le climat continental** de l'Est de la France, et l'**influence maritime** de la Manche. Cependant, le climat local est très **nuancé** sur la zone d'étude, notamment du fait d'une **variété de conditions topographiques et altitudinales**. Ainsi, la Pointe de Givet et les plaines sont globalement plus chaudes et sèches que les hauteurs du massif Ardennais, qui reçoivent davantage de précipitations.

En termes de **géologie**, la zone d'étude est relativement diversifiée. On y trouve :

- **L'Ardenne primaire**, relief accidenté caractérisé par un socle hercynien constitué de roches métamorphiques pauvres
- Les **Crêtes pré-ardennaises et la Thiérache ardennaise** au Sud-ouest qui présentent des roches sédimentaires beaucoup plus récentes
- Une enclave sédimentaire calcaire au niveau de la **pointe de Givet**.

La **topographie** qui en découle est variée, avec des zones de plateau, des vallées encaissées, des éperons rocheux et des collines en périphérie du massif. Au sein de la zone d'étude, l'altitude varie globalement entre 100 et 500m (504m pour le point culminant côté français).

Cette diversité climatique, géologique et topographique explique en partie la **diversité d'habitats naturels** présents sur le périmètre d'étude. Ainsi, le massif Ardennais est globalement couvert par la **forêt**, parfois très anthropisée (présence de plantations d'espèces allochtones) mais qui présente aussi des habitats d'intérêt communautaire (forêt à caractère montagnard, ...). La Thiérache, le plateau de Rocroi et la pointe de Givet, quant à eux, concentrent des habitats agropastoraux, dont certains très patrimoniaux comme les **pelouses calcaires**.

L'érosion du massif ardennais a également abouti à la formation d'éperons rocheux et autres **milieux rupicoles**, qui hébergent des espèces emblématiques comme le Grand-Duc d'Europe, ou le Faucon pèlerin.

Cependant, ce sont bien les **zones humides** qui constituent l'originalité du plateau et des vallées ardennaises. Milieux aujourd'hui menacés par les aménagements et la déprise agricole, les tourbières, bois tourbeux, prés humides et autres rizières sont le joyau naturel de la pointe des Ardennes. Ces habitats, patrimoniaux pour la plupart, s'inscrivent dans le grand contexte hydrographique de la Meuse et de ses affluents que sont la Semoy et la Sormonne. Riches en espèces, ils abritent aussi bien une **flore d'exception** avec 5 plantes protégées

nationalement et 9 régionalement (Dactylorhiza sphagnicola, deux espèces de Drosera, Trientalis europea,...) qu'une faune rare et menacée. La Cigogne noire aime s'y nourrir, tout comme l'Engoulevent d'Europe ou la Bécasse des bois. S'y ajoutent 6 espèces d'amphibiens, 5 de reptiles et 4 espèces d'insectes protégées nationalement qui participent grandement à l'intérêt écologique de ces habitats humides.

Ces derniers font aujourd'hui l'objet de mesures de protection et de restauration, notamment par l'intermédiaire du réseau **Natura 2000**, et à travers la politique du **Parc naturel régional des Ardennes**, qui promeut un développement local respectueux des enjeux environnementaux.

### 1.3 Contexte socio-économique de la zone d'étude

Sur le périmètre d'étude, l'essentiel du tissu urbain et de l'activité économique est **concentré autour du réseau hydrographique**, et en particulier le long de la Meuse.

Si par le passé l'agriculture et la sylviculture contribuaient fortement à l'économie locale, c'est bien **l'industrie** qui, depuis un siècle, représente le centre névralgique de l'économie ardennaise. Cette industrie, alimentée par la production de charbon de bois sur le massif ardennais, s'est développée dans la vallée jusqu'à la crise du secteur métallurgique dans les années 80. Aujourd'hui, elle représente toujours une part non négligeable de l'activité économique, même si les services et le tourisme ont pris le relais dans ce contexte difficile.

En parallèle, du fait de conditions topographiques plus favorables, une **activité agricole** extensive perdure sur le plateau de Rocroi, tandis que la Thiérache est caractérisée par une mosaïque de bocage et de grandes cultures. La **sylviculture** est toujours active sur le massif, où l'on identifie une majorité de forêts domaniales ou communales, bénéficiant du régime forestier. Dans ces forêts, on pratique aussi des loisirs comme la **chasse au grand gibier**, ou encore l'affouage<sup>1</sup> qui génère un lien fort entre les habitats et la forêt. Le Sanglier est d'ailleurs l'un des symboles majeurs des Ardennes.

La **balance démographique** est quant à elle **déficitaire**, malgré un fort taux de natalité. Les crises successives subies par l'industrie locale, en accentuant le taux de chômage, ont abouti à une émigration importante. Cependant, les communes rurales sont moins touchées, car le cadre de vie demeure attractif, notamment pour les urbains en quête d'un retour à la nature.

---

<sup>1</sup> Voir Glossaire

## 1.4 Définition des zones humides et législation

Les zones humides sont des milieux de vie remarquables possédant une végétation et une faune spécifique. La loi sur l'eau de 1992 établit une définition des zones humides. En vue de préparer cette loi, une première définition scientifique des zones humides a été proposée en 1991 sous la direction du Muséum national d'histoire naturelle.

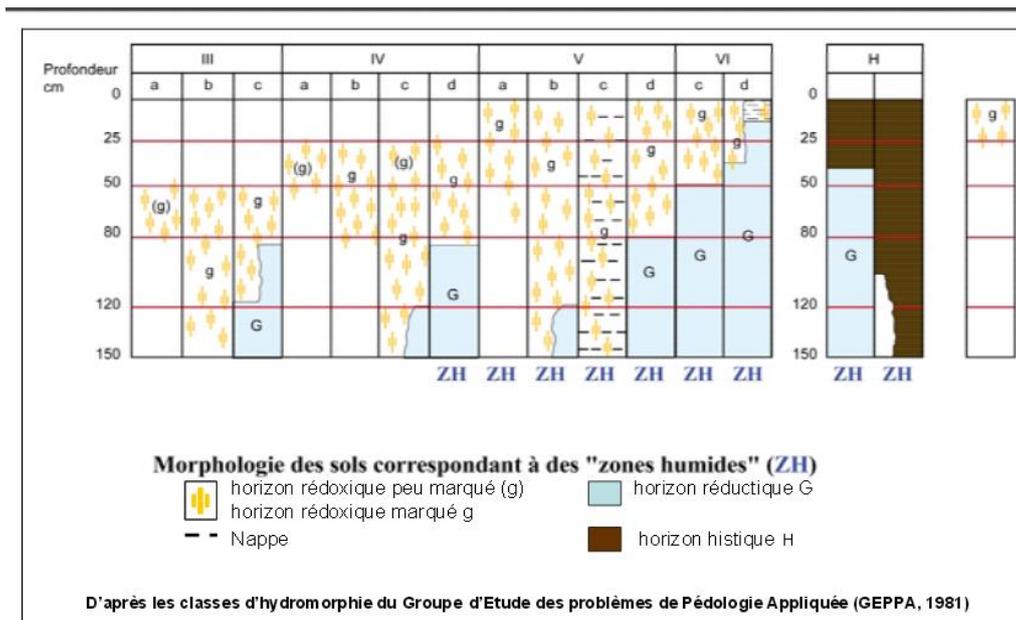
Selon la loi sur l'eau de 1992, les zones humides sont des « terrains, exploités ou non, habituellement inondés ou gorgés d'eau douce, salée ou saumâtre de façon permanente ou temporaire ; la végétation, quand elle existe, y est dominée par des plantes hygrophiles pendant au moins une partie de l'année ».

Cette définition a été précisée en annexe de l'arrêté du 24 juin 2008 mis à jour par celle de l'arrêté du 1<sup>er</sup> octobre 2009<sup>2</sup>, donnant deux critères permettant de délimiter les zones humides : l'hydromorphie du sol et la présence de plantes hygrophiles. Il suffit que l'un des deux critères soit rempli pour que la zone soit considérée comme humide.

### ➤ Critère relatif à l'hydromorphie du sol

Les sols considérés comme correspondant à des zones humides sont :

- les histosols
- les réductisols
- les sols possédant des traits rédoxiques débutant à moins de 25 cm de la surface du sol
- les sols possédant des traits rédoxiques débutant à moins de 50 cm de la surface du sol et se prolongeant par des traits réductiques entre 80 et 120 cm de profondeur.



<sup>2</sup> Figure 2 : Classes d'hydromorphie. Les classes Vb, Vc, Vd, VI, H correspondent à des sols de zones humides, les classes IVd et Va peuvent être exclues par le préfet de région après avis du conseil scientifique régional du patrimoine naturel

➤ **Critère relatif à la flore**

- Présence et abondance d'espèces hygrophiles définies dans l'annexe 2.1. de l'arrêté
- Présence d'habitats caractéristiques des zones humides définis dans l'annexe 2.2. de l'arrêté.

## **1.5 Présentation des 3 études et de leurs objectifs**

Dans le cadre du module de dernière année « Bassin versant et ressource en eau » dispensé par l'école d'ingénieur d'AgroParisTech, centre de Nancy, les étudiants en domaine d'approfondissement « Gestion des milieux naturels » doivent acquérir les compétences nécessaires à la gestion de la ressource hydrique. C'est dans ce cadre qu'un travail de 5 semaines est réalisé sur le territoire commun au PNRA et au bassin versant de la Meuse en collaboration avec l'ONF, le PNRA et l'EPAMA.

Trois études intimement imbriquées sont menées en parallèle.

### **1.5.1 Localisation des zones humides forestières potentielles ;**

En milieu forestier, la présence du couvert rend impossible l'utilisation de photos aériennes ou satellites pour la prélocalisation des zones humides. Les inventaires de ces zones passent donc essentiellement par des prospections sur l'intégralité du territoire qui les rendent très coûteux. L'objectif de cette étude est d'établir une méthodologie de prélocalisation des zones humides forestières.

### **1.5.2 Mise en place d'un outil d'aide à la hiérarchisation des zones humides ;**

Chaque zone humide identifiée est importante. Cependant, les moyens d'actions étant limités, il est intéressant de prioriser l'intervention selon les enjeux du territoire, la fonctionnalité des zones humides et leur vulnérabilité. L'objectif de cette étude est de construire un outil opérationnel d'aide à la hiérarchisation des zones humides.

### **1.5.3 Élaboration d'une stratégie de restauration du marais des Hauts Buttés.**

La commune de Monthermé constitue l'une des zones d'action prioritaires du PNR puisqu'elle est propriétaire d'une zone humide d'intérêt : le marais des Hauts-Buttés. Les objectifs de cette étude sont de comprendre et analyser la situation actuelle du marais, de caractériser son fonctionnement hydrologique et d'analyser les services passés et actuels rendus par le site. Ces éléments doivent permettre d'aboutir à la proposition de prémices de plan de gestion.

## II. Objectifs de l'étude

En tant que milieux à forts enjeux patrimoniaux, écologiques et sociaux, il est important de bien connaître les zones humides dans leur fonctionnement et leur localisation. Malheureusement, dans le cas du PNR Ardennes/Bassin de la Meuse la majorité du territoire est forestière ce qui rend plus compliquée la localisation des zones humides. En effet, à cause du couvert forestier, il est impossible de les situer par photo interprétation et il est nécessaire de réaliser des inventaires en plein si l'on veut les cartographier, ce qui devient rapidement très coûteux.

L'objectif de cette étude est donc de concevoir une méthode de prélocalisation des zones humides, en s'appuyant sur des critères pédologiques et dans la mesure du possible floristiques, en milieux ouverts et fermés. Ceci devrait permettre aux gestionnaires d'optimiser la prospection sur le terrain en connaissant mieux les zones à forts enjeux. Le modèle de prélocalisation qui sera développé dans un premier temps doit pouvoir être reproduit sur d'autres territoires. Il devra rendre compte des probabilités de présence des zones humides sur le territoire d'étude mais aussi des niveaux d'engorgement du sol. La cartographie des zones humides potentielles ainsi obtenue devra permettre d'alimenter les réflexions sur la gestion et l'exploitation de ces zones.

## III. Matériels et méthode

### 3.1 Données à disposition

Afin d'établir un modèle statistique de probabilité de présence des zones humides sur le périmètre d'étude, il est nécessaire d'avoir un certain nombre de points où les variables que l'on envisage de faire intervenir dans notre modèle sont connues. Pour cela, il est décidé d'utiliser les 282 relevés de l'Inventaire forestier national (IFN) et les 22 relevés EcoPlant (réalisés par le LERFoB)<sup>3</sup> effectués sur la zone en guise de jeu de calibration du modèle statistique de localisation des zones humides. Ces données sont exclusivement forestières. Il avait été envisagé de prendre également en compte les relevés de la base Sophy pour la calibration afin d'avoir des données en milieu ouvert. Cette base de données ne renseigne pas la pédologie et il n'était donc pas possible de classer les points en tant que zone humide ou non comme cela a été fait pour les points en milieu forestier.

Les variables explicatives choisies pour entrer dans le modèle sont issues de différents organismes. Les données climatiques (telles que le bilan hydrique, les précipitations et la réserve

---

<sup>3</sup> Relevés floristiques réalisés à l'échelle nationale

utile maximale) proviennent du LERFoB, comme les variables topographiques (MNT) ou mixtes (indice de Moore), tandis que la géologie est fournie par le BRGM.

L'un des objectifs de l'étude est de valider les prédictions du modèle par des données floristiques. Pour cela, les relevés des bases IFN, EcoPlant et Sophy présents sur le périmètre d'étude ont été retenus puisque la flore qui leur est associée est renseignée.

Enfin, dans le cadre de la validation, la cartographie des zones humides potentielles obtenue à partir du modèle sera comparée à celle obtenue par l'EPAMA qui nous a mis à disposition ses résultats. Une comparaison sera également faite avec les cartes d'Etat-major (cartographiant les zones « difficilement franchissables », souvent corrélées à la présence de zones humides) et celle de la DREAL qui cartographie les zones humides avérées déjà recensées.

## 3.2 Méthode de régression logistique

### ➤ Calibration du modèle

Le jeu de données formé fournit sur chaque relevé la variable à expliquer (présence ou absence de zone humide) et les différentes variables explicatives. L'utilisation d'un modèle statistique va nous permettre de comprendre lesquelles de ces variables environnementales influencent la localisation des zones humides de façon significative. Pour expliquer une variable binaire du type présence/absence, la méthode statistique adaptée est celle de la régression logistique.

À l'issue du travail sur logiciel statistique, une équation est obtenue donnant la probabilité de trouver une zone humide sur n'importe quel point de la carte en fonction des variables environnementales sélectionnées. L'équation est de forme binomiale :

$$Y = a + b * Variable\ 1 + c * Variable\ 2 + \dots$$

Où {a,b,c...} sont des coefficients de régression issus du logiciel.

Cette équation est transformée à l'aide de la formule suivante afin d'obtenir une probabilité comprise entre 0 et 1.

$$P = \frac{1}{1 + e^{(-Y)}}$$

Où P est donc la probabilité de trouver une zone humide en fonction des variables.

Cette équation sera utilisée dans un logiciel de cartographie (ici ArcGis) afin d'obtenir la carte de localisation potentielle des zones humides, la formule étant calculée sur chaque pixel de 50x50m.

Il est possible d'étudier ensuite les caractéristiques du modèle en calculant le poids des différentes variables dans le modèle et leur significativité, notamment grâce à des anovas (analyse de la variance).

➤ **Validation statistique du modèle**

La campagne de terrain va permettre de se déplacer sur des points judicieux de la carte et de déterminer si oui ou non ces points sont en zone humide. Ces données sont compilées dans une table pour comparer avec ce que notre modèle avait prédit comme probabilité.

La première étape de modélisation permet d’obtenir une carte avec des probabilités comprises entre 0 et 1 de trouver une zone humide, mais si on veut obtenir une carte binaire de présence/absence des zones humides, il faut fixer une valeur seuil de probabilité permettant de discriminer ce qui est une zone humide et ce qui n’en est pas une. Grâce à la phase de terrain il est possible de tester de nombreuses valeurs de seuil et d’estimer laquelle est optimale. Il est possible d’automatiser cette phase de travail avec un logiciel de programmation (ici R a été utilisé). Pour chaque valeur de seuil testée, on peut réaliser une matrice de confusion et calculer différents paramètres présentés ci-dessous.

Tableau 1: Matrice de confusion d'un modèle de régression logistique pour un seuil de probabilité donné

Observations vs Prédictions	0 (on prédit une zone non humide)	1 (on prédit une zone humide)
0 (on n’observe pas de zone humide sur le terrain)	<b>a</b> = nombre d’absences bien prédites	<b>b</b> = nombre de présences fausses prédites
1 (on observe une zone humide)	<b>c</b> = nombre de présences ratées par le modèle	<b>d</b> = nombre de présences bien prédites

*Spécificité (capacité du modèle à bien prédire une absence) :  $Sp = a / (a + c)$*

*Sensibilité (capacité du modèle à bien prédire une présence) :  $Ss = d / (b + d)$*

*Taux de succès (capacité du modèle à faire une bonne prédiction):  $S = (a + d) / (a + b + c + d)$*

Pour chaque valeur seuil, on obtient une matrice de confusion et on peut calculer un couple { $Ss$  ;  $Sp$ }. Cela permet d’obtenir la courbe  **$Ss = f(1 - Ss)$** , aussi appelée « courbe ROC ». L’aire sous cette courbe (appelé aussi AUC, *Area Under Curve*) est un indice de performance du modèle. Lorsque l’AUC est supérieure à 0.7, le modèle est bon.

On gardera comme seuil celui optimisant à la fois la spécificité et la sensibilité. Les pixels de la carte ayant une probabilité supérieure au seuil seront classés en zone humide potentielle et ceux dont la probabilité est inférieure en zone non humide.

### 3.3 Sélection des variables et construction du jeu de données

#### 3.3.1 Variable à expliquer

Pour pouvoir obtenir un modèle statistique de prélocalisation des zones humides, il est nécessaire de définir au préalable si les placettes retenues pour la calibration du modèle via la régression logistique sont ou non situées en zones humides. Pour cela, on se base sur des critères pédologiques. En effet, les zones humides peuvent être définies d'après des critères pédologiques et/ou floristiques. En raison de la quantité restreinte de données sur la flore, les données floristiques seront utilisées pour la validation du modèle et non la calibration.

En ce qui concerne les placettes IFN, tous les points dont le type de sol est noté « Tourbe », « Gley », « Stagnogley » ou « Pseudogley podzolique » sont considérés comme étant en zone humide, d'après la définition qu'en donne l'IFN (*Documentation relative aux données écologiques – Points forêt – Campagnes 2005 à 2011*, téléchargeable sur le site de l'IFN). Sur le même principe, les types de sol « Ranker à moder-mor », « Lithosol sur éboulis » ou « Sol brun calcaire », sont automatiquement classés en zone non humide. Pour ce qui est des placettes dont le type de sol est « Pseudogley », « Sol lessivé hydromorphe », « Sol brun », « Sol brun acide », « Sol brun hydromorphe » ou « Sol brun lessivé », elles sont classées en zone humide lorsque la profondeur d'apparition d'oxydation est inférieure ou égale à 25 cm (d'après la définition légale des zones humides<sup>4</sup>).

Pour ce qui est des placettes de la base EcoPlant, elles sont classées en zone humide à partir du moment où il y a observation de décoloration du sol et/ou trace de réduction du fer dans les 25 premiers cm.

Nous disposons donc d'une donnée binaire de présence/absence des zones humides sur 304 placettes. La surface couverte dépasse celle de la zone d'étude, mais pour calibrer un modèle il est important d'avoir le maximum de données possibles tout en respectant l'homogénéité du milieu (grands ensembles géologiques notamment ici). Cette donnée sera la variable à expliquer dans le modèle.

#### 3.3.2 Variables potentiellement explicatives

En termes de variables environnementales, nous disposons de différentes cartes du LERFoB sur la zone d'étude qui pourraient être intéressantes pour le modèle ; nous avons finalement sélectionné l'altitude, l'indice de Moore, la réserve utile maximale (RUM), les précipitations

---

<sup>4</sup> Voir définition dans l'introduction

annuelles et le bilan hydrique climatique annuel. Nous avons extrait les valeurs de ces différentes cartes pour les 304 placettes (fonction « Extract MultiValue to Point » d'ArcGis).

Tableau 2: Caractéristiques des données sur les variables écologiques provenant du LERFoB

Donnée	Définition	Commentaire	Résolution
<b>Hydrologie</b>			
<b>bhc61_13</b>	Bilan hydrique climatique annuel	P – ETP	France, 1 km France, 50m
<b>p61mr13_50</b>	Précipitations annuelles	rééchantillonnage bilinéaire de précipitations au pas du km	France, 50m
<b>Ru_am</b>	Réserve utile maximale	Calcul d'Al Majou	France, 1km
<b>moore_2</b>	Indice de Moore	= (flow accumulation * surfacepixel) / tan(pente)	France, 50m France, 50m
<b>Topographie</b>			
<b>MNT</b>	<b>Altitude</b>	<b>Raster</b>	<b>France, 50m</b>

Une variable explicative prenant en compte la perméabilité pourrait également s'avérer pertinente. Elle est construite en compilant la géologie précisée sur les cartes au 1/50 000ème et une donnée du BRGM, l'indice de développement et persistance des eaux (IDPR). Cet indice rend compte de la capacité du sol à laisser s'infiltrer ou ruisseler les eaux de surface. Il découle de données du MNT et résulte du rapport entre la distance au plus proche cours d'eau réel et la distance au plus proche cours d'eau calculé, pondéré par la pente (voir *Carte de vulnérabilité intrinsèque simplifiée des eaux souterraines du bassin Rhin Meuse et de la région Lorraine*, BRGM/RP-56539-FR, 2010). Ne disposant pas de la donnée numérisée, nous avons utilisé la donnée publique et l'avons numérisé e nous-mêmes en simplifiant les classes. Quatre classes ont été établies, de 1 (ruissellement des eaux de surface) à 4 (infiltration des eaux). Concernant la géologie, les différents substrats indiqués sur les cartes au 1/50 000ème ont été classés en trois groupes : perméable, imperméable, non renseigné.

Tableau 3 : Attribution des classes pour l'indice compilant la géologie et l'IDPR (PP : perméable, P : plutôt perméable, I : plutôt imperméable, II : imperméable)

Géologie \ IDPR	IDPR		
	1 et 2	3 et 4	Non renseigné
Perméable	P	PP	PP
Imperméable	II	I	II
Non renseigné	I	P	N/A (8 placettes)

### 3.3.3 Sélection des variables explicatives par le modèle

Pour la régression logistique, le logiciel de programmation R a été utilisé en travaillant sur une table reprenant pour chaque placette la variable à expliquer (présence/absence de zone humide) et les différentes variables explicatives.

Un modèle statistique performant doit expliquer au mieux la variable considérée tout en étant parcimonieux ; il y a donc un travail de sélection des variables à réaliser afin de voir lesquelles expliquent le mieux la répartition des zones humides et lesquelles risqueraient d'alourdir inutilement le modèle. Il existe dans R des fonctions pour automatiser la sélection des variables et on obtient dans ce cas un modèle certes performant sur le plan statistique, mais qui n'a pas forcément de sens écologique. Nous allons donc sélectionner les variables une à une pour s'assurer que leur intervention dans le modèle est pertinente.

Dans la sélection de variables, différents paramètres interviennent. Le « d2 » (déviance du modèle testé par rapport au modèle saturé, c'est-à-dire le modèle dans lequel on a tous les paramètres) est un bon indicateur. Il donne une idée de l'apport relatif d'une variable au modèle total et on prendra en priorité la variable avec le « d2 » le plus élevé.

Dans un second temps, nous avons considéré les corrélations, c'est-à-dire que nous avons vérifié avant d'ajouter des variables qu'elles n'étaient pas « redondantes » avec une variable déjà sortie. Par exemple, si le bilan hydrique apparaît mais que dans les étapes précédentes on a déjà ajouté les précipitations et la RUM, on va calculer les corrélations entre les variables : si celles-ci sont trop élevées, on n'ajoutera pas le bilan hydrique.

Enfin, nous affichons les courbes de réponse du modèle à la variable avant de l'ajouter. Ces courbes présentent la probabilité de trouver une zone humide en fonction de la variable considérée. C'est notamment lors de cette étape que nous avons constaté que l'IDPR ne pouvait pas entrer dans notre modèle car les courbes de réponses étaient incohérentes : on trouvait que la probabilité de trouver une zone humide augmentait avec l'infiltration de l'eau dans le sol. Il est donc capital d'afficher ces courbes.

Par ailleurs, lorsqu'une variable est sélectionnée, nous étudions la possibilité de prendre son carré en compte dans le modèle (on compare alors les « d2 » et les courbes de réponse entre le modèle simple  $v_1$  et le modèle quadratique  $v_1+v_1^2$ ). Voici en exemple la première étape pour le modèle que nous avons réalisé.

Tableau 4 : Tableau de la déviance pour la première étape de construction du modèle

var.eco	dev0	dev	dev.diff	d2
RUM	278.214	249.0039	29.210046	0.10499129
Indice de Moore	278.214	252.0586	26.155439	0.09401195
Indice de perméabilité	278.214	261.3212	16.892775	0.06071864
Précipitations	278.214	269.0487	9.165337	0.03294348
Altitude	278.214	269.3042	8.909834	0.03202511
BHC	278.214	269.3168	8.897219	0.03197977

D'après le Tableau 4, on sélectionnerait en première variable du modèle la RUM (« d2 » le plus élevé). On compare à présent les « d2 » des modèles simple et quadratique (Figure 4) et on affiche les courbes de réponse (Figure 3).

Les courbes de réponses sont cohérentes : la probabilité de trouver une zone humide augmente lorsque la RUM augmente. Cependant dans le cas de la variable au carré, un seuil est atteint pour des valeurs autour de 130mm, ce qui n'a pas de signification écologique a priori. C'est donc la RUM sous sa forme simple qui est conservée en tant que première variable explicative.

En fin de première étape, on a donc un modèle de type :  $Y = a + b \cdot RUM$  (les coefficients sont donnés par R).

```
> d2simple
[1]0.09770811
> d2carre
[1]0.1088186
```

Figure 4 : Calcul sur R des « d2 » pour les modèles simple et quadratique

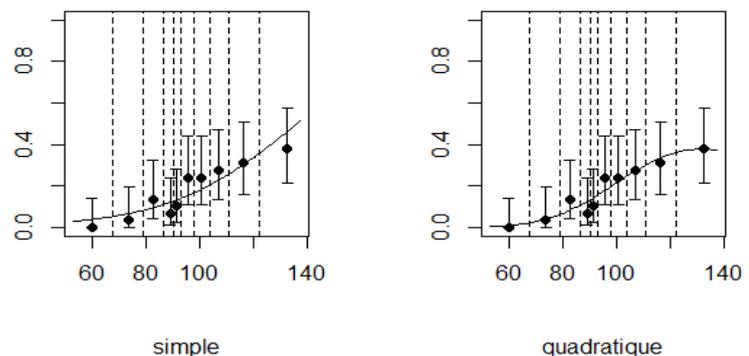


Figure 3 : Courbe de réponse du modèle à la RUM (probabilité de trouver une zone humide en fonction de la RUM, modèle simple à gauche et modèle quadratique à droite)

On progresse ainsi en ajoutant les variables une à une. On s'arrête lorsque les variables sélectionnées n'apportent plus assez d'information au modèle et ne feraient que l'alourdir : si le « d2 » n'augmente plus beaucoup notamment. La réalisation d'une anova sur le modèle est également intéressante car elle donne la significativité des différentes variables. L'équation finale obtenue ainsi que le calcul des probabilités de présence sont présentés dans la partie « Résultats » de ce rapport.

## 3.4 Méthodologie dans l'utilisation des données flore et stations

### 3.4.1 Objectifs

L'existence de données concernant la flore et les stations forestières de la zone d'étude constitue un atout pour la validation du modèle et des prédictions obtenues. Les données disponibles sont les suivantes :

- relevés flore de l'IFN, de la base EcoPlant et de la base Sophy
- cartographie des stations par l'ONF

Un catalogue des stations forestières de l'Ardenne primaire (Drapier, 1989), détaillant la typologie des stations, est également disponible pour compléter l'analyse faite sur la flore. Il nous a permis de déterminer le niveau d'hydromorphie des relevés.

En raison des restrictions observées sur l'utilisation de la méthode de bio-indication du niveau d'engorgement (voir partie « Discussion » de ce rapport), il a été décidé d'utiliser ces données dans le cadre de la **validation** du modèle plutôt que de les intégrer à celui-ci en tant que variables potentiellement explicatives de la présence des zones humides.

L'objectif est donc d'utiliser les informations que ces données peuvent fournir sur la répartition avérée de zones humides sur le territoire d'étude, afin de confirmer ou d'infirmer les prédictions réalisées par le modèle statistique. Dans un second temps, si les données sont suffisantes, ce travail peut également permettre de caractériser le type de zones humides prédites par un niveau d'engorgement.

### 3.4.2 Méthode

#### a. Flore

L'apport des données de la flore repose sur le principe de la bio-indication. Ce principe suppose que la présence ou l'état d'espèces (ici végétales) en des points précis du territoire renseigne sur la valeur de certaines variables écologiques en ces points. Par exemple, la présence d'espèces dites acidiphiles comme la Myrtille (*Vaccinium myrtillus*) indiquent un pH faible pour les sols sur lesquels elles se trouvent.

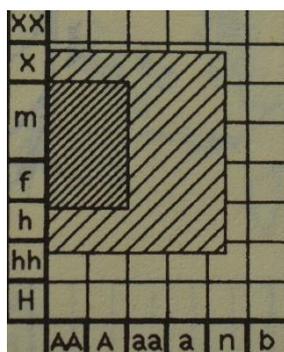


Figure 5 : Schéma décrivant les conditions optimales pour le développement de l'espèce (acidité horizontalement, humidité verticalement), ici *Vaccinium myrtillus*

Les variables bio-indiquées qui présentent un intérêt dans cette étude sont les variables d'engorgement (permanent et/ou temporaire) des sols. La méthode utilisée est la méthode de valeurs indicatrices établie par Jean-Claude Gégout (2006). Cette méthode permet de calculer, à partir des valeurs indicatrices propres à chaque espèce végétale, la valeur de différentes variables écologiques sur des placettes de relevés forestiers.

Les valeurs indicatrices d'engorgement propres à chaque espèce ont été calculées à partir des données de 5400 relevés de la base EcoPlant, qui comportent des descriptions des sols, ainsi que grâce à la formule de fidélité à une classe (Chytry, 2002), qui attribue à chaque espèce un indice selon sa fidélité à une classe de valeurs bio-indiquées.

L'indice d'engorgement permanent (respectivement temporaire) a été calculé à partir de la fréquence de chaque espèce sur les sols à horizon réductique (respectivement rédoxique) à une profondeur inférieure à 40 cm d'une part, et dans les autres horizons d'autre part. La méthode de Jean-Claude Gégout indique également des seuils permettant de classer les valeurs d'engorgement obtenues, selon le classement suivant (Tableau 5) :

*Tableau 5 : Classes d'humidité correspondant aux valeurs d'engorgement bio-indiquées*

Classe	Valeurs d'engorgement permanent bio-indiquées	Valeurs d'engorgement temporaire bio-indiquées	Niveau d'humidité
1	0.1 à 0.2	0.10 à 0.13	Très humide
2	0.05 à 0.099	0.05 à 0.099	Humide
3	0 à 0.049	0 à 0.049	Frais
4	-0.049 à -0.0001	-0.049 à -0.0001	Moyen
5	-0.11 à -0.05	-0.07 à -0.05	Sec

Pour chacune des placettes de relevés disponibles sur le territoire d'étude, les valeurs bio-indiquées pour les engorgements temporaire et permanent sont donc calculées.

Pour chaque placette, l'écart-type des valeurs d'engorgement bio-indiquées est également calculé afin d'attribuer un poids différent aux relevés selon leur homogénéité. Ainsi, les relevés floristiques les plus homogènes, donc les plus fiables, se voient attribuer un poids plus important dans la représentation cartographique (Annexe 2).

#### ➤ **Evaluation de la bio-indication**

Avant d'interpréter les résultats obtenus par l'intermédiaire de cet outil, il convient d'en estimer la pertinence sur le secteur d'étude. En effet, la méthode de bio-indication utilisée a été calibrée avec des données nationales, et pourrait donc être inadaptée pour une zone plus restreinte.

Au niveau des placettes de calibration du modèle (IFN et EcoPlant), une comparaison entre le niveau d'hydromorphie réel et le niveau d'engorgement bio-indiqué par la flore est donc réalisée. Le caractère temporaire ou permanent du niveau d'hydromorphie traduisant des réalités de

fonctionnement hydrique bien différentes, il convient de les considérer séparément. À cet effet, les descriptions du sol sont utilisées pour classer les placettes en trois catégories (Tableau 6).

Tableau 6 : Classement des placettes en fonction des types de sols

Drainé	Hydromorphie temporaire	Hydromorphie permanente
Sols bruns	Pseudogley podzolique	Tourbe
Sol lessivé hydromorphe	Pseudogley (prof. hyd. < 25 cm)	Gley
Lithosol sur éboulis		Stagnogley
Ranker à moder-mor		
Pseudogley (prof. hyd. > 25 cm)		

Ces niveaux d'hydromorphie réels sont ensuite comparés aux valeurs d'engorgement bio-indiquées par la flore pour chaque placette (Figure 6).

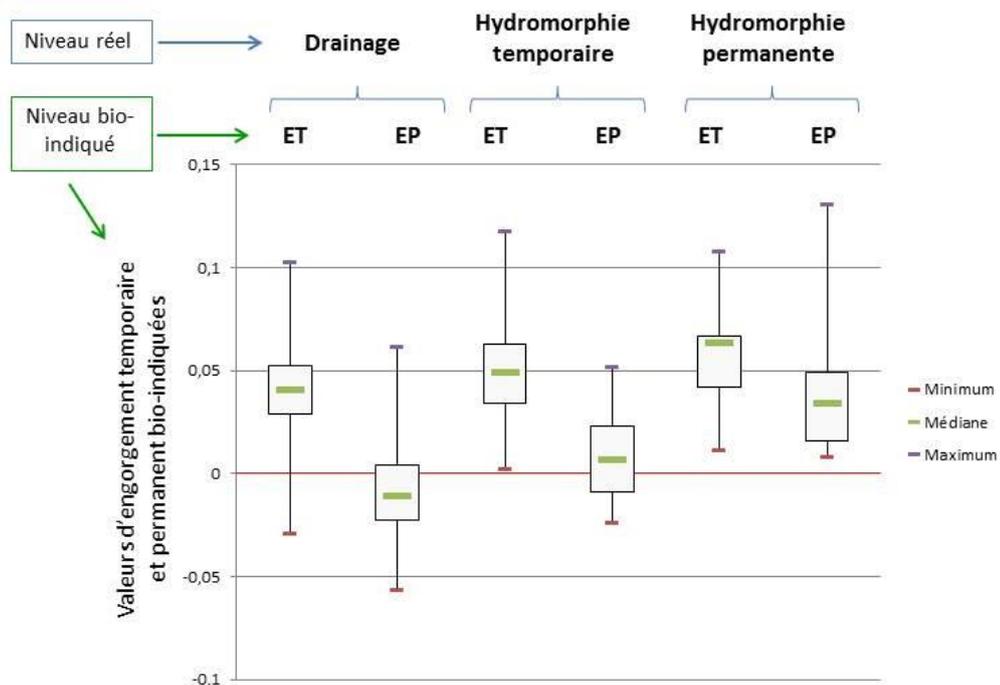


Figure 6 : Comparaison du niveau d'engorgement bio-indiqué par la flore

Sur les sols drainés, la bio-indication donne des valeurs d'engorgement temporaire qui paraissent un peu trop élevées par rapport à la réalité. Cette tendance pourrait être expliquée par la méthode de classement des niveaux d'hydromorphie réels, avec une profondeur seuil de 25 cm en accord avec la définition légale, tandis que la méthode de bio-indication se base sur une profondeur seuil de 40 cm. Cela conduit à une tendance à classer dans la catégorie « sol drainé » des sols que la bio-indication classerait en hydromorphie temporaire, ce qui peut expliquer la valeur étonnamment

haute du niveau d'engorgement temporaire bio-indiqué. En revanche, les valeurs négatives d'engorgement permanent concordent avec le drainage.

Sur les sols à hydromorphie temporaire, les valeurs bio-indiquées correspondent à ce qui serait attendu : un niveau d'engorgement temporaire positif, et un niveau d'engorgement permanent nul.

En revanche, sur les sols à hydromorphie permanente, on constate une incohérence pour l'engorgement temporaire bio-indiqué. En effet, on s'attendrait à un niveau d'engorgement temporaire nul, et la valeur positive observée pourrait être expliquée par le faible nombre de

On remarque donc que la bio-indication dans cette application paraît peu performante. Les résultats de la comparaison des valeurs bio-indiquées avec les prédictions du modèle de localisation des zones humides sont donc à considérer avec précaution.

placettes correspondant à un engorgement permanent réel.

## **b. Stations**

### ➤ **Etape 1 : Niveau d'humidité attribué par l'IFN**

Grâce au catalogue des stations de l'Ardenne primaire (Drapier, 1989), un niveau d'engorgement est attribué à chaque unité stationnelle. Disposant de la correspondance entre les codes de stations de l'ONF et les unités stationnelles du catalogue de l'IFN, on peut alors représenter la répartition des niveaux d'humidité de l'IFN sur la cartographie des stations (Annexe 3, carte 1).

Tableau 7 : Correspondance entre les codes de stations et les niveaux d'humidité

Code IFN	Code ONF	Niveau d'humidité attribué par le catalogue de stations
US0	C2G	Très humide
US31	C2K	
US03	C2I	Humide
US04	C2I	
US06	C2D	
US07	C2D	
US08	C2D	
US25	C2E	Frais
US22	C2J	
US33	C2I	Moyen
US13	C2H	
US14	C2E	
US13	C2H	
US24	C2E	
US16	C2C	
US26	C2C	
US17	C2A	
US27	C2A	
US18	C2B	
US28	C2B	
US29	C2F	

➤ **Etape 2 : Bio-indication**

Dans le catalogue de stations, une liste d'espèces diagnostiques de chaque unité stationnelle est également disponible et permet donc d'appliquer la méthode précédente des valeurs indicatrices à l'échelle des stations. Ainsi, pour chaque unité stationnelle définie par le catalogue, une valeur bio-indiquée d'engorgement lui est associée.

### ➤ **Etape 3 : Comparaison**

La comparaison de la cartographie des niveaux d'engorgement des stations établie par les deux méthodes précédentes a montré l'existence de plusieurs incohérences (Annexe 3). En particulier, on constate que le niveau d'engorgement bio-indiqué par la flore apparaît souvent comme moins élevé que le niveau attribué par l'IFN, à l'exception des zones très humides.

Dès lors, en raison des réserves émises sur la pertinence de la méthode de bio-indication à l'échelle de la station (faible nombre d'espèces diagnostiques fournies par le catalogue des stations, incertitudes sur certaines espèces dont seul le genre en français était disponible), cet outil a été abandonné au profit de l'attribution de niveaux d'engorgement (classes d'humidité) réalisé par l'IFN. C'est donc la cartographie issue de cette méthode qui a été retenue pour la comparaison future avec les prédictions du modèle (Annexe 3, carte 1).

#### **c. Flore de zone humide**

L'arrêté du 24 juin 2008 précise une liste d'espèces hygrophiles caractéristiques des milieux humides. Les espèces relevées sur les placettes IFN, EcoPlant et Sophy et appartenant à cette liste d'espèces hygrophiles sont donc notées. Ce travail permet d'obtenir une carte de répartition du nombre d'espèces hygrophiles par placette (Annexe 4), qu'il sera intéressant de comparer à la cartographie prédictive du modèle par la suite.

### **3.5 Méthodologie de validation par une phase de terrain**

L'étude comporte une phase de terrain afin de valider le modèle de localisation des zones humides potentielles. Les points de relevés de terrain sont choisis de façon à ce qu'ils soient représentatifs de notre zone d'étude et de ses variations topographiques. Des transects (en moyenne de 2 km) sont placés sur les gradients des zones potentiellement humides à moins humides prédites par le modèle construit. Nous avons également comparé la carte établie par notre modèle à la carte des stations forestières, et choisi des points de relevés de terrain dans les secteurs où il y a incohérence entre le niveau d'humidité indiqué par ces deux cartes.

Notre modélisation reposant sur des données de flore obtenues en milieu fermé (relevés de l'IFN et relevés EcoPlant réalisés par le LERFoB), il est intéressant de vérifier la validité de notre modèle en milieu ouvert également. Les transects sont donc placés en milieu ouvert et en milieu fermé.

Au total 73 relevés de terrain sont effectués, environ 5 points par transect (cf Carte 3 des relevés de terrain effectués). Les données stationnelles et pédologiques sont relevées (fiche terrain en Annexe 1). La reconnaissance des espèces de la flore n'est pas réalisable car la phase de terrain a lieu en période hivernale.

### ➤ Protocole de terrain

Des trous de tarière de 50 cm de profondeur au moins sont réalisés afin de vérifier la présence de traces d'hydromorphie. Les points de relevés de terrain sont alors répartis en 6 classes de niveau d'engorgement, selon des critères pédologiques tels l'épaisseur de tourbe ou la présence de décoloration ou d'oxydation (voir clé de classes d'engorgement en Annexe 7). Les classes 0 et 1 sont celles avec peu ou pas d'engorgement, les points appartenant à cette classe sont considérés comme n'appartenant pas à des zones humides

## IV. Résultats

### 4.1 Cartographie des zones humides potentielles

Afin d'obtenir une cartographie de la probabilité de présence de zone humide sur le territoire de l'étude, une régression logistique a été réalisée à partir des données présentées précédemment. Celle-ci aboutit à la définition d'une équation qui traduit la probabilité de présence :

$$P_{\text{présence}} = \frac{1}{1 + e^{-Y}}$$

Où  $Y = -6,809401 + 0,025186 * RUM + 0,625293 * Moore - 0,003224 * prec + 0,044453 * BHC$

Les variables sélectionnées et leur apport relatif au modèle sont présentés dans le tableau ci-dessous :

Tableau 8 : Apport relatif des variables au modèle

Variables intervenant dans le modèle					
	Réserve utile maximale	Précipitations moyennes annuelles	Bilan hydrique	Indice de Moore	Indice géologie-perméabilité
Apport relatif au modèle	14.0%	0.3%	0.4%	21.8%	31.2%

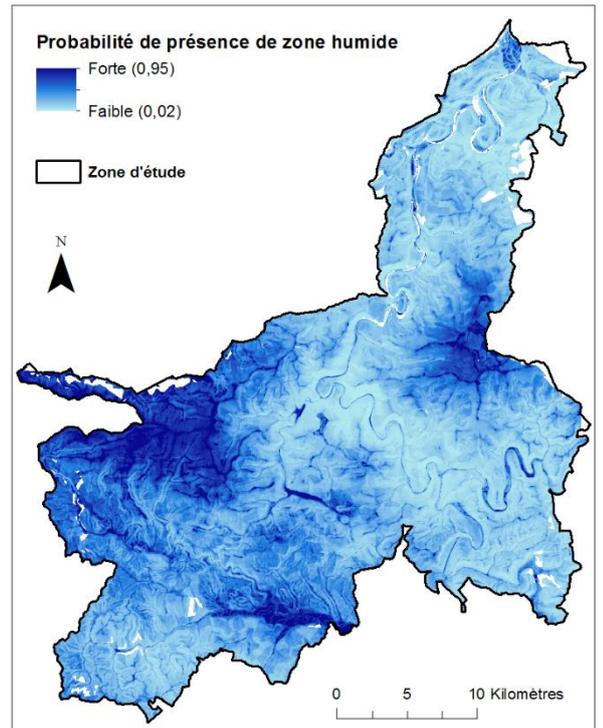
L'équation est valable pour les zones définies comme « plutôt imperméables » (code I) par l'indice de perméabilité croisant la nature géologique et l'IDPR. Pour les autres zones, il est nécessaire d'ajouter les valeurs suivantes qui tiennent alors compte de la perméabilité :

Tableau 9 : Coefficients à intégrer à l'équation de probabilité de présence de zone humide selon l'indice de perméabilité

II (imperméable)	- 0,647997
P (plutôt perméable)	1,263839
PP (perméable)	- 0,075712

Le calcul de cette probabilité est alors réalisé sous SIG pour chaque pixel de 50x50m et la carte suivante est obtenue. Elle représente les zones humides potentielles sur le territoire d'étude. Cependant, la valeur de chaque pixel prise indépendamment des autres pixels ne peut être analysée. Seules les tendances communes à plusieurs pixels sont pertinentes.

Les valeurs absolues de probabilité ne sont pas à analyser en tant que telles car elles sont directement liées au modèle et à ses incertitudes. Néanmoins, la comparaison relative de ces valeurs permet de distinguer des espaces où la probabilité de trouver des zones humides est plus forte qu'ailleurs.



Carte 1 : Probabilité de présence de zone humide sur la zone d'étude

#### 4.1.1 Validation du modèle par comparaison avec la présence réelle de zone humide sur les placettes de calibration

Avant toute autre validation du modèle, on peut déjà effectuer une première vérification des résultats obtenus en comparant les probabilités prédites au niveau des placettes de la base de données avec la situation réelle sur ces points.

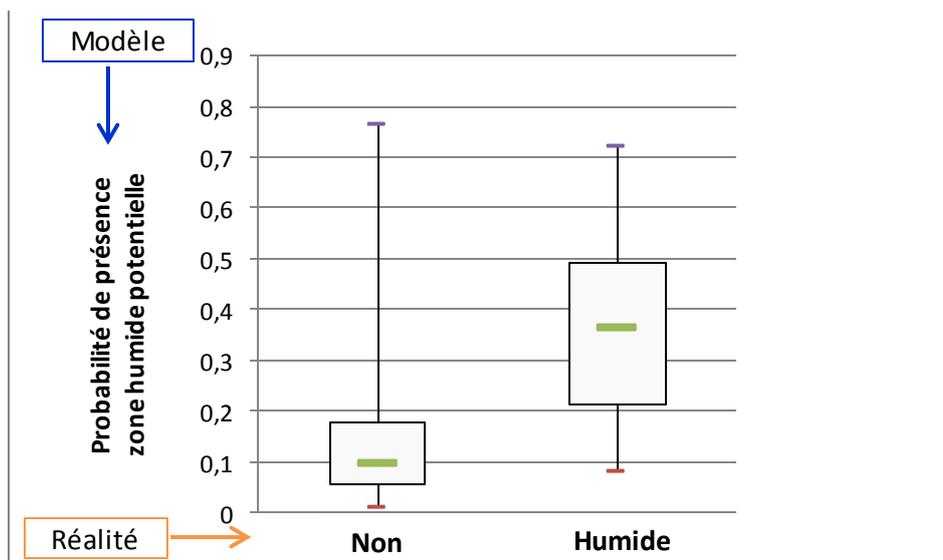


Figure 7 : Comparaison entre la probabilité de présence de zone humide obtenue dans le modèle

Cette représentation indique que le modèle paraît correspondre à la situation réelle observée sur les placettes de la base de données. En effet, pour les placettes effectivement drainées, on observe des probabilités de présence de zone humide globalement faibles (autour de 0,1), à l'exception de quelques valeurs absurdes. Pour les zones effectivement humides, les probabilités sont globalement plus élevées.

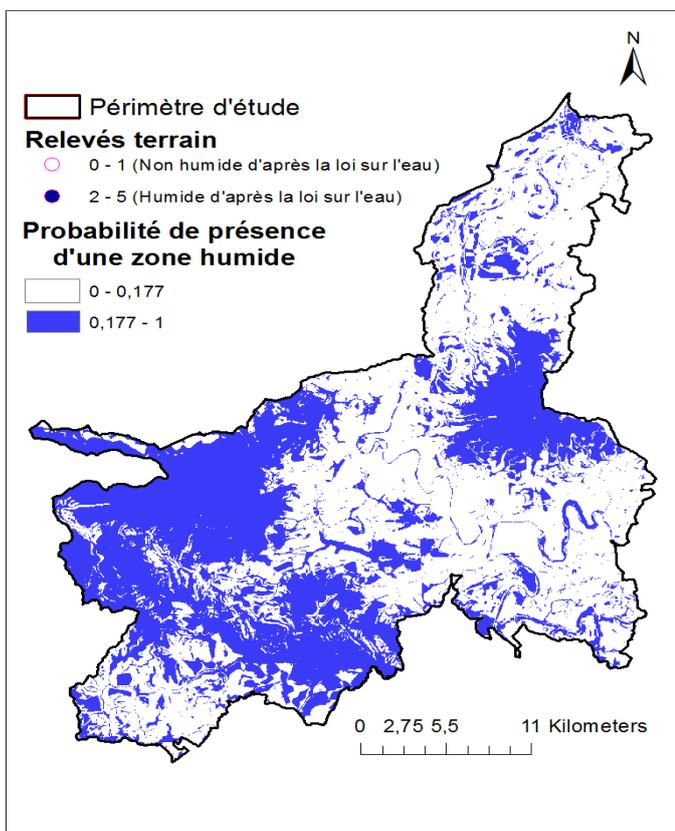
Cette comparaison constitue un premier indicateur de qualité du modèle.

## 4.2 Résultats – Phase de validation (terrain)

### 4.2.1 Carte des probabilités de présence de zone humide

Statistiquement, il est possible de déterminer le seuil de probabilité (seuil ROC) qui sépare une zone probablement humide d'une zone probablement non humide. Ce seuil a été déterminé à partir de nos placettes terrain : il est de 0,177. Les zones de notre carte avec une probabilité supérieure à ce seuil sont donc considérées comme étant en zone humide. On obtient alors la Carte 2.

En parallèle de ce seuil, les performances du modèle sont calculées : l'AUC est de 0,74, ce qui indique une performance moyenne du modèle. La sensibilité et la spécificité confirment cela, avec respectivement 72% et 78%.



D'après cette carte, notre modèle prédit 44% de la surface de notre zone d'étude comme étant potentiellement humide. Ce résultat est intéressant dans la mesure où il permet de réduire la surface à parcourir dans l'hypothèse d'une campagne de terrain pour délimiter les zones humides. Cette campagne serait donc moins coûteuse, aussi bien en temps qu'en argent, qu'une campagne systématique sur l'ensemble de la zone.

*Carte 2 : Carte des probabilités de présence d'une zone humide sur la zone d'étude en tenant compte du seuil statistique*

## 4.2.2 Cartographie des placettes terrain selon leur classe de zone humide

Les placettes ont été représentées sur la carte issue de notre modèle selon leur classe de zone humide (Carte 3). Dans l'ensemble, la cohérence entre les deux est plutôt bonne, mais il existe aussi plusieurs situations où elle est moins exacte.

Par exemple, on peut observer une bonne corrélation entre les observations terrain et notre prédiction au nord de Rocroi, et une zone plus problématique vers Monthermé où les observations de terrain ne correspondent pas à notre modèle.

Au final, le modèle a 75% de chance de bien prédire l'état d'un pixel (zone humide ou non humide).

Afin d'essayer de préciser notre modèle vis-à-vis de l'engorgement notamment, nous avons repris les classes de nos placettes terrain pour les étudier plus précisément.

Pour chaque classe, nous avons fait la moyenne des probabilités de présence prédites par notre modèle pour chaque placette terrain. La Figure 8 montre ces résultats.

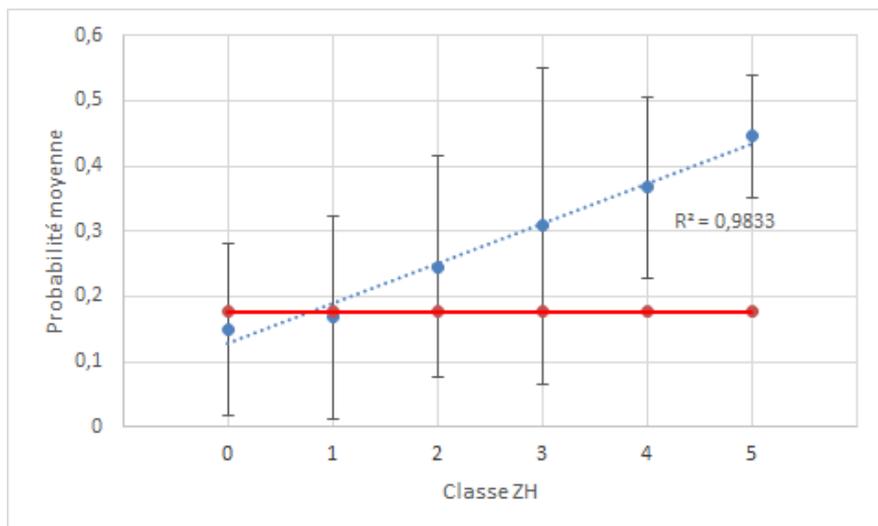
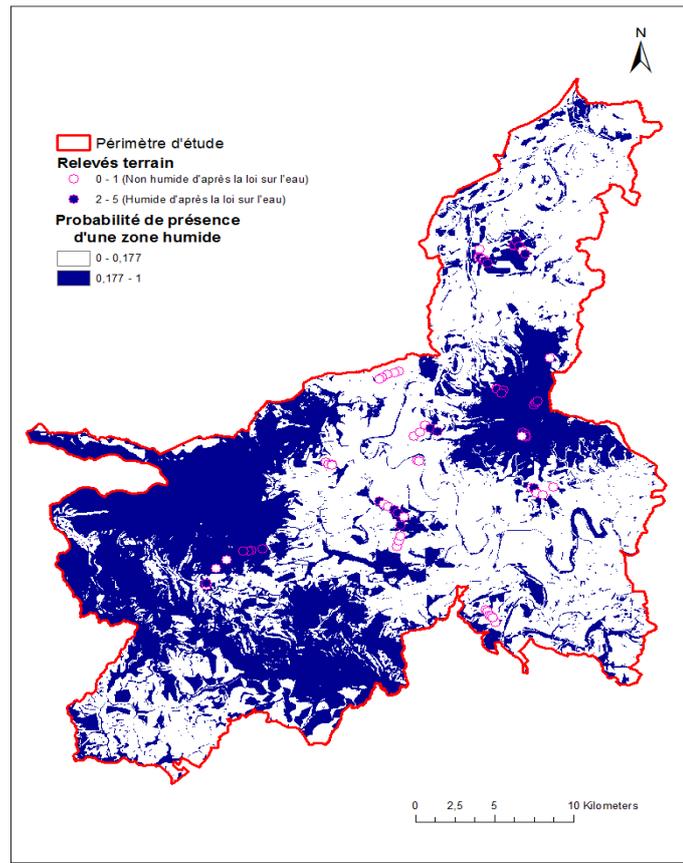


Figure 8 : Graphique des moyennes des probabilités de présence par classe d'engorgement

Les classes 0 et 1 se retrouvent bien comme n'étant pas humide, contrairement aux classes 2 à 5 qui le sont d'après la loi sur l'eau.

Des seuils de probabilité ont ensuite été déterminés entre les différentes classes ; pour cela, la moyenne des valeurs obtenues pour chaque classe a été faite. Par exemple, pour le seuil entre les classes 0 et 1 (voir la Figure 9), le calcul suivant a été réalisé :  $\frac{0,150+0,168}{2} = 0,16$  qui correspond à la moyenne des probabilités moyennes des classes 0 et 1.

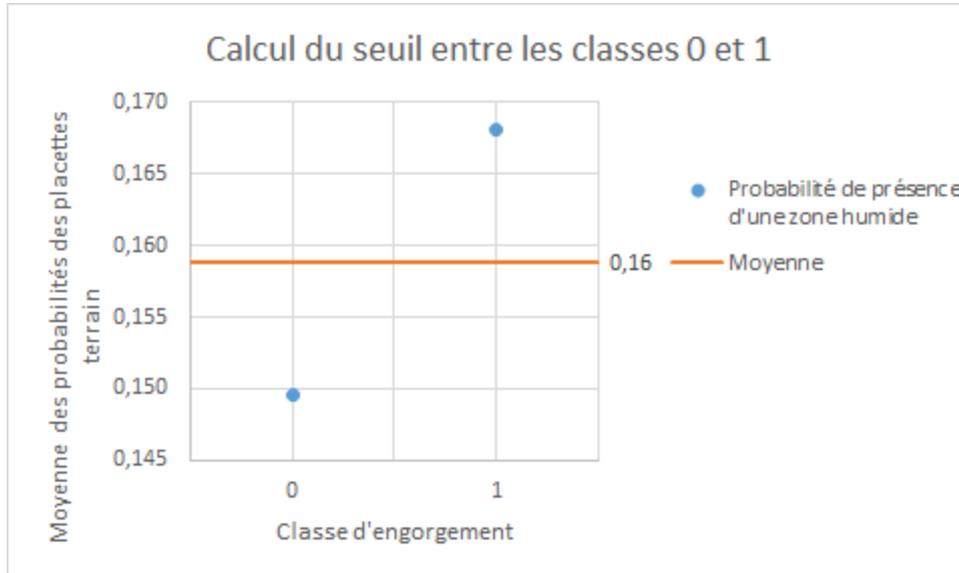


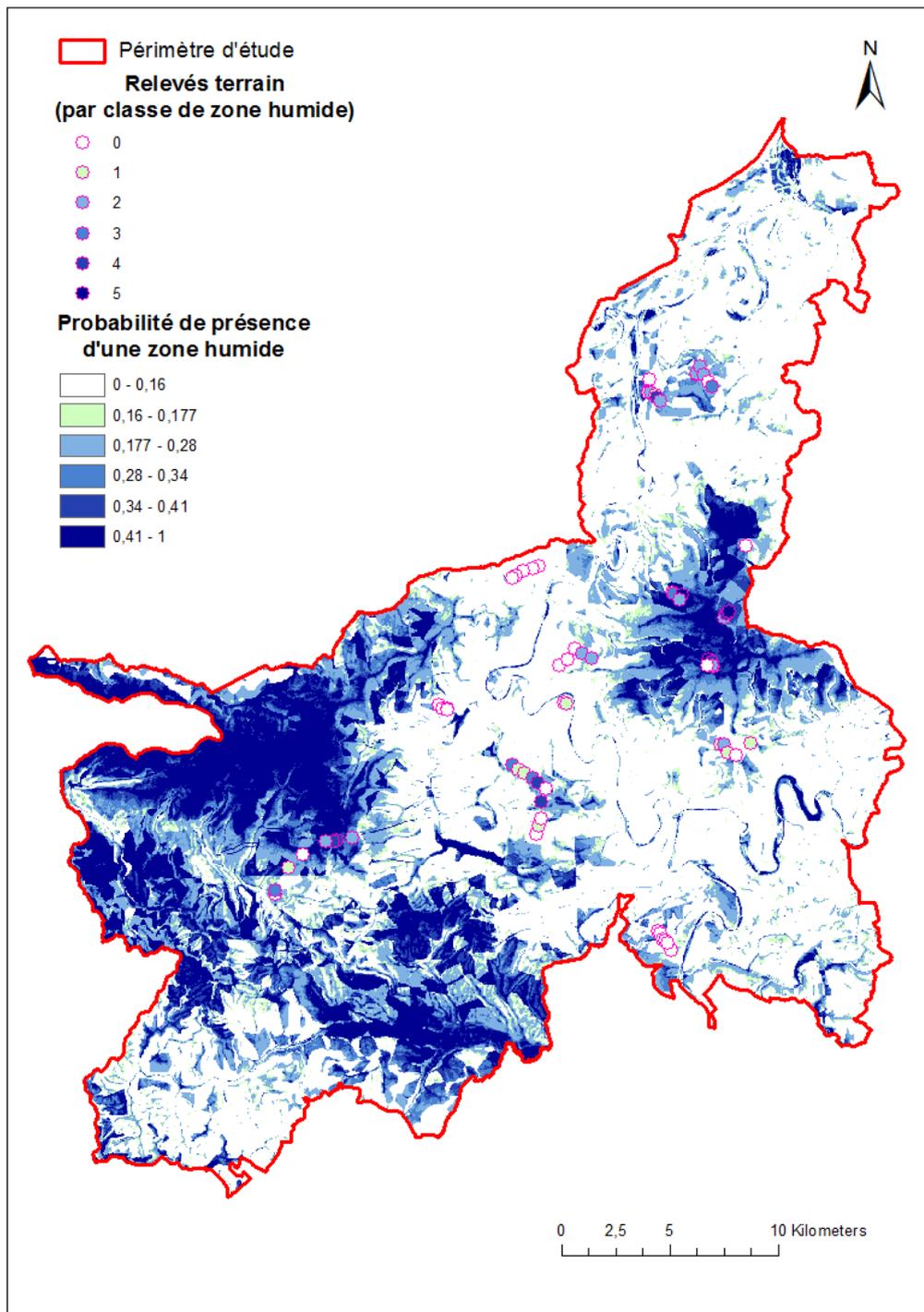
Figure 9 : Graphique présentant le calcul du seuil entre les classes 0 et 1

Pour le seuil entre les classes 1 et 2, le seuil ROC défini précédemment a été repris. Puis les autres seuils ont été calculés comme le premier. Les résultats sont présentés dans le Tableau 10.

Tableau 10 : Les seuils des classes d'engorgement

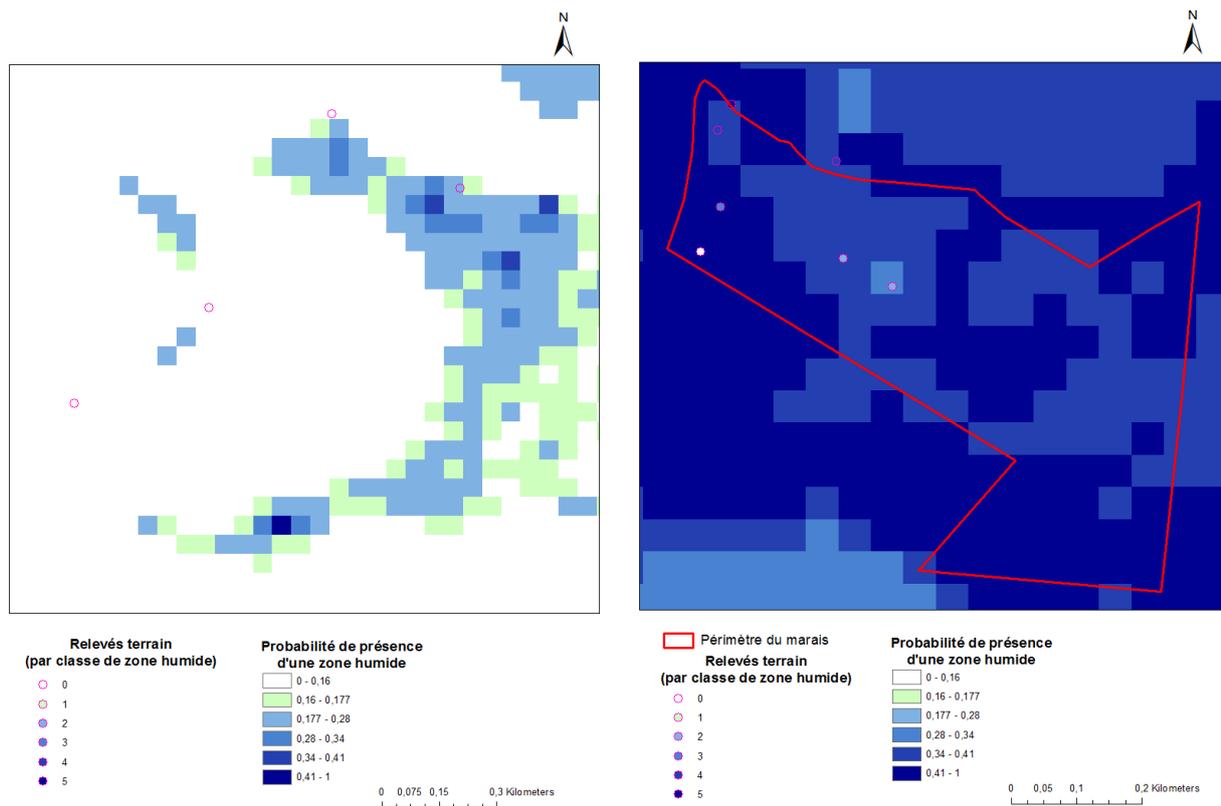
	Classes 0 à 1	Classes 1 à 2	Classes 2 à 3	Classes 3 à 4	Classes 4 à 5
<b>Valeur seuil</b>	0,16	0,177 (seuil ROC)	0,28	0,34	0,41

La carte issue de notre modèle a donc été modifiée en insérant ces différents seuils (voir Carte 4). Les placettes terrain ont également été positionnées sur cette nouvelle carte, tout en gardant une hiérarchisation selon leur classe déterminée sur le terrain.



*Carte 4 : Carte des probabilités de présence d'une zone humide en fonction des classes d'engorgement, et avec les placettes terrain*

De nouveau, on retrouve une assez bonne cohérence globale, malgré quelques points où le modèle est peu efficace. On peut voir sur la Carte 5 un exemple où la cohérence est bonne (carte de gauche, il est même possible d'observer des Hauts-Buttés ne correspondent pas à la prédiction de notre modèle).



Carte 5 : Zoom de la carte des probabilités de présence en fonction des classes d'engorgement avec les placettes  
Le Tableau 11 donne le pourcentage de surface de notre zone d'étude dans chaque classe.

Tableau 11 : Pourcentage de surface prédite pour chaque classe d'engorgement

Classe de zone humide	0	1	2	3	4	5
Surface de notre zone d'étude	51%	5%	15%	6%	6%	17%

On retrouve les 44% de surface classée en zone humide par notre modèle à travers les classes 2 à 5, qui sont surtout classés dans les deux extrêmes, c'est-à-dire les sols avec plus de 30cm de tourbe (classe 5) ou ceux avec des taches d'oxydation ou de réduction dans les 25 premiers centimètres (classe 2).

#### 4.2.3 Carte des probabilités de présence d'une zone humide et leurs liens avec l'engorgement

D'après la Figure 8, un lien linéaire peut être établi entre les moyennes des classes : la régression linéaire donne un  $R^2$  de 0,9833, ce qui est très significatif. L'hypothèse d'une corrélation linéaire entre la probabilité de présence d'une zone humide en un point et le degré d'engorgement de la zone humide potentielle sur ce même point peut donc être faite. En clair, plus une zone a une forte probabilité d'être une zone humide, plus son engorgement serait important si une zone humide est effectivement présente. Ce résultat n'est pas surprenant dans la mesure où la quasi-totalité des

variables insérées dans le modèle témoignent d'une quantité d'eau en un point, comme les précipitations, le bilan hydrique, l'indice de Moore...

La Carte 4 peut donc être vue comme une carte décrivant l'engorgement de la zone d'étude.

Le modèle prédit 44% de la surface de notre zone d'étude comme étant potentiellement humide. Dans le cadre d'une prospection de terrain pour localiser les zones humides, cela permettrait de réduire la surface à parcourir en comparaison avec une campagne systématique sur l'ensemble de la zone.

Le modèle a 75% de chance de bien prédire la présence d'une zone humide.

Une corrélation linéaire a été montrée entre les classes d'engorgement et la probabilité de présence de zones humides : le modèle permet donc d'appréhender l'engorgement d'une zone.

### 4.3 Comparaison avec les données existantes

#### 4.3.1 Comparaison avec les données flore et stations

La carte de probabilité de présence de zones humides potentielles peut être comparée aux cartes de niveaux d'engorgement bio-indiqués (placettes) ou de niveau d'humidité par station.

##### a. Flore

La carte globale de comparaison du modèle avec les niveaux d'engorgement bio-indiqués est présentée en Annexe 5. Dans un souci de simplification, les engorgements temporaire et permanent sont sommés dans cette représentation, comme l'autorise la méthode de calcul de valeurs indicatrices employée. Dans les zones où le modèle prédit une forte probabilité de présence de zone humide, la bio-indication par la flore semble cohérente. Cependant, dans les zones intermédiaires ou de faible probabilité, les valeurs bio-indiquées sont plus disparates et sans logique apparente. La Figure 10 ci-dessous permet de visualiser des points de concordance et les contradictions.

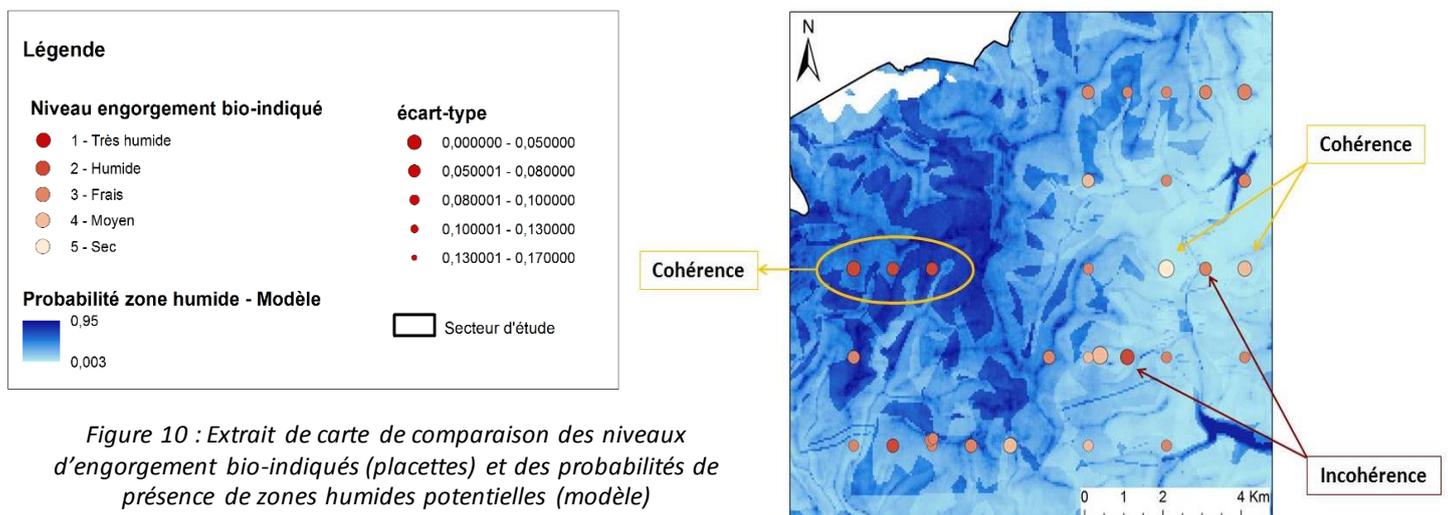


Figure 10 : Extrait de carte de comparaison des niveaux d'engorgement bio-indiqués (placettes) et des probabilités de présence de zones humides potentielles (modèle)

## b. Stations

La carte globale de superposition entre les probabilités de présence obtenues par le modèle et les niveaux d'humidité par station est présentée en Annexe 6. On constate une cohérence d'ensemble entre les deux cartes. Cependant, des zones telles que celles identifiées sur la Figure 12 présentent des anomalies. En effet, sur la zone [1], le modèle prédit une probabilité faible de présence de zone humide tandis que le catalogue de station attribue un niveau d'humidité élevé ; sur la zone [2], à l'inverse, le catalogue de station indique un niveau d'humidité faible, incohérent avec la probabilité élevée du modèle.

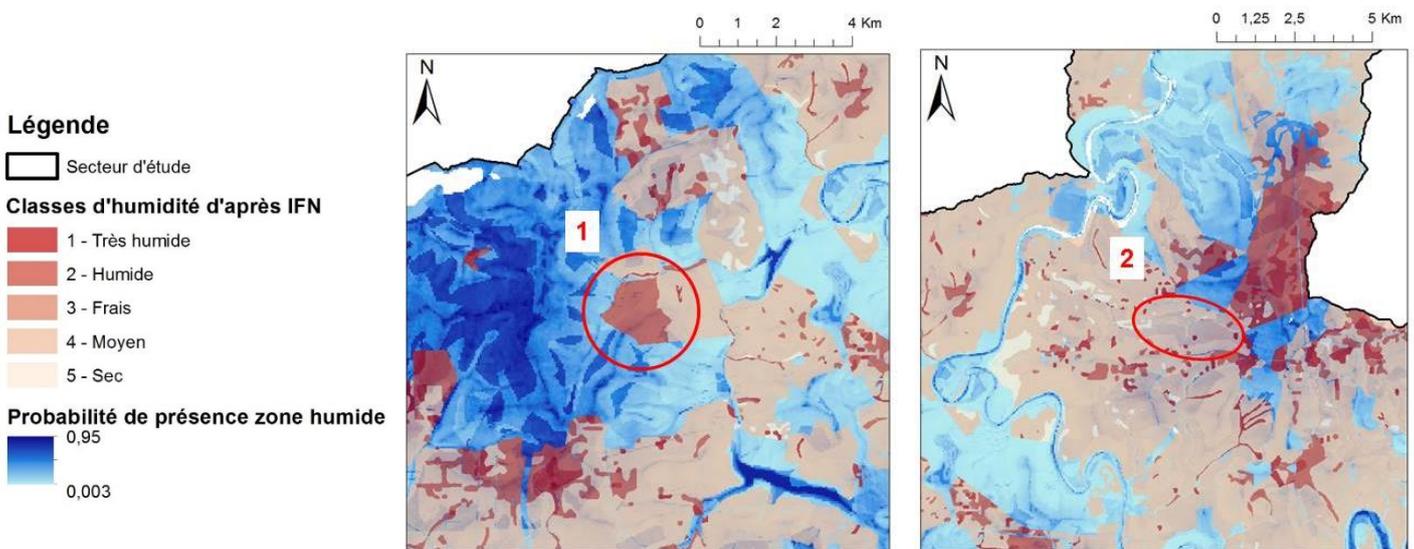


Figure 11 : Extraits de cartes de comparaison des probabilités de présence de zone humide (modèle) et des niveaux d'humidité des stations (IFN)

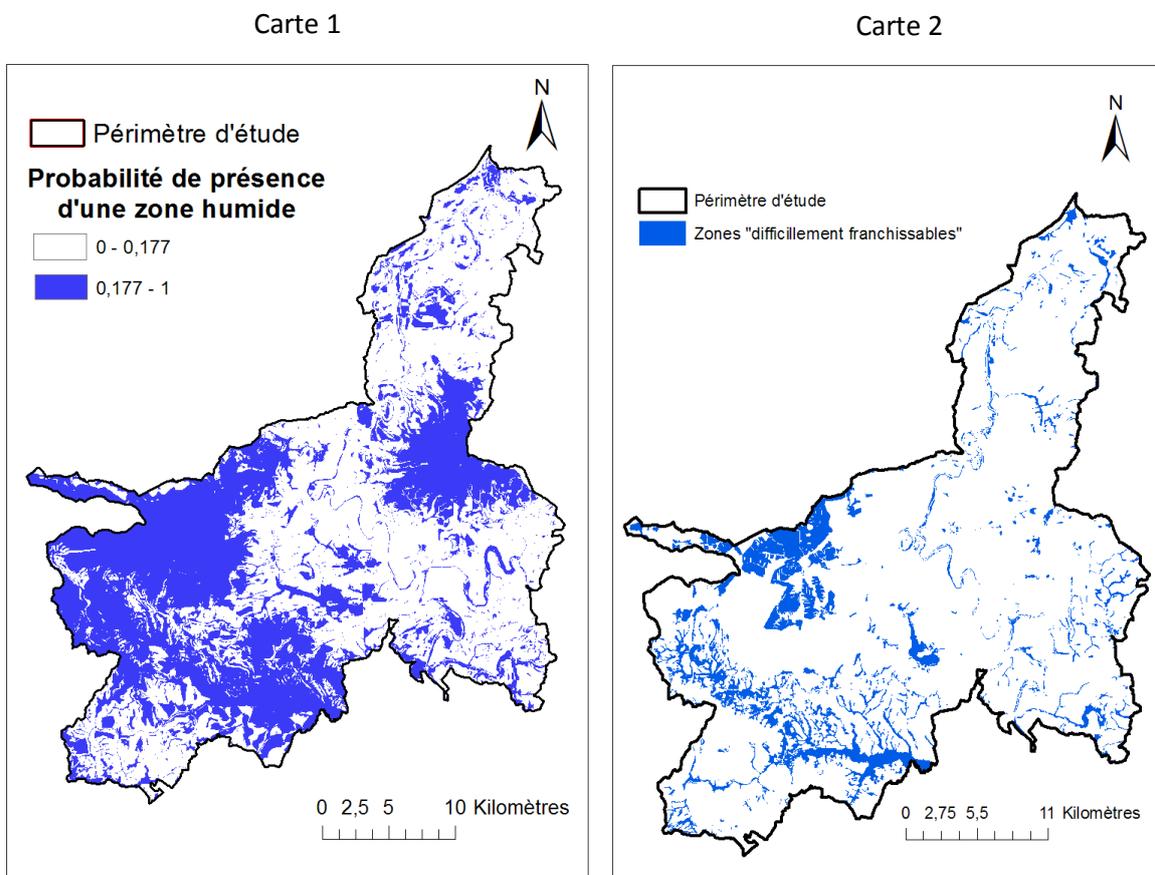
### 4.3.2 Comparaison avec les autres cartes et modèles

Dans le cadre de la validation, il a semblé intéressant de se pencher sur la robustesse du modèle en comparant la cartographie prédictive des zones humides avec des données déjà existantes.

Dans un premier temps, la carte issue du modèle est comparée avec celle de l'Etat-major (Carte 6). Cette dernière carte date du XIX<sup>ème</sup> siècle et représente les zones « difficilement franchissables par les armées ». Ces zones correspondent notamment aux cours d'eau et aux zones humides, mais pas uniquement.

Via la fonction « Zonal Histogram » de Arcgis (« Spatial Analyst tool » → « Zonal » → « Zonal Histogram »), on constate que 73% des zones « difficilement franchissables » sont prédites en tant que zone humide d'après le modèle statistique. En effet, une forte correspondance existe entre les deux cartes et ce même au niveau de zones humides plutôt ponctuelles ( Figure 13 **Erreur ! Source du renvoi introuvable.**, Carte 1). Il convient toutefois de relativiser ces observations puisqu'il existe de

grandes zones « difficilement franchissables » qui ne sont pas prédites en tant que zones humides pas le modèle (Figure 13, Carte 2). Cela peut s'expliquer notamment par le fait que les zones « difficilement franchissables » ne désignent pas que les zones humides.



Carte 6: Comparaison entre le modèle de probabilité de présence de zone humide (à gauche) et les zones "difficilement franchissables par les armées" selon les cartes de l'Etat-major (à droite)

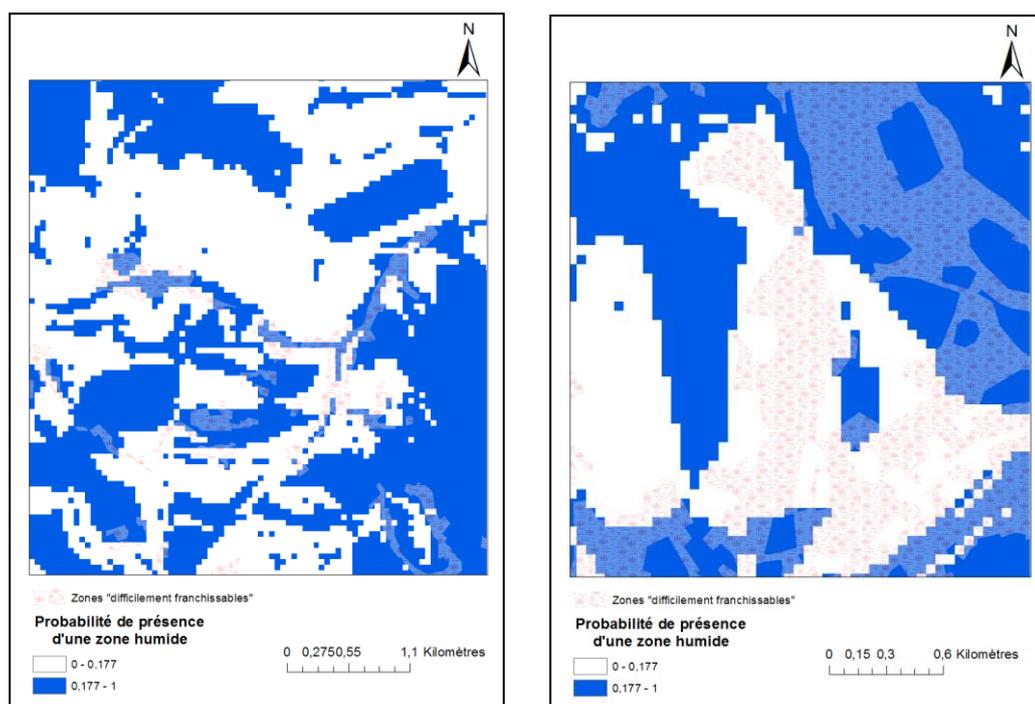
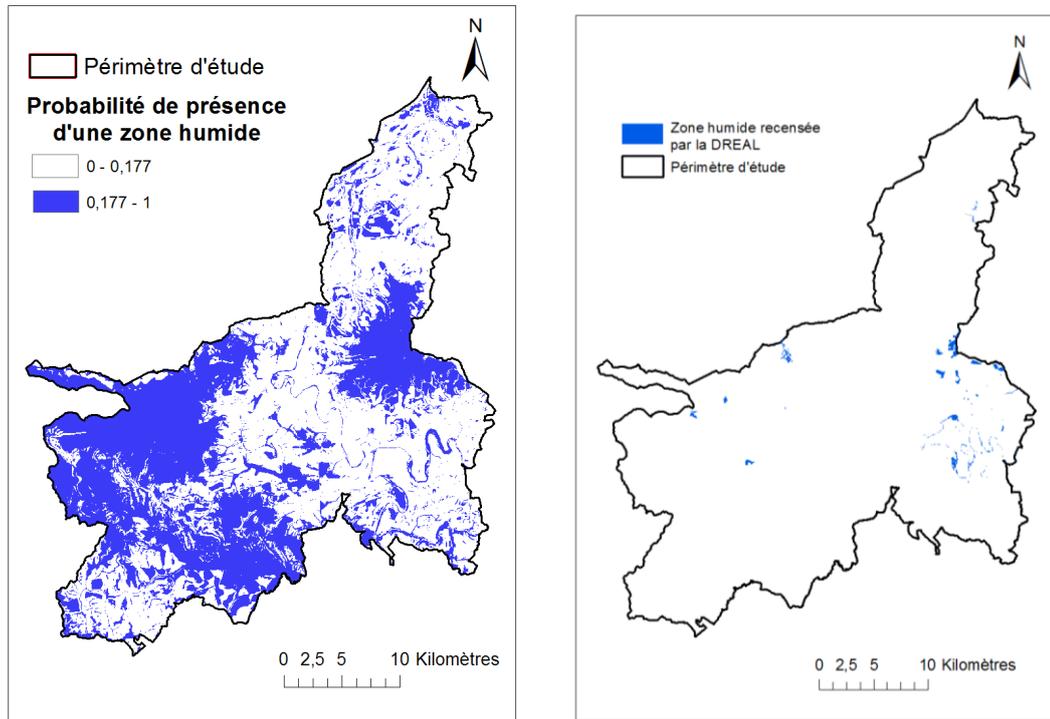


Figure 12 : Correspondance entre les zones « difficilement franchissables » et les probabilités de présence d'une zone humide

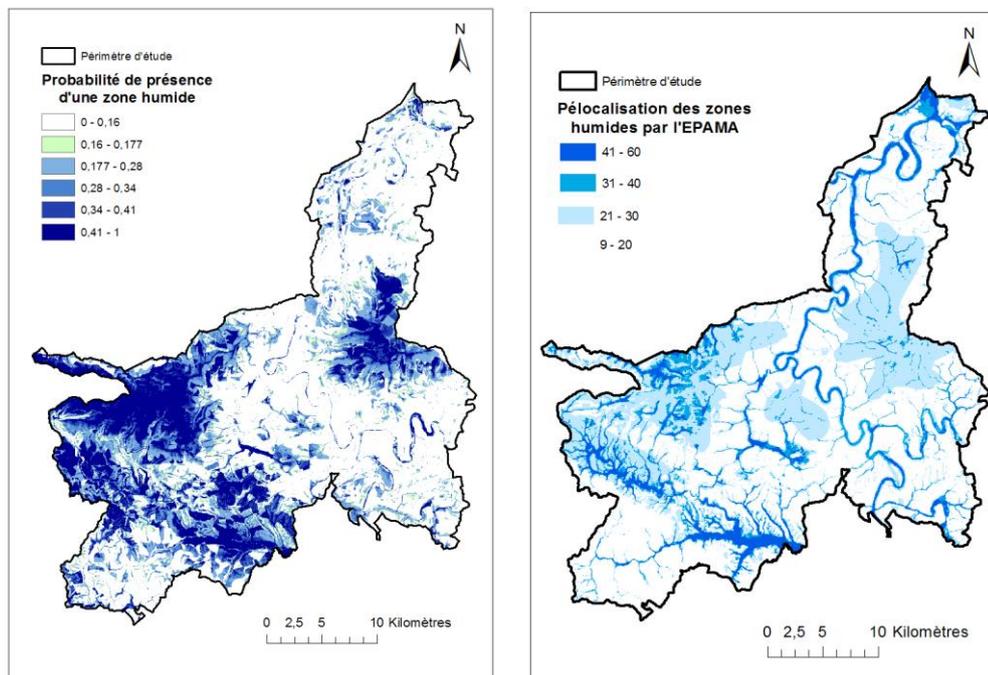
Dans un deuxième temps, la carte prédictive issue du modèle a été comparée à la carte des zones humides recensées (selon la loi sur l'eau) par la DREAL (Carte 7).



Carte 7 : Comparaison entre le modèle de probabilité de présence de zone humide (à gauche) et les zones humides recensées par la DREAL (à droite)

Malgré le peu de zones humides recensées, on remarque qu'elles sont globalement bien prédites par le modèle statistique. En effet, grâce à la fonction « Zonal Histogram » de Arcgis (« Spatial Analyst tool » → « Zonal » → « Zonal Histogram »), on constate que 71 % des zones humides cartographiées par la DREAL sont bien prédites par le modèle.

Enfin, la cartographie prédictive du modèle peut être comparée à celle du modèle établi par l'EPAMA (Carte 9).



Carte 8 : Comparaison entre le modèle de probabilité de présence de zone humide (à gauche) et le modèle établi par l'EPAMA (à droite)

On constate que les cours d'eau ressortent beaucoup plus sur la carte de l'EPAMA, ce qui pourrait être dû aux imprécisions de notre modèle au niveau des cours d'eau (voir Partie « Discussion »). Cela peut s'expliquer par le fait que la façon de construire notre modèle et les variables considérées sont différentes, ce qui aboutit à une cartographie différente mais qui pourrait être complémentaire de celle de l'EPAMA.

On remarque cependant que les grands ensembles humides sont globalement localisés aux mêmes endroits quel que soit le modèle. La gradation au sein des zones humides n'est toutefois pas la même et peut s'expliquer par le fait que les classes choisies pour hiérarchiser les zones humides ne sont pas les mêmes. Le tableau ci-après présente la correspondance entre les classes du modèle de l'EPAMA et de notre modèle.

Tableau 12 : Nombre de placettes terrain pour chaque classe de notre modèle et du modèle de l'EPAMA

		Classes de notre modèle					Total	
		0	1	2	3	4		5
Classes du modèle de l'EPAMA	09-20	22	7	9	1	1	0	40
	21-30	6	4	5	5	5	2	27
	31-40	1	1	0	0	2	0	4
	41-60	0	0	1	0	0	1	2
Total		29	12	15	6	8	3	73

On remarque que la première classe du modèle de l'EPAMA correspond assez bien à notre classe 0, et que les deux sont plutôt classées en zone non humide. Cependant, les autres classes ne correspondent pas. En effet, la façon de les construire est très différente, et les choix effectués ne sont pas les mêmes : il n'est donc pas anormal que les classes ne correspondent pas.

En conclusion, malgré des différences entre le modèle statistique et les différentes sources de cartographie des zones humides, on observe de nombreuses similitudes qui permettent de considérer la fiabilité du modèle comme satisfaisante.

## V. Discussion, limites et perspectives

### 5.1 Facteurs d'imprécision dans la modélisation

Si le modèle établi semble avoir une certaine robustesse statistique, il convient toutefois d'en aborder les limites et les perspectives d'amélioration.

En ce qui concerne les variables intégrées au modèle, la vectorisation manuelle de l'IDPR ainsi que la simplification des 20 classes originelles en 4 classes entraîne une approximation non négligeable de cette variable. De plus, malgré des recherches bibliographiques approfondies, il a

parfois été impossible de déterminer le caractère perméable ou imperméable de certaines couches géologiques ce qui a eu pour conséquence une perte d'information géologique. Par conséquent, la variable « indice », qui regroupe la donnée IDPR et la donnée géologique, ne peut être considérée comme fiable à 100% ce qui affaiblit inévitablement le modèle.

Dans une perspective d'amélioration du modèle il pourrait être intéressant de prendre en compte d'autres variables comme par exemple des variables liées au peuplement (essence majoritaire, densité, vitesse de croissance...). En effet, on peut supposer que le peuplement a un impact non négligeable sur la ressource en eau du milieu (prélèvement par les racines...). On pourra également s'intéresser à d'autres variables de topographie telles que le confinement ou la concavité, pouvant avoir des conséquences sur les apports latéraux en eau.

Concernant la validation du modèle, il est nécessaire de considérer les données pédologiques acquises sur le terrain avec une certaine réserve. En effet, notamment dans des contextes d'hydromorphie peu prononcée, il peut subsister des erreurs d'interprétation à l'origine d'incohérences entre le modèle et la réalité du terrain. Étant donné les conditions saisonnières défavorables, les informations des relevés pédologiques n'ont pas pu être complétées par des relevés floristiques. Toutefois, lors de campagnes de terrains ultérieures, il serait intéressant de prêter attention à la flore caractéristique des zones humides et ainsi lever, éventuellement, certaines de ces incohérences.

Au niveau des cours d'eau, les prédictions établies par le modèle doivent être prises en compte avec d'extrêmes précautions. À ces endroits, la géologie mentionne souvent la présence « d'alluvions » ce qui ne permet pas de préciser la perméabilité du sol. De plus, les cours d'eau constituent très souvent des limites de polygones : la vectorisation de l'IDPR y est très incertaine.

Le fonctionnement des zones humides près des cours d'eau est généralement différent : il serait plutôt lié à la nappe d'eau accompagnatrice qui s'infiltré dans des couches perméables pour alimenter la zone humide. D'autres phénomènes liés à la microtopographie (pente forte, encaissement) peuvent empêcher la formation de zones humides. Par ailleurs, à proximité des cours d'eau la plupart des paramètres (perméabilité, géologie) peuvent varier très fortement sur une faible distance. Le modèle pourrait être amélioré à ces endroits via la mise en place de zone « buffer ». Il ne prédirait alors la présence de zones humides qu'à partir d'une certaine distance du cours d'eau combinée à une certaine altitude au-dessus du lit mineur du cours d'eau afin de ne pas prendre en compte les mouvements de nappe.

### 5.1.1 Flore

Comme il a été constaté dans les résultats, la bio-indication ne semble pas concorder avec les prédictions établies dans le modèle. Si le modèle peut être à perfectionner, on peut également

questionner la pertinence de l'utilisation de la méthode de bio-indication sur cette zone d'étude. En effet, la calibration de cet outil ayant été réalisée à l'échelle nationale, son application à l'échelle locale pourrait entraîner des biais dus aux conditions écologiques locales. Les précipitations annuelles de la zone d'étude, supérieures à la moyenne nationale, pourraient être à l'origine d'un décalage des valeurs d'engorgement bio-indiquées.

La qualité de l'utilisation de la méthode de bio-indication peut néanmoins être améliorée. En premier lieu, comme l'autorise la méthode de bio-indication employée, la somme des engorgements temporaire et permanent a été utilisée dans la carte de comparaison au modèle. Cependant, comme il a été constaté lors de l'évaluation de la bio-indication (cf 2.4.2.), il peut être plus pertinent de distinguer ces deux types d'engorgement. Dès lors, on peut même considérer qu'il serait intéressant d'intégrer cette distinction à la construction du modèle, en remplaçant l'indice binaire de présence/absence de zone humide par un indice à trois modalités (drainé, temporaire, permanent).

D'autre part, l'utilisation d'une autre série de valeurs indicatrices, si possible plus adaptée au contexte local, pourrait permettre de réduire le biais dû à l'utilisation d'une calibration nationale.

### 5.1.2 Stations

Les incohérences observées entre la carte des niveaux d'humidité des stations forestières et les probabilités de présence de zone humide prédites par le modèle illustrent la difficulté d'intégrer le caractère humide des stations dans la typologie. Le catalogue utilisé, établi en 1989 par Drapier, n'intègre probablement pas suffisamment de paramètres pédologiques, topographiques ou floristiques permettant de distinguer les stations selon un critère d'humidité. Une évolution de la méthode utilisée pour la typologie des stations pourrait constituer un atout pour l'identification des zones humides forestières.

## 5.2 Limites dans l'utilisation des résultats

Il est important de distinguer les zones humides potentielles prédites par le modèle des zones humides avérées, réellement observées sur le terrain. Le modèle peut prédire la présence d'une zone humide sans que celle-ci puisse effectivement être observée sur le terrain. Effectivement, le modèle ne prend pas en compte les phénomènes non naturels qui impactent les zones humides comme l'anthropisation par exemple. Ainsi, il convient de bien garder à l'esprit que seule une campagne de terrain pourra permettre de lever ces doutes.

Par ailleurs, les incohérences entre les données terrain et le modèle peuvent s'expliquer, en milieu ouvert, par le fait que le modèle a été calibré sur des données entièrement forestières et serait donc moins adapté aux milieux ouverts.

De plus, au-delà d'être le seul moyen de valider le modèle convenablement, une campagne de terrain permettra d'augmenter le jeu de calibration du modèle (tout en gardant une partie des relevés pour la validation) et ainsi de l'affiner.

### **5.3 Extrapolation du modèle**

L'un des objectifs de cette étude était d'élaborer un modèle de prélocalisation des zones humides qui soit reproductible sur d'autres zones d'études.

Étant donné que le modèle statistique établi n'est basé que sur des variables climatiques, géologiques et topographiques, il est tout à fait envisageable de le reproduire sur d'autres aires géographiques où les données sont disponibles. À condition bien évidemment d'acheter ces données auprès du LERFoB et du BRGM au préalable. De plus, il faut que la zone ait été un minimum inventoriée par l'IFN (ou tout autre organisme fiable) afin d'avoir des données pédologiques d'entrée qui permettront d'expliquer la variable binaire de présence/absence de zone humide.

## **VI. Conclusion**

L'objectif premier de cette étude était le développement d'une méthode de prélocalisation des zones humides sur le territoire du PNR des Ardennes. Un modèle statistique basé sur des critères pédologiques a été construit. Les zones humides potentielles ont alors pu être localisées et hiérarchisées selon un potentiel niveau d'engorgement du sol. La possibilité d'utiliser les critères floristiques pour la définition des zones humides a été étudiée mais ne s'est pas avérée concluante dans l'immédiat. Une méthode similaire se basant sur des données floristiques pourrait être appliquée, à condition de vérifier que les données utilisées en entrée permettent de définir la présence de zones humides.

Comme attendu, un tel modèle est reproductible pour l'étude d'autres territoires. Son fonctionnement global resterait identique et son efficacité serait assurée grâce à une calibration à l'échelle locale. Les avantages ainsi que les limites de la méthode ont été précisés et doivent être pris en compte par le gestionnaire pour une utilisation cohérente.

Enfin, la prélocalisation des zones humides amène à mieux connaître ces espaces à forts enjeux. Leur hiérarchisation permet d'orienter et de limiter les prospections sur le terrain, qui restent néanmoins nécessaires à la localisation des zones humides avérées.

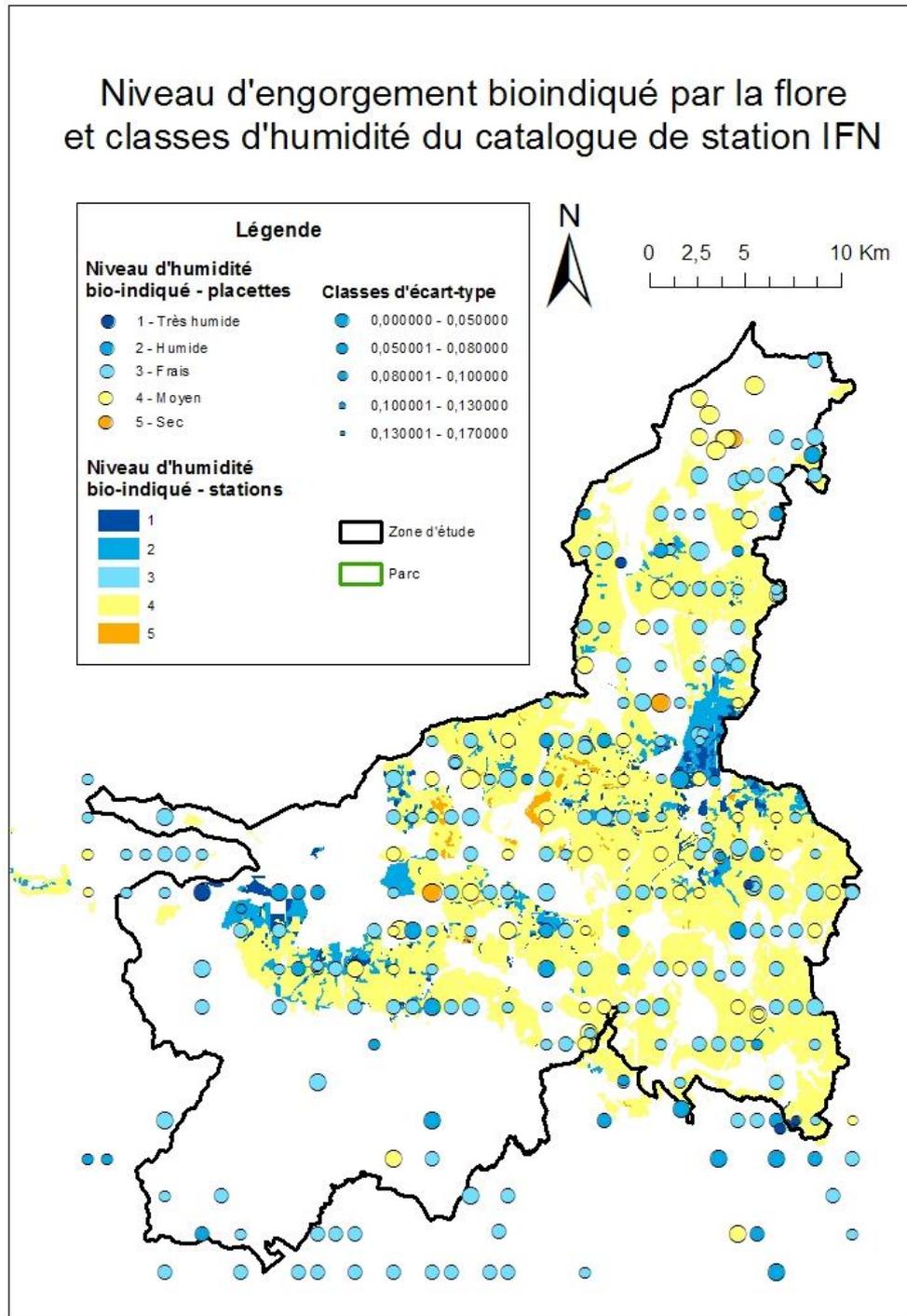
Rappelons que ce projet avait initialement une vocation pédagogique. Il a ainsi permis d'appréhender les divers enjeux liés aux zones humides à travers l'utilisation d'outils techniques comme les modèles statistiques ou les SIG. Cette étude fut également l'occasion de répondre aux attentes de commanditaires et donc de se confronter à des exigences professionnelles.

## Annexes

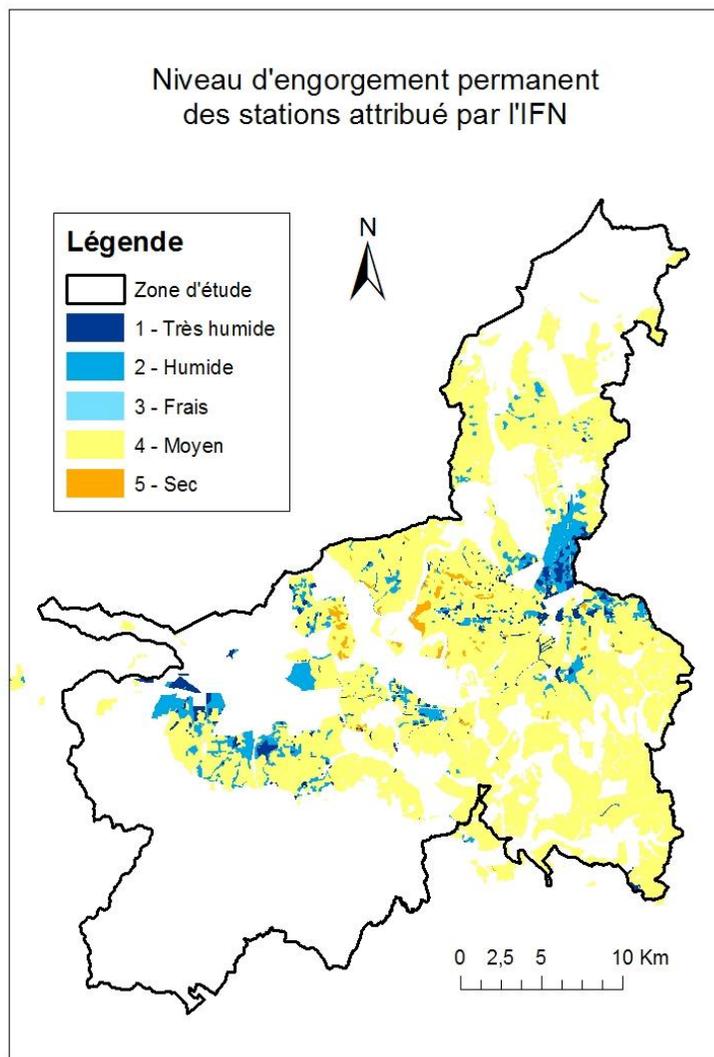
Annexe 1 : Fiche terrain utilisée pour la validation du modèle de prélocalisation des zones humides	39
Annexe 2 : Carte des niveaux d'engorgement bio-indiqués par la flore (selon la méthode de Jean-Claude Gégout) et classes d'humidité du catalogue de station IFN .....	40
Annexe 3 : Comparaison entre les cartographies des niveaux d'engorgement des stations forestières, attribués par l'IFN et bio-indiqués par la méthode de Jean-Claude Gégout .....	41
Annexe 4 : Carte du nombre d'espèce hygrophiles par placette d'après l'arrêté de 2008.....	42
Annexe 5 : Comparaison entre les niveaux d'engorgement bio-indiqués par la flore (placettes) et les probabilités de présence de zone humide (modèle).....	43
Annexe 6 : Carte de comparaison entre la répartition prédite du modèle des zones humides potentielles et le niveau d'humidité des stations d'après l'IFN .....	44
Annexe 7 : Clé de classes d'engorgement utilisée sur le terrain.....	45
Annexe 8 : Classes de perméabilité des couches géologiques d'après les notices du BRGM .....	46
Annexe 9 : Classes de perméabilité simplifiées de l'IDPR.....	47



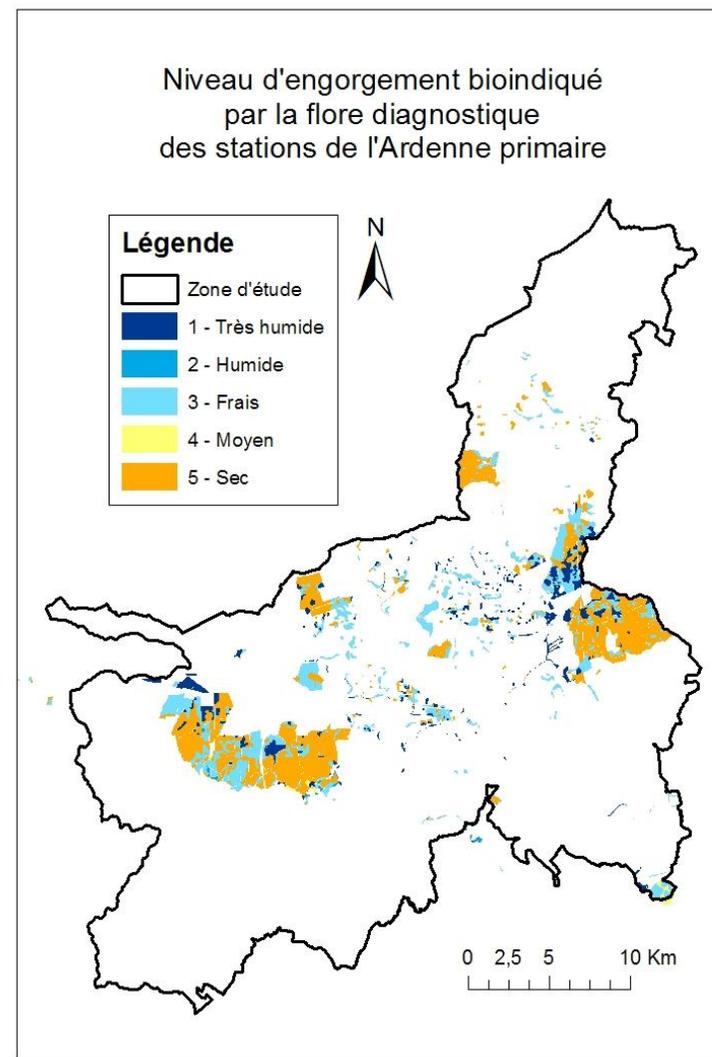
Annexe 2 : Carte des niveaux d'engorgement bio-indiqués par la flore (selon la méthode de Jean-Claude Gégout) et classes d'humidité du catalogue de station IFN

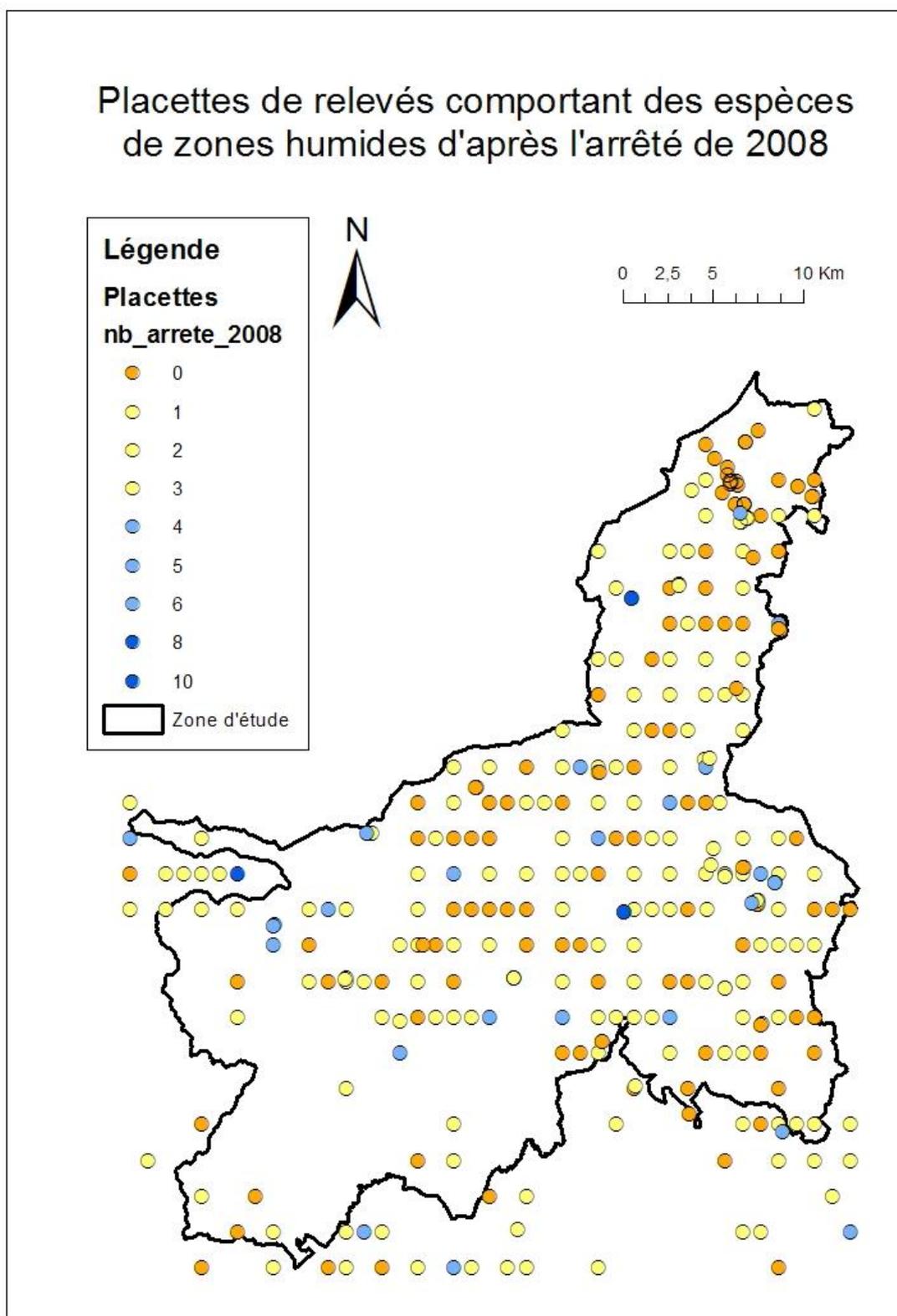


Carte 1

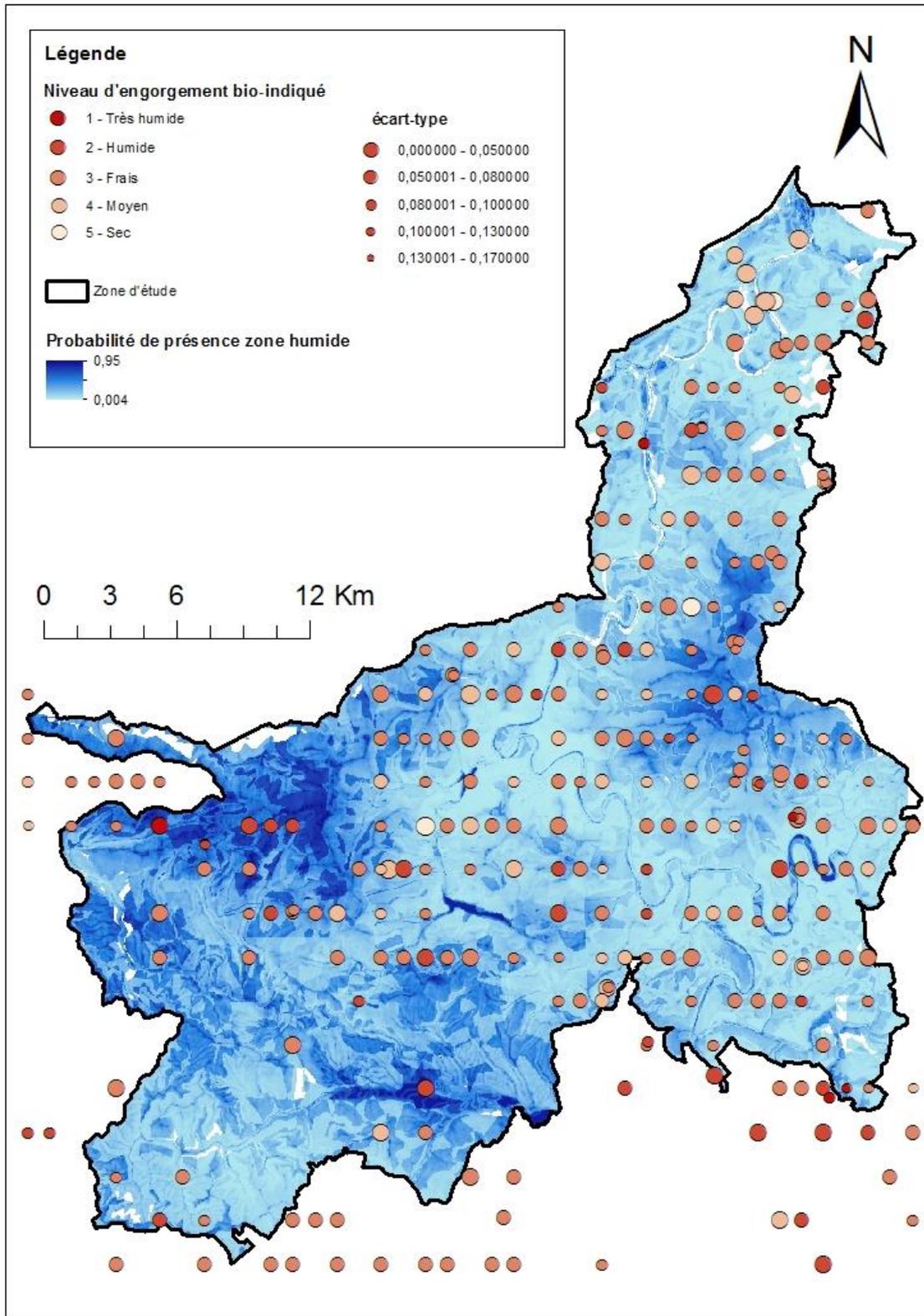


Carte 2

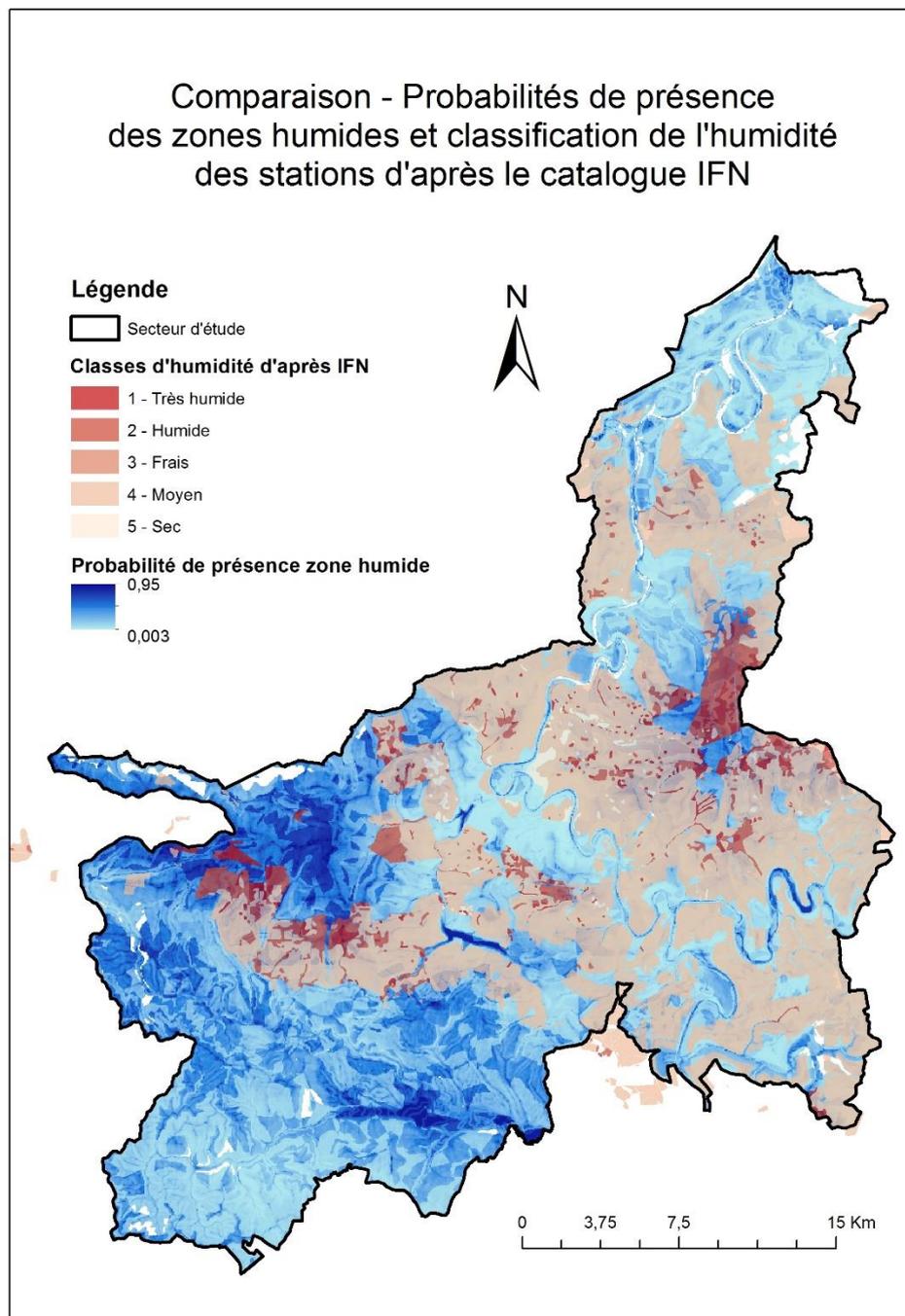




Annexe 5 : Comparaison entre les niveaux d'engorgement bio-indiqués par la flore (placettes) et les probabilités de présence de zone humide (modèle)



Annexe 6 : Carte de comparaison entre la répartition prédite du modèle des zones humides potentielles et le niveau d'humidité des stations d'après l'IFN



Annexe 7 : Clé de classes d'engorgement utilisée sur le terrain

Horizon histique (organique et fibreux : tourbe) très épais ( $\geq 30$ cm) -----	5
<i>Si non</i>	
Horizon histique épais ( <b>11 à 29 cm</b> ) -----	4
<i>Si non</i>	
Horizon histique <b>absent</b> et <b>décoloration complète en surface</b> ou peu épais ( $\leq 10$ cm) -----	3
<i>Si non</i>	
Taches rouille et grises sans décoloration complète dans les <b>25 premiers cm</b> -----	2
<i>Si non</i>	
Taches rouille et grises sans décoloration complète entre <b>25 et 50 cm</b> -----	1
<i>Si non</i>	
<b>Aucun de ces critères</b> -----	0

Annexe 8 : Classes de perméabilité des couches géologiques d'après les notices du BRGM (1 : perméable – 2 : imperméable – 0 : pas d'information)

Notation	Classe perméabilité
b1	1
b2	2
b2b	2
b3b	2
b3c	2
b3d	2
C	0
C1	1
C2	0
C2(3)	0
C2a	0
CF	0
d1a	2
d1a'	0
d1b	2
d1c	2
d1d	2
d2a	0
d2b	1
d2c	1
d3a	2
d3b	0
d3c	2
d4c1	0
d4d	0
d6d	0
e2b	0
Fp	1
Fv3	0
Fy	1
Fz	1
Fz/Fy	1
hydro	0
l3a	0
l3b	0
l3c	1
l4	0

Notation	Classe perméabilité
l5	0
j1a	1
j1b	1
j2a	1
j2b	1
j2c	1
j3	0
j4a	0
K4	2
K4b2	0
K4c1	0
K4d	0
K4e	0
K4f	0
K4g	0
K4h	0
K4i	0
K4k	0
K4l	0
K4y	0
k5a	2
k5b	2
k5c	0
l3c	0
l4a	0
l5	0
l6a	0
l6b	0
LP	0
LP/l1-2	0
n3	1
R1	2
R1/d1a	2
R1/d1b	2
R1/d1b2	2
R1/d2a	2
R1/d2b	2

*Annexe 9 : Classes de perméabilité simplifiées de l'IDPR*

<b>IDPR</b>	<b>Classe perméabilité</b>	
0 – 800 (4 classes)	4	Forte infiltration
800 – 1200 (2 classes)	3	↓
1200 – 1600 (2 classes)	2	↓
1600 – 2000 (2 classes)	1	Refus d'infiltration