

Projet de Fin d'Etudes

PFE Semestre 9 & 10

**Quel outil d'analyse utiliser pour
expliquer la répartition des
Gomphidae en Loire ?**



LOISEAU Marie

2017-2018

Directeur de recherche :

Mme BOISNEAU Catherine

QUEL OUTIL D'ANALYSE UTILISER POUR EXPLIQUER LA REPARTITION DES *GOMPHIDAE* EN LOIRE ?

Encadré par :

Boisneau Catherine

Réalisé par :

Loiseau Marie

2017/2018

AVERTISSEMENT

Cette recherche a fait appel à des lectures, enquêtes et interviews. Tout emprunt à des contenus d'interviews, des écrits autres que strictement personnel, toute reproduction et citation, font systématiquement l'objet d'un référencement.

L'auteur (les auteurs) de cette recherche a (ont) signé une attestation sur l'honneur de non plagiat.

FORMATION PAR LA RECHERCHE, PROJET DE FIN D'ETUDES EN GENIE DE L'AMENAGEMENT ET DE L'ENVIRONNEMENT

La formation au génie de l'aménagement et de l'environnement, assurée par le département aménagement et environnement de l'Ecole Polytechnique de l'Université de Tours, associe dans le champ de l'urbanisme, de l'aménagement des espaces fortement à faiblement anthropisés, l'acquisition de connaissances fondamentales, l'acquisition de techniques et de savoir faire, la formation à la pratique professionnelle et la formation par la recherche. Cette dernière ne vise pas à former les seuls futurs élèves désireux de prolonger leur formation par les études doctorales, mais tout en ouvrant à cette voie, elle vise tout d'abord à favoriser la capacité des futurs ingénieurs à :

- Accroître leurs compétences en matière de pratique professionnelle par la mobilisation de connaissances et de techniques, dont les fondements et contenus ont été explorés le plus finement possible afin d'en assurer une bonne maîtrise intellectuelle et pratique,
- Accroître la capacité des ingénieurs en génie de l'aménagement et de l'environnement à innover tant en matière de méthodes que d'outils, mobilisables pour affronter et résoudre les problèmes complexes posés par l'organisation et la gestion des espaces.

La formation par la recherche inclut un exercice individuel de recherche, le projet de fin d'études (P.F.E.), situé en dernière année de formation des élèves ingénieurs. Cet exercice correspond à un stage d'une durée minimum de trois mois, en laboratoire de recherche, principalement au sein de l'équipe Ingénierie du Projet d'Aménagement, Paysage et Environnement de l'UMR 6173 CITERES à laquelle appartiennent les enseignants-chercheurs du département aménagement.

Le travail de recherche, dont l'objectif de base est d'acquérir une compétence méthodologique en matière de recherche, doit répondre à l'un des deux grands objectifs :

- Développer toute ou partie d'une méthode ou d'un outil nouveau permettant le traitement innovant d'un problème d'aménagement
- Approfondir les connaissances de base pour mieux affronter une question complexe en matière d'aménagement.

Afin de valoriser ce travail de recherche nous avons décidé de mettre en ligne sur la base du Système Universitaire de Documentation (SUDOC), les mémoires à partir de la mention bien.

REMERCIEMENTS

Je souhaite remercier Madame Catherine Boisneau pour m'avoir encadrée tout au long de ce Projet de Fin d'études. Je vous remercie pour votre disponibilité, pour les réponses données à mes interrogations et pour l'ensemble des documents fournis et explications nécessaires à la bonne réalisation de ce projet. Je suis également reconnaissante de votre aide pour la rédaction de ce rapport.

Je souhaite également remercier Madame Nina Richard pour m'avoir fourni les données récoltées lors de l'échantillonnage des *Gomphidae*.

Je remercie sincèrement Madame Clémence Chaudron pour avoir pris le temps de répondre à mes interrogations sur différents tests statistiques et pour m'avoir aidée à les implémenter sur le logiciel R.

Enfin je remercie mes amis de promotion et mes colocataires avec qui j'ai pu échanger sur mon projet et passer une agréable année avec ses moments de détente, nécessaires à un travail de qualité.

SOMMAIRE

Avertissement.....	
Remerciements.....	
Sommaire.....	1
Lexique.....	2
Sigles.....	2
Introduction.....	3
1. Ecologie des <i>Gomphidae</i> et menaces.....	4
2. Utilisation du logiciel Maxent.....	10
3. Tests statistiques multivariés.....	15
4. Quel outil est le plus approprié pour expliquer la répartition des <i>Gomphidae</i> en Loire ?.....	22
Conclusion.....	24
Bibliographie.....	25
Table des figures.....	28
Table des tableaux.....	28
Annexe 1 : Exemple d'une fiche de données présentant les variables environnementales mesurées et le comptage des invertébrés.....	29
Annexe 2 : Présentation de l'interface de travail du logiciel MaxEnt.....	31
Annexe 3 : Résultats pouvant être obtenus par le logiciel MaxEnt.....	32
Annexe 4 : Tableau brut des données à charger dans R (avec la variable qualitative substrat) ...	35
Annexe 5 : Tableau brut des données à charger dans R (avec les sous-variables quantitatives subs_sgr et subst_s).....	38
Annexe 6 : Script langage R.....	41
Annexe 7 : Démarche générale de l'étude.....	42
Table des matières.....	43

LEXIQUE

Diapause : phase d'arrêt du développement d'un individu pendant des périodes défavorables de l'environnement.

Entropie : correspond à une caractérisation statistique de l'information manquante ([Venditti, 1998 in Kato 2012](#))

Granulométrie : mesure de la taille des particules telles que les sédiments, détermination de leur forme et étude de leur répartition.

Lentique : correspond à des eaux lentes, voire stagnantes et s'oppose à « lotique ».

Lotique : correspond à des eaux courantes.

Sténoèce : espèce ayant une faible amplitude de niche écologique ; les individus sont dits spécialistes.

Traits écologiques : caractéristiques biologiques d'une espèce

Valence écologique : capacité d'une espèce (végétale ou animale) à tolérer les variations plus ou moins grandes des facteurs écologiques d'un milieu différent de son milieu naturel et de le coloniser.

SIGLES

ACM : Analyse des Correspondances Multiples

ACP : Analyse en Composantes Principales

AFC : Analyse Factorielle des Correspondances

AFM : Analyse Factorielle Multiple

CEN : Conservatoire d'Espaces Naturels

CLA : Contrat pour la Loire et ses Annexes 2015-2020

OPIE : Office pour les Insectes et leur Environnement

PNA : Plan National d'Actions

ROC : Receiver Operating Characteristic, courbe obtenue avec le logiciel MaxEnt permettant d'évaluer la capacité de discrimination du modèle.

R-TEMUS : Restauration du lit et Trajectoires Ecologiques, Morphologiques et d'Usages en Basse-Loire

SIC : Site d'intérêt Communautaire

UICN : Union Internationale pour la Conservation de la Nature

INTRODUCTION

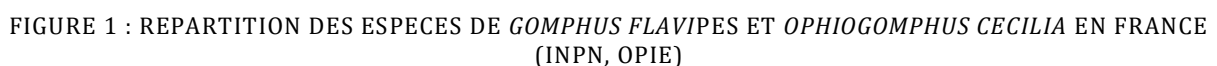
La Loire, bien que connue sous l'appellation de « dernier fleuve sauvage », n'a pas été épargnée par les grands travaux d'aménagements des cours d'eau du 19 et 20^{ème} siècle. Le chenal de navigation entre les Ponts-de-Cé et Nantes en est un vestige. Les conséquences directes telles que l'incision importante du lit principal ou l'abaissement de la ligne d'eau entraînent une fragilisation des ponts et des berges mais aussi une déconnexion du lit principal avec ses bras secondaires, ses boires et zones humides adjacentes. La prise de conscience des problèmes sociétaux et environnementaux causés a engendré la mise en place du Contrat pour la Loire et ses Annexes 2015-2020 ayant pour principales missions de rééquilibrer le lit mineur et de restaurer les fonctions hydro-écologiques des annexes hydrauliques.

Le programme de recherche R-TEMUS (Restauration du lit et Trajectoires Ecologiques, Morphologiques et d'USages en Basse-Loire), porté par l'université de Tours, s'inscrit également dans ce contrat. Il présente dans son volet fonctionnement biotique une étude sur les larves de *Gomphidae*, famille d'Anisoptères reconnue vulnérable dans une grande partie des fleuves européens suite à l'aménagement des cours d'eau des siècles derniers. Cette famille est tout particulièrement intéressante car son micro-habitat larvaire optimal est lié à des zones de profondeur peu importante où le courant est ralenti (Gouraud, 2015). Nous les retrouvons notamment derrière les épis. La destruction de ceux-ci dans le cadre du programme de rééquilibrage du lit de la Loire pourrait avoir des conséquences néfastes à la survie des espèces de gomphes. *Gomphus flavipes* (Charpentier, 1825) et *Ophipogomphus cecilia*, (Geoffroy in Fourcoy, 1785) présents sur le secteur, sont particulièrement importants à protéger de par leur intérêt patrimonial. L'étude des variables environnementales influençant la répartition de ces deux espèces en Loire apparaît comme un préalable nécessaire pour mieux estimer les conséquences des travaux de restauration sur leur dynamique et ainsi de mieux orienter les mesures de conservation pouvant s'appliquer.

La répartition des communautés écologiques sur le territoire et les variables environnementales qui régissent leurs interactions sont évaluées depuis de nombreuses années par des méthodes statistiques. Avec le temps, de nouvelles méthodes d'analyse ont émergé permettant notamment de modéliser spatialement la distribution des populations. La méthode du maximum d'entropie, dite MaxEnt, en est un exemple. Ce logiciel fait actuellement partie des outils de modélisation spatiale les plus performants. Néanmoins est-il aussi accessible que ne le sont les tests statistiques ? Chaque outil d'analyse présente des avantages, des inconvénients et des conditions d'utilisation bien spécifiques. Il apparaît nécessaire d'étudier ces différents outils afin de déterminer le plus pertinent d'entre eux pour expliquer la répartition des *Gomphidae* en Loire.

Le présent document est introduit par le contexte du projet avec une brève explication de l'écologie des *Gomphidae* et des menaces qui perturbent leur environnement. Les actions préalablement menées pour échantillonner des individus entre Saint-Mathurin-sur-Loire et Oudon seront également présentées. Le fonctionnement général du logiciel MaxEnt sera expliqué et sa compatibilité avec les données récoltées sur le terrain sera testée. Ensuite, l'influence des variables environnementales sur la répartition des individus sera évaluée par plusieurs tests statistiques. Un bilan de ces méthodes permettra de déterminer l'outil d'analyse le plus pertinent. Les perspectives pour la suite de l'étude concluront ce rapport.

La famille des *Gomphidae* appartient au sous-ordre des Anisoptères. Elle présente ainsi les caractéristiques spécifiques aux Odonates, c'est-à-dire quatre ailes membraneuses, de tailles égales, ainsi que de courtes antennes. Les individus sont reconnaissables par leurs gros yeux globuleux, qui, à l'exception des autres familles d'Anisoptères, ne se rejoignent pas en un point (OPIE, PNA Odonates). Quatre genres et neuf espèces sont observables sur le territoire national. Parmi elles, *Gomphus flavipes* et *Ophiogomphus cecilia* présentent un enjeu patrimonial et sont considérées comme « prioritaires » selon le Plan National d'Actions en faveur des Odonates.



Le gomphe à pattes jaunes (*Gomphus flavipes*) et le gomphe serpentín (*Ophiogomphus cecilia*) sont particulièrement intéressants à étudier de par leur caractère sténoèce. Ces derniers ont des niches écologiques à faible amplitude signifiant que les perturbations de leur milieu ont des conséquences directes sur leur survie. Les individus sont particulièrement sensibles à l'altération de leur habitat (Gouraud, 2015). De plus, leur cycle de vie est dit amphibiotique (Figure 2). Leur phase larvaire, longue de 2 à 5 ans et segmentée de 11 à 15 stades se fait dans le milieu aquatique. Pendant cette période, les individus vivent dans le sédiment grâce à leur corps aplati et aux éperons de leurs tibias atrophiés leur permettant de s'enfouir dans un substrat à granularité relativement fine. Afin d'être le moins visible possible des proies et des prédateurs, la stratégie adoptée par les individus est de s'enfouir dans le sédiment en ayant seulement la tête et le dernier segment abdominal à la surface. Les larves se nourrissent principalement d'autres invertébrés tels que les Chironomidae, les Oligochètes et les Amphipodes (Merlet & Houard, 2012). Après l'émergence, les individus vivent en milieu terrestre. Ainsi leur présence est représentative d'une bonne qualité de l'écosystème en général : de la qualité de l'eau à la qualité des sédiments en passant par un bon maintien de l'intégrité des berges et de la ripisylve. D'après Burcher & Smock, 2002, les larves fouisseuses sont de bons bio-indicateurs de la qualité du substrat.

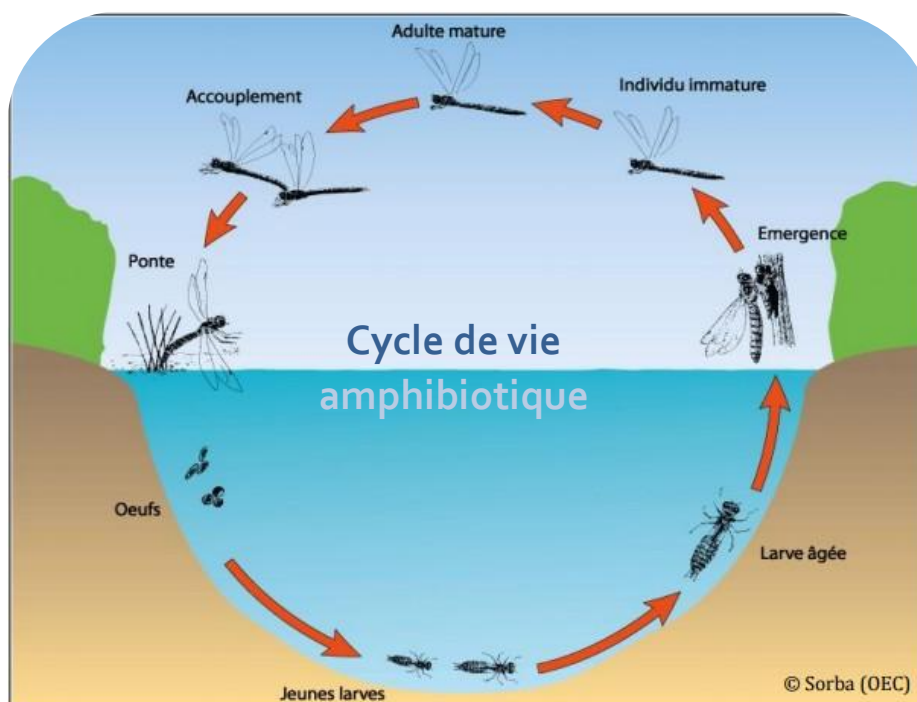


FIGURE 2 : SCHEMA DU CYCLE DE VIE DES ODONATES (SORBA, OEC IN BERQUIER C., 2015)

1.2 Habitats favorables

Les larves de *Gomphus flavipes* vivent dans les courants lents dans un sédiment à granularité fine : sableux, vaseux ou argileux. Elles affectionnent les bancs de sable et zones de ralentissement où un dépôt de sédiments fins se produit (PNA *Gomphus flavipes*, 2014). Elles sont donc souvent observables derrière des zones d'encombres ou d'aménagements dans le lit mineur du cours d'eau tels que les épis (Figure 3) (Heidemann & Seidenbusch, 2002 in Gouraud, 2015). Les larves sont sensibles à des phénomènes de variation de température. Elle doit être supérieure à 17°C pour permettre une éclosion des œufs. Des modifications de la composition chimique de l'eau peuvent également avoir des influences négatives sur la phase larvaire ; le faible pouvoir de dispersion de l'espèce la rendant d'autant plus sensible à la modification de ces paramètres. Il faut néanmoins noter des phénomènes de dérive larvaire passive, au gré des

crues, assez importants chez les larves de gomphes à pattes jaunes (PNA *Gomphus flavipes*, 2014). Au moment des émergences, le maintien de l'intégrité des berges et la présence de haies et de prairies à proximité directe des cours d'eau sont essentielles pour la phase de maturation et d'alimentation des adultes.



FIGURE 3 : PHOTO D'UN EPIS DE NAVIGATION SUR LA LOIRE, ZONE FAVORABLE AU DEVELOPPEMENT LARVAIRE DES INDIVIDUS (GIP ESTUAIRE)

Le gomphe serpentín préfère des eaux courantes, de bonne qualité et bien oxygénée (Suhling & Müller, 1996 in Merlet F. & Houard X., 2012). Les zones de micro-habitats favorables pour les larves sont caractérisées par un substrat meuble, à granularité plus grossière, allant de sables à graviers (Merlet F. & Houard X., 2012) et une eau peu profonde. *Ophiogomphus cecilia* semble un peu moins exigeant dans le choix de ses milieux de développement que *Gomphus flavipes* (Gouraud, 2015). Par exemple, la largeur des cours d'eau ne semble pas être un facteur influençant le développement de l'espèce. Elle pourrait aussi bien être présente dans le cours principal des fleuves que sur ses bras secondaires. Néanmoins, dans le bassin de la Loire, malgré de nombreuses observations d'exuvies sur les affluents du fleuve, les populations connues seraient essentiellement observées sur la Loire elle-même (Lett *et al.*, 2001). La végétation aquatique ne serait également pas un facteur très explicatif de leur présence selon Dommanget, (2004, in Merlet & Houard, 2012).

1.3 Menaces et statuts de protection

De manière générale, les principales menaces affectant les gomphes sont liées d'une part à la pollution de l'eau et aux modifications de sa composition chimique et d'autre part à la rectification des cours d'eau avec la canalisation ou encore la chenalisation. Les *Gomphidae* ont souffert de ces perturbations les deux siècles derniers (Grand & Boudot, 2006) ; leurs effectifs ont drastiquement chutés. Ainsi, la disparition croissante des espèces de *Gomphidae* a été prise en considération et des statuts de protection leur ont été affectés. *Gomphus flavipes* et *Ophiogomphus cecilia* étaient auparavant considérés comme « en danger » en Europe et en France sur les listes rouges des Libellules (UICN). Les deux espèces sont aujourd'hui classées comme « préoccupation mineure » au nom de cette liste ; l'UICN considérant que les populations sont désormais stables à l'échelle européenne. Elles bénéficient néanmoins d'un statut de protection sur l'ensemble du territoire national au nom de l'article 2 de l'arrêté du 23 avril 2007, intitulé « les insectes protégés sur le territoire et les modalités de leur protection », et figurent aux annexes II et IV de la directive européenne Habitats-Faune-Flore.

Etant particulièrement sensibles aux altérations de leurs habitats, les populations de *Gomphidae* peuvent être menacées dès la réalisation de travaux dans le lit mineur du cours d'eau. Sur le bassin ligérien, dans le cadre du Contrat pour la Loire et ses Annexes (CLA), la restauration de plus d'une vingtaine d'annexes hydrauliques situées en Loire aval est prévue. Les épis de navigation seront retirés permettant une remobilisation des sédiments piégés derrière ceux-ci et dans les bras secondaires (Rodrigues, 2016). Néanmoins, ces zones sont favorables au développement des individus de *Gomphidae*. La destruction de ces ouvrages peut alors être une véritable menace pour la survie de l'espèce dans le secteur aval de la Loire. A contrario, la reconnexion du lit avec ses annexes hydrauliques pourrait être un point positif pour les individus qui finissent parfois leur croissance dans ces zones après une crue. Il est donc intéressant d'avoir plus de connaissances quant aux exigences des individus afin d'estimer si une compensation est envisageable entre les zones d'habitats favorables au développement larvaire des *Gomphidae*, qui vont être détruites par les travaux, et les nouveaux créés à leur issue, via la reconnexion des annexes hydrauliques. L'étude des variables environnementales influençant le plus leur répartition en milieu naturel est donc un élément clé pour anticiper au mieux les conséquences engendrées par les futures modifications du lit mineur.

1.4 Traits écologiques permettant une modélisation de leur répartition

Les *Gomphidae* sont beaucoup plus exigeants en termes d'habitats que d'autres insectes : ils sont particulièrement sensibles à la qualité et aux conditions de leur milieu. La température de l'eau, la profondeur mais surtout la vitesse du courant et la granulométrie du substrat semblent être des traits écologiques à forte influence sur la présence des individus. En effet, passant la première partie de leur vie enfouis dans le sédiment, les caractéristiques du substrat sont particulièrement importantes : *Gomphus flavipes* appréciera particulièrement un substrat très fin avec des débris organiques à la surface des sédiments tandis qu'*Ophiogomphus cecilia* affectionnera davantage des plages sableuses ou sablo-graveleuses (Gouraud, 2015).

De plus, la vitesse du courant est un élément essentiel de leur habitat, à prendre en considération. La préférence des gomphes à pattes jaune pour les zones à courant lent est une caractéristique majeure de leur habitat de prédilection. Un courant plus fort, limitant le colmatage interstitiel, sera plus apprécié par les gomphes serpent.

La vitesse du courant ainsi que la nature du substrat sont donc deux variables environnementales particulièrement explicatives de la répartition des individus. La faible amplitude de la niche écologique des *Gomphidae* devrait théoriquement permettre l'obtention d'un modèle prédictif plus fiable (Thomas B et al., 2015). De plus, de par leur pouvoir de dispersion assez limité et leur comportement grégaire, la modélisation de leur répartition permettrait de rendre compte du degré de fragmentation des habitats et éventuellement d'observer l'isolement de certaines populations.

1.5 Protocole d'échantillonnage des *Gomphidae* : un exemple sur la Loire

Afin de juger l'impact potentiel des travaux de restauration du lit mineur pour le CLA sur les populations de *Gomphidae*, cinq stations d'échantillonnage ont été choisies entre Oudon et Saint Mathurin-sur-Loire afin d'améliorer les connaissances globales sur l'écologie de ces espèces au sein de la Loire. Ces sites ont été choisis de par leur proximité avec des épis de navigation qui vont se voir retirés lors de futures opérations de restauration. De plus, des exuvies ont été observées et récoltées dans ces zones, lors de précédentes études réalisées par le Conservatoire d'Espaces Naturels (CEN). Dans la continuation de cette étude, le but de la campagne menée à partir de 2016 est d'acquérir plus de savoirs autant sur la composition des populations larvaires que sur les caractéristiques de leurs niches écologiques ou bien sur leurs comportements en Loire (Rodrigues, 2016).

Des campagnes d'échantillonnage des larves de *Gomphidae* ainsi que le relevé d'un certain nombre de données environnementales (Annexe 1) ont été effectuées en septembre 2016 et en juin 2017 sur cinq stations prédéfinies (Figure 4). La station d'Oudon, située le plus à l'aval, est particulièrement intéressante de par sa localisation à l'interface entre la limite aval de distribution des gomphes et la limite amont de la marée dynamique. Les stations d'Ancenis, Varades et le Fresne-sur-Loire ont été sélectionnées pour leur proximité avec des épis d'enlèvement. La station de Saint-Mathurin-sur-Loire a, quant à elle, un rôle de station de référence, n'étant pas localisée dans une zone soumise à travaux. Les trois sites aval sont classés comme site d'intérêt communautaire (SIC) Natura 2000 « Vallée de la Loire de Nantes aux Ponts-de-Cé et ses annexes » pour leur richesse d'habitats.



FIGURE 4 : LOCALISATION DES STATIONS D'ÉCHANTILLONNAGE SUR LA ZONE D'ÉTUDE

A chaque station, l'échantillonnage a été réalisé selon deux transects transversaux (Figure 5) contenant 10 points de prélèvements chacun espacés de 15 mètres, et d'un transect longitudinal, espacé de moins d'1 mètre de la berge, contenant 20 points de prélèvements. Un prélèvement de substrat a donc été réalisé à chaque point à l'aide d'un haveneau pour les zones peu profondes ou d'un cône berthois pour les zones de profondeur plus importante. Les échantillons ont été lavés, observés et triés en laboratoire. Les larves de *Gomphidae* ont été identifiées jusqu'à l'espèce et la majorité des autres invertébrés a été identifiée jusqu'au genre. Les coordonnées GPS en Lambert 93 de chaque point d'échantillonnage ont été notées. De plus, la nature du substrat en fonction de sa granulométrie a été observée et la présence de matière organique ou de colmatage a été identifiée. La hauteur d'eau (cm), la température (°C), la vitesse de courant (m/s) mais aussi la concentration en oxygène (mg/L) sont autant d'éléments qui ont été observés à chaque point.

Tous les échantillons de la campagne 2016 n'ayant pas été encore identifiés, la présente étude s'appuiera uniquement sur les données récoltées sur les stations d'Oudon, de Varades et de Saint-Mathurin-sur-Loire. De plus, malgré la forte présence d'exuvies sur ces zones, répertoriées par le CEN, le nombre de larves dénombrées sur ces stations est très faible. Afin d'estimer avec le plus d'exactitude possible l'employabilité du logiciel MaxEnt, les données concernant le genre *Caenis* seront utilisées pour des essais de modélisation.

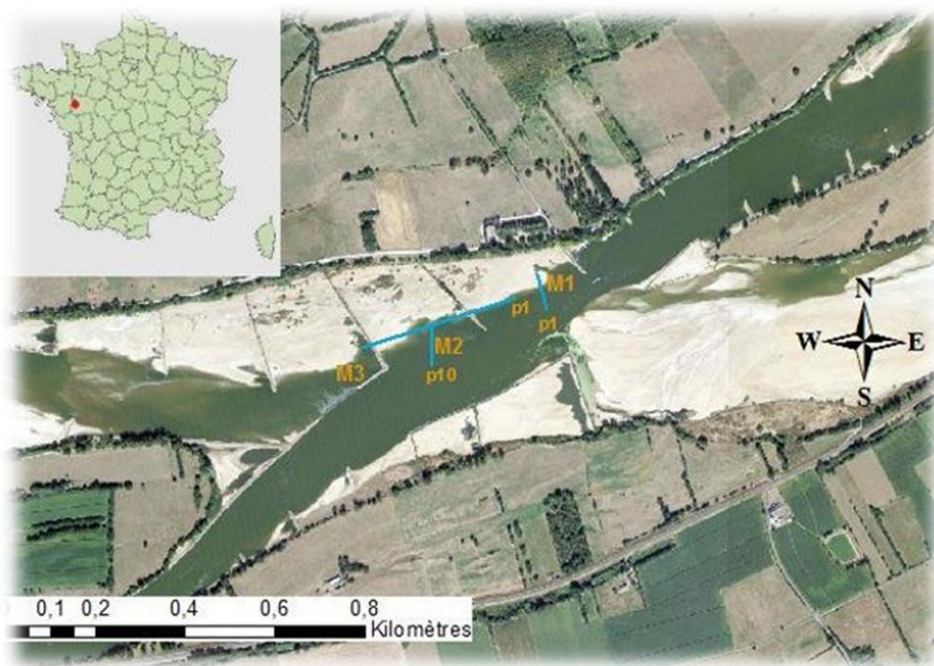


FIGURE 5 : VUE AERIENNE DE LA STATION SATELLITE DE VARADES ET POSITIONNEMENT DES TRANSECTS (DONNEES : NINA RICHARD ; REALISATION : KEVIN SILLIAU)

2. UTILISATION DU LOGICIEL MAXENT

2.1 Principe général de la méthode

La méthode MaxEnt, ou maximum d'entropie, correspond d'après [Kato \(2012\)](#), « à la loi de probabilité permettant de prendre en compte un maximum d'incertitudes dans le processus de modélisation ». En d'autres termes, cette méthode permet de prédire la distribution spatiale potentielle d'une espèce animale ou végétale à partir de données de présence et de variables environnementales connues caractérisant la niche écologique de l'espèce ([Phillips *et al.*, 2006](#)). Il est aujourd'hui fondamental d'avoir de meilleures connaissances sur l'écologie des espèces et leur distribution géographique afin de mettre en place des plans de conservation plus fonctionnels ([Elith *et al.*, 2006](#)).

Disponible depuis 2004, cette méthode a été, dès lors, intensivement utilisée pour les modélisations de la distribution des espèces ([Elith *et al.*, 2011](#)). Elle est devenue populaire pour deux principales raisons : elle a généralement une meilleure précision de prédiction que les autres méthodes et elle est relativement facile d'utilisation. Elle est ainsi devenue très compétitive avec les méthodes les plus performantes de par, également, la diversité des sujets qu'elle peut traiter à travers des applications écologiques, d'évolution, de conservation ou encore de biosécurité ([Thomas *et al.*, 2015](#)). Cet outil a notamment été utilisé pour mettre en exergue des corrélations d'occurrence entre les espèces, pour cartographier des distributions actuelles ou pour prédire leurs évolutions dans le temps et dans l'espace ([Elith *et al.*, 2006](#)). Plus de 1000 études scientifiques ayant utilisées MaxEnt comme outil d'analyse de données ont été publiées entre 2006 et 2013 ([Merow *et al.*, 2013](#)).

Le programme MaxEnt présente l'avantage de nécessiter seulement des données de présence et des données environnementales (précipitation, température). Il ne nécessite pas de données d'absence comme c'est le cas pour de nombreuses méthodes ([Elith *et al.*, 2011](#)). Ces données peuvent être très difficiles à obtenir aussi bien sur le terrain que dans les bases de données historiques. Par exemple, les muséums ou les herbarium présentent une source très importante de données d'occurrence de présence mais non d'absence. La majorité de ces bases données est issue de plus d'un siècle d'investissements publics et privés ayant permis de collecter de très nombreuses données d'occurrence d'espèces ([Elith *et al.*, 2011](#)). La méthode MaxEnt permet de maximiser l'utilité de telles ressources d'informations. De plus, cet outil permet l'utilisation de données aussi bien quantitatives que qualitatives. Par ailleurs, MaxEnt présente quelques faiblesses. Il induit une trop forte autocorrélation spatiale dans les données d'observation pouvant fausser le modèle et introduire des biais dans les prédictions ([Kato, 2012](#)). Il est aussi important de noter que ce modèle de distributions d'espèces ne représente qu'une partie de la niche potentielle ; la modélisation est limitée à la disponibilité des variables environnementales. Elle ne prend également pas compte les interactions biotiques entre les niches ([Luque *et al.*, 2012](#)).

2.2 Fonctionnement de l'outil

D'après les termes de [S. J. Phillips *et al.*, \(2006\)](#), le fonctionnement de l'outil MaxEnt se définit de la manière suivante : « Sur un ensemble de pixels X (correspondant aux pixels de la zone d'étude), on prend plusieurs points de présence d'un phénomène (x_1, \dots, x_m) issus d'une distribution de probabilités inconnue π . » Ainsi, deux types de données sont nécessaires : les

coordonnées de présence de l'espèce étudiée ainsi que les variables environnementales caractérisant l'habitat des individus.

- Les données d'observation contenant le nom de l'espèce et les coordonnées de présence (longitude et latitude) des individus observés doivent être renseignées dans un fichier au format .csv.
- Les variables environnementales doivent préalablement être rasterisées afin d'être chargées dans le logiciel sous un format ascii (.asc). L'ensemble de ces variables fait référence à une même aire géographique, l'aire d'étude, fragmentée en une grille de pixels. Chacune d'entre elles doit couvrir la même surface au pixel près.

Une fois ces deux types de données créés, ils peuvent être chargés directement dans le logiciel MaxEnt (Phillips, 2017) (Annexe 2): le fichier des données de présence est ouvert dans la commande « samples » tandis que le dossier contenant l'ensemble des contraintes environnementales rasterisées est ouvert dans la commande « layers ». Cet outil performant propose de nombreux paramètres différents afin d'obtenir un traitement des données le plus adapté possible aux attentes de la modélisation. Néanmoins, d'après Merow *et al.* (2013), l'emploi de ces variantes n'est pas toujours bien expliqué dans les études c'est pourquoi de nombreux utilisateurs se servent uniquement des paramètres par défaut alors que des paramètres plus spécifiques seraient plus adaptés à leur besoin. La fonction « run » permet ensuite au logiciel de lancer un algorithme (Figure 6) qui quantifie la proximité de chaque pixel avec l'habitat occupé (Kato, 2012) afin de renseigner une probabilité d'occurrence spatialisée.

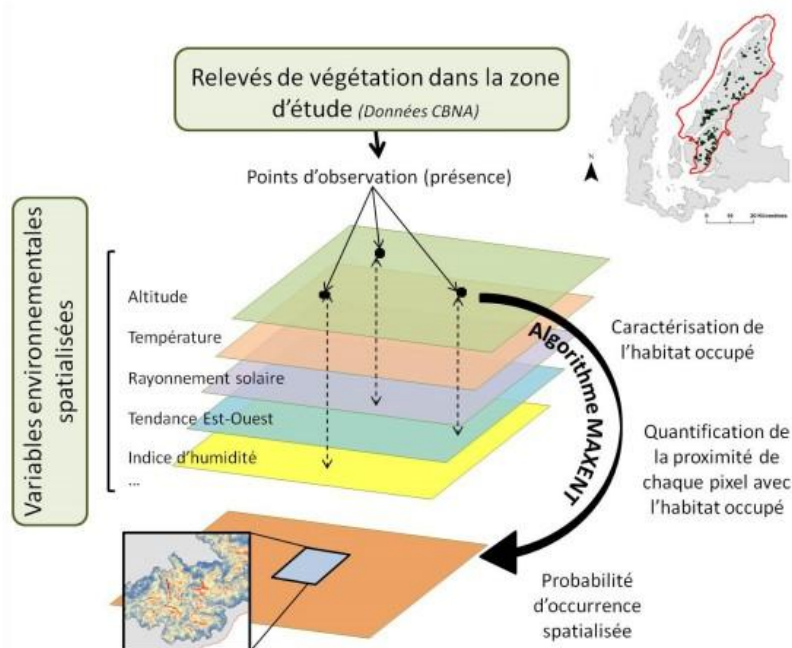


FIGURE 6 : ALGORITHME DE FONCTIONNEMENT DU LOGICIEL MAXENT

Le principe général du résultat fourni par le modèle est que la meilleure estimation d'une distribution est celle qui est la moins contraignante pour l'espèce (Kato, 2012). Ainsi, Luque *et al.*, (2012) écrivent que les contraintes sont définies après comparaison de la distribution des valeurs des variables environnementales aux points d'observation avec leur distribution pour des points pris aléatoirement sur la zone ». Si les valeurs des variables en un point donné correspondent davantage aux conditions aux points d'observation qu'aux points pris au hasard, la probabilité d'occurrence prédite pour l'espèce sera élevée (Phillips *et al.*, 2006). En effet, il

existe une corrélation entre les données de présence d'une espèce à un point donné et les conditions environnementales à ce point. Si ces conditions sont réunies à un autre endroit sur la zone, les chances que des individus de l'espèce donnée soient également présents à cet endroit sont importantes.

Le logiciel fourni plusieurs types de résultats sous forme de tableaux, de cartes ou de graphiques (Phillips, 2017) (Annexe 3). Afin d'évaluer la robustesse du modèle et ainsi sa fiabilité, une validation des résultats est nécessaire. Elle se fait via deux premières courbes : une courbe d'omission et de prédilection qui permet d'évaluer une erreur de prédiction du modèle qui indiquerait l'absence d'un individu alors que dans les faits il est présent (Luque *et al.*, 2012), et la courbe ROC (Receiver Operating Characteristic) qui permet d'évaluer la capacité de discrimination du modèle. L'algorithme fournit aussi une carte de répartition spatiale des espèces avec un dégradé de couleurs indiquant si la probabilité de présence (liées aux conditions) est importante ou non. Enfin, afin de savoir quelles variables influencent le plus le modèle, un tableau présentant le pourcentage de contribution des variables, un graphique (si l'option jackknife est sélectionnée) ou encore des courbes de relations entre les variables sont créés. Ces résultats sont importants afin de prendre connaissance des paramètres environnementaux ayant l'impact le plus significatif sur les conditions favorables d'un habitat et indirectement, sur la répartition des espèces sur le territoire.

2.3 Conditions d'utilisation

Malgré une prise en main du logiciel relativement facile, comme tout logiciel, MaxEnt requiert des conditions d'utilisation spécifiques :

- (1) Les données de présence doivent être le plus indépendantes possibles les unes des autres. Dans les faits, les points d'observation devraient être définis aléatoirement sur l'ensemble de la zone d'étude (Elith *et al.*, 2006) afin d'avoir une vision plus exhaustive des variables environnementales les plus favorables au développement des individus.
- (2) La sélection des variables doit être faite en fonction de leur disponibilité dans la zone d'étude et sur le territoire national (reproductibilité) ainsi que leur pertinence par rapport à l'écologie des espèces étudiées (Luque *et al.*, 2012). Les valeurs des variables choisies doivent pouvoir être accessibles, notamment dans des bases de données, en tout point de la zone d'étude. En effet, il est indispensable de pouvoir comparer les conditions environnementales aux points d'observations des individus avec ces mêmes conditions sur l'ensemble du territoire afin d'estimer leur probabilité de présence sur d'autres sites. Il est par exemple possible d'utiliser des données issues de la BD CarTHAgE (Base de Données sur la CARTographie THématique des AGences de l'eau et du ministère chargé de l'environnement), des données d'occupation du sol via le Corine Land Cover 2012 (Union européenne – SoeS) ou bien des données bioclimatiques renseignées sur la base de données WorldClim.
- (3) Le choix de l'échelle d'étude est très important et dépend des variables environnementales choisies. Par exemple, si les données d'observation sont comparées à des données bioclimatiques, une échelle nationale devra *a minima* être utilisée afin d'obtenir des résultats fiables (Phillips *et al.*, 2004). Même si aucune surface optimale n'a été définie dans la littérature de par sa forte dépendance aux caractéristiques propres à chaque étude, nous pouvons tout de même admettre que la majorité des modélisations utilisant l'outil MaxEnt travaille à des échelles relativement étendues (de la région au continent).

2.4 Employabilité du logiciel MaxEnt pour l'analyse de la répartition des *Gomphidae*

L'objectif de la première partie de cette étude était d'évaluer l'employabilité du logiciel MaxEnt pour traiter les données acquises suite à l'échantillonnage de *Gomphidae* dans le cadre du Contrat pour la Loire et ses Annexes. Ainsi, une comparaison de ces données avec les conditions d'utilisation du logiciel a été faite (Figure 7). Des incohérences ont alors été observées.

Tout d'abord, les prélèvements ont été réalisés selon des transects. Les points sont donc proches les uns des autres et non indépendants et aléatoires (Figure 7 A et B). La condition de points d'observation pris au hasard n'est donc pas respectée. En effet, les stations d'échantillonnage ont toujours été choisies dans des zones particulières : au niveau des épis de navigation. Les éventuelles données de présence des individus seront donc disponibles seulement à ces points alors qu'un échantillonnage au hasard sur l'ensemble de la zone aurait pu révéler la présence de *Gomphidae* à d'autres endroits. Un biais important s'installe quant à l'évaluation des variables environnementales pouvant expliquer la présence des larves de *Gomphidae*.

Ensuite, une partie des données environnementales recueillies sur la zone d'étude n'est pas accessible, en tout point, sur l'ensemble de la zone. En d'autres termes nous possédons seulement les données mesurées aux points d'observation (Figure 7B) et nulle part ailleurs sur la zone. Bien que certaines données, telles que la température de l'eau, pourraient être accessibles via des bases de données mises en ligne par des organismes comme les Agences de l'eau, les données correspondant à la vitesse de courant ou à la granularité du sédiment ne sont, par exemple, pas encore disponibles. Nous avons précédemment montré que ces deux paramètres étaient certainement les plus explicatifs de la présence des individus. Ils sont donc indispensables pour acquérir des estimations de distribution des espèces cohérentes. N'ayant pas ces données, il est impossible d'estimer la présence des individus sur l'ensemble de la zone en fonction des caractéristiques collectées aux points d'observation.

Enfin, en l'absence de surface optimale d'étude, nous pouvons uniquement émettre l'hypothèse que la campagne d'échantillonnage des *Gomphidae* de septembre 2016 a été réalisée sur une zone trop restreinte pour fournir des probabilités d'occurrence fiables. En effet, il semble exister un biais trop important entre des données d'observation échantillonnées très localement, à l'échelle de stations, et les variables environnementales mesurées, en général, à l'échelle nationale. Si des mesures de vitesse de courant ou de granularité du substrat étaient réalisées, elles seraient certainement faites sur l'ensemble du réseau hydrographique de la Loire. Il serait alors intéressant de travailler à cette échelle-là pour estimer avec plus d'exactitude la répartition des *Gomphidae*.

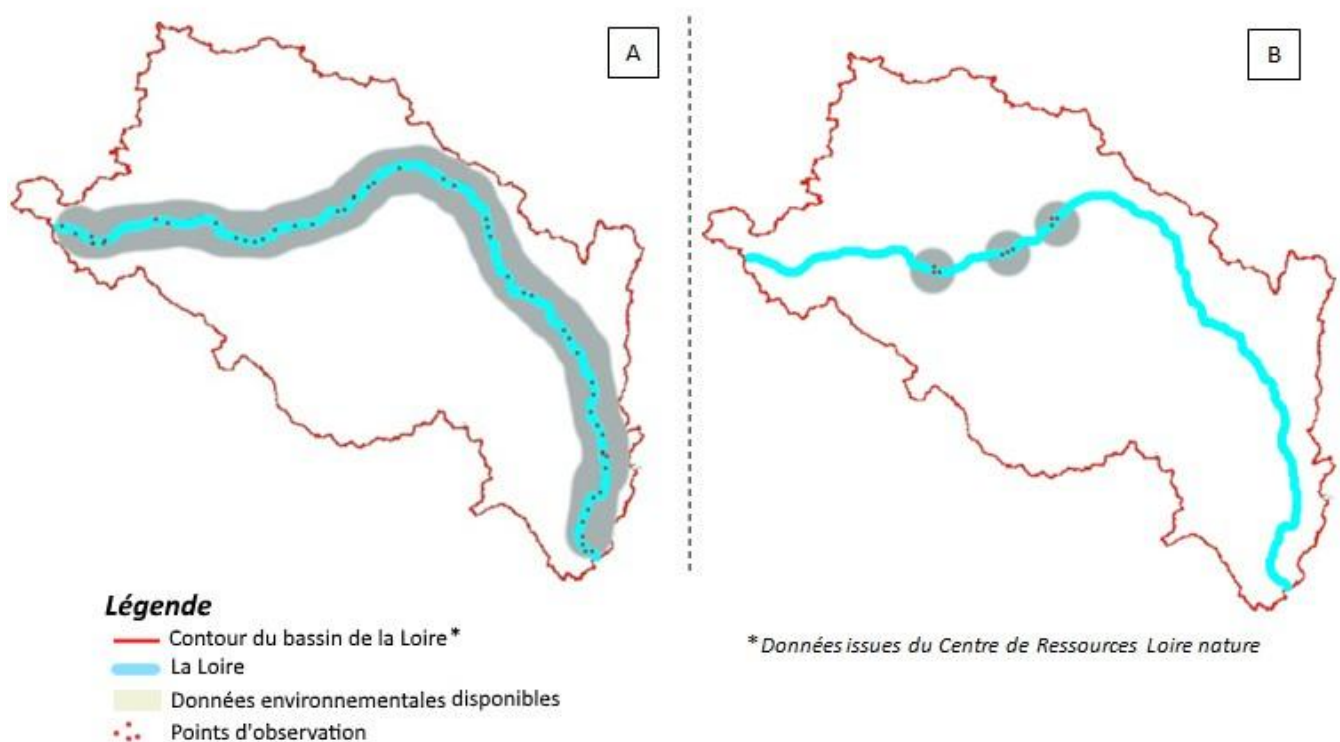


FIGURE 7 : SCHEMAS PRESENTANT LES DONNEES THEORIQUES REQUISES D'APRES LES CONDITIONS D'UTILISATION DU LOGICIEL (A) ET LES DONNEES REELLEMENT RECOLTEES PAR L'ETUDE (B).

Au vu du protocole d'échantillonnage mis en place et du choix des variables environnementales mesurées lors de la campagne de prélèvements de septembre 2016, l'utilisation du logiciel MaxEnt pour le traitement des données collectées ne semblent par être adapté. Trop de contraintes d'utilisation du logiciel ne correspondent pas aux conditions particulières de prélèvements et aux attentes de la modélisation pour cette étude. En effet, le logiciel pourrait être utilisé si l'objectif était de cartographier la répartition des individus en fonction de critères bioclimatiques ou d'occupation des sols par exemple. Ces données, étant disponibles en ligne, pourraient être spatialisées pour la zone d'étude et comparées avec les valeurs aux points d'observation. Néanmoins, l'objectif est tout autre puisqu'il concerne l'influence de paramètres plus caractéristiques du micro-habitat larvaire des *Gomphidae* tels que la vitesse du courant ou la nature du substrat. L'utilisation de l'outil MaxEnt, dans le but d'estimer la répartition des individus en fonction de ces paramètres, devra attendre la mise à disposition de telles données prises quotidiennement sur l'ensemble de la zone.

3. TESTS STATISTIQUES MULTIVARIÉS

Le logiciel de modélisation MaxEnt n'étant pas adapté au traitement des données de l'étude, des méthodes statistiques multidimensionnelles ont été testées afin de répondre à la question suivante : quelles variables environnementales influencent le plus l'abondance des individus du genre *Caenis* ? Ceux-ci appartiennent à l'ordre des Ephéméroptères dont ils se distinguent par leur unique paire d'ailes, tenues en position horizontale, et de surface importante par rapport à leur corps (Brulin, 2007). Dix espèces sont présentes en France se développant dans la majorité des rivières et plans d'eau du pays. Ce genre a été utilisé, à défaut des *Gomphidae*, pour les tests statistiques car le nombre plus important d'individus récoltés permet d'obtenir une vision plus cohérente des résultats obtenus. Si ceux-ci sont satisfaisants, les tests pourront dans un second temps être utilisés sur les populations de *Gomphidae*.

Les différents tests statistiques envisagés seront expliqués brièvement dans cette partie. Le test Hill Smith et des tests complémentaires seront présentés plus en détail ainsi que sa méthode d'utilisation sur les données de terrain et les résultats obtenus.

3.1 Choix du test

Les statistiques descriptives multidimensionnelles représentent l'ensemble des méthodes statistiques permettant de traiter simultanément un nombre quelconques de variables. Ces méthodes sont dites descriptives du fait qu'elles ne supposent aucun modèle sous-jacent de type probabiliste (Baccini, 2010). En d'autres termes, il n'est pas nécessaire de prouver que les variables suivent une loi normale. Plusieurs types de tests existent permettant de traiter soit uniquement des variables quantitatives, telle que l'ACP (Analyse en Composantes Principales) ou bien uniquement des variables qualitatives comme l'AFC (Analyse Factorielle des Correspondances). Des tests mixtes permettent de traiter aussi bien les variables qualitatives que les variables quantitatives. L'AFM, ou Analyse Factorielle Multiple, ou encore le test Hill Smith (Hill & Smith, 1976) en sont des exemples (Hervé, 2010). Pour chaque test statistique des conditions d'utilisation spécifiques doivent être respectées.

Le cas d'étude présent consiste à évaluer l'implication de variables environnementales qualitatives et quantitatives sur l'abondance des individus du genre *Caenis*. En effet, la hauteur d'eau, la vitesse du courant mais aussi le type de substrat sont autant de variables répertoriées sur le terrain qu'il semble intéressant de mettre en relation avec le nombre d'individus collectés. Afin de choisir le test statistique le plus adapté à l'étude en cours, plusieurs d'entre eux ont été la source de recherches. Leurs modalités d'application, leurs avantages et leurs inconvénients tout comme leur adéquation avec les attentes de la présente étude ont été évalués (Tableau 1).

TABEAU 1 : RECAPITULATIF DES TESTS STATISTIQUES ENVISAGÉS.

Tests	Types de variables traitées	Méthodes	Compatibilité
ACP	Nombre quelconque de variables quantitatives	Résumer un grand ensemble de données par le calcul de coefficients de corrélation linéaire entre chaque variable initiale et chaque facteur retenu.	Compatible partiellement : omission du traitement des données qualitatives dans l'analyse
AFC	2 variables	ACP particulière réalisée	Non compatible : le cas

	qualitatives	sur les profils associés à la table de contingence croisant les deux variables qualitatives considérées.	présent étudie une seule variable qualitative. Non prise en compte des données quantitatives.
ACM	Nombre quelconque de variables qualitatives	Equivalent d'une ACP mais pour des variables qualitatives.	Non compatible : le cas présent étudie une seule variable qualitative. Non prise en compte des données quantitatives.
AFM	Variables quantitatives et qualitatives	Synthèse de l'ACP et l'ACM permettant d'analyser simultanément plusieurs tableaux de variables (données homogènes à l'intérieur de chaque).	Compatible : traite dans un même modèle variables quantitatives et qualitatives.
Modèle linéaire	Variables quantitatives et qualitatives	Mise en relation d'une variable à expliquer Y avec p variables quantitatives ou qualitatives. Obtention de coefficients de corrélation et de p-value pour chaque variable.	Compatibilité à vérifier : le modèle doit suivre une loi normale, respecter différents tests sur les résidus et les données doivent être indépendantes.
Hill Smith	Variables quantitatives et qualitatives	Association d'une ACP et d'une ACM permettant de traiter des données multivariées mixtes.	Compatible : traite dans un même modèle variables quantitatives et qualitatives.

Suite aux recherches préliminaires sur les différents tests statistiques multivariés, trois d'entre eux semblent pouvoir convenir particulièrement au cas d'étude : les modèles linéaires, l'AFM et le test Hill Smith. En effet, bien que l'ACP puisse être utilisée pour l'analyse d'une partie des données (quantitatives), les tests cités ci-dessus sont plus complets et permettent de traiter des variables mixtes. L'ACP ne pourrait être utilisée qu'après quantification des variables qualitatives. Le modèle linéaire est quant à lui une méthode intéressante qui permet de connaître la variable qui influence le plus l'abondance des individus (Chouquet, 2010). Néanmoins, son utilisation exige le respect préalable d'un certain nombre de tests. Afin d'avoir une première vision de l'influence potentielle des variables par rapport à l'abondance des individus, ce test ne sera pas réalisé dans cette étude. Cependant il serait intéressant de l'implémenter ultérieurement pour avoir plus de détails quant à la puissance des interactions entre les données environnementales. L'AFM et l'analyse Hill Smith sont des techniques plus globales qui permettent de savoir si les variables sont corrélées ou non avec l'abondance. Ces tests étant assez similaires, seul le test Hill Smith a été utilisé. Il pourrait néanmoins être intéressant de réaliser une AFM avec les mêmes données pour s'assurer de la cohérence des résultats obtenus. Faute de temps, cette analyse complémentaire n'a pas été effectuée.

3.2 Le test Hill Smith

L'analyse Hill et Smith (Hill & Smith, 1976) doit son nom à ses deux créateurs, Mark Oliver Hill et Anthony John Edwin Smith. Publié en 1976, ce test statistique permet d'étudier les relations multivariées d'un ensemble de variables mixtes continues, ordinales et/ou qualitatives

distinctes (Hervé, 2010). L'objectif de cette analyse est de synthétiser l'information en réduisant le nombre de dimensions. Ainsi, une lecture visuelle et simple des relations entre les variables du jeu de données est possible. Le test Hill Smith s'apparente à une technique mixte entre l'ACP et l'ACM (Chessel & Lascaux, 2007). Il intègre les deux techniques pour mettre en relation des variables dont le format diverge. Malgré la puissance de cette technique et son utilisation certaine pour un grand nombre d'études, elle est peu citée dans la littérature.

L'analyse d'Hill Smith exige différentes conditions d'utilisation qui sont expliquées dans son application avec les données de terrain. Un certain nombre de résultats sont fournis par cette méthode, et notamment une projection graphique des variables sur un plan factoriel.

*3.3 Application du test Hill Smith aux données récoltées sur le genre *Caenis**

Le test Hill Smith a été utilisé pour estimer l'influence des variables citées ci-après, sur l'abondance des individus du genre *Caenis* : la granularité (qualitatif), la vitesse du courant (quantitatif), la concentration en dioxygène (quantitatif), la hauteur d'eau (quantitatif). L'étude concerne trois stations (Saint-Mathurin, Varades et Oudon) pour lesquelles le dénombrement d'individus échantillonnés est terminé. Pour réaliser l'analyse, le logiciel de traitement statistique R couplé à R-Studio a été utilisé. Cet outil puissant, notamment dans l'analyse de fonctions statistiques multivariées, nécessite un « langage » informatique particulier.

Etape 1 : Création du tableau de données

Tout d'abord, il est nécessaire de classer les données dans un tableau au format .csv avec un séparateur point-virgule. Les entêtes du tableau ne doivent pas contenir de signes particuliers tels que des accents, des espaces ou des points. Dans notre cas, les colonnes correspondant à l'identification de la station, à l'abondance, à la vitesse du courant, au taux de dioxygène dans l'eau, à la hauteur d'eau et au type de substrat ont été créées (Annexe 4). De plus, chaque identifiant de station doit être unique. Il est de type 01P1. La première lettre correspond au lieu de la station (O pour Oudon, M pour Varades et SM pour Saint-Mathurin), le numéro suivant correspond au numéro du transect et le Px fait référence au numéro du point d'échantillonnage sur ce transect.

Lors des prélèvements, des problèmes ou des oublis ont dû se produire puisque la valeur de certaines données n'était pas renseignée dans le tableau. Cela était notamment le cas de quelques valeurs de vitesse de courant, d'une valeur du taux de dioxygène et d'une valeur de type de substrat. Le test ne traitant pas un tableau aux données manquantes, l'ensemble des valeurs des points d'échantillonnage contenant une valeur absente a été supprimé de l'échantillon final. De plus, les substrats vaseux et limoneux étant trop peu représentés, les résultats étaient faussés par leur présence dans le tableau. Les points contenant ce type de substrat ont également été supprimés. Le tableau final est visible en annexe (Annexe 4).

Etape 2 : Traitement des données sur R

Le tableau de données a été chargé dans le logiciel R-studio. La fonction `read.csv2(« titre.csv », header=TRUE, desc= « , »)` a ensuite permis de le charger dans l'interface de travail (Champely, 2005). Le caractère quantitatif de certaines variables a été renseigné par la fonction `as.numeric`, tandis que le caractère qualitatif de la variable « substrat » a été défini par la fonction `as.factor`. Une fois réalisé, la colonne « identification » a été retirée du tableau car celle-ci n'est pas une variable. Le test Hill Smith a ensuite été lancé grâce à la fonction associée : `hill<-dudi.hillsmith(titre,scannf = FALSE)`. Des fonctions complémentaires ont permis d'obtenir les coordonnées des variables et des représentations graphiques via les fonctions `scatter(hill)` et `s.label(hill$co)` (Champely, 2005). Le script entier est disponible en annexe (Annexe 6).

Étape 3 : Analyse bivariée des variables environnementales avec la variable « abondance »

Afin d'aller plus loin dans l'analyse, les variables quantitatives ont été comparées une à une avec la variable « abondance » par les tests de Pearson, de Kendall et de Spearman (Motulsky, 2007). Ces tests permettent de définir si les variables sont corrélées ou non avec l'abondance, si cette corrélation est positive ou négative et si elle est significative. Pour ce faire, la fonction cor.test a été utilisée en renseignant l'une des trois méthodes.

Le test de Fisher et du chi-deux (Motulsky, 2007) ont été utilisés pour comparer et évaluer une éventuelle corrélation entre l'abondance des individus et la variable « substrat ».

Enfin, la variable « substrat » a été divisée en deux variables « subs_sgr » (sable grossier) et « subst_s » (sable) dont les données ont été transformées en données de présence ou d'absence (0 ou 1) (Annexe 5). Ces variables sont donc devenues quantitatives et ont été comparées individuellement avec l'abondance par les tests de Pearson, Kendall et Spearman.

3.4 Résultats obtenus

Test Hill Smith

Le pourcentage d'inertie fourni par le modèle indique que 24,34 % de la variabilité des données est expliquée par les deux premiers axes. Cette valeur n'est pas très élevée mais est relativement satisfaisante pour un travail avec des données écologiques. De plus, le test fournit les coordonnées des variables projetées sur le plan. Enfin, il transmet des schémas de corrélation permettant d'indiquer de quelles manières les variables évoluent les unes par rapport aux autres. Ces rapports ne sont néanmoins pas quantifiés par le test Hill Smith. Le graphique créé par le modèle étant peu lisible, il a été refait sur XLSTAT via les coordonnées projetées des variables sur le plan (Figure 8).

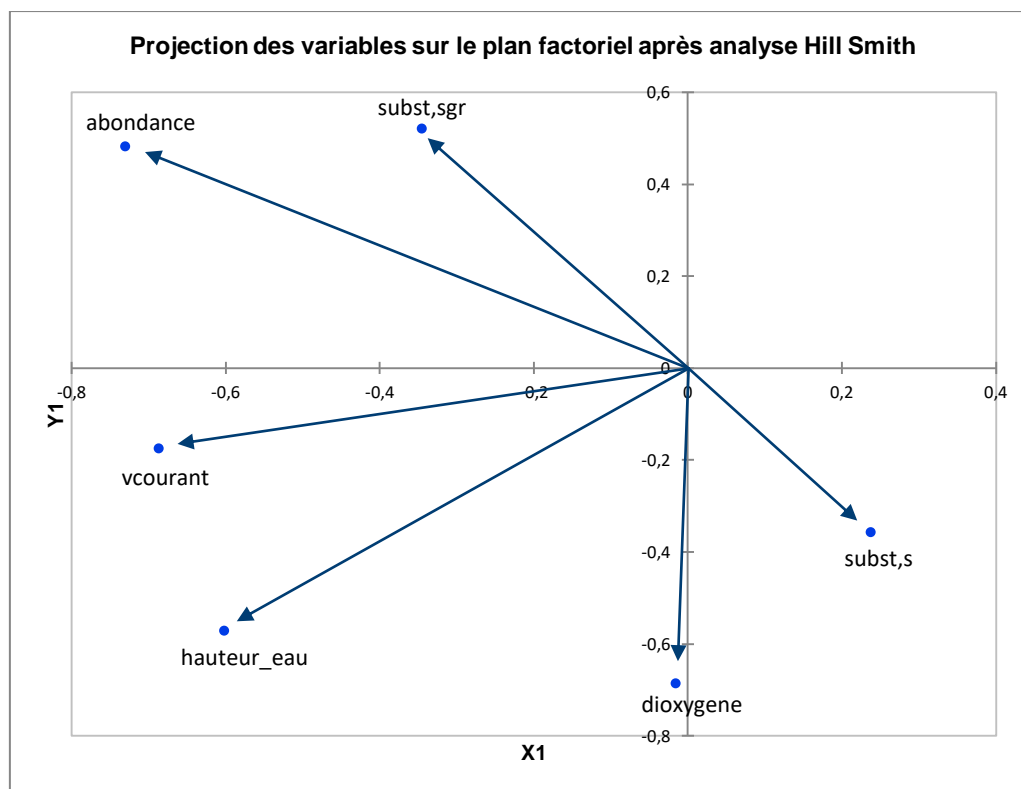


FIGURE 8 : PROJECTION DES VARIABLES SUR LE PLAN FACTORIEL

L'abondance des individus apparaît corrélée positivement avec un substrat de type sable grossier et avec la vitesse du courant. En d'autres termes cela signifierait que plus la vitesse du courant augmente et plus l'abondance est importante. A contrario, l'abondance semble corrélée négativement avec le taux de dioxygène dans l'eau et un substrat de type sable. L'orientation des axes rend difficile l'interprétation des relations entre l'abondance et la hauteur d'eau.

Tests complémentaires

Pour plus de justesse dans les interprétations, les résultats du taux de corrélation et de son taux de significativité, entre l'abondance et les différentes variables quantitatives, obtenus par les tests de Spearman et de Fisher ont été reportés dans le tableau 2. Les résultats fournis par les tests de Spearman, de Kendall et de Pearson étant presque identiques, seuls ceux issus du test de Spearman ont été renseignés. Il en est de même entre les résultats apportés par les tests de Fisher et du chi-deux.

TABLEAU 2 : RESULTATS DES TESTS DE SPEARMAN ET DE FISHER

Variables	p-value	Taux de corrélation	Test
Abondance vs vcourant	0.0001075	0.3689582	Spearman
Abondance vs hauteur_eau	0.0001388	0.3634079	Spearman
Abondance vs dioxygene	0.5991	-0.05188592	Spearman
Abondance vs substrat	0.8063	0.07516329	Fisher

D'après ces tests, il existe une corrélation positive (significative car $p\text{-value} < 0,05$), mais néanmoins faible, entre la vitesse du courant et l'abondance ainsi qu'entre la hauteur d'eau et l'abondance. L'abondance des *Caenis* augmente donc avec la hauteur d'eau et la vitesse du courant. Néanmoins, il n'y a pas de lien entre leur abondance et la concentration en dioxygène ($p\text{-value} > 0,05$). Enfin, d'après le test de Fisher, il n'y a pas de corrélation significative entre l'abondance et le type de substrat. Cependant, le graphique fourni par le test d'Hill Smith laisse penser que les individus ont une forte préférence pour les substrats de type sable grossier. Afin d'aller plus loin, les données qualitatives de la variable substrat ont été transformées en deux sous-variables quantitatives : « subs_sgr » (sable grossier) et « subst-s » (sable). Ces données ont été comparées une à une avec la variable abondance en utilisant les mêmes tests que ceux présentés précédemment (Tableau 3).

TABLEAU 3 : RESULTATS DES TESTS DE KENDALL ENTRE L'ABONDANCE ET LE TYPE DE SUBSTRAT

Variables	p-value	Taux de corrélation	Test
Abondance vs subs-sgr	0.01099	0.2172595	Kendall
Abondance vs subst_s	0.01099	-0.2172595	Kendall

L'abondance est corrélée positivement, mais faiblement, et de manière significative avec un substrat de type sable grossier. A l'inverse, l'abondance est corrélée négativement avec un substrat de type sable. Les individus semblent particulièrement affectionner un substrat plus grossier. Ces résultats sont en adéquation avec les observations faites par l'analyse Hill Smith. Dans les stations au substrat limoneux ou vaseux (qui avaient été retirées de l'analyse), seuls

quelques individus étaient présents, soit moins de 1% de l'échantillonnage total. En effet, 64% des individus ont été récoltés sur un substrat à sable grossier contre 35% récoltés sur un substrat sableux. Il semble que le test de Fisher ne donnait pas de résultats significatifs pour la variable « substrat » au vu de la grande divergence du nombre d'individus en fonction du type de substrat en question.

Pour conclure, le test statistique Hill Smith a permis de comparer au moyen d'une unique analyse, la corrélation générale de l'ensemble des variables, qu'elles soient quantitatives ou qualitatives, sans quantification. De premières hypothèses ont pu être émises et renforcées par des tests complémentaires permettant de mettre en exergue les corrélations positives et négatives entre les variables. Ainsi, les *Caenis* semblent se développer dans les milieux à vitesse de courant relativement élevée et à hauteur d'eau relativement importante sur un substrat de type sable grossier. Ces observations seront par la suite comparées avec les recherches scientifiques ayant déjà été réalisées sur le genre *Caenis*. Ainsi, si les résultats sont cohérents avec ceux référencés dans la littérature, l'emploi de ces tests statistiques pourra être validé.

3.5 Mise en relation des résultats obtenus avec l'écologie du genre *Caenis*, et validité du modèle

Les larves de *Caenis* peuplent une grande diversité de milieux aquatiques : les flaques, les marais, les lacs ou encore les rivières (Brulin, 2007). Leur ubiquité rend difficile l'attribution de leur présence ou de leur absence à un type particulier d'habitats. De plus, selon les espèces, les milieux occupés sont différents. Tandis que les larves de *C. lactea* affectionnent particulièrement les eaux stagnantes et se tiennent la plupart du temps sur un support immergé tel que les tiges ou les feuilles des végétaux ou encore sur du bois mort (Brulin, 2007), les larves de *C. robusta* et de *C. horaria* restent en suspension dans un film à la surface des mares ou sous les frondes des lentilles d'eau. Les espèces plus rhéophiles se retrouvent plus facilement dans les graviers. Plusieurs études scientifiques ont néanmoins été réalisées afin de définir au mieux les facteurs environnementaux les plus explicatifs de la présence des individus appartenant au genre *Caenis*. Lynda D. Corkum (1989) a notamment mené une étude sur les larves de *Caenis* sur 40 sites dans les rivières de l'Alberta au Canada. Ses recherches ont montré que les larves se trouvaient le plus souvent sur un substrat stable, dans des rivières profondes et étroites. Plusieurs scientifiques auraient préalablement montré (Bishop, 1973, in Corkum, 1989) que la répartition des individus dépendaient plus de la stabilité du substrat que de la taille des particules. Les larves de *Caenis* ne seraient donc pas très exigeantes vis-à-vis de la granularité du substrat. Au cours de ces études, elles ont majoritairement été trouvées dans des milieux riches en espèces végétales aquatiques. Quant à la vitesse du courant, les individus furent trouvés dans des zones aussi bien lentes que lotiques.

Ainsi, les *Caenis* sont majoritairement trouvés dans un milieu profond (Corkum, 1989). Ceci est en adéquation avec les résultats obtenus par l'analyse Hill Smith. Le type de substrat est très variable en fonction de l'espèce et ne semble pas être un critère de choix pour les individus qui peuvent occuper des milieux à granularité variée tant que le substrat est relativement stable. La vitesse de courant ne semble pas être non plus un paramètre déterminant quant à la répartition des individus puisqu'ils peuvent être retrouvés aussi bien dans des eaux stagnantes que courantes. Ainsi, la cohérence entre les résultats obtenus et les observations préalablement faites et ayant été renseignées dans la littérature est à modérer. La grande facilité d'adaptation des individus aux conditions environnementales de leur milieu ne permet pas d'affirmer que les résultats obtenus par les tests statistiques soient déterminants pour expliquer leur présence quel que soit le milieu. Néanmoins, vis-à-vis du cas d'étude présent, l'analyse Hill Smith et les tests statistiques qui lui ont été associés forment un modèle intéressant puisqu'ils fournissent des résultats clairs vis-à-vis des variables environnementales ayant une influence sur

l'abondance des individus. En effet, dans les conditions de prélèvements de cette étude, c'est-à-dire à l'arrière d'épis en Loire, une vitesse de courant importante, une hauteur d'eau élevée et un substrat de type sable grossier favorisent la présence de ces individus. Ce résultat n'est néanmoins pas exhaustif puisque dans d'autres conditions, d'autres paramètres pourraient apparaître comme plus favorables pour les larves de *Caenis*, ceux-ci ayant une grande valence écologique.

Une étude (Buffagni *et al.*, 2009) réunissant les observations faites et publiées dans plus de 2800 articles scientifiques de toute l'Europe, sur 344 espèces et sous-espèces d'Ephéméroptères, répertorie la présence des individus de chaque espèce en fonction des caractéristiques du milieu. D'après la catégorie « type de substrat » de cet inventaire, les populations de *Caenidae* ne sont pas spécifiques à un seul substrat. Selon les espèces les substrats préférentiels sont différents et ne sont pas uniques, c'est-à-dire que les individus d'une même espèce ont été retrouvés dans différents substrats. Les types de substrat dans lesquels les individus ont été le plus trouvés sont dans le sable, dans la vase, dans les macrophytes ou encore dans du gravier grossier. Le substrat n'est donc pas un élément pertinent pour expliquer la présence des larves de *Caenis*. Une autre étude (Ghani *et al.*, 2016) réalisée sur l'influence de la variation des conditions environnementales sur la diversité et l'abondance de plusieurs genres d'éphéméroptères au nord de la Péninsule Malaisienne a montré que la présence des *Caenidae* était fortement liée à des pH élevés et des températures de l'eau importantes. Cet exemple illustre le fait que des paramètres environnementaux différents selon la zone d'étude peuvent être explicatifs de la présence des individus de cette famille.

3.6 Bilan du modèle utilisé

L'analyse Hill Smith est un outil très intéressant à utiliser qui présente de grands avantages. Elle est relativement facile à implémenter et fournit des résultats numériques et graphiques clairs. De plus, elle permet de traiter des données qualitatives et quantitatives simultanément et donne une bonne vision des relations s'exerçant entre les variables. Néanmoins quelques limites à cette analyse et à son utilisation dans cette étude peuvent être soulevées.

Tout d'abord, cette analyse ne permet pas de quantifier les relations entre les variables et l'abondance. Elle a besoin d'être couplée à d'autres tests afin d'obtenir des taux de corrélation et de significativité. D'autres outils, tels que les modèles linéaires pourraient être intéressants à mettre en place afin de combler cette combinaison nécessaire de plusieurs tests statistiques. En effet, de tels modèles quantifient directement les corrélations détectées entre les variables analysées. L'analyse Hill Smith est donc intéressante pour un premier aperçu des corrélations existantes mais risque d'être insuffisante pour des études plus approfondies.

Ensuite, afin de s'assurer que l'analyse statistique réalisée soit en adéquation avec les traits écologiques appréciés par les individus et décrits dans la littérature, il aurait été intéressant de comparer les résultats obtenus avec les exigences d'un genre plus sélectif que ne l'est *Caenis*. La valence écologique de cette espèce est tellement large qu'il est difficile de savoir si le modèle a présenté les principales corrélations entre l'abondance et les différentes variables environnementales. Il serait intéressant d'identifier les individus échantillonnés jusqu'à l'espèce afin d'avoir des informations plus précises quant à leurs exigences, celles-ci variant d'une espèce à l'autre.

Enfin, ce test statistique pourra être utilisé uniquement lorsque le nombre d'individus de *Gomphidae* sera suffisamment important pour permettre d'acquérir des résultats significatifs. De nouvelles campagnes de prélèvements, avec d'éventuelles modifications des protocoles d'échantillonnage sont nécessaires afin d'obtenir des abondances suffisantes et pouvoir ainsi déterminer les variables environnementales influençant le plus la répartition de ces individus.

4. QUEL OUTIL EST LE PLUS APPROPRIÉ POUR EXPLIQUER LA RÉPARTITION DES *GOMPHIDAE* EN LOIRE ?

Cette étude a permis de tester différents outils d'analyse afin d'interpréter de la manière la plus juste possible, la répartition des *Gomphidae* en Loire. Nous rappelons que le projet d'enlèvements d'épis en Loire aval engendrera des conséquences directes sur les habitats actuels des populations de *Gomphidae* et notamment *Ophiogomphus cecilia* et *Gomphus flavipes*, espèces patrimoniales retrouvées derrière ces aménagements. Ces outils présentent des avantages, des inconvénients et des spécificités qu'il est intéressant de déterminer afin de choisir le test le plus adapté à l'étude en question.

Tout d'abord, le logiciel MaxEnt, outil intensivement utilisé de nos jours pour modéliser la répartition potentielle d'espèces, animales ou végétales, sur une zone donnée, a été testé. La puissance de cet outil est dans sa capacité à procurer des cartes de répartition potentielle uniquement à partir de données de présence et de variables environnementales. Le rendu visuel obtenu permet d'acquérir une vision très claire des lieux pouvant être favorables à une espèce en fonction des conditions environnementales de la zone. Néanmoins, son emploi nécessite des conditions d'utilisation particulières comme notamment une indépendance des données les unes par rapport aux autres et un accès à des bases de données renseignant les valeurs des variables à une grande échelle. En d'autres termes, les valeurs de la vitesse du courant ou de la granularité du substrat sur tout le territoire auraient été nécessaires afin que le logiciel puisse modéliser la répartition potentielle des individus en fonction de ces variables. N'ayant pas recours à de telles données, une incompatibilité entre l'emploi du logiciel MaxEnt et les données recueillies et disponibles pour l'étude concernant les *Gomphidae* a été soulevée. En effet, les données d'observations non prises aléatoirement sur le territoire prospecté et le nonaccès à des valeurs de variables environnementales en tout point de la zone ont été les principaux problèmes ne permettant pas l'emploi du logiciel. Toutefois, cet outil informatique reste intéressant à conserver en mémoire si l'étude de la répartition des individus en fonction de paramètres bioclimatiques par exemple était envisagée, ces informations étant disponibles dans des bases de données nationales. Des modifications dans les protocoles d'échantillonnage seraient néanmoins nécessaires afin d'obtenir des données indépendantes.

Ces premières observations ont entraîné une recherche alternative d'outil d'analyse pouvant permettre d'expliquer la répartition des *Gomphidae*. Un certain nombre de tests statistiques présente un fort pouvoir explicatif des relations s'entretenant entre différentes variables. Afin d'être en cohérence avec ces outils d'analyse, l'axe de recherches a évolué en se demandant quelles variables influençaient le plus la répartition des populations. En effet, en connaissant les exigences environnementales particulières de ces dernières, il est envisageable de chercher des endroits répondant à ces critères afin de compenser la destruction de leurs habitats. Au vu de cette interrogation plusieurs tests statistiques ont été évalués afin de déterminer les plus pertinents. L'analyse Hill Smith est apparue comme la plus intéressante à utiliser de par le fait qu'elle puisse traiter simultanément des données quantitatives et qualitatives. La vitesse du courant, la hauteur d'eau, le type de substrat ainsi que le taux de dioxygène ont été les variables choisies pour l'analyse. Le nombre de *Gomphidae* récolté sur le terrain étant très faible lors de la campagne de prélèvements, les tests statistiques ont été faits sur le genre *Caenis*, plus

représenté dans l'échantillonnage. Le test Hill Smith a été effectué sur le logiciel R, fournissant un aperçu global des corrélations s'entretenant entre l'abondance et les diverses variables environnementales testées. Cette analyse a été complétée par d'autres tests tels que Spearman ou Fisher afin d'obtenir une quantification des corrélations entre les variables prises une à une avec l'abondance. Le taux de corrélation et le taux de significativité fournis par le test entre chaque couple de variables ont permis d'identifier une corrélation positive entre l'abondance des *Caenis* et, la hauteur d'eau, la vitesse du courant, et un substrat de type sable grossier. D'après les recherches complémentaires sur l'écologie des larves de *Caenis*, il s'est avéré que le caractère opportuniste et très variable en fonction des espèces de ce genre, a rendu difficile l'affirmation que les tests aient montré les corrélations les plus importantes entre l'abondance et les conditions du milieu. Afin de s'assurer de la justesse de ces tests statistiques, il serait intéressant de les implémenter avec des jeux de données correspondants à des individus ayant une valence écologique moins large. Il a été montré au début de ce projet que les espèces *Ophiogomphus cecilia* et *Gomphus flavipes*, populations au cœur des préoccupations du Contrat pour la Loire et ses Annexes, avaient plus d'exigences écologiques quant à leur milieu. L'application des tests statistiques cités précédemment à de telles populations d'individus devraient permettre d'obtenir des résultats plus fiables quant aux variables explicatives de leur répartition sur le territoire. La combinaison de ces tests statistiques est donc prometteuse pour la suite de l'étude même si cela laisse entendre que de nouvelles campagnes de prélèvements des *Gomphidae* sont nécessaires afin d'obtenir un échantillon suffisamment grand pour pouvoir être traité par des méthodes statistiques.

Même si l'utilisation du logiciel MaxEnt pour l'étude en question est apparue impossible au vu du manque d'informations essentielles à son implémentation, l'analyse Hill Smith et les tests qui lui ont été associés semblent intéressants pour comprendre les corrélations s'entretenant entre les variables et ainsi expliquer les conditions favorables aux individus. Cette étude n'est pas exhaustive et une recherche d'outils d'analyse complémentaires peut être envisagée. En effet, d'autres tests statistiques ou d'autres logiciels informatiques pourraient être testés. L'AFM et les modèles linéaires restent des méthodes statistiques qu'il serait intéressant de tester puisqu'ils pourraient apporter de nouvelles informations et ainsi compléter celles obtenues par l'analyse Hill Smith. De plus, le programme XLSTAT-Ecology, extension de XLSTAT, permet l'analyse de nombreuses méthodes spécifiques à l'écologie. Il serait intéressant de se renseigner sur ce programme afin d'en déceler les pistes d'étude pouvant être intéressantes pour comprendre la répartition des individus. Faute de temps, ces outils d'analyse n'ont pas pu être testés mais restent des axes de recherches intéressants pour la suite. L'implémentation de différents tests est la voie la plus efficace pour déterminer celui qui est le plus en adéquation avec les objectifs recherchés.

CONCLUSION

Dans le cadre du Contrat pour la Loire et ses Annexes, il est important d'étudier, avant les travaux d'enlèvements d'épis prévus sur la Loire aval, la répartition des *Gomphidae* en Loire. *Gomphus flavipes* et *Ophiogomphus cecilia*, espèces d'intérêt patrimonial, sont inféodées tout particulièrement au bassin ligérien. Afin d'améliorer les mesures de conservation de ces espèces vulnérables et ainsi de mettre en place des mesures compensatoires à la perte d'habitats engendrée par les travaux, il est indispensable d'acquérir plus de connaissances quant aux facteurs environnementaux caractérisant le plus leurs zones de prédilection pendant la phase larvaire. Les objectifs de cette étude ont donc été de déterminer quels outils d'analyse de la répartition des *Gomphidae* étaient les plus pertinents. Plusieurs d'entre eux ont été testés dans le but de mettre en relation l'abondance des individus échantillonnés et les valeurs de différentes variables environnementales mesurées aux mêmes points d'échantillonnage (Annexe 7).

L'employabilité du logiciel MaxEnt a tout d'abord été testée puis réfutée de par une incompatibilité entre ses conditions d'utilisation et les données recueillies sur le terrain ou accessibles dans des bases de données. Les tests statistiques testés, tels que l'analyse Hill Smith et des tests de Pearson ou de Fisher, ont quant à eux montré un potentiel explicatif intéressant des variables environnementales influençant la répartition des individus dans leur milieu. Ces outils ont ensuite été mis en commun afin d'évaluer leurs avantages, leurs faiblesses et leur compatibilité avec le cas d'étude présent. Des pistes de recherche via l'utilisation de d'autres tests statistiques ou de nouveaux programmes informatiques ont été évoquées afin d'approfondir ce projet.

Pour conclure, cette étude aboutit à une piste d'analyse des données très prometteuse puisque l'analyse Hill Smith et les tests associés permettent l'acquisition d'une vision claire des relations s'entretenant entre les variables environnementales et l'abondance des individus d'une population. Ces tests statistiques sont relativement simples d'utilisation. Nous pouvons espérer qu'ils apporteront des réponses quant aux exigences écologiques des *Gomphidae*, permettant ainsi de mieux cerner les zones qui peuvent leur être favorables. Pour s'en assurer il faudra par la suite réaliser de nouvelles campagnes d'échantillonnage afin d'obtenir un nombre suffisant d'individus et ainsi acquérir des résultats exploitables par les tests statistiques mis en avant dans le présent document. Il sera nécessaire de comparer les résultats statistiques obtenus avec les observations reportées dans la littérature concernant les conditions écologiques des *Gomphidae* afin de s'assurer de leur cohérence. Dans l'attente de ces nouvelles campagnes, la recherche d'outils d'analyse pertinents peut être poursuivie afin d'étudier le plus large panel de solutions.

BIBLIOGRAPHIE

Baccini Alain (2010) – Statistique Descriptive Multidimensionnelle. Institut de Mathématiques de Toulouse – Université Paul Sabatier. 33 pages.

Bensettiti F. & Gaudillat V., (2004) - Cahiers d'habitats Natura 2000. Connaissance et gestion des habitats et des espèces d'intérêt communautaire. Espèces animales. *La documentation française*, Tome 7, p. 289-292.

Berquier Cyril (2015) - Étude écologique et patrimoniale du peuplement des odonates de Corse appliquée à la conservation des espèces et des zones humides à enjeux. *Rapport de thèse. Université de Corse-Pascal Paoli*. 102 pages

Bishop, J. E., 1973. Limnology of a small Malayan river Sungai Gombak. Dr. W. Junk B.V. Publishers, The Hague, 485 pp.

Brulin Michel (2007) – Du côté des Ephéméroptères : gros plan sur les CAENIS. *INSECTES* n°144

Buffagni A., Cazzola M., López-Rodríguez MJ., Alba-Tercedor J. & Armanini DG. (2009) - Distribution and ecological preferences of European freshwater organisms. Volume 3. Ephemeroptera. *Pensoft Publishers* 257 pages.

Burcher C.L. & Smock L.A. (2002) - Habitat distribution, dietary composition and life history characteristics of odonate nymphs in a blackwater coastal plain stream. *American Midland Naturalist*. 148. P.75-89.

Champely Stéphane (2005) – Introduction à l'analyse multivariée (factorielle) sous R. 57 pages.

Chessel D. & Lascaux J.M (2007) – Entre ACP et AM : l'analyse de Hill et Smith. *ADE-4*. 30 pages.

Chouquet C. (2010) – Modèles linéaires. Laboratoire de Statistique et Probabilités - Université Paul Sabatier – Toulouse. 50 pages

Cokrum Lynda D. (2007) – Habitat characterization of the morphologically similar mayfly larvae *Caenis* and *Tricirithides* (Ephemeroptera). *Hydrobiologia* 179 : 103-109

Dommanget J.-L. (2004) – *Ophiogomphus cecilia* (Geoffroy in Fourcroy, 1785). In : Bensettiti F. & Gaudillat V. (2004). Cahiers d'habitats Natura 2000. Connaissance et gestion des habitats et des espèces d'intérêt communautaire. Tome 7. *Espèces animales. La Documentation française*. Tome 7, 353 pages

Elith, J., Graham, C. H., Anderson, R. P., Dudi'k, M., Ferrier, S., Guisan, A., Hijmans, R. J., Huettmann, F., Leathwick, J. R., Lehmann, A., Li, J., Lohmann, L. G., Loiselle, B. A., Manion, G., Moritz, C., Nakamura, M., Nakazawa, Y., Overton, J. McC., Peterson, A. T., Phillips, S. J., Richardson, K. S., Scachetti-Pereira, R., Schapire, R. E., Sobero'n, J., Williams, S., Wisz, M. S. and Zimmermann, N. E. (2006) - Novel methods improve prediction of species' distributions from occurrence data. *Ecography* 29: 129 - 151.

Elith J., Steven J. Phillips, Hastie T., Dudik M., Chee Y.E. & Yates C. J., (2011) - A statistical explanation of MaxEnt for ecologists. *Diversity and distributions* , 17(1) : 43-57.

- Ghani WMHWA., Rawi CSM., Hamid SA., Al-Shami SA., Ahmad AH. & Hassan ANN. (2016)** - Variation in environmental conditions influences diversity and abundance of Ephemeroptera in forest streams of northern Peninsular Malaysia. *Tropical Ecology* 57(3): 489-501
- Gouraud C., (2015)** – Ecologie larvaire de deux *Gomphidae* de Loire moyenne : *Gomphus flavipes* et *Ophiogomphus cecilia*. Département d'Ingénierie des Milieux Aquatiques et des Corridors Fluviaux - Université François Rabelais & Ecole Polytechnique de Tours, 55 pp
- Grand D., & Boudot J.-P., (2006)** – Les libellules de France, Belgique et Luxembourg. Biotope, Mèze. *Collection Parthénopé*, 480 pages.
- Heidemann H., & Seidenbursch R., (2002)** - Larves et exuvies des libellules de France et d'Allemagne (sauf de Corse). *Société française d'odonatologie*, 416 pages.
- Hervé Maxime (2010)** – Aide mémoire de statistique appliqué à la biologie – Construire son étude et analyser ses résultats à l'aide du logiciel R. pp. 19-29
- Hill, M.O. & Smith, A.J.E. (1976)** - Principal component analysis of taxonomic data with multi-state discrete characters. *Taxon* : 25, 249-255.
- Kato Y. (2012)** - Méthodologie pour le calcul de probabilité d'implantation de l'habitat dispersé par entropie maximale. *Rapport Caribsat, programme INTERREG Caraïbe IV*, 12 p.
- Lett J.-M., Cloupeau R., Pratz J.-L. & Male-Malherbe E. (2001)** - Liste commentée des Odonates de la région Centre (Départements du Cher, de l'Eure-et-Loir, de l'Indre, de l'Indre-et-Loire, du Loir-et-Cher et du Loiret). *Martinia*. Tome 17, fascicule 4. Pages 123-168.
- Luque S., Redon M., Isenmann M., Sanz T., (2012)** – Prédiction de la distribution d'alliances de végétation des milieux ouverts d'altitude à l'aide de l'approche dite du maximum d'entropie. – Programme CarHAB – volet modélisation, 77p
- Merlet F. & Houard X. (2012)** - Synthèse bibliographique sur les traits de vie du Gomphe serpent (*Ophiogomphus cecilia* (Geoffroy in Fourcroy, 1785)) relatifs à ses déplacements et à ses besoins de continuités écologiques. *Office pour les insectes et leur environnement & Service du patrimoine naturel du Muséum national d'Histoire naturelle*, Paris. 8 p.
- Merow C., Smith M. J, and Silander J. A (2013)** - A practical guide to MaxEnt for modeling species' distributions: what it does, and why inputs and settings matter. *Ecography* 36: 1058–1069.
- Motulsky H.J. (2007)** – Biostatistique, Une approche intuitive – *Sciences & Méthodes*. Edition boeck. pp. 325-330 et 367-379
- Office Pour les Insectes et leur Environnement.** Plan National d'Actions en faveur des Odonates. [en ligne ; consulté le 10.10.17] Disponible sur : <https://odonates.pnaopie.fr>
- Phillips S.J., (2017)** - A Brief Tutorial on Maxent. [online ; consulted on 22.10.17] Available from URL: http://biodiversityinformatics.amnh.org/open_source/maxent/
- Phillips S.J, Dudik M., Schapire R.E, (2004)** – A maximum entropy approach to species distribution modelling. Modelling of species distributions with Maxent : new extensions and a comprehensive evaluation. *Proceedings of the Twenty-First International Conference on Machine Learning*, pages 655-662.

Phillips S.J., Robert P. Anderson & Robert E. Schapire, (2006) – Maximum entropy modelling of species geographic distributions – *Ecological Modelling*, Vol 190/3-4 pp 231-259

PNA Odonates (2014) – *Gomphus flavipes*. [en ligne ; consulté le 10.10.17]. Disponible sur : <https://odonates.pnaopie.fr/especes/Gomphidae/gomphus-flavipes/>

PNA Odonates (2014) – *Ophiogomphus cecilia*. [en ligne ; consulté le 10.10.17] Disponible sur : <https://odonates.pnaopie.fr/especes/Gomphidae/ophiogomphus-cecilia/>

Radosavljevic A. & Anderson R. P. (2014) - Making better Maxent models of species distributions: complexity, overfitting and evaluation. *J. Biogeogr.*, 41: 629–643.

Rodrigues S. (2016) – Restauration du lit et Trajectoires Ecologiques, Morphologiques et d'Usages en Basse Loire (R-TEMUS). UMR CNRS CITERES 7324 – Ecole polytechnique de l'université François-Rabelais de Tours, 58pages.

Suhling F. & Müller O. (1996) - Die Flußjungfern Europas. *Gomphidae*. Die Neue Brehm-Bücherei, Westarp-Wissenschaften. Vol.628, 237 pages.

Thomas B., Lutz S. & Rideau C. (2015) - Modélisation de la répartition des campagnols aquatiques du genre *Arvicola* en Normandie. *Groupe Mammalogique Normand, Agence de l'Eau Seine-Normandie*, 62 p. + annexes.

Venditti V., (1998) - Aspects du principe de maximum d'entropie en modélisation statistique. *Université Joseph Fourier*.

TABLE DES FIGURES

Figure 1 : Répartition des espèces de <i>gomphus flavipes</i> et <i>ophiogomphus cecilia</i> en france (inpn, opie)	4
Figure 2 : Schéma du cycle de vie des odonates (Sorba, OEC in Berquier C., 2015)	5
Figure 3 : Photo d'un épis de navigation sur la loire, zone favorable au développement larvaire des individus (GIP ESTUAIRE)	6
Figure 4 : Localisation des stations d'échantillonnage sur la zone d'étude	8
Figure 5 : Vue aerienne de la station satellite de varades et positionnement des transects (donnees : Nina RICHARD ; realisation : Kevin SILLIAU)	9
Figure 6 : Algorithme de fonctionnement du logiciel MaxEnt	11
Figure 7 : Schémas présentant les donnees théoriques requises d'après les conditions d'utilisation du logiciel (A) et les données réellement récoltées par l'étude (B).....	14
Figure 8 : Projection des variables sur le plan factoriel	18

TABLE DES TABLEAUX

Tableau 1 : Récapitulatif des tests statistiques envisages.....	15
Tableau 2 : Résultats des tests de Spearman et de Fisher	19
Tableau 3 : Résultats des tests de Kendall entre l'abondance et le type de substrat.....	19

ANNEXE 1 : EXEMPLE D'UNE FICHE DE DONNEES
PRESENTANT LES VARIABLES ENVIRONNEMENTALES
MESUREES ET LE COMPTAGE DES INVERTEBRES

Station de St Mathurin sur Loire : transect
transversal A (SM1)

Date	13/09/2016
Opérateurs	YG/NR
conditions météo	SOLEIL

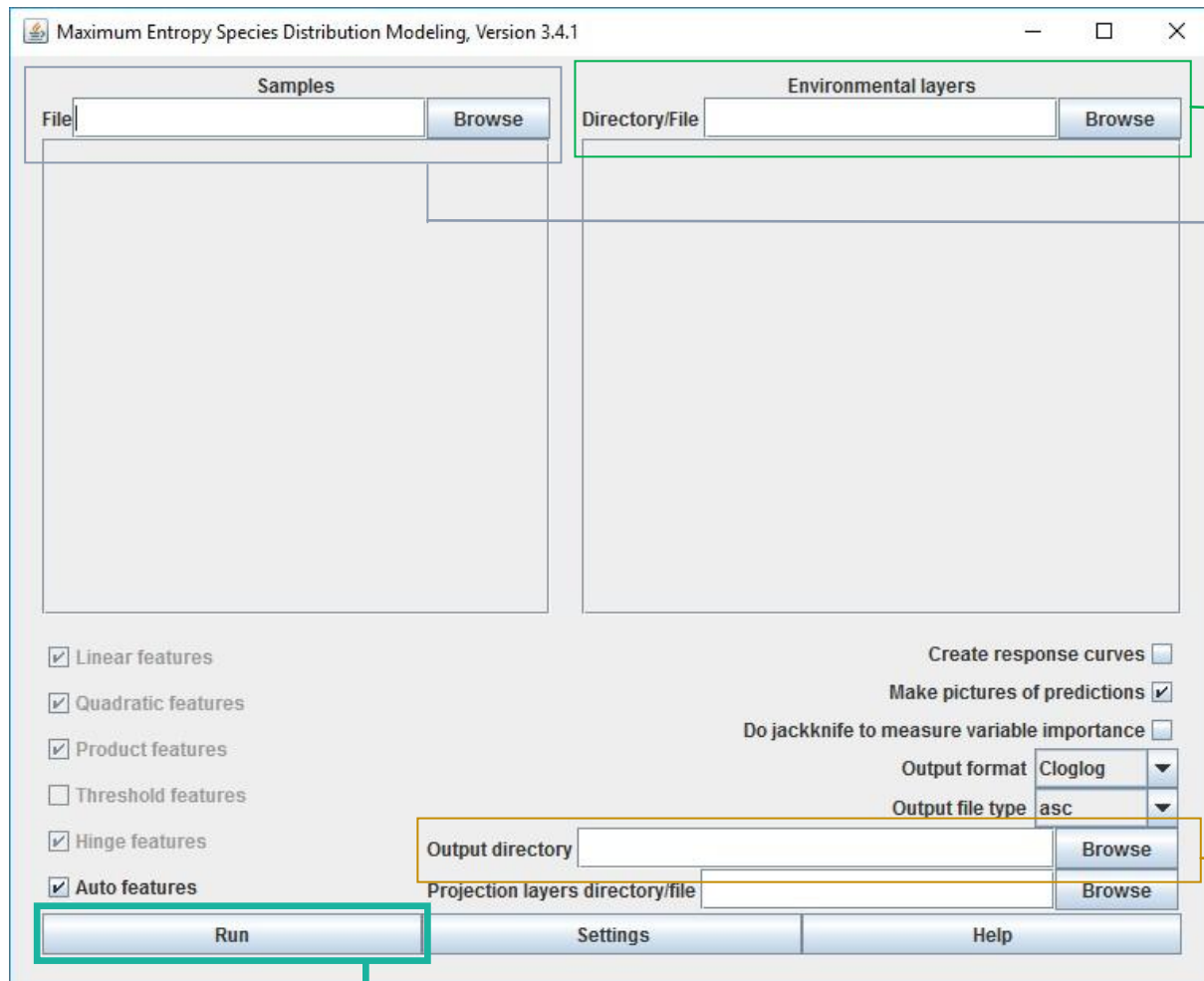
heure début	15H14
Temps - air	32°C 14H11 A OMBRE

Numéro Prélèvement	SM1-P1-0916	SM1-P2-0916	SM1-P3-0916	SM1-P4-0916	SM1-P5-0916
Coordonnées GPS	448890	448890	448889	448888	448887
	6705903	6705898	6705893	6705889	6705884
Distance à la berge (m)	57	51	45	39	33
Heure prélèvement	15H14				
Hauteur d'eau (m)	35	50	65	80	80
Vitesse de courant 0,8h (m/s)	0,2	0,3	0,4	0,4	0,3
Vitesse de courant au plus près du fond (m/s)					
O2 au plus près du fond (mg/l)	10,12	9,19	9,15	9	8,93
O2 au plus près du fond (%age saturation)	122	110	109	107	106
O2 au plus 0,8h (mg/l)					
O2 au plus 0,8h (%age saturation)					
T°C eau au plus près du fond	24,2	23,7	23,5	23,5	23,4
T°C eau en surface					
Vase					
Limons < 2mm					
Sables < 2mm					
Sables grossiers 2mm <x < 2cm	*	*	*	*	*
Pierres et galets > 2cm					
Blocs ou Roche mère					
Stabilité du substrat (stable ou meuble)	MEUBLE	MEUBLE	MEUBLE	MEUBLE	MEUBLE
Végétaux aquatiques (espèce, recouvrement % surface cercle de 50cm diam autour du point)		ALGUES VERTES	ALGUES VERTES	ALGUES VERTES	ALGUES VERTES
Ombrage porté de 0 à 10%					
Ombrage porté de 11 à 40%					
Ombrage porté de 41 à 80%					
Ombrage porté de 81 à 100%					
Pas de colmatage	*	*	*	*	*
Colmatage léger (gros éléments enfouis à 25% de leur hauteur)					

Colmatage important (25%< colmatage< 75%)					
Colmatage complet (gros éléments enfouis à plus de 75 % de leur hauteur)					
Présence de MO grossière > 1cm (oui/non)	OUI	OUI	OUI	OUI	OUI
Nature de MO grossière > 1cm (feuilles, racines, branches...)	ALGUES+MO	ALGUES+MO	ALGUES+MO	ALGUES+MO	ALGUES+MO
Préleveur (haveneau ou cone berthois)					

Ephéméroptères					
Caenis	7	12	36	7	2
Baetidae/Ameletidae					
<i>Ephemera sp</i>					
Ephéméroptères n.d.	9	4	1	3	5
Anisoptères					
<i>Onychogomphus forcipatus</i>					
<i>Ophiogomphus cecilia</i>			1		

ANNEXE 2 : PRESENTATION DE L'INTERFACE DE TRAVAIL DU LOGICIEL MAXENT



Données
environnementales
sous format ascii
(.asc)

Données de présence
sous format .csv

Dossier de destination

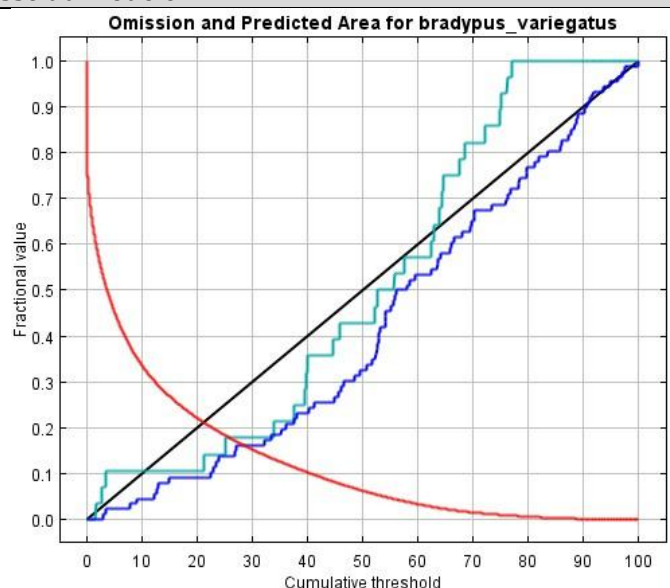
Mise en marche

ANNEXE 3 : RESULTATS POUVANT ETRE OBTENUS PAR LE LOGICIEL MAXENT

Evaluer la robustesse du modèle

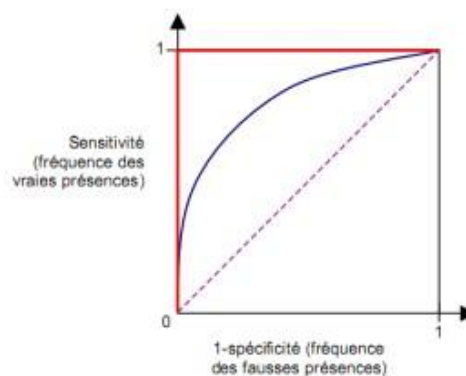
Courbe d'omission et de prédilection :

Le taux d'omission évalue la probabilité d'une erreur de prédiction du modèle qui indiquerait l'absence d'un individu alors que dans les faits il est présent. Plus le taux d'omission est faible et plus le modèle est performant (Luque *et al.*, 2012). L'omission théorique est représentée par la droite $a=0,5$.



Courbe ROC (Receiver Operating Characteristic) :

Permet d'évaluer la capacité de discrimination du modèle. Elle a pour but de déterminer si le modèle est aléatoire ou bien si les prédictions statistiques sont proches de la réalité (Luque *et al.*, 2012). Il va falloir s'intéresser à la valeur de l'aire sous la courbe ROC appelée AUC (Area Under the Curve) (Radosavljevic & Anderson, 2014). Cette valeur est comprise entre 0 et 1. Une valeur de 0,5 correspond à un modèle aléatoire (Kato, 2012), fortement influencé par le hasard. Une valeur se rapprochant de 1 correspond davantage à un modèle qui prédit la réalité (Thomas *et al.*, 2015). Ainsi, une valeur de l'AUC élevé indique que les prédictions du modèle sont statistiquement fiables.

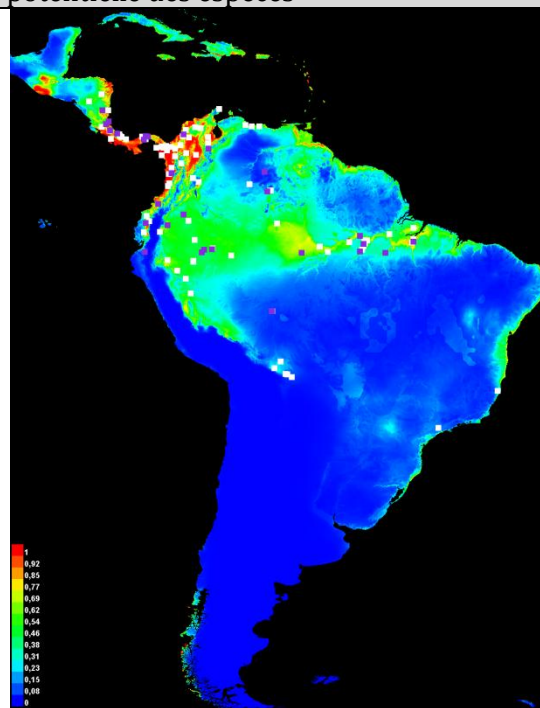


- Courbe ROC représentant le pire des cas avec un fort effet du hasard (AUC de 0,5 associée à une capacité de discrimination très faible)
- Courbe ROC représentant le meilleur des cas (AUC de 1 associée à une capacité de discrimination très forte, le modèle donne des prédictions exactes)
- Courbe ROC représentant la capacité de discrimination d'un modèle étudié (AUC compris entre 1 et 0,5)

Modélisation de la répartition spatiale potentielle des espèces

Carte de répartition :

Le logiciel fournit une carte présentant la répartition potentielle des espèces en fonction des conditions environnementales (Radosavljevic & Anderson, 2014). Les différentes couleurs utilisées indiquent si la probabilité de présence est importante ou non. Dans l'exemple ci-contre, les couleurs chaudes correspondent à des aires avec de bonnes prédictions.



Détermination des variables influençant le plus le modèle

Tableau pourcentage de contribution/importance des permutations :

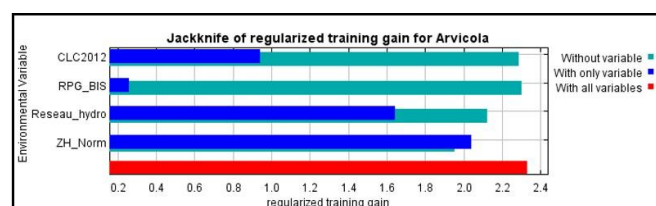
L'importance d'une variable pour expliquer la répartition d'une espèce est définie après une permutation des valeurs de la variable entre les points d'observation. Il y a ensuite une comparaison de la capacité de discrimination avec chaque permutation. Plus cette capacité est affectée par les permutations, plus la variable est explicative de la répartition de l'espèce (Luque *et al.*, 2012). Le résultat est exprimé en pourcentage. Plus ce pourcentage est important, plus la variable associée est explicative.

Le pourcentage de contribution des variables indique uniquement les poids respectifs des variables qui interviennent dans le modèle. Il ne suffit pas pour interpréter l'importance relative de la variable pour l'espèce (Luque *et al.*, 2012). Il dépend uniquement du cheminement suivi par l'algorithme pour aboutir au meilleur résultat.

Variable	Percent contribution	Permutation importance
pre6190_17	33.7	4.2
pre6190_110	29.4	7.1
h_dem	10.1	10.7
tmn6190_ann	8.5	30.6
tmx6190_ann	5.5	19.9
frs6190_ann	3.5	14.9
vap6190_ann	2.5	1.4
pre6190_11	1.9	2.6
dtr6190_ann	1.5	6.2
pre6190_14	1.5	1.2
pre6190_ann	0.7	0
ttmp6190_ann	0.7	0.1
ecoreg	0.4	0.8
cld6190_ann	0.1	0.1

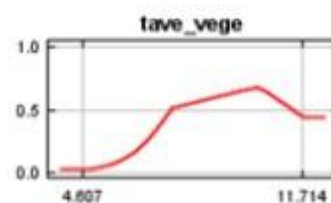
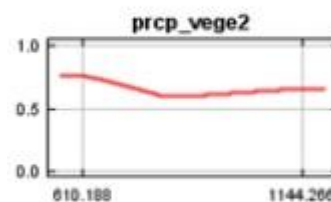
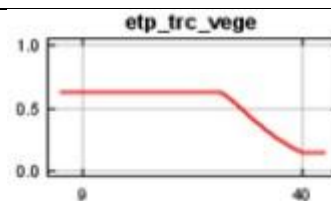
Jackknife :

Le poids de chaque variable est calculé si l'option « jackknife » est choisie. Il est représenté sous forme d'un graphique à plusieurs barres horizontales correspondant aux différentes variables. Plus la valeur du gain est important, plus la variable aura un impact significatif sur la répartition.



Courbes de relation entre les variables environnementales et la probabilité d'occurrence :

Elles sont un point de contrôle de la justesse des prédictions ; elles vérifient si les prédictions fortes correspondent bien à des valeurs des variables favorables à l'espèce étudiée.



ANNEXE 4 : TABLEAU BRUT DES DONNEES A CHARGER DANS
R (AVEC LA VARIABLE QUALITATIVE SUBSTRAT)

identification	abondance	vcourant	dioxygene	hauteur_eau	substrat
01P1	0	0	5,7	0	s
01P2	2	0	7,37	25	s
01P3	6	0	7,9	40	s
01P4	20	0	8,06	100	s
01P5	12	0	8,08	65	s
01P6	5	0,1	8,05	120	s
01P7	43	0,05	8,02	250	s
01P8	33	0	8,04	210	s
01P9	60	0,2	8,08	210	sgr
01P10	3	0,5	8,24	280	sgr
03P5	10	0,1	7,94	205	s
03P6	0	0,2	7,83	90	s
03P7	13	0,3	7,88	140	s
03P8	32	0,1	7,84	250	s
03P9	22	0,3	7,93	330	s
03P10	7	0,4	7,77	350	s
04P11	0	0	8,38	35	s
04P12	0	0	9,02	35	s
04P13	0	0	8,75	35	s
04P14	0	0	8,8	35	s
04P15	0	0	8,8	35	sgr
04P16	0	0,2	8,17	35	s
04P17	0	0,2	8,16	35	s
04P18	0	0,2	8,41	35	sgr
04P19	0	0	9,2	35	s
04P20	0	0	8,13	35	s
M1P1	14	0,7	8,38	300	s
M1P2	1	0,7	8,43	290	s
M1P3	0	0,6	8,41	200	s
M1P4	9	0,4	8,41	180	s
M1P5	22	0,15	8,38	185	sgr
M1P7	0	0	9,28	5	sgr
M1P8	2	0	11,18	60	s
M1P9	3	0	11,76	50	sgr
M1P10	0	0	14,8	15	s
M2P1	0	0	9,66	30	s
M2P2	0	0	8,79	65	s
M2P3	0	0	8,55	75	s
M2P4	6	0	8,36	95	s
M2P5	3	0	8,23	145	s

M2P6	0	0	7,91	3,55	s
M2P7	0	0,1	7,92	3,45	s
M2P8	58	0,2	7,95	4	sgr
M2P9	10	0,3	7,99	4	sgr
M2P10	31	0,4	7,94	4	sgr
M3P1	0	0	9,8	50	sgr
M3P2	0	0	9,63	15	sgr
M3P3	0	0	10,02	10	s
M3P4	0	0	10,86	20	s
M3P5	0	0	9,62	40	sgr
M3P6	3	0	14,92	25	s
M3P7	0	0	8,75	20	s
M3P8	1	0	11	25	s
M3P9	1	0	8,85	25	s
M3P10	0	0	8,47	30	s
M3P11	1	0	12,25	30	sgr
M3P12	0	0	10,59	38	sgr
M3P13	1	0	8,17	30	sgr
M3P14	0	0	8,7	15	sgr
M3P15	10	0	8,1	40	sgr
M3P16	0	0	8,25	20	sgr
M3P17	3	0	8,84	40	s
M3P18	9	0	8,97	20	s
M3P19	0	0	4,36	45	s
M3P20	0	0	8,35	15	s
SM1P1	7	0,2	10,12	35	sgr
SM1P2	12	0,3	9,19	50	sgr
SM1P3	36	0,4	9,15	65	sgr
SM1P4	7	0,4	9	80	sgr
SM1P5	2	0,3	8,93	80	sgr
SM1P6	0	0,1	8,92	90	sgr
SM1P7	10	0,05	90,3	110	s
SM1P8	1	0	11,42	90	s
SM1P9	0	0	11,7	55	s
SM1P10	3	0	9,38	30	s
SM2P1	24	0,1	10,28	80	sgr
SM2P2	1	0,1	8,89	80	sgr
SM2P3	0	0,2	8,49	100	sgr
SM2P4	1	0,2	8,13	85	sgr
SM2P5	1	0,1	7,94	60	s
SM2P6	0	0,1	7,95	60	s
SM2P7	0	0,05	8,13	60	s
SM2P8	1	0,05	7,92	80	s
SM2P9	9	0,05	8,08	100	s
SM2P10	24	0,05	8,02	135	s

SM3P1	1	0,3	9,11	70	s
SM3P2	8	0,3	9,07	70	sgr
SM3P3	1	0,2	9,01	45	sgr
SM3P4	1	0,3	8,96	60	s
SM3P5	48	0,5	9,01	50	sgr
SM3P6	14	0,4	9,08	25	sgr
SM3P7	21	0,5	9,06	30	sgr
SM3P8	5	0,5	9,01	30	sgr
SM3P9	4	0	8,86	70	s
SM3P10	0	0	9,57	50	s
SM3P11	35	0,4	8,55	50	sgr
SM3P12	2	0,1	8,87	60	sgr
SM3P13	13	0,1	13,4	35	sgr
SM3P14	9	0	11,42	50	sgr
SM3P15	24	0	12,27	45	sgr
SM3P16	13	0	12,47	45	sgr
SM3P17	1	0	10,31	70	sgr
SM3P18	0	0,1	10,54	55	sgr
SM3P19	33	0	11,8	40	sgr
SM3P20	61	0	10,63	45	sgr

ANNEXE 5 : TABLEAU BRUT DES DONNEES A CHARGER DANS
R (AVEC LES SOUS-VARIABLES QUANTITATIVES SUBS_SGR ET
SUBST_S)

abondance	vcourant	dioxygene	hauteur_eau	subs_sgr	subst_s
0	0	5,7	0	0	1
2	0	7,37	25	0	1
6	0	7,9	40	0	1
20	0	8,06	100	0	1
12	0	8,08	65	0	1
5	0,1	8,05	120	0	1
43	0,05	8,02	250	0	1
33	0	8,04	210	0	1
60	0,2	8,08	210	1	0
3	0,5	8,24	280	1	0
10	0,1	7,94	205	0	1
0	0,2	7,83	90	0	1
13	0,3	7,88	140	0	1
32	0,1	7,84	250	0	1
22	0,3	7,93	330	0	1
7	0,4	7,77	350	0	1
0	0	8,38	35	0	1
0	0	9,02	35	0	1
0	0	8,75	35	0	1
0	0	8,8	35	0	1
0	0	8,8	35	1	0
0	0,2	8,17	35	0	1
0	0,2	8,16	35	0	1
0	0,2	8,41	35	1	0
0	0	9,2	35	0	1
0	0	8,13	35	0	1
14	0,7	8,38	300	0	1
1	0,7	8,43	290	0	1
0	0,6	8,41	200	0	1
9	0,4	8,41	180	0	1
22	0,15	8,38	185	1	0
0	0	9,28	5	1	0
2	0	11,18	60	0	1
3	0	11,76	50	1	0
0	0	14,8	15	0	1
0	0	9,66	30	0	1
0	0	8,79	65	0	1
0	0	8,55	75	0	1
6	0	8,36	95	0	1

3	0	8,23	145	0	1
0	0	7,91	3,55	0	1
0	0,1	7,92	3,45	0	1
58	0,2	7,95	4	1	0
10	0,3	7,99	4	1	0
31	0,4	7,94	4	1	0
0	0	9,8	50	1	0
0	0	9,63	15	1	0
0	0	10,02	10	0	1
0	0	10,86	20	0	1
0	0	9,62	40	1	0
3	0	14,92	25	0	1
0	0	8,75	20	0	1
1	0	11	25	0	1
1	0	8,85	25	0	1
0	0	8,47	30	0	1
1	0	12,25	30	1	0
0	0	10,59	38	1	0
1	0	8,17	30	1	0
0	0	8,7	15	1	0
10	0	8,1	40	1	0
0	0	8,25	20	1	0
3	0	8,84	40	0	1
9	0	8,97	20	0	1
0	0	4,36	45	0	1
0	0	8,35	15	0	1
7	0,2	10,12	35	1	0
12	0,3	9,19	50	1	0
36	0,4	9,15	65	1	0
7	0,4	9	80	1	0
2	0,3	8,93	80	1	0
0	0,1	8,92	90	1	0
10	0,05	90,3	110	0	1
1	0	11,42	90	0	1
0	0	11,7	55	0	1
3	0	9,38	30	0	1
24	0,1	10,28	80	1	0
1	0,1	8,89	80	1	0
0	0,2	8,49	100	1	0
1	0,2	8,13	85	1	0
1	0,1	7,94	60	0	1
0	0,1	7,95	60	0	1
0	0,05	8,13	60	0	1
1	0,05	7,92	80	0	1
9	0,05	8,08	100	0	1

24	0,05	8,02	135	0	1
1	0,3	9,11	70	0	1
8	0,3	9,07	70	1	0
1	0,2	9,01	45	1	0
1	0,3	8,96	60	0	1
48	0,5	9,01	50	1	0
14	0,4	9,08	25	1	0
21	0,5	9,06	30	1	0
5	0,5	9,01	30	1	0
4	0	8,86	70	0	1
0	0	9,57	50	0	1
35	0,4	8,55	50	1	0
2	0,1	8,87	60	1	0
13	0,1	13,4	35	1	0
9	0	11,42	50	1	0
24	0	12,27	45	1	0
13	0	12,47	45	1	0
1	0	10,31	70	1	0
0	0,1	10,54	55	1	0
33	0	11,8	40	1	0
61	0	10,63	45	1	0

```

caenis <- read.csv2("caenis.csv",header=TRUE,dec=",")
head(caenis)
caenis$substrat<-as.factor(caenis$substrat)
caenis$hauteur_eau<-as.numeric(caenis$hauteur_eau)
caenis$dioxygene<-as.numeric(caenis$dioxygene)
caenis$vcourant<-as.numeric(caenis$vcourant)
caenis$abondance<-as.numeric(caenis$abondance)
library(ade4)
library(adegraphics)
caenis2<-caenis[,-c(1)]#permet de retirer la colonne identification
head(caenis2)
hill <- dudi.hillsmith(caenis2,scannf = FALSE)
summary(hill)
a<-hill$co ### récupérer coordonnées des variables
b<-hill$li ### récupérer coordonnées des individus
scatter(hill) ### graphique variables + tous les individus
s.label(hill$co)### graphique variables

#### mesure corrélation entre 1 variable qualitative et 1 variable quantitative (Rapport de
corrélation)

library(BioStatR)
eta2(caenis$abondance,caenis$substrat)####calcul rapport corrélation
fisher.test(caenis$abondance,caenis$substrat)##test de fisher
tableau<-table(caenis$abondance,caenis$substrat)
chisq.test(tableau)##test du chi2
caenis2

#### mesure corrélation entre 1 variable quantitative et 1 variable quantitative
cor.test(caenis$abondance,caenis$hauteur_eau,method="kendall")
cor.test(caenis$abondance,caenis$hauteur_eau,method="pearson")
cor.test(caenis$abondance,caenis$hauteur_eau,method="spearman")

```

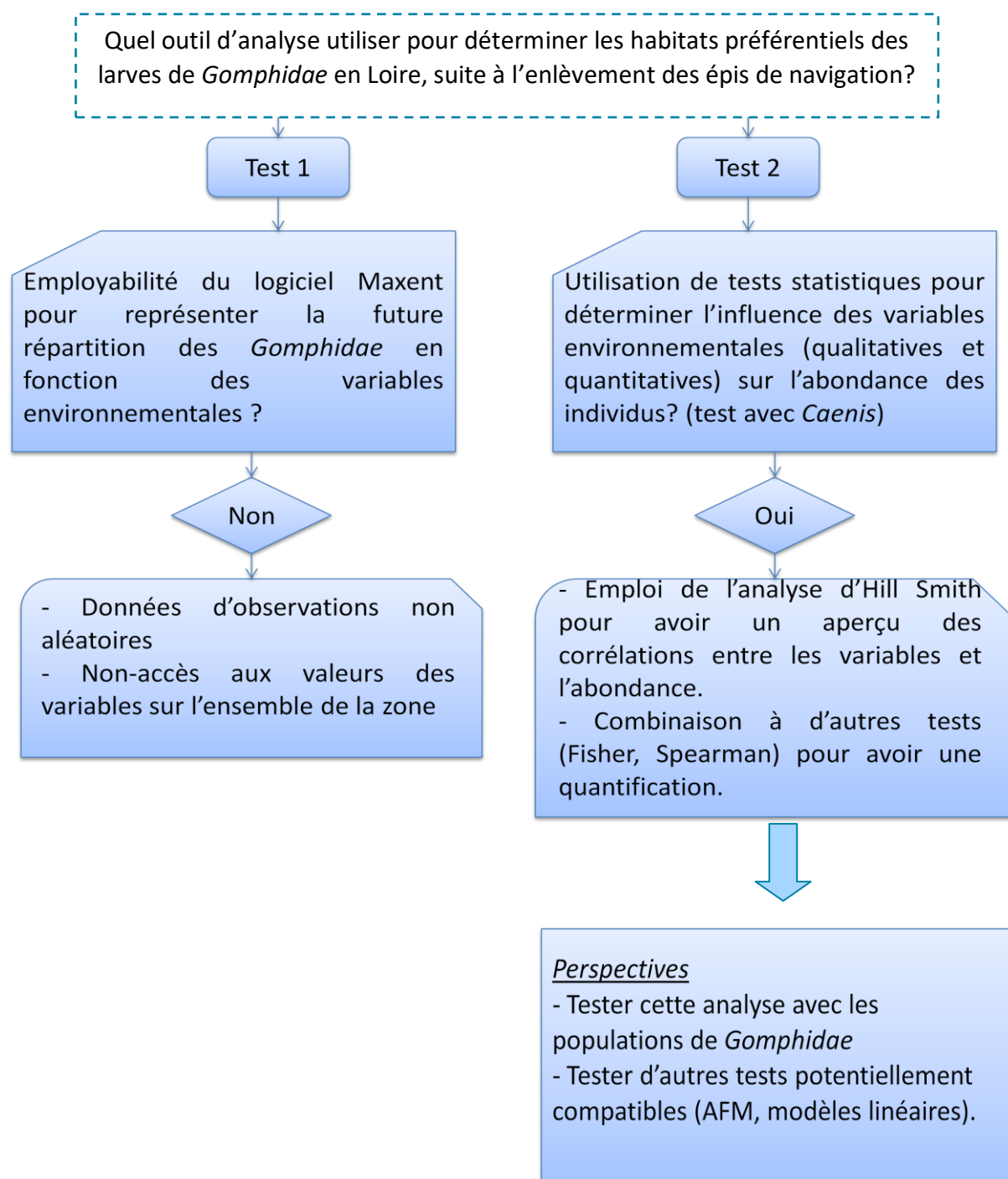



TABLE DES MATIERES

Avertissement.....	
Remerciements.....	
Sommaire.....	1
Lexique.....	2
Sigles	2
Introduction	3
1. Ecologie des <i>Gomphidae</i> et menaces	4
1.1 Description des espèces.....	4
1.2 Habitats favorables.....	5
1.3 Menaces et statuts de protection	6
1.4 Traits écologiques permettant une modélisation de leur répartition	7
2. Utilisation du logiciel Maxent	10
2.1 Principe général de la méthode	10
2.2 Fonctionnement de l'outil.....	10
2.3 Conditions d'utilisation	12
2.4 Employabilité du logiciel MaxEnt pour l'analyse de la répartition des <i>Gomphidae</i>	13
3. Tests statistiques multivariés.....	15
3.1 Choix du test.....	15
3.2 Le test Hill Smith.....	16
3.3 Application du test Hill Smith aux données récoltées sur le genre <i>Caenis</i>	17
3.4 Résultats obtenus	18
3.5 Mise en relation des résultats obtenus avec l'écologie du genre <i>Caenis</i> , et validité du modèle.....	20
3.6 Bilan du modèle utilisé	21
4. Quel outil est le plus approprié pour expliquer la répartition des <i>Gomphidae</i> en Loire ?	22
Conclusion.....	24
Bibliographie	25
Table des figures.....	28
Table des tableaux	28
Annexe 1 : Exemple d'une fiche de données présentant les variables environnementales mesurées et le comptage des invertébrés	29
Annexe 2 : Présentation de l'interface de travail du logiciel MaxEnt.....	31
Annexe 3 : Résultats pouvant être obtenus par le logiciel MaxEnt	32
Annexe 4 : Tableau brut des données à charger dans R (avec la variable qualitative substrat) ...	35
Annexe 5 : Tableau brut des données à charger dans R (avec les sous-variables quantitatives subs_sgr et subst_s).....	38

Annexe 6 : Script langage R.....	41
Annexe 7 : Démarche générale de l'étude	42
Table des matières.....	43

Equipe DATE

*Dynamiques et Actions Territoriales et
Environnementales*

35 allée Ferdinand de Lesseps

BP 30553

37205 TOURS cedex 3

Directeur de recherche :

BOISNEAU Catherine

LOISEAU Marie

Projet de Fin d'Etudes

DA5 : 2017-2018

PFE Semestre 9 &10

**Quel outil d'analyse utiliser pour expliquer la répartition des
Gomphidae en Loire ?**

Résumé : Dans le cadre du Contrat pour la Loire et ses Annexes 2015-2020, des travaux de restauration dans le lit mineur de la Loire sont prévus afin notamment de rétablir une continuité sédimentaire entravée par les grands travaux d'aménagement du 19 et 20^{ème} siècle. Deux espèces de *Gomphidae*, *Gomphus flavipes* et *Ophiogomphus cecilia*, particulièrement inféodées au bassin ligérien, présentent un intérêt patrimonial. L'enlèvement des épis de navigation aura des conséquences sur la phase larvaire de ces espèces qui affectionnent particulièrement les zones à courant d'eau ralenti. Une campagne d'échantillonnage a été réalisée sur cinq stations entre Oudon et Saint-Mathurin-sur-Loire en septembre 2016. Dans le but d'évaluer plus précisément les impacts des travaux et ainsi mettre en place des mesures de conservation adaptées, il est important de connaître l'outil d'analyse le plus pertinent pour expliquer la répartition des individus. Le logiciel MaxEnt, outil de modélisation de la distribution des individus d'une espèce animale ou végétale donnée, ainsi que des tests statistiques multivariés ont été implémentés. Alors que des incompatibilités ont été décelées entre les conditions d'utilisation du logiciel et les données de terrain récoltées, l'analyse Hill Smith couplée à des tests de Spearman et de Fisher est apparue comme un outil pertinent. Cette combinaison a permis de mettre en avant une corrélation positive entre l'abondance des *Caenis* et, la hauteur d'eau, la vitesse du courant et un substrat de type grossier. Ces tests sont prometteurs et pourront être appliqués sur les populations de *Gomphidae* afin de déterminer les variables environnementales régissant le plus leur répartition.

Mots Clés : *Gomphidae*, Loire, MaxEnt, répartition, tests statistiques, Hill Smith.