UNIL | Université de Lausanne

Faculté des Géosciences et de l'Environnement

# Modélisation du pH du sol et son influence sur la distribution des espèces et des communautés végétales dans les Alpes vaudoises

Travail de Master ès Sciences en Biogéosciences

Coralie Boehler

**Sous la direction des Professeurs Pascal Vittoz et Antoine Guisan, et de Anne Dubuis, doctorante**
**Département d'Ecologie et Evolution**

Université de Lausanne - juin 2012

# Table des matières

**RÉSUMÉ** ..... 3
MOTS CLEFS ..... 3

ABSTRACT ..... 4
KEYWORDS ..... 4

**INTRODUCTION** ..... 5

**MODÉLISATION DU pH DU SOL** ..... 11
MATERIEL ET METHODES ..... 11
1. Site de l'étude ..... 11
2. Echantillonnage du pH ..... 12
3. Prélèvements sur le terrain ..... 12
4. Analyse des échantillons ..... 12
5. Recherche de prédicteurs pour le modèle du pH ..... 12
6. Elaboration du modèle pour le pH ..... 15
7. Evaluation du modèle ..... 15
8. Carte du pH ..... 16

RESULTATS ..... 16
1. Données de pH ..... 16
2. Données de géologie et utilisation du sol ..... 18
3. Prédicteurs conservés dans le modèle de pH ..... 21
4. Evaluation du modèle ..... 22
5. Carte du pH ..... 23

DISCUSSION ..... 25

**MODÉLISATION DE LA DISTRIBUTION DES ESPÈCES ET COMMUNAUTÉS VÉGÉTALES** ..... 30
MATERIEL ET METHODES ..... 30
1. Relevés phytosociologiques ..... 30
2. Prédicteurs topo-climatiques ..... 30
3. Données de pH ..... 31
4. Elaboration d'un jeu de données de calibration et d'un jeu de données indépendant ..... 31
5. Modèles de distribution d'espèces ..... 31
6. Pouvoir prédictif et importance des variables ..... 32
7. Comparaison de l'amélioration de l'AUC obtenue et des traits biologiques et valeurs écologiques de Landolt ..... 33
8. Comparaison des communautés prédites et observées ..... 33

RESULTATS ..... 35
1. Pouvoir prédictif et importance des variables ..... 35
2. Traits biologiques, valeurs de Landolt et AUC ..... 38
3. Comparaison des communautés prédites et observées ..... 39

DISCUSSION ..... 41

**CONCLUSION ET PERSPECTIVES** ..... 44

**REMERCIEMENTS** ..... 44

**REFERENCES BIBLIOGRAPHIQUES** ..... 45

**ANNEXES** ..... 50

# RÉSUMÉ

**But de l'étude:** En écologie, les modèles de distribution des espèces visent à prédire la répartition potentielle des plantes sur la base d'observations de présences des espèces et de variables liées à leur milieu. La plupart des études utilisent des variables topo-climatiques mais les données édaphiques sont rarement intégrées aux modèles. Le but de cette étude consiste d'abord à élaborer une carte du $pH_{H2O}$ à partir d'un modèle basé sur des échantillons prélevés sur le terrain et ensuite d'évaluer l'effet de l'intégration des données de pH tirées de cette nouvelle carte dans les modèles de distribution des espèces et des communautés végétales.
**Site de l'étude:** Alpes vaudoises (Préalpes vaudoises), Suisse
**Méthodes:** 429 données de $pH_{H2O}$ ont été obtenues après échantillonnage sur le terrain et analyse en laboratoire. Ces données ont été couplées à des variables influençant potentiellement l'évolution du sol lors de l'élaboration d'un modèle sous forme de modèle de régression linéaire généralisé (Generalised Linear Model : GLM) permettant d'obtenir des valeurs de pH pour chaque surface de $25m^2$ dans les Alpes vaudoises. Ces valeurs de pH ont ensuite été ajoutées en tant que prédicteur dans 4 types de modèles différents (GLM, Generalised Boosting Model : GBM, Generalised Additive Model : GAM et Random Forest : RF) permettant d'estimer la distribution des espèces végétales. Au total, 2 sets de 4 modèles sont créés pour chaque espèce, un set topo-climatique seul et un set où le pH est ajouté aux prédicteurs topo-climatiques. L'amélioration du pouvoir prédictif des modèles due au pH est testée pour chaque espèce, pour les espèces regroupées en fonctions de leurs caractéristiques écologiques et fonctionnelles ainsi que pour les communautés de plantes prédites en additionnant les modèles pour chaque espèce.
**Résultats:** Le modèle obtenu pour le pH du sol n'est pas assez précis et peine à prédire les valeurs extrêmes. Par la suite, l'ajout de ces valeurs de pH aux prédicteurs topo-climatiques pour la modélisation des plantes n'a pas engendré d'amélioration consistante, que ce soit au niveau des espèces individuelles ou de celles-ci regroupées en fonction de leurs caractéristiques écologiques et fonctionnelles. La prédiction de la composition des communautés n'est pas non plus améliorée par l'addition du pH comme prédicteur dans les modèles des espèces.
**Conclusions principales:** D'autres prédicteurs doivent être ajoutés au modèle du pH pour obtenir une carte plus proche de la réalité. Les valeurs obtenues dans cette étude ne sont pas assez précises et n'ont pas permis d'améliorer les modèles de distribution des espèces de plantes. Les prédictions obtenues pour les espèces et les communautés sont très proches de celles du modèle topo-climatique seul, le pH n'ayant amené aucune information supplémentaire. La carte obtenue permet cependant d'avoir une première idée de la distribution du pH dans les Alpes vaudoises.

## MOTS CLEFS

Communautés végétales ; données de présence/absence ; flore alpine ; GAM ; GBM ; GLM ; modèles de distribution des espèces ; pH ; Préalpes ; RF ; sol ; variables édaphiques ; variables topo-climatiques.

# ABSTRACT

**Aim:** In plant ecology, species distribution models (SDM) are used to predict the plants potential distribution. These models link species presences and absences with environmental variables and allow obtaining distribution maps for the studied species. Most studies use topo-climatic variables as predictors and edaphic data are seldom taken into account in the models. The main purpose of this study is to build up a map of $pH_{H2O}$ using samples taken from the field and to test if pH data from this map improve the models for the species distribution and communities' composition.
**Location:** Western Swiss Alps.
**Methods:** To predict pH, 429 $pH_{H2O}$ data were obtained from laboratory analysis and joined to variables reflecting the factors involved in soil development in a Generalised Linear Model (GLM). This model was then projected on the whole study area. The pH data from this map were added as a predictor in four different modelling techniques (GLM, Generalised Boosting Model: GBM, Generalised Additive Model: GAM and Random Forest: RF) to predict the plants species distribution. 2 sets of 4 models have been created for each species, one with topo-climatic predictors only and one with the pH predictor in addition to the topo-climatic ones. The amelioration in the model predictive power due to the pH is tested for each species and for the species grouped according to their ecological and functional characteristics. The quality of community composition predicted by stacking species distribution map without and with pH is also compared.
**Results:** The pH model was not accurate enough and was unable to predict extreme values. The addition of pH values originating from the map in the models for the species did not improve their predictive power in a consistent way. No clear tendency was found between the species ecological or functional characteristics and the change in models predictive power caused by pH. Finally, prediction of communities' composition was not ameliorated by adding pH in the models.
**Main conclusions:** Other predictors have to be added in the pH model in order to get a reliable map of pH. The pH data obtained from the model are not precise enough and did not improve the plants distribution models. The predictions for the species and communities are very close to the ones of the topo-climatic only models, involving that the pH data did not bring any supplementary information. The map of pH drawn in this study enables us anyway to have a first idea of the pH distribution in the western Swiss Alps.

## KEYWORDS

Edaphic variables ; Generalised Additive Models (GAM) ; Generalised Boosting Models (GBM) ; Generalised Linear Models (GLM) ; mountain flora ; pH ; presence/absence data ; Random Forest ; soil ; species communities ; Species Distribution Models (SDM) ; topo-climatic variables ; western Swiss Alps.

# INTRODUCTION

Connaître le fonctionnement de la Vie c'est espérer pouvoir mieux la préserver. La protection des espèces doit ainsi passer par une meilleure compréhension des conditions écologiques qui leur sont liées dans les écosystèmes afin de pouvoir réduire les perturbations qu'elles subissent. Ainsi, dans le domaine de l'écologie, la connaissance de la distribution des espèces végétales fournit de nombreux renseignements concernant les conditions écologiques qui leur sont favorables. Le concept de niche écologique, introduit par Hutchinson en 1957 (Hutchinson, 1957), s'applique aux conditions du milieu dans lesquelles une espèce peut se trouver. Plus précisément, la niche fondamentale concerne l'ensemble des conditions théoriquement favorables à une espèce. Dans la réalité cependant, les espèces ne sont pas seules mais au contraire en compétition les unes avec les autres, ce qui réduit le nombre d'habitats qui leur sont potentiellement propices. Ainsi, bien qu'un site puisse présenter les conditions favorables à une espèce, la présence d'autres espèces peut l'empêcher de s'y implanter. D'autres facteurs, historiques notamment, peuvent également diminuer le nombre de sites occupés. Seule une partie de la niche fondamentale présente ainsi réellement l'espèce en question, appelée niche réalisée.

La connaissance exhaustive des sites occupés par les espèces sur le terrain n'est évidemment pas concevable. Traditionnellement, les biologistes effectuent des relevés phytosociologiques leur permettant d'acquérir des informations sur les conditions environnementales favorables aux plantes ainsi que sur les associations entre espèces. Ces relevés donnent des renseignements ponctuels mais ne permettent pas d'avoir une idée de la composition de la végétation à l'échelle d'une région. Avec le développement des systèmes informatiques et statistiques, les données des relevés phytosociologiques ont pu être intégrées dans des processus de modélisation, permettant d'extrapoler des données pour des zones plus étendues. Les modèles de distribution des espèces (Species Distribution Models : SDM *;* Guisan et Zimmermann, 2000; Guisan et Thuiller, 2005) visent à prédire la répartition potentielle des espèces sur la base de données de présences des espèces et de données concernant leur milieu en s'appuyant sur le concept de la niche écologique. Les caractéristiques des sites où l'espèce a été observée sont ainsi relevées, permettant de prédire d'autres endroits présentant les mêmes conditions où l'espèce pourrait potentiellement être trouvée. On a ainsi une idée de la répartition potentielle de l'espèce, pouvant être cartographiée (Guisan et Zimmermann, 2000 ; Ferrier et Guisan, 2006). En plus des espèces individuelles, les modèles de distribution peuvent également être appliqués de plusieurs façons dans le cas des communautés (Ferrier et Guisan, 2006), ce qui permet de prédire la richesse spécifique et la composition de celles-ci.

Depuis leur apparition dans les années 1970, le nombre d'études utilisant les SDMs est en progression croissante (Guisan et Thuiller, 2005). Ils sont utilisés à des fins de recherche, par exemple dans le but de tester des hypothèses biogéographiques (Huntley et al., 2004 ; McPherson et al., 2004); de protection et de gestion des réserves naturelles (Pearce et Lindenmayer, 1998 ; Ferrier, 2002 ; Araùjo et al., 2004) et également dans le but de prédire la réponse des espèces face aux changements climatiques (Thomas et al., 2004 ; Thuiller, 2004).

La plupart des études utilisent des variables topo-climatiques pour qualifier les conditions environnementales et élaborer les modèles de prédiction des espèces. L'altitude, la pente, certaines caractéristiques climatiques telles que la température, les précipitations par exemple sont ainsi régulièrement employées et présentent de bons résultats (Fischer, 1990 ; Moore et al., 1991 ; Brzeziecki et al., 1993 ; Brown, 1994 ; Guisan et al., 1998 ; Guisan et al., 1999). Selon Guisan et Zimmermann (2000), le recours à des variables de ressources au sens de Austin (Austin, 1980) en plus des variables topo-climatiques pourrait cependant encore améliorer les modèles puisque ces variables ont une influence plus directe sur les plantes via leur nutrition. Les caractéristiques édaphiques notamment semblent prometteuses puisqu'elles conditionnent l'approvisionnement en eau et en nutriments des plantes (Gobat, Aragno, Matthey, 2010). Le sol est également l'habitat de nombreuses espèces de bactéries et champignons qui peuvent faciliter ou au contraire nuire aux végétaux. On sait par exemple que certaines bactéries appelées Plant Growth Promoting Rhizobacteria (PGPR; Kloepper et Schroth, 1978) stimulent la santé et la croissance des plantes (Glick, 1995) alors que d'autres en sont de parfaits parasites. Les champignons ont également des effets contradictoires : certains sont parasites et profitent de leur hôte alors que d'autres s'associent aux plantes en formant des mycorhizes, leur prodiguant de l'eau et des minéraux.

Jusqu'à dernièrement, peu d'études ont fait appel aux prédicteurs édaphiques pour envisager la niche potentielle des espèces. Les bases de données concernant les sols ne sont d'ailleurs souvent pas encore suffisamment précises et étendues pour permettre leur utilisation dans le cadre de la modélisation de distribution d'espèces. Dans leur article paru en 2005, Coudun et Gégout ont étudié la réponse écologique de plantes herbacées forestières face au pH de la partie supérieure du sol dans deux régions de France. Pour ce faire, ils ont analysé le $pH_{H2O}$ de 2'402 échantillons de sol puis ont construit un modèle de régression logistique pour 46 espèces. La réponse des espèces a été étudiée par leur optimum et leur amplitude écologiques. Dans chacune des deux régions étudiées, les espèces ont réagi au pH. Les optima et amplitudes des espèces en termes de pH étaient similaires pour ces régions, des valeurs de pH comprises entre 6 et 7 représentant l'optimum pour le plus grand nombre d'entre elles. Les auteurs ont ainsi conclu à une relative stabilité de l'évolution des espèces en termes de conditions d'acidité entre le nord-ouest et le nord-est de la France.

En 2006, Coudun et al. ont testé l'importance respective de prédicteurs climatiques et édaphiques sur la modélisation de la distribution de *Acer campestre* en France. Ils ont testé 150 prédicteurs climatiques et 6 variables édaphiques. Les variables finalement conservées dans leur modèle sont les précipitations en automne et l'indice d'évapotranspiration potentielle de Thornthwaite pour les prédicteurs climatiques et le $pH_{H2O}$ comme prédicteur édaphique. Selon leurs résultats, les prédictions pour *Acer campestre* ont été améliorées en ajoutant le prédicteur pH au modèle climatique, avec un succès de prédiction atteignant 71%.

Coudun et Gégout ont alors cherché en 2007 à prédire la répartition en termes d'abondance de *Vaccinium myrtillus* en France. Pour ce faire, ils ont utilisé 2'905 relevés d'abondance, 65 prédicteurs climatiques et 2 prédicteurs édaphiques élaborés à partir d'indices de composition végétale. Le modèle a été testé sur un jeu de données indépendant en

observant la différence entre valeurs prédites et valeurs observées. Le modèle final, comprenant la température moyenne annuelle, le pH et le rapport carbone/azote comme prédicteurs a donné de très bons résultats, avec un indice AUC de 0.9 environ, confirmant que l'utilisation de variables liées à la capacité nutritive du sol est importante dans les modèles de distribution des espèces.

Récemment, une autre tentative d'évaluation de la pertinence de différents indicateurs édaphiques dans les modèles de distribution d'espèces a été menée par Dubuis et al. (in revision) de l'Université de Lausanne. Les auteurs ont échantillonné 252 sites dans les Alpes vaudoises dont ils ont analysé le $pH_{H2O}$, l'azote total, le phosphore, la teneur en argile, en sable et en limon ainsi que le rapport carbone/azote. La contribution de ces variables à l'amélioration des prédictions par les modèles a ensuite été testée en utilisant 4 types de modèles différents. Au final, le pH et la teneur en azote se sont révélés des prédicteurs particulièrement pertinents puisqu'ils ont amélioré les prédictions pour 47% et 25% des espèces respectivement.

Au vu des résultats de ces études, le pH semble être un prédicteur intéressant pour permettre d'améliorer la qualité des modèles de distribution des espèces. Il reflète en effet la capacité d'échange cationique et le taux de saturation du sol, indiquant ainsi la capacité du sol à assurer la nutrition végétale (Gobat, Aragno, Matthey, 2010). Le pH est considéré comme une variable de ressource puisqu'il détermine l'assimilation des plantes en minéraux (Duchaufour, 1989) mais il peut aussi être considéré comme variable directe puisqu'il a un effet physiologique sur la croissance des plantes (Austin, 1980). Le pH détermine également la concentration du sol en métaux lourds, l'activité des microorganismes et de la faune du sol ainsi que la forme d'humus (Castrignano et al., 2011). Par conséquent, certaines espèces végétales étant plutôt calcicoles ou acidophiles, la teneur en protons mesurée par le pH du sol semble un déterminant important de la distribution de ces espèces.

L'intégration des données de pH dans les modèles de distribution d'espèces pose cependant un défi. En effet, comme pour les relevés phytosociologiques, les valeurs de pH résultant de prélèvements sur le terrain sont des données ponctuelles qui nécessitent d'être extrapolées avant de pouvoir être intégrées aux modèles pour une région entière. Pour produire une carte continue du pH du sol, différentes techniques ont été utilisées dans des études antérieures, dont le kriegage (Castrignano et al., 2011), l'Inverse Distance Weighting (IDW) et la méthode dite « splines » (Schloeder et al., 2001), une nouvelle méthode basée sur le théorème des surfaces dénommée High Accuracy Surface Modelling (HASM ; Shi et al., 2009 ; Shi et al., 2011) ainsi que la méthode de classification Random Forest (Wiesmeier et al., 2011). Les types de terrains concernés par ces études diffèrent par leur taille et leur amplitude altitudinale, si bien qu'aucune méthode généralisable ne s'en dégage.

Cette étude vise d'abord à élargir le nombre d'échantillons de pH dans les Alpes vaudoises pour faire suite au travail de Dubuis et al. (in revision) afin d'élaborer une carte du pH pour toute la zone d'étude grâce à un modèle. Les données de pH obtenues à l'aide du modèle sont ensuite ajoutées aux prédicteurs topo-climatiques couramment utilisés dans les modèles de distribution d'espèces afin de déterminer si ce prédicteur apporte une amélioration de leur pouvoir prédictif.

Ce travail est subdivisé en deux parties (Fig. 1) : une première partie concerne la modélisation du pH du sol. Des prédicteurs reflétant les facteurs influençant l'évolution du sol sont recherchés afin d'élaborer le meilleur modèle possible. Sur la base de ce modèle, une carte du pH pour toute la zone d'étude est obtenue. Etant donné que les Alpes vaudoises sont majoritairement calcaires, l'évolution du sol sur ce type de substrat devrait engendrer des sols plutôt acides en plaine et dans les fonds de vallées alors que ceux-ci devraient être plus basiques sur les sommets.

La deuxième partie de cette étude vise ensuite à modéliser la distribution des espèces en ajoutant les données de pH obtenues à partir de la carte du pH aux variables topo-climatiques généralement utilisées. L'amélioration des prédictions est testée par l'amélioration du pouvoir prédictif du modèle pour chaque espèce et pour les espèces regroupées en fonction de leurs traits fonctionnels et valeurs écologiques (Landolt, 2010). Finalement, les communautés prédites en cumulant les modèles des espèces sont comparées aux communautés présentes sur le terrain, telles que décrites dans les relevés phytosociologiques. Etant donné que le pH est un indicateur des conditions de nutrition végétale et qu'il a déjà été démontré comme étant un prédicteur important pour modéliser la distribution de certaines espèces (Coudun et al., 2006 ; Coudun et Gégout, 2007 ; Dubuis et al., in revision), les modèles devraient être améliorés par l'ajout de ce prédicteur. En conséquence, les communautés végétales prédites devraient être plus proches de celles observées sur le terrain que des prédictions faites sans incorporer le pH dans les modèles.

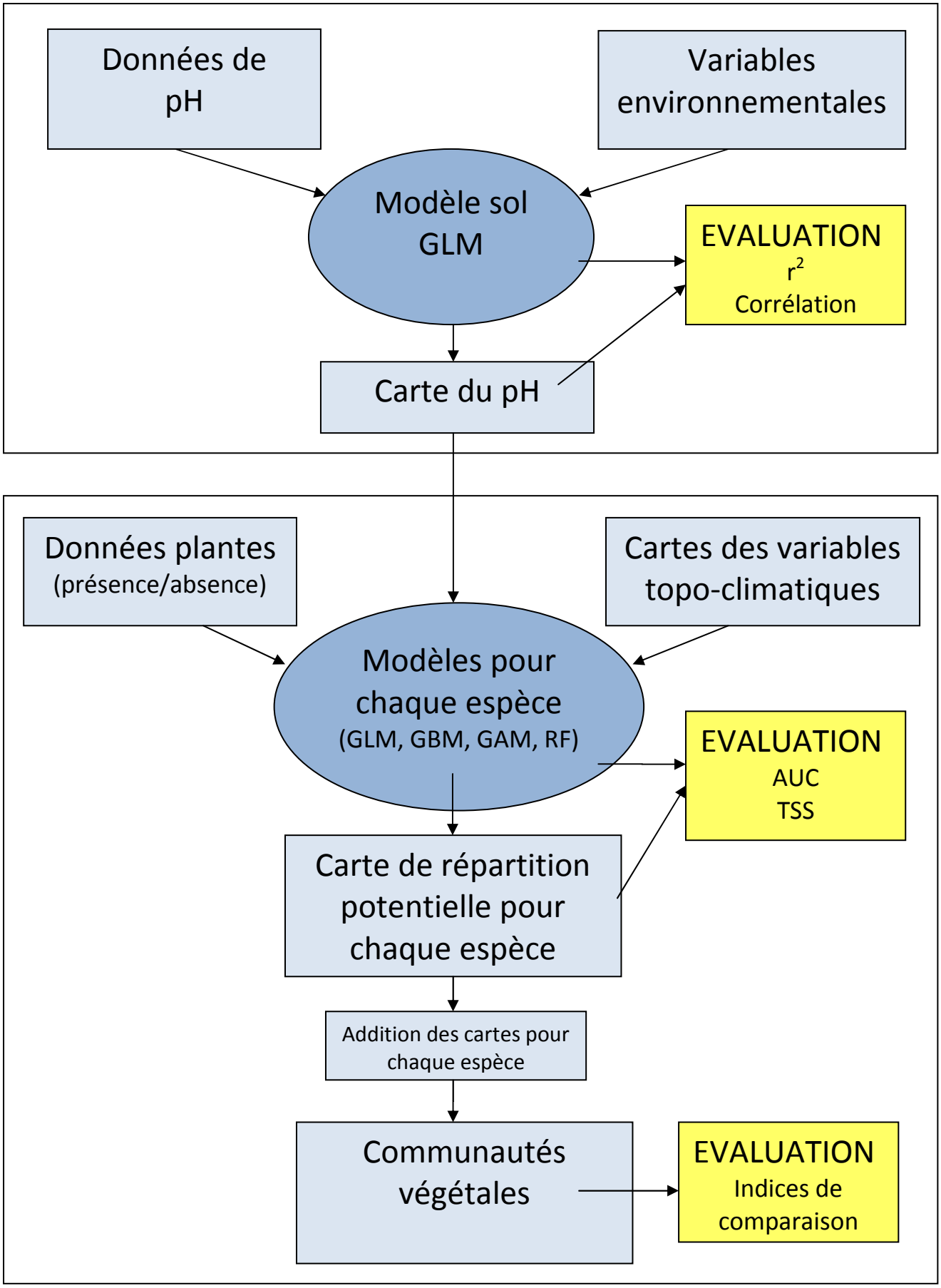



***Fig. 1 :*** *Schéma général des étapes de modélisation dans cette étude*

# MODÉLISATION DU pH DU SOL

## MATERIEL ET METHODES

### 1. Site de l'étude

Le site de cette étude se trouve dans le Canton de Vaud en Suisse et comprend les Alpes occidentales (46°10' à 46°30'N; 6°50' à 7°10'E) (Fig. 2). La superficie totale de cette zone est d'environ 700 km$^2$ avec des altitudes comprises entre 375 mètres à Montreux et 3'210 mètres au sommet du massif des Diablerets. Le climat est tempéré avec des températures et des précipitations moyennes annuelles variant de 8°C et 1'200 mm à 600 mètres à -5°C et 2'600 mm à 3'000 mètres (Bouët, 1985). Les roches de la région sont surtout calcaires mais on y trouve aussi des marbres, schistes, évaporites et dolomies. Les sols présentent une grande diversité allant de sols bruns profonds du type BRUNISOLS à des sols bruns lessivés (NEOLUVISOLS) et des sols beaucoup plus fins reposant directement sur le substrat tels que des LITHOSOLS selon le Référentiel Pédologique (Baize et Girard, 2009). La végétation locale subit une forte pression anthropique : les prairies de basse et moyenne altitude sont généralement fertilisées et fauchées ou pâturées et les hauteurs sont utilisées comme pâturages d'été (Randin et al., 2009a).

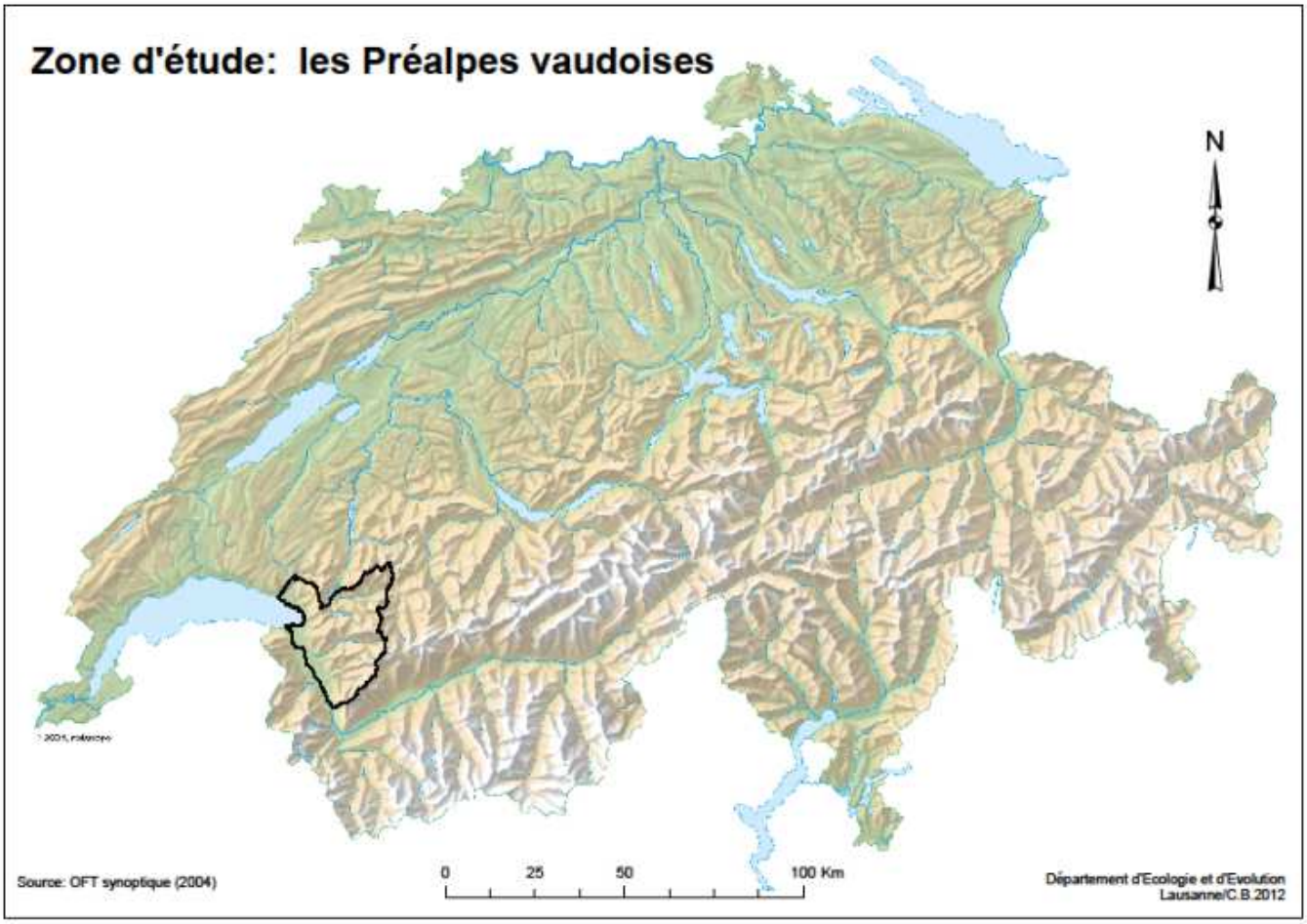



***Fig. 2:*** *Localisation de la zone d'étude entourée d'un trait noir*

## 2. Echantillonnage du pH

Pour permettre par la suite de modéliser la distribution des espèces, les prélèvements des échantillons de sol dont le $pH_{H2O}$ sera analysé doivent correspondre aux sites des relevés de végétation ayant eu lieu entre 2002 et 2009. Avant cette étude, 242 données de pH étaient déjà disponibles pour ces relevés grâce à des prélèvements et analyses effectués par Dubuis et al. durant l'année 2009. Un échantillonnage aléatoire stratifié selon l'altitude, la pente et l'orientation ainsi que l'accessibilité a alors été appliqué sur les relevés de végétation pas encore échantillonnés pour finalement définir 200 sites de prélèvements représentant toutes les strates. Afin que les données de pH correspondent à l'endroit exact des relevés de végétation, un GPS (Global Positioning System) de la marque Trimble® a été employé, permettant de retrouver les relevés de végétation avec une précision de ± 1 mètre.

## 3. Prélèvements sur le terrain

L'échantillonnage des sols a eu lieu durant les mois de juin à septembre 2011 et a été fait de la même manière que l'échantillonnage précédent (Dubuis et al., in revision). Sur chaque site, 50 à 100g de sol ont été prélevés dans les 10 premiers centimètres du solum en prenant soin d'enlever la litière, puis transportés dans des sacs en plastique portant le numéro du relevé de végétation ainsi que la date d'échantillonnage. Pour chaque site, les coordonnées et l'altitude indiquées par le GPS ont été relevées sur une fiche descriptive, de même que la pente, l'orientation et le type de formation végétale observables.

## 4. Analyse des échantillons

Dès le retour du terrain, les échantillons ont été mis séparément à sécher à l'air ambiant dans des bacs en plastique ouverts pendant une semaine au moins. Ils ont ensuite été tamisés à 2mm, permettant de séparer la terre fine (soit la fraction inférieure à 2mm) du squelette (fraction supérieure à 2mm) qui ont été conservées séparément dans deux sacs en plastique portant la date et le numéro du relevé. Le $pH_{H2O}$ des échantillons a été déterminé en laboratoire au mois d'octobre au moyen d'un pH-mètre de la marque WTW® selon la méthode de Kohler, Kosakyan et Scherrer (2010) (voir Annexe 1). Pour chaque échantillon, 10 grammes de terre fine ont été mélangés à 25ml d'eau distillée puis le mélange a été laissé à macérer et agité régulièrement pendant 2 heures avant la manipulation. Le pH-mètre a alors été calibré au moyen de solutions tampon (pH 4.0 et 7.1) puis l'électrode a été placée dans la suspension de sol permettant d'obtenir une valeur de $pH_{H2O}$ au centième. Des réplicas sur 10% des échantillons ont été effectués quelques jours plus tard avec le même appareil afin de contrôler la constance des mesures.

## 5. Recherche de prédicteurs pour le modèle du pH

Le sol est connu pour être influencé par 5 grands groupes de facteurs écologiques ; à savoir le climat, le matériel minéral parental, les êtres vivants et la matière organique qu'ils produisent, le relief et le temps, soit l'âge du sol (Gobat, 2010 :187). Des prédicteurs entrant dans chacune de ces catégories ont été recherchés pour élaborer le modèle du pH (Table 1).

***Table 1 :*** *Prédicteurs testés pour la modélisation du pH*

| Variables | Unités | Détails | Provenance |
|---|---|---|---|
| Précipitations | mm · année $^{-1}$ | Précipitations moyennes annuelles | MeteoSuisse / WSL |
| Températures | degrés Celsius | Température moyenne annuelle | MeteoSuisse / WSL |
| Altitude | mètres | Altitude au-dessus du niveau de la mer | OFT - MNT25 |
| Pente | degrés | Degré d'inclinaison de la pente | OFT - MNT25 |
| Position topographique | sans unité | Surfaces concaves ou convexes | OFT - MNT25 |
| Géologie | sans unité | Classes de géologie simplifiées | Atlas géologique de la Suisse |
| Utilisation du sol | sans unité | Classes d'utilisation du sol simplifiées | OFS / OFT |

Pour le climat, des données concernant les précipitations et les températures moyennes annuelles entre 1961-1990 sont disponibles par les stations météorologiques suisses (MeteoSuisse) sous forme de données qui ont été interpolées avec l'aide d'un modèle numérique de terrain à 25m par l'Institut fédéral de recherches sur la forêt, la neige et le paysage (WSL ; Zimmermann et Kienast, 1999). Des données concernant le relief, notamment l'altitude, la pente et la position topographique, ont été dérivées du modèle numérique de terrain à 25 mètres.

Pour le matériel minéral parental, des données concernant la géologie de la région sont disponibles dans l'Atlas géologique de la Suisse au 1 :25'000 ainsi que sous forme numérique. Les cartes de l'Atlas géologique de la Suisse sont basées sur des informations stratigraphiques telles que l'époque de dépôt des différentes roches et non pas sur des informations lithologiques telles que la composition de la roche. Une reclassification en fonction du type de roche dominant a ainsi été nécessaire pour mettre en évidence le type de substrat à partir duquel le sol a évolué. Celle-ci s'est largement inspirée de la légende de la carte de géologie simplifiée (GT_CH) créée par l'Office Fédéral de la Statistique en tâchant de retranscrire au mieux le caractère acide ou basique des roches (Table 2). Elle a cependant dû être adaptée à la zone d'étude, qui présente des types de roches ne figurant pas dans la légende de la carte GT_CH. Les étages géologiques des différentes cartes de la région ont tout d'abord été regroupés par ordre chronologique puis attribués à une classe sur la base des informations contenues dans les notices explicatives des cartes géologiques. Les polygones des cartes géologiques numérisées ont alors été identifiés et recodés, permettant d'obtenir finalement un carte raster avec une résolution de 25 mètres. Pour la zone de Château-d'Oex, la carte géologique au 1 :25'000 n'est pas disponible à l'heure actuelle et les données utilisées proviennent de la carte GT_CH présentant une résolution de 100 mètres.

***Table 2 :*** *Classes de la carte GT_CH et classes utilisées dans cette étude (classes_cb)*

| N° | Classes GT_CH | Classes_cb |
|---|---|---|
| 1 | Lacs | Lacs, cours d'eau |
| 2 | Glaciers | Glaciers, dépôts glaciaires récents |
| 3 | Silt, glaise, loess, moraine de fond, moraine de surface | Silt, glaise, loess, moraine de fond, moraine de surface |
| 4 | Limons argileux, argiles | Moraine calcaire (glaciaire local) |
| 5 | Graviers et sables (dépôt glaciaire) | Alluvions (graviers, sables, argiles) |
| 6 | Graviers et sables (dépôt actuel) | Marais, charbon |
| 7 | Gravier anguleux, blocs, éboulis de pente | Gravier anguleux, blocs, éboulis de pente |
| 8 | Marnes avec insertions de grès faiblement solidifiées | Dépôts artificiels, remblais |
| 9 | Marnes et schistes argileux | Marnes et schistes argileux |
| 10 | Grès rouge sans calcaire, schistes argileux sablonneux | Grès, schistes argileux sablonneux |
| 11 | Argiles maigres à grasses, ferreuses | Graviers et sables (dépôt fluviatile) |
| 12 | Marnes avec insertions de grès moyennement solidifiées | Limons argileux, argiles |
| 13 | Conglomérats, faiblement à moyennement solidifiés, avec grès et dépôts | Conglomérats, faiblement à moyennement solidifiés, avec grès et dépôts |
| 14 | Conglomérats, faiblement à moyennement solidifiés, avec grès et dépôts | Gypse |
| 15 | Conglomérats, brèches | Conglomérats, brèches |
| 16 | Schistes argileux, phyllite calcaire | Schistes argileux |
| 17 | Schistes marneux, phyllite calcaire, grès | Schistes marneux, schistes calcaires |
| 18 | Phyllites calcaires, schistes lustrés calcaires | Silex |
| 19 | Calcaires massifs | Calcaires massifs |
| 20 | Calcaires, calcaires sablonneux, marneux, schisteux | Calcaires, calcaires sablonneux, marneux, schisteux |
| 21 | Dépôts de schistes marneux et de marnes | Dépôts de schistes marneux et de marnes |
| 22 | Dolomites et cornieules | Dolomites et cornieules |
| 23 | Granits, diorites quartzeuses, syénites quartzeuses, diorites | Granites, diorites quartzeuses, syénites quartzeuses, diorites |
| 24 | Porphyres quartzeux, porphyrites, tufs porphyriques | Porphyres quartzeux, porphyrites, tufs porphyriques |
| 25 | Quartzite | Quartzites |
| 26 | Gneiss à deux micas ou à biotite | Gneiss à deux micas ou à biotite |
| 27 | Conglomérats et brèches schisteuses | Conglomérats et brèches schisteuses |
| 28 | Gneiss à séricite | Amphibolites |
| 29 | Schistes verts | Calcaires métamorphiques |
| 30 | Serpentines | Serpentines, roches ophiolitiques |

Concernant les êtres vivants, des données d'utilisation du sol, reflétant dans une certaine mesure le type de végétation dominant, sont disponibles par l'Office Fédéral de la Statistique (OFS, utilisation du sol en 72 catégories) et l'Office Fédéral de la Topographie (OFT, carte Vector25). Les données les plus précises ont été trouvées dans le modèle Vector25 de l'OFT (surfaces primaires). Certaines zones manquaient cependant d'informations concernant le type d'utilisation du sol, ce qui a pu être résolu en comparant ces zones avec la carte d'utilisation du sol de l'OFS (utilisation du sol en 72 catégories). Les catégories ont été simplifiées en tenant compte du type de végétation dominant.

Malheureusement, aucun prédicteur concernant le temps, à savoir l'âge du sol, n'a pu être trouvé.

Les données ont toutes été converties en format raster puis un masque a été appliqué, permettant de conserver uniquement les milieux ouverts et non urbanisés sur lesquels porte notre étude. Les valeurs des variables ont alors été extraites pour chaque pixel de 25 mètres.

## 6. Elaboration du modèle pour le pH

Le logiciel utilisé pour les traitements statistiques est R 2.14 (R Development Core Team, R Foundation for Statistical Computing, Vienne, AUT) avec la librairie rms (Regression Modelling Strategies ; Harrell, 2012). Les analyses cartographiques ont été réalisées avec ArcGis 10 (ESRI 2010).

Afin de ne conserver que les prédicteurs les plus pertinents dans le modèle du pH, les corrélations entre prédicteurs ont été observées (voir Annexe 3). Les prédicteurs présentant une faible corrélation entre eux (inférieure à ± 0.7) ont été identifiés et utilisés dans les analyses pour réaliser la modélisation du pH. Le type de modèle choisi est un modèle de régression linéaire généralisé (GLM) qui permet, selon Randin et al., d'obtenir une bonne transférabilité des données à d'autres situations (Randin et al., 2006). Pour l'élaboration d'un GLM, la distribution de la variable à modéliser doit tout d'abord être définie ; dans le cas du $pH_{H2O}$, cette distribution suit une loi normale ou loi gaussienne. Les différents prédicteurs sont ensuite ajoutés au modèle sous forme linéaire, ou quadratique si la variation du pH en fonction du prédicteur est susceptible de dessiner une parabole.

Par la suite, une analyse de la variance (ANOVA) avec un test du $Chi^2$ a été effectuée pour déterminer l'importance des prédicteurs dans l'explication de la variance du modèle. Etant donné que les sols sont moins profonds en altitude et que le substrat est plus proche de la surface, le pH du sol est davantage influencé par la géologie en montagne qu'en plaine, ce qui pourrait améliorer la qualité du modèle. Le même modèle a ainsi été testé en ne considérant respectivement que les relevés situés à plus de 1'000 mètres, 1'500 mètres et 2'000 mètres.

## 7. Evaluation du modèle

Pour déterminer la qualité du modèle on a mesuré son coefficient de détermination $R^2$ (Cox & Snell, 1989, Nagelkerke, 1991) qui reflète sa qualité d'ajustement. La qualité d'ajustement correspond à la différence entre la variance des données et la variance des valeurs produites par le modèle. Elle s'étend de 0 à 1 ; zéro indiquant un mauvais ajustement et 1 une excellente correspondance du modèle aux données. L'évaluation du pouvoir prédictif du modèle a été estimée en calculant la corrélation entre valeurs observées et valeurs prédites. Pour obtenir une évaluation pseudo-indépendante, on emploie une procédure dite de « split-sample ». Durant cette procédure, le jeu de données est séparé aléatoirement en deux partitions : 70% pour élaborer le modèle qui est ensuite testé sur le 30% des données restantes, le tout étant répété 100 fois avec une partition aléatoire différente des données. On observe après chaque étape la corrélation entre le pH observé du 30% des données et le pH prédit par le modèle. Plus la valeur obtenue est proche de 1, meilleure est la corrélation entre les données et le modèle, une valeur de 0 étant une absence de corrélation et une valeur de -1 une corrélation négative. La valeur finale d'évaluation du pouvoir prédictif est obtenue en faisant la moyenne des 100 valeurs de corrélation.

8. Carte du pH

Le modèle final a été projeté avec les valeurs des prédicteurs extraites pour toute la zone, permettant d'obtenir une carte de pH pour les Alpes vaudoises. Sur la base de cette carte, les données de pH prédites par le modèle ont ensuite été extraites pour les sites des relevés, ce qui permet de comparer le pH mesuré sur le terrain et le pH prédit, notamment en termes de variance, d'écart-type et de corrélation. Les valeurs ont également été extraites pour toute la zone, permettant d'obtenir une valeur de $pH_{H2O}$ prédite pour chaque pixel de 25 mètres.

# RESULTATS

1. Données de pH

Certains relevés ont dû être retirés des analyses car ils se trouvaient dans des catégories de géologie et d'utilisation du sol trop peu représentées. Le nombre d'échantillons récoltés durant l'été 2011 s'élève finalement à 187 (voir Annexe 2). Ils ont été complétés par 242 mesures déjà existantes, pour atteindre au total 429 valeurs de pH correspondant aux relevés de végétation (Fig. 3). Ces relevés se répartissent entre les altitudes de 432 et 2'708 mètres, avec une majorité comprise entre 800 et 2'400 mètres d'altitude (Fig. 4).

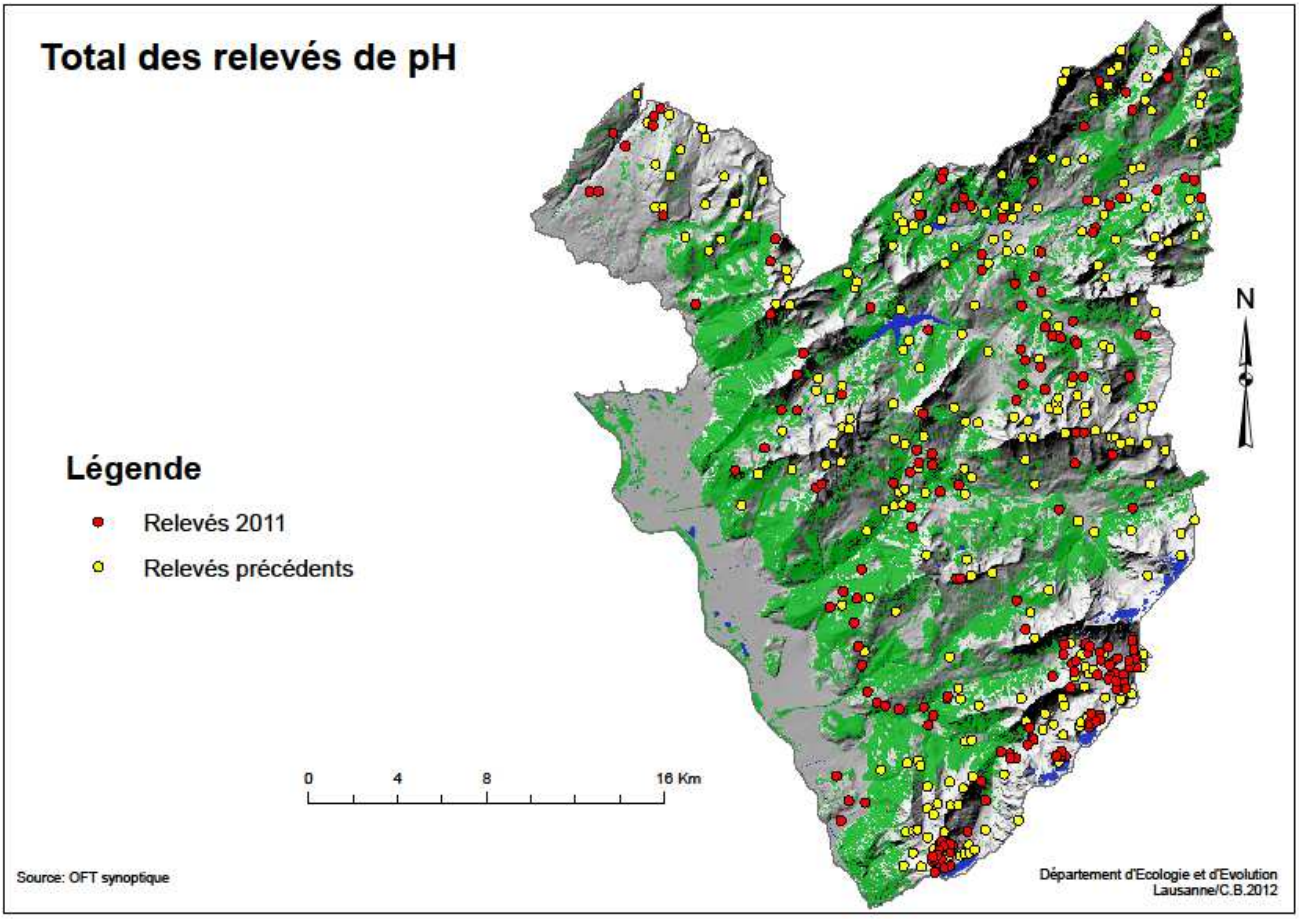



***Fig. 3:*** *Localisation des échantillons de pH. En rouge les relevés effectués durant l'été 2011, en jaune les relevés précédents. Les surfaces en vert représentent les forêts, celles en bleu les lacs et glaciers.*

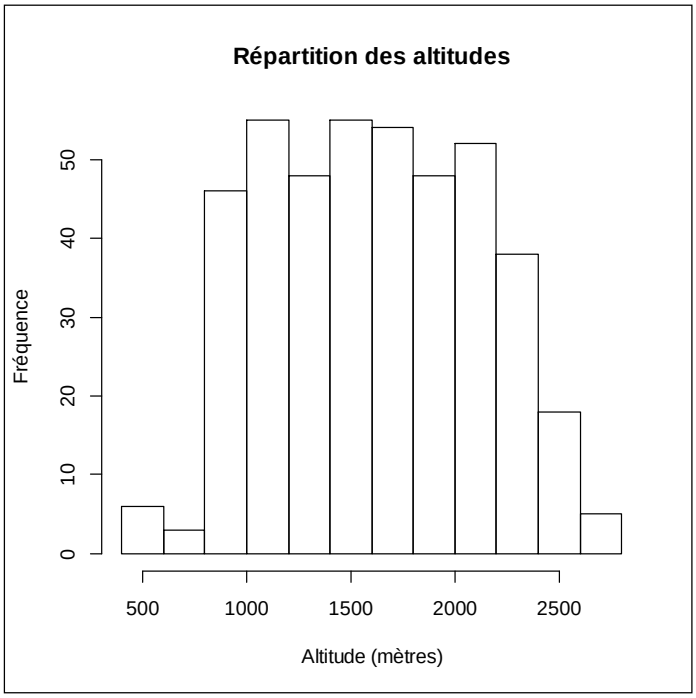



***Fig. 4:*** *Histogramme de l'altitude des prélèvements de pH.*

Les valeurs de $pH_{H2O}$ mesurées en laboratoire sont comprises entre 4.0 et 9.0, avec une médiane à 6.2 et un écart-type de 1.13 (Fig. 5). Les réplicas effectués quelques jours plus tard pour 10% des échantillons ont donné des valeurs consistantes.

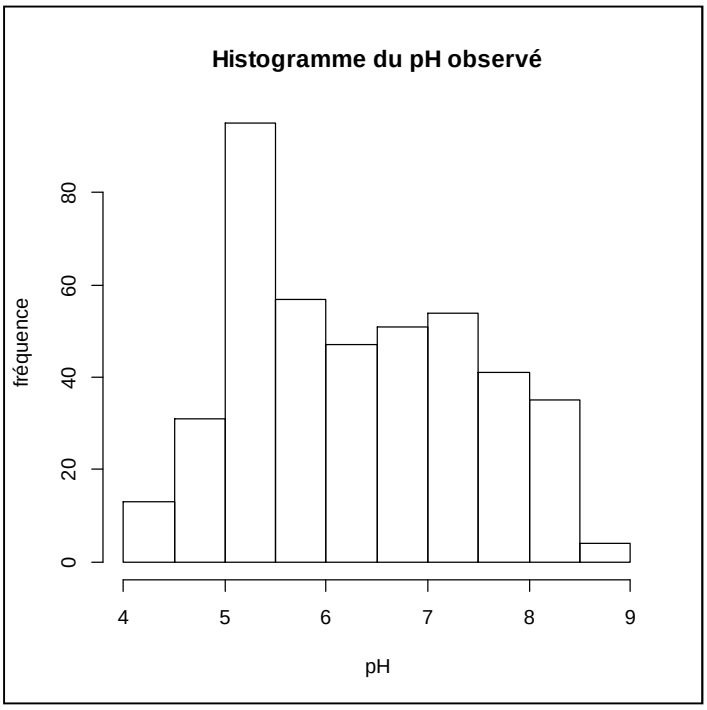



***Fig. 5 :*** *Histogramme du pH observé sur les 429 échantillons.*

## 2. Données de géologie et utilisation du sol

Après codage et reclassification, les données concernant la géologie des Alpes vaudoises ont été regroupées en 27 catégories en fonction du type de roche dominant. Sur la carte suivante (Fig. 6), les couleurs bleues représentent des zones à caractère basique comprenant notamment des roches telles que des calcaires, schistes marneux et conglomérats et les couleurs rouges des zones acides contenant des grès, granites, quartzites ou des schistes sablonneux. Les zones en gris sont des zones avec des roches mixtes, notamment des éboulis ou de la moraine qui sont formés de différents types de roches.

Sur cette carte, les gneiss du Massif des Aiguilles Rouges sont clairement visibles en rouge foncé au niveau de Martigny tout comme les nappes de flysch du domaine piémontais au niveau du Lac de l'Hongrin et autour des Tours d'Aï (en orange vif). Sur la montagne du Chamossaire, la nappe du Niesen du domaine valaisan affleure et présente des turbidites de flysch, ce qui explique que ce sommet soit aussi représenté en rouge. Au niveau du massif des Diablerets, la nappe du Wildhorn est constituée d'une série de dolomies, calcaires, marbres et schistes recouverts de flysch (Esher et al., 1997), représentée de ce fait par des couleurs rouges et bleues sur notre carte. En ligne droite en direction du nord-est à partir de Villeneuve se trouve la nappe des Préalpes médianes, essentiellement formée de calcaires, dolomies et cornieules, représentée en bleu. Le bassin lémanique constitué d'anciens dépôts glaciaires (moraine) est représenté en gris. De même, la plaine du Rhône, formée majoritairement d'alluvions est elle aussi colorée de gris alors que les fonds de vallées, recouverts de moraine calcaire lors de la dernière glaciation, sont en bleu clair.

Pour l'utilisation du sol, une classification en 9 catégories a été retenue, à savoir : les éboulis et pierriers (classe 1); rochers (classe 2); marais (classe 3); forêts (classe 4); buissons et forêt clairsemée (classe 5) ; alpages, pâturages (classe 6); prairies et végétation herbacée improductive (classe 7); infrastructures et exploitations humaines (classe 8) et les lacs, rivières et glaciers (classe 9) (Fig. 7).

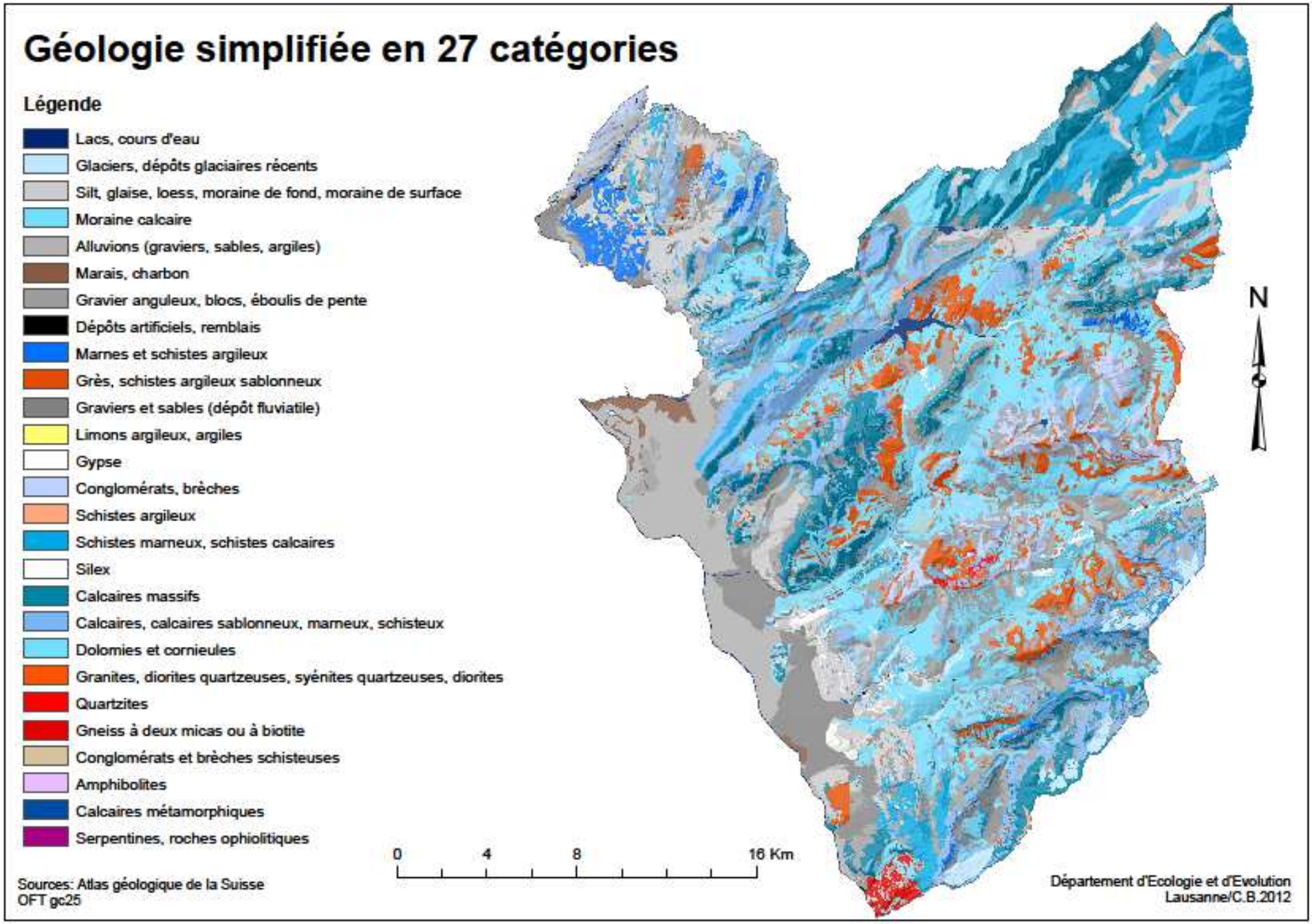



***Fig.6 :*** *Carte de la géologie simplifiée des Alpes vaudoises*

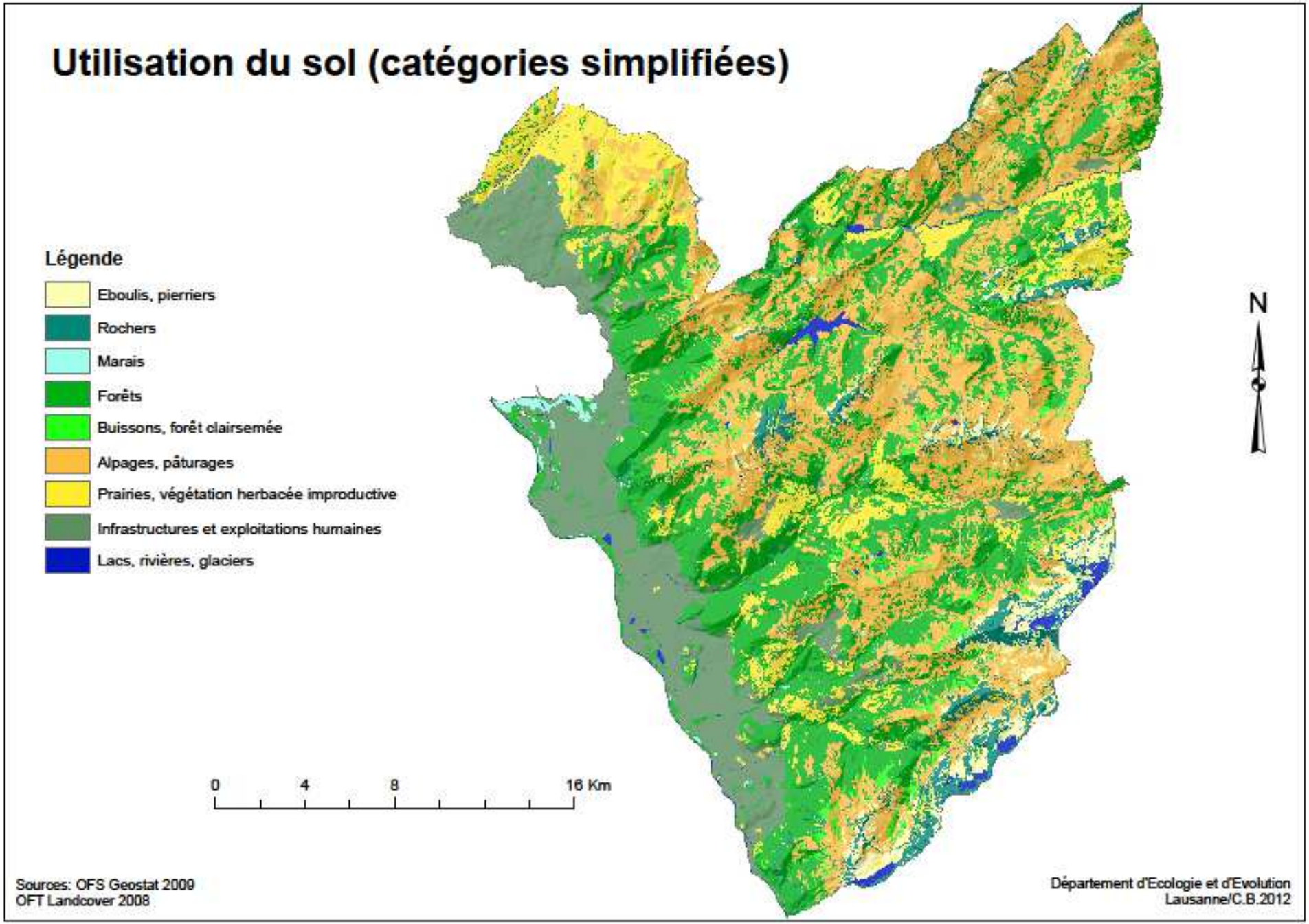



***Fig.7:*** *Carte d'utilisation du sol simplifiée des Alpes vaudoises*

## 3. Prédicteurs conservés dans le modèle de pH

Le modèle de pH final comprend les prédicteurs suivants :

- La géologie simplifiée
- La pente
- L'altitude
- La position topographique
- L'utilisation du sol simplifiée

Formule : pH ~ as.factor(géologie) + pol(altitude) [(2)] + pente + position topographique + as.factor(utilisation du sol).

Le pH est lié à ces prédicteurs comme en témoignent les graphiques ci-dessous (Fig. 8).

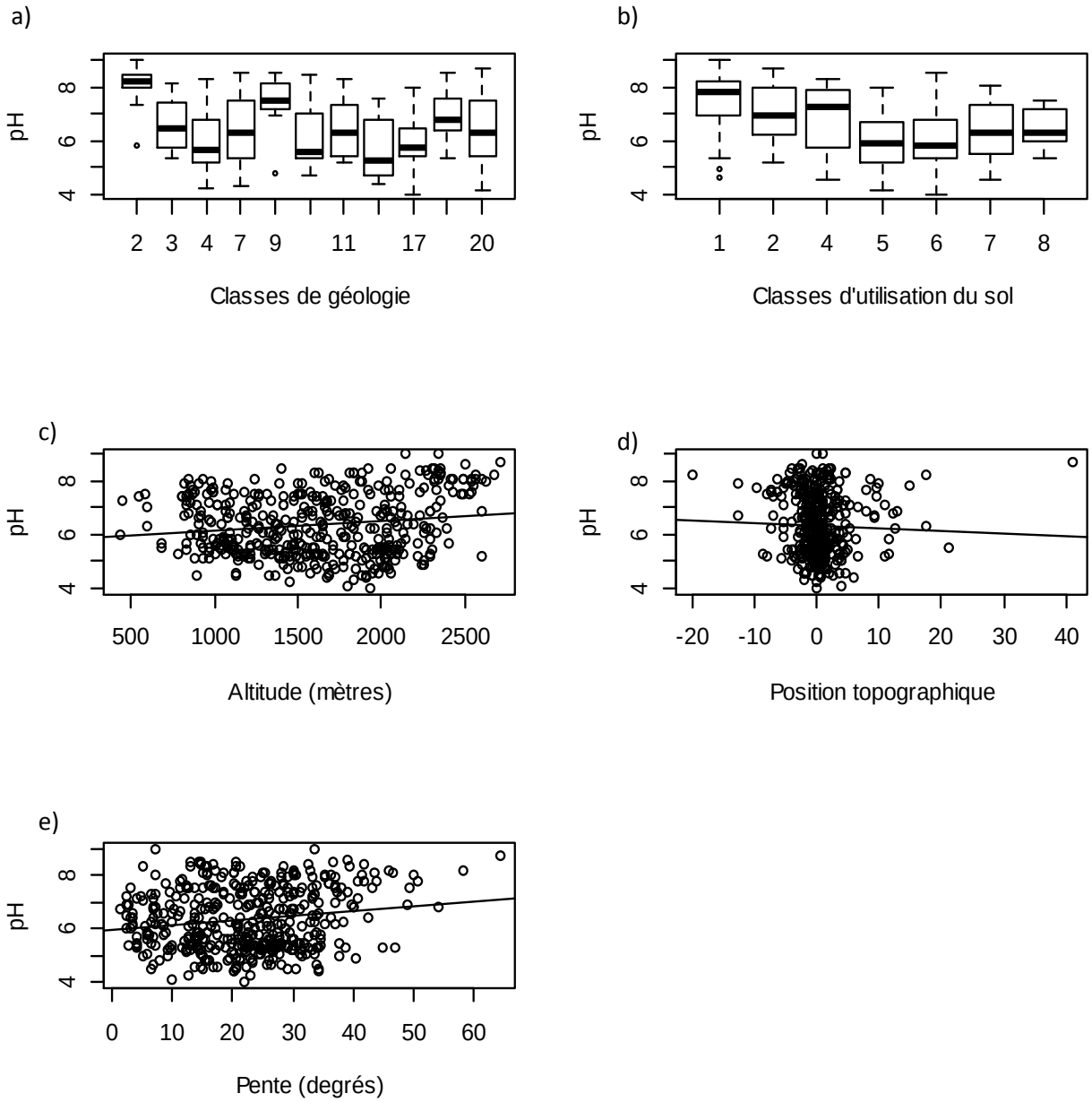



***Fig. 8 :*** *Diagrammes en boîte et en nuage de points du pH en fonction des classes de géologie (a) (voir Table 2 pour la description des numéros), des classes d'utilisation du sol (b), de l'altitude (c), de la position topographique (d) et de la pente (e) .*

Les sols issus des classes de géologie 2 (dépôts glaciaires récents) et 9 (marnes et schistes argileux) présentent une moyenne de pH élevée alors que les sols issus des classes 10 (grès, schistes argileux sablonneux) et 15 (conglomérats, brèches) ont une moyenne basse. Au niveau de l'utilisation du sol, les classes 5 (buissons, forêt clairsemée) et 6 (alpages, pâturages) présentent en moyenne des valeurs de pH plus basses que les autres catégories d'utilisation du sol. Quant à la relation avec la topographie, on s'aperçoit que le pH a tendance à augmenter avec l'altitude et l'inclinaison de la pente.

Selon les résultats de l'ANOVA, la géologie est le prédicteur expliquant le plus de variance dans le modèle du pH. Elle est suivie de l'utilisation du sol, de l'altitude et de la pente, la position topographique se trouvant en dernier (Table 3).

***Table 3 :*** *Résultats de l'analyse de la variance et test du Chi²*

| | Degrés de liberté | Variance expliquée | Degrés de liberté résiduels | Variance résiduelle | Pr(>Chi) |
|---|---|---|---|---|---|
| Variance initiale | NA | NA | 427 | 553.420 | NA |
| Géologie | 10 | 111.570 | 417 | 441.850 | 2.24E-21 |
| Altitude | 2 | 25.441 | 415 | 416.409 | 9.38E-07 |
| Pente | 1 | 10.982 | 414 | 405.427 | 0.0005 |
| Position topo. | 1 | 4.165 | 413 | 401.262 | 0.0330 |
| Utilisation du sol | 6 | 28.264 | 407 | 372.998 | 2.72E-05 |

## 4. Evaluation du modèle

La qualité d'ajustement ($R^2$) obtenue pour le modèle comprenant tous les relevés est de 0.29 ; le pouvoir prédictif déterminé en calculant la corrélation entre pH observé et pH prédit par le modèle après la procédure split-sample est de 0.5 (Table 4).

***Table 4 :*** *Résultats de l'évaluation des modèles*

| | Totalité des relevés | > 1000 mètres | > 1'500 mètres | >2'000 mètres |
|---|---|---|---|---|
| $r^2$ | 0.29 | 0.3 | 0.39 | 0.51 |
| corrélation | 0.5 | 0.49 | 0.57 | 0.65 |

Bien que la qualité d'ajustement soit meilleure pour le modèle élaboré avec les relevés situés à plus haute altitude, la projection du modèle devrait se faire seulement sur une zone plus petite. En conséquence, le nombre de relevés de végétation à disposition pour la deuxième partie de ce travail serait aussi réduit, ce qui diminuerait le nombre d'espèces

avec suffisamment d'occurrences pour être modélisées ainsi que la qualité des modèles. La totalité des données a donc été conservée pour élaborer le modèle.

## 5. Carte du pH

La projection du modèle sur l'ensemble de la zone d'étude produit des données de pH comprises entre 3.0 et 11.9. La majorité des valeurs prédites se trouvent entre 5.8 (1er quartile) et 6.7 (3ème quartile) (Fig. 9).

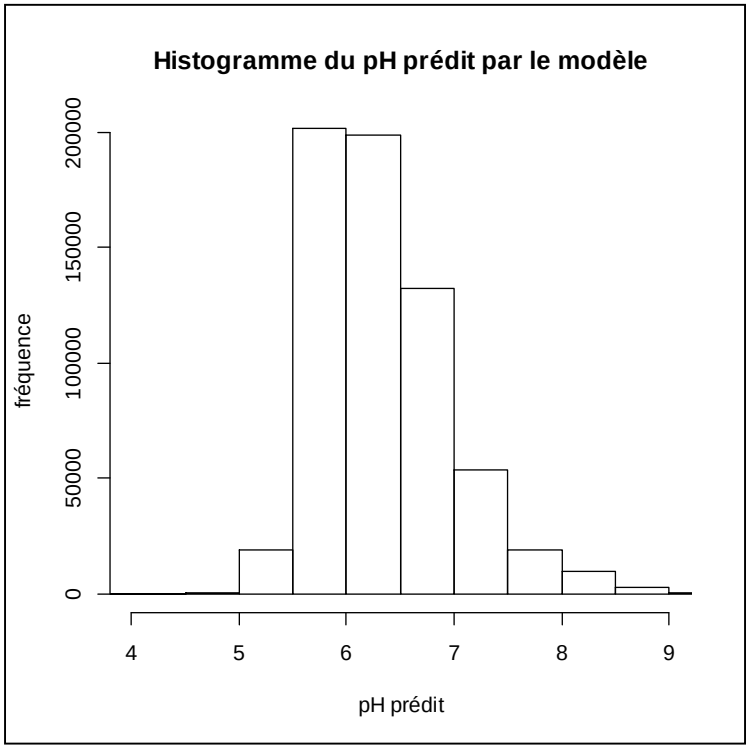



***Fig.9:*** *Histogramme des valeurs de pH prédites par le modèle pour toute la zone d'étude*

La carte de pH obtenue ne concerne que les milieux ouverts puisque les zones boisées ou urbanisées ont été soustraites à l'aide d'un masque. La légende est comprise entre un pH de 3.0 (acide en rouge) et un pH de 9.0 (basique en bleu) (Fig. 10). Sur cette carte, on remarque que, de manière générale, les sommets sont plutôt basiques avec des valeurs de pH entre 7 et 9 alors que les fonds de vallées sont plus acides, le pH prédit étant en général autour de 5.

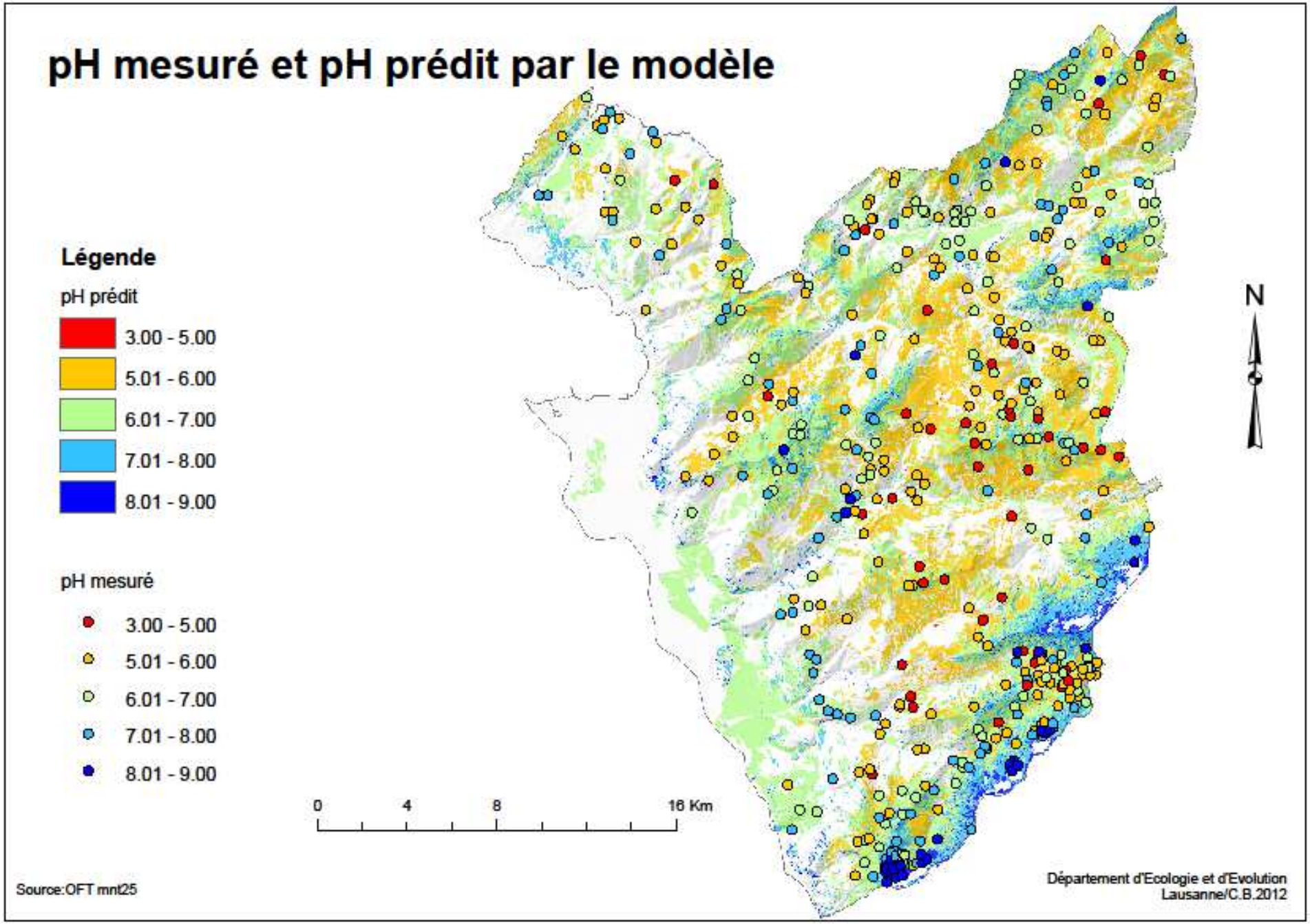





***Fig.10 :*** *Carte du pH mesuré et prédit par le modèle pour toute la zone d'étude*

La comparaison des valeurs de pH observées et des valeurs de pH prédites par le modèle à l'emplacement des relevés donne une corrélation de 0.52 (Fig. 11). Les écarts-types sont de 1.13 pour le pH observé et de 0.66 pour le pH prédit, les variances respectivement de 1.29 et 0.43.

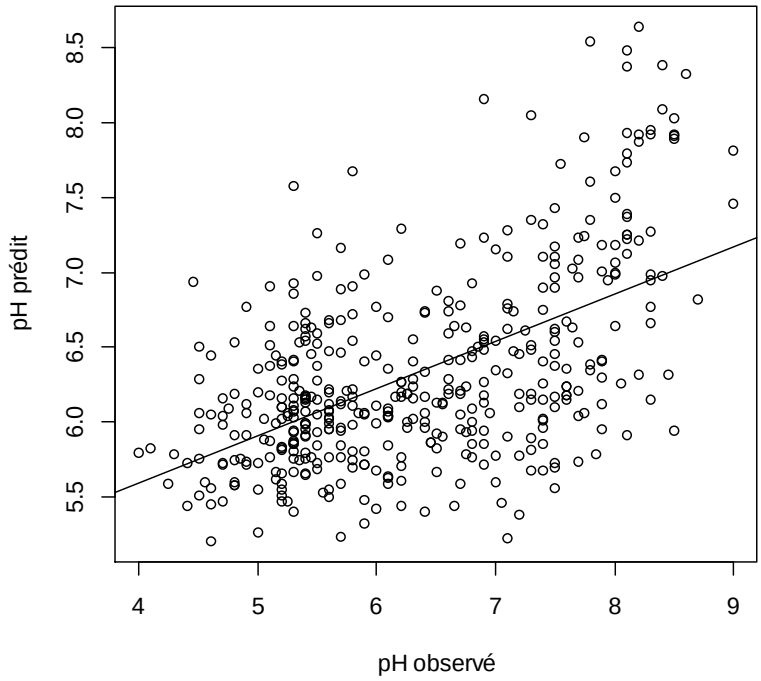



***Fig. 11 :*** *Valeurs et droite de régression entre pH observé sur le terrain et pH prédit par le modèle.*

## DISCUSSION

Les valeurs de pH mesurées en laboratoire montrent que l'on peut trouver des horizons de surface très basiques (pH de 9.0) mais également très acides (pH de 4.0) dans les Alpes vaudoises malgré la prédominance des roches calcaires dans cette région. Bien que la géologie semble liée en partie aux valeurs de pH observées, on s'aperçoit, en examinant les diagrammes du pH en fonction des prédicteurs (Fig. 8), que l'utilisation du sol, l'altitude et, dans une moindre mesure, la pente et la position topographique ont également une influence sur les valeurs observées. Ceci vient confirmer que le pH est une variable fortement dépendante du milieu et que la recherche de prédicteurs influençant l'évolution du sol est une étape nécessaire à sa modélisation.

Bien que le modèle ait calculé quelques valeurs de pH extrêmes (pH de 3.0 et de 11.0), peu probables dans la réalité, la majorité des valeurs se situe dans une fourchette tout-à-fait concevable (entre 5.8 et 6.7). On s'aperçoit cependant que le modèle a tendance à surévaluer les sols acides puisqu'il n'y a que très peu de valeurs basses (pH < 5) et à sous-évaluer les sols calcaires (pH > 8) réduisant ainsi la dispersion des valeurs.

Conformément aux théories pédologiques actuelles, dans lesquelles des processus d'acidification ont lieu avec l'augmentation de la profondeur et du degré d'évolution des sols calcaires (Gobat, Aragno, Matthey, 2010), les altitudes plus élevées où le sol est plus mince ont obtenu des valeurs basiques et la plaine et les fonds de vallées où le sol est plus important, des valeurs plus acides. Bien que leurs distributions soient différentes, les valeurs de pH prédites par le modèle sont corrélées aux valeurs observées et cette corrélation s'améliore en altitude.

Malgré le fait que la géologie soit la variable la plus importante dans l'explication de la variance du modèle, l'information apportée par ce prédicteur n'est pas suffisante pour permettre d'obtenir une bonne qualité d'ajustement. Il en est de même pour l'utilisation du sol, pourtant située en deuxième position dans l'ANOVA. D'autres prédicteurs doivent donc encore être ajoutés au modèle pour permettre de l'améliorer. Etant donné que les modèles calibrés sur les relevés de haute altitude ont obtenu une meilleure qualité d'ajustement, une variable liée à la profondeur du sol aurait peut-être permis d'améliorer ces résultats. En effet, étant donné que les sols sont moins profonds en altitude et donc que le substrat est plus proche de la surface, le pH du sol est davantage influencé par la géologie en montagne qu'en plaine. L'ajout d'un prédicteur lié à la profondeur du solum aurait pu permettre de nuancer l'influence de la géologie en plaine et au contraire de l'amplifier sur les hauteurs et probablement d'obtenir une meilleure qualité d'ajustement. Dans une future étude, il pourrait ainsi être intéressant d'effectuer des sondages à la tarière pour mesurer la profondeur du sol et intégrer ces données au modèle. Des informations concernant les formations superficielles recouvrant les roches auraient également pu améliorer le modèle puisque le pH est influencé par ces dernières (Herbauts et Tanghe, 1987). Suite à l'orogenèse alpine, d'autres dépôts plus récents, notamment des moraines et des loess ont été déposés par les glaciers, les cours d'eau et les vents sur les roches (Spaltenstein, 1984). Les roches affleurent ainsi rarement à nu mais sont au contraire recouvertes d'autres dépôts ce qui diminue leur influence sur la pédogenèse. Ces dépôts ne figurent malheureusement pas sur les cartes géologiques et aucun inventaire de ces formations n'est actuellement disponible pour toute la zone d'étude. Enfin, une variable liée à la fertilisation des prairies qui engendre une augmentation du pH par ajout d'ions ammoniums aurait également pu améliorer la qualité d'ajustement du modèle.

Plusieurs remarques d'ordre méthodologique peuvent aussi être soulevées. Premièrement, étant donné que le pH varie entre les différents horizons du sol, en réponse à des processus de lixiviation, de remontées capillaires, de chélation, la considération de l'horizon supérieur uniquement n'est probablement pas suffisante pour élaborer une carte de pH réellement représentative des sols sur le terrain. Le creusement de profils pédologiques permettant d'observer le sol sur toute sa profondeur aurait amené davantage d'informations mais aurait également considérablement augmenté le temps de réalisation nécessaire à cette étude. Il s'agit également d'une démarche qui nécessite l'accord du propriétaire du terrain et l'accès aux registres fonciers pour le déterminer. Deuxièmement, la question de la représentativité d'une mesure unique pour un pixel de 25 mètres peut être posée puisque la zone d'étude se trouve dans un contexte montagneux où les transitions entre milieux sont abruptes. Les conditions rencontrées sur une surface de 4x4m ne reflètent ainsi pas forcément celles de surfaces plus grandes. Il faut cependant préciser que le choix de la démarche adoptée dans cette étude a été opéré lors de la planification du terrain en fonction du temps et du budget à disposition, ainsi qu'en fonction de l'étendue de la zone d'étude. Enfin le nombre de

prélèvements (429) est relativement faible pour une surface de 700km$^2$ et la détermination de leur emplacement a dû être choisie conformément aux sites des relevés de végétation. L'échantillonnage n'a ainsi pas été effectué selon le type de matériau parental qui est pourtant un des facteurs majeurs de l'évolution du sol, et aucun point de prélèvement n'a malheureusement eu lieu dans les catégories de roches acides.

On peut également discuter de la technique de modélisation utilisée dans notre étude. Dans la littérature, pour prédire des caractéristiques édaphiques, d'autres techniques d'interpolation ont été testées telles que le kriegeage, la méthode Random Forest et la méthode High Accuracy Surface Modelling (HASM). En 2011, Castrignano et al. ont testé différentes techniques de kriegage afin d'élaborer une carte du pH pour une doline de 15'000m$^2$ au nord de l'Italie. Les auteurs ont utilisé 110 échantillons de pH et 5 types de modèles différents (multi-linear regression, ordinary kriging, regression kriging, kriging with external drift et multi-collocated cokriging). Ils ont ensuite fait appel à un test de validation croisée pour comparer les différentes techniques de modélisation. Au final, l'interpolation du pH sur la base des mesures voisines par les modèles regression kriging et kriging with external drift s'est révélée être une meilleure méthode que la méthode de régression linéaire mais les modèles n'ont pas montré de grandes différences entre eux. Il faut cependant noter que la zone étudiée par ces auteurs est très restreinte et présente peu de dénivelé, ce qui rend possible l'approche par kriegage. Au vu de l'étendue et de l'amplitude altitudinale de notre zone d'étude et, étant donné que le pH y varie fortement, l'utilisation de ces modèles n'est ainsi pas recommandée. Certains points qui paraissent spatialement proches présentent en effet des conditions d'altitude et de pente qui peuvent être bien différentes, ce que l'approche de kriegage ne permet pas de prendre en considération. Le recours à une variable liée à la géologie nécessite de plus des limites tranchées qui sont impossibles dans l'approche du kriegeage. Leur modèle de régression linéaire n'a de plus obtenu qu'une valeur $r^2$ de 0.33, ce qui est inférieur à celle obtenue dans notre étude.

Dans une étude parue en 2011, Wiesmeier et al. ont cherché à prédire non pas le pH, mais les stocks de carbone organique, de carbone total, d'azote total et de soufre total. 120 profils de sols ont été creusés dans une région de Mongolie intérieure en Chine et les valeurs de carbone organique dissous, de carbone total, d'azote total, de densité et de pH ont été analysées en laboratoire. Les auteurs ont ensuite couplé ces données à différents prédicteurs, l'utilisation du sol, le type de sol selon, l'unité géologique et 12 variables topographiques pour construire leurs modèles en utilisant la méthode Random Forest (RF) et la méthode Classification and Regression Trees (CART) pour tester la pertinence des prédicteurs. Selon leurs résultats, l'utilisation du sol s'est révélée être le prédicteur le plus pertinent dans les modèles. Les modèles sous forme de RF ont obtenu des résultats satisfaisants puisqu'ils expliquent entre 42-62% de la variance pour les caractéristiques édaphiques testées. Les cartes obtenues pour la zone d'étude sont qualifiées par les auteurs « d'acceptables », qui en concluent que l'approche RF couplée à CART est prometteuse pour de futures études. Il faut cependant noter que les stocks élémentaires sont des propriétés du sol moins fluctuantes que l'acidité du sol, ce qui rend leur modélisation plus facile. De plus, l'amplitude altitudinale de leur zone est réduite, avec des altitudes comprises entre 1'010 et 1'609m et leur résolution plus grossière (pixels de 90x90m) que celle de cette étude. Leur démarche est ainsi difficilement applicable dans notre cas, où la variable modélisée suit une dynamique plus rapide et où la variation altitudinale est nettement plus élevée.

Une nouvelle technique pour modéliser le pH du sol a été proposée en 2009 par Shi et al. Basée sur le théorème des surfaces, leur technique HASM (High Accuracy Surface Modelling) avait auparavant déjà donné de bons résultats lors de la construction de modèles de terrain ou d'interpolations climatiques. Pour tester la pertinence de leur méthode, ils ont utilisé 150 échantillons de sol prélevés dans la province de Jiangxi en Chine pour élaborer une carte du pH pour toute la région. Les résultats de l'approche avec le modèle HASM ont alors été comparés avec ceux d'autres méthodes d'interpolation généralement utilisées, le krigeage, la méthode IDW et la méthode « splines ». Au final, leur modèle s'est révélé meilleur que les autres techniques de modélisation puisqu'il a présenté des valeurs d'erreur moyenne et absolue plus basses. Contrairement à la méthode du kriegage qui a tendance à lisser les valeurs, à la méthode « splines » qui accentue les extrêmes et à la méthode IDW qui crée des discontinuités, la carte obtenue avec leur modèle semblait également présenter des valeurs de pH plus réalistes. Dans une seconde étude (Shi et al., 2011), les auteurs ont cherché à améliorer leur modèle HASM en y intégrant une variable liée à l'utilisation du sol pour prédire différentes propriétés édaphiques dont le pH. L'amélioration des prédictions a été testée en comparant les résultats obtenus avec ceux du modèle HASM sans utilisation du sol ainsi qu'avec différentes autres techniques de kriegeage généralement utilisées auxquelles ils ont à chaque fois ajouté les données d'utilisation du sol. Selon leurs résultats, l'intégration de données d'utilisation du sol a permis d'améliorer les modèles pour la majorité des caractéristiques édaphiques, le meilleur modèle étant finalement le modèle HASM avec utilisation du sol. Ainsi, pour le pH, le pouvoir prédictif calculé sur la base de l'erreur du modèle est de 68.89%. Par rapport à leur précédente étude, l'ajout de la variable « utilisation du sol » a permis d'intégrer des discontinuités spatiales au modèle, ce qui le rend plus proche des conditions rencontrées sur le terrain. Malheureusement, comme le disent les auteurs, ce modèle est difficile d'accès pour des non-mathématiciens et doit encore être amélioré pour rendre son utilisation plus abordable. Aucun logiciel statistique n'est équipé à l'heure actuelle des fonctions statistiques nécessaires. Cette démarche aurait pourtant été intéressante dans le cadre de notre étude puisque l'amplitude altitudinale de leur zone d'étude est relativement élevée (entre 42 et 1'204m.) et que les résultats obtenus pour la variable de pH sont bons.

Le choix d'un modèle sous forme de GLM dans cette étude semble justifié puisque l'approche par kriegage est difficilement applicable et que l'utilisation de la méthode HASM n'est pas encore possible. La méthode RF aurait pu être testée mais ce type de modèle s'ajuste très précisément aux données, ce qui nous fait craindre que son pouvoir prédictif ait pu être moins bon que celui du GLM. Notre démarche s'inscrit à la suite de plusieurs études mais elle s'en distingue par l'étendue et le dénivelé des Alpes vaudoises. La résolution choisie est également plus fine et la recherche de prédicteurs pour tous les facteurs influençant l'évolution du sol s'est révélée intéressante pour la modélisation du pH du sol. Bien que le modèle obtenu dans cette étude doive encore être amélioré, il permet déjà de se faire une bonne idée de la variabilité spatiale des valeurs de pH de l'horizon supérieur du sol dans les Alpes vaudoises.

# MODÉLISATION DE LA DISTRIBUTION DES ESPÈCES ET COMMUNAUTÉS VÉGÉTALES

## MATERIEL ET METHODES

### 1. Relevés phytosociologiques

Entre 2002 et 2009, 912 relevés de végétation ont été effectués dans la zone d'étude. Ils ont tous eu lieu dans des milieux ouverts (prairies, pâturages, pelouses et éboulis) selon un échantillonnage aléatoire stratifié basé sur l'altitude, la pente et l'orientation du terrain (Hirzel & Guisan, 2002). L'autocorrélation spatiale a été évitée en respectant une distance de 200 mètres entre les relevés. Pour chacun de ces relevés, un inventaire complet avec l'abondance des espèces (code d'abondance-dominance de Braun-Blanquet) de la végétation sur une parcelle de 4m$^2$ est disponible. Chaque relevé a aussi été attribué à une association végétale au sens de la phytosociologie sigmatiste (Braun-Blanquet, 1928).

### 2. Prédicteurs topo-climatiques

Dans les modèles, 3 variables climatiques et 2 variables topographiques ont été utilisées comme prédicteurs topo-climatiques (Table 5). Les variables climatiques comprennent les degrés jours, l'indice d'humidité et la radiation solaire globale annuelle ; la topographie a été appréhendée par la pente et la position topographique. Des valeurs positives de position topographique correspondent à des bosses et crêtes sur le terrain, des valeurs négatives à des dépressions et vallées.

***Table 5 :*** *Prédicteurs topo-climatiques utilisés pour la modélisation des espèces végétales*

| Variables | Unités | Détails |
|---|---|---|
| Degrés jours | °C · jour ·année $^{-1}$ | Somme des jours multipliés par la température si > 0°C |
| Indice d'humidité | mm ·jour $^{-1}$ | Moyenne mensuelle de l'humidité du sol (précipitations – évapotranspiration) journalière des mois de juin-août, soit durant la période de croissance |
| Radiation solaire globale annuelle | kJ ·m$^{-2}$·jour$^{-1}$ | Somme de la moyenne mensuelle de la radiation solaire journalière |
| Pente | degrés | Degré d'inclinaison de la pente |
| Position topographique | sans unité | Surfaces concaves ou convexes |

Les données climatiques proviennent des stations météorologiques suisses interpolées à l'aide d'un modèle numérique de terrain. Les données topographiques ont été dérivées du modèle numérique de terrain. Ces données sont disponibles à l'UNIL et ont été calculées par l'Institut fédéral de recherche sur la forêt, la neige et le paysage (WSL), suivant la démarche de Zimmermann & Kienast (1999). Leur résolution est de 25 mètres.

Ces prédicteurs sont présumés être d'une grande importance écophysiologique pour les plantes (Guisan & Zimmermann, 2000 ; Pearson et al., 2002 ; Körner, 2003 ; Guisan & Thuiller, 2005) et ont déjà donné des résultats positifs lors de précédentes modélisations sur la même zone d'étude (Engler et al., 2009 ; Randin et al., 2009c, Pellissier et al., 2010). Les deux variables topographiques sont des variables indirectes au sens de Guisan & Zimmermann (2000) ou Guisan & Thuiller (2005). La pente est à l'origine de processus gravitaires alors que la position topographique concerne plutôt l'accumulation d'eau, la persistance de la neige, l'enrichissement en azote ou la protection contre le vent (Randin et al., 2009a).
Tous ces prédicteurs sont disponibles sous forme de cartes en format raster. Les valeurs ont ainsi pu être extraites pour chacun des sites des 912 relevés de végétation.

## 3. Données de pH

Les données de pH pour les 912 relevés de végétation ont été dérivées de la carte issue du modèle créé dans la première partie de ce travail. Les relevés, pour lesquels le modèle n'a pas pu prédire de valeur, car situés dans des zones urbanisées ou de forêts, ont été enlevés, ce qui a réduit le nombre de relevés à 897.

## 4. Elaboration d'un jeu de données de calibration et d'un jeu de données indépendant

Les valeurs obtenues pour les 5 prédicteurs topo-climatiques et pour le pH ont été couplées aux données de présence/absence des espèces des relevés de végétation. L'entier des données a alors été séparé en deux partitions : un jeu de données de calibration (606 relevés réalisés de 2002 à 2007) sur lequel ont été basés les modèles et un jeu de données indépendant (290 relevés réalisés en 2009) qui a servi à l'évaluation des modèles créés. Les deux jeux de données ont été échantillonnés indépendamment, toutes les strates sont représentées dans les deux cas.

Dans le cadre de cette étude, les espèces présentant plus de 20 occurrences au sein des 912 relevés ont été utilisées pour calibrer les modèles de distribution des espèces. Les données d'espèces arbustives (telles que *Acer pseudoplatanus* et *Picea abies*, par exemple) ayant été recensées dans les milieux ouverts sous forme de semis ont été supprimées car elles ne sont pas représentatives de la réelle distribution de l'espèce. Au total, 207 espèces ont été modélisées.

## 5. Modèles de distribution d'espèces

La modélisation de la distribution des espèces végétales a été élaborée en utilisant la librairie BIOMOD (Thuiller et al., 2009) dans le logiciel R. Cette librairie permet de faire différents types de modèles ; ceux choisis pour cette étude sont :

- Generalised Linear Models (GLM ; McCullagh and Nelder ,1989)
- Generalised Additive Models (GAM ; Hastie and Tibshirani, 1990)
- Generalised Boosting Models (GBM ; Ridgeway, 1999; Friedman, Hastie and Tibshirani, 2000; Friedman, 2001)
- Random Forest (RF ; Breiman , 2001 ; Breiman, 2002)

Il s'agit de deux modèles de régression (GAM et GLM) et de deux modèles basés sur des arbres de classification (Random Forest et GBM). Pour chaque espèce végétale, ces 4 modèles ont été élaborés sur la base des prédicteurs topo-climatiques seuls et sur les prédicteurs topo-climatiques et le pH, soit 2 sets de 4 modèles.

## 6. Pouvoir prédictif et importance des variables

L'évaluation du pouvoir prédictif a eu lieu sur le jeu de données indépendant en utilisant deux indices :

- Area Under the Curve (AUC ; Fielding & Bell, 1997)
- True Skill Statistic (TSS ; Allouche et al., 2006)

L'indice AUC représente l'aire sous la courbe d'un graphique dont les axes sont la sensibilité et la spécificité du modèle. La sensibilité mesure le nombre de présences correctement prédites, la spécificité le nombre d'absences correctement prédites par le modèle. Les valeurs de l'indice AUC sont comprises entre 0 et 1, un indice entre 0 à 0.5 indiquant que les résultats du modèle sont à l'opposé de la réalité, un indice entre 0.5 et 0.7 que les prédictions du modèle sont faibles, un indice supérieur à 0.7 que les résultats sont utilisables, supérieur à 0.8 que le modèle est bon et un indice de 1 que le modèle est parfait. L'indice TSS compare le nombre de prédictions justes en soustrayant les prédictions correctes pouvant être attribuées au hasard sur la base d'une table de contingence. Les seuils sont compris entre -1 et +1, +1 indiquant une correspondance parfaite du modèle alors qu'une valeur de 0 ou moins indique que la performance n'est pas meilleure que si elle était purement aléatoire.

L'amélioration apportée par la variable de pH au pouvoir prédictif des modèles a été appréhendée en observant la différence entre la moyenne des indices AUC et TSS des sets de modèles topo-climatiques et topo-climatique avec pH pour chaque espèce du jeu de données indépendant. La corrélation entre amélioration de l'AUC et du TSS a ensuite été observée pour déterminer si le changement de la qualité des modèles est constant pour les deux indices. La moyenne du changement de l'AUC et TSS due au pH pour l'ensemble des espèces a également été calculée afin de déterminer l'impact de ce prédicteur dans les modèles. Pour nous permettre de comparer les résultats avec ceux de Dubuis et al., le pourcentage d'espèces avec une amélioration de l'AUC supérieure à 0.02 a été calculé ainsi que la moyenne de cette amélioration. Enfin, un exemple de projection pour l'espèce dont le modèle a été le plus amélioré par les données de pH a été élaboré.

L'importance respective des variables dans les modèles a été calculée en faisant varier de manière aléatoire une variable à la fois et en calibrant le modèle avec cette variable mélangée, les autres étant maintenues constantes. On observe ensuite la corrélation entre les résultats obtenus et le modèle initial. L'importance de la variable est alors déterminée en soustrayant la corrélation calculée à 1. Plus la valeur est élevée, plus le prédicteur est important dans le modèle.

## 7. Comparaison de l'amélioration de l'AUC obtenue et des traits biologiques et valeurs écologiques de Landolt

Les valeurs écologiques de Landolt (Landolt et al., 2010) utilisées sont : la réaction (R), l'humus (H), les nutriments (N), l'humidité (F), la dispersité (D), la lumière (L), la température (T) et la continentalité (K). Concernant les valeurs de Landolt, l'échelle s'étend en général de 1 pour des conditions de faible valeur de l'indicateur à 5 pour des conditions de forte valeur. Ainsi par exemple pour la réaction, l'échelle est comprise entre 1 pour les espèces poussant sur des sols très acides où la réaction est faible à 5 pour des sols basiques où la réaction est élevée.

Pour les traits biologiques, la hauteur de végétation (vegH), la surface foliaire spécifique (SLA) et le contenu en matière sèche des feuilles (LDMC) ont été utilisés. La hauteur de végétation est la distance entre l'organe photosynthétique le plus élevé et le sol. Elle est généralement associée à la vigueur et à la fécondité de la plante (Cornelissen et al., 2003). La surface foliaire spécifique est la surface sur une face d'une feuille fraîche divisée par sa masse sèche après séchage au four. Elle est en général bien corrélée au taux de croissance ou à la capacité photosynthétique de la plante (Cornelissen et al., 2003). Le contenu en matière sèche des feuilles est calculé comme la masse d'une feuille séchée au four divisée par sa masse fraîche lorsque saturée en eau. Ce trait biologique est généralement corrélé négativement avec le taux de croissance potentielle et positivement avec la durée de vie foliaire. Ces valeurs correspondent à la moyenne calculée à partir de 10 individus pour chaque espèce, toutes échantillonnées dans la zone d'étude (Pottier et al., in press).

L'amélioration de l'AUC obtenue pour chaque espèce du jeu de données indépendant a été associée aux traits fonctionnels et valeurs écologiques de Landolt (Landolt et al., 2010) afin de déterminer si certains types d'espèces sont sensiblement mieux prédits par les données de pH. Pour les traits biologiques, une régression linéaire a été appliquée sur les valeurs de l'amélioration de l'indice AUC obtenue pour chaque espèce par l'ajout des données de pH. Les espèces ont également été groupées selon leurs valeurs écologiques et les diagrammes en boîte de l'amélioration de l'indice AUC due au pH ont été observés en fonction de ces valeurs.

## 8. Comparaison des communautés prédites et observées

Pour chaque espèce et chaque relevé du jeu de données indépendant, une moyenne des probabilités issue des 4 modèles (GLM, GBM, GAM et RF) a été calculée pour les sets de modèles topo-climatiques et topo-climatique avec pH en pondérant chaque prédiction par l'AUC obtenue lors de l'évaluation du modèle.

Pour obtenir des données en présence-absences binaires pour chaque espèce et chaque point à partir des probabilités prédites par le modèle, sans dépendre d'un seuil qui risquerait de surestimer le nombre d'espèces prédit, on a utilisé la fonction rbinom (Kachitvichyanukul et Schmeiser, 1988) dans R pour tirer des présences ou absences d'une distribution binomiale basée sur la valeur de probabilité prédite. Ce procédé à été répété 1'000 fois, résultant en 1'000 matrices différentes de communautés prédites (Dubuis et al., 2011). Ces matrices de communautés ont été comparées aux espèces observées sur la base d'une table de contingence (voir la Table 6 pour un exemple); une telle table contient, pour chaque communauté, le nombre d'espèces correctement prédites comme présentes, les espèces

prédites en trop, les espèces correctement prédites comme absentes et le nombre d'espèces prédites comme absentes alors qu'elles sont présentes en réalité.

Ces quatre valeurs permettent de calculer différents indices pour évaluer la qualité de prédiction de la communauté, à savoir : le taux d'espèces surestimées, le taux d'espèces sous-estimées, le succès de prédiction, la sensibilité, la spécificité et la similarité (estimée avec l'indice de Sorensen). La sensibilité (1) mesure la proportion des espèces observées qui ont été correctement prédites. La spécificité (2) mesure la proportion d'espèces prédites comme absentes dans la communauté et qui n'ont effectivement pas été observées. La similarité (3) évalue la ressemblance de composition entre la communauté observée sur le terrain et celle prédite par le modèle de communauté.

***Table 6 :*** *Exemple de tableau de contingence.*

| | **Plante observée** | **Absence de la plante** |
|---|---|---|
| **Présence prédite** | A<br>Vrais positifs | B<br>Faux positifs |
| **Absence prédite** | C<br>Faux négatifs | D<br>Vrais négatifs |

$$\text{Sensibilité} = \frac{A}{A+C} \qquad (1)$$

$$\text{Spécificité} = \frac{D}{D+B} \qquad (2)$$

$$\text{Similarité (indice de Sorensen)} = \frac{2A}{2A+B+C} \qquad (3)$$

Les indices obtenus pour les communautés assemblées sur la base des modèles topo-climatiques et ceux obtenus pour les communautés assemblées sur la base des modèles topo-climatiques avec pH ont ensuite été comparés, ainsi que les diagrammes en nuages de points de ces indices en fonction de l'altitude, de la pente et de la position topographique pour déterminer s'il y a une amélioration des prédictions en fonction des caractéristiques environnementales.

L'erreur entre richesse spécifique prédite par les modèles avec et sans pH et la richesse spécifique observée a également été calculée ainsi que sa distribution en fonction de l'altitude, de la pente et de la position topographique.

Tous ces indices ont été calculés pour les 1'000 communautés prédites par retirage pour chacun des 290 relevés du jeu de données indépendant puis la moyenne des 1'000 communautés a été calculée pour les analyses.

# RESULTATS

## 1. Pouvoir prédictif et importance des variables

L'amélioration moyenne de l'indice AUC due à l'ajout du pH dans le modèle s'élève à 0.0073, celle de l'indice TSS à 0.0091 (Table 7) (voir aussi Annexe 4). Etant donné la très forte corrélation entre ces deux indices ($r^2$ de 0.85) (Fig. 12), on ne présente ici que les résultats pour l'indice AUC.

***Table 7 :*** *Valeurs des indices AUC et TSS pour les 2 sets de modèles*

| | Modèle topo-climatique | Modèle TC + pH |
|---|---|---|
| Indice AUC moyen | 0.7865 | 0.7939 |
| Indice TSS moyen | 0.5140 | 0.5232 |

Le nombre d'espèces présentant une amélioration de leur indice AUC supérieure à 0.02 s'élève à 43, soit 20,7% des espèces testées. La moyenne de l'amélioration pour ces espèces est de 0.0393 (voir Annexe 5). L'espèce la mieux prédite par l'ajout des données de pH est *Valeriana montana* avec une amélioration de 0.1293 (Fig. 13).

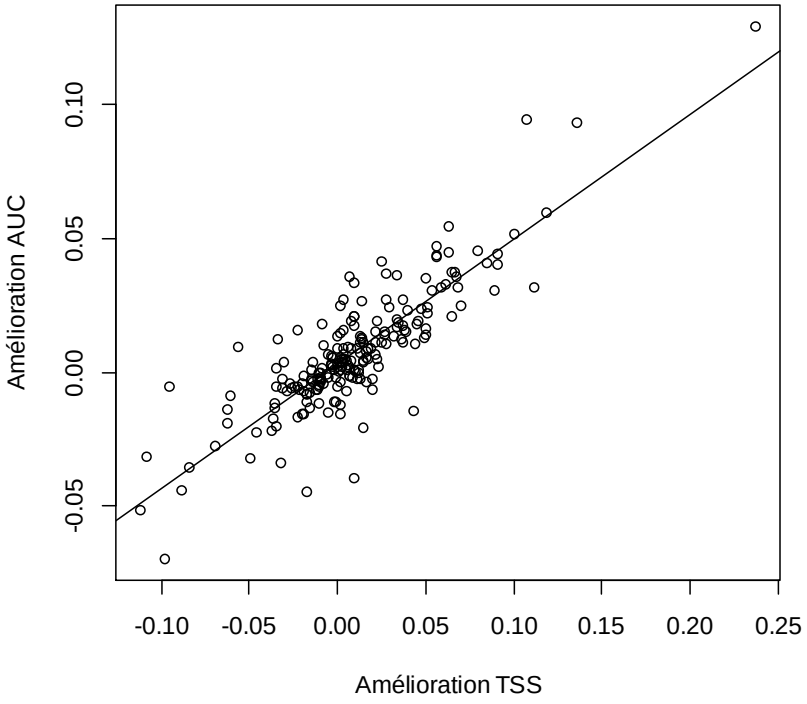


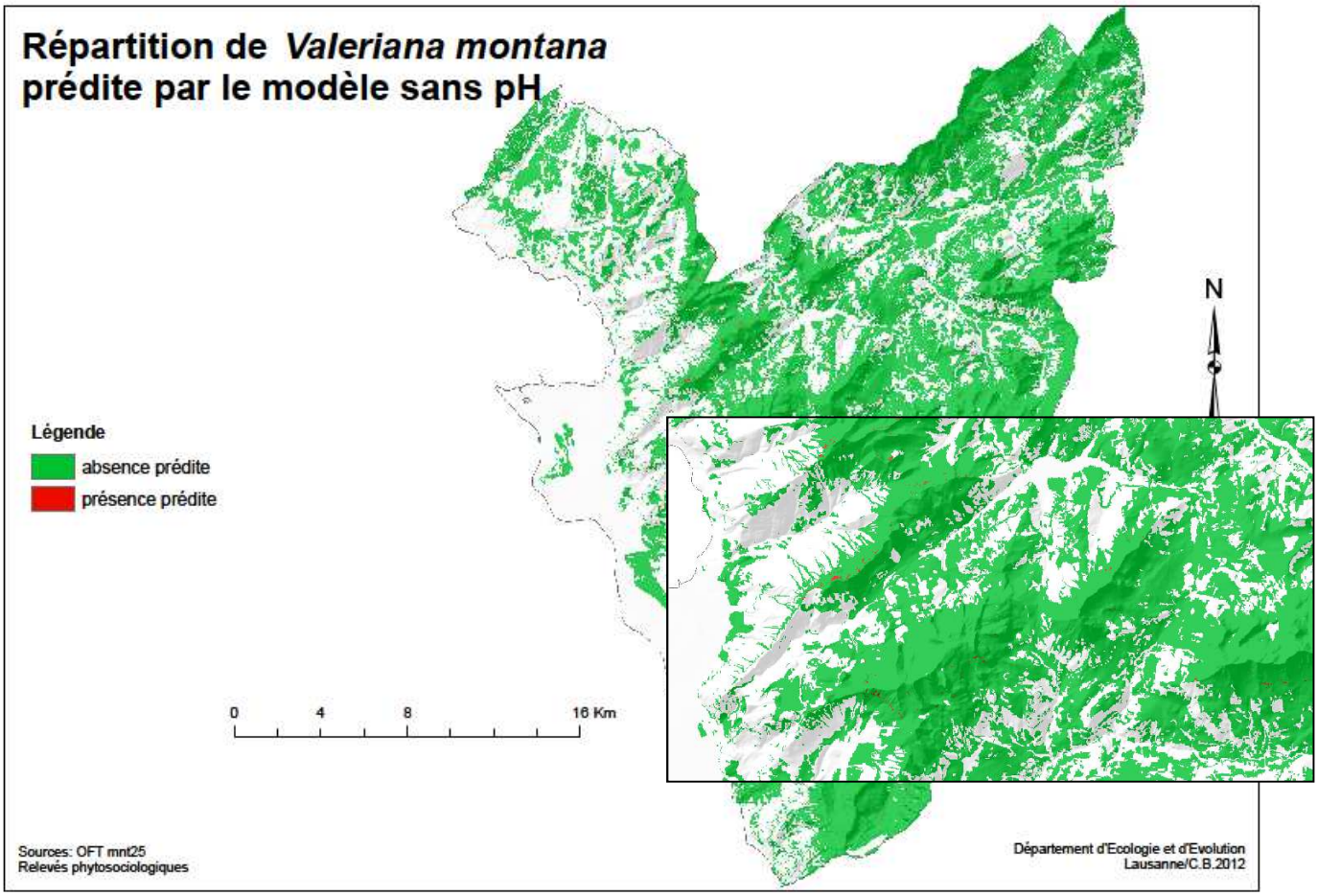

***Fig. 12 :*** *Graphique de la corrélation entre indices AUC et TSS.*

Concernant l'importance des prédicteurs dans les modèles, la variable degrés jours est la plus importante avec une moyenne de 0.4360, suivie de l'humidité avec une moyenne de 0.2270 et de la pente 0.2093 (Table 8). Le prédicteur pH vient en quatrième position avec une importance moyenne de 0.1966, largement devant la radiation (0.0878) et la position topographique (0.0828). Dans les modèles GBM et RF, le pH est le troisième prédicteur le plus important, devançant l'humidité dans le GBM et la pente dans le RF. La radiation et la position topographique ont très peu d'influence sur les prédictions, comme l'attestent les valeurs obtenues.

***Table 8 :*** *Importance des prédicteurs dans les modèles et moyenne*

| Prédicteur | GAM | GBM | GLM | RF | Moyenne |
|---|---|---|---|---|---|
| Degrés jours | 0.5292 | 0.3656 | 0.5497 | 0.2997 | 0.4360 |
| Pente | 0.2327 | 0.2057 | 0.2399 | 0.1591 | 0.2093 |
| Humidité | 0.2635 | 0.1628 | 0.3021 | 0.1797 | 0.2270 |
| Radiation | 0.0755 | 0.0910 | 0.0855 | 0.0991 | 0.0878 |
| Position topo. | 0.0772 | 0.0793 | 0.0938 | 0.0809 | 0.0828 |
| pH | 0.2041 | 0.2129 | 0.1979 | 0.1715 | 0.1966 |

La comparaison entre les cartes de répartition pour *Valeriana montana* (Fig. 13) intégrant ou pas les données de pH montre davantage de zones favorables à cette espèce prédites avec les données de pH. Il s'agit en effet d'une espèce poussant sur des terrains instables tels que des éboulis et pierriers, de préférence calcaires (Lauber et Wagner, 2007). La distribution prédite par le modèle topo-climatique seul semble trop restreinte par rapport aux conditions favorables à cette espèce et la considération des données de pH permet d'obtenir davantage de précision concernant ce type de terrains puisque les données d'utilisation du sol y ont été intégrées.



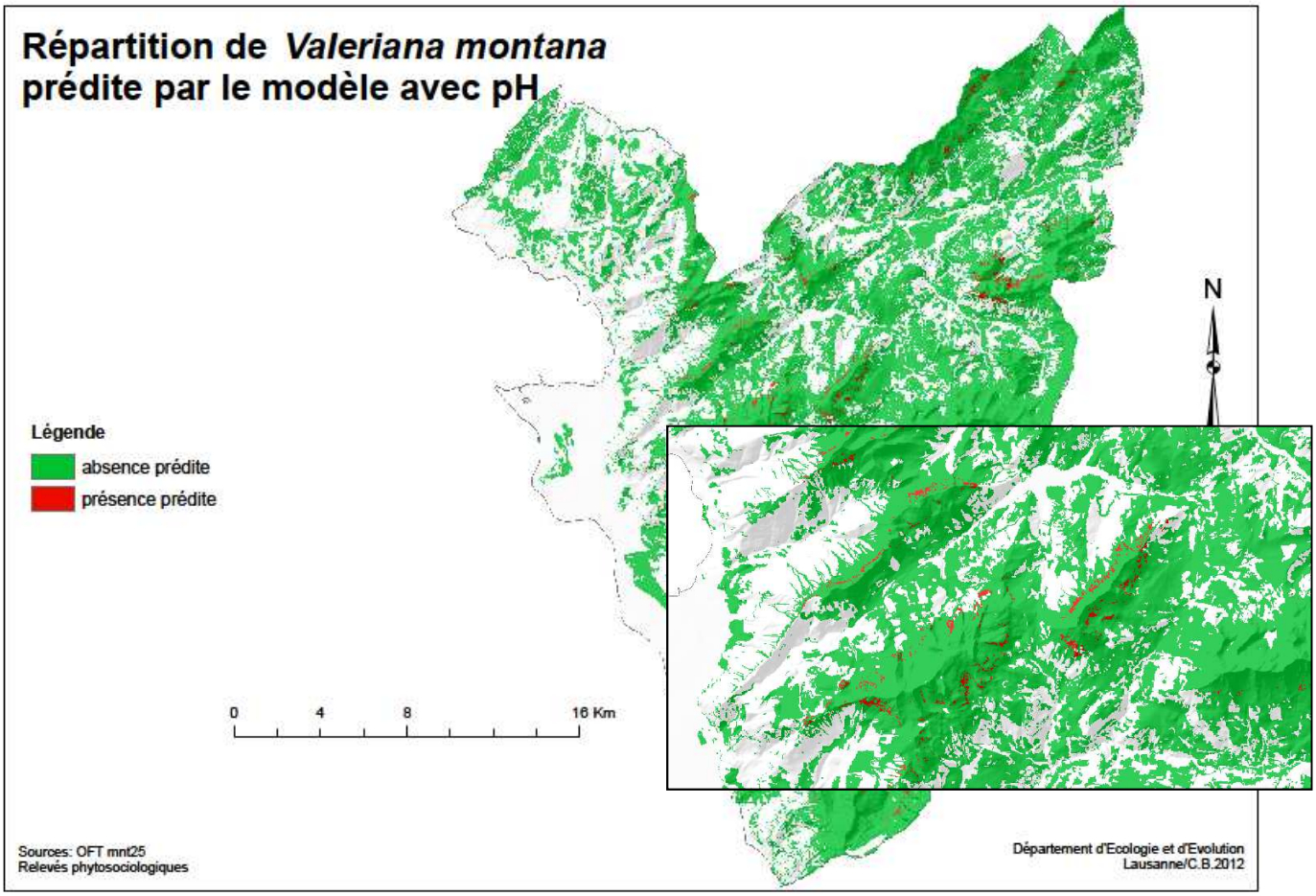




***Fig. 13 :*** *Cartes de répartition du set de modèles topo-climatique et topo-climatique avec pH pour Valeriana montana*

## 2. Traits biologiques, valeurs de Landolt et AUC

L'observation de l'amélioration de l'indice AUC due au pH en fonction des traits biologiques ne montre aucune tendance d'amélioration des prédictions en fonction de ces traits vu la dispersion des points et la pente presque nulle des graphiques obtenus (Fig. 14) (voir aussi Annexe 6). En moyenne, les données de pH n'améliorent pas les prédictions puisque les valeurs d'amélioration obtenues sont proches de zéro.

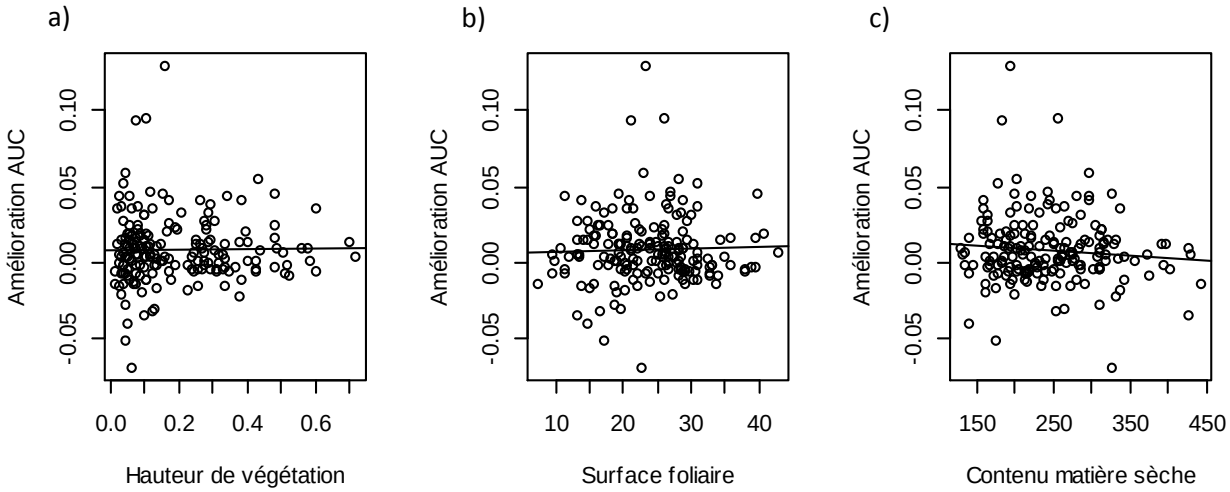



***Fig. 14 :*** *Diagrammes en nuage de points de l'amélioration de l'indice AUC en fonction de la hauteur de végétation (a), de la surface foliaire (b) et du contenu en matière sèche (c) des espèces du jeu de données indépendant.*

Pour les valeurs écologiques de Landolt, seuls les diagrammes pour la réaction (R), l'humidité (F) et l'humus (H) sont présentés ici (Fig. 15), les autres pouvant être consultés dans les annexes 7 et 8. Excepté pour l'humidité, où l'on s'aperçoit que les espèces poussant sur des sols secs (valeur de 1) sont mieux prédites par le modèle intégrant les données de pH, les autres classes de conditions ne présentent pas d'amélioration distincte, les moyennes se situant presque toujours au même niveau du graphique. La classe d'humidité 5 n'est représentée par une seule espèce, ce qui explique le trait unique du diagramme.

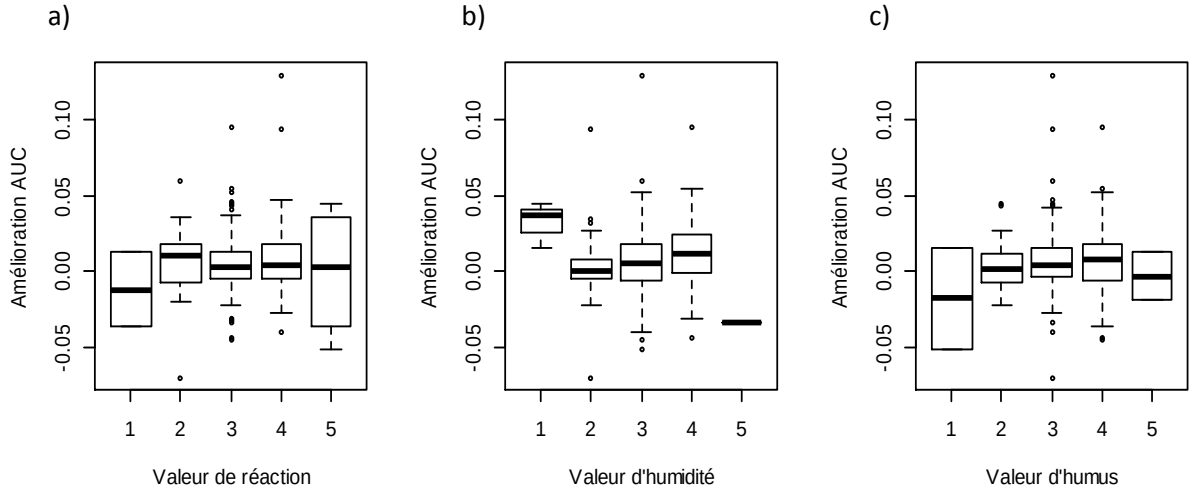



***Fig. 15 :*** *Indice AUC en fonction des valeurs de réaction (a), d'humidité (b) et d'humus (c).*

## 3. Comparaison des communautés prédites et observées

La comparaison entre les communautés prédites et observées sur le terrain se fait au moyen des indices élaborés à partir du tableau de contingence. L'indice le plus pertinent dans notre cas est l'indice de similarité puisqu'il prend toutes les présences et absences en considération. En comparant le modèle topo-climatique seul au modèle complété des données de pH, on s'aperçoit que les valeurs obtenues pour la similarité sont très légèrement supérieures pour le modèle intégrant les données de pH (Fig. 16c) (voir aussi Annexe 9). La valeur de similarité obtenue pour le modèle sans pH est de 0.312, celle du modèle avec pH de 0.328. Concernant la sensibilité, la moyenne du modèle avec pH est également légèrement supérieure (0.335 par rapport à 0.326 du modèle sans pH) alors que celle de la spécificité est presque identique (0.9152 contre 0.9153 pour le modèle avec pH).

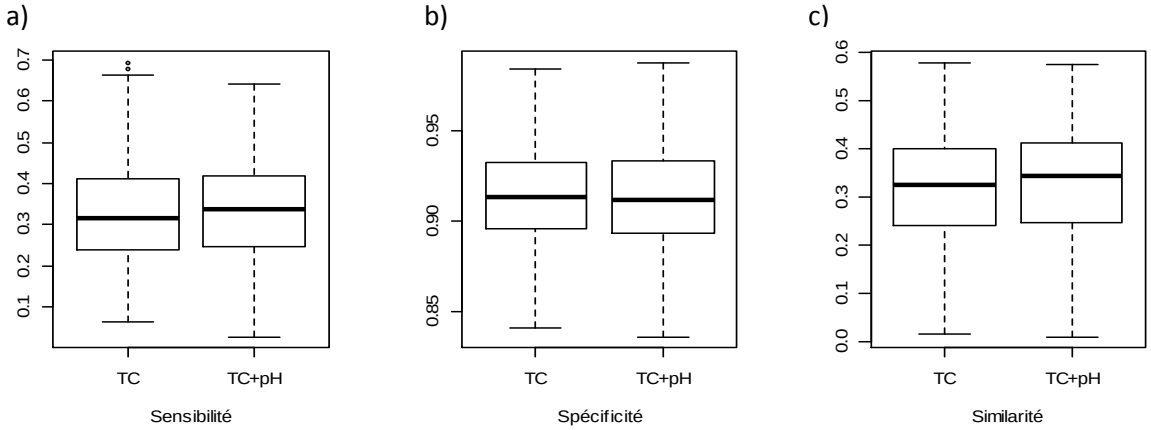



***Fig. 16 :*** *Diagrammes en boîte de la sensibilité (a), de la spécificité (b) et de la similarité (c) moyennes des 1'000 communautés de chaque relevé du jeu de données indépendant des modèles topo-climatiques et topo-climatiques avec pH.*

Pour les deux modèles, l'observation de ces mêmes indices en fonction des caractéristiques topographiques (Fig. 17, voir aussi Annexe 10) montre que la sensibilité et la similarité diminuent avec l'augmentation de l'altitude, de la pente et de la position topographique alors que la spécificité s'élève avec l'augmentation de celles-ci. Les modèles sont donc moins bons pour prédire les présences lorsque l'altitude ou la pente augmentent alors qu'au contraire, les absences sont de mieux en mieux prédites.

En comparant les graphiques des indices des deux modèles en fonction de l'altitude, on s'aperçoit qu'il y a une très légère amélioration pour le modèle avec pH puisque les points en rouge – un point correspondant à l'indice pour un relevé du jeu de données indépendant – et la droite de régression en rouge du modèle avec pH sont sensiblement plus hauts que ceux figurant en bleu (Fig. 17, voir aussi Annexe 10). Cette même tendance peut être observée en examinant les droites de régression des modèles en fonction de la pente et de la position topographique mais là également la variation est faible (Annexes 11, 12).

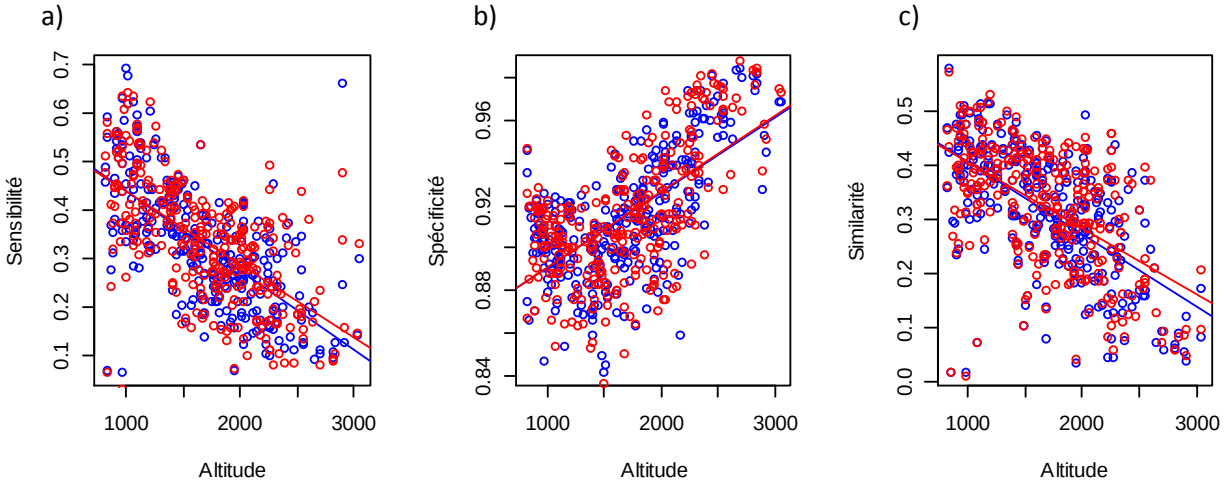



***Fig. 17 :*** *Diagrammes en nuage de points des modèles topo-climatique (en bleu) et topo-climatique avec pH (en rouge) pour la sensibilité (a), la spécificité (b) et la similarité (c) moyennes des 1'000 communautés de chaque relevé du jeu de données indépendant en fonction de l'altitude. La droite de régression pour les modèles topo-climatiques avec et sans pH est la même pour la spécificité.*

L'observation de l'erreur dans la prédiction du nombre d'espèces par rapport aux espèces réellement comptabilisées sur le terrain (richesse spécifique) montre qu'en moyenne les deux modèles évaluent plutôt bien la richesse spécifique des communautés puisque celle-ci est proche de 0 (Fig. 18) mais il y a peu de changement entre le modèle topo-climatique et celui complété des données de pH.

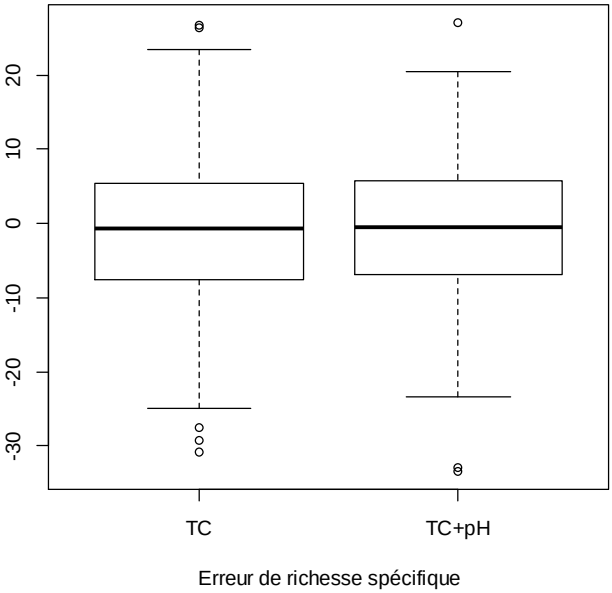



***Fig. 18 :*** *Diagrammes en boîte de l'erreur moyenne des 1'000 communautés de chaque relevé du jeu de donnée indépendant entre richesse spécifique observée et prédite par les modèles topo-climatique et topo-climatique avec pH.*

Enfin, la distribution de cette même erreur en fonction de l'altitude, de la pente et de la position topographique, montre là encore qu'il y a très peu de différence entre les deux modèles, les droites de régression étant quasiment identiques (Fig. 19, voir aussi Annexe 13). Les valeurs obtenues sont proches de 0 dans les deux cas, le nombre d'espèces prédit par les modèles étant très proche du nombre comptabilisé lors des relevés phytosociologiques. Avec l'augmentation de l'altitude, on s'aperçoit que les modèles ont tendance à sous-évaluer le nombre d'espèces dans les communautés.

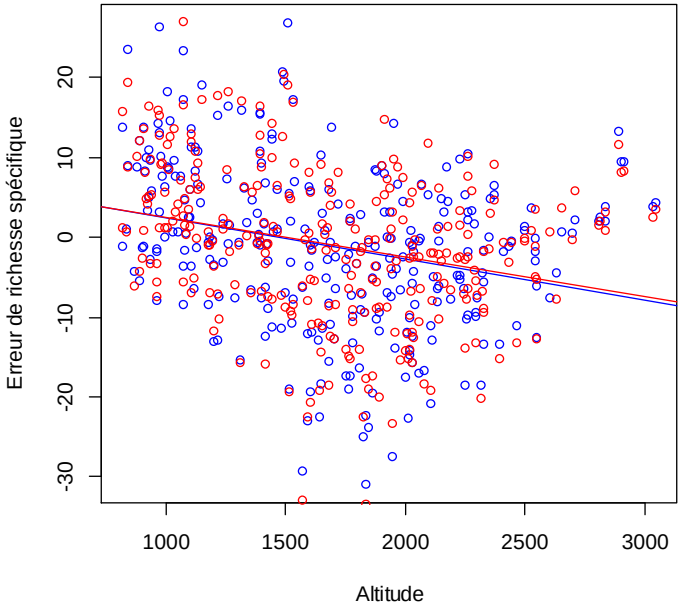



***Fig. 19 :*** *Diagramme en nuage de points de l'erreur de richesse spécifique moyenne pour chaque relevé du jeu de données indépendant des modèles topo-climatique (en bleu) et topo-climatique avec pH (en rouge) en fonction de l'altitude.*

## DISCUSSION

Suite à l'élaboration de la carte du pH pour la zone d'étude, les données de pH obtenues ont été ajoutées aux prédicteurs topo-climatiques couramment utilisés afin de déterminer si les projections des espèces et des communautés ont été améliorées en intégrant une variable édaphique dans les modèles de distribution des espèces. L'observation des changements de valeurs d'AUC et TSS montre qu'il y a effectivement eu une légère amélioration des prédictions pour certaines espèces en ajoutant les données de pH aux données topo-climatiques usuelles. Le nombre d'espèces concernées par une augmentation de l'indice AUC supérieure à 0.02 reste cependant modeste (43) puisqu'il concerne moins d'un quart de toutes les espèces et que cette amélioration est de 0.0393 en moyenne. Une seule espèce (*Valeriana montana*) a obtenu une augmentation supérieure à 0.1. Le nombre d'espèces pour lesquelles les données de pH ont au contraire détérioré les prévisions est de 77, soit 37% des espèces avec une moyenne de détérioration de l'AUC de 0.0119. La pertinence de l'intégration des données de pH dans les modèles topo-climatiques dépend donc des

espèces modélisées, certaines obtenant une projection plus proche de la réalité alors que d'autres s'en éloignent au contraire.

Lors de la comparaison de l'amélioration de l'indice AUC en fonction des traits biologiques et valeurs écologiques de Landolt, seules les espèces poussant sur des sols secs présentant une valeur d'indice d'humidité de 1 ont été légèrement mieux prédites. Ce résultat peut s'expliquer du fait, qu'étant donné que les précipitations y sont moins abondantes, les processus de lessivage et de lixiviation provoquant l'acidification des sols sont moins prononcés sur les sols secs et on y trouve des valeurs de pH plus élevées. En intégrant des données de pH, les espèces poussant sur ce type de terrain peuvent donc être légèrement mieux prédites que par les modèles topo-climatiques traditionnels. Seules 2 espèces présentant une valeur écologique d'humidité de 1 figurent cependant parmi les espèces présentant une amélioration de leur indice AUC supérieure à 0.02, ce qui limite la validité de ce résultat. En définitive, aucune amélioration des prédictions pour une catégorie de conditions écologiques particulières n'a pu être déterminée.

Les résultats de cette étude pour les modèles des espèces sont décevants par rapport à ce que nous attendions suite au travail de Dubuis et al. (in revision). En effet, les auteurs avaient trouvé une amélioration de 47% des espèces testées avec un indice d'AUC moyen de 0.05 en ajoutant les données de pH aux modèles. Dans notre cas, seuls 20% des prédictions des espèces ont été améliorées, avec une moyenne de l'indice AUC plus basse. Les espèces améliorées dans leur études étaient typiquement acidophiles, telles que *Vaccinium myrtillus* et *Nardus stricta* alors que dans nos résultats, l'espèce avec la plus grande amélioration est plutôt calcicole (Lauber et Wagner, 2007). Les auteurs ont également trouvé une relation entre l'amélioration des modèles et les conditions écologiques préférentielles des espèces ce qui n'est pas le cas ici. Ces différences de résultats peuvent s'expliquer du fait que Dubuis et al. ont utilisé des valeurs de pH mesurées directement et non pas déduites d'une modélisation. Ces valeurs correspondent donc au pH réel sur le terrain, ce qui apporte probablement une information plus précise qui a permis d'améliorer les modèles de distribution d'espèces. Dans notre étude, les valeurs de pH utilisées ont été dérivées d'un modèle qui n'a pas obtenu une qualité d'ajustement très élevée. L'information contenue dans ces données est donc moins pertinente et n'a malheureusement pas produit d'amélioration suffisante pour les prédictions des espèces. Notons toutefois que dans notre cas le modèle a été testé sur un jeu de données indépendant et non pas par validation croisée comme l'ont fait Dubuis et al., ce qui rend son évaluation plus stricte.

Etant donné que les prédictions pour les espèces n'ont pas été améliorées par l'ajout du prédicteur pH dans le modèle, la prédiction des communautés peut difficilement être meilleure. On observe cependant une légère amélioration de la moyenne des indices de sensibilité et de similarité pour le modèle avec pH, ce qui nous laisse penser que le pH peut être un prédicteur intéressant mais que le modèle élaboré dans cette étude n'est pas suffisamment précis pour engendrer de nette amélioration. Pour les deux sets de modèles, la sensibilité et la similarité diminuent avec l'augmentation de l'altitude, de la pente et de la position topographique alors que la spécificité s'élève. Les présences sont ainsi de moins en moins bien prédites alors que la prédiction des absences s'améliore au contraire. Ceci vient probablement du fait que les espèces sont moins nombreuses dans ces conditions et que le nombre d'occurrences est trop faible pour obtenir une bonne calibration des modèles. Les

modèles vont donc plutôt prédire des absences, et donc la spécificité du modèle des communautés s'améliore.

Au niveau de la richesse spécifique, nous avons vu qu'en moyenne le nombre d'espèces prédites dans les communautés par nos modèles est plutôt bon, avec une erreur de prédiction de richesse spécifique proche de 0 mais que les modèles intégrant les données de pH n'ont pas été meilleurs que les modèles topo-climatiques traditionnels pour prédire les communautés végétales. Le nombre d'espèces prédites par les modèles a tendance à être sous-évalué avec l'augmentation de l'altitude, probablement également en raison d'une mauvaise calibration des modèles des espèces à plus haute altitude.

Dans notre cas, et contrairement à Peppler-Lisbach et Schröder (2004), les données de pH n'ont pas amélioré les prédictions pour les communautés. En effet, dans leur étude les auteurs ont mis en évidence l'importance du pH pour la modélisation des communautés végétales. Pour ce faire, ils ont élaboré un modèle de régression linéaire pour chaque espèce faisant partie des communautés de *Nardus* en couplant les données de présence-absence des ces espèces à des prédicteurs reflétant l'utilisation du sol, les conditions climatiques et les caractéristiques édaphiques. Les prédictions des espèces ont ensuite été utilisées afin de prédire des communautés sur la base des probabilités de présence des espèces. Selon leurs résultats, 138 espèces ont été modélisées avec succès, les prédicteurs les plus souvent utilisés étant les précipitations et le pH. Au niveau des communautés, les auteurs ont analysé le taux de prédictions correctes, la sensibilité et la spécificité ainsi que la richesse spécifique. Leurs résultats ont révélé des valeurs conformes à la réalité, avec un indice de sensibilité de 0.67 et un indice de spécificité de 0.90. Dans cette étude, les auteurs n'ont cependant modélisé que des communautés de Nardus et sur la base d'espèces propres à ces communautés, ce qui explique que leurs résultats soient meilleurs que les nôtres, dans lesquels la sensibilité n'a été que de 0.33 et la spécificité de 0.91. Dans notre cas en effet, nous avons cherché à appréhender toutes les communautés possibles sur la base des 207 espèces modélisées, ce qui augmente considérablement le risque d'erreur du modèle.

En définitive, l'amélioration des prédictions au niveau des espèces est mitigée étant donné que moins d'un quart d'entre elles seulement a obtenu une amélioration de l'indice AUC supérieure à 0.02. La considération des traits fonctionnels et valeurs écologiques de Landolt n'a pas permis d'expliquer les résultats obtenus. Enfin, concernant les communautés prédites par le modèle, les résultats sont bons mais n'ont pas été améliorés par l'ajout des données de pH dans les projections.

# CONCLUSION ET PERSPECTIVES

Contrairement aux études précédentes, dont celle de Dubuis et al. (in revision), l'ajout de données de pH dans les modèles de distribution d'espèces n'a pas donné les résultats attendus. Les modèles intégrant les données de pH ont amélioré les prédictions pour certaines espèces mais leur nombre reste faible et la différence par rapport au modèle topo-climatique est minime. La raison vient probablement du modèle élaboré pour obtenir les données de pH qui n'a pas obtenu une qualité d'ajustement très élevée. Les résultats incertains d'un modèle étant ici utilisés pour élaborer un deuxième modèle, l'erreur ne peut en être qu'augmentée. Aussi, les données de pH n'ont-elles pas été très pertinentes dans les modèles testés sur les espèces végétales. Ceux-ci ont au contraire donné plus d'importance aux prédicteurs degrés-jours, humidité et pente ce qui fait qu'au final le modèle contenant les données de pH est très semblable au modèle topo-climatique simple. La ressemblance des résultats obtenus le confirme.

Les modèles de distribution nécessitent encore d'être améliorés pour permettre d'obtenir des prédictions fiables concernant les espèces et communautés végétales. Dans cette étude, le set de modèles topo-climatique n'a ainsi obtenu qu'un indice AUC moyen de 0.78, ce qui n'est pas suffisant pour pouvoir le qualifier de bon modèle. Nous avons vu que le pH est une variable difficile à modéliser et que le modèle élaboré dans cette étude n'a pas permis d'améliorer les prédictions. A l'avenir, il pourrait ainsi être intéressant de tester d'autres variables édaphiques lors de la modélisation des espèces végétales. La profondeur du sol, mesurée à l'aide de sondages à la tarière et modélisée à l'aide de la pente et de la position topographique, est par exemple une propriété du sol plus stable que le pH et pourrait donner des informations quant à la stabilité et au degré d'évolution du sol, permettant également d'appréhender le pH de manière indirecte.

# REMERCIEMENTS

Je tiens en particulier à remercier le Dr Pascal Vittoz, Anne Dubuis et Loïc Pellissier doctorants au Département d'Ecologie et Evolution pour leurs aide et conseils ainsi que toutes les personnes qui m'ont accompagnée ou m'ont soutenue durant mon terrain.

# REFERENCES BIBLIOGRAPHIQUES


Akaike, H. (1973). Maximum likelihood identification of Gaussian autoregressive moving average models. *Biometrika, 60,* 255-265.

Allouche, O., Tsoar, A. et Kadmon, R. (2006). Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS). *Journal of Applied Ecology, 43,* 1223-1232.

Anderson, R.P., Peterson, A.T. et Gomez-Laverde, M. (2002). Using niche-based GIS modelling to test geographic predictioins of competitive exclusion and competitive release in South American pocket mice. *Oikos, 98,* 3-16.

Araùjo, M.B. et al. (2004). Would climate change drive species out of reserves? An assessment of existing reserve-selection methods. *Global Change Biology, 10,* 1618-1626.

Austin, M.P. (1980). Searching for a model for use in vegetation analysis. *Vegetatio, 42,* 11-21.

Austin, M.P. et Heylinger, P.C. (1989). Vegetation survey design for conservation: gradsec sampling of forests in north-east New South Wales. *Biological Conservation, 50,* 13-32.

Austin, M.P., Nicholls, A.O. Et Margules, C.R. (1990). Measurement of the realized qualitative environmental niches of five *Eucalyptus* species. *Ecological Monographs, 60,* 161-177.

Baize, D. et Girard, M.-C. (2009). *Référentiel pédologique 2008.* Versailles: Éditions Quae.

Bouët, M. (1985). *Climat et météorologie de la Suisse romande*. Lausanne: Payot.

Braun-Blanquet, J. (1928). *Pflanzensociologie. Grundzüge der Vegetationskunde* (1ère éd.). Berlin.

Breiman, L. (2001). Random Forests. *Machine Learning, 45(1),* 5-32.

Breiman, L. (2002). Manual On Setting Up, Using, And Understanding Random Forests V3.1.

Brown, D.G. (1994). Predicting vegetation types at treeline using topography and biophysical disturbance variables. *Journal of Vegetable Science, 5,* 641-656.

Brzeziecki, B., Kienast, F. et Wildi, O. (1993). A simulated map of the potential natural forest vegetation of Switzerland. *Journal of Vegetable Science, 4,* 499-508.

Castrignano, A. et al. (2011). Using Digital Elevation Model to Improve Soil pH Prediction in an Alpine Doline. *Pedosphere, 21(2),* 259-270.

Cornelissen, J.H.C. et al. (2003). A handbook of protocols for standardised and easy measurement of plant functional traits worldwide. *Australian Journal of Botany, 51,* 335-380.

Coudun, C. et Gégout, J.-C. (2005). Ecological behaviour of herbaceous forest species along à pH gradient: a comparion between oceanic and semicontinental regions in northern France. *Global Ecology and Biogeography, 14,* 263-270.

Coudun, C. et 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.

Coudun, C. et al. (2006). Soil nutritional factor improve models of plant species distribution: an illustration with *Acer campestre* (L.) in France. *Journal of Biogeography, 33,* 1750-1763.

Cox, D.R. et Snell, E.J. (1989). *The Analysis of Binary Data* (2e éd.). London: Chapman and Hall.

Delarze, R. et Gonseth, Y. (2008). *Guide des milieux naturels de Suisse* (2e éd.). Bussigny : Rossolis.

Dubuis, A. et al. (2011). Predicting spatial patterns of plant species richness: a comparison of direct macroecological and species stacking modelling approaches. *Diversity and Distributions, 17(6),* 1122-1131.

Dubuis, A. et al. (in revision).Edaphic variables improve predictions of plant species distributions in a mountainous landscape. *Journal of Vegetation Science.*

Duchaufour, P. (1989). Pédologie et groupes écologiques. I. Rôle du type d'humus et du pH. *Bulletin d'écologie, 20,* 1-6.

Engler, R. et al. (2009). Predicting future distributions of mountain plants under climate change: does dispersal capacity matter? *Ecography, 32*, 34-45.

Esher, A. et al. (1997). Geological framework and structural evolution of the Western Swiss-Italian Alps. In *Deep sturcture of Switzerland - Results of NFP-20.*, pp. 205-222. Birkhäuser Basel.

ESRI (2010). *ArcGIS 10*. Redlands UK: Environmental Systems Research Institute, Inc.

Ferrier, S. (2002). Mapping spatial pattern in biodiversity for regional conservation planning. Where to from here? *Syst. Biol., 51,* 331-363.

Ferrier, S. et Guisan, A. (2006). Spatial modelling of biodiversity at the community level. *Journal of Applied Ecology, 43,* 393-404.

Fielding, A.H. et Bell, J.F. (1997). A review of methods for the assessment of prediction errors in conservation presence-absence models. *Environmental Conservation, 24,* 38-49.

Fischer, H.S. (1990). Simulating the distribution of plant communities in an alpine landscape. *Coenoses, 5,* 37-43.

Friedman, J.H. (2001). Greedy function approximation: A gradient boosting machine. *Annals of Statistics, 29,* 1189-1232.

Friedman, J.H., Hastie, T.J., and Tibshirani, R. (2000). Additive logistic regression: a statistical view of boosting. *Annals of Statistics, 28,* 337-374.

Glick, B. R. (1995). The enhancement of plant growth by free-living bacteria. *Canadian Journal of Microbiology, 41,* 109-117

Gobat, J.-M., Aragno, M. et Matthey, W. (2010). *Le sol vivant* (3e éd.). Lausanne : Presse Polytechniques et Universitaires Romandes.

Gough, L. et al. (2000). Vascular plant species richness in Alaskan arctic tundra: the importance of soil pH. *Journal of Ecology, 88,* 54-66.

Guisan, A. (2003). Simuler la répartition géographique des espèces et de la végétation (ou "si De Candolle avait eu un ordinateur... "). *Saussurea, 33,* 79-99.

Guisan, A., Theurillat, J.-P. et Kienast, F. (1998). Predicting the potential distribution of plant species in an alpine environment. *Journal of Vegetable Science, 9,* 65-74.

Guisan, A. et Thuiller, W. (2005). Predicting species distribution : offering more than simple habitat models. *Ecology Letters, 8,* 993-1009.

Guisan, A., Weiss, S.B. et Weiss, A.D. (1999). GLM versus CCA spatial modelling of plant species distribution. *Plant Ecology, 143,* 107-122.

Guisan, A. et Zimmermann, N.E. (2000). Predictive habitat distribution models in ecology. *Ecological Modelling, 135,* 147-186.

Harrel, F.E.Jr (2012). *Regression Modelling Strategies*. Redlands UK: Environmental Systems Research Institute, Inc.

Hastie, T. et Tibshirani, R. (1990). *Generalized Additive Models*. London: Chapman and Hall.

Herbauts, J. et Tanghe, M. (1987). Relations entre formations superficielles, sols et associations forestières sur la cuesta bajocienne de Lorraine belge. L'exemple du bois de la côte à Virton-Saint-Mard. *Bulletin de la Société Royale de Botanique de Belgique, 120, 93-105.*

Hirzel, A. et Guisan, A. (2002). Which is the optimal sampling strategy for habitat suitability modelling. *Ecological Modelling, 157,* 331-341.

Hutchinson, G.E. (1957). Concluding remarks. *Cold Spring Harbor Symposia on Quantitative Biology, 22,* 415–427.

Huntley, B. et al. (2004). The performance of models relating species geographical distributions to climate is independant of trophic level. *Ecology Letters, 7,* 417-426.

Kachitvichyanukul, V. et Schmeiser, B.W. (1988). Binomial random variate generation. *Communications of the ACM*, *31*, 216–222.

Kloepper, J. W. et Schroth, M. N. (1978). *Plant growth-promoting rhizobacteria on radishes* in *Proc. of the 4th Internat. Conf. on Plant Pathogenic Bacter (Vol. 2).* Angers: INRA, Station de Pathologie Végétale et Phytobactériologie, p. 979-882. .

Kohler, R. , Kosakyan, A. et Scherrer, L. (2010). *Travaux pratiques analyses pédologiques & matière organique.* Neuchâtel : Laboratoire Sol et Végétation et Biologie du Sol.

Körner, C. (2003). *Alpine plant life* (2e éd.). Berlin: Springer.

Landolt, E. et al. (2010). *Flora Indicativa. Ecological indicator values and biological attributes of the Flora of Switzerland and the Alps.* Bern: Haupt Verlag.

Lauber, K. et Wagner, G. (2007). *Flora Helvetica, Flore illustrée de Suisse* (3e éd.).Berne : Haupt.

McBratney, A.B., Mendonça Santos, M.L. et Minasny,B. (2003). On digital soil mapping. *Geoderma, 117,*3-52.

McCullagh, P. et Nelder, J.A. (1989). *Generalized Linear Models* (2e éd.). London : Chapman and Hall.

McKenzie, N.J. et Ryan, P.J. (1999). Spatial prediction of soil properties using environmental correlation. *Geoderma, 89,* 67-94.

McPherson, J.M., Jetz, W. et Rogers, D.J. (2004). The effects of species' range sizes on the accuracy of distribution models: ecological phenomenon or statistical artefact? *Journal of Applied Ecology, 41,* 811-823.

Moore, D.M., Lees, B.G. et Davey, S.M. (1991). A new method for predicting vegetation distributions using decision tree analysis in a geographic information system. *Environmental Management, 15,* 59-71.

Nagelkerke, N.J.D. (1991). A note on a general definition of the coefficient of determination. *Biometrika, 78,* 691-692.

Pearce, J. et Lindenmayer, D. (1998). Bioclimatic analysis to enhance reintroduction biology of the endangered helmeted honeyeater (*Lichenostomus melanops cassidix*) in southeastern Australia. *Restor. Ecol., 6,* 238-243.

Pearson, R.G. et al. (2002). SPECIES: a spatial evaluation of climate impact on the envelope of species. *Ecological Modelling, 154,* 289-300.

Pellissier, L. et al. (2010). Spatial pattern of floral morphology: possible insight into the effects of pollinators on plant distributions. *Oikos, 119,* 1805-1813.

Peppler-Lisbach, C. et Schröder, B. (2004). Predicting the species composition of *Nardus stricta* communities by logistic regression modelling. *Journal of Vegetation Science, 15,* 623-634.

Pottier et al. (in press). Accuracy of plant assemblage prediction from species distribution models varies along environmental gradients.

R Development Core Team. R 2.14. Vienne : R Foundation for Statistical Computing.

Randin, C.F et al. (2006). Are niche-based species distribution models transferable in space? *Journal of Biogeography, 33,* 1689-1703.

Randin, C.F. et al. (2009a). Land use improves spatial predictions of mountain plant abondance but not presence-absence. *Journal of Vegetation Science, 20,* 996-1008.

Randin, C.F. et al. (2009b). Introduction of Snow and Geomorphic Disturbance Variables into Predictive Models of Alpine Plant Distribution in the Western Swiss Alps. *Arctic, Antarctic, and Alpine Research, 41,* 347-361.

Randin, C.F. et al. (2009c). Climate change and plant distribution: local models predict high-elevation persistence. *Global Change Biology, 15,* 1557-1569.

Ridgeway, G. (1999). The state of boosting. *Computing Science and Statistics, 31,* 172-181.

Schloeder, C.A., Zimmermann, N.E. et Jacobs, M.J. (2001). Division S-8 – Nutrient Management & Soil & Plant Analysis. Comparison of Methods for Interpolating Soil Properties Using Limited Data. *Soil Science Society of America Journal, 65,* 470-479.

Shi, W. et al. (2009). Surface modelling of soil pH. *Geoderma, 150,* 113-119.

Shi, W. et al. (2011). Surface modelling of soil properties based on land use information. *Geoderma, 162,* 347-357.

Spaltenstein, H. (1984). *Pédogenèses sur calcaire dur dans les Hautes Alpes calcaires*. Lausanne: EPFL

Thomas, C.D. et al. (2004). Extinction risk from climate change. *Nature, 427,* 145-147.

Thuiller, W. (2004). Patterns and uncertainties of species' range shifts under climate change. *Global Change Biology, 10, 2020-2027.*

Thuiller, W. et al. (2009). BIOMOD - a platform for ensemble forecasting of species distributions. *Ecography, 32,* 369-373.

Université de Neuchâtel (2007). *Protocole de description des sols.* Neuchâtel : Institut de botanique.

Violle, C. et al. (2007). Let the concept of trait be functional ! *Oikos, 116,* 882-892.

Weiher, E. et al. (1999). Challenging Theophrastus: A common core list of plant traits for functional ecology. *Journal of Vegetation Science, 10,* 609-620.

Wiesmeier, M. et al. (2010). Digital mapping of soil organic matter stocks using Random Forest modelling in a semi-arid steppe ecosystem. *Plant Soil, 340,* 7-24.

Zimmermann, N.E. et Kienast, F. (1999). Predictive mapping of alpine grasslands in Switzerland: Species versus community approach. *Journal of Vegetation Science, 10,* 469-482.

**Cartes topographiques et géologiques :**

Carte nationale de la Suisse au 1 :25'000. Feuille 1244 *Châtel-St-Denis.* Wabern : Office fédéral de topographie.

Carte nationale de la Suisse au 1 :25'000. Feuille 1245 *Château-d'Oex.* Wabern : Office fédéral de topographie.

Carte nationale de la Suisse au 1 :25'000. Feuille 1264 *Montreux.* Wabern : Office fédéral de topographie.

Carte nationale de la Suisse au 1 :25'000. Feuille 1265 *Les Mosses.* Wabern : Office fédéral de topographie.

Carte nationale de la Suisse au 1 :25'000. Feuille 1284 *Monthey.* Wabern : Office fédéral de topographie.

Carte nationale de la Suisse au 1 :25'000. Feuille 1285 *Les Diablerets.* Wabern : Office fédéral de topographie.

Carte nationale de la Suisse au 1 :25'000. Feuille 1304 *Val d'Illiez.* Wabern : Office fédéral de topographie.

Carte nationale de la Suisse au 1 :25'000. Feuille 1305 *Dent de Morcles.* Wabern : Office fédéral de topographie.

Carte tectonique de la Suisse au 1 :50'000. Wabern : Office fédéral de topographie.

Badoux, H. (1965). *Montreux.* Feuille 47 de l'Atlas géologique de la Suisse au 1 :25'000. Berne : Commission géologique suisse.

Badoux, H. et al. (1960). *Monthey.* Feuille 37 de l'Atlas géologique de la Suisse au 1 :25'000. Berne : Commission géologique suisse.

Badoux, H. et al. (1971). *Dent de Morcles.* Feuille 58 de l'Atlas géologique de la Suisse au 1 :25'000. Berne : Commission géologique suisse.

Badoux, H., Gabus, J.H. et Mercanton, C.H. (1990). *Les Diablerets.* Feuille 88 de l'Atlas géologique de la Suisse au 1 :25'000. Berne : Commission géologique suisse.

Gagnebin, E. et al. (1934). *Val d'Illiez.* Feuille 8 de l'Atlas géologique de la Suisse au 1 :25'000. Berne : Commission géologique suisse.

Lombard, A. et al. (1974). *Les Mosses.* Feuille 64 de l'Atlas géologique de la Suisse au 1 :25'000. Berne : Commission géologique suisse.

Weidmann, M. et al. (1993). *Châtel-St-Denis.* Feuille 92 de l'Atlas géologique de la Suisse au 1 :25'000. Berne : Commission géologique suisse.

# ANNEXES

1. Protocole de mesure du pH (Kohler, Kosakyan et Scherrer, 2010).
2. Valeurs de pH mesurées (2011)
3. Corrélations entre prédicteurs
4. Amélioration AUC et TSS moyenne des 4 modèles due au pH
5. Tableau des espèces avec amélioration de l'AUC supérieure à 0.02 et leurs indices et traits biologiques
6. Amélioration des indices AUC et TSS en fonction des traits biologiques
7. Amélioration de l'indice AUC en fonction des valeurs écologiques de Landolt
8. Amélioration de l'indice TSS en fonction des valeurs écologiques de Landolt
9. Indices d'évaluation des communautés
10. Régressions linéaires des indices pour les modèles topo-climatique et topo-climatique avec pH en fonction de l'altitude
11. Régressions linéaires des indices pour les modèles topo-climatique et topo-climatique avec pH en fonction de la pente
12. Régressions linéaires des indices pour les modèles topo-climatique et topo-climatique avec pH en fonction de la position topographique
13. Diagrammes de l'erreur de richesse spécifique en fonction de la pente, de l'altitude et de la position topographique

# 1. Protocole de mesure du pH



10. Après refroidissement en dessiccateur, peser à nouveau le creuset et les cendres. Noter sa masse **"B"**.

### Récupération

Le sol est à descendre au sous-sol dans les bacs bleus à côté de la chambre climatisée

## 5. Le pH

### Principes

De manière générale, le pH représente la concentration de ions $H^+$ libres dans une solution. Il s'exprime selon une échelle de 0 à 14. Les valeurs faibles indiquent l'acidité (vinaigre, jus de citron), les valeurs hautes la basicité (savon) tandis que la neutralité se situe aux environs de 7.

Dans le sol, deux types de pH peuvent être mesurés ($pH_{H2O}$) et $pH_{KCl}$).

Ainsi, on distingue l'acidité active du sol qui représente les ions H+ libres dans la solution de sol et donc capables d'agir dans sur l'électrode du pH-mètre dès que l'on réhumecte le sol. Ces ions sont cependant en équilibre avec les $H^+$ fixés sur le complexe argilo-humique (soit la partie "solide" du sol) et dont la quantité est 1'000 à 100'000 fois plus grande. Ces $H^+$ fixés représentent l'acidité potentielle du sol.

*Le pH $H_2O$*

Il rend compte de l'acidité libre ou acidité active, c'est-à-dire de l'activité des ions $H_3O^+$ (protons) présents à l'état dissocié dans la phase aqueuse du sol. On obtient donc la mesure du pH de l'eau du sol et des éléments qui y sont dissout.

Ces ions en solution sont en équilibre dynamique avec les protons fixés sur le complexe organo-minéral (appelé aussi complexe adsorbant ou complexe d'échange) du sol qui agit comme un acide faible :

$$HA\ (sol) + H_2O \leftrightarrow A^-\ (sol) + H_3O^+$$

### Méthode

*a) Matériel*

- ✓ Gobelets plastiques ou béchers de 100 ml.
- ✓ Baguette de verre.
- ✓ pH-mètre avec électrode pH.

*b) Réactifs*

- ✓ Eau désionisée
- ✓ Solutions tampon pH = 4.00 et pH = 7.00. *Attention, les solutions tampon doivent impérativement être à température ambiante lors du calibrage du pH-mètre. Il ne faut donc pas oublier de les sortir à l'avance du frigo où elles sont conservées.*



- ✓ KCl 1 M: dissoudre 74.55 g de KCl en poudre dans 1 litre d'eau désionisée.

### Mode opératoire

*a) Echantillon de départ :*

- ✓ Sol séché à l'air et tamisé à 2 mm (= terre fine), ou
- ✓ Tourbe fraîche ou tourbe séchée à l'air, éviter les grosses racines et les branchages.

*b) Blancs*

Il n'est pas nécessaire de faire de blancs, mais il faut bien calibrer le pH-mètre.

*c) Procédure*

11. **Placer** de la terre fine dans un gobelet plastique (ou un bécher de 100 ml).
12. **Ajouter** environ **50 ml d'eau désionisée**. L'important est ici de travailler avec un mélange sol/liquide dans la **proportion 1 : 2.5**. Par exemple, si vous ne pouvez travailler que sur 10 g de sol séché à l'air, il vous faudra ajouter 25 ml d'eau désionisée.
13. Bien **homogénéiser** le mélange sol-eau ou tourbe-eau en brassant énergiquement avec une baguette de verre. Veiller à rincer la baguette de verre avec de l'eau désionisée entre les échantillons afin d'éviter de polluer ceux-ci!
14. **Laisser** macérer au moins **1h** (2h sont nécessaire pour les sols calcaires) en remuant de temps à autre avec la baguette.
15. **Etalonner** le pH-mètre avec les solutions tampons pH 7 et pH 4.
16. **Mesurer le pH** en plaçant l'électrode dans la suspension de sol et en prenant toujours soin d'agiter légèrement le milieu avec un poisson de petite taille. **Noter** la valeur du pH après stabilisation.
17. Après chaque mesure, rincer l'électrode à l'eau désionisée et l'essuyer avec du papier.

### Récupération

- ✓ Verser délicatement un maximum de solution (eau ou KCl) à l'évier sous flux d'eau important puis jeter le restant de sol à la poubelle.

## 6. La granulométrie en 5 fractions, sans décarbonatation

### Principe

Le but de l'analyse granulométrique est de déterminer la distribution pondérale des différentes fractions texturales de la partie minérale d'un sol.

Les particules minérales, après destruction des agrégats, se répartissent dans diverses catégories de grosseur regroupées selon l'échelle internationale illustrée à la figure 17.1.

## 2. Valeurs de pH mesurées (échantillons 2011)

| N° relevé | X | Y | pH |
|---|---|---|---|
| 10 | 583602.0 | 147106.0 | 6.4 |
| 12 | 580501.0 | 127198.0 | 8.2 |
| 13 | 580218.0 | 151855.3 | 6.1 |
| 15 | 570801.2 | 135749.7 | 6.1 |
| 20 | 578730.7 | 123459.3 | 8.3 |
| 22 | 568601.0 | 124902.0 | 7.4 |
| 26 | 571496.8 | 135100.8 | 5.2 |
| 27 | 565430.7 | 137557.8 | 6.5 |
| 32 | 572200.0 | 124598.0 | 5.8 |
| 34 | 578931.4 | 125661.6 | 5.5 |
| 35 | 564731.4 | 137559.6 | 5.9 |
| 38 | 557146.5 | 150027.3 | 5.5 |
| 39 | 576433.0 | 142861.3 | 5.5 |
| 43 | 577487.0 | 121970.0 | 9 |
| 44 | 578526.4 | 126655.7 | 5.2 |
| 48 | 577830.3 | 139060.5 | 7.3 |
| 49 | 571828.7 | 118061.6 | 7.5 |
| 53 | 571573.0 | 123826.0 | 6 |
| 58 | 574603.9 | 122198.4 | 7.5 |
| 60 | 577927.5 | 140659.6 | 5.3 |
| 64 | 578015.9 | 140559.6 | 5.6 |
| 65 | 578310.0 | 139049.6 | 5.4 |
| 66 | 568733.1 | 142158.9 | 5.3 |
| 67 | 568127.0 | 129099.0 | 6.2 |
| 69 | 576598.5 | 141304.5 | 7.3 |
| 74 | 577300.1 | 140801.2 | 4.5 |
| 76 | 576052.6 | 147851.5 | 7 |
| 78 | 564227.3 | 144264.6 | 5.3 |
| 79 | 572025.2 | 148258.6 | 5.3 |
| 80 | 573900.5 | 119998.2 | 5.2 |
| 81 | 578330.9 | 150339.6 | 6.7 |
| 85 | 572917.1 | 147112.0 | 6.5 |
| 87 | 571123.0 | 124203.0 | 7.3 |
| 88 | 572598.0 | 129998.0 | 5.3 |
| 89 | 576130.3 | 143557.7 | 6.3 |
| 92 | 570603.0 | 132302.0 | 5.4 |
| 99 | 562633.0 | 134856.9 | 5.7 |
| 107 | 580834.4 | 140954.8 | 6 |
| 109 | 575005.8 | 122110.3 | 6.2 |
| 111 | 576598.0 | 138496.0 | 5.3 |
| 112 | 577718.2 | 125065.3 | 6.4 |
| 114 | 576899.1 | 140897.0 | 5.2 |
| 122 | 573730.7 | 144569.4 | 5.8 |
| 129 | 577995.0 | 136555.4 | 5.7 |
| 132 | 571331.0 | 123367.0 | 6 |
| 134 | 578331.4 | 136550.5 | 6.9 |
| 138 | 572791.0 | 130001.0 | 5.8 |
| 149 | 582903.0 | 148007.0 | 7.5 |
| 154 | 579624.1 | 135552.7 | 5.9 |

| N° relevé | X | Y | pH |
|---|---|---|---|
| 167 | 577940.6 | 135156.6 | 4.6 |
| 169 | 563938.0 | 135856.1 | 5.2 |
| 173 | 578831.6 | 145753.9 | 5.3 |
| 178 | 572330.5 | 117065.6 | 8.1 |
| 179 | 568343.0 | 126102.0 | 7.4 |
| 180 | 583296.0 | 147898.0 | 6.9 |
| 187 | 578726.3 | 145546.7 | 5.1 |
| 190 | 581115.0 | 140916.4 | 5.5 |
| 193 | 575531.3 | 142258.1 | 6.1 |
| 194 | 557699.9 | 149446.4 | 5.7 |
| 197 | 576003.2 | 122808.3 | 7.3 |
| 198 | 576073.8 | 122726.4 | 7.4 |
| 199 | 575893.6 | 123272.4 | 6.8 |
| 207 | 578499.0 | 146996.0 | 7.1 |
| 209 | 570501.0 | 134777.5 | 6.8 |
| 213 | 568200.0 | 126924.0 | 7.7 |
| 214 | 575305.0 | 138002.0 | 5.9 |
| 216 | 567493.0 | 129402.0 | 5.6 |
| 218 | 572204.0 | 124700.0 | 5.6 |
| 220 | 568323.0 | 130396.0 | 6.1 |
| 225 | 579500.0 | 146755.0 | 7.6 |
| 227 | 567424.3 | 138262.1 | 7.4 |
| 230 | 558928.0 | 150370.2 | 7.2 |
| 231 | 580002.0 | 147097.0 | 5.8 |
| 232 | 577445.5 | 127014.6 | 8.3 |
| 237 | 559275.4 | 151124.0 | 7.1 |
| 238 | 571029.2 | 146356.1 | 6.4 |
| 240 | 571868.0 | 133903.0 | 5 |
| 247 | 573103.4 | 118598.5 | 7.3 |
| 248 | 580701.0 | 126300.0 | 7.7 |
| 249 | 582124.0 | 152551.6 | 6.8 |
| 255 | 575601.0 | 138701.0 | 6 |
| 266 | 580523.5 | 151059.2 | 6.1 |
| 267 | 565435.8 | 139163.3 | 6.4 |
| 268 | 567200.0 | 121106.0 | 6 |
| 271 | 575301.0 | 128999.0 | 5.6 |
| 272 | 568503.0 | 119900.0 | 6.3 |
| 276 | 567737.0 | 119994.0 | 7 |
| 282 | 567394.0 | 119097.0 | 7.3 |
| 283 | 575299.8 | 121901.7 | 7.6 |
| 290 | 580391.8 | 139059.1 | 6.2 |
| 291 | 556095.9 | 147396.3 | 7.4 |
| 292 | 556500.7 | 147401.7 | 7.5 |
| 297 | 569403.0 | 124255.0 | 7.5 |
| 299 | 566900.0 | 128700.0 | 7.4 |
| 300 | 564447.3 | 145258.7 | 7.5 |
| 302 | 575801.0 | 122502.0 | 7.9 |
| 303 | 560867.8 | 142309.2 | 5.3 |
| 304 | 574998.5 | 121895.7 | 6.9 |

| N° relevé | X | Y | pH |
|---|---|---|---|
| 311 | 575500.0 | 140300.0 | 6.9 |
| 315 | 559399.1 | 146299.4 | 7.5 |
| 316 | 575702.0 | 139799.0 | 7.2 |
| 322 | 576401.0 | 139495.0 | 6.4 |
| 328 | 575699.0 | 127699.0 | 5.6 |
| 332 | 572432.0 | 117965.9 | 6.2 |
| 333 | 570529.0 | 133202.0 | 4.6 |
| 339 | 564233.3 | 141861.9 | 7.8 |
| 341 | 571831.3 | 117159.9 | 8.1 |
| 342 | 558978.9 | 150775.4 | 5.4 |
| 346 | 565724.8 | 140131.0 | 6.9 |
| 350 | 571926.9 | 117360.6 | 8.1 |
| 352 | 570934.4 | 146352.4 | 5.6 |
| 353 | 579732.9 | 125754.5 | 5 |
| 354 | 580129.9 | 125663.9 | 5.3 |
| 358 | 579599.3 | 126198.6 | 4.6 |
| 359 | 580600.0 | 126401.0 | 8 |
| 360 | 566306.2 | 134077.5 | 7.1 |
| 365 | 579498.0 | 126100.0 | 6.9 |
| 367 | 579515.6 | 125598.1 | 5.8 |
| 371 | 569772.6 | 134297.1 | 5.9 |
| 375 | 571084.2 | 137391.7 | 7.4 |
| 388 | 571931.2 | 118164.7 | 7.5 |
| 391 | 577198.0 | 133097.0 | 4.5 |
| 392 | 572335.3 | 117664.5 | 6.9 |
| 393 | 572032.1 | 116962.5 | 8.6 |
| 395 | 577300.0 | 122000.0 | 8.5 |
| 396 | 577300.0 | 122202.0 | 8.4 |
| 403 | 572330.9 | 118043.3 | 6.2 |
| 404 | 577104.0 | 121999.0 | 8.5 |
| 408 | 573712.5 | 120867.6 | 7.3 |
| 409 | 577402.2 | 126569.4 | 8 |
| 418 | 573241.8 | 146739.2 | 6.6 |
| 422 | 572020.0 | 117928.8 | 7.5 |
| 429 | 581627.7 | 147460.8 | 5.6 |
| 432 | 580700.0 | 126000.0 | 5.6 |
| 433 | 569002.0 | 124403.0 | 7.6 |
| 434 | 566522.1 | 134237.7 | 6.8 |
| 441 | 571830.6 | 117755.3 | 8.1 |
| 452 | 571531.4 | 117254.5 | 8.2 |
| 455 | 571620.7 | 117358.5 | 8.1 |
| 458 | 575233.2 | 143258.6 | 5.3 |
| 460 | 571633.8 | 116761.4 | 8.2 |
| 471 | 572527.7 | 146661.4 | 5.7 |
| 478 | 572705.1 | 134195.4 | 5.5 |
| 479 | 571730.2 | 117561.7 | 6.9 |
| 480 | 574689.1 | 146200.3 | 6.9 |
| 481 | 570900.0 | 135200.0 | 5.6 |
| 487 | 571933.4 | 147960.6 | 5.4 |

| N° relevé | X | Y | pH |
|---|---|---|---|
| 488 | 580526.9 | 133166.3 | 7.4 |
| 489 | 579033.1 | 152354.6 | 5.6 |
| 491 | 580598.0 | 126794.0 | 7 |
| 493 | 579435.1 | 125462.4 | 4.8 |
| 496 | 571498.3 | 135601.9 | 5.3 |
| 505 | 567999.0 | 127999.0 | 5.7 |
| 506 | 571529.3 | 117458.1 | 8.1 |
| 509 | 571328.6 | 141154.0 | 6.4 |
| 510 | 570000.0 | 124099.0 | 7.8 |
| 513 | 576429.7 | 144659.1 | 5.6 |
| 515 | 577832.4 | 141558.2 | 6.6 |
| 516 | 573731.9 | 143864.2 | 7.1 |
| 520 | 579424.6 | 126908.0 | 6.9 |
| 524 | 579057.6 | 126674.0 | 7.5 |
| 525 | 578575.9 | 123290.9 | 7.8 |
| 529 | 579045.5 | 123851.1 | 7.7 |
| 532 | 578547.2 | 123508.8 | 8.4 |
| 535 | 578410.7 | 127026.8 | 8.3 |
| 537 | 578371.4 | 126967.8 | 8.3 |
| 538 | 578409.6 | 127054.0 | 8.3 |
| 542 | 576937.1 | 125579.8 | 5.9 |
| 545 | 580387.0 | 126010.0 | 5.3 |
| 546 | 580130.7 | 125296.8 | 6.5 |
| 557 | 579851.0 | 126386.0 | 5.4 |
| 559 | 579099.5 | 126315.5 | 5.4 |
| 560 | 579803.2 | 124995.6 | 5.7 |
| 566 | 577807.0 | 126142.8 | 5.4 |
| 567 | 580711.0 | 126347.0 | 6.1 |
| 570 | 580222.9 | 125027.6 | 5.4 |
| 573 | 579775.8 | 125347.5 | 5.6 |
| 575 | 580613.0 | 126331.0 | 6.2 |
| 577 | 580380.3 | 126283.0 | 5.1 |
| 579 | 577897.1 | 125796.1 | 7.1 |
| 588 | 578655.5 | 123912.0 | 5.4 |
| 594 | 579078.0 | 123673.0 | 7.3 |
| 601 | 578637.1 | 123439.8 | 8.5 |
| 602 | 578885.4 | 123563.3 | 8.1 |
| 610 | 577964.2 | 126264.0 | 4.7 |
| 613 | 578584.1 | 126937.4 | 8.3 |

## 3. Corrélations entre prédicteurs

| | pH | Pente | Altitude | Position topo. | Géologie | Utilisation sol | Précipitations | Températures |
|---|---|---|---|---|---|---|---|---|
| pH | 1.00 | 0.12 | 0.13 | -0.12 | 0.01 | -0.26 | 0.09 | -0.13 |
| Pente | | 1.00 | 0.19 | -0.01 | 0.28 | -0.16 | 0.17 | -0.20 |
| Altitude | | | 1.00 | 0.14 | 0.24 | -0.61 | **0.94** | **-0.99** |
| Position topo. | | | | 1.00 | 0.02 | -0.09 | 0.14 | -0.14 |
| Géologie | | | | | 1.00 | -0.07 | 0.19 | -0.25 |
| Utilisation sol | | | | | | 1.00 | -0.60 | 0.61 |
| Précipitations | | | | | | | 1.00 | **-0.94** |
| Températures | | | | | | | | 1.00 |

## 4. Amélioration AUC et TSS moyennes des 4 modèles due au pH

| Espèce | Amelioration AUC | Amelioration TSS |
|---|---|---|
| *Valeriana_montana* | 0.12925 | 0.23725 |
| *Alchemilla_glabra_aggr* | 0.09450 | 0.10700 |
| *Hieracium_bifidum_aggr* | 0.09350 | 0.13625 |
| *Potentilla_aurea* | 0.05950 | 0.11850 |
| *Ranunculus_aconitifolius* | 0.05475 | 0.06325 |
| *Crepis_aurea* | 0.05200 | 0.10050 |
| *Trifolium_badium* | 0.04700 | 0.05600 |
| *Phyteuma_spicatum* | 0.04575 | 0.08000 |
| *Festuca_rubra_aggr* | 0.04500 | 0.06300 |
| *Globularia_cordifolia* | 0.04425 | 0.09075 |
| *Phleum_rhaeticum* | 0.04375 | 0.05625 |
| *Campanula_cochleariifolia* | 0.04350 | 0.05650 |
| *Adenostyles_glabra* | 0.04150 | 0.02525 |
| *Deschampsia_cespitosa* | 0.04075 | 0.08450 |
| *Phyteuma_orbiculare* | 0.04050 | 0.09050 |
| *Astrantia_major* | 0.03775 | 0.06625 |
| *Anthyllis_vulneraria_sl* | 0.03750 | 0.06525 |
| *Poa_supina* | 0.03700 | 0.02750 |
| *Salix_herbacea* | 0.03625 | 0.03400 |
| *Ligusticum_mutellina* | 0.03600 | 0.06750 |
| *Rumex_alpestris* | 0.03575 | 0.00700 |
| *Galium_anisophyllon* | 0.03500 | 0.05025 |
| *Hieracium_villosum_aggr* | 0.03325 | 0.00925 |
| *Gentiana_purpurea* | 0.03275 | 0.06175 |
| *Anthoxanthum_odoratum_aggr* | 0.03175 | 0.05850 |
| *Euphrasia_hirtella* | 0.03175 | 0.11150 |
| *Leontodon_hispidus_sl* | 0.03175 | 0.06875 |
| *Luzula_sylvatica* | 0.03075 | 0.08900 |
| *Ranunculus_nemorosus_aggr* | 0.03075 | 0.05375 |
| *Thlaspi_repens* | 0.02750 | 0.02725 |
| *Silene_vulgaris_sl* | 0.02700 | 0.00350 |
| *Trollius_europaeus* | 0.02700 | 0.03675 |
| *Alchemilla_xanthochlora_aggr* | 0.02650 | 0.01400 |
| *Parnassia_palustris* | 0.02500 | 0.00200 |
| *Pulsatilla_alpina_sl* | 0.02500 | 0.06975 |
| *Gentiana_lutea* | 0.02450 | 0.05100 |
| *Plantago_atrata_sstr* | 0.02425 | 0.02975 |
| *Dactylorhiza_fuchsii* | 0.02375 | 0.04750 |
| *Trifolium_pratense_sl* | 0.02325 | 0.03975 |
| *Poa_cenisia* | 0.02200 | 0.05125 |
| *Phleum_hirsutum* | 0.02125 | 0.00950 |
| *Senecio_doronicum* | 0.02125 | 0.00950 |
| *Cerastium_fontanum_sl* | 0.02100 | 0.06475 |
| *Veratrum_album_subsp_lobelianum* | 0.01975 | 0.03350 |
| *Prunella_grandiflora* | 0.01950 | 0.00750 |
| *Viola_biflora* | 0.01950 | 0.02275 |
| *Scabiosa_lucida* | 0.01900 | 0.04575 |
| *Hieracium_lactucella* | 0.01875 | 0.03475 |
| *Geum_montanum* | 0.01800 | 0.04525 |
| *Solidago_virgaurea_sl* | 0.01800 | -0.00900 |
| *Crocus_albiflorus* | 0.01775 | 0.03725 |

| | | |
|---|---|---|
| *Luzula_multiflora* | 0.01775 | 0.00975 |
| *Plantago_alpina* | 0.01725 | 0.03400 |
| *Polygala_alpestris* | 0.01625 | 0.05050 |
| *Linaria_alpina_sstr* | 0.01600 | -0.02250 |
| *Aposeris_foetida* | 0.01575 | 0.00350 |
| *Chaerophyllum_hirsutum_aggr* | 0.01575 | 0.03800 |
| *Agrostis_capillaris* | 0.01500 | 0.03875 |
| *Helianthemum_nummularium_sl* | 0.01500 | 0.02700 |
| *Soldanella_alpina* | 0.01500 | 0.02200 |
| *Carex_pallescens* | 0.01450 | 0.00150 |
| *Cirsium_acaule* | 0.01425 | 0.02650 |
| *Hypericum_maculatum_sl* | 0.01400 | 0.05050 |
| *Vicia_sativa_sl* | 0.01375 | 0.00000 |
| *Geranium_sylvaticum* | 0.01350 | 0.01250 |
| *Phleum_pratense* | 0.01350 | 0.03225 |
| *Alchemilla_conjuncta_aggr* | 0.01275 | 0.01375 |
| *Helictotrichon_versicolor* | 0.01275 | 0.04950 |
| *Leucanthemum_vulgare_aggr* | 0.01250 | 0.03600 |
| *Vaccinium_myrtillus* | 0.01250 | 0.01350 |
| *Agrostis_stolonifera* | 0.01225 | -0.03400 |
| *Cerastium_latifolium* | 0.01150 | 0.01425 |
| *Clinopodium_vulgare* | 0.01150 | 0.03750 |
| *Veronica_aphylla* | 0.01150 | 0.02475 |
| *Plantago_major_sl* | 0.01125 | 0.02200 |
| *Primula_elatior_sstr* | 0.01125 | 0.01275 |
| *Ranunculus_repens* | 0.01075 | 0.02800 |
| *Euphrasia_salisburgensis* | 0.01050 | 0.04450 |
| *Lotus_corniculatus_aggr* | 0.01025 | -0.00800 |
| *Ranunculus_acris_sl* | 0.00950 | 0.00575 |
| *Campanula_barbata* | 0.00950 | -0.05675 |
| *Campanula_scheuchzeri* | 0.00925 | 0.00025 |
| *Lolium_perenne* | 0.00925 | 0.00350 |
| *Festuca_quadriflora* | 0.00900 | 0.01925 |
| *Potentilla_erecta* | 0.00900 | 0.01725 |
| *Helictotrichon_pubescens* | 0.00875 | 0.00775 |
| *Rumex_acetosa* | 0.00875 | 0.01075 |
| *Centaurea_jacea_sstr* | 0.00800 | 0.01650 |
| *Achillea_millefolium* | 0.00750 | 0.01300 |
| *Medicago_lupulina* | 0.00700 | 0.00325 |
| *Carex_ferruginea* | 0.00650 | 0.02200 |
| *Trisetum_flavescens* | 0.00650 | -0.00500 |
| *Bellis_perennis* | 0.00600 | -0.00375 |
| *Cirsium_spinosissimum* | 0.00575 | 0.00200 |
| *Crepis_biennis* | 0.00550 | -0.00225 |
| *Nardus_stricta* | 0.00550 | 0.01650 |
| *Prunella_vulgaris* | 0.00550 | -0.00250 |
| *Geum_rivale* | 0.00525 | 0.00400 |
| *Poa_alpina* | 0.00525 | 0.00275 |
| *Saxifraga_oppositifolia* | 0.00525 | 0.02250 |
| *Festuca_ovina_aggr* | 0.00500 | 0.00350 |
| *Polygala_vulgaris_sl* | 0.00500 | 0.01750 |
| *Carduus_defloratus_sl* | 0.00450 | 0.00800 |
| *Saxifraga_aizoides* | 0.00450 | 0.00525 |
| *Sesleria_caerulea* | 0.00425 | 0.01500 |
| *Trifolium_repens_sstr* | 0.00400 | 0.01450 |
| *Linum_catharticum* | 0.00400 | 0.00275 |
| *Arrhenatherum_elatius* | 0.00375 | 0.00225 |
| *Taraxacum_officinale_aggr* | 0.00375 | -0.01400 |
| *Aster_bellidiastrum* | 0.00375 | -0.03050 |
| *Thymus_pulegioides_sstr* | 0.00325 | -0.00375 |
| *Veronica_arvensis* | 0.00325 | 0.00000 |
| *Veronica_serpyllifolia_sl* | 0.00275 | 0.00625 |
| *Hippocrepis_comosa* | 0.00250 | -0.00325 |
| *Cynosurus_cristatus* | 0.00225 | 0.00250 |
| *Euphorbia_cyparissias* | 0.00225 | 0.00025 |
| *Pedicularis_verticillata* | 0.00200 | -0.00300 |
| *Calamagrostis_varia* | 0.00200 | 0.02325 |
| *Stellaria_graminea* | 0.00150 | -0.00875 |
| *Hieracium_murorum_aggr* | 0.00150 | -0.03500 |
| *Knautia_dipsacifolia_sstr* | 0.00150 | 0.00675 |
| *Carex_sempervirens* | 0.00100 | 0.00925 |
| *Plantago_media* | 0.00100 | 0.00550 |
| *Sanguisorba_minor_sl* | 0.00100 | 0.01225 |

| | | |
|---|---|---|
| *Carum_carvi* | 0.00075 | -0.01500 |
| *Luzula_campestris* | 0.00075 | 0.00175 |
| *Veronica_chamaedrys* | 0.00075 | -0.00200 |
| *Festuca_pratensis_sl* | 0.00050 | 0.01025 |
| *Rhinanthus_alectorolophus* | 0.00050 | 0.00100 |
| *Briza_media* | 0.00000 | 0.01175 |
| *Dactylis_glomerata* | -0.00025 | -0.00975 |
| *Pritzelago_alpina_sstr* | -0.00025 | -0.01075 |
| *Daucus_carota* | -0.00075 | -0.00625 |
| *Holcus_lanatus* | -0.00125 | 0.00725 |
| *Festuca_violacea_aggr* | -0.00150 | 0.00650 |
| *Thymus_praecox_subsp_polytrichus* | -0.00150 | -0.01875 |
| *Carex_caryophyllea* | -0.00175 | -0.00550 |
| *Ranunculus_alpestris* | -0.00175 | -0.00075 |
| *Doronicum_grandiflorum* | -0.00200 | 0.00825 |
| *Galium_pumilum* | -0.00200 | -0.00950 |
| *Campanula_rotundifolia* | -0.00225 | -0.01500 |
| *Crepis_pyrenaica* | -0.00225 | 0.02000 |
| *Hypochaeris_radicata* | -0.00225 | 0.01300 |
| *Luzula_alpinopilosa* | -0.00225 | -0.03150 |
| *Plantago_lanceolata* | -0.00225 | 0.01100 |
| *Primula_veris_sl* | -0.00225 | -0.00925 |
| *Ajuga_reptans* | -0.00250 | -0.01125 |
| *Poa_trivialis_sl* | -0.00275 | -0.01450 |
| *Tragopogon_pratensis_sl* | -0.00325 | -0.01000 |
| *Vicia_sepium* | -0.00350 | -0.00925 |
| *Bromus_erectus_sstr* | -0.00375 | 0.00200 |
| *Lathyrus_pratensis* | -0.00375 | -0.01400 |
| *Poa_pratensis* | -0.00375 | 0.01600 |
| *Pimpinella_major* | -0.00425 | -0.00800 |
| *Carex_sylvatica* | -0.00425 | -0.02000 |
| *Laserpitium_latifolium* | -0.00425 | -0.02725 |
| *Heracleum_sphondylium_sl* | -0.00475 | -0.01075 |
| *Dryas_octopetala* | -0.00500 | -0.01025 |
| *Ranunculus_bulbosus* | -0.00525 | -0.03500 |
| *Anthriscus_sylvestris* | -0.00550 | 0.00000 |
| *Bromus_hordeaceus* | -0.00550 | -0.09550 |
| *Polygonum_bistorta* | -0.00550 | -0.02475 |
| *Ranunculus_montanus_aggr* | -0.00550 | -0.02300 |
| *Campanula_rhomboidalis* | -0.00600 | -0.02625 |
| *Silene_acaulis* | -0.00600 | -0.03100 |
| *Sedum_atratum* | -0.00625 | -0.01250 |
| *Vicia_cracca_sstr* | -0.00625 | -0.01125 |
| *Knautia_arvensis* | -0.00650 | 0.02000 |
| *Polygonum_viviparum* | -0.00650 | -0.02150 |
| *Saxifraga_paniculata* | -0.00675 | -0.01875 |
| *Arabis_alpina_sl* | -0.00700 | 0.00525 |
| *Homogyne_alpina* | -0.00725 | -0.02825 |
| *Trifolium_thalii* | -0.00750 | -0.01600 |
| *Brachypodium_pinnatum* | -0.00800 | -0.01750 |
| *Myosotis_alpestris* | -0.00800 | -0.01775 |
| *Leontodon_autumnalis* | -0.00850 | -0.06075 |
| *Poa_minor* | -0.01075 | -0.01775 |
| *Polygala_chamaebuxus* | -0.01075 | -0.00075 |
| *Cruciata_laevipes* | -0.01100 | -0.00200 |
| *Gentiana_campestris_sstr* | -0.01150 | -0.01025 |
| *Pimpinella_saxifraga_aggr* | -0.01175 | -0.03550 |
| *Veronica_alpina* | -0.01200 | 0.00175 |
| *Euphrasia_minima* | -0.01325 | -0.01600 |
| *Salix_retusa* | -0.01350 | -0.03600 |
| *Vaccinium_vitis.idaea* | -0.01400 | -0.06250 |
| *Androsace_chamaejasme* | -0.01425 | 0.04325 |
| *Glechoma_hederacea_sstr* | -0.01475 | -0.00525 |
| *Gentiana_verna* | -0.01525 | 0.00175 |
| *Galium_album* | -0.01550 | -0.02025 |
| *Trifolium_medium* | -0.01550 | -0.01925 |
| *Carex_montana* | -0.01650 | -0.02300 |
| *Carlina_acaulis_subsp_caulescens* | -0.01700 | -0.03650 |
| *Carex_nigra* | -0.01875 | -0.06200 |
| *Arnica_montana* | -0.02025 | -0.03450 |
| *Galium_megalospermum* | -0.02075 | 0.01425 |
| *Carex_flacca* | -0.02175 | -0.03700 |
| *Myosotis_arvensis* | -0.02225 | -0.04625 |

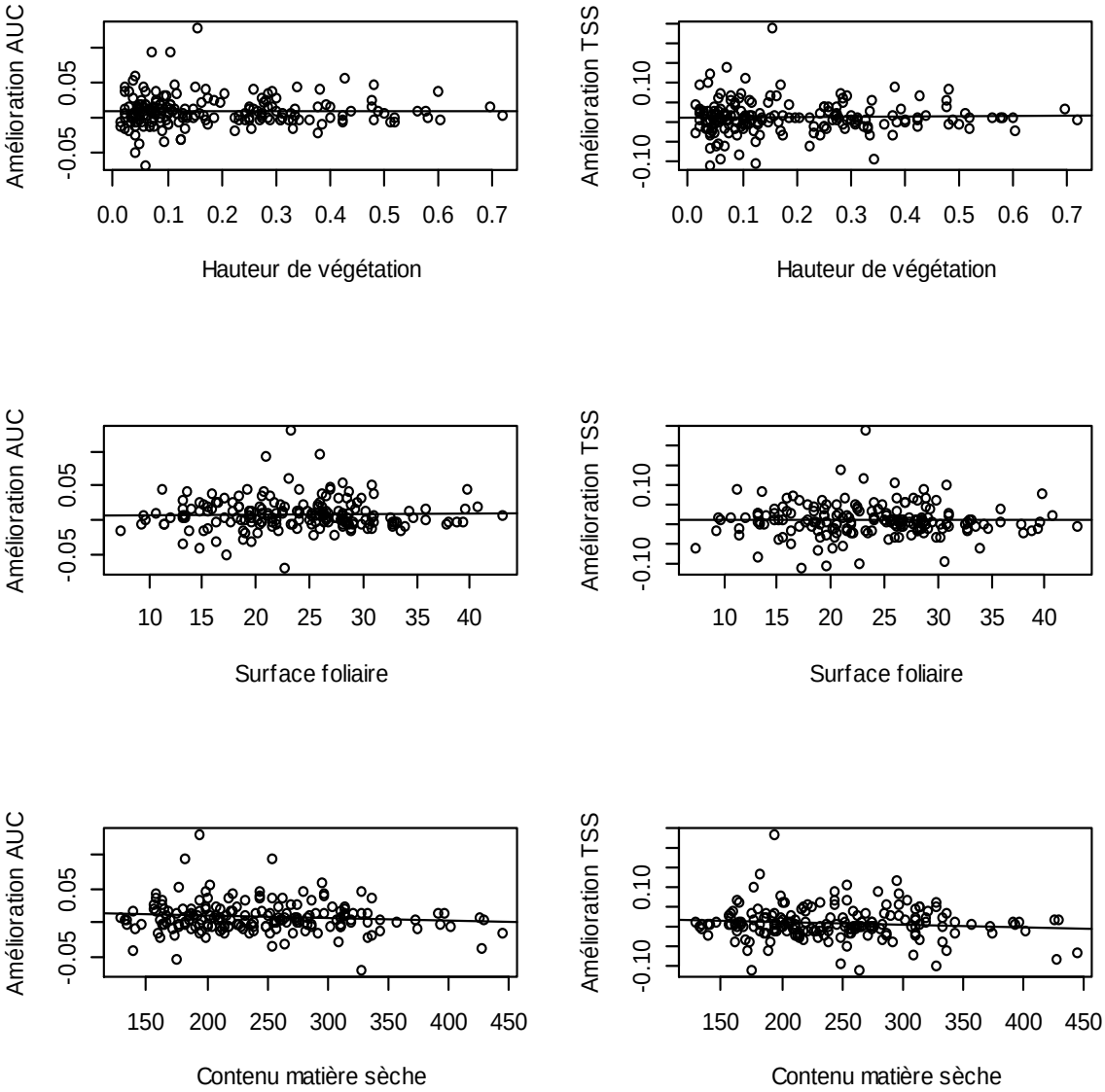
| | | |
|---|---|---|
| *Potentilla_crantzii* | -0.02725 | -0.06900 |
| *Bartsia_alpina* | -0.03125 | -0.10825 |
| *Thesium_alpinum* | -0.03225 | -0.04975 |
| *Alchemilla_coriacea_aggr* | -0.03350 | -0.03200 |
| *Vaccinium_gaultherioides* | -0.03550 | -0.08425 |
| *Primula_auricula* | -0.03975 | 0.00925 |
| *Cardamine_pratensis* | -0.04400 | -0.08875 |
| *Colchicum_autumnale* | -0.04450 | -0.01725 |
| *Gypsophila_repens* | -0.05150 | -0.11150 |
| *Agrostis_rupestris* | -0.06950 | -0.09775 |

## 5. Tableau des espèces avec amélioration de l'AUC supérieure à 0.02 et leurs indices et traits biologiques

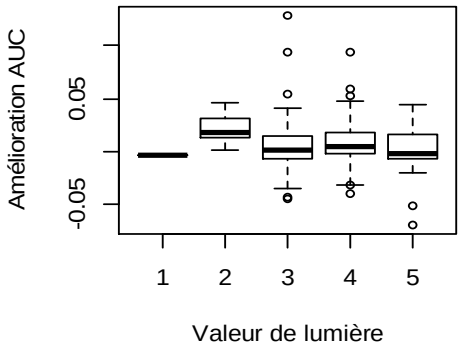
| Espèce | Amelio.AUC | Amelio.TSS | vegH | SLA | LDMC | D | F | H | K | L | N | R | T |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *Valeriana montana* | 0.1293 | 0.2373 | 0.156 | 23.192 | 193.696 | 2 | 3 | 3 | 3 | 3 | 2 | 4 | 2 |
| *Alchemilla glabra aggr* | 0.0945 | 0.1070 | 0.104 | 25.89 | 254.395 | 4 | 4 | 4 | 2 | 3 | 3 | 3 | 2 |
| *Hieracium bifidum aggr* | 0.0935 | 0.1363 | 0.072 | 21.004 | 182.113 | 3 | 2 | 3 | 3 | 4 | 2 | 4 | 2 |
| *Potentilla aurea* | 0.0595 | 0.1185 | 0.041 | 22.99 | 295.073 | 4 | 3 | 3 | 3 | 4 | 2 | 2 | 2 |
| *Ranunculus aconitifolius* | 0.0548 | 0.0633 | 0.428 | 28.048 | 202.039 | 5 | 4 | 4 | 2 | 3 | 4 | 3 | 2 |
| *Crepis aurea* | 0.0520 | 0.1005 | 0.038 | 30.848 | 177.597 | 4 | 3 | 4 | 3 | 4 | 4 | 3 | 2 |
| *Trifolium badium* | 0.0470 | 0.0560 | 0.114 | 26.925 | 243.083 | 4 | 3 | 3 | 3 | 4 | 3 | 4 | 2 |
| *Phyteuma spicatum* | 0.0458 | 0.0800 | 0.482 | 39.69 | 199.526 | 4 | 3 | 4 | 3 | 2 | 3 | 3 | 3 |
| *Festuca rubra aggr* | 0.0450 | 0.0630 | 0.152 | 19.241 | 327.025 | 4 | 3 | 3 | 3 | 4 | 3 | 3 | NA |
| *Globularia cordifolia* | 0.0443 | 0.0908 | 0.023 | 11.298 | 279.426 | 1 | 1 | 2 | 4 | 5 | 2 | 5 | 2 |
| *Phleum rhaeticum* | 0.0438 | 0.0563 | 0.339 | 27.017 | 296.665 | 4 | 3 | 3 | 3 | 4 | 4 | 3 | 2 |
| *Campanula cochleariifolia* | 0.0435 | 0.0565 | 0.055 | 23.924 | 221.307 | 2 | 3 | 2 | 3 | 4 | 2 | 4 | 2 |
| *Adenostyles glabra* | 0.0415 | 0.0253 | 0.258 | 20.736 | 157.854 | 2 | 3 | 3 | 3 | 2 | 3 | 4 | 2 |
| *Deschampsia cespitosa* | 0.0408 | 0.0845 | 0.381 | 13.597 | 296.144 | 5 | 4 | 3 | 3 | 3 | 4 | 3 | 3 |
| *Phyteuma orbiculare* | 0.0405 | 0.0905 | 0.172 | 28.619 | 244.156 | 5 | 3 | 4 | 3 | 4 | 2 | 4 | 2 |
| *Astrantia major* | 0.0378 | 0.0663 | 0.294 | 26.231 | 243.286 | 4 | 3 | 3 | 3 | 3 | 3 | 4 | 2 |
| *Anthyllis vulneraria sl* | 0.0375 | 0.0653 | 0.077 | 15.952 | 201.005 | 3 | 1 | 3 | 4 | 4 | 2 | 3 | 4 |
| *Poa supina* | 0.0370 | 0.0275 | 0.03 | 30.98 | 263.28 | 4 | 4 | 3 | 3 | 5 | 4 | 3 | 2 |
| *Salix herbacea* | 0.0363 | 0.0340 | 0.021 | 18.305 | 336.02 | 4 | 4 | 4 | 2 | 5 | 2 | 2 | 1 |
| *Ligusticum mutellina* | 0.0360 | 0.0675 | 0.058 | 21.366 | 253.517 | 4 | 3 | 3 | 3 | 4 | 3 | 3 | 2 |
| *Rumex alpestris* | 0.0358 | 0.0070 | 0.599 | 26.533 | 158.254 | 4 | 4 | 4 | 3 | 3 | 4 | 3 | 2 |
| *Galium anisophyllon* | 0.0350 | 0.0503 | 0.119 | 20.598 | 218.766 | 3 | 2 | 3 | 3 | 4 | 2 | 3 | 2 |
| *Hieracium villosum aggr* | 0.0333 | 0.0093 | 0.205 | 28.555 | 194.11 | 3 | 3 | 3 | 3 | 4 | 2 | 4 | 2 |
| *Gentiana purpurea* | 0.0328 | 0.0618 | 0.287 | 17.012 | 232.161 | 4 | 3 | 3 | 2 | 4 | 2 | 2 | 2 |
| *Leontodon hispidus sl* | 0.0318 | 0.0688 | 0.099 | 27.712 | 162.633 | 4 | 3 | 3 | 3 | 4 | 3 | 3 | 3 |
| *Euphrasia hirtella* | 0.0318 | 0.1115 | NA | NA | NA | 3 | 2 | 4 | 4 | 4 | 2 | 3 | 2 |
| *Anthoxanthum odoratum aggr* | 0.0318 | 0.0585 | 0.096 | 29.87 | 285.803 | 4 | 3 | 3 | 3 | 4 | 3 | 2 | 3 |
| *Ranunculus nemorosus aggr* | 0.0308 | 0.0538 | NA | NA | NA | 5 | 3 | 3 | 3 | 3 | 2 | 3 | 3 |
| *Luzula sylvatica* | 0.0308 | 0.0890 | NA | NA | NA | 4 | 3 | 4 | 2 | 2 | 2 | 2 | 4 |
| *Thlaspi repens* | 0.0275 | 0.0273 | 0.035 | 13.211 | 155.845 | 2 | 3 | 2 | 3 | 5 | 2 | 5 | 1 |
| *Trollius europaeus* | 0.0270 | 0.0368 | 0.274 | 14.732 | 268.633 | 5 | 4 | 4 | 3 | 4 | 3 | 3 | 2 |
| *Silene vulgaris sl* | 0.0270 | 0.0035 | 0.303 | 29.37 | 163.041 | 3 | 2 | 3 | 3 | 3 | 2 | 3 | 3 |
| *Alchemilla xanthochlora aggr* | 0.0265 | 0.0140 | 0.173 | 21.646 | 275.026 | NA | NA | NA | NA | NA | NA | NA | NA |
| *Pulsatilla alpina sl* | 0.0250 | 0.0698 | 0.281 | 16.443 | 303.897 | 3 | 3 | 3 | 3 | 4 | 3 | 4 | 2 |
| *Parnassia palustris* | 0.0250 | 0.0020 | 0.053 | 26.041 | 198.494 | 5 | 4 | 3 | 3 | 4 | 2 | 4 | 2 |

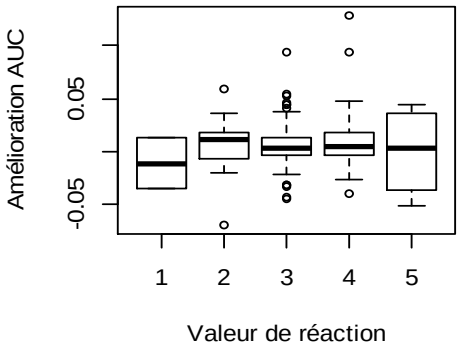
| | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *Gentiana lutea* | 0.0245 | 0.0510 | 0.477 | 17.841 | 224.338 | 4 | 3 | 3 | 2 | 4 | 3 | 4 | 2 |
| *Plantago atrata sstr* | 0.0243 | 0.0298 | 0.079 | 16.178 | 208.457 | 4 | 3 | 4 | 3 | 4 | 3 | 4 | 2 |
| *Dactylorhiza fuchsii* | 0.0238 | 0.0475 | NA | NA | NA | NA | NA | NA | NA | NA | NA | NA | NA |
| *Trifolium pratense sl* | 0.0233 | 0.0398 | 0.187 | 24.326 | 259.637 | 4 | 3 | 3 | 3 | 3 | 3 | 3 | 3 |
| *Poa cenisia* | 0.0220 | 0.0513 | 0.079 | 22.234 | 313.466 | 2 | 3 | 2 | 2 | 4 | 2 | 4 | 2 |
| *Senecio doronicum* | 0.0213 | 0.0095 | 0.196 | 15.017 | 156.761 | 4 | 3 | 3 | 3 | 4 | 2 | 4 | 2 |
| *Phleum hirsutum* | 0.0213 | 0.0095 | 0.28 | 25.731 | 313.518 | 4 | 3 | 3 | 3 | 4 | 3 | 4 | 2 |
| *Cerastium fontanum sl* | 0.0210 | 0.0648 | 0.164 | 28.85 | 164.129 | 4 | 3 | 3 | 2 | 3 | 4 | 3 | 2 |

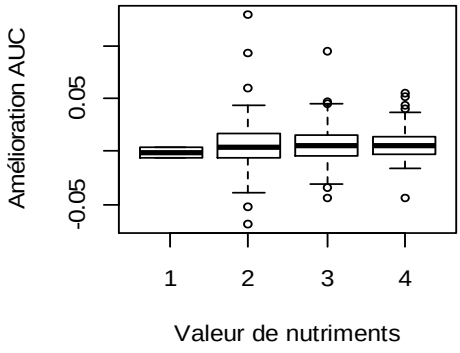
## 6. Amélioration des indices AUC et TSS en fonction des traits biologiques

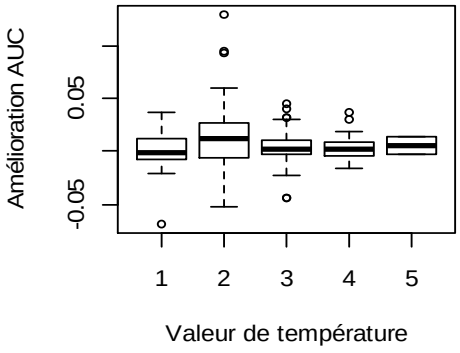


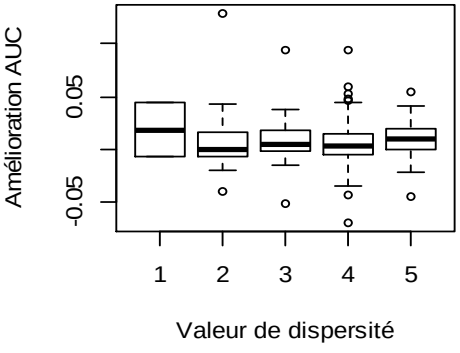
## 7. Amélioration de l'indice AUC en fonction des valeurs écologiques de Landolt

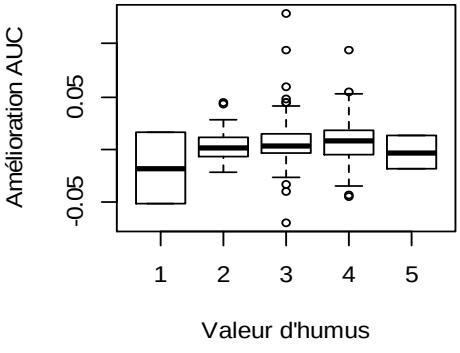


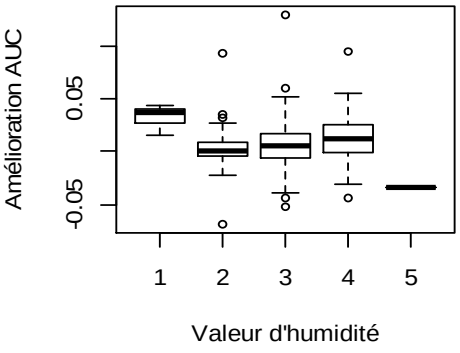



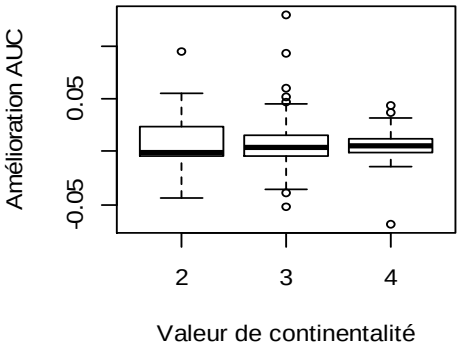



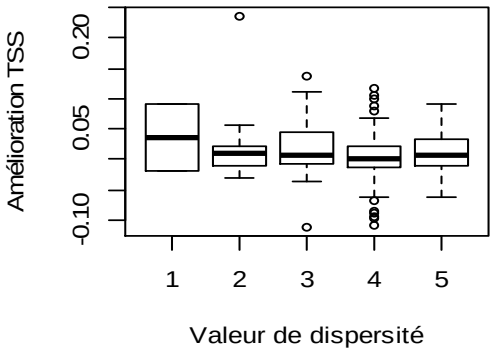



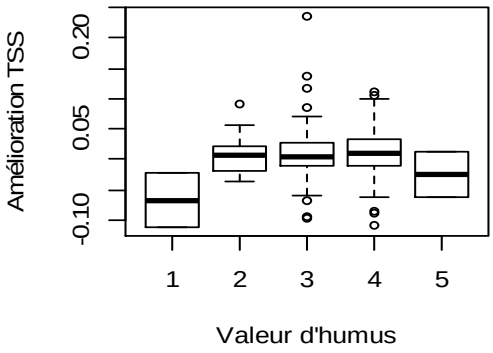



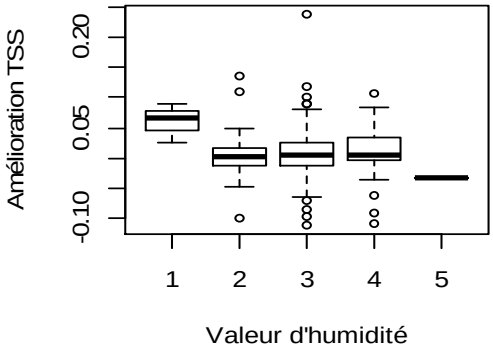



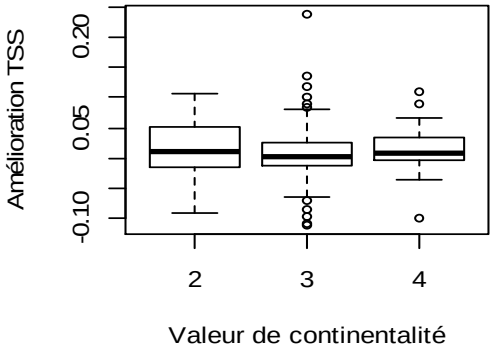



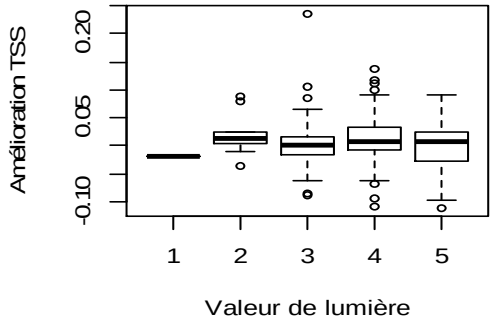
## 8. Amélioration de l'indice TSS en fonction des valeurs écologiques de Landolt

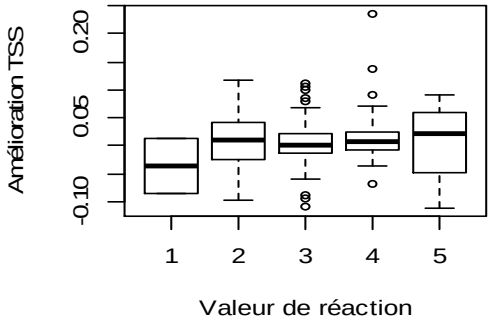


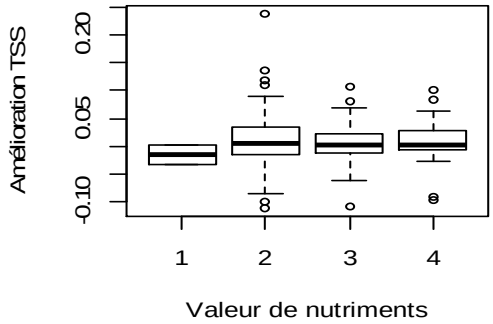



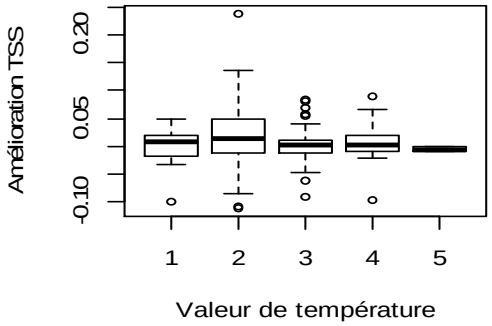



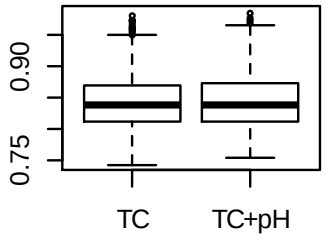



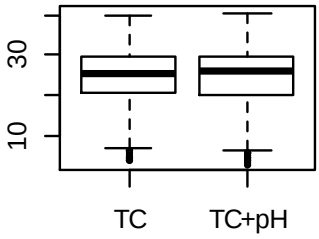



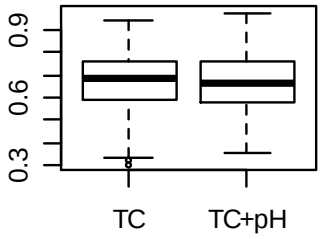
## 9. Indices d'évaluation des communautés



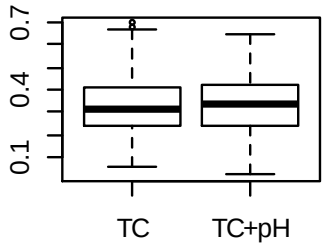

Succès de prédiction



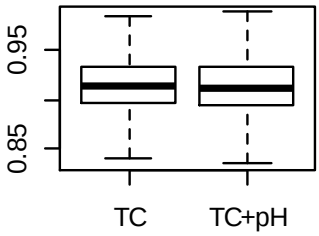

Richesse spécifique



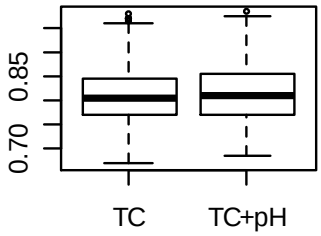

Présences sous-estimées



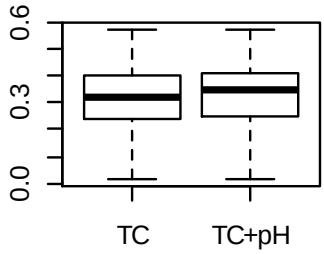

Sensibilité



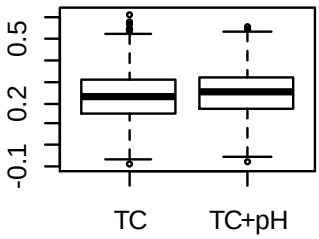

Spécificité



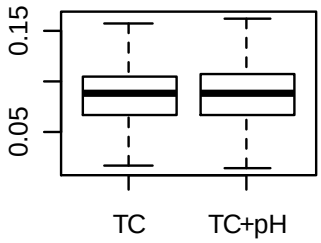

Kappa pondéré




Similarité

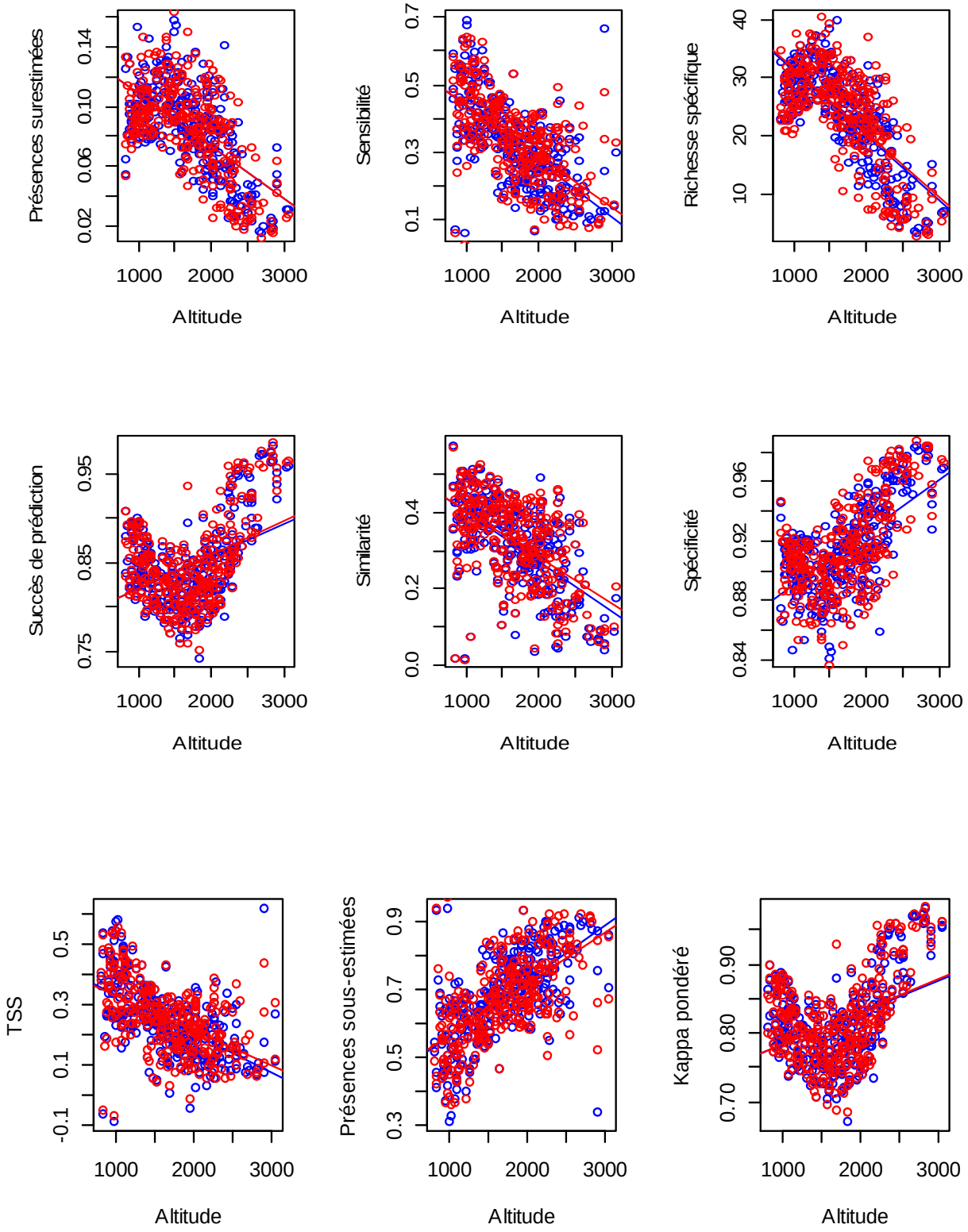



TSS

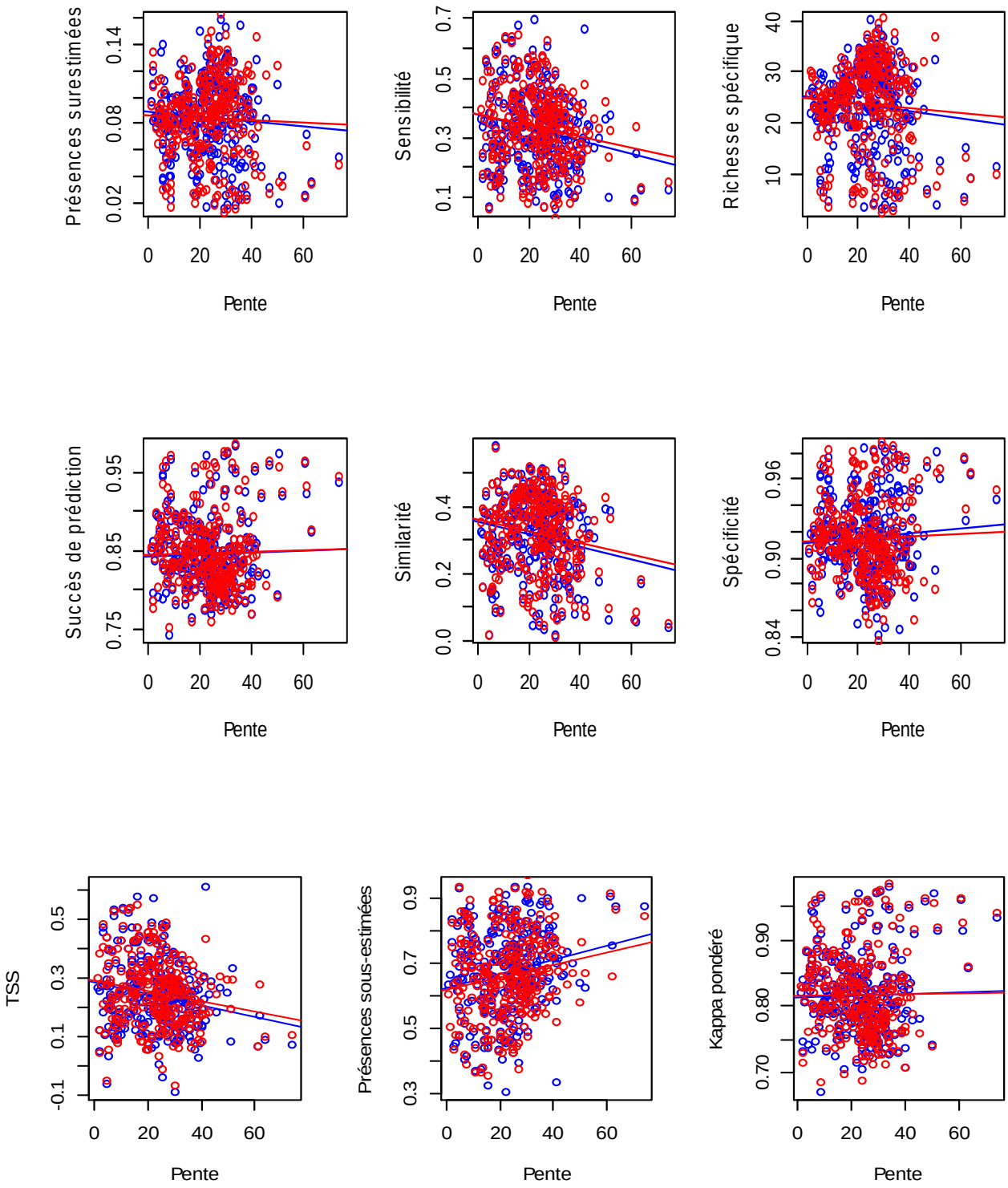



Présences surestimées

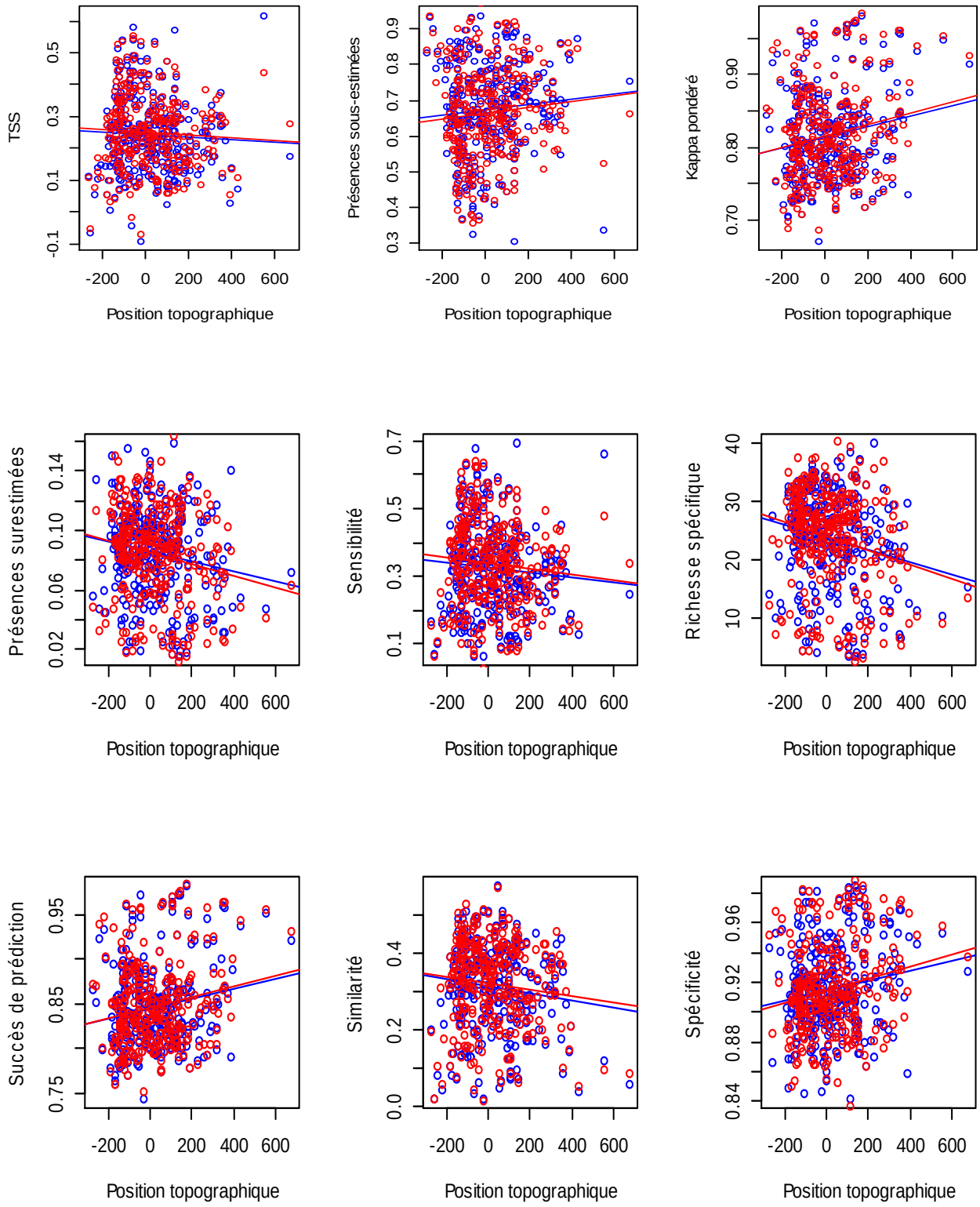
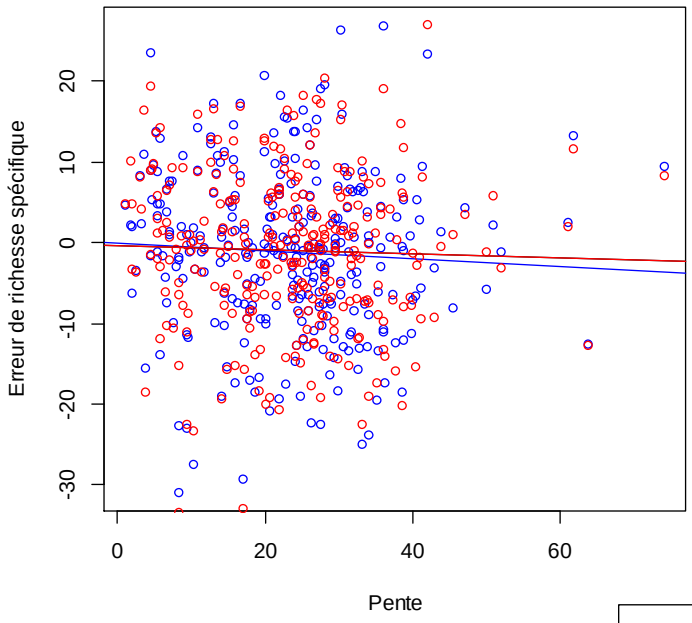
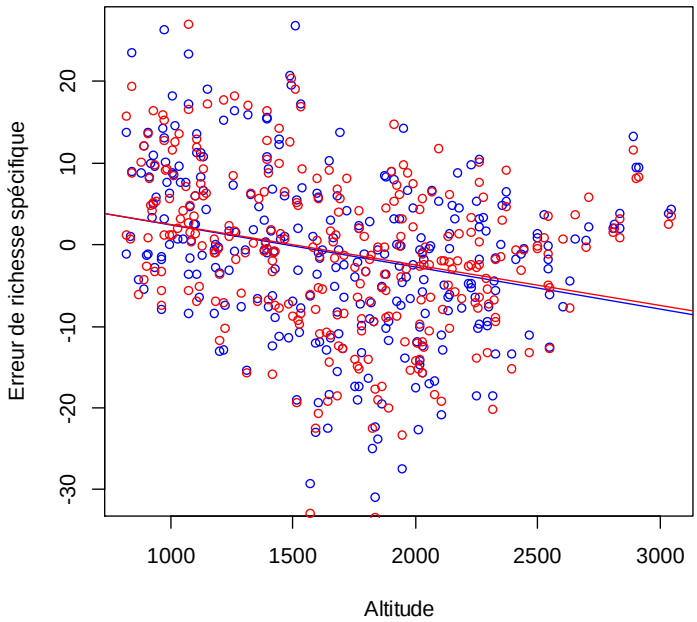
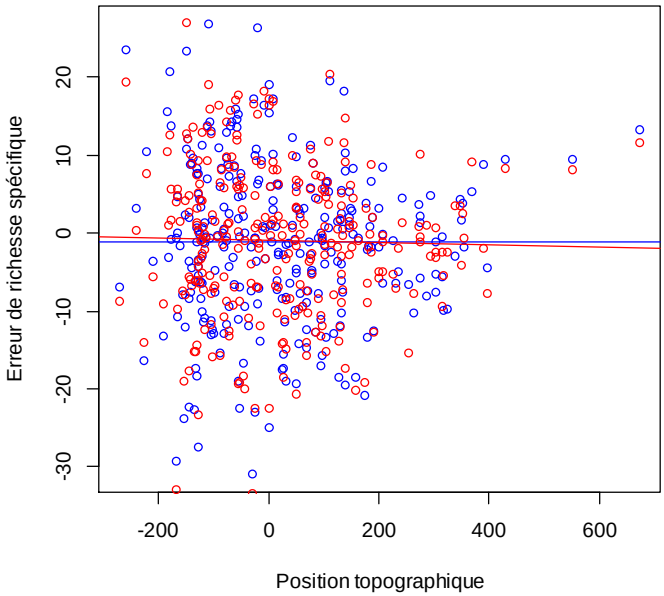
**10. Régressions linéaires des indices pour les modèles topo-climatique et topo-climatique + pH en fonction de l'altitude**



**11. Régressions linéaires des indices pour les modèles topo-climatique et topo-climatique + pH en fonction de la pente**



## 12. Régressions linéaires des indices pour les modèles topo-climatique et topo-climatique + pH en fonction de la position topographique



## 13. Diagrammes de l'erreur de richesse spécifique en fonction de la pente, de l'altitude et de la position topographique