

## **Première exploration des données issus du programme de science participative Lichens GO!**



Benjamin Canut, Université Toulouse III, ENSAT Toulouse.

Master 2 Biologie Végétale, EVE (Ecologie Végétale et Environnement)

Rapport de stage de Master 2, 1 mars 2021 au 31 août 2021.

Encadré par Laure Turcati <sup>1</sup>, Romain Lorrillière <sup>2</sup>, Sébastien Turpin <sup>3</sup>, Simon Benateau <sup>3</sup>, Yannick Agnan <sup>4</sup>.

Tutrice de stage : Nathalie Séjalon-Delmas <sup>5</sup> ; Correctrice : Patricia Jargeat <sup>6</sup>.

<sup>1</sup> Ingénieure de recherche en sciences participatives, Observatoire des Sciences de l'Univers, Sorbonne Université.

<sup>2</sup> Post-doc écologue et analyse quantitative, CESCO, Museum National d'Histoire Naturelle.

<sup>3</sup> Chef de projet sur Vigie-nature école, CESCO, Museum National d'Histoire Naturelle.

<sup>4</sup> Chargé de cours, Earth and Life Institute, Université catholique de Louvain.

<sup>5</sup> Maître de conférences, laboratoire EDB, Université Toulouse III Paul Sabatier.

<sup>6</sup> Maître de conférences, LRSV, Université Toulouse III Paul Sabatier.

## **Remerciement**

Je souhaite remercier en premier lieu les différents encadrants de mon stage qui m'ont apporté un grand soutien et de nouvelles connaissances. Ainsi, je remercie chaleureusement Romain Lorrillière, Laure Turcati, Simon Benateau, Sébastien Turpin et Yannick Agnan pour leur aide et expertise. Je remercie également tous les membres du laboratoire CESCO pour leur accueil et leurs conseils. Merci à Gilles Platner pour son accompagnement sur le terrain et son expertise du terrain et du sujet. Je remercie tous les participants du programme Lichens GO! pour leur travail et leur implication dans le projet.

## **Table des matières**

<b>Résumé</b>	<b>3</b>
Abstract	4
<b>Introduction</b>	<b>5</b>
<b>Matériel et Méthodes</b>	<b>9</b>
Données de lichens	9
Protocole Lichens GO! de 2017	10
Nouveau protocole Lichens GO!	11
Donnée de pollution atmosphérique	12
Sélection des critères de pollution	12
Sélection des sites d'échantillonnages de lichens 2021	13
Nettoyage des données	15
Traits d'espèces	15
Analyses de données	16
Modèles statistiques	16
Etude des traits d'espèces	17
Diversité bêta	17
<b>Résultats</b>	<b>18</b>
Analyses des 4 indices : LDV, Richesse, Shannon, Somme des fréquences	18
Etude du point de rupture de la pente	20
Relation entre les traits de vie et la pollution	22
Diversité bêta	24
<b>Discussion</b>	<b>25</b>
<b>Conclusion</b>	<b>28</b>
<b>Bibliographie</b>	<b>29</b>
<b>Annexe</b>	<b>34</b>

## Résumé

La pollution atmosphérique en milieu urbain est une des préoccupations majeures pour la santé et l'environnement. Les projets de sciences participatives ont pour rôle de faire avancer les connaissances et sensibiliser les citoyens aux problèmes sanitaires, écologiques et climatiques. Lichens GO! est un projet de science participative pour quantifier la pollution de l'air via les lichens. Le but de ce stage est de valider un indicateur de la qualité de l'air adéquat au protocole du projet. L'étude se base sur des relevés lichens fait selon le protocole à proximité de stations de relevé de la pollution de l'air d'Airparif. Les tests se basent sur des modèles non linéaires entre : 4 indicateurs de la qualité de l'air (LDV, indice de Shannon, richesse spécifique et somme des fréquences) et la médiane de NO<sub>2</sub> en µg/m<sup>3</sup> ; entre 6 traits d'espèces des lichens (ensoleillement, humidité, pH, eutrophisation, altitude et poléotolérance) et la pollution. Une étude de la diversité bêta par l'indice de Sorensen a également été réalisée pour analyser la composition des communautés entre les sites de faibles et de forts taux de NO<sub>2</sub>. Les résultats montrent que les 4 indicateurs de la qualité de l'air ont une relation négative avec le taux de pollution jusqu'à un certain seuil (31 µg/m<sup>3</sup> NO<sub>2</sub>), après le seuil un plateau est atteint. Les traits d'espèces ont mis en évidence un gradient d'urbanité où l'on retrouve des espèces fortement adaptées à des milieux fortement pollués. Enfin, la composition des communautés tend à être plus similaire avec l'augmentation du taux de NO<sub>2</sub>. Les indicateurs de pollutions sélectionnées peuvent être utilisés pour la quantification de la pollution. Cependant, ils ne permettent de la quantifier que sur une échelle de pollution assez fine (0 à ~40µg/m<sup>3</sup>). Il est donc nécessaire d'y associer les traits d'espèces et la composition des communautés pour une meilleure quantification de la pollution. A partir de ces résultats il est possible de proposer un indicateur de pollution aux participants et de faire évoluer le projet.

## Abstract

The role of citizen-science projects is to advance knowledge and raise awareness of health, environmental and climate issues. Urban air pollution is one of the major concerns for health and the environment. Lichens GO! is a participatory science project to quantify air pollution via lichens. The goal of this internship is to validate an air quality indicator suitable for the project protocol. The study is based on lichen surveys done according to the project protocol in the vicinity of Airparif air pollution monitoring stations. The tests are based on non-linear models between: 4 air quality indicators (LDV, Shannon index, species richness and sum of frequencies) and the median of NO<sub>2</sub> in  $\mu\text{g}/\text{m}^3$ ; between 6 lichen species traits (solar, aridity, pH, eutrophication, altitude and poleotolerance) and pollution. Also, a test of beta diversity by Sorensen's index has been done to analyze the community composition between low and high NO<sub>2</sub> levels. The results show that the 4 air quality indicators have a negative relationship with the pollution level up to a certain threshold value (31  $\mu\text{g}/\text{m}^3$  NO<sub>2</sub>), after the threshold value, diversity levelled off is reached. Species traits revealed an urbanity gradient, species are strongly adapted to highly polluted environments. Finally, community composition tended to be similar with high NO<sub>2</sub> levels. The selected pollution indicators can be used to quantify pollution. However, they can only be used to quantify pollution on a fairly fine scale (0 to  $\sim 40\mu\text{g}/\text{m}^3$ ). It is therefore necessary to associate species traits and community composition for a better quantification of the pollution. From these results it is possible to propose a pollution indicator to the participants and to make the project evolve.

## Introduction

Depuis des dizaines d'années, la pollution atmosphérique est un sujet au cœur des préoccupations de la santé publique, écologiques et climatiques (Ehrlich, 1966 ; Darrall, 1988), notamment en milieu urbain avec des villes de plus en plus peuplées et une circulation urbaine de plus en plus dense (Costemalle, 2020). La pollution atmosphérique en France est responsable de près de 40 000 décès par an pour les particules fines et de 7 000 par an pour le  $\text{NO}_2$  (Santé Publique France, 2021). La forte industrialisation, la combustion des énergies fossiles, l'importance des grandes villes dans l'exportation et l'importation des marchandises, la production de déchets et leur incinération renforcent ces émissions. Les sources de pollution sont donc diverses et bien ancrées dans notre société. De plus, ces sources de pollutions rejettent différents composés dans l'air (Fig. 1). Les polluants atmosphériques les plus courants sont les particules fines (Particulate Matter), les oxydes d'azote ( $\text{NO}_x$ ), les oxydes de carbone ( $\text{CO}_x$ ), le dioxyde de soufre ( $\text{SO}_2$ ), les composés organiques volatiles (COV) et les hydrocarbures aromatiques polycycliques (HAP) (Site Airparif ; Elichegaray et al., 2008).

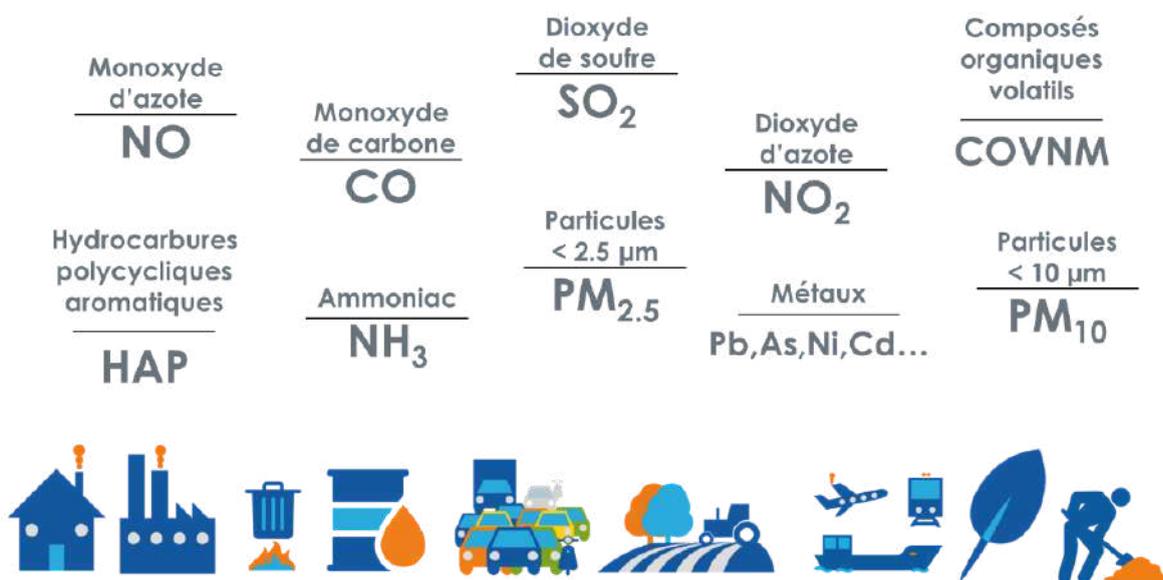


Figure 1. Visualisation des différents polluants et des sources de pollution. (Airparif).

Tous ces polluants proviennent de notre mode de vie actuel et de la façon dont nous gérons la consommation des énergies et le traitement de nos déchets. Malheureusement, ces polluants ont un impact significatif sur les populations locales, qu'elles soient animales ou végétales. Pour certains végétaux, une concentration en polluants trop élevée affecte leur capacité de reproduction (Cuinica et al., 2013) ce qui réduit considérablement la dispersion et le maintien de l'espèce. Pour les animaux, incluant l'humain, des maladies pulmonaires et cardiaques sont développées à cause de la pollution atmosphérique (Yang & Omaye, 2008, Kim et al., 2020).

Au vu des conséquences écologiques et sanitaires actuelles de la pollution atmosphérique (Eymard et al., 2021), les autorités publiques prennent des mesures lors de fortes concentrations de polluants afin de préserver la santé humaine et la biodiversité (Le droit de l'environnement). Par exemple, des seuils de pollution sont définis afin d'agir au mieux en cas d'épisodes extrêmes de pollution (seuil à  $40 \mu\text{g}/\text{m}^3$  annuel pour le  $\text{NO}_2$  ; Directive 2008/50/CE du Parlement européen et du Conseil du 21 mai 2008). Pour quantifier la pollution de l'air, des stations de prélèvements ont été mises en service pour capter et mesurer avec précision la pollution en direct. Les suivis de pollution par ces stations sont faits par des associations agréées de surveillance de la qualité de l'air (AASQA) dans toutes les régions de France métropolitaine (Eymard et al., 2021). Ces organismes créés dans les années 1970 sont chargés de suivre l'évolution de la pollution, d'alerter la population en cas d'excès de pollution, d'aider dans l'amélioration des connaissances et d'accompagner les décideurs sur les actions à mettre en place contre la pollution. Un inconvénient majeur de cette méthode est qu'elle est coûteuse à mettre en place et ne capte la pollution que dans une zone restreinte. Il est donc difficile d'envisager d'installer ces stations partout pour avoir un aperçu très fin dans chaque zone de France.

Afin de renforcer la détection et la quantification de la pollution atmosphérique dans les milieux urbains et naturels, l'utilisation de la bioindication est devenue complémentaire à la station de mesure fixe. La bioindication est apparue dans les années 70 et repose sur l'analyse des communautés biologiques dans un milieu afin de juger la qualité de celui-ci (Ellenberg, 1974). Pour appliquer la bioindication, il est nécessaire d'avoir des espèces bioindicatrices, c'est-à-dire des organismes vivants (animaux ou végétaux), dont la niche écologique (Hutchinson, 1957) est connue et qui permet grâce à un suivi de connaître l'effet d'une perturbation abiotique (Holt & Miller, 2011). Depuis 1970, les bioindicateurs sont largement utilisés dans le contrôle des conditions environnementales afin d'alerter en cas de perturbations biotiques et abiotiques (Burger, 2006). Pour la plupart des milieux, des bioindicateurs peuvent être utilisés, de nombreuses études démontrent l'efficacité de l'utilisation de ces bioindicateurs pour quantifier la pollution. Par exemple, la qualité de l'eau peut être évaluée par les biofilms (Chaumet et al., 2019), par les macroinvertébrés benthiques (Patang et al., 2018) ou par les algues (Barinova & Smith, 2019). La qualité du sol par les bactéries (Santorufò et al., 2012) ou les mycorhizes (Wang et al., 2020). De façon générale, l'étude de la composition des communautés en fonction des changements environnementaux est utilisée pour quantifier l'impact de ces changements. Pour la qualité de l'air, les lichens sont les plus couramment utilisés pour apporter un indice de pollution (Loppi, 2019). Les lichens sont des organismes symbiotiques qui regroupent un champignon (ascomycètes) et

une algue ou une cyanobactérie (Fuisting, 1868 ; Conti & Cecchetti, 2001), ils appartiennent au règne des Fungi. Dans cette symbiose, le champignon a un rôle de support, et de protection pour l'algue contre les conditions extérieures. Il capte également l'eau et les minéraux essentiels à leur développement. L'algue permet grâce à la photosynthèse un apport de sucre au champignon (Hill, 1994). Leur symbiose et leur forme complexe leur permet d'être cosmopolites et de coloniser tous types de milieux. Les lichens se développent sur un large panel de surface comme les troncs, les roches, les murs, les toitures ou autres (Prieto Lamas et al., 1995 ; De Albuquerque et al., 2018). Malgré leur capacité de colonisation, ils n'en restent pas moins des organismes sensibles aux changements environnementaux, tels qu'une modification de l'humidité ambiante, du pH du substrat et de la pollution (Conti & Cecchetti, 2001). Même si certaines espèces sont connues pour avoir une grande valence écologique, certaines espèces sont quant à elles sensibles à des changements assez fins. Cette diversité de sensibilité pour les changements abiotiques font des lichens de bons indicateurs de pollution (Nash III, 2008). L'absence de cuticule et de stomate sur la partie extérieure du lichens fait qu'il n'y a pas de barrières naturelles contre les composants toxiques et les polluants (Conti & Cecchetti, 2001). La partie fongique du lichen agit donc comme une éponge et absorbe toutes les particules environnantes comme les particules et gaz polluants. Il existe diverses méthodes permettant une estimation de la qualité de l'air via les lichens. D'abord, par la bioaccumulation, qui est l'analyse des concentrations de composés chimiques provenant de l'environnement dans des organismes vivants. Ces méthodes mesurent la concentration de polluants présents dans les lichens. Les lichens sont récoltés puis analysés en laboratoires pour permettre de voir les concentrations des différents composés dans l'organisme (Jeran et al., 2002, Frati et al., 2007). Ces méthodes permettent de détecter la pollution de façon précise mais sont très intrusives et détériorent les milieux. Enfin, les méthodes de bioindications utilisent le calcul d'un indice, cette méthode est non intrusive et ne perturbe pas le milieu car elle se base essentiellement sur l'observation des individus. Des études menées par Asta et son équipe visent à quantifier la pollution atmosphérique en calculant un indice issu de la fréquence des espèces présent sur les sites appelé Valeur de Diversité Lichéniques (Asta et al., 2002). D'autres indices sont aussi utilisés comme par exemple l'IAP (Index of Atmospheric Purity) se basant sur la fréquence des individus mais aussi sur leur tolérance face au stress de la pollution atmosphérique (Kricke & Loppi, 2002). Enfin, des études se basent sur la richesse en espèces, la diversité alpha ou bêta pour évaluer la pollution atmosphérique (Matos et al., 2019).

Par ailleurs, avec l'augmentation des soucis liés à l'environnement et la santé publique et l'implication directe des citoyens à ces problématiques, mais aussi de l'avancée des

connaissances scientifiques dans ces domaines, le gouvernement décida en 2013 d'impliquer les citoyens et la recherche dans des projets communs (Houlier & Merilhou-Goudard, 2016 ; Boeuf et al., 2012). Les sciences participatives sont alors créées, le but est d'impliquer des citoyens volontaires sur des projets de recherches ou problématiques actuels (Teyssède & Couvet, 2011). Grâce à ces projets, les participants peuvent élargir leur connaissance et avoir une première sensibilisation à la recherche et au domaine d'étude. Les chercheurs ou la structure porteurs du projet essaient en contrepartie d'accompagner au mieux les participants dans leur recherche, à répondre à leurs interrogations et à leurs faire des retours sur les analyses et résultats obtenus grâce à leur investissement. La force des projets de sciences participatives en écologie réside dans la récolte d'un grand nombre de données par les observateurs permettant de faire des analyses statistiques par la suite. La récolte des données s'effectue selon un protocole bien défini et rigoureux afin d'éviter tout biais lors de l'échantillonnage fait par de nombreux opérateurs. C'est un potentiel gain de temps et de ressources pour les chercheurs qui peuvent alors se focaliser sur l'analyse des données. Par exemple, en écologie, les sciences participatives peuvent permettre de faire un suivi de la biodiversité et de voir son évolution dans le temps, mais aussi de comprendre son déclin (Boeuf et al., 2012 ; Teyssède & Couvet, 2011). La science participative ne cesse de se développer au travers de nombreux projets dans différents domaines (Houlier & Merilhou-Goudard, 2016).

Parmi les programmes de recherches et de sciences participatives en France, Vigie-Nature portés par le MNHN proposent des projets de sciences participatives portant sur le suivi d'escargot, d'oiseaux, de chauve-souris et d'autres espèces (<https://www.vigienature.fr/>). Le projet de science participative Lichens GO! est créé en 2017 par PartiCitaE, un observatoire scientifique sur des projets urbains portés par Sorbonne Université (<http://www.particitae.upmc.fr/fr/index.html>), et Vigie-Nature École, une branche de Vigie-Nature qui se consacre sur les projets de sciences participatives à destination des écoles du primaire et secondaire (<https://www.vigienature-ecole.fr>). Ce projet a pour but d'étudier la qualité de l'air par l'observation des lichens. Le protocole de ce programme se base sur un allègement de l'AFNOR NF 16413 de biosurveillance de la qualité de l'air. Ce projet est en collaboration avec Tela Botanica, le réseau des botanistes français amateurs (<https://www.tela-botanica.org/>), qui aide à l'élaboration d'un outil de saisie des données et à la mise en place des outils d'animations.

Le stage s'est déroulé au sein du Centre d'Ecologie et des Sciences de la Conservation (CESCO) qui est une unité mixte de recherche (UMR) rattachée au Muséum national d'histoire naturelle (MNHN) de Paris et sous tutelle du CNRS, de Sorbonne Université et du

MNHN. Les sujets abordés au CESCO sont axés sur toutes les questions de conservation et de compréhension des processus de biodiversités, mais aussi sur l'accompagnement des transformations sociales nécessaire pour la conservation des milieux et des espèces. Les domaines de recherches peuvent aller de la recherche appliquée à la conservation en passant par des études sociales. L'UMR est composée de deux pôles : un pôle recherche et un pôle sciences participatives. Le CESCO aborde toutes les notions qui sont liées de manière directe ou indirecte à l'écologie et la conservation. Au sein de la plateforme de sciences participatives se trouve le programme Vigie-Nature regroupant divers projets de sciences participatives dont Lichens GO!.

L'objectif de ce stage est d'analyser les jeux de données Lichens GO! de 2017 et de 2021 afin de valider un indice de pollution qui serait adéquat au protocole du projet et ensuite de proposer d'éventuelles améliorations au niveau du protocole. Pour ce faire, des analyses de la relation entre les données issues du protocole Lichens GO! et des relevés de pollution de NO<sub>2</sub> sont utilisés afin de tester la validité du protocole. Plusieurs indices sont utilisés pour quantifier la pollution à partir des lichens, tels que la richesse spécifique, la diversité alpha et bêta, la LDV et la somme des fréquences des espèces. Ces indices seront comparés à différentes valeurs de pollution tels que la médiane, le 3<sup>e</sup> quartile, le nombre d'heures au-dessus du seuil de pollution.

## **Matériel et Méthodes**

### *Données de lichens*

Pour cette étude, différents jeux de données de lichens sont à disposition. Tout d'abord, le jeu de données Lichens GO! de 2017 qui rassemble 34 sites et 102 arbres. Les relevés se sont fait dans la région Île-de-France. Parmi ces sites, 27 sont à proximité de stations de relevés de pollution Airparif. Ensuite, un jeu de données Lichens GO! de 2021 relevé au mois de juin, avec 24 sites et 72 arbres correspondant aux sites fait en 2017 et étant les plus proches de stations de relevés de Airparif. Enfin, il y a la base du projet Lichens GO! qui regroupe les observations des participants, cela représente 86 sites et 258 arbres, les sites sont répartis dans toute la France.

Les sites d'observations retenus dans le protocole Lichens GO! doivent respecter plusieurs critères. Le site doit avoir au moins trois arbres alignés et espacés d'une distance comprise entre 2 et 10 mètres. En effet, à moins de 2 mètres de distance entre les arbres, on risque d'avoir une ambiance de sous-bois et d'observer d'autres espèces de lichens que celles ciblées par le protocole. Au-dessus de 10 mètres, il est compliqué de considérer les différents arbres comme des répliquats. Les écorces des espèces ne doivent pas être à pH acide ou se renouveler régulièrement (platanes, bouleaux). Les arbres doivent faire une circonférence d'au moins 30 centimètres, avec un tronc droit et une hauteur sous la première branche à au moins 2 mètres du sol afin d'avoir un écoulement de l'eau sur le tronc le plus uniforme possible (Fig. 2). Pour la collecte des données, trois faces de chaque arbre sont échantillonnées grâce à une grille de cinq quadrats alignés de 10 centimètres chacun. Cet échantillonnage permet d'avoir accès à la fréquence des espèces et avoir un proxy assez fiable de l'abondance sans demander aux participants de dénombrer les individus de chaque espèce.

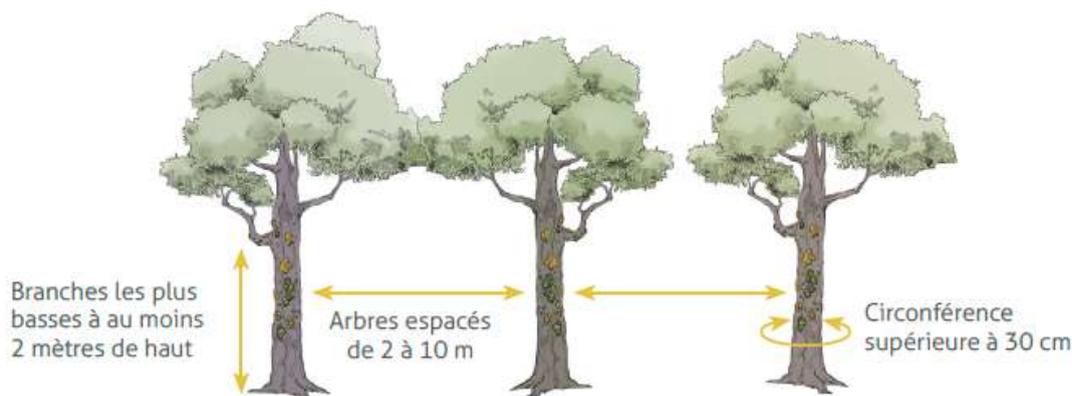
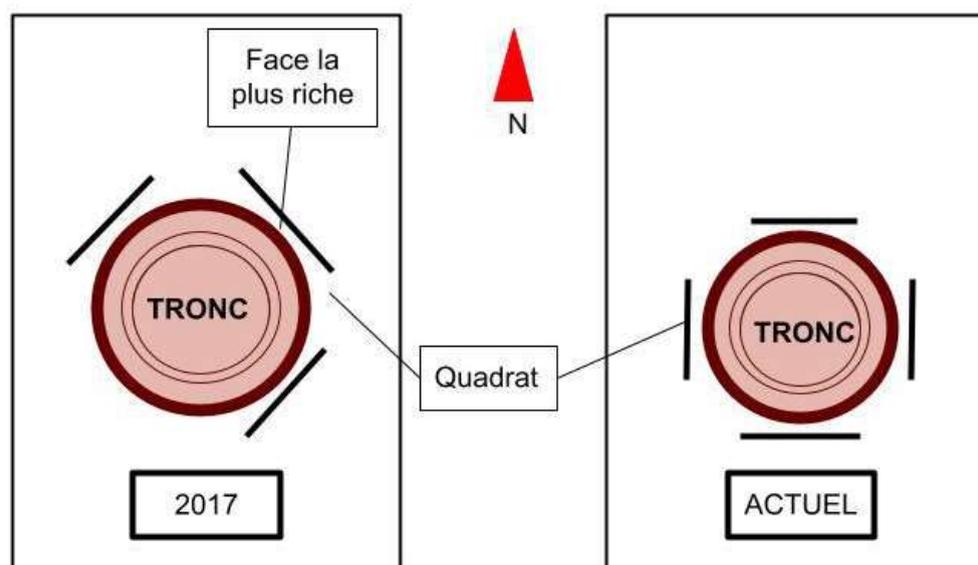


Figure 2. Illustration de la sélection des arbres du protocole Lichens GO! (Lichens GO!)

La première face observée est celle ayant le plus de diversité spécifique estimée par l'observateur, les deux autres faces sont celles se situant de part et d'autre de la première. La base de la grille est placée à 1 mètre du sol sur la face du tronc et les espèces de lichens y sont observées. Sur une fiche de terrain, les coordonnées géographiques des arbres, leur espèce, l'exposition des faces selon les points cardinaux, l'ombrage du tronc y sont notés. Les espèces de lichens sont observées à la loupe de botaniste dans les quadrats et leur absence ou présence dans chaque quadrat y est noté. La détermination des lichens est basée sur le guide de détermination des lichens de France (Van Haluwyn et al., 2009). Les données détaillées issues des observations sont rentrées dans une base de données avec les observations pour chaque quadrat de chaque face de chaque arbre, avec la localisation GPS des arbres.

Le protocole reste très similaire à la première version. Le nombre d'arbres, les essences, l'espacement et la morphologie des troncs restent inchangés. En revanche, la circonférence passe de 30 centimètres à 40 centimètres minimum, le nombre de faces de 3 à 4 faces et la position des faces est définie par leur exposition selon les quatre points cardinaux (nord, est, sud et ouest ; Fig. 3).



Les lichens sont déterminés selon une clé d'identification mise en place pour le projet par PartiCitaE et Vigie-Nature école (Turcati & Turpin, 2018) afin de se concentrer uniquement sur les espèces les plus communes et faciles à identifier. La clé de détermination regroupe 41 espèces ou groupes d'espèces et se base sur des critères morphologiques présents chez les lichens comme la forme du thalle, la couleur du thalle, la présence ou absence de structures morphologiques (cils, isidies, soralies). Le type de quadrat reste inchangé et l'absence/présence de lichens est toujours utilisée. Pour la saisie des données, les participants doivent se rendre sur la page Internet de saisie sur le site de Tela Botanica<sup>1</sup> afin de saisir leurs données informatiquement.

<sup>1</sup>(<http://www.particitae.upmc.fr/fr/suivez-les-lichens/saisie-de-donnees-lichens-go.html>)

## *Donnée de pollution atmosphérique*

Les données sont issues du site officiel de l'association Airparif (<https://www.airparif.asso.fr>). Ces données proviennent des stations permanentes qui mesurent la concentration des polluants dans l'atmosphère de façon continue et la concentration est moyennée à l'heure. L'étude se concentre uniquement sur les relevés horaires en NO<sub>2</sub> en µg/m<sup>3</sup> sur les 27 stations présentes en région Île-de-France. Le jeu de données comporte les relevés allant du 1<sup>er</sup> janvier 2014 au 31 décembre 2020. Cette fourchette de temps est nécessaire car les lichens ne réagissent pas immédiatement aux changements, mais après une certaine période de temps (que nous avons estimé ici de trois à cinq ans). De plus, en complément des données de pollution, les coordonnées GPS des différentes stations sont référencées.

## *Sélection des critères de pollution*

De nombreuses méthodes sont utilisées pour calculer un indice de pollution via les lichens. Une revue de la littérature a permis de sélectionner les indices les plus couramment utilisés. De plus, dans un but de vulgarisation, les indices doivent être simples à expliquer et à mettre en place pour les participants (Schiller et al., 2001). Quatre indices ont été retenus pour leur efficacité et leur simplicité : la richesse spécifique, la diversité Shannon, l'indice de pollution Valeur de la Diversité Lichéniques (Asta et al., 2002) et la somme des fréquences des espèces. Le choix d'utiliser la somme des fréquences des espèces résulte de l'analyse de l'indice de pollution LDV qui semble être aussi une formule calculant la somme des fréquences de façon un peu plus complexe qu'une somme des fréquences classiques. De plus, pour tester l'hétérogénéité des populations au sein des sites, une analyse de la diversité bêta est effectuée pour comparer les différents groupes entre eux et observer si les groupes sont différents entre eux. Pour calculer la diversité bêta, l'indice de Sorensen est utilisé, indice qui se base sur l'absence ou la présence de l'espèce au sein d'un site.

Pour vérifier si les indices lichéniques sélectionnés sont corrélés aux mesures physiques de pollution, une étude préalable de ces données de pollution est nécessaire. Les relevés sur 6 ans de pollutions peuvent être exploités de différentes manières : ici la recherche porte sur les critères les plus pertinents. Le but était de trouver des indices qui reflètent au mieux les périodes de pollutions cycliques et régulières et de ne pas prendre des indices issus d'événements particuliers pour analyser au mieux l'effet de la pollution sur les différents

indicateurs lichéniques. Une des premières analyses a permis de savoir si la pollution en NO<sub>2</sub> était constante dans une journée et dans une année. Pour mettre en évidence les tendances de pollution, les analyses se concentrent uniquement sur les mesures ayant dépassé la valeur seuil définie par le code de l'environnement, à savoir 40 µg/m<sup>3</sup>. Tout d'abord, les analyses graphiques ont permis de mettre en évidence des heures et mois plus propices à la pollution. La moyenne en NO<sub>2</sub> n'a pas été prise en compte car la pollution atmosphérique dépend de nombreux paramètres, tels que la température, les activités humaines, les événements rares. Les valeurs issues des moyennes peuvent être fortement influencées par des épisodes courts et spontanés de forte pollution et ne sont donc pas un reflet d'une constance de pollution. La médiane a alors été utilisée car moins sensible aux épisodes de pollution rares et isolées. Ensuite, les autres indices de pollutions utilisés pour les analyses sont : le 3<sup>e</sup> quartile, le nombre d'heures consécutives au-dessus d'un certain seuil et le nombre d'heures cumulées au-dessus du seuil de pollution normal (40 µg/m<sup>3</sup>). La valeur minimum et la valeur maximum de NO<sub>2</sub> devaient être utilisées dans les analyses, cependant comme ces valeurs dépendent aussi d'évènements parfois soudains et rares, elles peuvent apporter des biais dans les analyses et n'ont donc pas été utilisées.

Enfin, seuls les résultats des tests faits en fonction de la médiane en NO<sub>2</sub> seront présentés car le nombre d'heures au-dessus du seuil de pollution et le nombre d'heures consécutif au-dessus du seuil ne donnent pas de résultats significatifs et le 3<sup>ème</sup> quartile suit la même tendance que la médiane.

### *Sélection des sites d'échantillonnages de lichens 2021*

En 2017, 24 relevés de lichens ont été faits au plus proche des stations de relevés de pollution Airparif (Fig. 4). Afin d'actualiser les données et d'appliquer le protocole Lichens GO! en vigueur, les sites ont été ré-échantillonnés. De plus, ces nouvelles données pourront être comparées, après standardisation aux données de 2017 pour observer potentiellement l'effet de changement de pollution. Les sites à échantillonner sont sélectionnés par une analyse de distance spatiale entre les stations et les sites avec le package R *sf* et la fonction *nearestpoint*. La distance entre les stations de relevé de pollution et les sites Lichens GO! a été calculée par la méthode du plus proche voisin. Pour obtenir de meilleurs résultats et limiter l'effet de la distance dans les analyses futures, seuls les sites à moins de 200 mètres des stations ont été sélectionnés dans la région Île-de-France. J'ai réalisé ces relevés durant deux semaines en juin 2021 (Fig. 5).

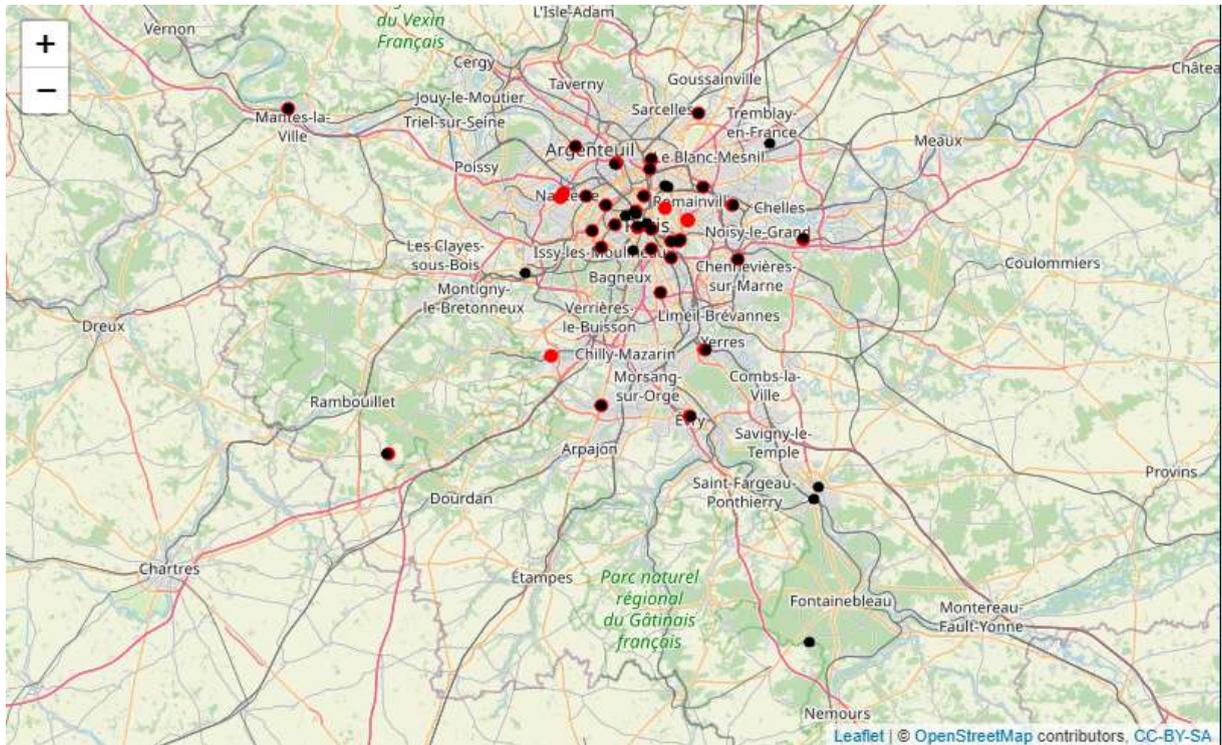


Figure 4. Distribution des sites de 2017 sur la région Île-de-France. Les points rouges représentent les sites de lichens et les points noirs les stations de relevés de la pollution.

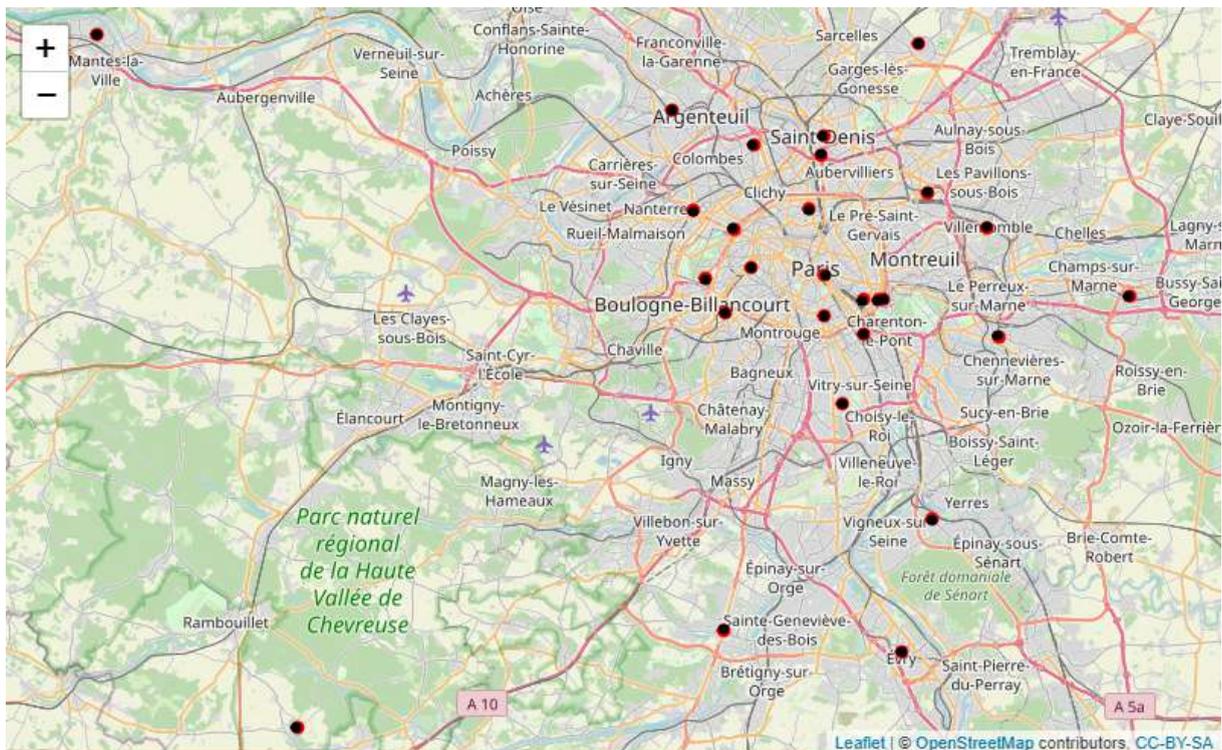


Figure 5. Distribution des sites échantillonnés en 2021 sur la région Île-de-France. Les points rouges représentent les sites de lichens et les points noirs les stations de relevés de la pollution.

## *Nettoyage des données*

Une fois toutes les données récupérées, une des principales activités fut de corriger et nettoyer les bases de données. Il y avait à disposition pour ce projet 8 bases de données différentes dont 3 comportaient les relevés de lichens. Dans un premier temps, l'objectif a été de corriger une anomalie qui est survenue durant une courte période sur le site de saisie de données de Tela Botanica. En effet, une partie des données étaient inutilisables car certaines observations de lichens n'étaient pas affiliées à un arbre précis du site d'étude. Une partie de ces données a pu être corrigée grâce à l'utilisation des coordonnées GPS. Pour le reste des données, un mail fut envoyé aux participants dont les données étaient incomplètes afin de récupérer leur fiche terrain et corriger manuellement les données manquantes. Ensuite, le format des tableaux fut réarrangé grâce au package *tidyr* et *dplyr* afin de faciliter les analyses et supprimer les informations non essentielles ou doubles. Pour les données de relevés de pollution, le jeu de données a aussi été réorganisé pour faciliter les analyses après importation sous R.

## *Traits d'espèces*

En plus des quatre indices (LDV, richesse spécifique, indice de Shannon et somme des fréquences des espèces) pour quantifier la pollution via les lichens, des analyses porteront sur la relation entre la proportion d'espèces nitrophiles et la pollution en NO<sub>2</sub>, ainsi que sur l'impact de la pollution sur 6 traits de vie des lichens : le pH, la poléotolérance, l'irradiation solaire, l'eutrophisation, l'altitude et l'aridité. Pour chaque espèce, un indice de chacun des traits de vie a été associé selon les recherches de Nimis et Martellos en 2007. Pour homogénéiser les indices, les valeurs initiales des 6 traits de vie ont été calculées pour correspondre à une échelle de 0 à 1. Puis, pour chaque site d'études, les valeurs ont été pondérées par la fréquence des espèces au sein de chaque site. Enfin, pour chaque espèce, une échelle de nitrophilie fut ajoutée selon le catalogue de Claude Roux (Roux, 2017). Au total, huit catégories de nitrophilie furent créés « Non nitrophile », « Non\_Peu », « Non\_Moyen », « Peu\_Moyen », « Moyen\_Tres », « Nitrophile », « Nitrotolerant » et « Heminitrophile ». Pour cette étude, seules les classes « Nitrophile » et « Nitrotolérant » ont été utilisées car selon la littérature, seules les espèces nitrophiles sont étudiées pour quantifier la pollution de l'air (Gadson et al., 2010).

Les analyses ont toutes été menées sous le logiciel open source R (v. 4.0.4, 2021). Les librairies utilisées pour cette étude sont : *dplyr* (Wickham et al., 2021), *tidyr* (Wickham, 2021), pour le nettoyage des données. La librairie *sf* (Pebesma, 2018), pour calculer la distance entre des points géographiques. La librairie *mgcv* (Wood et al., 2016), pour des modèles non linéaires de type *gam* ; *lmtest*, pour des modèles linéaires. La librairie *ade4* et *car*, pour analyser les modèles linéaires. La librairie *factoextra* (Kassambara & Mundt, 2020), pour les analyses en composantes principales. La librairie *corrplot* (Wei & Simko, 2017), pour créer des matrices de corrélations. La librairie *vegan* (Oksanen et al., 2020), pour calculer les indices de diversités alpha et bêta. La librairie *ggplot2* (Wickham, 2016), pour créer les différentes représentations graphiques. La librairie *leaflet* (Cheng et al., 2021), pour créer des cartes.

### *Modèles statistiques*

Les modèles statistiques sont utilisés pour tester la relation entre les 4 indicateurs de la qualité de l'air et la médiane en NO<sub>2</sub>. Pour les jeux de données de 2017 seul, puis 2017 et 2021 combinés, un poids est ajouté aux modèles qui représente la distance entre le site d'étude et la station de relevé de pollution la plus proche. Ce poids est calculé de la façon suivante :  $1/\text{distance}^2$ . Pour les modèles de la base de données 2017 et 2021 combinés, un effet "année" est ajouté. Les modèles linéaires sont d'abord testés mais ne sont pas significatifs. Après vérification graphique, les modèles non linéaire de type *gam* du package *mgcv* sont utilisés.

Pour tester les résultats obtenus par les modèles non linéaires, un test d'application linéaire par morceaux est utilisé pour mettre en évidence les différentes pentes des modèles par la fonction *piecewise.linear* de la librairie *SiZer* (Sonderegger, 2020). Enfin, chaque pente est testée par des modèles linéaires. Ce test d'application linéaire par morceaux se base uniquement sur les jeux de données 2017 et 2021 combinés car ce sont les données des sites de lichens le plus proche des stations et cela permet de faire des tests sur un jeu de données plus grands.

### *Etude des traits d'espèces*

Pour étudier la relation entre les traits d'espèces et la pollution en NO<sub>2</sub>, une ACP a permis de déterminer si les sites peuvent être différenciés par les 6 traits de vie (poléotolérance, ensoleillement, pH, eutrophisation, altitude, aridité). L'ACP est effectuée sur le jeu de données 2017 et 2021 combinés pour avoir un jeu de données plus robuste. La fonction *dudi.pca* de la librairie *ade4* est utilisée pour produire l'ACP, les représentations graphiques sont effectuées avec la librairie *factoextra*. Enfin, les valeurs du premier axe de l'ACP sont récupérées et utilisées afin de produire un indice de pollution basé sur les traits de vie et étudier sa relation avec la médiane en NO<sub>2</sub> dans un modèle non linéaire. Lors des analyses, le premier axe de l'ACP sera défini comme un gradient d'urbanité.

### *Diversité bêta*

La diversité bêta est calculée pour observer le changement de composition d'espèces de lichens entre les sites. Cette diversité est calculée sur le jeu de données de 2021 pour avoir seulement les données d'une année avec le nouveau protocole, et selon le point de rupture qui délimite deux groupes à différentes teneur de NO<sub>2</sub>. La dissimilarité des sites à pu être étudiée via l'indice de Sorensen. Cet indice se base sur la présence ou absence des espèces dans chaque site et fait une comparaison deux à deux pour calculer la dissimilarité entre les sites. L'indice de Sorensen est calculé grâce à la fonction *vegdist* du package *vegan*.

Pour confirmer que les deux groupes de NO<sub>2</sub> diffèrent sur l'indice de Sorensen, le test de comparaison de moyenne Kruskal-Wallis est utilisé. En effet, les conditions de validité ne sont pas respectées pour effectuer une ANOVA, le test non paramétrique est donc utilisé.

## Résultats

### Analyses des 4 indices : LDV, Richesse, Shannon, Somme des fréquences

Les premiers tests sont portés sur l'analyse des quatre indices de diversités pour quantifier la pollution. Pour pouvoir déterminer les indices qui réagissent au mieux à la pollution, différents modèles sont testés sur les différents jeux de données (Annexe 1). Les résultats dans le jeu de données combiné (2017 et 2021) et la base de 2017 seul montrent une relation non linéaire entre les quatre indices de qualité de l'air et les concentrations de pollutions. Pour la base de données de 2021, seule la relation entre la richesse spécifique et la médiane en  $\text{NO}_2$  est significative et non linéaire. Pour la plupart des modèles, on retrouve les mêmes tendances graphiques : la courbe montre une baisse rapide des indicateurs de biodiversités avant d'atteindre un plateau (Fig. 6).

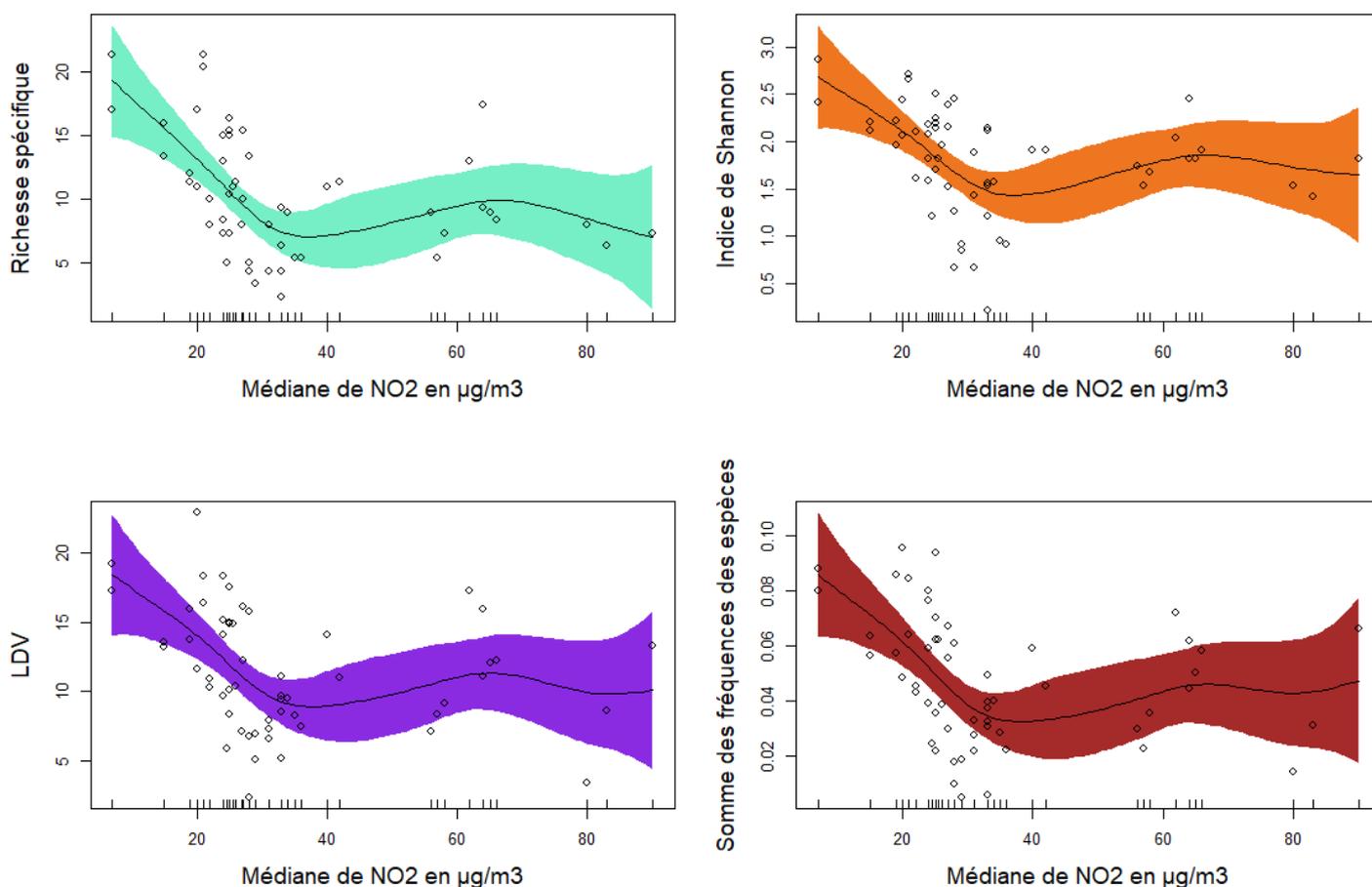


Figure 6. Relation entre les 4 indicateurs de la qualité de l'air et la médiane de  $\text{NO}_2$ . Modèle GAM du jeu de données 2017 et 2021. Les barres au niveau de l'axe des abscisses et les points représentent les sites.

Les modèles sur le jeu de données des participants Lichens GO! donnent des résultats différents. Pour trois indices (LDV, indice de Shannon et richesse spécifique), les résultats montrent une baisse de la diversité avec une augmentation de la pollution en NO<sub>2</sub> comme par exemple entre Shannon et la médiane de NO<sub>2</sub> (Fig. 7).

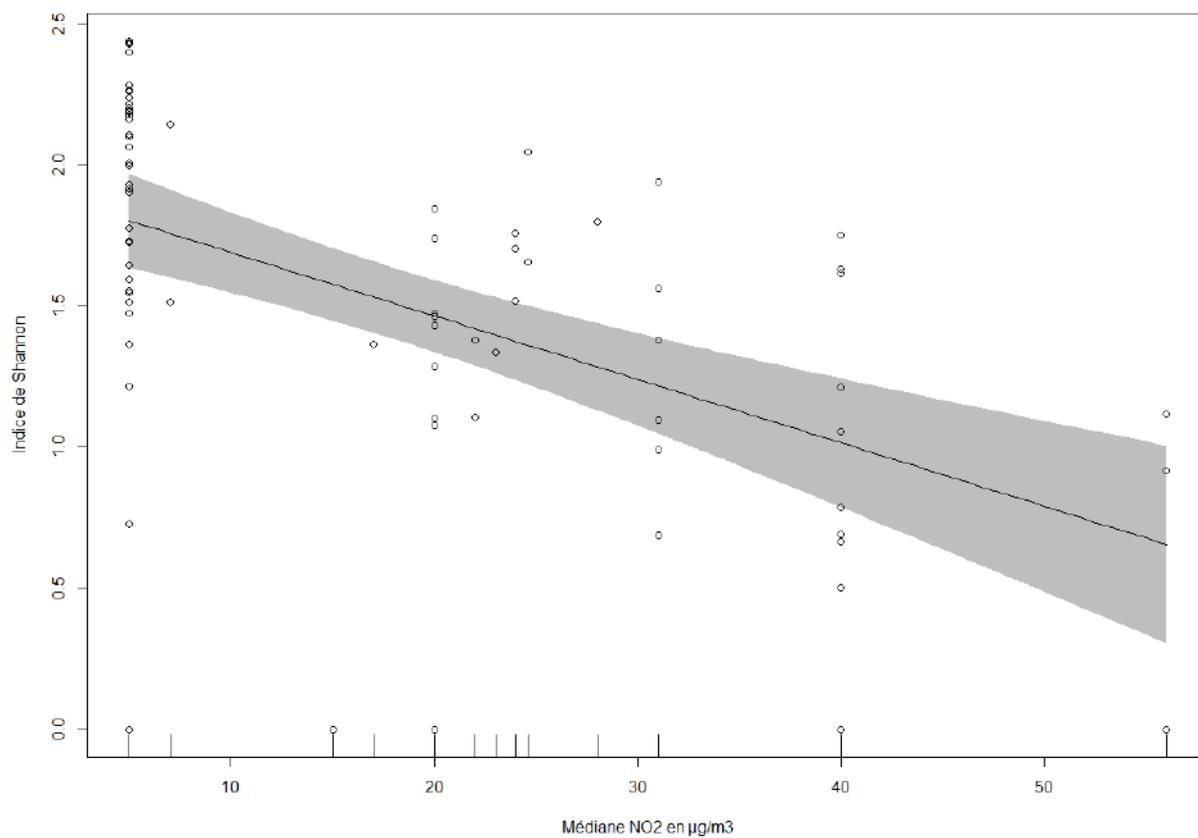


Figure 7. Relation entre l'indice de Shannon et la médiane en NO<sub>2</sub> (p-value < 0.001, R<sup>2</sup> : 26,8 %).

Modèle linéaire sur le jeu de données des participants Lichens GO!.

### Etude du point de rupture de la pente

Sur les graphiques des modèles non linéaires, la forme de la courbe laisse à suggérer une rupture de pente qui départage le jeu de données en deux parties. Ici, le test présenté se base sur les données de 2017 et 2021 regroupées et sur la relation entre la LDV et la médiane en  $\text{NO}_2$ , car c'est le modèle non linéaire le plus significatif. Deux groupes semblent alors se distinguer du graphique, avec un premier groupe en dessous de  $32 \mu\text{g}/\text{m}^3$  de  $\text{NO}_2$  et un groupe au-dessus (Fig. 8).

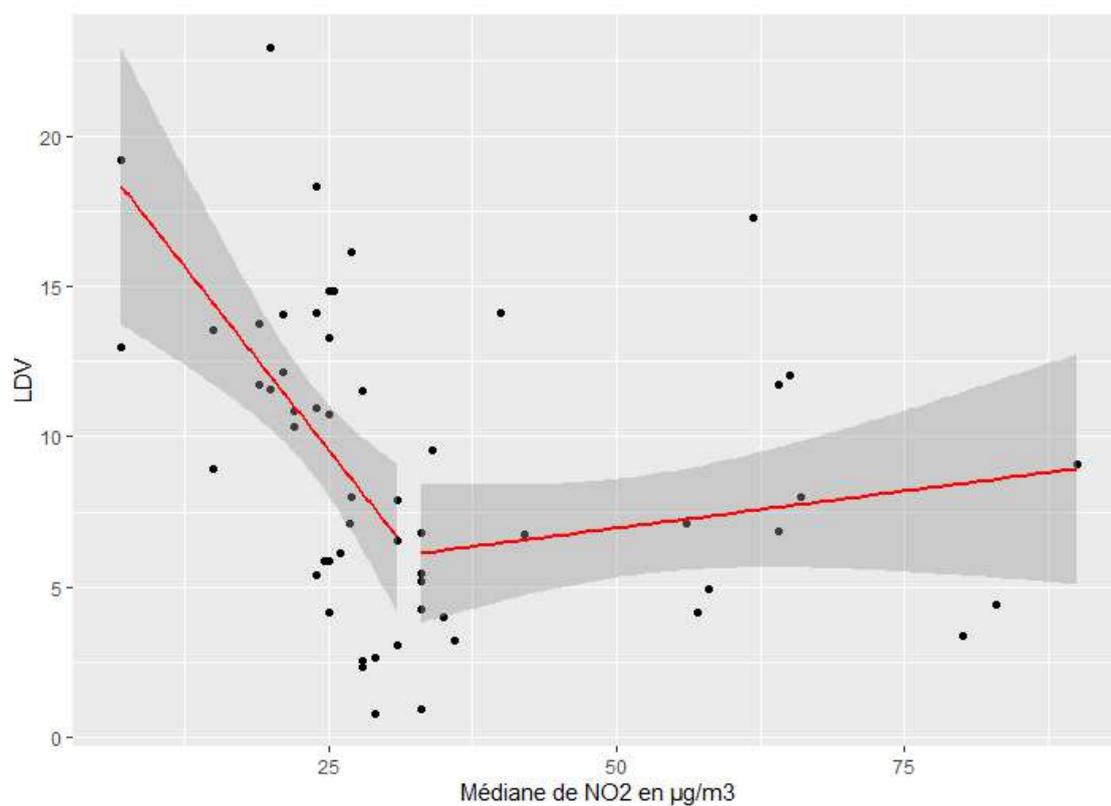


Figure 8. Représentation de la rupture de pente avec les deux droites de régressions. Jeu de données 2017 et 2021 combinés.

Ensuite, les deux régressions linéaires sont testées. Le modèle linéaire avant le point d'inflexion montre qu'il y a bien une relation négative entre l'indice de Shannon et la médiane en NO<sub>2</sub> (p-value < 0,001, R<sup>2</sup> : 42 %). Il n'y a cependant aucune relation significative après le point d'inflexion. Le résultat est identique pour les trois autres indices de pollution (Fig. 9).

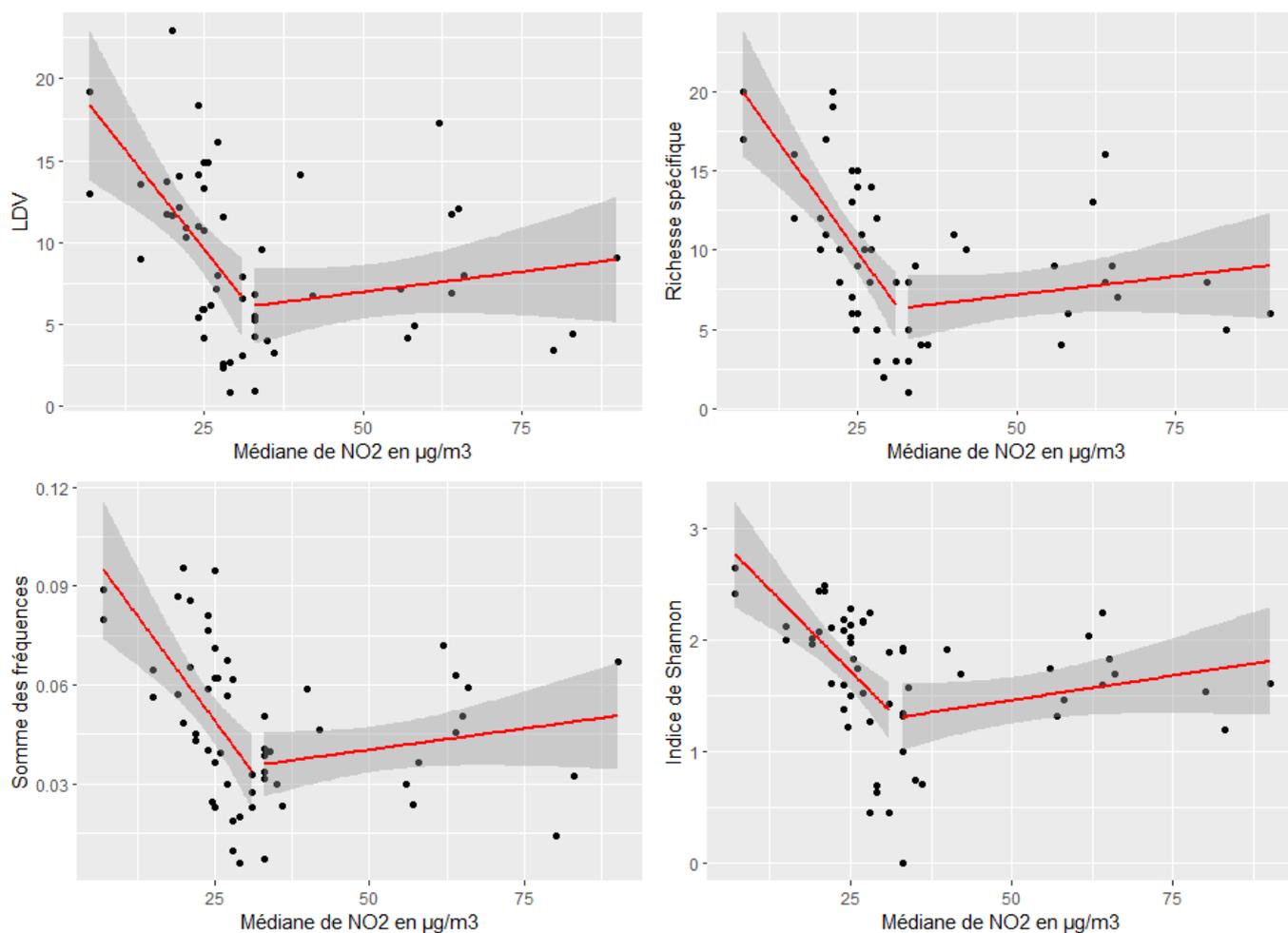


Figure 9. Représentation des régressions linéaires entre les différents indicateurs de la qualité de l'air et la pollution de NO<sub>2</sub>. Jeu de données combiné 2017 et 2021.

## Relation entre les traits de vie et la pollution

Sur le jeu de données combinant 2017 et 2021, le premier axe de l'ACP (52,6 % de la variance) regroupe l'aridité, l'eutrophisation, l'altitude et la poléotolérance (Fig. 10). L'eutrophisation est la variable la mieux représentée sur le plan vectoriel du premier axe, tandis que le deuxième axe (20,0 % de la variance) est défini par le pH et l'ensoleillement.

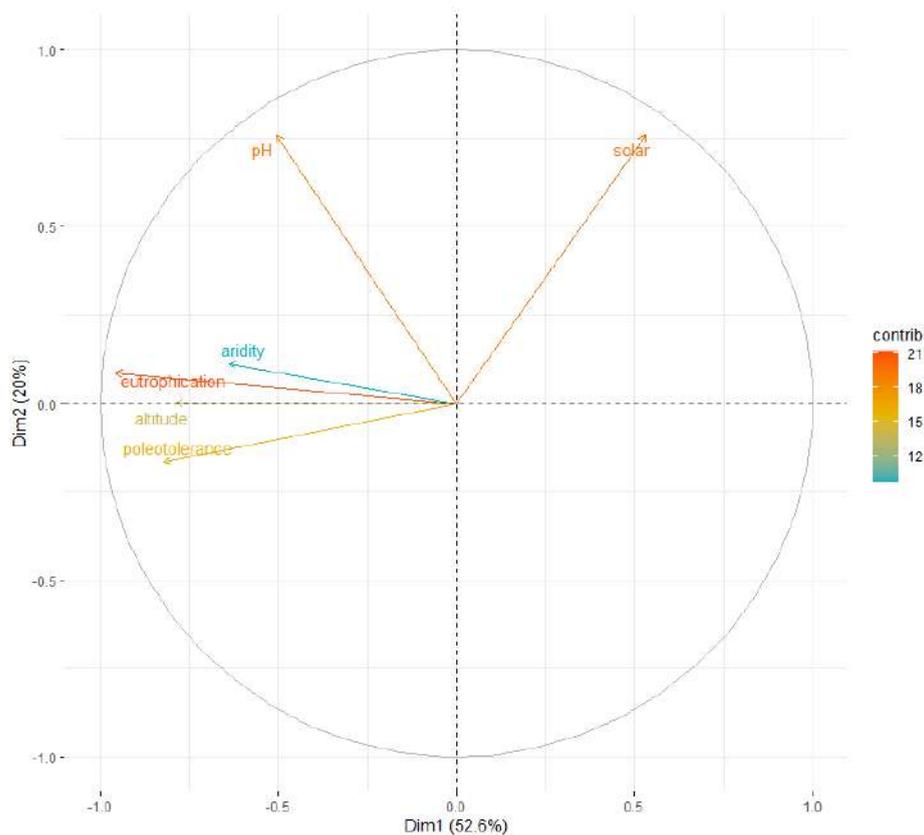


Figure 10. Cercle de corrélation des 6 traits de vie issus de l'ACP. Jeu de données 2017 et 2021 combinées.

Ensuite les valeurs du premier axe de l'ACP sont extraites et testées dans un modèle non linéaire. Le résultat montre une corrélation positive (p-value : 0.000983 \*\*\* ;  $R^2$  : 32,6%) entre la médiane en  $\text{NO}_2$  et les valeurs du premier axe représentant un gradient d'urbanité (Fig. 11).

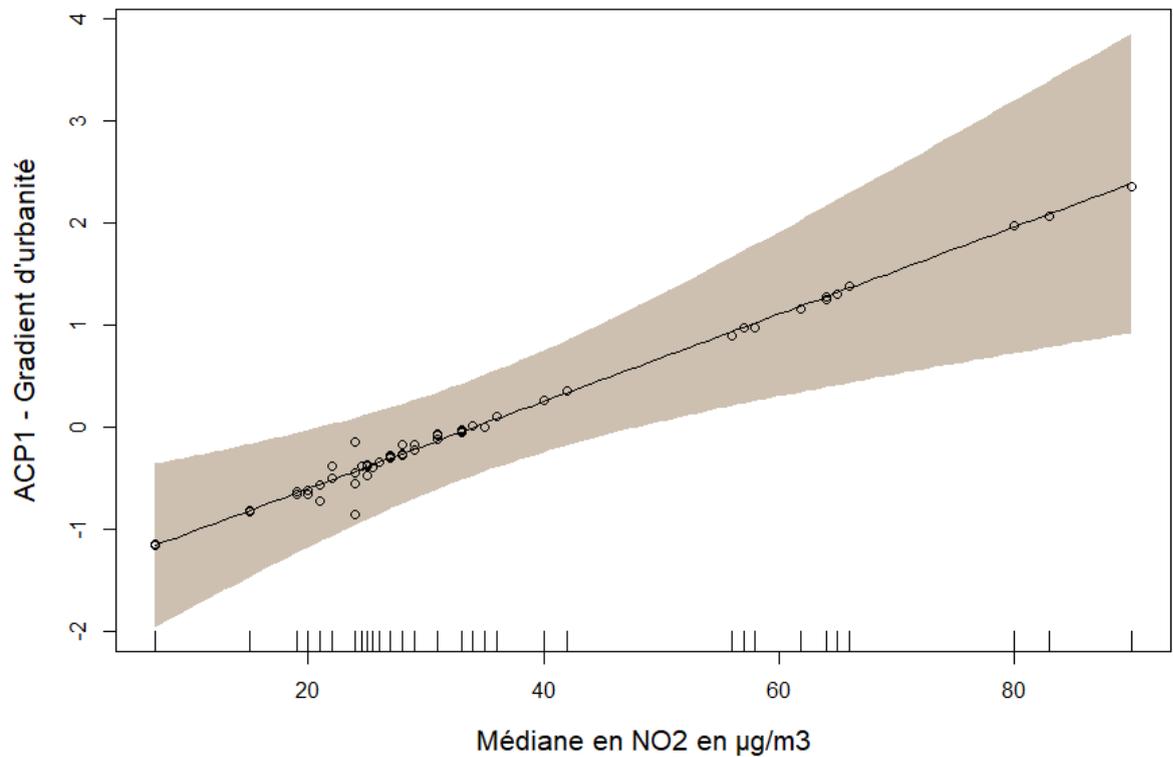


Figure 11. Relation entre le gradient d'urbanité issu du premier axe et la médiane de  $\text{NO}_2$  (p-value : 0.000983 \*\*\* ;  $R^2$  : 32,6%). Jeu de données 2017 et 2021 combinés.

L'analyse de la composition des communautés en fonction du seuil de NO<sub>2</sub> établis par les modèles non linéaires (31 µg/m<sup>3</sup>), montre qu'il y a bien une différence entre les deux catégories de sites (p-value = 2,5e<sup>-06</sup>), les sites exposés à un taux de NO<sub>2</sub> plus faible ont tendance à être différent tandis que les sites exposés à un taux plus élevé sont plus similaires (Fig. 12).

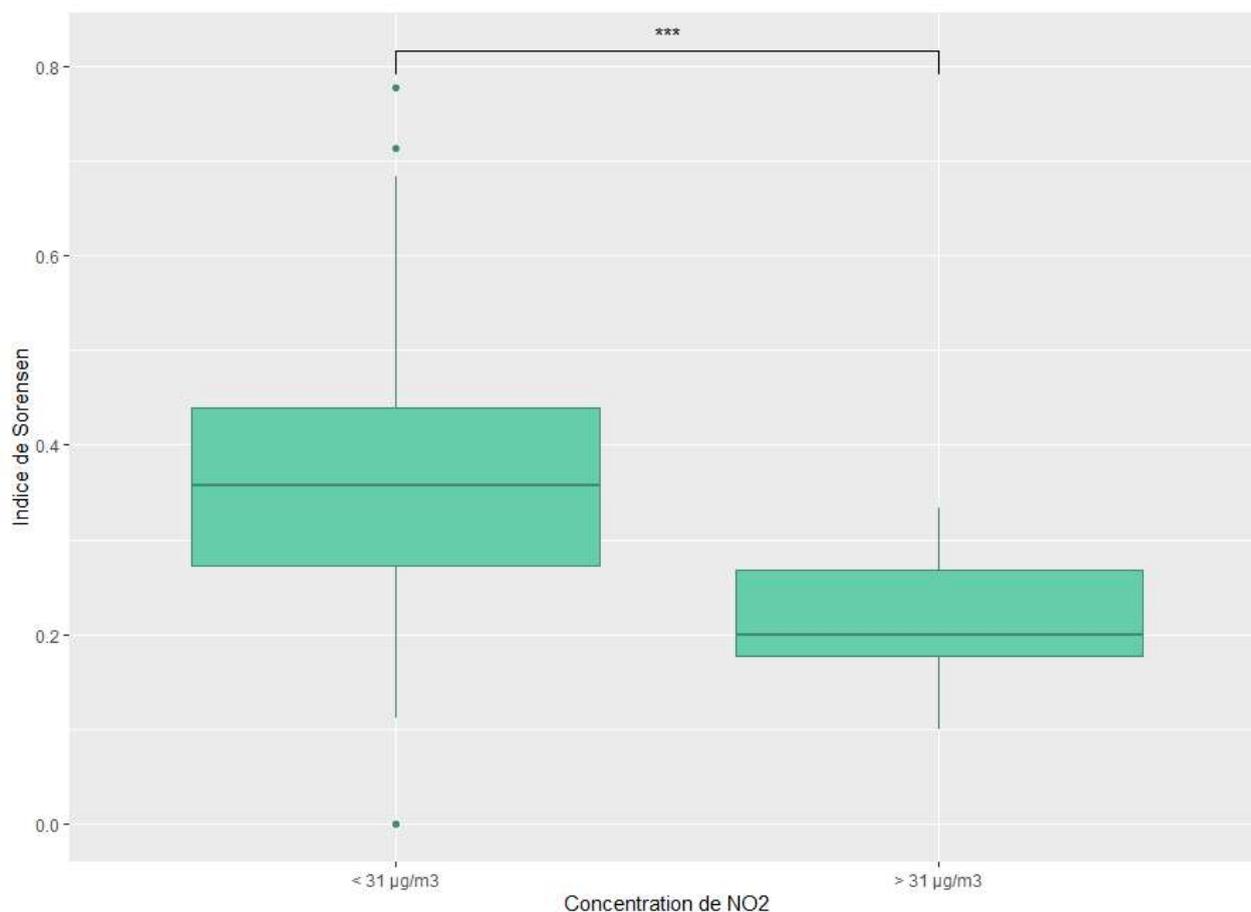


Figure 12. Différence de l'indice de dissimilarité de Sorensen entre les deux groupes de sites de lichens avant et après 32 µg/m<sup>3</sup> de NO<sub>2</sub> (p-value = 2,5e<sup>-06</sup>).

## Discussion

La pollution atmosphérique est à l'origine de nombreux problèmes environnementaux et sanitaires. Dans le but d'approfondir les connaissances et de sensibiliser les citoyens sur les connaissances dans ce domaine des projets de sciences participatives sont élaborés. Dans le cadre du projet de sciences participatives Lichens GO!, le but de ce projet est de suivre la qualité de l'air via les lichens sur les milieux urbains et semi-urbains. L'objectif de ce stage est de définir un indicateur de la qualité de l'air à partir des données issus du programme et qui conviendra au protocole du projet.

Lors des tests des quatre indicateurs de la qualité de l'air sélectionnés (LDV, indice de Shannon, richesse spécifique et somme des fréquences des espèces), les modèles linéaires n'ont pas donné de résultats significatifs. Après vérification de la répartition des données, les modèles non linéaires sont utilisés et confirment la présence d'une relation non linéaire entre les indicateurs de la qualité de l'air et la médiane en NO<sub>2</sub>. Cependant, seule la richesse spécifique en fonction de la médiane en NO<sub>2</sub> est significative sur les différents jeux de données. Grâce au test *piecewise.linear* les résultats montrent la présence d'un seuil de NO<sub>2</sub> (31 µg/m<sup>3</sup>) qui délimite les données en deux parties. La présence de ce seuil n'est pas connue dans la littérature. Le seuil est proche du seuil définis par les autorités gouvernementales sur la pollution de l'air pour le NO<sub>2</sub> (40 µg/m<sup>3</sup>). Ce seuil est le reflet d'une valeur limite à partir de laquelle la composition des communautés est différente.

L'analyse de la première pente, mise en évidence par le *piecewise.linear*, montre une régression linéaire de l'indice de la qualité de l'air en fonction de la médiane en NO<sub>2</sub>. En effet, la diversité spécifique tend à diminuer avec l'augmentation du taux de NO<sub>2</sub> (Fig. 8 et 9). Il est important de noter que malgré la non significativité lors des modèles non linéaires sur le jeu de données de 2021 pour trois des indices de la qualité de l'air (LDV, somme des fréquences et indice de Shannon), les résultats du *piecewise.linear* sont similaires et significatifs pour les quatre indices. Ce résultat conforte notre hypothèse de base où la diversité de lichens diminue en fonction de l'augmentation de NO<sub>2</sub>. En effet, des études ont déjà montré que, par exemple, la richesse spécifique chez les lichens est fortement corrélée avec les taux de polluants présents dans l'atmosphère (Davies et al., 2007 ; Matos et al., 2019 ; Stevens et al., 2004 ; Sterzyńska et al., 2018). Les espèces sensibles à la pollution comme *Evernia prunastri* tendent à disparaître avec des taux de polluants élevés (Lackovičová et al., 2013) faisant baisser la diversité spécifique des milieux.

Cependant, au-delà du seuil, les résultats montrent une pente différente de la première. A partir de ce seuil, la rupture de pente est fortement marquée. De plus, la seconde pente après le seuil de 31 µg/m<sup>3</sup> n'est pas significative, donc la diversité de lichens ne diminue pas

et n'augmentent pas avec des concentrations de NO<sub>2</sub> plus élevées. Ce qui laisse à suggérer qu'un plateau est atteint. Les indices de diversités ne semblent alors pas suffisant pour quantifier la pollution atmosphérique après une certaine concentration de NO<sub>2</sub>. A partir de ce point la diversité est probablement composée d'espèces adaptées aux types de milieux fortement anthropisés et pollués (Gadson et al., 2010). Cette seconde pente est à analyser avec précaution car le nombre de sites après le point de rupture est faible et peut ne pas être suffisant pour en dégager une tendance. Pour pallier ce problème, il serait intéressant d'échantillonner d'autres sites de lichens proches de stations de pollutions à fort taux de NO<sub>2</sub>.

Les indicateurs de pollutions sélectionnées peuvent quantifier le niveau de pollution à une échelle assez fine (entre 0 µg/m<sup>3</sup> et environ 40 µg/m<sup>3</sup>) mais au-delà de ce point, toute interprétation doit se faire avec prudence. Cependant, en Île-de-France, peu de stations de pollution dépassent les 40 µg/m<sup>3</sup> et ces indicateurs peuvent donc être utilisés. Il n'est donc pas impossible de quantifier et d'avoir une idée du taux de pollution lors du relevé d'un site. A partir de ces indices, il est envisageable de créer différentes méthodes pour quantifier au mieux cette pollution, comme par exemple en créant des classes de pollutions en lien avec la richesse spécifique obtenue sur le site d'étude afin de savoir approximativement le taux de pollution présent. Le jeu de données Lichens GO! possède assez d'observations pour créer une échelle basée sur les relevés en utilisant le site le plus pauvre et le plus riche comme échelle basse et échelle haute. Cette échelle sera donc en constante évolution avec l'apport de nouvelles données qui pourraient délimiter une nouvelle échelle. Ce système est déjà utilisé dans des études avec d'autres indicateurs comme la LDV (Boonpeng et al., 2018 ; Estrabou et al., 2011). La LDV est un indice largement répandu et utilisé dans les études de pollution (Castello & Skert, 2004 ; Llop et al., 2011, Loppi, 2019). Le calcul de la LDV étant aussi fortement corrélé à la somme des fréquences, il serait donc aussi envisageable d'utiliser la somme des fréquences plutôt que la LDV. En effet, la somme des fréquences serait plus facile à expliquer aux observateurs et peut-être un bon proxy du recouvrement des espèces.

En plus des indices de pollution, le gradient d'urbanité révèle aussi son utilité dans la quantification de la pollution. Les sites ayant des taux de NO<sub>2</sub> plus élevés semblent être des sites dont les communautés de lichens sont adaptées à des taux d'eutrophisation plus élevés, mais aussi de pollutions, avec une préférence pour les pH acides et les milieux plus arides. Ces résultats confortent les hypothèses de bases où les espèces dans des milieux plus exposés à la pollution y sont plus résistantes. Dans les zones à fortes teneurs en NO<sub>x</sub>, on retrouve des effets antagonistes comme l'eutrophisation qui entrave la sensibilité de certaines espèces (Davies et al., 2007). Ce gradient d'urbanité peut venir en complément des indicateurs de la qualité de l'air pour apporter de la robustesse dans la quantification de la pollution atmosphérique. Malgré une potentielle relation entre le taux d'espèces nitrophiles et la teneur

en NO<sub>2</sub> dans la littérature (Gadsdon et al., 2010), les données ne montrent pas de relation entre nitrophilie et taux de NO<sub>2</sub>.

La diversité bêta et de la composition des communautés qui sont présentes dans les sites, montrent que les sites avec des taux de NO<sub>2</sub> plus élevés (> 31 µg/m<sup>3</sup>) ont des compositions d'espèces similaires. Les sites ayant des taux de NO<sub>2</sub> en dessous du seuil de 31 µg/m<sup>3</sup> possèdent une large diversité d'espèces. Malheureusement, il semble difficile d'utiliser ce critère pour différencier les sites avec un faible taux de NO<sub>2</sub> et un fort taux de NO<sub>2</sub> car les espèces présentes dans les milieux à fort NO<sub>2</sub> peuvent se retrouver dans des sites à faible NO<sub>2</sub> et inversement (Annexe 2). Pour pouvoir utiliser ce critère et différencier les sites et quantifier la pollution en NO<sub>2</sub> des analyses complémentaires sont nécessaires. Il faudrait analyser plus précisément les espèces présentes dans les sites et sélectionner les espèces indicatrices de faibles, moyens et forts taux de NO<sub>2</sub>, comme par exemple en utilisant *Evernia prunastri* comme indicateur de basse pollution (Blasco et al., 2008).

Au vu des différents résultats, il est possible de donner une explication à la valeur seuil observé sur la figure 7. En effet, jusqu'au point de rupture la diversité spécifique diminue par la perte des espèces sensibles à la pollution tel que les lichens fruticuleux. Ensuite, après cette limite de pollution, il est possible d'imaginer que seules les espèces adaptées au fort taux de pollution résistent. Ce résultat peut être validé par l'étude de la bêta diversité qui montre la même composition au sein des sites au-dessus du seuil.

Il est important de rappeler que ce stage porte sur un seul polluant. Il est compliqué de savoir si les résultats obtenus durant ces analyses sont uniquement dus à la pollution au NO<sub>2</sub>. En effet, en milieu urbain, il existe de nombreux polluants atmosphériques qui sont susceptibles d'altérer la diversité des lichens tels que les HAPs ou le SO<sub>2</sub> (Nimis et al., 1991 ; Blasco et al., 2006). De plus, il existe peu d'études sur le rayon d'action et la circulation des polluants en ville. De plus, la concentration des polluants atmosphériques peut varier en fonction des vents, de la température, de l'humidité relative et de la topographie urbaine (Wallace et al., 2010 ; Ilten & Selici, 2008). De ce fait, un site avec une grande diversité de lichens peut être à plusieurs mètres d'un site très pollué.

Les sciences participatives ne cessent de prendre de l'ampleur en France (Houllier & Merilhou-Goudard, 2016). Les projets de sciences participatives permettent un double bénéfice, un bénéfice pour la recherche et un bénéfice pour le social. En effet, les projets permettent l'avancement des connaissances, grâce par exemple à des suivis réguliers de populations pour quantifier l'effet des changements globaux (Julliard, 2017). Mais aussi d'impliquer au plus les citoyens aux problématiques sanitaires et écologiques sur leur environnement (Julliard, 2017). Un des nombreux enjeux des sciences participatives est de

toujours impliquer les participants aux projets et de maintenir leur attachement aux projets. Auquel cas les projets peuvent voir le nombre de participants décroître dans le temps.

C'est pour cela que des travaux comme ce stage sont nécessaires pour faire vivre les projets et donner une nouvelle impulsion d'engouement aux participants. Ce travail permet aussi de montrer l'efficacité et l'intérêt du projet. De plus, les premières analyses sont concluantes et un indice de la qualité de l'air peut être mis en place et permettre aux participants d'avoir une estimation de la pollution atmosphérique immédiatement lors de l'observation des lichens. Ensuite, il est envisageable d'implémenter cette étude de stage dans l'application Galaxy Bricks. Galaxy Bricks est une application développée par le programme Vigie-Nature École et permet aux participants, notamment les enseignants et les élèves, de faire des analyses statistiques simplifiées sur les données récoltées. Cet outil permet de sensibiliser et d'enseigner aux élèves la démarche scientifique au niveau de l'analyse de données (problématique, données nécessaires, analyses pertinentes). Avec le travail fourni lors de ce stage, des premières analyses sur Lichens GO! seront proposés aux participants. Enfin, grâce aux résultats, d'autres analyses sont possibles notamment sur la composition des communautés mais aussi sur l'amélioration du protocole.

## **Conclusion**

Les résultats montrent que les 4 indicateurs de la qualité de l'air ont une relation négative avec le taux de pollution jusqu'à un certain seuil ( $31 \mu\text{g}/\text{m}^3 \text{NO}_2$ ), après le seuil un plateau est atteint. Les traits d'espèces ont mis en évidence un gradient d'urbanité où l'on retrouve des espèces fortement adaptées à des milieux fortement pollués. Enfin, la composition des communautés tend à être similaire avec des taux de  $\text{NO}_2$  élevés. Les indicateurs de pollutions sélectionnées peuvent être utilisés pour la quantification de la pollution. Cependant, ils ne permettent de la quantifier que sur une échelle de pollution assez fine (0 à  $\sim 40 \mu\text{g}/\text{m}^3$ ). Il est donc nécessaire d'y associer les traits d'espèces et la composition des communautés pour une meilleure quantification de la pollution. A partir de ces résultats il est possible de proposer un indicateur de pollution aux participants et de faire évoluer le projet. Pour approfondir les résultats et les analyses, il serait intéressant d'échantillonner d'autres sites de lichens proches des stations de relevé de pollution. Comme la composition lichénique des sites dépend de nombreux facteurs, il serait alors intéressant de prendre en compte plusieurs polluants.

## Bibliographie

Asta J., Erhardt W., Ferretti M., Fornasier F., Kirschbaum U., Nimis P.L., Purvis O.W., Pirintsos S., Scheidegger C., Van Haluwyn C. & Wirth W. (2002) Mapping lichen diversity as an indicator of environmental duality. In: Nimis P.L., Scheidegger C., Wolseley, P.A. (eds.): *Monitoring with Lichens – Monitoring Lichens. Kluwer Academic, Dordrecht*, 273–279.

Barinova S., Smith, T. (2019). Algae diversity and ecology during a summer assessment of water quality in the Abraham Lincoln Birthplace National Historical Park, USA. *Diversity*, 11: 206-226.

Blasco M., Domeño C., Nerín, C. (2006) Use of lichens as pollution biomonitors in remote areas: comparison of PAHs extracted from lichens and atmospheric particles sampled in and around the Somport tunnel (Pyrenees). *Environ. Sci. Technol.*, 40: 6384-6391.

Blasco M., Domeño C., Nerín C. (2008) Lichens biomonitoring as feasible methodology to assess air pollution in natural ecosystems: Combined study of quantitative PAHs analyses and lichen biodiversity in the Pyrenees Mountains. *Anal. Bioanal. Chem.*, 391: 759-771.

Bœuf G., Allain Y-M., Bouvier M. (2012) L'apport des sciences participatives à la connaissance de la biodiversité en France. *La Lettre de l'OCIM*, 8-18.

Boonpeng C., Sriviboon C., Polyiam W., Sangiamdee D., Watthana S., Boonpragob K. (2018) Assessing atmospheric pollution in a petrochemical industrial district using a lichen-air quality index (LiAQI). *Ecol. Indic.*, 95: 589-594.

Burger J. (2006) Bioindicators: A Review of Their Use in the Environmental Literature 1970–2005.

Castello M., Skert N. (2005) Evaluation of lichen diversity as an indicator of environmental quality in the North Adriatic submediterranean region. *Sci. Total Environ.*, 336: 201-214.

Chaumet B., Morin S., Hourtané O., Artigas J., Delest B., Eon M., Mazzella N. (2019) Flow conditions influence diuron toxicokinetics and toxicodynamics in freshwater biofilms. *Sci. Total Environ.*, 652: 1242-1251.

Cheng J., Karambelkar B., Xie Y. (2021) leaflet: Create Interactive Web Maps with the JavaScript 'Leaflet' Library. R package version 2.0.4.1. <https://CRAN.R-project.org/package=leaflet>

Costemalle, V. (2020). Bayesian Probabilistic Population Projections for France. *Economics and Statistics*, 520-52: 29-47.

- Conti M., Cecchetti G. (2001) Biological monitoring: lichens as bioindicators of air pollution assessment - A review. *Environ. Pollut.*, 114: 471-492.
- Cuinica L. G., Abreu I., Esteves da Silva J. (2014) Effect of air pollutant NO<sub>2</sub> on *Betula pendula*, *Ostrya carpinifolia* and *Carpinus betulus* pollen fertility and human allergenicity. *Environ. Pollut.*, 186: 50-55.
- Darrall N. M. (1988) The effect of air pollutants on physiological processes in plants. *Plant Cell Environ.*, 12: 1-30.
- Davies L., Bates J., Bell J., James P., Purvis O. (2007) Diversity and sensitivity of epiphytes to oxides of nitrogen in London. *Environ. Pollut.* 146: 299–310.
- De Albuquerque M., Putzke J., Schünemann A., Vieira F., Victoria F., Pereira, A. (2018) Colonisation of stranded whale bones by lichens and mosses at Hennequin Point, King George Island, Antarctica. *Polar Rec.*, 54: 29-35.
- Van Haluwyn C., Asta J., Gavériaux J-P. (2009) Guide des lichens de France: lichens des arbres. *Belin, Paris* : 241 pages.
- Elichegaray C., Bouallala S., Maitre A., Ba M. (2008) État et évolution de la pollution atmosphérique. *Rev. mal. respir.*, 26: 191-206.
- Ellenberg H. (1974) Indicator values of vascular plants in central Europe. *Scr. Geobot.*, 9.
- Estrabou C., Filippini E., Soria J.P., Schelotto G., Rodriguez J.M. (2011) Air quality monitoring system using lichens as bioindicators in Central Argentina. *Environ. Monit. Assess.*, 182: 375–383.
- Eymard L., Turcati L., Payan S. (2021) Pollution de l'air en ville : cartographie, microcapteurs et sciences participatives. *The conversation*.
- Fрати L., Santoni S. Nicolardi V., Gaggi C., Brunialti G., Guttova A., Gaudino S., Pati A., Pirintsos S.A., Loppi S. (2007) Lichen biomonitoring of ammonia emission and nitrogen deposition around a pig stockfarm. *Environ. Pollut.*, 146: 311-316.
- Fuisting W. (1867) Zur entwicklungsgeschichte der Pyrenomyceten. *Botanisches Zentralblatt*, 25: 177-311.
- Gadsdon S.R., Dagley J.R., Wolseley P.A., Power S.A. (2010) Relationships between lichen community composition and concentrations of NO<sub>2</sub> and NH<sub>3</sub>. *Environ. Pollut.*, 158: 2553-2560.

Gombert S., Asta J., Seaward M.R.D. (2003) Correlation between the nitrogen concentration of two epiphytic lichens and the traffic density in an urban area. *Environ. Pollut.*, 123: 281-290.

Hill D. J. (1994) The nature of the symbiotic relationship in lichens. *Endeavour*, 18: 96-103.

Holt E. A., Miller S. W. (2011) Bioindicators: using organisms to measure. *Nature*, 3: 8-13.

Houllier F., Merilhou-Goudard J. B. (2016) Les sciences participatives en France: Etats des lieux, bonnes pratiques et recommandations. *Les sciences participatives en France* : 124 pages.

Hutchinson, G.E. (1957) Concluding remarks. Cold Spring Harbor Symposium. *Quant. Biol.*, 22, 415–427.

Ilten N., Selici A.T. (2008) Investigating the impacts of some meteorological parameters on air pollution in Balikesir, Turkey. *Environ. Monit. Assess.*, 140: 267-277.

Jeran Z., Jaimovi R., Batič F., Mavsar R. (2002) Lichens as integrating air pollution monitors. *Environ. Pollut.*, 120: 107-113.

Julliard R. (2017) Science participative et suivi de la biodiversité: l'expérience Vigie-Nature. *Nat. Sci. Soc.*, 25: 412-417.

Kassambara A., Mundt F. (2020) factoextra: Extract and Visualize the Results of Multivariate Data Analyses. R package version 1.0.7. <https://CRAN.R-project.org/package=factoextra>

Kim, D., Chen Z., Zhou L-F., Huang S-X. (2018) Air pollutants and early origins of respiratory diseases. *Chronic Dis. Transl. Med.*, 4: 75-94.

Kricke R., Loppi S. (2002) Bioindication: The I.A.P. Approach. In: Nimis P. L., Scheidegger C., Wolseley P. (eds): *Monitoring with Lichens — Monitoring Lichens. Kluwer academic Publishers*, 21-37.

Lackovičová A., Guttova A., Backor M., Pisut P., Pisut I. (2013) Response of Evernia prunastri to urban environmental conditions in Central Europe after the decrease of air pollution. *The Lichenologist*, 45: 89-100.

Llop E., Pinho P., Matos P., Pereira M.J., Branquinho C. (2011) The use of lichen functional groups as indicators of air quality in a Mediterranean urban environment. *Ecol. Indic.*, 13: 215-221.

Loppi S. (2019) May the diversity of epiphytic lichens be used in environmental forensics?. *Diversity*, 11: 36-49.

Losacco C., Perillo A. (2018) Particulate matter air pollution and respiratory impact on humans and animals. *Environ. Sci. Pollut. Res.*, 34: 33901-33910.

Matos P., Vieira J., Rocha B., Branquinho C., Pinho P. (2019) Modeling the provision of air-quality regulation ecosystem service provided by urban green spaces using lichens as ecological indicators. *Sci. Total Environ.*, 665: 521-530.

Nash III T.H. (2008) Lichens biodiversity. Second edition. *Cambridge university, New York* : 486 pages.

Nimis P.L., Lazzarin G., Gasparo D. (1991) Lichens as bioindicators of air pollution by SO<sub>2</sub> in the Veneto Region. *Int. j.*, 11: 3-76

Oksanen J., Blanchet F.G., Friendly M., Kindt R., Legendre P., McGlenn D., Minchin P. R., O'Hara R.B., Simpson G.L., Solymos P., Stevens M.H.H., Szoecs E., Wagner H. (2020) vegan: Community Ecology Package. R package version 2.5-7. <https://CRAN.R-project.org/package=vegan>

Opal (2013) The OPAL air survey : Take Part. (<https://www.imperial.ac.uk/opal/surveys/airsurvey/>)

Patang F., Soegianto A., Hariyanto S. (2018) Benthic macroinvertebrates diversity as bioindicator of water quality of some rivers in East Kalimantan, Indonesia. *Int. J. Ecol.*

Pebesma E., (2018) Simple Features for R: Standardized. Support for Spatial Vector Data. *The R Journal*, 10: 439-446.

Prieto Lamas B., Rivas Brea M.T., Silva Hermo B.M. (1995) Colonization by lichens of granite churches in Galicia (northwest Spain). *Sci. Total. Environ.*, 167: 343–351.

Santorufu L., Van Gestel C. A., Rocco A., Maisto G. (2012) Soil invertebrates as bioindicators of urban soil quality. *Environ. Pollut.*, 161: 57-63

Schiller A., Hunsaker C.T., Kane M.A., Wolfe A.K., Dale V.H., Suter G.W., Russell C.S., Pion G., Jensen M.H., Konar V.C., (2001) Communicating Ecological Indicators to Decision Makers and the Public. *Conserv. Ecol.* 5.

Sonderegger D. (2020) SiZer: Significant Zero Crossings. R package version 0.1-7.

<https://CRAN.R-project.org/package=SiZer>

Stevens C. J., Dise N. B., Mountford J. O., Gowing D. J. (2004) Impact of nitrogen deposition on the species richness of grasslands. *Science*, 303: 1876-1879.

Sterzyńska M., Nicia P., Zadrozny P., Fiera C., Shrubovych, J., Ulrich W. (2018) Urban springtail species richness decreases with increasing air pollution. *Ecol. Indic.*, 94: 328-335.

Teyssède A., Couvet D. (2011) Biodiversité et science participative, de la recherche à la gestion des écosystèmes. *SFE<sup>2</sup>*, 1-8.

Turcati L., Turpin S. et al. (2018) Livret de participation au programme Lichens Go !. *PartiCitaE, Paris* : 20 pages.

Wallace J., Corr D., Kanaroglou P. (2010) Topographic and spatial impacts of temperature inversions on air quality using mobile air pollution surveys. *Sci. Total Environ.*, 408: 5086-5096.

Wang Q., Lu H., Chen J., Jiang Y., Williams M. A., Wu S., Li J., Liu J., Yang G., Yan, C. (2020) Interactions of soil metals with glomalin-related soil protein as soil pollution bioindicators in mangrove wetland ecosystems. *Sci. Total Environ.*, 709: 136051-136060.

Wei T., Simko V. (2017) R package "corrplot": Visualization of a Correlation Matrix (Version 0.84). Available from <https://github.com/taiyun/corrplot>.

Weldon J., Grandin U. (2021) Weak recovery of epiphytic lichen communities in Sweden over 20 years of rapid air pollution decline. *The Lichenologist*, 53: 203–213.

Wickham H. (2016) ggplot2: Elegant Graphics for Data Analysis. *Springer-Verlag New York*, 2016.

Wickham H. (2021) tidyr: Tidy Messy Data. R package version 1.1.3. <https://CRAN.R-project.org/package=tidyr>

Wickham H., François R., Henry L., Müller K. (2021) dplyr: A Grammar of Data Manipulation. R. package version 1.0.5. <https://CRAN.R-project.org/package=dplyr>

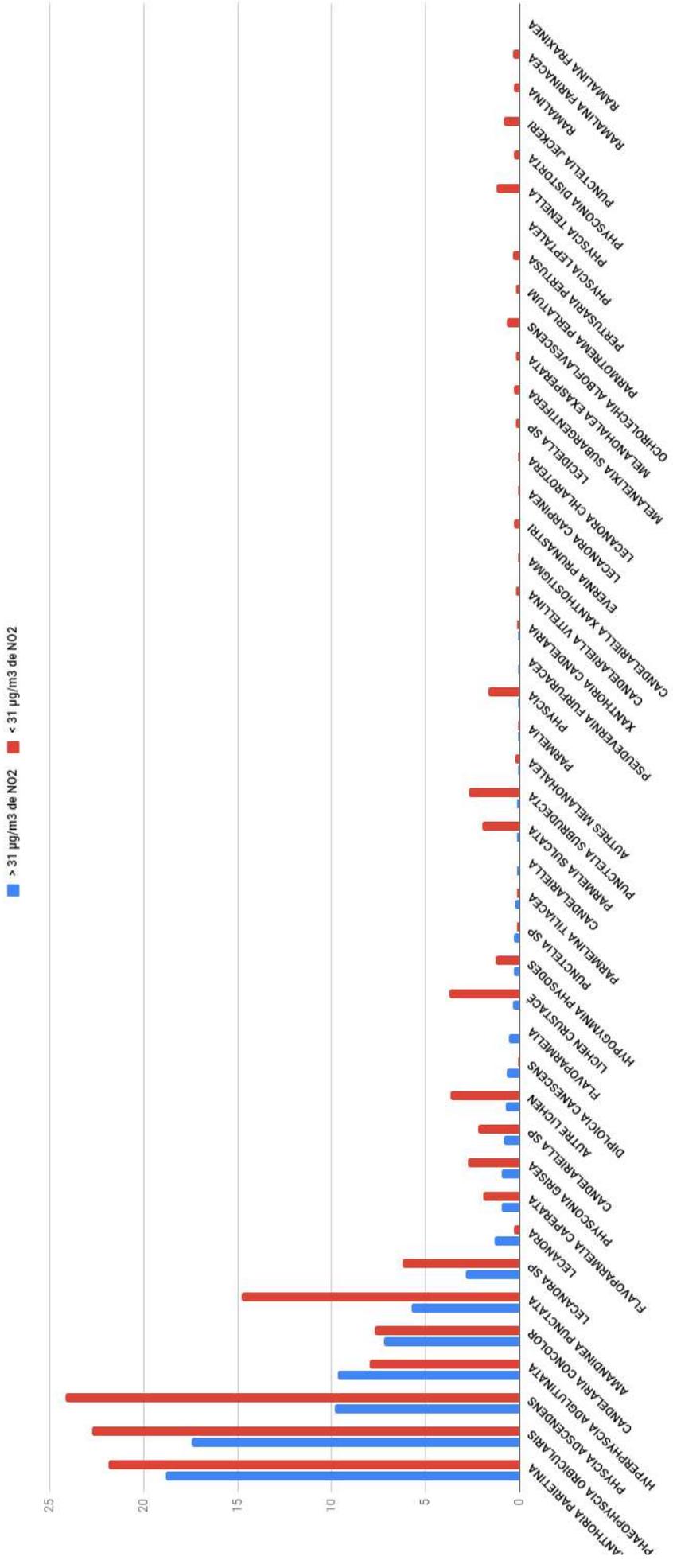
Wood S.N., Pya N., Saefken B. (2016) Smoothing parameter and model selection for general smooth models (with discussion). *J. Amer. Statist. Assoc.*, 111: 1548-1575.

Yang W., Omaye S. (2009) Air pollutants, oxidative stress and human health. *Mutat. Res.*, 674: 45–54.



	LDV ~ Médiane du nombre d'heures en continue au dessus du seuil de NO <sub>2</sub>	0.292	/
	Richesse ~ Médiane NO2	0.0483*	37.4%
	Richesse ~ 3e Quartiles de NO2	0.029*	40.1%
	Richesse ~ Nombre d'heures au dessus du seuil de NO <sub>2</sub>	0.298	/
	Richesse ~ Médiane du nombre d'heures en continue au dessus du seuil de NO <sub>2</sub>	/	/
	Shannon ~ Médiane NO2	0.113	/
	Shannon ~3e Quartiles de NO2	0.07	/
	Shannon ~ Nombre d'heures au dessus du seuil de NO <sub>2</sub>	0.298	/
	Shannon ~ Médiane du nombre d'heures en continue au dessus du seuil de NO <sub>2</sub>	0.751	/
	Somme des fréquences ~ Médiane en NO2	0.201	/
	Somme des fréquences ~ 3e Quartiles de NO2	0.147	/
	Somme des fréquences ~ Nombre d'heures au dessus du seuil de NO <sub>2</sub>	0.551	/
	Somme des fréquences ~ Médiane du nombre d'heures en continue au dessus du seuil de NO <sub>2</sub>	/	/
<b>Lichens GO! 2017 + 2021 (N=58)</b>	LDV ~ Médiane NO2	7.17e-05***	54%
	LDV ~ 3e Quartiles de NO2	8.58e-05***	57.6%
	LDV ~ Nombre d'heures au dessus du seuil	0.144	/
	LDV ~ Médiane du nombre d'heures en continue au dessus du seuil	/	/
	Richesse ~ Médiane NO2	8.07e-05***	45 %
	Richesse ~ 3e Quartiles de NO2	6.71e-05***	47.7 %
	Richesse ~ Nombre d'heures au dessus du seuil	0.192	/

	Richesse ~ Médiane du nombre d'heures en continue au dessus du seuil	0.09	/
	Shannon ~ Médiane NO2	0.00164 **	38.5 %
	Shannon ~3e Quartiles de NO2	0.00166 **	39.7 %
	Shannon ~ Nombre d'heures au dessus du seuil	0.531	/
	Shannon ~ Médiane du nombre d'heures en continue au dessus du seuil	0.306	/
	Somme des fréquences ~ Médiane en NO2	0.00018 ***	39.1%
	Somme des fréquences ~ 3e Quartiles de NO2	0.000477 ***	49.6 %
	Somme des fréquences ~ Nombre d'heures au dessus du seuil	0.452	/
	Somme des fréquences ~ Médiane du nombre d'heures en continue au dessus du seuil	/	/



Annexe 2. Moyenne du nombre d'individus par espèce en fonction du taux de NO2. Jeu de données 2017 et 2021 combinés.

