

Convention 2005 – 2008 ADEME / IFN / ENGREF n°0562C0029  
**Programme « Biondicateurs de qualité des sols »**

**VALIDATION DE BIOINDICATEURS FLORISTIQUES POUR LA SURVEILLANCE DE  
L'ETAT NUTRITIONNEL DES SOLS FORESTIERS FRANÇAIS A PARTIR DES  
DONNEES DE L'INVENTAIRE FORESTIER NATIONAL.**

Rapport final



UMR INRA – ENGREF 1092  
Laboratoire d'Étude de la Ressource  
Forêt Bois (LERFOB)  
Équipe Écologie Forestière  
14 rue Girardet, CS 14216,  
540432 Nancy Cedex

Jean-Claude GEGOUT

**NOVEMBRE 2008**



# SOMMAIRE

<b>1</b>	<b>INTRODUCTION .....</b>	<b>7</b>
<b>2</b>	<b>QUALITÉ DE PRÉDICTION DES VARIABLES NUTRITIONNELLES DU SOL PAR LE CARACTÈRE BIOINDICATEUR DE LA FLORE .....</b>	<b>9</b>
2.1	EVALUATION DE LA QUALITÉ DES PRÉDICTIONS RÉALISÉES PAR BIOINDICATION.....	11
2.1.1	<i>Introduction .....</i>	<i>11</i>
2.1.2	<i>Matériel et méthodes.....</i>	<i>11</i>
2.1.3	<i>Résultats : qualité de la bioindication par la flore.....</i>	<i>13</i>
2.2	EFFICACITÉ DE VALEURS INDICATRICES DE PH ÉTABLIES À L'ÉCHELLE NATIONALE PAR RAPPORT À DES VALEURS INDICATRICES RÉGIONALES .....	14
2.2.1	<i>Introduction .....</i>	<i>14</i>
2.2.2	<i>Comparaison des valeurs indicatrices définies dans le nord-est et le nord-ouest de la France.....</i>	<i>15</i>
2.2.3	<i>Comparaison des prédictions réalisées avec des valeurs indicatrices régionales et des valeurs indicatrices nationales .....</i>	<i>16</i>
2.3	VARIABILITÉ GÉOGRAPHIQUE DE LA QUALITÉ DE PRÉDICTION DU PH EN FRANCE .....	16
2.3.1	<i>Matériel et méthodes.....</i>	<i>16</i>
2.3.2	<i>Résultats .....</i>	<i>19</i>
2.4	UTILISATION DES VALEURS INDICATRICES POUR ESTIMER LA QUALITÉ NUTRITIONNELLE DES SOLS.....	20
<b>3</b>	<b>EVALUATION DE LA BAISSÉ DE QUALITÉ DES PRÉDICTIONS LIÉE À L'UTILISATION DE RELEVÉS FLORISTIQUES INCOMPLETS EN VUE DE L'UTILISATION DES DONNÉES DE L'INVENTAIRE FORESTIER NATIONAL .....</b>	<b>23</b>
3.1	INTRODUCTION .....	25
3.2	MATÉRIEL ET MÉTHODES.....	25
3.3	QUALITÉ DES PRÉDICTIONS ET PROPORTION D'ESPÈCES DU RELEVÉ UTILISÉES POUR LA BIOINDICATION DU PH.....	26
3.4	QUALITÉ DES PRÉDICTIONS ET NOMBRE D'ESPÈCES DU RELEVÉ UTILISÉES POUR LA PRÉDICTION DU PH...	27
3.5	POSSIBILITÉ DE PRÉDICTION TOUTE L'ANNÉE DES VARIABLES DU SOL AVEC LA FLORE.....	30
3.5.1	<i>Méthode de détermination de la flore hivernale .....</i>	<i>30</i>
3.5.2	<i>Prédictions réalisées.....</i>	<i>31</i>
3.5.3	<i>Qualité des prédictions réalisées avec des relevés simulés hivernaux ou sans bryophytes</i>	<i>32</i>
3.6	CONCLUSIONS SUR LA POSSIBILITÉ D'UTILISER DES RELEVÉS FLORISTIQUES PARTIELS POUR RÉALISER LES PRÉDICTIONS DES VARIABLES DU SOL.....	34
<b>4</b>	<b>POSSIBILITÉ DE CARTOGRAPHIE DU PH DES SOLS FORESTIERS À PARTIR DES VALEURS INDICATRICES D'ECOPLANT ET DES RELEVÉS FLORISTIQUES DE L'INVENTAIRE FORESTIER NATIONAL .....</b>	<b>35</b>
4.1	INTRODUCTION .....	37
4.2	MÉTHODE DE SPATIALISATION DU PH À L'ÉCHELLE DU TERRITOIRE FRANÇAIS.....	37
4.3	RÉSULTATS : LA CARTE DES PH DES SOLS FORESTIERS FRANÇAIS.....	38
4.3.1	<i>La carte des pH.....</i>	<i>38</i>
4.3.2	<i>Précision de la carte.....</i>	<i>41</i>
<b>5</b>	<b>CONCLUSION ET PERSPECTIVES .....</b>	<b>43</b>
	<b>ANNEXE : ARTICLES SCIENTIFIQUES, PUBLIÉS ENTRE 2005 ET 2008, QUI UTILISENT LES DONNÉES DE LA BASE DE DONNÉES ECOPLANT OU SES VALEURS INDICATRICES. ....</b>	<b>47</b>

## LISTE DES FIGURES

<b>Figure 1.</b>	Courbes de réponse et valeurs indicatrices de 5 espèces végétales vis-à-vis du pH dans les forêts françaises. ....	11
<b>Figure 2.</b>	Démarche de prédiction d'une variable écologique à partir des valeurs indicatrices des espèces et de relevés floristiques. ....	12
<b>Figure 3.</b>	Distribution des distances minimales entre les 261 placettes du jeu de validation et les placettes utilisées pour modéliser les valeurs indicatrices des espèces. Les quantiles 10, 25, 50, 75 et 90 de la distribution des 261 distances minimales sont représentées. Les distances élevées assurent l'indépendance entre données de calibration des valeurs indicatrices et données de validation. ....	13
<b>Figure 4.</b>	Distribution sur le territoire français des 261 placettes de validation des valeurs indicatrices des plantes utilisées pour estimer la qualité nutritionnelle des sols forestiers. ....	13
<b>Figure 5.</b>	Localisation des 2404 relevés d'EcoPlant utilisés pour comparer le caractère indicateur des plantes entre le nord-est et le nord-ouest de la France. Les pH acides sont indiqués en noirs, les pH neutres sont représentés par un cercle blanc. ....	15
<b>Figure 6.</b>	Valeurs indicatrices pour le pH de 46 espèces forestières dans le nord-est et le nord-ouest de la France. ....	15
<b>Figure 7.</b>	Carte de répartition des substrats géologique des placettes utilisées pour analyser la qualité de prédiction en fonction du type de substrat siliceux acide (rouge), peu acide (beige) et calcaire (bleu). ....	17
<b>Figure 8.</b>	Carte des pH de l'horizon A du sol de 3601 placettes forestières d'EcoPlant mesurés en laboratoire (haut) et bioindiqués par la flore des placettes (bas). ....	18
<b>Figure 9.</b>	Ecarts entre pH bioindiqués par la flore des placettes et pH mesurés. (bleu : pH bioindiqué > pH mesuré, rouge : pH bioindiqué < pH mesuré, gris : écarts de pH < 0.5 unité) ....	19
<b>Figure 10.</b>	Représentation des pH et C/N estimés sur 12 453 placettes issues des bases de données Sophy et EcoPlant. Les placettes indiquées en grisé sont situées au nord de la France, celles en noir sont situées au sud de la France. Une courbe de tendance est superposée au nuage. ....	21
<b>Figure 11.</b>	Histogramme des nombres d'espèces des 261 relevés utilisés pour étudier l'effet de relevés incomplets sur la qualité de prédiction. ....	25
<b>Figure 12.</b>	Variation de l'écart (REQM) entre pH mesuré et pH bioindiqué par la flore pour des proportions croissantes d'espèces des relevés prises en compte pour réaliser les prédictions. ....	26
<b>Figure 13.</b>	Variation de la corrélation ( $R^2$ ) entre pH mesurés et pH bioindiqués par la flore pour des proportions croissantes d'espèces des relevés prises en compte pour réaliser les prédictions. ....	27
<b>Figure 14.</b>	Evolution de la REQM [a] et du R2 [b, c] entre pH mesurés et bioindiqués avec le nombre d'espèces du relevé utilisées pour réaliser la prédiction. Les fig. 14a et b sont réalisés avec les 261 relevés de validation. Le graphe 14c est réalisé avec les 575 relevés d'EcoPlant comprenant au moins 30 espèces. ....	28
<b>Figure 15.</b>	Evolution, selon le niveau d'acidité mesurée du sol, de la REQM entre pH mesurés et pH prédits avec un nombre croissant d'espèces du relevé prises en compte (Tot : totalité des espèces prises en compte). Les valeurs de REQM présentées pour chaque nombre d'espèces étudié correspondent à une moyenne de 50 prédictions réalisées après une sélection aléatoire des espèces. ....	29
<b>Figure 16.</b>	Relation entre pH mesurés sur les 261 placettes de validation et pH estimés par la flore. En haut à gauche, relevés complets. Dans les autres graphes, relevés partiels (croix noires) comparés aux relevés complets. Voir les espèces incluses dans chaque simulation dans le tableau 7. ....	33
<b>Figure 17.</b>	Méthode de spatialisation des variables bioindiquées à partir des valeurs indicatrices d'EcoPlant et des placettes de l'Inventaire Forestier National. ....	37
<b>Figure 18.</b>	Représentation du pH estimé par les valeurs indicatrices d'EcoPlant sur les 104 375 placettes réalisées par l'Inventaire Forestier National en France entre 1989 et 2004. ....	39
<b>Figure 19.</b>	Carte des pH de l'horizon A des sols forestiers français estimés à partir des valeurs indicatrices d'EcoPlant et des 104 375 placettes réalisées par l'Inventaire Forestier National en France entre 1989 et 2004. ....	40
<b>Figure 20.</b>	Relation, pour les 261 placettes du jeu de validation, entre le pH mesuré et le pH spatialisé, obtenu par interpolation et du pH bioindiqué sur les placettes de l'Inventaire Forestier National. La droite $y=x$ est ajoutée au graphe. ....	41

## LISTE DES TABLEAUX

<b>Tableau 1.</b>	Données utilisées pour le calcul des valeurs indicatrices à l'échelle de la France. . . . .	11
<b>Tableau 2.</b>	Qualité d'estimation par la bioindication des variables liées aux capacités de nutrition du sol. Les données employées sont les 261 placettes du jeu de validation. . . . .	14
<b>Tableau 3.</b>	Lien entre valeurs indicatrices nationales et régionales et qualités des prédictions réalisées à partir de VI nationales et régionales. . . . .	16
<b>Tableau 4.</b>	Modèles, à l'échelle de la France et par catégories de substrats géologiques, de prédiction des pH mesurés et bioindiqués par la longitude, la latitude et l'altitude. Les coefficients des régressions multiples sont présentés dans les 4 premières colonnes. Le R <sup>2</sup> et la REQM des régressions sont présentés ensuite. . . . .	19
<b>Tableau 5.</b>	Variation spatiale des pH à l'échelle de la France, à substrat et altitude fixés, selon les gradients nord-sud et est-ouest (les valeurs sont estimées avec les paramètres des équations du tableau 4). . . . .	20
<b>Tableau 6.</b>	Types biologiques de Ellenberg et Landolt. . . . .	30
<b>Tableau 7.</b>	Synthèse des classes d'espèces utilisées pour simuler des prédictions au cours des différentes périodes de l'année et avec différents niveaux de qualité. Le nombre d'espèces appartenant à chaque classe est mentionné dans le tableau. . . . .	31
<b>Tableau 8.</b>	Caractéristiques principales des prédictions de pH du sol (horizon A) réalisées avec des relevés floristiques complets et partiels simulant l'absence de bryophytes, des relevés exhaustifs réalisés en période hivernale et des relevés rapides réalisés en période hivernale. . . . .	33



## 1 INTRODUCTION

La bioindication de l'humidité ou des capacités de nutrition du sol par la présence ou l'absence des plantes est utilisée depuis longtemps pour évaluer la qualité des milieux forestiers (Becker & Le Goff, 1988; Cajander, 1926; Duchaufour, 1948), pour choisir des essences de reboisement adaptées aux conditions écologiques des sites, ou pour prédire les capacités de production des peuplements (Lahti, 1995; Nieppola, 1993). Dans un contexte d'acidification et d'eutrophisation à long terme des écosystèmes, les plantes sont également utilisées pour détecter et analyser les changements des caractéristiques édaphiques sous l'influence des activités anthropiques (Falkengren-Grerup, 1989; Thimonier et al., 1994). Le caractère bioindicateur des plantes a enfin été utilisé pour comprendre la variabilité de la diversité dans les forêts européennes (Dupre & Diekmann, 1998). Le système de bioindication mis au point par Ellenberg pour l'Europe centrale, basé sur le préférendum des espèces, est le plus fréquemment utilisé pour aborder ces problématiques (Ellenberg et al., 1992). Au niveau international, les systèmes de bioindication utilisés à large échelle spatiale, utilisant la présence/absence des plantes, reposent tous sur une évaluation du caractère indicateur des espèces établie de façon empirique à partir de l'expérience des naturalistes.

Les méthodes de modélisation de la distribution des espèces mises au point à la fin du 20<sup>ème</sup> siècle permettent maintenant de définir une méthode complètement formalisée de bioindication à partir de la présence/absence des espèces. Une première étape consiste à modéliser la distribution de chaque espèce vis-à-vis de la variable étudiée en établissant une courbe de probabilité de présence de l'espèce pour les différentes valeurs prises par la variable dans les conditions naturelles. Cette courbe est généralement réalisée à partir d'inventaires floristiques effectués sur de nombreux sites pour lesquels on relève également la valeur de la variable étudiée. Les courbes de réponse permettent d'établir le caractère indicateur des espèces et il est ensuite possible d'utiliser celui-ci pour prédire la variable uniquement à partir des inventaires floristiques établis sur les placettes. L'une des méthodes les plus simples et efficaces pour définir le caractère indicateur des plantes consiste à définir un indice numérique synthétique pour chaque espèce. L'indice retenu, établi à partir de la courbe de réponse de l'espèce et appelé classiquement « valeur indicatrice », est la valeur du facteur pour laquelle la probabilité de présence de l'espèce est maximale. Il est ensuite possible de réaliser une estimation du facteur par bioindication à partir de l'inventaire floristique des sites : la valeur estimée correspond à la moyenne des valeurs indicatrices des espèces présentes sur chaque site.

Les premières valeurs indicatrices établies en France ont été élaborées pour 157 espèces forestières dans le massif vosgien vis-à-vis des paramètres trophiques du sol (Gégout, 1995). Contrairement aux approches empiriques méditerranéennes, le caractère bioindicateur des espèces a été déterminé à partir des courbes de réponse des espèces. Dans le cadre d'un contrat ADEME, la valeur indicatrice de plus de 500 espèces de la flore française a ensuite été déterminée à l'échelle du territoire national pour des variables édaphiques et climatiques grâce à la base de données EcoPlant (Gégout et al., 2002).

La détermination du caractère bioindicateur des plantes étant réalisée, l'objectif de ce projet financé par l'ADEME est de :

- Déterminer la qualité de prédiction du pH, de la nutrition minérale et la nutrition azotée des sols forestiers français à partir du caractère bioindicateur des espèces végétales ;
- Evaluer la baisse de qualité des prédictions liée à l'utilisation de relevés floristiques incomplets en vue de l'utilisation des données de l'Inventaire Forestier National ;
- Cartographier l'état nutritionnel des sols forestiers français à l'aide des 80 000 placettes de l'Inventaire Forestier National.

Ce projet s'inscrit dans le cadre du besoin prioritaire « surveillance de la qualité des sols » défini par l'ADEME dans son appel à proposition.

Ce rapport final de la convention vise à présenter les résultats obtenus pour les trois volets. L'annexe du rapport indique les articles scientifiques publiés durant la durée de la convention qui utilisent la valeur indicatrice des espèces.

## **2 QUALITE DE PREDICTION DES VARIABLES NUTRITIONNELLES DU SOL PAR LE CARACTERE BIOINDICATEUR DE LA FLORE**



## 2.1 EVALUATION DE LA QUALITE DES PREDICTIONS REALISEES PAR BIOINDICATION

### 2.1.1 Introduction

Des valeurs indicatrices des variables nutritionnelles du sol ont été établies à partir de la base EcoPlant pour les principales espèces de la flore forestière française au début des années 2000 (Gégout et al., 2000). Elles permettent d'élargir le système de bioindication mis en place pour le massif Vosgien (Gégout et al., 2003). Cependant, la qualité de prédiction des variables prédites par ce système de bioindication établi à l'échelle de la France n'a pas été vérifiée par des données mesurées. L'objet de cette partie est d'utiliser un jeu de données représentatif des forêts françaises pour évaluer la qualité de prédiction obtenue par bioindication des variables nutritionnelles avec la flore des sites.

### 2.1.2 Matériel et méthodes

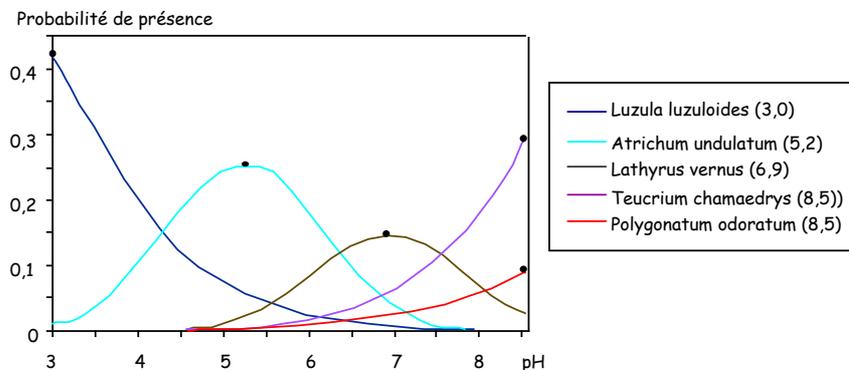
#### Variables nutritionnelles étudiées et méthode de bioindication

Des valeurs indicatrices ont été calculées pour le  $pH_{eau}$ , le rapport S/T et le rapport C/N de l'horizon A du sol. Le rapport S/T est calculé, après extraction au pH du sol des cations échangeables, comme le rapport  $(Ca+Mg+K+Na)/CEC_{pHsol}$ . La CEC est mesurée en sol calcaire ou calculée comme étant la somme des cations échangeables  $(Ca+Mg+K+Na+Al+H)$  en sol acide. Le rapport C/N est le rapport carbone organique/azote total. Ces variables ont été mesurées sur des échantillons de sols prélevés au moment de la réalisation des relevés floristiques (tableau 1).

**Tableau 1.** Données utilisées pour le calcul des valeurs indicatrices à l'échelle de la France.

Variable	Nb de relevés utilisés	Gamme étudiée	Seuil d'occurrences	Nb d'espèces traitées
$PH_{eau}$	3 835	3 – 8.5	10	568
S/T	2 670	0 – 100	10	518
C/N	3 547	10 – 40	10	557

La réponse des espèces aux facteurs a été modélisée par l'élaboration d'une courbe de probabilité de présence de chaque espèce le long de chaque facteur à l'aide de modèles logistiques (figure 1).

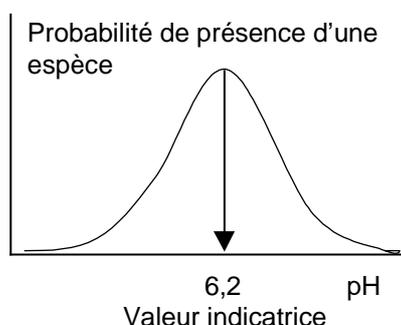


**Figure 1.** courbes de réponse et valeurs indicatrices de 5 espèces végétales vis-à-vis du pH dans les forêts françaises

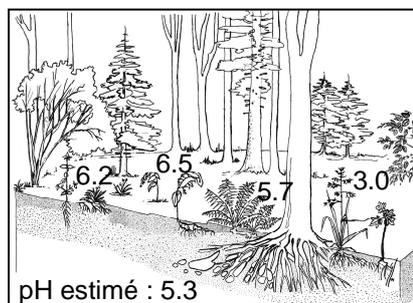
Le valeur indicatrice de chaque espèce pour chaque variable de l'état nutritionnel du sol a été résumé par la valeur de la variable correspondant au maximum de la courbe de réponse de

l'espèce (Ter Braak et Looman, 1986). La prédiction d'une variable du sol à l'aide du caractère bioindicateur de la flore a été effectué en réalisant la moyenne des valeurs indicatrices des espèces présentes sur chaque placette (figure 2).

### 1 - Valeur indicatrice des espèces



### 2 - Prédiction des variables du sol avec les valeurs indicatrices



Estimation du pH, C/N, et d'autres variables écologiques avec les valeurs indicatrices des espèces du relevé

**Figure 2.** démarche de prédiction d'une variable écologique à partir des valeurs indicatrices des espèces et de relevés floristiques.

### Mesure de la qualité de prédiction

Deux mesures ont été utilisées pour évaluer la qualité d'estimation des différentes approches de bioindication. La première est la racine de l'erreur quadratique moyenne des prédictions :

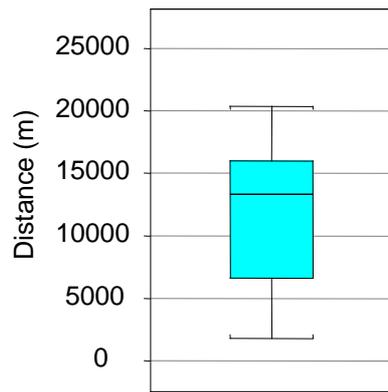
$$RMSE = \sqrt{\frac{1}{n} \cdot \sum_{j=1}^n (x_j - \hat{x}_j)^2}$$

Cette erreur prend en compte le biais des prédictions par rapport aux valeurs mesurées et l'imprécision des prédictions autour de ce biais.

La seconde mesure de qualité utilisée est le carré de la corrélation linéaire  $R^2$  entre les valeurs mesurées et prédites de la variable. Le  $R^2$  est indépendant de tout biais des prédictions qui peut être corrigé par une régression linéaire. Il présente également l'avantage de permettre les comparaisons entre les qualités de prédiction de différentes variables.

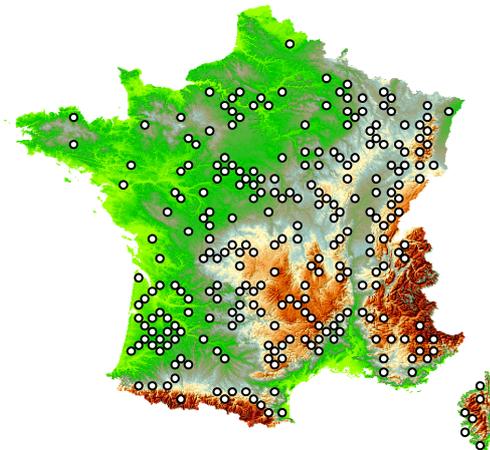
### Données de validation

Un jeu de 261 relevés du réseau européen (Badeau 1998), répartis en quinconce sur une grille de 16 km de côté placée sur la France entière, a été utilisé pour estimer la qualité des prédictions réalisées par les valeurs indicatrices. Ce jeu de validation est différent du jeu de calibration et ne repose pas sur les mêmes sources de données. Cet avantage permet d'éviter une trop forte corrélation temporelle avec les données de calibration, qui aurait pu être à l'origine d'une surestimation de la qualité de la bioindication (Araujo et al., 2006). La distance minimale moyenne entre ces 261 relevés et les placettes utilisées pour calculer les valeurs indicatrices de niveau trophique est égale à 12.6 km (figure 3). Ces distances, très importantes et supérieures à 1 km pour plus de 90 % des placettes, assurent une absence ou quasi-absence d'autocorrélation spatiale entre les données utilisées pour la calibration et la validation des modèles, ce qui assure une estimation sans biais spatial de la qualité d'estimation.



**Figure 3.** Distribution des distances minimales entre les 261 placettes du jeu de validation et les placettes utilisées pour modéliser les valeurs indicatrices des espèces. Les quantiles 10, 25, 50, 75 et 90 de la distribution des 261 distances minimales sont représentées. Les distances élevées assurent l'indépendance entre données de calibration des valeurs indicatrices et données de validation.

La répartition systématique des placettes de validation sur l'ensemble des forêts du territoire français constitue un autre avantage de ces données (figure 4). L'erreur de prédiction des variables du milieu par la végétation de ces 261 placettes est une estimation non biaisée de l'erreur de prédiction attendue pour toute placette sélectionnée de façon aléatoire dans les forêts françaises. Elle correspond donc à l'erreur moyenne attendue en n'importe quel point forestier du territoire français.



**Figure 4.** Distribution sur le territoire français des 261 placettes de validation des valeurs indicatrices des plantes utilisées pour estimer la qualité nutritionnelle des sols forestiers.

### 2.1.3 Résultats : qualité de la bioindication par la flore

Les écarts entre valeurs mesurées et estimées, mesurés par la REQM, sont de l'ordre de 0.8 unité de pH, 20 % de S/T ou 5 unités de C/N (tableau 2). Ces écarts correspondent à l'amplitude moyenne des classes de niveau trophique définies pour ces variables (Baize & Girard, 1992; Duchaufour, 1988). La REQM représente la différence entre une mesure réalisée sur quelques grammes de terre issus d'un échantillon prélevé à un moment donné sur l'horizon A d'une fosse, et une estimation réalisée à partir d'un relevé floristique réalisé sur 400 m<sup>2</sup>. Il existe entre les deux sources de données une différence d'ordre spatial : la flore intègre tout l'espace de la placette (en général 400 m<sup>2</sup>) alors que l'échantillon ne représente que la zone de la fosse. La variabilité spatiale du pH de l'horizon A au sein des placettes a été testée à partir

d'un sous-jeu de 27 placettes d'EcoPlant situées en milieu siliceux dans le massif vosgien (de Boutray, 2003). La REQM des mesures, estimée à partir du pH mesuré sur 5 échantillons répartis systématiquement sur la placette, est égale à 0.30 unité. Cette REQM entre pH mesurés atteint 0.47 en milieu calcaire en Meuse. Elle a été estimée à partir du pH mesuré sur 4 échantillons répartis aux 4 points cardinaux à 15 m du centre des placettes en mesurant l'écart entre une série de mesure et la moyenne des 3 autres. Cette variabilité intra placette de 0.3 à 0.5 unité de pH explique une partie des différences observées entre pH mesuré sur une fosse et pH bioindiqué. La mesure est également réalisée à un moment donné alors que le relevé floristique, qui contient une forte proportion d'espèces pérennes, est représentatif d'une période plus longue. C'est en général à des informations représentatives d'un domaine spatial et d'un laps de temps assez vaste, que l'utilisateur cherche à avoir accès.

**Tableau 2.** Qualité d'estimation par la bioindication des variables liées aux capacités de nutrition du sol. Les données employées sont les 261 placettes du jeu de validation.

Variable	Régression logistique	
	REQM	R2
pH	0.81	0.61
S/T (%)	20.9	0.62
C/N	4.58	0.49
Moyenne		0.57

La qualité de prédiction des variables, lorsqu'elle est réalisée à l'échelle de la France, est comparable à celle réalisée dans les Vosges à partir de valeurs indicatrices établies pour ce massif (Gégout et al., 2003). A l'échelle internationale, des estimations réalisées par des valeurs indicatrices établies par des courbes de réponses n'ont été comparées à des valeurs mesurées que pour le pH (Wamelink et al., 2005). Selon les approches et les territoires étudiés, l'erreur quadratique moyenne de prédiction, qui varie de 0.6 à 2.0 unités, est cohérente avec les valeurs trouvées pour la France. La précision des prédictions obtenues par des valeurs indicatrices nationales reste à comparer par rapport à la précision obtenue par l'ajustement de valeurs indicatrices calibrées à l'échelle locale.

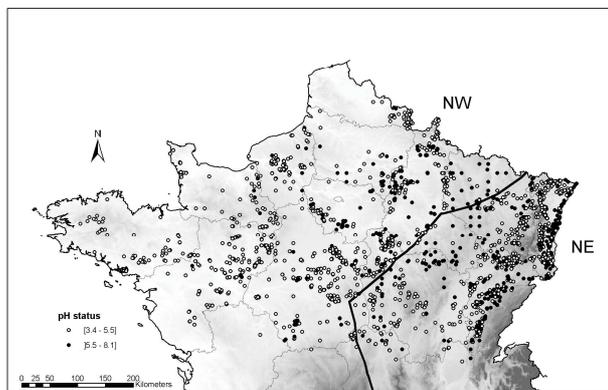
## 2.2 EFFICACITE DE VALEURS INDICATRICES DE PH ETABLIES A L'ECHELLE NATIONALE PAR RAPPORT A DES VALEURS INDICATRICES REGIONALES

### 2.2.1 Introduction

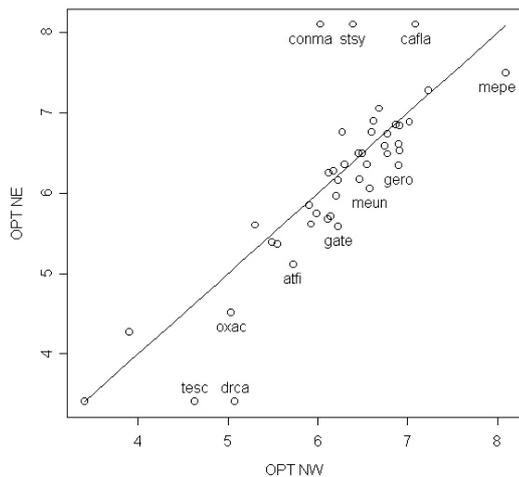
La variabilité spatiale de la réponse d'une espèce aux gradients écologiques peut être liée à sa variabilité génétique (présence d'écotypes), à la présence d'espèces compétitrices différentes selon les régions (Diekmann & Lawesson, 1999; Prinzing et al., 2002), ou à des valeurs différentes de co-facteurs importants pour l'espèce analysée et corrélés au gradient étudié. Cependant, l'existence et l'importance d'une variabilité géographique du comportement des espèces restent peu clairs et elles n'avaient jamais été étudiées avec des jeux de données conséquents et des informations floristiques et écologiques mesurées simultanément sur les mêmes sites. Si la variabilité spatiale de la réponse des espèces s'avérait importante, elle conduirait à une variabilité régionale de leur caractère indicateur. Cette variabilité aurait pour conséquence une détérioration des prédictions de l'estimation de la qualité des sols quand on passe d'une prédiction régionale à une prédiction nationale.

## 2.2.2 Comparaison des valeurs indicatrices définies dans le nord-est et le nord-ouest de la France

La variabilité du caractère indicateur des plantes a été testé en comparant les valeurs indicatrices déterminées dans le nord-est à partir de 1049 placettes situées à plus de 300 km des côtes et dans le nord-ouest de la France à partir de 1353 placettes situées à moins de 300 km des côtes (Coudun & Gegout, 2005). La localisation des placettes utilisées, issues de la base de données EcoPlant, est mentionnée dans la figure 5. La comparaison des valeurs indicatrices entre les deux régions montre que le caractère indicateur acidiphile ou neutrophile des espèces ne change pratiquement pas entre le nord-est et le nord-ouest de la France (figure 6). Le coefficient de corrélation ( $R^2$ ) entre les valeurs indicatrices des 46 espèces les plus fréquentes dans les régions atteint 0.78. Plus des  $\frac{3}{4}$  de ces espèces ont des différences de valeurs indicatrices inférieures à 0,5 unité de pH entre le nord-est et le nord-ouest de la France. Ces résultats indiquent une homogénéité globale des valeurs indicatrices à l'échelle du nord de la France qui semblent justifier le calcul du caractère bioindicateur des plantes à l'échelle de vastes espaces géographiques.



**Figure 5.** Localisation des 2404 relevés d'EcoPlant utilisés pour comparer le caractère indicateur des plantes entre le nord-est et le nord-ouest de la France. Les pH acides sont indiqués en noirs, les pH neutres sont représentés par un cercle blanc.



**Figure 6.** Valeurs indicatrices pour le pH de 46 espèces forestières dans le nord-est et le nord-ouest de la France.

### 2.2.3 Comparaison des prédictions réalisées avec des valeurs indicatrices régionales et des valeurs indicatrices nationales

Les valeurs indicatrices (VI) calculées à l'échelle nationale à partir de 3835 placettes réparties sur l'ensemble du pays, et celles calculées pour le nord-est et le nord-ouest de la France ont été utilisées pour réaliser des prédictions dans chaque région (tableau 3). Le  $R^2$  entre valeurs de pH prédites par la flore est très élevé : il dépasse 0.97 entre les pH prédits avec les VI calibrées régionalement et les pH prédits avec les VI calibrées au niveau national (tableau 3, ligne VI France). Cette très forte corrélation signifie une grande stabilité des valeurs prédites à partir du caractère indicateur de la flore indépendamment du fait que celui-ci ait été calibré au niveau régional ou au niveau national.

	Nb placet.		VI_France	VI_NE	VI_NO
VI France		R2	1.00	0.98	0.97
Predictions NE	1051	REQM	0.84	0.86	
		R2	0.62	0.63	
Predictions NO	1353	REQM	0.83		0.88
		R2	0.58		0.57

**Tableau 3.** lien entre valeurs indicatrices nationales et régionales et qualités des prédictions réalisées à partir de VI nationales et régionales

Les comparaisons entre pH prédits et mesurés montrent un écart (estimé par la racine de l'erreur quadratique moyenne ou REQM) légèrement plus faible lorsque les estimations sont réalisées avec les valeurs indicatrices nationales que lorsqu'elles sont réalisées avec les VI régionales et ceci pour les 2 régions. Le  $R^2$  entre estimations et mesures est semblable lorsque les prédictions sont réalisées avec les VI régionales ou avec les VI nationales.

Il n'y a donc pas de perte de qualité des prédictions de pH lorsqu'on passe de valeurs indicatrices régionales à des valeurs indicatrices nationales dans la comparaison qui a été réalisée. Il est probable que ces conclusions, établies pour le nord de la France, puissent être généralisées à la France entière.

## 2.3 VARIABILITE GEOGRAPHIQUE DE LA QUALITE DE PREDICTION DU pH EN FRANCE

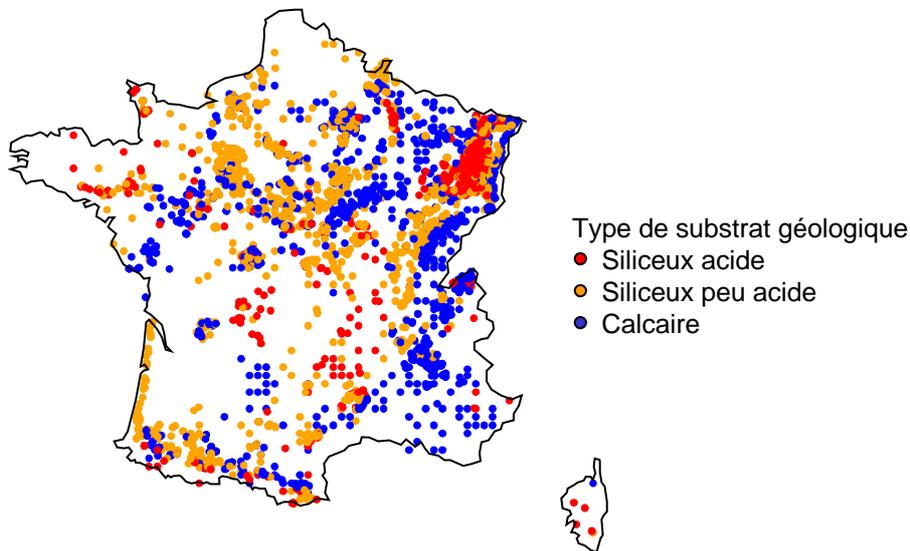
Les valeurs indicatrices déterminées à l'échelle nationale sont aussi efficaces que des valeurs indicatrices du nord-est ou du nord-ouest de la France pour estimer le pH dans ces deux régions. L'objet de cette nouvelle partie est d'observer les gradients de variation géographique des pH bioindiqués à l'échelle nationale et de vérifier que ces gradients correspondent aux gradients réels de variation des pH. Pour répondre à cet objectif, des modèles de variation spatiale des pH ont été établis en fonction de la longitude et de la latitude. Pour éviter des biais climatiques ou édaphiques, l'altitude et les substrats ont été pris en compte dans les modèles.

### 2.3.1 Matériel et méthodes

Les 3601 relevés d'EcoPlant contenant des mesures de pH réalisées en laboratoire ont été utilisés pour étudier la variation spatiale à large échelle des pH bioindiqués. Outre un relevé floristique complet et une mesure de pH, ces relevés sont localisés en longitude (X en coordonnées Lambert II) et en latitude (Y en coordonnées Lambert II). L'altitude et le substrat géologique de ces relevés sont également utilisés pour modéliser le pH. Les substrats géologiques étant très variés, il ont été regroupés en trois classes, « acide », « peu acide » et « carbonatés » selon les modalités suivantes :

- Substrats acides : roches cristallines grenues ou microgrenues, roches volcaniques massives acides (rhyolites, k eratophyres ...), roches m etamorphiques non carbonat ees non shisteuses (gneiss, migmatites), roches s edimentaires coh erentes siliceuses (gr es  a ciment non calcaire, gaize, flysch non calcaire). Cette classe contient 790 relev es avec des mesures de pH.
- Substrats peu acides : roches volcaniques massives non acides (basaltes, trachytes, and esites ...) et pyroclastiques (br eches, cendres), roches m etamorphiques non carbonat ees shisteuses, roches s edimentaires meubles siliceuses (argiles, limons, argiles  a silex, alluvions, colluvions, sables,  eboulis, pal eosols). Cette classe contient 1297 relev es.
- Substrats carbonat es : roches form ees ou contenant du calcaire. Cette classe contient 1147 placettes.

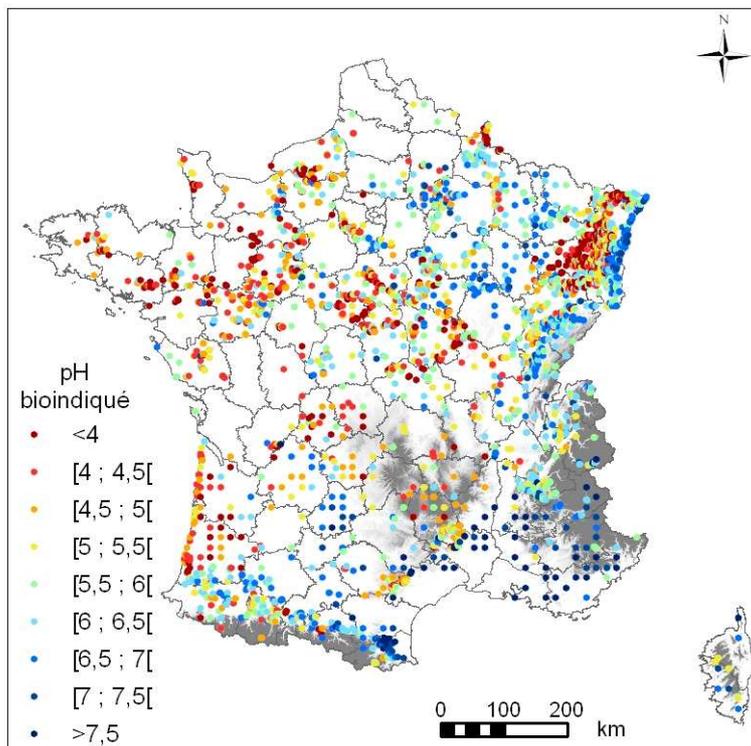
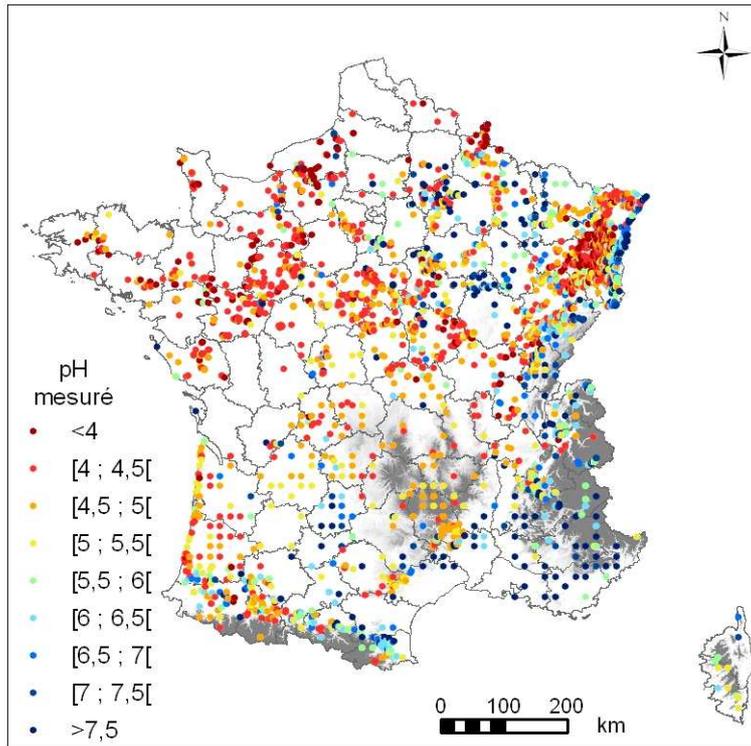
La distribution spatiale des substrats des placettes, indiqu ee dans la figure 7, met en  evidence les grands massifs siliceux acides (Bretagne, Vosges, Massif central) ainsi que les principaux territoires calcaires (Jura, essentiel de la zone m editerran eenne, et aur eoles des bassins parisiens et aquitains).



**Figure 7.** Carte de r epartition des substrats g eologique des placettes utilis ees pour analyser la qualit e de pr ediction en fonction du type de substrat siliceux acide (rouge), peu acide (beige) et calcaire (bleu).

Le pH de l'horizon A du sol a  et e bioindiqu e sur ces 3601 placettes  a l'aide des relev es floristiques en r ealisant la moyenne des valeurs indicatrices des esp eces pr esentes sur chaque placette (Figure 8).

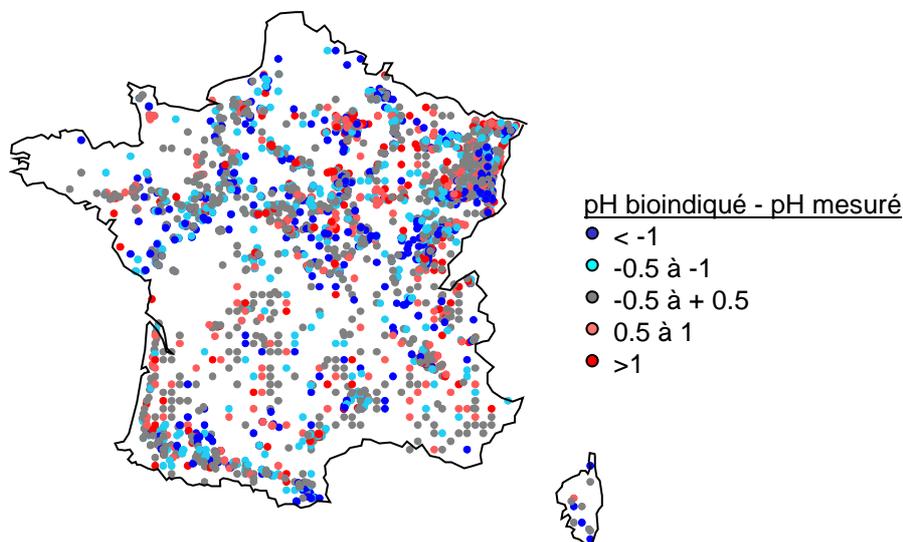
Les pH bioindiqu es et mesur es ont ensuite  et e mod eliss es pour chaque classe de substrats et pour tous les relev es par des mod eles incluant la longitude, la latitude et l'altitude de la placette comme variables explicatives. Les mod eles ont  et e r ealis es par r egression lin eaire multiple.



**Figure 8.** carte des pH de l'horizon A du sol de 3601 placettes forestières d'EcoPlant mesurés en laboratoire (haut) et bioindiqués par la flore des placettes (bas).

### 2.3.2 Résultats

La distribution spatiale des écarts entre valeurs bioindiquées et valeurs mesurées ne montre pas de structure particulière : les différences de pH négatives et positives sont à peu près également réparties sur l'ensemble du territoire Français (figure 9).



**Figure 9.** Ecarts entre pH bioindiqués par la flore des placettes et pH mesurés. (bleu : pH bioindiqué > pH mesuré, rouge : pH bioindiqué < pH mesuré, gris : écarts de pH < 0.5 unité)

Pour les pH mesurés comme pour les pH bioindiqués, il existe des gradients nationaux à substrat et altitude contrôlés (tableaux 4 et 5) :

- Les modèles montrent une augmentation de plus d'une unité de pH quand on va du nord vers le sud de la France. Cette augmentation concerne tous les substrats.
- Il existe également une augmentation moyenne des pH quand on se dirige de l'ouest vers l'est. Cette augmentation est nulle pour les substrats acides et elle augmente progressivement pour dépasser 1.5 unités de pH pour les substrats calcaires.

D'autre part, il existe à position géographique et substrat contrôlés, une baisse du pH avec l'augmentation d'altitude de l'ordre de 0.5 à 1 unité (tableau 4) de pH pour 1000 mètres.

Substrats	Cste	X (km)	Y (km)	Z (km)	R <sup>2</sup>	REQM	Nb sites
Acides (pH mesurés)	9.2	6.41E-05	-1.92E-03	-0.70	0.19	0.72	790
Peu acides (pH mesurés)	7.4	1.08E-03	-1.42E-03	-0.64	0.10	0.82	1297
Calcaire (pH mesurés)	9.2	3.26E-03	-2.33E-03	-0.61	0.19	1.24	1147
Tous substrats (pH mesurés)	9.3	2.08E-03	-2.40E-03	-0.83	0.12	1.19	3601
Acide (pH bioindiqués)	10.0	-6.28E-05	-2.25E-03	-0.42	0.14	1.02	789
Peu acides (pH bioindiqués)	8.0	1.69E-03	-1.67E-03	-1.05	0.12	0.94	1297
Calcaire (pH bioindiqués)	9.2	2.73E-03	-2.13E-03	-0.65	0.26	0.84	1147
Tous substrats (pH bioindiqués)	9.8	2.05E-03	-2.49E-03	-0.98	0.14	1.12	3600

**Tableau 4.** Modèles, à l'échelle de la France et par catégories de substrats géologiques, de prédiction des pH mesurés et bioindiqués par la longitude, la latitude et l'altitude. Les coefficients des régressions multiples sont présentés dans les 4 premières colonnes. Le R2 et la REQM des régressions sont présentés ensuite.

Tous les coefficients sont significatifs au seuil de 1/1000 sauf ceux liés à la variable X pour les pH mesurés et bioindiqués sur substrats acides.

A substrat et altitude égaux, les modèles mettent donc en évidence des pH minimums dans le nord-ouest de la France et maximum dans le sud-est, la différence pouvant dépasser 2 unités de pH.

Cette variation spatiale des pH à l'échelle de la France est comparable pour les pH bioindiqués et pour les pH mesurés (tableau 5) :

- les variables significatives sont les mêmes dans les modèles. L'effet de la longitude, non significatif avec les pH mesurés sur substrat acide ne l'est pas non plus avec les pH bioindiqués. Par contre, cette variable a un effet significatif dans les autres cas pour les pH mesurés et bioindiqués. Pour tous les autres cas, tous les effets sont significatifs tant pour les pH mesurés que pour les pH bioindiqués ;
- La variation de l'écart de pH d'un substrat à l'autre entre nord et sud de la France et entre est et ouest est cohérente entre valeurs mesurées et bioindiquées : les écarts sont forts dans les deux cas entre nord et sud pour les milieux acides et calcaires, ils sont plus faibles pour les substrats peu acides. De même l'écart s'accroît entre est et ouest quand on passe des milieux acides aux milieux calcaires et ce gradient peut être observé sur les mesures et les valeurs bioindiquées.
- Enfin les écarts moyens entre nord et sud et entre est et ouest sont comparables pour les variables mesurées et les variables bioindiquées. Sur les 8 écarts de pH liés à la géographie estimés, 6 ne diffèrent pas plus de 0.1 unités de pH entre valeurs mesurées et bioindiquées et deux diffèrent de 0.4 unité de pH. Les deux cas concernent le gradient est-ouest et dans un cas la valeur bioindiquée est plus élevée, dans l'autre, c'est la valeur mesurée.

Variation géographique	pH mesuré		pH bioindiqué	
	Nord->Sud	Ouest->Est	Nord->Sud	Ouest->Est
Substrats acides	1.2	0.0	1.3	0.0
Substrats peu acides	0.9	0.6	1.0	1.0
Substrats calcaire	1.4	2.0	1.3	1.6
Tous substrats	1.4	1.2	1.5	1.2

Valeur retenues pour les calculs : dX=dY=600km

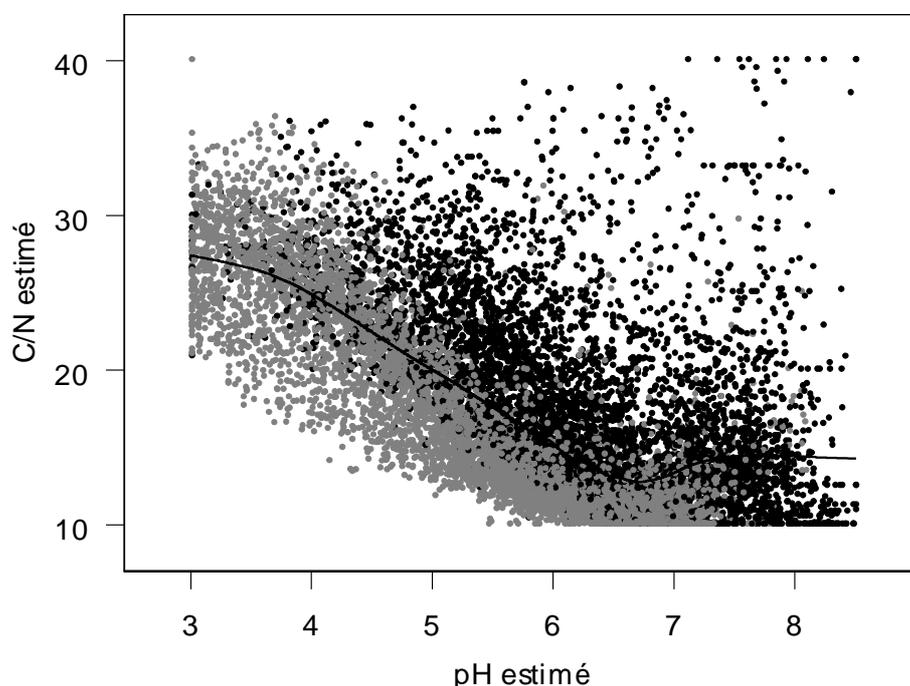
**Tableau 5.** Variation spatiale des pH à l'échelle de la France, à substrat et altitude fixés, selon les gradients nord-sud et est-ouest (les valeurs sont estimées avec les paramètres des équations du tableau 4).

Cette analyse montre qu'il est possible d'utiliser des valeurs indicatrices déterminées à l'échelle nationale pour prédire les pH forestiers du territoire français. L'erreur de prédiction est homogène sur l'ensemble du territoire français et il n'existe pas de biais de prédiction lié à la localisation en France.

## 2.4 UTILISATION DES VALEURS INDICATRICES POUR ESTIMER LA QUALITE NUTRITIONNELLE DES SOLS

Les valeurs indicatrices ont été utilisées pour prédire les variables de milieu sur de nombreuses placettes comportant un relevé floristique. Un exemple de prédiction du pH et du C/N à partir des valeurs indicatrices est fourni dans la figure 10 pour plus de 12 000 placettes réparties dans toute la France et issues des bases de données Sophy et EcoPlant. Ce nombre de placettes est quatre fois plus important que le nombre de placettes forestières comportant les mesures de ces deux variables. La figure 10 montre la relation bien connue entre les deux variables : le pH est corrélé négativement au C/N. Ce gradient, opposant les sols acides à mauvaise nutrition azotée aux sols peu acides à bonne nutrition azotée, correspond au gradient de richesse minérale connu en écologie forestière. C'est le gradient de niveau trophique mentionné dans la

Flore Forestière Française (Rameau et al., 1989; Rameau et al., 1993) où il est explicité par les humus et l'acidité du sol.



**Figure 10.** Représentation des pH et C/N estimés sur 12 453 placettes issues des bases de données Sophy et EcoPlant. Les placettes indiquées en grisé sont situées au nord de la France, celles en noir sont situées au sud de la France. Une courbe de tendance est superposée au nuage.

Le grand nombre de placettes met également en évidence un optimum de nutrition azotée (représentée par une valeur minimale de C/N prédit) pour les pH compris entre 6 et 7. Ce lien, logique, n'avait jamais été représenté. La remontée moyenne du C/N pour les sols carbonatés et surtout la possibilité d'avoir, pour des pH élevés, des C/N très élevés, est également très bien mis en évidence par la bioindication floristique. Les sols concernés sont très mal connus des pédologues ce qui est bien montré par la nomenclature des humus qui est mal établie pour ces conditions. Selon les auteurs on parle de mor ou moder calcique ou carbonaté, d'amphimull, de tangel voire de mulls à litière épaisse. Il est remarquable que la flore puisse indiquer ces conditions particulières de sol. Cet élément, ajouté à la relative faiblesse de la relation existant entre pH et C/N bioindiqués, montre l'utilité de distinguer le gradient de nutrition azotée du gradient d'acidité dans la définition d'un « niveau trophique » des placettes. Cette distinction est essentielle dans le cadre de l'évolution temporelle des écosystèmes sous l'effet des activités anthropiques qui peuvent conduire à un découplage entre niveaux d'azote et d'acidité des sols (Diekmann & Dupre, 1997). Ce découplage existe déjà au niveau spatial si on distingue les sols du nord et du sud de la France (figure 10). Ici encore, la flore permet d'illustrer un phénomène méconnu et non montré : non seulement les sols du nord de la France ont des pH plus acides et des C/N plus faibles en moyenne que les sols du sud de la France, mais à pH égal, la flore indique un C/N plus élevé de 1 à 5 unités (la différence varie selon les valeurs de pH) dans la moitié sud de la France par rapport à la moitié nord.



### **3 EVALUATION DE LA BAISSÉ DE QUALITÉ DES PREDICTIONS LIÉE À L'UTILISATION DE RELEVÉS FLORISTIQUES INCOMPLETS EN VUE DE L'UTILISATION DES DONNÉES DE L'INVENTAIRE FORESTIER NATIONAL**



### 3.1 INTRODUCTION

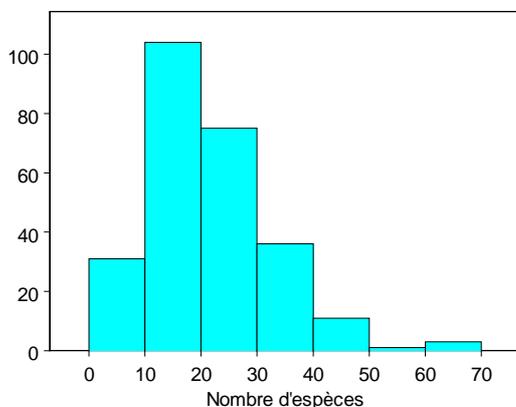
Les qualité de prédiction a toujours été évaluée à partir d'inventaires floristiques complets réalisés pendant la saison de végétation par des opérateurs compétents en botanique. Des problèmes de détermination des espèces, un temps limité de réalisation des relevés ou des inventaires effectués hors de la saison de végétation peuvent conduire à des relevés incomplets. L'utilisation de ces relevés pour estimer des variables écologiques est liée au lien entre qualité de prédiction et caractère complet de l'inventaire floristique réalisé sur les placettes. Cette problématique est d'autant plus importante que l'inventaire forestier national réalise des milliers de relevés floristiques annuellement en un temps limité et tout au long de l'année. La possibilité de prédire correctement les variables du sol avec des relevés incomplets ouvrirait la voie à l'utilisation de ces milliers de relevés.

La première voie employée pour analyser l'effet de relevés floristiques incomplets sur la qualité de prédiction a été d'utiliser une part variable des espèces de relevés exhaustifs pour réaliser les prédictions et d'observer comment cette part modifiait la qualité de prédiction. Cette voie a été abordée en observant la relation entre proportion d'espèces du relevé utilisée et qualité de prédiction d'une part et en observant la relation entre nombre d'espèces du relevé utilisé et qualité de prédiction d'autre part. L'utilisation d'espèces visibles tout au long de l'année a constitué la seconde approche pour estimer la possibilité de prédire les variables du sol avec des relevés incomplets.

### 3.2 MATERIEL ET METHODES

Le jeu de 261 relevés du réseau européen (Badeau 1998, figure 4), répartis en quinconce sur une grille de 16 km de côté placée sur la France entière, a été utilisé pour estimer la qualité des prédictions réalisées par les valeurs indicatrices.

Le nombre d'espèces contenues dans ces relevés varie de 1 à 61 avec une majorité de placettes contenant entre 10 et 20 espèces (figure 11). Cette forte variabilité conduit à différencier fortement les approches selon qu'on sélectionne une proportion ou un nombre d'espèces de chaque relevé pour analyser l'effet de relevés incomplets sur la qualité de prédiction.



**Figure 11.** Histogramme des nombres d'espèces des 261 relevés utilisés pour étudier l'effet de relevés incomplets sur la qualité de prédiction

La sélection d'une partie des espèces d'un relevé pour étudier l'effet de relevés partiels sur la qualité de prédiction pose le problème du choix de ces espèces qui va influencer sur le résultat.

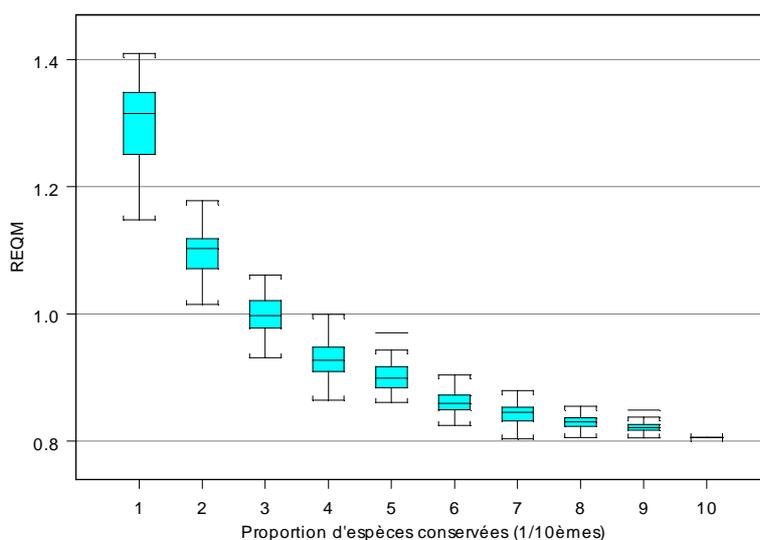
Pour pallier ce problème, un programme de sélection d'espèces et de prédiction des variables a été réalisé qui repose sur les principes suivants :

- sélection aléatoire des espèces au sein des relevés ;
- 50 répétitions du tirage aléatoire du nombre (ou de la proportion) d'espèces retenu au sein de chaque relevé ;
- prédiction de la variable pour chacun de ces 50 simulations de relevés partiels.

Cette procédure permet d'éliminer l'effet d'une sélection particulière sur la qualité de prédiction. Le pH est la variable du milieu utilisée pour analyser la qualité des prédictions en fonction de l'exhaustivité des relevés. La racine de l'erreur quadratique moyenne et le R<sup>2</sup> entre mesures et estimations ont été utilisés pour évaluer la qualité de prédiction du pH

### 3.3 QUALITE DES PREDICTIONS ET PROPORTION D'ESPECES DU RELEVÉ UTILISEES POUR LA BIOINDICATION DU PH

L'évolution de la qualité de prédiction mesurée par la REQM et le R<sup>2</sup> a été testée avec des proportions décroissantes d'espèces des relevés retenues pour réaliser les prédictions. Les proportions retenues varient de 100 à 10 % des espèces des relevés avec des estimations réalisées chaque 10 %. La croissance du R<sup>2</sup> et la décroissance de la REQM varient de façon exponentielle avec l'augmentation de la proportion d'espèces des relevés conservée (figures 12 et 13).

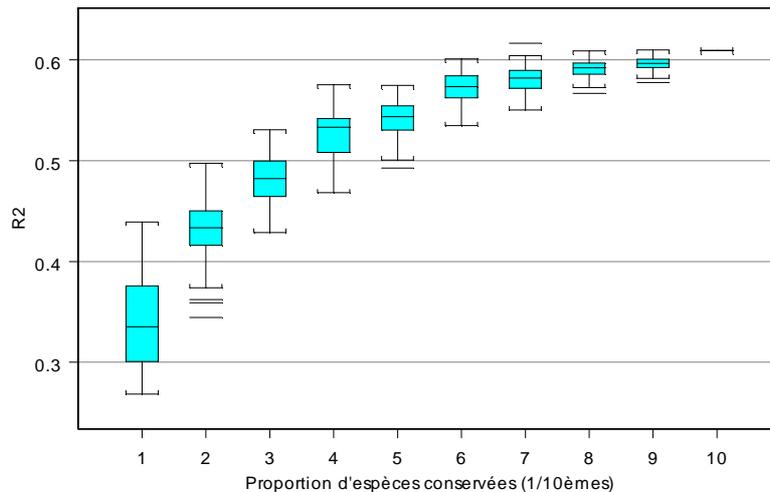


**Figure 12.** Variation de l'écart (REQM) entre pH mesuré et pH bioindiqué par la flore pour des proportions croissantes d'espèces des relevés prises en compte pour réaliser les prédictions.

Trois gammes de proportions peuvent être identifiées :

- Lorsque la proportion d'espèces conservées est inférieure au 1/3 des espèces contenues dans les relevés, la qualité de prédiction est dégradée. Le R<sup>2</sup> est inférieur à 0.5 et l'écart entre pH mesurés et estimés par la flore est supérieur à 1 ;
- Lorsque la proportion d'espèces conservées est comprise entre 1/3 et 2/3 du total des espèces des relevés, la qualité de prédiction est acceptable ;
- Au delà de 2/3 des espèces des relevés conservées, la qualité de prédiction est élevée et varie peu.

Ces résultats sont cohérents avec ceux de Ewald (2003) qui montre, à partir de 93 relevés effectués dans les Alpes bavaroises, une faible dégradation de la qualité de prédiction quand au moins 50 % des espèces des relevés sont conservées.

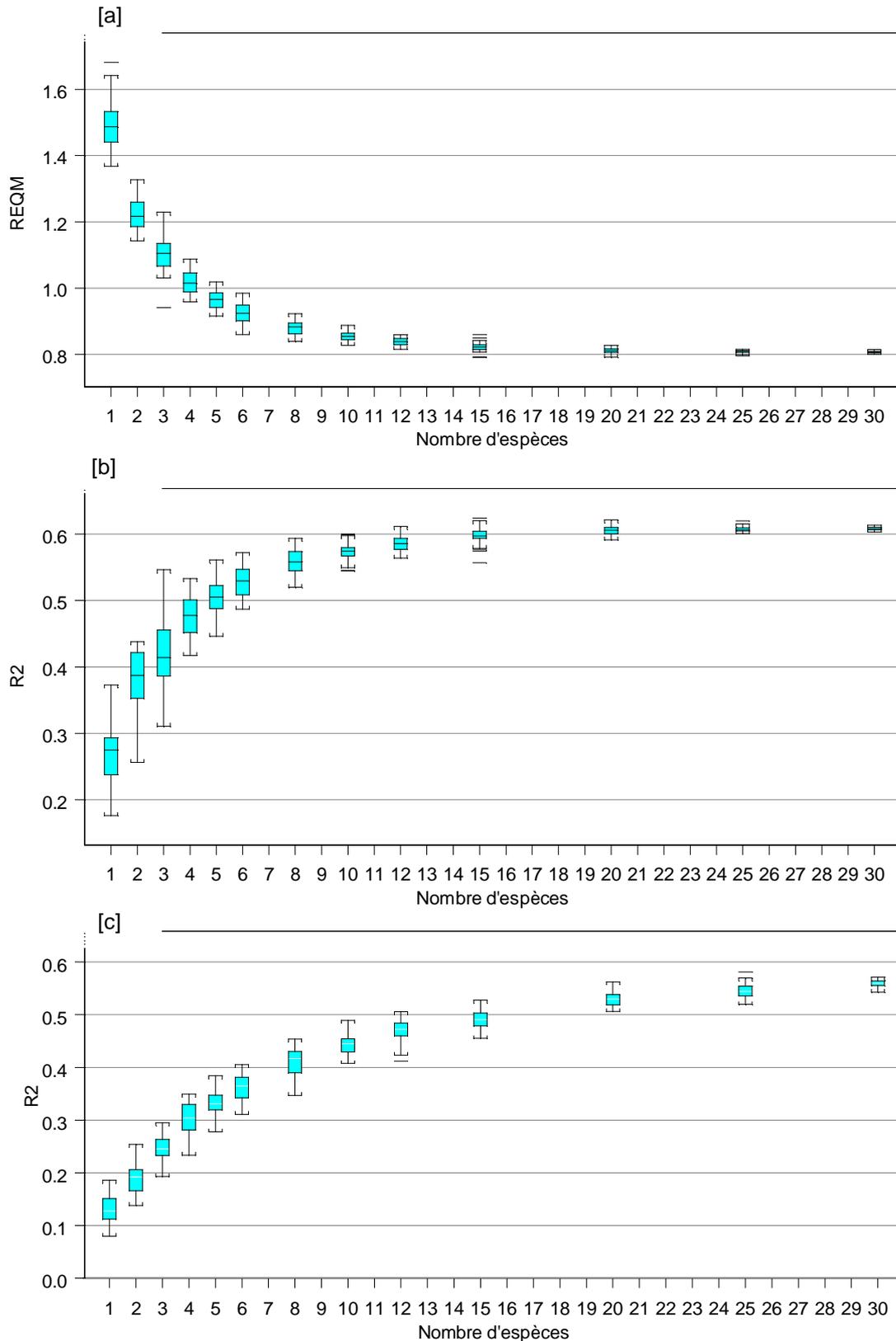


**Figure 13.** Variation de la corrélation ( $R^2$ ) entre pH mesurés et pH bioindiqués par la flore pour des proportions croissantes d'espèces des relevés prises en compte pour réaliser les prédictions.

Pour une même proportion d'espèces des relevés sélectionnées, la qualité de prédiction peut être très variable notamment pour les faibles proportions. Pour une proportion donnée, la variabilité des  $R^2$  est inférieure à 0.1 unité lorsque plus de 50% des espèces des relevés sont conservées. Ceci signifie que pour un même taux d'espèces conservées, la qualité de prédiction est variable lorsque le taux est faible alors qu'elle est constante lorsque le taux est élevé. Il faut noter que les gammes de  $R^2$  ou de REQM observées pour chaque proportion étudiée, décroissent en partie du fait d'un artefact méthodologique : lors des tirages aléatoires, les espèces sélectionnées sont d'autant plus communes d'un tirage à l'autre (ce qui implique des valeurs bioindiquées obligatoirement proche) que la proportion d'espèces choisie est importante.

### 3.4 QUALITE DES PREDICTIONS ET NOMBRE D'ESPECES DU RELEVÉ UTILISEES POUR LA PREDICTION DU PH

La qualité de prédiction augmente progressivement avec le nombre d'espèces prises en compte pour prédire le pH (Figure 14a et b). L'augmentation est très rapide lorsque 1 à 6 espèces sont utilisées, elle augmente ensuite modérément entre 6 et 12 espèces. Elle augmente peu au delà de 12 espèces.

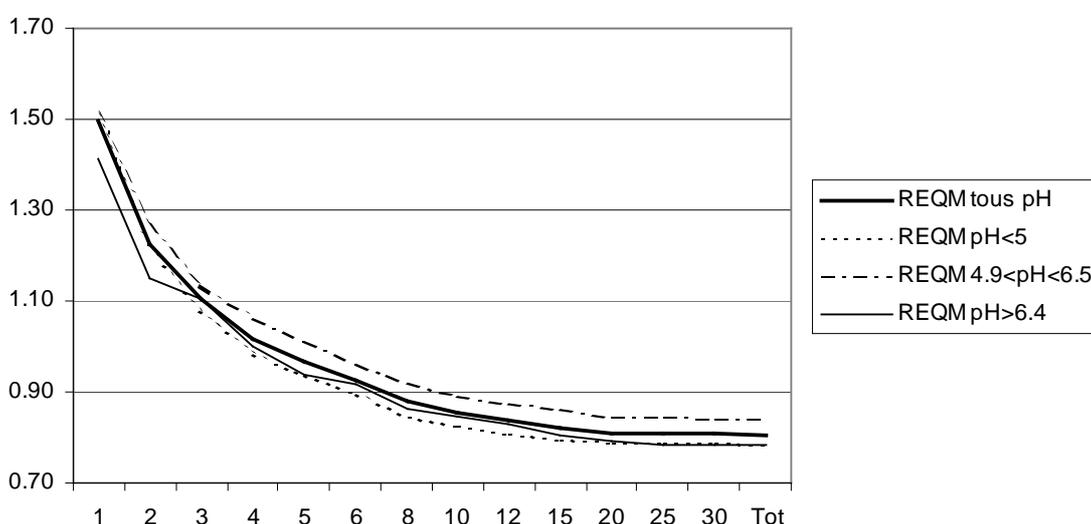


**Figure 14.** Evolution de la REQM [a] et du R2 [b, c] entre pH mesurés et bioindiqués avec le nombre d'espèces du relevé utilisées pour réaliser la prédiction. Les fig. 14a et b sont réalisés avec les 261 relevés de validation. Le graphe 14c est réalisé avec les 575 relevés d'EcoPlant comprenant au moins 30 espèces.

L'augmentation de qualité avec le nombre d'espèces prises en compte est cependant limitée par le nombre d'espèces contenues dans les relevés. De nombreux relevés comprennent en

effet de 10 à 20 espèces ce qui conduit à une stabilité de la qualité de prédiction pour les relevés concernés dès que ce nombre d'espèces est atteint. Pour éviter ce biais, le test a été recommencé sur les 575 relevés d'EcoPlant qui contiennent 30 espèces et plus (figure 14c). Ce jeu de placettes possède l'inconvénient de ne pas être représentatif des relevés présents en France et d'avoir été utilisé pour caculer la valeur indicatrice des espèces, mais il est peu biaisé pour observer la progression de la qualité des prédictions avec le nombre d'espèces utilisées pour les réaliser. L'augmentation de qualité de prédiction observée avec ce jeu suit une trajectoire comparable à celle observée avec les 261 placettes représentatives des forêts françaises. La qualité de prédiction se stabilise pour un nombre d'espèces un peu plus élevé (entre 15 et 20) que celui observé précédemment.

Le nombre d'espèces dépendant du niveau d'acidité (14 espèces en moyenne à pH 4, 26 à pH 7), la variation de la qualité de prédiction avec le nombre d'espèces a été observée en fonction du niveau d'acidité du milieu (figure 15) afin de voir si la fonction de croissance de qualité observée avec le nombre d'espèces était identique pour les milieux acides, neutres ou basiques.



**Figure 15.** Evolution, selon le niveau d'acidité mesurée du sol, de la REQM entre pH mesurés et pH prédits avec un nombre croissant d'espèces du relevé prises en compte (Tot : totalité des espèces prises en compte). Les valeurs de REQM présentées pour chaque nombre d'espèces étudié correspondent à une moyenne de 50 prédictions réalisées après une sélection aléatoire des espèces.

La croissance de qualité de prédiction avec le nombre d'espèces du relevé prises en compte s'accroît de façon presque identique sur milieux acides ou basiques, ce qui permet de ne pas avoir à considérer différents milieux pour fixer des seuils de nombre d'espèces à retenir pour obtenir un niveau de qualité minimum.

En conclusion, l'augmentation de qualité de prédiction observée avec le nombre d'espèces est très cohérente avec celle observée avec la proportion d'espèces du relevé utilisées pour réaliser les prédictions. Cette augmentation de qualité est très forte pour les faibles nombres (1 à 6) ou proportions (0 à 30 %) d'espèces utilisées pour réaliser les prédictions. Elle diminue progressivement pour devenir faible à nulle au delà de 60 % ou une quinzaine d'espèces présentes dans les relevés. Cette règle est valide quel que soit le niveau d'acidité du sol.

### 3.5 POSSIBILITE DE PREDICTION TOUTE L'ANNEE DES VARIABLES DU SOL AVEC LA FLORE

En hiver le tapis végétal est beaucoup moins riche en espèces que le reste de l'année, cependant de nombreuses espèces continuent à être identifiables. Afin d'évaluer les possibilités de prédiction toute l'année, il est nécessaire de tester la qualité des prédictions réalisées à partir de relevés floristiques hivernaux. Ne disposant pas de relevés phytoécologiques réalisés en hiver, nous utilisons les 261 relevés représentatifs des forêts françaises n'ayant pas été utilisés pour calibrer les valeurs indicatrices. Pour ces relevés, seules les espèces identifiables en hiver doivent être conservées afin de simuler des conditions hivernales. Il est pour cela nécessaire de déterminer les espèces identifiables en hiver au sein des 869 espèces indicatrices. Dans leurs travaux sur la valeur indicatrice des espèces, Ellenberg et Landolt donnent des indications sur le type biologique des espèces, Ellenberg précise en plus la persistance des feuilles au cours de l'année. Ces informations concernent les flores d'Allemagne et de Suisse dont beaucoup d'espèces sont communes à celles de la flore française. Ces informations ont été utilisées et complétées à l'aide de la flore forestière française pour les espèces non traitées par Ellenberg ou Landolt afin de déterminer les espèces pouvant être déterminées en hiver.

#### 3.5.1 Méthode de détermination de la flore hivernale

##### Utilisation des types biologiques

Ellenberg et Landolt utilisent des types biologiques proches de ceux définis par Raunkiaer qui peuvent être utilisés pour définir les espèces visibles en hiver. Ces types sont récapitulés dans le tableau 6.

Ellenberg	Landolt	Raunkiaer
Phanérophytes	Arbres sempervirents Arbres caducifoliés	PHANEROPHYTES: <i>plantes ligneuses dont les bourgeons se situent à plus de 50cm du sol.</i>
Nanophanérophytes	Arbustes sempervirents Arbustes caducifoliés	
Chaméphytes	Chaméphytes ligneux Chaméphytes herbacés	CHAMEPHYTES: <i>plantes vivaces dont les bourgeons se situent à moins de 50cm du sol.</i>
Hémicryptophytes	Hémicryptophytes	HEMICRYPTOPHYTES: <i>bourgeons de renouvellement situés au niveau du sol.</i>
Géophytes	Géophytes	GEOPHYTES: <i>organes de renouvellement enfouis dans le sol.</i>
Thérophytes	Thérophytes Thérophytes à graines peu persistantes	THEROPHYTES: <i>passent la saison défavorable sous forme de graines.</i>
-	Mousses	-

**Tableau 6.** types biologiques de Ellenberg et Landolt.

Les types biologiques sont basés sur la disposition par rapport au sol, des organes assurant la pérennité de la plante (bourgeons, bulbes,...) pendant la saison défavorable. On peut donc définir que les espèces dont les bourgeons sont en hiver situés au dessus du sol sont identifiables. Les espèces concernées sont les phanérophytes et les chaméphytes. On peut également identifier les mousses.

## Permanence des feuilles au cours de l'année

Il s'agit ensuite de déterminer au sein des géophytes, hémicryptophytes et thérophytes les espèces qui sont identifiables en hiver par leur feuilles.

Ellenberg propose quatre types de feuilles :

- Les feuilles sempervirentes ;
- Les feuilles persistantes l'hiver et remplacées chaque printemps ;
- Les feuilles vertes l'été ;
- Les feuilles vertes de printemps (espèces vernaies).

Les espèces dont les feuilles sont sempervirentes en hiver sont facilement identifiables toute l'année. Les espèces dont les feuilles sont persistantes peuvent être identifiées toute l'année, mais avec difficulté l'hiver. Les autres espèces ne sont pas identifiables en hiver.

### 3.5.2 Prédictions réalisées

Quatre séries de prédictions ont été réalisées avec les 248 placettes de validation du réseau Européen contenant un inventaire floristique complet effectué durant la période de végétation. Le type biologique de Raunkier et le caractère de persistance des feuilles ont été utilisés pour sélectionner les espèces à retenir dans les prédictions simulant différentes saisons de relevés (hiver ou saison de végétation) ou des niveaux variables de qualité du relevé floristique (relevé complet ou relevé rapide ; tableau 7).

Persistance feuille/ Type biologique	Année	Persistance hiver	Printemps	Été	Non renseign.	Total
PhanérophYTE	32	3		87	7	<b>129</b>
Chaméphyte	40	18		18	38	<b>114</b>
Liane ligneuse	1		1	9	1	<b>12</b>
Hémicryptophyte	11	114 <sup>+</sup>	1*	233*	69*	<b>428</b>
Géophyte	1	8 <sup>+</sup>	13*	47*	15*	<b>84</b>
Thérophyte		6 <sup>+</sup>	2*	16*	4*	<b>28</b>
Non renseigné	1	1 <sup>+</sup>			16*	<b>18</b>
Mousse	55 <sup>+</sup>					<b>55</b>
<b>Total</b>	<b>141</b>	<b>150</b>	<b>17</b>	<b>410</b>	<b>150</b>	<b>868</b>

**Tableau 7.** Synthèse des classes d'espèces utilisées pour simuler des prédictions au cours des différentes périodes de l'année et avec différents niveaux de qualité. Le nombre d'espèces appartenant à chaque classe est mentionné dans le tableau.

Les espèces utilisées dans les quatre simulations sont les suivantes : i - « relevés complets » : toutes les espèces ; ii - « relevés sans mousses » : toutes les espèces sauf les mousses ; iii - « relevés hivernaux complets » : toutes les espèces sauf celles notées d'une astérisque \* dont la case est grisée sombre ; iv - « relevés hivernaux aisés » : toutes les espèces sauf celles notées d'une astérisque ou d'un signe + dont les cases sont grisées sombre ou clair.

Ces prédictions, utilisent une partie des espèces des relevés de façon à simuler des relevés effectués par l'IFN en période de végétation et hors période de végétation avec deux niveaux de qualité :

- La première série de prédictions, appelée « Relevés complets » correspond à des relevés exhaustifs effectués pendant la saison de végétation. Elle utilise l'ensemble des espèces des relevés ;
- La deuxième série de prédictions utilise l'ensemble des végétaux vasculaires des relevés, c'est à dire toutes les espèces des relevés sauf les mousses. Cette série est dénommée « Relevés sans mousses » ;
- La troisième série de prédictions simule des relevés réalisés en période hivernale et utilise l'ensemble des espèces visibles à cette période. Les catégories d'espèces utilisées sont précisées dans le tableau 6. Ces catégories sont définies à partir du caractère ligneux et de

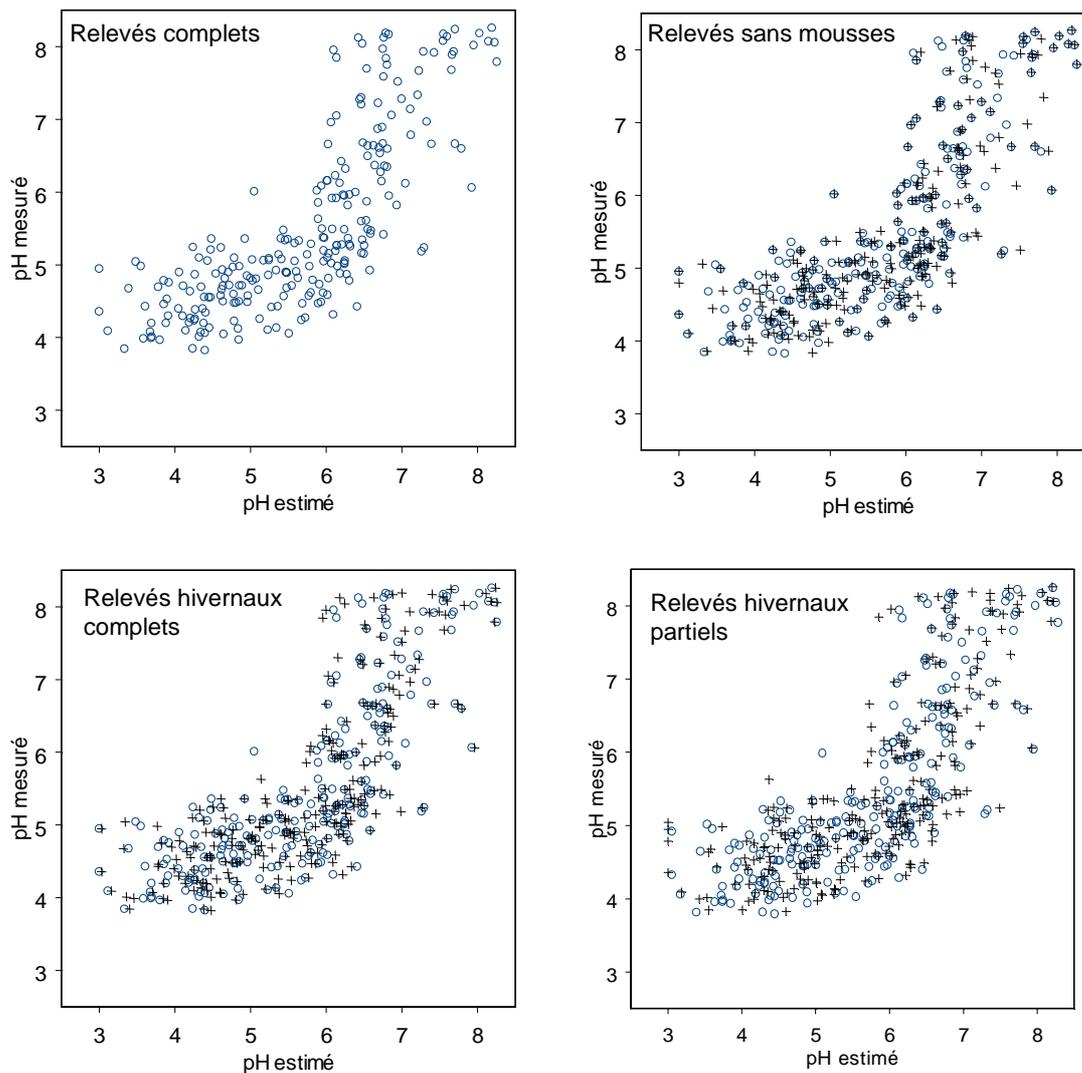
- la permanence des feuilles en hiver des végétaux. Cette simulation est dénommée « relevés hivernaux complets » ;
- La quatrième et dernière série de prédictions réalisées simule également des relevés hivernaux. Par rapport à la précédente, elle utilise seulement les espèces reconnaissables relativement facilement (cf. tableau 6) : les mousses et les herbacées persistantes mais non vertes sont éliminées en plus des espèces non visibles l'hiver. Cette simulation est dénommée « Relevés hivernaux aisés ».

### **3.5.3 Qualité des prédictions réalisées avec des relevés simulés hivernaux ou sans bryophytes**

La comparaison des valeurs mesurées et prédites avec les quatre simulations est indiquée dans la figure 16. Avec les relevés complets, les valeurs prédites sont plutôt inférieures aux valeurs mesurées pour les pH très acides. Elles ont ensuite plutôt tendance à être supérieures jusqu'à pH 6 pour enfin être à nouveau être inférieures en milieu calcaire. Cette caractéristique de la bioindication du pH par les valeurs indicatrices est générale que les valeurs indicatrices qu'elles proviennent d'EcoPlant (Gégout et al., 2003) ou d'Ellenberg (Schaffer et Sykora, 2000). Le problème est en partie méthodologique lié à la méthode de prédiction par une seule valeur synthétique du comportement de l'espèce plutôt que par la courbe de réponse entière (Gégout et al., 2003). Elle est pour partie également liée au fait que les relevés traduisent, à travers le pH bioindiqué, un changement de nutrition minérale non corrélé linéairement au pH. Les valeurs de pH bioindiquées sont ainsi plus corrélées au S/T (mesuré au pH du sol) de l'horizon A que ne l'est le pH mesuré.

Les valeurs prédites avec une partie seulement des espèces sont très semblables aux prédictions réalisées à partir des relevés complets et elles présentent les mêmes biais par rapport aux valeurs mesurées (figure 16). Le  $R^2$  entre les valeurs prédites avec des relevés complets et avec des relevés utilisant seulement 31 % d'espèces fréquentes représentant 53 % des occurrences dans les relevés est ainsi supérieur à 0.9 (tableau 8).

Les qualités de prédiction se dégradent légèrement lorsque les mousses ne sont pas utilisées pour réaliser les prédictions. Ces résultats sont globalement cohérents avec ceux de Berges (Aubert et al, 2007) observés sur des jeux de données régionaux ou locaux qui montrent un faible effet de la non utilisation des mousses. Les prédictions sont également très faiblement dégradées pour les simulations de relevés hivernaux complets. Ces résultats corroborent également ceux de Berges (Aubert et al, 2007) obtenus sur des jeux de données locaux et régionaux. Elles se dégradent de façon un peu plus importante pour les simulations de relevés hivernaux faits assez rapidement qui n'utiliseraient que les ligneux et les espèces toujours vertes (donc avec des feuilles vertes en hiver).



**Figure 16.** Relation entre pH mesurés sur les 261 placettes de validation et pH estimés par la flore. En haut à gauche, relevés complets. Dans les autres graphes, relevés partiels (croix noires) comparés aux relevés complets. Voir les espèces incluses dans chaque simulation dans le tableau 7.

Type de relevé simulé	Relevés complets	R. sans mousses	hivernaux complets	hivernaux aisés
Nombre d'espèces avec V.I.	868	813	452	268
Proportion des espèces (%)	100	94	52	31
Nombre moyen d'espèces par relevé	20.7	18.8	15.3	11.1
Proportion d'espèces conservées par relevé (%)	100	90	74	53
R2 avec pH estimés par les relevés complets	1	0.97	0.98	0.92
R2 avec pH mesurés	0.61	0.60	0.59	0.56
REQM	0.81	0.84	0.81	0.88

**Tableau 8.** Caractéristiques principales des prédictions de pH du sol (horizon A) réalisées avec des relevés floristiques complets et partiels simulant l'absence de bryophytes, des relevés exhaustifs réalisés en période hivernale et des relevés rapides réalisés en période hivernale.

### **3.6 CONCLUSIONS SUR LA POSSIBILITE D'UTILISER DES RELEVES FLORISTIQUES PARTIELS POUR REALISER LES PREDICTIONS DES VARIABLES DU SOL**

Une méthode robuste, qui repose sur la simulation de relevés partiels par tirage aléatoire d'espèces dans des relevés complets, a été mise au point pour étudier la possibilité d'utiliser des relevés floristiques incomplets pour bioindiquer le pH des sites forestiers. Les tests réalisés montrent que la conservation de 60 % des espèces des relevés ne conduit à pratiquement aucune dégradation de la qualité de la bioindication. Par contre, celle-ci se dégrade fortement lorsque le taux d'espèces conservées représente moins du 1/3 du nombre total d'espèces présentes sur le site. En nombre absolu d'espèces présentes dans les relevés, les prédictions de pH sont de faible qualité lorsque moins de 6 espèces sont utilisées, elles sont de qualité acceptable lorsque les prédictions sont réalisées avec 7 à 12 - 15 espèces. Au delà, la qualité de prédiction s'améliore peu avec le nombre d'espèces présentes dans les relevés. Ces résultats sont valides quel que soit le niveau d'acidité des sols.

La bonne qualité de prédiction des variables nutritionnelles du sol obtenue à partir de relevés floristiques incomplets est un élément très important pour une large utilisation de la méthode de bioindication par les valeurs indicatrices des espèces végétales. La réalisation d'un relevé complet sur 400 m<sup>2</sup> dure environ 1 heure, un bon nombre d'espèces étant trouvées rapidement, alors que les dernières trouvées sur la placette nécessitent un temps long de recherche. L'inventaire de 15 à 20 espèces permet de réaliser un gain de temps important sur la durée de réalisation du relevé qui peut être réduite à une durée comprise entre 10 et 30 minutes (plus en milieu acide à faible nombre d'espèces et moins en milieu à pH élevé riche en espèces). Ce temps réduit ouvre la voie à une utilisation possible de la méthode par les gestionnaires forestiers dans le cadre de la gestion courante.

Les résultats montrent également une faible dégradation de la qualité de la bioindication lorsque seule la végétation hivernale est utilisée pour prédire le pH. Ces résultats permettent d'étendre la période de travail à toute l'année avec une qualité de prédiction peu dégradée en hiver si les relevés sont réalisés soigneusement par des personnels ayant une bonne connaissance de la flore. La qualité de prédiction en hiver par la flore sera moyenne en cas de relevés réalisés par un non spécialiste.

Les relevés floristiques de l'Inventaire Forestier National sont réalisés toute l'année par des équipes qui passent un temps limité à les réaliser. Les tests effectués à partir des placettes d'EcoPlant montrent qu'il est possible d'utiliser la majorité de ces relevés pour prédire le pH des sols forestiers français.

**4 POSSIBILITE DE CARTOGRAPHIE DU PH DES SOLS FORESTIERS A PARTIR DES VALEURS INDICATRICES D'ECOPLANT ET DES RELEVES FLORISTIQUES DE L'INVENTAIRE FORESTIER NATIONAL.**

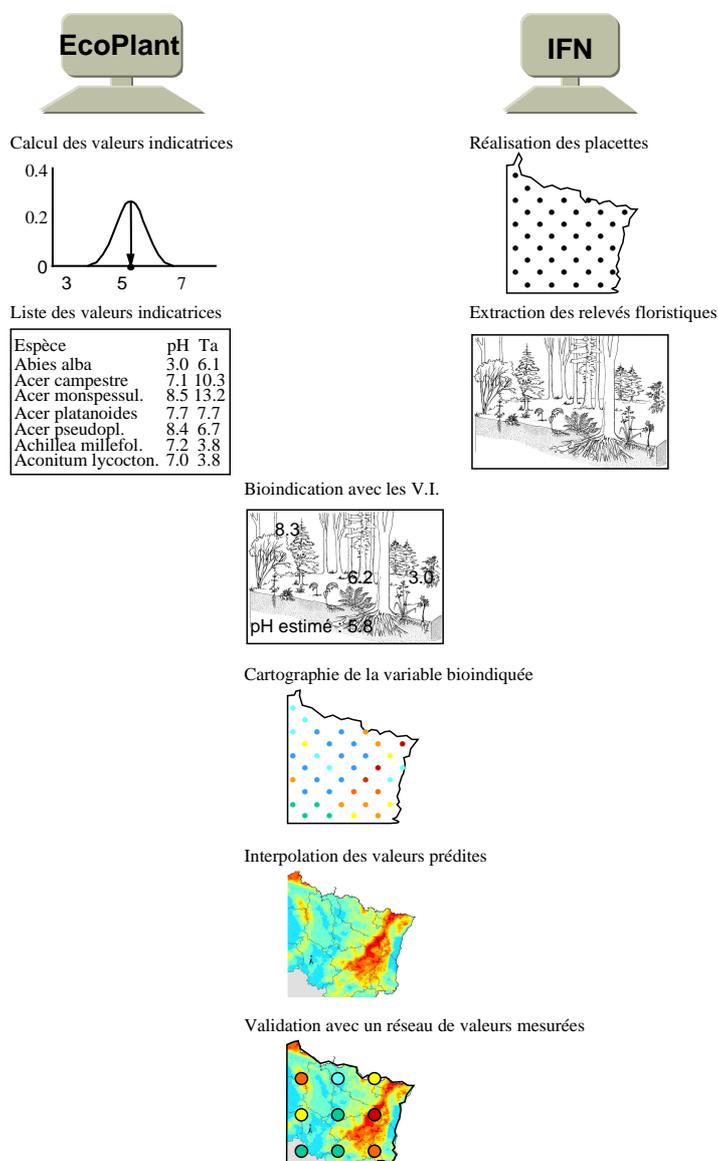


## 4.1 INTRODUCTION

L'ensemble du travail réalisé montre que la méthode des valeurs indicatrices est une approche efficace pour estimer les paramètres nutritionnels du sol à l'échelle des forêts françaises. Il démontre en outre que des inventaires floristiques incomplets ou réalisés à toute période de l'année peuvent être utilisés pour réaliser les prédictions des facteurs écologiques à partir de la flore. Il semble donc cohérent d'utiliser les relevés floristiques de l'IFN pour prédire le pH des sols forestiers français.

## 4.2 METHODE DE SPATIALISATION DU pH A L'ECHELLE DU TERRITOIRE FRANÇAIS

Le principe de la spatialisation des variables de milieu en utilisant conjointement les valeurs indicatrices issues d'EcoPlant et les données de l'IFN est simple (figure 17).



**Figure 17.** Méthode de spatialisation des variables bioindiquées à partir des valeurs indicatrices d'EcoPlant et des placettes de l'Inventaire Forestier National.

L'IFN échantillonne chaque année 7000 à 10000 placettes de façon systématique dans l'espace et dans le temps. Les placettes sont sélectionnées sur une grille kilométrique dont l'origine est tirée aléatoirement. L'échantillonnage, défini à l'échelle départementale jusqu'en 2004, est depuis 2005, défini à l'échelle nationale. L'information issue des placettes depuis 2005 peut être analysée annuellement à l'échelle du territoire français, ou par regroupement des placettes réalisées pendant 5 ou 10 ans pour obtenir des résultats plus précis ou valides à l'échelle départementale. Un relevé des principales conditions écologiques, des caractéristiques dendrométriques des peuplements et des espèces végétales présentes est effectué sur chaque placette. La prise de données floristiques et écologiques est effectuée par deux personnes sur un rayon de 15m autour du centre de la placette pendant une durée de l'ordre de 20 minutes. Les relevés étant réalisés tout au long de l'année, seuls ceux effectués pendant la saison de végétation contiennent les espèces herbacées ne subsistant pas l'hiver. Au total, 104 375 placettes régulièrement réparties sur le territoire français ont été réalisées par l'IFN entre 1989 et 2004 et ont été utilisées pour réaliser la carte des pH. L'ensemble de ces placettes correspond à une densité de 1 placette par km<sup>2</sup> de surface forestière environ.

La valeur des paramètres nutritionnels est bioindiquée, pour chaque placette de l'IFN, en réalisant la moyenne des valeurs indicatrices des espèces présentes sur le site. L'effet faible du caractère incomplet des relevés sur la précision d'estimation étant confirmée, on peut attendre une bonne précision de la carte obtenue. La spatialisation du pH des sols forestiers à l'échelle de l'ensemble du territoire est réalisée par interpolation des valeurs de pH estimées sur toutes les placettes à l'aide du krigeage qui constitue une méthode d'interpolation classique. Le jeu de données utilisé pour valider la carte, qui comprend 261 placettes régulièrement réparties sur le territoire français, est le même que celui utilisé pour la validation des valeurs indicatrices.

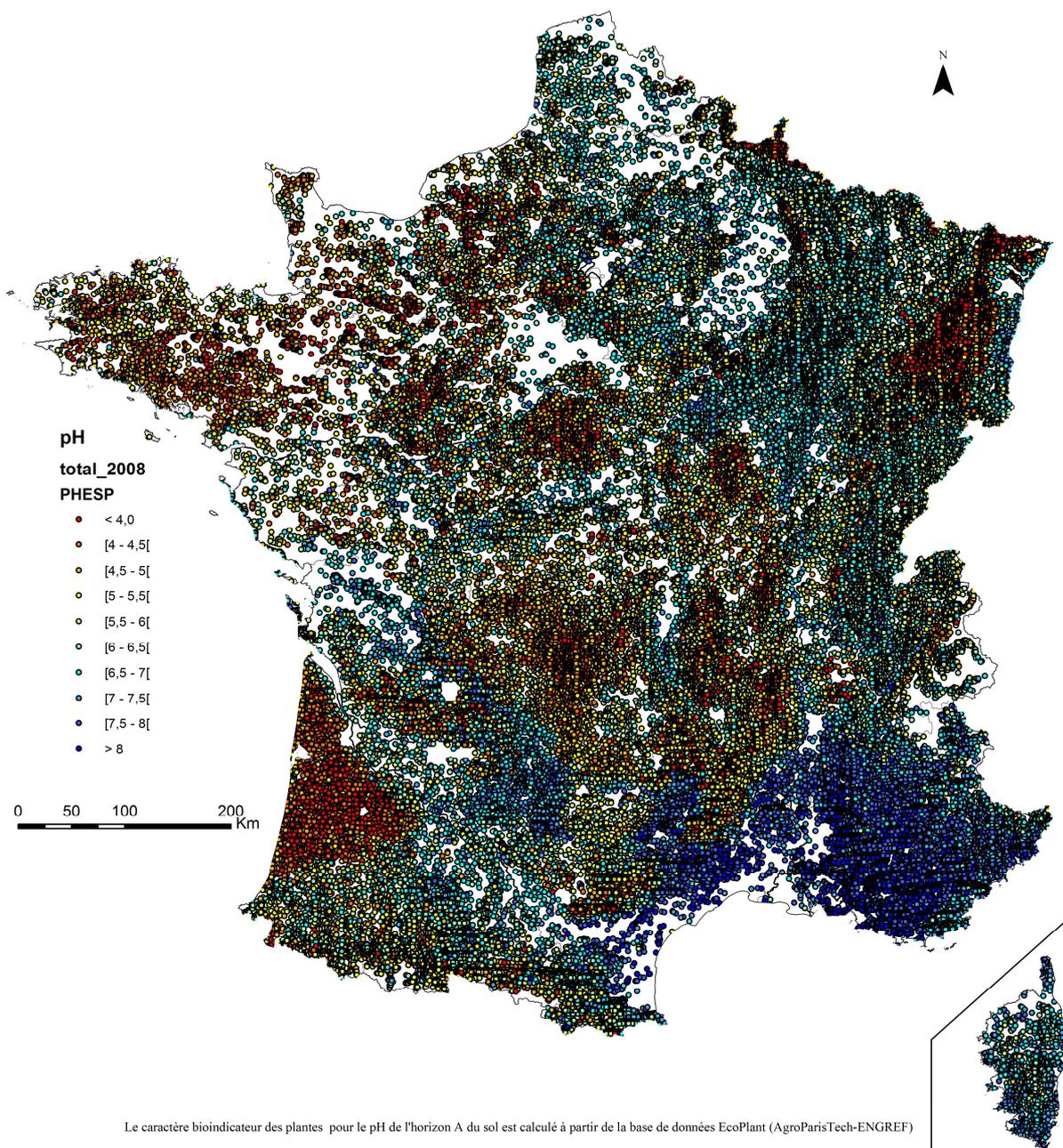
#### **4.3 RESULTATS : LA CARTE DES PH DES SOLS FORESTIERS FRANÇAIS**

##### **4.3.1 La carte des pH**

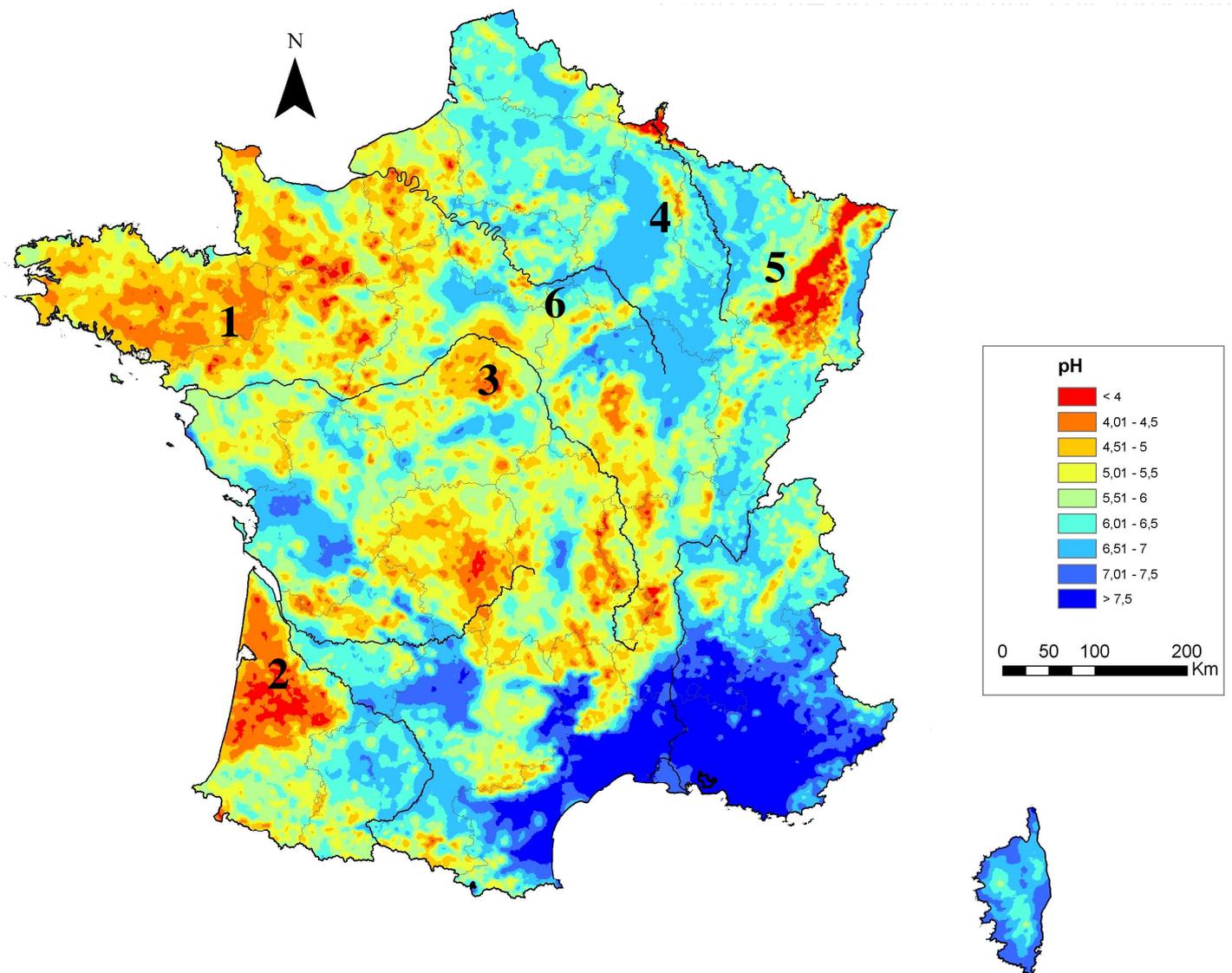
La représentation du pH des 104 375 placettes réalisées en France illustre l'énorme densité des placettes de l'IFN et la variabilité spatiale des pH estimés par la Flore (figure 18). La carte laisse clairement apparaître des régions siliceuses de France où les placettes ont des pH faibles (Bretagne, Sologne, Massif Vosgien, Massif central, Landes) ou les auréoles du bassins parisien à pH élevés.

La spatialisation des pH en tout point du territoire par la méthode du krigeage fournit une première carte qui montre la distribution spatiale des pH des sols forestiers français à l'échelle de l'ensemble du territoire national (figure 19). La validité des pH indiqués ne s'applique qu'aux espaces boisés qui sont non amendés et peu anthropisés contrairement aux espaces agricoles voisins. La carte obtenue est cependant cohérente avec la carte cantonale du pH des horizons de surfaces réalisée à partir d'analyses de sol de milieux agricoles distribués sur l'ensemble de la France (Saby et al., 2004).

Il est possible d'illustrer la précision spatiale de la carte obtenue par bioindication en observant qu'elle met en évidence des territoires à sols acides de surfaces importante –Bretagne (fig. 19 – 1), Landes (fig. 19 – 2), Sologne (fig. 19 – 3)- mais aussi des zones acides de faible surface comme faible l'Argonne (fig. 19 – 4), les alluvions de la Moselle (fig. 19 – 5) ou la forêt de Fontainebleau (fig. 19 – 6).



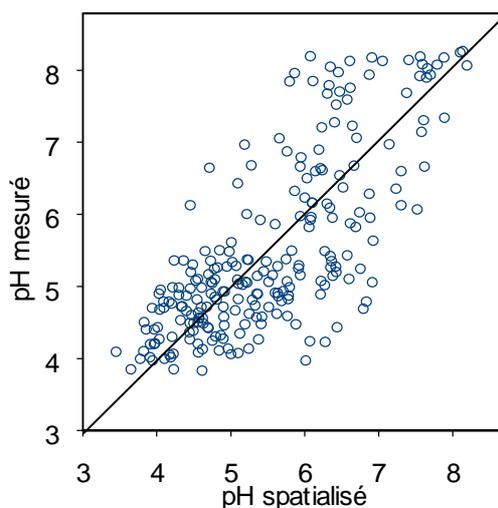
**Figure 18.** Représentation du pH estimé par les valeurs indicatrices d'EcoPlant sur les 104 375 placettes réalisées par l'Inventaire Forestier National en France entre 1989 et 2004.



**Figure 19.** Carte des pH de l'horizon A des sols forestiers français estimés à partir des valeurs indicatrices d'EcoPlant et des 104 375 placettes réalisées par l'Inventaire Forestier National en France entre 1989 et 2004.

### 4.3.2 Précision de la carte

La précision d'estimation du pH fourni par la carte a été analysée avec les 261 placettes du jeu de validation par la mise en relation du pH mesuré avec le pH spatialisé (figure 20). Le  $R^2$  entre les pH mesurés (pH eau de l'horizon A) et les pH spatialisés est égal à 0.58 alors qu'il était égal à 0.61 lorsqu'il était évalué directement avec la flore des placettes. La REQM est égale à 0.81 ce qui correspond au même écart entre pH mesuré et pH bioindiqué que lorsque le pH est estimé directement sur un site à partir de sa flore. Cette faible diminution du  $R^2$  signifie d'une part que les relevés floristiques de l'IFN semblent suffisants pour réaliser de bonnes estimations d'acidité des sols, et d'autre part, que la méthode d'estimation du pH des sols forestiers, à partir du pH des placettes environnantes, est une approche efficace.



**Figure 20.** Relation, pour les 261 placettes du jeu de validation, entre le pH mesuré et le pH spatialisé, obtenu par interpolation du pH bioindiqué sur les placettes de l'Inventaire Forestier National. La droite  $y=x$  est ajoutée au graphe.

Le  $R^2$  plus faible et la REQM identique de l'estimation par la carte, comparée à l'estimation ponctuelle, signifie que la moindre précision de l'estimation par la carte est compensée par une diminution du biais d'estimation. Cette qualité inattendue de la précision de la carte des pH provient du fait que les aléas de la composition floristique des placettes sont éliminés lorsque les pH bioindiqués sont moyennés à l'échelle paysagère à partir du pH des placettes environnantes. La forte densité de placettes de l'IFN, proche de 1 placette/km<sup>2</sup> en forêt permet en effet de réaliser des estimations en tout point du territoire à partir de plusieurs placettes proches.



## 5 CONCLUSION ET PERSPECTIVES

Le projet sur la validation de bioindicateurs floristiques pour la surveillance de l'état nutritionnel des sols a permis de répondre aux trois principaux objectifs de l'étude : i – la qualité de prédiction des paramètres de la nutrition minérale par les valeurs indicatrices des plantes a été déterminée. Le  $R^2$  entre valeurs mesurées et estimées est proche de 0.6 ; ii – la baisse de qualité de prédiction liée à l'utilisation de relevés incomplets a été estimée : elle devient importante lorsque moins de 6 espèces sont utilisées pour réaliser les prédictions. Cette valeur permet d'utiliser la plupart des relevés réalisés en hiver ; iii – l'état nutritionnel des sols a été cartographié à partir de la cartographie du pH. Au lieu de 80 000 placettes prévues initialement, ce sont plus de 100 000 placettes qui ont été utilisées pour réaliser cette cartographie.

La méthodologie employée pour réaliser la carte du pH est développée depuis très peu de temps. La précision obtenue pour cette spatialisation peut être améliorée selon trois voies. La première est celle de l'amélioration des méthodes de prédiction avec la flore. La connaissance de l'écologie d'un plus grand nombre d'espèces ou la prédiction avec des méthodes moins frustrées que celles des valeurs indicatrices pourrait améliorer légèrement la qualité de prédiction. La seconde voie est acquise avec la réalisation continue de nouveaux relevés floristiques par de multiples organismes et notamment l'IFN : la densification des points (7000 nouvelles placettes par an pour l'IFN) va conduire à une amélioration « mécanique » de la qualité des cartes réalisées. La troisième voie consiste à améliorer les méthodes d'interpolation utilisées. La modélisation des variables édaphiques bioindiquées par des variables topographiques et/ou géologiques associées à des techniques de co-krigeage est l'une des voies d'amélioration prometteuses de la spatialisation du pH des sols forestiers (Chaplot et al., 2000).

La méthodologie développée pour spatialiser le pH des sols peut être employée pour spatialiser les autres variables de sol et de climat pour lesquelles des valeurs indicatrices ont été calculées. Cependant la réalisation de ces cartes nécessite une étude préalable pour chaque variable de façon à estimer comme pour le pH, la précision des prédictions, l'absence de biais géographique et la possibilité d'utiliser des relevés incomplets. Les cartes de variables nutritionnelles produites devraient fournir une information nouvelle à l'échelle et avec la précision obtenue pour la carte du pH. Les cartes des variables climatiques produites par bioindication seront intéressantes en elles-mêmes mais aussi, et peut-être surtout, par comparaison avec les cartes produites par interpolation des variables mesurées. L'écart entre valeurs mesurées et bioindiquées pourrait prendre sens s'il était structuré spatialement. Ainsi, des différences régionales marquées pourraient signaler une erreur d'interpolation des variables mesurées ou indiquer une flore en déséquilibre avec le climat régional. Des cartes d'altitude bioindiquée pourraient permettre de définir des étages bioclimatiques équivalents d'un massif à l'autre : la délimitation géographique de ces étages reposerait sur une altitude bioindiquée unique sur l'ensemble du territoire français.

Un changement d'échelle et une cartographie à l'échelle d'une région naturelle ou du paysage avec une très forte densité de placettes (de l'ordre de 1 par ha ou dizaine d'ha) permettrait de spatialiser à fine résolution les variables sans nécessiter d'instrumentation très lourde de postes météorologiques ou de mesures de conditions de sols. Les cartes produites à cette échelle, complètement originales, pourraient montrer la variabilité des conditions de climat ou de sol avec le relief ou d'autres paramètres locaux. La production de telles cartes en quantité importante pourrait être réalisée par les gestionnaires des forêts principalement publiques lors

de la réalisation des aménagements forestiers. Un travail important, d'une part sur la validation de la méthode à l'échelle locale, et d'autre part sur l'optimisation de la durée de réalisation des relevés floristiques, est nécessaire pour une réelle mise en œuvre de cette cartographie par le gestionnaire. La pertinence de la cartographie des stations forestières, qui a toujours été réalisée à l'échelle locale, laisse augurer une bonne application à fine échelle de la méthode de cartographie quantitative développée à partir des valeurs indicatrices d'EcoPlant. La faible diminution de la précision de prédiction due à des relevés incomplets, comme les tentatives de prédiction directe des variables avec un petit nombre d'espèces par des méthodes originales (la régression par arbre et les forêts de Breiman, Pierrat, comm. pers.) laissent penser que des relevés floristiques réalisés très rapidement à l'échelle locale pourront rester efficaces pour prédire des variables du milieu.

Le champ de la cartographie des variables de milieu par la combinaison de la bioindication et des très grandes bases de données est un nouveau champ de recherche et d'application utile et prometteur. Il fournira des cartes des facteurs écologiques de base qui pourront être utilisées isolément ou comme des briques indispensables à la progression des connaissances dans différents domaines. Parmi ceux-ci, la modélisation prédictive de la distribution des espèces et la spatialisation des capacités de production des essences forestières sont deux voies qui ont été explorées et publiées dans plusieurs revues internationales (voir annexe).

Le couplage des valeurs indicatrices d'EcoPlant et des 7000 à 10000 relevés annuels de l'IFN ouvre la possibilité de mettre en place une surveillance des niveaux d'acidité et d'azote des sols spatialisée à l'échelle de la France. Des cartes réalisées avec 30 000 à 100 000 placettes pourraient être réalisées chaque 5 à 10 ans et diffusées auprès d'un large public. La comparaison de ces cartes d'une période à l'autre devrait mettre en évidence la réponse de la végétation forestière aux évolutions climatiques et nutritionnelles des forêts françaises. Cette évolution de la végétation, productrice d'énergie et élément moteur des cycles biogéochimiques, constitue un élément essentiel de l'évolution des écosystèmes. En ce sens, les valeurs indicatrices couplées aux données de l'Inventaire Forestier National constituent un outil de monitoring novateur et important pour évaluer l'évolution des conditions écologiques des écosystèmes forestiers français.

## Bibliographie

- Araujo, A.P., Pearson, R.G., Thuiller, W., & Erhard, M. (2006) Validation of species-climate impact models under climate change. *Journal of Biogeography*.
- Baize, D. & Girard, M., C." (1992) *Référentiel Pédologique, principaux sols d'Europe AFES*, INRA, Paris.
- Becker, M. & Le Goff, N. (1988) Diagnostic stationnel et potentiel de production. *Rev. For. Fr.*, 40, 29-43.
- Cajander, A.K. (1926) The theory of forest types. *Acta For. Fenn.*, 29, 1:108.
- Chaplot, V., Walter, C., Curmi, P., & Hollier-Larousse, A. (2000) The use of auxiliary geophysical data to improve a soil-landscape model. *Soil Science*, 165, 961-970.
- Coudun, C. & Gegout, J.C. (2005) Ecological behaviour of herbaceous forest species along a pH gradient: a comparison between oceanic and semicontinental regions in northern France. *Global Ecology and Biogeography*, 14, 263-270.
- de Boutray, A. (2003) Etude de la variabilité du pH et du C/N dans les sapinières vosgiennes, Ecole nationale du Génie Rural, des Eaux et des Forêts. Mémoire stage court de 3ème année de la Formation des Ingénieurs Forestiers.
- Diekmann, M. & Dupre, C. (1997) Acidification and eutrophication of deciduous forests in northwestern Germany demonstrated by indicator species analysis. *Journal of Vegetation Science*, 8, 855-864.
- Diekmann, M. & Lawesson, J.E. (1999) Shifts in ecological behaviour of herbaceous forest species along a transect from northern Central to North Europe. *Folia Geobotanica*, 34, 127-141.
- Duchaufour, P. (1948) Recherches écologiques sur la chênaie atlantique française. *Annales de l'école nationale des eaux et forêts*, 11, 1-332.
- Duchaufour, P. (1988) *Abrégé de Pédologie*.
- Dupre, C. & Diekmann, M. (1998) Prediction of occurrence of vascular plants in deciduous forests of south Sweden by means of Ellenberg indicator values. *Applied Vegetation Science*, 1, 139-150.
- Ellenberg, H., Weber, H.E., Düll, R., Wirth, V., Werner, W., & Paulißen, D. (1992) Zeigerwerte von Pflanzen in Mitteleuropa. In *Scripta geobotanica*, Vol. 18, pp. 1-248.
- Ewald, J. (2003) The sensitivity of Ellenberg indicator values to the completeness of vegetation relevés. *Basic and Applied Ecology*, 6, 507-513.
- Falkengren-Grerup, U. (1989) Soil acidification and its impact on ground vegetation. *Ambio*, 18, 179-183.
- Gégout, J.C., Coudun, C., Brisse, H., & Berges, L. (2002). Comportement écologique des espèces forestières vis-à-vis du climat et du sol en France. Application à l'évaluation des charges critiques d'acidité et d'azote. ENGREF, Laboratoire d'Etude des Ressources FORêt-Bois. Rapport final de la convention de recherche ADEME/ENGREF n°99 62 003, Nancy.
- Gégout, J.C., Hervé, J.C., Houllier, F., & Pierrat, J.C. (2003) Prediction of forest soil nutrient status using vegetation. *Journal of Vegetation Science*, 14, 55-62.
- Lahti, T. (1995) Understorey vegetation as an indicator of forest site potential in southern Finland. *Acta Forestalia Fennica*, 246, 68.
- Nieppola, J. (1993) Understorey plants as indicators of site productivity in *Pinus sylvestris* L. stands. *Scandinavian Journal of Forest Research*, 8, 49-65.

- Pinto, P.E. & Gegout, J.C. (2005) Assessing the nutritional and climatic response of temperate tree species in the Vosges Mountains. *Annals of Forest Science*, 62, 761-770.
- Prinzing, A., Durka, W., Klotz, S., & Brandl, R. (2002) Geographic variability of ecological niches of plant species: are competition and stress relevant? In *Ecography*, Vol. 25, pp. 721-729.
- Rameau, J.C., Mansion, D., & Dumé, G. (1989) Flore forestière française - guide écologique illustré - tome 1 : plaines et collines.
- Rameau, J.C., Mansion, D., & Dumé, G. (1993) Flore forestière française - guide écologique illustré - tome 2 : montagnes Institut pour le Développement, Paris.
- Saby, N., Schwartz, C., Walter, C., Arrouays, D., Lemerrier, B., Roland, N., & Squidant, H. (2004) Base de Données des Analyses de Terre : procédure de collecte et résultats de la campagne 1995-2000. *Etude et Gestion des Sols*, 11, 235-253.
- Seynave, I., Gégout, C., Hervé, J.C., & Dhote, J.F. (2006) Facteurs écologiques et productions du Hêtre en France. *Forêt entreprise*, 167, 41-45.
- Seynave, I., Gégout, J.C., Hervé, J.C., Dhote, J.F., Drapier, J., Bruno, E., & Dumé, G. (2005) Picea abies site index prediction by environmental factors and understorey vegetation: a two-scale approach based on survey databases. *Canadian Journal of Forest Research- Revue Canadienne De Recherche Forestiere*, 35, 1669-1678.
- Thimonier, A., Dupouey, J.L., Bost, F., & Becker, M. (1994) Simultaneous eutrophication and acidification of a forest ecosystem in North-East France. *New Phytologist*, 126, 533-539.
- Wamelink, G.W.W., Goedhart, P.W., Van Dobben, H.F., & Berendse, F. (2005) Plant species as predictors of soil pH: replacing expert judgement by hard data. In *Journal of Vegetation Science*.

**ANNEXE : ARTICLES SCIENTIFIQUES, PUBLIES ENTRE 2005 ET 2008, QUI UTILISENT LES DONNEES DE LA BASE DE DONNEES ECOPLANT OU SES VALEURS INDICATRICES.**

Articles utilisant les informations de la base EcoPlant

- Gégout J.C., Coudun C., Bailly G., Jabiol B. (2005). EcoPlant: A forest site database linking floristic data with soil and climate variables. *Journal of Vegetation Science*, 16 (2) :257-260.
- Coudun C., Gégout J.C. (2005). Ecological behaviour of herbaceous forest species along a pH gradient: a comparison between oceanic and semicontinental regions in northern France. *Global Ecology & Biogeography*, 14 (3) : 263-270.
- Coudun C., Gégout J.C. (2006). The derivation of species response curves with logistic regression is sensitive to sampling intensity and curve characteristics. *Ecological Modelling*, 199 : 164-175.
- Lenoir J., Gégout J.C., Marquet P.A., de Ruffray P., Brisse H., (2008). A Significant Upward Shift in Plant Species Optimum Elevation During the 20<sup>th</sup> Century. *Science*, 320, 1768-1771.

Articles utilisant les valeurs indicatrices d'EcoPlant liées à la nutrition du sol pour prédire la distribution des espèces

- Pinto P.E., Gégout J.C. (2005). Assessing the nutritional and climatic response of temperate forest tree species in the Vosges Mountains. *Annals of Forest Science*, 62 : 761-770.
- Coudun C., Gégout J.C., Piedallu C., Rameau J.C. (2006). Soil nutritional factors improve plant species distribution models: an illustration with *Acer campestre* L. in France. *Journal of Biogeography*, 33 : 1750-1763.
- Coudun C., Gégout J.C. (2007). Quantitative prediction of the distribution and abundance of *Vaccinium myrtillus* with climatic and edaphic factors. *Journal of Vegetation Science*, 18: 517-524.
- Marage D., Gégout J.C., (in press) Importance of soil nutrients in the distribution of forest communities on a large geographical scale. *Global Ecology and Biogeography*.
- Piedallu C., Gégout J.C. (2008) Assessment of efficient topographic solar radiation to improve large-scale plant distribution models. *Agricultural and forest meteorology*, 148, 1696-1706.

Articles utilisant les valeurs indicatrices d'EcoPlant liées à la nutrition du sol pour prédire la production des essences forestières

- Seynave I., Gégout J.C., Hervé J.C., Dhôte J.F., Drapier J., Bruno E., Dumé G. (2005). *Picea abies* site index prediction by environmental factors and understorey vegetation: a two-scale approach based on survey databases. *Canadian Journal of Forest Research*, 35(7) :1669-1678
- Seynave, I., Gégout J.C., Hervé J.C., Dhôte J.F. (2008) Is spatial distribution of *Fagus sylvatica* limited by its potential growth? *Journal of Biogeography*, 35, 1851-1862.

## Résumé

La base de données EcoPlant qui comprend plus de 5000 placettes possédant un inventaire floristique et des analyses de sol réalisées en laboratoire a été utilisée pour définir le caractère indicateur des espèces de la flore forestière française vis-à-vis du pH de surface du sol, du rapport C/N (indicateur de nutrition azotée) et du rapport S/T (indicateur de nutrition minérale et de toxicité aluminique). L'objectif de ce projet était : i - de déterminer l'efficacité du système indicateur créé pour prédire le pH, le C/N et le S/T de surface en tout site forestier du territoire français à partir de sa flore ; ii – d'évaluer si des relevés floristiques incomplets ou réalisés pendant la période hivernale restaient utilisables pour prédire les conditions nutritionnelles des sols ; iii – de déterminer si les valeurs indicatrices d'EcoPlant pouvaient être utilisées pour prédire et spatialiser les conditions de pH des sols forestiers français à partir des milliers de placettes réalisées par l'inventaire forestier national (IFN).

La qualité de prédiction par la flore a été déterminée à partir d'un jeu de validation de 261 placettes représentatives des forêts françaises et indépendantes des données de calibration du système bioindicateur. La racine de l'erreur quadratique moyenne (REQM) entre mesures et prédictions est égale à 0.81 unité pour le pH, 20.9 % pour le S/T et 4.58 unités pour le C/N. Les  $R^2$ , égaux respectivement à 0.61, 0.62 et 0.49 montrent une bonne capacité de prédiction des variables nutritionnelles par la flore. Une méthode par tirage aléatoire d'espèces dans des inventaires floristiques exhaustifs a été mise au point pour tester la qualité de prédiction des variables par des relevés incomplets. Les prédictions sont de qualité mauvaise avec des relevés comprenant moins de 6 espèces, assez bonne entre 6 et 12 espèces et très bonne au delà de 12 espèces. Ces seuils absolus sont valides pour tous les milieux qu'ils soient acides ou calcaires. La prédiction par bioindication du pH après élimination dans les relevés des espèces ne subsistant pas en hiver a permis de montrer qu'il est possible de réaliser des prédictions avec la flore hivernale.

La qualité des prédictions réalisées avec les valeurs indicatrices d'EcoPlant et la possibilité d'utiliser des relevés incomplets ou effectués en hiver ont rendu possible l'utilisation des relevés de l'IFN pour réaliser une carte du pH des sols forestiers français. Les valeurs indicatrices d'EcoPlant ont été utilisées pour prédire le pH sur 104 375 placettes de l'IFN régulièrement réparties en France. Une interpolation du pH des placettes par krigeage a été utilisée pour produire une estimation du pH en tout point du territoire et réaliser la carte du pH de surface des sols forestiers français au pas du  $\text{km}^2$ . La précision de cette carte a été testée avec les 261 placettes du jeu de validation. Le  $R^2$  entre le pH de la carte et le pH mesuré est égal à 0.58 et l'erreur moyenne est égale à 0.81. La précision de prédiction est donc similaire à celle obtenue ponctuellement avec le relevé floristique d'un site. La définition de la carte est importante puisqu'elle permet de visualiser des zones acides de petites taille comme l'Argonne, les alluvions de la Moselle ou la forêt de fontainebleau.

La méthodologie mise en place pour le pH doit pouvoir être utilisée, après une phase de validation, pour réaliser des cartes à haute résolution et sur l'ensemble de la France de la nutrition azotée. L'IFN produisant environ 7000 placettes par an, il sera probablement possible de réaliser des cartes des conditions nutritionnelles des sols par bioindication régulièrement dans le temps et de se servir de l'outil créé pour suivre à long terme l'impact des facteurs environnementaux sur les écosystèmes forestiers français.